Technische Universität München

Department of Mathematics

REGIME SWITCHING OF HYDROLOGICAL TIME SERIES USING COPULAS

Joshwa Georgekutty

Supervisor: Prof. Claudia Czado, Ph.D.

Master's Thesis

September 16, 2025

Acknowledgments

I want to express my deepest gratitude to Professor Claudia Czado for their invaluable guidance, constructive feedback, and continuous encouragement throughout this thesis.
I also acknowledge the use of OpenAI's ChatGPT in assisting with phrasing suggestions (abstract), improving the readability of the text, and generating LaTeX code for specific tables. All analyses, interpretations, and conclusions presented in this report are entirely my own.

Abstract

Hydrological time series are shaped by complex, nonlinear dynamics that vary across different regimes of flow and weather conditions. Traditional linear models struggle to capture such regime-dependent behavior, motivating the use of flexible copula-based approaches. This thesis develops a regime-switching vine copula framework to analyze three hydrological variables measured at the Noce River's Ott station: water level (height), temperature, and electrical conductivity. The data consist of over 20,000 hourly observations (2021–2023), supplemented by precipitation inputs from two relevant locations: the Malga reservoir, which regulates upstream flow, and the Mezzo station, a local meteorological gauge.

The analysis proceeds in three steps. First, Markov-switching autoregressive models identify high- and low-regime states for each hydrological variable. Second, regime-specific copula models capture how dependencies among water level, temperature, and conductivity change across joint states. Third, vine copula regressions link each hydrological variable to lagged precipitation inputs from Malga and Mezzo, providing a flexible representation of conditional dependence.

The results show that overall pairwise associations are modest, but regime-dependent patterns are important. Internal dependencies among height, temperature, and conductivity are strongest when water levels and temperatures are high but conductivity is low. Incorporating precipitation enhances the model fit, with vine regressions indicating that the Mezzo station, particularly at the two-hour lag, acts as a consistent external driver, whereas Malga precipitation shows a more limited, regime-dependent influence.

Overall, the study demonstrates that regime-switching vine copulas offer a powerful tool for uncovering subtle dependence structures in hydrological systems. They provide new insights into how local precipitation and managed storage interact with river dynamics, and offer a pathway for improved hydrological modeling and risk assessment under changing climatic conditions.

Contents

List of Figures

List of Tables

Introduction

Hydrological systems are inherently complex, governed by interactions among climatic forcing, river basin characteristics, and anthropogenic regulation. Their dynamics are nonlinear and often regime-dependent: the statistical properties of variables such as water level, temperature, and electrical conductivity may shift markedly between high- and low-flow conditions or between meteorological states. Capturing such structural breaks is a central challenge in hydrological time series analysis. Traditional approaches, which assume stationarity or rely on linear dependence, are often too restrictive to represent the intricate temporal and cross-variable dependencies found in real-world data.
This thesis applies such a framework to the *Noce River* in northern Italy, using a dataset of over 20,000 hourly observations collected between 2021 and 2023. The study focuses on three hydrological variables measured at the Ott station: water level (height), temperature, and electrical conductivity. These are analyzed in conjunction with precipitation inputs from two external sites: the Malga reservoir, which regulates upstream storage and flow, and the Mezzo station, which provides local meteorological measurements. By integrating both river dynamics and external rainfall drivers, the analysis seeks to uncover how precipitation and management practices affect hydrological variability across regimes.
The methodological contribution is threefold. First, Markov-switching autoregressive models are used to classify the time series into high- and low-regimes for each hydrological variable. Second, regime-specific vine copula models are estimated to characterize dependence structures among the variables themselves. Third, vine copula regression models are applied to link each hydrological response variable to precipitation covariates from Malga and Mezzo, yielding flexible regime-dependent conditional dependence estimates.
The results demonstrate that while pairwise dependencies are often weak in magnitude, they vary systematically across regimes. Internal associations between water level, temperature, and conductivity are strongest when water levels and temperatures are elevated but conductivity is low. Incorporating precipitation covariates improves model fit substantially, with precipitation from the Mezzo site emerging as the most consistent external driver, especially at short time lags, while Malga precipitation exerts more occasional and regime-specific effects. These findings illustrate the importance of combining regime-switching with flexible copula models to capture the nuanced dependence structures that govern hydrological dynamics.
Overall, this study contributes to hydrological time series modeling by demonstrating how copula-based regime-switching methods can provide richer insights into nonlinear dependence, regime-specific variability, and the role of external drivers. Beyond methodological value, these insights offer practical relevance for water resource management and risk assessment under increasingly variable climate conditions.

Part I

Mathematical Background

1 Model selection and evaluation

In this part, we provide a variety of potential models and use well-established information-theoretic criteria to choose the best specification. In particular, we compare model performance using the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). The maximized log-likelihood, which we define below, is the foundation of both criterion. The following definitions are based on the works of Georgii (2015), Fahrmeir et al. (2013), and Rossi (2018).

Definition (Log-likelihood function). *Let $f(x;\boldsymbol{\theta})$ be a probability model with parameter space Θ. For a sample $x_1,\dots,x_n$ the likelihood function for a model $\mathcal{M}$ is given by the joint probability density*

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

Under the assumption of independent observations, this becomes

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

where f_i is the density of X_i. The log-likelihood *is then*

$$\ell_{\mathcal{M}}(\boldsymbol{\theta}) = \log L_{\mathcal{M}}(\boldsymbol{\theta}).$$

Definition (Akaike's Information Criterion (AIC)). *For a model $\mathcal{M}$ with k parameters (where k denotes the total number of free parameters in the model) and maximized log-likelihood $\ell_{\mathcal{M}}(\widehat{\boldsymbol{\theta}})$, the Akaike Information Criterion is defined by*

$$\mathrm{AIC}_{\mathcal{M}} = -2\,\ell_{\mathcal{M}}(\widehat{\boldsymbol{\theta}}) \;+\; 2k.$$

Definition (Bayesian Information Criterion (BIC)). *With the same notation and sample size n, the Bayesian Information Criterion is defined by*

$$\mathrm{BIC}_{\mathcal{M}} = -2\,\ell_{\mathcal{M}}(\widehat{\boldsymbol{\theta}}) \;+\; \log(n)\,k,$$

where k is the number of model parameters.

2 Stochastic Process and Time Series

In this section, we present the essential building blocks for analyzing temporal data. We begin by defining a *stochastic process* and its realization as a *time series*, and then explore the *autocovariance* and *autocorrelation* functions that capture temporal dependence. The following definitions are based on the book by Shumway and Stoffer (2017).

Definition (Stochastic Process). *Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and T an index set (often "time," e.g. $\mathbb{Z}$ or $\mathbb{R}$). A* stochastic process *is a family of random variables*

$$\{\,X_t\colon \Omega \to \mathbb{R}\,\}_{t\in T}, \quad \textit{where } X_t(\omega) \;=\; X(\omega,t).$$

- *For each fixed $t \in T$, the map $\omega \mapsto X_t(\omega)$ is a random variable on $(\Omega,\mathcal{F},\mathbb{P})$.*
- *For each fixed $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is called a* sample function *or* realization.
- *The set of all sample functions is the* ensemble*; in time series work we typically observe just one realization $\{X_t(\omega_0)\}$.*

Definition (Time Series). *A* time series *is simply one realization of a stochastic process observed at discrete times* $t = 1, 2, 3, \ldots$. *In other words, for some fixed* $\omega_0 \in \Omega$ *we write*

$$x_t \;=\; X_t(\omega_0), \quad t = 1, 2, 3, \ldots$$

and so the observed sequence is

$$(x_1, x_2, x_3, \ldots).$$

Remark. *In practice we often impose additional structure on* $\{X_t\}$*, for example:*

- ***Stationarity:*** *for any* $k \in \mathbb{N}$ *and any* $t_1, \ldots, t_k$ *in the index set, the joint distribution of*

  $$\big(X_{t_1}, \ldots, X_{t_k}\big)$$

  depends only on the differences $(t_i - t_j)$*, not on the absolute times.*
- ***Ergodicity:*** *time-averages along a single realization converge (in probability) to ensemble averages.*
- ***Indexing:*** *although* T *can be continuous, most data arrive at equally-spaced points so one takes* $T = \mathbb{Z}$ *or* $\mathbb{N}$.

Definition (White noise). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* T *an index set. A* white noise *process is a family*

$$\{\, W_t \colon \Omega \to \mathbb{R} \,\}_{t \in T}$$

of uncorrelated random variables satisfying

$$\mathbb{E}[W_t] = 0, \quad \big(W_t\big) = \sigma_w^2 \in (0, \infty).$$

We write

$$W_t \sim \mathrm{WN}(0, \sigma_w^2).$$

If moreover the W_t *are independent and each* $W_t \sim \mathcal{N}(0, \sigma_w^2)$*, then* $\{W_t\}$ *is called* Gaussian white noise*, denoted*

$$W_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_w^2).$$

Remark. *White noise serves as the innovation term in many classical time-series models:*

- *An* autoregressive *model of order p, AR(p), takes the form*

  $$X_t \;=\; \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} \;+\; W_t,$$

  where $\{W_t\}$ *is white noise.*
- *A* moving-average *model of order q, MA(q), writes*

  $$X_t \;=\; W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}.$$

- *Combining both yields an ARMA*(p, q) *model, and including differencing yields ARIMA*(p, d, q)*.*

Time series analysis hinges on the fact that observations from a stochastic process are interdependent across time. We can capture this dependence either through explicit models or by measuring covariances and correlations. Because we observe an entire sequence rather than a single pair of points, we end up with not just one covariance or correlation but a whole series known respectively as the *autocovariance* and *autocorrelation* functions.

Definition (Autocovariance function). *Let* $\{X_t\}_{t\in\mathbb{Z}}$ *be a stochastic process with mean function* $\mu_t = \mathbb{E}[X_t]$*. The* autocovariance function *is*

$$\gamma(s,t) \;=\; \mathrm{Cov}(X_s, X_t) \;=\; \mathbb{E}\big[(X_s-\mu_s)(X_t-\mu_t)\big], \quad \forall\, s,t.$$

In particular, when $s = t$ *this reduces to the variance:*

$$\gamma(t,t) = \mathbb{E}\big[(X_t-\mu_t)^2\big] = \mathrm{Var}(X_t).$$

If $\gamma(s,t) = 0$*, then* X_s *and* X_t *are uncorrelated, though other forms of dependence may still exist.*

Definition (Autocorrelation Function). *Let* $\{X_t\}$ *be a stochastic process with autocovariance* $\gamma(s,t) = \mathrm{Cov}(X_s, X_t)$*. The* autocorrelation function *is*

$$\rho(s,t) \;=\; \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\,\gamma(t,t)}}, \qquad -1 \le \rho(s,t) \le 1.$$

It measures the linear association between X_s *and* X_t*, normalized to lie in* $[-1,1]$*.*

Definition (Partial Autocorrelation Function). *For a* stationary *process* $\{X_t\}$*, the* partial autocorrelation at lag h*, denoted* ϕ_{hh}*, is the correlation between* X_{t+h} *and* X_t *after removing the linear effects of the intervening observations* $\{X_{t+1},\dots,X_{t+h-1}\}$*. Equivalently:*

$$\phi_{11} = \mathrm{Corr}\big(X_{t+1}, X_t\big) = \rho(1), \quad \phi_{hh} = \mathrm{Corr}\big(X_{t+h} - \widehat{X}_{t+h},\, X_t - \widehat{X}_t\big).$$

where

$$\widehat{X}_{t+h} = \beta_1 X_{t+h-1} + \beta_2 X_{t+h-2} + \cdots + \beta_{h-1} X_{t+1}, \quad \widehat{X}_t = \beta_1 X_{t+1} + \beta_2 X_{t+2} + \cdots + \beta_{h-1} X_{t+h-1},$$

are the best linear predictors of X_{t+h} *and* X_t*, respectively, based on the* $h-1$ *intervening observations* $\{X_{t+1},\dots,X_{t+h-1}\}$*, with the coefficients* $\{\beta_j\}$ *chosen to minimize the mean-squared prediction error:*

$$\{\beta_j\} \;=\; \arg\min_{\gamma_1,\dots,\gamma_{h-1}} E\Big[\big(X_{t+h} - \sum_{j=1}^{h-1} \gamma_j\, X_{t+h-j}\big)^2\Big].$$

3 Time Series Decomposition

Real-world time series often exhibit systematic trends, seasonal patterns, or both. By isolating these deterministic components from the observed data, we can more clearly analyze the underlying random fluctuations.
In this section, we explain how to split a time series into its deterministic trend and seasonal components, allowing for the analysis of the remaining random fluctuations more effectively (Huang and Petukhina, 2022).

3.1 Deterministic Trend and Seasonal Components

A *deterministic trend* T_t in a series $\{X_t\}$ satisfies:

- It represents a gradual, systematic change over time.
- It can be modeled as a smooth, non-repeating function of time, $T(t)$.

The deterministic seasonal component S_t of a time series $\{X_t\}$ can be characterized by:

- Systematic, repeating fluctuations that recur at regular intervals.
- A constant, known period p often called the (minimum) seasonal period, over which the pattern repeats.
- Consistency in the shape, timing, and magnitude of the cycle from one period to the next.

This seasonal behavior, also referred to as seasonal variation or seasonal effect, typically originates from calendar-driven cycles, weather patterns, decision-making schedules, and related external factors.
After removing T_t and S_t, the remaining series $R_t = X_t - T_t - S_t$ consists of the irregular or noise component.

3.2 Additive vs. Multiplicative Decomposition

Two common frameworks for separating these components are:

1. **Additive model:**
$$X_t = T_t + S_t + R_t.$$
Use when the amplitude of seasonal swings does not depend on the overall level of the series.

2. **Multiplicative model:**
$$X_t = T_t \times S_t \times R_t.$$
Appropriate if seasonal fluctuations scale proportionally with the series level (and $X_t \neq 0$).

4 Hidden Markov Models (HMMs)

Hidden Markov models (HMMs) are models in which the distribution that generates an observation depends on the state of an underlying and unobserved Markov process. They provide flexible general-purpose models for univariate and multivariate time series, especially for discrete-valued series, including categorical series and series of counts. The following definitions are based on the book by Zucchini, MacDonald, and Langrock (2016).

4.1 Markov Chain and Markov Property

A sequence of discrete random variables

$$\{S_t\}_{t\in\mathbb{N}}, \quad S_t \in \{1,\dots,m\},$$

is called a *Markov chain* if it satisfies the *Markov property*, namely that the distribution of the next state depends only on the present state and not on the past history. In set-wise form, this reads: for every $t \in \mathbb{N}$, every $i \in \{1,\dots,m\}$, and all measurable subsets $B_1,\dots,B_{t-1},B \subseteq \{1,\dots,m\}$,

$$\Pr\big(S_{t+1} \in B \;\big|\; S_t = i,\; S_{t-1} \in B_{t-1},\, \dots,\, S_1 \in B_1\big) = \Pr\big(S_{t+1} \in B \;\big|\; S_t = i\big). \tag{4.1}$$

Equivalently, in point-wise form, for all $s_1,\dots,s_{t+1} \in \{1,\dots,m\}$,

$$\Pr\big(S_{t+1} = s_{t+1} \;\big|\; S_t = s_t,\, S_{t-1} = s_{t-1},\, \dots,\, S_1 = s_1\big) = \Pr\big(S_{t+1} = s_{t+1} \;\big|\; S_t = s_t\big). \tag{4.2}$$

Here, the statement (4.2) is commonly referred to as the *Markov property.*

Transition probabilities In a Markov chain on the state space $\{1,\dots,m\}$, the one-step transition probabilities are

$$P_{ij} \;=\; \Pr\big(S_{t+1} = j \mid S_t = i\big), \quad i,j \in \{1,\dots,m\}, \quad \sum_{j=1}^{m} P_{ij} = 1 \quad \forall\, i. \tag{4.3}$$

More generally, the t-step transition probability is defined by

$$P_{ij}^{(t)} := \Pr\big(S_{s+t} = j \mid S_s = i\big), \quad i,j \in \{1,\dots,m\},\; t \in \mathbb{N}. \tag{4.4}$$

If the transition probabilities

$$\Pr\big(S_{s+t} = j \mid S_s = i\big)$$

remain the same for all starting times s, the chain is called *homogeneous*; otherwise it is *non-homogeneous.* In what follows we assume homogeneity and therefore write

$$\gamma_{ij}(t) := \Pr\big(S_{s+t} = j \mid S_s = i\big), \quad i,j \in \{1,\dots,m\}, \tag{4.5}$$

where $\gamma_{ij}(t)$ makes no reference to s. Collecting all $\gamma_{ij}(t)$ into a $m\times m$ matrix gives the *transition matrix*

$$\boldsymbol{\Gamma}(t) = \big[\gamma_{ij}(t)\big]_{i,j=1}^{m} \;\in\; \mathbb{R}^{m\times m},$$

whose (i,j) entry is $\gamma_{ij}(t)$.

The *unconditional* probabilities of being in each state at time t, $\Pr(S_t = j)$, can be collected into the row vector

$$\mathbf{u}(t) = \big(\Pr(S_t = 1), \Pr(S_t = 2), \ldots, \Pr(S_t = m)\big), \quad t \in \mathbb{N}.$$

Here $\mathbf{u}(1)$ is the *initial distribution* of the chain. Since the chain is homogeneous, the distribution evolves by right-multiplying with the one-step transition matrix $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}(1)$:

$$\mathbf{u}(t+1) \;=\; \mathbf{u}(t)\,\boldsymbol{\Gamma}. \tag{4.6}$$

Equation (4.6) simply states that the probability of being in state j at time $t+1$ is the weighted sum

$$u_j(t+1) = \sum_{i=1}^{m} u_i(t)\,\Gamma_{ij}.$$

Stationary distributions A Markov chain with transition matrix $\boldsymbol{\Gamma}$ is said to have a *stationary distribution* $\boldsymbol{\delta}$ (a row vector with non-negative entries) if it satisfies

$$\boldsymbol{\delta}\,\boldsymbol{\Gamma} = \boldsymbol{\delta} \quad \text{and} \quad \sum_{i=1}^{m} \delta_i = 1.$$

The first of these requirements expresses stationarity, the second is the requirement that $\boldsymbol{\delta}$ is indeed a probability distribution.
Since

$$\mathbf{u}(t+1) = \mathbf{u}(t)\,\boldsymbol{\Gamma},$$

and if we choose $\mathbf{u}(1) = \boldsymbol{\delta}$ with $\boldsymbol{\delta}\,\boldsymbol{\Gamma} = \boldsymbol{\delta}$, it follows by induction that

$$\mathbf{u}(t) = \boldsymbol{\delta} \quad \text{for all } t.$$

so that the chain remains in the same distribution at every step. We call a Markov chain initialized in this way a *stationary Markov chain.*

An *irreducible* Markov chain has a unique, strictly positive, stationary distribution. For a derivation of this result, see Norris (1997, Theorem 1.7) or Karlin and Taylor (1981, Chapter 1).

4.2 Hidden Markov Models

Let

$$\{S_t\}_{t\in\mathbb{N}},\ S_t \in \{1,\ldots,m\}, \quad \{X_t\}_{t\in\mathbb{N}},\ X_t \in \mathcal{X},$$

where $\mathcal{X}$ is a finite observation alphabet. Write the history vectors

$$\mathbf{S}^{(t)} = (S_1,\ldots,S_t), \quad \mathbf{X}^{(t)} = (X_1,\ldots,X_t).$$

An HMM is specified by the following three conditional-independence relations:

1. **Initial distribution:**

$$\Pr(S_1 = i) = \delta_i, \quad i \in \{1,\ldots,m\}, \quad \sum_{i=1}^{m} \delta_i = 1. \tag{4.7}$$

2. **Hidden Markov property:**

$$\Pr\big(S_t = s_t \;\big|\; \mathbf{S}^{(t-1)} = \mathbf{s}^{(t-1)}\big) = \Pr\big(S_t = s_t \;\big|\; S_{t-1} = s_{t-1}\big), \qquad t = 2, 3, \dots, \quad s_i \in \{1, \dots, m\}. \tag{4.8}$$

   Here we recall that

$$\mathbf{S}^{(t-1)} = (S_1, S_2, \dots, S_{t-1}), \quad \mathbf{s}^{(t-1)} = (s_1, s_2, \dots, s_{t-1}).$$

3. **Conditional-independence of observations:**

$$\Pr\big(X_t = x_t \;\big|\; \mathbf{X}^{(t-1)} = \mathbf{x}^{(t-1)}, \mathbf{S}^{(t)} = \mathbf{s}^{(t)}\big) = \Pr\big(X_t = x_t \;\big|\; S_t = s_t\big), \qquad t \in \mathbb{N}. \tag{4.9}$$

In other words:

- The hidden-state process $\{S_t\}$ satisfies the Markov property.
- The observation process $\{X_t\}$ is state-dependent: given the entire hidden history only the current state S_t matters.

If $\{S_t\}$ has m states, we call $\{X_t\}$ an m-state HMM.

State-dependent distributions For discrete observations, define for each $i = 1, \dots, m$:

$$p_i(x) \;=\; \Pr(X_t = x \mid S_t = i), \quad x \in \mathcal{X}, \quad \sum_{x \in \mathcal{X}} p_i(x) = 1.$$

That is, p_i is the probability mass function of X_t given $S_t = i$.
For discrete-valued observations X_t, define

$$u_i(t) := \Pr(S_t = i), \quad i = 1, \dots, m, \; t \in \mathbb{N},$$

and collect these into the row vector $\mathbf{u}(t) = (u_1(t), \dots, u_m(t))$. Then the marginal distribution of X_t is

$$\Pr(X_t = x) = \sum_{i=1}^{m} \Pr(S_t = i)\, \Pr(X_t = x \mid S_t = i) = \sum_{i=1}^{m} u_i(t)\, p_i(x).$$

This expression can be written in matrix form as

$$\Pr(X_t = x) = \mathbf{u}(t) \begin{pmatrix} p_1(x) & & 0 \\ & \ddots & \\ 0 & & p_m(x) \end{pmatrix} \mathbf{1} = \mathbf{u}(t)\, \mathbf{P}(x)\, \mathbf{1}, \tag{4.10}$$

where $\mathbf{P}(x) = \operatorname{diag}\big(p_1(x), \dots, p_m(x)\big)$ and $\mathbf{1} = (1, \dots, 1)^\top$.
It follows from (4.6) that

$$\mathbf{u}(t) = \mathbf{u}(1)\, \mathbf{\Gamma}^{t-1}, \quad \text{hence} \quad \Pr(X_t = x) = \mathbf{u}(1)\, \mathbf{\Gamma}^{t-1}\, \mathbf{P}(x)\, \mathbf{1}. \tag{4.11}$$

Equation (4.11) holds for any homogeneous chain (not necessarily stationary). If the chain is stationary with $\mathbf{u}(1) = \boldsymbol{\delta}$ satisfying $\boldsymbol{\delta}\, \mathbf{\Gamma} = \boldsymbol{\delta}$, then

$$\Pr(X_t = x) = \boldsymbol{\delta}\, \mathbf{P}(x)\, \mathbf{1}. \tag{4.12}$$

4.3 Likelihood of an HMM

Here we consider the likelihood of a hidden Markov model (HMM) in general. Suppose we observe a sequence

$$x_1, x_2, \dots, x_T$$

generated by an m-state HMM with initial distribution $\boldsymbol{\delta}$, one-step transition matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{m\times m}$, and state-dependent probability mass functions $p_i(\cdot)$. We denote the likelihood of observing $(x_1, \dots, x_T)$ by

$$L_T \;=\; \Pr(X_1 = x_1,\, X_2 = x_2, \dots, X_T = x_T).$$

Proposition 1. *Let $x_1, \dots, x_T$ be the observed outputs of an m-state HMM with initial distribution $\boldsymbol{\delta}$, one step transition matrix $\boldsymbol{\Gamma} \in \mathbb{R}^{m\times m}$, and state-dependent PMFs $p_i(\cdot)$. Then the likelihood of observing the sequence $(x_1, \dots, x_T)$ is*

$$L_T = \Pr(X_1 = x_1, \dots, X_T = x_T) = \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\boldsymbol{\Gamma}\,\mathbf{P}(x_2)\,\cdots\,\boldsymbol{\Gamma}\,\mathbf{P}(x_T)\,\mathbf{1}. \tag{4.13}$$

If $\boldsymbol{\delta}$ is stationary (so $\boldsymbol{\delta}\,\boldsymbol{\Gamma} = \boldsymbol{\delta}$), then equivalently

$$L_T = \Pr(X_1 = x_1, \dots, X_T = x_T) = \boldsymbol{\delta}\,\boldsymbol{\Gamma}\,\mathbf{P}(x_1)\,\cdots\,\boldsymbol{\Gamma}\,\mathbf{P}(x_T)\,\mathbf{1}. \tag{4.14}$$

Proof. We give the argument for discrete observations only. For the continuous-observation case, see for example Cappé et al. (2005). First,

$$L_T = \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)}\big) = \sum_{s_1,\dots,s_T=1}^{m} \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)},\, \mathbf{S}^{(T)} = \mathbf{s}^{(T)}\big).$$

By the HMM assumptions (4.8) and (4.9), the joint probability factors as

$$\begin{aligned}\Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)},\, \mathbf{S}^{(T)} = \mathbf{s}^{(T)}\big) = \Pr(S_1 = s_1) \prod_{k=2}^{T} \Pr\big(S_k = s_k \mid S_{k-1} = s_{k-1}\big) \\ \times \prod_{k=1}^{T} \Pr\big(X_k = x_k \mid S_k = s_k\big).\end{aligned} \tag{4.15}$$

Hence

$$L_T = \sum_{s_1,\dots,s_T} \Big(\delta_{s_1}\,\gamma_{s_1,s_2}\,\cdots\,\gamma_{s_{T-1},s_T}\Big)\Big(p_{s_1}(x_1)\,\cdots\,p_{s_T}(x_T)\Big),$$

which by standard matrix-sum identities is exactly (4.13):

$$L_T = \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\boldsymbol{\Gamma}\,\mathbf{P}(x_2)\,\cdots\,\boldsymbol{\Gamma}\,\mathbf{P}(x_T)\,\mathbf{1}.$$

Finally, if $\boldsymbol{\delta}$ is the stationary distribution ($\boldsymbol{\delta}\,\boldsymbol{\Gamma} = \boldsymbol{\delta}$), then $\boldsymbol{\delta}\,\mathbf{P}(x_1) = \boldsymbol{\delta}\,\boldsymbol{\Gamma}\,\mathbf{P}(x_1)$, and one obtains (4.14). □

A very simple but crucial consequence of the matrix expression for the likelihood is the so-called *forward algorithm*, which allows one to compute the likelihood recursively. This recursive computation is fundamental not only for efficient likelihood evaluation and parameter estimation, but also for forecasting future observations, decoding the most likely hidden-state sequence, and performing model-checking diagnostics.

4.4 Forward Probabilities

To derive the forward algorithm for an m-state HMM, recall that $\boldsymbol{\delta} = (\delta_1, \dots, \delta_m)$ is the initial state–distribution row vector, $\boldsymbol{\Gamma} = (\gamma_{ij})$ is the $m \times m$ transition-probability matrix, and $\mathbf{P}(x_t) = \operatorname{diag}\big(p_1(x_t), \dots, p_m(x_t)\big)$ is the diagonal emission-probability matrix at time t.
We define the forward probability row vector $\boldsymbol{\alpha}_t$ by

$$\boldsymbol{\alpha}_t = (\alpha_t(1), \dots, \alpha_t(m)) = \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\boldsymbol{\Gamma}\,\mathbf{P}(x_2)\cdots\boldsymbol{\Gamma}\,\mathbf{P}(x_t), \quad t = 1, 2, \dots, T.$$

Equivalently,

$$\boldsymbol{\alpha}_t = \boldsymbol{\delta}\,\mathbf{P}(x_1)\prod_{s=2}^{t}\big(\boldsymbol{\Gamma}\,\mathbf{P}(x_s)\big). \tag{4.16}$$

It follows immediately that for $t = 1, 2, \dots, T-1$,

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{\alpha}_t\,\boldsymbol{\Gamma}\,\mathbf{P}(x_{t+1}),$$

or in scalar form,

$$\alpha_{t+1}(j) = \left(\sum_{i=1}^{m}\alpha_t(i)\,\gamma_{ij}\right)p_j(x_{t+1}), \quad j = 1, \dots, m.$$

Proposition 2. *For $t = 1, 2, \dots, T$ and $j = 1, 2, \dots, m$,*

$$\alpha_t(j) = \Pr\big(\mathbf{X}^{(t)} = \mathbf{x}^{(t)},\, S_t = j\big).$$

Proof. For $t = 1$, since $\boldsymbol{\alpha}_1 = \boldsymbol{\delta}\,\mathbf{P}(x_1)$,

$$\alpha_1(j) = \delta_j\,p_j(x_1) = \Pr(S_1 = j)\,\Pr\big(X_1 = x_1 \mid S_1 = j\big) = \Pr(X_1 = x_1,\, S_1 = j),$$

so the claim holds at $t = 1$.
Now assume it holds at time t. Then

$$\begin{aligned}
\alpha_{t+1}(j) &= \sum_{i=1}^{m}\alpha_t(i)\,\gamma_{ij}\,p_j(x_{t+1}) \\
&= \sum_{i=1}^{m}\Pr\big(\mathbf{X}^{(t)} = \mathbf{x}^{(t)},\, S_t = i\big)\ \Pr(S_{t+1} = j \mid S_t = i)\ \Pr(X_{t+1} = x_{t+1} \mid S_{t+1} = j) \\
&\quad \text{(by the induction hypothesis and the two HMM factorization properties (4.7),(4.8))} \\
&= \sum_{i=1}^{m}\Pr\big(\mathbf{X}^{(t+1)} = \mathbf{x}^{(t+1)},\, S_t = i,\, S_{t+1} = j\big) \\
&\quad \text{(law of total probability over } S_t\text{)} \\
&= \Pr\big(\mathbf{X}^{(t+1)} = \mathbf{x}^{(t+1)},\, S_{t+1} = j\big),
\end{aligned}$$

as required. □

4.5 Backward probabilities

Define the *backward probability vector* $\boldsymbol{\beta}_t$ by

$$\beta_t(i) := \Pr\big(X_{t+1} = x_{t+1}, \dots, X_T = x_T \mid S_t = i\big), \quad i = 1, \dots, m. \tag{4.17}$$

Conditioning on the next hidden state S_{t+1} and using the HMM factorization,

$$\beta_t(i) = \sum_{j=1}^{m} \Pr(S_{t+1} = j \mid S_t = i)\, \Pr(X_{t+1} = x_{t+1} \mid S_{t+1} = j) \\ \times \Pr(X_{t+2} = x_{t+2}, \ldots, X_T = x_T \mid S_{t+1} = j).$$

which simplifies to

$$\beta_t(i) = \sum_{j=1}^{m} \gamma_{ij}\, p_j(x_{t+1})\, \beta_{t+1}(j).$$

In matrix form, with $\mathbf{P}(x) = \operatorname{diag}(p_1(x), \ldots, p_m(x))$, this recursion is

$$\boldsymbol{\beta}_t = \boldsymbol{\Gamma}\,\mathbf{P}(x_{t+1})\,\boldsymbol{\beta}_{t+1}.$$

Properties of Forward and Backward Probabilities We now establish a result relating the forward and backward probabilities $\alpha_t(i)$ and $\beta_t(i)$ to the joint probability $\Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)},\, S_t = i\big)$. This will be used in the EM algorithm for HMMs.

For each $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, m$,

$$\alpha_t(i)\,\beta_t(i) \;=\; \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)},\, S_t = i\big). \tag{4.18}$$

Summing over i gives

$$\sum_{i=1}^{m} \alpha_t(i)\,\beta_t(i) = \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)}\big) = L_T. \tag{4.19}$$

Moreover, from the factorized form of the likelihood and the definitions $\boldsymbol{\alpha}_t = \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\boldsymbol{\Gamma} \cdots \mathbf{P}(x_t)$ and $\boldsymbol{\beta}_t = \boldsymbol{\Gamma}\,\mathbf{P}(x_{t+1}) \cdots \mathbf{P}(x_T)\,\mathbf{1}$, we see that

$$L_T = \boldsymbol{\alpha}_t\,\boldsymbol{\beta}_t \;\in\; \mathbb{R}. \tag{4.20}$$

In applying the EM algorithm to HMMs for state identification and parameter estimation, we will also need the following two properties.

Proposition 3. *Firstly, for* $t = 1, \ldots, T$,

$$\Pr\big(S_t = j \mid \mathbf{X}^{(T)} = \mathbf{x}^{(T)}\big) = \frac{\alpha_t(j)\,\beta_t(j)}{L_T};$$

and secondly, for $t = 2, \ldots, T$,

$$\Pr\big(S_{t-1} = j,\, S_t = k \mid \mathbf{X}^{(T)} = \mathbf{x}^{(T)}\big) = \frac{\alpha_{t-1}(j)\,\gamma_{jk}\,p_k(x_t)\,\beta_t(k)}{L_T}.$$

For the full proof see Zucchini et al. (2016, Appendix B).

4.6 Expectation–Maximization Algorithm for Parameter Estimation in HMMs

To handle the unobserved state trajectory

$$\mathbf{S}^{(T)} = (S_1, \ldots, S_T),$$

we introduce, for each $t = 1, \dots, T$, the binary indicator row vector in m dimension

$$\mathbf{u}(t) = \big(u_1(t), \dots, u_m(t)\big)^\top \in \{0,1\}^m, \quad u_j(t) = \begin{cases} 1, & S_t = j, \\ 0, & \text{otherwise,} \end{cases}$$

and, for each $t = 2, \dots, T$, the binary indicator $m \times m$ matrix

$$\mathbf{v}(t) = \big(v_{jk}(t)\big)_{j,k=1}^{m} \in \{0,1\}^{m \times m}, \quad v_{jk}(t) = \begin{cases} 1, & S_{t-1} = j \text{ and } S_t = k, \\ 0, & \text{otherwise.} \end{cases}$$

As described by Zucchini et al. (2016), the EM algorithm for estimating Hidden Markov Models proceeds as follows:

Algorithm 1 EM Algorithm for Hidden Markov Models

Require: Observations $\mathbf{x}^{(T)} = (x_1, \dots, x_T)$, number of states m, tolerance ε, maximum iterations $I_{\max}$

Output: Parameter estimates $\boldsymbol{\theta} = (\boldsymbol{\delta}, \boldsymbol{\Gamma}, \{\boldsymbol{\phi}_j\}_{j=1}^m)$

Initialize: choose starting values $\boldsymbol{\theta}^{(0)}$

for $i = 1, \dots, I_{\max}$ **do**

  E-step: compute for $t = 1, \dots, T$, $j, k = 1, \dots, m$:

$$\widehat{u}_j^{(i)}(t) = \frac{\alpha_t^{(i-1)}(j)\, \beta_t^{(i-1)}(j)}{L_T^{(i-1)}}, \quad \widehat{v}_{jk}^{(i)}(t) = \frac{\alpha_{t-1}^{(i-1)}(j)\, \gamma_{jk}^{(i-1)}\, p_k(x_t; \boldsymbol{\phi}_k^{(i-1)})\, \beta_t^{(i-1)}(k)}{L_T^{(i-1)}},$$

  where $L_T^{(i-1)} = \Pr(\mathbf{x}^{(T)}; \boldsymbol{\theta}^{(i-1)})$.

  M-step: update parameters

$$\delta_j^{(i)} = \widehat{u}_j^{(i)}(1), \quad \gamma_{jk}^{(i)} = \frac{\sum_{t=2}^{T} \widehat{v}_{jk}^{(i)}(t)}{\sum_{\ell=1}^{m} \sum_{t=2}^{T} \widehat{v}_{j\ell}^{(i)}(t)}, \quad \boldsymbol{\phi}_j^{(i)} = \arg\max_{\boldsymbol{\phi}} \sum_{t=1}^{T} \widehat{u}_j^{(i)}(t)\, \log p_j(x_t; \boldsymbol{\phi}).$$

  if $\|\boldsymbol{\theta}^{(i)} - \boldsymbol{\theta}^{(i-1)}\| < \varepsilon$ **then**

    break

  end if

end for

return $\boldsymbol{\theta}^{(i)}$

5 Copula

We present the idea of a copula in this section. Next, we examine important copula frameworks and constructs, such as vine copula models. Every definition adheres to the methodology presented in Czado (2019).

Definition (Copula)**.** *A multivariate distribution function C on the d-dimensional hypercube $[0,1]^d$ with uniformly distributed marginals is called a d-dimensional copula. If C is absolutely continuous, its copula density c is obtained by taking the d-th order partial derivative*

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

Any marginal distribution can be combined with a copula to create new multivariate distribution functions using the following theorem.

Definition (Sklar's Theorem)**.** *Let $\mathbf{X} = (X_1, \dots, X_d)$ be a d-dimensional random vector with joint distribution function F and marginals F_i, $i = 1, \dots, d$. Then there exists a copula C such that*

$$F(x_1, \dots, x_d) = C\big(F_1(x_1), \dots, F_d(x_d)\big).$$

If F is absolutely continuous with density f and marginals densities f_i, then the joint density factorises as

$$f(x_1, \dots, x_d) = c\big(F_1(x_1), \dots, F_d(x_d)\big)\; f_1(x_1) \cdots f_d(x_d),$$

where c is the copula density. Moreover, when all F_i are continuous the copula is unique, and one may recover it (and its density) by

$$C(u_1, \dots, u_d) = F\big(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\big),$$

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

Definition (Probability Integral Transform)**.** *Let $X \sim F$ be a continuous random variable and let x denote its realization. Then*

$$u := F(x)$$

is called the probability integral transform *(short PIT) at x.*

Remark (Distribution of the probability integral transform)**.** *Assume $X \sim F$. Then $U := F(X)$ is uniformly distributed on* $[0,1]$. *Indeed,*

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

where F^{-1} denotes the quantile function.

If one wishes to convert the provided data into copula data, the following definition is required. For instance, the probability integral transform can be used to convert original data into copula data (u-scale). It is then possible to transform this data to standard normal margins (z-scale). It is also possible to return the altered data to its original x-scale form. We now provide a more thorough definition of the scales discussed here.

Definition (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*
$$Z_i := \Phi^{-1}(U_i) \;=\; \Phi^{-1}\big(F_i(X_i)\big), \quad i = 1, 2,$$
with density
$$g(z_1, z_2) = c\big(\Phi(z_1), \Phi(z_2)\big)\, \phi(z_1)\, \phi(z_2).$$

We have so far used the multivariate distribution's copula and marginal distributions to express it. The bivariate conditional distribution will be obtained in the next step by mixing the marginals with a suitable copula.

Definition (Bivariate conditional distribution and density)**.** *The bivariate conditional distribution function and density are defined by*
$$\begin{aligned} F_{1|2}(x_1 \mid x_2) &= \frac{\partial}{\partial u_2} C_{12}\big(F_1(x_1), u_2\big)\bigg|_{u_2 = F_2(x_2)} \\ &= \frac{\partial}{\partial F_2(x_2)} C_{12}\big(F_1(x_1), F_2(x_2)\big), \\ f_{1|2}(x_1 \mid x_2) &= c_{12}\big(F_1(x_1), F_2(x_2)\big)\, f_2(x_2). \end{aligned}$$

We can also obtain the conditional distribution functions of a bivariate copula distribution. These are called **h-functions** and are given by
$$h_{1|2}(u_1 \mid u_2) := \frac{\partial}{\partial u_2} C_{12}(u_1, u_2), \qquad h_{2|1}(u_2 \mid u_1) := \frac{\partial}{\partial u_1} C_{12}(u_1, u_2).$$

Moreover, we introduce the **tail dependence coefficients**, which measure the strength of extreme co-movements. For a bivariate random vector with copula C, the **upper tail dependence coefficient** is defined as
$$\lambda^{\text{upper}} = \lim_{t \to 1^-} \Pr\big(X_2 > F_2^{-1}(t) \mid X_1 > F_1^{-1}(t)\big) = \lim_{t \to 1^-} \frac{1 - 2t + C(t,t)}{1 - t},$$
and the **lower tail dependence coefficient** as
$$\lambda^{\text{lower}} = \lim_{t \to 0^+} \Pr\big(X_2 \le F_2^{-1}(t) \mid X_1 \le F_1^{-1}(t)\big) = \lim_{t \to 0^+} \frac{C(t,t)}{t}.$$

The set of **bivariate Archimedean copulas** defined in Czado (2019) represents a family of parametric copulas sharing the common form
$$C(u_1, u_2) = \varphi^{[-1]}\big(\varphi(u_1) + \varphi(u_2)\big),$$
where φ is a continuous, strictly decreasing, convex function satisfying $\varphi(1) = 0$, and its generalized inverse is given by
$$\varphi^{[-1]}(t) := \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0), \\ 0, & \varphi(0) \le t < \infty. \end{cases}$$

Also, some rotations of the copula are usually considered to accommodate larger ranges of dependence:

Family	Copula	Parameter	Tail dependence
Clayton	$\big(u_1^{-\delta} + u_2^{-\delta} - 1\big)^{-1/\delta}$	$0 < \delta < \infty$	Lower
Gumbel	$\exp\Big(-\big[(-\ln u_1)^{\delta} + (-\ln u_2)^{\delta}\big]^{1/\delta}\Big)$	$\delta \geq 1$	Upper
Frank	$-\dfrac{1}{\delta}\ln\Big(1 + \dfrac{(e^{-\delta u_1} - 1)(e^{-\delta u_2} - 1)}{e^{-\delta} - 1}\Big)$	$\delta \in \mathbb{R} \setminus \{0\}$	None
Joe	$1 - \Big[(1-u_1)^{\delta} + (1-u_2)^{\delta} - (1-u_1)^{\delta}(1-u_2)^{\delta}\Big]^{1/\delta}$	$\delta \geq 1$	Upper

Bivariate Archimedean copulas with single parameter

- 270°: $C_{270}(u_1, u_2) := C(1 - u_2,\, u_1)$
- 180°: $C_{180}(u_1, u_2) := C(1 - u_1,\, 1 - u_2)$
- 90°: $C_{90}(u_1, u_2) := C(u_2,\, 1 - u_1)$

5.1 Vine copula

The vine copula decomposes a multivariate density into products of conditional densities. Let $(X_1, \ldots, X_d)$ be a random vector with joint distribution $F_{1,\ldots,d}$ and density $f_{1,\ldots,d}$. Then one has

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

For the trivariate case $d = 3$, let X_1, X_2, X_3 have joint density $f_{1,2,3}$. We may write

$$
f_{1,2,3}(x_1, x_2, x_3) = f_{3|1,2}(x_3 \mid x_1, x_2)\; f_{2|1}(x_2 \mid x_1)\; f_1(x_1).
$$

To determine the conditional component $f_{3|1,2}(x_3 \mid x_1, x_2)$, we first form the bivariate conditional density of (X_1, X_3) given $X_2 = x_2$:

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

Hence the conditional density of X_3 given $(X_1, X_2) = (x_1, x_2)$ is

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

The remaining bivariate conditionals are

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

Substituting all components into the trivariate decomposition yields the full joint density

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

After examining the case $d = 3$, we present three classes of vine copula decompositions: regular, canonical, and drawable. We first define the regular vine (R-vine) distribution and introduce the necessary graph-theoretic concepts before presenting regular vine copulas. Finally, we give the definition of the normal vine copula.

Definition (Graph theoretic background)**.** *Here we look at some graph theory basics.*

1. A graph *is a pair* $G = (N, E)$ *of sets such that*

$$E \subseteq \big\{\{x, y\} : x, y \in N\big\}.$$

2. *Elements of* E *are referred to as edges of the graph* G*, while those of* N *are called nodes.*

3. *The number of neighbors of a node* $v \in N$ *is its* degree, *denoted* $d(v)$*.*

4. *A graph* $P = (N, E)$ *is called a* path *if*

$$N = \{v_0, v_1, \ldots, v_k\}, \quad E = \big\{\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{k-1}, v_k\}\big\}.$$

5. *A* cycle *is a path with* $v_0 = v_k$*.*

6. *A graph* $T = (N, E)$ *is a* tree *if any two nodes of* T *are connected by a unique path in* T*. Equivalently, every node has at least one connection to another node and there are no cycles.*

From these concepts, we define a regular vine (R-vine) tree sequence as follows.

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

1. *For each* $j = 1, \ldots, d-1$*, the graph* $T_j = (N_j, E_j)$ *is a connected tree.*

2. T_1 *has node set* $N_1 = \{1, \ldots, d\}$ *and edge set* E_1*.*

3. *For* $j \geq 2$*,* T_j *has node set* $N_j = E_{j-1}$ *and edge set* E_j*.*

4. *Proximity condition: For* $j = 2, \ldots, d-1$ *and every edge* $\{a, b\} \in E_j$*, it holds that* $|a \cap b| = 1$*.*

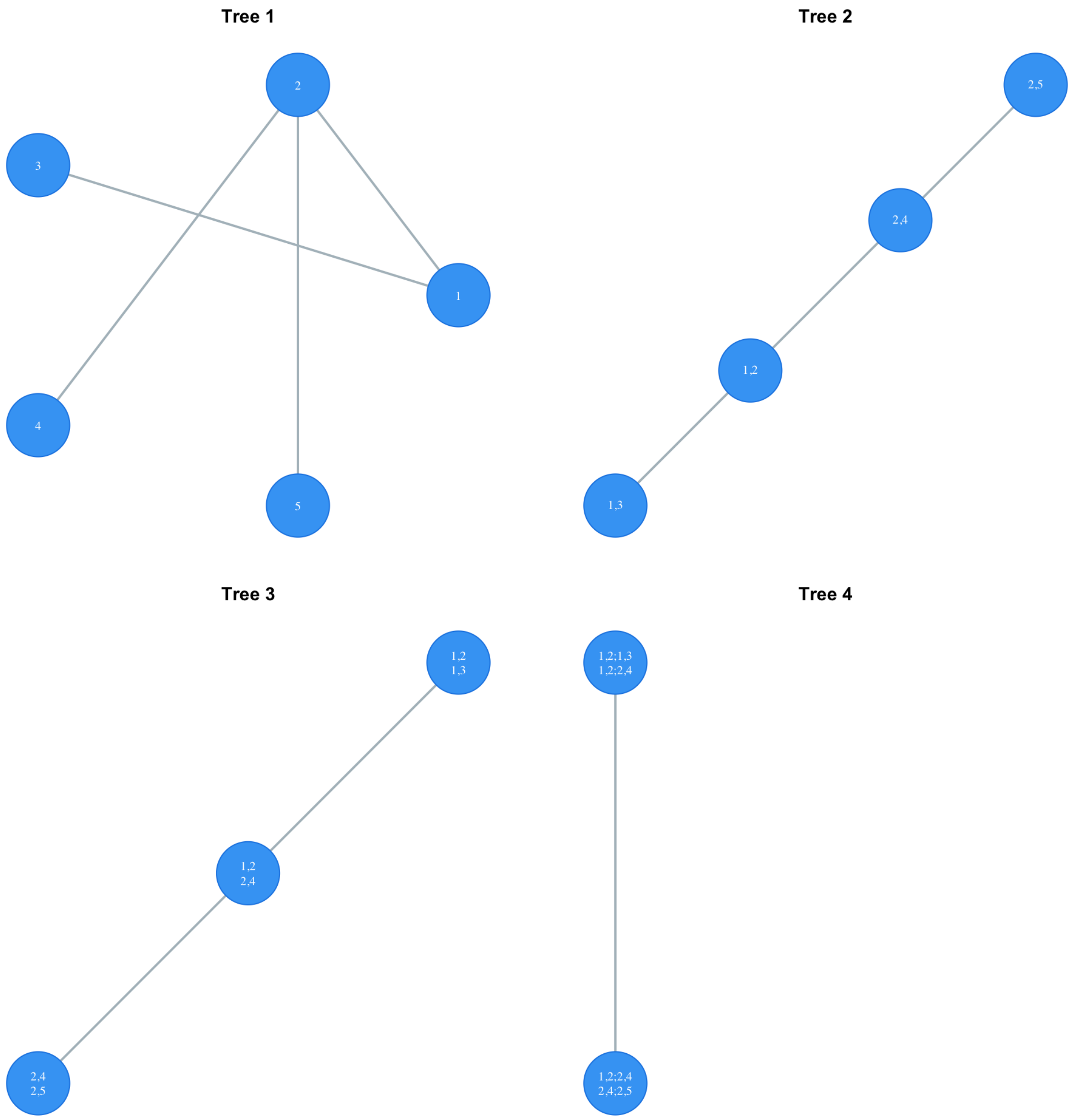


Figure 5.1: Example of a R-vine tree sequence with five elements

Example (R-vine tree sequence)**.** *Let* $d = 5$*. An R-vine tree sequence* $V = (T_1, T_2, T_3, T_4)$ *can be constructed as follows:*

- $T_1 = (N_1, E_1)$ *with*

$$N_1 = \{1, 2, 3, 4, 5\}, \quad E_1 = \{\{1, 2\}, \{1, 3\}, \{2, 4\}, \{2, 5\}\}.$$

  (This is neither a simple path nor a star.)

- $T_2 = (N_2, E_2)$ *with*

$$N_2 = E_1 = \{\{1, 2\}, \{1, 3\}, \{2, 4\}, \{2, 5\}\},$$

  and choose edges linking pairs that share a common node:

$$E_2 = \{\{\{1, 2\}, \{1, 3\}\}, \{\{1, 2\}, \{2, 4\}\}, \{\{2, 4\}, \{2, 5\}\}\}.$$

- $T_3 = (N_3, E_3)$ *with*

$$N_3 = E_2 = \{\{\{1, 2\}, \{1, 3\}\}, \{\{1, 2\}, \{2, 4\}\}, \{\{2, 4\}, \{2, 5\}\}\},$$

and link those that share a common edge from T_2:

$$E_3 = \big\{\{\{\{1,2\},\{1,3\}\},\ \{\{1,2\},\{2,4\}\}\},\ \{\{\{1,2\},\{2,4\}\},\ \{\{2,4\},\{2,5\}\}\}\big\}.$$

- $T_4 = (N_4, E_4)$ *with*

$$N_4 = E_3 = \{\{\{\{1,2\},\{1,3\}\},\{\{1,2\},\{2,4\}\}\},\ \{\{\{1,2\},\{2,4\}\},\{\{2,4\},\{2,5\}\}\}\},$$

and the single edge

$$E_4 = \Big\{\{\{\{\{1,2\},\{1,3\}\},\{\{1,2\},\{2,4\}\}\},\ \{\{\{1,2\},\{2,4\}\},\{\{2,4\},\{2,5\}\}\}\}\Big\}.$$

In each tree T_j, the nodes are the edges of T_{j-1}, and edges in T_j connect exactly those pairs of nodes whose corresponding edges in T_{j-1} share a common element.

Definition (Complete union and conditioned sets)**.** *Let $E_1,\dots,E_d$ be the edge sets of the trees in an R-vine, and let $N_1 = \{1,\dots,d\}$ be the set of all nodes in the first tree. For any edge $e \in E_i$ with $i \geq 2$, define its* complete union

$$A_e \;=\; \big\{\, j \in N_1 : \exists\, e_1 \in E_1,\dots,e_{i-1} \in E_{i-1} \text{ such that } j \in e_1 \in e_2 \in \cdots \in e_{i-1} \in e \big\}.$$

If $e = \{a,b\}$, its conditioning set *is*

$$D_e \;:=\; A_a \,\cap\, A_b,$$

and the two conditioned sets *are*

$$C_{e,a} \;:=\; A_a \setminus D_e, \qquad C_{e,b} \;:=\; A_b \setminus D_e.$$

We then write

$$C_e \;=\; C_{e,a} \cup C_{e,b},$$

and represent the edge e in the vine by the triple

$$e = \big(C_{e,a},\, C_{e,b};\, D_e\big).$$

Definition (Regular vine distribution)**.** *Let $\mathbf{X} = (X_1,\dots,X_d)$ be a d-dimensional random vector. A distribution F for $\mathbf{X}$ is called a* regular vine (R-vine) distribution *if there exists a triplet $(\mathcal{F},\, V,\, \mathcal{B})$ satisfying:*

1. *$\mathcal{F} = (F_1,\dots,F_d)$ is the collection of marginal distribution functions of $X_1,\dots,X_d$.*
2. *V is an R-vine tree sequence on the set of indices $\{1,\dots,d\}$.*
3. *$\mathcal{B} = \{\, C_e : e \in E_i,\ i = 1,\dots,d-1\}$ assigns to each edge e in tree T_i of the vine a bivariate (symmetric) copula C_e.*
4. *For each edge $e = \{a,b\} \in E_i$ with conditioning set D_e, the copula $C_{C_{e,a},C_{e,b};D_e}$ corresponds to the conditional distribution of $(X_{C_{e,a}}, X_{C_{e,b}})$ given $X_{D_e} = x_{D_e}$.*

Under these specifications, the joint density of $\mathbf{X}$ admits the factorization

$$f(x_1,\dots,x_d) = \prod_{i=1}^{d} f_i(x_i) \;\times \prod_{i=1}^{d-1} \prod_{e\in E_i} c_e\big(F_{C_{e,a}|D_e}(x_{C_{e,a}} \mid x_{D_e}),\ F_{C_{e,b}|D_e}(x_{C_{e,b}} \mid x_{D_e})\big),$$

where $e \in E_i$, $i = 1,\dots,d-1$ 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

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

Definition (Regular vine copula)**.** *A* regular vine (R-vine) copula *is an R-vine distribution on $[0,1]^d$ whose marginal distributions are all Uniform$[0,1]$.*

Now we consider two important subclasses of R-vines, namely the *drawable vine* (D-vine) and the *canonical vine* (C-vine). First, we examine their respective tree sequences; next, we present the form of their joint densities; and finally, we illustrate each with concrete examples.

Definition (D-vine tree sequence)**.** *A regular vine tree sequence $V = (T_1, \ldots, T_{d-1})$ is called a* drawable vine *(or D-vine) if, in each tree $T_j = (N_j, E_j)$ and for every node $n \in N_j$,*

$$\left|\{\, e \in E_j : n \in e\}\right| \leq 2.$$

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

$$f_{1,\ldots,d}(x_1, \ldots, x_d) = \left[\prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,(i+j);\,(i+1),\ldots,(i+j-1)}\right] \cdot \left[\prod_{k=1}^{d} f_k(x_k)\right]. \tag{5.1}$$

where the joint density $f_{1,\ldots,d}$ is decomposed into products of the marginal densities f_k and the pair-copula densities

$$c_{i,(i+j);D}\big(\cdot, \cdot; x_D\big),$$

with $D = \{i+1, \ldots, i+j-1\} \subset \{1, \ldots, d\}$.

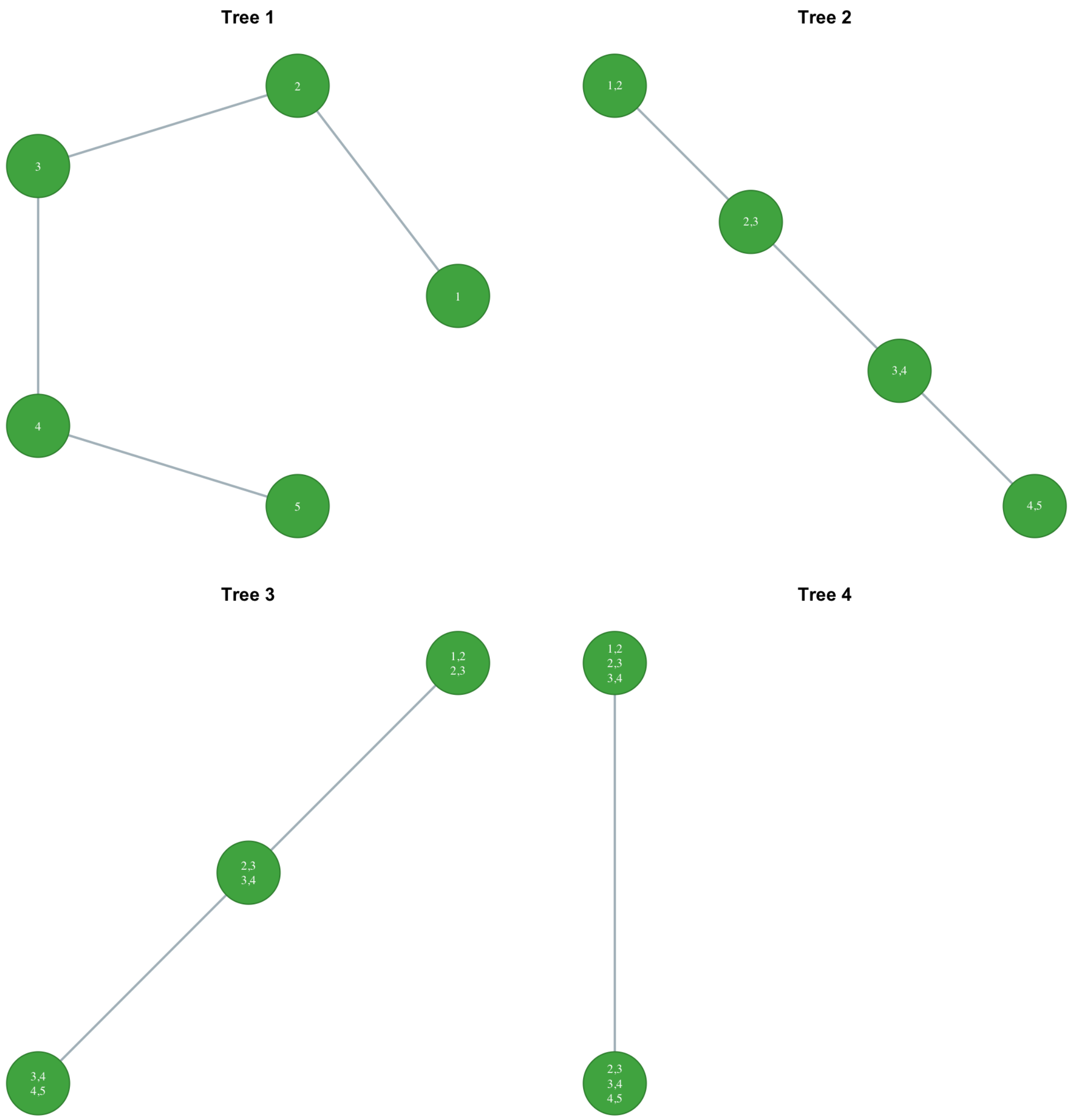


Figure 5.2: Example of a D-vine tree sequence with five elements

Definition (C-vine tree sequence)**.** *A regular vine tree sequence* $V = (T_1, \ldots, T_{d-1})$ *is called a* canonical vine *(or C-vine) if, in each tree* $T_i = (N_i, E_i)$*, there exists a unique* root node $r_i \in N_i$ *that is incident to every edge in* E_i*; that is,*

$$\left|\{\, e \in E_i \mid r_i \in e \}\right| = d - i.$$

Definition (Canonical vine (C-vine) density)**.** *The joint density* $f_{1,\ldots,d}(x_1, \ldots, x_d)$ *admits the C-vine decomposition*

$$f_{1,\ldots,d}(x_1, \ldots, x_d) = \left[\prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,(j+i)\,;,1,\ldots,(j-1)}\right] \cdot \left[\prod_{k=1}^{d} f_k(x_k)\right], \tag{5.2}$$

and we call this a canonical vine distribution*.*

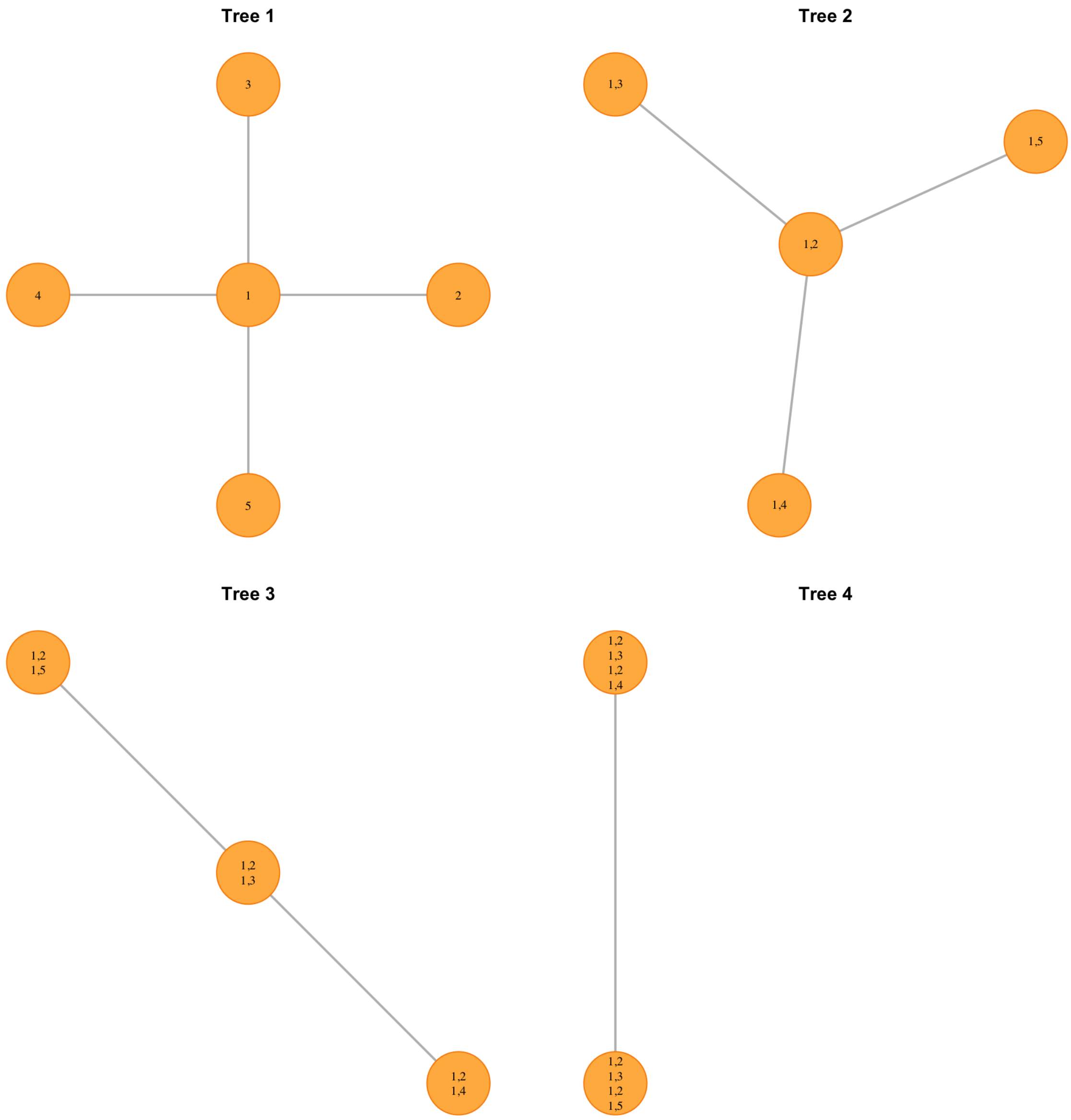


Figure 5.3: Example of a C-vine tree sequence with five elements

Definition (Simplifying assumption for C- and D-vines)**.** *A C- or D-vine satisfies the simplifying assumption if, for every pair of indices* i, j *and every conditioning set* $D \subset \{1, \ldots, d\}$*,*

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

for all x_D. When this holds, the corresponding C- or D-vine density reduces to the simplified form used in (4.1) and (4.2).

5.2 Simulating from vine copula

Here we present the general procedure to simulate from an R-vine copula. The approach for C- and D-vines is analogous (see Czado (2019)). The idea is to draw $u_1, \ldots, u_d$ recursively from the univariate conditional copulas implied by the joint R-vine. In particular:

1. Sample $w_j \sim U(0,1)$ independently for $j = 1, \ldots, d$.
2. Set
$$u_1 = w_1.$$
3. For $j = 2, \ldots, d$, compute
$$u_j = C^{-1}_{j|1,\ldots,j-1}\big(w_j \mid u_1, \ldots, u_{j-1}\big).$$

In expanded form, one writes

$$\begin{aligned}
u_1 &= w_1, \\
u_2 &= C^{-1}_{2|1}(w_2 \mid u_1), \\
&\ \vdots \\
u_d &= C^{-1}_{d|1,\ldots,d-1}(w_d \mid u_1, \ldots, u_{d-1}).
\end{aligned}$$

The univariate conditional distributions $C_{e_a|e_b;D_e}$ are obtained by differentiating the corresponding pair-copulas. For an edge $e = (e_a, e_b; D_e)$,

$$\begin{aligned}
C_{e_a|e_b;D_e}(w_1 \mid w_2; \theta_{e_a,e_b;D_e}) &= \frac{\partial}{\partial w_2} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}), \\
C_{e_b|e_a;D_e}(w_2 \mid w_1; \theta_{e_a,e_b;D_e}) &= \frac{\partial}{\partial w_1} C_{e_a,e_b;D_e}(w_1, w_2; \theta_{e_a,e_b;D_e}).
\end{aligned}$$

5.3 Selection and Parameter Estimation of Vine Copulas

In this section we describe a stage-wise procedure, based on the Dißmann et al. (2013) algorithm, to choose and fit a regular vine model. The approach proceeds tree by tree:

1. **Compute pairwise weights.** For every candidate edge $\{i, j\}$ with $1 \leq i < j \leq d$, calculate a measure of dependence $W_{i,j}$. These weights must also respect the proximity condition for subsequent trees.
2. **Build the maximum spanning tree.** Using the weights $W_{i,j}$, find the maximum spanning tree on the complete graph of nodes $\{1, \ldots, d\}$ and estimate the corresponding parameters $\boldsymbol{\theta}_{a_e,b_e;\,D_e}$ for the selected copula,
$$C_{a_e,b_e;\,D_e}\,.$$
3. **Generate pseudo-observations.** Recall that for a bivariate copula $C_{a_e,b_e;D_e}$ with true parameter vector $\boldsymbol{\theta}_{a_e,b_e;D_e}$, the two conditional distribution functions are
$$C_{a_e|b_e}\big(u \mid v;\, \boldsymbol{\theta}_{a_e,b_e;D_e}\big) = \frac{\partial}{\partial v}\, C_{a_e,b_e;D_e}\big(u, v;\, \boldsymbol{\theta}_{a_e,b_e;D_e}\big),$$
$$C_{b_e|a_e}\big(v \mid u;\, \boldsymbol{\theta}_{a_e,b_e;D_e}\big) = \frac{\partial}{\partial u}\, C_{a_e,b_e;D_e}\big(u, v;\, \boldsymbol{\theta}_{a_e,b_e;D_e}\big).$$

Then, for each fitted copula $C_{a_e,b_e;D_e}$ with estimate $\widehat{\boldsymbol{\theta}}_{a_e,b_e;D_e}$, transform the original data $(x_1,\ldots,x_T)$ into pseudo-observations

$$u_{k|b_e} = C_{a_e|b_e}\big(u_{k,b_e} \mid u_{k,a_e};\, \widehat{\boldsymbol{\theta}}_{a_e,b_e;D_e}\big), \quad u_{k|a_e} = C_{b_e|a_e}\big(u_{k,a_e} \mid u_{k,b_e};\, \widehat{\boldsymbol{\theta}}_{a_e,b_e;D_e}\big), \quad k=1,\ldots,n.$$

4. **Proceed to the next tree.** Use the pseudo-observations just generated to compute new pairwise weights and repeat steps 1–3 for tree T_2, and so on, until all $d-1$ trees are fitted.

To estimate the parameters $\boldsymbol{\theta}_e$ for edge

$$e = (a_e,\, b_e;\, D_e)$$

in tree T_i is based on the pseudo-observations. In particular, $\boldsymbol{\theta}_e$ is estimated by maximising

$$\widehat{\boldsymbol{\theta}}_e \;=\; \arg\max_{\boldsymbol{\theta}_e} \prod_{k=1}^{n} c_{a_e,b_e;D_e}\Big(u_{k,a_e|D_e},\, u_{k,b_e|D_e}\;;\; \boldsymbol{\theta}_e\Big)\,, \tag{5.3}$$

To compare different R-vine models one typically uses penalized log-likelihood criteria, namely the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). For an R-vine copula$(\mathcal{V},\mathcal{B})$ with fitted parameter vector $\boldsymbol{\theta}$ and n pseudo-observations $\mathbf{u}_1,\ldots,\mathbf{u}_n$, these are defined as

$$\mathrm{AIC}_{\mathrm{RV}} = -2\sum_{k=1}^{n} \ln\big(\ell_k(\widehat{\boldsymbol{\theta}};\ \mathbf{u}_k)\big) \;+\; 2K \qquad \mathrm{BIC}_{\mathrm{RV}} = -2\sum_{k=1}^{n} \ln\big(\ell_k(\widehat{\boldsymbol{\theta}};\ \mathbf{u}_k)\big) \;+\; \ln(n)\,K$$

where K is the number of model parameters and $\ell_k(\widehat{\boldsymbol{\theta}};\mathbf{u}_k)$ is the corresponding likelihood of the R-vine copula,

$$\ell_k(\boldsymbol{\theta};\mathbf{u}_k) := \prod_{i=1}^{d-1} \prod_{e\in E_i} c_{a_e,b_e;D_e}\Big(C_{a_e|D_e}\big(u_{k,a_e} \mid \boldsymbol{u}_{k,D_e}\big),\, C_{b_e|D_e}\big(u_{k,b_e} \mid \boldsymbol{u}_{k,D_e}\big)\Big).$$

5.4 Zero-inflated margins in vine copula models

The following development is taken from Funk et al. (2025). In many environmental applications (e.g. precipitation), margins exhibit a point mass at zero, which standard vine-copula methods cannot accommodate directly. Funk et al. introduce a mixture-margin and randomized transform as follows:

Mixture-margin model For each zero-inflated margin X_j, write

$$F_j(x) = p_{j,0}\, I\{x=0\} + (1-p_{j,0})\, F_{j,c}(x)\, I\{x>0\}, \quad p_{j,0} = \Pr(X_j = 0), \tag{5.4}$$

where $F_{j,c}(x) = \Pr(X_j \le x \mid X_j > 0)$ is a continuous CDF on $(0,\infty)$.

Randomized pseudo-observations Define for each observation $X_{j,k}$, where

$$j = 1,\ldots,d \quad \text{and} \quad k = 1,\ldots,n.$$

$$U_{j,k} = \begin{cases} V_{j,k}\, p_{j,0}, & X_{j,k} = 0, \\ p_{j,0} + (1-p_{j,0})\, F_{j,c}(X_{j,k}), & X_{j,k} > 0, \end{cases} \quad V_{j,k} \sim \mathrm{Uniform}(0,1). \tag{5.5}$$

Zeros are thus mapped uniformly into $[0, p_{j,0}]$, so that

$$U_{j,k} \sim \mathrm{Uniform}(0,1) \quad \text{and} \quad u_{j,k} \text{ denotes the observed realization of } U_{j,k}.$$

Zero-inflated vine density Let $\boldsymbol{u}_k = (u_{1,k}, \ldots, u_{d,k})$. Then the joint density is

$$\begin{aligned} f(x_1, \ldots, x_d) = & \prod_{j=1}^{d} \big[p_{j,0}\, \mathbf{1}\{x_j = 0\} + (1 - p_{j,0})\, f_{j,c}(x_j)\, \mathbf{1}\{x_j > 0\} \big] \\ & \times \prod_{i=1}^{d-1} \prod_{e \in E_i} c_e\big(u_{C_{e,a}|D_e}(x),\, u_{C_{e,b}|D_e}(x); \theta_e\big). \end{aligned} \tag{5.6}$$

where $c_e(\,\cdot, \cdot\,; \boldsymbol{\theta}_e)$ is the pair-copula density on edge e, evaluated at the realized conditional pseudo-observations $u_{C_{e,a}|D_e,k}$ and $u_{C_{e,b}|D_e,k}$

Part II

Noce Data Analysis

6 Data description

The Noce dataset comprises hydrological data collected at two distinct monitoring stations: "ott", the primary gauging station, and "mz", the gauging and meteorological stations located at Mezzolombardo. The dataset contains $20,569$ hourly observations from August 23, 2021, through December 28, 2023, structured into 7 variables: timestamps (`date`), water levels (`Noce_mz_h(m)`, `Adige_sl_h(m)`), `Noce_ott_h(m)`), temperature (`Noce_ott_t(°C)`) and electrical conductivity (`Noce_ott_ec(mS/cm)`). Data recorded at the "ott" station represents hourly point measurements, whereas data from the "mz" station has been aggregated into hourly means from observations taken every 15 minutes.

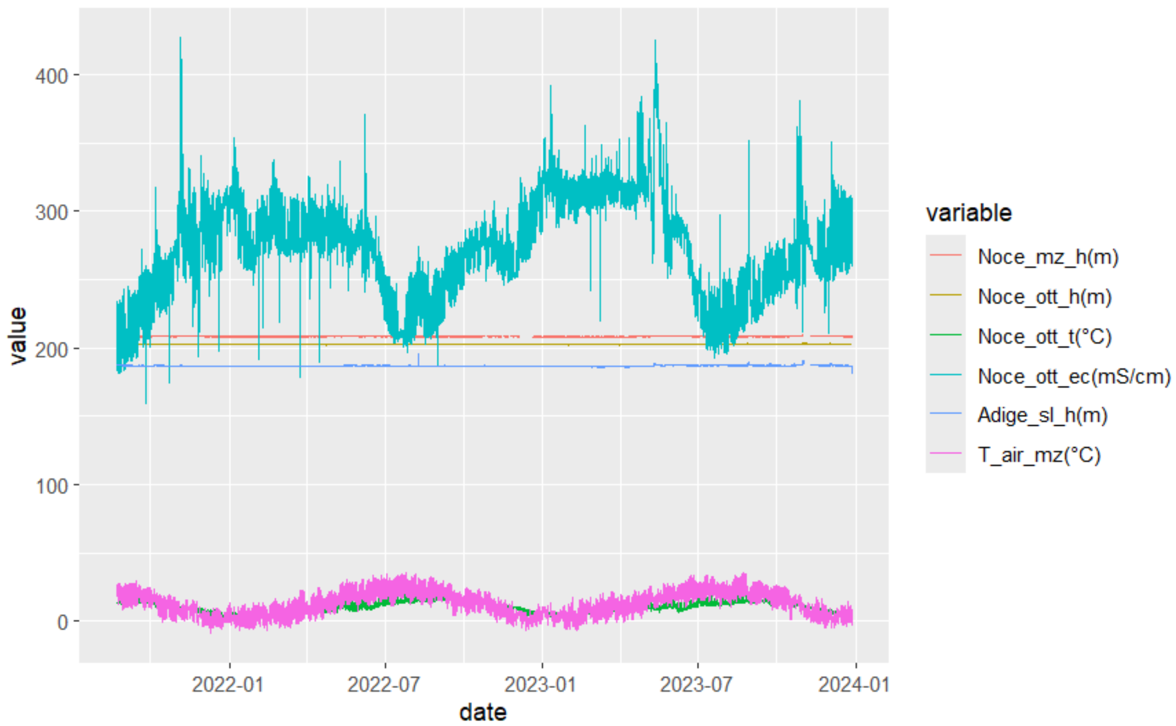


Figure 6.1: Time series plot of variable measurements from the Noce dataset.

7 Decompose selected variables

Here we analyse only ott variables (ott.h, ott.t, and ott.ec). We extract the timestamps along with the measurements of water level, temperature, and electrical conductivity and organize them into individual time series. Each series is subsequently cleaned by removing missing values and decomposed into key components: a long-term trend, seasonal pattern(s), and a residual component that reflects random fluctuations. Also, we set the decomposition frequency to 24 and 168 to capture the inherent daily and weekly cycles in our observations. We first transform each cleaned `ott` measurement series into a multiple-seasonal time series that captures both daily and weekly cycles and remove any missing entries, then apply a decomposition that enforces stable periodic patterns while smoothing the long-term trend over approximately one week. Finally, we isolate the residual component to obtain de-seasonalized data by using the `mstl()` function from the forecast package.

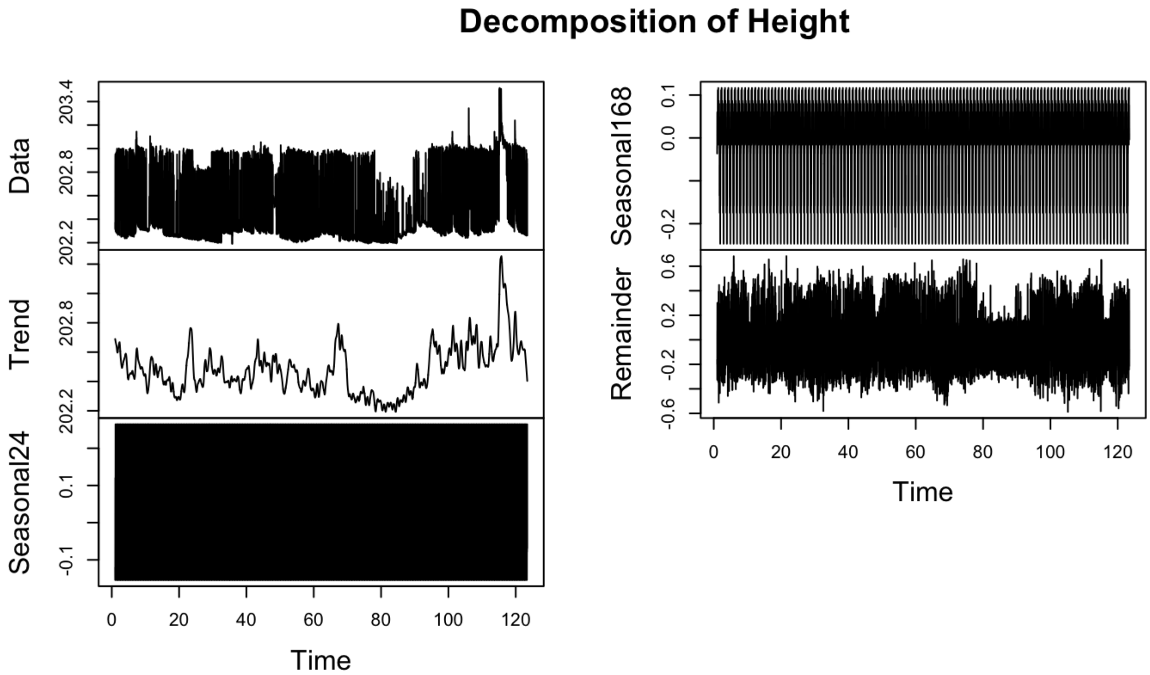


Figure 7.1: Decomposition plot of the water level measurement (`ott.h`), treated as an `msts` series, using the `mstl()` function from the forecast package.

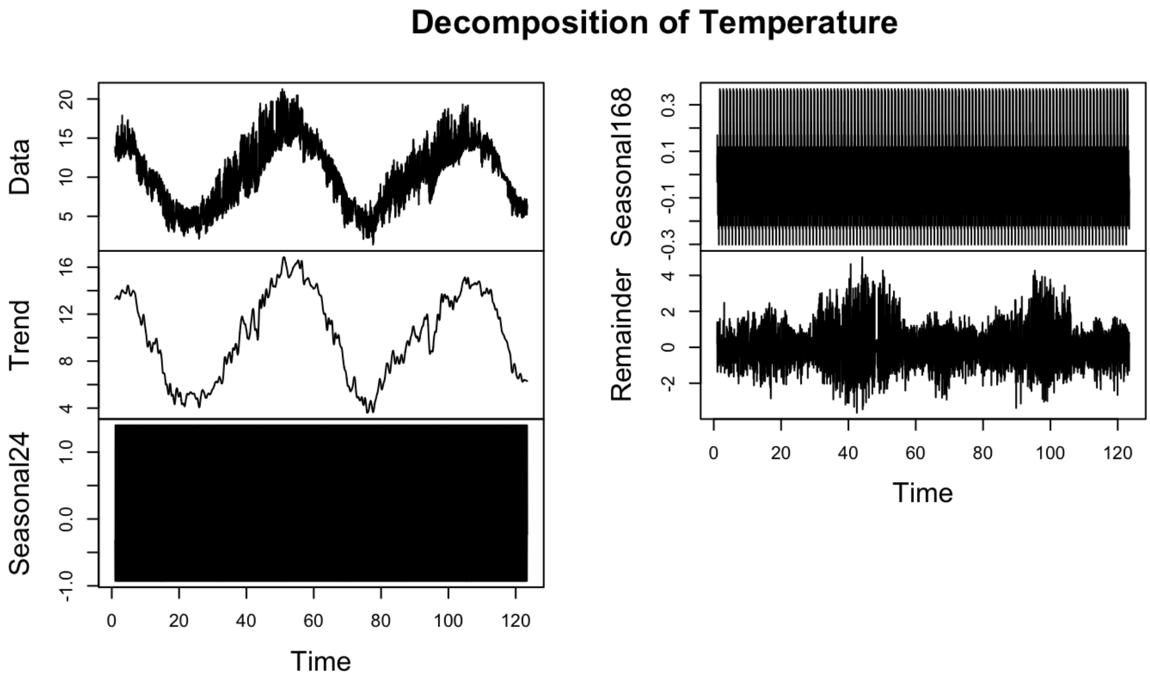


Figure 7.2: Decomposition plot of temperature (`ott.t`), treated as an `msts` series, using the `mstl()` function from the forecast package.

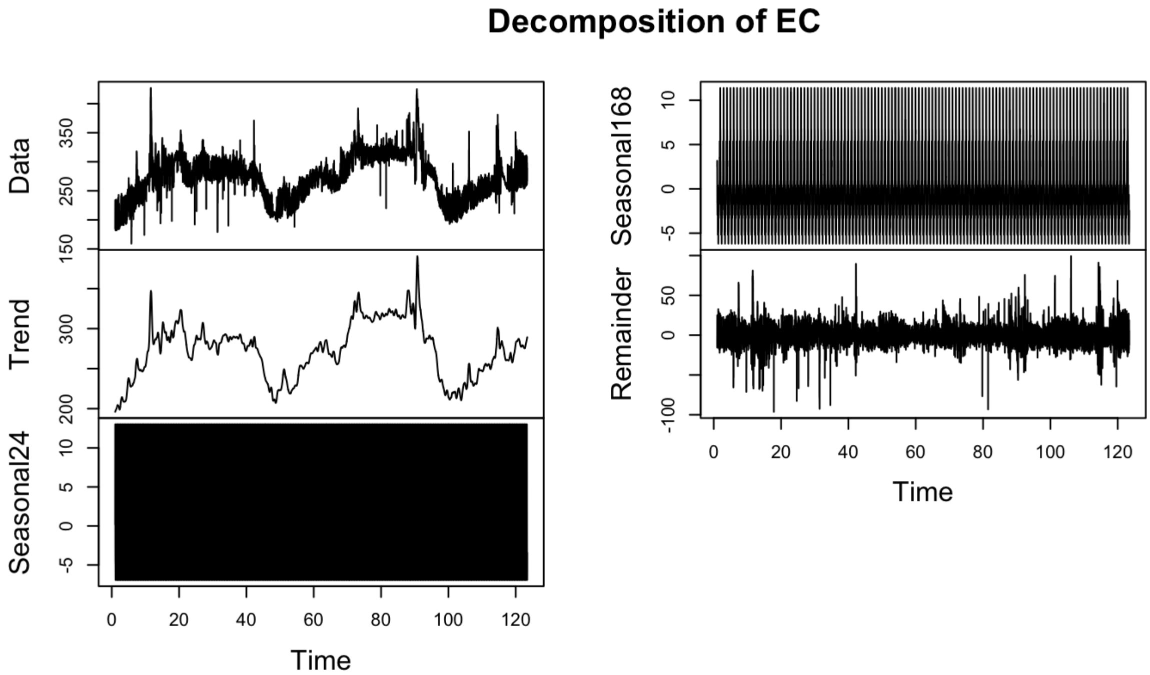


Figure 7.3: Decomposition plot of electrical conductivity (`ott.ec`), treated as an `msts` series, using the `mstl()` function from the forecast package.

8 Deseasonalization and Detrending

After careful consideration, we removed only the seasonality and trend components from the temperature (`ott.t`) and electrical conductivity (`ott.ec`) series, while leaving the water level (`ott.h`) unchanged. This decision is justified, since the water level in the Noce River is highly managed and does not exhibit the same natural seasonal-trend structure as the other variables.

From now on, we adopt the following notational conventions:

Notation	**Meaning**
`ott.h`	unchanged
`ott.t.d`	$\equiv$ `ott.t.deseasonalized`
`ott.ec.d`	$\equiv$ `ott.ec.deseasonalized`

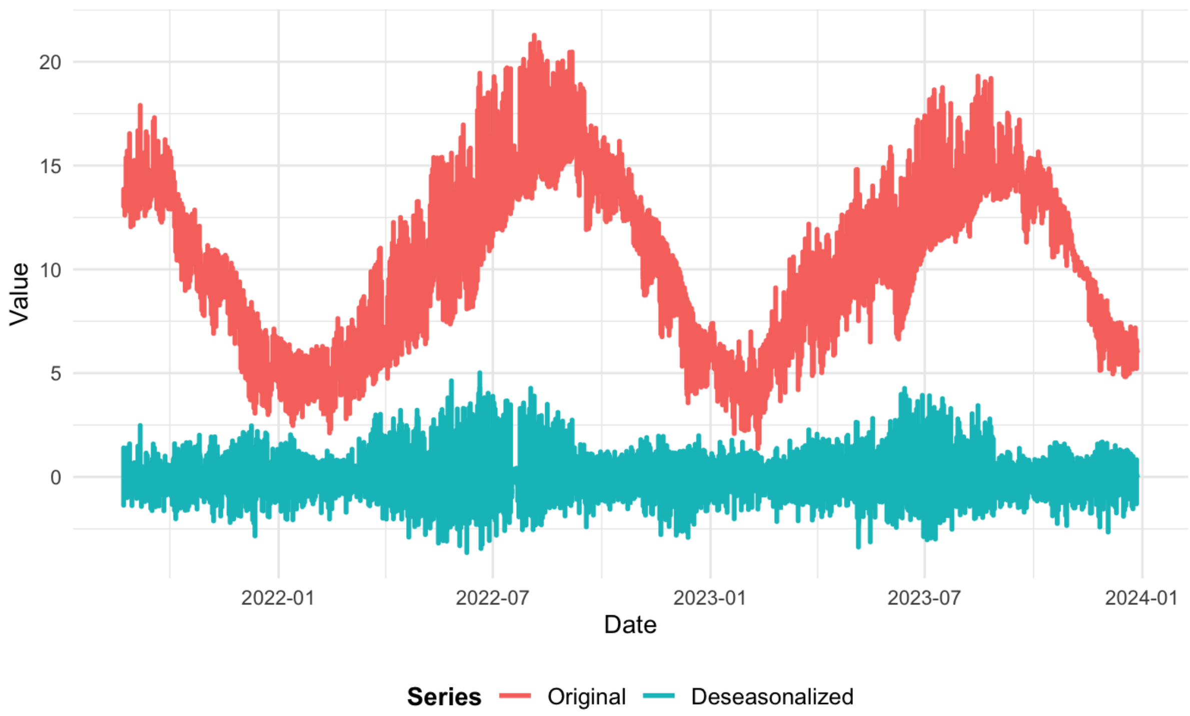


Figure 8.1: Comparison of the original temperature (`ott.t`) and deseasonalized temperature (`ott.t.d`) series from the Noce River, with the original observations shown in red and the deseasonalized values shown in teal.

Figure 8.2: Comparison of the original electrical conductivity (`ott.ec`) and deseasonalized electrical conductivity (`ott.ec.d`) series from the Noce River, with the original observations shown in red and the deseasonalized values shown in teal.

After removing the sesonality and trend components, examining the ACF and PACF plots for the scaled variables provides a clear view of any remaining temporal dependencies. The ACF reveals how each series correlates with its own past values at different lags, while the PACF focuses on the additional correlation contributed by each lag once the effects of earlier lags are taken into account.

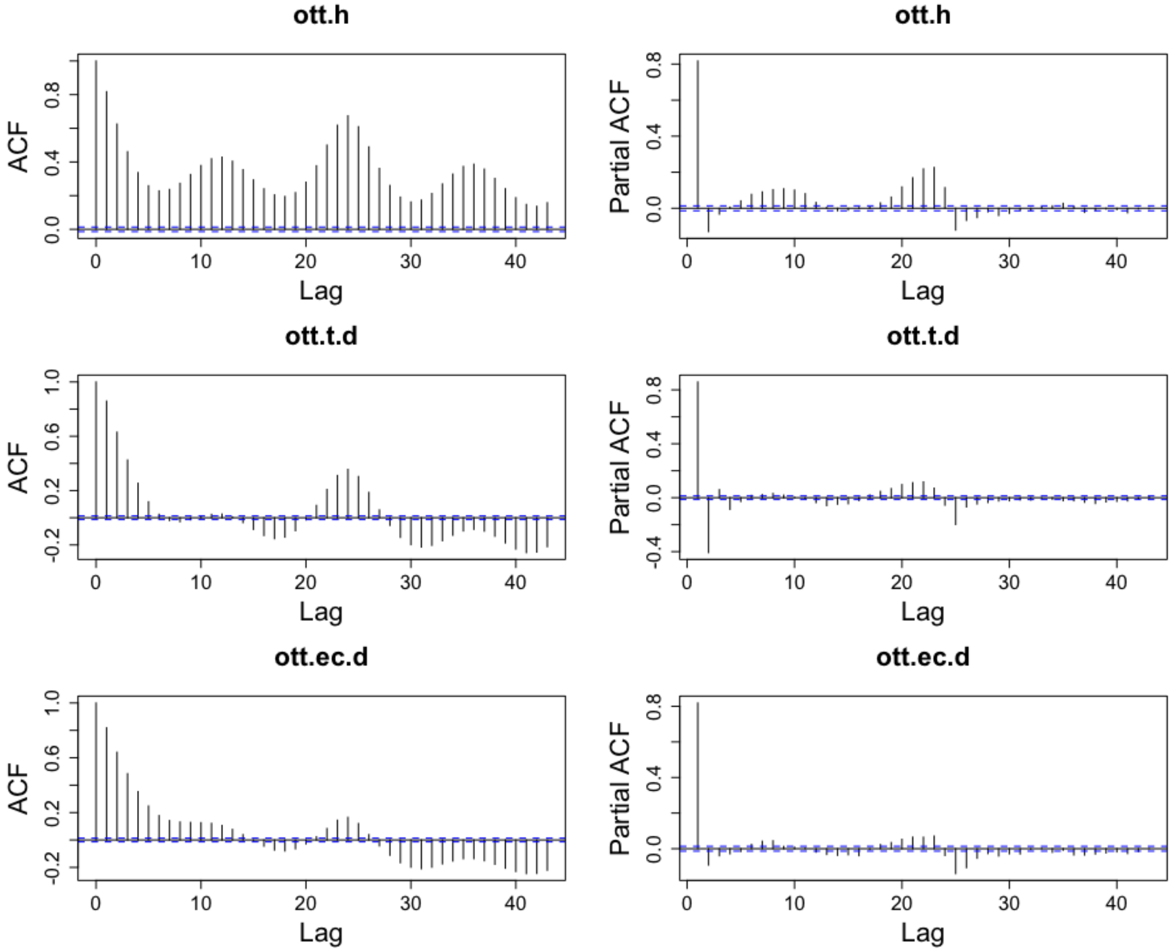


Figure 8.3: Autocorrelation (left column) and partial autocorrelation (right column) functions for the original water level series (`ott.h`) and the deseasonalized temperature (`ott.t.d`) and electrical conductivity (`ott.ec.d`) series from the Noce River.

9 Regime Switching using `MSwM` Package

In this section, we demonstrate how to fit Markov-switching models to our univariate time series using the `MSwM` package. We begin by specifying a baseline model through `lm()`, then invoke `msmFit()` to layer on a finite-state Markov process, choosing the number of regimes and the coefficients allowed to switch, and estimate parameters through the expectation maximization algorithm. Once the model is fitted, we employ `plotDiag()` to inspect residual behavior in each regime, and `msmResid()` to generate state-specific summary statistics.

9.1 Baseline Model

For each observed series $y_t^{(v)}$ (with $v \in \{\mathrm{h}, \mathrm{t}, \mathrm{ec}\}$) we begin with an intercept-only regression:

$$y_t^{(v)} = \mu^{(v)} + \varepsilon_t^{(v)}, \quad \varepsilon_t^{(v)} \overset{\text{iid}}{\sim} \mathcal{N}\big(0, \sigma^{2\,(v)}\big).$$

Fitting this in `R` via `lm(y~1)` yields the pooled estimates

$$\hat{\mu}^{(v)} = \bar{y}^{(v)}, \quad \hat{\sigma}^{2\,(v)} = \frac{1}{T}\sum_{t=1}^{T}\big(y_t^{(v)} - \bar{y}^{(v)}\big)^2.$$

These serve as starting values for the EM routine of the Markov-switching model.

Variable	Estimate	Std_Error	Residual_SD	N_Observations
ott.h	202.479	0.002	0.290	20568
ott.t.d	0.000	0.006	0.931	20568
ott.ec.d	-0.002	0.086	12.382	20568

Table 1: **Baseline (intercept-only)** estimates for `ott.h`, `ott.t.d`, and `ott.ec.d`, rounded to three decimal places.

Notes.

- *Estimate* $\hat{\mu}^{(v)} = \bar{y}^{(v)} = \frac{1}{T}\sum_{t=1}^{T} y_t^{(v)}$.
- *Std_Error*

$$\widehat{\mathrm{SE}}\big(\hat{\mu}^{(v)}\big) = \frac{\hat{\sigma}^{(v)}}{\sqrt{T}} \quad \text{where} \quad \hat{\sigma}^{(v)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\big(y_t^{(v)} - \bar{y}^{(v)}\big)^2}.$$

- *Residual_SD*

$$\hat{\sigma}^{(v)} = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\big(y_t^{(v)} - \bar{y}^{(v)}\big)^2},$$

  the sample standard deviation of the residuals.

- *N_Observations* T, the total number of time points.

9.2 Two-Regime Mean- and Variance-Switching Model without Autoregression (MV-AR(0))

For each variable $v \in \{h, t, ec\}$, we let

$$S_t \in \{1, 2\} \quad \text{follow a two-state Markov chain with transition matrix} \quad \mathbf{P} = \big(p_{ij}\big)_{i,j=1}^{2},$$

where

$$p_{ij} = \Pr\big(S_t = j \mid S_{t-1} = i\big), \quad \sum_{j=1}^{2} p_{ij} = 1 \quad (i, j = 1, 2).$$

Conditioned on the current regime $S_t = i$, each random observation follows,

$$Y_t^{(v)} \mid \{S_t = i\} \;\sim\; \mathcal{N}\big(\mu_i^{(v)},\, (\sigma_i^{(v)})^2\big).$$

Since we set the autoregressive order to zero ($p = 0$), the full parameter vector for variable v is

$$\boldsymbol{\theta}^{(v)} = \big(\mu_1^{(v)},\, \mu_2^{(v)},\, (\sigma_1^{(v)})^2,\, (\sigma_2^{(v)})^2,\, p_{11},\, p_{12},\, p_{21},\, p_{22}\big).$$

We fit these mean–variance switching models separately to the water level (`ott.h`), temperature (`ott.t.d`), and conductivity (`ott.ec.d`) series. Table 2 reports, for each variable v, the estimated regime means $\mu_i^{(v)}$, regime standard deviations $\sigma_i^{(v)}$, and transition probabilities p_{ij} $(i, j = 1, 2)$.

Variable	μ_1	σ_1	μ_2	σ_2	p_{11}	p_{12}	p_{21}	p_{22}
ott.h	202.284	0.0525	202.841	0.184	0.924	0.142	0.076	0.858
ott.t.d	0.030	0.436	-0.039	1.320	0.937	0.082	0.063	0.918
ott.ec.d	2.183	20.826	-0.809	6.871	0.915	0.031	0.085	0.969

Table 2: **MV-AR(0)** model estimates and transition-probability matrices for `ott.h`, `ott.t.d`, and `ott.ec.d`. Regime 1 has parameters (μ_1, σ_1) and Regime 2 has (μ_2, σ_2); $p_{ij} = \Pr(S_t = j \mid S_{t-1} = i)$.

9.3 Residual Diagnostic Plots for MV-AR(0) Models

To assess the adequacy of our multivariate AR(0) specifications, we examine several residual diagnostics for each variable: `ott.h`, `ott.t.d` and `ott.ec.d` series. For each fitted model we use the `plotDiag()` function from the `MSwM` package to generate:

- ACF/PACF of the standardized residuals,
- Normal Q–Q plots (to assess the normality assumption).

9.3.1 Height Model (`ott.h`)

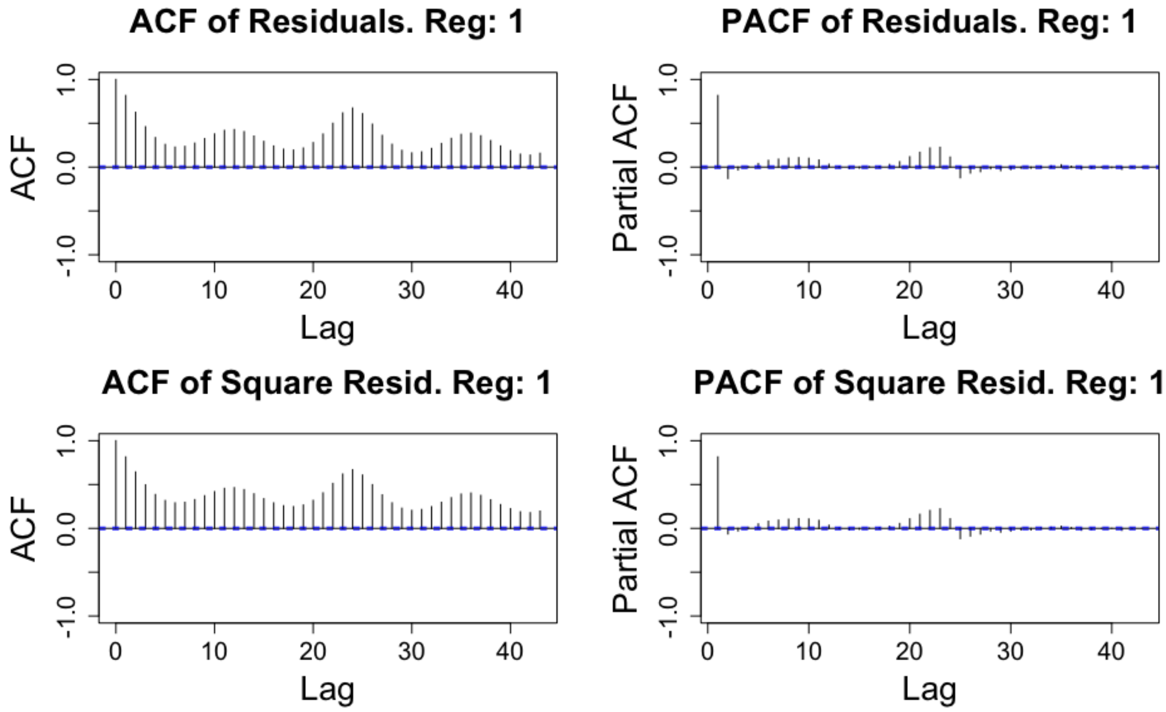


Figure 9.1: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(0) height model** (`ott.h`), generated via the `plotDiag()` function in the `MSwM` package.

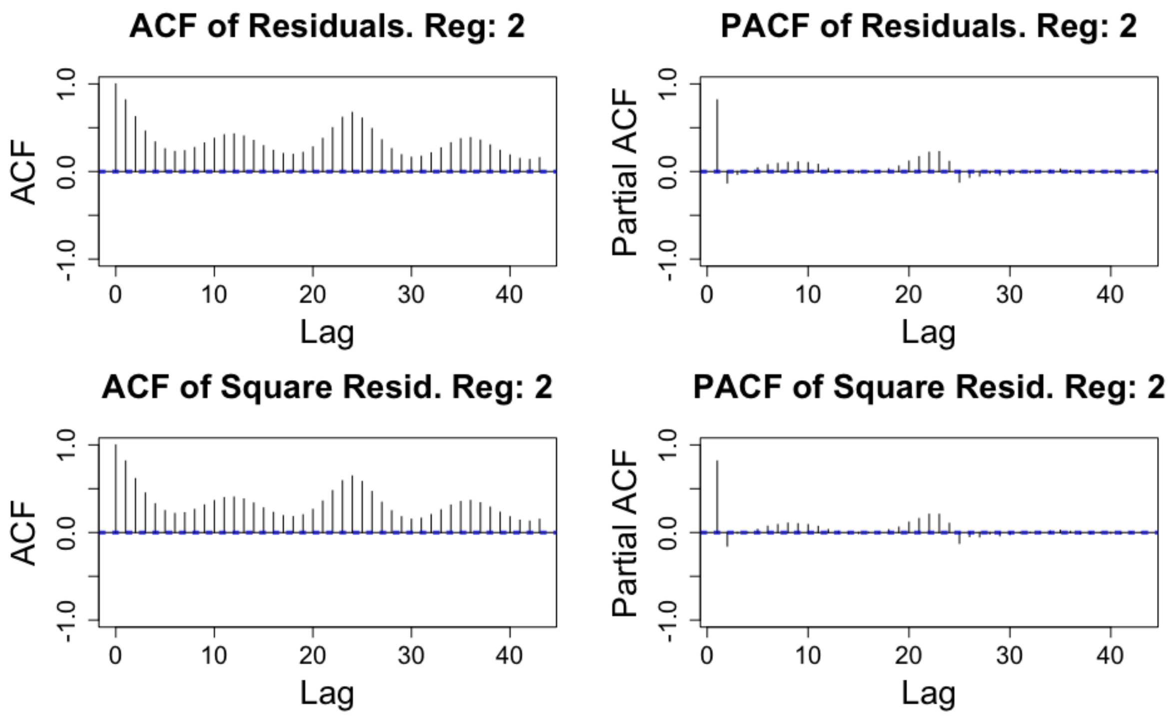


Figure 9.2: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(0) height model** (`ott.h`), generated via the `plotDiag()` function in the `MSwM` package.

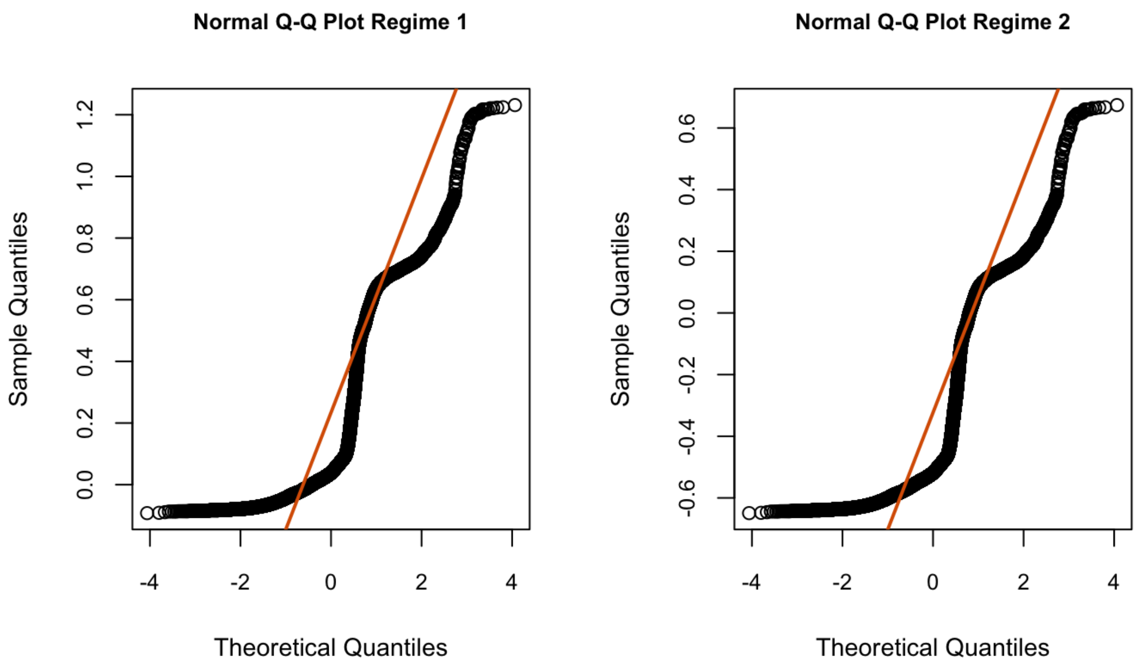


Figure 9.3: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(0) height model** (`ott.h`), created using the `plotDiag()` function in the `MSwM` package.

9.3.2 Deseasonalized Temperature (`ott.t.d`)

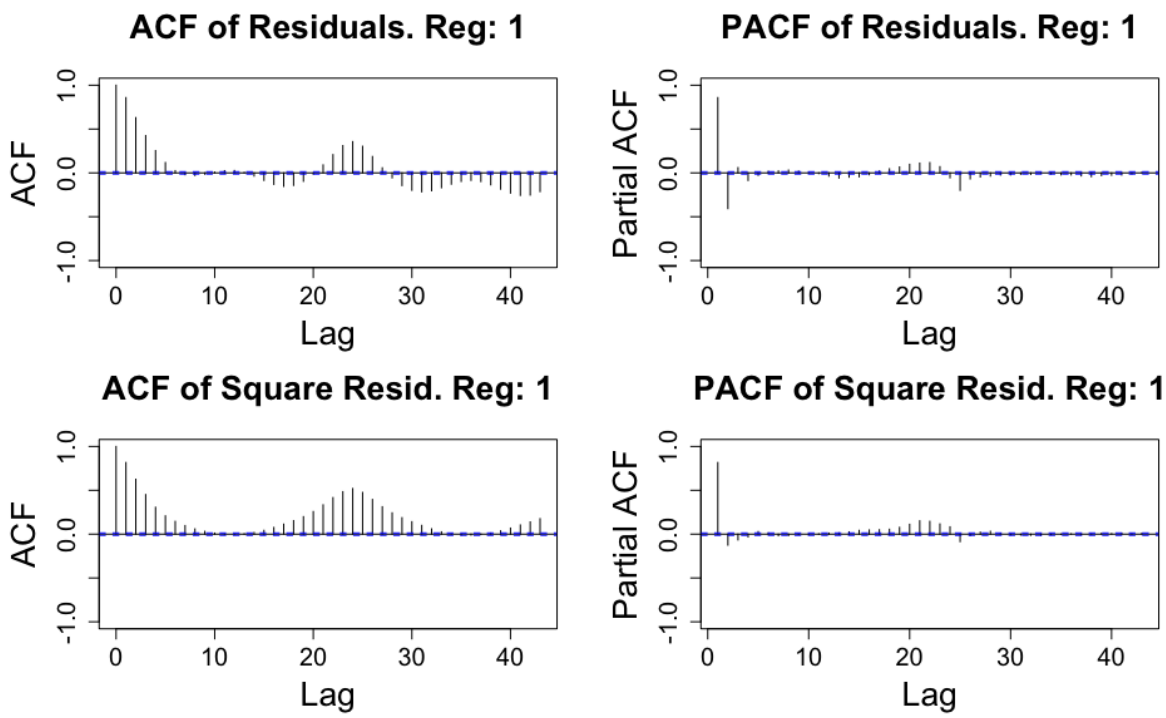


Figure 9.4: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(0) temperature model** (`ott.t.d`), generated via the `plotDiag()` function in the `MSwM` package.

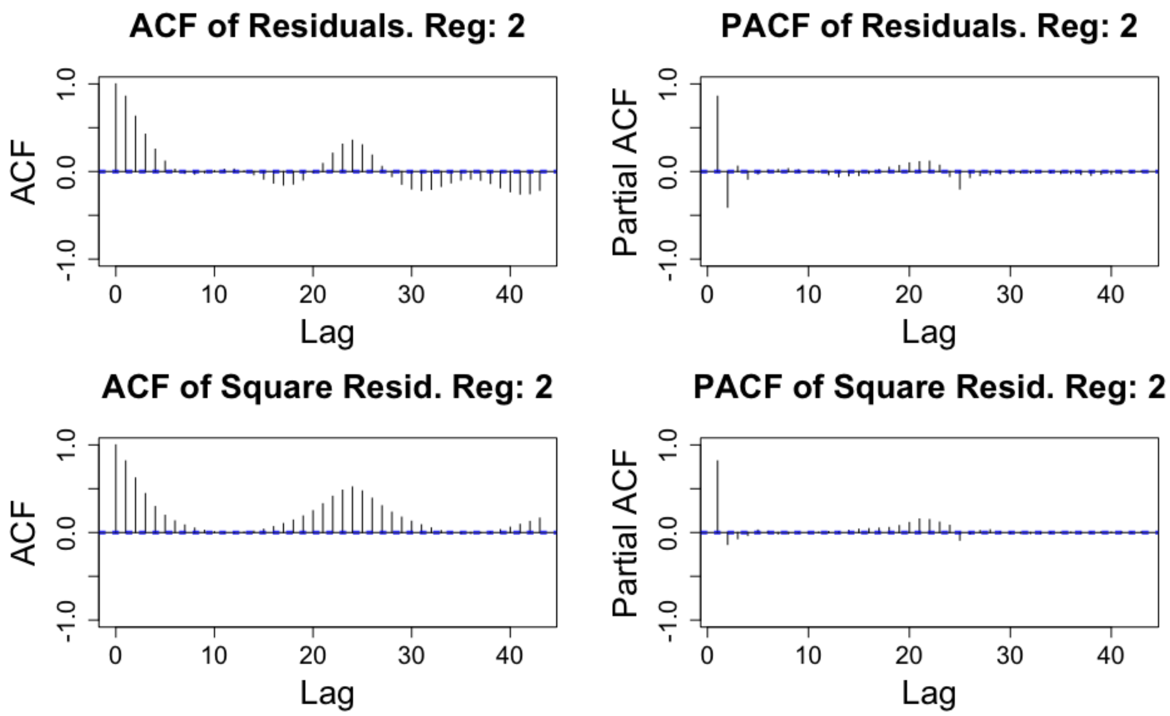


Figure 9.5: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(0) temperature model** (`ott.t.d`), generated via the `plotDiag()` function in the `MSwM` package.

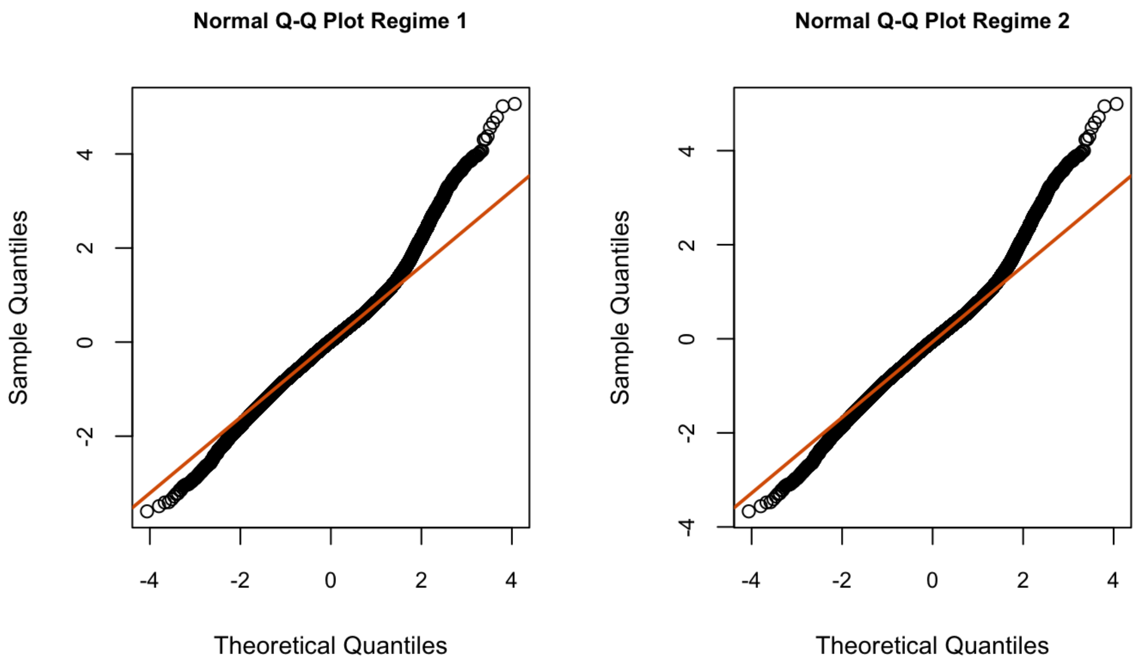


Figure 9.6: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(0) temperature model** (`ott.t.d`), created using the `plotDiag()` function in the `MSwM` package.

9.3.3 Deseasonalized Electrical Conductivity (`ott.ec.d`)

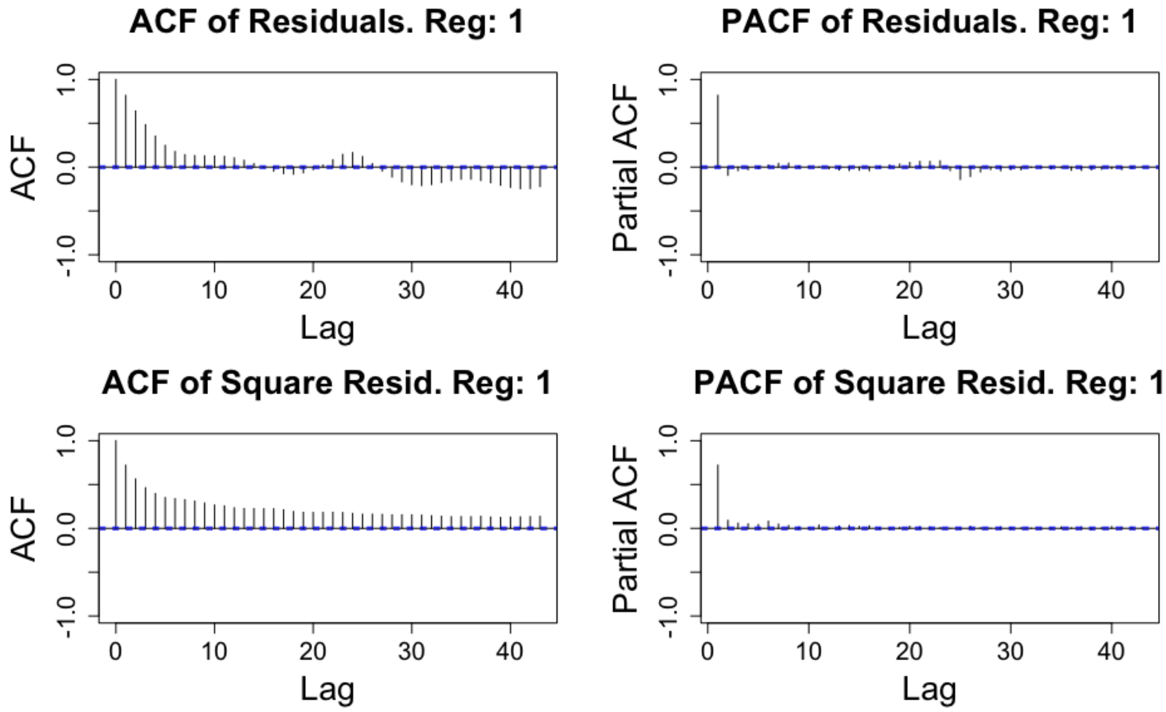


Figure 9.7: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(0) ec model** (`ott.ec.d`), generated via the `plotDiag()` function in the `MSwM` package.

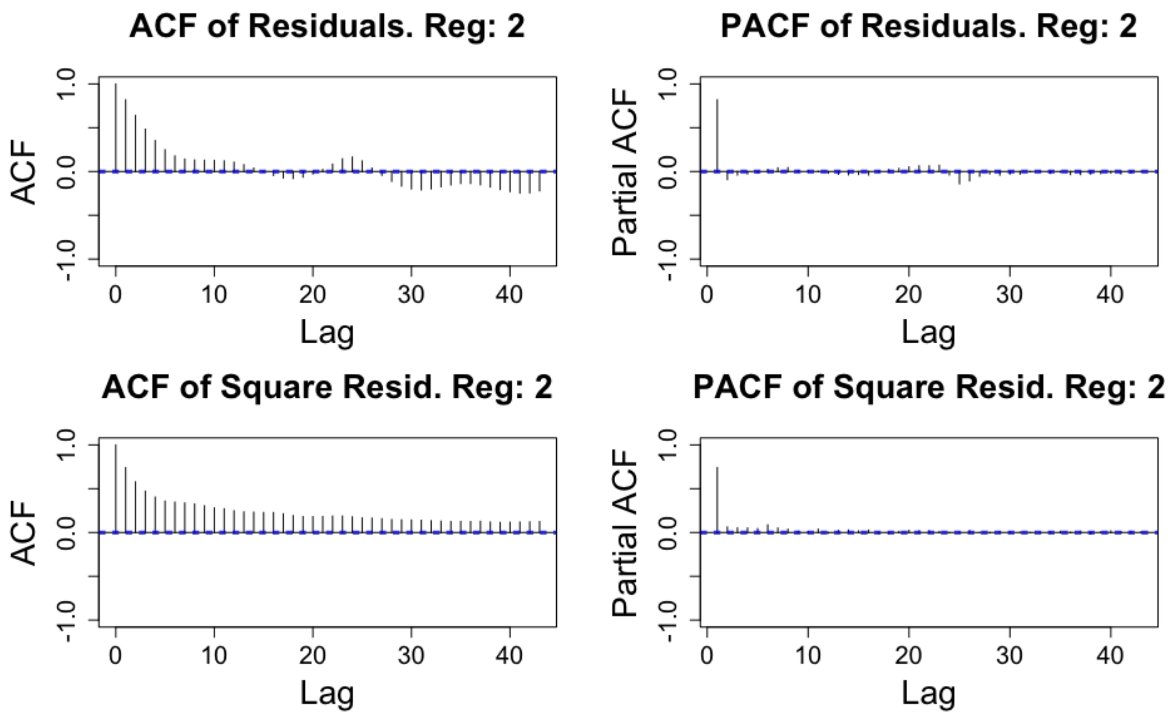


Figure 9.8: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(0) ec model** (`ott.ec.d`), generated via the `plotDiag()` function in the `MSwM` package.

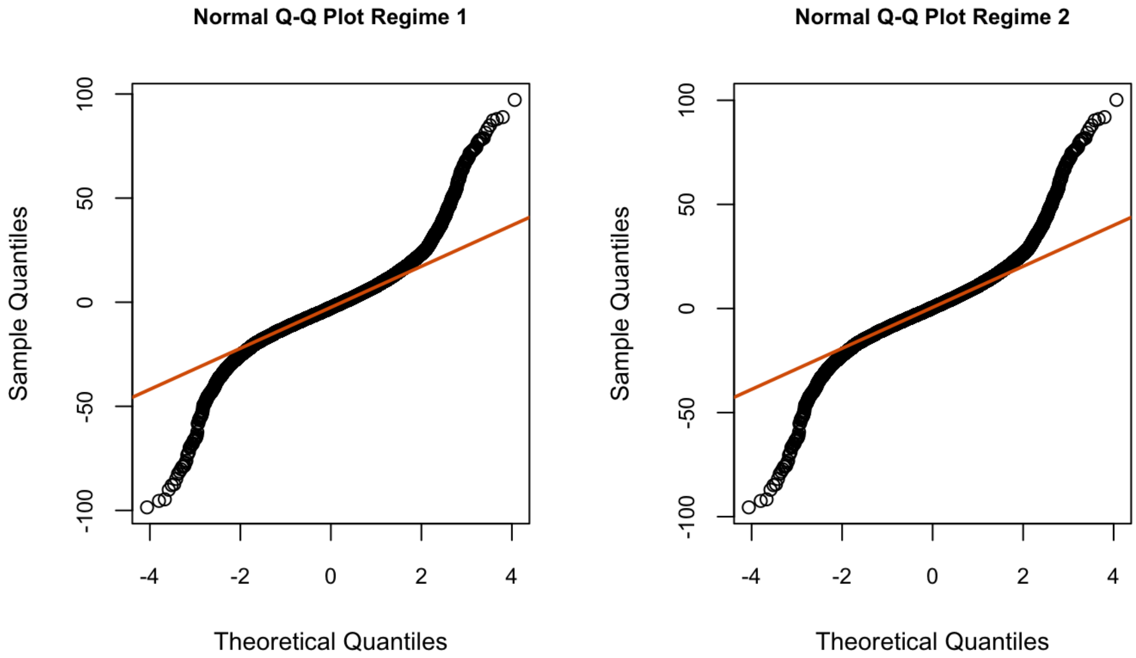


Figure 9.9: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(0) ec model** (`ott.ec.d`), created using the `plotDiag()` function in the `MSwM` package.

9.4 Two-Regime Mean and Variance Switching AR(1) Model - (MV-AR(1))

For each observed series $y_t^{(v)}$ (with $v \in \{\mathrm{h}, \mathrm{t}, \mathrm{ec}\}$) we use the same intercept-only linear models as the baseline models;

$$y_t^{(v)} \;=\; \mu^{(v)} + \varepsilon_t^{(v)}, \quad \varepsilon_t^{(v)} \overset{\text{iid}}{\sim} \mathcal{N}\big(0, \sigma^{2\,(v)}\big).$$

For each of the three variables, height (`ott.h`), temperature (`ott.t.d`) and electricity conductivity (`ott.ec.d`), we estimate a two-regime Markov-switching AR(1) model. Let $S_t \in \{1, 2\}$ be the unobserved regime at time t, evolving as a first-order Markov chain. Conditional on $S_t = i$, the process satisfies

$$y_t \;=\; \mu_i \;+\; \phi_i\, y_{t-1} \;+\; \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_i^2).$$

Hence both the conditional mean parameters μ_i, ϕ_i and the innovation variance σ_i^2 switch across the two regimes.
The hidden regime evolves according to

$$\Pr\big(S_t = j \mid S_{t-1} = i\big) \;=\; p_{ij}, \quad \sum_{j=1}^{2} p_{ij} = 1, \quad i, j \in \{1, 2\}.$$

Thus, each series has ten parameters:

$$\boldsymbol{\theta} = \big(\mu_1, \mu_2,\ \phi_1, \phi_2,\ \sigma_1^2, \sigma_2^2,\ p_{11}, p_{12}, p_{21}, p_{22}\big).$$

- μ_i is the regime-specific mean.
- ϕ_i is the regime-specific AR(1) term.

- σ_i^2 is the regime-specific variance.
- $\mathbf{P} = (p_{ij})$ is the Markov transition matrix.
- One autoregressive lag is included ($p = 1$).

Series	$\hat{\mu}_1$	$\hat{\phi}_1$	$\hat{\sigma}_1$	$\hat{\mu}_2$	$\hat{\phi}_2$	$\hat{\sigma}_2$	p_{11}	p_{12}	p_{21}	p_{22}
ott.h	133.747	0.340	0.260	0.251	0.999	0.011	0.655	0.150	0.345	0.850
ott.t.d	0.074	0.708	0.871	−0.017	0.960	0.195	0.736	0.085	0.264	0.915
ott.ec.d	−0.385	0.539	13.073	0.191	0.967	2.923	0.643	0.100	0.357	0.900

Table 3: Estimated parameters for the **MV-AR(1)** models. For each series, Regime 1 has intercept μ_1, AR(1) coefficient ϕ_1, and residual standard deviation σ_1, while Regime 2 has μ_2, ϕ_2, and σ_2. The entries p_{ij} form the transition-probability matrix for S_t.

9.5 Residual Diagnostic Plots for MV-AR(1) Models

9.5.1 Height Model (`ott.h`)

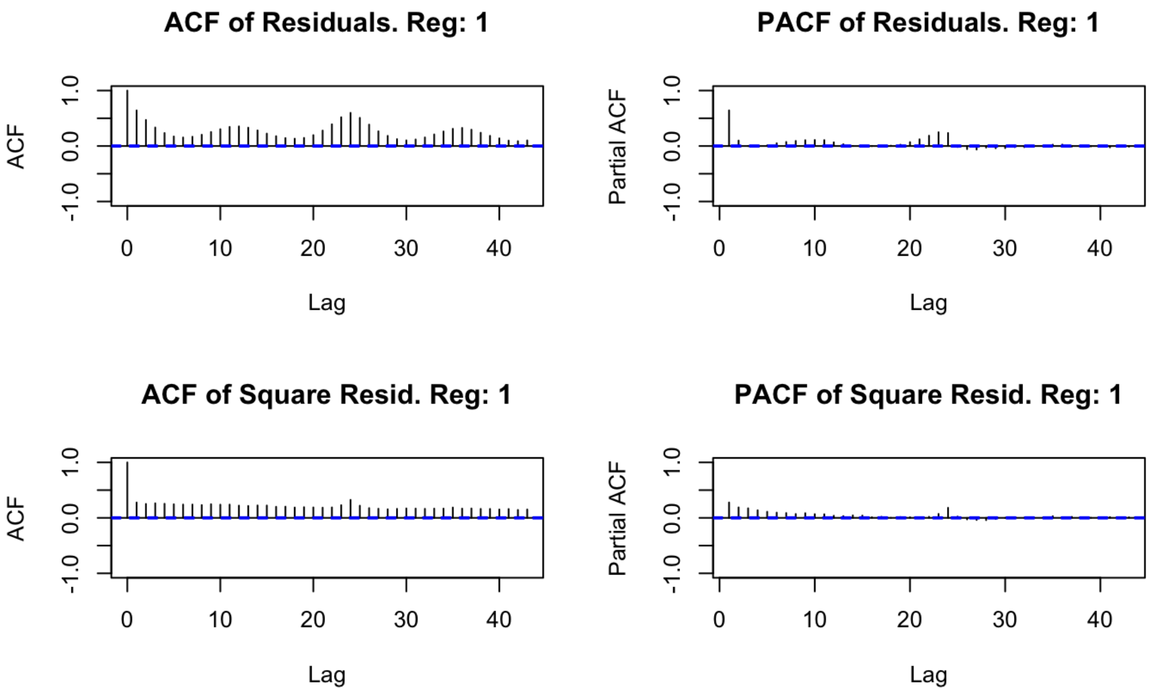


Figure 9.10: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(1) height model** (`ott.h`), generated via the `plotDiag()` function in the `MSwM` package.

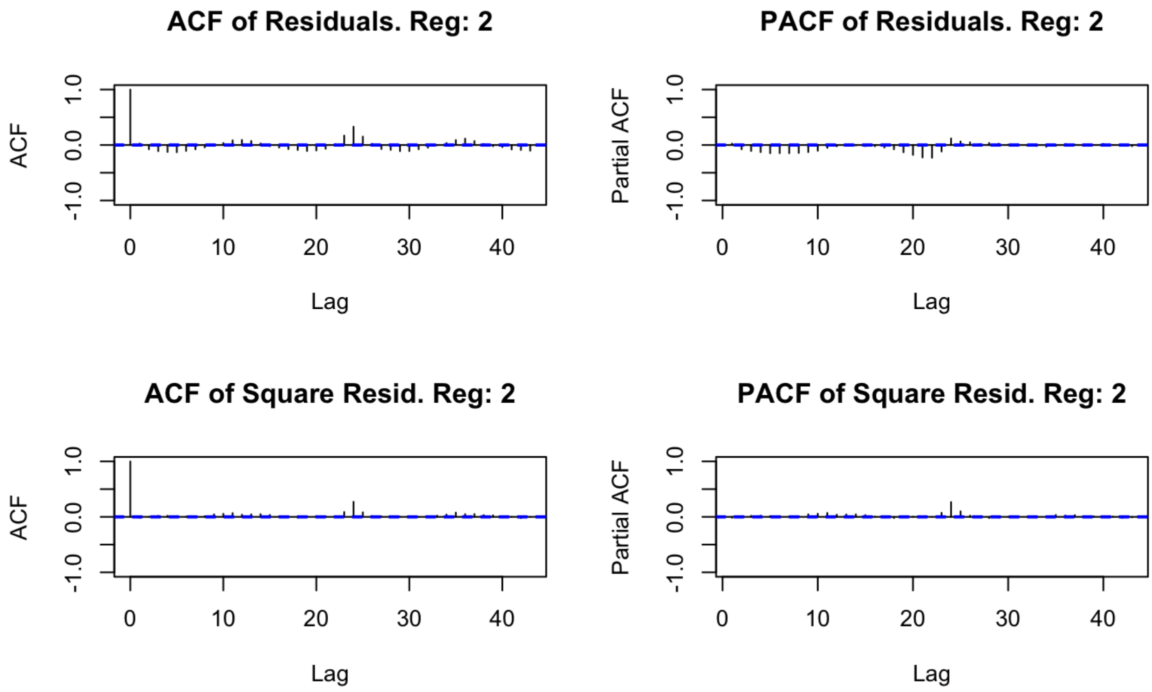


Figure 9.11: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(1) height model** (`ott.h`), generated via the `plotDiag()` function in the `MSwM` package.

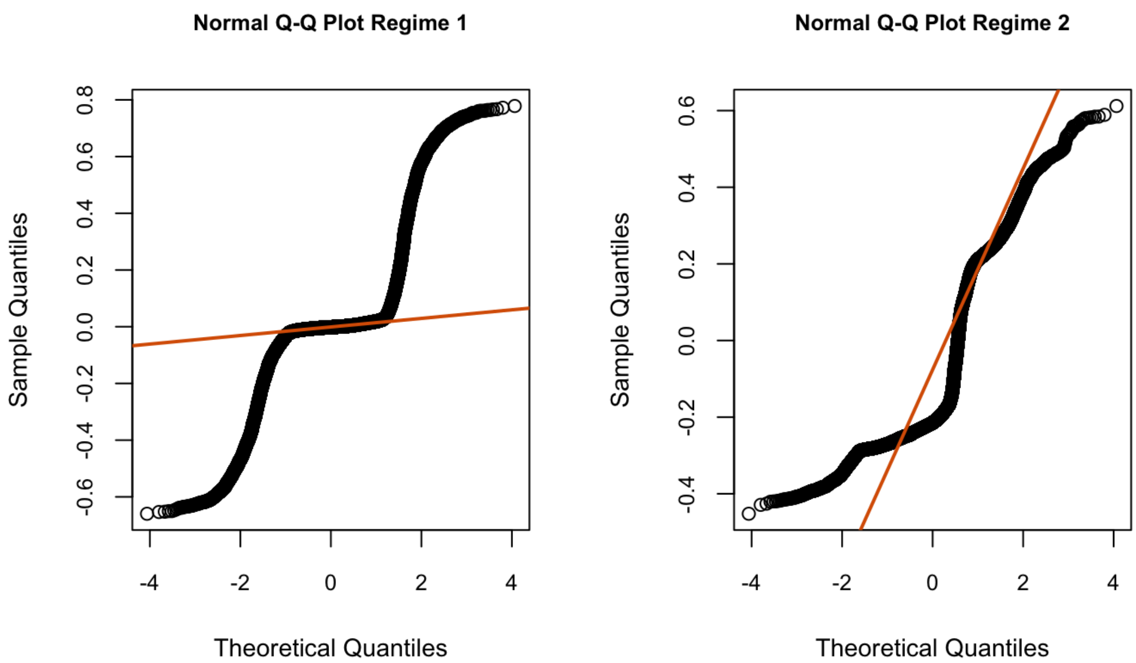


Figure 9.12: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(1) height model** (`ott.h`), created using the `plotDiag()` function in the `MSwM` package.

9.5.2 Deseasonalized Temperature (`ott.t.d`)

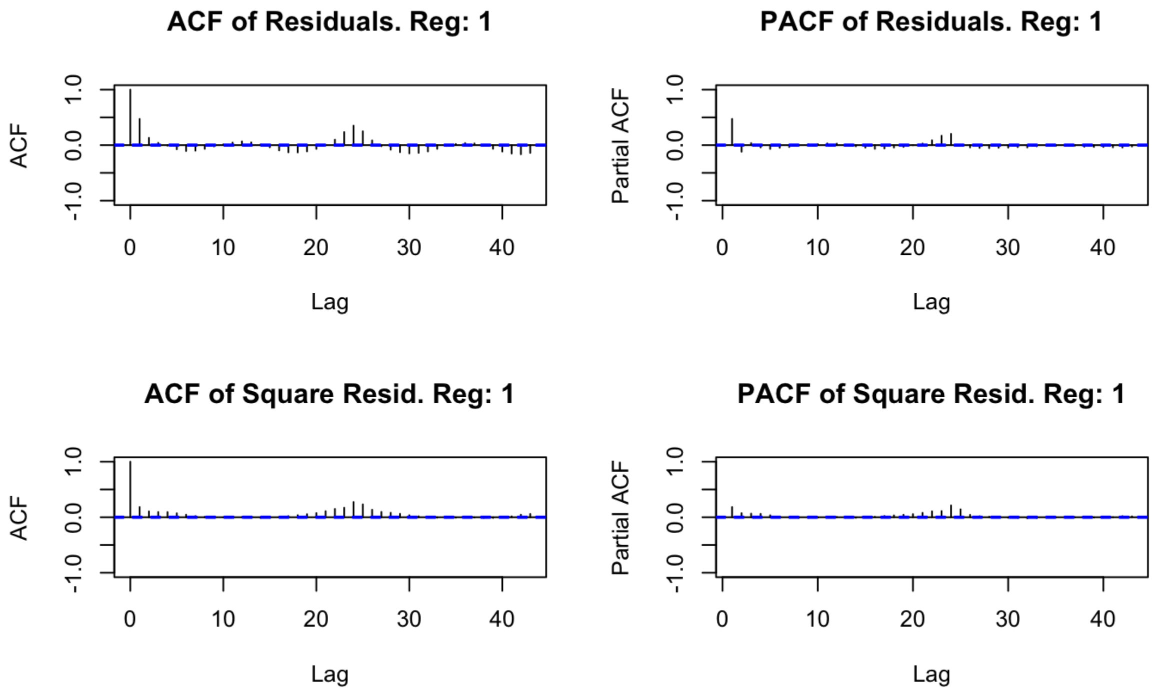


Figure 9.13: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(1) temperature model (`ott.t.d`)**, generated via the `plotDiag()` function in the `MSwM` package.

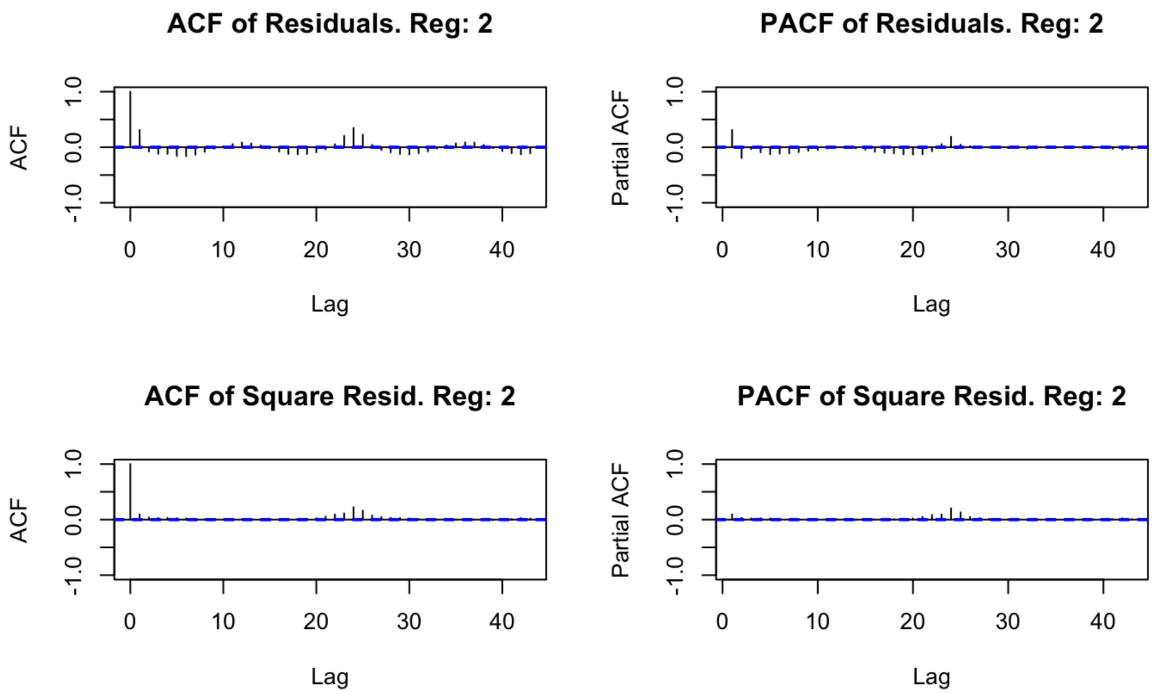


Figure 9.14: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(1) temperature model (`ott.t.d`)**, generated via the `plotDiag()` function in the `MSwM` package.

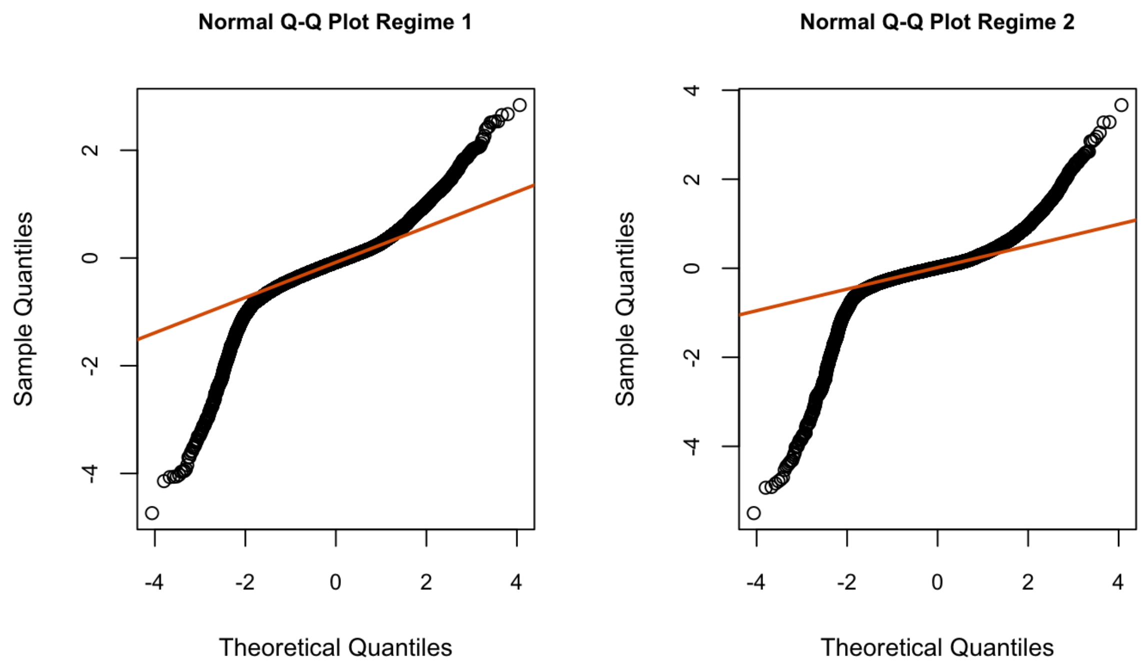


Figure 9.15: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(1) temperature model** (`ott.t.d`), created using the `plotDiag()` function in the `MSwM` package.

9.5.3 Deseasonalized Electrical Conductivity (`ott.ec.d`)

Figure 9.16: ACF and PACF diagnostics for the **Regime-1 residuals** of the **MV-AR(1) ec model** (`ott.ec.d`), generated via the `plotDiag()` function in the `MSwM` package.

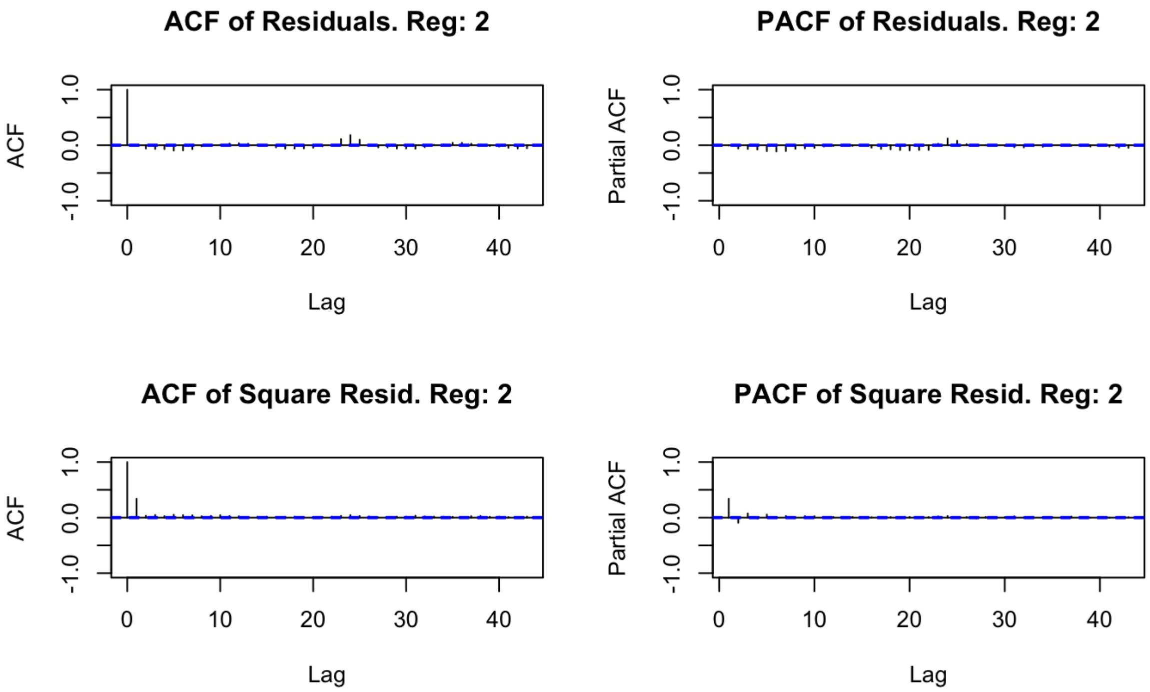


Figure 9.17: ACF and PACF diagnostics for the **Regime-2 residuals** of the **MV-AR(1) ec model** (`ott.ec.d`), generated via the `plotDiag()` function in the `MSwM` package.

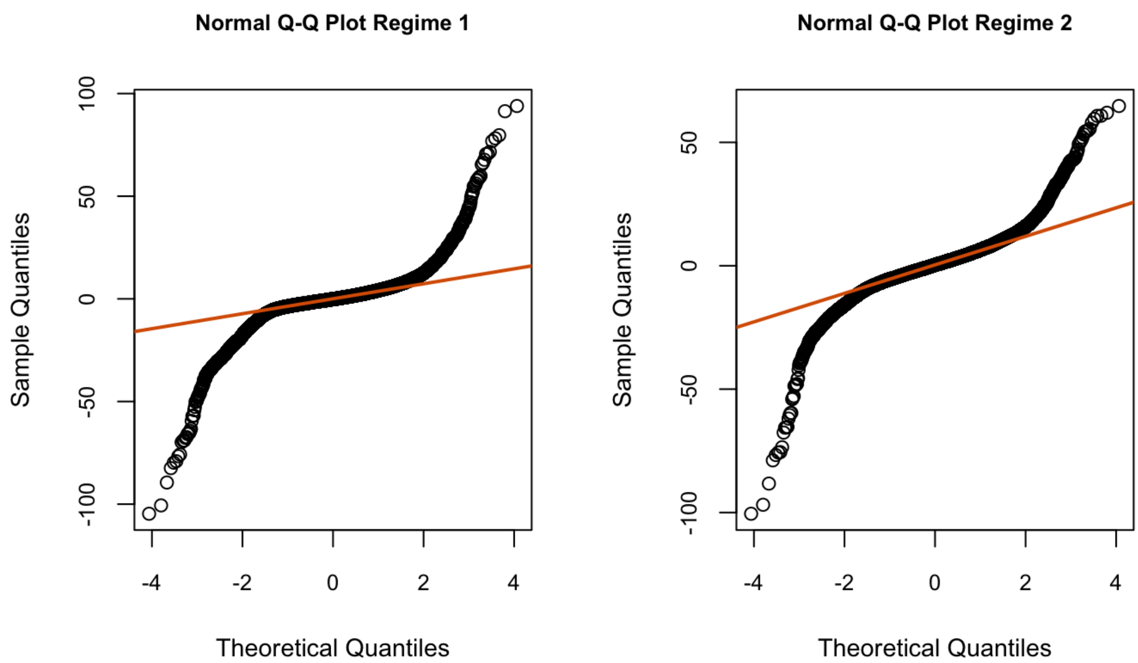


Figure 9.18: Normal Q–Q plots of the standardized residuals for **Regime 1 and Regime 2** from the **MV-AR(1) ec model** (`ott.ec.d`), created using the `plotDiag()` function in the `MSwM` package.

10 Comparison of Regime-Switching Model Fits

In this section, we summarize and compare the goodness-of-fit statistics for all Markov-switching models fitted to the three ott variables (`ott.h`, `ott.t.d`, `ott.ec.d`) using MSwM package under two dynamic specifications (no autoregression and AR(1)).

Variable (AR order)	Log-likelihood	AIC	BIC
ott.h.AR(0)	16 150.64	−32 297.28	−32 262.01
ott.h.AR(1)	36 005.86	−72 003.72	−71 932.44
ott.t.AR(0)	−24 263.02	48 530.04	48 566.76
ott.t.AR(1)	−6629.84	13 267.68	13 338.13
ott.ec.AR(0)	−76 654.11	153 347.22	153 312.68
ott.ec.AR(1)	−61 992.87	123 992.74	124 064.19

In every case the AR(1) specification yields substantially higher log-likelihoods and lower information criteria, indicating that including a single autoregressive lag greatly enhances model fit. We therefore adopt the **MV-AR(1)** (two-regime mean and variance witching AR(1)) models as the preferred specification for subsequent analysis.

11 Regime-Specific Data Splits

At this point, we convert each univariate AR(1) model's underlying smoothed state probabilities into distinct regime labels, which we subsequently merge into a single joint-regime factor. We can acquire distinct data subsets for every potential regime configuration by dividing the dataset based on this joint component. Our copula analysis uses these regime-specific subsamples as input, allowing us to investigate how the dependence structure between the series changes under various joint situations.

We first split the full dataset into eight disjoint subsets, one for each joint regime combination inferred from the MV-AR(1) models. For each series, we classify its two regimes as "High" or "Low" based on which regime's estimated intercept ($\hat{\mu}$) from the AR(1) fit is larger. Within each subset, we apply a kernel density estimate to perform a probability integral transform, mapping each variable's observations onto the unit interval required for copula estimation.

Table 4: Estimated means, standard deviations, and regime labels for each variable

Variable	Regime 1	Regime 2
`ott.h`	Higher mean: 133.7781, SD = 0.2602	Lower mean: 0.2513, SD = 0.0105
`ott.t.d`	Higher mean: 0.0737, SD = 0.8707	Lower mean: -0.0174, SD = 0.1944
`ott.ec.d`	Lower mean: -0.3861, SD = 13.0706	Higher mean: 0.1901, SD = 2.9227

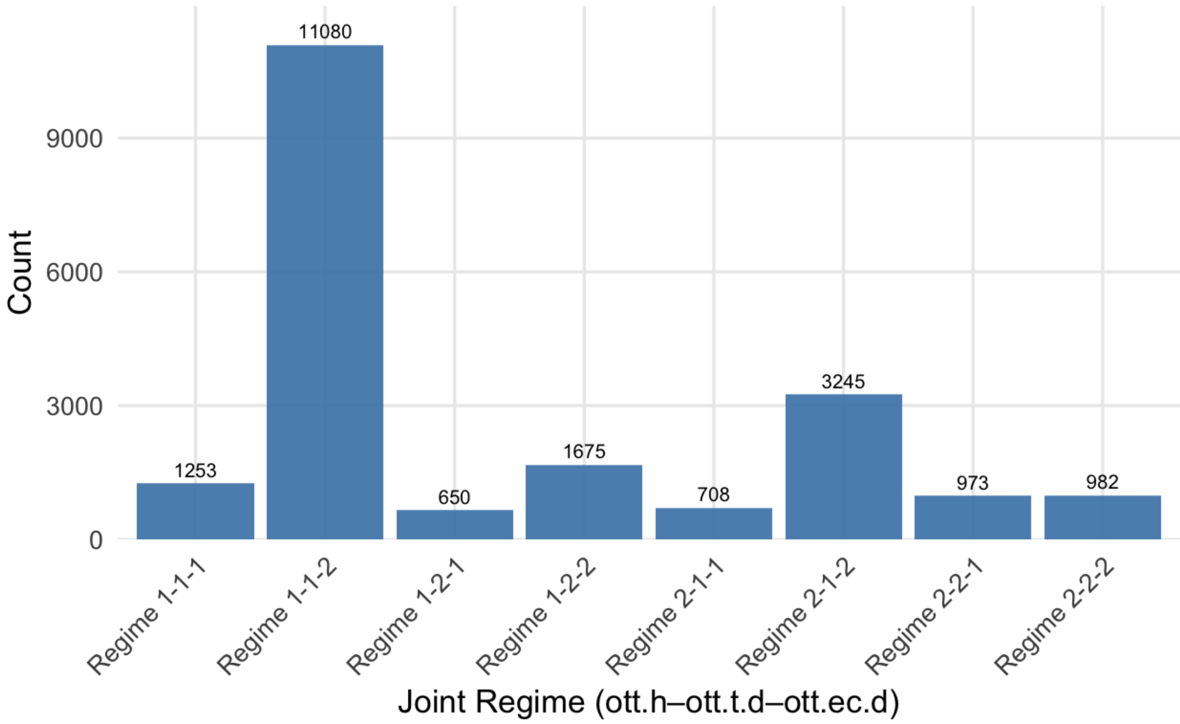


Figure 11.1: Frequency of observations for joint regimes

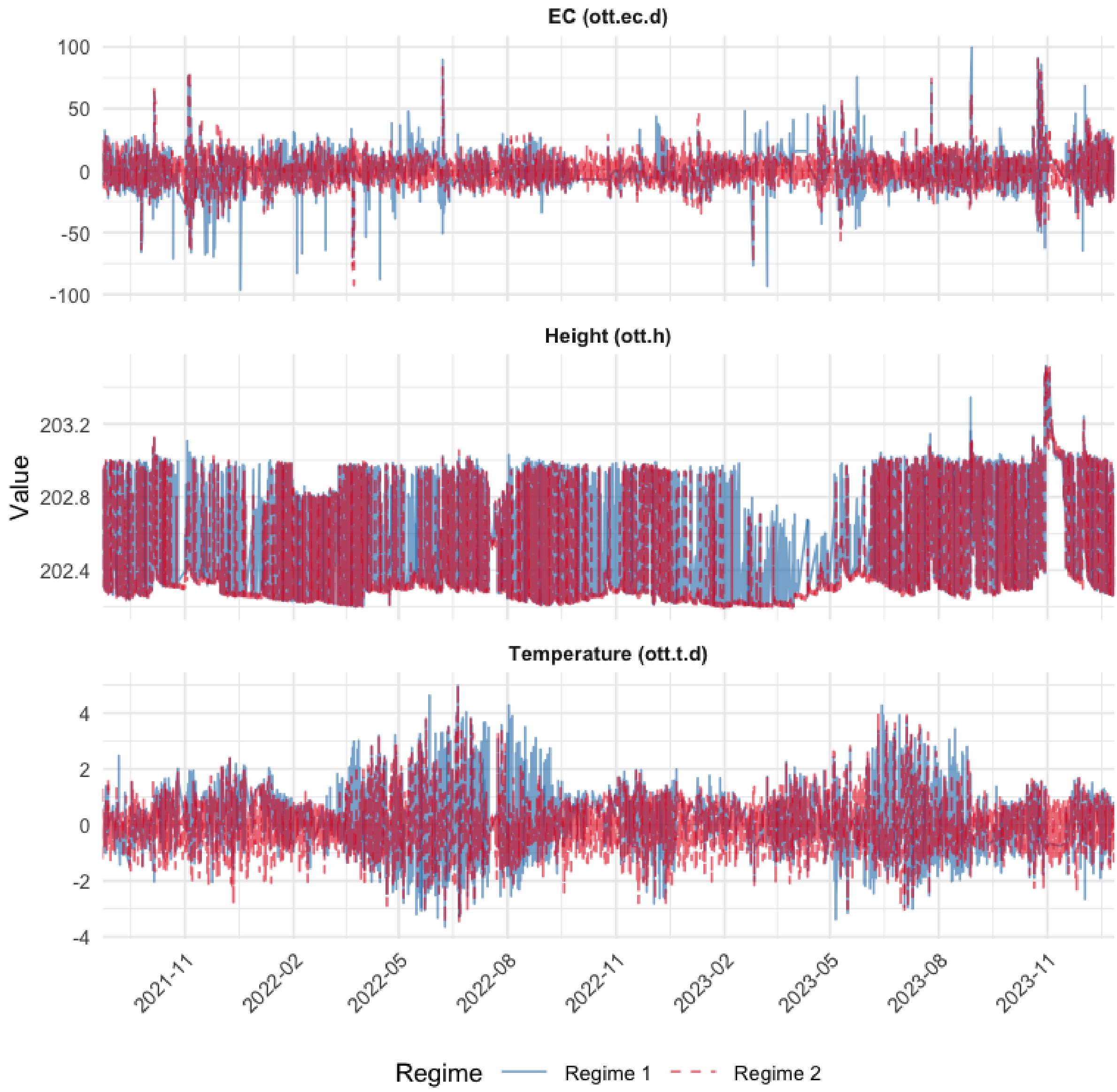


Figure 11.2: Time-series of height (`ott.h`), deseasonalized temperature (`ott.t.d`), and deseasonalized electrical conductivity (`ott.ec.d`), each colored by its inferred two-state regime.

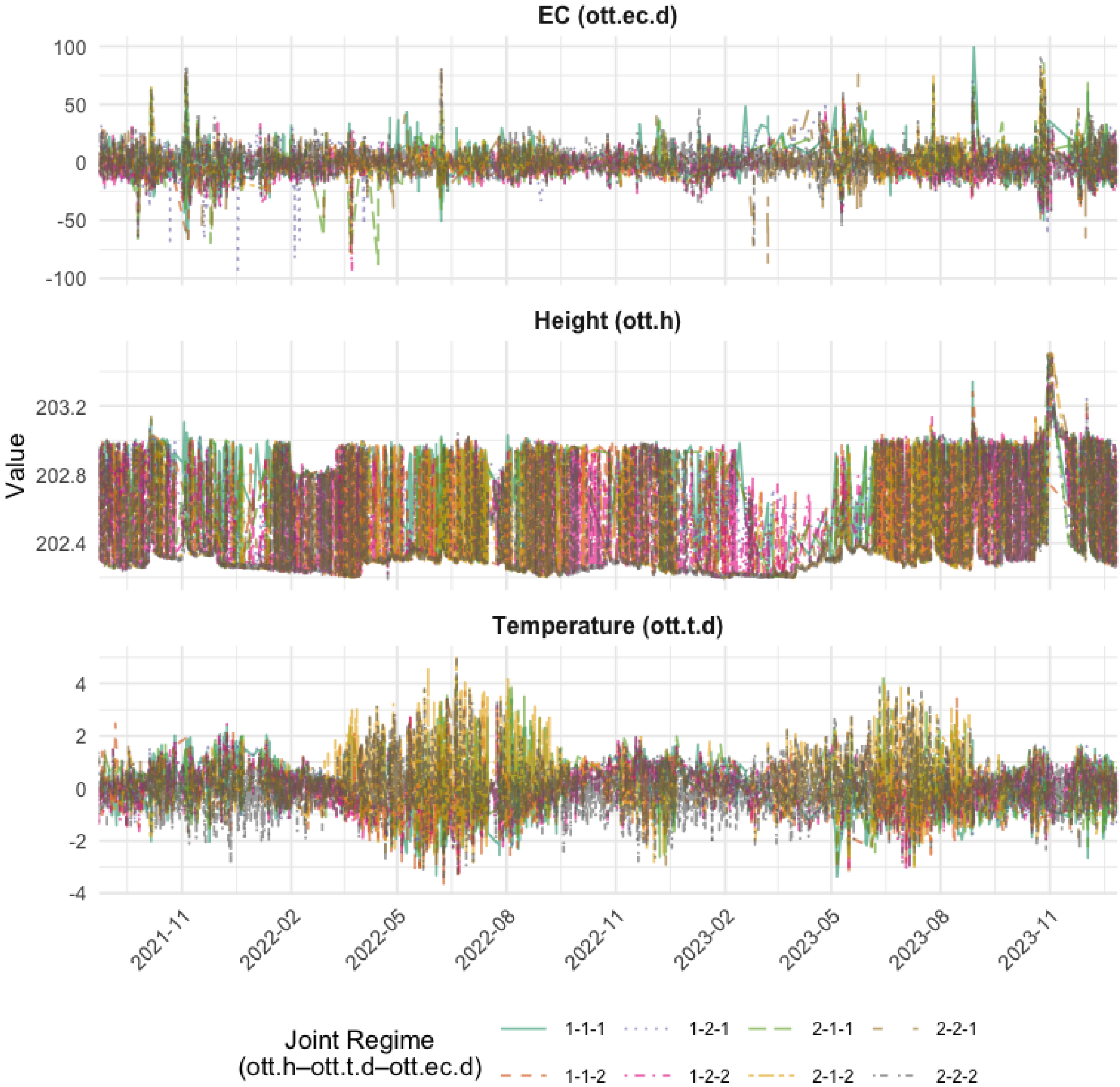


Figure 11.3: Time series of height (`ott.h`), deseasonalized temperature (`ott.t.d`), and deseasonalized electric conductivity (`ott.ec.d`), each colored by the inferred joint regime combination (ott.h–ott.t.d–ott.ec.d).

12 Copula Analysis

We conduct three complementary copula analyses on each regime-specific dataset:

1. **OTT Copula Analysis.** We fit a vine copula to the three OTT variables (`ott.h`, `ott.t.d` and `ott.ec.d`) in isolation, to characterize the dependence structure among the OTT measurements under each of the eight joint regimes.

2. **Joint Copula Analysis.** We extend the OTT-only model by including six additional precipitation variables alongside the three OTT variables, and fit a vine copula to all nine variables simultaneously in each of the eight joint regimes, capturing cross-dependencies between OTT and rainfall.

3. **Regime Specific Vine Regressions.** For each regime dataset, we perform three separate vine-regression analyses, each pairing one OTT variable with the six rain variables. This yields regime-specific conditional dependence models of each OTT variable given the rainfall covariates. The vine regressions are performed using the `vinereg_zi()` described in Sulejmani et al. (2025).

12.1 OTT Copula Analysis

For each of the eight joint regimes (i.e., all combinations of the two-state assignments in height, temperature, and electrical conductivity), we directly fitted a vine copula with `vine()` (from the `rvinecopulib` package). Table 5 summarizes the results for each joint regime.

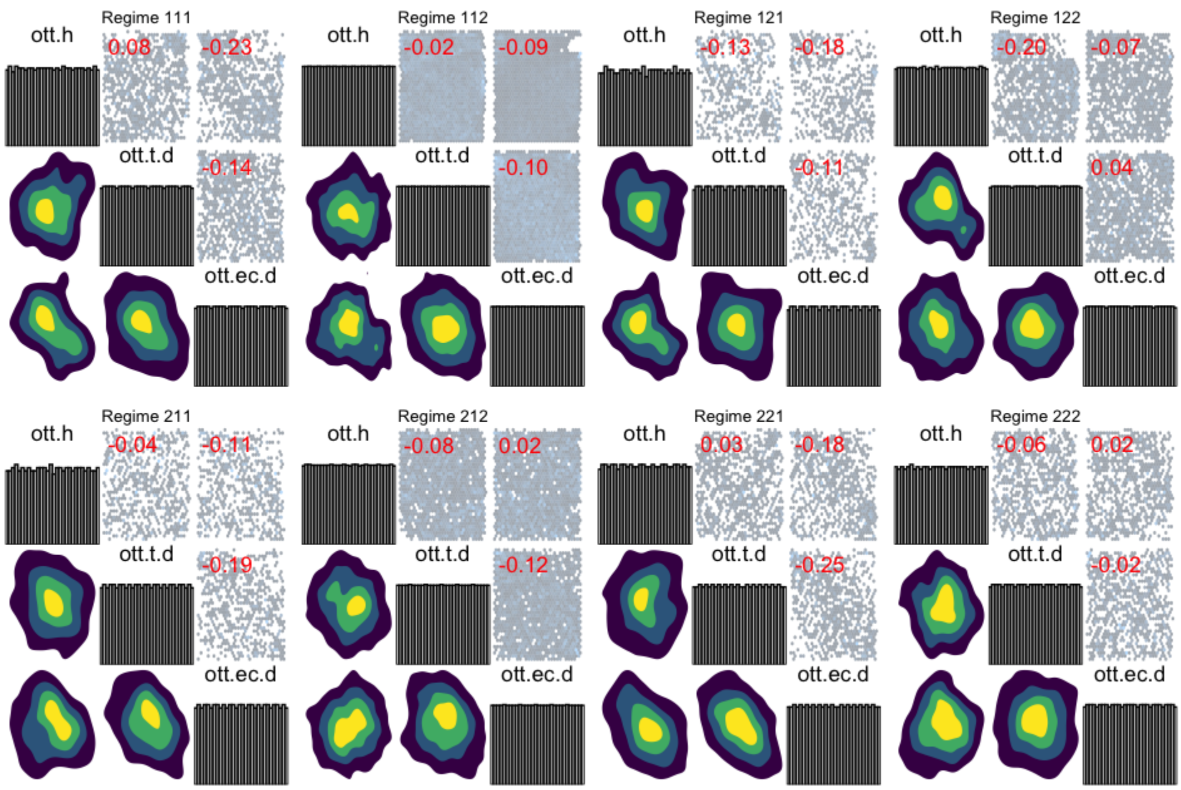


Figure 12.1: Pairs copula plots of ott variables across each regime.

Table 5: Regime-specific copula fits of OTT copula analysis (without precipitation variables). H/L labels correspond to higher/lower mean regimes for (`ott.h`, `ott.t.d`, `ott.ec.d`).

Regime summary	**Tree**	**Edge**	**Copula**	**Par1**	**Par2**	τ
Regime 1–1–1 (H–H–L) (n=1253)	1	(1,3)	tawn	0.34	0.56	–0.19
	1	(2,3)	BB8	2.08	0.87	–0.25
	2	(1,2);3	Indep	0.00	0.00	0.00
Regime 1–1–2 (H–H–H) (n=11080)	1	(1,2)	Gaussian	–	–	–0.062
	1	(2,3)	Joe	1.10	–	–0.048
	2	(1,3);2	Clayton	0.11	–	0.053
Regime 1–2–1 (H–L–L) (n=650)	1	(1,3)	tawn	0.30	0.65	–0.14
	1	(2,3)	BB8	1.84	0.85	–0.19
	2	(1,2);3	Frank	–	–	–0.063
Regime 1–2–2 (H–L–H) (n=1675)	1	(1,2)	Frank	–	–	–0.074
	1	(2,3)	Clayton	0.25	–	–0.111
	2	(1,3);2	Indep	0.00	0.00	0.000
Regime 2–1–1 (L–H–L) (n=708)	1	(2,1)	Clayton	0.27	–	–0.12
	1	(1,3)	BB8	1.89	0.83	–0.20
	2	(2,3);1	Gaussian	–	–	–0.16
Regime 2–1–2 (L–H–H) (n=3245)	1	(2,1)	BB8	1.78	0.92	–0.223
	1	(1,3)	Clayton	0.14	–	–0.067
	2	(2,3);1	Indep	0.00	0.00	0.000
Regime 2–2–1 (L–L–L) (n=973)	1	(1,3)	BB8	3.35	0.57	–0.235
	1	(2,3)	Frank	–	–	–0.142
	2	(1,2);3	Frank	0.47	–	0.052
Regime 2–2–2 (L–L–H) (n=982)	1	(1,3)	tll	[30×30 grid]	–	–0.094
	1	(2,3)	BB8	1.60	0.72	–0.099
	2	(1,2);3	Clayton	0.087	–	–0.042

Variable mapping:

1 $\leftrightarrow$ `ott.h`, 2 $\leftrightarrow$ `ott.t.d`, 3 $\leftrightarrow$ `ott.ec.d`

Conclusions In Table 4 we define each variable's "Regime 1" state and "Regime 2" state. Thus, for example, "Regime 1–1–1" corresponds to high-height, high-temperature, and low-conductivity, whereas "Regime 2–2–2" is low-height, low-temperature, and high-conductivity. From the copula fits in Table 5, we draw the following conclusions: the strongest negative dependence between height and conductivity ($\tau \approx -0.19$) and between temperature and conductivity ($\tau \approx -0.25$) occurs under Regime 1–1–1, indicating a pronounced inverse relationship when height and temperature are high but conductivity is low; mixed-state regimes (e.g., Regime 1–2–1 or Regime 2–1–2) display much weaker or negligible dependence (often defaulting to independence in Tree 2, $|\tau| < 0.12$); in Regime 1–1–2 the height–conductivity edge even flips to slight positive dependence ($\tau \approx +0.05$) when conductivity is high; and under Regime 2–2–2 the negative association persists but is milder ($\tau \approx -0.09$). Moreover, the vine structure itself varies across regimes: each regime selects the Tree 1 pairings (and thus the Tree 2 conditioning) that best capture its own dependence pattern.

12.2 Joint Copula Analysis

Incorporation of Zero Inflated Rainfall Variables For each regime-specific dataset, we now append six-hourly zero-inflated precipitation series derived from the Copernicus satellite grid and spatially filtered to two target locations on the managed Noce River: the Malga reservoir (which regulates flow via upstream lake storage) and the Mezzo station (the nearest air-temperature and rainfall gauge). These combined sources capture both storage-controlled runoff and local meteorological variability for subsequent copula estimation. Specifically, for each joint regime $r \in \{1, \ldots, 8\}$, we add

$$\texttt{malga.tp}^{(r)},\ \texttt{malga.tp1}^{(r)},\ \texttt{malga.tp2}^{(r)},\quad \texttt{mezzo.tp}^{(r)},\ \texttt{mezzo.tp1}^{(r)},\ \texttt{mezzo.tp2}^{(r)},$$

where "tp" denotes the total precipitation contemporaneously every hour, and "tp1", "tp2" are the lags of one and two hours, respectively.

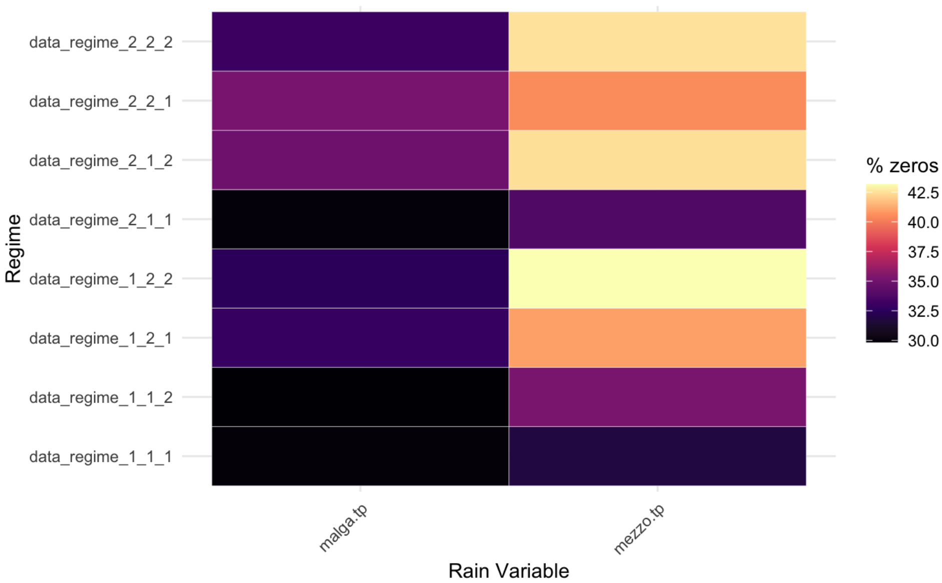


Figure 12.2: Heatmap of zero-inflation in hourly rainfall across eight regimes. Each cell shows the percentage of observations equal to zero for a given rain variable and regime, highlighting which regimes/variables exhibit the strongest zero-mass.

The exploratory pairs-plot of pseudo-observations, detailed summary tables for the fitted R-vine, and fitted contour plots are provided in the Appendix. Please see Appendix A for detailed results.

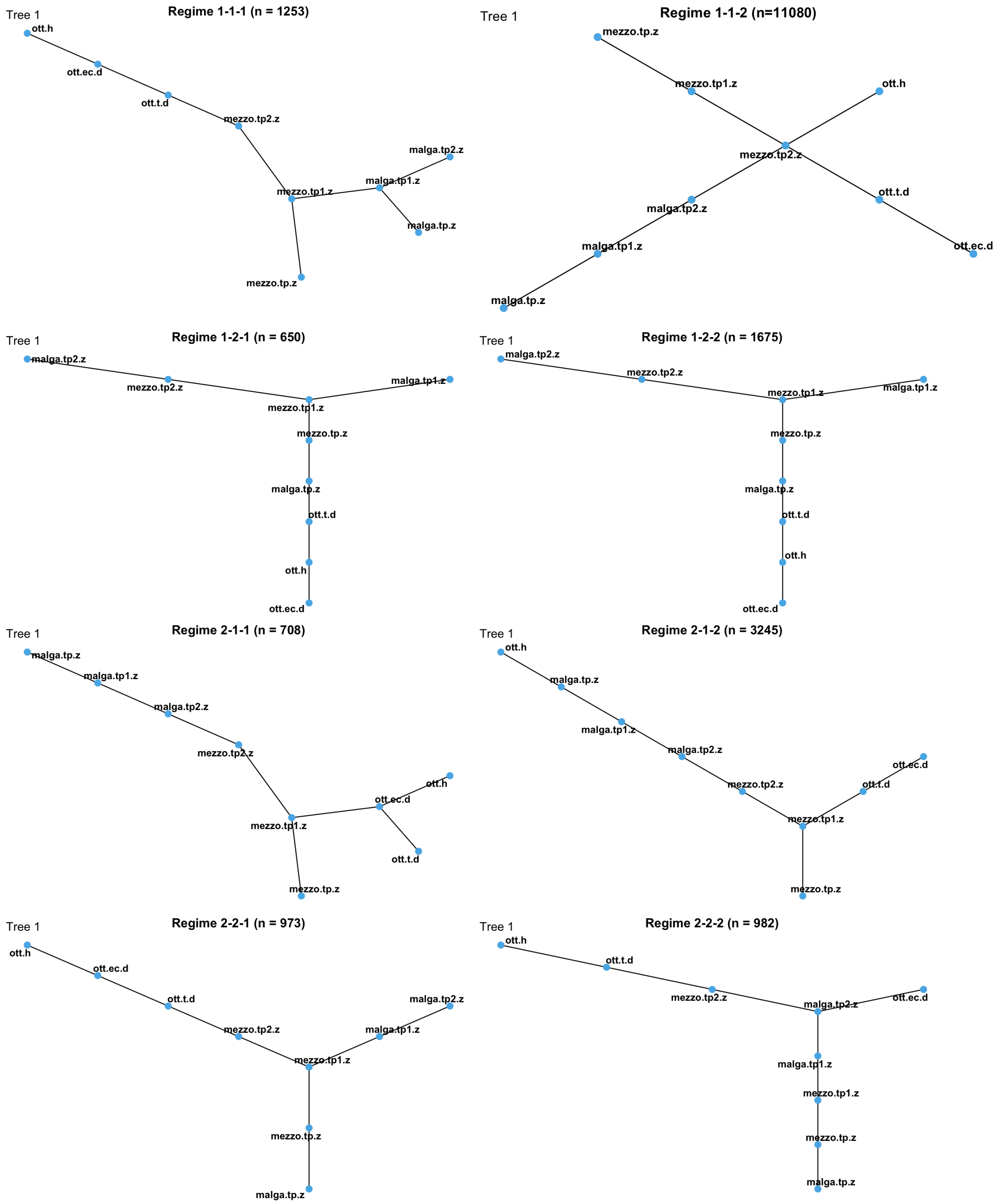


Figure 12.3: Tree 1 plots of joint copula analysis (with precipitation variables) for all eight regimes.

Edge	1-1-1 H-H-L	1-1-2 H-H-H	1-2-1 H-L-L	1-2-2 H-L-H	2-1-1 L-H-L	2-1-2 L-H-H	2-2-1 L-L-L	2-2-2 L-L-H
mezzo.tp – mezzo.tp1	✓	✓	✓	✓	✓	✓	✓	×
mezzo.tp1 – mezzo.tp2	✓	✓	✓	✓	✓	✓	✓	×
malga.tp1 – malga.tp2	✓	✓	×	×	✓	✓	✓	✓
malga.tp2 – mezzo.tp2	×	✓	✓	✓	✓	✓	×	✓
malga.tp1 – mezzo.tp1	✓	×	✓	✓	×	×	✓	✓
ott.ec.d – ott.h	✓	×	✓	✓	✓	×	✓	×
ott.ec.d – ott.t.d	✓	✓	×	×	✓	✓	✓	×
malga.tp – malga.tp1	✓	✓	×	×	✓	✓	×	×
malga.tp – mezzo.tp	×	×	✓	✓	×	×	✓	✓
mezzo.tp2 – ott.t.d	✓	✓	×	×	×	×	✓	✓
ott.h – ott.t.d	×	×	✓	✓	×	×	×	✓
malga.tp – ott.t.d	×	×	✓	✓	×	×	×	×
malga.tp – ott.h	×	×	×	×	×	✓	×	×
malga.tp2 – ott.ec.d	×	×	×	×	×	×	×	✓
mezzo.tp – mezzo.tp2	✓	×	×	×	×	×	×	×
mezzo.tp1 – ott.ec.d	×	×	×	×	✓	×	×	×
mezzo.tp1 – ott.t.d	×	×	×	×	×	✓	×	×
mezzo.tp2 – ott.h	×	✓	×	×	×	×	×	×

Table 6: Presence (✓) or absence (×) of each Tree 1 edge across the eight regimes from the joint copula analyses (with precipitation), sorted by decreasing frequency of occurrence.

Conclusions on OTT variables Regimes are defined by the states of *water level, temperature,* and *electrical conductivity.* The height–conductivity edge (`ott.ec.d -- ott.h`) is selected in every regime where conductivity is *low*, and otherwise drops out except in one case (high water, low temperature). The temperature–conductivity edge (`ott.ec.d -- ott.t.d`) is selected in all *high-temperature* regimes and, among *low-temperature* regimes, only when all three variables are *low.* The water–temperature edge (`ott.h -- ott.t.d`) appears only under *low temperature* (and not in every such regime) and disappears whenever temperature is *high.* Cross-site links to OTT variables are sparse and switch on only in narrow windows (e.g., `mezzo.tp2.z -- ott.t.d` when water level and temperature share the same state). Overall, across the eight regimes, the within-OTT edges occur in 5/8 (water–conductivity), 5/8 (temperature–conductivity), and 3/8 (water–temperature) cases, respectively.

13 Comparison of OTT (without precipitation) and Joint Copula Analyses (with precipitation)

In this section, we summarize the log-likelihood, model complexity, and information-criterion metrics for both the OTT-only and joint copula models across all eight regimes. Table 7 presents these results side by side, allowing for a direct comparison of fit quality (log-likelihood), number of parameters, and predictive performance penalties (AIC and BIC).

Table 7: OTT-only vs. Joint Copula Analysis across Eight Regimes

Regime	**Analysis**	**Log Likelihood**	**Num. Params.**	**AIC**	**BIC**
Regime 1-1-1 (n=1253)	OTT-only	-5468.45	81.65	11100.20	11519.32
	Joint	33114.80	299.83	-65629.94	-64090.85
Regime 1-1-2 (n=11080)	OTT-only	-42933.49	395.25	86657.49	89547.93
	Joint	318636.6	975.16	-635322.8	-628191.5
Regime 1-2-1 (n=650)	OTT-only	-3230.10	96.43	6653.07	7084.78
	Joint	17187.33	250.38	-33873.88	-32752.92
Regime 1-2-2 (n=1675)	OTT-only	-6674.53	148.09	13645.23	14448.40
	Joint	45963.52	392.88	-91141.27	-89010.43
Regime 2-1-1 (n=708)	OTT-only	-3589.18	27.35	7233.07	7357.87
	Joint	18447.79	107.39	-36680.80	-36190.83
Regime 2-1-2 (n=3245)	OTT-only	-14353.41	261.44	29229.69	30820.50
	Joint	84484.69	506.68	-167956.02	-164872.95
Regime 2-2-1 (n=973)	OTT-only	-4900.51	27.83	9856.67	9992.47
	Joint	23411.17	272.65	-46277.06	-44946.45
Regime 2-2-2 (n=982)	OTT-only	-4604.33	126.66	9461.98	10081.31
	Joint	25355.37	308.92	-50092.90	-48582.40

Across all regimes, the Joint models report much larger log-likelihoods and smaller AIC/BIC; however, these quantities are computed for a *higher-dimensional* random vector (OTT variables *plus* precipitation) than in the OTT-only fits. Hence, even though the sample size per regime is the same, the likelihoods are not on the same scale, and the information criteria are not directly comparable. For illustration, in Regime 1–1–1, the Joint fit shows $\ell : -5468 \to 33115$ and AIC $: \approx 11100 \to -65630$ with parameters $82 \to 300$; this reflects added flexibility for the enlarged system rather than superiority over the OTT-only model. We therefore refrain from calling the Joint model "better" and instead view it as complementary; its added value for the OTT variables should be judged from the regime-wise edge structure and, where possible, out-of-sample scores computed on the same OTT target.

14 Individual Copula Regression of OTT Variables on Rainfall

In this section, we fit a separate R-vine regression model for each OTT variable, treating it as the response and the six zero-inflated rainfall series as predictors. Specifically, for each regime we estimate vine-regression models of the form

$$\texttt{ott.h} \sim \texttt{malga.tp} + \texttt{mezzo.tp} + \texttt{malga.tp1} + \texttt{malga.tp2} + \texttt{mezzo.tp1} + \texttt{mezzo.tp2},$$

and similarly for `ott.t.d` and `ott.ec.d`. All models are selected by AIC over the parametric family set (including zero-inflated margins for the rainfall covariates), using `vinereg_zi()` method described in Sulejmani et al. (2025).
Thus, for each regime (e.g. 1–1–1) we obtain three separate vine-regression fits, one per OTT response, each capturing the dependence of OTT on the six rainfall series in a flexible pair-copula regression framework. The summary tables are provided below. The fitted contour plots are presented in Appendix B.

Table 8: Copula summary for `ott.h` across all regimes. H/L labels correspond to higher/lower mean regimes for (`ott.h`, `ott.t.d`, `ott.ec.d`).

Regime summary	**Tree**	**Edge**	**Conditioning**	**Copula**	**Par1**	**Par2**	τ
Regime 1–1–1 (H–H–L) (n=1253) Vine structure order: **1 2 5 4 7 6 3** cll: -4802.37 caic: 9650.31 cbic: 9767.26 edf: 22.8	1	(1,2)	—	Frank	-0.28	—	-0.03
	2	(1,5)	2	Frank	-0.27	—	-0.03
	3	(1,4)	5, 2	Frank	-0.25	—	-0.03
	4	(1,7)	4, 5, 2	t	-0.01	50.00	-0.01
	5	(1,6)	7, 4, 5, 2	Frank	-0.20	—	-0.02
	6	(1,3)	6, 7, 4, 5, 2	Clayton	0.01	—	0.01
Regime 1–2–1 (H–L–L) (n=650) Vine structure order: **1 7 4 6 2 3** cll: -2587 caic: 5200 cbic: 5258 edf: 12.9	1	(1,7)	—	Frank	-0.23	—	-0.03
	2	(1,4)	7	Frank	-0.22	—	-0.02
	3	(1,6)	4, 7	Frank	-0.21	—	-0.02
	4	(1,2)	6, 4, 7	Frank	-0.21	—	-0.02
	5	(1,3)	2, 6, 4, 7	Frank	-0.20	—	-0.02
Regime 1–2–2 (H–L–H) (n=1675) Vine structure order: **1 2 5 4 6 7 3** cll: -Inf caic: Inf cbic: Inf edf: 13.5	1	(1,2)	—	Frank	-0.14	—	-0.02
	2	(1,5)	2	Frank	-0.15	—	-0.02
	3	(1,4)	5, 2	Frank	-0.13	—	-0.01
	4	(1,6)	4, 5, 2	Frank	-0.13	—	-0.01
	5	(1,7)	6, 4, 5, 2	Clayton	0.01	—	0.00
	6	(1,3)	7, 6, 4, 5, 2	Frank	-0.13	—	-0.01
Regime 2–1–1 (L–H–L) (n=708) Vine structure order: **1 2 4 7 3** cll: -2147.44 caic: 4329.4 cbic: 4408.1 edf: 17.25	1	(1,2)	—	t	-0.04	25.42	-0.02
	2	(1,4)	2	Frank	-0.32	—	-0.04
	3	(1,7)	4, 2	t	0.05	20.14	0.03
	4	(1,3)	7, 4, 2	t	-0.03	11.09	-0.02
Regime 2–1–2 (L–H–H) (n=3245) Vine structure order: **1 5 2 4 3 7 6** cll: -13134.33 caic: 26305.6 cbic: 26418.3 edf: 18.5	1	(1,5)	—	Joe	1.00	—	0.00
	2	(1,2)	5	Frank	-0.19	—	-0.02
	3	(1,4)	2, 5	Frank	-0.19	—	-0.02
	4	(1,3)	4, 2, 5	Frank	-0.17	—	-0.02
	5	(1,7)	3, 4, 2, 5	Joe	1.00	—	0.00
	6	(1,6)	7, 3, 4, 2, 5	Frank	-0.16	—	-0.02
Regime 2–2–1 (L–L–L) (n=973) Vine structure order: **1 4 7 5 2 6** cll: -1416.52 caic: 2863.35 cbic: 2954.6 edf: 17.65	1	(1,4)	—	t	-0.04	15.12	-0.03
	2	(1,7)	4	Joe	1.01	—	0.01
	3	(1,5)	7, 4	Frank	-0.30	—	-0.03
	4	(1,2)	5, 7, 4	Frank	-0.20	—	-0.02
	5	(1,6)	2, 5, 7, 4	Joe	1.01	—	0.00
Regime 2–2–2 (L–L–H) (n=982) Vine structure order: **1 3 5 4 2 6** cll: -3232.9 caic: 6500 cbic: 6585 edf: 17.41	1	(1,3)	—	t	-0.02	30.05	-0.01
	2	(1,5)	3	Joe	1.00	—	0.00
	3	(1,4)	5, 3	Frank	-0.20	—	-0.02
	4	(1,2)	4, 5, 3	Frank	-0.18	—	-0.02
	5	(1,6)	2, 4, 5, 3	Joe	1.00	—	0.00

Variable mapping: 1 = `ott.h`, 2 = `malga.tp`, 3 = `mezzo.tp`, 4 = `malga.tp1`, 5 = `malga.tp2`, 6 = `mezzo.tp1`, 7 = `mezzo.tp2`.

Conclusions on rainfall influence for *ott.h*. Due to convergence and numerical difficulties for Regime 1–1–2 (n=11080), the vine regression analysis could not be completed within a reasonable time. Excluding Regime 1–1–2 and measuring influence by the magnitude of Kendall's τ for each pair-copula involving `ott.h`, the variable with the highest influence is `malga.tp1` **(4)**, which attains the largest and most consistent association (including the maximum observed $|\tau| = 0.04$ in Regime 2–1–1). The next strongest is `malga.tp` **(2)**, typically showing small negative dependence ($|\tau| \approx 0.02$–0.03). All other variables are weaker: `mezzo.tp` **(3)** and `malga.tp2` **(5)** are generally around $|\tau| \approx 0.01$–0.02, while `mezzo.tp1` **(6)** and `mezzo.tp2` **(7)** are often near independence (several Joe copulas with $\tau \approx 0$). Overall, across the considered regimes, the associations with **ott.h** are uniformly small ($|\tau| \leq 0.04$), indicating weak univariate effects relative to any multivariate structure. This outcome is consistent with the regime construction. Since precipitation patterns already define regimes, much of the precipitation effect has been absorbed at the regime level, leaving only weak within-regime precipitation dependence. In particular, in high-precipitation regimes, rainfall no longer explains water level variation because the regime definition already conditions this effect out.

Table 9: Regime-specific summaries for `ott.h` with precipitation variables

Regime	var	edf	cll	caic	cbic	p_value
111	ott.h	15.8	-4820.04	9671.65	9.75e+03	NA
111	malga.tp	1.0	3.79	-5.57	-4.42e-01	0.00592
111	malga.tp2	1.0	3.57	-5.14	-7.57e-03	0.00753
111	malga.tp1	1.0	3.07	-4.14	9.97e-01	0.01324
111	mezzo.tp2	2.0	3.38	-2.76	7.51e+00	0.03405
111	mezzo.tp1	1.0	1.89	-1.79	3.35e+00	0.05163
111	mezzo.tp	1.0	1.97	-1.94	3.19e+00	0.04708
121	ott.h	7.94	-2592.92	5201.707	5237.24	NA
121	mezzo.tp2	1.00	1.32	-0.643	3.83	0.104
121	malga.tp1	1.00	1.23	-0.462	4.01	0.117
121	mezzo.tp1	1.00	1.13	-0.266	4.21	0.132
121	malga.tp	1.00	1.11	-0.211	4.27	0.137
121	mezzo.tp	1.00	1.06	-0.124	4.35	0.145
122	ott.h	7.54	-Inf	Inf	Inf	NA
122	malga.tp	1.00	1.28	-0.564	4.86	0.1093
122	malga.tp2	1.00	1.42	-0.831	4.59	0.0925
122	malga.tp1	1.00	1.15	-0.291	5.13	0.1302
122	mezzo.tp1	1.00	1.12	-0.240	5.18	0.1345
122	mezzo.tp2	1.00	1.26	-0.527	4.90	0.1119
122	mezzo.tp	1.00	1.06	-0.116	5.31	0.1457
211	ott.h	10.3	-2191.12	4402.7	4449.52	NA
211	malga.tp	2.0	7.56	-11.1	-2.00	5.19e-04
211	malga.tp1	1.0	2.65	-3.3	1.26	2.13e-02
211	mezzo.tp2	2.0	12.15	-20.3	-11.17	5.29e-06
211	mezzo.tp	2.0	21.31	-38.6	-29.50	5.54e-10
212	ott.h	12.5	-13165.36	26355.73	26431.835	NA
212	malga.tp2	1.0	11.03	-20.07	-13.984	2.63e-06
212	malga.tp	1.0	5.01	-8.01	-1.927	1.56e-03
212	malga.tp1	1.0	4.71	-7.43	-1.342	2.14e-03
212	mezzo.tp	1.0	3.58	-5.17	0.915	7.42e-03
212	mezzo.tp2	1.0	3.45	-4.90	1.184	8.61e-03
212	mezzo.tp1	1.0	3.24	-4.48	1.609	1.09e-02
221	ott.h	11.7	-1439.10	2901.5123	2958.40	NA
221	malga.tp1	2.0	14.87	-25.7463	-15.99	3.47e-07
221	mezzo.tp2	1.0	2.83	-3.6590	1.22	1.74e-02
221	malga.tp2	1.0	2.62	-3.2341	1.65	2.21e-02
221	malga.tp	1.0	1.21	-0.4288	4.45	1.19e-01
221	mezzo.tp1	1.0	1.05	-0.0937	4.79	1.48e-01
222	ott.h	11.4	-3250.96	6524.766	6580.60	NA
222	mezzo.tp	2.0	8.76	-13.525	-3.75	0.000156
222	malga.tp2	1.0	5.10	-8.205	-3.32	0.001400
222	malga.tp1	1.0	1.59	-1.179	3.71	0.074610
222	malga.tp	1.0	1.29	-0.583	4.31	0.108017
222	mezzo.tp1	1.0	1.32	-0.640	4.25	0.104217

Table 10: Copula summary for `ott.t.d` across all regimes. H/L labels correspond to higher/lower mean regimes for (`ott.h`, `ott.t.d`, `ott.ec.d`).

Regime summary	Tree	Edge	Conditioning	Copula	Par1	Par2	τ
Regime 1–1–1 (H–H–L) (n=1253) Vine structure order: **1 2 7 5 6** cll: -1305.37 caic: 2639 cbic: 2712 edf: 14.3	1	(1,2)	—	BB8	1.74	0.86	0.18
	2	(1,7)	2	Frank	0.32	—	0.04
	3	(1,5)	7, 2	Clayton	0.08	—	-0.04
	4	(1,6)	5, 7, 2	Frank	0.32	—	0.04
Regime 1–2–1 (H–L–L) (n=650) Vine structure order: **1 4 2 5** cll: -1033.9 caic: 2091 cbic: 2144 edf: 11.81	1	(1,4)	—	Tawn	0.30	1.00, 2.04	0.21
	2	(1,2)	4	Joe	1.05	—	0.03
	3	(1,5)	2, 4	Clayton	0.09	—	-0.05
Regime 1–2–2 (H–L–H) (n=1675) Vine structure order: **1 6 7 5 4 2 3** cll: -2765 caic: 5560 cbic: 5641 edf: 14.9	1	(1,6)	—	Tawn	0.30	0.76, 1.70	0.16
	2	(1,7)	6	Clayton	0.17	—	-0.08
	3	(1,5)	7, 6	Clayton	0.12	—	0.05
	4	(1,4)	5, 7, 6	Clayton	0.05	—	-0.02
	5	(1,2)	4, 5, 7, 6	Clayton	0.06	—	0.03
	6	(1,3)	2, 4, 5, 7, 6	Clayton	0.04	—	-0.02
Regime 2–1–1 (L–H–L) (n=708) Vine structure order: **1 6** cll: -807 caic: 1633.5 cbic: 1677.6 edf: 9.7	1	(1,6)	—	t	0.08	6.75	0.05
Regime 2–1–2 (L–H–H) (n=3245) Vine structure order: **1 5 7 3** cll: -3672 caic: 7377 cbic: 7473 edf: 15.7	1	(1,5)	—	BB8	1.32	0.95	-0.12
	2	(1,7)	5	t	0.01	11.62	0.00
	3	(1,3)	7, 5	Clayton	0.03	—	-0.01
Regime 2–2–1 (L–L–L) (n=973) Vine structure order: **1 7** cll: -1427 caic: 2875 cbic: 2926 edf: 10.43	1	(1,7)	—	Tawn	0.42	0.33, 2.79	-0.18
Regime 2–2–2 (L–L–H) (n=982) Vine structure order: **1 7 4 5** cll: -1418 caic: 2856 cbic: 2904 edf: 10	1	(1,7)	—	Frank	-1.07	—	-0.12
	2	(1,4)	7	Clayton	0.07	—	0.03
	3	(1,5)	4, 7	Joe	1.04	—	-0.03

Variable mapping: 1 = `ott.t.d`, 2 = `malga.tp`, 3 = `mezzo.tp`, 4 = `malga.tp1`, 5 = `malga.tp2`, 6 = `mezzo.tp1`, 7 = `mezzo.tp2`.

Conclusions for `ott.t.d`. Considering all pair-copulas that involve variable 1 and excluding Regime 1–1–2 (no fitted pairs), the most influential variables by sample-size–weighted $|\tau|$ are `mezzo.tp1` **(6)** ($|\tau| \approx 0.10$, typically positive; peak $\tau = +0.16$ in Regime 1–2–2), followed by `malga.tp` **(2)** ($|\tau| \approx 0.08$, peak $\tau = +0.18$ in Regime 1–1–1) and `malga.tp2` **(5)** ($|\tau| \approx 0.07$, typically negative; peak $\tau = -0.12$ in Regime 2–1–2). `malga.tp1` **(4)** shows a moderate average effect ($|\tau| \approx 0.06$) but attains the *single strongest* regime-level link ($\tau = +0.21$ in Regime 1–2–1). `mezzo.tp2` **(7)** is mostly negative ($|\tau| \approx 0.06$; $\tau = -0.18$ in Regime 2–2–1), while `mezzo.tp` **(3)** is weak (near zero). Overall, dependence for `ott.t.d` is stronger than for `ott.h`, with the largest effects captured by Tawn/BB8 copulas, indicating some asymmetric tail behavior.

Table 11: Regime-specific summaries for `ott.t.d` with precipitation variables

Regime	var	edf	cll	caic	cbic	p_value
111	ott.t.d	9.3	-1360.27	2739.139	2786.885	NA
111	malga.tp	2.0	49.16	-94.329	-84.062	4.45e-22
111	mezzo.tp2	1.0	1.17	-0.334	4.800	1.27e-01
111	malga.tp2	1.0	3.41	-4.825	0.309	8.99e-03
111	mezzo.tp1	1.0	1.19	-0.377	4.756	1.23e-01
121	ott.t.d	6.82	-1060.73	2135.09	2165.61	NA
121	malga.tp1	3.00	22.62	-39.24	-25.81	8.24e-10
121	malga.tp	1.00	1.72	-1.43	3.04	6.39e-02
121	malga.tp2	1.00	2.58	-3.17	1.31	2.30e-02
122	ott.t.d	6.89	-2829.64	5673.06	5710.44	NA
122	mezzo.tp1	3.00	26.51	-47.01	-30.74	1.82e-11
122	mezzo.tp2	1.00	21.15	-40.30	-34.87	7.84e-11
122	malga.tp2	1.00	10.61	-19.21	-13.79	4.11e-06
122	malga.tp1	1.00	2.15	-2.29	3.13	3.83e-02
122	malga.tp	1.00	2.40	-2.79	2.63	2.85e-02
122	mezzo.tp	1.00	1.53	-1.05	4.37	8.05e-02
211	ott.t.d	7.65	-812.46	1640.22	1675.11	NA
211	mezzo.tp1	2.00	5.32	-6.64	2.48	0.00489
212	ott.t.d	10.7	-3749.00	7519.455	7584.73	NA
212	malga.tp2	2.0	69.40	-134.794	-122.62	7.27e-31
212	mezzo.tp2	2.0	5.42	-6.841	5.33	4.42e-03
212	mezzo.tp	1.0	1.20	-0.409	5.68	1.21e-01
221	ott.t.d	7.44	-1452.3	2919.4	2955.7	NA
221	mezzo.tp2	3.00	24.7	-43.5	-28.8	1.03e-10
222	ott.t.d	6.93	-1435.68	2885.23	2919.13	NA
222	mezzo.tp2	1.00	13.23	-24.46	-19.57	2.69e-07
222	malga.tp1	1.00	2.44	-2.88	2.01	2.71e-02
222	malga.tp2	1.00	1.88	-1.77	3.12	5.23e-02

Table 12: Copula summary for `ott.ec.d` across all regimes. H/L labels correspond to higher/lower mean regimes for (`ott.h`, `ott.t.d`, `ott.ec.d`).

Regime summary	**Tree**	**Edge**	**Conditioning**	**Copula**	**Par1**	**Par2**	τ
Regime 1–1–1 (H–H–L) (n=1253)	1	(1,3)	—	Clayton	0.16	—	-0.07
Vine structure order: **1 3 5 6**	2	(1,5)	3	Clayton	0.06	—	0.03
cll: -5249.08	3	(1,6)	5, 3	Frank	-0.32	—	-0.04
caic: 10525.85							
cbic: 10596.89							
edf: 13.9							
Regime 1–2–1 (H–L–L) (n=650)	1	(1,5)	—	Gumbel	1.08	—	0.07
Vine structure order: **1 5**							
cll: -2662.60							
caic: 5354.75							
cbic: 5420.88							
edf: 14.8							
Regime 1–2–2 (H–L–H) (n=1675)	1	(1,5)	—	Clayton	0.11	—	0.05
Vine structure order: **1 5 4**	2	(1,4)	5	Clayton	0.04	—	0.02
cll: -6036.11							
caic: 12091.11							
cbic: 12142.34							
edf: 9.5							
Regime 2–1–1 (L–H–L) (n=708)	1	(1,3)	—	Joe	1.13	—	-0.07
Vine structure order: **1 3 5 6**	2	(1,5)	3	Clayton	0.15	—	0.07
cll: -2910.24	3	(1,6)	5, 3	Frank	-0.72	—	-0.08
caic: 5850.88							
cbic: 5920.25							
edf: 15.2							
Regime 2–1–2 (L–H–H) (n=3245)	1	(1,7)	—	t	-0.00	6.21	-0.00
Vine structure order: **1 7 6 3**	2	(1,6)	7	Clayton	0.03	—	0.01
cll: -11755.37	3	(1,3)	6, 7	Joe	1.02	—	-0.01
caic: 23559.04							
cbic: 23706.01							
edf: 24.2							
Regime 2–2–1 (L–L–L) (n=973)	1	(1,3)	—	t	0.12	6.61	0.08
Vine structure order: **1 3 5 6 7**	2	(1,5)	3	Frank	0.42	—	0.05
cll: -3955.16	3	(1,6)	5, 3	Joe	1.09	—	-0.05
caic: 7935.00	4	(1,7)	6, 5, 3	Joe	1.08	—	0.04
cbic: 7995.23							
edf: 12.3							
Regime 2–2–2 (L–L–H) (n=982)	1	(1,7)	—	t	0.06	5.55	0.04
Vine structure order: **1 7 3 6**	2	(1,3)	7	Gumbel	1.03	—	-0.03
cll: -3573.06	3	(1,6)	3, 7	Joe	1.12	—	0.06
caic: 7181.42							
cbic: 7267.75							
edf: 17.7							

Variable mapping: 1 = `ott.ec.d`, 2 = `malga.tp`, 3 = `mezzo.tp`, 4 = `malga.tp1`, 5 = `malga.tp2`, 6 = `mezzo.tp1`, 7 = `mezzo.tp2`.

Conclusions for `ott.ec.d`. Excluding Regime 1–1–2, the variables with the largest influence on `ott.ec.d` (by sample-size–weighted $|\tau|$ over all pair-copulas involving variable 1) are **`malga.tp2` (5)** ($|\tau| \approx 0.05$, mostly positive; max $|\tau| = 0.07$), followed by **`mezzo.tp` (3)** ($|\tau| \approx 0.039$, typically slightly negative; max $|\tau| = 0.08$ in Regime 2–2–1) and **`mezzo.tp1` (6)** ($|\tau| \approx 0.034$, usually negative; max $|\tau| = 0.08$ in Regime 2–1–1). Remaining links are weak: `malga.tp1` (4) appears once with $|\tau| \approx 0.02$, and `mezzo.tp2` (7) is small and mostly positive ($|\tau| \approx 0.015$, peak $|\tau| = 0.04$). `malga.tp` (2) is not selected in any regime. Overall, dependence levels are modest ($|\tau| \leq 0.08$); stronger pairs use Clayton/Gumbel/t/Frank copulas, while Joe copulas occur near independence.

Table 13: Regime-specific summaries for `ott.ec.d` with precipitation variables

Regime	var	edf	cll	caic	cbic	p_value
111	ott.ec.d	10.8	-5263.75	10549.17	10604.81	NA
111	mezzo.tp	1.0	11.13	-20.26	-15.13	2.38e-06
111	malga.tp2	1.0	2.38	-2.77	2.37	2.90e-02
111	mezzo.tp1	1.0	1.15	-0.30	4.83	1.29e-01
121	ott.ec.d	13.8	-2665.63	5358.79	5420.45	NA
121	malga.tp2	1.0	3.02	-4.05	0.43	0.0139
122	ott.ec.d	7.45	-6045.16	12105.213	12145.60	NA
122	malga.tp2	1.00	7.81	-13.611	-8.19	7.78e-05
122	malga.tp1	1.00	1.25	-0.494	4.93	1.14e-01
211	ott.ec.d	12.2	-2926.06	5876.53	5932.206	NA
211	mezzo.tp	1.0	9.02	-16.05	-11.483	2.16e-05
211	malga.tp2	1.0	2.93	-3.85	0.709	1.55e-02
211	mezzo.tp1	1.0	3.87	-5.74	-1.181	5.39e-03
212	ott.ec.d	20.2	-11781.88	23604.072	23726.69	NA
212	mezzo.tp2	2.0	23.24	-42.474	-30.30	8.10e-11
212	mezzo.tp1	1.0	1.42	-0.845	5.24	9.17e-02
212	mezzo.tp	1.0	1.85	-1.709	4.38	5.41e-02
221	ott.ec.d	7.34	-3976.95	7968.58	8004.40	NA
221	mezzo.tp1	2.00	9.76	-15.52	-5.76	5.78e-05
221	malga.tp2	1.00	1.52	-1.04	3.85	8.15e-02
221	mezzo.tp	1.00	9.00	-16.00	-11.12	2.21e-05
221	mezzo.tp2	1.00	1.51	-1.03	3.86	8.20e-02
222	ott.ec.d	13.7	-3587.36	7202.042	7268.81	NA
222	mezzo.tp2	2.0	8.87	-13.746	-3.97	0.00014
222	mezzo.tp	1.0	1.38	-0.759	4.13	0.09673
222	mezzo.tp1	1.0	4.06	-6.114	-1.22	0.00439

15 Overall Conclusions from the Copula Analyses

Taken together, the three complementary copula analyses reveal that dependence structures among the OTT variables and their links to precipitation are both regime-dependent and generally weak to modest in magnitude. The OTT-only models show that internal relationships among height, temperature, and conductivity are strongest when height and temperature are high but conductivity is low, while mixed-state regimes display near-independence. The joint copula models demonstrate clear statistical gains from incorporating precipitation, yielding much higher likelihoods and substantially lower information criteria across all regimes, though at the cost of larger parameter counts.

Finally, across regimes (excluding 1–1–2), the vine regressions show that rainfall influences are *variable-specific* rather than uniformly dominated by one site. For `ott.h`, `malga.tp1` has the largest and most consistent association (max $|\tau| \approx 0.04$), with `malga.tp` next and the `mezzo` gauges generally weak or near-independent. For `ott.t.d`, however, the dominant driver is `mezzo.tp1` (sample-size weighted $|\tau| \approx 0.10$, peak $\tau \approx +0.16$), followed by `malga.tp` and `malga.tp2`; `malga.tp1` shows a moderate average effect but attains the single strongest regime-level link ($\tau \approx +0.21$). For `ott.ec.d`, `malga.tp2` is most influential ($|\tau| \approx 0.05$), with `mezzo.tp` and `mezzo.tp1` also contributing; `malga.tp` is not selected and remaining links are weak. Overall, pairwise dependencies are modest (roughly $|\tau| \leq 0.04$ for `ott.h`, $|\tau| \leq 0.08$ for `ott.ec.d`, and $|\tau| \approx 0.10$ for `ott.t.d`), consistent with the regime definitions absorbing much of the precipitation signal.

Appendices

A Joint Copula Analyses

To better understand the dependence structure in our data, we transformed the datasets into pseudo-observations, scaling the values between 0 and 1 to place them on the copula scale. For each dataset, we then constructed a matrix of pairwise plots: histograms along the diagonal to show the marginal distributions, scatter plots in the upper panels with Kendall's τ reported as a measure of monotonic dependence, and contour plots in the lower panels based on Gaussian-transformed margins to illustrate the joint density patterns.

A.1 Regime 1-1-1 (n=1253)

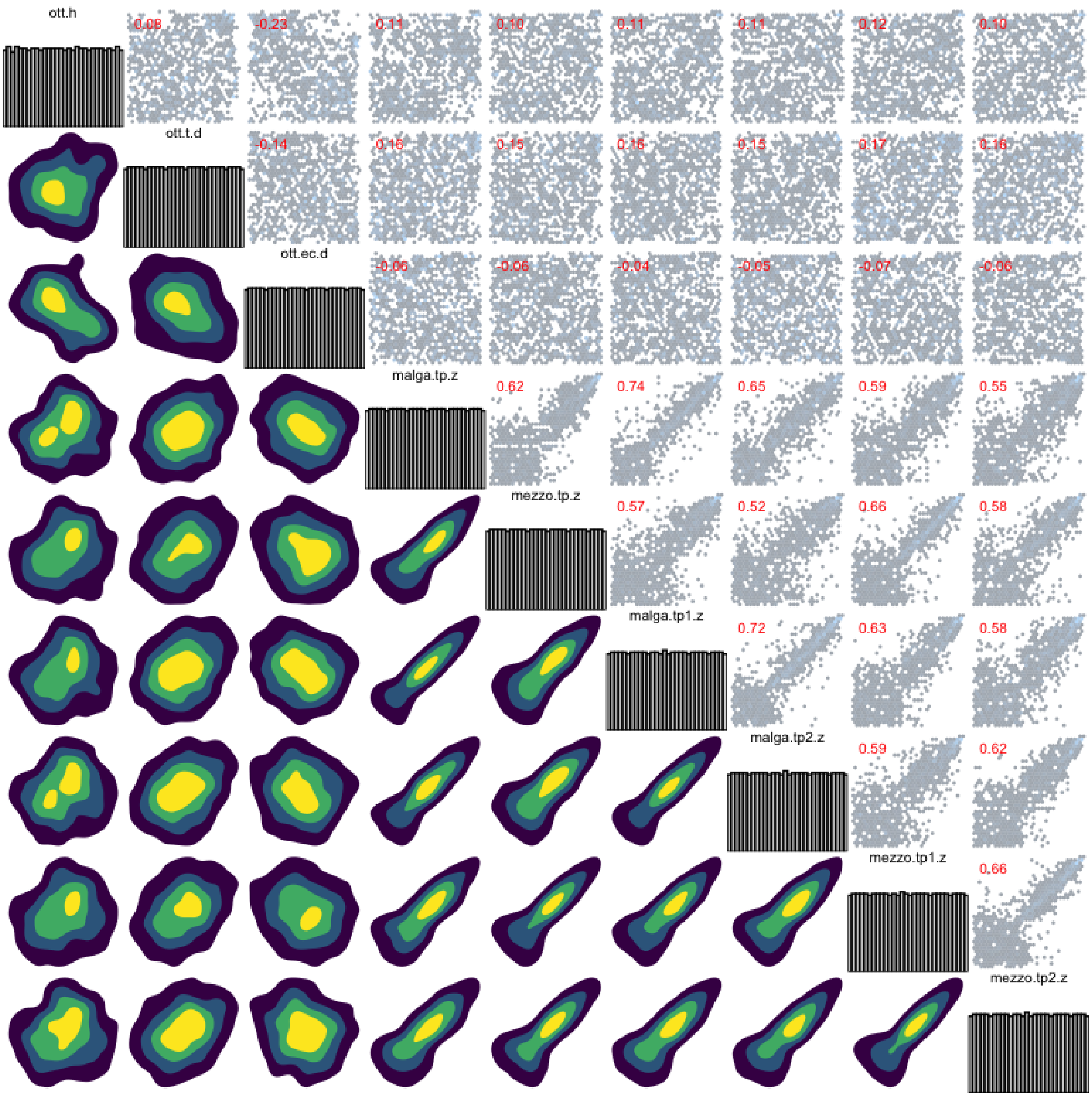


Figure A.1: Pairwise scatterplot matrix of pseudo-observations for regime 1–1–1. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 14: Pair–copula family selection summary for Regime 1-1-1 (n=1253). All tree–edge combinations with non–independence families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(7,6)	—	d,d	tll	0	[30x30 grid]	36.64	0.69	1208.70
1	2	(4,6)	—	d,d	tll	0	[30x30 grid]	38.92	0.69	1227.92
1	3	(1,3)	—	c,c	frank	0	–2.18	1.00	– 0.23	80.95
1	4	(6,8)	—	d,d	tll	0	[30x30 grid]	24.94	0.60	857.77
1	5	(3,2)	—	c,c	frank	0	–1.35	1.00	– 0.15	27.11
1	6	(5,8)	—	d,d	tll	0	[30x30 grid]	29.71	0.63	1074.48
1	7	(2,9)	—	c,d	frank	0	1.67	1.00	0.18	38.85
1	8	(8,9)	—	d,d	tll	0	[30x30 grid]	31.06	0.64	1090.90
3	2	(4,5)	(8,6)	d,d	tll	0	[30x30 grid]	38.43	0.42	392.02
3	3	(1,9)	(2,3)	c,d	joe	0	1.13	1.00	0.07	21.42
3	6	(5,2)	(9,8)	d,c	gumbel	0	1.05	1.00	0.05	3.63
4	4	(6,2)	(9,5,8)	d,c	joe	0	1.09	1.00	0.05	6.58
5	1	(7,9)	(5,4,8,6)	d,d	tll	0	[30x30 grid]	42.10	0.36	315.82

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

Conclusions for Regime 1–1–1. The pair–copula family selection results for Regime 1–1–1 highlight a clear contrast between the rainfall series and the OTT variables. The Malga and Mezzo precipitation variables, especially their lagged versions, exhibit extreme dependence with Kendall 's τ values around 0.6–0.7, reflecting high redundancy and persistence within rainfall records. The flexible nonparametric TLL copulas mostly capture these strong links. By contrast, the OTT variables (`ott.h`, `ott.t.d`, `ott.ec.d`) show only weak associations with one another, for example a modest negative dependence between water level and conductivity ($\tau \approx -0.23$) and between temperature and conductivity ($\tau \approx -0.15$). Direct OTT–rainfall links are also weak, with the strongest being temperature and `mezzo.tp2` ($\tau \approx 0.18$), while most others are close to independence. Overall, in this regime, the dependence structure is dominated by strong internal correlations among precipitation series, whereas OTT variables display only small pairwise associations with each other and with rainfall.

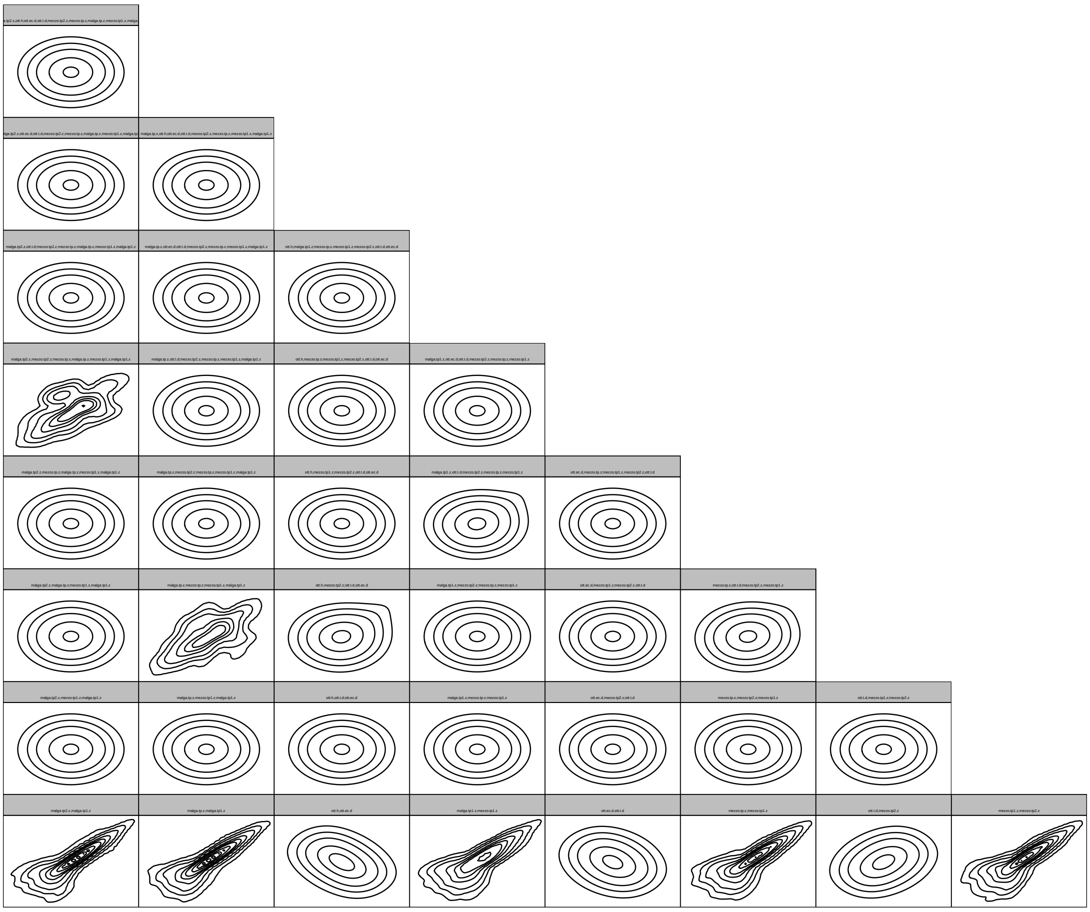


Figure A.2: Fitted pair-copula contour plots for all variable pairs in Regime 1–1–1, shown on the Gaussian scale.

A.2 Regime 1-1-2 (n=11080)

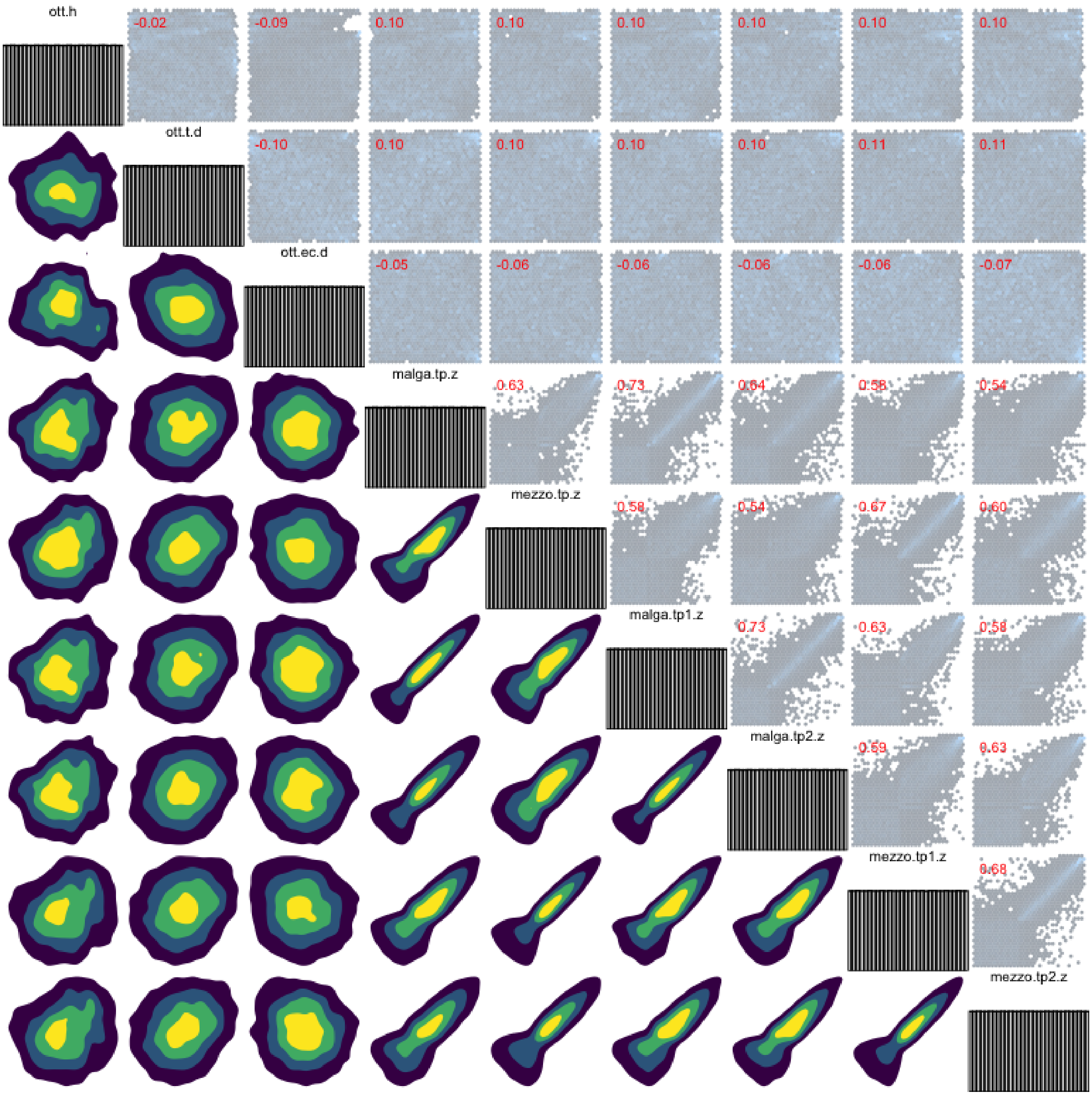


Figure A.3: Pairwise scatterplot matrix of pseudo-observations for regime 1–1–2. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 15: Pair–copula family selection summary for Regime 1-1-2 (n=11080). All tree–edge combinations with non–independence families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(1,9)	–	c,d	bb8	0	(1.86, 0.67)	2.00	0.12	191.91
1	2	(3,2)	–	c,c	bb8	90	(1.66, 0.70)	2.00	− 0.10	129.61
1	3	(4,6)	–	d,d	tll	0	[30×30 grid]	118.33	0.70	11 097.93
1	4	(2,9)	–	c,d	tawn	180	(0.30, 1.00, 1.71)	3.00	0.18	231.37
1	5	(6,7)	–	d,d	tll	0	[30×30 grid]	121.90	0.70	10 671.09
1	6	(5,8)	–	d,d	tll	0	[30×30 grid]	97.36	0.65	9698.31
1	7	(7,9)	–	d,d	tll	0	[30×30 grid]	77.85	0.61	7786.39
1	8	(8,9)	–	d,d	tll	0	[30×30 grid]	102.92	0.65	9584.56
2	1	(1,2)	–	c,c	frank	0	(−0.37)	1.00	− 0.04	21.92
2	2	(3,9)	–	c,d	clayton	270	(0.16)	1.00	− 0.07	116.10
2	4	(2,8)	–	c,d	t	0	(0.03, 9.29)	2.00	0.02	45.37
3	1	(1,3)	(2,9)	c,c	tll	0	[30×30 grid]	108.49	− 0.08	843.39
3	4	(2,5)	(8,9)	c,d	t	0	(0.02, 10.24)	2.00	0.01	36.63
3	5	(6,8)	(9,7)	d,d	tll	0	[30×30 grid]	133.17	0.41	3138.91
4	1	(1,8)	(3,2,9)	c,d	clayton	180	(0.06)	1.00	0.03	11.30
4	4	(2,7)	(5,8,9)	c,d	t	0	(−0.00, 12.79)	2.00	− 0.00	20.34
5	1	(1,5)	(8,3,2,9)	c,d	clayton	180	(0.04)	1.00	0.02	6.08
5	3	(4,5)	(8,9,7,6)	d,d	tll	0	[30×30 grid]	114.68	0.43	4128.72
5	4	(2,6)	(7,5,8,9,3)	c,d	t	0	(0.02, 22.36)	2.00	0.01	10.69
6	1	(1,7)	(5,8,3,2,9)	c,d	clayton	180	(0.05)	1.00	0.03	10.53
6	3	(4,2)	(5,8,9,7,6)	d,c	gumbel	0	(1.02)	1.00	0.02	6.03
7	1	(1,6)	(7,5,8,3,2,9)	c,d	frank	0	(0.20)	1.00	0.02	4.86

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

Conclusions for Regime 1–1–2 (OTT variables). For this regime, the OTT variables exhibit only weak mutual dependence: conductivity with temperature shows $\tau \approx -0.10$, while height with conductivity has $\tau \approx -0.08$. Other OTT–rainfall edges are negligible (mostly $\tau \approx 0$). Overall, OTT variables appear nearly independent in this regime, with no strong pairwise associations.

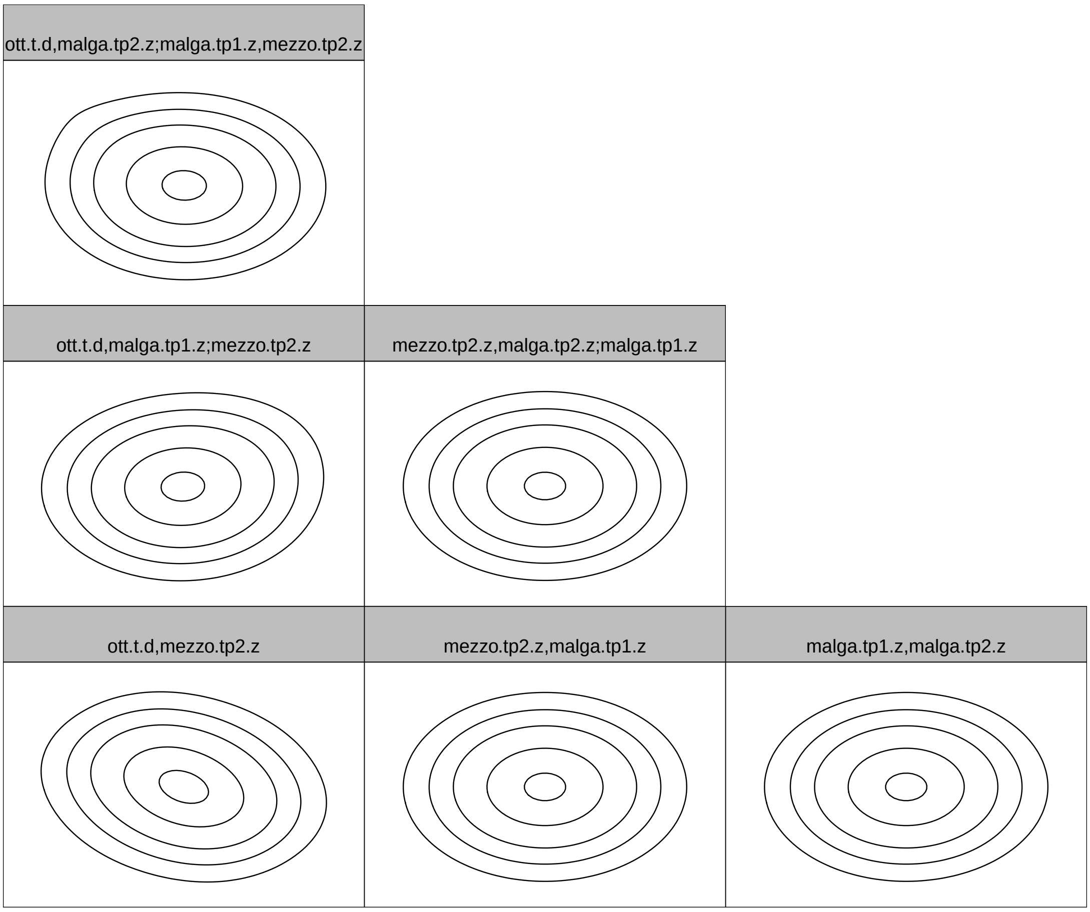


Figure A.4: Fitted pair-copula contour plots for all variable pairs in Regime 1–1–2.

A.3 Regime 1-2-1 (n=650)

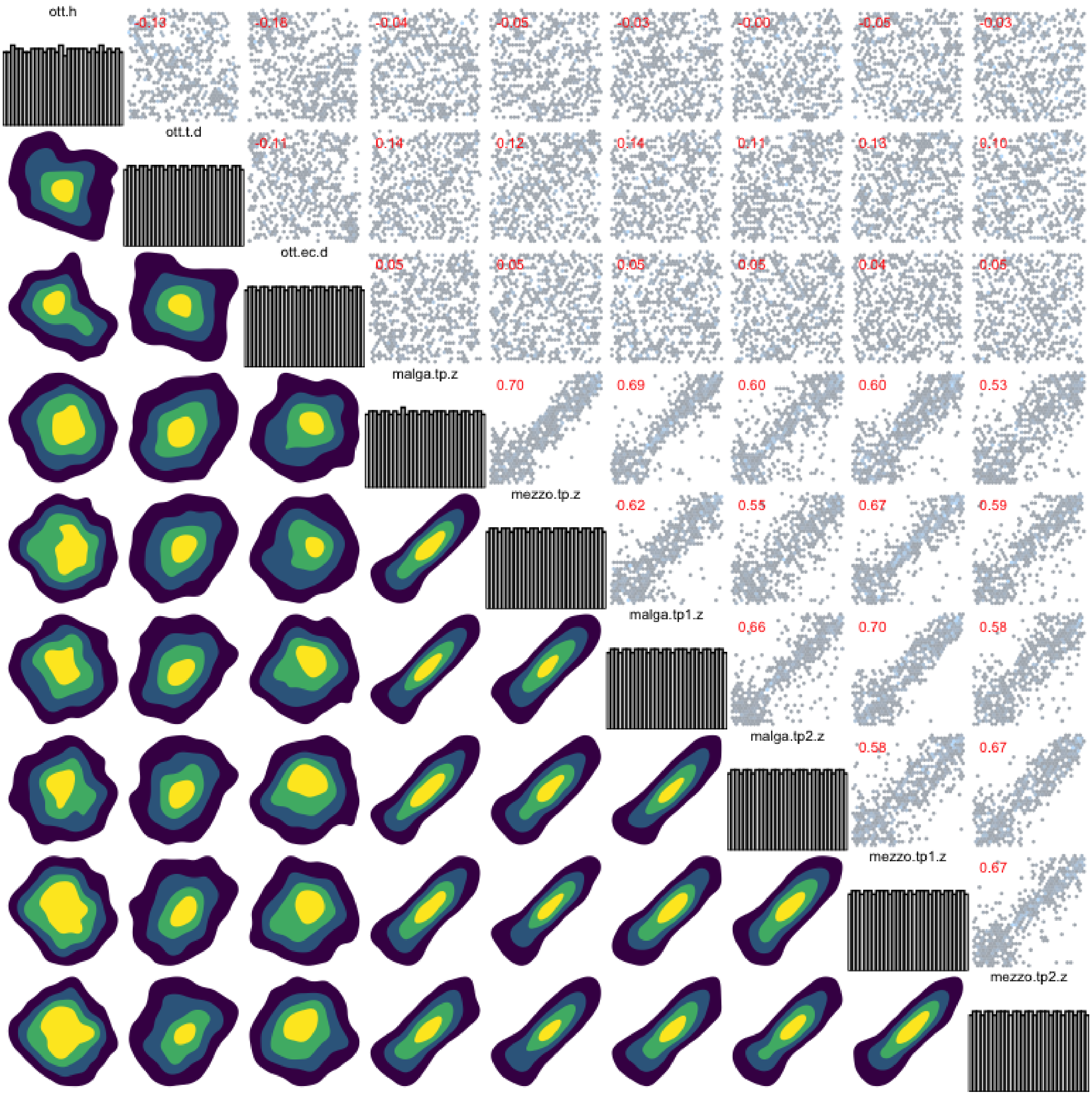


Figure A.5: Pairwise scatterplot matrix of pseudo-observations for regime 1–2–1. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 16: Pair–copula family selection summary for Regime 1-2-1 (n=650). All tree–edge combinations with non-independent families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(3,1)	—	c,c	bb8	90	(1.86, 0.85)	2.00	−0.20	38.79
1	2	(1,2)	—	c,c	clayton	270	0.33	1.00	−0.14	17.06
1	3	(2,4)	—	c,d	frank	0	1.35	1.00	0.15	15.83
1	4	(4,5)	—	d,d	tll	0	[30x30 grid]	26.74	0.65	477.05
1	5	(6,8)	—	d,d	tll	0	[30x30 grid]	22.99	0.63	502.83
1	6	(5,8)	—	d,d	tll	0	[30x30 grid]	33.70	0.62	474.84
1	7	(7,9)	—	d,d	tll	0	[30x30 grid]	25.74	0.63	460.97
1	8	(8,9)	—	d,d	tll	0	[30x30 grid]	33.39	0.62	458.07
2	1	(3,2)	(1)	c,c	gaussian	0	–0.25	1.00	−0.16	19.92
3	1	(3,4)	(2,1)	c,d	frank	0	0.67	1.00	0.07	3.70
3	4	(4,6)	(8,5)	d,d	tll	0	[30x30 grid]	27.89	0.39	256.87
4	5	(6,7)	(9,5,8)	d,d	tll	0	[30x30 grid]	30.99	0.41	252.18

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

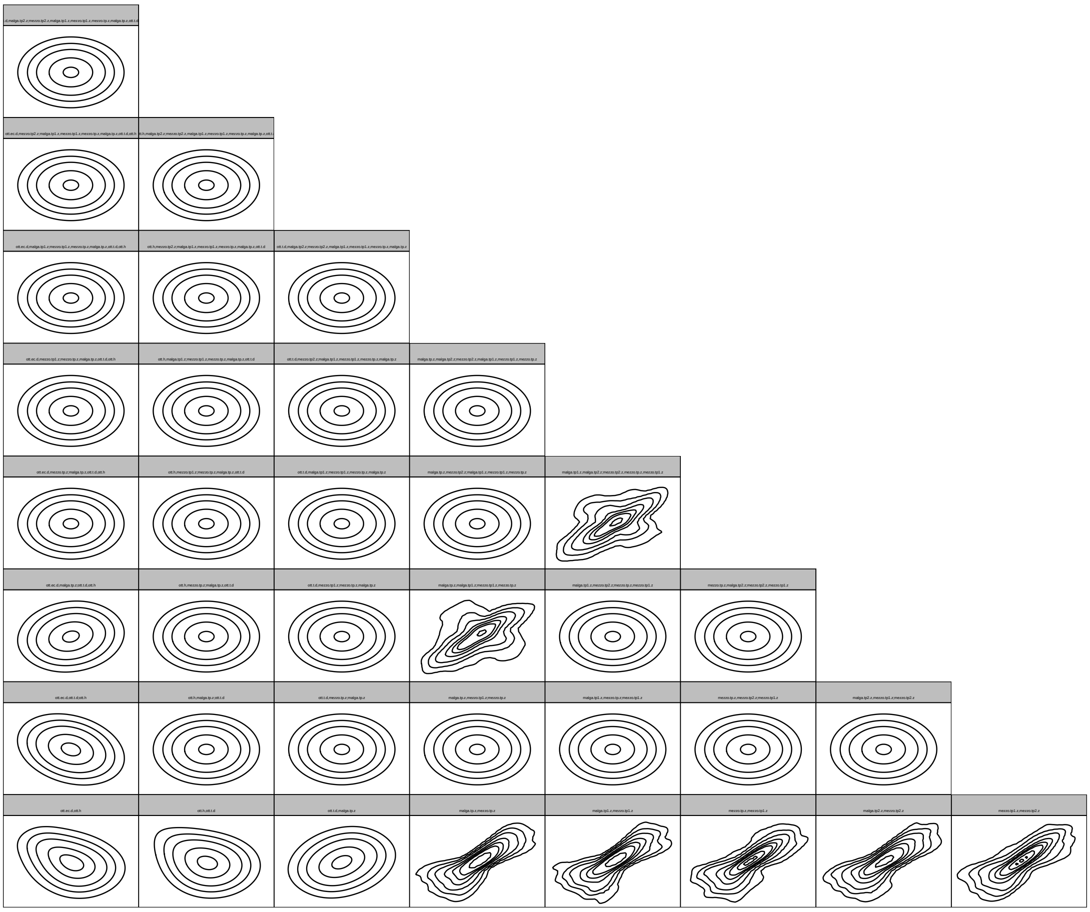


Figure A.6: Fitted pair-copula contour plots for all variable pairs in Regime 1–2–1.

A.4 Regime 1-2-2 (n=1675)

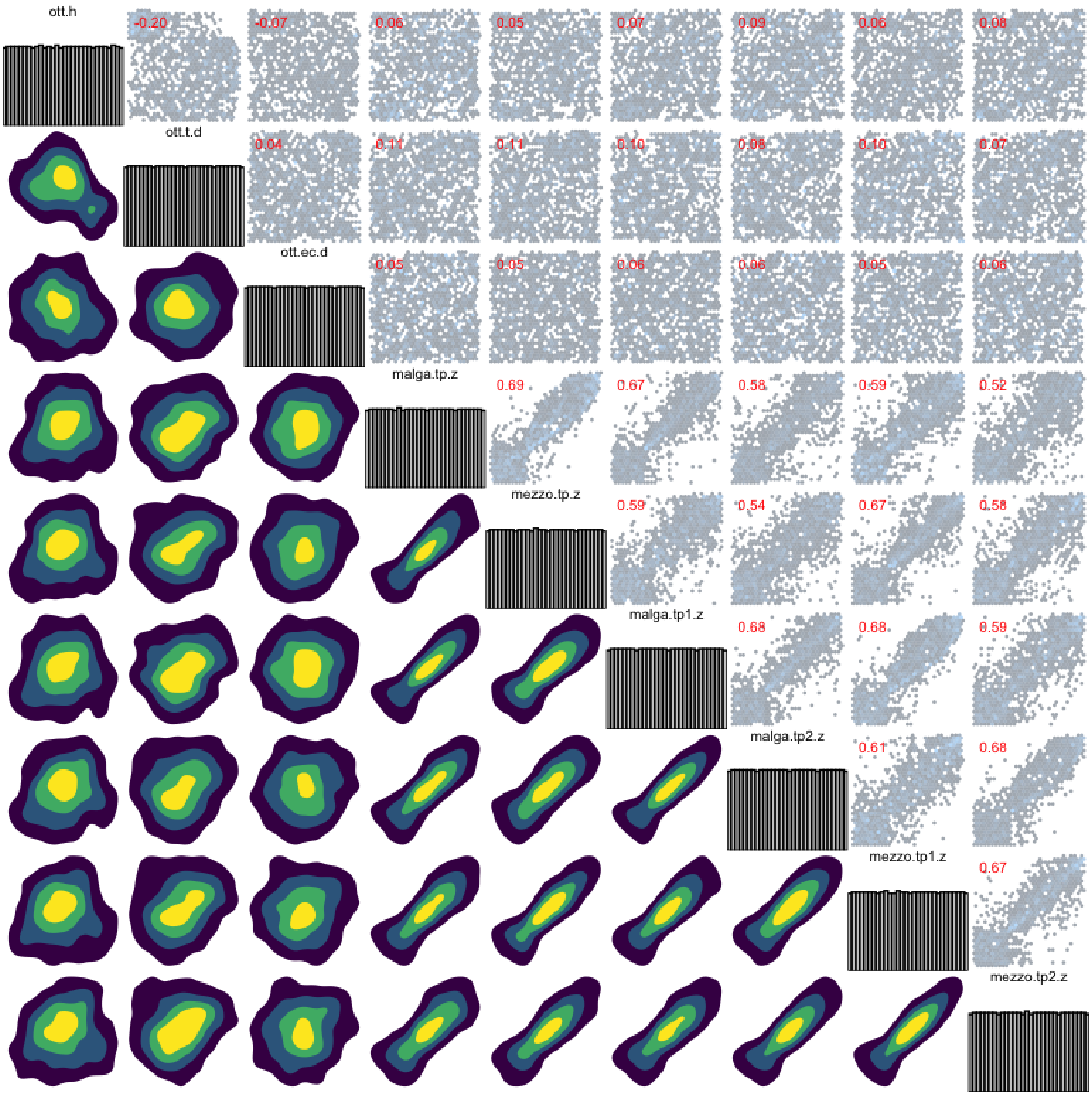


Figure A.7: Pairwise scatterplot matrix of pseudo-observations for regime 1–2–2. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 17: Pair–copula family selection summary for Regime 1-2-2 (n=1675). All tree–edge combinations with non–independence families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(3,1)	—	c,c	clayton	270	0.15	1.00	−0.07	14.03
1	2	(1,2)	—	c,c	bb8	270	(1.76, 0.93)	2.00	−0.23	129.12
1	3	(2,4)	—	c,d	bb8	0	(1.61, 0.77)	2.00	0.12	31.49
1	4	(4,5)	—	d,d	tll	0	[30x30 grid]	40.19	0.65	1233.78
1	5	(5,8)	—	d,d	tll	0	[30x30 grid]	51.86	0.63	1189.23
1	6	(8,6)	—	d,d	tll	0	[30x30 grid]	42.95	0.65	1269.27
1	7	(6,7)	—	d,d	tll	0	[30x30 grid]	52.71	0.64	1221.70
1	8	(7,9)	—	d,d	tll	0	[30x30 grid]	39.48	0.65	1219.48
2	2	(1,4)	(2)	c,d	joe	180	1.21	1.00	0.11	27.74
3	1	(3,4)	(2,1)	c,d	gaussian	0	0.08	1.00	0.05	5.42
3	4	(4,6)	(8,5)	d,d	tll	0	[30x30 grid]	27.89	0.39	256.87
3	6	(8,9)	(7,6)	d,d	tll	0	[30x30 grid]	48.33	0.47	723.27
5	3	(2,7)	(6,8,5,4)	c,d	t	0	(−0.08, 8.72)	2.00	−0.05	10.78
6	2	(1,7)	(6,8,5,4,2)	c,d	frank	0	0.54	1.00	0.06	4.51

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

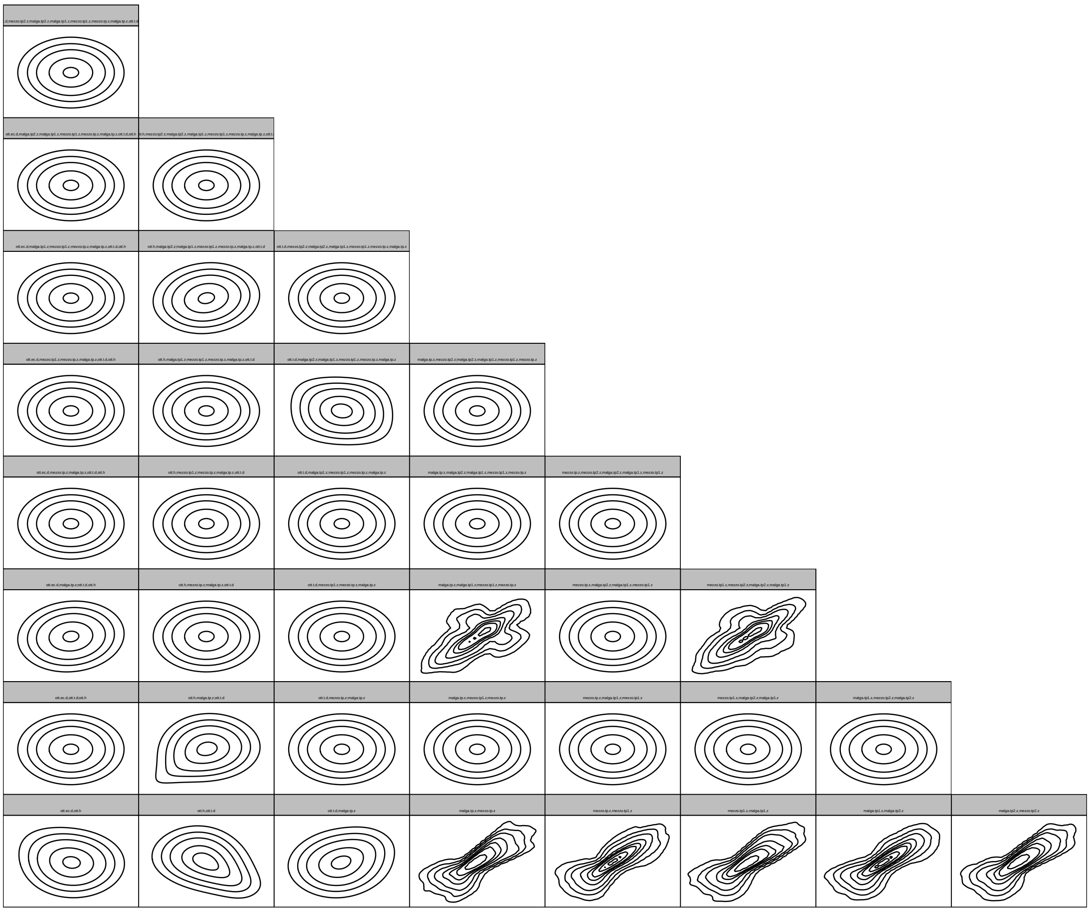


Figure A.8: Fitted pair-copula contour plots for all variable pairs in Regime 1–2–2.

A.5 Regime 2-1-1 (n=708)

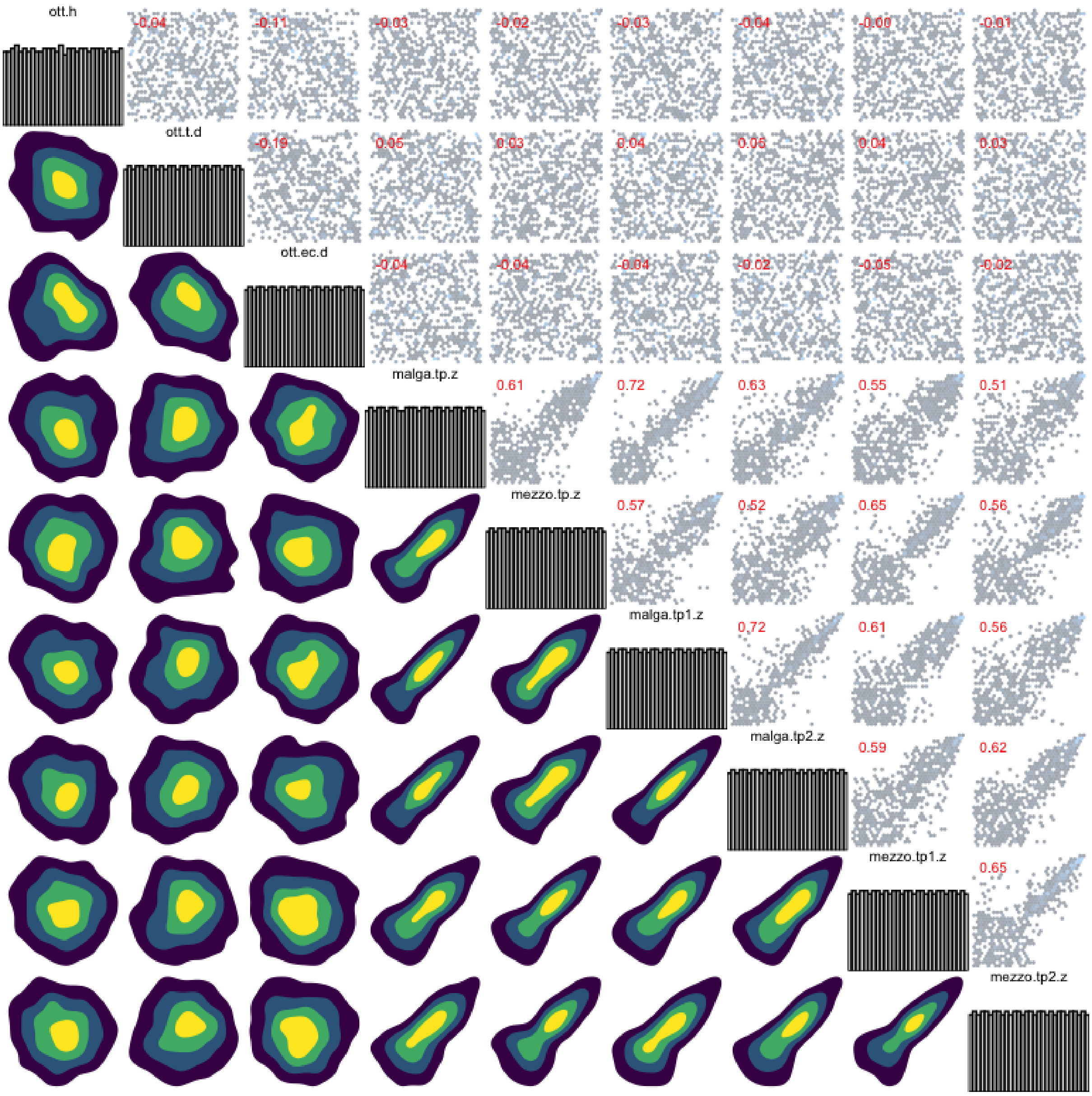


Figure A.9: Pairwise scatterplot matrix of pseudo-observations for regime 2–1-1. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 18: Pair–copula family selection summary for Regime 2-1-1 (n=708). All tree–edge combinations with non-independent families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(1,3)	—	c,c	tawn	270	(0.30, 0.75, 2.21)	3.00	−0.20	28.15
1	2	(2,3)	—	c,c	bb8	270	(1.89, 0.84)	2.00	−0.20	34.44
1	3	(4,6)	—	d,d	tll	0	[30x30 grid]	31.36	0.68	681.41
1	4	(3,8)	—	c,d	clayton	270	0.12	1.00	−0.06	7.74
1	5	(6,7)	—	d,d	tawn	180	(0.99, 0.97, 7.00)	3.00	0.83	827.12
1	6	(5,8)	—	d,d	bb6	180	(4.70, 2.50)	2.00	0.86	666.80
1	7	(7,9)	—	d,d	bb7	180	(6.00, 2.45)	2.00	0.77	637.78
1	8	(8,9)	—	d,d	bb6	180	(3.74, 3.30)	2.00	0.88	724.73
2	1	(1,2)	(3)	c,c	frank	0	–0.55	1.00	−0.06	3.29
2	3	(4,7)	(6)	d,d	bb1	180	(0.00, 1.07)	2.00	0.07	13.93
2	5	(6,9)	(7)	d,d	bb7	0	(1.06, 0.17)	2.00	0.11	41.49
2	6	(5,9)	(8)	d,d	gumbel	180	1.14	1.00	0.12	38.59
2	7	(7,8)	(9)	d,d	frank	0	2.25	1.00	0.24	73.33
3	3	(4,9)	(7,6)	d,d	bb7	0	(1.11, 0.25)	2.00	0.16	37.02
3	5	(6,8)	(9,7)	d,d	t	0	(0.55, 49.80)	2.00	0.37	138.37
3	6	(5,7)	(9,8)	d,d	tawn	180	(0.30, 1.00, 1.50)	3.00	0.15	15.45
4	3	(4,8)	(9,7,6)	d,d	bb1	180	(0.00, 1.02)	2.00	0.02	8.46
4	4	(3,7)	(9,5,8)	c,d	clayton	180	0.13	1.00	0.06	4.00

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

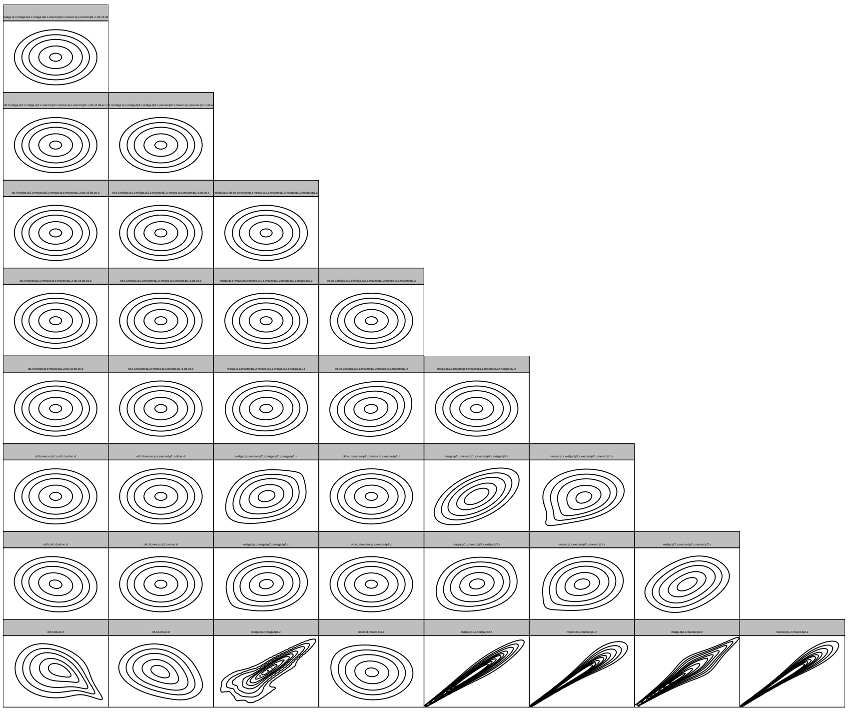


Figure A.10: Fitted pair-copula contour plots for all variable pairs in Regime 2–1–1.

A.6 Regime 2-1-2 (n=3245)

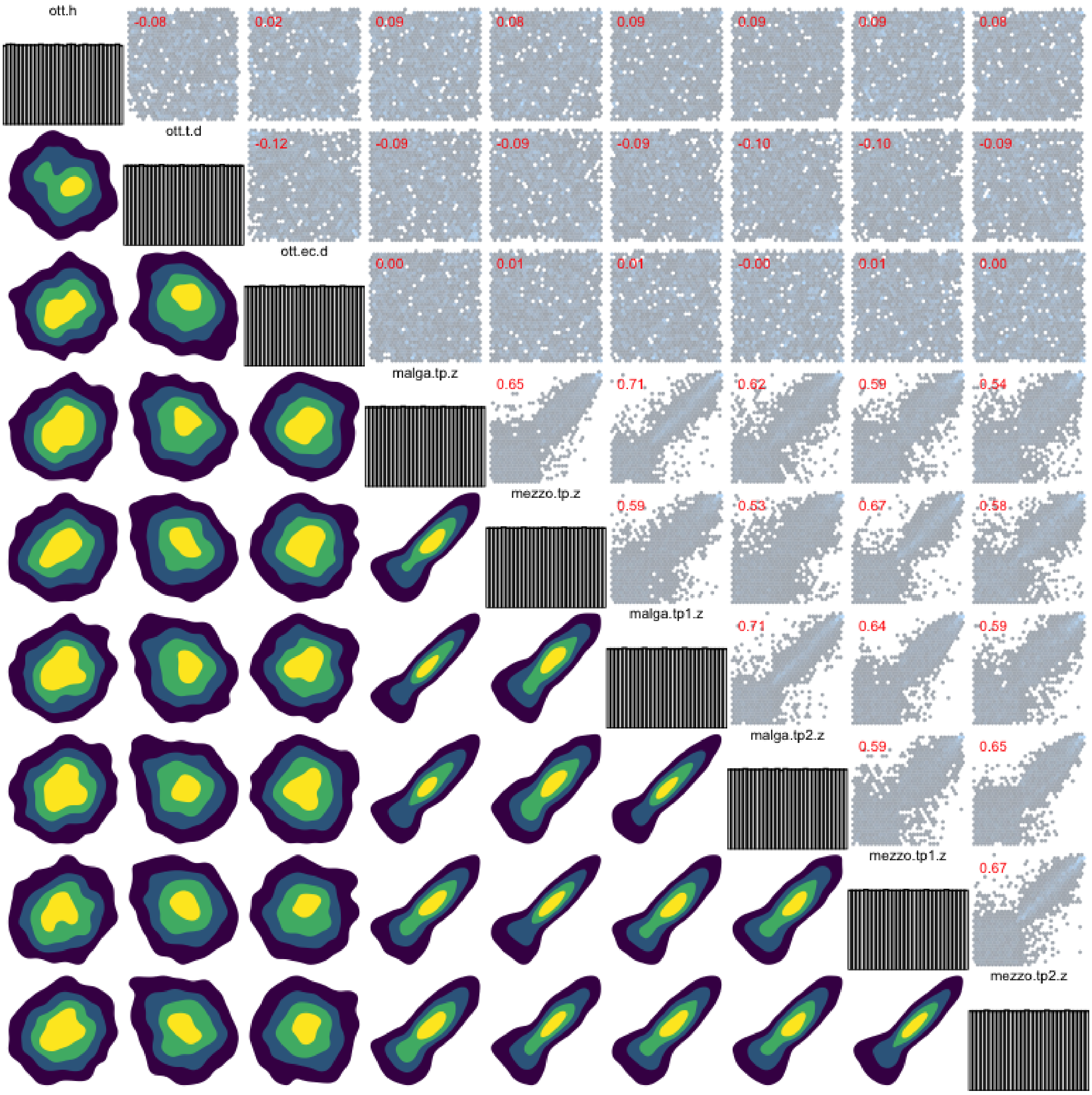


Figure A.11: Pairwise scatterplot matrix of pseudo-observations for regime 2–1–2. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 19: Pair–copula family selection summary for Regime 2-1-2 (n=3245). All tree–edge combinations with non–independence families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	1	(1,4)	—	c,d	gumbel	0	1.10	1.00	0.09	48.63
1	2	(3,2)	—	c,c	bb8	90	(1.30, 0.96)	2.00	−0.12	73.59
1	3	(2,8)	—	c,d	bb8	90	(1.32, 0.94)	2.00	−0.11	62.65
1	4	(4,6)	—	d,d	tll	0	[30x30 grid]	65.39	0.69	2905.25
1	5	(6,7)	—	d,d	tll	0	[30x30 grid]	68.04	0.68	2871.61
1	6	(5,8)	—	d,d	tll	0	[30x30 grid]	60.10	0.64	2706.38
1	7	(7,9)	—	d,d	tll	0	[30x30 grid]	45.19	0.62	2346.71
1	8	(8,9)	—	d,d	tll	0	[30x30 grid]	60.20	0.65	2853.30
2	1	(1,6)	(4)	c,d	clayton	180	0.06	1.00	0.03	5.16
2	2	(3,8)	(2)	c,d	t	0	(−0.02, 5.50)	2.00	−0.01	35.23
2	3	(2,9)	(8)	c,d	joe	90	1.07	1.00	−0.04	14.95
3	3	(2,7)	(9,8)	c,d	clayton	270	0.06	1.00	−0.03	4.86
3	5	(6,8)	(9,7)	d,d	tll	0	[30x30 grid]	72.56	0.42	1200.87
5	3	(2,4)	(6,7,9,8)	c,d	clayton	270	0.09	1.00	−0.04	9.47
5	4	(4,5)	(8,9,7,6)	d,d	tll	0	[30x30 grid]	67.54	0.43	1186.44
6	1	(1,2)	(8,9,7,6,4)	c,c	frank	0	–0.46	1.00	−0.05	10.95

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

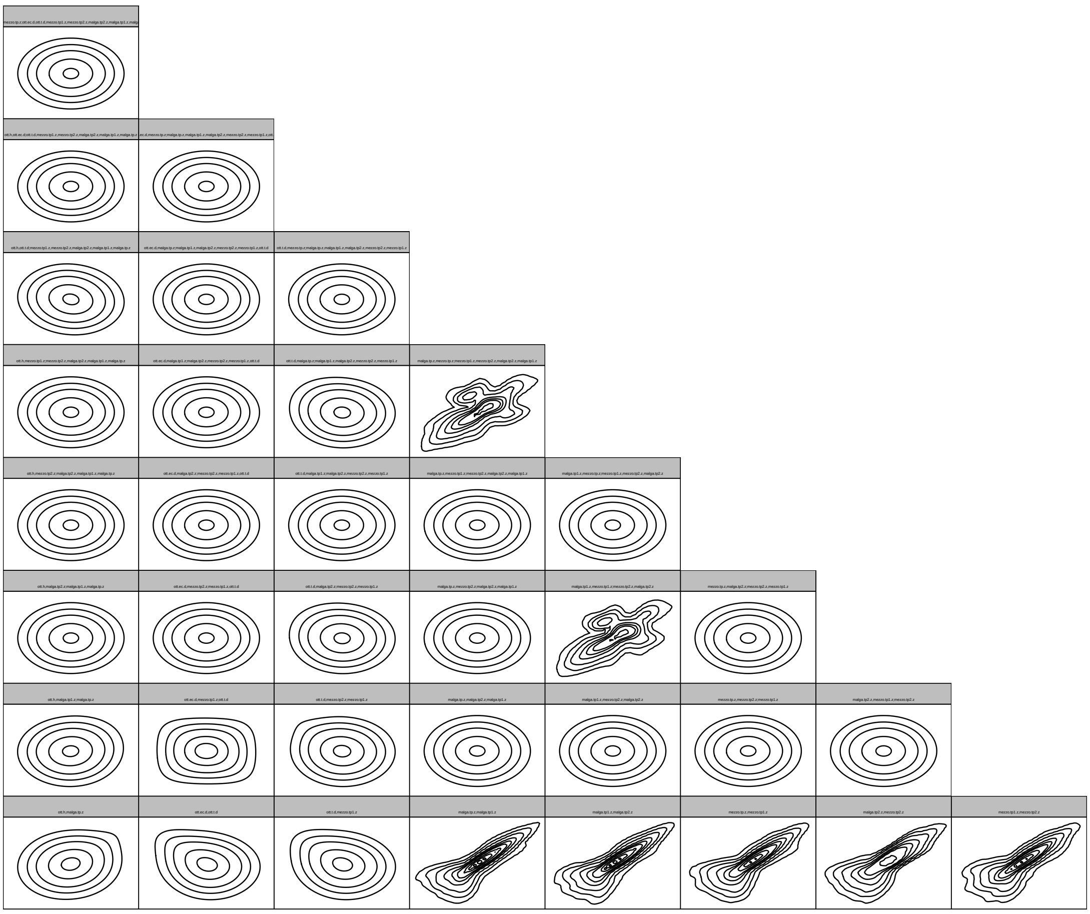


Figure A.12: Fitted pair-copula contour plots for all variable pairs in Regime 2–1–2.

A.7 Regime 2-2-1 (n=973)

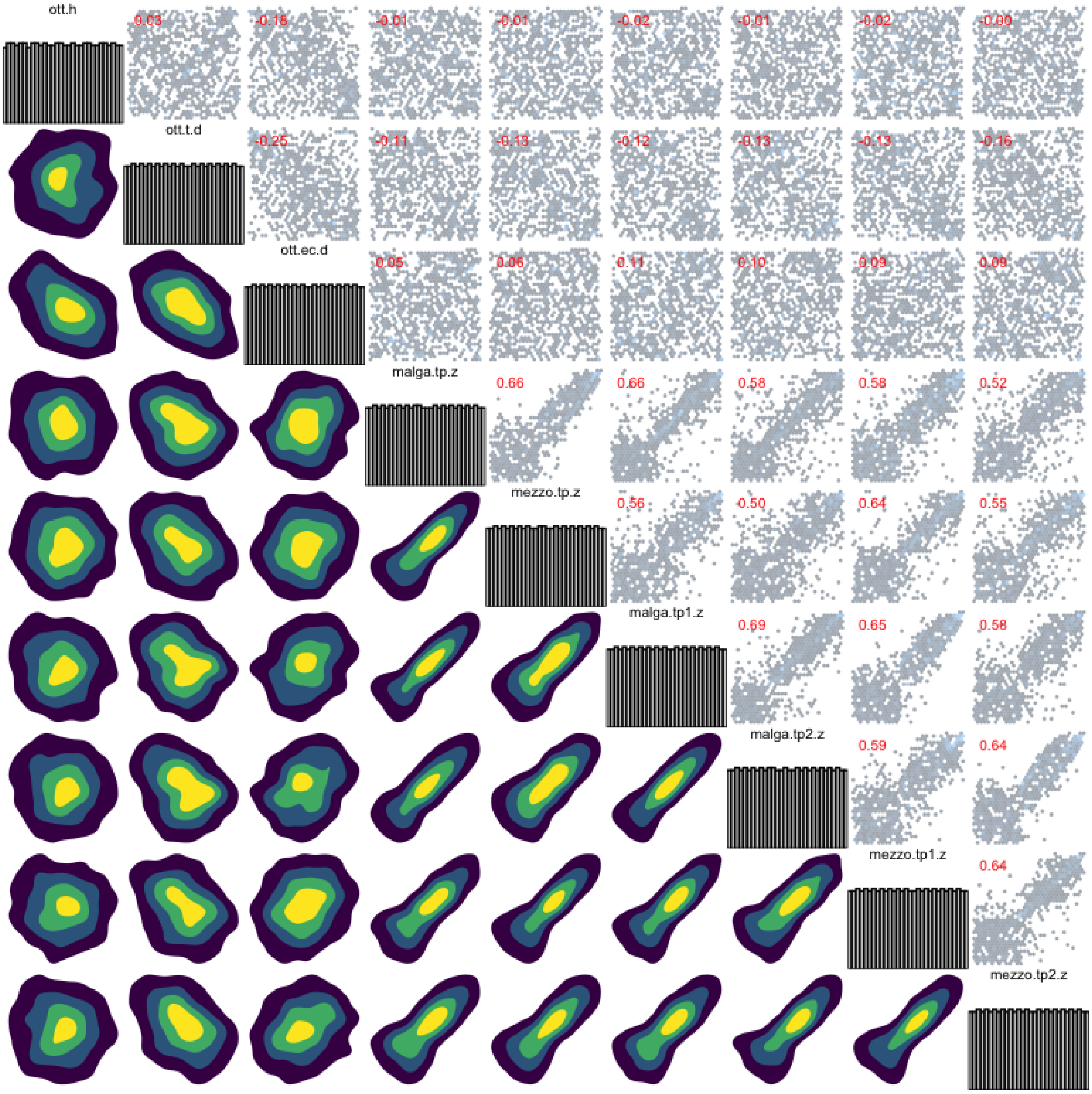


Figure A.13: Pairwise scatterplot matrix of pseudo-observations for regime 2–2–1. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 20: Pair–copula family selection summary for Regime 2–2–1 (n=973). Non-independence families only; values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log–lik
1	1	(4,5)		d,d	tll	0	–	23.68	0.63	731.24
1	2	(7,6)		d,d	tll	0	–	36.55	0.65	739.72
1	3	(1,3)		c,c	tawn	270	–	3.00	0.20	57.78
1	4	(6,8)		d,d	tll	0	–	24.33	0.56	205.97
1	5	(3,2)		c,c	bb8	0	–	2.00	−0.26	74.27
1	6	(5,8)		d,d	tll	0	–	33.94	0.60	644.21
1	7	(2,9)		c,d	frank	0	–	0.00	−1.15	26.89
1	8	(8,9)		d,d	tll	0	–	34.29	0.60	662.96
3	2	(7,9)	(8,5)	d,d	tll	0	–	36.61	0.45	366.56
4	1	(4,6)	(9,8,5)	d,d	tll	0	–	29.92	0.41	397.15
5	5	(7,5)	(6,9,8)	c,d	gaussian	0	–	1.00	−0.08	4.70
5	4	(6,3)	(2,9,7)	c,d	frank	0	–	1.00	0.08	4.66

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

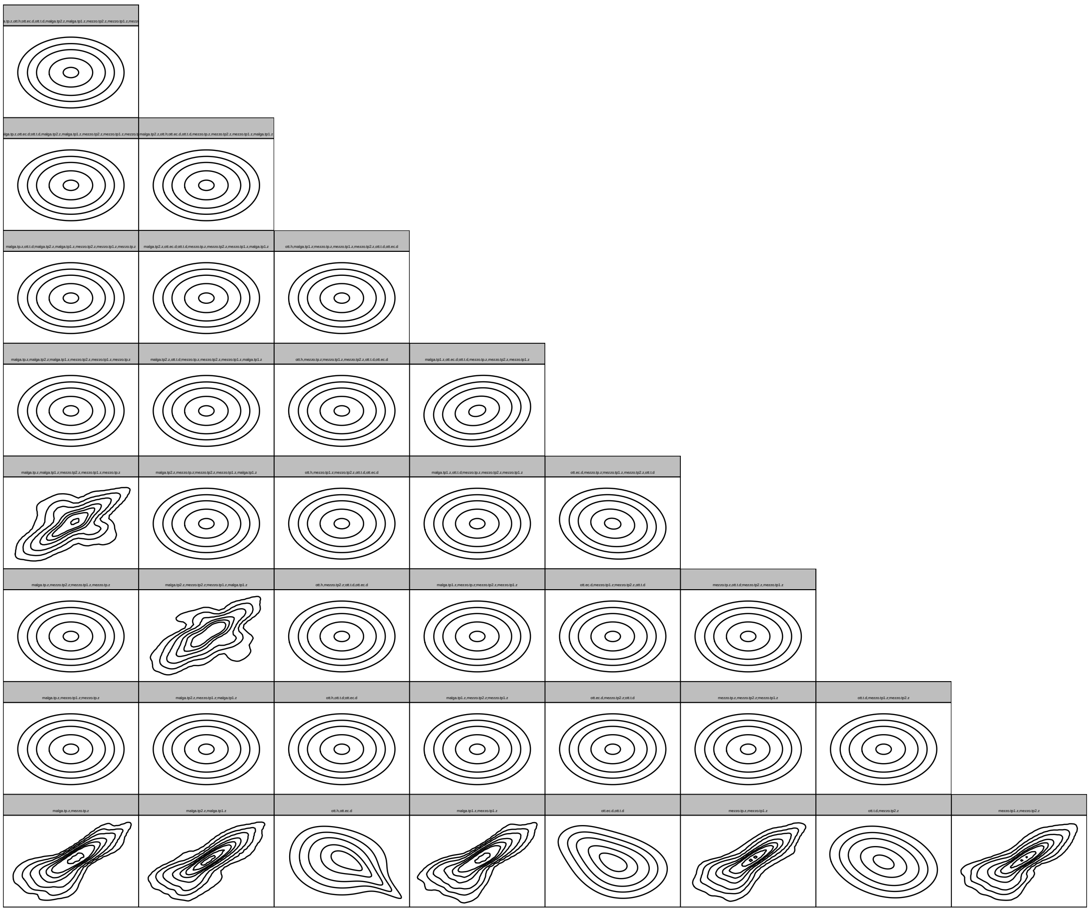


Figure A.14: Fitted pair-copula contour plots for all variable pairs in Regime 2–2–1.

A.8 Regime 2-2-2 (n=982)

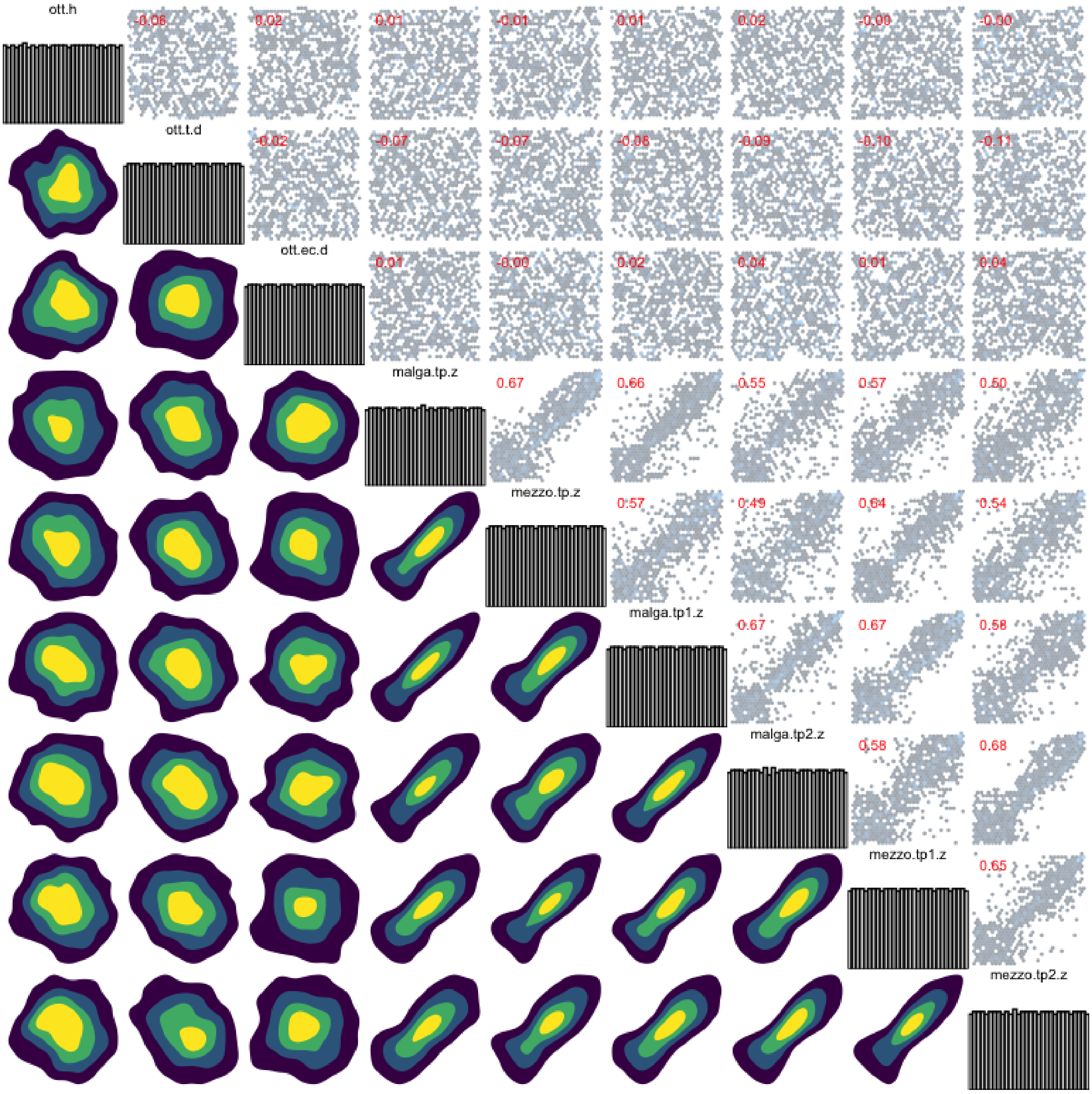


Figure A.15: Pairwise scatterplot matrix of pseudo-observations for regime 2–2–2. Each off-diagonal panel displays the corresponding bivariate scatterplot with empirical contours, illustrating the pairwise dependence within this regime.

Table 21: Pair–copula family selection summary for Regime 2-2-2 (n=982). All tree–edge combinations with non–independence families, values rounded to two decimals.

Tree	Edge	Conditioned	Conditioning	Var. Types	Family	Rot.	Parameters	df	τ	Log-lik
1	2	(4,5)	—	d,d	tll	0	[30x30 grid]	30.32	0.64	712.16
1	3	(5,8)	—	d,d	tll	0	[30x30 grid]	37.20	0.61	624.95
1	4	(8,6)	—	d,d	tll	0	[30x30 grid]	30.85	0.65	743.23
1	5	(1,2)	—	c,c	gaussian	0	–0.08	1.00	–0.05	3.63
1	6	(6,7)	—	d,d	tll	0	[30x30 grid]	44.50	0.63	707.12
1	7	(2,9)	—	c,d	frank	0	–1.00	1.00	–0.11	13.07
1	8	(7,9)	—	d,d	tll	0	[30x30 grid]	32.75	0.64	765.73
3	2	(4,6)	(8,5)	d,d	tll	0	[30x30 grid]	42.43	0.47	439.80
3	4	(8,9)	(7,6)	d,d	tll	0	[30x30 grid]	39.55	0.46	390.76
3	5	(1,7)	(9,2)	c,d	clayton	180	0.12	1.00	0.05	6.30
6	3	(5,1)	(2,9,7,6,8)	d,c	clayton	180	0.07	1.00	0.04	5.32
7	1	(3,2)	(9,4,5,8,6,7)	c,c	joe	90	1.07	1.00	–0.04	3.85
8	1	(3,1)	(2,9,4,5,8,6,7)	c,c	clayton	0	0.12	1.00	0.06	4.94

Variable mapping:
1 = `ott.h`, 2 = `ott.t.d`, 3 = `ott.ec.d`, 4 = `malga.tp`, 5 = `mezzo.tp`, 6 = `malga.tp1`, 7 = `mezzo.tp1`, 8 = `malga.tp2`, 9 = `mezzo.tp2`.

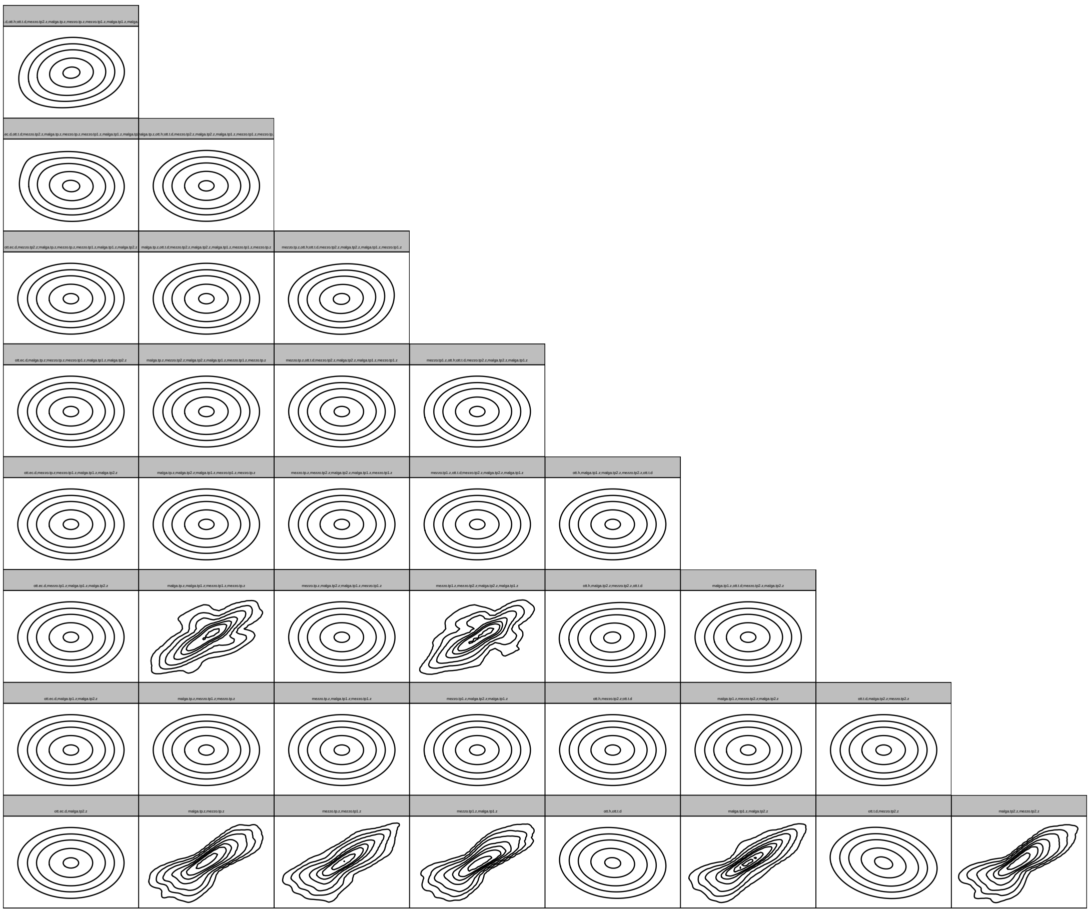


Figure A.16: Fitted pair-copula contour plots for all variable pairs in Regime 2–2–2.

B Appendix for Section 15 (Individual Regression Analyses)

To further examine the dependence structure across variables under different regimes, contour plots of the fitted copula models are presented. These plots provide a visual representation of the joint distribution implied by the estimated copulas and allow for a direct comparison of dependence patterns across regimes. These plots indicate that most variable pairs exhibit weak dependence, as evidenced by the nearly circular and symmetric contour lines. This suggests that the joint distribution in this regime is close to independence, with little evidence of asymmetric or strong tail dependence.

B.1 Regime 1-1-1

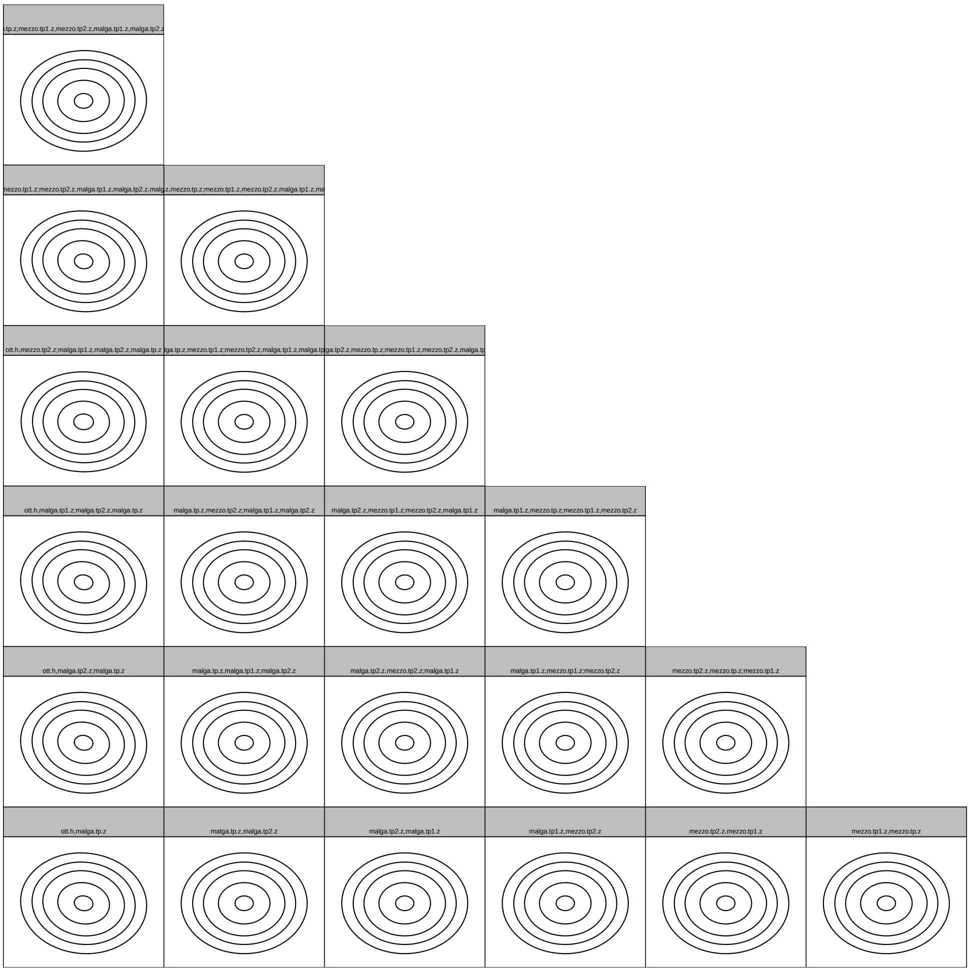
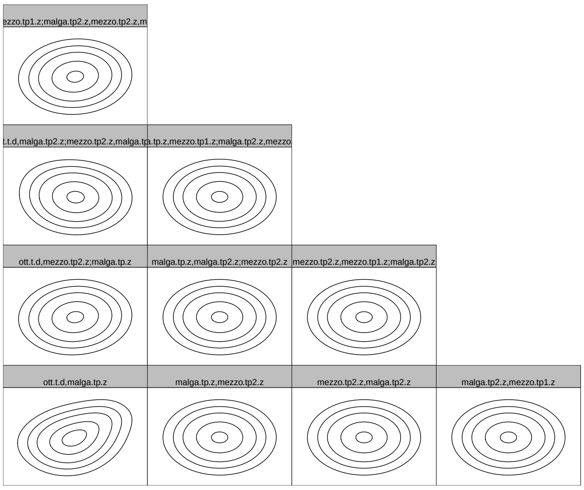
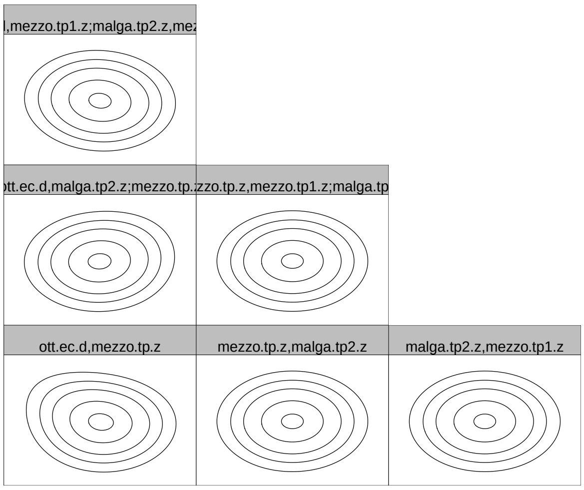


Figure B.1: Contour plot of the fitted vine copula for ott.h in regime 1-1-1



(a) Contour plot for `ott.t.d` in regime 1-1-1



(b) Contour plot for `ott.ec.d` in regime 1-1-1

Figure B.2: Contour plots of the fitted vine copulas for regime 1-1-1.

B.2 Regime 1-2-1

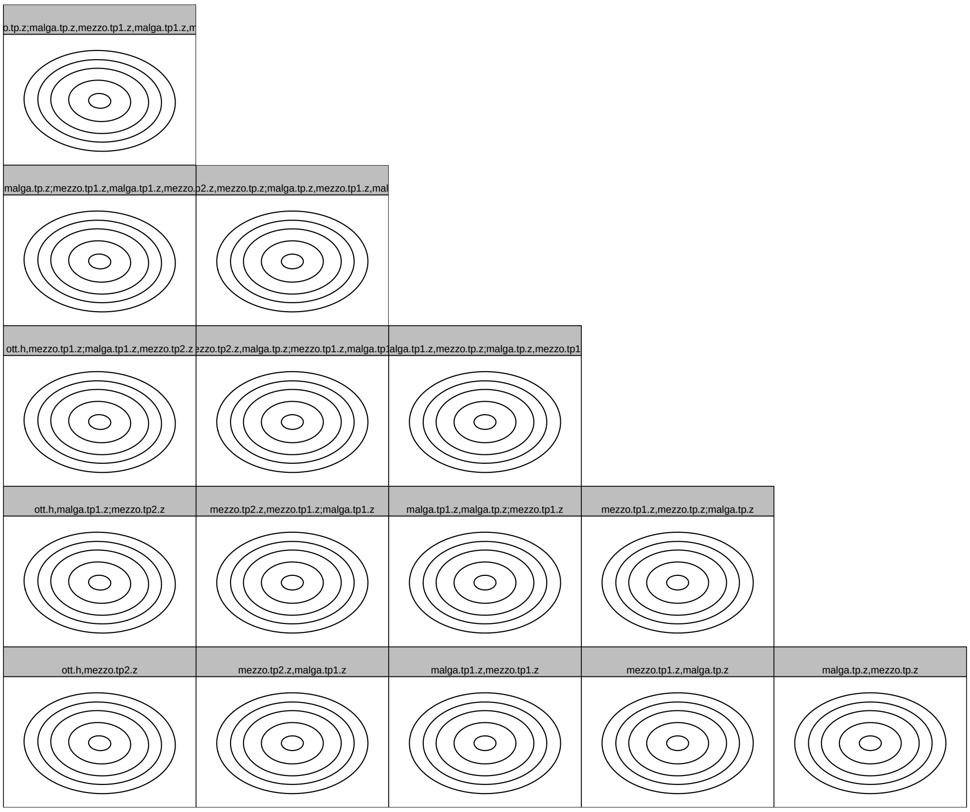
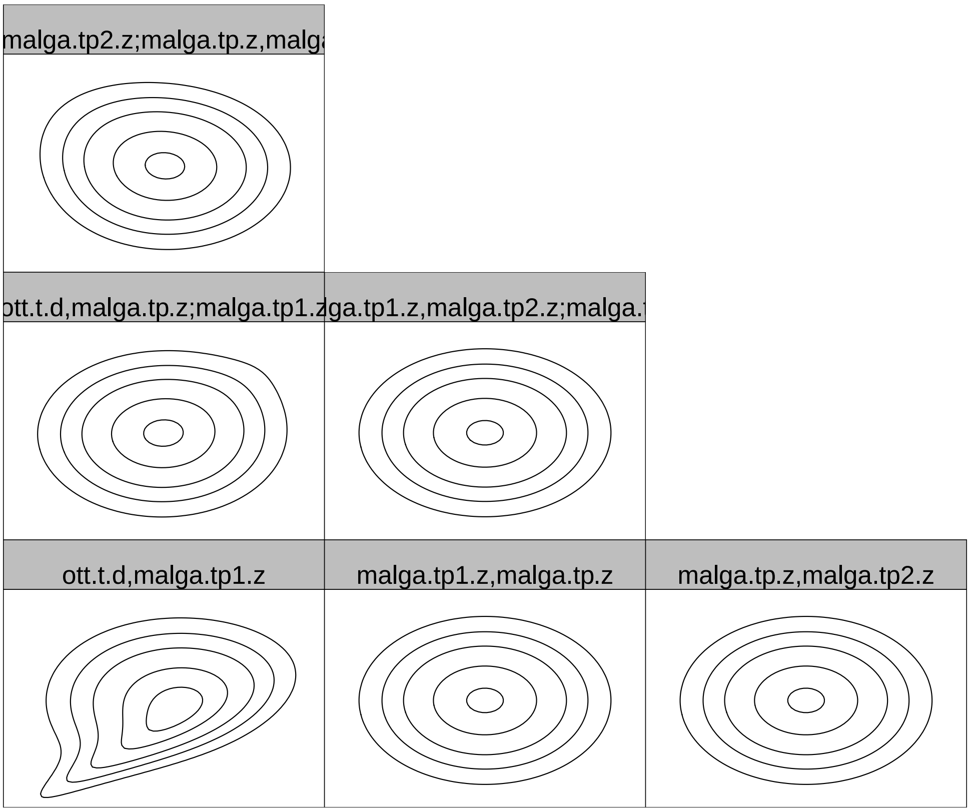
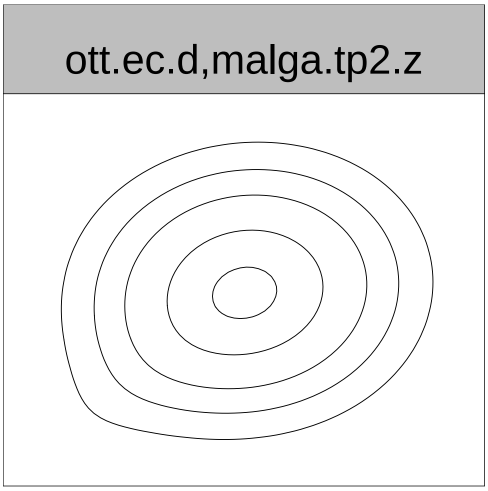


Figure B.3: Contour plot of the fitted vine copula for `ott.h` in regime 1–2–1.



(a) Contour plot for `ott.t.d` in regime 1–2–1.



(b) Contour plot for `ott.ec.d` in regime 1–2–1.

B.3 Regime 1-2-2

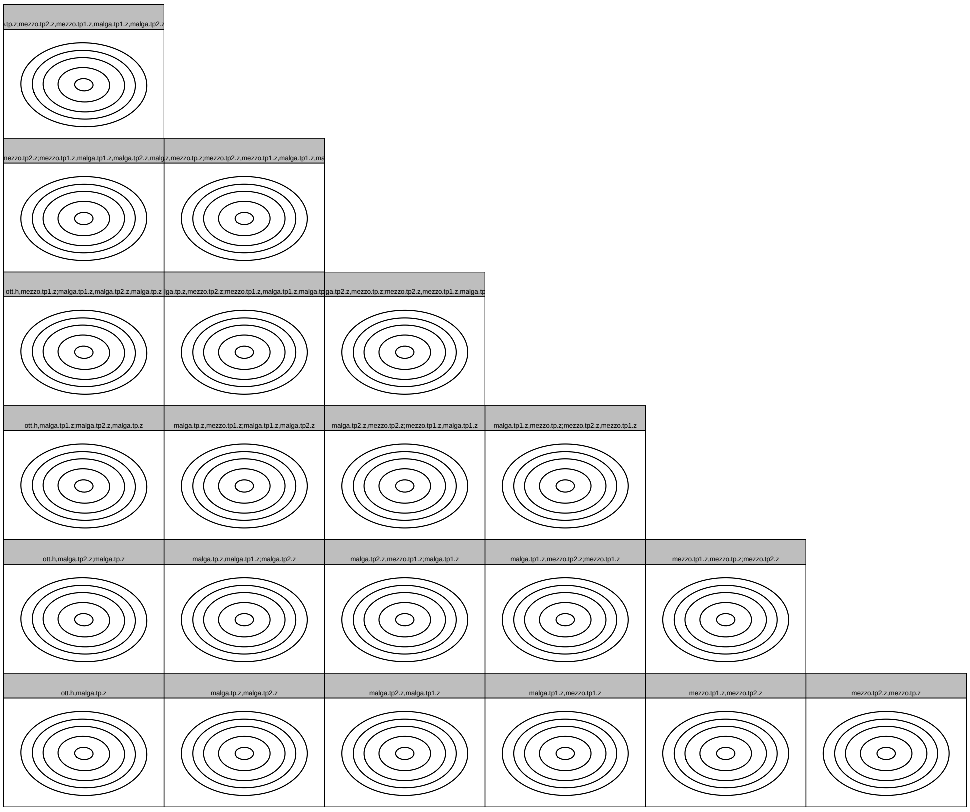


Figure B.5: Contour plot of the fitted vine copula for `ott.h` in regime 1–2–2.

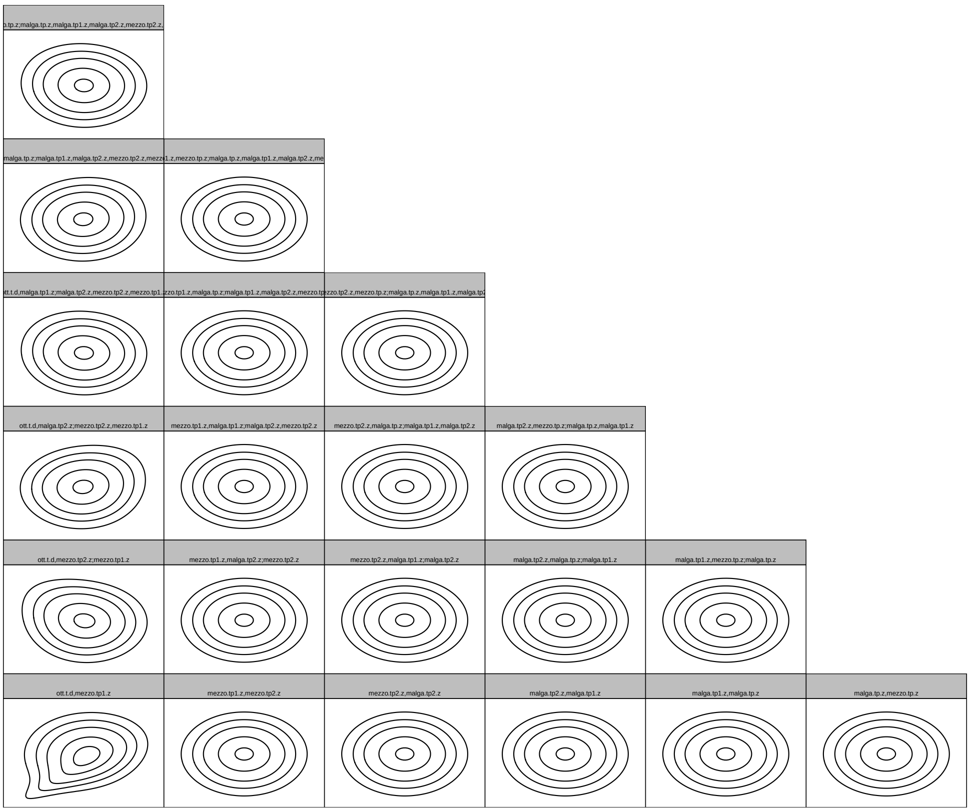


Figure B.6: Contour plot of the fitted vine copula for `ott.t.d` in regime 1–2–2.

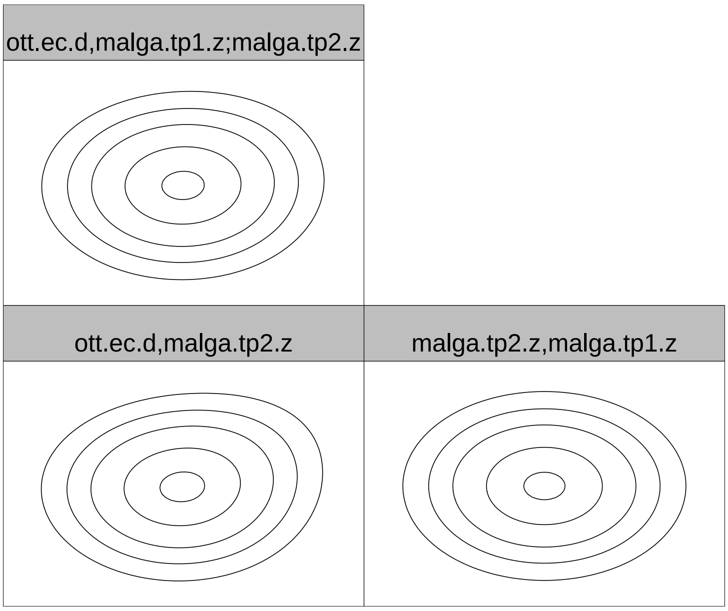


Figure B.7: Contour plot of the fitted vine copula for `ott.ec.d` in regime 1–2–2.

B.4 Regime 2-1-1

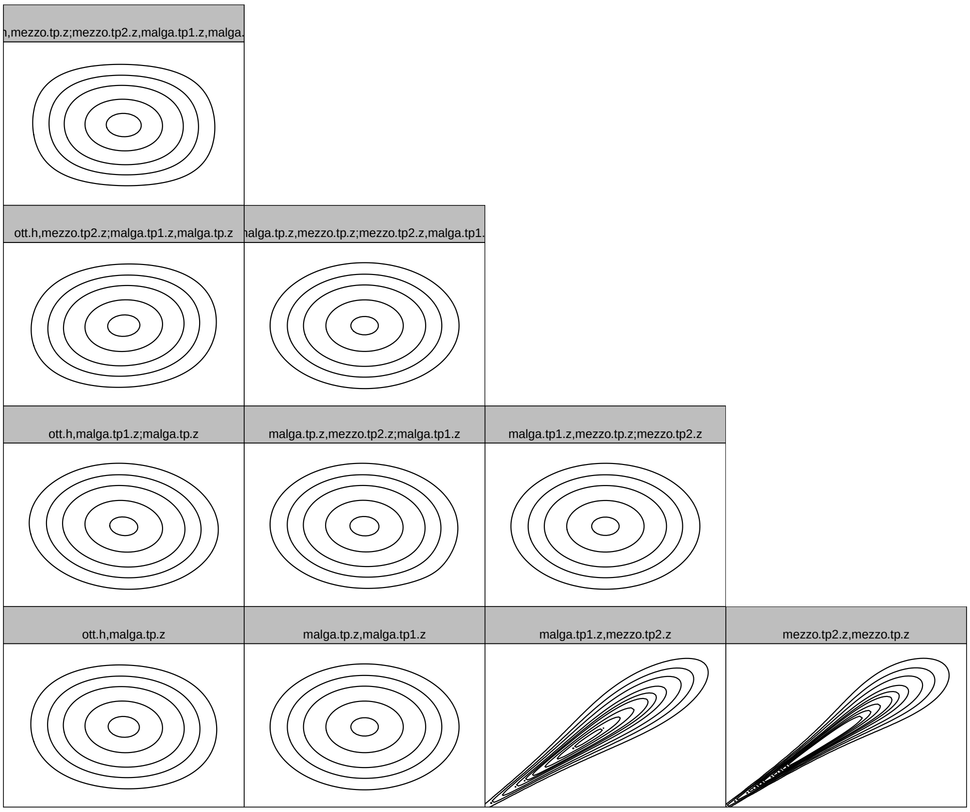


Figure B.8: Contour plot of the fitted vine copula for `ott.h` in regime 2–1–1.

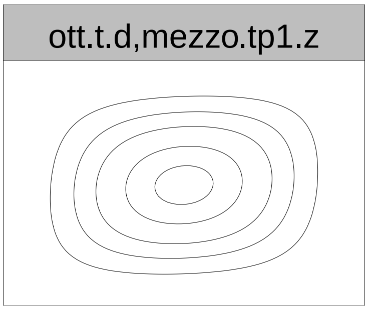


Figure B.9: Contour plot of the fitted vine copula for `ott.t.d` in regime 2–1–1.

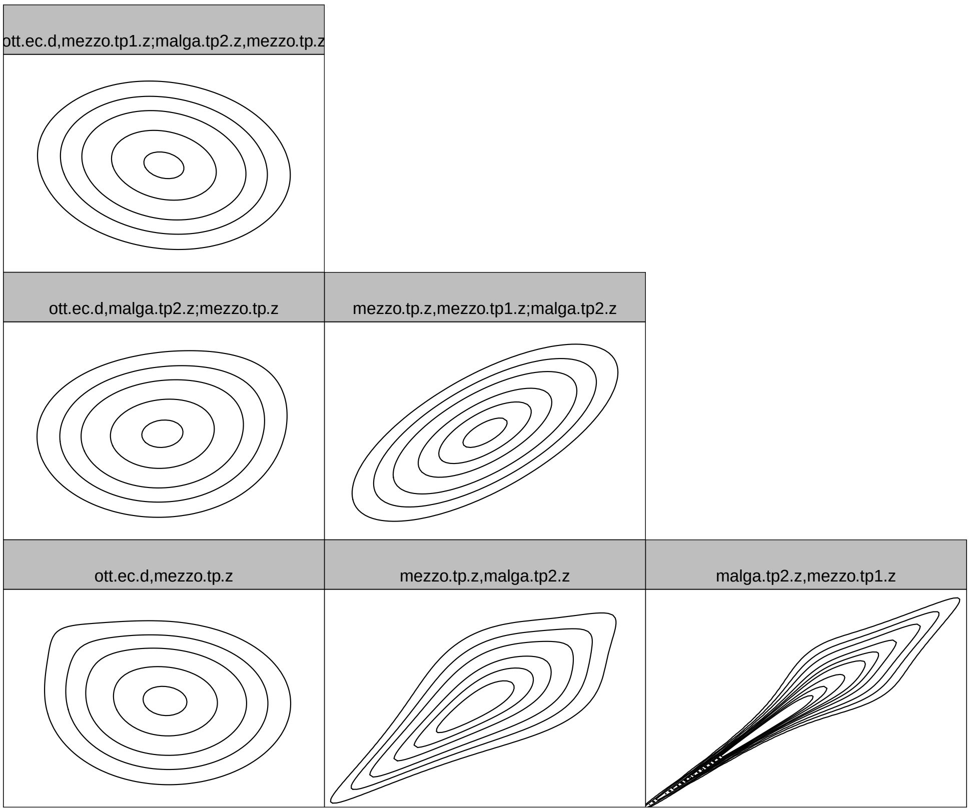


Figure B.10: Contour plot of the fitted vine copula for `ott.ec.d` in regime 2–1–1.

B.5 Regime 2-2-1

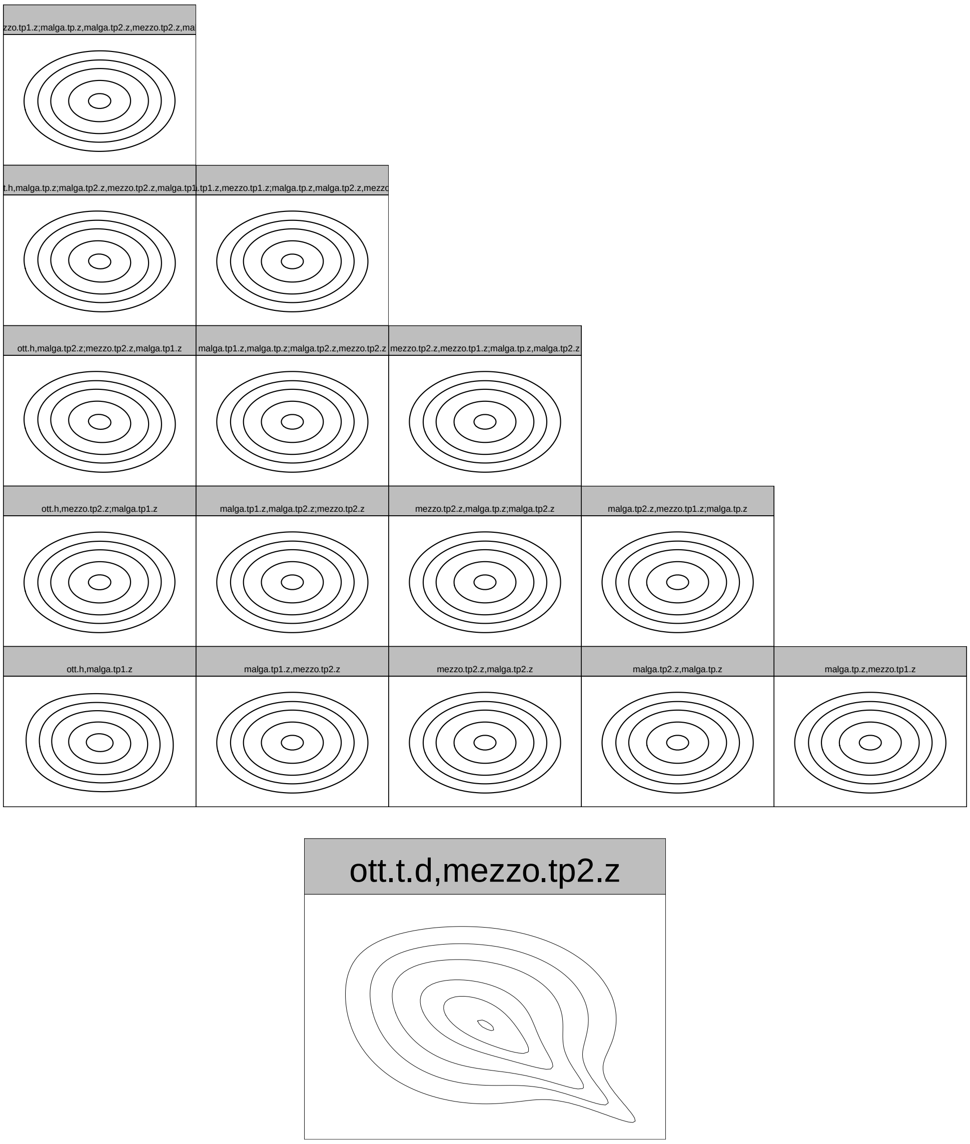


Figure B.11: Contour plots of the fitted vine copula for `ott.h` (top) and `ott.t.d` (bottom) in regime 2–2–1.

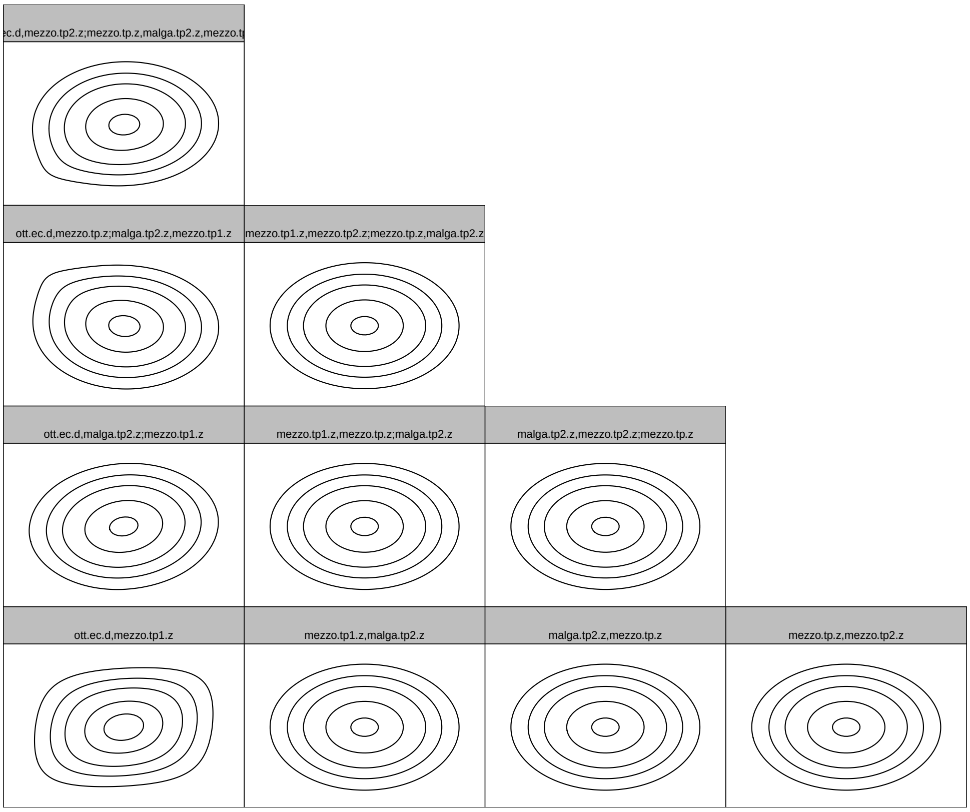


Figure B.12: Contour plot of the fitted vine copula for `ott.ec.d` in regime 2–2–1.

B.6 Regime 2-1-2

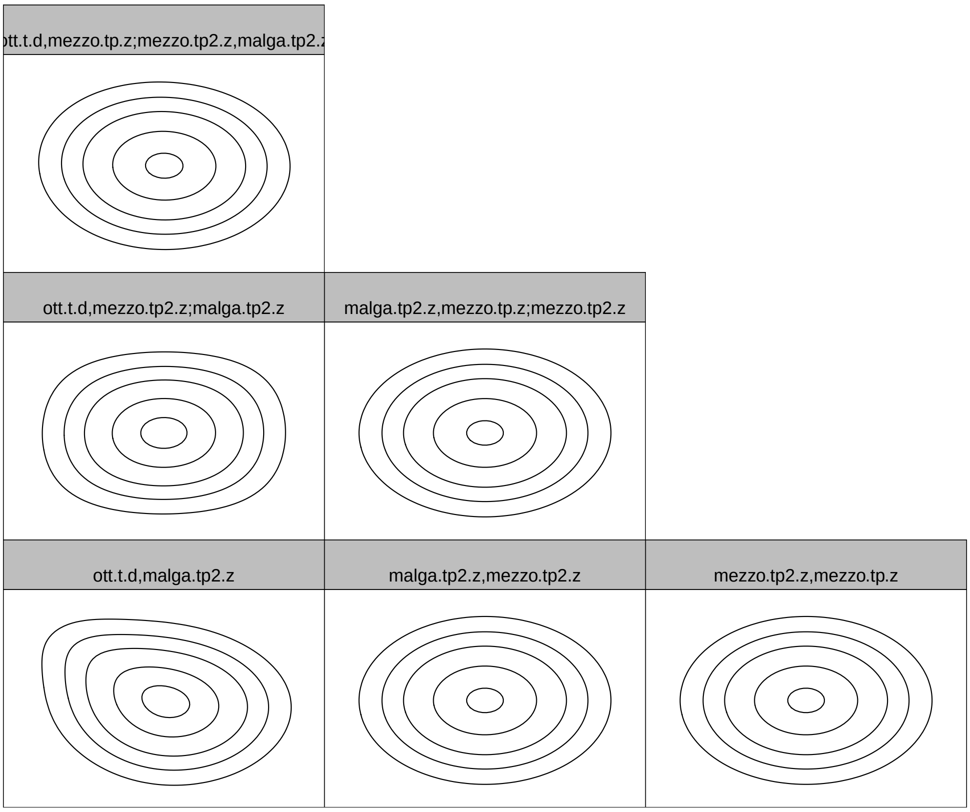


Figure B.13: Contour plot of the fitted vine copula for `ott.t.d` in regime 2–1–2.

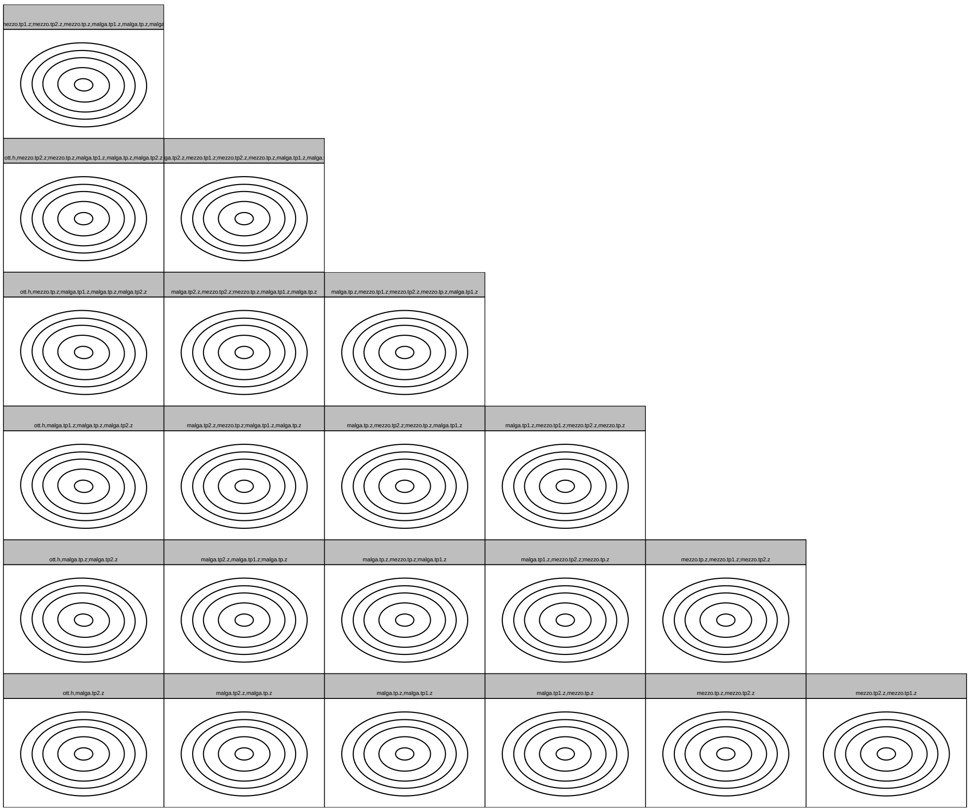


Figure B.14: Contour plot of the fitted vine copula for `ott.h` in regime 2–1–2.

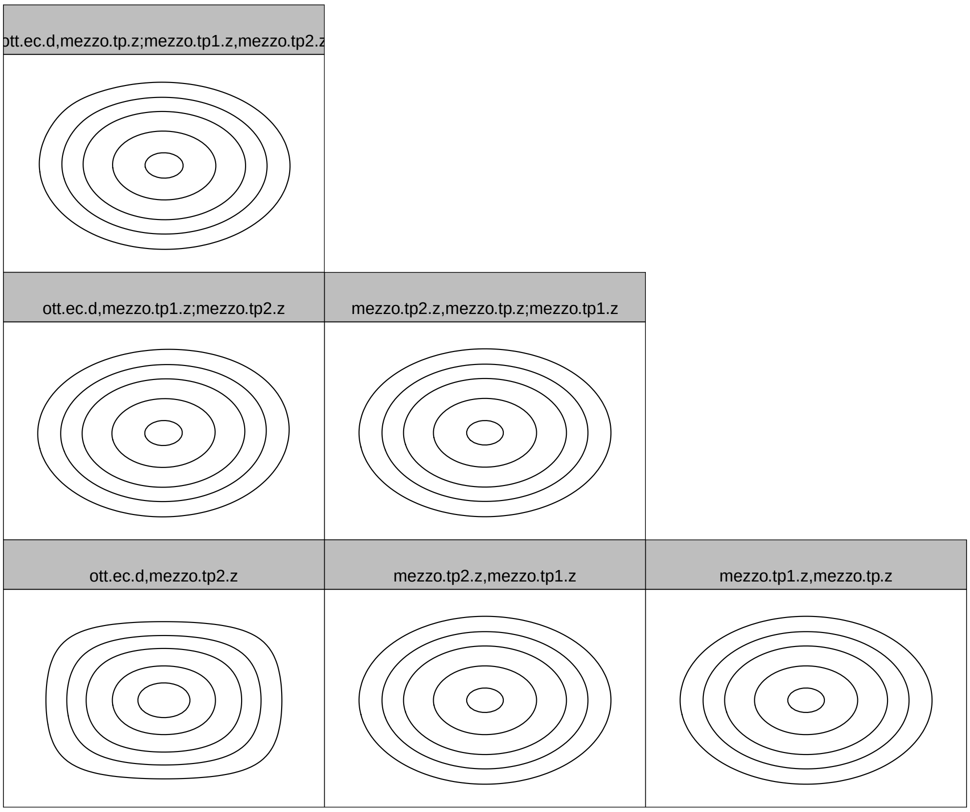


Figure B.15: Contour plot of the fitted vine copula for `ott.ec.d` in regime 2–1–2.

B.7 Regime 2-2-2

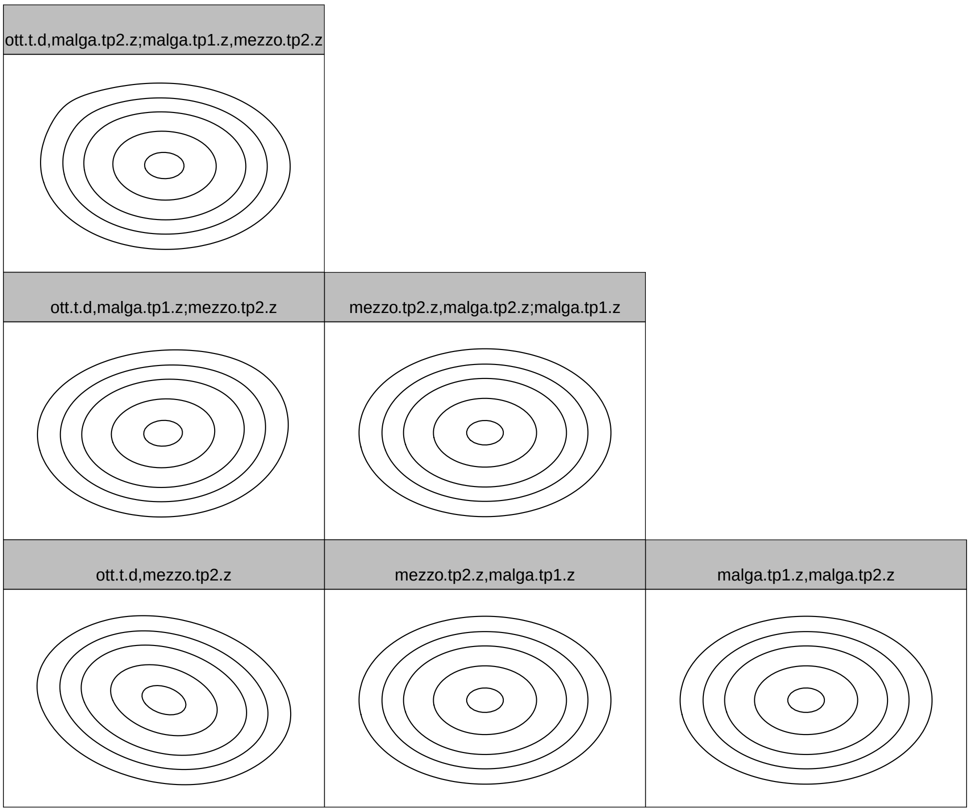


Figure B.16: Contour plot of the fitted vine copula for `ott.t.d` in regime 2–2–2.

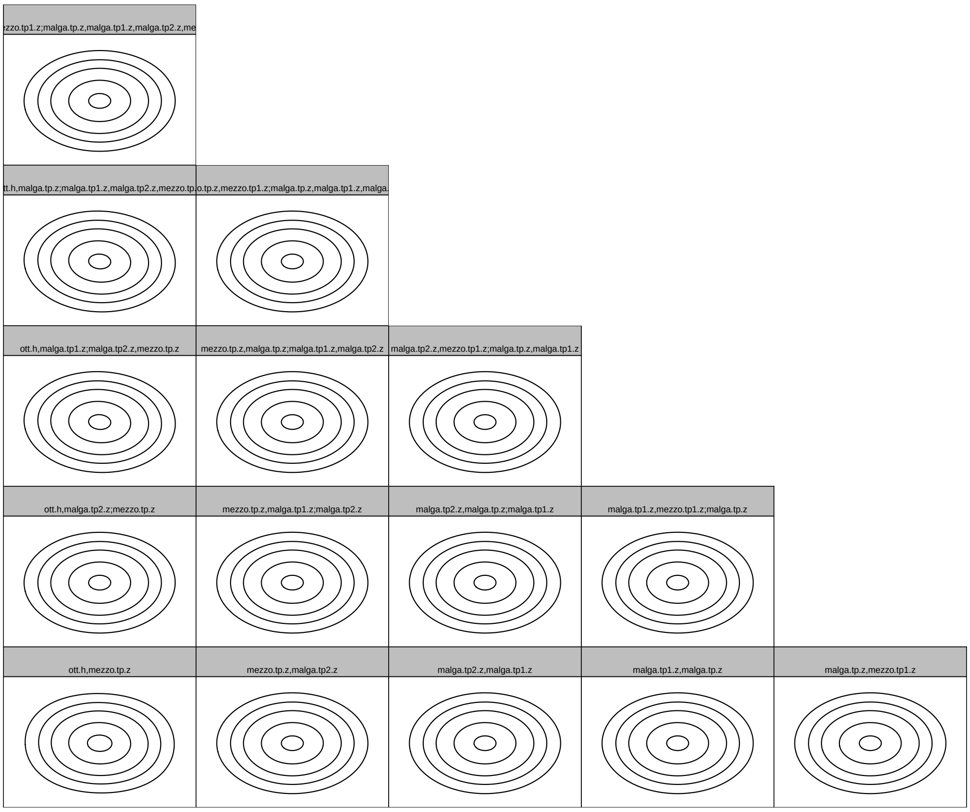


Figure B.17: Contour plot of the fitted vine copula for `ott.h` in regime 2–2–2.

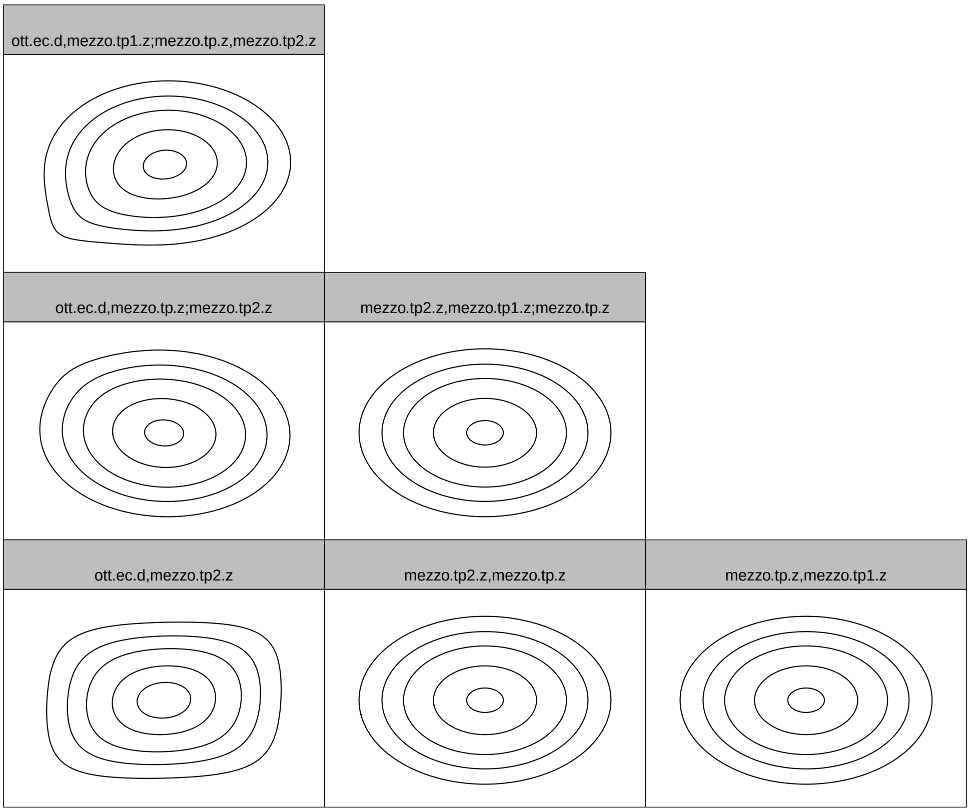


Figure B.18: Contour plot of the fitted vine copula for `ott.ec.d` in regime 2–2–2.

References

Olivier Cappé, Éric Moulines, and Thomas Rydén. *Inference in Hidden Markov Models*. Springer Series in Statistics. Springer, New York, 2005. ISBN 978-0-387-40264-2. doi: 10.1007/0-387-28982-8.

Claudia Czado. *Analyzing Dependent Data with Vine Copulas: A Practical Guide with R*, volume 222 of *Lecture Notes in Statistics*. Springer, Cham, 2019. ISBN 978-3-030-13784-7. doi: 10.1007/978-3-030-13785-4.

Claudia Czado. Analyzing the lynx data using regime switching with r package MSwM. Unpublished manuscript; public link not found, n.d.a.

Claudia Czado. Regime switching to explain hydropeaking on the noce river. Technical report, Technical University of Munich, n.d.b. Unpublished/uncatalogued; public link not found.

Ludwig Fahrmeir, Thomas Kneib, Stefan Lang, and Brian Marx. *Regression: Models, Methods and Applications*. Springer, Berlin; Heidelberg, 2013. ISBN 978-3-642-34332-2. doi: 10.1007/978-3-642-34333-9.

Henri Funk, Ralf Ludwig, Helmut Küchenhoff, and Thomas Nagler. Towards more realistic climate model outputs: A multivariate bias correction based on zero-inflated vine copulas. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, page qlaf044, 2025. doi: 10.1093/jrsssc/qlaf044. Preprint version: arXiv:2410.15931 (2024).

Hans-Otto Georgii. *Stochastik: Eine Einführung in die Wahrscheinlichkeitstheorie und Statistik*. De Gruyter, Berlin; Boston, 5 edition, 2015. ISBN 978-3-11-035963-3. doi: 10.1515/9783110359701.

Changquan Huang and Alla Petukhina. *Applied Time Series Analysis and Forecasting with Python*. Statistics and Computing. Springer, Cham, 2022. ISBN 978-3-031-13583-5. doi: 10.1007/978-3-031-13584-2.

Samuel Karlin and Howard M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, New York, 1981. ISBN 978-0-12-398650-4.

Brett T. McClintock, Roland Langrock, Olivier Gimenez, Emmanuelle Cam, David L. Borchers, Richard Glennie, and Toby A. Patterson. Uncovering ecological state dynamics with hidden markov models. *Ecology Letters*, 23(12):1878–1903, 2020. doi: 10.1111/ele.13610.

Théo Michelot. *Analysing time series data with hidden Markov models in `hmmTMB`*, 2025. URL `https://theomichelot.r-universe.dev/hmmTMB/doc/readme`. R package vignette.

J. R. Norris. *Markov Chains*. Number 2 in Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 1997. ISBN 978-0-521-48181-6.

Barbara Rossi. *Advanced Econometric Methods*. MIT Press, 2018.

Josep A. Sánchez-Espigares and Alberto López-Moreno. Mswm examples. R package vignette, Dept. of Statistics and Operations Research, UPC–BarcelonaTech, 2021. URL `https://cran.r-project.org/web/packages/MSwM/vignettes/examples.pdf`.

Majla Sulejmani, Ferdinand Buchner, and Claudia Czado. D-vine regression with zero-inflated variables. 2025.

Walter Zucchini, Iain L. MacDonald, and Roland Langrock. *Hidden Markov Models for Time Series: An Introduction Using R*. Number 150 in Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, 2 edition, 2016. ISBN 978-1-4822-5383-2. doi: 10.1201/b20790.