Supplementary Material for “Vine Copula Models for Climate Trend Analysis using Canadian Temperature Data”

Haoxin Zhuang[1], Liqun Diao[*1], and Grace Y. Yi[2]

[1]Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada

[2]Department of Statistical and Actuarial Science, Department of Computer Science, University of Western Ontario, 1151 Richmond Street, London, ON, Canada,

December 10, 2020

Contents

[*]Corresponding author. Email: l2diao@uwaterloo.ca

1 Copula and Vine Copula Model

Copula (Joe, 1997; Nelsen, 2007) is a powerful and flexible tool to model the multivariate distribution and it allows separate models for marginal distribution and dependence structure. To cope with the restrictions of multivariate copula, a graphical model, *vine copula*, (Joe, 1997; Bedford and Cooke, 2002; Aas et al., 2009) was developed based on density decomposition and bivariate copulas and it can model the multivariate distribution flexibly.

Following Sklar (1959), the joint cumulative distribution of a d-dimensional random vector $(Y_1, \ldots, Y_d)$ can be written as

$$F(y_1, \ldots, y_d) = C(F_1(y_1), \ldots, F_d(y_d))$$

where $F_1(\cdot), \ldots, F_d(\cdot)$ are the marginal cumulative distribution function of $Y_1, \ldots, Y_d$ respectively, $C(\cdot)$ is the copula function defined on $[0, 1]^d$ as $C(u_1, \ldots, u_d) = P(U_1 \leq u_1, \ldots, U_d \leq u_d)$, and $U_1, \ldots, U_d$ are all random variable distributed as Uniform(0,1). Note that if the marginal distributions are all continuous, the copula C always uniquely exists.

In order to flexibly model the association relation in high dimension settings using copula model, the concept of vine copula was proposed by Bedford and Cooke (2002). After Aas et al. (2009) proposed the pair-copula construction idea, the regular vine (R-Vine) copula can be used to decompose a d-dimensional multivariate distribution into $d(d-1)/2$ bivariate distributions. This kind of decomposition enjoys convenience in parameters estimation and flexibility in modeling the dependence structure between random variables. Bedford and Cooke (2002) gave the definition of vine and R-Vine.

Definition 1. *$\mathcal{V}$ is a* vine *on d elements if:*

1. *$\mathcal{V}$=$(T_1, \ldots, T_m)$;*
2. *T_1 is a tree with nodes $N_1 = 1, \ldots, d$ and a set of edges denoted E_1;*
3. *For $i = 2, \ldots, m$, T_i is a tree with nodes $N_i \subset N_1 \cup E_1 \cup E_2 \cup \cdots \cup E_{i-1}$ and edge set E_i.*

A vine $\mathcal{V}$ is a regular vine *on d elements if:*

1. *$m = d - 1$;*
2. *T_i is a connected tree with edge set E_i and node set $N_i = E_{i-1}$, with the number of $N_i = d - (i - 1)$ for $i = 1, \ldots, n$, where the number of N_i is the cardinality of the set N_i;*
3. *The proximity condition holds: for $i = 2, \ldots, d - 1$, if $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ are two nodes in N_i connected by an edge (recall $a_1, a_2, b_1, b_2 \in N_{i-1}$), then the number of $a \cap b = 1$.*

Canonical vine (C-Vine) is a commonly used special case of R-Vine copula. Suppose the random variables are labeled as $1, \ldots, d$, and the C-Vine has trees $T_1, \ldots, T_{d-1}$. For formal definition of trees and other related concepts, refer to Bedford and Cooke (2002). For $i = 1, \ldots, d - 1$, T_i has $d + 1 - i$ nodes and $d - i$ edges. In T_1, the $d - 1$ edges represent the dependence between random variable 1 and other $d - 1$ random variables. In T_2, the $n - 2$ edges represent the dependence relation between random variable 2 and other remaining $d - 2$ random variables conditional on random variable 1. Similarly, for T_i, the $d - i$ edges represent the dependence relation between random variable i and other remaining $d - i$ random variables conditional on random variable $1, \ldots, i-1$. As a result, the n-dimensional joint density function of $U_1, \ldots, U_d \in [0, 1]^d$, $f(u_1, \ldots, u_d)$ can be decomposed as

$$f(u_1, \ldots, u_d; \theta) = \prod_{i=1}^{d-1} \prod_{k=i+1}^{d} c_{ik}(u_{i|\mathcal{D}_{ik}}, u_{k|\mathcal{D}_{ik}}; \theta_{ik})$$

where $c_{ik}(\cdot, \cdot)$ denotes the bivariate copula density of random variables i and k conditional on the conditioning set $\mathcal{D}_{ik} = \{u_1, \ldots, u_{i-1}\}$ and $\theta = \{\theta_{ik} : i = 1, \ldots, d-1; k = i+1, \ldots, d\}$ denotes the

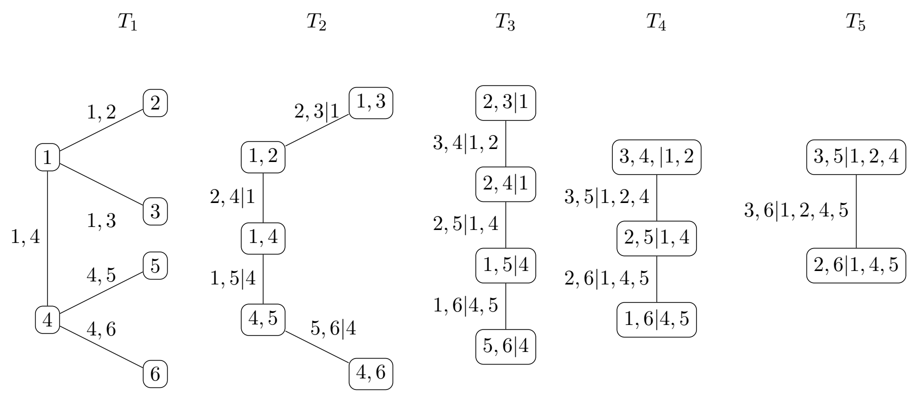


Figure 1: R-Vine with 6 random variables

parameter vector. Note that when $i = 1$, the conditioning set is empty. The conditional terms $u_{i|\mathcal{D}_{ik}}$ and $u_{k|\mathcal{D}_{ik}}$ can be calculated by applying the following formulas iteratively

$$u_{p|q} = \frac{\partial C_{pq}(u_p, u_q)}{\partial u_q},$$
$$u_{q|p} = \frac{\partial C_{pq}(u_p, u_q)}{\partial u_p}.$$

Example 1. *In Figure 1, each node in a specific vine tree represents a random variable or an edge in the previous tree. Each edge in a specific vine tree corresponds to the (conditional) bivariate copula that connects the variables of two nodes. Then the joint density of the six random variables* *can be expressed as*

$$
\begin{aligned}
f(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4,\varepsilon_5,\varepsilon_6) =& f(\varepsilon_1)\cdot f(\varepsilon_2)\cdot f(\varepsilon_3)\cdot f(\varepsilon_4)\cdot f(\varepsilon_5)\cdot f(\varepsilon_6)\\
&\cdot c_{12}(u_1,u_2)\cdot c_{13}(u_1,u_3)\cdot c_{14}(u_1,u_4)\cdot c_{45}(u_4,u_5)\cdot c_{46}(u_4,u_6)\\
&\cdot c_{23}(u_{2|1},u_{3|1})\cdot c_{24}(u_{2|1},u_{4|1})\cdot c_{15}(u_{1|4},u_{5|4})\cdot c_{56}(u_{5|4},u_{6|4})\\
&\cdot c_{34}(u_{3|12},u_{4|12})\cdot c_{25}(u_{2|14},u_{5|14})\cdot c_{16}(u_{1|45},u_{6|45})\\
&\cdot c_{35}(u_{3|124},u_{5|124})\cdot c_{26}(u_{2|145},u_{6|145})\\
&\cdot c_{36}(u_{3|1245},u_{6|1245})\\
=& f(\varepsilon_1)\cdot f(\varepsilon_2)\cdot f(\varepsilon_3)\cdot f(\varepsilon_4)\cdot f(\varepsilon_5)\cdot f(\varepsilon_6)\\
&\cdot \{c_{12}(u_1,u_2)\cdot c_{13}(u_1,u_3)\cdot c_{45}(u_4,u_5)\cdot c_{46}(u_4,u_6)\\
&\quad\cdot c_{23}(u_{2|1},u_{3|1})\cdot c_{56}(u_{5|4},u_{6|4})\}\\
&\cdot \{c_{14}(u_1,u_4)\cdot c_{24}(u_{2|1},u_{4|1})\cdot c_{15}(u_{1|4},u_{5|4})\\
&\quad\cdot c_{34}(u_{3|12},u_{4|12})\cdot c_{25}(u_{2|14},u_{5|14})\cdot c_{16}(u_{1|45},u_{6|45})\\
&\quad\cdot c_{35}(u_{3|124},u_{5|124})\cdot c_{26}(u_{2|145},u_{6|145})\cdot c_{36}(u_{3|1245},u_{6|1245})\}
\end{aligned}
$$

For the expression following the first equation, we start from the marginal distribution. The expressions in the five rows under the marginal distribution represent the (conditional) bivariate copulas in T_1 to T_5 respectively. The expression following the second equation is reorganized as the same form as in (3) in the main text. The conditioning set of each edge is clearly indicated in the expression. For example, the conditioning set of 3 and 4 is $\mathcal{D}_{34} = \{\varepsilon_1, \varepsilon_2\}$ from $c_{34}(u_{3|12}, u_{4|12})$ in the expression. And the conditional term $u_{3|12} = F(\varepsilon_3|\varepsilon_1, \varepsilon_2)$ is calculated as

$$
\begin{aligned}
F(\varepsilon_2|\varepsilon_1) &= \frac{\partial F(\varepsilon_1,\varepsilon_2)}{\partial F_1(\varepsilon_1)} = \frac{\partial C_{12}(F_1(\varepsilon_1),F_2(\varepsilon_2))}{\partial F_1(\varepsilon_1)} = \frac{\partial C_{12}(u_1,u_2)}{\partial u_1}\\
F(\varepsilon_3|\varepsilon_1) &= \frac{\partial F(\varepsilon_1,\varepsilon_3)}{\partial F_1(\varepsilon_1)} = \frac{\partial C_{13}(F_1(\varepsilon_1),F_3(\varepsilon_3))}{\partial F_1(\varepsilon_1)} = \frac{\partial C_{13}(u_1,u_3)}{\partial u_1}\\
F(\varepsilon_3|\varepsilon_1,\varepsilon_2) &= \frac{\partial F(\varepsilon_2|\varepsilon_1,\varepsilon_3|\varepsilon_1)}{\partial F(\varepsilon_2|\varepsilon_1)} = \frac{\partial C_{23}(F(\varepsilon_2|\varepsilon_1),F(\varepsilon_3|\varepsilon_1))}{\partial F(\varepsilon_2|\varepsilon_1)} = \frac{\partial C_{23}(u_{2|1},u_{3|1})}{\partial u_{2|1}}
\end{aligned}
$$

Next, we use an example with a 4-variable C-Vine to illustrate how we conduct prediction with our proposed model.

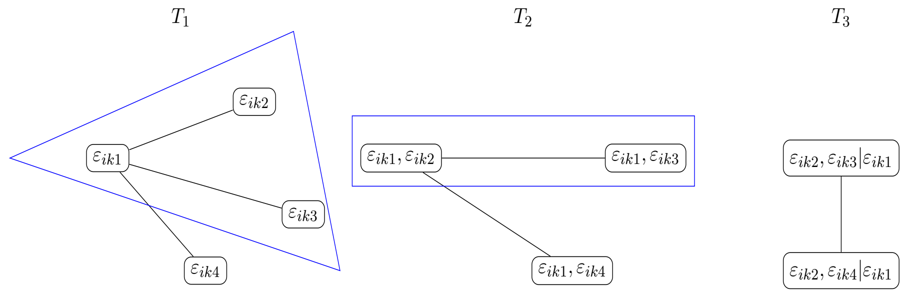


Figure 2: 3-variable C-Vine nested in original 4-variable C-Vine

Example 2. *In Figure 2, suppose that we already have the observations y_{ik1}, y_{ik2} y_{ik3}, the marginal parameter vector $(\hat{\beta}, \hat{\omega})$and the vine copula parameter vector $\hat{\theta}$. We want to predict the value of y_{ik4}. First, we calculate the marginal mean of y_{ikl} for $l = 1, 2, 3, 4$,*

$$\hat{\mu}_{ikl} = g_l^{-1}(x_{ikl}^T \hat{\beta}_l)$$

The copula data for $l = 1, 2, 3$ are calculated as,

$$\hat{\varepsilon}_{ikl} = y_{ikl} - \hat{\mu}_{ikl},$$
$$\hat{u}_{ikl} = F_l(\hat{\varepsilon}_{ikl}; \hat{\omega}_l).$$

Then the conditional distribution of ε_{ik4} can be expressed as

$$f(\varepsilon_{ik4}|\hat{\varepsilon}_{ik1}, \hat{\varepsilon}_{ik2}, \hat{\varepsilon}_{ik3}) = f_4(\varepsilon_{ik4}; \hat{\omega}_4) \cdot \prod_{r=1}^{3} c_{kr,k4}(\hat{u}_{ikr|\mathcal{D}_{ikr,ik4}}, u_{ik4|\mathcal{D}_{ikr,ik4}}; \hat{\theta}_{ikr,ik4})$$

where the conditional terms are calculated as

$$\begin{aligned}
\hat{u}_{ik2|ik1} &= \frac{\partial c_{k1,k2}(\hat{u}_{ik1}, \hat{u}_{ik2}; \hat{\theta}_{ik1,ik2})}{\partial \hat{u}_{ik1}} \\
\hat{u}_{ik3|ik1} &= \frac{\partial c_{k1,k3}(\hat{u}_{ik1}, \hat{u}_{ik3}; \hat{\theta}_{ik1,ik3})}{\partial \hat{u}_{ik1}} \\
u_{ik4|ik1} &= \frac{\partial c_{k1,k4}(\hat{u}_{ik1}, F_4(\varepsilon_{ik4}; \hat{\omega}_4); \hat{\theta}_{ik1,ik4})}{\partial \hat{u}_{ik1}} \\
\hat{u}_{ik3|ik1,ik2} &= \frac{\partial c_{k2,k3}(\hat{u}_{ik2|ik1}, \hat{u}_{ik3|ik1}; \hat{\theta}_{ik2,ik3})}{\partial \hat{u}_{ik2|ik1}} \\
u_{ik4|ik1,ik2} &= \frac{\partial c_{k2,k4}(\hat{u}_{ik2|ik1}, u_{ik4|ik1}; \hat{\theta}_{ik2,ik4})}{\partial \hat{u}_{ik2|ik1}}
\end{aligned}$$

And the predicted value of y_{ik4} is

$$\hat{y}_{ik4} = \int_{-\infty}^{\infty} \varepsilon_{ik4} f(\varepsilon_{ik4}|\hat{\varepsilon}_{ik1}, \hat{\varepsilon}_{ik2}, \hat{\varepsilon}_{ik3}) d\varepsilon_{ik4} + \hat{\mu}_{ik4}$$

2 Simulation

In this section, we provide detailed description of the simulation studies. The simulation studies consist of four parts: validity and efficiency, robustness, copula selection and prediction (subject extrapolation), in which we examine the finite sample performance of the proposed composite likelihood under simultaneous and two-stage estimation procedure under different scenarios. The four parts will be elaborated in the Sections 2.1-2.4, respectively.

Table 1: Copula functions and the values of the dependence parameters in the dependence structure within each time block

Bivariate Variable	$\varepsilon_{ik1}, \varepsilon_{ik2}$	$\varepsilon_{ik1}, \varepsilon_{ik3}$	$\varepsilon_{ik1}, \varepsilon_{ik4}$	$\varepsilon_{ik2}, \varepsilon_{ik3}\|\varepsilon_{ik1}$	$\varepsilon_{ik2}, \varepsilon_{ik4}\|\varepsilon_{ik1}$	$\varepsilon_{ik3}, \varepsilon_{ik4}\|\varepsilon_{ik1}, \varepsilon_{ik2}$
Copula Function	Clayton	Gumbel	Gaussian	Frank	Gaussian	Frank
			Strong Dependence			
Kendall's Tau	0.7	0.7	0.7	0.6	0.6	0.5
Dependence Parameter	$\theta_{k1,k2} = 4.67$	$\theta_{k1,k3} = 3.33$	$\theta_{k1,k4} = 0.89$	$\theta_{k2,k3} = 7.93$	$\theta_{k2,k4} = 0.81$	$\theta_{k3,k4} = 5.74$
			Moderate Dependence			
Kendall's Tau	0.4	0.4	0.4	0.3	0.3	0.2
Dependence Parameter	$\theta_{k1,k2} = 1.33$	$\theta_{k1,k3} = 1.67$	$\theta_{k1,k4} = 0.59$	$\theta_{k2,k3} = 2.92$	$\theta_{k2,k4} = 0.45$	$\theta_{k3,k4} = 1.86$

Table 2: Copula functions and the values of dependence parameters in dependence structure within time blocks for strong and moderate dependence settings

Bivariate Variables	$\varepsilon_{i11}, \varepsilon_{i12}$	$\varepsilon_{i11}, \varepsilon_{i13}$	$\varepsilon_{i11}, \varepsilon_{i14}$	$\varepsilon_{i12}, \varepsilon_{i13} \vert \varepsilon_{i11}$	$\varepsilon_{i12}, \varepsilon_{i14} \vert \varepsilon_{i11}$	$\varepsilon_{i13}, \varepsilon_{i14} \vert \varepsilon_{i11}, \varepsilon_{i12}$
Copula Function	Clayton	Gumbel	Gaussian	Frank	Gaussian	Frank
Strong Dependence	4.67	3.33	0.89	7.93	0.81	5.74
Moderate Dependence	1.33	1.67	0.59	2.92	0.45	1.86

Bivariate Variables	$\varepsilon_{i21}, \varepsilon_{i22}$	$\varepsilon_{i21}, \varepsilon_{i23}$	$\varepsilon_{i21}, \varepsilon_{i24}$	$\varepsilon_{i22}, \varepsilon_{i23} \vert \varepsilon_{i21}$	$\varepsilon_{i22}, \varepsilon_{i24} \vert \varepsilon_{i21}$	$\varepsilon_{i23}, \varepsilon_{i24} \vert \varepsilon_{i21}, \varepsilon_{i22}$
Copula Function	Joe	Clayton	Gumbel	Joe	Clayton	Joe
Strong Dependence	5.46	4.67	3.33	3.83	3.00	2.86
Moderate Dependence	2.22	1.33	1.67	1.77	0.86	1.44

Bivariate Variables	$\varepsilon_{i31}, \varepsilon_{i32}$	$\varepsilon_{i31}, \varepsilon_{i33}$	$\varepsilon_{i31}, \varepsilon_{i34}$	$\varepsilon_{i32}, \varepsilon_{i33} \vert \varepsilon_{i31}$	$\varepsilon_{i32}, \varepsilon_{i34} \vert \varepsilon_{i31}$	$\varepsilon_{i33}, \varepsilon_{i34} \vert \varepsilon_{i31}, \varepsilon_{i32}$
Copula Function	Frank	Gumbel	Gaussian	Frank	Gumbel	Frank
Strong Dependence	11.41	3.33	0.89	7.93	2.50	5.74
Moderate Dependence	4.16	1.67	0.59	2.92	1.43	1.86

Bivariate Variables	$\varepsilon_{i41}, \varepsilon_{i42}$	$\varepsilon_{i41}, \varepsilon_{i43}$	$\varepsilon_{i41}, \varepsilon_{i44}$	$\varepsilon_{i42}, \varepsilon_{i43} \vert \varepsilon_{i41}$	$\varepsilon_{i42}, \varepsilon_{i44} \vert \varepsilon_{i41}$	$\varepsilon_{i43}, \varepsilon_{i44} \vert \varepsilon_{i41}, \varepsilon_{i42}$
Copula Function	Joe	Clayton	Gumbel	Joe	Clayton	Joe
Strong Dependence	5.46	4.67	3.33	7.93	2.50	5.74
Moderate Dependence	2.22	1.33	1.67	2.92	1.43	1.86

Bivariate Variables	$\varepsilon_{i51}, \varepsilon_{i52}$	$\varepsilon_{i51}, \varepsilon_{i53}$	$\varepsilon_{i51}, \varepsilon_{i54}$	$\varepsilon_{i52}, \varepsilon_{i53} \vert \varepsilon_{i51}$	$\varepsilon_{i52}, \varepsilon_{i54} \vert \varepsilon_{i51}$	$\varepsilon_{i53}, \varepsilon_{i54} \vert \varepsilon_{i51}, \varepsilon_{i52}$
Copula Function	Clayton	Joe	Frank	Joe	Clayton	Joe
Strong Dependence	11.41	3.33	0.89	3.83	3.00	2.86
Moderate Dependence	4.16	1.67	0.59	1.77	0.86	1.44

2.1 Validity and Efficiency

In this subsection, we explore the validity and efficiency loss of the proposed composite likelihood method relative to the likelihood-based methods. We introduce various simulation settings in Section 2.1.1, describe evaluation metrics in Section 2.1.2, and report the results in Section 2.1.3.

2.1.1 Simulation Settings

We consider the scenarios where the sample size $n = 500$ or 1000, the number of time blocks is $a = 4$ and the number of time points in each time block is $b = 4$. The covariates x_{ikl} are independently generated from a uniformly distribution on $[0,5]$ for $i = 1, \dots, n$, $k = 1, \dots, a$ and $l = 1, \dots, b$. Suppose that the marginal model is

$$Y_{ikl} = \beta_{0l} + \beta_{1l} x_{ikl} + \beta_{2l} k + \varepsilon_{ikl}, \tag{1}$$

where $\varepsilon_{ikl} \sim N(0, \sigma_l^2)$, for $i = 1, \dots, n$, $k = 1, 2, 3, 4$ and $l = 1, 2, 3, 4$. We set the values of the marginal parameters as $\eta_l = (\beta_{0l}, \beta_{1l}, \beta_{2l}, \sigma_l)^{\mathrm{T}} = (l, l+1, l+2, 2)^{\mathrm{T}}$ for $l = 1, 2, 3, 4$.

In this subsection, we assume the error terms bear the R-Vine structure as demonstrated in Figure 1 in the main text and we further assume the conditional independence in tree structure T_4 and beyond for simplicity. We consider two scenarios where the dependence is either strong or moderate. For the scenario of strong or moderate dependence, the (conditional) bivariate copulas connecting the time blocks in T_1, T_2 and T_3 are all Gaussian(0.8) or Gaussian(0.5), a Gaussian copula with the parameter value shown in the brackets. More specifically, the bivariate copula functions and their corresponding parameter values for the C-Vine structure within each time block are given in Table 1. In the scenario of strong dependence, the Kendall's Taus of the bivariate copulas in T_1, T_2 and T_3 are set to be 0.7, 0.6 and 0.5, respectively; in that of moderate dependence, they are set to be 0.4, 0.3 and 0.2, respectively. The values of the dependence parameters are set to reach the desired degree of dependence. We generate the error terms ε_i from joint density (4) in the main text, in which the marginal distribution is normal and the dependence structure is the previously specified R-Vine; the values of Y_{ikl} are determined by (1).

The simulation is repeated 500 times. We compare the performance of the following four estimation methods:

(1) Method 1: full likelihood using simultaneous estimation procedure,

(2) Method 2: full likelihood using two-stage estimation procedure,

(3) Method 3: composite likelihood using simultaneous estimation procedure described in Section 3.1,

(4) Method 4: composite likelihood using two-stage estimation proceduredescribed in Section 3.2.

Note that the first stage of Method 2 and 4 are both using the marginal likelihood and essentially provide the same estimates for marginal parameters.

2.1.2 Evaluation Metrics

The following five evaluation metrics are used to evaluate different aspects of the estimators obtained by using the four estimation methods.

- *Empirical Bias (EBias)*: The difference between the average of the estimated values from 500 simulations and the true value of the parameters;
- *Empirical Standard Error (ESE)*: The sample standard deviation of the 500 estimates;
- *Asymptotic Standard Error (ASE)*: The average of 500 estimated asymptotic standard deviation of the estimators;
- *Empirical Coverage Probability (ECP)*: The proportion of the 500 confidence intervals that contain the true parameter value;
- *Asymptotic Efficiency (Efficiency)*: The ratio of the asymptotic variance of an estimator obtained from Method 2,3 or 4 relative to those of Method 1.

2.1.3 Simulation Results

We report the simulation results which include *EBias*, *ESE*, *ASE*, *ECP* and *Efficiency* for the four estimation methods. For the setting of strong dependence and $n = 500$, Table 3 summarizes the results for marginal parameters and dependence parameters. The results for strong dependence and $n = 1000$ are reported in Tables 4 and those for moderate dependence and $n = 500, 1000$ are summarized in Tables 5 and 6.

The results in Table 3 show that when dependence is strong and sample size is 500, the finite sample biases for the estimates of the marginal parameters η obtained from all four estimation methods are fairly small, ASEs and ESEs are close to each other, and ECP is close to the nominal level 95%. These results suggest that the proposed composite likelihood methods (i.e. Method 3 and 4) yield consistent estimates. However, these methods may incur noticeable efficiency loss; Method 3 is more efficient than Method 4, as expected. Similar patterns are observed for the estimates of the dependence parameters within blocks θ, as shown in Table 3.

As expected, the performance of the four methods becomes better as the sample size increases, as displayed in Tables 4 and 6. Regarding the efficiency loss incurred by the composite likelihood methods, it becomes less severe when the dependence among the response components is weaker, as illustrated in Tables 5 and 6. We notice that the efficiency loss remains stable as the sample size increases by exploring the performance under the settings with $n = 1000$. Generally speaking, the efficiency loss of using the simultaneous composite likelihood (i.e., Method 3) is mild to moderate for within-block parameters θ, while the computational time is significantly reduced compared to using the simultaneous full likelihood (i.e., Method 1). The efficiency loss of coefficient estimators β using the two-stage estimation procedure is obviously more severe by further ignoring dependence structure within blocks. In Table 3, the two-stage estimation procedures based on full likelihood and composite likelihood (Method 2 and 4) suffer from a similar amount of efficiency loss when estimating the dependence parameters θ, suggesting that the efficiency loss is mainly due to the variation introduced from the first stage when estimating marginal parameters, and is not aggravated much by making working conditional independence assumptions. Under the moderate dependence setting, the two-stage procedure still leads to significant efficiency loss on marginal parameters, but comparable and mild efficiency loss on dependence parameters using both full likelihood and composite likelihood as shown in Tables 5 and 6.

In summary, all four methods are valid and provide consistent results for the estimation of parameters of the models. Simultaneous composite likelihood provides consistent estimates for all within-block parameters with moderate efficiency loss, even when the sample size is small and dependence is strong, while the two-stage estimation procedure of full likelihood and composite likelihood could introduce biases and significant efficiency loss under the strong dependence structure, although it can greatly speed up the estimation process.

Table 3: Simulation results using the four estimation methods: strong dependence and $n = 500$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	0.317	0.424	0.179	0.141	−0.008	−0.017	0.021	0.010	−0.032	−0.015	−0.012	0.035	−0.364	−0.332	−0.422	−0.354	−0.118	−0.072	−0.017	1.963	−0.012	0.744
	ESE	0.064	0.065	0.067	0.068	0.004	0.003	0.004	0.006	0.007	0.006	0.007	0.010	0.031	0.032	0.032	0.034	0.180	0.076	0.004	0.219	0.008	0.184
	ASE	0.065	0.066	0.066	0.070	0.003	0.003	0.003	0.006	0.007	0.006	0.007	0.010	0.031	0.032	0.032	0.034	0.177	0.071	0.004	0.220	0.007	0.182
	ECP	0.958	0.963	0.948	0.956	0.938	0.938	0.928	0.948	0.954	0.950	0.950	0.948	0.954	0.948	0.948	0.946	0.951	0.951	0.951	0.949	0.951	0.954
Method 2: Full likelihood Two−stage Estimation	EBias*	0.745	0.316	−0.107	0.389	0.158	0.010	0.096	−0.130	−0.081	−0.011	0.039	0.060	−0.910	−0.799	−1.057	−0.920	−10.607	−3.791	−0.233	−21.440	−0.364	−6.405
	ESE	0.119	0.116	0.121	0.123	0.031	0.030	0.032	0.033	0.013	0.012	0.011	0.010	0.053	0.054	0.057	0.058	0.221	0.110	0.007	0.266	0.009	0.218
	ASE	0.119	0.119	0.119	0.119	0.031	0.031	0.031	0.031	0.013	0.012	0.011	0.011	0.056	0.055	0.058	0.060	0.228	0.107	0.006	0.302	0.009	0.246
	ECP	0.953	0.958	0.943	0.948	0.945	0.940	0.943	0.938	0.958	0.945	0.958	0.963	0.953	0.945	0.950	0.950	0.928	0.940	0.930	0.913	0.933	0.953
	Efficiency	0.299	0.309	0.308	0.343	0.012	0.008	0.012	0.036	0.278	0.264	0.423	0.827	0.305	0.336	0.298	0.319	0.601	0.441	0.460	0.533	0.613	0.550
Method 3: Composite likelihood Simultaneous Estimation	EBias*	0.256	0.378	0.166	0.047	0.038	−0.022	−0.004	0.011	−0.041	0.005	0.029	0.076	−0.709	−0.733	−0.838	−0.773	−0.256	−0.472	−0.052	1.418	−0.051	0.310
	ESE	0.086	0.088	0.090	0.090	0.006	0.005	0.005	0.006	0.016	0.017	0.018	0.019	0.039	0.042	0.044	0.046	0.241	0.111	0.006	0.241	0.009	0.185
	ASE	0.089	0.091	0.091	0.094	0.006	0.005	0.005	0.006	0.017	0.018	0.019	0.020	0.041	0.044	0.045	0.048	0.235	0.105	0.006	0.252	0.008	0.185
	ECP	0.950	0.963	0.958	0.963	0.948	0.950	0.935	0.945	0.965	0.965	0.960	0.953	0.958	0.958	0.960	0.963	0.943	0.948	0.935	0.956	0.955	0.953
	Efficiency	0.535	0.530	0.526	0.554	0.345	0.330	0.455	0.980	0.166	0.118	0.145	0.241	0.556	0.516	0.493	0.494	0.567	0.458	0.501	0.766	0.760	0.974
Method 4: Composite likelihood Two−stage Estimation	EBias*	0.745	0.316	−0.107	0.389	0.158	0.010	0.096	−0.130	−0.081	−0.011	0.039	0.060	−0.910	−0.799	−1.057	−0.920	−6.816	−2.362	−0.157	−19.722	−0.461	−7.777
	ESE	0.119	0.116	0.121	0.123	0.031	0.030	0.032	0.033	0.013	0.012	0.011	0.010	0.053	0.054	0.057	0.058	0.253	0.129	0.007	0.274	0.010	0.222
	ASE	0.119	0.119	0.119	0.119	0.031	0.031	0.031	0.031	0.013	0.012	0.011	0.011	0.056	0.055	0.058	0.060	0.249	0.121	0.007	0.306	0.010	0.247
	ECP	0.953	0.958	0.943	0.948	0.945	0.940	0.943	0.938	0.958	0.945	0.958	0.963	0.953	0.945	0.950	0.950	0.935	0.935	0.940	0.910	0.918	0.950
	Efficiency	0.299	0.309	0.308	0.343	0.012	0.008	0.012	0.036	0.278	0.264	0.423	0.827	0.305	0.336	0.298	0.319	0.505	0.347	0.393	0.517	0.545	0.547

[1] EBias*=EBias×10^2

Table 4: Simulation results using the four estimation methods: strong dependence and $n = 1000$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	−0.135	−0.069	−0.156	−0.195	−0.013	−0.001	−0.002	0.002	0.003	−0.007	< 0.001	0.031	−0.041	0.006	−0.069	−0.008	0.753	0.077	0.001	1.080	−0.007	0.573
	ESE	0.047	0.048	0.049	0.051	0.002	0.002	0.002	0.004	0.005	0.004	0.005	0.007	0.022	0.022	0.023	0.024	0.131	0.053	0.003	0.163	0.005	0.132
	ASE	0.046	0.047	0.047	0.049	0.002	0.002	0.002	0.004	0.005	0.004	0.005	0.007	0.022	0.022	0.023	0.024	0.125	0.050	0.003	0.155	0.005	0.129
	ECP	0.948	0.943	0.938	0.945	0.963	0.955	0.955	0.933	0.958	0.940	0.943	0.955	0.943	0.958	0.945	0.943	0.940	0.960	0.950	0.953	0.938	0.943
Method 2: Full likelihood Two−stage Estimation	EBias*	−0.098	0.408	−0.317	0.316	0.018	−0.145	0.077	−0.166	−0.008	−0.020	0.045	0.039	−0.153	−0.112	−0.207	−0.133	−4.422	−1.612	−0.096	−11.068	−0.181	−1.911
	ESE	0.087	0.085	0.082	0.086	0.023	0.021	0.021	0.021	0.009	0.008	0.008	0.008	0.040	0.040	0.042	0.043	0.160	0.077	0.004	0.189	0.006	0.142
	ASE	0.084	0.084	0.084	0.084	0.022	0.022	0.022	0.022	0.009	0.008	0.008	0.008	0.040	0.039	0.042	0.042	0.159	0.075	0.004	0.195	0.006	0.154
	ECP	0.928	0.950	0.955	0.933	0.943	0.943	0.963	0.958	0.953	0.940	0.950	0.940	0.938	0.948	0.938	0.948	0.928	0.940	0.935	0.913	0.940	0.956
	Efficiency	0.298	0.306	0.306	0.342	0.012	0.008	0.012	0.036	0.279	0.264	0.423	0.829	0.299	0.329	0.293	0.313	0.617	0.451	0.478	0.633	0.677	0.701
Method 3: Composite likelihood Simultaneous Estimation	EBias*	−0.010	0.058	−0.137	−0.174	−0.017	−0.033	−0.004	0.002	−0.011	−0.006	0.014	0.042	−0.154	−0.114	−0.186	−0.122	0.734	0.034	−0.010	0.743	−0.022	0.463
	ESE	0.064	0.067	0.067	0.069	0.004	0.003	0.004	0.004	0.011	0.012	0.013	0.014	0.029	0.031	0.033	0.035	0.177	0.078	0.004	0.186	0.006	0.134
	ASE	0.063	0.064	0.064	0.066	0.004	0.003	0.004	0.004	0.012	0.013	0.013	0.014	0.029	0.031	0.032	0.034	0.167	0.074	0.004	0.178	0.006	0.130
	ECP	0.938	0.935	0.938	0.938	0.948	0.953	0.935	0.930	0.955	0.955	0.955	0.963	0.933	0.933	0.930	0.935	0.943	0.953	0.935	0.940	0.953	0.953
	Efficiency	0.535	0.528	0.524	0.553	0.349	0.331	0.456	0.996	0.166	0.117	0.144	0.240	0.549	0.509	0.487	0.488	0.562	0.455	0.499	0.760	0.755	0.977
Method 4: Composite likelihood Two−stage Estimation	EBias*	−0.098	0.408	−0.317	0.316	0.018	−0.145	0.077	−0.166	−0.008	−0.020	0.045	0.039	−0.153	−0.112	−0.207	−0.133	−2.564	−0.898	−0.059	−10.215	−0.229	−2.577
	ESE	0.087	0.085	0.082	0.086	0.023	0.021	0.021	0.021	0.009	0.008	0.008	0.008	0.040	0.040	0.042	0.043	0.184	0.091	0.005	0.201	0.007	0.145
	ASE	0.084	0.084	0.084	0.084	0.022	0.022	0.022	0.022	0.009	0.008	0.008	0.008	0.040	0.039	0.042	0.042	0.177	0.086	0.005	0.204	0.007	0.155
	ECP	0.928	0.950	0.955	0.933	0.943	0.943	0.963	0.958	0.953	0.940	0.950	0.940	0.938	0.948	0.938	0.948	0.940	0.938	0.948	0.925	0.935	0.956
	Efficiency	0.298	0.306	0.306	0.342	0.012	0.008	0.012	0.036	0.279	0.264	0.423	0.829	0.299	0.329	0.293	0.313	0.503	0.343	0.396	0.583	0.599	0.694

[1] EBias*=EBias×10^2

Table 5: Simulation results using the four estimation methods: moderate dependence and $n = 500$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	0.133	0.151	0.363	−0.550	0.012	0.017	0.081	0.050	0.008	0.028	0.007	0.157	−0.386	−0.288	0.520	−0.276	0.230	−0.372	−0.066	−0.272	−0.023	0.279
	ESE	0.093	0.093	0.102	0.113	0.011	0.011	0.013	0.022	0.024	0.022	0.025	0.032	0.034	0.039	0.040	0.036	0.101	0.071	0.018	0.141	0.017	0.139
	ASE	0.093	0.097	0.099	0.115	0.011	0.012	0.012	0.021	0.024	0.023	0.025	0.032	0.036	0.041	0.041	0.036	0.099	0.069	0.016	0.146	0.018	0.140
	ECP	0.955	0.960	0.943	0.955	0.950	0.950	0.925	0.955	0.948	0.965	0.940	0.950	0.965	0.948	0.953	0.935	0.930	0.940	0.958	0.953	0.950	0.950
Method 2: Full likelihood Two−stage Estimation	EBias*	0.273	0.307	−0.618	−0.291	−0.125	0.033	0.205	−0.021	−0.017	−0.128	0.047	0.147	−0.524	−0.510	−0.720	−0.381	−1.032	−1.162	−0.181	−1.724	−0.097	−0.023
	ESE	0.126	0.127	0.133	0.132	0.031	0.030	0.032	0.033	0.029	0.027	0.029	0.033	0.040	0.045	0.048	0.039	0.106	0.077	0.019	0.142	0.017	0.139
	ASE	0.127	0.130	0.131	0.131	0.031	0.031	0.031	0.031	0.029	0.028	0.029	0.033	0.042	0.046	0.047	0.038	0.104	0.074	0.017	0.148	0.018	0.141
	ECP	0.955	0.965	0.940	0.955	0.955	0.938	0.955	0.923	0.943	0.955	0.940	0.950	0.950	0.948	0.938	0.935	0.938	0.963	0.955	0.948	0.953	0.958
	Efficiency	0.529	0.556	0.573	0.762	0.135	0.146	0.142	0.471	0.661	0.690	0.753	0.938	0.745	0.769	0.756	0.882	0.911	0.881	0.905	0.966	0.977	0.990
Method 3: Composite likelihood Simultaneous Estimation	EBias*	−0.241	0.246	0.323	−0.576	0.086	−0.048	0.037	0.051	−0.007	−0.025	0.054	0.180	−0.470	−0.452	−0.633	−0.352	0.230	−0.448	−0.096	−0.517	−0.053	0.246
	ESE	0.098	0.100	0.109	0.116	0.013	0.015	0.015	0.022	0.025	0.023	0.026	0.032	0.037	0.043	0.044	0.038	0.111	0.078	0.019	0.145	0.017	0.139
	ASE	0.100	0.105	0.106	0.117	0.013	0.015	0.014	0.022	0.025	0.025	0.027	0.033	0.039	0.044	0.044	0.037	0.107	0.073	0.017	0.151	0.018	0.141
	ECP	0.945	0.960	0.928	0.960	0.958	0.958	0.928	0.948	0.958	0.960	0.943	0.955	0.960	0.943	0.938	0.930	0.935	0.955	0.960	0.943	0.958	0.953
	Efficiency	0.861	0.847	0.875	0.952	0.718	0.621	0.700	0.960	0.858	0.854	0.880	0.950	0.878	0.865	0.882	0.942	0.849	0.894	0.927	0.932	0.987	0.996
Method 4: Composite likelihood Two−stage Estimation	EBias*	0.273	0.307	−0.618	−0.291	−0.125	0.033	0.205	−0.021	−0.017	−0.128	0.047	0.147	−0.524	−0.510	−0.720	−0.381	−0.741	−1.073	−0.165	−1.885	−0.120	−0.042
	ESE	0.126	0.127	0.133	0.132	0.031	0.030	0.032	0.033	0.029	0.027	0.029	0.033	0.040	0.045	0.048	0.039	0.111	0.079	0.019	0.144	0.017	0.139
	ASE	0.127	0.130	0.131	0.131	0.031	0.031	0.031	0.031	0.029	0.028	0.029	0.033	0.042	0.046	0.047	0.038	0.108	0.075	0.017	0.151	0.018	0.141
	ECP	0.955	0.965	0.940	0.955	0.955	0.938	0.955	0.923	0.943	0.955	0.940	0.950	0.950	0.948	0.938	0.935	0.938	0.965	0.963	0.950	0.953	0.953
	Efficiency	0.529	0.556	0.573	0.762	0.135	0.146	0.142	0.471	0.661	0.690	0.753	0.938	0.745	0.769	0.756	0.882	0.846	0.853	0.891	0.934	0.980	0.990

[1] EBias*=EBias$\times 10^2$

Table 6: Simulation results using the four estimation methods: moderate dependence and $n = 1000$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*	−0.185	−0.001	−0.309	−0.253	−0.078	−0.015	−0.002	−0.011	0.066	−0.023	0.030	0.108	−0.005	0.125	−0.077	0.046	0.691	0.044	0.025	0.106	−0.012	0.449
	ESE	0.064	0.070	0.071	0.083	0.008	0.008	0.008	0.016	0.016	0.016	0.018	0.023	0.026	0.029	0.030	0.027	0.074	0.052	0.013	0.107	0.013	0.105
	ASE	0.066	0.069	0.070	0.081	0.008	0.008	0.008	0.015	0.017	0.017	0.018	0.023	0.026	0.029	0.029	0.026	0.070	0.049	0.012	0.103	0.012	0.099
	ECP	0.965	0.940	0.945	0.948	0.955	0.960	0.963	0.933	0.953	0.943	0.955	0.963	0.948	0.960	0.948	0.935	0.945	0.953	0.955	0.945	0.958	0.950
Method 2: Full likelihood Two−stage Estimation	EBias*	−0.666	0.295	−0.742	0.081	0.072	−0.085	0.140	−0.159	0.105	−0.068	0.098	0.127	−0.072	0.078	−0.175	0.023	0.116	−0.306	0.019	−0.601	−0.039	0.328
	ESE	0.085	0.091	0.089	0.095	0.022	0.021	0.021	0.022	0.020	0.020	0.020	0.024	0.030	0.033	0.035	0.029	0.078	0.055	0.013	0.106	0.013	0.105
	ASE	0.090	0.092	0.092	0.093	0.022	0.022	0.022	0.022	0.020	0.020	0.020	0.024	0.030	0.033	0.034	0.027	0.074	0.053	0.012	0.104	0.013	0.099
	ECP	0.970	0.963	0.960	0.953	0.963	0.953	0.963	0.958	0.958	0.958	0.948	0.945	0.950	0.955	0.950	0.930	0.945	0.943	0.955	0.948	0.960	0.940
	Efficiency	0.528	0.553	0.573	0.759	0.134	0.144	0.140	0.470	0.662	0.690	0.753	0.937	0.739	0.760	0.750	0.873	0.898	0.871	0.896	0.973	0.973	0.994
Method 3: Composite likelihood Simultaneous Estimation	EBias*	−0.070	0.241	−0.244	−0.217	−0.071	−0.082	−0.019	−0.011	0.024	−0.057	0.028	0.094	−0.045	0.057	−0.114	0.021	0.693	0.015	0.015	−0.031	−0.023	0.455
	ESE	0.070	0.076	0.077	0.086	0.009	0.010	0.010	0.016	0.017	0.017	0.019	0.024	0.028	0.031	0.033	0.028	0.081	0.056	0.013	0.111	0.013	0.105
	ASE	0.071	0.075	0.075	0.083	0.009	0.011	0.010	0.015	0.018	0.018	0.019	0.023	0.027	0.031	0.031	0.026	0.076	0.052	0.012	0.107	0.013	0.099
	ECP	0.958	0.945	0.943	0.945	0.955	0.950	0.938	0.935	0.960	0.960	0.968	0.965	0.945	0.950	0.928	0.920	0.935	0.950	0.950	0.935	0.958	0.945
	Efficiency	0.864	0.847	0.876	0.949	0.727	0.615	0.699	0.989	0.861	0.854	0.882	0.951	0.873	0.857	0.877	0.934	0.844	0.888	0.918	0.934	0.981	0.996
Method 4: Composite likelihood Two−stage Estimation	EBias*	−0.666	0.295	−0.742	0.081	0.072	−0.085	0.140	−0.159	0.105	−0.068	0.098	0.127	−0.072	0.078	−0.175	0.023	0.244	−0.274	−0.012	−0.696	−0.049	0.321
	ESE	0.085	0.091	0.089	0.095	0.022	0.021	0.021	0.022	0.020	0.020	0.020	0.024	0.030	0.033	0.035	0.029	0.082	0.057	0.013	0.110	0.013	0.105
	ASE	0.090	0.092	0.092	0.093	0.022	0.022	0.022	0.022	0.020	0.020	0.020	0.024	0.030	0.033	0.034	0.027	0.077	0.053	0.012	0.106	0.013	0.099
	ECP	0.970	0.963	0.960	0.953	0.963	0.953	0.963	0.958	0.958	0.958	0.948	0.945	0.950	0.955	0.950	0.930	0.943	0.950	0.948	0.933	0.958	0.940
	Efficiency	0.528	0.553	0.573	0.759	0.134	0.144	0.140	0.470	0.662	0.690	0.753	0.937	0.739	0.760	0.750	0.873	0.835	0.842	0.882	0.939	0.976	0.993

[1] EBias*=EBias×10^2

2.2 Robustness

In this section, we examine the robustness of the simultaneous and two-stage composite likelihood estimation procedures (i.e., Method 3 and Method 4 in Section 2.1) in contrast to the counterparts based on full likelihood formulation (i.e., Method 1 and Method 2 in Section 2.1).

2.2.1 Simulation Settings

The simulation studies are the same settings as those in Section 2.1.1. To examine how the four methods behave when the dependence structure connecting different time blocks is misspecified, we simulate data from settings where all (conditional) bivariate copulas connecting the time blocks are all specified as Frank(7.93) for strong dependence and Frank(2.92) for moderate dependence setting, respectively, but we assume them to be Gaussian copula functions for model fitting.

2.2.2 Simulation Results

We report the performance of the four estimation methods in terms of the same evaluation metrics as described in Section 2.1.2. The results for the strong dependence or moderate dependence and $n = 500$ or $n = 1000$ are summarized in Tables 7-10.

Under simultaneous estimation procedure, the full likelihood fails to provide consistent estimators for both marginal and dependence parameters, with non-ignorable empirical biases, gaps between ASEs and ESEs, and discrepancies between the ECPs and the 95% nominal level. These patterns are not improved by increasing the sample size, while they are less severe for a weaker dependence. The full likelihood based two-stage estimation provides inefficient yet valid estimators for marginal parameters but invalid results for the dependence parameters. Both simultaneous and two-stage estimation procedures based on the proposed composite likelihood function (Method 3 and 4) provide valid results for both marginal parameters $(\beta^{\mathrm{T}}, \omega^{\mathrm{T}})^{\mathrm{T}}$ and dependence parameters θ within time blocks; estimators using Method 3 incur less finite sample biases and are more efficient than Method 4. Obviously, the proposed composite likelihood provide robustness with respect to misspecification of dependence structure linking time blocks.

Table 7: Simulation results using the four estimation methods when block-connecting structure is misspecified: strong dependence and $n = 500$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	4.035	7.165	4.057	4.476	−0.002	−0.020	0.015	0.005	−0.070	−0.034	−0.047	−0.005	−9.798	−7.674	−9.916	−9.548	−59.683	−15.620	−1.371	−27.232	0.025	5.794
	ESE	0.090	0.090	0.094	0.094	0.005	0.004	0.004	0.006	0.010	0.010	0.011	0.013	0.041	0.047	0.047	0.051	0.249	0.115	0.008	0.244	0.008	0.184
	ASE	0.069	0.073	0.072	0.076	0.005	0.004	0.004	0.006	0.011	0.012	0.012	0.014	0.034	0.035	0.036	0.038	0.180	0.083	0.006	0.220	0.008	0.185
	ECP	0.793	0.726	0.778	0.810	0.945	0.971	0.924	0.948	0.953	0.971	0.968	0.945	0.236	0.432	0.277	0.344	0.232	0.623	0.601	0.790	0.950	0.927
Method 2: Full likelihood Two−stage Estimation	EBias*	0.884	0.378	−0.016	0.375	−0.200	0.004	0.105	−0.099	−0.106	−0.059	0.005	0.029	−0.685	−0.643	−0.954	−0.804	−42.505	−11.992	−0.885	−27.852	−0.531	−0.137
	ESE	0.120	0.118	0.122	0.124	0.031	0.029	0.032	0.034	0.015	0.014	0.012	0.013	0.048	0.051	0.056	0.057	0.237	0.121	0.008	0.284	0.009	0.222
	ASE	0.119	0.120	0.120	0.120	0.031	0.031	0.031	0.031	0.015	0.014	0.013	0.013	0.049	0.052	0.057	0.058	0.350	0.186	0.011	0.392	0.015	0.269
	ECP	0.955	0.945	0.950	0.938	0.945	0.940	0.945	0.940	0.948	0.960	0.958	0.950	0.958	0.948	0.963	0.955	0.764	0.900	0.821	0.904	0.956	0.961
Method 3: Composite likelihood Simultaneous Estimation	EBias*	0.354	0.374	0.250	0.097	0.019	−0.027	−0.006	0.012	−0.092	−0.040	−0.035	0.014	−0.580	−0.617	−0.729	−0.662	−0.515	−0.601	−0.050	1.335	−0.028	0.561
	ESE	0.088	0.090	0.091	0.093	0.006	0.005	0.005	0.006	0.017	0.018	0.019	0.020	0.039	0.043	0.045	0.047	0.224	0.107	0.006	0.216	0.007	0.184
	ASE	0.085	0.088	0.091	0.091	0.006	0.005	0.005	0.006	0.016	0.017	0.018	0.020	0.037	0.041	0.043	0.046	0.224	0.101	0.006	0.231	0.008	0.184
	ECP	0.960	0.963	0.955	0.960	0.940	0.945	0.943	0.945	0.958	0.955	0.963	0.963	0.960	0.958	0.953	0.948	0.943	0.955	0.945	0.955	0.945	0.958
Method 4: Composite likelihood Two−stage Estimation	EBias*	0.884	0.378	−0.016	0.375	−0.200	0.004	0.105	−0.099	−0.106	−0.059	0.005	0.029	−0.685	−0.643	−0.954	−0.804	−6.902	−2.448	−0.156	−18.732	−0.418	−7.898
	ESE	0.120	0.118	0.122	0.124	0.031	0.029	0.032	0.034	0.015	0.014	0.012	0.013	0.048	0.051	0.056	0.057	0.239	0.123	0.007	0.251	0.009	0.223
	ASE	0.119	0.120	0.120	0.120	0.031	0.031	0.031	0.031	0.015	0.014	0.013	0.013	0.049	0.052	0.057	0.058	0.240	0.117	0.006	0.288	0.009	0.245
	ECP	0.955	0.945	0.950	0.938	0.945	0.940	0.945	0.940	0.948	0.960	0.958	0.950	0.958	0.948	0.963	0.955	0.940	0.943	0.943	0.920	0.913	0.945

[1] EBias*=EBias$\times 10^2$

Table 8: Simulation results using the four estimation methods when block-connecting structure is misspecified: strong dependence and $n = 1000$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	3.976	7.162	4.118	4.606	−0.005	0.003	−0.005	−0.010	−0.007	−0.011	−0.029	< 0.001	−9.440	−7.182	−9.389	−9.014	−58.638	−14.465	−1.290	−27.151	0.018	5.767
	ESE	0.067	0.069	0.069	0.071	0.003	0.003	0.003	0.004	0.007	0.008	0.008	0.010	0.031	0.036	0.035	0.038	0.187	0.081	0.005	0.188	0.006	0.134
	ASE	0.047	0.049	0.049	0.052	0.003	0.003	0.003	0.004	0.008	0.008	0.008	0.009	0.023	0.024	0.024	0.026	0.121	0.055	0.004	0.155	0.005	0.130
	ECP	0.715	0.595	0.735	0.718	0.971	0.964	0.951	0.932	0.971	0.958	0.955	0.945	0.087	0.243	0.123	0.152	0.113	0.482	0.337	0.612	0.945	0.930
Method 2: Full likelihood Two−stage Estimation	EBias*	−0.180	0.493	−0.280	0.304	0.051	−0.141	0.092	−0.140	0.005	−0.037	0.031	0.036	−0.152	−0.056	−0.246	−0.150	−37.137	−9.330	−0.720	−18.848	0.667	4.154
	ESE	0.087	0.085	0.082	0.087	0.023	0.021	0.021	0.021	0.011	0.010	0.009	0.009	0.036	0.038	0.042	0.043	0.174	0.088	0.005	0.210	0.007	0.146
	ASE	0.085	0.085	0.085	0.085	0.022	0.022	0.022	0.022	0.011	0.010	0.009	0.009	0.035	0.037	0.040	0.041	0.209	0.108	0.006	0.232	0.008	0.159
	ECP	0.938	0.960	0.958	0.940	0.943	0.950	0.955	0.955	0.955	0.960	0.948	0.945	0.943	0.928	0.943	0.938	0.475	0.881	0.731	0.875	0.833	0.952
Method 3: Composite likelihood Simultaneous Estimation	EBias*	0.171	0.196	0.093	0.024	−0.003	−0.027	−0.011	0.002	−0.081	−0.060	−0.058	−0.026	−0.133	−0.093	−0.157	−0.090	0.437	−0.050	−0.012	1.020	−0.007	0.607
	ESE	0.065	0.069	0.069	0.071	0.004	0.003	0.004	0.004	0.011	0.012	0.013	0.014	0.028	0.030	0.032	0.035	0.168	0.074	0.004	0.172	0.006	0.135
	ASE	0.062	0.064	0.064	0.066	0.004	0.003	0.004	0.004	0.012	0.012	0.013	0.014	0.028	0.030	0.032	0.034	0.159	0.071	0.004	0.163	0.005	0.130
	ECP	0.943	0.940	0.945	0.948	0.950	0.950	0.935	0.933	0.963	0.963	0.953	0.953	0.925	0.943	0.940	0.938	0.938	0.945	0.945	0.950	0.955	0.953
Method 4: Composite likelihood Two−stage Estimation	EBias*	−0.180	0.493	−0.280	0.304	0.051	−0.141	0.092	−0.140	0.005	−0.037	0.031	0.036	−0.152	−0.056	−0.246	−0.150	−2.782	−0.952	−0.062	−9.450	−0.200	−2.754
	ESE	0.087	0.085	0.082	0.087	0.023	0.021	0.021	0.021	0.011	0.010	0.009	0.009	0.036	0.038	0.042	0.043	0.176	0.087	0.005	0.191	0.006	0.146
	ASE	0.085	0.085	0.085	0.085	0.022	0.022	0.022	0.022	0.011	0.010	0.009	0.009	0.035	0.037	0.040	0.041	0.170	0.084	0.005	0.190	0.006	0.154
	ECP	0.938	0.960	0.958	0.940	0.943	0.950	0.955	0.955	0.955	0.960	0.948	0.945	0.943	0.928	0.943	0.938	0.948	0.945	0.940	0.930	0.938	0.953

[1] EBias*=EBias$\times 10^2$

Table 9: Simulation results using the four estimation methods when block-connecting structure is misspecified: moderate dependence and $n = 500$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	−1.573	0.412	−1.902	−1.362	0.016	−0.016	0.082	0.044	−0.006	−0.004	−0.024	0.161	−2.284	−0.022	−2.062	−0.834	−3.260	−1.688	−0.382	−4.933	0.004	0.116
	ESE	0.097	0.096	0.107	0.115	0.012	0.013	0.014	0.022	0.026	0.025	0.028	0.033	0.030	0.038	0.037	0.036	0.101	0.067	0.017	0.144	0.017	0.139
	ASE	0.096	0.103	0.103	0.117	0.012	0.013	0.012	0.021	0.027	0.028	0.028	0.034	0.033	0.038	0.038	0.035	0.095	0.065	0.016	0.147	0.018	0.140
	ECP	0.947	0.960	0.937	0.952	0.955	0.957	0.915	0.955	0.947	0.972	0.947	0.952	0.905	0.960	0.912	0.932	0.930	0.940	0.948	0.920	0.967	0.952
Method 2: Full likelihood Two−stage Estimation	EBias*	0.163	0.264	−0.853	−0.428	−0.136	0.041	0.273	0.014	0.005	−0.157	0.057	0.149	−0.396	−0.391	−0.597	−0.384	−3.294	−0.548	−0.211	−5.544	−0.090	0.065
	ESE	0.130	0.130	0.134	0.134	0.031	0.030	0.032	0.033	0.033	0.031	0.032	0.035	0.034	0.040	0.043	0.037	0.101	0.071	0.018	0.143	0.017	0.140
	ASE	0.131	0.134	0.134	0.133	0.031	0.031	0.031	0.031	0.033	0.032	0.033	0.035	0.036	0.041	0.042	0.036	0.099	0.068	0.016	0.149	0.018	0.141
	ECP	0.955	0.958	0.945	0.950	0.960	0.940	0.943	0.920	0.940	0.953	0.953	0.958	0.960	0.935	0.933	0.945	0.930	0.943	0.960	0.910	0.968	0.955
Method 3: Composite likelihood Simultaneous Estimation	EBias*	−0.257	0.144	−0.326	−0.602	0.072	−0.040	0.051	0.051	−0.020	−0.020	0.027	0.175	−0.370	−0.320	−0.534	−0.360	0.312	−0.591	−0.083	−0.635	−0.069	0.289
	ESE	0.098	0.100	0.111	0.117	0.013	0.015	0.016	0.022	0.027	0.026	0.029	0.034	0.032	0.038	0.039	0.036	0.104	0.070	0.018	0.142	0.017	0.140
	ASE	0.101	0.107	0.107	0.118	0.013	0.015	0.014	0.021	0.027	0.028	0.029	0.034	0.033	0.038	0.039	0.035	0.101	0.067	0.016	0.149	0.018	0.140
	ECP	0.953	0.965	0.940	0.953	0.958	0.953	0.920	0.948	0.953	0.965	0.950	0.960	0.950	0.945	0.940	0.940	0.938	0.958	0.950	0.953	0.943	0.953
Method 4: Composite likelihood Two−stage Estimation	EBias*	0.163	0.264	−0.853	−0.428	−0.136	0.041	0.273	0.014	0.005	−0.157	0.057	0.149	−0.396	−0.391	−0.597	−0.384	−0.650	−1.197	−0.150	−2.028	−0.137	0.000
	ESE	0.130	0.130	0.134	0.134	0.031	0.030	0.032	0.033	0.033	0.031	0.032	0.035	0.034	0.040	0.043	0.037	0.104	0.072	0.018	0.141	0.017	0.140
	ASE	0.131	0.134	0.134	0.133	0.031	0.031	0.031	0.031	0.033	0.032	0.033	0.035	0.036	0.041	0.042	0.036	0.101	0.069	0.016	0.149	0.018	0.141
	ECP	0.955	0.958	0.945	0.950	0.960	0.940	0.943	0.920	0.940	0.953	0.953	0.958	0.960	0.935	0.933	0.945	0.940	0.943	0.948	0.948	0.945	0.955

[1] EBias*=EBias×10^2

Table 10: Simulation results using the four estimation methods when block-connecting structure is misspecified: moderate dependence and $n = 1000$

Methods	Metrics	Marginal Parameters																Dependence Parameters					
		β_{01}	β_{02}	β_{03}	β_{04}	β_{11}	β_{12}	β_{13}	β_{14}	β_{21}	β_{22}	β_{23}	β_{24}	σ_1	σ_2	σ_3	σ_4	$\theta_{k1,k2}$	$\theta_{k1,k3}$	$\theta_{k1,k4}$	$\theta_{k2,k3}$	$\theta_{k2,k4}$	$\theta_{k3,k4}$
Method 1: Full likelihood Simultaneous Estimation	EBias*[1]	−1.607	0.441	−1.853	−1.010	−0.076	−0.019	0.002	−0.015	0.062	−0.018	0.017	0.109	−2.074	0.279	−1.793	−0.548	−3.042	−1.333	−0.318	−0.4.337	0.023	0.288
	ESE	0.067	0.073	0.073	0.086	0.008	0.009	0.008	0.016	0.018	0.018	0.020	0.024	0.023	0.027	0.027	0.026	0.073	0.047	0.012	0.107	0.013	0.105
	ASE	0.068	0.073	0.073	0.083	0.009	0.009	0.009	0.015	0.019	0.020	0.020	0.024	0.024	0.027	0.027	0.025	0.067	0.046	0.011	0.104	0.013	0.099
	ECP	0.947	0.940	0.947	0.940	0.952	0.962	0.967	0.932	0.955	0.960	0.945	0.965	0.862	0.950	0.887	0.935	0.925	0.945	0.942	0.910	0.955	0.947
Method 2: Full likelihood Two−stage Estimation	EBias*	−0.788	0.291	−0.835	0.023	0.082	−0.065	0.156	−0.145	0.133	−0.081	0.120	0.137	−0.079	0.117	−0.201	−0.006	−2.347	0.316	−0.069	−4.106	−0.017	0.344
	ESE	0.088	0.092	0.090	0.097	0.022	0.021	0.021	0.022	0.023	0.022	0.023	0.025	0.026	0.029	0.031	0.027	0.075	0.050	0.012	0.106	0.013	0.105
	ASE	0.093	0.095	0.095	0.095	0.022	0.022	0.022	0.022	0.023	0.023	0.023	0.025	0.026	0.029	0.030	0.026	0.071	0.049	0.012	0.105	0.013	0.099
	ECP	0.958	0.968	0.958	0.953	0.968	0.958	0.965	0.958	0.948	0.950	0.943	0.943	0.958	0.950	0.965	0.930	0.950	0.953	0.960	0.935	0.955	0.943
Method 3: Composite likelihood Simultaneous Estimation	EBias*	−0.059	0.235	−0.164	−0.184	−0.074	−0.066	−0.022	−0.010	0.014	−0.060	0.009	0.087	−0.055	0.069	−0.114	−0.006	0.220	−0.120	0.013	0.058	−0.034	0.467
	ESE	0.069	0.075	0.076	0.087	0.009	0.010	0.010	0.016	0.018	0.019	0.020	0.025	0.024	0.027	0.029	0.027	0.077	0.050	0.012	0.108	0.013	0.105
	ASE	0.071	0.076	0.076	0.084	0.009	0.011	0.010	0.015	0.019	0.020	0.021	0.024	0.023	0.027	0.027	0.025	0.072	0.048	0.011	0.105	0.012	0.099
	ECP	0.963	0.958	0.940	0.948	0.960	0.965	0.933	0.938	0.963	0.960	0.958	0.950	0.943	0.958	0.948	0.923	0.950	0.953	0.960	0.935	0.955	0.948
Method 4: Composite likelihood Two−stage Estimation	EBias*	−0.788	0.291	−0.835	0.023	0.082	−0.065	0.156	−0.145	0.133	−0.081	0.120	0.137	−0.079	0.117	−0.201	−0.006	0.217	−0.415	−0.015	−0.591	−0.057	0.303
	ESE	0.088	0.092	0.090	0.097	0.022	0.021	0.021	0.022	0.023	0.022	0.023	0.025	0.026	0.029	0.031	0.027	0.076	0.051	0.012	0.108	0.013	0.106
	ASE	0.093	0.095	0.095	0.095	0.022	0.022	0.022	0.022	0.023	0.023	0.023	0.025	0.026	0.029	0.030	0.026	0.072	0.049	0.012	0.105	0.013	0.099
	ECP	0.958	0.968	0.958	0.953	0.968	0.958	0.965	0.958	0.948	0.950	0.943	0.943	0.958	0.950	0.965	0.930	0.935	0.950	0.958	0.940	0.955	0.940

[1] EBias*=EBias$\times 10^2$

2.3 Copula Selection

In this subsection, we aim to explore the capacity of the proposed copula selection procedure in Section 4 in the main text and examine how frequently we can select the correct copula forms for C-Vine structure within the time blocks.

2.3.1 Simulation Setting

We simulate data from the same setting as that in Section 2.1.1. We evaluate the performance for copula selection under both the strong and moderate dependence settings, and $n = 500$ or 1000. The simulation is repeated 500 times.

2.3.2 Copula Set and Evaluation Metrics

For simplicity, we construct a set of candidate copula functions including the commonly-used copulas in the Archimedean family (Clayton, Gumbel, Frank and Joe copula), Gaussian copula and t copula. The *mis-selected rate* of a copula function is used to evaluate the copula selection performance, which is computed as the number of times for which the copula function is incorrectly selected divided by the number of simulations.

2.3.3 Simulation Results

We report the mis-selected rates for the six (conditional) bivariate copulas in Table 11, where the correct forms are specified in Table 1.

Table 11: Mis-selected rates for copula functions within each block

Degree of Dependence	Sample Size	$\varepsilon_{ik1},\varepsilon_{ik2}$	$\varepsilon_{ik1},\varepsilon_{ik3}$	$\varepsilon_{ik1},\varepsilon_{ik4}$	$\varepsilon_{ik2},\varepsilon_{ik3}\|\varepsilon_{ik1}$	$\varepsilon_{ik2},\varepsilon_{ik4}\|\varepsilon_{ik1}$	$\varepsilon_{ik3},\varepsilon_{ik4}\|\varepsilon_{ik1},\varepsilon_{ik2}$
Strong Dependence	500	0.264	0	0.008	0	0	0
	1000	0.192	0	0.006	0	0	0
Moderate Dependence	500	0.182	0	0	0.002	0.002	0.024
	1000	0.074	0	0	0	0.002	0.006

The mis-selected rates for all the (conditional) bivariate copulas are close to 0, except for the bivariate copula between ε_{ik1} and ε_{ik2}, for which the true form is a Clayton copula. The mis-selected rates of all (conditional) bivariate copulas drop, as the sample size increases or the dependence becomes weaker. The mis-selected rate for the Clayton copula drops from 26.4% to 19.2% by increasing the sample size from 500 to 1000 in the scenario of a strong dependence and drops even more dramatically from 18.2% to 7.4% for the scenario of a moderate dependence. Generally speaking, we are confident with the proposed copula selection method with fairly low mis-selected rates.

2.4 Prediction

2.4.1 Subject Extrapolation and Time Extrapolation

Two kinds of prediction are of our interest: *subject extrapolation* and *time extrapolation*. We explain the meaning of two kinds of predictions, how we create the training and test set and how we conduct prediction in both cases.

- *Subject Extrapolation*: We are interested in predicting the value of the response for a new subject at a past or current time point. We partition the data by subjects, use 90% of the subjects as the training set, denoted by $\{(y_i^{\mathrm{T}}, x_i^{\mathrm{T}})^{\mathrm{T}} : i = 1, \ldots, 450\}$, and reserve 10% of the subjects as the test set, denoted by $\{(y_i^{\mathrm{T}}, x_i^{\mathrm{T}})^{\mathrm{T}} : i = 451, \ldots, 500\}$. The training set is used to fit a model, which is utilized to predict y_{ikl} for a subject from the test set using its covariate information and responses from the first $l-1$ time points in the kth time block.

- *Time Extrapolation*: We are interested in predicting the response value for a subject at a future time point. We partition the data by time points, use the time points from the first four blocks as the training set, denoted by $\{(y_{ikl}^{\mathrm{T}}, x_{ikl}^{\mathrm{T}})^{\mathrm{T}} : i = 1, \ldots, 500; k = 1, 2, 3, 4; l = 1, 2, 3, 4\}$, and reserve the time points in the fifth block as the test set, denoted by $\{(y_{ikl}^{\mathrm{T}}, x_{ikl}^{\mathrm{T}})^{\mathrm{T}} : i = 1, \ldots, 500; k = 5; l = 1, 2, 3, 4\}$. The training set is used to fit a model, which is utilized to predict y_{ikl} for a time point in time block $k = 5$, based on the covariate information and the first $l-1$ time points in the 5th time block.

2.4.2 Evaluation Metrics

Let $y_{ikl}^{(r)}$ denote the response value of the ith subject at the lth time point in the kth time block from the rth independent dataset and let $\hat{y}_{ikl}^{(r)}$ be the corresponding predicted value. We consider the following two evaluation metrics:

- *Mean Absolute Error (MAE)*: the mean of the absolute difference between the predicted value and the true value over all time points in the test set across 200 simulations. To evaluate subject extrapolation, the MAE is computed as

$$\frac{1}{200 \cdot 50 \cdot 5 \cdot 4} \sum_{r=1}^{200} \sum_{i=451}^{500} \sum_{k=1}^{5} \sum_{l=1}^{4} |\hat{y}_{ikl}^{(r)} - y_{ikl}^{(r)}|;$$

to evaluate time extrapolation, it is computed by

$$\frac{1}{200 \cdot 500 \cdot 4} \sum_{r=1}^{200} \sum_{i=1}^{500} \sum_{l=1}^{4} |\hat{y}_{i5l}^{(r)} - y_{i5l}^{(r)}|.$$

The model that provides a smaller MAE is preferred.

- *Percentage Outperformance*: Percentage outperformance of Model 1 versus Model 2 is calculated as the number of times that Model 1 provides a smaller MAE than Model 2, divided by the number of time points in the test set and then averaged over 200 simulations. If percentage outperformance is over 50%, Model 1 provides better prediction accuracy than Model 2. Let $\hat{y}_{ikl}^{(1r)}$ and $\hat{y}_{ikl}^{(2r)}$ be the predicted values from Models 1 and 2, respectively. To evaluate subject extrapolation, percentage outperformance is computed as

$$\frac{1}{200 \cdot 50 \cdot 5 \cdot 4} \sum_{r=1}^{200} \sum_{i=451}^{500} \sum_{k=1}^{5} \sum_{l=1}^{4} I(|\hat{y}_{ikl}^{(1r)} - y_{ikl}^{(r)}| \leq |\hat{y}_{ikl}^{(2r)} - y_{ikl}^{(r)}|);$$

to evaluate time extrapolation, it is computed as

$$\frac{1}{200 \cdot 500 \cdot 4} \sum_{r=1}^{200} \sum_{i=1}^{500} \sum_{l=1}^{4} I(|\hat{y}_{i5l}^{(1r)} - y_{i5l}^{(r)}| \leq |\hat{y}_{i5l}^{(2r)} - y_{i5l}^{(r)}|).$$

Percentage outperformance is more robust than the MAE, which may be sensitive to extreme prediction values.

2.4.3 Additional Simulation Results

Table 12: MAEs of different models for subject extrapolation under the proposed scenarios (prediction standard error in the brackets)

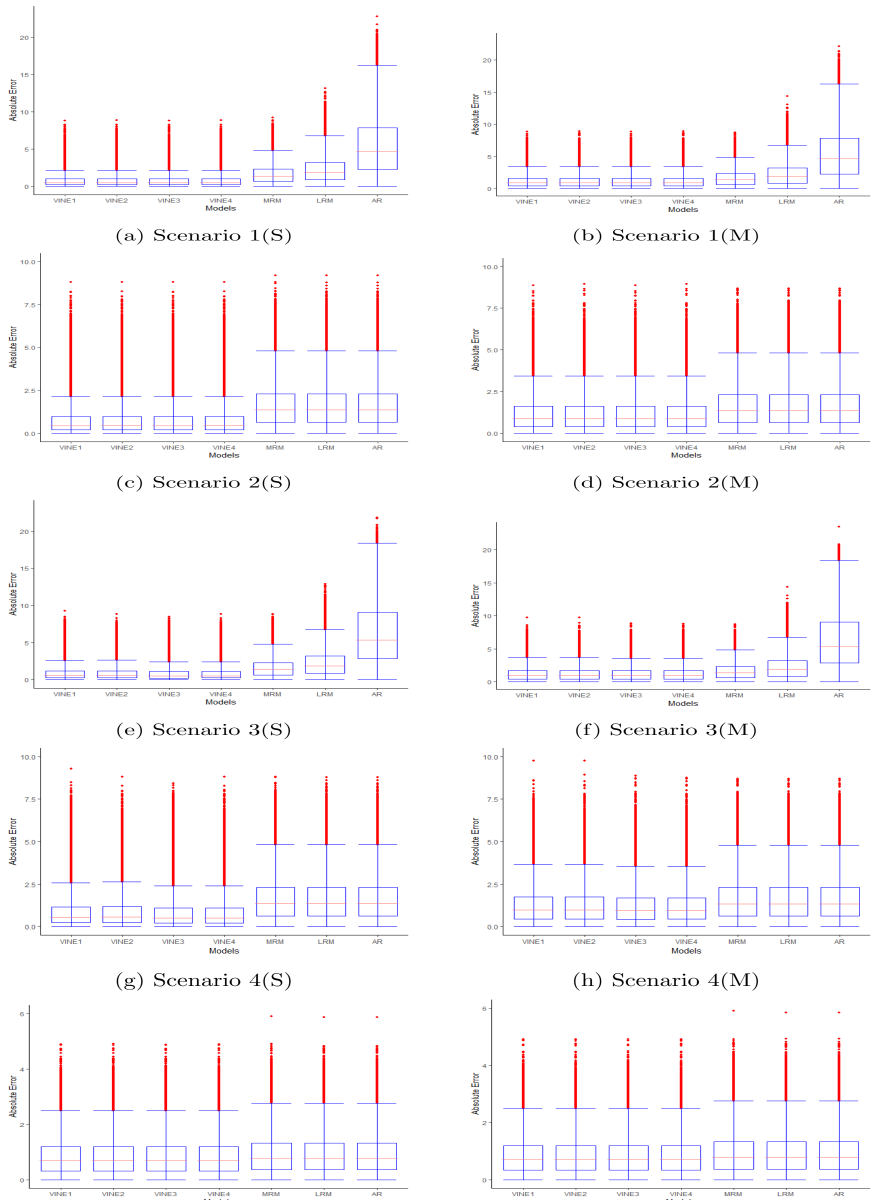
	VINE1	VINE2	VINE3	VINE4	MRM	LRM	AR
Scenario 1(S)	0.761 (1.158)	0.762 (1.159)	0.767 (1.159)	0.767 (1.160)	1.598 (1.977)	2.249 (2.841)	5.331 (6.550)
Scenario 1(M)	1.145 (1.486)	1.145 (1.487)	1.147 (1.487)	1.147 (1.488)	1.604 (1.980)	2.252 (2.843)	5.332 (6.549)
Scenario 2(S)	0.761 (1.158)	0.762 (1.159)	0.767 (1.159)	0.767 (1.160)	1.598 (1.977)	1.598 (1.977)	1.599 (1.975)
Scenario 2(M)	1.145 (1.486)	1.145 (1.487)	1.147 (1.487)	1.147 (1.488)	1.604 (1.980)	1.604 (1.980)	1.606 (1.978)
Scenario 3(S)	0.871 (1.216)	0.889 (1.270)	0.831 (1.217)	0.834 (1.270)	1.599 (1.986)	2.248 (2.856)	6.098 (7.336)
Scenario 3(M)	1.232 (1.555)	1.235 (1.572)	1.199 (1.555)	1.201 (1.572)	1.605 (1.986)	2.253 (2.858)	6.100 (7.337)
Scenario 4(S)	0.871 (1.216)	0.889 (1.270)	0.831 (1.217)	0.835 (1.271)	1.600 (1.986)	1.600 (1.985)	1.601 (1.985)
Scenario 4(M)	1.232 (1.555)	1.235 (1.572)	1.199 (1.555)	1.201 (1.573)	1.606 (1.986)	1.606 (1.986)	1.607 (1.985)
Scenario 5	0.830 (1.038)	0.830 (1.039)	0.830 (1.038)	0.830 (1.039)	0.923 (1.153)	0.923 (1.154)	0.922 (1.152)
Scenario 6	0.830 (1.038)	0.830 (1.039)	0.830 (1.038)	0.830 (1.039)	0.923 (1.153)	0.922 (1.154)	0.922 (1.152)

S: strong dependence setting; M: moderate dependence setting

Table 13: Simulation results for subject extrapolation and time extrapolation in terms of percentage outperformance VINE4 versus the other models

	Subject Extrapolation			Time Extrapolation		
	MRM	LRM	AR	MRM	LRM	AR
Scenario 1(S)	0.618	0.814	0.909	0.868	0.847	0.938
Scenario 1(M)	0.558	0.725	0.874	0.761	0.780	0.937
Scenario 2(S)	0.618	0.744	0.745	0.868	0.776	0.836
Scenario 2(M)	0.558	0.637	0.665	0.761	0.636	0.745
Scenario 3(S)	0.611	0.807	0.945	0.868	0.840	0.919
Scenario 3(M)	0.552	0.719	0.904	0.748	0.764	0.938
Scenario 4(S)	0.611	0.725	0.725	0.868	0.707	0.710
Scenario 4(M)	0.534	0.665	0.665	0.748	0.618	0.620
Scenario 5	0.545	0.547	0.535	0.692	0.693	0.533
Scenario 6	0.545	0.536	0.534	0.692	0.550	0.534

S: strong dependence setting; M: moderate dependence setting



(a) Scenario 1(S)

(b) Scenario 1(M)

(c) Scenario 2(S)

(d) Scenario 2(M)

(e) Scenario 3(S)

(f) Scenario 3(M)

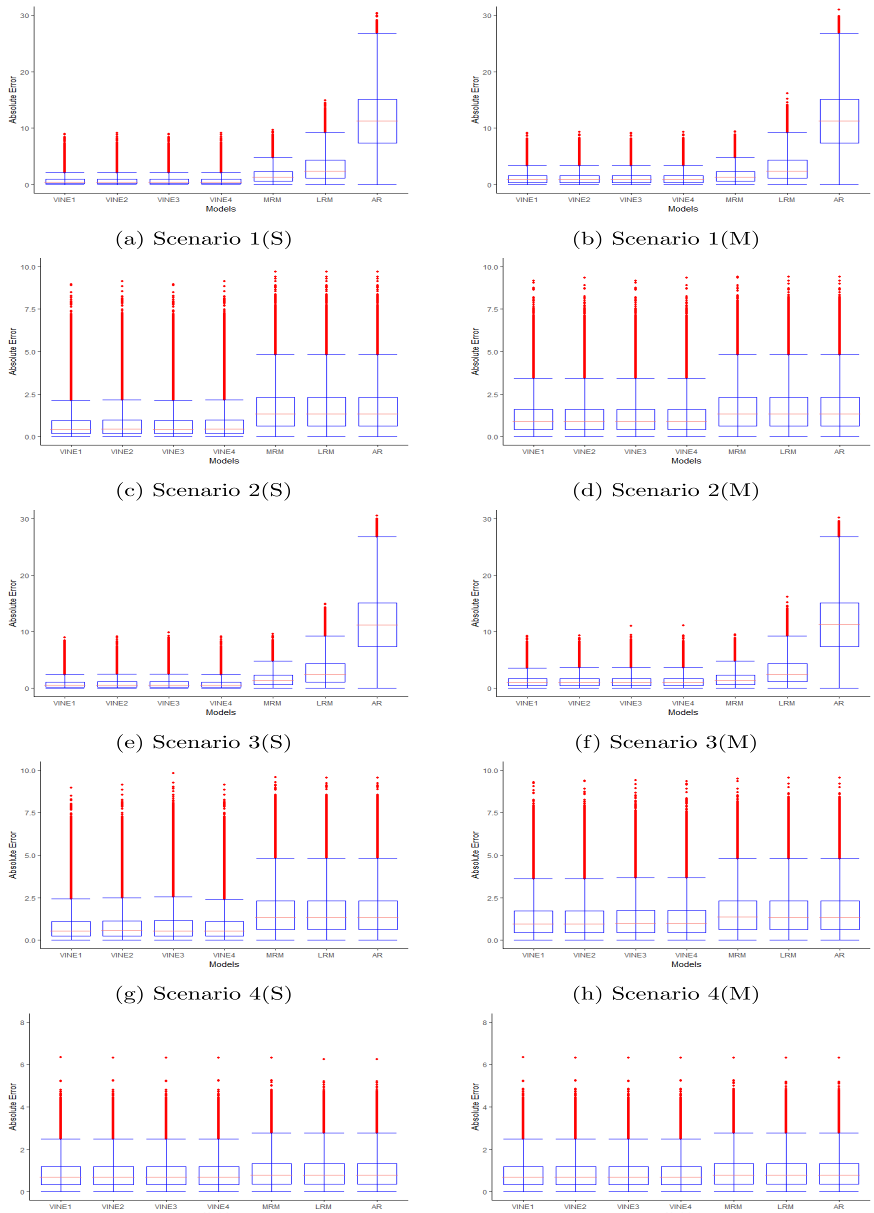
(g) Scenario 4(S)

(h) Scenario 4(M)

(i) Scenario 5

(j) Scenario 6

Figure 3: Boxplots of MAEs of different models for subject extrapolation under the six scenarios



(a) Scenario 1(S) (b) Scenario 1(M)

(c) Scenario 2(S) (d) Scenario 2(M)

(e) Scenario 3(S) (f) Scenario 3(M)

(g) Scenario 4(S) (h) Scenario 4(M)

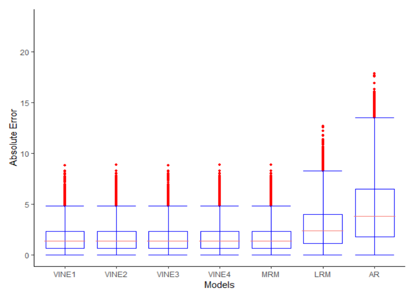
(i) Scenario 5 (j) Scenario 6

Figure 4: Boxplots of MAEs of different models for time extrapolation under the six scenarios

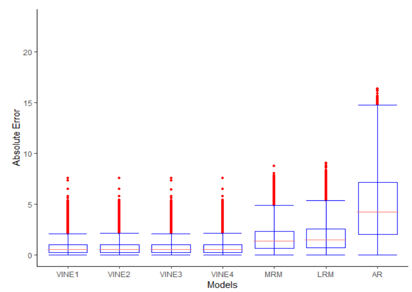
2.4.4 MAEs by Time Points for Subject Extrapolation

MAE by time points for subject extrapolation for the lth time point is computed by

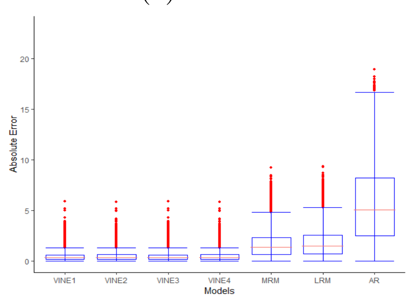
$$\frac{1}{200 \cdot 50 \cdot 5} \sum_{r=1}^{200} \sum_{i=451}^{500} \sum_{k=1}^{5} |\hat{y}_{ikl}^{(r)} - y_{ikl}^{(r)}|.$$



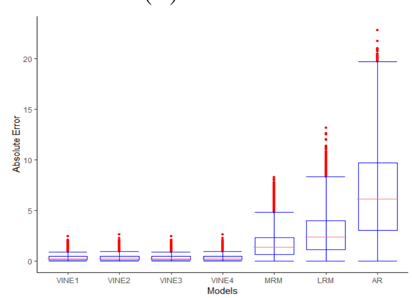
(a) $l = 1$



(b) $l = 2$

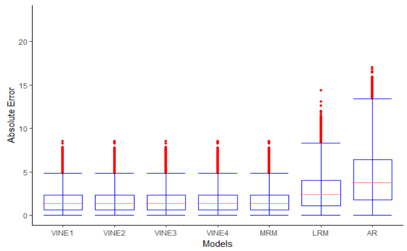


(c) $l = 3$

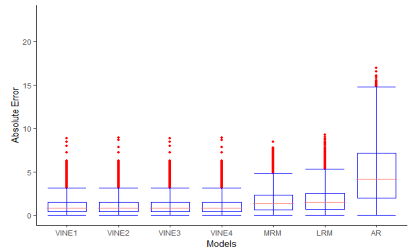


(d) $l = 4$

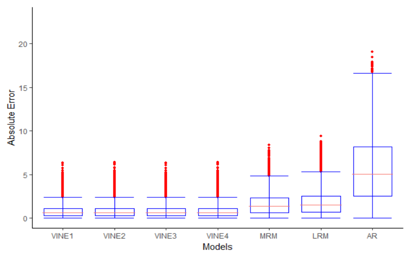
Figure 5: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 1(S)



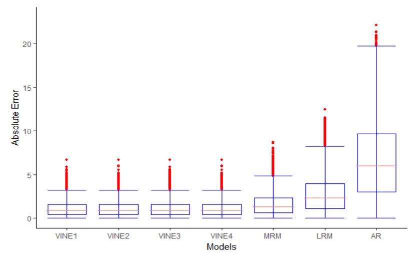
(a) $l = 1$



(b) $l = 2$

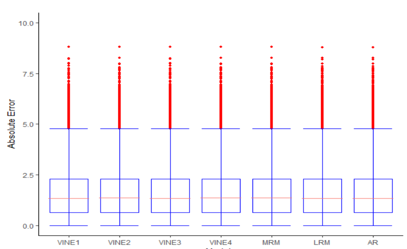


(c) $l = 3$

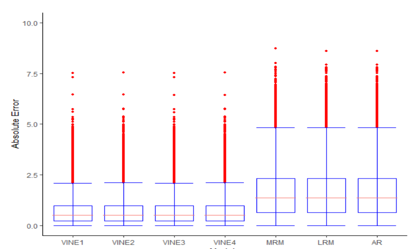


(d) $l = 4$

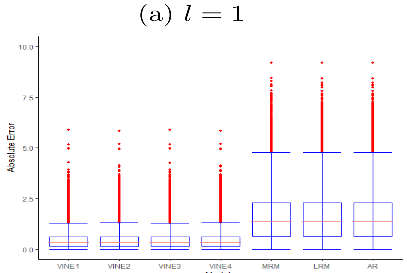
Figure 6: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 1(M))



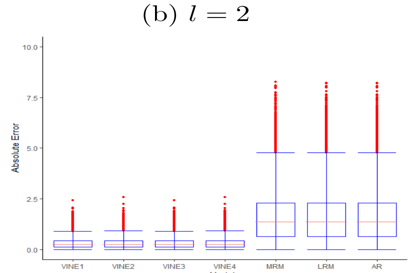
(a) $l = 1$



(b) $l = 2$

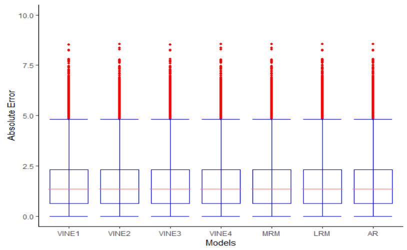


(c) $l = 3$

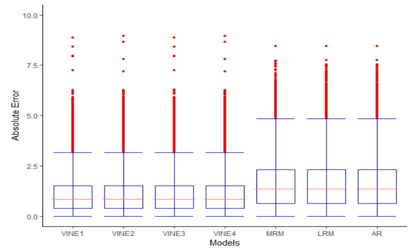


(d) $l = 4$

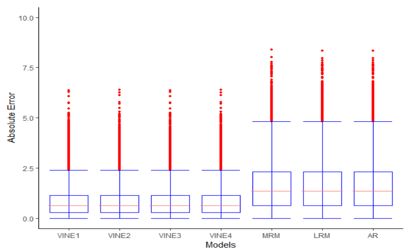
Figure 7: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 2(S)



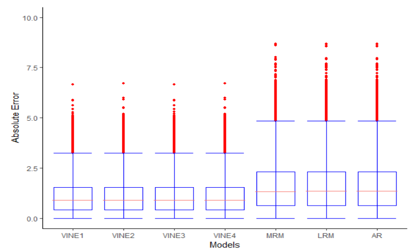
(a) $l = 1$



(b) $l = 2$

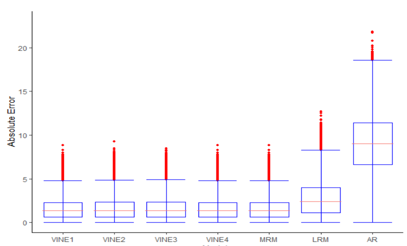


(c) $l = 3$

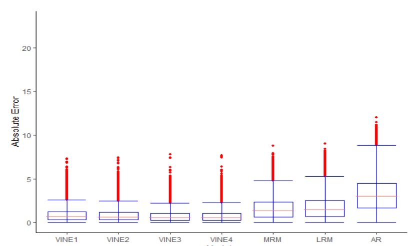


(d) $l = 4$

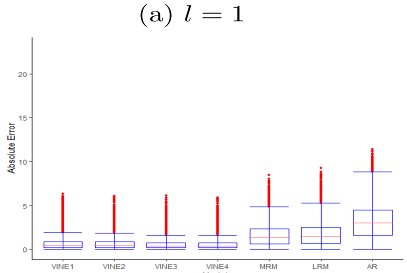
Figure 8: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 2(M)



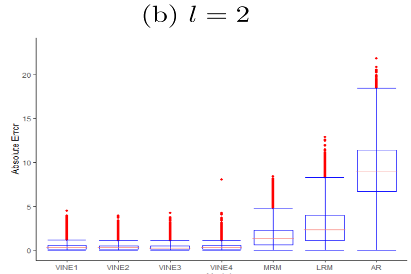
(a) $l = 1$



(b) $l = 2$

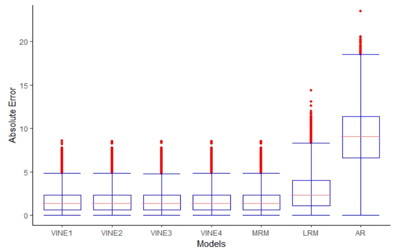


(c) $l = 3$

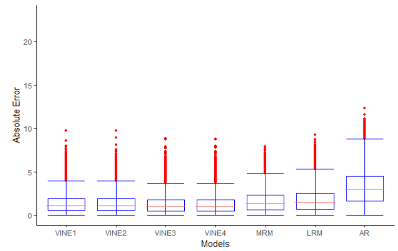


(d) $l = 4$

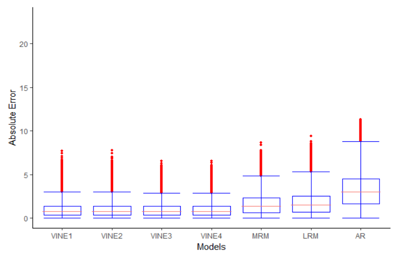
Figure 9: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 3(S)



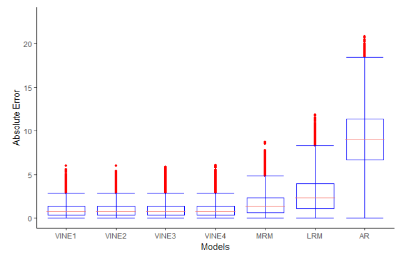
(a) $l = 1$



(b) $l = 2$

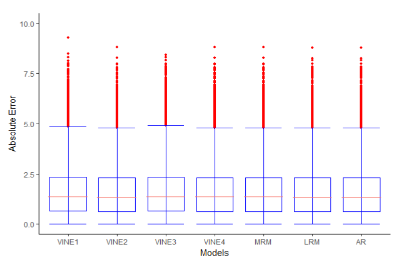


(c) $l = 3$

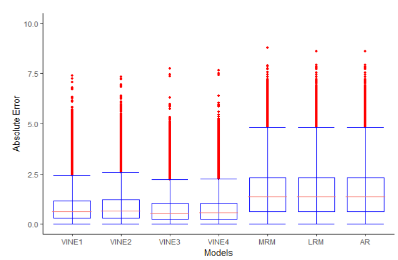


(d) $l = 4$

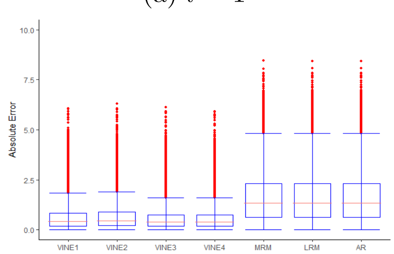
Figure 10: Subject extrapolation:boxplots for prediction results in 4 time points in Scenario 3(M)



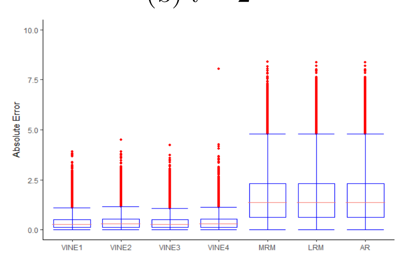
(a) $l = 1$



(b) $l = 2$

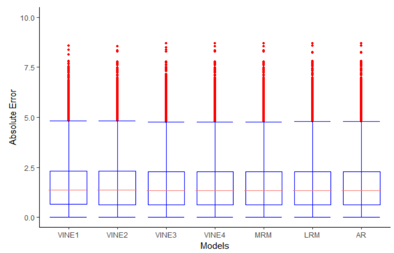


(c) $l = 3$

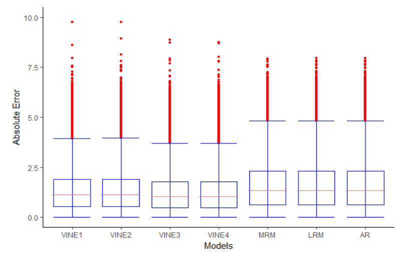


(d) $l = 4$

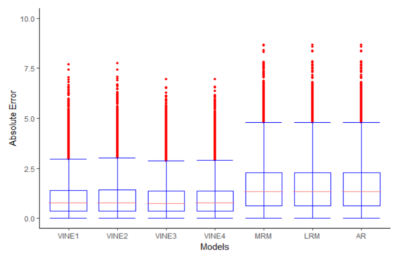
Figure 11: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 4(S)



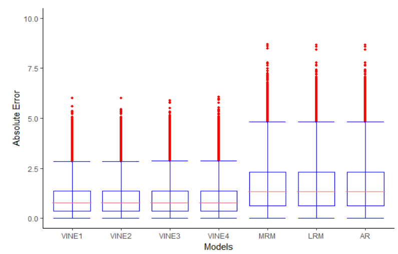
(a) $l = 1$



(b) $l = 2$

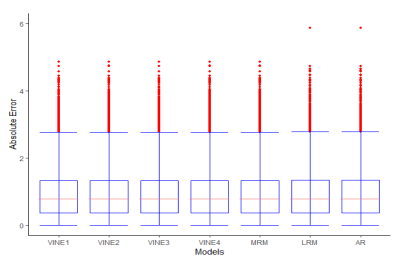


(c) $l = 3$

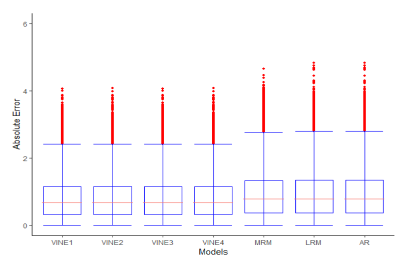


(d) $l = 4$

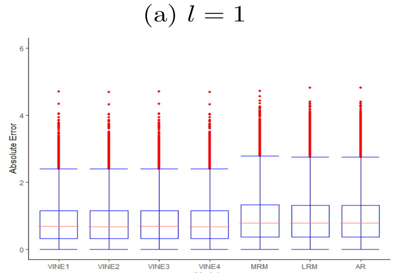
Figure 12: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 4(M)



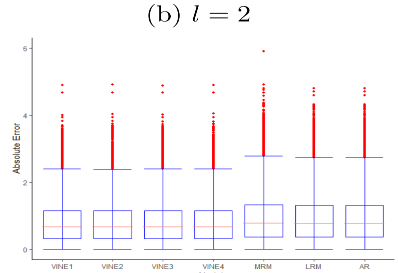
(a) $l = 1$



(b) $l = 2$

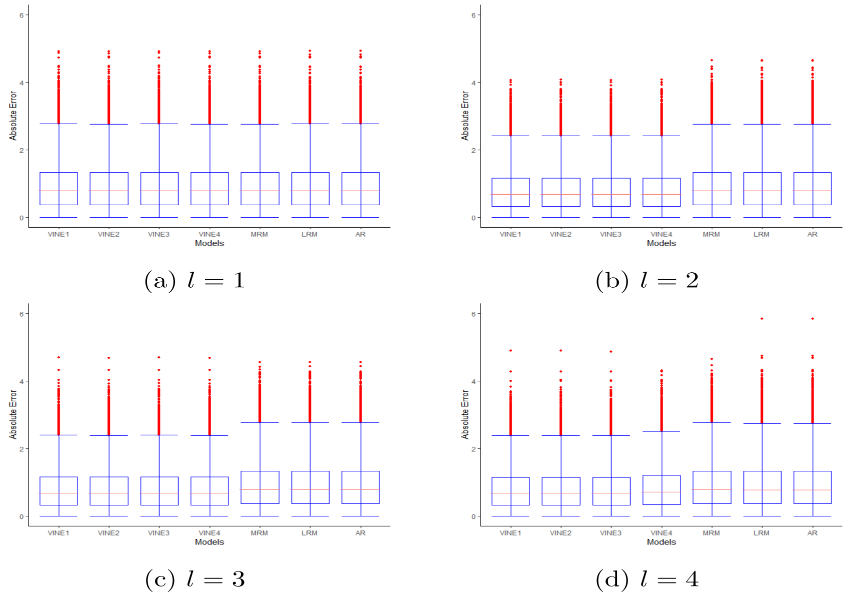


(c) $l = 3$



(d) $l = 4$

Figure 13: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 5



(a) $l = 1$ (b) $l = 2$

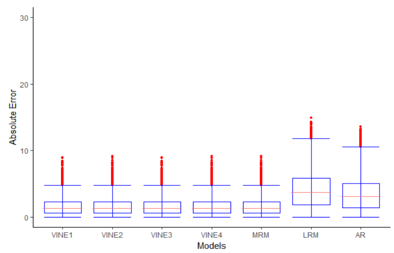
(c) $l = 3$ (d) $l = 4$

Figure 14: Subject extrapolation: boxplots for prediction results in 4 time points in Scenario 6

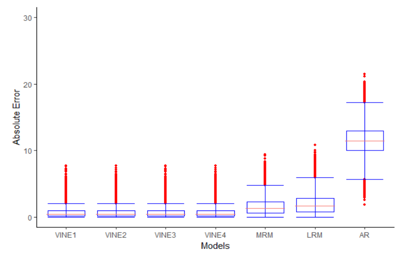
2.4.5 MAEs by Time Points for Time Extrapolation

MAE by time points for time extrapolation for the lth time point is computed by

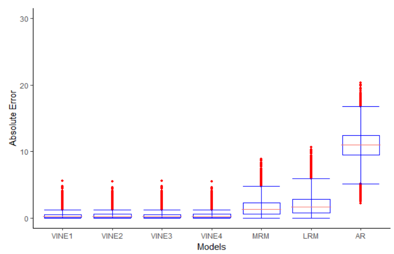
$$\frac{1}{200 \cdot 500} \sum_{r=1}^{200} \sum_{i=1}^{500} |\hat{y}_{i5l}^{(r)} - y_{i5l}^{(r)}|.$$



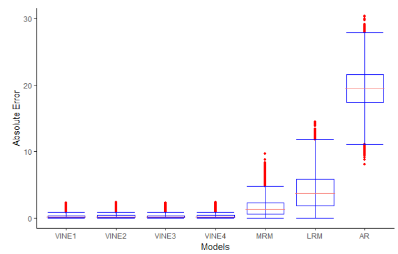
(a) $l = 1$



(b) $l = 2$

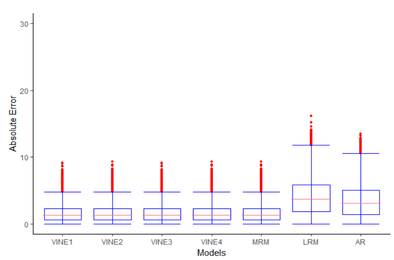


(c) $l = 3$

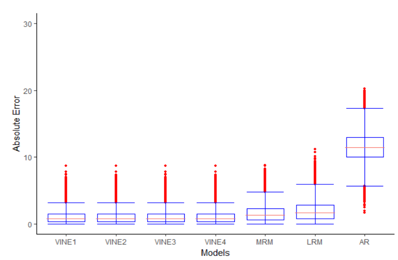


(d) $l = 4$

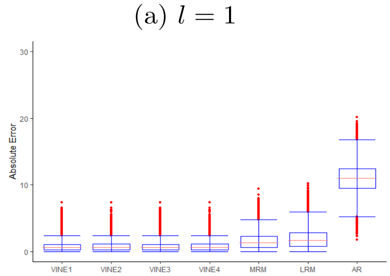
Figure 15: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 1(S)



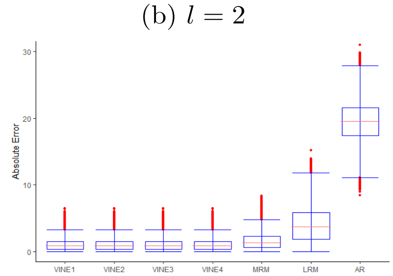
(a) $l = 1$



(b) $l = 2$

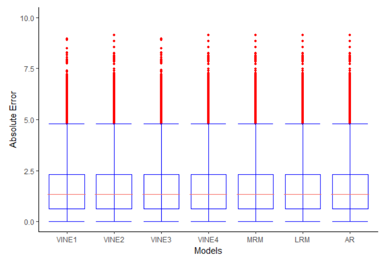


(c) $l = 3$

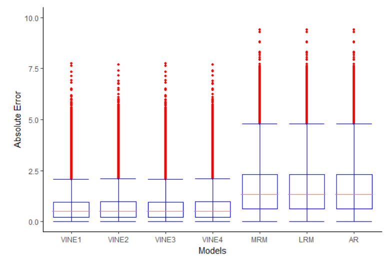


(d) $l = 4$

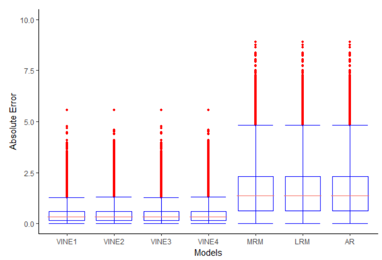
Figure 16: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 1(M)



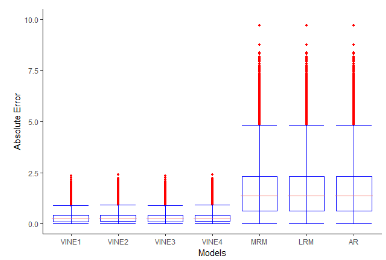
(a) $l = 1$



(b) $l = 2$

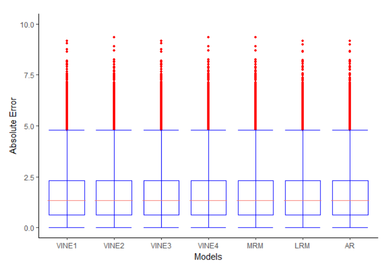


(c) $l = 3$

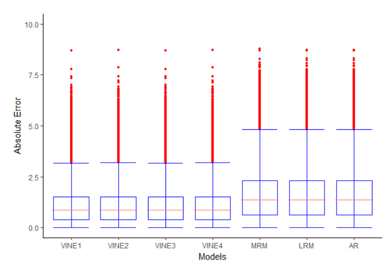


(d) $l = 4$

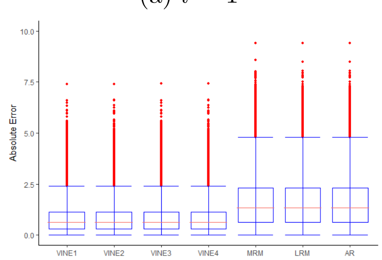
Figure 17: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 2(S)



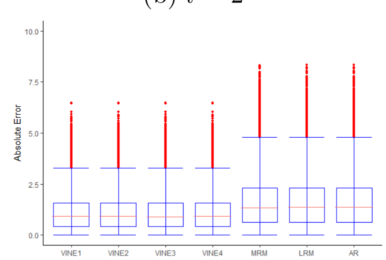
(a) $l = 1$



(b) $l = 2$

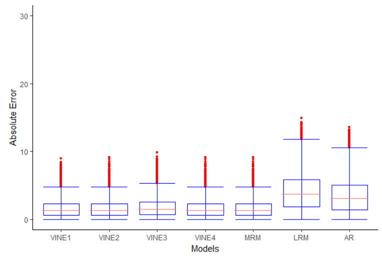


(c) $l = 3$

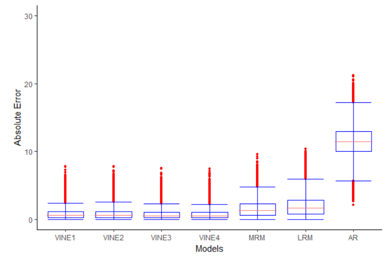


(d) $l = 4$

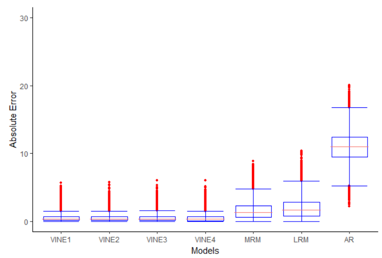
Figure 18: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 2(M)



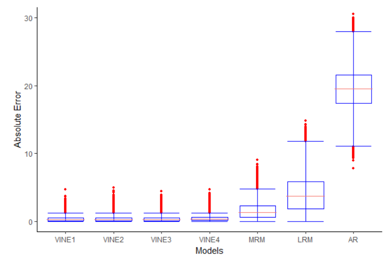
(a) $l = 1$



(b) $l = 2$

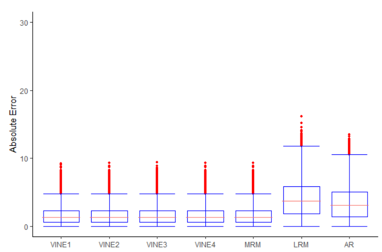


(c) $l = 3$

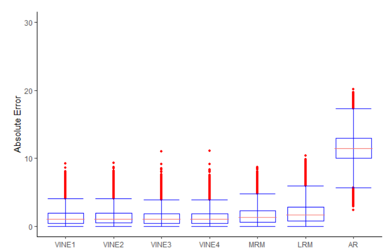


(d) $l = 4$

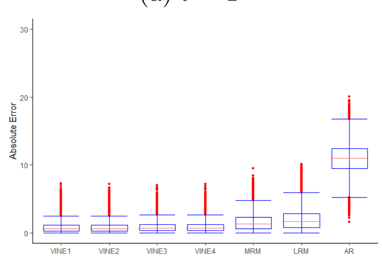
Figure 19: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 3(S)



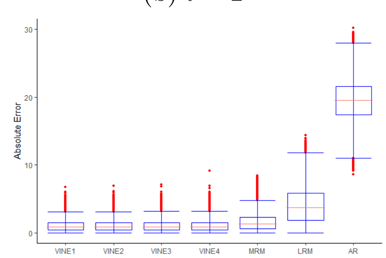
(a) $l = 1$



(b) $l = 2$

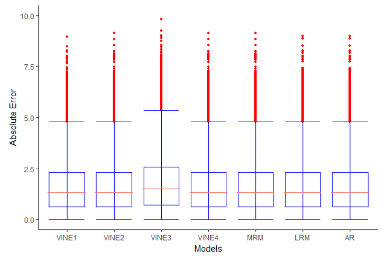


(c) $l = 3$

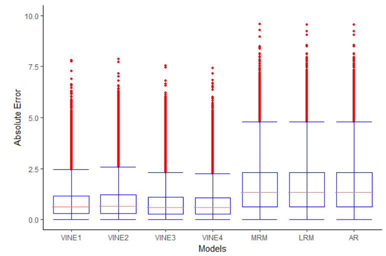


(d) $l = 4$

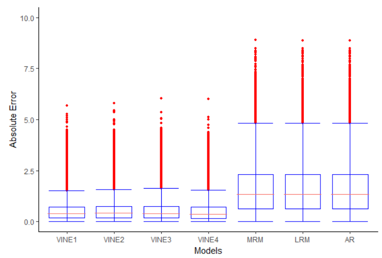
Figure 20: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 3(M)



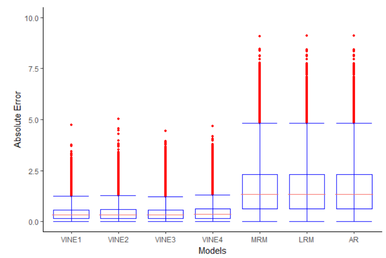
(a) $l = 1$



(b) $l = 2$

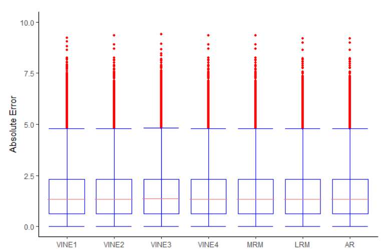


(c) $l = 3$

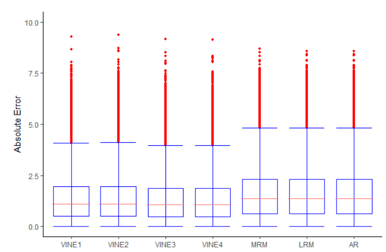


(d) $l = 4$

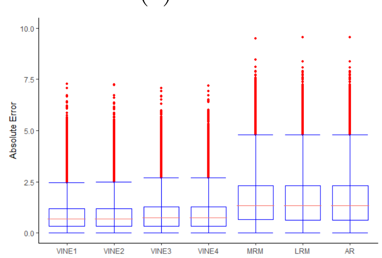
Figure 21: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 4(S)



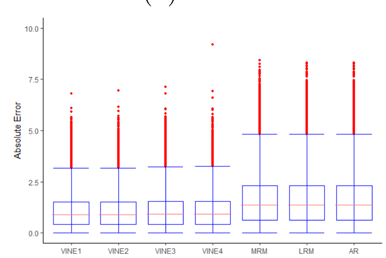
(a) $l = 1$



(b) $l = 2$

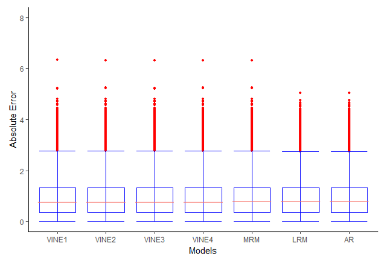


(c) $l = 3$

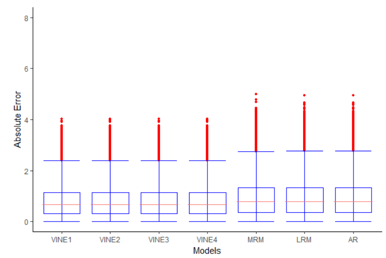


(d) $l = 4$

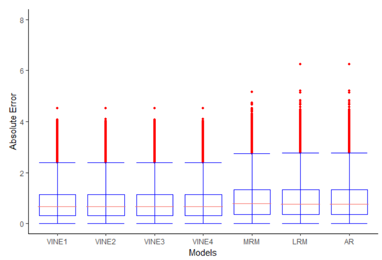
Figure 22: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 4(M)



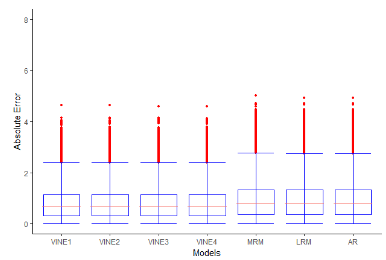
(a) $l = 1$



(b) $l = 2$

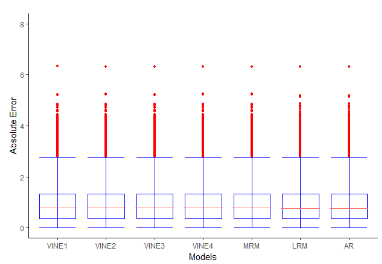


(c) $l = 3$

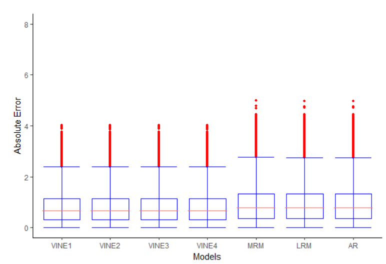


(d) $l = 4$

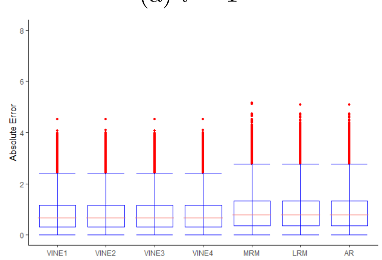
Figure 23: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 5



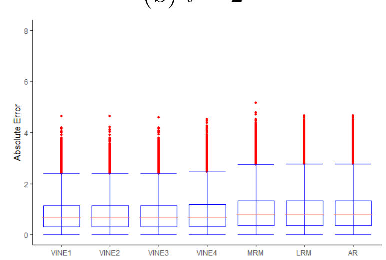
(a) $l = 1$



(b) $l = 2$



(c) $l = 3$



(d) $l = 4$

Figure 24: Time extrapolation: boxplots for prediction results in 4 time points in Scenario 6

3 Data Analysis

3.1 Dataset Description

Table 14: Location Information of 47 Observation Stations

ID	Name	Latitude	Longitude	Elevation	Group	ID	Name	Latitude	Longitude	Elevation	Group
1	LANSDOWNE HOUSE	52.23	−87.88	255	Training	25	BROCKVILLE	44.60	−75.67	96	Training
2	PICKLE LAKE	51.45	−90.22	386	Training	26	CORNWALL	45.02	−74.75	64	Training
3	RED LAKE	51.07	−93.80	386	Training	27	KINGSTON	44.22	−76.60	93	Training
4	FORT FRANCES	48.65	−93.43	342	Training	28	MORRISBURG	44.92	−75.18	82	Training
5	MINE CENTRE	48.80	−92.60	361	Training	29	OTTAWA	45.38	−75.72	79	Training
6	DRYDEN	49.78	−92.83	413	Training	30	OWEN SOUND	44.58	−80.93	179	Training
7	KENORA	49.78	−94.37	406	Training	31	RIDGETOWN	42.45	−81.88	206	Training
8	CAMERON FALLS	49.15	−88.35	233	Training	32	VINELAND	43.17	−79.42	79	Training
9	GERALDTON	49.78	−86.93	349	Training	33	WELLAND	43.00	−79.27	175	Training
10	THUNDER BAY	48.37	−89.33	199	Training	34	WINDSOR	42.27	−82.97	190	Training
11	HORNEPAYNE	49.20	−84.77	335	Training	35	LONDON	43.03	−81.15	278	Training
12	SAULT STE MARIE	46.48	−84.52	192	Training	36	WOODSTOCK	43.13	−80.77	282	Training
13	WAWA	47.97	−84.78	287	Training	37	BELLEVILLE	44.15	−77.40	76	Training
14	CHAPLEAU	47.82	−83.35	447	Training	38	HAMILTON	43.17	−79.93	238	Training
15	SUDBURY	46.62	−80.80	348	Training	39	ORANGEVILLE	43.92	−80.08	412	Training
16	EARLTON	47.70	−79.85	243	Training	40	TORONTO	43.67	−79.40	113	Training
17	IROQUOIS FALLS	48.75	−80.67	259	Training	41	HALIBURTON	45.03	−78.53	330	Training
18	KAPUSKASING	49.42	−82.47	227	Training	42	PETERBOROUGH	44.23	−78.37	191	Training
19	MOOSONEE	51.27	−80.65	10	Training						
20	SMOKY FALLS	50.07	−82.17	183	Training	43	BIG TROUT LAKE	53.83	−89.87	224	Validation
21	TIMMINS	48.57	−81.38	295	Training	44	SIOUX LOOKOUT	50.12	−91.90	383	Validation
22	MADAWASKA	45.50	−77.98	316	Training	45	BEATRICE	45.13	−79.40	297	Validation
23	NORTH BAY	46.37	−79.42	370	Training	46	HARROW	42.03	−82.90	182	Validation
24	GORE BAY	45.88	−82.57	194	Training	47	ATITOKAN	48.8	−91.58	442	Validation

3.2 Model Fitting Results

Table 15: Estimates of first parameters of the copula functions in the C-Vine structure obtained by the two−stage estimation procedure (standard error in the bracket)

Tree \ Month	2	3	4	5	6	7	8	9	10	11	12
1	3.035(25.076)	1.647(6.797)	0.311(0.063)	0.120(1.253)	–	0.253(0.922)	0.375(0.390)	−0.735(1.636)	−0.248(0.065)	–	–
2		1.804(5.513)	0.544(0.717)	–	0.101(0.136)	0.055(0.449)	0.201(0.049)	1.796(0.275)	1.152(0.267)	1.519(0.268)	1.405(0.201)
3			0.156(5.391)	0.419(1.622)	–	0.100(3.417)	−1.728(2.472)	–	1.059(0.586)	1.392(0.432)	0.023(0.908)
4				1.934(10.769)	0.230(0.636)	0.268(0.254)	–	–	−0.188(0.563)	1.678(1.094)	0.209(0.065)
5					–	0.191(2.059)	1.461(1.274)	–	0.106(0.421)	–	1.080(0.045)
6						0.548(0.077)	1.352(0.254)	1.962(0.795)	0.884(0.583)	−1.113(0.080)	–
7							1.108(0.168)	0.316(0.484)	−1.133(0.222)	1.136(0.335)	−1.037(0.184)
8								–	1.334(0.621)	–	−0.224(0.158)
9									1.216(0.131)	1.581(0.518)	–
10										1.611(2.412)	−0.138(0.171)
11											1.688(0.570)

Table 16: Estimates of second parameters of the copula functions in the C-Vine structure obtained by the two−stage estimation procedure (standard error in the bracket)

Tree \ Month	2	3	4	5	6	7	8	9	10	11	12
1	0.078(0.952)	0.263(3.607)	–	–	–	–	–	–	–	–	–
2		0.860(1.060)	–	–	–	–	–	0.115(0.069)	–	0.256(0.301)	0.084(0.060)
3			3.674(28.822)	–	–	–	0.075(0.081)	–	0.164(0.888)	0.271(2.377)	5.077(9.447)
4				0.248(2.377)	12.932(56.992)	–	–	–	–	0.819(0.703)	–
5					–	–	0.168(0.578)	–	–	–	–
6						–	0.525(0.329)	–	–	–	–
7							0.310(1.036)	–	–	–	–
8								–	0.348(0.702)	–	10.113(24.569)
9									–	0.152(0.225)	–
10										0.054(0.076)	–
11											0.958(0.080)

3.3 Prediction Results

The prediction results for subject extrapolation are summarized in Table 17. Both MAEs and percentage outperformances suggest that the proposed R-Vine model estimated using the composite likelihood method can provide a lot more precise prediction than the other three conventional models.

Table 17: Prediction results for subject extrapolation (prediction standard error in the brackets)

	MAE				Percentage Outperformance		
Name	VINE4	MRM	LRM	AR	VINE4 vs MRM	VINE4 vs LRM	VINE4 vs AR
BIG TROUT LAKE	**1.916 (1.979)**	2.085 (2.185)	4.127 (4.056)	3.144(2.673)	0.604	0.760	0.708
SIOUX LOOKOUT	**1.908 (2.013)**	2.243 (2.185)	3.840 (4.056)	2.901 (2.857)	0.642	0.717	0.725
BEATRICE	**1.441 (1.949)**	1.555 (2.185)	2.641 (4.056)	1.753 (2.557)	0.625	0.708	0.646
HARROW	**1.568 (1.939)**	1.658 (2.185)	2.798 (4.056)	1.683 (2.629)	0.563	0.667	0.542
ATITOKAN	**1.685 (1.975)**	1.923 (2.185)	3.582 (4.056)	2.304 (2.729)	0.646	0.792	0.646
Average	**1.704 (1.971)**	1.893 (2.185)	3.398 (4.056)	2.357 (2.689)	0.616	0.729	0.653

The prediction results of the 37 stations for time extrapolation in 2018 are summarized in Table 18. The VINE4 method provides the smallest MAE for 14 stations, MRM for 16 stations and AR for 2 stations. The VINE4 has the smallest average MAE for all the stations. Since the dependence between months within each year is moderate, the advantage of the VINE4 method versus other models is limited, which agrees with our findings in Section 5. We also find that the MAEs of VINE4 are less variant. However, the MAEs based on other methods give prediction with extremely large MAEs in some occasions (results not shown here).

Table 18: Prediction result for time extrapolation in year 2018 (prediction standard error in the brackets)

	MAE					Percentage Outperformance			
Name	VINE4	MRM	LRM	AR	SARIMA	VINE4 vs MRM	VINE4 vs LRM	VINE4 vs AR	VINE4 vs SARIMA
LANSDOWNE HOUSE	**2.132 (2.064)**	2.222 (2.193)	4.262 (4.013)	3.210 (2.479)	2.335 (2.482)	0.583	0.740	0.688	0.602
PICKLE LAKE	2.252 (2.015)	**2.218 (2.193)**	4.085 (4.013)	3.256 (2.901)	2.406 (2.767)	0.583	0.726	0.726	0.627
RED LAKE	2.233 (1.998)	**2.154 (2.193)**	3.825 (4.013)	3.037 (2.551)	2.297 (2.590)	0.548	0.702	0.702	0.560
FORT FRANCES	2.511 (1.956)	2.636 (2.193)	3.339 (4.013)	2.541 (2.957)	**2.057 (2.596)**	0.667	0.648	0.537	0.430
MINE CENTRE	2.239 (1.961)	**2.218 (2.193)**	3.341 (4.013)	2.267 (2.649)	2.320 (2.431)	0.575	0.658	0.525	0.565
DRYDEN	2.117 (1.986)	**2.027 (2.193)**	3.785 (4.013)	2.735 (2.816)	2.387 (2.578)	0.491	0.713	0.620	0.600
KENORA	**2.096 (1.997)**	2.112 (2.193)	3.684 (4.013)	2.573 (3.128)	2.538 (2.603)	0.567	0.725	0.600	0.594
CAMERON FALLS	**1.994 (1.954)**	2.054 (2.193)	3.110 (4.013)	2.092 (2.578)	2.479 (2.303)	0.611	0.648	0.509	0.702
GERALDTON	2.184 (2.037)	**2.164 (2.193)**	3.502 (4.013)	2.532 (2.722)	2.374 (2.545)	0.583	0.694	0.542	0.560
THUNDER BAY	1.925 (1.940)	2.006 (2.193)	3.255 (4.013)	2.060 (3.162)	**1.779 (2.258)**	0.630	0.676	0.528	0.475
SAULT STE MARIE	1.961 (1.988)	2.048 (2.193)	3.211 (4.013)	2.311 (2.643)	**1.799 (2.082)**	0.567	0.592	0.617	0.520
WAWA	**1.976 (1.938**	2.240 (2.193)	2.882 (4.013)	2.353 (2.802)	2.452 (2.705)	0.702	0.583	0.619	0.635
CHAPLEAU	1.852 (1.960)	**1.832 (2.193)**	3.022 (4.013)	1.944 (2.760)	1.958 (2.169)	0.597	0.667	0.528	0.550
SUDBURY	1.894 (2.038)	1.831 (2.193)	3.118 (4.013)	**1.823 (2.559)**	2.055 (2.176)	0.467	0.692	0.467	0.642
EARLTON	**1.912 (2.030)**	1.919 (2.193)	3.203 (4.013)	2.014 (2.594)	2.097 (2.302)	0.542	0.650	0.567	0.584
KAPUSKASING	1.953 (2.032)	**1.888 (2.193)**	3.523 (4.013)	2.192 (2.821)	2.104 (2.413)	0.508	0.700	0.517	0.535
MOOSONEE	**1.974 (2.069)**	2.174 (2.193)	4.018 (4.013)	2.676 (2.991)	2.075 (2.280)	0.643	0.702	0.607	0.552
TIMMINS	1.985 (2.024)	**1.913 (2.193)**	3.342 (4.013)	2.072 (3.023)	2.189 (2.385)	0.500	0.675	0.458	0.550
MADAWASKA	2.221 (1.998)	2.472 (2.193)	3.189 (4.013)	2.305 (2.589)	**1.646 (2.053)**	0.667	0.650	0.547	0.395
NORTH BAY	1.850 (1.921)	**1.712 (2.193)**	2.675 (4.013)	1.737 (2.945)	1.836 (2.050)	0.467	0.583	0.450	0.520
GORE BAY	**1.740 (2.002)**	1.776 (2.193)	3.242 (4.013)	2.139 (2.601)	2.196 (1.976)	0.536	0.667	0.631	0.675
BROCKVILLE	**1.674 (1.938)**	1.737 (2.193)	2.814 (4.013)	1.889 (2.796)	1.952 (2.137)	0.512	0.690	0.528	0.550
CORNWALL	1.676 (2.019)	**1.632 (2.193)**	2.909 (4.013)	1.919 (2.614)	2.018 (2.248)	0.491	0.667	0.602	0.646
KINGSTON	**1.776 (2.000)**	1.805 (2.193)	2.904 (4.013)	1.930 (2.769)	1.788 (1.905)	0.611	0.611	0.565	0.510
OTTAWA	1.737 (1.986)	**1.687 (2.193)**	2.728 (4.013)	1.734 (2.689)	1.725 (2.231)	0.509	0.648	0.463	0.476
RIDGETOWN	**2.094 (1.996)**	2.201 (2.193)	3.329 (4.013)	2.485 (2.491)	2.216 (2.010)	0.573	0.677	0.604	0.570
VINELAND	1.804 (2.003)	**1.682 (2.193)**	3.204 (4.013)	2.024 (2.712)	1.752 (1.768)	0.476	0.690	0.548	0.510
WELLAND	1.891 (2.017)	**1.817 (2.193)**	2.985 (4.013)	2.203 (2.657)	1.821 (1.871)	0.533	0.617	0.583	0.557
WINDSOR	2.041 (2.031)	1.914 (2.193)	2.678 (4.013)	**1.864 (2.748)**	1.927 (2.338)	0.500	0.537	0.491	0.520
LONDON	**1.800 (2.023)**	1.898 (2.193)	2.700 (4.013)	1.920 (2.624)	2.000 (1.925)	0.529	0.593	0.565	0.574
WOODSTOCK	1.722 (1.961)	**1.529 (2.193)**	2.370 (4.013)	1.685 (2.731)	1.828 (2.018)	0.472	0.556	0.536	0.545
BELLEVILLE	**1.885 (2.053)**	2.023 (2.193)	3.160 (4.013)	1.922 (2.600)	1.949 (1.911)	0.556	0.583	0.667	0.630
HAMILTON	1.801 (2.037)	**1.778 (2.193)**	2.899 (4.013)	2.009 (2.450)	1.788 (1.902)	0.542	0.583	0.575	0.542
ORANGEVILLE	**1.748 (1.984)**	1.831 (2.193)	2.767 (4.013)	1.983 (2.306)	2.035 (2.412)	0.639	0.556	0.625	0.642
TORONTO	1.788 (1.849)	**1.755 (2.193)**	2.828 (4.013)	2.016 (2.728)	1.958 (2.387)	0.556	0.648	0.546	0.520
HALIBURTON	**1.880(2.104)**	1.955 (2.193)	2.888 (4.013)	1.922 (2.512)	1.939 (2.104)	0.594	0.594	0.565	0.580
PETERBOROUGH	**1.880 (2.038)**	2.010 (2.193)	2.861 (4.013)	2.085 (2.366)	1.998 (2.029)	0.583	0.575	0.550	0.545
Average	**1.951 (2.014)**	1.969 (2.193)	3.179 (4.013)	2.202 (2.707)	2.056 (2.242)	0.560	0.646	0.568	0.562

In addition, we also try to predict the temperature value of the last three seasons in a year, given the temperature values in the first season, i.e. months 1-3. For the prediction of temperature in the month $l > 4$, we plug in our prediction in months from 4 to $l - 1$ and combine with the true temperature in months 1-3 to form the temperature information in the previous months.

Table 19: Prediction results for subject extrapolation of month 4-12, given the first 3 months (prediction standard error in the brackets)

Name	MAE				Percentage Outperformance		
	VINE4	MRM	LRM	AR	VINE4 vs MRM	VINE4 vs LRM	VINE4 vs AR
BIG TROUT LAKE	**1.721 (1.599)**	1.858 (1.907)	3.838 (4.056)	2.688 (2.739)	0.525	0.764	0.736
SIOUX LOOKOUT	**1.836 (1.616)**	2.010 (1.907)	3.629 (4.056)	2.777 (2.797)	0.578	0.711	0.789
BEATRICE	**1.322 (1.584)**	1.388 (1.907)	1.962 (4.056)	1.605 (2.629)	0.569	0.625	0.708
HARROW	1.586 (1.590)	**1.524 (1.907)**	2.151 (4.056)	1.571 (2.592)	0.458	0.611	0.472
ATITOKAN	**1.500 (1.595)**	1.612 (1.907)	3.055 (4.056)	2.027 (2.447)	0.556	0.792	0.681
Average	**1.593 (1.597)**	1.679 (1.907)	2.927 (4.056)	2.134 (2.641)	0.537	0.701	0.677

Table 20: Prediction results for time extrapolation of month 4-12, given the first 3 months (prediction standard error in the brackets)

	MAE					Percentage Outperformance			
Name	VINE4	MRM	LRM	AR	SARIMA	VINE4 vs MRM	VINE4 vs LRM	VINE4 vs AR	VINE4 vs SARIMA
LANSDOWNE HOUSE	**1.900 (1.564)**	1.945 (2.541)	3.894 (4.013)	2.580 (2.710)	2.498 (2.567)	0.514	0.764	0.694	0.665
PICKLE LAKE	2.027 (1.638)	**1.999 (2.541)**	3.976 (4.013)	2.976 (3.010)	2.357 (2.578)	0.508	0.746	0.730	0.654
RED LAKE	1.867 (1.549)	**1.827 (2.541)**	3.581 (4.013)	2.701 (2.597)	2.250 (2.487)	0.524	0.730	0.683	0.625
FORT FRANCES	1.865 (1.550)	2.149 (2.541)	2.751 (4.013)	**1.842 (2.878)**	2.094 (2.733)	0.704	0.654	0.481	0.580
MINE CENTRE	1.764 (1.562)	1.852 (2.541)	2.826 (4.013)	**1.724 (2.659)**	2.122 (2.559)	0.633	0.689	0.478	0.657
DRYDEN	1.952 (1.568)	**1.897 (2.541)**	3.609 (4.013)	2.704 (3.207)	2.441 (2.715)	0.556	0.741	0.679	0.634
KENORA	**1.823 (1.566)**	1.861 (2.541)	3.386 (4.013)	2.246 (3.155)	2.408 (2.735)	0.578	0.789	0.589	0.628
CAMERON FALLS	1.795 (1.572)	1.826 (2.541)	2.500 (4.013)	1.740 (2.635)	**1.730 (2.399)**	0.580	0.580	0.506	0.480
GERALDTON	1.798 (1.679)	**1.785 (2.541)**	3.074 (4.013)	1.878 (2.337)	2.145 (2.329)	0.593	0.685	0.444	0.575
THUNDER BAY	1.736 (1.567)	1.841 (2.541)	2.573 (4.013)	1.776 (2.655)	**1.627 (2.387)**	0.605	0.642	0.516	0.420
SAULT STE MARIE	**1.709 (1.572)**	1.771 (2.541)	2.412 (4.013)	2.104 (2.937)	1.754 (2.229)	0.600	0.533	0.633	0.547
WAWA	**1.912 (1.701)**	1.991 (2.541)	2.131 (4.013)	2.001 (2.816)	2.125 (2.248)	0.603	0.476	0.572	0.625
CHAPLEAU	1.538 (1.492)	1.496 (2.541)	2.498 (4.013)	**1.479 (3.068)**	1.895 (2.174)	0.611	0.685	0.426	0.615
SUDBURY	1.696 (1.577)	1.634 (2.541)	2.531 (4.013)	**1.592 (2.870)**	2.064 (2.257)	0.556	0.656	0.456	0.585
EARLTON	1.654 (1.572)	1.665 (2.541)	2.564 (4.013)	**1.583 (2.869)**	2.013 (2.367)	0.522	0.622	0.422	0.735
KAPUSKASING	1.739 (1.569)	**1.641 (2.541)**	2.984 (4.013)	1.751 (2.668)	2.097 (2.367)	0.500	0.733	0.500	0.685
MOOSONEE	**1.822 (1.564)**	1.980 (2.541)	3.283 (4.013)	2.089 (2.623)	2.771 (2.405)	0.651	0.730	0.540	0.694
TIMMINS	1.727 (1.574)	1.653 (2.541)	2.755 (4.013)	**1.639 (3.061)**	2.201 (2.425)	0.533	0.644	0.467	0.605
MADAWASKA	2.097 (1.571)	2.301 (2.541)	2.431 (4.013)	2.111 (2.696)	**1.749 (2.098)**	0.644	0.589	0.512	0.450
NORTH BAY	1.577 (1.603)	**1.508 (2.541)**	2.282 (4.013)	1.622 (2.808)	1.967 (1.857)	0.556	0.622	0.533	0.585
GORE BAY	**1.486 (1.574)**	1.524 (2.541)	2.366 (4.013)	1.979 (2.679)	1.811 (2.035)	0.524	0.587	0.698	0.628
BROCKVILLE	**1.481 (1.562)**	1.513 (2.541)	1.984 (4.013)	1.711 (2.974)	1.823 (1.987)	0.528	0.604	0.540	0.580
CORNWALL	1.470 (1.594)	**1.403 (2.541)**	2.274 (4.013)	1.771 (2.908)	1.896 (2.036)	0.457	0.654	0.593	0.632
KINGSTON	**1.561 (1.576)**	1.583 (2.541)	2.214 (4.013)	1.801 (3.096)	1.626 (1.972)	0.544	0.617	0.630	0.560
OTTAWA	1.493 (1.570)	**1.388 (2.541)**	2.081 (4.013)	1.488 (2.853)	1.649 (2.105)	0.469	0.617	0.481	0.554
RIDGETOWN	**1.740 (1.569)**	1.784 (2.541)	2.608 (4.013)	2.465 (2.655)	1.869 (2.136)	0.569	0.667	0.694	0.580
VINELAND	1.598 (1.563)	**1.491 (2.541)**	2.498 (4.013)	1.838 (2.691)	1.563 (1.841)	0.429	0.746	0.540	0.510
WELLAND	1.570 (1.597)	1.503 (2.541)	2.160 (4.013)	2.042 (3.126)	**1.494 (1.924)**	0.489	0.600	0.667	0.450
WINDSOR	1.740 (1.620)	**1.547 (2.541)**	1.872 (4.013)	1.585 (2.718)	1.847 (1.975)	0.407	0.444	0.506	0.550
LONDON	**1.604 (1.578)**	1.746 (2.541)	1.915 (4.013)	1.719 (2.798)	1.820 (2.004)	0.644	0.580	0.556	0.575
WOODSTOCK	1.437 (1.491)	**1.411 (2.541)**	1.702 (4.013)	1.725 (3.003)	1.928 (1.985)	0.481	0.444	0.593	0.695
BELLEVILLE	**1.687 (1.576)**	1.755 (2.541)	2.261 (4.013)	1.704 (2.853)	2.026 (1.967)	0.519	0.556	0.593	0.610
HAMILTON	1.538 (1.577)	**1.486 (2.541)**	2.098 (4.013)	1.827 (2.662)	1.769 (1.969)	0.467	0.567	0.611	0.585
ORANGEVILLE	**1.438 (1.520)**	1.475 (2.541)	1.887 (4.013)	1.692 (2.728)	1.958 (2.154)	0.537	0.556	0.630	0.654
TORONTO	1.527 (1.595)	**1.466 (2.541)**	1.983 (4.013)	1.685 (2.614)	2.035 (1.936)	0.543	0.630	0.556	0.620
HALIBURTON	**1.651 (1.578)**	1.750 (2.541)	2.105 (4.013)	1.760 (2.890)	2.066 (2.156)	0.569	0.556	0.500	0.615
PETERBOROUGH	**1.797 (1.588)**	1.813 (2.541)	2.128 (4.013)	1.930 (3.258)	1.843 (2.079)	0.500	0.522	0.511	0.505
Average	**1.698 (1.578)**	1.710 (2.541)	2.545 (4.013)	1.915 (2.832)	1.979 (2.221)	0.547	0.629	0.561	0.593

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198.

Bedford, T. and Cooke, R. M. (2002). Vines–a new graphical model for dependent random variables. *Ann. Statist.*, 30(4):1031–1068.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts.* Chapman and Hall/CRC.

Nelsen, R. B. (2007). *An introduction to copulas.* Springer Science & Business Media.

Sklar, A. (1959). Functions de repartition an dimension set leursmarges. *Publications de L'Institut de Statistique de L'Universite de Paris*, 8:229–231.