



Supplement of

Characterizing Climate-Driven Shifts in Chilean Rainfall Regimes with a Hybrid Hidden Markov–Copula Framework

Mauricio Herrera-Marín et al.

Correspondence to: Mauricio Herrera-Marín (mherrera@udd.cl)

The copyright of individual parts of the supplement might differ from the article licence.

1 Supplementary Information:
2 Characterizing Climate-Driven Shifts in Chilean
3 Rainfall Regimes
4 with a Hybrid Hidden Markov–Copula Framework

5 **Contents**

6	S1 nHMM Calibration Diagnostics	2
7	S1.1 PIT histograms	3
8	S1.2 KS/CvM p-values for Rosenblatt transform uniformity	3
9	S2 DTW Zoning: Methodology and Full CVI Table	3
10	S2.1 DTW distance and Ward clustering	3
11	S2.2 Complete CVI table	4
12	S3 GPD Tail Diagnostics	4
13	S3.1 Mean-excess plots (tail-stability diagnostic)	4
14	S3.2 Full GPD parameter table (all zones and states)	4
15	S4 Vine-Copula Selection and Goodness-of-Fit	5
16	S4.1 Full vine structure table (all zones)	5
17	S4.2 Simplifying assumption check	5
18	S5 Multivariate Scoring (ES/VS)	6
19	S6 Areal CRPS and Event Verification	7
20	S7 State–Index Relationships: Rank Correlations	7
21	S8 Out-of-Sample Verification Details	9
22	S8.1 PIT histograms and CRPS distributions (MJJA 2022)	9
23	S8.2 Spatial generalization: non-medoid stations	9
24	S8.3 Rolling-origin hindcast (2012–2021)	9
25	S9 Full Model-Selection Grid and Sensitivity Analyses	10
26	S9.1 Full BIC grid (30 configurations)	10
27	S9.2 Sensitivity: vine family set, truncation, GPD threshold	14

Organization.

- **S1:** nHMM calibration diagnostics — PIT histograms and KS/CvM p -values by zone and state.
- **S2:** DTW zoning — full methodology and complete CVI table.
- **S3:** GPD tail diagnostics — mean-excess plots and full per-zone parameter estimates.
- **S4:** Vine-copula selection — structure summaries and Rosenblatt transform diagnostics.
- **S5:** Multivariate scoring (ES/VS) and station-set definitions.
- **S6:** Areal CRPS and event verification (Brier scores).
- **S7:** State–index relationships: rank correlations and robustness checks.
- **S8:** Out-of-sample verification details — MJJA 2022 and rolling-origin hindcasts 2012–2021.
- **S9:** Full model-selection grid ($K \times C$, 30 configurations) and sensitivity analyses.

Note on arid zones (2, 4). As discussed in the main text (Sect. 4.1), Zones 2 and 4 exhibit anomalous areal-CRPS values because the vine ensemble occasionally generates large positive precipitation draws in near-zero regimes. This is a genuine limitation of the current framework in arid environments; interval-based diagnostics (Tables 6 and S6) are the preferred evaluation tool for these zones.

S1 nHMM Calibration Diagnostics

This section amplifies the validation framework described in main Sect. 2.4.1. Conditional PIT residuals $U_{t,m} = F_{m|s}(R_{t,m})$ are computed using Viterbi-decoded states (TRAIN: 1980–2021). Under correct specification they should be i.i.d. $\text{Unif}(0, 1)$ [1, 2]. Tables S1–S5 report KS and Cramér–von Mises (CvM) p -values applied to the first five Rosenblatt-transform coordinates of each state–zone combination.

Reading guide.

$p \gtrsim 0.10$: no evidence against uniformity. $p \lesssim 0.05$: possible dependence misfit. States 1 and 2 consistently show $p \approx 0$ across most zones in the first two Rosenblatt components, reflecting the relatively sparse training samples ($n \approx 845$ –1269 days) and the non-Gaussian dependence structure captured by the Archimedean vine families. States 3, 4, and 5 show broadly adequate calibration across the rainfall-active zones (5–6), with most components passing at the 10 % level.

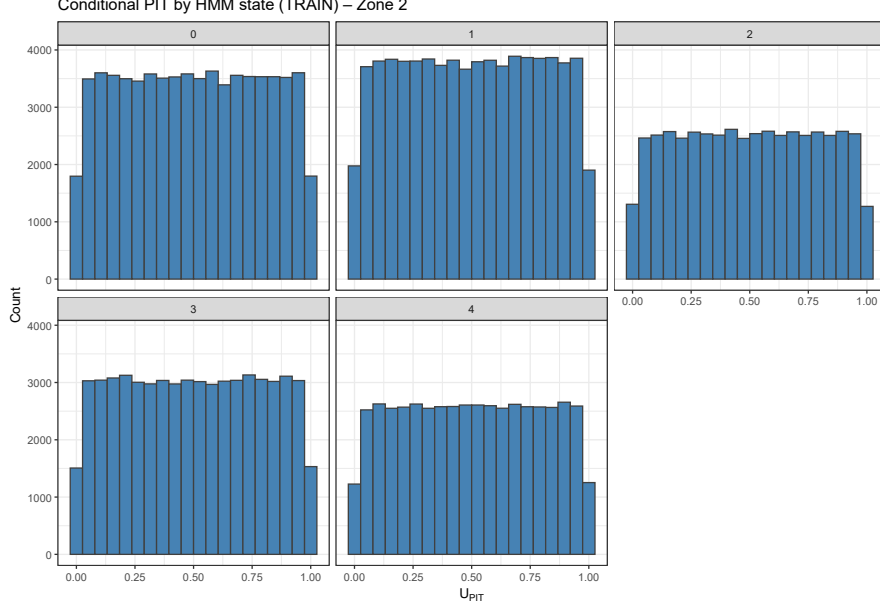


Fig. S1: Conditional PIT histograms by nHMM state (TRAIN) for Zone 02. Bars: empirical frequencies; dashed line: $\text{Unif}(0, 1)$.

S1.1 PIT histograms

S1.2 KS/CvM p-values for Rosenblatt transform uniformity

S2 DTW Zoning: Methodology and Full CVI Table

S2.1 DTW distance and Ward clustering

Let X_i and X_j be two standardized MJJA precipitation series. The DTW cumulative cost satisfies:

$$D(m, n) = |x_{i,m} - x_{j,n}| + \min\{D(m-1, n), D(m, n-1), D(m-1, n-1)\}, \quad (\text{S1})$$

with $D(0, 0) = 0$ and boundary conditions $D(m, 0) = D(0, n) = \infty$ [3]. Ward's minimum variance criterion [4] is applied to the pairwise DTW distance matrix; the medoid of zone Z_z is $\hat{X}_z = \arg \min_{X_i \in Z_z} \sum_{X_j \in Z_z} D_{\text{DTW}}(X_i, X_j)$.

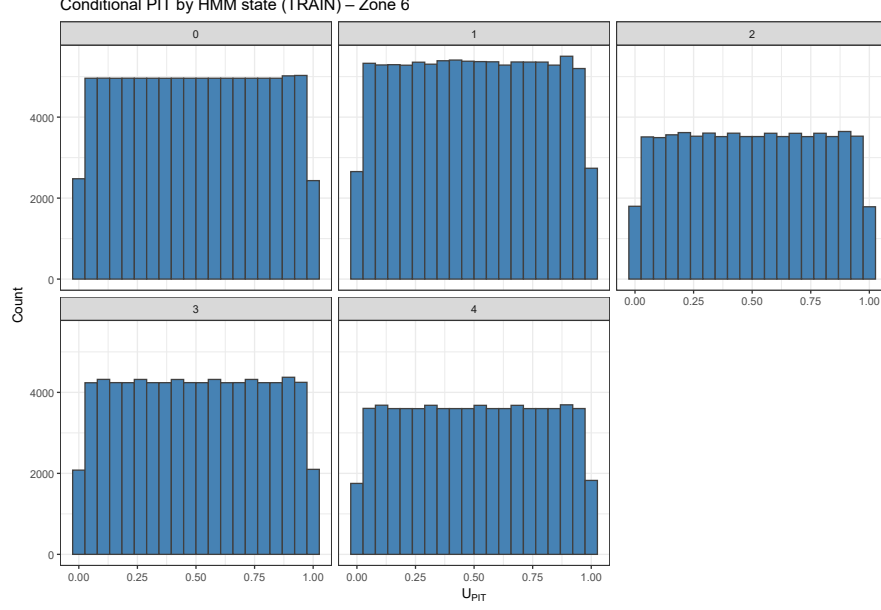


Fig. S2: As Fig. S1 for Zone 06 (wet southern zone). PIT histograms are broadly flat for states 0, 3, 4; states 1 and 2 show mild U-shapes consistent with the heavy-tail GPD fit and sparse training samples.

S2.2 Complete CVI table

S3 GPD Tail Diagnostics

S3.1 Mean-excess plots (tail-stability diagnostic)

For a valid GPD fit above u^* , the mean-excess function $e(u) = \mathbb{E}[X - u \mid X > u]$ should be approximately linear above the threshold. A positive slope indicates $\hat{\xi} > 0$ (heavy tail); a near-zero slope indicates $\hat{\xi} \approx 0$; a negative slope indicates $\hat{\xi} < 0$ (bounded tail).

S3.2 Full GPD parameter table (all zones and states)

Table S7 reports GPD estimates across all zones. Bootstrap 95% CIs use $B = 500$ parametric resamples of each medoid's own exceedances (not neighbour data), guaranteeing that the CI contains $\hat{\xi}$ by construction. The sign contrast of $\hat{\xi}$ between dry states (0, 3, 4: bounded tails) and episodic-frontal states (1, 2: heavy tails) is consistent across zones and motivates the state-specific tail treatment. Entries marked “no GPD” had fewer than 25 exceedances above u^* .

Table S1: KS and CvM p -values for first 5 Rosenblatt components (Zone 2, TRAIN).

State	Comp	KS p	CvM p	n	Notes
0	1	0.000	0.000	1179	Residual misfit
0	2	0.000	0.000	1179	
0	3	0.000	0.000	1179	
0	4	0.000	0.000	1179	
0	5	0.000	0.000	1179	
1	1	0.000	0.000	1269	
1	2	0.000	0.000	1269	
1	3	0.000	0.000	1269	
1	4	0.000	0.000	1269	
1	5	0.000	0.000	1269	
2	1	0.000	0.000	845	
2	2	0.000	0.000	845	
2	3	0.000	0.000	845	
2	4	0.000	0.000	845	
2	5	0.000	0.000	845	
3	1	0.000	0.000	1013	
3	2	0.000	0.000	1013	
3	3	0.000	0.000	1013	
3	4	0.000	0.000	1013	
3	5	0.000	0.000	1013	
4	1	0.000	0.000	860	
4	2	0.000	0.000	860	
4	3	0.000	0.000	860	
4	4	0.000	0.000	860	
4	5	0.000	0.000	860	

S4 Vine-Copula Selection and Goodness-of-Fit

S4.1 Full vine structure table (all zones)

Key patterns.

Archimedean families (Clayton, Gumbel, Frank) dominate throughout, reflecting asymmetric and non-Gaussian spatial dependence. Gaussian pairs are rare ($< 10\%$) in the northern arid zones (2–3) and increase toward the south (Zones 5–7), consistent with more Gaussian-like precipitation dependence under persistent frontal forcing. Wet states (3, 4) tend to have higher Gaussian percentages than dry states, likely because the heavier precipitation totals under frontal conditions produce more elliptically distributed spatial fields.

S4.2 Simplifying assumption check

The simplifying assumption (pair-copula parameters independent of conditioning values) was examined by comparing pair families and parameter estimates across conditioning-value terciles on low-order edges. Family assignments and parameter magnitudes were stable across terciles in all zones and states; no systematic ordering was detected that would indicate a strong violation. This finding is consistent with

Table S2: As Table S1 for Zone 3.

State	Comp	KS p	CvM p	n	Notes
0	1–5	0.000	0.000	1179	All comps: residual misfit
1	1–5	0.000	0.000	1269	
2	1	0.353	0.481	845	Adequate
2	2	0.174	0.372	845	Adequate
2	3	0.802	0.821	845	Adequate
2	4	0.000	0.000	845	Misfit
2	5	0.007	0.026	845	Borderline
3	1–5	0.000	0.000	1013	
4	1	0.003	0.015	860	Borderline
4	2	0.108	0.223	860	Adequate
4	3	0.989	0.994	860	Adequate
4	4	0.000	0.000	860	Misfit
4	5	0.000	0.000	860	Misfit

Table S3: As Table S1 for Zone 5 (southern, storm-dominated).

State	Comp	KS p	CvM p	n	Notes
0	1–5	0.000	0.000	1179	
1	1–5	0.000	0.000	1269	
2	1	0.000	0.000	845	
2	2	0.004	0.021	845	Borderline
2	3	0.000	0.000	845	
2	4	0.017	0.013	845	Borderline
2	5	0.000	0.024	845	
3	1	0.408	0.496	1013	Adequate
3	2	0.297	0.467	1013	Adequate
3	3	0.768	0.700	1013	Adequate
3	4	0.481	0.522	1013	Adequate
3	5	0.006	0.011	1013	Borderline
4	1	0.401	0.495	860	Adequate
4	2	0.382	0.523	860	Adequate
4	3	0.001	0.003	860	Misfit
4	4	0.938	0.972	860	Adequate
4	5	0.120	0.114	860	Adequate

the expected attenuation of conditional dependence at higher tree levels and supports retaining the simplifying assumption for the truncated ($\leq$ Tree 2) R-vines.

S5 Multivariate Scoring (ES/VS)

Energy scores (ES) and variogram scores (VS) are computed on fixed station sets per state (constant $d \geq 2$ across all test days) to ensure stable multivariate dimension during scoring. Due to the areal-CRPS anomalies in Zones 2 and 4 (see main

Table S4: As Table S1 for Zone 6 (wet southern).

State	Comp	KS p	CvM p	n	Notes
0	1	0.651	0.609	1179	Adequate
0	2	0.408	0.473	1179	Adequate
0	3	0.447	0.529	1179	Adequate
0	4	0.066	0.165	1179	Adequate
0	5	0.365	0.403	1179	Adequate
1	1–5	0.000	0.000	1269	
2	1	0.024	0.066	845	Borderline
2	2	0.000	0.000	845	
2	3	0.013	0.076	845	Borderline
2	4	0.003	0.013	845	Borderline
2	5	0.000	0.000	845	
3	1	0.256	0.274	1013	Adequate
3	2	0.503	0.416	1013	Adequate
3	3	0.168	0.236	1013	Adequate
3	4	0.113	0.106	1013	Adequate
3	5	0.382	0.457	1013	Adequate
4	1	0.037	0.121	860	Borderline
4	2	0.263	0.301	860	Adequate
4	3	0.751	0.655	860	Adequate
4	4	0.055	0.045	860	Borderline
4	5	0.372	0.691	860	Adequate

102 text, Sect. 3.5.2 and Supplement S6), ES/VS diagnostics for those zones should be
103 interpreted with caution.

104 S6 Areal CRPS and Event Verification

105 *Zones 5 and 6 (rainfall-active).*

106 The vine ensemble yields statistically significant improvements in mean areal-CRPS:
107 Zone 5: $2.67 \rightarrow 3.25$ (18 % reduction; $p_{\text{DM}} < 0.001$); Zone 6: $3.47 \rightarrow 4.71$ (26 %
108 reduction; $p_{\text{DM}} < 0.001$). Pinball losses at $\tau \in \{0.90, 0.95\}$ are consistently lower for
109 the vine model (see main Table 7), confirming skill at upper quantiles.

110 *Zones 2 and 4 (near-zero precipitation).*

111 The areal-CRPS anomalies (CRPSS = -2152 and -214 respectively) are caused by
112 outlier ensemble draws under near-zero precipitation: when essentially all observations
113 are zero but the vine occasionally samples large positive values, the areal mean is
114 inflated and CRPS diverges. The interval-based diagnostics (main Table 6) are the
115 appropriate tool for these zones.

116 S7 State–Index Relationships: Rank Correlations

117 *Effective sample size.*

118 Under AR(1) persistence with lag-1 autocorrelations r_x and r_y : $N_{\text{eff}} = N(1 -$
119 $r_x r_y)/(1 + r_x r_y)$ [5]. Pearson’s r is tested via Fisher’s z -transform; nonparametric

Table S5: As Table S1 for Zone 7.

State	Comp	KS p	CvM p	n	Notes
0	1	0.051	0.164	1179	Adequate
0	2	0.190	0.445	1179	Adequate
0	3	0.000	0.000	1179	
0	4	0.002	0.003	1179	Borderline
0	5	0.031	0.156	1179	Borderline
1	1	0.037	0.055	1269	Borderline
1	2–5	0.000	0.000	1269	
2	1–5	0.000	0.000	845	
3	1	0.857	0.907	1013	Adequate
3	2	0.049	0.076	1013	Adequate
3	3	0.259	0.180	1013	Adequate
3	4	0.053	0.027	1013	Borderline
3	5	0.192	0.128	1013	Adequate
4	1	0.877	0.812	860	Adequate
4	2	0.141	0.152	860	Adequate
4	3	0.000	0.000	860	
4	4	0.005	0.006	860	Borderline
4	5	0.093	0.198	860	Adequate

Table S6: All DTW zoning candidates ($G \in \{5, \dots, 10\}$, two linkage variants). $\uparrow$: larger is better; $\downarrow$: smaller is better. **Bold:** selected.

Config.	Sil $\uparrow$	CH $\uparrow$	DB $\downarrow$	DB* $\downarrow$	Dunn $\uparrow$	COP $\downarrow$	Rank
ward.D2, 7	0.390	186.4	0.743	0.899	0.012	0.241	1
ward.D2, 6	0.393	201.6	0.810	0.957	0.012	0.271	2
ward.D, 6	0.401	212.9	0.787	1.113	0.016	0.259	3
ward.D2, 5	0.390	235.4	0.816	1.098	0.011	0.291	4
ward.D2, 4	0.381	265.7	0.888	1.102	0.011	0.341	5
ward.D, 5	0.376	224.3	0.843	1.128	0.016	0.306	6
ward.D, 4	0.362	263.8	0.905	1.141	0.016	0.342	7
ward.D, 7	0.381	193.8	0.748	1.315	0.007	0.236	8

percentile CIs from the stationary block bootstrap ($\ell \propto N^{1/3}$) [6]. Multiple testing controlled by Benjamini–Hochberg FDR at $q = 0.10$ [7].

Findings.

Signs and spatial patterns of all correlations are unchanged across Pearson, Spearman, and Kendall analyses. Median absolute difference between Pearson and Spearman < 0.05 . Dry-state occupancy f_1 retains its positive association with SOI and negative associations with ONI/Niño 3.4 after FDR correction. These results confirm that the teleconnections reported in the main text are monotonic but not strictly linear, motivating the non-Gaussian copula component of the framework.

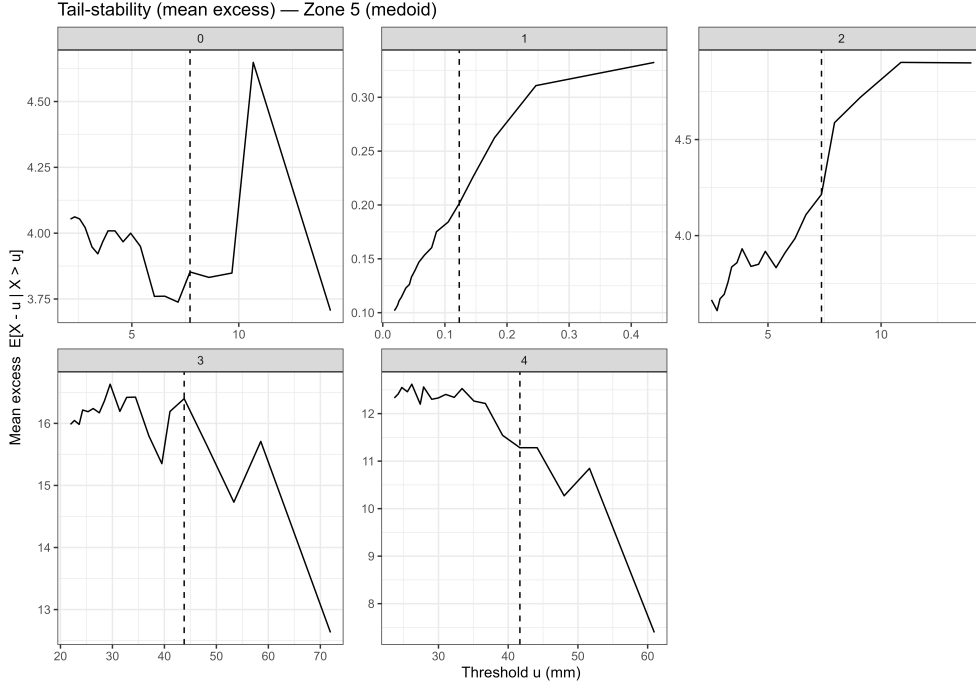


Fig. S3: Mean-excess plots at the Zone 5 medoid by nHMM state. Linear behaviour above the dashed threshold u^* confirms GPD adequacy. States 1–2: positive slope ($\hat{\xi} > 0$, heavy tail). States 0, 3, 4: near-zero or negative slope ($\hat{\xi} \leq 0$, bounded or near-exponential tail).

S8 Out-of-Sample Verification Details

S8.1 PIT histograms and CRPS distributions (MJJA 2022)

S8.2 Spatial generalization: non-medoid stations

To verify that the regime–copula structure transfers within zones (not an artifact of medoid selection), MJJA 2022 forecasts were verified at three additional non-medoid stations per zone. Distributions of POC and CRPSS at non-medoid stations mirror medoid behaviour; no systematic degradation was observed in Zones 5–7. Detailed station-level results are available in the project repository.

S8.3 Rolling-origin hindcast (2012–2021)

A rolling-origin hindcast was performed for each year $y \in \{2012, \dots, 2021\}$ (train 1980–($y - 1$), forecast MJJA y). Median CRPSS is positive throughout for Zones 5–6, with interannual fluctuations tracking ENSO phase (higher skill during strong El Niño or La Niña years) but preserving overall gains across the decade.

Table S7: GPD parameter estimates, all zones and states (medoid station).
 u^* : 0.90-quantile threshold; 95 % boot CI for $\hat{\xi}$. “no GPD”: $n_{\text{exc}} < 25$.

Zone	State	u^*	$\hat{\sigma}$	$\hat{\xi}$	CI_L	CI_U	n_{exc}	Tail
2	0	—	—	—	—	—	15	no GPD
	1	1.85	1.84	+0.368	−0.20	+0.80	30	heavy
	2	11.32	8.00	−0.003	−0.49	+0.21	68	~exp
	3	—	—	—	—	—	18	no GPD
	4	15.87	19.60	−0.096	−0.36	+0.07	68	bounded
3	0	0.62	0.90	−0.052	−0.63	+0.43	32	bounded
	1	0.86	1.72	−0.094	−0.67	+0.21	33	bounded
	2	14.14	8.31	+0.204	−0.06	+0.44	82	heavy
	3	3.03	2.88	−0.005	−0.31	+0.22	68	~exp
	4	46.69	19.79	−0.034	−0.29	+0.16	86	bounded
4	0	0.19	0.21	+0.921	+0.37	+1.39	52	heavy
	1	0.13	0.14	+0.179	−0.39	+0.91	34	heavy
	2	10.78	7.96	+0.006	−0.29	+0.26	81	~exp
	3	16.05	12.67	+0.064	−0.21	+0.25	96	heavy
	4	57.06	23.30	−0.254	−0.92	−0.08	86	bounded
5	0	7.72	3.92	−0.018	−0.30	+0.17	113	bounded
	1	0.12	0.11	+0.458	−0.20	+0.80	60	heavy
	2	7.37	3.72	+0.121	−0.19	+0.40	77	heavy
	3	43.80	19.09	−0.161	−0.37	0.00	102	bounded
	4	41.68	12.53	−0.112	−0.53	+0.06	86	bounded
6	0	49.75	16.75	−0.116	−0.36	+0.01	118	bounded
	1	14.20	6.50	+0.361	+0.11	+0.63	100	heavy
	2	16.40	9.53	+0.234	−0.06	+0.56	75	heavy
	3	46.63	11.60	−0.038	−0.20	+0.15	102	bounded
	4	35.57	8.15	−0.016	−0.28	+0.19	85	bounded
7	0	9.40	5.30	−0.050	−0.20	+0.08	117	bounded
	1	11.66	8.01	+0.069	−0.17	+0.29	123	heavy
	2	6.26	3.72	+0.183	−0.20	+0.49	81	heavy
	3	6.75	4.31	−0.163	−0.41	+0.12	101	bounded
	4	5.16	4.33	−0.082	−0.30	+0.19	84	bounded

142 S9 Full Model-Selection Grid and Sensitivity 143 Analyses

144 S9.1 Full BIC grid (30 configurations)

Table S8: Selected vine structures (all zones and states). Type: R-vine; Trunc: truncation level; %Gauss: % Gaussian pairs; %Arch: % Archimedean pairs (Clayton, Gumbel, Frank); n : training days.

Zone	State	Trunc.	%Gauss	%Arch	n
2	0	2	0.0	11.7	1179
2	1	2	0.0	7.2	1269
2	2	2	0.9	23.1	845
2	3	2	0.0	16.2	1013
2	4	2	2.7	16.2	860
3	0	2	0.0	10.6	1179
3	1	2	0.0	6.8	1269
3	2	2	1.5	28.5	845
3	3	2	3.1	15.3	1013
3	4	2	5.4	38.8	860
4	0	2	0.0	16.2	1179
4	1	2	0.0	12.1	1269
4	2	2	2.0	27.3	845
4	3	2	5.2	37.1	1013
4	4	2	17.3	11.2	860
5	0	2	3.2	20.4	1179
5	1	2	0.0	15.9	1269
5	2	2	4.5	36.5	845
5	3	2	7.8	22.7	1013
5	4	2	21.9	21.2	860
6	0	2	5.4	24.8	1179
6	1	2	2.7	18.2	1269
6	2	2	6.7	33.6	845
6	3	2	3.4	23.1	1013
6	4	2	14.2	33.1	860
7	0	2	7.6	20.0	1179
7	1	2	4.7	23.6	1269
7	2	2	7.6	35.2	845
7	3	2	6.5	27.1	1013
7	4	2	8.6	39.0	860

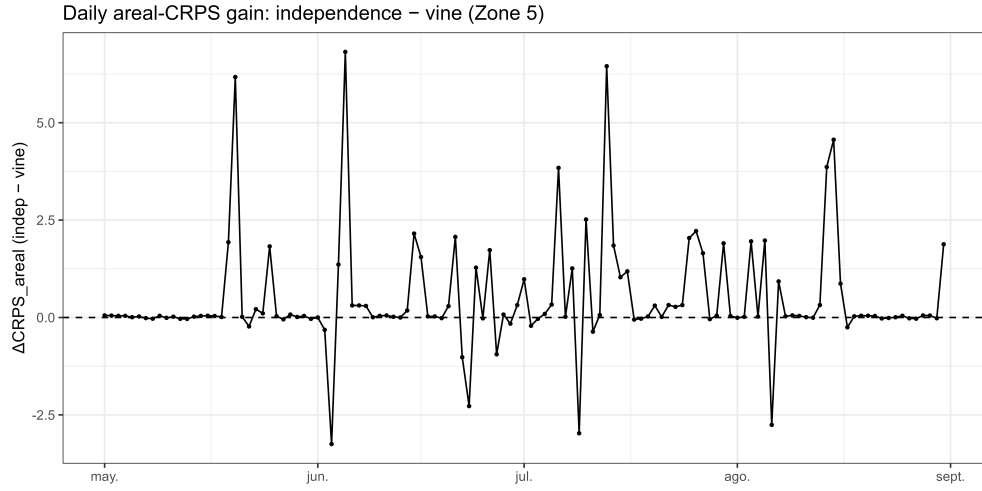


Fig. S4: Daily areal-CRPS gains (independence – vine) for Zone 05. Positive values favour the vine model; the median gain is positive throughout MJJA 2022 and concentrated in storm-event days.

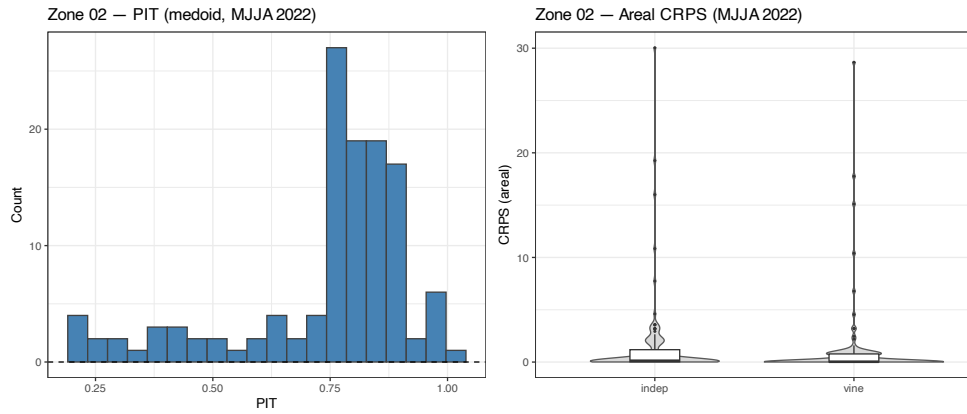


Fig. S5: PIT histograms (medoid) and areal-CRPS distributions (vine vs. independence) for Zone 05, MJJA 2022. PIT is broadly consistent with $\text{Unif}(0, 1)$; the vine ensemble yields systematically lower CRPS.

Table S9: Full model-selection grid ($K \times C$, 30 configurations). Lower AIC/BIC is better. N/p : identifiability ratio ($N_{\text{obs}} = 2,386,692$). **Bold:** adopted configuration. [†]: BIC minimum within $K = 5$ (see main Sect. 2.5 footnote).

Config.	K	C	log-lik	AIC	BIC	p	N/p
5hs5dg	5	5	-5,108,541	10,275,650	10,483,644	29,284	82
5hs3dg [†]	5	3	-5,151,741	10,332,950	10,437,601	14,734	162
5hs4dg	5	4	-5,122,235	10,288,488	10,444,811	22,009	108
5hs6dg	5	6	-5,102,387	10,277,893	10,537,558	36,559	65
5hs7dg	5	7	-5,099,291	10,286,251	10,597,588	43,834	54
5hs8dg	5	8	-5,092,477	10,287,172	10,650,181	51,109	47
5hs2dg	5	2	-5,398,125	10,811,168	10,864,146	7,459	320
6hs3dg	6	3	-5,006,775	10,048,940	10,174,621	17,695	135
6hs4dg	6	4	-4,967,902	9,988,655	10,176,342	26,425	90
6hs2dg	6	2	-5,258,452	10,534,833	10,598,508	8,965	266
4hs4dg	4	4	-5,286,428	10,608,046	10,733,017	17,595	136
4hs3dg	4	3	-5,326,106	10,675,762	10,759,395	11,775	203
4hs5dg	4	5	-5,278,886	10,604,603	10,770,911	23,415	102
4hs6dg	4	6	-5,271,477	10,601,423	10,809,069	29,235	82
4hs7dg	4	7	-5,264,731	10,599,573	10,848,556	35,055	68
4hs8dg	4	8	-5,265,315	10,612,381	10,902,702	40,875	58
4hs2dg	4	2	-5,597,664	11,207,238	11,249,534	5,955	401
3hs4dg	3	4	-5,495,957	11,018,280	11,111,914	13,183	181
3hs5dg	3	5	-5,485,567	11,006,229	11,130,867	17,548	136
3hs6dg	3	6	-5,477,117	10,998,060	11,153,700	21,913	109
3hs3dg	3	3	-5,537,114	11,091,865	11,154,496	8,818	271
3hs7dg	3	7	-5,473,218	10,998,992	11,185,636	26,278	91
3hs8dg	3	8	-5,471,321	11,003,929	11,221,575	30,643	78
3hs2dg	3	2	-5,807,709	11,624,325	11,655,953	4,453	536
2hs5dg	2	5	-5,719,541	11,462,449	11,545,429	11,683	204
2hs4dg	2	4	-5,733,220	11,483,987	11,546,298	8,773	272
2hs6dg	2	6	-5,712,327	11,453,841	11,557,490	14,593	164
2hs7dg	2	7	-5,708,621	11,452,248	11,576,565	17,503	136
2hs3dg	2	3	-5,762,719	11,537,163	11,578,806	5,863	407
2hs8dg	2	8	-5,705,178	11,451,182	11,596,168	20,413	117
2hs2dg	2	2	-6,094,110	12,194,127	12,215,101	2,953	808

145 S9.2 Sensitivity: vine family set, truncation, GPD threshold

146 We assessed three vine family sets, three truncation levels ($\text{trunc} \in \{2, 3, 4\}$), and
147 three GPD thresholds ($u^* \in \{0.85, 0.90, 0.95\}$) for Zones 2 and 5.

148 *Summary.*

149 Family sets $\mathcal{F}_A = \{\text{Gaussian, Clayton, Gumbel, Frank}\}$ and $\mathcal{F}_C = \mathcal{F}_A \cup \{\text{Joe}\}$ per-
150 form near-identically; Gaussian-only ($\mathcal{F}_B$) degrades dependence preservation in wetter
151 states and southern zones. Threshold $u^* \in [0.90, 0.95]$ yields lower CRPS in heavier-
152 tail zones. Truncation level 2 is sufficient; higher truncation adds complexity with
153 negligible scoring improvement. We retain $\mathcal{F}_A$ with $u^* = 0.90$ and truncation 2 as the
154 main specification.

155 References

- 156 [1] Dawid, A.P.: Present position and potential developments: Some personal views:
157 Statistical theory: The prequential approach. *Journal of the Royal Statistical*
158 *Society. Series A* **147**(2), 278–292 (1984) <https://doi.org/10.2307/2981683>
- 159 [2] Gneiting, T., Raftery, A.E.: Strictly proper scoring rules, prediction, and estima-
160 tion. *Journal of the American Statistical Association* **102**(477), 359–378 (2007)
161 <https://doi.org/10.1198/016214506000001437>
- 162 [3] Berndt, D.J., Clifford, J.: Using dynamic time warping to find patterns in time
163 series. *KDD workshop* **10**(16), 359–370 (1994)
- 164 [4] Ward, J.H.: Hierarchical grouping to optimize an objective function. *Journal of*
165 *the American Statistical Association* **58**(301), 236–244 (1963) [https://doi.org/10.](https://doi.org/10.1080/01621459.1963.10500845)
166 [1080/01621459.1963.10500845](https://doi.org/10.1080/01621459.1963.10500845)
- 167 [5] Pyper, B.J., Peterman, R.M.: Comparison of methods to account for autocorrela-
168 tion in correlation analyses of fish data. *Canadian Journal of Fisheries and Aquatic*
169 *Sciences* **55**(9), 2127–2140 (1998) <https://doi.org/10.1139/f98-104>
- 170 [6] Politis, D.N., Romano, J.P.: The stationary bootstrap. *Journal of the Ameri-*
171 *can Statistical Association* **89**(428), 1303–1313 (1994) [https://doi.org/10.1080/](https://doi.org/10.1080/01621459.1994.10476870)
172 [01621459.1994.10476870](https://doi.org/10.1080/01621459.1994.10476870)
- 173 [7] Benjamini, Y., Hochberg, Y.: Controlling the false discovery rate: A practical and
174 powerful approach to multiple testing. *Journal of the Royal Statistical Society:*
175 *Series B* **57**(1), 289–300 (1995)

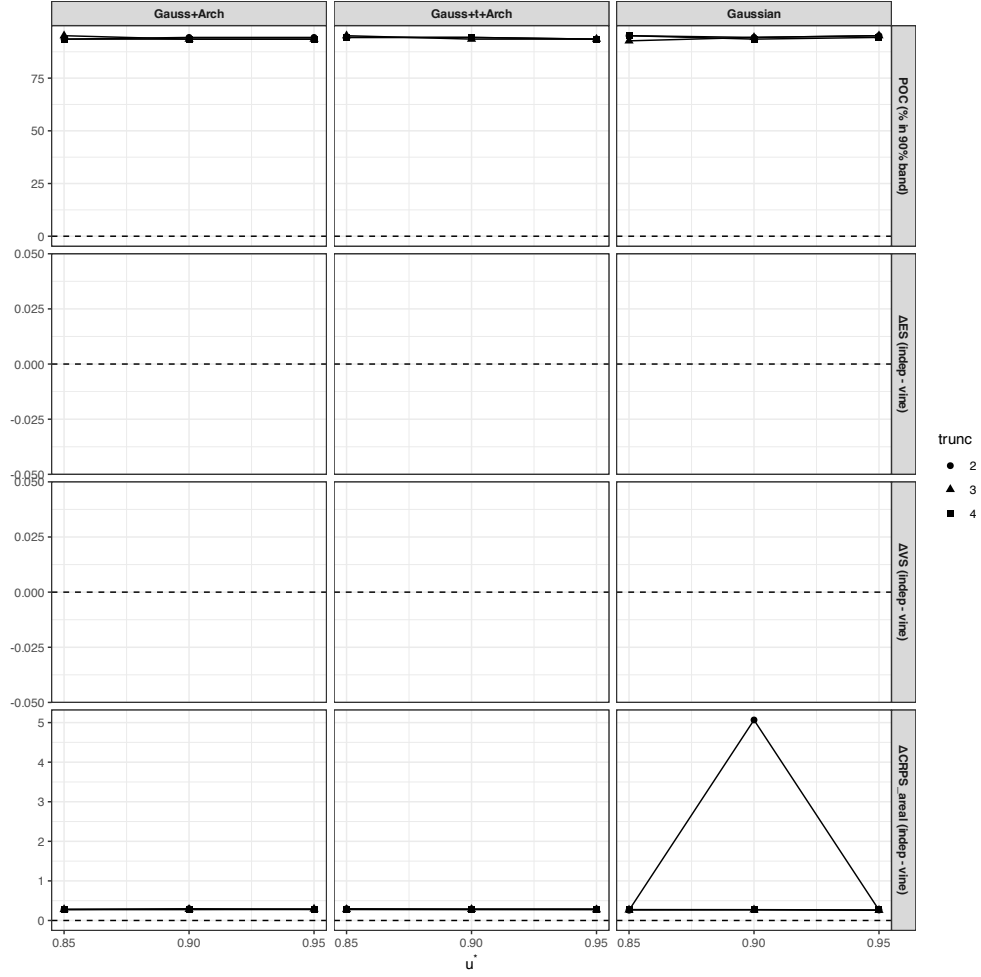


Fig. S6: Sensitivity of coverage (POC), multivariate scores (ΔES , ΔVS), and areal-CRPS gain to vine family set, truncation level, and GPD threshold (Zone 2 shown; Zone 5 shows qualitatively identical patterns).