Tail-Aware Density Forecasting of Locally Explosive Time Series: A Neural Network Approach

Online Appendix

Elena Dumitrescu
University Paris-Panthéon-Assas, CRED, 75006 Paris, France
and
Julien Peignon
CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine
PSL University, Paris, France
and
Arthur Thomas
Université Paris-Dauphine, Université PSL
LEDa, CNRS, IRD, 75016 Paris, France

February 10, 2026

This Online Appendix contains supplementary material for the main paper. Section 1 describes the alternative forecasting methods used in our comparative analysis. Section 2 gathers the auxiliary theoretical results underlying the benchmark computations. Section 3 provides additional tables and figures for the Monte Carlo simulations, in particular the results for longer forecast horizons ($h = 2$ and $h = 5$), while Section 4 provides additional tables for the empirical applications.

1 Alternative Conditional Density Forecasting Methods

To assess the performance of our proposed MDN approach, we compare it against a diverse set of established conditional density forecasting methods. These benchmarks span different methodological paradigms: the nonparametric Nadaraya-Watson kernel density estimator (Nadaraya 1964, Watson 1964), the simulation-based approach of Lanne et al. (2012) and the closed-form predictive densities of Gourieroux & Jasiak (2025), both designed for mixed causal-noncausal AR processes, and the learning-based FlexZBoost method of Dalmasso et al. (2020).

1.1 Nadaraya-Watson

The Nadaraya-Watson kernel density estimator (Rosenblatt 1969) serves as our main benchmark method. This classical nonparametric approach employs kernel smoothing to construct continuous density estimates without imposing parametric assumptions on the underlying distribution. It provides a baseline for conditional density estimation through direct estimation of joint and marginal probability densities.

From a time series $(X_t)_{t=1}^T$ and a forecast horizon h, the conditional density is obtained through the Bayes' formula:

$$\hat{p}_h(X_{t+h}|X_t) = \frac{\hat{p}_h(X_t, X_{t+h})}{\hat{p}(X_t)},$$

where the joint and marginal densities are:

$$\hat{p}_h(X_t, X_{t+h}) = \frac{1}{T-h} \sum_{i=1}^{T-h} K_{b_{\text{joint}}} \left((X_t, X_{t+h}) - (X_i, X_{i+h}) \right),$$
$$\hat{p}(X_t) = \frac{1}{T-1} \sum_{i=1}^{T-1} K_{b_{\text{marginal}}}(X_t - X_i).$$

We employ a gaussian kernel, with bandwidths $(b_{\text{joint}}, b_{\text{marginal}})$ selected according to Silverman's rule.

1.2 Lanne et al. (2012)

The method proposed by Lanne et al. (2012) relies on a simulation-based approach for computing point and density forecasts in mixed causal-noncausal autoregressive processes. Given an observed series $(X_t)_{t=1}^T$, it predicts X_{T+h} by approximating the noncausal component through Monte Carlo simulations.

The forecast of X_{T+h} conditional on the observed data is computed recursively:

$$E_T(X_{T+h}) = \phi_1 E_T(X_{T+h-1}) + \cdots + \phi_r E_T(X_{T+h-r}) + E_T(v_{T+h}),$$

where $E_T(\cdot) = E[\cdot \mid X_1, \ldots, X_T]$ denotes the conditional expectation given all observations up to time T, and r is the number of lag coefficients in the causal part of the model. The noncausal component $v_{T+h} = \varphi(B^{-1})^{-1}\epsilon_{T+h}$ is approximated by:

$$v_{T+h} \approx \sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j},$$

with β_j derived from the power series expansion of $\varphi(B^{-1})^{-1}$ and M chosen sufficiently large. This approach entails model risk due to the necessity of specifying the innovation density g, as well as estimation risk since the causal and noncausal parameters $\phi_1, \ldots, \phi_r$ and $\psi_1, \ldots, \psi_s$ (from which the coefficients β_j are derived), together with the parameters of g, must all be estimated from the observed data.

The point forecast is computed via Monte Carlo simulations:

$$E_T\left(\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}\right) \approx \frac{\sum_{i=1}^{N}\left(\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}^{(i)} \prod_{k=1}^{s} g(e_{T-s+k}^{(i)}; \lambda)\right)}{\sum_{i=1}^{N}\left(\prod_{k=1}^{s} g(e_{T-s+k}^{(i)}; \lambda)\right)},$$

where N sequences of simulated innovations, $\{\epsilon_{T+1}^{(i)}, \ldots, \epsilon_{T+M}^{(i)}\}_{i=1}^{N}$, are drawn from the error distribution. To derive the predictive density, they first obtain the empirical conditional cumulative distribution function (CDF) of X_{T+h}. For this, they replace the sum $\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j}$ with the indicator function $\mathbf{1}(\sum_{j=0}^{M-h} \beta_j \epsilon_{T+h+j} \leq x)$ in the numerator of the expectation (in Eq. 1.2) for a predefined equispaced grid of values $x_1, \ldots, x_K$. This grid spans the relevant range of the predicted variable, typically constructed from appropriate lower and upper quantiles of the empirical distribution, and serves as the discrete support on which the CDF is numerically evaluated. The predictive probability density function (PDF), $\hat{p}(X_{T+h}|X_1, \ldots, X_T)$, is then immediately obtained by numerical differentiation of the empirical CDF.

While Lanne et al. (2012) originally developed their forecasting algorithm using Student's t-distributed innovations, the methodology extends to α-stable distributions. The framework requires the innovation density g to be absolutely continuous with respect to the Lebesgue measure, with a density that is positive for all $x \in \mathbb{R}$ and twice continuously differentiable. The α-stable distributions satisfy these regularity conditions: for any $\alpha \in (0, 2]$, they admit a smooth Lebesgue density $g_\alpha(x) > 0$ for all $x \in \mathbb{R}$ (see Samorodnitsky & Taqqu 1996, Nolan 2020). Moreover, efficient algorithms exist both for evaluating the α-stable density numerically and for simulating α-stable random variables, making the Monte Carlo procedure in Eq. (1.2) directly applicable. In our simulation study, we implement this α-stable variant.

1.3 Gourieroux & Jasiak (2025)

Gourieroux & Jasiak (2025) develop closed-form predictive densities for mixed causal-noncausal Vector Autoregressive (VAR) processes. Their method relies on the spectral decomposition of the autoregressive companion matrix to identify latent causal and noncausal state components. While the original framework is designed for multivariate processes, we apply it here in a univariate setting.

For a purely noncausal AR(1) process, the one-step-ahead predictive density given X_t admits the closed-form expression:

$$p(X_{t+1}|X_t) = \frac{1}{|\psi_1|} \frac{l(X_{t+1})}{l(X_t)} \, g\left(X_{t+1} - \frac{1}{\psi_1} X_t\right),$$

where l denotes the stationary marginal density of the process and g is the density of $\eta_t = -\frac{1}{\psi_1}\varepsilon_t$. This methodology involves both model risk, as it requires specifying the density g, and estimation risk, as ψ_1 and the parameters of g must be estimated from the data. In practice, the stationary density l is approximated via kernel density estimation. In our implementation, we employ a Gaussian kernel with bandwidth selected according to Silverman's rule. For longer horizon forecasts, the method proceeds iteratively: first compute the predictive density $p(X_{t+1}|X_t)$ and extract its mode, then forecast X_{t+2} conditional on X_{t+1} set to this mode, and repeat until the desired horizon h is reached.

1.4 FlexZBoost

FlexZBoost (Dalmasso et al. 2020) is a learning-based approach to conditional density forecasting, which makes it a direct competitor for our MDN approach. The method uses the FlexCode framework (Izbicki 2017) by employing gradient boosting as the underlying regression mechanism.[1] Its theoretical foundation rests on the representation of the conditional density through an orthonormal basis expansion:

$$p_h(X_{t+h}|\mathbf{X}_t) = \sum_{j=1}^{B} \beta_j(\mathbf{X}_t)\phi_j(X_{t+h}),$$

where $\{\phi_j(X_{t+h})\}_{j=1}^B$ constitutes an orthonormal basis system, and $\{\beta_j(\mathbf{x}_t)\}_{j=1}^B$ represent feature-dependent expansion coefficients. In practice, we follow Dalmasso et al. (2020) and employ a cosine basis with 31 basis functions. This formulation transforms the density estimation problem into B independent regression tasks, wherein the coefficients are obtained by regressing the orthogonal projections $\phi_j(X_{t+h})$ onto the predictor space $\mathbf{X}_t$.

2 Auxiliary Theoretical Results

This section gathers the theoretical foundations underlying the benchmark computations used throughout our Monte Carlo analysis. First, we present the methodology of Fries (2022) for computing predictive conditional moments of MARMA processes driven by α-stable innovations, which serves as the theoretical ground truth for evaluating forecasting performance. Second, we recall the closed-form expression for the predictive density of the

[1]More specifically, the XGBoost algorithm (Chen & Guestrin 2016) is used.

noncausal Cauchy AR(1) process derived by Gourieroux & Zakoian (2017), which provides a unique setting where the full conditional distribution, not just its moments, can be compared against estimated densities.

2.1 Computing Predictive Conditional Moments

This subsection summarizes the methodology of Fries (2022) for computing predictive conditional moments of MARMA processes with α-stable errors. Consider the two-sided moving average representation $X_t = \sum_{k\in\mathbb{Z}} a_k \varepsilon_{t+k}$, where the coefficients satisfy $\sum_{k\in\mathbb{Z}} |a_k|^s < \infty$ for some $s \in (0,\alpha) \cap [0,1]$, ensuring strict stationarity (Rosenblatt 2000). A key insight from Fries (2022) is that noncausal processes can admit finite conditional moments up to order $\lfloor 2\alpha + 1 \rfloor$ despite having infinite marginal variance, provided that the process is sufficiently anticipative. For $\alpha \in (0,2) \setminus \{1\}$ and $p \in \{1,2,3,4\}$ with $p < 2\alpha + 1$, the conditional moments $\mathbb{E}[X^p_{t+h}|X_t = x]$ exist and take the form:

$$\mathbb{E}[X_{t+h}|X_t = x] = \kappa_1 x + \frac{a(\lambda_1 - \beta_1\kappa_1)}{1 + a^2\beta_1^2} g(x), \tag{1}$$

$$\mathbb{E}[X^p_{t+h}|X_t = x] = \kappa_p x^p + \frac{a x^{p-1}(\lambda_p - \beta_1\kappa_p)}{1 + a^2\beta_1^2} g(x) - R_p(x), \quad p \in \{2,3,4\}, \tag{2}$$

where $a = \tan(\pi\alpha/2)$, the function $g(x)$ involves the marginal density $f_{X_t}(x)$ and integrals of the form

$$\mathcal{H}(n,\boldsymbol{\theta};x) = \int_0^{+\infty} e^{-\sigma_1^\alpha u^\alpha} u^{n(\alpha-1)} \Big[\theta_1 \cos\Big(ux - a\beta_1\sigma_1^\alpha u^\alpha\Big) + \theta_2 \sin\Big(ux - a\beta_1\sigma_1^\alpha u^\alpha\Big)\Big]\, du, \tag{3}$$

and $R_p(x)$ denotes remainder terms involving $\mathcal{H}$ evaluated at orders $n = 2,\ldots,p$ with specific coefficient vectors $\boldsymbol{\theta}$ depending on $\alpha, \beta_1, \kappa_1, \ldots, \kappa_p, \lambda_1, \ldots, \lambda_p$.

The parameters appearing in (1)-(3) are defined as follows:

$$\sigma_1^\alpha = \sigma^\alpha \sum_{k\in\mathbb{Z}} |a_k|^\alpha, \qquad \beta_1 = \beta \frac{\sum_{k\in\mathbb{Z}} a_k^{\langle\alpha\rangle}}{\sum_{k\in\mathbb{Z}} |a_k|^\alpha},$$

$$\kappa_p = \frac{\sum_{k\in\mathbb{Z}} |a_k|^\alpha \left(\frac{a_{k-h}}{a_k}\right)^p}{\sum_{k\in\mathbb{Z}} |a_k|^\alpha}, \qquad \lambda_p = \beta \frac{\sum_{k\in\mathbb{Z}} a_k^{\langle\alpha\rangle} \left(\frac{a_{k-h}}{a_k}\right)^p}{\sum_{k\in\mathbb{Z}} |a_k|^\alpha},$$

where $y^{\langle\alpha\rangle} = \text{sign}(y)|y|^\alpha$ denotes the signed power function. The coefficients $(a_k)_{k\in\mathbb{Z}}$ of the MA(∞) representation can be recovered from the MARMA polynomial structure via partial fraction decomposition (see Fries 2022, Equation 3.7). For practical computation, these infinite sums are truncated, which yields accurate approximations since (a_k) decays geometrically for MARMA processes.

2.2 Closed-Form Predictive Density for the Noncausal Cauchy AR(1)

For the special case of a purely noncausal AR(1) process with Cauchy errors ($\alpha = 1$, $\beta = 0$), Gourieroux & Zakoian (2017) derive a closed-form expression for the causal predictive density. Consider the model

$$(1 - \psi_1 F)X_t = \varepsilon_t, \qquad \varepsilon_t \overset{\text{i.i.d.}}{\sim} \text{Cauchy}(0, \sigma),$$

where $|\psi_1| < 1$ ensures stationarity. The predictive density at horizon h is given by:

$$p(X_{t+h} \mid X_t) = \frac{1}{\pi\sigma_h} \cdot \frac{1}{1 + \left(\frac{X_t - \psi_1^h X_{t+h}}{\sigma_h}\right)^2} \cdot \frac{\sigma^2 + (1 - |\psi_1|)^2 X_t^2}{\sigma^2 + (1 - |\psi_1|)^2 X_{t+h}^2},$$

where the horizon-dependent scale parameter is

$$\sigma_h = \sigma \frac{1 - |\psi_1|^h}{1 - |\psi_1|}.$$

This closed-form expression enables direct comparison of estimated predictive densities against the true conditional distribution, providing a rigorous benchmark for density forecast evaluation that complements the moment-based comparisons available for general α-stable MARMA processes.

More generally, for any stable mixed-causal process (MAR) with a single anticipative root, the conditional distribution during extreme events follows a binomial law: the future trajectory can either continue its explosive path or collapse abruptly, with the probability of crashing at horizon h determined by the persistence parameter (see Gourieroux & Zakoian 2017, Fries 2022, de Truchis et al. 2025, Gourieroux et al. 2025). In the same vein, de Truchis et al. (2025) and Gourieroux et al. (2025) showed that the conditional distribution of the MAR model with at least two anticipative roots is 0,1-valued, *i.e.*, the conditional predictive density converges to a Dirac measure. This implies that, unlike the noncausal AR(1) where only crash probabilities can be predicted, the AR(2) specification allows for exact prediction of the bubble peak, the crash date, and the post-crash value. Properly accounting for these specificities of non-causal models is critical for ensuring accurate forecasts and strong predictive performance enspecially in the tails.

3 Additional Simulation results

In this Section, we include detailed results on the Model Confidence Set analysis summarized by an asterisk ($*$) in the tables of Section 3.4.2 in the main paper and then provide evidence of the robustness of our findings in Sections 3.3, 3.4.2 and 3.4.3 of the article to the choice of the forecasting horizon by setting $h = \{2, 5\}$.

3.1 Model Confidence Set

Table 1: Model Confidence Set Test Results: MAR(0, 1) Process, 1-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+1} \mid y_t]$			$\mathbb{E}[y_{t+1}^2 \mid y_t]$			$\mathbb{E}[y_{t+1}^3 \mid y_t]$			$\mathbb{E}[y_{t+1}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
	1.0	1.0000	0.0000	0.0000	1.0000	0.0002	0.0004	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	1.0000	0.0002	0.0002	–	–	–
Nadaraya-Watson	1.4	1.0000	1.0000	1.0000	1.0000	0.1312	0.1698	1.0000	0.0208	0.0228	–	–	–
	1.6	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	1.0000	0.0000	0.0000	1.0000	0.0000	0.0004
	1.8	1.0000	0.0792	0.4850	1.0000	0.0044	0.8030	1.0000	0.1902	1.0000	1.0000	0.2558	1.0000
	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0000	–	–	–
Lanne et al. (2012)	1.4	0.0000	0.0990	0.1102	0.0000	0.0000	0.0012	0.0000	0.0000	0.0002	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
	1.0	0.0000	0.0028	0.0018	0.0000	0.0330	0.0350	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0002	0.0002	–	–	–
Gourieroux and Jasiak (2025)	1.4	0.0000	0.6404	0.4834	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0000	0.0002	0.0018
	1.0	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
FlexZBoost	1.4	0.0000	0.6404	0.4834	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.0	0.0002	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0000	1.0000	1.0000	0.0004	1.0000	1.0000	0.0004	1.0000	1.0000	–	–	–
Mixture Density Network	1.4	0.0000	0.2064	0.1670	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0698	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0006	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.9898	0.0000	1.0000	0.3868

Notes: This Table reports MCS p-values from the Model Confidence Set procedure of Hansen et al. (2011) applied at the 90% confidence level ($\alpha = 0.10$). Values represent the probability threshold at which each model enters the set of superior forecasting methods. A model belongs to the $\text{MCS}_{90\%}$ if its p-value ≥ 0.10. Values in red indicate models included in the MCS (i.e., models with statistically indistinguishable superior performance). The loss function is the squared error of predictive moments relative to theoretical ones.

Table 2: Model Confidence Set Test Results: MAR(0, 2) Process, 1-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+1} \mid y_t]$			$\mathbb{E}[y_{t+1}^2 \mid y_t]$			$\mathbb{E}[y_{t+1}^3 \mid y_t]$			$\mathbb{E}[y_{t+1}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0234	0.0296	–	–	–
Nadaraya-Watson	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0006	0.0008	–	–	–
	1.6	1.0000	0.0000	0.0008	1.0000	0.0062	0.0502	1.0000	0.1858	0.3776	1.0000	1.0000	1.0000
	1.8	1.0000	0.1758	0.5138	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1.0	0.0000	0.0006	0.0018	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
FlexZBoost	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0008	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0020	0.0012	0.0000	0.0032	0.0020
	1.8	0.0000	0.0030	0.0000	0.0000	0.0962	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.0	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.1454	1.0000	1.0000	0.0000	1.0000	1.0000	0.0030	1.0000	1.0000	–	–	–
Mixture Density Network	1.4	0.4572	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	0.2448	0.1272
	1.8	0.0016	1.0000	1.0000	0.0000	0.4470	0.0092	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: For details on variable definitions and methodology, refer to Table 1.

Table 3: Model Confidence Set Test Results: MAR(1, 1) Process, 1-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+1} \mid y_t]$			$\mathbb{E}[y_{t+1}^2 \mid y_t]$			$\mathbb{E}[y_{t+1}^3 \mid y_t]$			$\mathbb{E}[y_{t+1}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
Nadaraya-Watson	1.4	1.0000	0.0004	0.0006	1.0000	0.0000	0.0000	1.0000	0.0002	0.0000	–	–	–
	1.6	1.0000	0.0000	0.0006	1.0000	0.0000	0.0002	1.0000	0.0000	0.0000	1.0000	0.0000	0.0006
	1.8	1.0000	0.7840	1.0000	1.0000	0.5364	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
FlexZBoost	1.4	0.0000	0.0004	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.0	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
Mixture Density Network	1.4	0.3582	1.0000	1.0000	0.0002	1.0000	1.0000	0.0446	1.0000	1.0000	–	–	–
	1.6	0.1802	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	1.0000	0.6182	0.0000	1.0000	0.5702	0.0000	0.6822	0.3766	0.0000	0.0734	0.0278

Notes: For details on variable definitions and methodology, refer to Table 1.

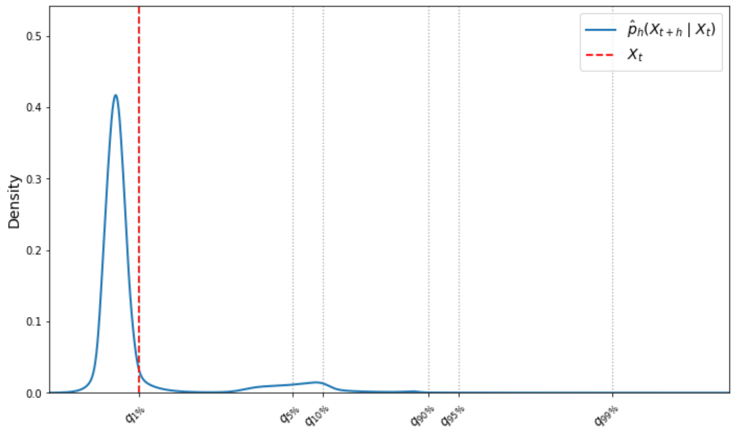
Table 4: Model Confidence Set Test Results: MARMA(1, 1, 1, 1) Process, 1-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+1} \mid y_t]$			$\mathbb{E}[y_{t+1}^2 \mid y_t]$			$\mathbb{E}[y_{t+1}^3 \mid y_t]$			$\mathbb{E}[y_{t+1}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
Nadaraya-Watson	1.4	1.0000	0.0002	0.0004	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	–	–	–
	1.6	1.0000	0.1940	0.3650	1.0000	0.0000	0.0004	1.0000	0.0980	0.1874	1.0000	0.0006	0.0024
	1.8	1.0000	0.0092	0.1660	1.0000	0.0162	0.8638	1.0000	0.2392	1.0000	1.0000	1.0000	1.0000
	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
FlexZBoost	1.4	0.0000	0.0002	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0000	0.0002	0.0002	0.0000	0.0006	0.0024
	1.8	0.0000	0.0000	0.0000	0.0000	0.0050	0.0158	0.0000	0.0256	0.0280	0.0000	0.0388	0.0340
	1.0	0.7244	1.0000	1.0000	0.0284	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0010	1.0000	1.0000	0.0350	1.0000	1.0000	0.0064	1.0000	1.0000	–	–	–
Mixture Density Network	1.4	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0036	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.8070	0.0000	0.3872	0.0340

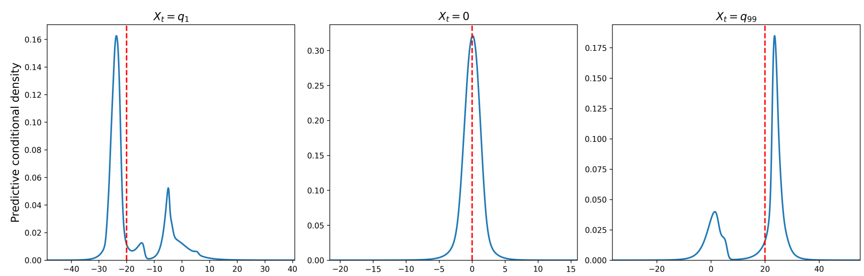
Notes: For details on variable definitions and methodology, refer to Table 1.

3.2 Bimodality Analysis and Sampled Trajectories (additional results to Section 3.3)

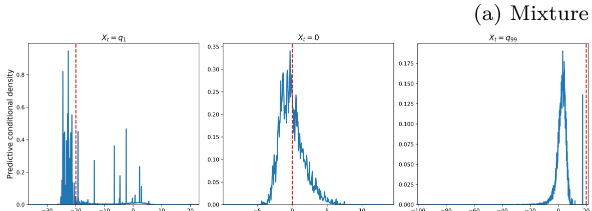
Figure 1: Animated GIF, 1-Step-Ahead Conditional Predictive Density of a MAR(0,1) Process Computed with Mixture Density Network



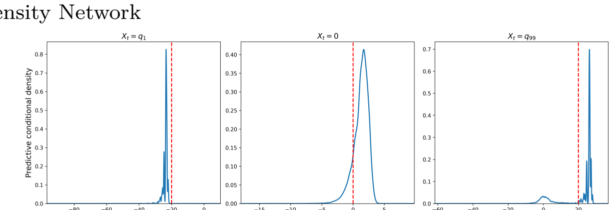
Notes: This animated figure requires a PDF reader with animation support and may not display correctly in all PDF viewers. The figure displays the estimated conditional predictive density $\hat{p}(X_{t+h}|X_t)$ for a purely noncausal MAR(0,1) process with α-stable innovations ($\alpha = 1.4$) across the full range of conditioning values X_t in the grid. The red dashed line indicates the current conditioning value X_t



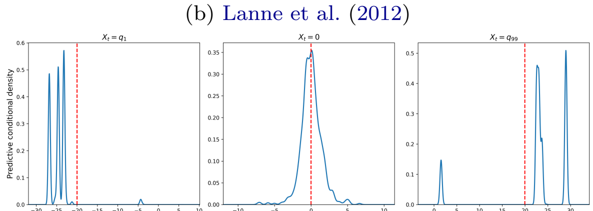
(a) Mixture Density Network



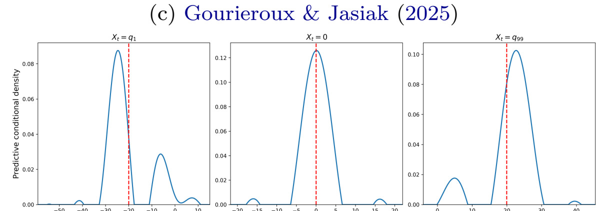
(b) Lanne et al. (2012)



(c) Gourieroux & Jasiak (2025)



(d) Nadaraya-Watson

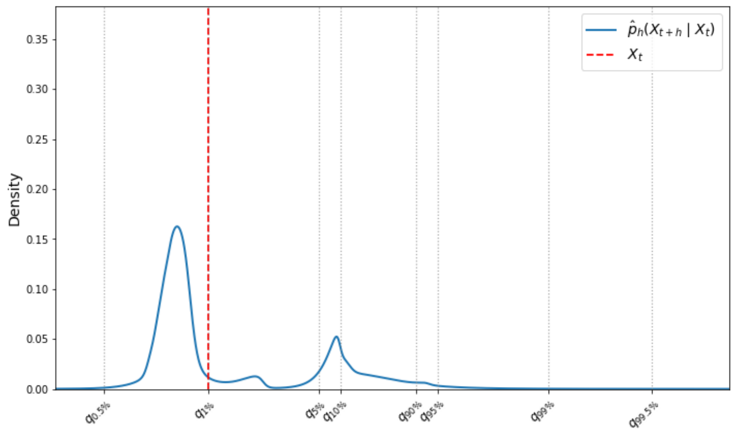


(e) FlexZBoost

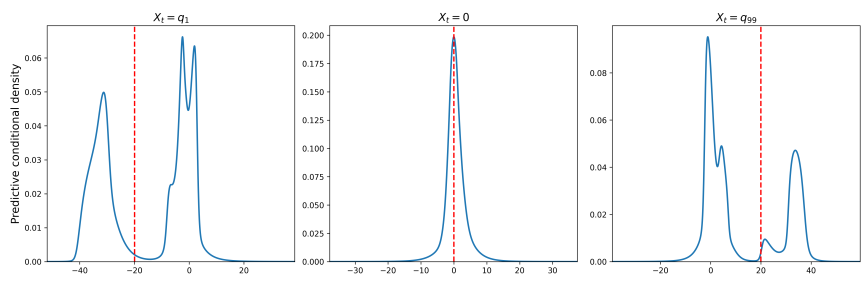
Figure 2: 2-Step-Ahead Conditional Predictive Density of a MAR(0,1) Process

Notes: This figure displays the estimated conditional predictive density $\hat{p}(X_{t+h}|X_t)$ for a purely noncausal MAR(0,1) process with α-stable innovations at three conditioning values: $X_t = q_1$ (left tail), $X_t = 0$ (center), and $X_t = q_{99}$ (right tail), where $q_.$ denotes the percentiles. The red dashed line indicates the conditioning value X_t. The tail-index is fixed at $\alpha = 1.4$.

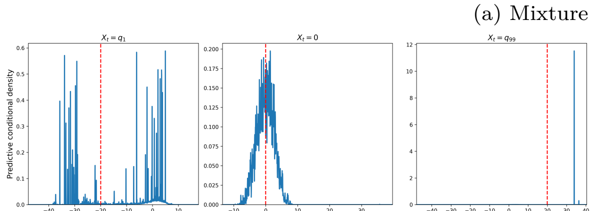
Figure 3: Animated GIF, 2-Step-Ahead Conditional Predictive Density of a MAR(0,1) Process Computed with Mixture Density Network



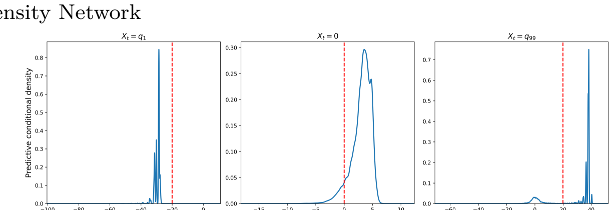
Notes: For details on variable definitions and methodology, refer to Figure 1.



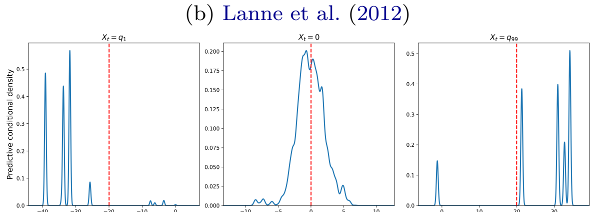
(a) Mixture Density Network



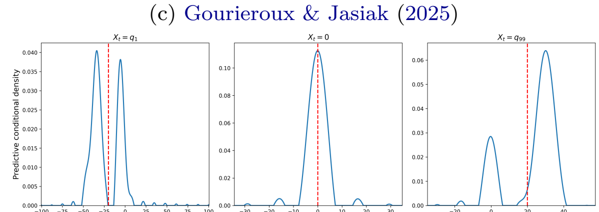
(b) Lanne et al. (2012)



(c) Gourieroux & Jasiak (2025)



(d) Nadaraya-Watson

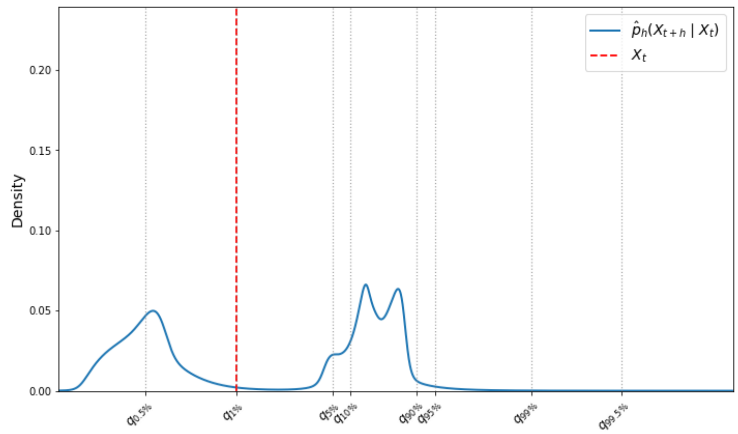


(e) FlexZBoost

Figure 4: 5-Step-Ahead Conditional Predictive Density of a MAR(0,1) Process

Notes: For details on variable definitions and methodology, refer to Figure 2.

Figure 5: Animated GIF, 5-Step-Ahead Conditional Predictive Density of a MAR(0,1) Process Computed with Mixture Density Network



Notes: For details on variable definitions and methodology, refer to Figure 1.

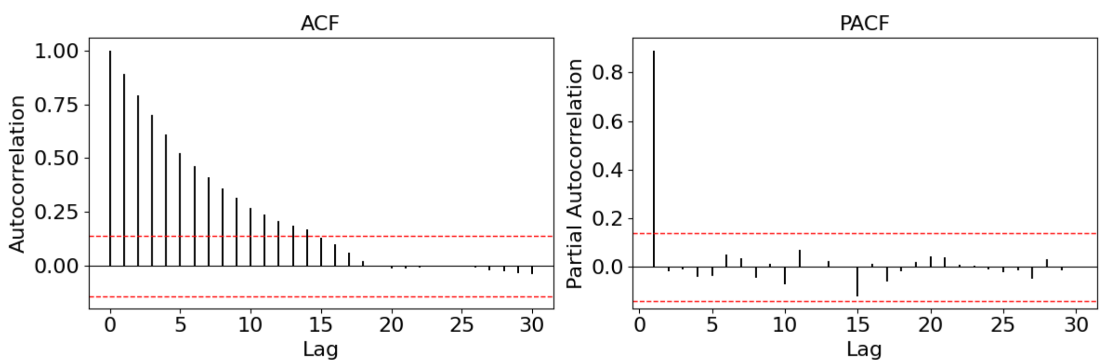


Figure 6: ACF - PACF of the MDN-sampled MAR(0,1) trajectory

Table 5: Estimated MAR(0,1) parameters from the MDN-sampled trajectory

Parameter	Estimate
ψ	0.887*** (0.033)
α	1.424*** (0.188)
β	0.297 (0.213)
σ	0.495*** (0.049)

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

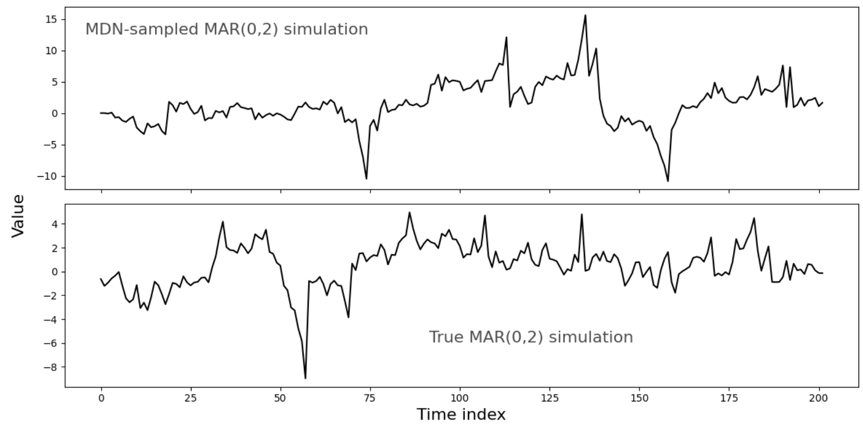


Figure 7: MDN-Sampled vs. True MAR(0,2) Trajectories
Notes: Top panel: trajectory generated by iteratively sampling from the MDN's one-step-ahead predictive density. Bottom panel: true MAR(0,2) simulation with $\alpha = 1.4$. The underlying MDN has been trained on a simulated sample of 5,000 realizations from the true MAR(0,2) specification with $\alpha = 1.4$.

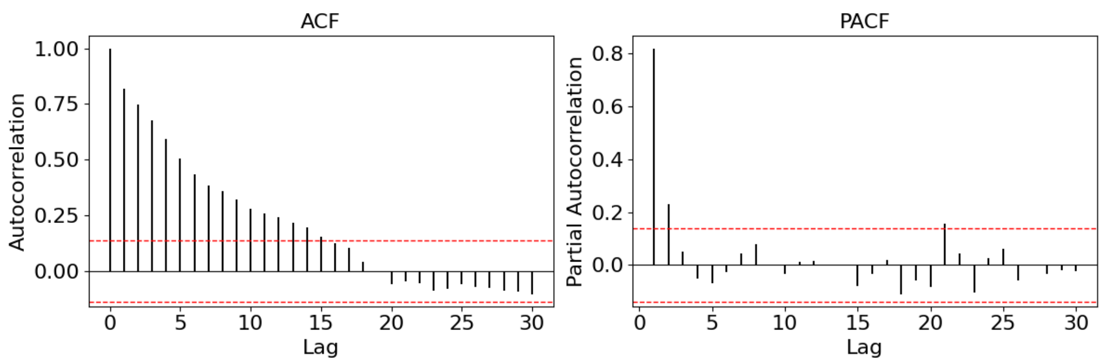


Figure 8: ACF - PACF of the MDN-sampled MAR(0,2) trajectory

Table 6: Estimated MAR(0,2) parameters from the MDN-sampled trajectory

Parameter	Estimate
ψ_1	0.665^{***} (0.067)
ψ_2	0.116^{*} (0.070)
α	1.475^{***} (0.191)
β	0.128 (0.202)
σ	0.651^{***} (0.065)

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

3.3 Predictive Moments Approach (additional results to Section 3.4.2)

Table 7: Root Mean Squared Error of Predictive Moments: MAR(0, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**0.649**	47.73	45.37	**20.67**	8081	7680	–	–	–	–	–	–
	1.2	**0.206**	11.33	10.42	**2.559**	594.3	546.7	**24.12**	3.255e+04	2.994e+04	–	–	–
	1.4	**0.115**	3.587	3.160	**0.703**	87.25	76.86	**3.973**	2256	1988	–	–	–
	1.6	**0.074**	1.238	1.025	**0.299**	18.27	15.11	**0.998**	246.7	204.0	**13.62**	3179	2629
	1.8	**0.059**	0.249	0.191	**0.114**	**1.824**	**1.368**	**0.334**	**13.49**	**10.10**	**2.438**	**98.94**	**74.08**
Lanne et al. (2012)	1.0	2.219	42.41	40.31	308.1	8418	8001	–	–	–	–	–	–
	1.2	0.423	13.03	11.99	11.13	720.1	662.6	**77.19**	3.853e+04	3.545e+04	–	–	–
	1.4	0.226	4.655	4.102	2.379	111.4	98.18	14.24	2546	2242	–	–	–
	1.6	**0.102**	1.153	0.956	0.979	17.65	14.61	**4.259**	235.9	195.1	**28.73**	3159	2612
	1.8	**0.051**	**0.152**	**0.119**	**0.441**	2.660	2.012	**1.652**	25.25	18.93	**7.370**	211.8	158.6
Gourieroux and Jasiak (2025)	1.0	9.917	26.59	25.46	248.3	**4203**	**3995**	–	–	–	–	–	–
	1.2	2.157	**5.156**	**4.818**	22.64	**223.9**	**206.2**	241.6	**1.175e+04**	**1.081e+04**	–	–	–
	1.4	0.917	**2.199**	**1.985**	5.341	**40.36**	**35.65**	31.38	**865.1**	**762.2**	–	–	–
	1.6	0.617	**1.046**	**0.932**	2.387	**10.94**	**9.151**	9.222	**130.8**	**108.3**	42.00	**1546**	**1278**
	1.8	0.375	0.357	0.365	1.169	2.461	1.999	3.577	18.69	14.19	12.21	**147.8**	**110.9**
FlexZBoost	1.0	2.206	**25.81**	**24.54**	2475	9855	9398	–	–	–	–	–	–
	1.2	0.792	6.277	5.783	96.96	597.5	551.0	347.7	5.265e+04	4.844e+04	–	–	–
	1.4	0.366	3.232	2.853	9.567	95.95	84.64	83.94	3049	2686	–	–	–
	1.6	0.271	1.376	1.148	2.220	24.27	20.11	29.45	407.5	337.5	779.3	7719	6399
	1.8	0.260	0.358	0.318	1.464	3.312	2.662	11.07	24.64	19.85	150.8	281.3	233.1
Mixture Density Network	1.0	**1.437**	**11.20**	**10.66**	**49.95**	**1199**	**1140**	–	–	–	–	–	–
	1.2	**0.358**	**3.424**	**3.153**	**5.574**	**155.1**	**142.7**	196.1	**1.007e+04**	**9263**	–	–	–
	1.4	**0.177**	**0.462**	**0.415**	**1.585**	**12.47**	**11.01**	**13.50**	**374.5**	**329.9**	–	–	–
	1.6	0.124	**0.634**	**0.529**	**0.973**	**6.715**	**5.581**	5.234	**83.05**	**68.75**	89.26	**1014**	**840.3**
	1.8	0.135	**0.217**	**0.186**	0.944	**1.707**	**1.423**	4.459	**14.88**	**11.52**	48.13	149.2	116.2

Notes: This Table reports the root mean squared error (RMSE) of estimated predictive moments relative to theoretical values. α denotes the tail index of the stable distribution. Predictive moments are evaluated over three spatial regions: Center $[q_{0.1}, q_{0.9}]$, Tails $[q_{0.01}, q_{0.1}] \cup [q_{0.9}, q_{0.99}]$, and Total $[q_{0.01}, q_{0.99}]$, where q_p represents the p-th quantile. Best method in red, second best in **bold black**.

Table 8: Model Confidence Set Test Results: MAR(0, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0002	0.0002	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.6	1.0000	0.0000	0.0064	1.0000	0.0000	0.0000	1.0000	0.0004	0.0002	1.0000	0.0014	0.0050
	1.8	0.2548	0.0032	0.0116	1.0000	0.3694	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Lanne et al. (2012)	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0040	0.0000	0.0000	–	–	–
	1.6	0.0028	0.0884	0.1180	0.0000	0.0000	0.0000	0.0000	0.0004	0.0002	0.0714	0.0000	0.0004
	1.8	1.0000	1.0000	1.0000	0.0000	0.0238	0.0090	0.0002	0.0004	0.0068	0.0004	0.0000	0.0002
Gourieroux and Jasiak (2025)	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0028	0.0030	0.0000	1.0000	1.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0002	0.0000	0.0028	0.0092
	1.8	0.0000	0.0000	0.0000	0.0000	0.0030	0.0000	0.0000	0.0004	0.0000	0.0000	0.0000	0.0002
FlexZBoost	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mixture Density Network	1.0	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0016	1.0000	1.0000	0.0068	1.0000	1.0000	0.0000	0.5836	0.6088	–	–	–
	1.4	0.0042	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0210	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	0.7304	0.0226	0.0000	1.0000	0.1066	0.0000	0.4720	0.0358	0.0000	0.0000	0.0000

Notes: This Table reports MCS p-values from the Model Confidence Set procedure of Hansen et al. (2011) applied at the 90% confidence level ($\alpha = 0.10$). Values represent the probability threshold at which each model enters the set of superior forecasting methods. A model belongs to the $\text{MCS}_{90\%}$ if its p-value ≥ 0.10. Values in red indicate models included in the MCS (i.e., models with statistically indistinguishable superior performance). The loss function is the squared error of predictive moments relative to theoretical ones. See Table 7 for corresponding RMSE values.

Table 9: Root Mean Squared Error of Predictive Moments: MAR(0, 2) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**0.327**	29.60	28.13	**3.139**	3044	2893	–	–	–	–	–	–
	1.2	**0.186**	7.053	6.489	**1.178**	**287.5**	**264.5**	**14.32**	**1.371e+04**	**1.261e+04**	–	–	–
	1.4	**0.091**	1.944	1.713	**0.347**	**36.00**	**31.71**	**1.623**	**742.3**	**653.9**	–	–	–
	1.6	**0.053**	**0.489**	**0.405**	**0.095**	4.481	**3.706**	**0.429**	**44.21**	**36.57**	**4.731**	**561.2**	**464.2**
	1.8	**0.040**	**0.107**	**0.084**	**0.102**	**0.467**	**0.356**	**0.168**	**2.635**	**1.976**	**0.703**	**18.60**	**13.93**
FlexZBoost	1.0	1.277	**14.24**	**13.54**	614.5	**2716**	**2589**	–	–	–	–	–	–
	1.2	0.573	**5.459**	**5.027**	29.36	334.3	307.8	73.93	2.430e+04	2.235e+04	–	–	–
	1.4	0.363	**1.916**	**1.697**	3.402	40.25	35.49	23.45	1057	931.1	–	–	–
	1.6	0.258	**0.465**	**0.411**	1.878	**4.446**	3.826	20.30	72.63	61.14	396.6	1506	1265
	1.8	0.205	0.368	0.307	0.874	2.314	1.826	4.440	17.90	13.72	42.63	153.7	118.4
Mixture Density Network	1.0	**0.737**	**13.00**	**12.35**	**35.31**	**1151**	**1094**	–	–	–	–	–	–
	1.2	**0.177**	**3.082**	**2.836**	**3.083**	**101.4**	**93.33**	**56.72**	**3916**	**3602**	–	–	–
	1.4	**0.114**	**1.016**	**0.897**	**1.224**	**17.09**	**15.06**	**9.612**	**288.7**	**254.4**	–	–	–
	1.6	**0.097**	0.591	0.492	**0.827**	**3.300**	**2.768**	**5.085**	**63.18**	**52.33**	**56.81**	**411.5**	**341.8**
	1.8	**0.076**	**0.197**	**0.156**	**0.649**	**0.642**	**0.644**	**3.233**	**5.399**	**4.570**	**25.78**	**48.24**	**39.92**

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 10: Model Confidence Set Test Results: MAR(0, 2) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.7282	0.0000	0.0000	1.0000	0.0002	0.0002	1.0000	0.0016	0.0016	–	–	–
	1.4	1.0000	0.0026	0.0030	1.0000	0.0004	0.0028	1.0000	0.0004	0.0008	–	–	–
	1.6	1.0000	0.9786	1.0000	1.0000	0.0062	0.0800	1.0000	1.0000	1.0000	1.0000	0.0748	0.1892
	1.8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
FlexZBoost	1.0	0.0000	1.0000	1.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0002	–	–	–
	1.6	0.0000	1.0000	0.3232	0.0000	0.0062	0.0010	0.0000	0.0482	0.0162	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mixture Density Network	1.0	0.0000	0.5762	0.6270	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	1.0000	1.0000	1.0000	0.0028	1.0000	1.0000	0.0312	1.0000	1.0000	–	–	–
	1.4	0.1258	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0004	0.7866	0.4288	0.0000	1.0000	1.0000	0.0000	0.1162	0.1022	0.0000	1.0000	1.0000
	1.8	0.0680	0.0000	0.0000	0.0000	0.0132	0.0000	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 9 for corresponding RMSE values.

Table 11: Root Mean Squared Error of Predictive Moments: MAR(1, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	0.842	54.89	52.16	29.23	1.046e+04	9942	–	–	–	–	–	–
	1.2	0.232	11.76	10.82	3.018	705.6	649.2	35.65	4.241e+04	3.902e+04	–	–	–
	1.4	0.127	3.873	3.413	0.818	106.5	93.83	5.471	3018	2658	–	–	–
	1.6	0.080	1.285	1.063	0.369	20.73	17.15	1.495	316.3	261.6	17.96	4631	3830
	1.8	0.064	0.279	0.213	0.130	2.515	1.885	0.393	21.57	16.15	2.794	179.3	134.3
FlexZBoost	1.0	2.424	28.40	27.00	3058	1.234e+04	1.177e+04	–	–	–	–	–	–
	1.2	0.866	6.879	6.338	119.5	757.2	698.2	516.2	7.308e+04	6.724e+04	–	–	–
	1.4	0.465	3.189	2.818	12.07	109.0	96.21	120.0	3802	3349	–	–	–
	1.6	0.289	1.461	1.219	2.279	27.65	22.90	21.76	495.4	409.9	526.3	9910	8201
	1.8	0.253	0.533	0.433	1.740	5.350	4.168	15.45	52.22	40.42	239.6	584.5	465.5
Mixture Density Network	1.0	1.583	7.698	7.333	62.87	878.5	835.1	–	–	–	–	–	–
	1.2	0.268	2.554	2.352	4.874	153.0	140.7	240.3	1.099e+04	1.011e+04	–	–	–
	1.4	0.195	0.737	0.656	1.525	18.40	16.23	15.13	572.0	503.9	–	–	–
	1.6	0.135	0.444	0.375	1.108	4.044	3.402	7.931	46.85	39.00	115.5	533.1	445.6
	1.8	0.098	0.294	0.230	0.837	2.939	2.269	4.058	30.49	22.98	48.13	343.2	258.9

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 12: Model Confidence Set Test Results: MAR(1, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.6	1.0000	0.0004	0.0028	1.0000	0.0000	0.0008	1.0000	0.0002	0.0010	1.0000	0.0006	0.0024
	1.8	1.0000	0.9888	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
FlexZBoost	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mixture Density Network	1.0	0.0002	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.1592	1.0000	1.0000	0.0366	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.0038	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0234	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0044	1.0000	0.5752	0.0000	0.3582	0.0210	0.0000	0.0100	0.0044	0.0000	0.0000	0.0002

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 11 for corresponding RMSE values.

Table 13: Root Mean Squared Error of Predictive Moments: MARMA(1, 1, 1, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**0.590**	32.00	30.41	**18.81**	**4448**	**4227**	–	–	–	–	–	–
	1.2	**0.171**	8.262	7.601	**1.489**	**260.5**	**239.7**	**12.81**	**1.340e+04**	**1.233e+04**	–	–	–
	1.4	**0.117**	3.848	3.390	**0.513**	68.64	60.47	**4.802**	**1735**	**1528**	–	–	–
	1.6	**0.067**	**1.248**	**1.033**	**0.187**	**8.975**	**7.423**	**1.119**	**98.39**	**81.37**	**10.84**	**1125**	**930.2**
	1.8	**0.051**	**0.526**	**0.395**	**0.128**	**2.708**	**2.029**	**0.371**	**14.91**	**11.17**	**1.526**	**89.14**	**66.74**
FlexZBoost	1.0	1.721	**21.61**	**20.55**	1767	5873	5609	–	–	–	–	–	–
	1.2	0.639	**4.934**	**4.546**	66.19	469.0	432.3	198.3	5.234e+04	4.816e+04	–	–	–
	1.4	0.362	**2.576**	**2.276**	6.652	**61.88**	**54.60**	64.18	2657	2341	–	–	–
	1.6	0.238	1.445	1.203	1.695	19.99	16.56	18.94	283.2	234.5	380.4	8777	7262
	1.8	0.219	0.662	0.516	0.864	4.351	3.307	5.816	44.27	33.37	63.45	439.7	331.9
Mixture Density Network	1.0	**0.877**	**5.987**	**5.697**	**40.34**	**632.3**	**601.1**	–	–	–	–	–	–
	1.2	**0.337**	**1.162**	**1.078**	**1.961**	**38.89**	**35.79**	**47.69**	**2831**	**2604**	–	–	–
	1.4	**0.244**	**0.570**	**0.515**	**1.817**	**7.700**	**6.837**	**19.49**	**229.6**	**202.5**	–	–	–
	1.6	**0.139**	**0.363**	**0.310**	**1.119**	**3.612**	**3.053**	**8.019**	**54.81**	**45.56**	**108.5**	**753.8**	**626.4**
	1.8	**0.106**	**0.187**	**0.156**	**0.801**	**1.761**	**1.422**	**3.911**	**15.69**	**12.03**	**35.38**	**146.6**	**112.2**

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 14: Model Confidence Set Test Results: MARMA(1, 1, 1, 1) Process, 2-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+2} \mid y_t]$			$\mathbb{E}[y_{t+2}^2 \mid y_t]$			$\mathbb{E}[y_{t+2}^3 \mid y_t]$			$\mathbb{E}[y_{t+2}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	1.0000	0.0000	0.0002	–	–	–
	1.6	1.0000	0.0000	0.0000	1.0000	0.0006	0.0094	1.0000	0.0212	0.0716	1.0000	0.5472	0.8434
	1.8	1.0000	0.0000	0.0086	1.0000	0.1394	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
FlexZBoost	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0000	0.0036	0.0058	0.0000	0.0106	0.0110
	1.8	0.0000	0.0000	0.0000	0.0000	0.1394	0.1298	0.0000	0.0522	0.0200	0.0000	0.0000	0.0000
Mixture Density Network	1.0	0.0036	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0002	1.0000	1.0000	0.0280	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	1.0000	1.0000	0.0000	1.0000	0.9328	0.0000	0.2890	0.0422	0.0000	0.0000	0.0000

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 13 for corresponding RMSE values.

Table 15: Root Mean Squared Error of Predictive Moments: MAR(0, 1) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**1.141**	71.94	68.37	**49.85**	1.563e+04	1.485e+04	–	–	–	–	–	–
	1.2	**0.352**	19.06	17.53	**4.398**	1293	1190	**61.79**	9.265e+04	8.524e+04	–	–	–
	1.4	**0.233**	6.043	5.324	**1.406**	183.7	161.9	**14.86**	5771	5084	–	–	–
	1.6	**0.159**	2.317	1.919	**0.675**	37.50	31.02	**4.083**	609.7	504.3	**69.22**	9728	8045
	1.8	**0.079**	**0.427**	**0.324**	**0.264**	**4.348**	**3.259**	**1.148**	57.23	42.85	**9.333**	738.6	553.0
Lanne et al. (2012)	1.0	4.204	51.83	49.28	926.6	1.329e+04	1.263e+04	–	–	–	–	–	–
	1.2	**0.506**	13.91	12.80	32.11	**1009**	**928.6**	**357.4**	**7.355e+04**	**6.767e+04**	–	–	–
	1.4	**0.248**	4.965	4.376	5.521	**152.4**	**134.3**	**42.34**	**4715**	**4154**	–	–	–
	1.6	**0.067**	**1.330**	**1.100**	**2.390**	**25.62**	**21.24**	**13.23**	**437.9**	**362.2**	**111.3**	**8117**	**6713**
	1.8	**0.033**	**0.120**	**0.092**	**0.932**	4.643	3.530	**4.163**	**49.35**	**37.05**	**26.60**	**501.6**	**375.9**
Gourieroux and Jasiak (2025)	1.0	30.28	36.06	35.54	1419	**5616**	**5356**	–	–	–	–	–	–
	1.2	7.755	15.94	14.97	135.9	1101	1014	2743	8.688e+04	7.993e+04	–	–	–
	1.4	3.058	8.019	7.211	22.23	243.2	214.5	215.0	8017	7063	–	–	–
	1.6	1.919	3.839	3.353	7.658	66.57	55.22	42.70	1204	995.8	237.5	2.156e+04	1.783e+04
	1.8	1.065	1.114	1.093	2.649	10.32	7.919	8.815	103.1	77.40	31.62	1035	775.0
FlexZBoost	1.0	2.427	**35.47**	**33.72**	2454	1.371e+04	1.305e+04	–	–	–	–	–	–
	1.2	0.969	**11.41**	**10.50**	93.06	1064	979.7	475.9	1.191e+05	1.096e+05	–	–	–
	1.4	0.642	**4.802**	**4.241**	8.841	177.3	156.2	124.5	6302	5552	–	–	–
	1.6	0.481	1.997	1.674	4.565	43.30	35.90	84.33	811.4	672.8	2642	1.985e+04	1.648e+04
	1.8	0.392	0.594	0.515	2.384	4.904	3.997	17.01	52.43	40.84	255.8	605.7	484.1
Mixture Density Network	1.0	**2.091**	**23.33**	**22.18**	**92.59**	**3325**	**3160**	–	–	–	–	–	–
	1.2	0.512	**5.074**	**4.672**	**10.97**	**277.5**	**255.3**	627.1	**2.122e+04**	**1.953e+04**	–	–	–
	1.4	0.287	**1.990**	**1.759**	**3.817**	**52.67**	**46.43**	47.33	**1586**	**1397**	–	–	–
	1.6	0.217	**0.995**	**0.832**	2.394	**14.65**	**12.19**	19.03	**256.2**	**212.1**	376.6	**4489**	**3719**
	1.8	0.213	0.427	0.350	1.793	**3.723**	**3.030**	11.61	**40.76**	**31.47**	145.7	**489.0**	**378.6**

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 16: Model Confidence Set Test Results: MAR(0, 1) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	–	–	–
	1.6	0.0000	0.0050	0.0084	1.0000	0.0006	0.0076	1.0000	0.0048	0.0162	1.0000	0.0084	0.0264
	1.8	0.0000	0.0000	0.0000	1.0000	0.5260	1.0000	1.0000	0.2638	1.0000	1.0000	0.2300	1.0000
Lanne et al. (2012)	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	–	–	–
	1.4	0.4140	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	–	–	–
	1.6	1.0000	1.0000	1.0000	0.0000	0.0042	0.0118	0.0000	0.0332	0.0686	0.0000	0.0084	0.0264
	1.8	1.0000	1.0000	1.0000	0.0000	0.0006	0.0008	0.0000	0.0314	0.2966	0.0000	0.2396	0.9590
Gourieroux and Jasiak (2025)	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.0006	0.0008	0.0000	0.0314	0.0936	0.0000	0.1122	0.1786
FlexZBoost	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0000	0.0000	0.0000	0.1074	0.0008	0.0000	0.0666	0.0936	0.0000	0.1122	0.0086
Mixture Density Network	1.0	0.0000	1.0000	1.0000	0.0076	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0090	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.4140	1.0000	1.0000	0.0000	1.0000	1.0000	0.0002	1.0000	1.0000	–	–	–
	1.6	0.0000	0.3778	0.2262	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	0.0000	0.0000	0.0000	1.0000	0.0198	0.0000	1.0000	0.6834	0.0000	1.0000	0.9590

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 15 for corresponding RMSE values.

Table 17: Root Mean Squared Error of Predictive Moments: MAR(0, 2) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**0.646**	60.19	57.21	**20.01**	1.135e+04	1.079e+04	–	–	–	–	–	–
	1.2	**0.311**	12.69	11.67	**3.962**	980.5	902.1	**83.75**	8.826e+04	8.120e+04	–	–	–
	1.4	**0.133**	**2.975**	**2.621**	**0.486**	**86.79**	**76.46**	**4.042**	**3268**	**2879**	–	–	–
	1.6	**0.073**	1.020	0.845	**0.290**	12.51	10.35	**1.451**	192.7	159.3	**45.88**	**3476**	**2875**
	1.8	**0.053**	**0.207**	**0.159**	**0.217**	**1.714**	**1.291**	**1.738**	**15.79**	**11.88**	**14.16**	**199.5**	**149.6**
FlexZBoost	1.0	1.672	**19.84**	**18.87**	598.1	**4347**	**4136**	–	–	–	–	–	–
	1.2	0.863	**6.588**	**6.070**	28.40	**561.1**	**516.4**	**297.8**	**5.433e+04**	**4.998e+04**	–	–	–
	1.4	0.497	3.192	2.822	6.064	112.8	99.44	123.2	4601	4053	–	–	–
	1.6	0.382	**0.757**	**0.662**	3.038	**9.996**	**8.442**	34.63	**165.7**	**138.4**	694.7	3904	3253
	1.8	0.280	0.439	0.378	1.341	3.241	2.584	8.245	29.56	22.80	84.25	297.1	229.3
Mixture Density Network	1.0	**1.464**	**17.64**	**16.77**	**115.0**	**2946**	**2800**	–	–	–	–	–	–
	1.2	**0.398**	**2.263**	**2.088**	**7.633**	**154.3**	**142.0**	329.2	**9495**	**8736**	–	–	–
	1.4	**0.244**	**1.306**	**1.156**	**3.448**	**43.70**	**38.53**	**40.85**	**1351**	**1191**	–	–	–
	1.6	**0.119**	**0.330**	**0.281**	**1.751**	**8.035**	**6.718**	**10.45**	**133.2**	**110.3**	**194.6**	**3032**	**2510**
	1.8	**0.099**	**0.262**	**0.207**	**0.829**	**1.549**	**1.283**	**3.112**	**16.29**	**12.37**	**32.38**	**194.6**	**147.3**

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 18: Model Confidence Set Test Results: MAR(0, 2) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0002	0.0002	1.0000	0.0000	0.0004	–	–	–
	1.4	1.0000	0.0000	0.0004	1.0000	0.0102	0.0168	1.0000	0.0094	0.0118	–	–	–
	1.6	1.0000	0.0000	0.0000	1.0000	0.0064	0.1176	1.0000	0.0164	0.0694	1.0000	0.9702	1.0000
	1.8	1.0000	1.0000	1.0000	1.0000	0.3926	1.0000	1.0000	1.0000	1.0000	1.0000	0.9716	1.0000
FlexZBoost	1.0	0.0000	0.0944	0.1002	0.0000	0.0000	0.0002	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002	0.0000	0.0004	–	–	–
	1.4	0.0000	0.0000	0.0002	0.0000	0.0102	0.0168	0.0000	0.0016	0.0044	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.2070	0.1456	0.0000	0.4300	0.2496	0.0000	0.9702	0.6976
	1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0008	0.0002
Mixture Density Network	1.0	0.0000	1.0000	1.0000	0.0006	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0146	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.0004	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0072	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	0.9982
	1.8	0.0092	0.0000	0.0000	0.0000	1.0000	0.0194	0.0000	0.8694	0.4522	0.0000	1.0000	0.4698

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 17 for corresponding RMSE values.

Table 19: Root Mean Squared Error of Predictive Moments: MAR(1, 1) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**1.406**	78.39	74.50	**65.66**	1.956e+04	1.859e+04	–	–	–	–	–	–
	1.2	**0.429**	21.34	19.63	**5.324**	1619	1489	**90.96**	**1.280e+05**	**1.178e+05**	–	–	–
	1.4	**0.262**	7.069	6.228	**1.539**	239.7	211.2	**18.37**	8331	7339	–	–	–
	1.6	**0.170**	2.572	2.129	**0.754**	**46.84**	**38.74**	**4.959**	**840.1**	**694.8**	**93.05**	**1.468e+04**	**1.214e+04**
	1.8	**0.085**	**0.479**	**0.363**	**0.302**	**5.606**	**4.202**	**1.581**	**79.92**	**59.83**	**13.37**	**1127**	**843.5**
FlexZBoost	1.0	2.579	**39.30**	**37.36**	3037	**1.660e+04**	**1.581e+04**	–	–	–	–	–	–
	1.2	1.242	**11.74**	**10.81**	114.1	**1336**	**1230**	739.6	1.562e+05	1.437e+05	–	–	–
	1.4	0.739	**5.023**	**4.438**	11.33	**205.4**	**181.0**	192.2	**8268**	**7284**	–	–	–
	1.6	0.523	**2.267**	**1.898**	5.377	55.86	46.30	103.7	1170	969.2	3413	3.149e+04	2.611e+04
	1.8	0.441	0.723	0.615	2.644	8.440	6.557	20.66	113.4	85.99	343.1	1568	1196
Mixture Density Network	1.0	**2.248**	**22.66**	**21.55**	**88.25**	**3912**	**3718**	–	–	–	–	–	–
	1.2	**0.614**	**5.118**	**4.715**	**13.53**	**341.2**	**314.0**	**637.2**	**2.851e+04**	**2.623e+04**	–	–	–
	1.4	**0.283**	**1.750**	**1.547**	**3.791**	**53.54**	**47.19**	**50.63**	**1767**	**1556**	–	–	–
	1.6	**0.237**	**1.022**	**0.856**	**2.750**	**18.30**	**15.21**	**23.14**	**380.9**	**315.3**	**494.6**	**8415**	**6965**
	1.8	**0.192**	**0.536**	**0.421**	**1.989**	**5.180**	**4.096**	**12.78**	**60.98**	**46.43**	**188.6**	**822.4**	**628.2**

Notes: For details on variable definitions and methodology, refer to Table 7.

Table 20: Model Confidence Set Test Results: MAR(1, 1) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0004	1.0000	0.0000	0.0000	–	–	–
	1.6	1.0000	0.0096	0.0216	1.0000	0.0006	0.0056	1.0000	0.0034	0.0092	1.0000	0.0108	0.0294
	1.8	1.0000	1.0000	1.0000	1.0000	0.7494	1.0000	1.0000	0.1408	1.0000	1.0000	0.1438	1.0000
FlexZBoost	1.0	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0002	0.0000	0.0002	0.0000	0.0000	0.0004	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1.8	0.0000	0.0006	0.0000	0.0000	0.0042	0.0002	0.0000	0.0032	0.0028	0.0000	0.0044	0.0010
Mixture Density Network	1.0	0.0036	1.0000	1.0000	0.0990	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0078	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.7260	1.0000	1.0000	0.0000	1.0000	1.0000	0.0012	1.0000	1.0000	–	–	–
	1.6	0.0220	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0000	0.4446	0.0412	0.0000	1.0000	0.0052	0.0000	1.0000	0.7536	0.0000	1.0000	0.4730

Notes: For details on variable definitions and methodology, refer to Table 8. See Table 19 for corresponding RMSE values.

Table 21: Root Mean Squared Error of Predictive Moments: MARMA(1, 1, 1, 1) Process, 5-Step-Ahead Forecasts

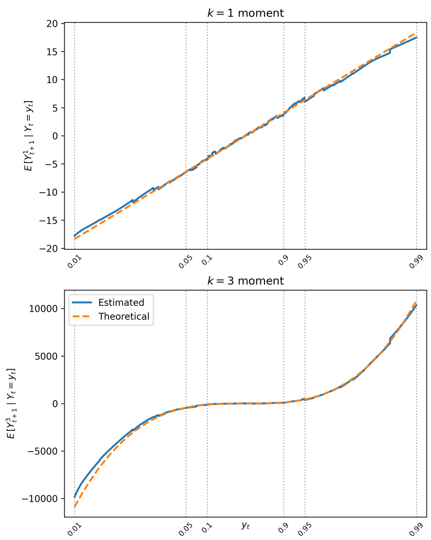
Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	**1.196**	62.23	59.14	**26.36**	**1.014e+04**	**9640**	–	–	–	–	–	–
	1.2	**0.328**	13.40	12.33	**3.038**	788.4	725.3	**62.72**	5.543e+04	5.099e+04	–	–	–
	1.4	**0.199**	4.897	4.315	**0.937**	132.4	116.6	**7.448**	**4080**	**3594**	–	–	–
	1.6	**0.087**	**1.815**	**1.502**	**0.224**	**22.89**	**18.93**	**1.756**	**310.9**	**257.2**	**24.82**	**4377**	**3620**
	1.8	**0.076**	**0.484**	**0.366**	**0.177**	**2.947**	**2.209**	**0.823**	**21.71**	**16.26**	**4.807**	**174.3**	**130.5**
FlexZBoost	1.0	2.129	**29.04**	**27.61**	1747	1.121e+04	1.067e+04	–	–	–	–	–	–
	1.2	0.946	**8.269**	**7.616**	63.64	**626.2**	**576.7**	287.9	**5.139e+04**	**4.728e+04**	–	–	–
	1.4	0.525	**4.014**	**3.545**	6.846	**109.7**	**96.73**	147.7	4537	3997	–	–	–
	1.6	0.355	2.008	1.673	3.500	31.19	25.87	55.75	440.6	365.7	1320	1.021e+04	8477
	1.8	0.284	0.906	0.704	**1.074**	7.723	5.825	7.401	74.10	55.69	84.15	740.3	557.0
Mixture Density Network	1.0	**1.023**	**13.55**	**12.88**	**121.2**	**3338**	**3173**	–	–	–	–	–	–
	1.2	**0.467**	**2.135**	**1.973**	**7.168**	**143.7**	**132.3**	**224.4**	**8459**	**7783**	–	–	–
	1.4	**0.233**	**1.032**	**0.915**	**4.314**	**31.35**	**27.69**	**47.73**	**1209**	**1065**	–	–	–
	1.6	**0.191**	**0.458**	**0.394**	**2.131**	**8.030**	**6.748**	**17.18**	**163.7**	**135.7**	**283.0**	**3511**	**2908**
	1.8	**0.128**	**0.156**	**0.144**	1.241	**2.412**	**1.984**	**5.718**	**19.27**	**14.91**	**75.77**	**230.1**	**179.4**

Notes: For details on variable definitions and methodology, refer to Table 7.

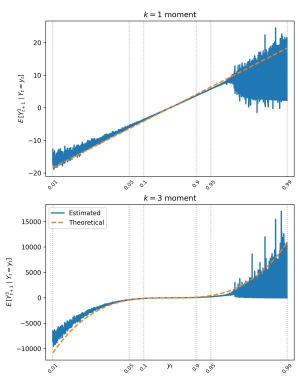
Table 22: Model Confidence Set Test Results: MARMA(1, 1, 1, 1) Process, 5-Step-Ahead Forecasts

Model	α	$\mathbb{E}[y_{t+5} \mid y_t]$			$\mathbb{E}[y_{t+5}^2 \mid y_t]$			$\mathbb{E}[y_{t+5}^3 \mid y_t]$			$\mathbb{E}[y_{t+5}^4 \mid y_t]$		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	0.8642	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	–	–	–
	1.4	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	1.0000	0.0000	0.0000	–	–	–
	1.6	1.0000	0.0000	0.0000	1.0000	0.0000	0.0002	1.0000	0.0010	0.0070	1.0000	0.1186	0.2564
	1.8	1.0000	0.0000	0.0004	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
FlexZBoost	1.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–	–	–	–
	1.2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	–	–	–
	1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0004	0.0002
	1.8	0.0000	0.0000	0.0000	0.0000	0.1674	0.0234	0.0000	0.0664	0.0306	0.0000	0.0012	0.0000
Mixture Density Network	1.0	1.0000	1.0000	1.0000	0.0120	1.0000	1.0000	–	–	–	–	–	–
	1.2	0.0960	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.4	0.2104	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	–	–	–
	1.6	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000
	1.8	0.0002	1.0000	1.0000	0.0000	0.8296	0.0234	0.0000	0.2250	0.0306	0.0000	0.0012	0.0000

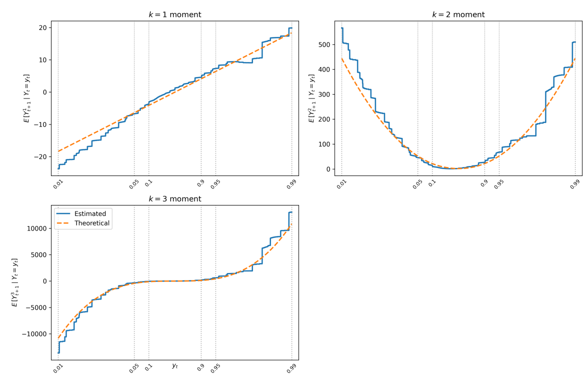
Notes: For details on variable definitions and methodology, refer to Table 8. See Table 21 for corresponding RMSE values.



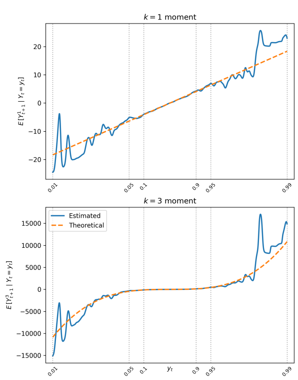
(a) Mixture Density Network



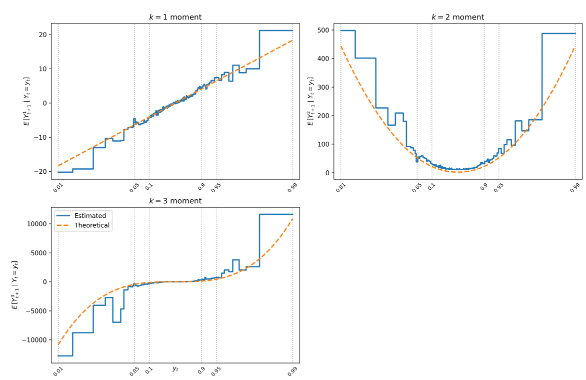
(b) Lanne et al. (2012)



(c) Gourieroux & Jasiak (2025)



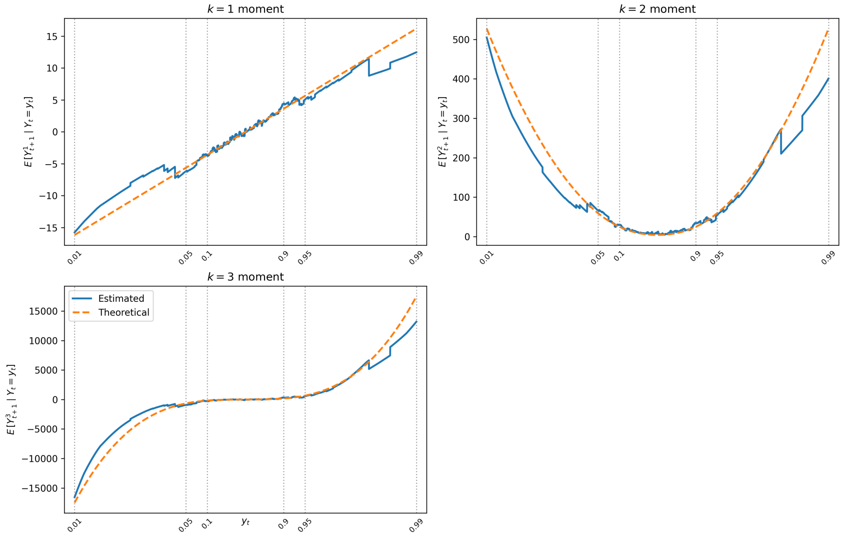
(d) Nadaraya-Watson



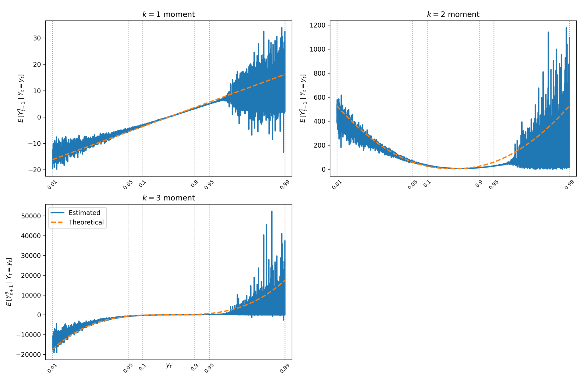
(e) FlexZBoost

Figure 9: Conditional Predictive Moment Accuracy: MAR(0,1) Process, 2-Step-Ahead Forecasts, with tail-index $\alpha = 1.4$

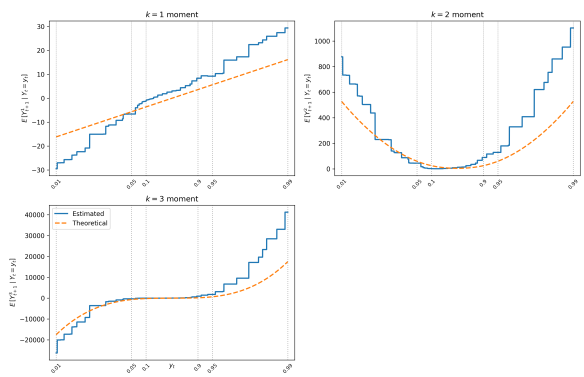
Notes: This figure displays the estimated predictive moments $\mathbb{E}[X_{t+h}^k \mid X_t]$ for $k \in \{1, 2, 3\}$ as a function of the conditioning variable X_t for a purely noncausal MAR(0,1) process with α-stable innovations. Each panel from (a) to (e) shows the results for a specific density forecasting method. Blue curves represent estimated moments, while orange curves show the corresponding theoretical ones.



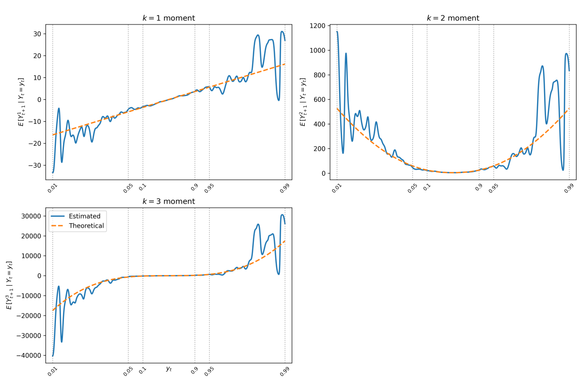
(a) Mixture Density Network



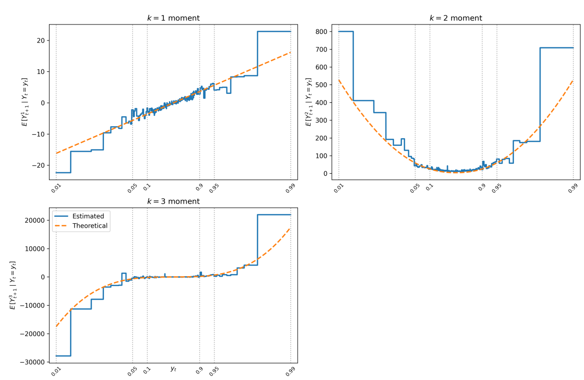
(b) Lanne et al. (2012)



(c) Gourieroux & Jasiak (2025)



(d) Nadaraya-Watson



(e) FlexZBoost

Figure 10: Conditional Predictive Moment Accuracy: MAR(0,1) Process, 5-Step-Ahead Forecasts, with tail-index $\alpha = 1.4$

Notes: For details on variable definitions and methodology, refer to Figure 9.

3.4 Comparison with realized outcomes (additional results to Section 3.4.3)

Let $\hat{p}_h(x \mid X_t)$ denote the estimated predictive density and $\hat{F}_h(x \mid X_t) = \int_{-\infty}^{x} \hat{p}_h(u \mid X_t)du$ the corresponding cumulative distribution function.

Logarithmic score:

$$\text{LogS} = -\log \hat{p}_h(X_{t+h} \mid X_t) \tag{4}$$

Continuous Ranked Probability Score:

$$\text{CRPS} = \int_{-\infty}^{\infty} \left(\hat{F}_h(x \mid X_t) - \mathbf{1}(X_{t+h} \leq x)\right)^2 dx \tag{5}$$

CDE loss:

$$\text{CDE} = \int_{-\infty}^{\infty} \hat{p}_h(x \mid X_t)^2 dx - 2\hat{p}_h(X_{t+h} \mid X_t) \tag{6}$$

Quantile score at level τ:

$$\text{QS}_\tau = (\tau - \mathbf{1}(X_{t+h} < \hat{q}_\tau))\,(X_{t+h} - \hat{q}_\tau) \tag{7}$$

where $\hat{q}_\tau$ satisfies $\hat{F}_h(\hat{q}_\tau \mid X_t) = \tau$.

Table 23: Density Forecast Performance Metrics: MAR(0, 1) Process, 2-Step-Ahead Forecasts

Model	α	CDE Loss			CRPS			Log Prob			QS 10%		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	-0.115	0.078	-0.092	1.712	5.874	2.880	-2.852	-7.174	-3.645	0.593	2.293	1.047
Lanne et al. (2012)	1.0	0.080	0.039	0.061	5.750	7.122	6.707	-4.380	-4.413	-4.996	1.154	1.163	1.444
Gourieroux and Jasiak (2025)	1.0	0.095	0.181	0.163	6.743	8.384	9.474	-4.532	-5.957	-5.449	2.994	4.939	3.790
FlexZBoost	1.0	-0.033	-0.014	-0.011	11.15	12.63	11.06	-3.500	-4.525	-4.826	1.169	5.669	5.489
Mixture Density Network	1.0	-0.085	-0.075	-0.076	1.917	4.174	3.055	-2.689	-3.015	-2.913	0.733	1.171	0.999
Nadaraya-Watson	1.2	-0.163	-0.167	-0.156	0.989	1.852	1.656	-2.086	-3.320	-2.527	0.346	0.481	0.500
Lanne et al. (2012)	1.2	-0.085	-0.029	-0.070	1.212	2.194	1.717	-2.435	-3.167	-2.817	0.385	0.452	0.489
Gourieroux and Jasiak (2025)	1.2	0.092	0.108	0.129	1.959	2.506	2.690	-3.132	-3.527	-3.597	0.790	0.926	0.970
FlexZBoost	1.2	-0.080	-0.076	-0.049	1.713	3.041	2.720	-2.593	-2.842	-3.228	0.554	1.187	1.164
Mixture Density Network	1.2	-0.184	-0.183	-0.164	0.958	1.927	1.589	-1.935	-2.204	-2.176	0.320	0.428	0.416
Nadaraya-Watson	1.4	-0.204	-0.310	-0.191	0.718	0.830	1.068	-1.815	-1.662	-2.111	0.216	0.224	0.270
Lanne et al. (2012)	1.4	-0.203	-0.296	-0.169	0.739	0.912	1.030	-1.732	-2.002	-2.133	0.219	0.294	0.296
Gourieroux and Jasiak (2025)	1.4	-0.068	-0.160	-0.010	0.990	1.046	1.406	-2.148	-2.101	-2.630	0.347	0.400	0.464
FlexZBoost	1.4	-0.171	-0.244	-0.133	0.880	1.008	1.198	-1.942	-2.052	-2.339	0.291	0.315	0.403
Mixture Density Network	1.4	-0.235	-0.360	-0.198	0.705	0.824	1.031	-1.613	-1.461	-1.886	0.213	0.238	0.281
Nadaraya-Watson	1.6	-0.258	-0.452	-0.230	0.581	0.483	0.745	-1.508	-1.051	-1.731	0.166	0.138	0.209
Lanne et al. (2012)	1.6	-0.247	-0.381	-0.205	0.593	0.491	0.762	-1.451	-1.197	-1.831	0.166	0.135	0.212
Gourieroux and Jasiak (2025)	1.6	-0.078	-0.216	-0.050	0.756	0.630	0.971	-1.859	-1.471	-2.235	0.257	0.237	0.325
FlexZBoost	1.6	-0.271	-0.431	-0.219	0.595	0.518	0.784	-1.579	-1.112	-1.823	0.171	0.153	0.223
Mixture Density Network	1.6	-0.283	-0.422	-0.229	0.581	0.527	0.764	-1.387	-1.129	-1.678	0.164	0.224	0.215
Nadaraya-Watson	1.8	-0.342	-0.599	-0.265	0.485	0.322	0.606	-1.207	-0.725	-1.487	0.138	0.102	0.174
Lanne et al. (2012)	1.8	-0.326	-0.421	-0.246	0.487	0.330	0.601	-1.210	-0.882	-1.506	0.134	0.092	0.167
Gourieroux and Jasiak (2025)	1.8	-0.185	-0.501	-0.143	0.565	0.355	0.696	-1.528	-0.826	-1.841	0.173	0.128	0.231
FlexZBoost	1.8	-0.313	-0.521	-0.240	0.502	0.372	0.633	-1.771	-2.097	-2.226	0.141	0.140	0.187
Mixture Density Network	1.8	-0.332	-0.594	-0.263	0.487	0.335	0.603	-1.214	-0.742	-1.507	0.137	0.122	0.175

Notes: This table reports density forecast performance metrics for different tail index values (α). CDE Loss (Conditional Density Estimation loss), CRPS (Continuous Ranked Probability Score), and QS 10% (Quantile Score at 10% level) are loss functions where lower values indicate better performance. Log Prob (Log Probability Score) is a scoring rule where higher values indicate better performance. Metrics are evaluated over three spatial regions: Center $[q_{0.1}, q_{0.9}]$, Tails $[q_{0.01}, q_{0.1}] \cup [q_{0.9}, q_{0.99}]$, and Total $[q_{0.01}, q_{0.99}]$, where q_p represents the p-th quantile. Best method in red, second best in **bold black**.

Table 24: Density Forecast Performance Metrics: MAR(0, 1) Process, 5-Step-Ahead Forecasts

Model	α	CDE Loss			CRPS			Log Prob			QS 10%		
		Center	Tails	Total	Center	Tails	Total	Center	Tails	Total	Center	Tails	Total
Nadaraya-Watson	1.0	-0.027	0.212	-0.010	2.553	10.71	4.790	-3.672	-11.26	-5.044	0.818	2.788	1.383
Lanne et al. (2012)	1.0	0.044	0.032	0.023	4.859	17.90	13.64	-4.105	-6.267	-6.171	1.381	3.430	3.658
Gourieroux and Jasiak (2025)	1.0	-0.038	0.200	0.162	4.068	27.80	26.86	-3.421	-7.376	-5.923	1.493	20.86	6.577
FlexZBoost	1.0	-0.033	-0.013	-0.010	3.636	14.42	11.84	-3.410	-4.571	-4.782	1.170	5.573	5.599
Mixture Density Network	1.0	-0.061	-0.041	-0.050	2.612	8.881	5.124	-2.969	-3.675	-3.338	0.890	1.969	1.461
Nadaraya-Watson	1.2	-0.086	-0.031	-0.071	1.454	3.890	2.741	-2.622	-5.334	-3.451	0.491	1.120	0.725
Lanne et al. (2012)	1.2	-0.041	-0.027	-0.039	1.928	4.377	2.940	-2.769	-3.693	-3.408	0.555	0.988	0.956
Gourieroux and Jasiak (2025)	1.2	0.023	0.083	0.144	3.089	5.384	7.136	-3.357	-4.305	-4.937	1.338	2.798	2.530
FlexZBoost	1.2	-0.077	-0.069	-0.045	1.773	4.070	3.270	-2.614	-3.171	-3.337	0.561	1.235	1.217
Mixture Density Network	1.2	-0.117	-0.078	-0.088	1.391	4.003	2.620	-2.320	-2.939	-2.755	0.454	0.897	0.745
Nadaraya-Watson	1.4	-0.150	-0.215	-0.110	0.968	1.602	1.622	-2.066	-2.385	-2.615	0.287	0.422	0.487
Lanne et al. (2012)	1.4	-0.119	-0.176	-0.092	1.104	1.668	1.610	-2.088	-3.155	-2.805	0.309	0.455	0.515
Gourieroux and Jasiak (2025)	1.4	0.042	-0.045	0.111	1.842	2.326	3.285	-2.931	-2.793	-4.081	0.834	1.234	1.388
FlexZBoost	1.4	-0.149	-0.221	-0.105	1.034	1.599	1.698	-2.018	-2.417	-2.569	0.315	0.503	0.521
Mixture Density Network	1.4	-0.170	-0.195	-0.118	0.959	1.785	1.642	-1.908	-2.079	-2.397	0.284	0.507	0.508
Nadaraya-Watson	1.6	-0.208	-0.291	-0.150	0.756	0.847	1.144	-1.661	-1.531	-2.135	0.220	0.249	0.369
Lanne et al. (2012)	1.6	-0.148	-0.296	-0.113	0.832	0.927	1.147	-1.818	-1.612	-2.323	0.223	0.257	0.365
Gourieroux and Jasiak (2025)	1.6	0.060	-0.015	0.122	1.312	1.318	2.103	-2.613	-2.147	-3.568	0.590	0.592	0.944
FlexZBoost	1.6	-0.204	-0.300	-0.141	0.784	0.972	1.179	-1.865	-3.097	-2.622	0.220	0.369	0.401
Mixture Density Network	1.6	-0.216	-0.315	-0.150	0.760	0.913	1.136	-1.637	-1.472	-2.103	0.217	0.356	0.375
Nadaraya-Watson	1.8	-0.267	-0.436	-0.182	0.604	0.586	0.876	-1.403	-0.998	-1.865	0.168	0.195	0.289
Lanne et al. (2012)	1.8	-0.204	-0.223	-0.136	0.634	0.596	0.875	-1.518	-1.661	-2.179	0.165	0.177	0.280
Gourieroux and Jasiak (2025)	1.8	-0.042	-0.131	0.024	0.855	0.768	1.246	-2.073	-1.262	-2.814	0.347	0.363	0.597
FlexZBoost	1.8	-0.244	-0.261	-0.155	0.638	0.867	0.908	-1.812	-4.552	-3.066	0.177	0.350	0.308
Mixture Density Network	1.8	-0.269	-0.440	-0.181	0.611	0.627	0.876	-1.411	-1.029	-1.884	0.169	0.240	0.291

Notes: For details on variable definitions and methodology, refer to Table 23.

4 Empirical Applications

4.1 Forecasting Natural Gas Prices in Real Time

Table 25: MSPE Ratios Relative to the No-Change Forecast

Horizon	Nadaraya-Watson	Gourieroux and Jasiak (2025)	Lanne et al. (2012)	FlexZBoost	Mixture Density Network
1	1.099	1.083	1.990	**1.081**	**0.928**
3	1.047	1.237	1.781	**0.923**	**0.878**
6	1.032	1.340	2.221	**0.959**	**0.799**
9	0.951	1.557	2.798	**0.659**	**0.746**
12	0.836	1.683	3.290	**0.693**	**0.778**
15	**0.717**	2.325	3.615	**0.755**	0.808
18	0.819	2.769	3.427	**0.806**	**0.805**
21	**0.861**	2.983	3.528	0.927	**0.813**
24	**0.897**	2.922	3.599	0.927	**0.825**

Values below 1 indicate improvements relative to the no-change forecast. Best method in red, second best in **bold black**.

Table 26: Diebold-Mariano Tests: MDN vs. No-change Forecast for the Real Henry Hub Spot Price

Horizon	p-value
1	0.3118
3	0.0292**
6	0.0664*
9	0.1496
12	0.2410
15	0.2011
18	0.0613*
21	0.0909*
24	0.1189

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4.2 Forecasting Inflation in Real Time

Table 27: Estimated MAR(0,2) parameters for US CPI inflation

Parameter	Real-Time In-Sample	Full Period Post-Revised
ψ_1	0.444*** (0.041)	0.440*** (0.041)
ψ_2	0.226*** (0.041)	0.247*** (0.041)
α	1.910*** (0.139)	1.902*** (0.139)
β	0.0009 (0.128)	0.002 (0.128)
σ	0.141*** (0.009)	0.136*** (0.009)

Note: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 28: Diebold-Mariano Tests: MDN vs. No-change Forecast for U.S. Inflation

Horizon	p-value
1	0.0000***
2	0.0000***
3	0.0000***
4	0.0002***
5	0.0011***
6	0.0022***
7	0.0003***
8	0.0001***
9	0.0003***
10	0.0050***
11	0.0393**
12	0.0010***

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

References

Chen, T. & Guestrin, C. (2016), XGBoost: A scalable tree boosting system, in 'Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining', pp. 785–794.

Dalmasso, N., Pospisil, T., Lee, A. B., Izbicki, R., Freeman, P. E. & Malz, A. I. (2020), 'Conditional density estimation tools in python and R with applications to photometric redshifts and likelihood-free cosmological inference', Astronomy and Computing **30**, 100362.

de Truchis, G., Fries, S. & Thomas, A. (2025), 'Forecasting extreme trajectories using seminorm representations'. Working paper.

Fries, S. (2022), 'Conditional moments of noncausal alpha-stable processes and the prediction of bubble crash odds', Journal of Business & Economic Statistics **40**(4), 1596–1616.

Gourieroux, C. & Jasiak, J. (2025), 'Nonlinear fore(back)casting and innovation filtering for causal-noncausal VAR models'.

Gourieroux, C., Lu, Y. & Robert, C.-Y. (2025), 'The causal-noncausal tail processes'.

Gourieroux, C. & Zakoian, J.-M. (2017), 'Local explosion modelling by non-causal process', Journal of the Royal Statistical Society Series B **79**(3), 737–756.

Hansen, P. R., Lunde, A. & Nason, J. M. (2011), 'The model confidence set', Econometrica **79**(2), 453–497.

Izbicki, R. (2017), 'Converting high-dimensional regression to high-dimensional conditional density estimation', Electronic Journal of Statistics **11**(2), 2800–2831.

Lanne, M., Luoto, J. & Saikkonen, P. (2012), 'Optimal forecasting of noncausal autoregressive time series', International Journal of Forecasting **28**(3), 623–631.

Nadaraya, E. A. (1964), 'On estimating regression', Theory of Probability & Its Applications **9**(1), 141–142.

Nolan, J. P. (2020), Univariate Stable Distributions: Models for Heavy Tailed Data, Springer Series in Operations Research and Financial Engineering, Springer, Cham.

Rosenblatt, M. (1969), Conditional probability density and regression estimators, in P. R. Krishnaiah, ed., 'Multivariate Analysis II', Academic Press, pp. 25–31.

Rosenblatt, M. (2000), Gaussian and Non-Gaussian Linear Time Series and Random Fields, Springer.

Samorodnitsky, G. & Taqqu, M. S. (1996), 'Stable non-gaussian random processes: Stochastic models with infinite variance', Bulletin of the London Mathematical Society **28**(5), 554–555.

Watson, G. S. (1964), 'Smooth regression analysis', Sankhyā: The Indian Journal of Statistics, Series A **26**(4), 359–372.