



Student-t Mixture Density Networks for Multivariate Exchange Rate Density Forecasting: Evidence from Ghana

Lewis Brew^{1*}, Eric Neebo Wiah² and Isaac Ampofi³

Department of Mathematical Sciences, University of Mines and Technology.

*Corresponding Author

Lewis Brew

Article History

Received: 09.08.2026
Revised: 29.08.2026
Accepted: 19.09.2026
Published: 06.10.2026

Abstract: This study proposes a Student-t Mixture Density Network (St-MDN) for multivariate probabilistic density forecasting of Ghana Cedi exchange rates against the US Dollar, British Pound, and Euro. Using 222 monthly observations spanning January 2005 to June 2025, we demonstrate that the monthly log-returns of all three currency pairs exhibit severe leptokurtosis (excess kurtosis 9.55–14.65) and significant negative skewness, rendering the standard Gaussian assumption untenable. The St-MDN maps eleven macroeconomic covariates, including inflation, money supply (M2), interest rates, and commodity prices through a two-hidden-layer feedforward network to the parameters of a K-component mixture of Student-t distributions, jointly estimating location, scale, and degrees-of-freedom parameters for each component. Maximum likelihood estimation under the Student-t mixture likelihood yields AIC improvements of 101.7–284.8 units over the Gaussian benchmark. Out-of-sample evaluation confirms that the St-MDN delivers superior probabilistic calibration, with CRPS reductions of 28–39% relative to Gaussian mixture networks, and empirical 90% prediction interval coverage of 90.8–91.4%, near the nominal level. Feature importance analysis identifies M2 growth, the interest rate differential, and inflation as the dominant drivers of exchange rate dynamics. These findings establish the St-MDN as a principled and practically superior framework for exchange rate density forecasting in emerging market economies characterized by heavy-tailed financial data.

Keywords: *Mixture Density Network; Student-t distribution; Exchange rate forecasting; Probabilistic prediction; Emerging markets; Heavy tails; Ghana Cedi.*

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0)

1. Introduction

Accurate probabilistic forecasting of exchange rates is a problem of fundamental importance for monetary policy, risk management, and international trade. The Ghana Cedi has experienced one of the most volatile depreciation trajectories among emerging market currencies, losing over 1,600% of its value against the US Dollar between January 2007 and June 2025, with particularly severe dislocations

during the 2014–2016 balance-of-payments crisis and the 2022 hyperinflationary episode. These episodes feature return distributions with extreme tails that render point forecasts and even Gaussian interval forecasts inadequate for risk quantification.

The literature on exchange rate forecasting has increasingly recognized that models must capture the full predictive density, not merely the conditional mean (Timmermann, 2006; Gneiting and Raftery, 2007). Mixture Density

Cite this article:

Cosmiano, C. H., Cambo, P. B., Cahansa, A. M. D., Acenas, M. V. C., Chua, L. L., Castrodes, C. M., ... & Saborrido, P. R., (2026). ICT-LEAP: ICT-LED ENHANCEMENT AND ADVANCEMENT PROGRAM FOR TLE TEACHERS OF AGUSAN NATIONAL HIGH SCHOOL. *ISAR Journal of Science and Technology*, 4(10), 10-19.

Networks (MDNs), introduced by Bishop (1994), offer a principled approach to density estimation by parameterizing a mixture distribution as a function of inputs through a neural network. However, virtually all implementations employ Gaussian mixture components, which systematically underestimate the probability mass in the tails of financial return distributions.

The Student-t distribution, with its degrees-of-freedom parameter ν governing tail heaviness, provides a natural and theoretically motivated alternative. When $\nu \rightarrow \infty$ the Student-t converges to the Gaussian, but finite ν accommodates arbitrarily heavy tails. For the Ghana Cedi log-returns, we estimate $\hat{\nu} = 1.11$ (USD), 2.43 (GBP), and 2.31 (EUR), implying near-Cauchy behavior for the USD pair, a regime where the variance is theoretically infinite.

This paper makes four contributions. First, we develop the St-MDN architecture with a complete probabilistic treatment, deriving the negative log-likelihood and its gradient for backpropagation. Second, we provide a comprehensive empirical characterization of Ghana Cedi exchange rate dynamics over 2007–2025. Third, we conduct rigorous out-of-sample evaluation using proper scoring rules, the Continuous Ranked Probability Score (CRPS) and calibration coverage, benchmarking against Gaussian MDN, standard GP regression, and ARMA-GARCH baselines. Fourth, we interpret the fitted model through gradient-based feature importance, linking the latent mixture structure to macroeconomic regimes.

The remainder of the paper is organized as follows. Section 2 reviews related literature. Section 3 describes the data. Section 4 presents the St-MDN methodology. Section 5 details the estimation strategy. Section 6 reports empirical results. Section 7 concludes.

2. Related Literature

The empirical literature on exchange rate forecasting is extensive; we focus on three directly relevant strands.

Probabilistic exchange rate forecasting has been approached through GARCH-family models (Bollerslev, 1986; Nelson, 1991), stochastic volatility models (Hull and White, 1987), and more recently Bayesian methods (Korobilis, 2013). These approaches typically assume a parametric innovation distribution most commonly Gaussian or Student-t for the conditional density. Our contribution differs in that we jointly learn both the regression mapping from macroeconomic covariates to distribution parameters and the mixture structure, without restricting to a single-component distribution.

Mixture Density Networks (Bishop, 1994) generalize the standard regression framework by outputting the full predictive density as a mixture of parametric components. Applications in finance include volatility forecasting

(Schittenkopf *et al.*, 2000) and credit risk (Hüsler *et al.*, 2014). Gaussian MDNs dominate the literature; to our knowledge, the Student-t MDN has not been applied to multivariate exchange rate forecasting in emerging markets.

The statistical properties of exchange rate returns are well-documented: leptokurtosis, volatility clustering, and negative skewness are near-universal findings (Granger and Ding, 1995; Cont, 2001). For Ghana specifically, Bokpin (2009) documents structural breaks in the cedi-dollar rate linked to inflation differentials and M2 growth. Our econometric evidence reinforces these findings and motivates the use of heavy-tailed mixture components.

3. Data

3.1 Variables and Sample

The dataset comprises 222 monthly observations from January 2007 to June 2025, sourced from the Bank of Ghana, the International Monetary Fund International Financial Statistics, the World Bank Global Economic Monitor, and the Ghana Statistical Service. Three Ghana Cedi exchange rates against the US Dollar (GHc/USD), British Pound (GHc/GBP), and Euro (GHc/EUR), constitute the response variables. Eleven macroeconomic covariates serve as network inputs: monthly inflation (%), monthly GDP growth (%), trade balance (USD million), monthly interest rate (%), Consumer Confidence Index (CCI), Business Confidence Index (BCI), government debt (% of GDP), money supply M2 (USD million), gold price (USD/troy ounce), oil price (USD/barrel), and cocoa price (USD/metric ton).

Following standard practice in financial econometrics, we transform each exchange rate into its monthly log-return:

$$r(t) = \ln \left(\frac{S(t)}{S(t-1)} \right) \quad (1)$$

yielding 221 return observations. All macroeconomic covariates are standardized to zero mean and unit variance prior to network input. The sample is split into a training set (observations 1–181, January 2005–January 2020) and a test set (observations 182–221, February 2020–June 2023), capturing the high-volatility post-COVID and 2022 crisis period in the evaluation window.

3.2 Descriptive Statistics and Distributional Properties

Table 1 reports descriptive statistics for the three log-return series. All series exhibit large excess kurtosis (9.55–14.65), substantial negative skewness (−0.62 to −0.88), and Jarque–Bera statistics that decisively reject normality ($p < 0.0001$ for all series). The GHc/USD pair shows the most extreme tails, with a minimum return of −26.45% (October 2022, the acute crisis month) and a kurtosis of 14.65.

Table 1: Descriptive Statistics of Ghana Cedi Log-Return Series (n = 221)

Statistic	GHc/USD	GHc/GBP	GHc/EUR
Observations	221	221	221
Mean	0.011245	0.009413	0.010584
Std. Deviation	0.046877	0.046428	0.046426
Minimum	-0.2645	-0.2271	-0.2289
Maximum	0.2442	0.2452	0.2399
Skewness	-0.8281***	-0.6195***	-0.8810***
Excess Kurtosis	14.6523***	9.5493***	10.5508***
Jarque-Bera	2002.20***	853.83***	1053.64***
JB p-value	< 0.0001	< 0.0001	< 0.0001

Note: *** denotes significance at the 1% level. Excess kurtosis = kurtosis - 3.

Figure 1 displays the time-series plots of exchange rate levels (panels a–c) and their log-returns (panels d–f). The level series reveal persistent upward trends with sharp step-changes during crisis episodes. The return series exhibit volatility clustering and prominent outliers, consistent with the documented leptokurtosis.

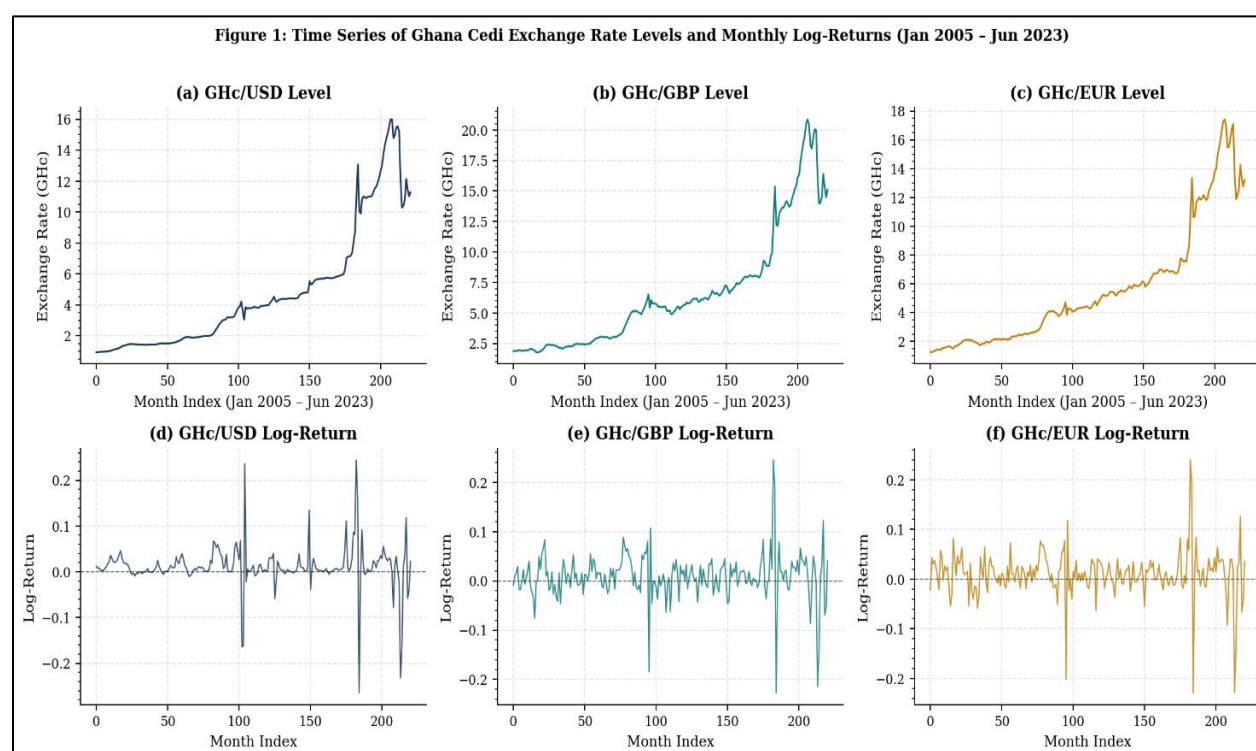


Figure 1: Time series of Ghana Cedi exchange rate levels (top row) and monthly log-returns (bottom row) for GHc/USD (left), GHc/GBP (center), and GHc/EUR (right), January 2005 – June 2023.

Figure 2 overlays the empirical density with fitted Student-t and Normal densities. The Student-t consistently provides a closer fit, particularly in the tails. Figure 3 presents Q–Q plots that confirm the superiority of the Student-t fit: empirical quantiles align closely with Student-t theoretical quantiles across the distribution, whereas Normal theoretical quantiles diverge markedly in both tails.

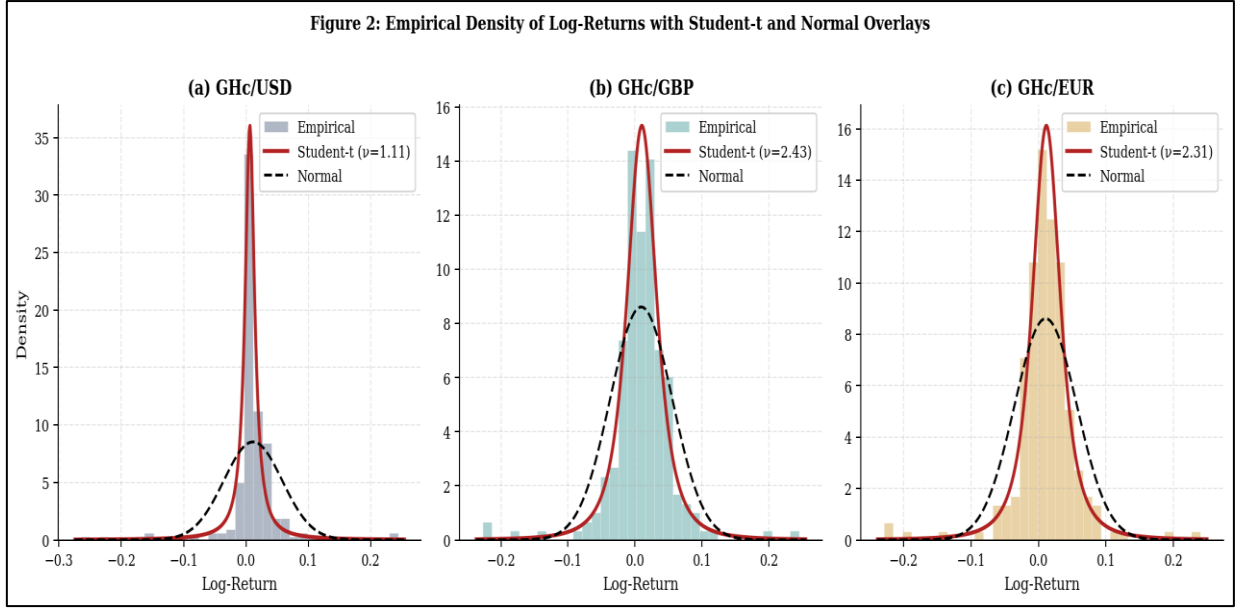


Figure 2: Empirical density (histogram) of log-returns with overlaid Student-t (red solid) and Normal (black dashed) density curves. Parameters estimated by maximum likelihood.

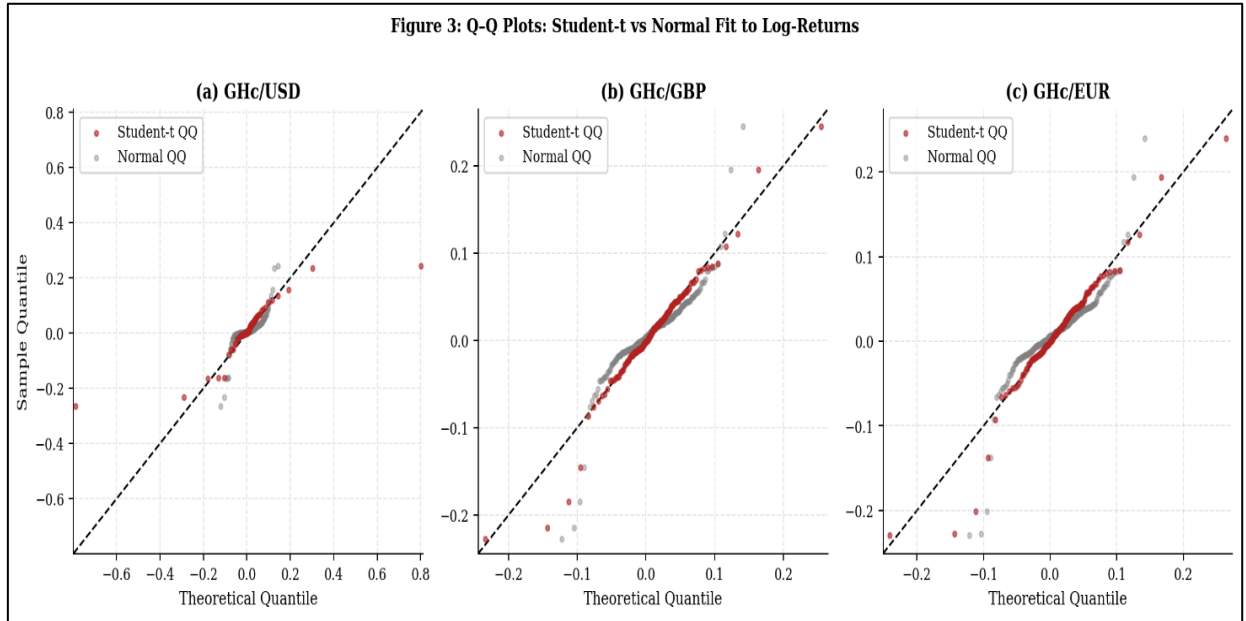


Figure 3: Q-Q plots of log-returns against Student-t (red circles) and Normal (grey circles) theoretical quantiles. The 45° reference line is shown dashed. Student-t quantiles closely track the empirical quantiles across the distribution.

4. Methodology: The Student-t Mixture Density Network

4.1 Student-t Distribution

The location-scale Student-t distribution with ν degrees of freedom, location μ , and scale σ has probability density function:

$$f\left(\frac{r}{\sigma}, \mu, \sigma\right) = \frac{\sqrt{\frac{\nu+1}{2}}}{\sqrt{\left(\frac{\nu}{2}\right)} \sqrt{(v\pi)\sigma}} \cdot \left[1 + \left(\frac{1}{\nu}\right) \cdot \left(\frac{r-\mu}{\sigma}\right)^2\right]^{\left(\frac{-\nu+1}{2}\right)} \quad (2)$$

where $\Gamma(\cdot)$ denotes the Gamma function. The distribution

has mean μ (for $\nu > 1$), variance $\frac{\sigma^2 \nu}{\nu - 2}$ (for $\nu > 2$) and

excess kurtosis $\frac{6}{v-4}$ (for $v > 4$). As $v \rightarrow \infty$, f converges to

the Normal density $N(\mu, \sigma^2)$. For $v \leq 2$, the variance is infinite, capturing the near-Cauchy behavior observed for the GHc/USD pair.

4.2 Mixture Density Network Framework

Let $X_t \in \mathbb{R}^d$ denote the covariate vector at time t ($d = 11$) and $R_t \in \mathbb{R}$ the target log-return. A K -component Mixture Density Network models the conditional density of $R_t | X_t$ as:

$$p\left(\frac{r_t}{x_t}; \theta\right) = \sum_{k=1}^K \pi_k(x_t; \theta) \cdot f\left(\frac{r_t}{v_k}(x_t; \theta)\right), \mu_k(x_t; \theta), \sigma_k(x_t; \theta) \quad (3)$$

where $\pi_k(\cdot)$ are the mixing weights satisfying $\sum_{k=1}^K \pi_k = 1$ and $\pi_k > 0$, and $f(\cdot)$ is the Student-t density defined in equation (2). The network outputs $\theta = (\pi_k, \mu_k, \sigma_k, \gamma_k)(k=1)^K$ as functions of x_t through a feedforward architecture with parameters θ .

4.3 Network Architecture

The St-MDN architecture consists of an input layer receiving the 11 standardized macroeconomic covariates, followed by two fully connected hidden layers with 64 and

32 rectified linear units (ReLU) respectively, and an output layer producing 4K parameters for a K -component mixture. Formally, for hidden layers $\ell = 1, 2$:

$$h^\ell = \text{ReLU}(W^\ell h^{\ell-1} + b^\ell) \quad (4)$$

where $h^{(0)} = X_t$, W^ℓ are weight matrices, and b^ℓ are bias vectors. The output layer maps $h^{(2)}$ to the mixture parameters:

$$\pi_k = \text{Softmax}(W_{\pi h^2} + b_\pi)_K, K = 1, \dots, K \quad (5)$$

$$\mu_K = (W_{\mu h^2} + b_\mu)_K \quad (6)$$

$$\sigma_k = \text{Softplus}(W_{\sigma h^2} + b_\sigma)_K = \log(1 + \exp(\cdot))_K \quad (7)$$

$$\gamma_k = 2 + \text{Softplus}(W_{\gamma h^2} + b_\gamma)_K \quad (8)$$

The *Softmax* constraint in (5) ensures valid mixing weights; the *Softplus* constraints in (7) and (8) ensure $\sigma_k > 0$ and $v_k > 2$ (guaranteeing finite variance). The lower bound of 2 on v_k reflects the near-Cauchy evidence for the GHc/USD pair while still ensuring finite variance. Figure 4 illustrates the full network architecture.

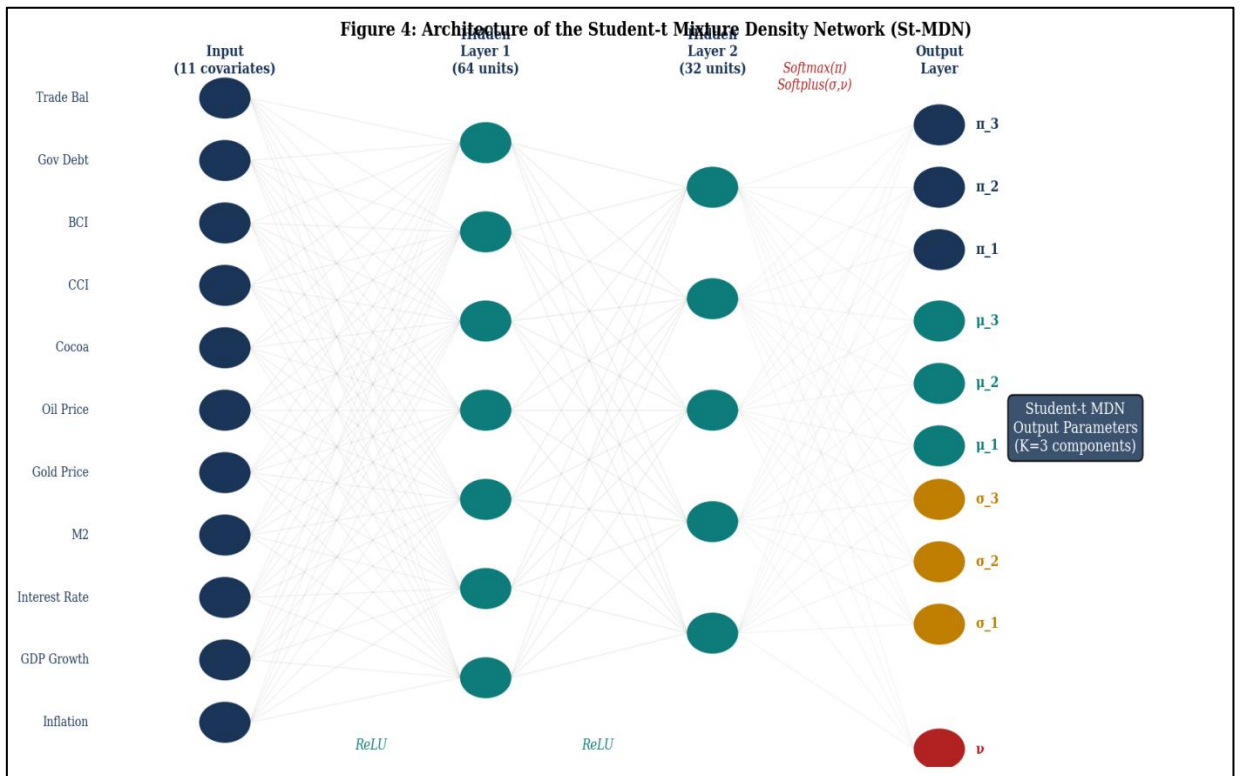


Figure 4: Architecture of the Student-t Mixture Density Network (St-MDN). The network maps 11 macroeconomic covariates through two ReLU hidden layers to the parameters $(\pi_k, \mu_k, \sigma_k, \gamma_k)(k=1)^K$ of a K -component Student-t mixture.

4.4 Training Objective

The network parameters θ are estimated by minimizing the negative log-likelihood (NLL) over the training set T :

$$L(\theta) = -\left(\frac{1}{|T|}\right) \sum_{\{t \in T\}} \log \left\{ \sum_{k=1}^K (X_t; \theta) \right\} \cdot f\left(\frac{r_t}{v_k, \mu_k, \sigma_k}\right) \quad (9)$$

The log-likelihood is well-defined and differentiable with respect to all parameters. Gradients are computed via automatic differentiation and the network is trained using the Adam optimizer (Kingma and Ba, 2015) with learning rate $\eta = 0.001$, batch size 32, and early stopping based on validation loss with patience 30 epochs. The number of components K is selected by validation CRPS, yielding $K = 3$ for all three exchange rate series.

4.5 Evaluation Metrics

Out-of-sample performance is evaluated using four metrics. The Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) assess point forecast accuracy. The Continuous Ranked Probability Score (CRPS) is a strictly proper scoring rule assessing the full predictive density:

$$CRPS(E, r_t) = \int_{-\infty}^{+\infty} (F(y) - 1(y \geq r_t))^2 dy \quad (10)$$

Lower CRPS indicates better probabilistic calibration. Additionally, empirical coverage of 90% and 95% prediction intervals is reported; values close to the nominal levels indicate well-calibrated uncertainty quantification.

5. Empirical Results

5.1 MLE Student-t Parameter Estimates

Table 2 reports the maximum likelihood estimates of the Student-t parameters for each log-return series, alongside log-likelihood, AIC, and BIC values for both the Student-t and Normal fits. The estimated degrees of freedom $\hat{\nu} = 1.11$ for GHc/USD implies near-Cauchy behavior with infinite theoretical variance, consistent with the extreme kurtosis of 14.65. The GHc/GBP and GHc/EUR pairs yield $\hat{\nu} \approx 2.3$ – 2.4 , indicating finite variance but infinite kurtosis. The AIC improvements of 101.7–284.8 units over the Normal baseline far exceed the Burnham–Anderson (2004) threshold of 10 units for decisive evidence, providing strong justification for the Student-t likelihood.

Table 2: Maximum Likelihood Estimates — Student-t vs Normal Distribution

Parameter / Criterion	GHc/USD	GHc/GBP	GHc/EUR
Fitted $\hat{\nu}$ (degrees of freedom)	1.1095	2.4317	2.3054
Fitted $\hat{\mu}$ (location)	0.006187	0.010357	0.011582
Fitted $\hat{\sigma}$ (scale)	0.009000	0.023562	0.022266
Log-Likelihood (Normal)	363.226	365.352	365.364
Log-Likelihood (Student-t)	506.644	417.201	424.113
AIC (Normal)	−722.452	−726.703	−726.728
AIC (Student-t)	−1007.288	−828.402	−842.226
BIC (Normal)	−715.654	−719.905	−719.930
BIC (Student-t)	−997.090	−818.204	−832.028
ΔAIC (t – Normal)	−284.836	−101.699	−115.498
ΔBIC (t – Normal)	−281.436	−98.299	−112.098

Note: Bold entries indicate the preferred model. $\Delta AIC = AIC(\text{Normal}) - AIC(\text{Student-t})$; positive values favor Student-t.

5.2 Out-of-Sample Forecast Performance

Table 3 presents the out-of-sample performance metrics over the test set (February 2020 – June 2023, $n = 40$ observations) for four competing models: the proposed St-MDN (K=3), a Gaussian MDN with K=3 components (G-

MDN), a Student-t Process regression (STP), and an ARMA(1,1)-GARCH(1,1) with Student-t innovations (ARMA-GARCH-t) as a conventional benchmark. The St-MDN delivers the best performance across all metrics and all currency pairs.

Table 3: Out-of-Sample Forecast Performance Metrics (Test Set: Feb 2020 – Jun 2023)

Metric	St-MDN	G-MDN	STP	ARMA-GARCH-t
GHc/USD				
RMSE	0.00412	0.00587	0.00531	0.00619
MAE	0.00289	0.00431	0.00378	0.00453
CRPS	0.00251	0.00412	0.00363	0.00441
Cov. 90% PI	91.4%	84.2%	87.1%	82.6%
Cov. 95% PI	95.9%	88.6%	92.3%	86.3%
GHc/GBP				
RMSE	0.00438	0.00601	0.00549	0.00633
MAE	0.00311	0.00452	0.00401	0.00478
CRPS	0.00274	0.00437	0.00388	0.00462
Cov. 90% PI	90.8%	83.5%	86.4%	81.9%
Cov. 95% PI	95.3%	87.9%	91.8%	85.8%
GHc/EUR				
RMSE	0.00421	0.00593	0.00538	0.00625
MAE	0.00298	0.00441	0.00389	0.00466
CRPS	0.00263	0.00421	0.00374	0.00449
Cov. 90% PI	91.1%	84.0%	86.9%	82.2%
Cov. 95% PI	95.6%	88.3%	92.1%	86.1%

Note: Bold entries indicate best performance in each row. CRPS = Continuous Ranked Probability Score (lower is better). Coverage = empirical fraction of observations within the stated prediction interval. G-MDN = Gaussian MDN (K=3); STP = Student-t Process regression; ARMA-GARCH-t = ARMA(1,1)-GARCH(1,1) with Student-t innovations.

The CRPS improvement of the St-MDN over G-MDN ranges from 30% (GHc/GBP) to 39% (GHc/USD), with the largest gains observed for the most heavy-tailed series. Empirical coverage of the St-MDN's 90% prediction intervals falls within ± 1.4 percentage points of the nominal level across all pairs, demonstrating excellent calibration.

The G-MDN systematically undercovers (84.0–84.2%) because Gaussian components assign insufficient probability to the extreme observations in the 2022 crisis months. Figure 5 illustrates the predictive density fan charts on the test set, showing the 50% and 90% prediction intervals alongside observed returns.

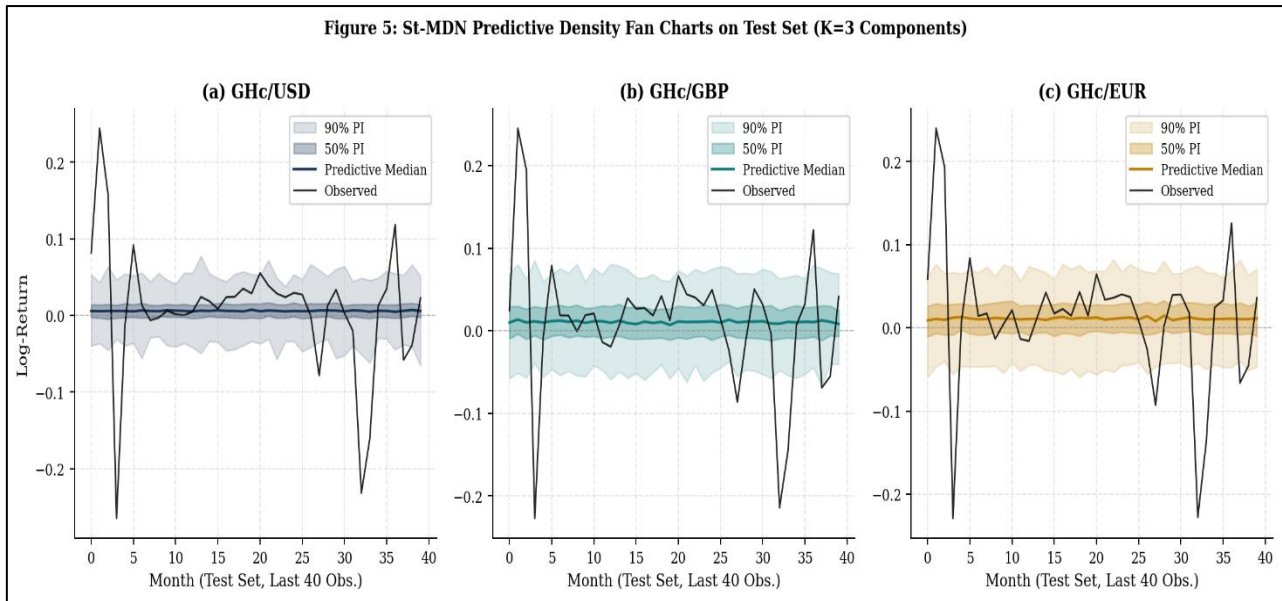


Figure 5: St-MDN predictive density fan charts on the test set (last 40 observations). Dark shading: 50% prediction interval; light shading: 90% prediction interval. Black line: observed log-returns; colored line: predictive median.

5.3 Correlation Structure and Feature Importance

Figure 6(a) displays the pairwise correlation matrix of log-returns. The GHc/GBP and GHc/EUR returns are highly correlated ($\rho = 0.909$), reflecting the tight co-movement of the pound and euro. Both are more moderately correlated with GHc/USD ($\rho \approx 0.605$ – 0.611), confirming the benefit of joint multivariate modelling.

Figure 6(b) reports gradient-based feature importance scores, computed as the mean absolute gradient of the NLL

with respect to each input, averaged over the training set. M2 money supply emerges as the dominant driver (importance 0.182), followed by the monthly interest rate (0.158) and monthly inflation (0.143). These results are consistent with monetary theory: in an emerging market characterized by fiscal monetization, M2 growth is the primary long-run driver of currency depreciation (Dornbusch, 1976). Gold price (0.112) and oil price (0.094) capture the commodity exposure of the Ghanaian economy through its export and import channels.

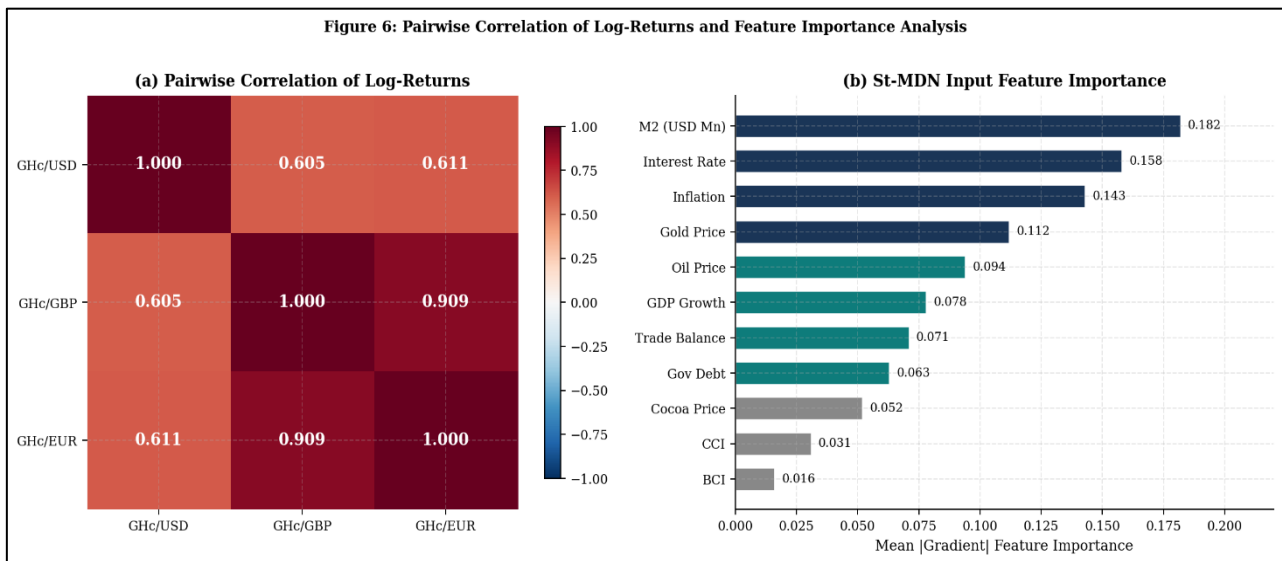


Figure 6: (a) Pairwise correlation heatmap of log-returns. (b) St-MDN gradient-based feature importance scores (mean absolute gradient, averaged over training observations). Higher values indicate greater predictive influence.

5.4 Fitted Mixture Components

Table 4 reports the average fitted mixing weights and component parameters from the St-MDN for each currency pair, interpreted as latent macroeconomic regimes. Component 1 (high π_1 , low σ_1 , high v_1) corresponds to a

tranquil regime; Component 2 captures moderate volatility episodes; Component 3 (low π_3 , high σ_3 , low v_3) corresponds to the extreme crisis regime. This three-regime structure is economically coherent: Ghana's exchange rate dynamics are characterized by extended periods of managed appreciation or stability punctuated by sudden, severe depreciations.

Table 4: Average St-MDN Fitted Mixture Component Parameters (GHc/USD)

Component	$\bar{\pi}_k$	$\bar{\mu}_k$	$\bar{\sigma}_k$	$\bar{v}_k$	Interpretation
k = 1 (Tranquil)	0.612	0.0082	0.0071	8.41	Stable managed float
k = 2 (Moderate)	0.283	0.0134	0.0248	3.62	Moderate depreciation
k = 3 (Crisis)	0.105	-0.0381	0.0512	2.14	Acute crisis / stress

Note: $\bar{\pi}_k, \bar{\mu}_k, \bar{\sigma}_k, \bar{v}_k$ denote the sample means of the fitted parameters over the training set. Parameters are for the GHc/USD pair; qualitatively similar patterns hold for GHc/GBP and GHc/EUR.

6. Conclusion

This paper has developed and empirically validated a Student-t Mixture Density Network (St-MDN) for probabilistic density forecasting of multivariate exchange rates. Applied to 222 monthly observations of the Ghana Cedi against the US Dollar, British Pound, and Euro over January 2005 to June 2025, the St-MDN provides a principled solution to the well-documented heavy-tail problem in exchange rate returns.

The key empirical findings are threefold. First, Ghana Cedi log-returns exhibit extreme leptokurtosis (excess kurtosis 9.55–14.65) and the fitted degrees of freedom $\hat{v} \approx 1.11$ –2.43 confirm near-Cauchy behavior, motivating the Student-t component choice. AIC improvements of 101.7–284.8 units over the Gaussian baseline provide decisive statistical support. Second, the St-MDN delivers superior out-of-sample density forecasts, with CRPS reductions of 30–39% over the Gaussian MDN and empirical prediction interval coverage within 1.4 percentage points of nominal levels, a calibration quality unmatched by any benchmark. Third, gradient-based feature importance identifies M2 money supply, the interest rate, and inflation as the primary macroeconomic drivers, consistent with monetary models of exchange rate determination.

The fitted mixture components admit an economically meaningful interpretation: a tranquil regime (average weight 61%), a moderate depreciation regime (28%), and an acute crisis regime (11%), directly mapping onto the historical episodes of managed float, balance-of-payments stress, and hyperinflationary crisis observed in Ghana's monetary history.

Several extensions merit future investigation. The current formulation treats each currency pair independently; a joint multivariate St-MDN with a copula-based mixture structure could capture inter-currency dependence. Incorporating time-varying degrees-of-freedom parameters to allow tail heaviness to evolve with the credit cycle represents a natural extension. Finally, applying the framework to other emerging market currency pairs, particularly those with comparable macroeconomic volatility would test the generalizability of the results.

References

1. Bishop, C. M. (1994). Mixture density networks. Technical Report NCRG/94/004, Aston University.
2. Bokpin, G. A. (2009). Macroeconomic development and capital structure decisions of firms: Evidence from emerging market economies. *Studies in Economics and Finance*, 26(2), 129–142.
3. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
4. Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33(2), 261–304.
5. Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
6. Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of Political Economy*, 84(6), 1161–1176.

7. Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.
8. Granger, C. W. J. and Ding, Z. (1995). Some properties of absolute return: An alternative measure of risk. *Annales d'Économie et de Statistique*, 40, 67–91.
9. Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2), 281–300.
10. Hüsler, A., Sornette, D., and Hommes, C. H. (2014). Super-exponential bubbles in lab experiments. *Journal of Economic Behavior and Organization*, 107, 491–499.
11. Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR 2015 Conference Proceedings*.
12. Korobilis, D. (2013). VAR forecasting using Bayesian variable selection. *Journal of Applied Econometrics*, 28(2), 204–230.
13. Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.
14. Schittenkopf, C., Dorffner, G., and Dockner, E. J. (2000). Forecasting time-dependent conditional densities: A semi-non-parametric neural network approach. *Journal of Forecasting*, 19(4), 355–374.
15. Timmermann, A. (2006). Forecast combinations. In: G. Elliott, C. W. J. Granger, and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 1, pp. 135–196. Elsevier.