Gaussian Process Conditional Copulas with Applications to Financial Time Series

José Miguel Hernández-Lobato[1], James Robert Lloyd[1] and Daniel Hernández-Lobato[2]

1. Cambridge University 2. Universidad Autónoma de Madrid

UNIVERSITY OF CAMBRIDGE

UNIVERSIDAD AUTÓNOMA DE MADRID

Neural Information Processing Systems Foundation

1. Introduction

Motivation: Understanding dependencies is a central problem in the analysis of **multi-variate financial time series**. These dependencies are traditionally described using co-variance matrices. However, a more general approach is to use **copulas**. Copulas are general models for dependence.

Problem: The use of copulas is inaccurate when the actual dependencies in the data are strongly influenced by other covariates. For example, dependencies can vary with time or be affected by observations of other time series.

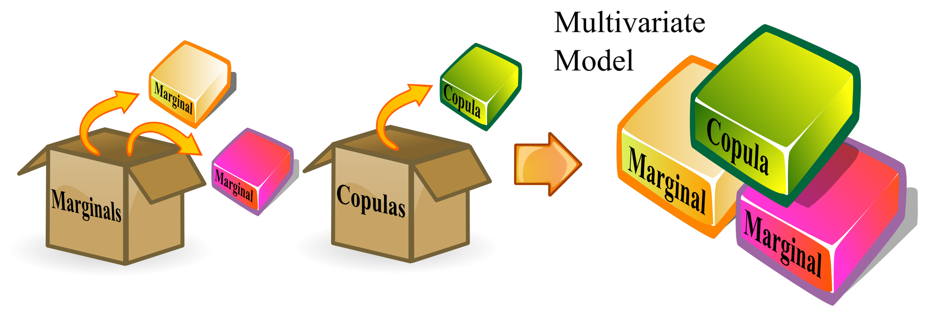
Solution: A probabilistic framework to estimate semi-parametric **conditional copulas**. We assume parametric copulas whose parameters are specified by unknown non-linear functions of arbitrary conditioning variables. The latent functions are approximated using **Gaussian processes** and **expectation propagation** for fast approximate inference.

2. Separating Marginals and Dependence

Sklar's Theorem: any joint distribution can be written in terms of its copula and its uni-variate marginals [Sklar, 1959]. The joint density is

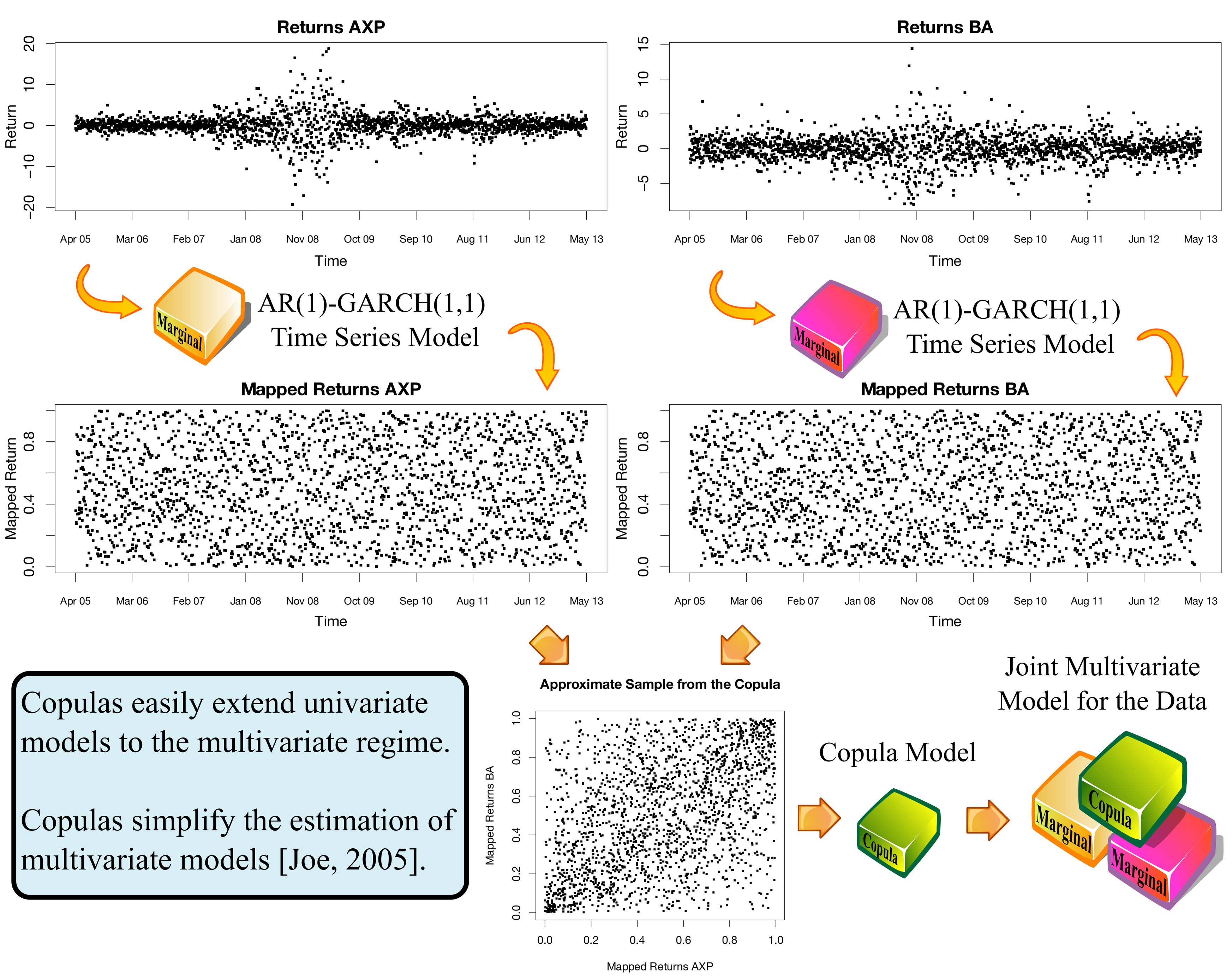
$$p(x_1, x_2) = \underbrace{c(u_1, u_2)}_{\text{copula}} \underbrace{p_1(x_1)\, p_2(x_2)}_{\text{marginals}},$$

where $u_i := P_i(x_i)$ and P_i denotes the marginal cdf for x_i.



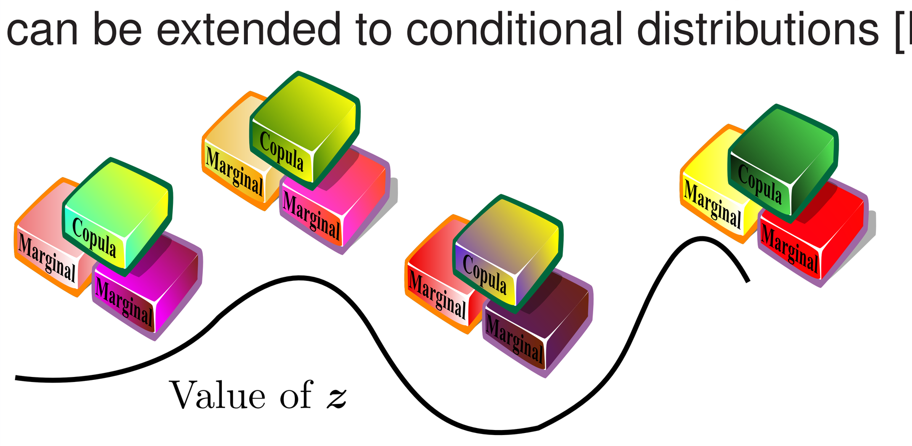
We can use copulas and marginals as **building blocks** for new multivariate models.

3. Additional Advantages of Copulas



4. Conditional Copulas

The distribution of x_1 and x_2 may depend on a covariate vector $\mathbf{z}$. The copula framework can be extended to conditional distributions [Patton, 2006].



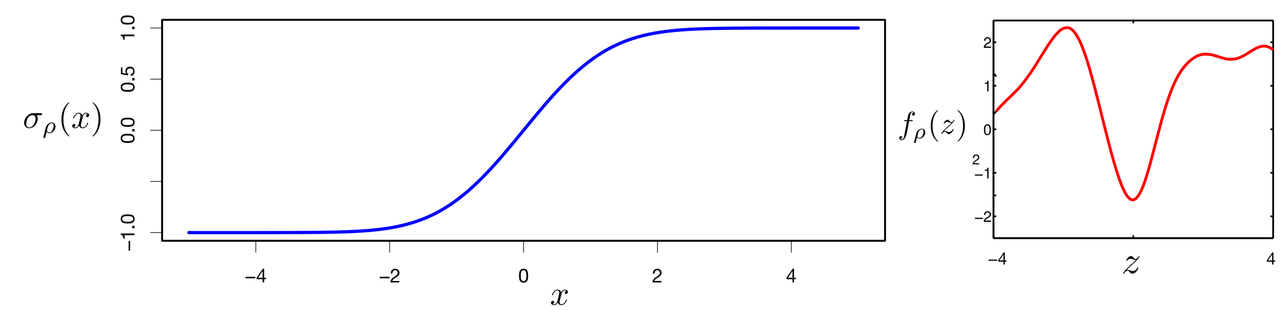
$$p(x_1, x_2|\mathbf{z}) = \underbrace{c(u_1, u_2|\mathbf{z})}_{\substack{\text{conditional}\\\text{copula}}} \underbrace{p_1(x_1|\mathbf{z})\, p_2(x_2|\mathbf{z})}_{\substack{\text{conditional}\\\text{marginals}}},$$

where $u_i := P_i(x_i|\mathbf{z})$ and P_i denotes the marginal cdf for x_i given $\mathbf{z}$.

5. Semiparametric Conditional Copulas and GPs

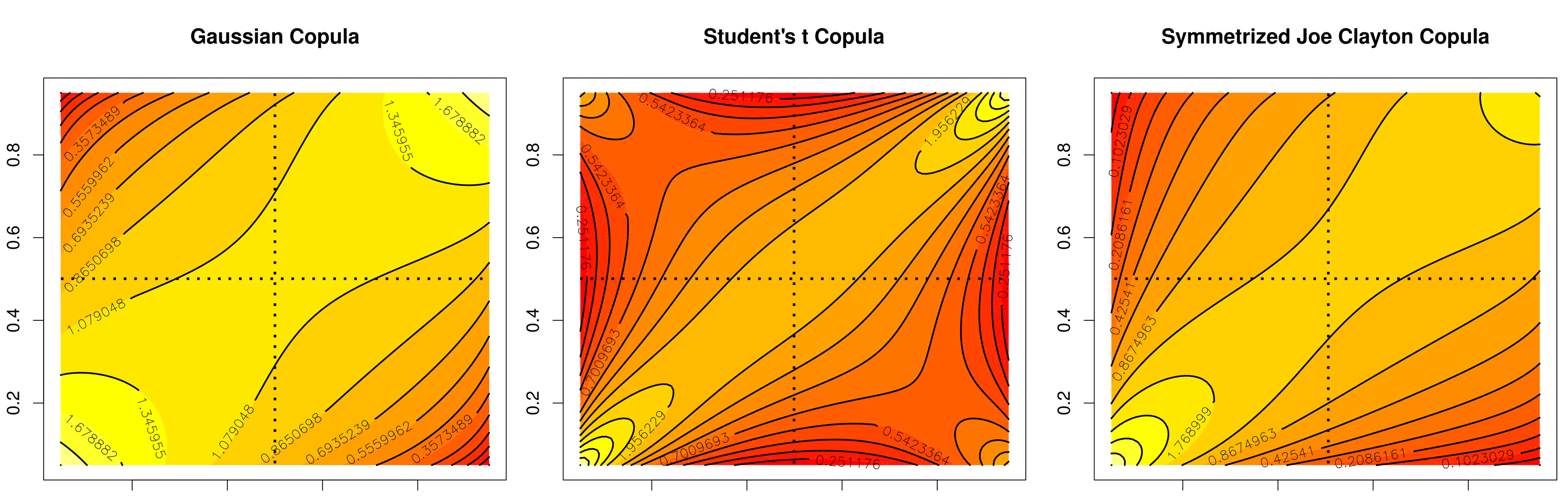
Main Assumption: $c(u_1, u_2|\mathbf{z})$ is a parametric copula model $c_{\text{par}}[u_1, u_2|\theta_1(\mathbf{z}), \ldots, \theta_k(\mathbf{z})]$ specified by k scalar parameters $\theta_1, \ldots, \theta_k$ that are functions of $\mathbf{z}$. We select $\theta_i(\mathbf{z}) = \sigma_i[f_i(\mathbf{z})]$, where f_i is a real function and σ_i is a link function that maps $\mathbb{R}$ to a set Θ_i of valid configurations for θ_i.

Example: conditional Student's t copula. $c_{\text{student}}[u_1, u_2|\rho(\mathbf{z}), \nu(\mathbf{z})]$, $\rho(\mathbf{z}) = \sigma_\rho[f_\rho(\mathbf{z})]$, $\nu(\mathbf{z}) = \sigma_\nu[f_\nu(\mathbf{z})]$, $\sigma_\rho(x) = 2\Phi(x) - 1$, $\sigma_\nu(x) = \exp(x)$, where Φ is the probit function.



We estimate the latent functions $f_1, \ldots, f_k$ using **Gaussian processes**.

6. Parametric Copula Models



Parameters: ρ Parameters: ρ, ν Parameters: τ_U, τ_L

7. Expectation Propagation for Approximate Inference

$$p(\mathbf{f}_1, \ldots, \mathbf{f}_k, \mathcal{D}_{U,V}|\mathcal{D}_\mathbf{Z}) = \overbrace{\left[\prod_{i=1}^n c_{\text{par}}[u_{1i}, u_{2i}|\sigma_1(f_{1i}), \ldots, \sigma_k(f_{ki})]\right]}^{\text{Copula Likelihood (Exact Factor)}} \overbrace{\left[\prod_{j=1}^k \mathcal{N}(\mathbf{f}_j|\mathbf{m}_\mathbf{Z}^{(j)}, \mathbf{V}_{\mathbf{Z},\mathbf{Z}}^{(j)})\right]}^{\text{Gaussian Prior}}$$

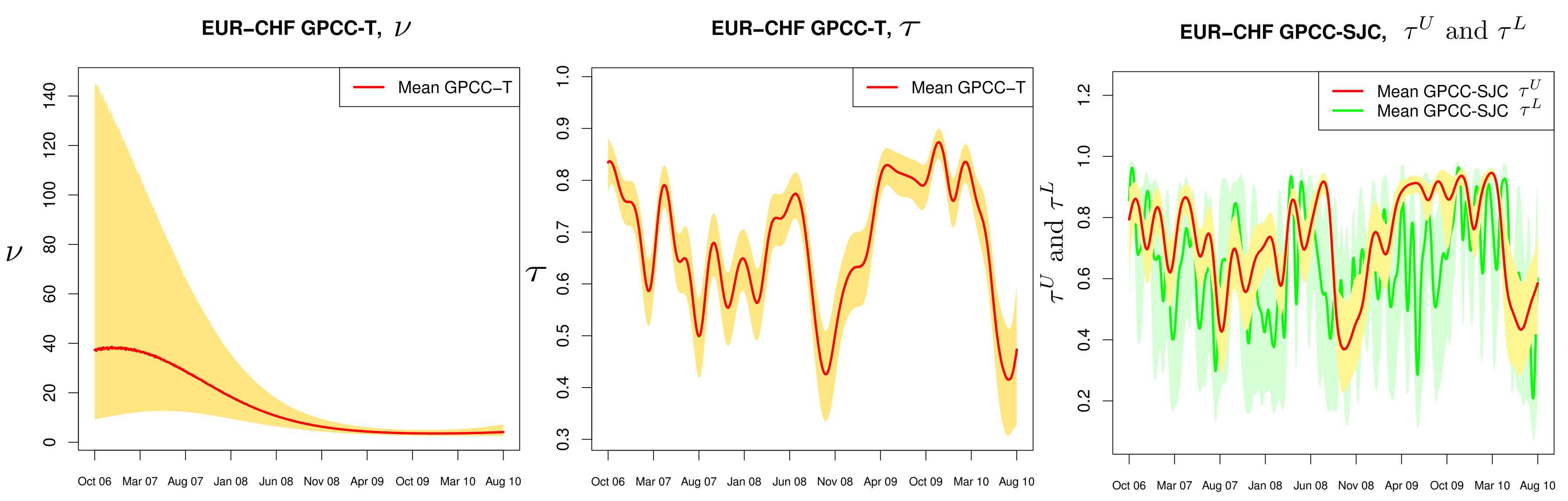
EP Projection $\qquad$ Exact

$$\mathcal{Q}(\mathbf{f}_1, \ldots, \mathbf{f}_k) = \underbrace{\left[\prod_{i=1}^n \tilde{k}_i \left[\prod_{j=1}^k \mathcal{N}(f_{ji}|, \tilde{m}_{ji}, \tilde{v}_{ji})\right]\right]}_{\text{Approximate Gaussian Factor}} \left[\prod_{j=1}^k \mathcal{N}(\mathbf{f}_j|\mathbf{m}_\mathbf{Z}^{(j)}, \mathbf{V}_{\mathbf{Z},\mathbf{Z}}^{(j)})\right]$$

where $\mathbf{f}_j = (f_{j1}, \ldots, f_{jk})^\top$. We run EP with sparse GPs [Naish-Guzman et al. 2008].

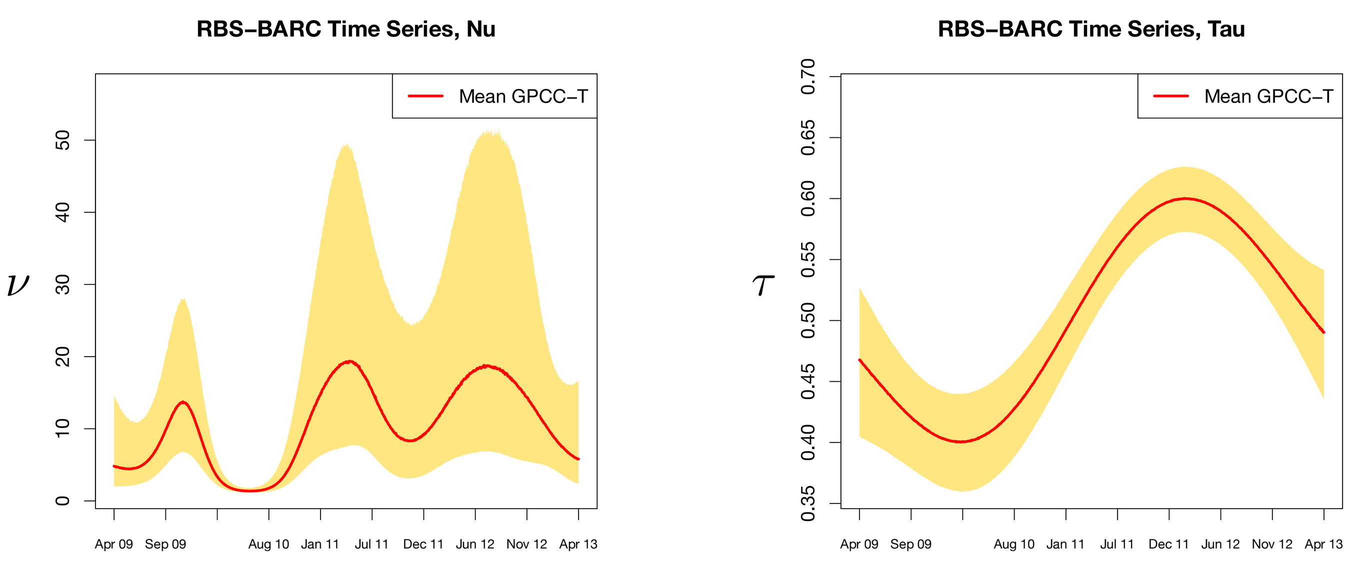
8. Results on Foreign Exchange Time Series

Method	AUD	CAD	JPY	NOK	SEK	EUR	GBP	NZD
GPCC-G	0.1260	0.0562	**0.1221**	0.4106	0.4132	0.8842	0.2487	0.1045
GPCC-T	**0.1319**	**0.0589**	0.1201	**0.4161**	**0.4192**	**0.8995**	**0.2514**	**0.1079**
GPCC-SJC	0.1168	0.0469	0.1064	0.3941	0.3905	0.8287	0.2404	0.0921
HMM	0.1164	0.0478	0.1009	0.4069	0.3955	0.8700	0.2374	0.0926
TVC	0.1181	0.0524	0.1038	0.3930	0.3878	0.7855	0.2301	0.0974
DSJCC	0.0798	0.0259	0.0891	0.3994	0.3937	0.8335	0.2320	0.0560
CONST-G	0.0925	0.0398	0.0771	0.3413	0.3426	0.6803	0.2085	0.0745
CONST-T	0.1078	0.0463	0.0898	0.3765	0.3760	0.7732	0.2231	0.0875
CONST-SJC	0.1000	0.0425	0.0852	0.3536	0.3544	0.7113	0.2165	0.0796



9. Results on Equity Exchange Time Series

Method	HD HON	AXP BA	CNW CSX	ED EIX	HPQ IBM	BARC HSBC	RBS BARC	RBS HSBC
GPCC-G	0.1247	0.1133	0.1450	**0.2072**	0.1536	0.2424	0.3401	0.1860
GPCC-T	**0.1289**	**0.1187**	**0.1499**	0.2059	**0.1591**	0.2486	**0.3501**	**0.1882**
GPCC-SJC	0.1210	0.1095	0.1399	0.1935	0.1462	0.2342	0.3234	0.1753
HMM	0.1260	0.1119	0.1458	0.2040	0.1511	**0.2486**	0.3414	0.1818
TVC	0.1251	0.1119	0.1459	0.2011	0.1511	0.2449	0.3336	0.1823
DSJCC	0.0935	0.0750	0.1196	0.1721	0.1163	0.2188	0.3051	0.1582
CONST-G	0.1162	0.1027	0.1288	0.1962	0.1325	0.2307	0.2979	0.1663
CONST-T	0.1239	0.1091	0.1408	0.2007	0.1481	0.2426	0.3301	0.1775
CONST-SJC	0.1175	0.1046	0.1307	0.1891	0.1373	0.2268	0.2992	0.1639



http://jhml.org/ jmh233@cam.ac.uk