Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Programa de Pós-Graduação em Estatística

SYMARFIMA: Um novo modelo dinâmico condicionalmente simétrico para séries temporais com estrutura de longa dependência

Helen da Silva Costa Benaduce

Porto Alegre, Março de 2021.

Dissertação submetida por Helen da Silva Costa Benaduce[1] como requisito parcial para a obtenção do título de Mestre em Estatística pelo Programa de Pós-Graduação em Estatística da Universidade Federal do Rio Grande do Sul.

Orientador:

Dr. Guilherme Pumi (PPGEst - UFRGS)

Comissão Examinadora:

Dra. Taiane Schaedler Prass (PPGEst - UFRGS)

Dr. Francisco José de Azevedo Cysneiros (PPGE - UFPE)

Dr. Marcio Valk (PPGEst - UFRGS)

Data de Apresentação: 31 de Março de 2021

[1]Bolsista do Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Agradecimentos

Agradeço:

- Aos meus pais Marlene e Flávio, pelas oportunidades, apoio e por todo esforço para que eu pudesse seguir estudando; ao meu irmão Victor pela amizade; à minha avó Chica por ser uma grande entusiasta dos meus estudos.
- Ao meu marido Guilherme, pelo amor, respeito, companheirismo diário e também por ser meu revisor de texto.
- Aos meus tios Marinês e Mauro, pela moradia e estrutura que me disponibilizaram em Porto Alegre.
- À Carla Daniela por me ajudar de tantas formas, desde a escolha do orientador até o acolhimento em Porto Alegre.
- Ao prof. Dr. Guilherme Pumi, pela orientação, dedicação, paciência e por ter compartilhado tanto conhecimento.
- À profa. Dra. Taiane, por todas as dicas, bibliografias e apontamentos tão necessários.
- Aos membros da banca, pela disponibilidade.
- À CAPES pela bolsa de estudos, indispensável para manter-me integralmente dedicada à pesquisa.
- À professores e funcionários do Programa de Pós-Graduação em Estatística da Univesidade Federal do Rio Grande do Sul

Resumo

Neste trabalho, introduzimos uma classe de modelos com distribuição condicional simétrica para dados de séries temporais com estrutura de longa dependência condicional, denominada modelo SYMARFIMA. No modelo proposto, a média condicional segue uma especificação ARFIMA(p, d, q), definida para acomodar uma estrutura de longa dependência, podendo ainda incluir um conjunto de covariáveis exógenas (aleatórias ou determinísticas) dependendo do tempo. A estimação dos parâmetros deste modelo é feita através do método de máxima verossimilhança parcial. Obtivemos condições de existência e estacionariedade para o modelo proposto. Obtivemos ainda a média incondicional, variância, estrutura de covariância e fórmulas fechadas para o vetor de escore e a matriz da informação de Fisher. Obtemos as propriedades assintóticas do estimador baseado em máxima verossimilhança parcial e estudamos testes de hipóteses, intervalos de confiança e previsão no contexto do modelo proposto. Além disso, é realizado a Simulação de Monte Carlo para estudar o comportamento do estimador em amostras finitas, bem como uma aplicação para dados reais.

Abstract

In this work we introduce a dynamical model for conditionally symmetric time series accommodating a long range dependent structure for the conditional mean. More specifically, the proposed model specify the underlying distribution of the time series, conditionally to its past, to be symmetric. The conditional mean is specified to accommodate a long range dependent structure, following an ARFIMA-like design, as well as a (possibly time dependent) set of regressors. We provide conditions for the existence and stationarity of the proposed model as well as closed formulas for its unconditional mean, variance and covariance structure. Parameter estimation is carried out via partial likelihood. The score vector and Hessian are obtained in closed forms. A finite sample monte carlo study of the proposed partial likelihood estimation is carried out and an application for a real data set is presented.

Índice

Capítulo 1

Introdução

Nos últimos anos, a literatura tratando de modelos de regressão envolvendo séries temporais não gaussianas evoluiu muito. Diversas abordagens têm sido investigadas na literatura. Uma das mais promissoras baseia-se na ideia de embutir modelos ARIMA clássicos dentro da estrutura flexível de modelos lineares generalizados (GLM)(McCullagh and Nelder, 1989). A ideia é simples: a estrutura de GLM trata do problema distribucional da série temporal permitindo ainda a presença de covariáveis, utilizadas para acomodar estruturas diversas, como heteroscedasticidade, tendências, etc.; enquanto que a estrutura de séries temporais acomoda a presença de corrrelação serial no processo. Embora simples, a abordagem é bastante geral. Esta combinação resulta na abordagem que hoje conhecemos pelo nome de *observation-driven*, nomenclatura introduzida em Cox et al. (1981). Modelos do tipo *observation-driven* são amplamente utilizados na prática na modelagem de séries temporais não-gaussianas na roupagem introduzida em Benjamin et al. (2003) e que hoje conhecemos como modelagem GARMA (ARMA generalizado). Uma das principais vantagens dos modelos GARMA, largamente explorado na literatura, é a capacidade de modelar variáveis aleatórias limitadas. Este tem sido o esteio de muitos modelos do tipo GARMA (ver, por exemplo, Kedem and Fokianos, 2002; Ferrari and Cribari-Neto, 2004; Rocha and Cribari-Neto, 2009; Bayer et al., 2017, 2018; Pumi et al., 2019, 2021, e referências ali contidas).

Uma particularidade pouco explorada dos modelos do tipo GARMA é a facilidade de incorporar características de alguma distribuição desejada no componente aleatório do modelo, como caudas pesadas. Ao mesmo tempo em que, no componente dinâmico, retêm qualquer estrutura de dependência desejada na resposta média condicional. Isso leva a uma estratégia de modelagem muito mais simples do que a abordagem usual, na qual particularidades da distribuição são inseridas em modelos lineares por meio do termo de erro. Além disso, a inferência condicional é naturalmente acomodada dentro desta estrutura, fornecendo uma ferramenta inferencial poderosa.

Esta particularidade foi explorada em Rocha and Cribari-Neto (2009) para modelar taxas e proporções, que são tipicamente assimétricas. Os autores propuseram o modelo beta autorregressivo de média móvel (βARMA), onde assume-se que a distribuição condicional da variável resposta segue uma distribuição beta e no componente sistemático permite um componente adicional autorregressivo de média móvel (ARMA), acomodando naturalmente assimetrias e também dispersões não constantes. Posteriormente, Pumi et al. (2019) introduziu o modelo βARFIMA, uma generalização do modelo βARMA considerando uma estrutura ARFIMA (autorregressivos de médias móveis fracionariamente integrado) capaz de acomodar uma estrutura de longa depêndencia no componente sistemático.

Em outra frente, mas ainda dentro da estrutura dos modelos GARMA, Maior and Cysneiros (2018) introduzem a classe de modelos SYMARMA. Tais modelos do tipo *observation-driven* são compostos de um componente aleatório que, condicionalmente aos valores anteriores do processo e um conjunto de covariáveis exógenas, possivelmente dependentes do tempo, que assume-se pertecer à classe de distribuições simétricas; e de um componente determinístico, tal como a média condicional, que assume uma estrutura semelhante ao ARMAX (Söderström and Stoica, 1989; Ljung, 1999), com erros seguindo um processo *martingale difference*. Quando a distribuição condicional é gaussiana, o modelo SYMARMA é equivalente a um modelo ARMAX gaussiano.

Semelhante à construção para distribuições da familia exponencial condicional em Benjamin et al. (2003) e, para família de distribuição beta condicional em Pumi et al. (2019), suponha que $\{Y_t\}_{t\in\mathbb{Z}}$ é um processo estocástico e assuma que a distribuição condicional de Y_t, dado o conjunto de informações passadas, denotado por $\mathscr{F}_{t-1} = \sigma\{Y_{t-1}, Y_{t-2}, \cdots, \mu_{t-1}, \mu_{t-2}, \cdots\}$, segue uma distribuição simétrica contínua, com densidade dada por

$$f(y_t|\mathscr{F}_{t-1}) = \frac{1}{\sqrt{\varphi}}\, g\biggl(\frac{(y_t-\mu_t)^2}{\varphi}\biggr), \qquad \forall y_t \in \mathbb{R}, \tag{1.1}$$

onde $\mu_t \in \mathbb{R}$ é um parâmetro de locação, $\varphi > 0$ é um parâmetro de escala e $g : (0, \infty) \to (0, \infty)$ satisfaz $\int_0^\infty u^{-\frac{1}{2}} g(u)du = 1$, conhecida como a função geradora da classe de distribuições simétricas. Quando existem, $\mathbb{E}(Y_t|\mathscr{F}_{t-1})$ é igual a μ_t e $\mathrm{Var}(Y_t|\mathscr{F}_{t-1}) = \xi\varphi$, onde $\xi > 0$ é uma constante que depende da função característica de Y_t dado $\mathscr{F}_{t-1}$.

Considere a especificação aditiva

$$\mu_t := \boldsymbol{x}_{t-1}'\boldsymbol{\beta} + \tau_t, \tag{1.2}$$

onde $\boldsymbol{x}_t \in \mathbb{R}^k$ é um conjuto k-dimensional de variáveis exógenas, e $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_k)'$ é um vetor de parâmetros desconhecidos. Em modelos do tipo GARMA, τ_t é responsável por acomodar a dependência serial no componente determinístico. Quando τ_t está ausente, temos a especificação usual de regressão de série temporais baseada em GLM (Kedem and Fokianos, 2002). Em casos não triviais, a especificação de τ_t é geralmente baseada em modelos clássicos de séries temporais como os modelos ARMA e ARFIMA, embora outras configurações menos comuns tenham sido propostas, como, por exemplo, em Pumi et al. (2021), no qual τ_t segue um processo caótico.

Maior and Cysneiros (2018) propuseram a classe de modelos SYMARMA (autoregressivo de médias móveis simétrico), que seguem uma estrutura semelhante ao GARMA (com a função de ligação sendo a identidade) incluindo um componente aleatório especificado por uma distribuição simétrica como em (1.1) e componente determinístico da forma (1.2), com um ARMA

$$\tau_t := \sum_{i=1}^{p} \phi_i \left(y_{t-i} - \boldsymbol{x}_{t-i}'\boldsymbol{\beta}\right) + \sum_{j=1}^{q} \theta_j r_{t-k}. \tag{1.3}$$

Em (1.3), $\boldsymbol{\phi} := (\phi_1, \cdots, \phi_p)'$ e $\boldsymbol{\theta} := (\theta_1, \cdots, \theta_q)'$ denotam os parâmetros autorregressivos e de média móvel, respectivamente. Embora a especificação (1.3) herde algumas das boas propriedades dos modelos ARMA, ela pode ser responsável apenas por estruturas de curta dependência para μ_t. Portanto, uma questão pertinente é: podemos melhorar o modelo SYMARMA de forma a reter suas boas propriedades, mas ao mesmo tempo permitir uma especificação de longa dependência para μ_t?

Essa pergunta é respondida afirmativamente no artigo intitulado "*SYMARFIMA: a Dynamical Model for Conditionally Symmetric Time Series with Long Range Dependence Mean Structure*", de autoria de Helen S.C. Benaduce e Guilherme Pumi, que encontra-se no Anexo 1 deste documento, a ser submetido para Retirar o nome da revista: a revista Statistical Modelling. O artigo é o principal produto de avaliação desta dissertação. A seguir, faremos um breve panorama do que apresentamos no artigo.

Objetivo

O artigo tem por principal objetivo propor e estudar em detalhes uma generalização para o modelo SYMARMA (Maior and Cysneiros, 2018) considerando uma estrutura ARFIMA no componente sistemático com intuito de permitir uma estrutura de longa dependência para a média condicional.

Novidades do artigo

Além da generalização através do componente sistemático, estendemos a abordagem utilizada na estimação dos parâmetros do modelo SYMARMA, feita via verossimilhança condicional, pela abordagem mais geral e flexível da verossimilhança parcial. Obtivemos as condições para a existência do modelo (que, indiretamente, apresenta condições para existência do SYMARMA), além de condições para a estacionariedade e os cálculos do primeiro e segundo momentos incondicionais. Estabelecemos condições para a consistência e normalidade assintótica da método da máxima verossimilhança parcial no contexto de modelos SYMARFIMA, estudamos testes de hipóteses, intervalos de confiança e previsão out-of-sample para o modelo proposto.

Suporte computacional

Na parte computacional, apresentamos a simulação de diferentes cenários de ajuste do modelo e por fim uma aplicação à dados reais. Todas as avaliações numéricas apresentadas no artigo foram implementadas em R (R Core Team, 2020), versão 4.0.3.

A organização desta dissertação é composta por uma breve revisão bibliográfica dos artigos fundamentais para a construção do nosso método, na qual introduzimos algumas ideia necessárias para o entendimento do trabalho; seguido por algumas conclusões e alternativas para possíveis trabalhos subsequentes e finalmente o artigo, sendo o produto principal desta dissertação que encontra-se no anexo.

Capítulo 2

Revisão Bibliográfica

Neste capítulo, apresentamos um breve resumo dos conceitos necessários para o estudo bem como de modelos que serviram de inspiração para o presente trabalho.

Modelos GARMA

A abordagem hoje conhecida como *observation-driven*, como apresentadas em Cox et al. (1981), pode ser explicada em termos simples. Suponha que μ_t denota um parâmetro que varia no tempo, observado no tempo t e seja Y_t uma variável aleatória observada no tempo t. Defina $\mathscr{F}_{t-1} := \sigma\{Y_{t-1}, Y_{t-2}, \cdots\}$ sendo a σ-álgebra gerada pelo passado observado até o tempo $t-1$. Em um modelo *observation-driven*, assumimos que

$$\mu_t = H(\mathscr{F}_{t-1}, \eta_t),$$

para alguma função mensurável H e uma sequência de variáveis aleatórias η_t que podem ser $\mathscr{F}_{t-1}$-mensuráveis ou exógenas. Modelos *observation-driven* naturalmente permitem inferência condicional mesmo quando estruturas complexas (representadas por H) são utilizadas. Com essa abordagem os modelos tipo *observation-driven* facilitam também a comparação e o diagnóstico de modelos.

A classe de modelos GARMA introduzidas em Benjamin et al. (2003) é do tipo *observation-driven*, a distribuição condicional é dependente no tempo t, dado o conjunto de informações passadas $\mathscr{F}_{t-1}$ e é modelada por alguma distribuição arbitrária, sendo que Benjamin et al. (2003) considera apenas distribuições da família exponencial. O pilar desta teoria é análogo aos clássicos modelos GLM, nos quais consideramos alguma distribuição de interesse para a variável resposta dado o passado, constituído por um componente aleatório,

$$Y_t|\mathscr{F}_{t-1} \sim f_t(\cdot; \boldsymbol{\nu}),$$

onde $\boldsymbol{\nu}$ é um parâmetro identificável da distribuição. O próximo passo é especificar uma estrutura para alguma quantidade condicional de interesse, como no caso clássico da média condicional $\mu_t := \mathbb{E}(Y_t|\mathscr{F}_{t-1})$, dependendo linearmente de um conjunto de covariáveis $\boldsymbol{x}_t$ e de um termo adicional, τ_t, responsável por uma possível dependência serial. Essa especificação constitui o componente sistemático, que pode ser escrito de forma simples via

$$g(\mu_t) = \boldsymbol{x}_t'\boldsymbol{\beta} + \tau_t.$$

Assim como nas abordagens GLM tradicionais, o preditor linear está relacionado a μ_t através de uma função de ligação, sendo a logit, probit e log-log complementar algumas das mais comumente utilizadas na literatura. O problema agora se torna a especificação de τ_t que melhor se adapta a aplicação de interesse ou que oferece a maior flexibilidade possível. Escolhas tradicionais são os clássicos modelos ARMA(p, q) e ARFIMA(p, d, q), embora hajam exemplos de estruturas mais específicas.

Modelos βARMA, KARMA e βARFIMA

O modelo introduzido em Rocha and Cribari-Neto (2009), construído sobre a classe de regressão beta descrita em Ferrari and Cribari-Neto (2004), propõe um modelo dinâmico para séries temporais limitadas no intervalo $(0, 1)$. No modelo denominado βARMA, que segue a estrutura do tipo GARMA, o componente aleatório é especificado por uma distribuição beta. Seu preditor linear, assim como em Benjamin et al. (2003), modela a média condicional através de uma função de ligação,

$$g(\mu_t) = \boldsymbol{x}_t'\boldsymbol{\beta} + \tau_t,$$

onde $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_l)'$ são os coeficientes relacionados às covariáveis. Por sua vez τ_t segue uma especificação ARMA(p, q), do tipo

$$\tau_t = \sum_{i=1}^{p} \phi_i \left(g(Y_{t-i}) - \boldsymbol{x}_{t-i}'\boldsymbol{\beta}\right) + \sum_{j=1}^{q} \theta_j r_{t-k},$$

onde p e q e $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_p)'$ e $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_p)'$ denotam a ordem e os coeficientes autorregressivos e de médias móveis, respectivamente.

Uma alternativa aos modelos βARMA é apresentado em Bayer et al. (2017), no qual os autores introduzem uma classe de modelos dinâmicos para séries temporais duplamente limitadas com distribuição condicional seguindo a distribuição Kumaraswamy, modelo entitulado KARMA (Kumaraswamy autorregressivo de médias móveis). Em construção semelhante a de Benjamin et al. (2003) para a variável dependente, assume-se ter uma distribuição Kumaraswamy condicional dado o conjunto de informação passado do processo. Nos modelos KARMA a mediana condicional é modelada por uma estrutura dinâmica contendo termos ARMA. Além da distribuição, uma distinção notável entre os modelos βARMA e KARMA é a estrutura contida no preditor linear: ao contrário do βARMA, no KARMA modela-se a mediana condicional. Isso mostra o quão flexível pode ser a abordagem do tipo GARMA e a vantagem em aplicá-la. Uma desvantagem em trabalhar com a distribuição Kumaraswamy é que, por não pertencer à família exponencial, a teoria assintótica para modelos do tipo GARMA que envolvem a distribuição Kumaraswamy se torna mais difícil de estabelecer. Um exemplo emblemático é apresentado em Pumi et al. (2020).

Em Pumi et al. (2019) propõe-se uma generalização do βARMA (Rocha and Cribari-Neto, 2009), considerando-se uma estrutura ARFIMA(p, d, q) para τ_t no formato

$$\tau_t = \sum_{i=1}^{p} \phi_i\big(g(Y_{t-i}) - \boldsymbol{x}_{t-i}'\boldsymbol{\beta}\big) + \sum_{k=1}^{\infty} c_k r_{t-k}, \tag{2.1}$$

onde os coeficientes c_k dependem dos parâmetros d e $\boldsymbol{\theta}$. Mais detalhes serão dados no artigo que é o produto final desta dissertação.

Modelos SYMARMA

Em Maior and Cysneiros (2018), os autores propuseram estender a classe de modelos ARMA gaussianos tradicionais para uma classe de modelos dinâmicos para séries temporais com distribuição simétrica condicional. A ideia foi seguir o enfoque dos modelos GARMA e βARMA para construir a classe dos modelos SYMARMA.

A abordagem do SYMARMA permite que uma distribuição simétrica qualquer seja utilizada na estrutura do componente aleatório. A motivação explora acomodar melhor observações atípicas, uma vez que a classe de distribuições simétrica contém distribuições com caudas mais pesadas ou mais leves do que a distribuição normal.

O modelo SYMARMA é definido pela componente aleatória (1.1) e pela componente dinâmica (1.2) com a especificação (1.3). Observe que $\mathrm{Var}(Y_t|\mathscr{F}_{t-1})$ depende de alguma distribuição de interesse através do ξ. Em Cysneiros et al. (2005) encontramos os valores de ξ para algumas distribuições, veja na Tabela 2.1 a seguir,

Tabela 2.1: Valor de ξ para algumas distribuições simétricas

Distribuição	ξ
Normal	1
t -Student	$\frac{\nu}{\nu-2}, \quad \nu > 2$
t-Student generalizada	$\frac{s}{r-2}, \quad s > 0, r > 2$
Logística-I	0,79569
Logística-II	$\pi^2/3$
Logística generalizada	$2\psi'(m)$
Exponencial potência	$2^{(1+k)}\frac{\Gamma(3[k+1]/2)}{\Gamma([k+1]/2)}$

A parametrização utilizada para a distribução t-Student, por exemplo, é dada da seguinte forma

$$f(x|\nu,\mu,\sigma) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma}\left(1+\frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}},$$

onde $\sigma := \sqrt{\varphi\xi}$ e ν são os graus de liberdade.

Embora apresente boas propriedades, o modelo SYMARMA só permite acomodar uma estrutura de curta dependência para μ_t, que pode não ser suficiente em algumas situações. Em virtude disso, a ideia é substituir a especificação ARMA(p,q) (1.3) por uma especificação ARFIMA(p,d,q) análoga à (2.1), a saber

$$\mu_t = \alpha + \boldsymbol{x}_t'\boldsymbol{\beta} + \sum_{j=1}^{p}\phi_j(y_{t-j} - \boldsymbol{x}_{t-j}'\boldsymbol{\beta}) + \sum_{k=1}^{\infty}c_k r_{t-k}. \tag{2.2}$$

O modelo proposto é especificado via (1.1) e (2.2). A diferença entre esta especificação e a especificação ARMA é a presença da soma infinita em (2.2) onde os coeficientes c_k carregam a informação relativa à parte média móvel e também à estrutura de longa dependência, parametrizada por d. Mais detalhes são dados no artigo anexo.

Máxima verossimilhança parcial

A estimação dos parâmetros no modelo SYMARMA é feita através de máxima verossimilhança condicional (CMLE), o que implica que apenas covariáveis determinísticas podem ser consideradas no modelo. Optamos pelo estimador de máxima verossimilhança parcial (PMLE) para os parâmetros no modelo SYMARFIMA e com isso obtemos algumas vantagens, como a liberdade de admitir diferentes comportamentos para a variável exógena $\boldsymbol{x}_t$.

Em Fokianos and Kedem (2004) temos uma abordagem que une os conceitos inferenciais de verossimilhança parcial (PL) e modelos lineares generalizados (GLM). Outra vantagem de utilizarmos PL é que não há a necessidade de modelar a distribuição conjunta da variável resposta e das covariáveis. Assim, a PL é baseada inteiramente na distribuição condicional, dado o conjunto de informações anteriores e das covariáveis anteriores, tipicamente reprensentada pela σ-álgebra gerada por valores passados. Estas σ-álgebras formam uma filtração natural no tempo, permitindo que o modelo incorpore covariaveis exógenas aleatórias, possivelmente com estrutura de dependência temporal, determinísticas, ou mistura entre eles. Assintóticamente falando, como a função de verossimilhança parcial é obtida em relação a uma filtração natural ao longo do tempo, isso permite-nos explorar a teoria de martingales para obter resultados assintóticos. Detalhes da estimação, avaliação do modelo, previsão e teoria assintótica baseados em PL no contexto do modelo proposto são apresentados no artigo anexo.

Capítulo 3

Conclusões e trabalhos futuros

Nesta dissertação, estendemos o modelo autorregressivo de médias móveis simétrico (SYMARMA). Pontualmente, propomos a substituição da especificação do tipo ARMAX por uma do tipo ARFIMAX na estrutura de μ_t. O modelo SYMARFIMA, assumindo que a distribuição condicional da variável resposta é alguma distribuição simétrica de interesse, comporta uma estrututa de longa dependência e, ainda, admite que variáveis exógenas possam ser determinísticas, estocásticas ou até mesmo séries temporais. Obtemos condições para a existência e estacionariedade do modelo SYMARFIMA e os cálculos do primeiro e segundo momentos incondicionais. Via máxima verossimilhança parcial apresentamos o procedimento de amostra finita para estimação do processo, derivamos a teoria assintótica, construímos testes de hipótese e intervalos de confiança dos estimadores. Construímos ainda o vetor score e derivamos a matriz de informação de Fisher condicional. Além disso, apresentamos em termos numéricos o procedimento para a realização da previsão.

Potenciais trabalhos futuros envolvem melhoria inferencial em pequenas amostras, estudo da robustez do modelo bem como podemos explorar aplicações do modelo SYMARFIMA.

Referências Bibliográficas

Bayer, F. M., Bayer, D. M., Pumi, G., 2017. Kumaraswamy autoregressive moving average models for double bounded environmental data. Journal of Hydrology 555, 385–396.

Bayer, F. M., Cintra, R. J., Cribari-Neto, F., 2018. Beta seasonal autoregressive moving average models. Journal of Statistical Computation and Simulation 88 (15), 2961–2981.

Benjamin, M. A., Rigby, R. A., Stasinopoulos, D. M., 2003. Generalized autoregressive moving average models. Journal of the American Statistical association 98 (461), 214–223.

Cox, D. R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., Lauritzen, S. L., 1981. Statistical analysis of time series: Some recent developments [with discussion and reply]. Scandinavian Journal of Statistics, 93–115.

Cysneiros, F. J. A., Paula, G. A., Galea, M., 2005. Modelos simétricos aplicados. Escola de Modelos de Regressão 9.

Ferrari, S. L. P., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. Journal of Applied Statistics 31 (7), 799–815.

Fokianos, K., Kedem, B., 2004. Partial likelihood inference for time series following generalized linear models. Journal of Time Series Analysis 25 (2), 173–197.

Kedem, B., Fokianos, K., 2002. Regression models for time series analysis. John Wiley & Sons.

Ljung, L., 1999. System Identification - Theory for the User. Prentice Hall.

Maior, V., Cysneiros, F., 02 2018. SYMARMA: a new dynamic model for temporal data on conditional symmetric distribution. Statistical Papers 59.

McCullagh, P., Nelder, J. A., 1989. Generalized Linear Models. Chapman & Hall / CRC, London.

Pumi, G., Prass, T. S., Souza, R. R. a., 2021. A dynamic model for double-bounded time series with chaotic-driven conditional averages. Scandinavian Journal of Statistics 48 (1), 68–86.

Pumi, G., Rauber, C., Bayer, F. M., 2020. Kumaraswamy regression model with Aranda-Ordaz link function. TEST 29, 1051 – 1071.

Pumi, G., Valk, M., Bisognin, C., Bayer, F. M., Prass, T. S., 2019. Beta autoregressive fractionally integrated moving average models. Journal of Statistical Planning and Inference 200, 196 – 212.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL `https://www.R-project.org/`

Rocha, A. V., Cribari-Neto, F., 2009. Beta autoregressive moving average models. Test 18 (3), 529–545.

Söderström, T., Stoica, P., 1989. System Identification. Prentice Hall International.

Anexo A

Artigo Benaduce e Pumi

SYMARFIMA: a dynamical model for conditionally symmetric time series with long range dependence mean structure

Helen da Silva Costa Benaduce [1], **Guilherme Pumi** [1],

[1] Programa de Pós-Graduação em Estatística, Statistics and Mathematics Department, Universidade Federal do Rio Grande do Sul, Brasil

Address for correspondence: Guilherme Pumi, Statistics and Mathematics Department, Universidade Federal do Rio Grande do Sul – 9500, Bento Gonçalves Avenue, Porto Alegre, Brasil.
E-mail: `guilherme.pumi@ufrgs.br`.
Phone: .
Fax: .

Abstract: In this work we introduce a dynamical model for conditionally symmetric time series accommodating a long range dependent structure for the conditional mean. More specifically, the proposed model specify the underlying distribution of the time series, conditionally to its past, to be symmetric. The conditional mean is specified to accommodate a long range dependent structure, following an ARFIMA-like design, as well as a (possibly time dependent) set of regressors. We provide conditions for the existence and stationarity of the proposed model as well as closed formulas for its unconditional mean, variance and covariance structure. Parameter estimation is carried out via partial likelihood. The score vector and Hessian are obtained in closed forms. A finite sample study of the proposed partial likelihood estimation is carried out. To show the usefulness of the proposed model, we present an application to a time series related to operations with clients in the financial sector.

Key words: dynamic models; long range dependent processes; symmetric distribution; time series analysis

1 Introduction

In recent years, there has been an increased interest in non-gaussian time series. Several approaches have been investigated in the literature. One of the most promising approaches is based on the idea of embedding classical ARIMA models within the flexible structure of generalized linear models. This approach yields a solid foundation for dynamical regression modeling of dependent data and is often called GARMA (generalized ARMA) models. GARMA models can be categorized as observation-driven models (Cox et al., 1981) and have been very successful in modeling non-gaussian time series. The main idea behind the approach dates back to the late 70's, but the name GARMA was firstly introduced in Benjamin et al. (2003). One of the main advantages of GARMA models that have been extensively explored in the literature is their ability to model bounded random variables. This has been the cornerstone of many GARMA-like models (see, for instance, Kedem and Fokianos, 2002; Ferrari and Cribari-Neto, 2004; Rocha and Cribari-Neto, 2009; Bayer et al., 2017, 2018; Pumi et al., 2019, 2021, and references therein).

A less explored facet of GARMA-like models is the ease to embed desired distributional features, such as heavy-tailed marginals, into the model's random component, while retaining any desired dependence structure on the conditional mean response. This leads to a much simpler modeling strategy than the usual approach of inserting distributional features through the error term in linear models. Furthermore, conditional inference is naturally accommodated within the framework, providing a powerful inferential tool.

This feature is explored in Rocha and Cribari-Neto (2009) to model rates and proportions, which are typically asymmetric. The authors proposed the beta autoregressive moving average model (βARMA), where it assumed that the conditional distribution of the response variable follows a beta distribution and in the systematic component it allows an additional autoregressive moving average component (ARMA), accommodating naturally asymmetries and also non-constant dispersions. Subsequently, in Pumi et al. (2019) the generalization of the βARMA model was sought to accommodate a long-dependent structure, with the modification of the additional component written in autoregressive terms and fractional moving averages (ARFIMA). The authors presented the autoregressive beta model and fractional moving averages (βARFIMA), models capable of detecting the long-range dynamics in the data, proving to be more adequate than the βARMA models in such cases.

Following the framework of GARMA models, Maior and Cysneiros (2018) introduces the class of SYMARMA models, composed of observational-driven models for which, conditionally on the past values of the process and a set of possibly dependent exogenous covariates, the random component is assumed to belong to the class of symmetrical distribution while the conditional mean assumed to follow an ARMAX-like structure (Söderström and Stoica, 1989; Ljung, 1999), with martingale difference errors. When the conditional distribution is Gaussian, the SYMARMA model is equivalent to a Gaussian ARMAX model.

Let $\{Y_t\}_{t\in\mathbb{Z}}$, be a stochastic process and assume that the conditional distribution of Y_t, given the past information, denoted by $\mathscr{F}_{t-1} = \sigma\{Y_{t-1}, Y_{t-2}, \cdots, \mu_{t-1}, \mu_{t-2}, \cdots\}$, follows a continuous symmetric distribution, with density given by

$$f(y_t; \mu_t, \varphi|\mathscr{F}_{t-1}) = \frac{1}{\sqrt{\varphi}}\, g\bigg(\frac{(y_t-\mu_t)^2}{\varphi}\bigg), \qquad \forall y_t \in \mathbb{R}, \tag{1.1}$$

where $\mu_t \in \mathbb{R}$ is a location parameter, $\varphi > 0$ is a scale parameter and $g : (0,\infty) \to (0,\infty)$ satisfy $\int_0^\infty u^{-\frac{1}{2}} g(u)du = 1$, known as the generating function from the class of symmetric distributions. When they exist, $\mathbb{E}(Y_t|\mathscr{F}_{t-1}) = \mu_t$ and $\mathrm{Var}(Y_t|\mathscr{F}_{t-1}) = \xi\varphi$, where $\xi > 0$ is a constant that depends on the characteristic function of Y_t given $\mathscr{F}_{t-1}$. If $\mathbb{E}(Y_t|\mathscr{F}_{t-1})$ does not exist, then μ_t is the symmetry point of f. Consider the additive specification

$$\mu_t := \boldsymbol{x}_t'\boldsymbol{\beta} + \tau_t, \tag{1.2}$$

where $\boldsymbol{x}_t \in \mathbb{R}^k$ is a k-dimensional set of exogenous variables, and $\boldsymbol{\beta} = (\beta_0, \cdots, \beta_k)'$ is a vector of unknown parameters. In observation-driven GARMA-like models, τ_t is responsible to accommodate serial dependence in the deterministic component. When τ_t is absent, we have the usual specification of GLM-based time series regression. In non-trivial cases, the specification of τ_t is usually based on classical time series models such as ARMA and ARFIMA models, although other less common configurations have been proposed, as, for instance, in Pumi et al. (2021), where τ_t follows a chaotic process.

Maior and Cysneiros (2018) proposed the class of SYMARMA (autoregressive and moving average symmetric) models, which follow a GARMA-like structure (with identity link) including a random component specified by a symmetric distribution as in (1.1) and deterministic component of the form (1.2), with an ARMA(p, q) stipulation for τ_t, namely

$$\tau_t := \sum_{i=1}^{p} \phi_i \left(y_{t-i} - \boldsymbol{x}_{t-i}'\boldsymbol{\beta}\right) + \sum_{j=1}^{q} \theta_j r_{t-k}. \tag{1.3}$$

In (1.3), $\boldsymbol{\phi} := (\phi_1, \cdots, \phi_p)'$ and $\boldsymbol{\theta} := (\theta_1, \cdots, \theta_q)'$, are autoregressive and moving average parameters, respectively. Although specification (1.3) inherits some of the nice properties of ARMA models, it can only account for short-range dependent structures for μ_t. Hence a natural question is: can we improve the SYMARMA model in such a way that it still retains its good properties, but allows a long-range dependent specification for μ_t? In this work we provide a positive answer to this question. The goal is to prescribe an ARFIMA(p,d,q) specification for τ_t, allowing for long range dependence in the conditional mean μ_t. We show that such generalization exists and is well-defined, derive conditions for stationarity, and its unconditional mean and covariance structure. Parameter estimation in the context of SYMARMA is carried on via conditional maximum likelihood approach. As result, only deterministic exogenous variables are allowed in the model. We propose the estimation via partial maximum likelihood (PMLE), which allows for the covariates to be deterministic, stochastic, or even time series. The score vector and Hessian matrix are derived and presented in closed form. The finite sample performance of the proposed parameter estimation is verified through a Monte Carlo simulation study.

The paper is organized as follows. In the next section we introduce the proposed model. In Section 3, we derive some theoretical properties for the proposed model, while in Section 4 we discuss inference via partial maximum likelihood. In Section 5 we present the asymptotic theory, hypotheses testing, confidence intervals and forecasting. In Section 6 we present a Monte Carlo simulation study to assess the finite sample performance of the PMLE in the proposed model. In Section 7 we apply our model to a real life time series related to operations with clients in the financial sector. Conclusions are presented in Section 8.

2 Proposed model

The most widely applied model for time series presenting long-range dependence is the class of ARFIMA models, introduced by Granger and Joyeux (1980) and Hosking (1981). Recall that a process $\{Z_t\}_{t\in\mathbb{Z}}$ is called an ARFIMA(p,d,q) if it is a weakly stationary solution of the set of difference equations

$$\phi(L)(1-L)^d Z_t = \theta(L)\varepsilon_t, \tag{2.1}$$

where L denotes the backward shift operator $L^k(Z_t) = Z_{t-k}$, for $k \in \mathbb{N}$, ε_t is an error term (usually taken as a white noise), $\phi(z)$ and $\theta(z)$ denote the AR and MA polynomials given respectively by

$$\phi(z) := -\sum_{i=0}^{p} \phi_i z^i, \quad \theta(z) := \sum_{j=0}^{q} \theta_j z^j, \quad \forall z \in \mathbb{C}, \tag{2.2}$$

where $\phi_0 = -1$ and $\theta_0 = 1$. We assume that the polynomials $\phi(z)$ and $\theta(z)$ present no common roots. The fractional term $(1-L)^d$ is defined by its binomial expansion,

$$(1-L)^{-d} = \sum_{k=0}^{\infty} \pi_k L^k, \quad \text{where} \quad \pi_k = \frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)} = \prod_{j=1}^{k} \frac{j-1+d}{j}, \ k \geq 1, \tag{2.3}$$

and $\pi_0 = 1$. In this work we consider $d \in (-1, 0.5)$, which imply that the resulting ARFIMA(p, d, q) is weakly stationary, provided that the polynomial $\phi(z)$ does not present roots in the unitary disk $\{z \in \mathbb{C} : |z| = 1\}$. Observe that when $d = 0$, (2.1) reduces to an ARMA(p, q) model. More details on the theory of ARFIMA processes can be found in Brockwell and Davis (1991) and Palma (2007).

Inspired by Maior and Cysneiros (2018) and Pumi et al. (2019), we propose a generalization of the SYMARMA model, which we call symmetric autoregressive fractionally integrated moving average models (SYMARFIMA, for short) by allowing τ_t in (1.2) to follow an ARFIMA(p, d, q)-like structure, with exogenous, possibly time dependent, random covariates. To motivate the specification of the proposed model, let $\{\boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ be a set of r-dimensional covariates (random or deterministic) and suppose that $\{Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}\}_{t\in\mathbb{Z}}$ satisfies an ARFIMA(p, d, q) specification of the form

$$\phi(L)(1 - L)^d \left(Y_t - [\omega + \boldsymbol{x}_t'\boldsymbol{\beta}]\right) = \theta(L)r_t, \quad |d| < 0.5., \tag{2.4}$$

where ω is a constant and $\boldsymbol{\beta} \in \mathbb{R}^r$ is a vector of unknown parameters associated with the covariates. Assume that the following conditions, endemic to the discussion of ARMA and ARFIMA, hold:

(a) The polynomials $\phi(z)$ and $\theta(z)$ have no common zeros;

(b) The polynomials $\phi(z)$ and $\theta(z)$ not have zeros in the closed unit disk.

Under (a) and (b), since $|d| < 0.5$, we can write

$$\phi(L)\left(Y_t - [\omega + \boldsymbol{x}_t'\boldsymbol{\beta}]\right) = (1 - L)^{-d}\theta(L)r_t. \tag{2.5}$$

Expanding $(1 - z)^{-d}\theta(z)$ in its Laurent's expansion, we obtain

$$(1 - L)^{-d}\theta(L)r_t = \sum_{k=0}^{\infty} c_k r_{t-k},$$

where

$$c_0 = 1 \quad \text{and} \quad c_k = \sum_{i=0}^{\min\{k,q\}} \theta_i \pi_{k-i}, \quad k > 0, \tag{2.6}$$

with π_k given in (2.3). Now, (2.5) can be rewritten as

$$\phi(L)(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}) = \alpha + \sum_{k=0}^{\infty} c_k r_{t-k} = \alpha + r_t + \sum_{k=1}^{\infty} c_k r_{t-k}.$$

where $\alpha = \phi(1)\omega$. If we define $\mu_t := Y_t - r_t$, then we can write that $\phi(L)(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}) = Y_t - \boldsymbol{x}_t'\boldsymbol{\beta} - \sum_{j=1}^{p} \phi_j(y_{t-j} - \boldsymbol{x}_{t-j}'\boldsymbol{\beta})$, to obtain that

$$\mu_t = \alpha + \boldsymbol{x}_t'\boldsymbol{\beta} - \sum_{j=1}^{p} \phi_j(y_{t-j} - \boldsymbol{x}_{t-j}'\boldsymbol{\beta}) + \sum_{k=1}^{\infty} c_k r_{t-k}. \tag{2.7}$$

The proposed SYMARFIMA(p, d, q) is specified by (1.1) and (2.7).

There are important differences between SYMARMA and SYMARFIMA models. In our framework, we allow the covariates in a SYMARFIMA model to be either random or deterministic (or a combination of both) and possibly time dependent. This flexibility is attained by considering a partial maximum likelihood estimation (PMLE) approach, as in Pumi et al. (2019) instead of the conditional maximum likelihood estimation (CMLE) applied in Maior and Cysneiros (2018). The PMLE approach allows for temporal conditional inference with respect to a filtration generated by the observer's knowledge at the time of observation. Working with such a filtration allows for the presence of possibly time dependent random covariates as well as deterministic ones (or any combination of them). From a computational point of view, there is no distinction between CMLE and PMLE. The asymptotic theory for the PMLE requires somewhat different assumptions than the CMLE and its proof require some specialization as well. For more details as well as an account of the asymptotic theory related to PMLE, see Fokianos and Kedem (2004) and references therein.

Observe that one main distinction between ARFIMA and SYMARFIMA models resides in the definition of the error term. In classical ARFIMA models, the process is defined unconditionally and it is driven by an exogenous white noise process while in the SYMARFIMA, the error term is recursively defined, in a residual fashion. Notice that, for all $t \in \mathbb{Z}$, r_t is $\mathscr{F}_{t-1}$-measurable and $\mathbb{E}(r_t|\mathscr{F}_{t-1}) = 0$, hence $\{(r_t, \mathscr{F}_{t-1})\}_{t\in\mathbb{Z}}$ forms a martingale difference sequence. Another small distinction between SYMARMA and SYMARFIMA models is the presence of α in (2.7), which could be included in $\boldsymbol{x}_t$ term as in (1.2). Explicitly naming the constant, however, helps simplifying the notation in the rest of the paper, but notice that α is not included in the autoregressive term. Some particular cases are easily spotted: a SYMARFIMA$(p, 0, q)$ with $\alpha = 0$ is equivalent to a SYMARMA(p, q) with the same parameters, while when (1.1) is Gaussian, then it is easy to show that the SYMARFIMA(p, d, q) is equivalent to a Gaussian ARFIMA(p, d, q).

3 Properties

In this section we examine some theoretical properties of the proposed SYMARFIMA process. We start by obtaining conditions for the existence of a SYMARFIMA process, which is related to the problem of when the decomposition

$$Y_t - \boldsymbol{x}_t'\boldsymbol{\beta} = \frac{\alpha}{\phi(1)} + \sum_{k=0}^{\infty} \lambda_k r_{t-k} \quad \forall t \in \mathbb{Z}, \tag{3.1}$$

holds where the λ_k's are the coefficients in the Laurent's expansion of $(1-z)^{-d}\phi(z)^{-1}\theta(z)$. According to Bondon and Palma (2007), under (a) and (b), the right hand side of (3.2) is L^2-convergent if, and only if, $d \in (-1, 1/2)$.

Since the family of distribution satisfying (1.1) is very broad, we need to split the study into two cases according to whether the distribution has heavy tails or not, which directly affects the existence of moments. Observe that $\mathbb{E}(Y_t|\mathscr{F}_{t-1})$ might not exist and, even if it does, $\mathbb{E}(Y_t)$ may not. In that case, μ_t is the symmetry point of (1.1) which is assumed to evolve over time according to (2.7). Specifying $r_t = y_t - \mu_t$ implies that r_t is an $\mathscr{F}_{t-1}$-mensurable error term and if $\mathbb{E}(Y_t|\mathscr{F}_{t-1})$ exists, $\{(r_t, \mathscr{F}_{t-1})\}_{t\in\mathbb{Z}}$ is a martingale difference. As a result, in this case $\{r_t\}_{t\in\mathbb{Z}}$ is (unconditionally) a symmetric and centered uncorrelated process with constant variance. In the next theorem we consider the case where Y_t has finite second moment.

Theorem 3.1 *If $\{Y_t, \boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ is a SYMARFIMA(p, d, q) process satisfying (a) and (b) and such that $\mathbb{E}(Y_t^2) < \infty$, then $\{Y_t\}_{t\in\mathbb{Z}}$ is well defined and (3.2) converges almost surely if, and only if $d < 0.5$.*

Proof: The proof follows the same lines as the proof of lemma 1.1 in Lopes and Prass (2013), with the necessary adaptation from the i.i.d. case. Suppose that $d < 0.5$. Under assumptions (a) and (b), the coefficients in (2.5) satisfy $\lambda_k = \frac{\phi(1)}{\theta(1)\Gamma(d)} k^{d-1} + O(k^{d-1})$, and $\sum_{k=0}^{\infty} \lambda_k^2 < \infty$. Upon noticing that the pair $\{r_t, \mathscr{F}_{t-1}\}_{t\in\mathbb{Z}}$ is a martingale difference, the result follows by the martingale convergence theorem (Hall and Heyde, 1980, theorem 2.5), and we conclude that (3.2) is well defined and converges almost surely to a random variable with finite moment. The proof of the converse follows the same argument as proof of lemma 1.1 in Lopes and Prass (2013) in view of the three series criterion for martingales (Hall and Heyde, 1980, theorem 26). ∎

A limitation of Theorem 3.2 is that it requires $\mathbb{E}(Y_t^2) < \infty$, which may not be fulfilled when (1.1) present heavy tails. The next theorem relaxes the moment conditions for existence, by imposing a restriction on the distribution's tail decay.

Theorem 3.2 *If $\{Y_t, \boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ is a SYMARFIMA(p,d,q) process satisfying (a) and (b) and such that*

$$P(|Y_t| > h|\mathscr{F}_{t-1}) \sim \kappa_0 h^{-a} \tag{3.2}$$

for some $a \in [1,2)$ and $\kappa_0 > 0$. Then $\{Y_t\}_{t\in\mathbb{Z}}$ is well defined and converges almost surely if, and only if, $(1-d)a > 1$. Furthermore, if $(1-d)a > 1$, then the series converges in L^s for all $0 < s < a$.

Proof: The proof is an adaptation of the proof in Prass (2013). First assume that $(1-d)a > 1$. Since $a \in [1,2)$ there exists $1 < \nu_0 < a$ such that $(1-d)a > (1-d)\nu > 1$ for all $\nu_0 \le \nu < a$. Since $\lambda_k = \frac{\phi(1)}{\theta(1)\Gamma(d)}k^{d-1} + O(k^{d-1})$ in (3.2), thus $\sum_{k=0}^{\infty}|\lambda_k|^r < \infty$. By Lemma 1 in Avram and Taqqu (1986), under the hypothesis,

$$\mathbb{E}\left(\left|\sum_{k=0}^{\infty}\lambda_k r_{t-k}\right|^{\nu}\right) \le 2\mathbb{E}(|r_t|^{\nu})\sum_{k=0}^{\infty}|\lambda_k|^r < \infty, \quad \text{for any} \quad 1 < \nu_0 \le \nu < a.$$

For $0 < s < \nu_0$ the proof is consequence of the inequality $\mathbb{E}(|X|^r)^{\frac{1}{r}} \le \mathbb{E}(|X|^s)^{\frac{1}{s}}$. Hence, (3.2) converges almost surely. Conversely, if (3.2) converges almost surely, by the 3 series theorem for martingale differences (Hall and Heyde, 1980, theorem 26), $\sum_{k=0}^{\infty} P(|\lambda_k r_{t-k}| \ge h|\mathscr{F}_{t-1})$ converges almost surely for all $h > 0$. The tail condition in (3.2) implies that

$$P(|r_{t-k}| \ge h|\lambda_k|^{-1}|\mathscr{F}_{t-1}) \sim \kappa_0 h^{-a}|\lambda_k|^a \sim \kappa_0 h^{-a} O(k^{d-1+a}) \sim \kappa_1 k^{(d-1)a}, \tag{3.3}$$

for all t. Hence, since $\sum_{k=0}^{\infty} P(|\lambda_k r_{t-k}| \ge h|\mathscr{F}_{t-1})$ converges, (3.3) implies that necessarily $a(1-d) < 1$. ∎

Observe that when (1.1) present heavy tails, the range of d for which the related SYMARFIMA(p,d,q) exists depend on the tails of the distribution.

Theorem 3.3 *Let $\{Y_t\}_{t\in\mathbb{Z}}$ be a SYMARFIMA(p,d,q) satisfying (a) and (b), for $d \in (-1, 0.5)$ and $\boldsymbol{x}_t := (x_{t1}, \cdots, x_{tr})'$. If $\mathbb{E}(x_{tk}) < \infty$, for all $t \in \mathbb{Z}$ and $k \in \{1, \cdots, r\}$ and $\mathbb{E}(Y_t)$ exists, then*

$$\mathbb{E}(Y_t) = \frac{\alpha}{\phi(1)} + \sum_{j=1}^{r}\beta_j\mathbb{E}(x_{tj}).$$

Proof: Under the hypothesis, (2.5) holds so we can write

$$\begin{aligned} Y_t &= \phi(L)^{-1}\alpha + \boldsymbol{x}_t'\boldsymbol{\beta} + \phi(L)^{-1}(1-L)^{-d}\theta(L)r_t \\ &= \frac{\alpha}{\phi(1)} + \boldsymbol{x}_t'\boldsymbol{\beta} + \psi(L)r_t. \end{aligned}$$

where $\psi(L) = \phi(L)^{-1}(1-L)^{-d}\theta(L)$. Then

$$\begin{aligned}\mathbb{E}(Y_t) &= \mathbb{E}\left(\frac{\alpha}{\phi(1)} + \boldsymbol{x}_t'\boldsymbol{\beta} + \psi(L)r_t\right) = \frac{\alpha}{\phi(1)} + \mathbb{E}(\boldsymbol{x}_t'\boldsymbol{\beta}) + \psi(L)\mathbb{E}(r_t)\\ &= \frac{\alpha}{\phi(1)} + \sum_{j=1}^{r}\beta_j\mathbb{E}(x_{tj}),\end{aligned}$$

since, under the hypothesis, $\mathbb{E}(r_t) = 0$. ∎

Theorem 3.4 *Let $\{Y_t, \boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ be a SYMARFIMA(p,d,q) satisfying (a), (b) and $\mathbb{E}(Y_t^2) < \infty$, for $d \in (-1, 0.5)$, where $\boldsymbol{x}_t := (x_{t1}, \cdots, x_{tr})'$ is a set of covariates. If the covariates are non-random, then*

$$\mathrm{Var}(Y_t) = \xi\varphi\sum_{k=0}^{\infty}\lambda_k^2.$$

If the covariates $\boldsymbol{x}_t$ are random and exogenous, with $\mathbb{E}(x_{tr}^2) < \infty$ then

$$\mathrm{Var}(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}) = \xi\varphi\sum_{k=0}^{\infty}\lambda_k^2.$$

Proof: Writing, $Y_t = \mu_t + r_t$ and since $\mathbb{E}(r_t) = 0$, we have

$$\begin{aligned}\mathrm{Var}\left(r_t\right) &= \mathbb{E}\left(r_t^2\right) = \mathbb{E}\left(\mathbb{E}\left(r_t^2|\mathscr{F}_{t-1}\right)\right) = \mathbb{E}\left(\mathrm{Var}\left(r_t|\mathscr{F}_{t-1}\right)\right) = \mathbb{E}\left(\mathrm{Var}\left(Y_t - \mu_t|\mathscr{F}_{t-1}\right)\right)\\ &= \mathbb{E}\left(\mathrm{Var}\left(Y_t|\mathscr{F}_{t-1}\right)\right) = \xi\varphi.\end{aligned}$$

Under the assumptions, $\sum_{k=1}^{\infty}\lambda_k^2 < \infty$, hence, form (3.2) we obtain

$$\mathrm{Var}\left(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}\right) = \mathrm{Var}\left(\frac{\alpha}{\phi(1)} + \sum_{k=0}^{\infty}\lambda_k r_{t-k}\right) = \xi\varphi\sum_{k=1}^{\infty}\lambda_i^2.$$

The result now follows in either condition. ∎

Theorem 3.5 *Let $\{Y_t, \boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ be a SYMARFIMA(p,d,q) satisfying (a), (b) and $\mathbb{E}(Y_t^2) < \infty$, for $d \in (-1, 0.5)$, where $\boldsymbol{x}_t := (x_{t1}, \cdots, x_{tr})'$ is a set of covariates. Then: If $\boldsymbol{x}_t$ is non-random, then*

$$\mathrm{Cov}\left(Y_t, Y_{t+h}\right) = \xi\varphi\sum_{k=0}^{\infty}\lambda_k\lambda_{k+h} \quad \textit{and} \quad \mathrm{Cor}\left(Y_t, Y_{t+h}\right) = \frac{\sum_{k=0}^{\infty}\lambda_k\lambda_{k+h}}{\sum_{i=0}^{\infty}\lambda_i^2}.$$

If $\boldsymbol{x}_t$ are random with $\mathbb{E}(x_{tr}^2) < \infty$ then

$$\mathrm{Cov}\left(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}, Y_{t+h} - \boldsymbol{x}_{t+h}'\boldsymbol{\beta}\right) = \xi\varphi\sum_{k=0}^{\infty}\lambda_k\lambda_{k+h},$$

and

$$\mathrm{Cor}\left(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}, Y_{t+h} - \boldsymbol{x}_{t+h}'\boldsymbol{\beta}\right) = \frac{\sum_{k=0}^{\infty}\lambda_k\lambda_{k+h}}{\sum_{i=0}^{\infty}\lambda_i^2}.$$

Proof: Under the assumptions, representation (3.2) holds so that

$$\mathrm{Cov}\left(Y_t - \boldsymbol{x}_t'\boldsymbol{\beta}, Y_{t+h} - \boldsymbol{x}_{t+h}'\boldsymbol{\beta}\right) = \mathrm{Cov}\left(\sum_{k=0}^{\infty}\lambda_k r_{t-i}, \sum_{i=0}^{\infty}\lambda_i r_{t+h-i}\right) = \xi\varphi\sum_{i=0}^{\infty}\lambda_i\lambda_{i+k}.$$

All results now follow trivially. ∎

Theorems 3.3 through 3.5 allow us to obtain stationarity conditions for the class of SYMARFIMA model.

Theorem 3.6 *Let $\{Y_t\}_{t\in\mathbb{Z}}$ be a well-defined SYMARFIMA(p,d,q), for $d\in(-1,0.5)$ and let $\{\boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ a set of random covariates. Then $\{Y_t\}_{t\in\mathbb{Z}}$ is stationary if, and only if, $\{\boldsymbol{x}_t\}_{t\in\mathbb{Z}}$ is stationary.*

Proof: Straightforward consequence of Theorems 3.3 through 3.5. ∎

4 Estimation

The log-likelihood function in the proposed SYMARFIMA(p,d,q) model resembles the log-likelihood in the SYMARMA(p,q), but the structure of μ_t is closer to the βARFIMA(p,d,q) model of Pumi et al. (2019). Let $\{(Y_t,\boldsymbol{x}_t)\}_{t=1}^n$ a sample (seen as random variables) from a SYMARFIMA(p,d,q) model, for $\boldsymbol{x}_t$ a set of l-dimensional covariates, with $\boldsymbol{x}_t=(x_{t1},\cdots,x_{tl})'$. Let $\boldsymbol{\gamma}=(\alpha,\boldsymbol{\beta}',d,\boldsymbol{\phi}',\boldsymbol{\theta}',\varphi)'$ denote the $(p+q+l+3)$-dimensional parameter vector, $\Omega\subseteq\mathbb{R}^{l+1}\times(-1,0.5)\times\mathbb{R}^{p+q}\times(0,\infty)$ and define

$$\ell_t(\boldsymbol{\gamma}) := \log(f(Y_t|\mathscr{F}_{t-1})) = \log\big(g(a_t)\big) - \frac{1}{2}\log(\varphi),$$

where $a_t=\frac{(Y_t-\mu_t)^2}{\varphi}$, for μ_t is defined in (2.7) and $\varphi>0$. The partial log-likelihood function is given by

$$\ell(\boldsymbol{\gamma}) = \sum_{t=1}^{n}\ell_t(\boldsymbol{\gamma}) = -\frac{n}{2}\log(\varphi) + \sum_{t=1}^{n}\log\big\{g(a_t)\big\}, \tag{4.1}$$

so that the partial maximum likelihood estimator of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}\in\Omega}{\mathrm{argmax}}(\ell(\boldsymbol{\gamma})).$$

4.1 Partial score vector

To construct the score vector, we need to obtain the derivative of $\ell(\boldsymbol{\gamma})$ given in (4.1), or, equivalently, the derivative of $\ell_t(\boldsymbol{\gamma})$. We start by noticing that,

$$\frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \mu_t} = -\frac{2g'(a_t)}{g(a_t)}\sqrt{\frac{a_t}{\varphi}},$$

and if $\gamma_j \neq \varphi$

$$\frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \gamma_j} = \frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \mu_t}\frac{\partial \mu_t}{\partial \gamma_j} = -\left[\frac{2g'(a_t)}{g(a_t)}\sqrt{\frac{a_t}{\varphi}}\right]\frac{\partial \mu_t}{\partial \gamma_j}, \tag{4.2}$$

since $P(g(a_t) \neq 0) = 1$. Upon observing the similarities between μ_t in (2.7) and the related structure in the βARFIMA(p, d, q) model. Identification of μ_t in (2.7) with η_t in equation (6) of Pumi et al. (2019), we obtain

$$\begin{aligned}
\frac{\partial \mu_t}{\partial \alpha} &= 1 - \sum_{k=1}^{\infty} c_k \frac{\partial \mu_{t-k}}{\partial \alpha};\\
\frac{\partial \mu_t}{\partial \beta_s} &= x_{ts} - \sum_{j=1}^{p} \phi_j x_{(t-j)s} - \sum_{k=1}^{\infty} c_k \frac{\partial \mu_{t-k}}{\partial \beta_s};\\
\frac{\partial \mu_t}{\partial d} &= \sum_{k=1}^{\infty}\left(r_{t-k}\sum_{i=0}^{\min\{k,q\}} \theta_i \pi_{k-i}[\psi(d+k-i) - \psi(d)] - c_k \frac{\partial \mu_{t-k}}{\partial d}\right);\\
\frac{\partial \mu_t}{\partial \phi_s} &= Y_{t-s} - \boldsymbol{x}'_{t-s}\boldsymbol{\beta} - \sum_{k=1}^{\infty} c_k \frac{\partial \mu_{t-k}}{\partial \phi_s};\\
\frac{\partial \mu_t}{\partial \theta_s} &= \sum_{k=s}^{\infty} \pi_{k-s} r_{t-k} - \sum_{k=1}^{\infty} c_k \frac{\partial \mu_{t-k}}{\partial \theta_s},
\end{aligned}$$

provided that the infinite series are convergent, where $\psi : (0, \infty) \to \mathbb{R}$ is the digamma function defined as $\psi(z) = \frac{d}{dz}\log\big(\Gamma(z)\big)$. Plugging these equations in (4.2), we obtain the desired derivatives. It remains to obtain the derivative with respect to φ. It is easy to show that

$$\frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \varphi} = -\frac{1}{\varphi}\left[\frac{1}{2} + \frac{g'(a_t)a_t}{g(a_t)}\right],$$

provided that $g(a_t) \neq 0$. Let $\boldsymbol{\rho} := (\alpha, \boldsymbol{\beta}', d, \boldsymbol{\phi}', \boldsymbol{\theta}')'$ so that $\boldsymbol{\gamma} = (\boldsymbol{\rho}', \varphi)$. Define

$$\boldsymbol{h}_1 := \left(\frac{\partial \ell_1(\boldsymbol{\gamma})}{\partial \mu_1}, \cdots, \frac{\partial \ell_n(\boldsymbol{\gamma})}{\partial \mu_n}\right)' \quad \text{and} \quad \boldsymbol{h}_2 := \left(\frac{\partial \ell_1(\boldsymbol{\gamma})}{\partial \varphi}, \cdots, \frac{\partial \ell_n(\boldsymbol{\gamma})}{\partial \varphi}\right)'$$

and let $D_{\boldsymbol{\rho}}$ be the $n \times (p + q + l + 2)$ matrix whose (i, j)th element is given by

$$[D_{\boldsymbol{\rho}}]_{i,j} := \frac{\partial \ell_i(\boldsymbol{\gamma})}{\partial \rho_j} \tag{4.3}$$

so that the score vector $U(\boldsymbol{\gamma})$ can be written as

$$U(\boldsymbol{\gamma}) = \big(U_{\boldsymbol{\rho}}(\boldsymbol{\gamma})', U_\varphi(\boldsymbol{\gamma})\big)' \quad \text{with} \quad U_{\boldsymbol{\rho}}(\boldsymbol{\gamma}) = D_{\boldsymbol{\rho}}'\boldsymbol{h}_1, \quad U_\varphi(\boldsymbol{\gamma}) = \mathbf{1}_n'\boldsymbol{h}_2,$$

where $\mathbf{1}_n := (1, \cdots, 1)' \in \mathbb{R}^n$. The PMLE is obtained as a solution of the non-linear system $U(\boldsymbol{\gamma}) = \mathbf{0}$. Since there are no closed formulas to solve this system, numeric optimization is required to obtain the PMLE.

4.2 Conditional information matrix

In this section we derive the Fisher's conditional information matrix which will be useful later on deriving the asymptotic properties of the partial maximum likelihood estimator for the proposed model. In this section we assume that g is such that $\mathbb{E}(Y_t^2|\mathscr{F}_{t-1})$ and $\mathbb{E}(Y_t^2)$ exist and are finite. Observe that

$$\frac{\partial \ell(\boldsymbol{\gamma})}{\partial \gamma_j} = \sum_{t=1}^n \left[\frac{\partial \ell_t(\mu_t, \varphi)}{\partial \mu_t}\frac{\partial \mu_t}{\partial \gamma_j} + \frac{\partial \ell_t(\mu_t, \varphi)}{\partial \varphi}\frac{\partial \varphi}{\partial \gamma_j}\right],$$

so that

$$\frac{\partial^2 \ell(\boldsymbol{\gamma})}{\partial \gamma_i \partial \gamma_j} = \sum_{t=1}^n \frac{\partial}{\partial \mu_t}\left[\frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \mu_t}\frac{d\mu_t}{\partial \gamma_j}\right]\frac{d\mu_t}{\partial \gamma_i} = \sum_{t=1}^n \left[\frac{\partial^2 \ell_t(\boldsymbol{\gamma})}{\partial \mu_t^2}\frac{\partial \mu_t}{\partial \gamma_j} + \frac{\partial \ell_t(\boldsymbol{\gamma})}{\partial \mu_t}\frac{\partial}{\partial \mu_t}\left(\frac{\partial \mu_t}{\partial \gamma_j}\right)\right]\frac{\partial \mu_t}{\partial \gamma_i}. \tag{4.4}$$

Observe that $\frac{\partial \ell_t(\gamma)}{\partial \mu_t}$ is antisymmetric, therefore $\mathbb{E}\left(\frac{\partial \ell_t(\gamma)}{\partial \mu_t}|\mathscr{F}_{t-1}\right) = 0$ almost surely. In view of this result, our task becomes to obtain a version for $\mathbb{E}\Big(\frac{\partial^2 \ell(\gamma)}{\partial \gamma_i \partial \gamma_j}|\mathscr{F}_{t-1}\Big)$, since all other terms in (4.4) are $\mathscr{F}_{t-1}$-measurable. We have

$$\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\gamma})}{\partial \gamma_i \partial \gamma_j}\bigg|\mathscr{F}_{t-1}\right) = \sum_{t=1}^n \mathbb{E}\left(\frac{\partial^2 \ell_t(\boldsymbol{\gamma})}{\partial \mu_t^2}\bigg|\mathscr{F}_{t-1}\right)\frac{\partial \mu_t}{\partial \gamma_i}\frac{\partial \mu_t}{\partial \gamma_j},$$

almost surely. Notice that

$$\frac{\partial^2 \ell_t(\boldsymbol{\gamma})}{\partial \mu_t^2} = -\frac{2}{\sqrt{\varphi}}\left[\frac{\partial a_t}{\partial \mu_t}\right]\frac{\partial}{\partial a_t}\left(\sqrt{a_t}\frac{g'(a_t)}{g(a_t)}\right) = \frac{4a_t}{\varphi}\left[\frac{g''(a_t)}{g(a_t)} - \frac{1}{g'(a_t)}\right] + \frac{2}{\varphi}\frac{g'(a_t)}{g(a_t)},$$

which is well defined since $g(a_t) \neq 0$ and $g'(a_t) \neq 0$ with probability 1. Observe that since g is symmetric around μ_t, the functions $\frac{g'(a_t)}{g(a_t)}$ and $\frac{1}{g'(a_t)}$ are both antisymmetric and, hence, $\mathbb{E}\Big(\frac{g'(a_t)}{g(a_t)}|\mathscr{F}_{t-1}\Big)$ and $\mathbb{E}\Big(\frac{a_t}{g'(a_t)}|\mathscr{F}_{t-1}\Big)$ are both 0, almost surely. We conclude that

$$\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\gamma})}{\partial \gamma_i \partial \gamma_j}\bigg|\mathscr{F}_{t-1}\right) = -\frac{4}{\varphi}\sum_{t=1}^n \mathbb{E}\left(\frac{g''(a_t)a_t}{g(a_t)}\bigg|\mathscr{F}_{t-1}\right)\frac{\partial \mu_t}{\partial \gamma_i}\frac{\partial \mu_t}{\partial \gamma_j}, \quad \text{a.s.}$$

Now, since

$$\frac{\partial}{\partial\varphi}\left[\frac{g'(a_t)a_t}{g(a_t)}\right] = \left[\frac{\partial a_t}{\partial\varphi}\right]\frac{\partial}{\partial a_t}\left[\frac{g'(a_t)a_t}{g(a_t)}\right] = \frac{a_t^2}{\varphi}\left[\frac{1}{g'(a_t)} - \frac{g''(a_t)}{g(a_t)}\right] - \frac{a_t g'(a_t)}{\varphi g(a_t)},$$

we have

$$\begin{aligned}\frac{\partial^2\ell_t(\boldsymbol{\gamma})}{\partial\varphi^2} &= \frac{1}{\varphi^2}\left[\frac{1}{2} + \frac{g'(a_t)a_t}{g(a_t)}\right] - \frac{1}{\varphi}\left(\frac{\partial}{\partial\varphi}\left[\frac{g'(a_t)a_t}{g(a_t)}\right]\right)\\ &= \frac{1}{\varphi^2}\left[\frac{1}{2} + \frac{g'(a_t)a_t}{g(a_t)}\right] + \frac{a_t^2}{\varphi^2}\left[\frac{g''(a_t)}{g(a_t)} - \frac{1}{g'(a_t)}\right] + \frac{a_t g'(a_t)}{\varphi^2 g(a_t)}\\ &= \frac{1}{\varphi^2}\left[\frac{1}{2} + \frac{g'(a_t)a_t}{g(a_t)} + \frac{a_t^2 g''(a_t)}{g(a_t)} - \frac{a_t^2}{g'(a_t)} + \frac{a_t g'(a_t)}{g(a_t)}\right].\end{aligned}$$

Since $\frac{g'(a_t)}{g(a_t)}$, $\frac{1}{g'(a_t)}$ and $\frac{g'(a_t)}{g(a_t)}$ are all antisymmetric, their expectations are zero. Hence

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\varphi^2}\middle|\mathscr{F}_{t-1}\right) = \frac{1}{\varphi^2}\left[\frac{n}{2} + \sum_{t=1}^{n}\mathbb{E}\left(\frac{a_t^2 g''(a_t)}{g(a_t)}\middle|\mathscr{F}_{t-1}\right)\right], \qquad \text{a.s.}$$

Finally, for $\gamma_i \neq \varphi$, we can write the derivative of (4.2) with respect to φ as

$$\frac{\partial^2\ell_t(\boldsymbol{\gamma})}{\partial\varphi\partial\gamma_i} = -\frac{1}{\varphi^{\frac{3}{2}}}\left[\frac{\partial\mu_t}{\partial\gamma_j}\right]\frac{\partial}{\partial\varphi}\left[\frac{g'(a_t)}{g(a_t)}(Y_t-\mu_t)\right],$$

and upon observing that $\frac{g'(a_t)}{g(a_t)}(Y_t-\mu_t)$ is symmetric, it follows that $\frac{\partial}{\partial\varphi}\left[\frac{g'(a_t)}{g(a_t)}(Y_t-\mu_t)\right]$ is antisymmetric and we conclude that

$$\mathbb{E}\left(\frac{\partial^2\ell_t(\boldsymbol{\gamma})}{\partial\varphi\partial\gamma_i}\middle|\mathscr{F}_{t-1}\right) = 0, \qquad \text{a.s.}$$

Let E_μ, $E_{\mu\varphi}$ and E_φ are $n\times n$ diagonal matrices for which the tth diagonal element is given by $[E_\mu]_{t,t} := -\mathbb{E}\Big(\frac{\partial^2\ell_t}{\partial\mu_t^2}\Big|\mathscr{F}_{t-1}\Big)$, and $[E_\varphi]_{t,t} = -\mathbb{E}\Big(\frac{\partial^2\ell_t}{\partial\varphi^2}|\mathscr{F}_{t-1}\Big)$. With this notation, the conditional Fisher information matrix based on a sample of size n can be written as

$$K_n(\boldsymbol{\gamma}) := \begin{pmatrix} K_{\boldsymbol{\rho},\boldsymbol{\rho}} & \mathbf{0}_s \\ \mathbf{0}_s' & K_{\varphi,\varphi}\end{pmatrix},$$

where $\mathbf{0}_s := (0,\cdots,0)' \in \mathbb{R}^{p+q+l+2}$, $K_{\boldsymbol{\rho},\boldsymbol{\rho}} := D_{\boldsymbol{\rho}}' E_\mu D_{\boldsymbol{\rho}}$, and $K_{\varphi,\varphi} := \mathbf{1}_n' E_\varphi \mathbf{1}_n$ for $D_{\boldsymbol{\rho}}$ given in (4.3).

5 Asymptotic theory

Asymptotic theory under the framework of partial maximum likelihood estimation for GARMA-like models was developed by Li (1991); Fokianos and Kedem (1998, 2004). The asymptotic theory for the SYMARFIMA(p,d,q) models falls under the scope of the general theory presented in Fokianos and Kedem (1998, 2004). The assumptions required for asymptotic existence/uniqueness, asymptotic consistency, and the normality of the partial likelihood estimator for SYMARFIMA(p,d,q) models are fundamentally the same in Pumi et al. (2019). Let $\{(Y_t,\boldsymbol{x}_t)\}_{t=1}^n$ be a sample from a SYMARFIMA(p,d,q) model, $Z_t := (1, Y_{t-1}, x_{t-1}, Y_{t-2}, x_{t-2}, \cdots)$ and let $\widehat{\boldsymbol{\gamma}}$ denote a solution of $U(\boldsymbol{\gamma}) = \mathbf{0}$. Besides conditions (a) and (b) given in Section 2, we shall assume that the following conditions, endemic to the discussion of partial maximum likelihood, hold.

(A) The parametric space Ω is an open set in $\mathbb{R}^{p+q+l+3}$ and the true parameter $\boldsymbol{\gamma}_0$ lies in Ω.

(B) For each t, Z_t almost surely belongs to a compact set $\Upsilon \subset \Omega$ and there exists $n_0 \in \mathbb{N}$ such that, for all $n > n_0$, $\sum_{t=1}^n \boldsymbol{Z}_t\boldsymbol{Z}_t'$ is almost surely positive. Additionally, assume that μ_t is almost surely well-defined for all $Z_t \in \Upsilon$ and $\boldsymbol{\gamma} \in \Omega$.

(C) There exists a probability measure λ in Ω such that $\int_\Omega \boldsymbol{t}\boldsymbol{t}'\lambda(d\boldsymbol{t})$ is positive definite and such that the weak convergence

$$\frac{1}{n}\sum_{t=1}^n I(Z_{t-1} \in A) \underset{n\longrightarrow\infty}{\longrightarrow} \lambda(A),$$

holds for all $\lambda-$continuity sets $A \subset \Omega$ under (2.7) with $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$.

Assumption A and B guarantees that the maximum likelihood problem is well-posed. Let $K_n(\boldsymbol{\gamma})$ denote the partial information matrix based on the given sample. Assumption B guarantee that $\frac{\partial^2\ell(\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}'}$ is continuous in $\boldsymbol{\gamma}$ while Assumption C assures the existence of a non-random information matrix, denoted by $K(\boldsymbol{\gamma})$, such that the weak convergence

$$\frac{K_n(\boldsymbol{\gamma})}{n} \underset{n\to\infty}{\longrightarrow} K(\boldsymbol{\gamma}), \quad \forall \boldsymbol{\gamma} \in \Omega,$$

holds and that $K(\boldsymbol{\gamma}_0)$ is a positive definite and invertible matrix. Finally, Assumptions A and B also imply that, as n increases, the partial information matrix will be eventually positive definite. See also the discussion in Fokianos and Kedem (1998, 2004) and references therein.

Theorem 5.1 *Under the assumptions (a), (b) and A-C, the probability that a locally unique maximum partial likelihood estimator exists in a neighborhood of $\boldsymbol{\gamma}_0$ tends to one. Furthermore, the estimator is consistent*

$$\widehat{\boldsymbol{\gamma}}_n \underset{n\to\infty}{\overset{P}{\longrightarrow}} \boldsymbol{\gamma}_0$$

and asymptotically normal

$$\sqrt{n}(\widehat{\boldsymbol{\gamma}}_n - \boldsymbol{\gamma}_0) \xrightarrow[n\to\infty]{d} N_{p+q+l+3}\big(\mathbf{0}, K(\boldsymbol{\gamma}_0)^{-1}\big).$$

The proof of Theorem 5.1 follows from the proof of theorem 3.1 in Fokianos and Kedem (1998) and theorem 1 in Fokianos and Kedem (2004). See also chapter 3 of Palma (2007).

5.1 Asymptotic tests and confidence intervals

Let γ_j and $\hat{\gamma}_j$ denote the jth component of the PMLE $\hat{\boldsymbol{\gamma}}$ based on a sample of size n from a SYMARFIMA(p,d,q) model and assume that assumptions (a), (b) and A-C are fulfilled. Theorem 5.1 offers a framework to construct asymptotic confidence intervals and to perform asymptotic hypothesis testing. The main ingredient for that is to observe that from Theorem 5.1, if n is large enough, the following approximation holds

$$\big[K_n(\hat{\boldsymbol{\gamma}})^{jj}\big]^{-\frac{1}{2}}(\hat{\gamma}_j - \gamma_j^*) \approx N(0,1), \tag{5.1}$$

where $K_n(\hat{\boldsymbol{\gamma}})^{jj}$ denotes the jth diagonal element of $K_n(\hat{\boldsymbol{\gamma}})^{-1}$. From this approximation, asymptotic confidence intervals can be derived. For $0 < \delta < 1/2$, a $100(1-\delta)\%$ confidence interval for γ_j is given by

$$\widehat{\gamma}_j \pm z_{1-\delta/2}\big(K_n(\widehat{\boldsymbol{\gamma}})^{jj}\big)^{1/2}.$$

Of particular interest in modeling is testing hypothesis of the form $H_0 : \gamma_j = \gamma_j^*$ against $H_1 : \gamma_j \neq \gamma_j^*$ for some γ_j^* given, tipically 0. To perform this test, we can apply (5.1) and construct the traditional Wald's z statistics by taking

$$z = \frac{\hat{\gamma}_j - \gamma_j^*}{\big(K_n(\widehat{\boldsymbol{\gamma}})^{jj}\big)^{1/2}},$$

which, under H_0, and for large enough n is approximately distributed as $N(0,1)$. Using the approximation (5.1) it is also possible to construct more general hypothesis test, as well as versions of well known test statistics such as the likelihood ratio or Rao's score tests. These statistics will be asymptotically distributed according to the limiting distribution of their counterparts under independence. The proof of such results is a combination of the techniques presented by Fahrmeir and Kaufmann (1985) with the results of Theorem 5.1.

5.2 Forecast

In possession of a PMLE estimate $\hat{\boldsymbol{\gamma}}$ based on a (numeric) sample $\{(y_t, \boldsymbol{x}_t)\}_{t=1}^n$ from a SYMARFIMA(p,d,q) model, forecasted values can be obtained straightforwardly.

The idea is recursively to construct estimates for μ_t by using (2.7) evaluated at the PMLE $\hat{\boldsymbol{\gamma}}$. Observe that the obtained sequence of estimates of μ_t, say $\{\hat{\mu}_t\}_{t=1}^n$, provides in-sample forecasts for y_t which can be used to estimate r_t. More specifically, h-step ahead forecast for y_t, denoted by $\hat{y}_n(h) := \hat{y}_{n+h}$, is obtained through

$$\hat{y}_n(h) := \widehat{\alpha} + \boldsymbol{x}'_{n+h}\widehat{\boldsymbol{\beta}} + \sum_{j=1}^{p} \widehat{\phi}_j\big([y_{n+h-j}]^* - \boldsymbol{x}'_{n+h-j}\widehat{\boldsymbol{\beta}}\big) + \sum_{k=1}^{m} \widehat{c}_k\left[\widehat{r}_{n+h-k}\right]^*,$$

where m is a user chosen truncation point for the MA(∞) representation in (2.7), $\hat{c}_k$ is obtained by applying the PMLE estimates $\hat{\theta}$ and $\hat{d}$ in expression (2.6),

$$[y_t]^* = \begin{cases} \hat{y}_n(t-n), & \text{if } t > n, \\ y_t, & \text{if } 0 \leq t \leq n, \end{cases}$$

and $\left[\widehat{r}_t\right]^* = (y_t - \hat{\mu}_t)I(0 \leq t \leq n)$. Starting the construction of $\hat{\mu}_t$ requires initialization. We initialize $y_t = r_t = 0$ and $\boldsymbol{x}_t = 0$, for $t \leq 0$. Also notice that when covariates are present in the model, h-step ahead forecasts require that future values of $\boldsymbol{x}_t$ up to time $n+h$ are provided.

6 Monte Carlo Study

In this section we study the finite sample performance of the proposed partial likelihood approach for parameter estimation in the context of SYMARFIMA models. For the simulation study, we simulate SYMARFIMA(p, d, q) models withe the following configurations:

- Underlying distributions: t_ν for $\nu \in \{4, 6\}$; sample sizes $n \in \{500, 1000\}$, $\varphi \in \{1, 5, 10\}$ and $\alpha = 20$.
- Long range dependence parameter $d \in \{0.2, 0.4\}$.
- SYMARFIMA$(1, d, 0)$ with $\phi \in \{0.2, 0.6\}$; SYMARFIMA$(0, d, 1)$ with $\theta \in \{0.2, 0.6\}$.
- SYMARFIMA$(1, d, 1)$ with $(\phi, \theta) \in \{(0.2, 0.6)\}$.
- For estimation purposes, m was taken as the sample size.
- In all cases a burn in of size 1,000 was applied.

The simulation results are shown in Tables 1 through 8. We observed that, for a SYMARFIMA$(1, d, 0)$ under distribution t_4 (Table 1), when $\phi = 0.2$, the model penalizes the estimation of ϕ by underestimating it and overestimating d. This behavior is also known from the literature of long range dependence processes. The results also shown that as φ increases this penalty is reduced/reversed. Also, observe that increasing φ increases the conditional variance of the process. This affects parameter estimation globally, except for the estimation of φ itself, as it is orthogonal to the other parameters. The impact of the penalty is most noticeable in the case of $\phi = 0.6$.

Another factor that affects the process's conditional variance is the degrees of freedom of the t-Student distribution. We observe that for a SYMARFIMA$(1, d, 0)$ under distribution t_6 (Table 2), the same behavior happens and the most problematic case again is $\phi = 0.6$. The estimates of d often exceeds 0.5 in this case. Comparing the results from the SYMARFIMA$(1, d, 0)$ and SYMARFIMA$(0, d, 1)$ side by side (Tables 1-4) we notice that the relationship between the estimate of d to the increase in φ is opposite though. For a SYMARFIMA$(1, d, 0)$ with the increase of φ we witnessed a decrease in the average estimate of d, while the opposite happens for the SYMARFIMA$(0, d, 1)$.

For a SYMARFIMA$(1, d, 1)$ under t-Student distribution with 4 and 6 degrees of freedom (Tables 5 and 8, respectively) estimation of the parameters is overall. In this case the relationship of ϕ, θ, d and α to φ needs further investigation.

Table 1: Simulation results for the SYMARFIMA$(1, d, 0)$ for $\nu = 4$

$n = 500$													
		$\phi = 0.2$						$\phi = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\phi}$	0.1319	0.0803	44.81	0.1142	0.0719	48.68	0.1306	0.1646	80.56	0.3371	0.0566	43.83
	$\hat{d}$	0.2401	0.0701	32.78	0.4553	0.0629	17.46	0.6142	0.1635	216.92	0.4899	0.0039	22.48
	$\hat{\alpha}$	21.7238	2.0378	11.34	22.4448	2.2904	14.02	45.9370	9.1687	133.20	35.8802	3.9055	79.42
	$\hat{\varphi}$	1.0014	0.1635	6.58	1.0046	0.1690	6.67	1.0695	0.1902	9.35	1.1278	0.1999	13.57
5	$\hat{\phi}$	0.1864	0.0817	30.33	0.1922	0.1244	38.00	0.5092	0.2063	30.36	0.4879	0.0499	19.37
	$\hat{d}$	0.1990	0.0694	24.51	0.3955	0.1135	15.97	0.2583	0.2035	88.19	0.4841	0.0349	22.36
	$\hat{\alpha}$	20.3896	2.1322	8.03	20.5451	3.9648	13.91	25.1276	11.6551	48.76	25.2005	4.4356	27.75
	$\hat{\varphi}$	4.9687	0.8377	6.61	5.0143	0.8842	7.08	5.0759	0.9035	7.34	5.1403	0.9211	7.20
10	$\hat{\phi}$	0.2678	0.2111	60.92	0.2463	0.2028	57.00	0.5521	0.1571	22.44	0.5107	0.0777	18.11
	$\hat{d}$	0.1297	0.1912	53.57	0.3476	0.1985	26.18	0.2239	0.1536	64.96	0.4697	0.0740	23.45
	$\hat{\alpha}$	18.3430	5.3326	15.73	19.0934	6.8898	20.67	22.5006	8.4159	34.48	24.0114	7.7649	28.57
	$\hat{\varphi}$	9.9207	2.0184	7.83	9.9436	1.8468	7.14	10.0866	1.7476	6.78	10.2453	1.9543	7.40
$n = 1000$													
		$\phi = 0.2$						$\phi = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\phi}$	0.1631	0.0527	26.41	0.1513	0.0513	29.47	0.6037	0.2044	27.89	0.4017	0.0444	33.05
	$\hat{d}$	0.2212	0.0440	19.62	0.4313	0.0426	10.96	0.1736	0.2001	85.98	0.4898	0.0042	22.46
	$\hat{\alpha}$	20.9267	1.3380	6.66	21.3576	1.5908	8.39	20.1192	11.1721	43.51	33.6155	3.8442	68.09
	$\hat{\varphi}$	1.0006	0.1161	4.72	1.0000	0.1192	4.67	1.0199	0.1413	5.57	1.0701	0.1336	7.93
5	$\hat{\phi}$	0.1937	0.0585	21.91	0.1998	0.1001	25.62	0.6045	0.1116	15.19	0.5052	0.0372	16.27
	$\hat{d}$	0.1990	0.0479	17.38	0.3933	0.0941	10.90	0.1798	0.1106	46.06	0.4842	0.0281	21.86
	$\hat{\alpha}$	20.1755	1.5169	5.74	20.2157	3.0234	9.48	19.7102	5.7331	22.71	23.9641	3.6503	21.06
	$\hat{\varphi}$	4.9818	0.6046	4.74	4.9971	0.6050	4.70	5.0351	0.6047	4.84	5.0713	0.6278	5.04
10	$\hat{\phi}$	0.3006	0.2516	68.43	0.2357	0.1769	41.18	0.5956	0.1004	13.50	0.5144	0.0542	15.90
	$\hat{d}$	0.1030	0.2278	60.56	0.3609	0.1751	18.80	0.1928	0.1002	40.29	0.4762	0.0510	22.06
	$\hat{\alpha}$	17.4864	6.3052	17.46	19.2945	6.1612	16.32	20.1893	5.2521	20.42	24.1513	8.9066	26.27
	$\hat{\varphi}$	9.9392	1.5877	5.94	9.9604	1.3337	5.08	10.0288	1.2001	4.73	10.1575	1.4696	5.32

Table 2: Simulation results for the SYMARFIMA$(1, d, 0)$ for $\nu = 6$

$n = 500$													
		$\phi = 0.2$						$\phi = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\phi}$	0.0828	0.0615	61.08	0.0712	0.0547	65.27	0.0477	0.0367	92.05	0.0801	0.0648	86.73
	$\hat{d}$	0.2745	0.0589	41.47	0.4889	0.0553	23.21	0.6852	0.0506	242.64	0.8606	0.0625	115.33
	$\hat{\alpha}$	22.9831	1.5828	15.56	23.6317	1.8757	18.44	49.1499	2.4150	145.75	49.6449	4.0685	148.35
	$\hat{\varphi}$	1.0039	0.1157	6.07	1.0075	0.1190	6.34	1.0622	0.1278	8.44	1.0572	0.1218	7.97
5	$\hat{\phi}$	0.1666	0.0775	33.78	0.1615	0.0965	39.84	0.2645	0.1950	59.22	0.2174	0.0888	63.98
	$\hat{d}$	0.2134	0.0681	26.90	0.4186	0.0843	16.22	0.4907	0.1981	161.84	0.7457	0.0857	86.95
	$\hat{\alpha}$	20.8821	2.0165	8.79	21.4842	3.3471	13.79	39.3935	11.8402	101.99	47.3630	8.7733	137.23
	$\hat{\varphi}$	4.9806	0.5941	6.34	5.0206	0.5769	6.20	5.1818	0.6294	7.19	5.1784	0.6200	7.14
10	$\hat{\phi}$	0.2049	0.1252	38.58	0.2118	0.1662	49.35	0.4260	0.1982	37.02	0.3323	0.1477	47.50
	$\hat{d}$	0.1840	0.1102	33.18	0.3756	0.1522	21.45	0.3391	0.1941	101.96	0.6372	0.1493	65.47
	$\hat{\alpha}$	19.9172	3.2301	10.15	20.1659	5.7821	17.50	29.7880	11.5650	61.13	40.4172	13.4987	107.21
	$\hat{\varphi}$	9.9436	1.1612	6.20	9.9795	1.1757	6.21	10.1800	1.1664	6.29	10.2717	1.1755	6.53
$n = 1000$													
		$\phi = 0.2$						$\phi = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\phi}$	0.1349	0.0555	36.83	0.1209	0.0539	42.36	0.2351	0.2449	67.86	0.1782	0.0485	70.34
	$\hat{d}$	0.2397	0.0471	25.70	0.4531	0.0468	15.13	0.5171	0.2407	186.98	0.7895	0.0418	97.48
	$\hat{\alpha}$	21.6427	1.4103	9.29	22.2271	1.6972	11.96	40.3463	13.5407	112.33	49.2153	3.9707	146.14
	$\hat{\varphi}$	1.0035	0.0851	4.41	1.0056	0.0849	4.52	1.0422	0.0909	5.96	1.0426	0.0852	5.68
5	$\hat{\phi}$	0.1826	0.0542	22.77	0.1781	0.0611	25.95	0.5665	0.1611	21.48	0.3605	0.1392	42.56
	$\hat{d}$	0.2083	0.0450	18.14	0.4110	0.0514	10.45	0.2135	0.1567	63.22	0.6156	0.1411	59.46
	$\hat{\alpha}$	20.4588	1.4036	5.91	20.9074	2.1963	9.17	21.9063	9.0073	33.59	38.8379	12.1127	98.73
	$\hat{\varphi}$	4.9992	0.4171	4.40	5.0159	0.4082	4.37	5.0510	0.4360	4.68	5.1248	0.4280	4.97
10	$\hat{\phi}$	0.2070	0.0940	26.64	0.2147	0.1417	34.62	0.5683	0.1306	17.92	0.4800	0.1445	26.38
	$\hat{d}$	0.1893	0.0823	22.57	0.3800	0.1351	15.22	0.2138	0.1278	51.78	0.5001	0.1447	37.23
	$\hat{\alpha}$	19.8252	2.4099	6.91	19.9933	4.0643	12.27	21.6824	7.0742	27.54	29.2395	12.2947	57.18
	$\hat{\varphi}$	9.9737	0.8410	4.43	9.9997	0.8491	4.46	10.0780	0.8297	4.39	10.1293	0.8366	4.50

Table 3: Simulation results for the SYMARFIMA$(0, d, 1)$ for $\nu = 4$

$n = 500$													
		$\theta = 0.2$						$\theta = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\theta}$	0.2081	0.0533	21.25	0.2047	0.0552	21.88	0.6038	0.0343	4.59	0.6017	0.0357	4.76
	$\hat{d}$	0.1883	0.0436	18.04	0.3941	0.0451	8.90	0.1910	0.0358	14.84	0.3952	0.0384	7.74
	$\hat{\alpha}$	20.0005	0.2533	1.00	20.00	0.9661	3.77	19.9969	0.3455	1.37	20.0053	1.4816	5.86
	$\hat{\varphi}$	0.9948	0.1720	6.87	0.9990	0.1667	6.56	0.9986	0.1676	6.77	0.9981	0.1693	6.80
5	$\hat{\theta}$	0.2006	0.0988	23.33	0.1380	0.0495	34.76	0.4607	0.0711	23.62	0.5602	0.0368	7.56
	$\hat{d}$	0.1950	0.0889	20.62	0.4705	0.0433	19.76	0.4445	0.1142	127.54	0.4792	0.0299	20.39
	$\hat{\alpha}$	19.9849	0.5656	2.23	19.9106	3.0499	11.58	19.7280	3.7936	12.40	20.0795	4.7589	18.89
	$\hat{\varphi}$	4.9943	0.8437	6.74	5.0224	0.8458	6.61	5.5210	1.0765	12.12	5.0832	0.8407	6.82
10	$\hat{\theta}$	0.1253	0.3096	61.11	0.1342	0.0461	35.06	0.4574	0.0648	24.27	0.5567	0.0357	7.96
	$\hat{d}$	0.2687	0.2905	54.90	0.4779	0.0341	20.61	0.4591	0.1000	134.55	0.4864	0.0175	21.73
	$\hat{\alpha}$	19.6775	3.4063	4.85	18.5714	7.1258	22.12	16.7552	9.9060	29.92	18.7976	9.0281	32.02
	$\hat{\varphi}$	9.8992	1.8453	6.98	10.1502	1.8191	7.18	11.1679	2.1275	12.99	10.1401	1.7753	7.09
$n = 1000$													
		$\theta = 0.2$						$\theta = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\theta}$	0.2045	0.0391	15.45	0.2001	0.0369	14.73	0.6022	0.0240	3.20	0.6015	0.0247	3.31
	$\hat{d}$	0.1943	0.0318	12.82	0.3980	0.0306	6.13	0.1946	0.0250	9.99	0.3976	0.0268	5.41
	$\hat{\alpha}$	19.9971	0.1852	0.73	20.0059	0.7717	3.03	20.0009	0.2415	0.95	20.0019	1.1877	4.69
	$\hat{\varphi}$	0.9976	0.1231	4.94	0.9988	0.1208	4.76	0.9972	0.1172	4.73	0.9991	0.1192	4.75
5	$\hat{\theta}$	0.1989	0.0846	17.29	0.1372	0.0379	33.16	0.4512	0.0589	25.10	0.5602	0.0291	7.16
	$\hat{d}$	0.1991	0.0791	14.61	0.4704	0.0409	19.13	0.4612	0.0919	133.29	0.4776	0.0315	19.92
	$\hat{\alpha}$	19.9695	0.4889	1.61	19.9829	2.5278	10.08	19.8801	3.4830	11.41	20.1329	4.2643	17.03
	$\hat{\varphi}$	5.0006	0.6268	4.96	5.0411	0.6234	5.00	5.5453	0.7750	11.46	5.0661	0.5841	4.74
10	$\hat{\theta}$	0.1414	0.2718	46.03	0.1332	0.0342	34.16	0.4477	0.0498	25.60	0.5560	0.0258	7.64
	$\hat{d}$	0.2545	0.2582	42.25	0.4763	0.0340	19.94	0.4720	0.0747	137.91	0.4857	0.0190	21.57
	$\hat{\alpha}$	19.5875	3.7084	4.59	19.7394	4.7192	15.59	17.2880	9.7648	27.23	19.7254	7.3513	26.80
	$\hat{\varphi}$	9.9364	1.2561	4.75	10.0750	1.2307	4.87	11.1361	1.4713	11.87	10.1259	1.2705	5.15

Table 4: Simulation results for the SYMARFIMA$(0, d, 1)$ for $\nu = 6$

$n = 500$													
		$\boldsymbol{\theta} = 0.2$						$\boldsymbol{\theta} = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\theta}$	0.2114	0.0571	23.50	0.2056	0.0583	22.94	0.6040	0.0389	5.21	0.5998	0.0393	5.24
	$\hat{d}$	0.1841	0.0482	20.26	0.3907	0.0493	9.89	0.1895	0.0397	16.11	0.3967	0.0406	8.03
	$\hat{\alpha}$	19.9869	0.1980	0.79	19.9718	0.7058	2.81	20.0019	0.2639	1.05	20.0090	1.1853	4.52
	$\hat{\varphi}$	0.9958	0.1151	6.09	0.9963	0.1180	6.32	0.9980	0.1166	6.21	1.0007	0.1168	6.21
5	$\hat{\theta}$	0.2121	0.0574	23.17	0.1248	0.0459	38.89	0.4531	0.0626	24.67	0.5542	0.0367	8.25
	$\hat{d}$	0.1860	0.0485	19.75	0.4842	0.0244	21.61	0.4636	0.0880	134.14	0.4892	0.0066	22.31
	$\hat{\alpha}$	20.0301	0.4361	1.74	19.9419	2.2465	8.57	19.7673	2.8548	9.78	19.9675	3.8175	15.18
	$\hat{\varphi}$	4.9831	0.5701	6.13	5.0194	0.5685	6.07	5.4324	0.6865	10.24	5.0503	0.6006	6.35
10	$\hat{\theta}$	0.1364	0.2909	58.57	0.1314	0.0458	35.88	0.4514	0.0562	24.96	0.5571	0.0384	8.04
	$\hat{d}$	0.2559	0.2739	52.37	0.4865	0.0197	22.01	0.4781	0.0603	140.19	0.4894	0.0104	22.50
	$\hat{\alpha}$	19.8635	1.3146	3.01	17.8426	6.7091	21.48	16.9322	7.7215	25.70	17.7527	8.7836	31.36
	$\hat{\varphi}$	9.9740	1.2584	6.53	10.1761	1.3159	6.97	10.9856	1.4214	11.11	10.1719	1.2144	6.51
$n = 1000$													
		$\boldsymbol{\theta} = 0.2$						$\boldsymbol{\theta} = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\theta}$	0.2051	0.0415	16.60	0.2038	0.0404	16.33	0.6018	0.0283	3.79	0.6000	0.0278	3.69
	$\hat{d}$	0.1928	0.0326	13.26	0.3946	0.0313	6.31	0.1952	0.0274	10.96	0.3986	0.0277	5.50
	$\hat{\alpha}$	19.9958	0.1339	0.53	19.9805	0.5608	2.20	19.9986	0.1756	0.70	20.0161	0.9543	3.72
	$\hat{\varphi}$	0.9965	0.0794	4.24	0.9986	0.0821	4.37	0.9986	0.0799	4.24	0.9994	0.0818	4.27
5	$\hat{\theta}$	0.2063	0.0399	15.97	0.1235	0.0337	38.73	0.4484	0.0477	25.30	0.5554	0.0271	7.67
	$\hat{d}$	0.1928	0.0329	13.56	0.4846	0.0213	21.47	0.4728	0.0697	137.25	0.4890	0.0082	22.27
	$\hat{\alpha}$	20.0153	0.2983	1.13	19.9831	1.8915	7.55	19.9354	2.4396	8.72	20.0049	3.4114	13.82
	$\hat{\varphi}$	4.9915	0.4165	4.43	5.0266	0.3926	4.19	5.4425	0.4996	9.39	5.0512	0.4118	4.47
10	$\hat{\theta}$	0.1607	0.2314	38.40	0.1265	0.0345	37.03	0.4453	0.0415	25.80	0.5544	0.0269	7.80
	$\hat{d}$	0.2358	0.2171	34.25	0.4870	0.0158	21.90	0.4815	0.0485	140.96	0.4900	0.0000	22.50
	$\hat{\alpha}$	19.8735	1.3767	2.16	18.7848	5.4909	16.13	17.7041	7.2445	22.03	19.1409	6.9811	24.26
	$\hat{\varphi}$	9.9770	0.8737	4.57	10.1048	0.8976	4.78	10.9387	1.0293	9.85	10.1346	0.8164	4.47

Table 5: Simulation results for the SYMARFIMA$(1, d, 1)$ for $\nu = 4$

		$n = 500$						$n = 1000$					
		$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.6$						$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
1	$\hat{\phi}$	0.0451	0.0767	79.51	0.0516	0.0747	76.45	0.0810	0.0699	61.73	0.0796	0.0692	62.55
	$\hat{\theta}$	0.6510	0.0411	9.33	0.6495	0.0409	9.15	0.6380	0.0327	7.08	0.6384	0.0330	7.14
	$\hat{d}$	0.2762	0.0552	40.90	0.4769	0.0549	20.72	0.2590	0.0439	31.86	0.4644	0.0449	17.19
	$\hat{\alpha}$	24.0564	2.0738	20.79	24.7281	3.0993	24.54	23.1187	1.8682	16.16	24.4702	3.1136	23.20
	$\hat{\varphi}$	1.0070	0.1766	6.85	1.0019	0.1775	7.04	1.0069	0.1210	4.83	1.0015	0.1173	4.64
5	$\hat{\phi}$	0.1679	0.1552	60.02	0.1398	0.1354	56.89	0.2270	0.1630	58.31	0.1630	0.0846	37.19
	$\hat{\theta}$	0.5901	0.2104	11.35	0.6031	0.1832	10.17	0.5931	0.0496	6.44	0.6066	0.0950	5.81
	$\hat{d}$	0.2246	0.1848	47.05	0.4463	0.1600	22.29	0.1779	0.1254	40.90	0.4245	0.0847	12.83
	$\hat{\alpha}$	21.0217	4.4607	16.87	23.6164	7.1746	29.71	19.0611	4.6682	16.41	22.5010	5.9469	23.24
	$\hat{\varphi}$	4.8992	0.9495	7.40	4.9738	0.9716	7.52	4.8127	0.8271	6.88	4.9870	0.6390	5.01
10	$\hat{\phi}$	0.2603	0.2830	103.01	0.2087	0.2212	77.26	0.2278	0.2944	103.56	0.1960	0.1631	50.85
	$\hat{\theta}$	0.5069	0.3831	23.78	0.4926	0.4257	26.58	0.5656	0.2605	14.07	0.5739	0.2281	10.32
	$\hat{d}$	0.2271	0.3862	114.35	0.4777	0.3612	48.25	0.2134	0.3266	100.44	0.4212	0.2213	24.35
	$\hat{\alpha}$	16.5803	13.8765	40.43	20.3392	19.5087	56.18	14.5278	15.7799	47.28	20.6782	22.1622	44.05
	$\hat{\varphi}$	9.5474	2.9107	11.05	9.8159	2.6427	9.58	9.5323	2.2873	8.81	9.9011	1.6801	6.08

Table 6: Simulation results for the SYMARFIMA$(1, d, 1)$ for $\nu = 6$

		$n = 500$						$n = 1000$					
		$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.6$						$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.6$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
	$\hat{\phi}$	0.0293	0.0602	85.48	0.0346	0.0656	83.60	0.0503	0.0580	75.11	0.0543	0.0606	73.61
	$\hat{\theta}$	0.6585	0.0399	10.18	0.6540	0.0412	9.73	0.6498	0.0305	8.53	0.6454	0.0312	7.97
1	$\hat{d}$	0.2861	0.0500	44.08	0.4862	0.0516	22.39	0.2766	0.0401	38.96	0.4781	0.0410	19.94
	$\hat{\alpha}$	24.3848	1.6200	21.95	24.7258	2.6438	24.10	23.8716	1.5325	19.43	24.6043	2.5190	23.37
	$\hat{\varphi}$	1.0054	0.1194	6.26	1.0058	0.1165	6.21	1.0042	0.0892	4.67	1.0045	0.0848	4.48
	$\hat{\phi}$	0.1167	0.1300	63.19	0.1131	0.1272	64.37	0.1634	0.1152	47.89	0.1352	0.0885	45.19
	$\hat{\theta}$	0.6252	0.0890	8.49	0.6255	0.1030	8.99	0.6115	0.0398	5.51	0.6209	0.0382	5.88
5	$\hat{d}$	0.2379	0.1103	40.33	0.4452	0.1094	20.35	0.2149	0.0798	30.88	0.4331	0.0595	13.90
	$\hat{\alpha}$	22.3920	3.7299	17.57	23.6955	5.9158	27.22	21.1031	3.3831	13.32	23.1792	4.9860	22.19
	$\hat{\varphi}$	4.9223	0.6321	6.80	5.0186	0.6414	6.70	4.9358	0.4680	5.02	5.0171	0.4088	4.33
	$\hat{\phi}$	0.2179	0.2518	90.92	0.1883	0.2094	77.73	0.2152	0.2726	95.49	0.1821	0.1451	48.86
	$\hat{\theta}$	0.5532	0.2908	17.05	0.5522	0.3148	18.16	0.5341	0.3326	18.56	0.6061	0.0828	6.12
10	$\hat{d}$	0.2185	0.3112	89.25	0.4390	0.2729	36.95	0.2504	0.3479	103.01	0.4044	0.1366	18.28
	$\hat{\alpha}$	19.5953	11.1540	31.01	21.5975	12.4198	39.25	18.6883	23.0726	45.48	21.6284	8.5225	29.16
	$\hat{\varphi}$	9.7244	1.7238	8.82	9.8719	1.4873	7.72	9.7190	1.6530	7.82	9.9096	0.9494	4.94

Table 7: Simulation results for the SYMARFIMA$(1, d, 1)$ for $\nu = 4$

		$n = 500$						$n = 1000$					
		$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.2$						$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.2$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
	$\hat{\phi}$	0.0201	0.0740	91.24	0.0239	0.0788	89.64	0.0390	0.0763	82.64	0.0442	0.0784	79.89
	$\hat{\theta}$	0.3220	0.0677	62.31	0.3180	0.0681	61.02	0.3067	0.0609	55.31	0.3038	0.0588	53.49
1	$\hat{d}$	0.2467	0.0473	27.89	0.4484	0.0497	14.48	0.2444	0.0376	24.94	0.4437	0.0382	12.29
	$\hat{\alpha}$	24.5586	1.8970	23.11	24.6016	2.4945	23.53	24.0795	1.9442	20.93	24.1679	2.4221	21.44
	$\hat{\varphi}$	0.5014	0.0857	49.86	0.4978	0.0840	50.22	0.5028	0.0600	49.72	0.4997	0.0604	50.03
	$\hat{\phi}$	0.1061	0.1333	67.61	0.0996	0.1410	71.34	0.1342	0.1065	51.08	0.1288	0.1147	54.81
	$\hat{\theta}$	0.2712	0.0944	48.48	0.2642	0.1439	53.79	0.2489	0.0736	35.42	0.2488	0.0895	38.96
5	$\hat{d}$	0.2163	0.0666	27.15	0.4295	0.1111	16.11	0.2131	0.0483	20.26	0.4193	0.0581	10.73
	$\hat{\alpha}$	22.4222	3.4814	17.53	23.0161	4.9269	22.60	21.7273	2.7660	13.33	22.3731	4.1174	18.31
	$\hat{\varphi}$	2.4952	0.4108	50.10	2.4995	0.4293	50.01	2.4925	0.2931	50.15	2.4984	0.2960	50.03
	$\hat{\phi}$	0.1785	0.1985	75.65	0.1984	0.2011	78.05	0.2256	0.1787	67.93	0.1820	0.1574	62.04
	$\hat{\theta}$	0.1726	0.3007	79.08	0.0442	0.4619	135.41	0.0121	0.4633	135.04	0.0598	0.4399	119.37
10	$\hat{d}$	0.2414	0.2507	58.69	0.5453	0.3859	54.98	0.3525	0.3871	106.98	0.5505	0.3538	46.41
	$\hat{\alpha}$	20.5608	5.0734	19.83	20.5879	7.5250	28.57	19.5213	4.9626	18.11	20.5422	8.6498	27.53
	$\hat{\varphi}$	4.9288	1.0044	50.71	4.8949	1.0218	51.05	4.9020	0.8691	50.98	4.9384	0.8613	50.62

Table 8: Simulation results for the SYMARFIMA$(1, d, 1)$ for $\nu = 6$

		$n = 500$						$n = 1000$					
		$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.2$						$\boldsymbol{\phi} = 0.2$ $\boldsymbol{\theta} = 0.2$					
φ	$\hat{\gamma}_i$	$d = 0.2$			$d = 0.4$			$d = 0.2$			$d = 0.4$		
		mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE	mean	sd.	MAPE
	$\hat{\phi}$	0.0154	0.0542	92.35	0.0163	0.0629	92.24	0.0237	0.0513	88.39	0.0284	0.0607	86.18
	$\hat{\theta}$	0.3265	0.0627	63.82	0.3248	0.0665	63.64	0.3185	0.0486	59.59	0.3138	0.0534	57.65
1	$\hat{d}$	0.2449	0.0508	28.25	0.4506	0.0514	14.81	0.2472	0.0350	25.14	0.4498	0.0376	13.41
	$\hat{\alpha}$	24.6302	1.3857	23.16	24.7814	2.0194	24.06	24.4302	1.2978	22.22	24.5505	1.9044	22.88
	$\hat{\varphi}$	0.6652	0.0804	33.48	0.6658	0.0789	33.42	0.6669	0.0540	33.31	0.6670	0.0559	33.30
	$\hat{\phi}$	0.0747	0.1138	71.72	0.0733	0.1372	79.16	0.1043	0.0969	57.70	0.0962	0.1116	64.06
	$\hat{\theta}$	0.2911	0.0882	52.67	0.2890	0.1058	56.57	0.2656	0.0689	39.24	0.2726	0.0783	44.53
5	$\hat{d}$	0.2258	0.0580	25.17	0.4278	0.0815	15.54	0.2229	0.0465	20.92	0.4246	0.0514	11.56
	$\hat{\alpha}$	23.2174	2.9601	18.43	23.5831	4.3786	23.24	22.4704	2.5073	14.81	23.1130	3.7164	19.77
	$\hat{\varphi}$	3.3296	0.3965	33.41	3.3242	0.3908	33.52	3.3272	0.2778	33.46	3.3287	0.2841	33.43
	$\hat{\phi}$	0.1799	0.2017	77.26	0.1738	0.2077	81.23	0.2172	0.1580	58.84	0.1705	0.1629	63.41
	$\hat{\theta}$	0.1189	0.3982	108.20	0.0486	0.4807	143.09	-0.0709	0.5305	174.52	0.0817	0.4260	113.54
10	$\hat{d}$	0.2884	0.3425	86.79	0.5637	0.3898	57.28	0.4401	0.4672	147.91	0.5391	0.3417	44.17
	$\hat{\alpha}$	20.5951	5.6027	20.15	21.0328	6.8369	25.65	19.9200	4.7010	16.82	21.2626	6.5255	22.70
	$\hat{\varphi}$	6.5163	0.9765	34.84	6.5525	1.0435	34.48	6.4686	0.9819	35.31	6.6109	0.7111	33.89

7 Application

In this section we present an application of the proposed model to a real data time series. We consider the time series of the operations with clients in the financial sector - purchases (henceforth called operations time series), which records the sum of the foreign currency purchase and sale contracts of commercial banks with the non-financial market. The commercial segment includes, since February 2014, export of goods, import of goods, back-to-back operations, international orders, adjustments in commercial transactions, acquisition of goods delivered abroad, acquisition of goods delivered in the country, nature originating in the exchange system, and it's regulation. The series studied consists of daily data, measured by millions of US dollars, comprising $n = 973$ observations from April 4, 2017 to February 26, 2021. The data from February 19, 2021 to February 26, 2021 was reserved for forecasting purposes. For better visualization and numerical stability, the data was divided by 1000. The data comes from the Central Bank in Brazil - Economic Department and were obtained through R's library BETS version 0.4.9 (Costa Ferreira et al., 2018).

Our main objective is to model and forecast the operation time series with both, the SYMARMA and the proposed SYMARFIMA model and provides grounds for comparison. Figure 1 present the time series plot of the data while Figure 2 show its autocorrelation and partial autocorrelation functions (ACF and PACF, respectively). No deterministic component nor seasonality is apparent from the plots. To verify its stationarity, we apply the Phillips-Perron and Dickey-Fuller tests. Both tests reject the null hypothesis of unitary roots with p-values smaller than 0.01. Hence we are ready to proceed modeling the time series with the SYMARMA and SYMARFIMA models, for which we apply a Box-Jenkins approach for model selection. A simple SYMARFIMA$(1, d, 0)$ with t_4 distribution sufficed to fit the data. Table 9 presents detailed information about the fitted model. It is well known in the literature of ARMA and ARFIMA models that an ARMA model can mimic the autocorrelation decay of an ARFIMA model as long as a sufficient number of autoregressive components are allowed in the model (Prass et al., 2012; Pumi et al., 2019). It is interesting to see that this is still the case for SYMARMA and SYMARFIMA. A good fit to the operations data required a SYMARMA$(5, 0)$. Detailed information about the fitted model is presented in Table 10. Observe that the values of the loglikelihood, AIC and BIC, are identical, for practical purposes, with a small advantage for the SYMARFIMA fit. Also, both models present similar values of φ.

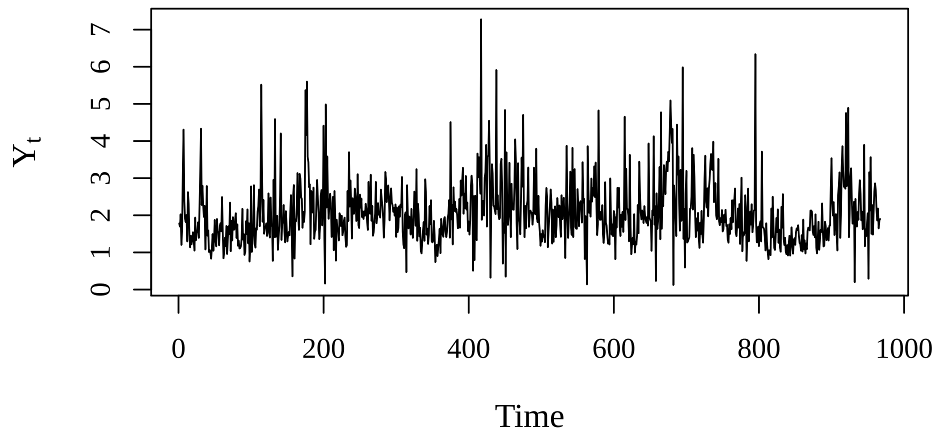
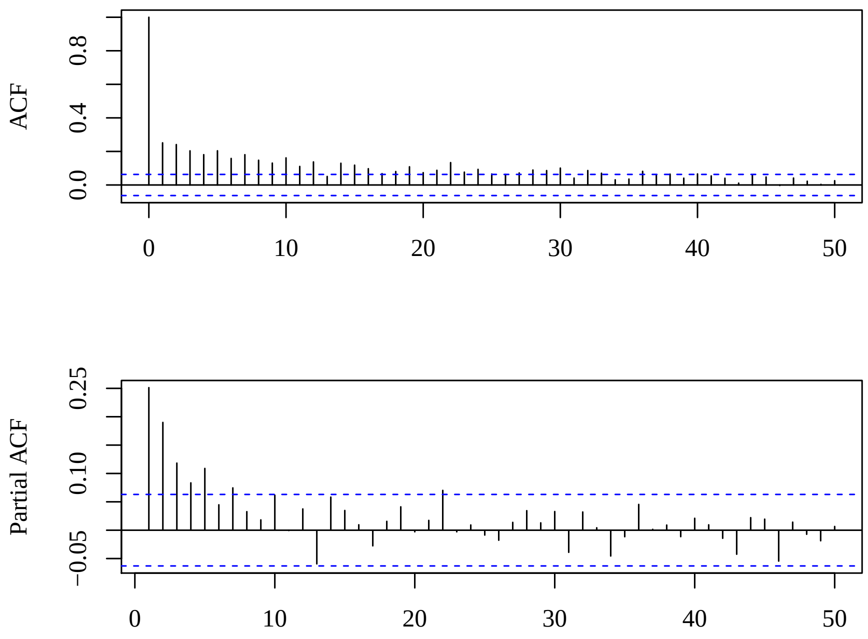


Figure 1: Plot of the operations with clients time series ($\times 1000$).



(a) ACF and PACF

Figure 2: ACF and PACF of the operations times series.

The residuals in both models (defined as $\hat{r}_t = y_t - \hat{\mu}_t$, see Section 5.2) were tested for the presence of serial correlation in the first 20 lags using a Ljung-Box test, corrected

Table 9: Fitted SYMARFIMA$(1, d, 0)$ with t_4 distribution for the operations time series.

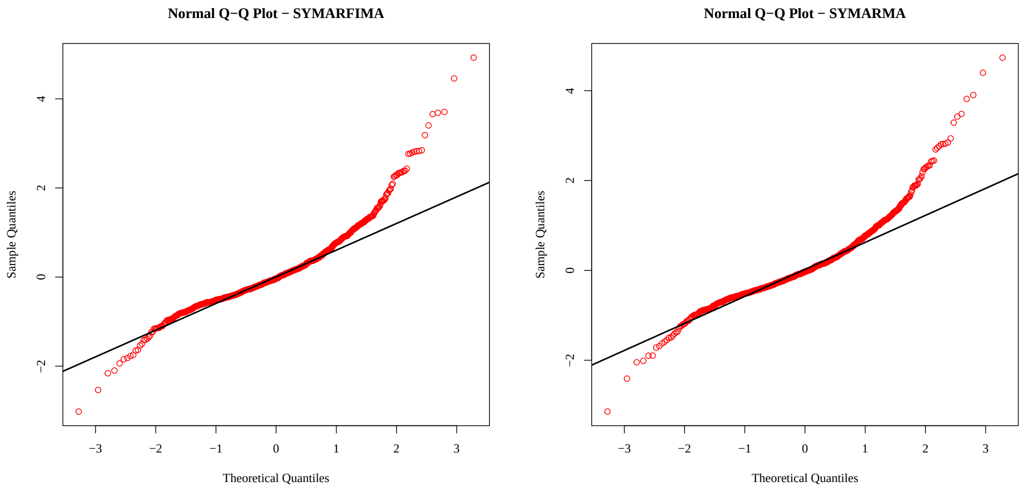
	Estimate	Std. Error	z stat.	Pr($>|z|$)
α	1.8878	0.0980	19.273	0.0000
d	0.2853	0.0297	9.600	0.0000
ϕ	-0.1140	0.0385	-2.964	0.0030
φ	0.2979	0.0183	16.295	-
Log-likelihood: -1061.9				
AIC: 2131.8	BIC: 2151.3			
Ljung-Box test (df = 19)			p-value = 0.529	

Table 10: Fitted SYMARFIMA$(5, 0)$ with t_4 distribution for the operations time series.

	Estimate	Std. Error	Pr($>|z|$)
intercept	1.8134	0.0543	0.0002
ϕ_1	0.1870	0.0260	0.0000
ϕ_2	0.1350	0.0262	0.0000
ϕ_3	0.0862	0.0264	0.0000
ϕ_4	0.0682	0.0262	0.0000
ϕ_5	0.1003	0.0259	0.0000
φ	0.3009	0.0182	-
Log-likelihood: -1062.6			
AIC: 2137.3	BIC: 2166.5		
Ljung-Box test (df = 15)			p-value = 0.347

for the number of fitted parameters, whose results are presented in Tables 9 and 10. Figure 3 show the residual QQ-plots for the residuals of both model. Since they are visually quite similar, we applied a two sample Kolmogorov-Smirnov test to check whether the residuals follow the same distribution. The test yielded p-value 0.996, so there is no evidence that the residuals are not identically distributed. From the QQ-plot we also conclude that there is a very strong departure from normality in the residuals, which was also verified upon applying a Shapiro-Wilk test to both residuals, yielding p-values 0 for all practical purposes.

The SYMARFIMA forecasts are about 20% more accurate than the SYMARMA ones, with a mean absolute error of forecast of 0.3157 and 0.3787 respectively and mean absolute percentage error of about 13.8% and 17.2%, respectively.



(a) SYMARFIMA$(1, d, 0)$ (b) SYMARMA$(5, 0)$

Figure 3: Residual QQ-plots from the (a) SYMARFIMA$(1, d, 0)$ and (b) SYMARMA$(5, 0)$ fit.

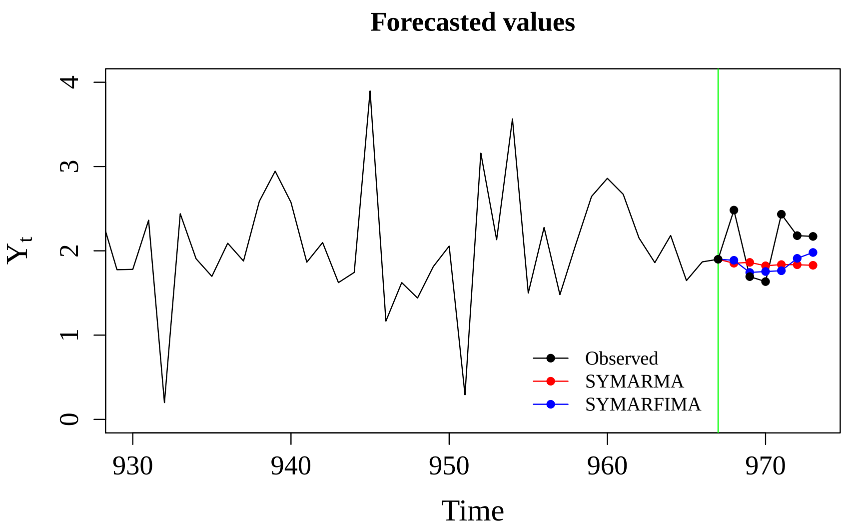


Figure 4: Forecasted values for the SYMARFIMA$(1, d, 0)$ and SYMARMA$(5, 0)$ fit.

8 Conclusion

In this work we introduce the class of SYMARFIMA models, which generalizes the class of SYMARMA models (Maior and Cysneiros, 2018) by allowing the conditional mean to present a long range dependence structure. We obtain conditions for existence and stationarity of the model as well as derive its mean, variance and covariance structure. We propose the use of a PMLE approach for parameter estimation, which allow the model to accommodate time dependent random covariates as well as deterministic ones. We provide closed forms for the score vector and the conditional Fisher matrix. We provide conditions for the consistency and asymptotic normality of the PMLE in the context of SYMARFIMA models, discuss hypothesis testing, confidence interval and forecasting. To assess the finite sample performance of the proposed estimation approach, we present a Monte Carlo simulation study. Finally, we present an application of the proposed methodology to a real data set.

Acknowledgments

H.S.C Benaduce acknowledges the financial support provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES Brazil. The authors are also thankful to Dra. Taiane Schaedler Prass for helpful comments and discussion.

References

Avram, F. and Taqqu, M. (1986). Weak convergence of moving averages with infinite variance. In E., E. and M.S., T., editors, *Dependence in Probability and Statistics*, volume 11 of *Progress in Probability and Statistics*. Birkhäuser, Boston, MA.

Bayer, F. M., Bayer, D. M., and Pumi, G. (2017). Kumaraswamy autoregressive moving average models for double bounded environmental data. *Journal of Hydrology*, **555**, 385–396.

Bayer, F. M., Cintra, R. J., and Cribari-Neto, F. (2018). Beta seasonal autoregressive moving average models. *Journal of Statistical Computation and Simulation*, **88**(15), 2961–2981.

Benjamin, M. A., Rigby, R. A., and Stasinopoulos, D. M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical Association*, **98**(461), 214–223.

Bondon, P. and Palma, W. (2007). A class of antipersistent processes. *Journal of Time Series Analysis*, **28**(2), 261–273.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer Science & Business Media.

Costa Ferreira, P., Speranza, T., and Costa, J. (2018). *BETS: Brazilian Economic Time Series*. URL `https://CRAN.R-project.org/package=BETS`. R package version 0.4.9.

Cox, D. R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., and Lauritzen, S. L. (1981). Statistical analysis of time series: Some recent developments [with discussion and reply]. *Scandinavian Journal of Statistics*, pages 93–115.

Fahrmeir, L. and Kaufmann, H. (1985). Consistency and asymptotic normality of the maximum likelihood estimator in generalized linear models. *The Annals of Statistics*, **1**(13), 342–368.

Ferrari, S. L. P. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, **31**(7), 799–815.

Fokianos, K. and Kedem, B. (1998). Prediction and classification of non-stationary categorical time series. *Journal of Multivariate Analysis*, **67**, 277–296.

Fokianos, K. and Kedem, B. (2004). Partial likelihood inference for time series following generalized linear models. *Journal of Time Series Analysis*, **25**(2), 173–197.

Granger, C. W. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, **1**(1), 15–29.

Hall, P. and Heyde, C. (1980). *Martingale limit theory and its application*. Probability and mathematical statistics. Academic Press, New York.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, **68**(1), 165–176.

Kedem, B. and Fokianos, K. (2002). *Regression models for time series analysis*. John Wiley & Sons.

Li, W. K. (1991). Testing model adequacy for some Markov regression models for time series. *Biometrika*, **78**(1), 83–89.

Ljung, L. (1999). *System Identification - Theory for the User*. Prentice Hall.

Lopes, S. R. and Prass, T. S. (2013). Seasonal FIEGARCH processes. *Computational Statistics & Data Analysis*, **68**, 262–295.

Maior, V. and Cysneiros, F. (2018). SYMARMA: a new dynamic model for temporal data on conditional symmetric distribution. *Statistical Papers*, **59**.

Palma, W. (2007). *Long-memory time series: theory and methods*, volume 662. John Wiley & Sons.

Prass, T. S., Bravo, J. M., Clarke, R. T., Collischonn, W., and Lopes, S. R. C. (2012). Comparison of forecasts of mean monthly water level in the Paraguay River, Brazil, from two fractionally differenced models. *Water Resources Research*, **48**(5). W05502.

Prass, T. S. (2013). *Estimação e previsão em processos SFIEGARCH: variância finita e infinita*. PhD thesis, Universidade Federal do Rio Grande do Sul, Porto Alegre.

Pumi, G., Valk, M., Bisognin, C., Bayer, F. M., and Prass, T. S. (2019). Beta autoregressive fractionally integrated moving average models. *Journal of Statistical Planning and Inference*, **200**, 196 – 212.

Pumi, G., Prass, T. S., and Souza, R. R. a. (2021). A dynamic model for double-bounded time series with chaotic-driven conditional averages. *Scandinavian Journal of Statistics*, **48**(1), 68–86.

Rocha, A. V. and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *TEST*, **18**(3), 529–545.

Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall International.