ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

DEPARTMENT OF STATISTICS
DOCTORAL PROGRAM

Modeling Multivariate Time Series

By
Tsamtsakiri Panagiota

Submitted to the Department of Statistics
of the Athens University of Economics and Business
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Athens, Greece
2023

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Διακριτές Πολυμεταβλητές Χρονοσειρές

Τσαμτσακίρη Παναγιώτα

Διδακτορική Διατριβή
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
για την απόκτηση
Διδακτορικού Τίτλου στη Στατιστική

Αθήνα
2023

DEDICATION

To my siblings Kiki and Iraklis

ACKNOWLEDGEMENTS

It is even more difficult and strange to me this because words are not enough to express my feelings and to thank all people who were close to me and they supported me with my thesis.

First of all i would like to express wholeheartedly my gratitude to my supervisor Prof. Dimitris Karlis who was the leader and he taught me on how to think, to search and to learn. I would like to thank him for his support, his aid, his countless revisions contributing by this way at the improvement of the thesis. I feel proud to be one of his students. Many thanks to Prof. Ioannis Ntzoufras for his support whenever i asked him for.

I owe a debt of gratitude to my parents Theodoros and Georgia for their whole-life support, valuable trust to this step of my life and infinite love. Words cannot describe how grateful i am to my siblings Kiki and Iraklis for always being there to encourage me for everything in my life. You are the best! I am much appreciated to my husband Bill for his patient, fostering and for being in my life.

Finally i wish to acknowledge State Scholarship Foundation in Greece for the scholarship that offered to me.

«This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (ΙΚΥ)»

ABSTRACT

Modeling Multivariate Time Series

by Tsamtsakiri Panagiota

The study of time series has become a subject of interest last years especially for count data with heteroskedasticity. While there are numerous research studies with univariate models, in multivariate case the literature is limited. The reason for the above is that there are countless difficulties at the construction and analysis of multivariate counting processes. There are two general categories of time series models with different constructions leading to different ways of their theoretical properties analysis. While considering heteroskedastic time series models from the category ergodic properties can be studied more easier than the the second one, methods for estimation are difficult to established.

This thesis deals with the construction of multivariate Integer Generalized Autoregressive Conditional Heteroskedastic(INGARCH) and Conditional Autoregressive Range(CARR) time series processes. The INGARCH(1,1) model has been also studied in multivariate case and implementations in 2 dimensions have been also presented recent years. The drawback was found at the restricted values of correlation coefficient boundaries depending on the way of model's construction. Based on a family of copulas a multivariate INGARCH(1,1) model and a CARR(1,1) model are studied considering interdependencies and self-dependencies respectively. According to model's complexity, its appropriateness and capability to study data with small sample sizes are examined and provided with simulations. H-steps ahead forecasting is considered by taking conditional expectation on volatilities and calculating marginal probability mass function.

Firstly, considering univariate INGARCH models where volatilities are linearly or log-linearly expressed offering more flexibility on conditions of stationarity respectively, a Bayesian Trans-dimensional Markov Chain Monte Carlo is provided. At a second stage a new alternative family of Sarmanov distribution is also presented in order to ameliorate boundaries of correlation coefficient and comparison with the known Sarmanov families are graphically discussed. A multivariate INGARCH(1,1) process is studied based on this alternative Sarmanov distribution and an implementation with daily crash counts on three towns in Netherlands is presented. A multivariate Conditional Aytoregressive Range(CARR(1,1)) model assuming an exponential distribution and reconstructing correlation coefficients boundaries is discussed. The proposed model is illustrated on bivariate series from the fields of renewable sources.

ΠΕΡΙΛΗΨΗ

Διακριτές Πολυμεταβλητές Χρονοσειρές

Τσαμτσακίρη Παναγιώτα

Η μελέτη των χρονοσειρών αποτελεί αντικείμενο έρευνας τα τελευταία χρόνια ειδικά των διακριτών μοντέλων με ετεροσκεδαστικότητα. Ενώ υπάρχουν πολλές μελέτες στην μονοδιάστατη περίπτωση , στην πολυδιάστατη η βιβλιογραφία είναι περιορισμένη. Ο λόγος για το παραπάνω είναι ότι υπάρχουν αρκετές δυσκολίες στην δομή και την ανάλυση πολυμταβλητών μοντέλων χρονοσειρών. Συμπεριλαμβάνοντας δύο γενικές κατηγορίες χρονοσειρών τα εμπόδια της πρώτης κατηγορίας σχετικά με την δομή και την ανάλυση, δεν είναι σημαντικές δυσκολίες της δεύτερης και το αντίστροφο. Λαμβάνοντας υπόψην τα παραπάνω ενώ στα ετεροσκεδαστικά μοντέλα η εργοδικότητα είναι εύκολο να μελετηθεί, η εφαρμογή μεθόδων για την εκτίμηση των παραμέτρων είναι δύσκολη. Επιπλέον όταν η σχετική συσχέτιση πρέπει να ληφθεί υπόψιν, η εκτίμηση του μοντέλου γίνεται περισσότερο πολύπλοκη.
Στην διατριβή αυτή η δομή ενός πολυμεταβλητού Integer Generalized Autoregressive Conditional Hteroskedastic(INGARCH) και ενός πολυμεταβλητών Conditional Autoregressive Range(CARR) μοντέλου χρονοσιερών μελετώνται. Το INGARCH(1,1) μοντέλο και εφαρμογές στις 2 διαστάσεις έχουν ήδη μελετηθεί τα τελευταία χρόνια. Το μειονέκτημα είναι ότι οι τιμές της συσχέτισης είναι περιορισμένες και αυτό προκύπτει απο την δομή του μοντέλου. Βασιζόμενοι σε μια οικογένεια copulas ένα πολυμεταβλητό INGARCH(1,1) και ένα CARR(1,1) μοντέλο μελετώνται θεωρώντας ότι τα volatilities εκφράζονται όχι μόνο με προηγούμενες τιμές τους και προηγουμενες τιμές της μεταβλητής αλλά και απο προηγούμενες τιμές άλλων μεταβλητών και παραμέτρων. Σύμφωνα με την πολυπλοκότητα του μοντέλου η καταλληλότητα και η ικανότητα για μελέτη δεδομένων όπου τα πλήθος τους δεν είναι μεγάλο, εξετάζεται με προσομοιώσεις. Πρόβλεψη h-βημάτων μπροστά σύμφωνα με την δεσμευμένη μέση τιμή στα volatilities γίνεται και οι περιθώριες κατανομές πυκνότητας υπολογίζονται.
Αρχικά θεωρώντας μονοδιάστατα INGARCH μοντέλα όπου τα volatilities εκράζονται γραμμικά ή λογαριθμικά γραμμικά προσφέροντας λιγότερη ή περρισότερη ελαστικότητα στους περιορισμούς των συνθηκών στασιμότητας αντίστοιχα, ένα Μπευζιανό Transdimensional Markov Chain Monte Carlo εφαρμόζεται. Έπειτα μια νέα οικογένεια κατανομών Sarmanov παρουσιάζεται με σκοπό να βελτιώσει τις τιμές των ορίων του συτελεστή συσχέτισης. Μια σύγκριση των τιμών του με τις ήδη υπάρχουσες παρουσιάζεται γραφικά. Ένα πολυμεταβλητό INGARCH(1,1) μοντέλο προτείνεται βασισμένο στην καινούρια οικογένεια copulas και μια εφαρμογή με δεδομένα απο τροχαία ατυχήματα σε τρείς πόλεις της Ολλανδίας παρουσιάζεται. Επίσης μελετάται ένα πολυμεταβλητό CARR(1,1) μοντέλο όπου η κατανομή είναι πολυδιάστατη εκθετική και τα όρια του συντελεστή συσχέτισης εξετάζονται. Οι εφαρμογές του προτεινόμενου μοντέλου είναι απο τον τομέα της βιώσιμης ανάπτυξης και των ανανεώσιμων πηγών ενέργειας.

Contents

List of Tables

List of Figures

List of abbreviations

ACP . . . . Autoregressive Conditional Poisson

AR . . . . AutoRegressive

ARCH . . . . . . AutoRegressive Conditional Heteroskedastic

ARMA . . . . . AutoRegressive Moving Average

BAR . . . . Bivariate AutoRegressive

BCP . . . . Bivariate Conditional Poisson

BIN . . . . BINomial

BNB . . . . Beta Negative Binomial

CARR . . . . . . Conditional AutoRegressive Range

CCC . . . . Conditional Constant Correlation

cdf . . . . Cumulative distribution function

CGINAR . . . . Combined Geometric INteger AutoRegressive

CPINAR . . . . Compound Poisson INteger AutoRegressive

DACP . . . . . . Double AutoRegressive Conditional Poisson

DCC . . . . Dynamic Conditional Correlation

DINGARCH . . Dynamic INteger Generalized AutoRegressive Conditional Heteroskedastic

DSTCC . . . . . Double Smooth Transition Conditional Correlation

EAR . . . . Exponential AutoRegressive

ECARR . . . . . Expoenntial Conditional AutoRegressive Range

ECCC . . . . . . Extended Conditional Constant Correlation

EGB2 . . . . . . Exponential Generalized Beta of second order

FCGARCH . . . Flexible Coefficient Generalized AutoRegressive Heteroskedastic

FGM . . . . Farlie-Gumbel-Morgenstern

FINGARCH . . Functional INGARCH

GARCH . . . . Generalized AutoRegressive Conditional Heteroskedastic

GEAR . . . . . . Generalized Exponential AutoRegressive

GED . . . . Generalized Error Distribution

GEV . . . . Gumbel Extreme Value

GFDCC . . . . . Quadratic Flexible Dynamic Conditional

GJR . . . . Glosten-Jagannnathan-Runkle

GLM . . . . Generalized Linear Models

GO-GARCH . . Generalized Orthogonal GARCH

GOF-GARCH . Generalized Orthogonal Factor GARCH

IGARCH . . . . Integrated Generalized AutoRegressive Conditional Heteroskedastic

i.i.d . . . . independent identically distributed

INAGARCH . . INteger Assymetric eneralized AutoRegressive Conditional Heteroskedastic

INBL . . . . . . INteger BiLinear

MA . . . . Moving Average

MCMC . . . . . Markov Chain Monte Carlo

MINARCH . . . Mixture INteger e AutoRegressive Conditional Heteroskedastic

MLE . . . . Maximum Likelihood Estimation

MSM . . . . . . Marginal Specific Modeld

NB . . . . Negative Binomial

NEAR . . . . . . New Exponential AutoRegressive

NIG . . . . Normal Inverse Gamma

NLAR . . . . . . New Laplace AutoRegressive

NMEAR . . . . New Multivariate Exponential AutoRegressive

OD . . . . Observation Driven

O-GARCH . . . Orthogonal Generalized AutoRegressive Conditional Heteroskedastic

PAGARCH . . . Periodic Asymmetric Generalized AutoRegressive Conditional Heteroskedastic

PDM . . . . . . Probability Distributed Model

pdf . . . . Probability Density Function

pgf . . . . Probability Generating Function

PIG . . . . Poisson Inverse Gauusian

pmf . . . . Probability Mass Function

PSINAR . . . . p-order Signed INteger AutoRegressive

PTARCH . . . . Power Tranformation Threshold AutoRegressive Conditional Heteroskedastic

QARCH . . . . . Quadratic AutoRegressive Conditional Heteroskedastic

QMLE . . . . . . Quasi Maximum Likelihood Estimation

RCAR . . . . . . Random Coefficient AutoRegressive

RCINAR . . . . Random Coefficient INteger AutoRegressive

RGS . . . . Regime Switching

RJMCMC . . . Reversible Jump Markov Chain Monte Carlo

RRINGARCH . Random Rounded INteger AutoRegressive Conditional Heteroskedastic

RrNGINGARCH Random Environment New

RSDC . . . . . . Regime Switching Dynamic Correlation

SINAR . . . . . Signed INteger AutoRegressive

STCC . . . . . . Smoothed Transition Conditional Correlation

STGARCH . . . Smooth Transition Generalized AutoRegressive Conditional Hteroskedastic

STINAR . . . . Skew True INtger AutoRegressive

TLAR . . . . . . Transposed Laplace AutoRegressive

TVCC . . . . . . Time Varying Conditional

VC-GARCH . . VArying Correlation Generalized AutoRegressive Conditional Heteroskedastic

ZI . . . . Zero Inflated

ZICP . . . . Zero Inflated Compound Poisson

Chapter 1

Introduction

1.1 Current models for time series

A series of observations $\{X_t\}_{t=0}^{T}$ is defined as a count series if $\{X_t\} \in \mathbb{N}_0$, $\forall$ t and as continuous series if $\{X_t\} \in \mathbb{R}$, $\forall$ t. Cox et al. (1981) classifies time-varying parameter models into two categories: observation driven(OD) models and parameter driven(PD) models. At the first one parameters depend on functions of lagged dependent values and past observations. An example of an observation driven model is:

$$X_t \mid x_{t-1}, y_t \sim Poisson(exp(y_t^T \beta + f(x_{t-1})))$$

where y_t is a p-dimensional row vector and β is a vector of regression coefficients and $f(\cdot)$ can be any function. Sometimes exogenous variables are needed on parameters expressions offering flexibility on the order of the model. Autoregressive(AR) and Integer valued AR(INAR) models considering different thinning operators proposed by Al-Osh and Alzaid (1987);McKenzie (1985), continuous ARMA, discrete ARMA(DARMA) and Integer valued ARMA(INARMA) models of Al-Osh and Alzaid (1988);Alzaid and Al-Osh (1990);Jacobs and Lewis (1978), Integer Autotregressive Conditional Heteroskedastic(INARCH) and Integer-valued Autoregressive Conditional Heteroskedastic models of Ferland et al. (2006):Fokianos et al. (2009);Fokianos and Tjøstheim (2011) are representative examples of OD models with applications to various disciplines such as finance Bollerslev (1986), epidemiology and public health Fokianos et al. (2009);Ferland et al. (2006);Fokianos and Tjøstheim (2011) On the other hand at parameter driven models parameters are stochastic processes that are subject to their own error. Furthermore at this category of models observed data are assumed to be independent given latent variables:

$$X_t \mid \epsilon_t, y_t \sim Poisson(exp(y_t^T \beta + \epsilon_t))$$

where y_t are defined as previously whilst ϵ_t is a latent process usually of a standard form as a classical time series model e.g. $\epsilon_t = \alpha\epsilon_{t-1} + r_t$ where r_t are i.i.d. $N(0, \sigma^2)$

Models studied by Zeger and Qaqish (1988) as an extension to the generalized linear model (GLM), stochastic volatility models proposed by Shephard (1995) and hidden Markov models MacDonald and Zucchini (1997) are examples at the class of PD models. On the other hand in a PD autocorrelation expressed by using latent variables. On the other hand at the category of OD models AR. OD and PD models have crucial different properties and usages. Parameter driven models are hard to estimate because likelihood function cannot be always written analytically and parameters estimation seems to be more challenging than these in OD models but stability properties including stationarity and mixing conditions are easier to established.

1.2 Continuous-valued time series models

1.2.1 Autoregressive time series models of order 1

The first stochastic model that was useful on the representation and study of series is the autoregressive model(AR) in which the current value of the process is expressed as a finite, linear aggregate of its previous values and a variable called innovation ϵ_t. The model is defined as:

$$X_t = \phi_1 X_{t-1} + \epsilon_t$$

or equivalently

$$\phi(B)X_t = \epsilon_t$$

where $\phi(B) = 1 - \phi_1 B$, B is a back shift operator defined by $BX_t = X_{t-1}$, ϵ_t is a white noise, a sequence of uncorrelated random variable with 0 mean and variance σ_ϵ^2. The above model is stationary under certain conditions. Another model expressed as a finite linear combination of white noise variables, proposed by GEORGE (1970) and defined as:

$$X_t = \epsilon_t - v_1\epsilon_{t-1}$$

or equivalently

$$X_t = v(B)\epsilon_t$$

where $v(B) = 1 - v_1(B)$ is a polynomial associated with the moving average(MA) model. Due to the expression of construction, the model is always stationary and it is capable to identify long-trends in data and smooth out short-term fluctuations. Numerous models have been studied and distinguished in two categories. By the mean of construction where a marginal distribution is firstly specified and then the error distribution is obtained, the first class called class of marginal specific models (MSM). In contrast in the second class of error specific models, the analysis of model properties is done by specifying a suitable distribution for the i.i.d. error variable ϵ_t without bothering on the exact form of the marginal distribution of the stationary time series X_t. A variety of the presented models is also considered to be part of random coefficient autoregressive models where the distribution of X_t derived from the distribution of error term ϵ_t.

The first non-Gaussian AR(1) model was the one with stationary exponential marginal distribution introduced by Gaver and Lewis (1980). Considering that ϵ_t is a sequence of i.i.d. non-negative random variables, chosen in such a way that X_t follows a specified stationary marginal distribution. That is, one has to find the form of the distribution of ϵ_t for a specified distribution of X_t considering properties of generating functions(g.f.) or characteristic functions (c.f.). This means that if the function of X_t is self-decomposable then there exists a well-defined distribution for the innovation ϵ_t as a consequence of $\phi_{X_t}(s) = \phi_{X_{t-1}}(\phi s)\phi_{\epsilon_t} s$ where ϕ is the characteristic function. Based on the above, Exponential Autoregressive EAR(1) model expressed as :

$$X_t = \epsilon_t + \begin{cases} \phi X_{t-1} & w.p. \quad \alpha \\ 0 & w.p. \quad 1-\alpha \end{cases}$$

The zero-defect of the EAR(1) model means that the relationship between the current observation and past observations decays exponentially over time and this was regarded as one of the drawbacks of that model. To overcome the problem Lawrance and Lewis (1985) constructed a transposed exponential autoregressive (TEAR) model keeping as marginals distribution exponential distributions. Moreover Lawrance and Lewis (1981) generalized the EAR(1) model by introducing other parameter ϕ known as new exponential autoregressive model (NEAR(1)). The benefit of this model is that it has an expression of simple random similar combinations of exponential independent and thus it is easy to estimate. A generalized version of the NEAR(1) model proposed by Sim (1990) in order to define a stationary sequence of exponential random variables. Efforts for constructions of Gamma AR models based on thinning operators have been made by Gaver and Lewis (1980); Sim (1990). The first Gamma AR model based on marginal specific models where marginal distributions are $Gamma(p, \lambda) = \frac{x^{p-1}\exp(-\frac{x}{\lambda})}{\lambda^p * \Gamma(p)}$, where $Gamma(p)$ is a Gamma function with $p, \lambda > 0$, shape and scale parameters respectively while at the second one the process defined as :

$$X_t = \lambda * X_{t-1} + \epsilon_t, \;\; t = 1, 2, \ldots$$

where $\lambda * X_t = \sum_{i=1}^{N(X)} E_i$, E_i are exponential random variables with parameter β and for each fixed value of X_t, $N(X)$ is a positive random variable with parameter $\lambda\beta x$ Pitt et al. (2002);Pitt and Walker (2005);Pitt and Walker (2006) suggested alternative methods of constructing the distribution of ϵ_t by specifying suitable conditional and marginal distributions. They introduced some latent variables to get the specified marginal distributions for the AR(1) sequences. A mixed exponential AR(1) model defined by Lawrance (1980) and a random coefficient AR(1) model with marginal mixed exponential distribution defined by Lawrance and Lewis (1985) where X_t follows an exponential distribution if the distribution of innovations is a mixture of exponential distributions. Similar to the case of EAR(1) model a Constant coefficient Laplace and a Random coefficient Laplace AR model referred as transposed Laplace AR(1)(TLAR) models also defined by Sim (1990). To bypass the problem of zero-defect in the LAR(1) process Sim (1993) studied an autoregressive process with symmetric Laplace marginal distributions with location 0 and scale parameter 1 as :

$$f(x) = \frac{1}{2} \exp(-|x|), \quad x \in \mathbb{R}.$$

The construction of an autoregressive time series model with stationary marginal Gumbel Extreme Value(GEV) distribution given by:

$$f(x; \mu, \sigma) = \exp\left\{ -\exp\left(-\frac{x-\mu}{\sigma} \right) \right\}, \;\; -\infty < x, \mu < \infty, \sigma > 0$$

was constructed to model extreme values of a dataset. This distribution is widely used in many areas such as engineering to model maximum wind speed, in finance to model daily loss or gain in stock prices. Based on the above distribution Toulemonde et al. (2010) and Balakrishna and Shiji (2014) constructed Gumbel's extreme value AR model summarizing the properties of a first order model. Many other stationary autoregressive time series processes also constructed considering each time different innovation's distribution. At the bibliography there are references of Cauchy AR process discussed by Tauchen and Pitts (1983), Logistic AR process constructed by Sim (1993) and Hyperbolic Secant autoregressive processes Rao and Johnson (1988). They showed that the autocorrelation function of stationary processes mentioned previously can decay slowly, which implies that the processes exhibits long-range dependence. Sometimes, suitable mixtures of more than one standard distribution may be a better choice compared to any single distribution. Normal-Laplace is one such distribution introduced by Reed and Jorgensen (2004), which is a convolution of normal and Laplace distributions. Jose et al. (2008) and Tomy and Jose (2009) discussed first-order autoregressive processes with normal-Laplace and generalized normal-Laplace stationary marginal distributions, where such processes provide the combined effect of Gaussian and non-Gaussian time series models. There are also other time series models with similar stationary marginal distributions in the literature. For example, Marshall–Olkin Esscher transformed Laplace distribution (George and George, 2013), mixed asymmetric Laplace distribution (Krishnan and George, 2017) and the references therein.

By combining moving average and autoregressive models, one gets what is known as autoregressive moving average(ARMA) model defined as:

$$X_t = \phi_1 X_{t-1} + \theta_1 \epsilon_{t-1}, \epsilon_t \sim N(0, \sigma^2)$$

where $\mathcal{F}_{t-1}$ is the σ-field algebra generated by $X_0, X_t, \quad t \in \mathbb{N}$ $\sigma_t = var(X_t \mid \mathcal{F}_{t-1}) = Var(\epsilon_t \mid \mathcal{F}_{t-1})$. The advantage of the above model is that it gives a more parsimonious model with relatively few unknown parameters. ARMA models with exponential stationary marginal distributions were introduced to study the sequences of non-negative random variables. For example, Jacobs (1978);Jacobs (1980) analysed the queuing networks using exponential ARMA(1, 1) model. Balakrishna and Ranganath (2015) analysed Bombay Stock Exchange data using ARMA(1, 1) model with generalized error distributed innovations. While the range of marginal distributions becomes too restricted then the innovations have non-standard distribution with discontinuities. Lewis et al. (1989) introduced a class of random coefficient models to generate a dependent sequence of $Gamma(p, \lambda)$ random variables where, the coefficients are i.i.d. Beta random variables.

The logical extension of AR(1) model to AR(2) has been introduced by GEORGE (1970) where an Autoregressive Integrated Moving Average(ARIMA) model was introduced and an AR(2) was included as a special case . Of course $(X_{t-1},\ X_{t-2})$ follows a bivariate normal distribution for a Gaussian AR model and hence it is easy to solve this problem. In order to avoid such complications in non-Gaussian cases, Lawrance and Lewis (1980) proposed an AR(2) modelbased on the work of Gaver and Lewis (1980),

Lawrance and Lewis (1985) studied simultaneously Exponential autoregressive models(EAR) of order 1 and 2. The NEAR(1) has been generalized to an NEAR(2) model identified as Random Lag Autoregressive RLAR(2) model incorporating randomness of the selection in terms. Numerous models have been studied in order to capture complex relationships between past and current values of time series providing flexibility at models forecasting. Such models have been studied by by

Jayakumar and Kuttykrishnan (2007) and Kim and Lynne (1997). Several models where the authors concentrate on the definition, the existence of the innovation distribution for specified marginal distributions and their second-order properties have been proposed in bibliography by Box et al. (2015), Trindade et al. (2010), Polasek and Pai (2004). Many other models were studied such autoregressive models with normal inverse Gaussian innovations(NIG), finite mixture autoregressive models Le et al. (1996);Wong and Li (2000);Wong et al. (2009);Nguyen et al. (2016) autoregressive models with slowly varying innovations,Davis and McCormick (1989);Feigin and Resnick (1992);Abraham and Balakrishna (1999);Ing and Yang (2014).

1.2.2 Heteroskedastic continuous-valued time series models

Extreme values in return series with financial data have been captured with heteroskedastic time series models. The idea for the construction of the first heteroskedastic model has been generated considering a non-standard ARMA model for ϵ_t^2, introduced by Engle (1982) well known as Autoregressive Conditional Heteroskedastic(ARCH) model and has the following form

$$\epsilon_t = \sigma_t \epsilon_{t-1}, \;\; \sigma_t^2 = \omega + \sum_{i=1}^{m} a_i \epsilon_{t-i}^2$$

where ϵ_t is the error term, σ_t^2 is the conditional variance of the error term at time t, $\omega > 0$ and $a_i \geq 0$ for $i \geq 1$. One of the weaknesses of the above model is that both positive and negative values of ϵ_{t-i}^2 provide similar effect on volatility as it depends on their previous values. Moreover based on the construction of volatility parameters $a_i, \;\; i = 1, 2, ...$ must be positive and considering values of series moments, values of a_i's are too restrictive. Bollerslev (1986) proposed a Generalised Autoregressive Conditional Heteroskedastic (GARCH) model in order to gives flexibility for the model of time-varying conditional variance assuming that the variance of the error term expressed not only of squared error term values at previous lags but also of the conditional variance at previous lags.

$$\sigma_t = \epsilon_t Z_t,$$

where

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} a_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \sigma_{t-j}^2.$$

After comparison with the ARCH model, GARCH model is more parsimonious because its volatility has terms of variance in past. One drawback of the model is that it takes into account only the movement of z_t. Conditonl of stationarity are some of the properties that have been discussed at the study of Bollerslev (1986) A potential disturbance was that the unconditional variance of the process to be modelled did not exist.Nelson (1991) also showed that when an Integrated Generalized Conditional Heteroskedastic (IGARCH) process starts at some infinite time point, its behaviour depends on the intercept ω and if the intercept is positive then the unconditional variance of the process grows linearly with time. Considering the fact that a GARCH model could be expressed as an ARMA model equivalently an IGARCH model could be expressed as an ARIMA model. An attempt to find which parameters are capable to decrease slowly autocorrelation function, was done by Baillie et al. (1996) and the models is known as first fractionally integrated GARCH(FIGARCH) model. In these models, the underlying volatility of the data is not directly observable and instead is modeled as a hidden variable that evolves over time according to some stochastic process. The purpose is to consider parameters using and modeling data at a high frequency within a single day named intra-daily data. Engle (2002) proposed a GARCH-X model including exogenous variables r_{t-l} as measures of intraday volatility. While the model was more flexible allowing modeling intradaily data, the shortcoming is that r_t is independent of σ_t. An attempt allowing flexibility on the range of parameters' values introduced by Nelson (1991) where volatility expressed by a log-linear form. Exponential Genaralized Conditional Heteroskedastic(EGARCH) and Exponential Genaralized Conditional Heteroskedastic with exogenous variables(EGARCH-X) are extensions of the ARCH/GARCH models and allows for asymmetric responses of volatility.

Another class of models assuming additional latent volatility processes for each exogenous variable considered at the volatility expression for modeling high frequency data within a single day , has been composed considering the models proposed by {Engle and Gallo (2006),Shephard and Sheppard (2010)}. Assuming numerous functions as density functions for error term, Engle and Gallo (2006) studied estimators consistency and asymptotic normality are consistent and asymptotically normal. The shortcoming of poor forecasting performance convey structural changes of the ARCH process leading to studies of other model proposed by Hamilton (1989), Diebold and Inoue (2001), Lamoureux and Lastrapes (1990), Hamilton and Susmel (1994),Engle and Lee (1999), Granger and Hyung (2004) and Medeiros and Veiga (2004). An important issue at this class of models was made by Hansen et al. (2012) who introduced models with endogenous variables to model intradaily data linked to the return variance by one more equation in the model. Furthermore based on the work of Nelson (1991) a Realized EGARCH(1,1) model had been also discussed. A Flexible Coefficient Generalized Autoregressive Conditional Heteroskedastic (FCGARCH) model with many features observed in empirical financial distributions including time-varying conditional variance, asymmetry, thick tails, and high peakedness. At the above models conditional distribution of y_t is an exponential generalized beta distribution of the second kind(EGBD2) introduced by McDonald (1991);McDonald (2008) and studied further from Wang et al. (2001).

Definition 1. *The exponential generalized beta distribution defined as*

$$EGBD2(z;\delta,\sigma,p,q)=\frac{\exp\left(\frac{p(z-\delta)}{\sigma}\right)}{|\sigma|B(p,q)\left(1+\exp\left(\frac{z-\delta}{\sigma}\right)^{p+q}\right)}$$

where δ is the location parameter that affects the mean of the distribution, σ reflects the scale of the density function, and p and q are shape parameters that together determine the skewness and kurtosis of the distribution.

An other class of GARCH models capable to accommodate asymmetries in the response have been studied by Glosten et al. (1993);Engle and Ng (1993);Sentana (1995). Construction, study and modeling of heteroskedastic models where not only variance but also standard deviation given by a volatility expression proposed by Ding et al. (1993);Taylor (2008);Schweer and Weiß (2014). Moreover considering special cases of GARCH model Duan (1997) constructed an augmented GARCH model. The first non-linear model named smooth transition GARCH model(STGARCH) proposed by Hagerud (1997);González-Rivera (1998);Anderson et al. (1999). Lanne and Saikkonen (2005) Zakoian (1994) and Rabemananjara and Zakoian (1993) proposed two other threshold models without a wide acceptance due to complex construction and behavior of standard deviation. Li and Li (1996) studied a double threshold ARCH(DTARCH) model where both conditional mean and conditional variance have a threshold-type structure as defined from Tong (1990). The observed problem at the DTARCH and STGARCH model is that when $k \to \infty$ then the models are not identifiable. To overcome the above problem Hagerud (1997) and González-Rivera (1998) proposed some criteria. Engle III and Sheppard

(2001) referred that at applications the assumption of GARCH models having constants parameters could not be appropriate when the series are too long. To bypass the problem, the assumption that parameters change at specific points of time, divide the series into subseries provided a time-varying GARCH(TV-GARCH) model. An other one way to model data containing breaks is the construction of Hidden Markov models or Markov switching models. Hamilton and Susmel (1994) studying a Markov-switching model in order to model very large shocks of stocks in 1987. Cai (1994) proposed an alternative construction of the above model. Furthermore Rydén et al. (1998) showed that a simplified model of this proposed by Cai (1994) was capable to model daily data with high frequency. Numerous other Markov-switching models have been also proposed by Granger and Ding (1995);Gray (1996);Klaassen (2002);Haas et al. (2004);Liu (2006);Lange and Rahbek (2009). An overview of ARCH/GARCH models had been given by Teräsvirta (2009). The increased automatization of financial markets and intradaily data in exchanges lead to study and modeling time between events. Moreover in the field of finance daily or monthly differences of $max(P_t) - min(P_t)$ where P_t is the logarithmic price of high-return assets given by a Geometric Brownian motion leads to the importance of a new heteroscedastic time series process. Considering the above, two models have been also discussed. The first namely Autoregressive Conditional Duration(ACD) model studied by Engle and Russell (1998) and Dufour and Engle (2000). Assuming that $\{t_0, t_1, \ldots, t_n\}$ be a sequence of arrival times and $x_i = t_i - t_{i-1}$ denotes the duration between events occurred at t_{i-1} and t_i. The model for durations X_i defined as:

$$X_t = \lambda_t \epsilon_t$$

where $\lambda_t = E(X_t \mid \mathcal{F}_{t-1})$ linked linearly with its previous values and past observations of x_t representing the time since the last event at time t and ϵ_t is a sequence of i.i.d. random variables. The second model known as Conditional Autoregressive Range(CARR) model given by:

$$R_t = \lambda_t \epsilon_t \tag{1.1}$$

where R_t is the range of the financial asset at time t, defined as the difference between the high and low prices over a given time period, λ_t is equal to conditional expectation of R_t up to time t representing the volatility and expressed linearly with range values of previous time periods and ϵ_t is a sequence of positive independent and identical random variables of the normalized range. The main difference between those two models is that while the time interval in the CARR model is constant in the ACD model is not. For both models numerous distributions for the error terms have been discussed. The standard ACD model is the Exponential ACD(1,1) which is the counterpart of the GARCH(1,1) stochastic process. Firstly it has been studied by Engle and Russell (1998) and then autocorrelation function and some other properties have been discussed by Bauwens and Giot (2000). Furthermore an alternative model allows for flexibility at parameter values of volatility construction discussed by Bauwens et al. (2001) where logarithmic scale of volatility defined similar to this of Fokianos and Tjøstheim (2011). Properties of ergodicity and stationarity of first order linear and log-linear Exponential ACD processes studied by Carrasco and Chen (2002) and Meitz and Teräsvirta (2006). studied a Generalized Gamma ACD model observing that the data fit the model better than other models. Zhang et al. (2001) introduced a Threshold Autoregressive Conditional duration model(TACD) based on the threshold GARCH model studied by Zakoian (1994);Rabemananjara and Zakoian (1993) in order to capture asymmetries in volatility. To model and forecast the dynamics of the price range, Chou (2005) studied the first Conditional Autoregressive Range(CARR) model that is very similar to the GARCH model of Bollerslev (1986) and the autoregressive conditional duration (ACD) model of Engle and Russell (1998). Then Xie and Wu (2017) proposed a Gamma CARR model and Chou (2006) studied an asymmetric CARR model. Chou (2005) and Chou et al. (2010) show that the range-based CARR model produces more accurate volatility Fernandes et al. (2005)and Chou et al. (2015) studied a CARR model when the range expressed from the difference max(Price)-min(Price) where prices are daily or monthly. At the second stochastic process studied the time between events. The main difference is that at the first model the time interval is fixed but at the second is not. Our study is concentrated at a bivariate CARR model where volatilities expressed linearly by their past values and past volatilities. Assume that $R_t = \lambda_t \epsilon_t$ where ϵ_t is a sequence of positive independent and identical random variables with mean equal to 1 and λ_t is equal to conditional expectation of R_t up to time t. A drawback with models in financial field is that autocorrelation function is dicreased up to time 5 and the order of the model becomes large. A way to simplify the model is to incorporate exogenous variables in the volatility. Chou (2005) made a brief discussion by proposing the CARRX model where R_t given again by (1.1) and $\lambda_t = d_0 + \sum_{i=1}^{p} a_i \lambda_{t-i} + \sum_{j=1}^{q} b_j R_{t-j} + \sum_{l=1}^{L} \gamma_{t-l} X_{t-1,l}$. Yatigammana et al. (2016) studied a model where errors follow a Weibull distribution. Similar to GARCH models Chiang et al. (2016) and Bauwens and Giot (2000) proposed lognormal log-CARR and log-ACD model respectively providing flexibility at the range of parameters values considering stationarity conditions. combine a CARR model with a DCC model studied from Engle (2002) using ranges to replace the GARCH volatilities in the first step of DCC

1.3 Integer-valued time series models

1.3.1 Integer-valued heteroskedastic time series models

The construction of integer-valued time series models was conducted from the necessity of modeling data as the number of incidences from a certain disease or the number of people waiting at a cashier in discrete time points. Numerous models for studying data of various disciplines such as sports, epidimiology, ecology and marketing proposed by Grunwald et al. (2000). The first integer-valued time series model was firstly introduced by {McKenzie (1985);Al-Osh and Alzaid (1987)}

Integer-valued heteroskedastic time series models have a vital role based on their construction where parameters or low order conditional moments expressed via past observations and past values assuming that those conditional moments change over time. Data with extreme values over time could be study by using heteroskedastic models. The first Integer-valued Autoregressive Conditional Heteroskedastic (INARCH) model given by :

$$X_t \mid \mathcal{F}_{T-1} \sim Poisson(\lambda_t) \tag{1.2}$$

where

$$\lambda_t = d_1 + a_1 \lambda_{t-1}.$$

$\mathcal{F}_{t-1}$ is the σ-algebra defined as $\{X_s, s \leq t\}$ and parameter $a_1 \in [0, 1]$. At the INARCH model dispersion of data derived from the construction of conditional moment λ_t. Coefficient (a_i) values affect moment λ_t at time t. Those values at volatility lead to data with extreme values. Furthermore a wider range at volatility permitted considering not only past values of its own but also past observations. The above model is the almost known Integer-valued Generalized Autoregressive Conditional Heteroskedastic(INGARCH) model and defined as:

$$X_t \mid \mathcal{F}_{T-1} \sim Poisson(\lambda_t) \tag{1.3}$$

where $\mathcal{F}_{t-1}$ is the σ-algebra defined as $\{X_s, s \leq t\}$ and

$$\lambda_t = d_1 + a_1 \lambda_{t-1} + b_1 X_{t-1}. \tag{1.4}$$

At the graphs below we present data studied from an integer valued autoregressive time series(INAR(1)) plot with Poisson innovations defined as:

$$X_t = a_1 \circ X_{t-1} + R_t$$

where

$$R_t \sim Poisson(\lambda) \;\; and \;\; a_1 \circ X = \sum_{i=1}^{X} Y_i$$

and data studied by an integer autoregressive heteroskedastic model(INGARCH(1,1)) :

$$X_t \mid \mathcal{F}_{T-1} \sim Poisson(\lambda_t) \tag{1.5}$$

where $\mathcal{F}_{t-1}$ is the σ-algebra defined as $\{X_s, s \leq t\}$ and

$$\lambda_t = d_1 + a_1\lambda_{t-1} + b_1 X_{t-1}$$

Furthermore OD models are a way to study data where autocorrelation function

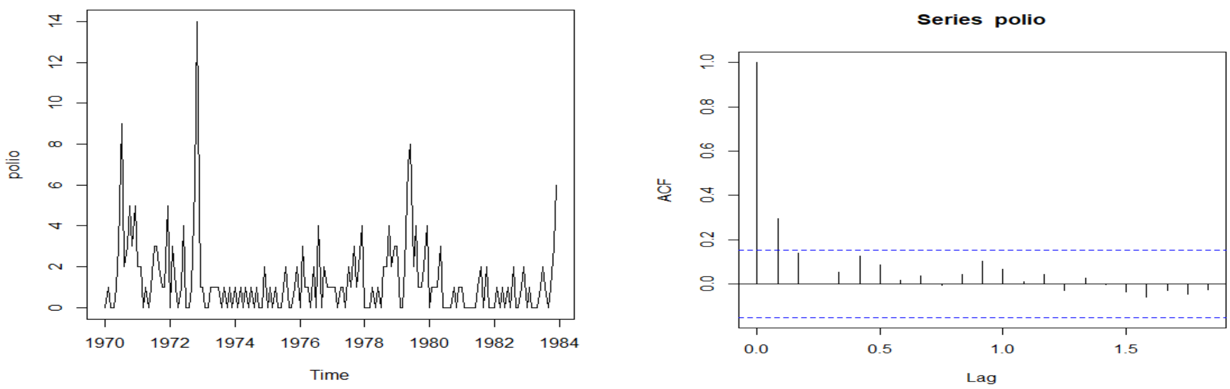


Figure 1.1: Series of polio dataset (left) and its ACF (right)

decays slowly. Rydberg and Shephard (2000) introduced an heteroskedastic time series process modeling the price of an asset at time t $p(t)$ given by

$$p(t) = p(0) + \sum_{i=1}^{N(t)} Z_i$$

where $N(t)$ is the number of trades recorded up to time t and Z_i is the change of price associated with the i^{th} transaction. An alternative expression for the price $p(t)$ given by:

$$p_t = p((t+1)\Delta) - p(t\Delta) = \sum_{i=1}^{[N(t+1)\Delta]} Z_i - \sum_{i=1}^{N(t)\Delta} Z_i = \sum_{i=N(t\Delta)+1}^{N[(t+1)\Delta]Z_i}$$

Considering the idea to "bin" out the time interval and count the number of transactions the first integer valued conditional autoregressive heteroskedastic model was constructed for the number of transactions that occurs in the interval $[n\Delta, (n+1)\Delta]$ and expressed as: $N_t = N[(t+1)\Delta] - N[t\Delta]$. Streett (2000) discussed the BIN(1,0) and BIN(1,1) models considering the first autoregressive heteroskedastic (ARCH(1)=GARCH(1,0)) process and the generalized heteroskedastic process (GARCH(1,1)) introduced by Bollerslev (1986) where data come from a normal distribution and variance conditional the past changes over time. Numerous applications have been also developed for several economic phenomena. An attempt for modeling count data with over or underdispersion Heinen (2003) studied two models by replacing the Poisson distribution with double Poisson distribution proposed by Efron (1986) and facing with the problem of normalizing constant's intractability and moments. At the first one, double autoregressive conditional Poisson model($DACP_1(1,1)$) variance is proportional to the mean and at the second one $DACP_2(1,1)$ variance expressed as a quadratic function. Furthermore Heinen (2003) considered only the expression for the number of transactions rename the model as autoregressive conditional Poisson(ACP) and studied some theoretical properties. Ferland et al. (2006) discriminated that at the ACP model when $a_1, b_1 > 0$ then the model could be expressed as an autoregressive moving average model of order 1(ARMA(1,1)) where volatility is expressed by (1.4). Additionally if $a_1 = b_1 = 0$ then the model of i.i.d. data has a mean $\lambda_t = b_0$. and when $a_1 = 0$ then the model reduced to an autoregressive model of order 1. A brief discussion for generalised double autoregressive conditional Poisson models has also been done from Heinen (2003). Zhu (2012a) examined properties of an INGARCH model assuming a Generalized Poisson distribution.

Definition 2. *Let $\{X_t\}$ be a time series of counts. Assuming that conditional on the past information the random variables X_i, $i = 1, \ldots, n$ are independent and the conditional distribution of X_t is specified by a generalised Poisson(GP) distribution:*

$$X_t \mid \mathcal{F}_{t-1} \sim \mathcal{GP}(\lambda_t^\star, k), \quad \frac{\lambda_t^\star}{1-k} = \lambda_t = d_1 + a_1\lambda_{t-1} + b_1 X_{t-1}$$

where $b_0 > 0$, $a_1, b_1 \geq 0$, max(-1,$-\lambda_t^\star/4$)<k<1, $\mathcal{F}_{t-1}$ is the σ-field generated by $\{X_s : s < t-1\}$. When $a_1 = 0$ then the model is called GP-INARCH(1) model and when k=0 then the model reduced to the INGARCH(1,1) model defined by (1.3)

On the other hand to model underdispersion a binomial $f(x=k) = \binom{n}{k} p^k (1-p)^{n-k}$, $n \in \mathbb{N}, p \in [0,1]$ and a negative binomial $f(x=r) = \binom{n+r-1}{n}(1-p)^n p^r$,

$r, n \in \mathbb{N},\ p \in [0,1]$ distributions have been proposed but differentiation with respect to parameter n is problematic and as a result difficulties with estimation have been provoked. Zhu (2011) constructed a univariate INGARCH Negative Binomial process considering that n is a known parameter. To estimate the parameter vector, maximum likelihood approach has been provided but for the choice of best initial values Yule-Walker and Conditional Least Squares methods have been also discussed. Chen et al. (2016) studied a Negative Binomial OD model with a time-varying conditional autoregressive mean where shape and scale parameters are considered unknown and Bayesian MCMC sampling used for parameters estimation. A Neyman type A OD model also discussed by Gonçalves et al. (2015) assuming a Compound Poisson law where its compounding distribution is also a Poisson law. Further study and analysis of an CP-INGARCH(1,1) time series process has been done from Silva (2016). Furthermore the same year Gonçalves et al. (2016) studied a zero-inflated integer-valued GARCH with general Compound Poisson deviates(ZICP-INGARCH(1,1)). A Zero-Inflated Compound Poisson Integer-valued GARCH model with orders 1 and 1, is calculated based on characteristic function of $X_t \mid \mathcal{F}_{t-1}$, $\forall t \in \mathbb{Z}$:

$$\Phi_{x_t|X_{t-1}}(X_t \mid \mathcal{F}_{t-1})(u) = \omega + (1-\omega)\exp\left\{\left(\frac{\lambda_t}{\phi_t'(0)}[\phi_t(u) - 1]\right) i\right\}, \quad u \in \mathbb{R},$$

λ_t defined by (1.4) for some constants $0 \leq \omega < 1$, $b_0 > 1$ and where $(\phi(t), t \in \mathbb{Z})$ is a family of characteristic functions on $\mathbb{R}$ and i denotes the imaginary unit.

The ZICP-INGARCH(1,1) model when $a_1 = 0$ is denoted as ZICP-INARCH(1) model. While $\omega = 0$ the above model reduced to a $CP - INGARCH(1,1)$ model. Moreover the model has exceptional forms of zero-inflated Poisson(ZP-INGARCH(1,1)) and negative binomial (NBINGARCH(1,1)) in case of $0 \leq a_1 < 1$

In case of numerous datasets with an excessive number of zeros, heteroskedastic models have been discussed by Zhu (2012c) based on the almost proposed zero-inflated Poisson distribution by Xie et al. (2001);Yang et al. (2011) and on weighted Poisson distribution proposed by Del Castillo and Pérez-Casany (1998);del Castillo and Pérez-Casany (2005) respectively. Based on the construction of volatility, parameter values are too restricted. For the above reason, Fokianos and Tjøstheim (2011) studied a new model where considering volatility's construction negative and positive parameters values are allowed. To overcome this problem another model based on theory of GLM has been proposed by Fokianos and Tjøstheim (2011) assuming

$$X_t \mid \mathcal{F}_{t-1} \sim Poisson(\lambda_t) \tag{1.6}$$

where

$$v_t = d_1 + \sum_{i=1}^{p} a_i v_{t-i} + \sum_{j=1}^{q} b_j \log(X_{t-j} + 1), \;\; v_t = \log(\lambda_t) \tag{1.7}$$

In this model, parameters d_1, a_i, b_j for any $i = 1, \dots, p$, $j = 1, \dots, q$ take values in $\mathbb{R}$ and both negative and positive correlations can occur. Certainly other specifications are possible. In particular, one may consider a model for the log-mean process by introducing $\log(X_{t-j} + u)$, where u is a constant. Fokianos and Tjøstheim (2011) presented results of the data analysis that do not indicate any gross deviations in terms of the mean square error of residuals for values of u varying from 1 to 10 with a step equal to 0.5. Bentarzi and Bentarzi (2017) studied a periodic INGARCH(1,1) model with the requirement of establishment of first and second order conditional moments. An alternative construction proposed by Xu et al. (2012) where not only conditional mean expressed similar to this of initial INGARCH model given by (1.3) but also conditional variance is assumed to have a constant ratio to the conditional mean. The new dispersed integer-valued autoregressive heteroskedastic (DINARCH) model is capable to study data with overdispersion or underdispersion and this flexibility can be reflected by the value of the defined ratio. General DINARCH(p) model, where the process $X_t \mid \mathcal{F}_{t-1} \sim P(\lambda_t)$, $\lambda_t = b_0 + \sum_{i=1}^{p} a_i X_{t-i}$, $E[X_t \mid \mathcal{F}_{t-1}] = \lambda_t$, $Var[X_t \mid \mathcal{F}_{t-1}] = \alpha\lambda_t$ where λ_t satisfies (1.3) and α is assumed constant. In case of $a < 1$ model is appropriate for data with overdispersion but datasets with underdispersion are used at the proposed DINGARCH(p,q) model. A negative binomial(NB), a double Poisson (DP) and a generalized Poisson(GP) DINARCH processes have been proposed based on Kotz et al. (2004);Efron (1986) and Gelfand and Dalal (1990);Consul and Jain (1973) respectively. Additionally considering exceptional cases such as this of a zero-inflated Poisson INGARCH or a weighted Poisson in case of numerous zeros have been also studied by Xie et al. (2001);Yang et al. (2011) and Zhu (2012a) respectively. The NB-DINARCH model is defined assuming a NB distribution $NB(r_t, \rho_t)$ with a pmf:

$$f(x_t \mid \mathcal{F}_{t-1}) = \frac{\Gamma(r_t + x_t)}{\Gamma(x_t + 1)\Gamma(r_t)} \rho_t^{x_t} (1 - \rho_t)^{r_t}$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$. First and second order moments defined as $E_{NB}(X_t \mid \mathcal{F}_{t-1}) = r_t\rho_t/(1 - \rho_t)$ and $Var_{NB}(x_t \mid \mathcal{F}_{t-1}) = r_t\rho_t/(1 - \rho_t)^2$ but considering the general equations from $DINGARCH$ model $r_t = \frac{\lambda_t}{1-\alpha}$ and $\rho_t = 1 - \frac{1}{\alpha}$ obtaining that conditional variance change over time. The DP-DINARCH model has been also studied by Heinen (2003) and was called $DACP_1$. At the GP-DINARCH model where $GP(\lambda^\star, k)$ referred previously after matching conditional moments with those at the general case $\lambda_t^\star = \frac{\lambda_t}{\sqrt{(\alpha}}$ and $k = 1 - \frac{1}{\alpha}$. An initiative idea for a mix construction of a beta-negative binomial integer-valued heteroskedastic(BNB-INGARCH) model that reduces to a Poisson -INGARCH or a NB-INGARCH model considering some restrictions and based on the study of Wang (2011) discussed by Gorgi (2020). A Beta Negative Binomial ($BNB(\lambda, r, \alpha)$) distribution is defined as:

$$f(x) = \frac{\Gamma(x+r)}{\Gamma(x+1)\Gamma(r)} \frac{B\{a+r, (a-1)\lambda/r + x\}}{B\{a, (a-1)\lambda/r\}}$$

with dispersion parameter $r > 0$, mean $\lambda > 0$ and tail parameter $a > 1$. When $a \to \infty$ the BNB converges to a NB distribution with success probability $\frac{\lambda}{r+\lambda}$ and in case that $r \to \infty$ the BNB converges to a Poisson distribution with mean λ. A hysteretic Poisson INGARCH process with advantages of modeling data with overdispersion and asymmetry studied by Truong et al. (2017). At the hysteretic Poisson INGARCH model a hysteretic variables responsible for dynamic switching modeling and and it can be an exogenous variable. Chen and Lee (2016) introduced generalized threshold Poisson INGARCH model with the main difference of relaxing the sharp transition in the parameter changes. An integer-valued process with positive and negative observations has been studied by Alomani et al. (2018) using a symmetric Skellam distribution and trying to model data with overdispersion with time-varying variance. An integer-valued asymmetric generalized autoregressive conditional heteroskedastic(INAGARCH) model was proposed by Hu and Andrews (2021) and defined as:

$$\begin{aligned}
X_t &= Z_t Y_t, \;\; Y_t \mid \mathcal{F}_{t-1} \sim Poisson(\lambda_t) \\
\lambda_t &= \left(\sqrt{1+4\eta_t} - 1\right)/2 \\
\eta_t &= a_{00} + \sum_{i=1}^{p} a_{0i} \left(|X_{t-i}| - \gamma_0 X_{t-i}\right)^2 + \sum_{j=1}^{q} \beta_{0j} \eta_{t-j},
\end{aligned} \tag{1.8}$$

where $\mathcal{F}_{t-1}$ is the σ-field , $a_{00} > 0, a_{1i} > 0, b_{1j} > 0$, $i = 1, \ldots, p$ and $j = 1, \ldots, q$, $p \geq 1$, $q \geq 0$. Furthermore, following Straumann (2005) parameter values a_{00}, a_{1i}, b_{1j} are all unique and the polynomials $a_0(x) = \sum_{i=1}^{p} a_{0i} x^i$ and $\beta_0(x) = \sum_{j=1}^{q} \beta_{0j} x^j$ have no common roots. The idea for the above construction based on the asymmetric generalized autoregressive conditional heteroskedastic(AGARCH) model where variables are continuously distributed in $\mathbb{R}$ studied by Ding et al. (1993). This model is appropriate to model data with positive and negative values while those observations can have asymmetric impact on conditional variance. The pmf of X_t is:

$$P(X_t = k \mid \mathcal{F}_{t-1}) = \begin{cases} \frac{1}{2}e^{-\lambda_t}\frac{\lambda_t^{|k|}}{|k|!}, & k \neq 0 \\ e^{-\lambda_t}, & k = 0 \end{cases}$$

which is symmetric about zero and exhibits exponential decay when $k \to \pm\infty$.

An heteroskedastic time series process with no assumptions on the relationship between conditional mean and variance included in the family of semiparametric models introduced by Liu and Yuan (2013) and called random rounded integer-valued autoregressive conditional heteroskedastic(RRINARCH) model. Other advantage except separateness between mean and variance is that it allows negative values at the series and at the autocorrelation function. Additionally autocorellation of the RRINARCH(p,q) model is similar to this of an AR(p) model. The construction of the above model based on two random rounding operators where the first one is a special case of probabilistic rounding introduced by M'Raïhi et al. (2000) defined as:

$$\odot_1(x, U) = \Delta(x) + \mathbb{I}(U \geq 1 + \Delta(x) - x), \;\; x \in \mathbb{R}$$
$$\odot_2(x, U) = \Delta(x^{1/2}) + \mathbb{I}(U \geq B(x)), \;\; x \in \mathbb{R}_+$$

where

1. $\mathbb{I}(A)$ is an indicator function of A
2. $\Delta(x) = max\{z \in \mathbb{Z} : z \leq x\}$
3. $B(x) = \frac{\left(\Delta x^{1/2}+1\right)^2 - x}{\left(\Delta\left(x^{1/2}\right)+1\right)^2 - \left(\Delta\left(x^{1/2}\right)\right)^2}$
4. U is a uniform random variable defined on the interval [0.1].

Considering that

$$\mathcal{F}(c) = \{X : E(X) = c, Var(X) < \infty\}$$

and

$$R(c) = (\Delta(c) + 1 - c)\,(\Delta(c) - c)\,.$$

Then the random rounding operator lies on the above collection $\mathcal{F}(c)$ and $Var(\odot_1(c, U)) = R(c)$ and RRINARCH(p,q) model has mean given by:

$$E(X_t \mid \mathcal{F}_{t-1}) = c + \sum_{i=1}^{p} \phi(j) X_{t-j}$$

and variance defined as:

$$Var(X_t \mid \mathcal{F}_{t-1}) = R(E(X_t \mid \mathcal{F}_{t-1})) + h_t$$

where h_t is a non-negative $\mathcal{F}_{t-1}$ measurable function. A generalised regime-switching integer -value(RGS-INGARCH(1,1)) GARCH(1,1) model defined by Lee and Hwang (2018) in order to model processes with time-varying dependent transition probabilities. A threshold integer-valued time series given by:

$$X_t \mid \mathcal{F}_{t-1} \sim Poisson(\lambda_t)$$

where λ_t given by (1.3) and

$$X_{t-1}^{(r)} = \begin{cases} X_{t-1} & \text{if } X_{t-1} > m_t \\ 0 & \text{if } X_{t-1} \leq m_t \end{cases}$$

$$X_{t-1}^{(l)} = \begin{cases} X_{t-1} & \text{if } X_{t-1} \leq m_t \\ 0 & \text{if } X_{t-1} > m_t \end{cases}$$

Parameter m_t is a time-varying threshold variable that determines the dynamic switching mechanism of the model. Another threshold model has been proposed by Liu (2012). The main difference with the previous model is that the threshold parameter at this model lies on the set of natural numbers without any dependence on time. Furthermore, an attempt for the construction of a mixture integer-valued autoregressive conditional heteroskedastic (MINARCH) model was given by Zhu et al. (2010) based on the idea of the general mixture model of Saikkonen (2007). The usefulness of the above model obtained at the modeling of data from bimodal marginal distributions as for example Canadian Lynx data studied from Tong (1990). A novel structure of an integer-valued heteroskedastic model to overcome the problem of equidispersion proposed by Barreto-Souza et al. (2022). Time-varying dispersion parameter was considered at this new structure in order to control conditional variance and to provide alternative and more flexible structure of volatility. The time varying dispresion integer-valued generalized autoregressive conditional heteroskedastic (tv-DINGARCH) model studied in case of Negative binomial distribution while a log-linear tv-INGARCH model with covariates has been also discussed.

Definition 3. *A $tv - DINGARCH(p_1, p_2, q_1, q_2)$ process $\{X_t\}_{t\in\mathbb{N}}$ is defined by $X_t \mid \mathcal{F}_{t-1} \sim MP(\lambda_t, \phi_t)$ with*

$$\lambda_t = f(X_{t-1}, \ldots, X_{t-p_1}, \lambda_{t-1}, \ldots, \lambda_{t-q_1}), \;\; \phi_t = f(X_{t-1}, \ldots, X_{t-p_2}, \phi_{t-1}, \ldots, \phi_{t-q_2})$$

where $\mathcal{F}_{t-1}$ and λ_0, ϕ_0 denoting some starting value, $f : \mathcal{N}^{p_1} \times (0,\infty)^{q_1} \to (0,\infty)$ and $g : \mathcal{N}^{p_2} \times (0,\infty)^{q_2} \to (0,\infty)$

A variable follows a mixed Poisson(MP) distribution if $X \mid Z = z \sim Poisson(\lambda z)$ and Z follows a non-negative distribution with $E(z) = 1$ and $Var(z) = \phi$, for $\lambda, \phi > 0$. In case of $X \sim MP(\lambda_t, \phi_t)$ then $E(X_t) = \lambda_t$ and $Var(X_t) = \lambda_t + \phi_t \lambda_t^2$. The random variable Z follows a Gamma distribution $G(a, b)$. studied two models the first one of a Negative binomial distribution and the second one of a Poisson inverse Gaussian(PIG) distribution. At the first case when $Z \sim G(\phi^{-1}, \phi^{-1})$ with $E(z) = 1$ and $Var(Z) = \phi$ then $X \sim NB(\lambda, \phi)$ given by:

$$f(x_t \mid \mathcal{F}_{t-1}) = \frac{\Gamma(x + \phi_t^{-1})}{x!\Gamma(\phi_t^{-1})} \left(\frac{\lambda_t \phi_t}{\lambda_t \phi_t + 1} \right)^x \left(\frac{1}{\lambda_t \phi_t + 1} \right)^{1/\phi_t}, \quad x \in \mathbb{N}_0$$

At the second model when $Z \sim IG(1, \phi)$ then $X \sim PIG(\lambda, \phi)$. An exceptional case of a log-linear tv-DINGARCH model also proposed without study of theoretical properties. While the construction based on this proposed before many years from Fokianos and Tjøstheim (2011), the main difference is obtained on the consideration of covariates at the expression of volatilities $\log(\lambda_t)$ and $\log(\phi_t)$. This necessity appeared after the idea of modeling Covid-19 data about deaths in Ireland. A variable with the name $\log -hosp$ which gives the number of hospitalized patients due to the disease, was included as covariate in order to model and study the effect to the mean and dispersion. The structure of the NB model is similar to this given previously in linear case but $\mu_t = \log(\lambda_t)$ and $v_t = \log(\phi_t)$ are defined as

$$\begin{cases} \mu_t = b_0 + b_1 \log(Y_{t-1} + 1) + b_2 \mu_{t-1} + \delta_1 (\log -hosp_t) \\ v_t = a_0 + a_1 \log(Y_{t-1} + 1) + a_2 v_{t-1} + \gamma_1 (\log -hosp_t) \end{cases}.$$

An INGARCH(1,1) model with Markovian covariates constructed by Liu (2012). Assuming a time homogeneous Markov chain Z_t given as a p-dimensional vector $Z_t = (Z_{1,t}, \ldots, Z_{p,t})$ that contains the covariates information at time t, the model defined as

$$X_t \mid \mathcal{F}_{t-1}^{X,Z}, Z_t \sim Poisson(e^{Z_t^T \gamma} \mu_t),$$

where

$$\mu_t = \delta + a\mu_{t-1} + \beta X_{t-1} e^{-Z_{t-1}^T \gamma},$$

$\mathcal{F}_{t-1}^{X,Z} = \sigma\{\mu_1, Y_s, Z_s, s \leq t\}$, $\mathcal{L}(Z_t \mid Z_{t-1})$ represents a distribution of the random variable Z_t. By this way a Negative Binomial INGARCH model with Markovian covariates has been also discussed. Clark and Dixon (2021) studied a spatially correlated self-excited INGARCH model assuming data model dependence and capturing spatial variation. A new model based on Poisson-Gamma conjugacy with a dynamic specification proposed by Gouriéroux and Lu (2019). The main feature of the above model is the flexibility offered by the construction providing homoskedastic or heteroskedastic models. At this study a Negative Binomial INAR(1) model and a Negative Binomial INARCH(1,1) model have been discussed. An overview about integer-valued generalized conditional heteroskedastic time series models given by Fokianos (2012) and recently a methodological review proposed by Davis et al. (2021).

1.4 Stationarity and Ergodicity of OD models

1.4.1 Ergodicity

At observation-driven process ergodicity can be examined by using Tweedie criteria under the assumption that the process is irreducible Markov chain. First idea, proposed by Rosenblatt (1956), is based on mixing condition satisfied by a sequence of random variables. Even though Tweedie (1976) has proposed irreducibility criteria, assumptions are not always satisfied. Meyn and Tweedie (1992);Meyn and Tweedie (1993);Meyn and Tweedie (1994) studied conditions for stability of discrete-time stochastic processes, of continuous-time stochastic processes respectively. At discrete time processes criteria of Harris recurrence, positive Harris recurrence and geometric recurrence and ergodicity were discussed. The chain is probabilistically stable if $\mathbf{\Phi}$ returns to sets of positive measure ('Harris recurrence') or if there is a unique invariant probability measure for $\mathbf{\Phi}$ ('positive Harris recurrence'). Considering the Markov chain $\mathbf{\Phi} = \{\Phi_0, \Phi_1, \ldots, \}$ evolving on a locally compact separable metric space X, whose Borel σ-algebra shall be denoted $\mathcal{B}$. With P_μ and E_μ denoted the probability and expectation conditional on Φ_0 having distribution μ and P_x and E_x when μ is concentrated at x. Criteria of ϕ-irreducibility are the most commonly used in cases of heteroskedastic time series models. If $\mathbf{\Phi}$ is positive Harris recurrent and t with invariant probability measure π and period $m \geq 1$, and if $\pi(f) < \infty$, then for any initial condition x satisfying

$$E_X \left[\sum_{k=1}^{\tau_A} f\left(\Phi_k\right) \right] < \infty$$

the distributions converge in f-norm:

$$\left\| \frac{1}{m} \sum_{i=1}^{m} P^{i+l}(x, \circ) - \pi \right\|_f \to 0$$

where frobenius norm defined as $||A||_F = \sqrt{(\sum_{i=1}^{m}\sum_{j=1}^{q}|a_{ij}|^2)}$ and A is an $n \times m$ matrix. Furthermore proved that if a Markov chain is aperiodic and geometrically recurrent then is geometrically ergodic. The main problems are when Markov chains are not irreducible or in general do not satisfy any mixing condition discussed by Doukhan and Doukhan (1994). Fonseca and Tweedie (2002) proposed stability condition and ergodic theorems for non-irreducible chains. Lu (1996) studied geometric ergodicity of an ARCH process while Rahbek (2003) studied stochastic properties and stationary solutions of multivariate ARCH processes. Francq and Zakoïan (2006) proposed mixing properties of a general class of GARCH (1, 1) models and before this paper Ling and McAleer (2002) discussed properties for stationarity and existence of moments for a class of GARCH models. Bibi and Aknouche (2008) examined stationarity and geometric ergodicity of a periodic GARCH processes. More generally Boussama et al. (2011) studied criteria for stationarity and geometric ergodicity of Baba, Engle, Kraft, and Kroner(BEKK) multivariate GARCH models by expressing GARCH processes as semi-polynomial Markov processes. Lescheb (2015) proposed necessary and sufficient conditions ensuring the existence of stationary solutions basing on the top Lyapunov exponent considering a periodic Asymmetric Generalized Autoregressive Conditional Heteroskedastic model(PAGARCH). Furthermore limit theorems and moment inequalities are difficult to obtain in many time series processes. Taking into account the above obstacle criteria of weak stationarity have been proposed and study. Meyn and Tweedie (1994) studied under which criteria Markov chains converge. Doukhan et al. (2012) introduced the more adapted weak dependence condition. The idea based on the asymptotic independence between "past" and "future". Considering a time series $X = (X_n)_{n\in\mathbb{N}}$ with values in a local topological space $E(E = \mathbf{R}^d)$. Assuming a variable $(X_{i_1}, X_{i_2}, \ldots, X_{i_u})$ for the past and one $(X_{j_1}, X_{j_2}, \ldots, X_{j_v})$ for the future where $i_1 \leq i_2 \leq \cdots \leq i_u \leq j_1 \leq j_2 \leq \cdots \leq j_v$, $u, v \in \mathbb{N}^\star$, independence of time series is calculated via independence between X_i and X_j given by:

$$\epsilon(r) = \sup_{D(X_i,X_j)\geq\ r} \sup_{(f,g)\in\mathcal{F}\times\mathcal{G}} \left\| Cov(f(x_i), g(x_j)) \right\|$$

where $D(X_i, X_j) = j_1 - i_u$. Many dependence criteria have been constructed considering mixing condition introduced by Rosenblatt (1956) or on martingales approximations(mixingales).

Definition 4. *Let $p \geq 1$ and let $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ be an increasing sequence of σ-algebras. The sequence $(X_n, \mathcal{F}_n)_{n\in Z}$ called an L^p mixingale if there exist non-negative sequences $(c_n)_{n\in\mathbb{Z}}$ and $(\phi(n))_{n\in\mathbb{Z}}$ such that $\psi(n) \to 0$ as $n \to \infty$ and for all integers*

$n \in \mathbb{Z},\ \ k \geq 0$

$$\begin{aligned} \left\| X_n - E(X_n \mid \mathcal{F}_{n+k}) \right\|_p &\leq c_n \psi(k+1) \\ \left\| E(X_n \mid \mathcal{F}_{n-k}) \right\| &\leq c_n \psi(k) \end{aligned}$$

However many classes of time series do not satisfy any mixing condition or limit theorems and moment inequalities are difficult to obtain. Bickel and Bühlmann (1999) and simultaneously Doukhan and Louhichi (1999) introduced a new idea of weak dependence. Let $\Lambda(\delta)$ be the set of Lipschitz functions from $\mathcal{X}$ to $\mathbb{R}$ with respect to the distance δ. For $f \in \Lambda(\delta)$, we denote this by $Lip(f)$, $f's$ Lipschitz constant.

$$\Lambda^{(1)}(\delta) = \left\{ f \in \Lambda(\delta) Lip(f) \leq 1 \right\}.$$

Furthermore considering a probability space (Ω, A, P) and a Polish space $\mathbf{X}$, two classes of functions $\mathcal{F}_u$, $\mathcal{G}_u$: $X_u \to \mathbb{R}$ have been also defined and $\mathcal{F} = \cup \mathcal{F}_u$ and $\mathcal{G} = \cup \mathcal{G}_u$ Assuming two random variables X and Y with values in $\mathcal{X}_u$ and $\mathcal{X}_v$ respectively and Ψ some function from $\mathcal{F} \times \mathcal{G}$ to $\mathbb{R}_+$, *the* $(\mathcal{F}, \mathcal{G}, \mathbb{R})$-dependence coefficient $\epsilon(X, Y)$ is given by:

$$\epsilon(X, Y) = sup_{f \in \mathcal{F}_u, g \in \mathcal{G}_u} = \frac{\left\| Cov\left(f(X), g(Y)\right) \right\|}{\Psi(f, g)}.$$

If $(X_n)_{n \in \mathbb{Z}}$ is a sequence of $\mathcal{X}$-valued random variables and $\Gamma(u, v, k)$ a set of $(i, j) \in \mathbb{Z}^u \times \mathbb{Z}^v$ such that $i_1 \leq \cdots \leq i_u \leq i_u + k \leq j_1 \leq \cdots \leq j_v$. The dependence coefficient $\epsilon(k)$ is defined by:

$$\epsilon(k) = sup_{u,v} sup_{(i,j) \in \Gamma(u,v,k)} \left(\epsilon \left(X_{i_1}, \ldots, X_{i_u} \right), \left(X_{j_1}, \ldots, X_{j_v} \right) \right).$$

The sequence $(X_n)_{n \in \mathbb{N}}$ is $(\mathcal{F}, \mathcal{G}, \Psi)$-dependent if the sequence $\epsilon(k))_{k \in \mathbb{N}}$ tends to zero. Doukhan et al. (2012) proposed numerous criteria of weak dependence like η, k,λ, $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\phi}$,θ and τ where a part of those constructed based on Rosenblatt (1956)'s criteria. The coefficient θ corresponds to

$$\Psi(f, g) = d_g ||f||_\infty Lip(g),$$

where f belongs to the set of bounded functions $\mathcal{F}_u$, $||f||_\infty = max f_i$. Moreover $(X_n)_{n \in \mathbb{Z}}$ be a sequence of R^k-valued random variables. We define the sequence $(Y_n)_{n \in \mathbb{Z}}$ by $Y_n = h(X_n)$. If X_n is θ-weak dependent then $(Y_n)_{n \in \mathbb{Z}}$ is also θ weak dependent and

$$\theta_Y(n) = \mathcal{O}(\theta(n)^{\frac{p-a}{p-1}}).$$

Definition 5. *Let $(\Omega, A, \mathbb{P})$ be a probability space and $\mathcal{F}$ be a σ-algebra on A. Let X be a random variable on $(\Omega, A, \mathbb{P})$ with values on $\mathbb{R}^m$ and $||X||_1 < \infty$ $\mathbb{X}$ be a polish space and δ be a metric on $\mathbb{X}$. For any L^P-integrable function X on $\mathbb{X}$:*

$$\theta_p(\mathcal{F}, X) = sup\{||E(f(X) \mid \mathcal{F}) - E(f(X))||_P / f \in \Lambda^{(1)}(\mathbb{R}^m)\} \tag{1.9}$$

and

$$\tau_p(\mathcal{F}, X) = \{||sup\{E(f(X) \mid \mathcal{F}) - E(f(X))\} / f \in \Lambda^{(1)}(\mathbb{R}^m)||\}. \tag{1.10}$$

The main difference between Tweedie and weak dependence criteria is that for the satisfaction of the second only the first order moments of the process must be infinite. Simpler inequalities for β, α and τ weak dependence coefficients have been also provided. Considering the probability space $(\Omega, A, \mathbb{P})$, where Ω is rich enough and X be another variable distributed as Y and independent of Y we can bound the coefficient τ as: $\tau(\mathcal{F}, Y) = ||Y - X||$. The above implies that there exists a variable $Y^\star$ measurable with respect $max\{\mathcal{F}, \sigma(Y), \sigma(\delta)\}$ independent of $\mathcal{F}$ distributed as Y

$$||Y - Y^\star||_1 = \tau_1(\mathcal{F}, Y). \tag{1.11}$$

It is easy to see that

$$\theta_P(\mathcal{F}, Y) \leq \tau_P(\mathcal{F}, Y). \tag{1.12}$$

The drawback is that weak dependence criteria provide stationarity of only first order moments. The usefulness of each one of those depend on the closed form of stochastic process. Fokianos et al. (2009) studied ergodicity of univariate INGARCH process considering a perturbed model.{Doukhan and Louhichi (1999),Dedecker et al. (2007)} overcome those difficulties by proposing weak dependence criteria for different types of processes. In case where recursive equation is expressed linearly weak dependence criteria can be used more easier than other cases. Furthermore the problem of ergodicity become more intensive when there are interdependencies between two or more stochastic processes. For an extended study of weak dependence, someone could see the work of Dedecker et al. (2007).

To sum up there are numerous criteria to provide ergodicity. The study of ergodicity is very important because via ergodic criteria, stationarity also provided. Even though there are too many weak dependence criteria at some cases, criteria proposed by Tweedie (1976) could be also used. There are studies as Fokianos et al. (2009) where stationarity proved using weak dependence criteria but a few years later at the same process stationarity has been also provided using Tweedie criteria by Tjøstheim (2012). The reason is that in some way Tweedie criteria are more difficult and complicated to use based on more theoretical aspect.

1.4.2 Consistency issue of Observation Driven(OD) Models

In the thesis we consider multivariate discrete and continuous OD models. Considering

$$X_t \mid \mathcal{F}_{t-1} \sim G^{\theta}(\lambda_t),$$

where (λ, d_λ) be a locally compact and separable metric space and $\lambda_t = \psi_{X_t}(\lambda_t)$ is a measurable function. Examination for the consistency of MLE based on theorem developed by Lee et al. (2016), providing stationarity and ergodicity of the process $\{X_t, t \geq 0\}$ and considering that second order partial derivatives are continuous and bounded. In the thesis ergodicity provided considering criteria proposed by Dedecker et al. (2007). We study numerical simulations considering a trivariate INGARCH(1,1) and a bivariate ECARR(1,1) model. Fokianos et al. (2009) provided ergodic properties assuming classical approaches as criteria proposed by Tweedie (1976) and β-mixing properties. The obstacle is the irreducibility of the Markov chain and for this reason Fokianos et al. (2009) considered a perturbed model. On the contrary in the continuous case of the ECARR model Fernandes et al. (2005) considering a trivariate model provided ergodicity via Tweedie criteria of the model without any disturbance. The drawback with the criteria of weak dependence is that only the first two conditional moments the process are stationary.

1.5 Multivariate heteroskedastic OD time series models

1.5.1 Bivariate processes for continuous data

In real world the construction of models taking into account properties of data expressed by two or more variables and examined for the correlation was proved necessary. The first bivariate autoregressive process is given by:

$$\mathbf{X}_t = \mathbf{A}_1 \mathbf{X}_{t-1} + \epsilon_t,$$

where $\mathbf{A}_1$ is a 2×2 matrix and the error process is a 2-dimensional zero mean white noise process with covariance matrix $E(\epsilon_t \epsilon_t^{'}) = \Sigma_\epsilon$ such that $\epsilon_t =\sim N(0, \Sigma_\epsilon)$. In similar way the expansion of the model at more than two dimensions has been also studied without any difficulty. The two-dimensional Vector Autoregressive(VAR(p)) model could be written as a 2p-dimensional VAR(1) model by stacking p consecutive X_t variables in a 2p-dimensional vector as:

$$X_t = \mathbf{A} X_{t-1} + \epsilon_t$$

where A referred as the companion matrix and $\mathbf{A}= \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_2 & 0 & \dots & 0 & 0 \\ 0 & I_2 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 0 & I_2 \end{bmatrix}$ and

$e_t= \begin{bmatrix} \epsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ Similarly to the univariate case the stable VAR process can be expressed as the weighted sum of past and present innovations. Considering the Wold decomposition theorem and under suitable conditions every K-dimensional nondeterministic zero mean stationary process y_t has an MA representation. A vector autoregressive moving average (VARMA) defined as:

$$\mathbf{X}_t = \mathbf{A}_1\mathbf{X}_{t-1} + \epsilon_t + \mathbf{B}_1\epsilon_t$$

where $\mathbf{A}_1$ and $\mathbf{B}_1$ are 2×2 matrices, $\mathbf{x}_t$ is a 2-dimensional vector of observations and e_t is a 2-dimensional vector of white noise error terms. At this model if the determinant of the MA operator of the VARMA process has all its roots outside the unit circle, the process has an equivalent pure VAR representation of possibly infinite order. A detailed introductory exposition of VARMA processes is provided by Lütkepohl (2005) and a more advanced treatment can be found in Deistler and Hannan (1988),Hsiao (1979);Hsiao (1981). Keating (2000) introduces the closely related idea of asymmetric VAR (AVAR) models, in which the lag length differs across variables such that the lag length is the same for each variable in all equations, but may differ across variables. The use of exogenous variables to a VAR model which include the same number of lags as the other equations but only lags of the exogenous variables have been also studied. The appearance of exogenous variables was not commonly used in VAR models but after few years financial market showed that those models were necessary. A better example is a small open economy that faces exogenous variation in world interest rates or in its terms of trade (see, e.g., Cushman and Zha (1997)). To represent the dynamics of the conditional variances and covariances, the construction of multivariate generalized autoregressive conditional heteroskedastic models(MGARCH) became useful. The drawback was that as the dimension of the model becomes high then the number of parameters increases and computational methods for parameters estimation are needed. On the other hand, the above fact is not a drawback because models with only a few parameters may not be able to capture the relevant dynamics in the covariance structure. An obligatory assumption is that the conditional covariance matrices implied by the model must be positive definite, but this is often infeasible in practice except the formula construction in way that positive definiteness is implied by the model structure.

The first GARCH model for the conditional covariance matrices was the so-called VEC model of Bollerslev et al. (1988) and an ARCH model of Engle et al. (1984). The lack at properties for consideration of positive definiteness in the construction of the model provide at the construction of more parsimonious models. Bollerslev et al. (1988) presented a simplified version assuming that the matrices are diagonal. A model that can be viewed as a restricted version of the VEC model is the Baba-EngleKraft-Kroner (BEKK) defined in Engle and Kroner (1995). It has the attractive property that the conditional covariance matrices are positive definite by construction. Engle and Kroner (1995) provide sufficient conditions for the two models, BEKK and VEC, to be equivalent. They also give a representation theorem that establishes the equivalence of diagonal VEC models and general diagonal BEKK models. A generalization of the univariate exponential GARCH model of Nelson (1991),where parameter restrictions to ensure positive definiteness are not needed proposed by Kawakatsu (2006). Engle et al. (1990) introduced the first factor GARCH model, assuming that the observations are generated by underlying factors that are conditionally heteroskedastic and possessing by this way a GARCH-type structure. Alexander and Chibumba (1997) proposed an Orthogonal GARCH (O-GARCH) model and Van der Weide (2002) extended the above model to a Generalized Orthogonal GARCH model (GO-GARCH).

Diebold and Nerlove (1989) propose a model similar to the previous one and Sentana (1998) made a comparison between the above models. Full Factor (FF–) GARCH model and Generalized Orthogonal Factor GARCH (GOF-GARCH) model have been studied by Vrontos et al. (2003) and Lanne and Saikkonen (2007) respectively. Correlation models are based on the decomposition of the conditional covariance matrix into conditional standard deviations and correlations. The simplest multivariate correlation model that is nested in the other conditional correlation models, is the Constant Conditional Correlation (CCC–) GARCH model of Bollerslev (1990). At varying-correlation GARCH(VC-GARCH) model proposed by Tse and Tsui (2002) conditional correlation is function of previous periods and a set of estimated correlations. Other multivariate GARCH models with transformations at Conditional correlation have been also studied by Kwan et al. (2005), Engle (2002), Billio et al. (2006), Cappiello et al. (2006), Silvennoinen and Teräsvirta (2005), Silvennoinen and Teräsvirta (2009a), Berben et al. (2005), Pelletier (2006), Silvennoinen and Teräsvirta (2009b), Silvennoinen and Teräsvirta (2009b), Fernandes et al. (2005).

1.5.2 Bivariate processes for discrete data

Definition for a joint conditional distribution is required due to necessity of modeling data in more than one dimension. In discrete case, the construction of a multivariate distribution is more complex than the continuous based on the study of theoretical properties like ergodicity and consistency of estimators. Furthermore while studies in case of homoskedastic models, at heteroskedastic models literature is restricted. For example while there is a multivariate negative binomial distribution studied by Andreassen (2013), there are no studies for multivariate Negative Binomial Integer Generalized Autoregressive Conditional Heteroskedastic(NB-INGARCH). A way in order to construct a multivariate model is the method of trivariate reduction introduced by Mardia (1970). Based on the above model a bivariate Poisson distribution had been also proposed by Liu (2012). Considering three mutually independent Poisson distributed random variables $X_{1t} \sim Poisson(\lambda_{1t} - \phi)$, $X_{2t} \sim Poisson(\lambda_{2t} - \phi)$ and $X_{3,t} \sim Poisson(\phi)$ then $Y_{1t} = X_{1t} + X_{3t}$ and $Y_{2t} = X_{2t} + X_{3t}$ have a joint bivariate Poisson distribution given by:

$$\begin{aligned} f(Y_{1t}, Y_{2t} \mid \mathcal{F}_{t-1}) &= P(Y_{1t} = m, Y_{2t} = l \mid \mathcal{F}_{t-1}) \\ &= \sum_{i=0}^{min(m,l)} P(X_1 = m - i)P(X_2 = l - i)P(X_3 = i) \\ &= e^{-(\lambda_{1t}+\lambda_{2t}-\phi)} \frac{(\lambda_{1t} - \phi)^m}{m!} \frac{(\lambda_{2t} - \phi)^l}{l!} \sum_{i=0}^{min(m,l)} \left(\frac{\phi}{(\lambda_{1t} - \phi)(\lambda_{2t} - \phi)} \right)^i \end{aligned}$$

Andreassen (2013) presented three simulation studies based on three methods the first one of setting initial values at λ_1 and $X_1 = (X_{1,1}, X_{2,1})$ and then using general structure for λ_t simulations from $BP(\lambda_{1,t}, \lambda_{2,t}, \phi)$ were provided. The second one based on generation of $X_{1,t} \sim Poisson(\lambda_{1,t} - \phi)$, $X_{2,t} \sim Poisson(\lambda_{2,t} - \phi)$ and $X_{3,t} \sim Poisson(\phi)$ and setting $Y = (X_{1,t} + X_{3,t}, X_{2,t} + X_{3,t})^T$ and the third one by using a closed form of a copula as usually an Archimedean copula or a Clayton copula. Based on method of trivariate reduction, the first bivariate heteroskedastic integer valued conditional heteroskedastic process constructed by . The effort was concentrated at the extension of the autoregressive conditional Poisson model of a vector autoregressive(AR)-type model for the conditional mean given as:

$$X_t \mid \mathcal{F}_{t-1}^X \sim BP(\lambda_\mathbf{t}) \tag{1.13}$$

where

$$\lambda_{\mathbf{t}} = \mathbf{d} + \mathbf{A}\lambda_{\mathbf{t-1}} + \mathbf{B}\mathbf{X}_{\mathbf{t-1}} \tag{1.14}$$

λ_{it}, $i = 1, \ldots, n$ is the number of variables and $t \in \mathbb{N}$ is volatility lag. Each λ_{it} does not depend only on its own past and this indicates that $\mathbf{A}$ and $\mathbf{B}$ are squared matrices with non zero elements. Providing that $||A||_1 + ||B||_1 < 1$ where $||A||_1 = max_{1 \leq j \leq p} \sum_{i=1}^{q} |a_{ij}|$, then there exists a unique solution (X_t, λ_t) to model (4.1) which is stationary and ergodic. The elements of **d,A,B** are assumed to be positive considering positivity of $\lambda_{i,t}$. The construction of the model has been done using the trivariate reduction method where the bivariate Poisson distribution given by:

The model is capable to capture dependence between the two time series Y_{1t}, Y_{2t} providing that $\phi \neq 0$. Another model proposed by Cui and Zhu (2018). The idea for the construction is based on Sarmanov distribution studied by Lakshminarayana et al. (1999) and given by:

$$f(x_1, x_2) = f(x_1)f(x_2)[1 + \omega\left(g_1\left(x_1\right)g_2\left(x_2\right) - E\left(g_1\left(x_1\right)\right)E\left(g_2\left(x_2\right)\right)\right)]$$

Cui and Zhu (2018) considered as marginal $f(x_i)$, $i = 1, 2$ $Poisson(\lambda_i)$ and $q(x_i) = e^{-y_i} - e^{-c\lambda_i}$, where $c = 1 - e^{-1}$

$$f(y_1, y_2) = \prod_{i=1}^{2} \frac{\lambda_i^{y_i}}{y_i!} e^{-\lambda_i} \left(1 + \delta \prod_{i=1}^{2} \left(e^{-y_i} - e^{-c\lambda_i}\right)\right),$$

where $c = 1 - e^{-1}$.

The proposed model was not well defined as mentioned from other researchers who proposed alternative constructions of multivariate distributions because while the structure of correlation coefficient is based on the equivalent study of Lakshminarayana et al. (1999), the way on how boundaries values have been also calculated missing or it is not clearly defined. Another attempt taking into advantage three families of copulas the family of Frank, the family of Gaussian and the family of FGM copula had been discussed by Cui et al. (2020) where bivariate Poisson distribution defined by:

$$f(x_1, x_2) = \frac{1}{Z(\lambda_1, \lambda_2, \omega)} \prod_{i=1}^{2} \frac{\lambda_i^{x_i}}{x_i!} e^{-\lambda_i} c(F_1(x_1), F_2(x_2)),$$

where $c(F_1(x_1), F_2(x_2))$ is one of the above mentioned copulas, ω be the parameter of dependence and normalizing constant defined as :

$$Z(\lambda_1, \lambda_2, \omega) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \frac{\lambda_1^{x_1} \lambda_2^{x_2}}{x_1! x_2!} e^{-(\lambda_1 + \lambda_2)} c(F_1(x_1), F_2(x_2)).$$

Based on Shmueli et al. (2005), series Z converges and both Z, Z^{-1} are bounded. Those motivations for the construction derived from the idea of Lakshminarayana et al. (1999) and copulas. At the above cases researchers deal with the problem of negative correlation. Although the difficulty is focused not only on the construction of a multivariate Poisson distribution that provides negative correlation but on the dependence structure which does not offer flexibility. Crucial attention is needed at parameters' estimation where classical approaches as maximum likelihood approximation seem to be time-consuming depending each time on the copulas' form. Cui and Zhu (2018) and Cui et al. (2020) decomposed the log-likelihood function based on maximization by parts and a modified version of this as alternative approaches with less complexity but with a bit high computational cost. Another approximation was presented by Fokianos et al. (2020) whose aim is an introduction of a copula via a vector of continuous random variables.

Considering the fact that copulas have a wide range of applications providing a general way for construction of multivariate distributions with continuous distributions, Fokianos et al. (2020) constructed a bivariate Poisson by introducing a copula form via a vector of continuous distributions and by utilizing fundamental properties of Poisson process. The method is flexible because it implies that each Y_i can be a Poisson process but the joint distribution is not mandatory a Poisson distribution.

1. Define initial values for parameter vector λ_0 and generate values of $U_l = (U_{1,l}, U_{2,l})$, $\ell = 1, 2, ..., k$ from a 2-dimensional copula. Then $U_{i,l}$ follow marginally the uniform distribution on (0,1) for i=1,2.

2. Consider the transformation $X_{i,l} = -\frac{\log U_{i,l}}{\lambda_{i,0}}$, $i = 1, 2$. Then the marginal distribution of $X_{i,l}$, $l = 1, 2, \ldots, k$ is exponential with parameter $\lambda_{i,0}$, $i = 1, 2$.

3. Define now taking K large enough $Y_{i,0} = max\left\{\sum_{l=1}^{K} X_{i,l} \leq 1\right\}$, $i = 1, 2$. Then $\mathbf{Y}_0$ is marginally a set of first values of a Poisson process with parameter λ_0.

Fokianos et al. (2020) with the same way study a bivariate log-linear Poisson INGARCH model given by:

$$\mathbf{Y}_t = \mathbf{N}_t(\mathbf{v}_t), \;\; \mathbf{v}_t = b_0 + \mathbf{A} v_{t-1} + \mathbf{B} \log(\mathbf{Y}_{t-1} + 1)$$

where $\mathbf{v}_t = \log \lambda_t$ and $\mathbf{1}_2$ is a 2-dimensional vector which consists of ones. Similarly to the linear model, the process N_t denotes a sequence of independent 2-

variate copula–Poisson processes which counts the number of events in $[0, exp(v_{1,t})] \times [0, exp(v_{2,t})]$.

At this model the author dealt with the problem of equidispersion. For this reason another model where data comes from a double Poisson distribution, was constructed by Heinen and Rengifo (2007) based on the definition given by:

$$X_t \mid \mathcal{F}_{t-1} \sim DP(\lambda_t, \phi), \;\; t = 1, 2, \ldots$$

where $E(X_t \mid \mathcal{F}_{t-1}) = \lambda_t$ and $Var(X_t \mid \mathcal{F}_{t-1}) = \frac{\lambda_t}{\phi}$. A brief discussion with distinction on the rank of matrices A and B has been also presented. The above models were uncorrelated and a way with copulas of multivariate time series processes also discussed. The main problem was the lack of uniqueness of multivariate distribution in discrete case. To bypass the problem Heinen (2003) proposed a way for variables continuousation by adding at a discrete variable Z a randomly generated variable U from a Uniform distribution in (0,1) expressed as :

$$Z^{\star} = Z + (U - 1)$$

The idea was based on Denuit and Lambert (2005) considering that this way of continuousation does not alter the concordance between the pairs of variables. A few years after the above study Heinen and Rengifo (2007) proposed an other one bivariate DPINGARCH model via the way of continuousation and using as tool for the construction of the bivariate distribution an archimedean copula. Assuming as previously a $DP(\lambda_t, \phi)$ where λ_t given by (1.4), the joint density of the counts in the double Poisson distribution is:

$$h(X_{1,t}, X_{2,t}, \theta.\Sigma) = \prod_{i=1}^{2} f_{DP}(X_i, \lambda_{i,t}, \phi_i) c(q_t; \Sigma).$$

f_{DP} denotes the double Poisson density as a function of the observation $Y_{i,t}$, $c(\cdot;\cdot)$ denotes the copula density of a multivariate normal and $\theta = (\mathbf{b}_0, \mathbf{A}, \mathbf{B})$, The $q_{i,t}$, gathered in the vector $q(t)$ are the normal quantiles of the $z_{i,t}$ and $z_{i,t}$ are the PIT of the continued extension of the count data under the marginal densities

$$z_{i,t} = F^{\star}_{i,t}(X^{\star}_{i,t}) = F_{i,t}(X_{i,t} - 1) + f_{i,t}(X_{i,t}) U_{i,t},$$

where $F^{\star}_{i,t}, F_{i,t}, f_{i,t}$ are the conditional cdf of the continued extension of the data, the conditional cdf and the conditional pdf respectively and $X^{\star}_{i,t}$ are the continued extension of the original count data.

A new development of a multi-factor dynamic copula model with time-varying loadings and observation driven dynamics proposed recently by Opschoor et al. (2021). The flexibility of the above structure and two other features as scalability to high dimensions and positivity of covariance matrix provide a model that allow simple methods of estimation and easy study of many theoretical properties. A new bivariate conditional Poisson(BCP) INGARCH model defined by based on the BCP distribution introduced from Piancastelli et al. (2020).

Definition 6. *A vector* (X_1, X_2) *follows a BCP distribution with parameters* $\lambda_1, \lambda_2 > 0$ *and* $\phi \in \mathbb{R}$ *if* $Z_1 \sim Poisson(\lambda_1)$ *and* $X_2 \mid X_1 = x_1 \sim Poisson(\mu_2 e^{\phi x_1})$ *where* $\mu_2 = \lambda_2 exp\{-\lambda_1(e^{\phi} - 1)\}$ *and the marginal of the Poisson distribution is a mixture Poisson distribution.*

The main difference is that at the study of Piancastelli et al. (2020) λ_2 replaced by the marginal mean of X_2 and the joint BCP given by:

$$f(x_1, x_2) = \prod_{i=1}^{2} \frac{\lambda_i^{x_i}}{x_i!} exp\Big\{-\lambda_1\left(1 + x_2\left(e^{\phi} - 1\right)\right) - \lambda_2 exp\left(-\lambda_1\left(e^{\phi} - 1\right) + \phi x_1\right) + \phi x_1 x_2\Big\}.$$

Parameter ϕ is appropriate for model's dependence considering as covariance and correlation respectively

$$cov(X_1, X_2) = \lambda_1 \lambda_2 (e^{\phi} - 1),$$

$$corr(X_1, X_2) = (e^{\phi} - 1)\sqrt{\frac{\lambda_1 \lambda_2}{1 + \lambda_2\left(e^{\lambda_1\left(e^{\phi}-1\right)^2}\right) - 1}}.$$

It is worth mentioned that this model is mathematically tractable and considering a wide range of values for ϕ a broad interval for correlation coefficient is also provided. At this paper a brief discussion on how unclear correlation coefficient's range had been constructed at the work of Cui and Zhu (2018) also presented. Piancastelli et al. (2020) studied a bivariate INGARCH(1,1) model and using a parametrization for the bivariate conditional Poisson distribution different from the almost proposed by Berkhout and Plug (2004) provide a wider range of contemporaneous correlation.

1.6 Objective & Structure of the thesis

While the above models have been extensively studied, main obstacles considering the lack of closed forms for construction of multivariate Poisson at more than two dimensions have been observed. Cui and Zhu (2018);Cui et al. (2020) studied bivariate INGARCH models with ideas based on the work of Lakshminarayana et al. (1999) and copulas. At the first study of Cui and Zhu (2018) distributions expressed based on the namely Farlie-Gumbel-Morgenstern copula while at the second one of Cui et al. (2020) at Frank and Gaussian copulas. At the above cases researchers deal with the problem of negative correlation. Although the difficulty is focused not only on the construction of a multivariate Poisson distribution that provides negative correlation but on the dependence structure which does not offer flexibility. Crucial attention is needed at parameters' estimation where classical approaches as maximum likelihood approximation seem to be time-consuming depending each time on the copulas' form. Cui and Zhu (2018) and Cui et al. (2020) decomposed the log-likelihood function based on maximization by parts and a modified version of this as alternative approaches with less complexity but with a bit high computational cost. Another approximation was presented by Fokianos et al. (2020) whose aim is an introduction of a copula via a vector of continuous random variables. Instead of quasi likelihood approximation at this study a bootstrap procedure also provided for copulas parameters. The limitation is that correlation boundaries at those multivariate models are restricted and this derives from the choice of copula which provide complex models. At the bivariate cases considering measures proposed by Meyn and Tweedie (2012), some theoretical aspects such as properties of stationarity and ergodicity are also provided. At the multivariate case theoretical aspects have been proved based on the similar way to this of univariate linear INGARCH Poisson model by Fokianos et al. (2009). Theoretical aspects are more easier to prove in case where volatilities at heteroskedastic models are linearly expressed than those of exponential volatilities expressions. It is worth mentioned the study of theoretical properties of a multivariate log-linear INGARCH Poisson model presented by Fokianos et al. (2020). The need of study is derived considering the vital lack of flexible Poisson distribution and regarding the flourishing literature of OD models with a limited research and progress of multivariate flexible heteroskedastic Poisson models.

The objective of the thesis is to develop and study fundamentally time series models leading to study and modeling multivariate time series models for count and continuous data. Apart from the construction of multivariate stochastic processes, we examine multivariate distributions that offer not only positive and negative correlation but also provide wider range for correlation coefficient boundaries. Furthermore specific cases of correlation coefficient in case of Poisson and Exponential distributions that arise from the construction of multivariate distributions have been discussed in detail. Methods for variable and model selection 'have been developed in order firstly to simplify a model considering less variables and than the proposed and to select among few models not necessarily from the same family. Those methods are not based on classical approaches but on bayesian methodologies. This is an important issue when there are studies with data that have been modelled using different processes but we need to examine which process is the best to fit the data. For this reason a bayesian model selection method will be developed to the next chapter among linear and log-linear INGARCH models.

In particular in chapter 2 using GLM models such as those proposed by Ferland et al. (2006) where observations are Poisson distributed and mean expressed linearly or log-linearly via its past values and past-observations, a bayesian model selection method is provided. Based on posterior models probabilities, posterior predictive distributions for volatilities have been also calculated. Efficiency and consistency of prior distribution are examined taking into account two criteria and providing a simulation study. Model selection method with the proposed priors is demonstrated through applications to two different datasets.

In chapter 3 a very interesting issue is discussed on the identification of multivariate distributions. The contribution is focused on the development of multivariate distributions providing negative correlation coefficient values. The construction of multivariate distribution is providing using a multiplicative factor responsible for variables dependence and allowing flexibility at correlation coefficient. Based on this structure, two case studies where random variables are Poisson and Exponentially distributed are discussed. Considering that out multivariate distribution is an alternative construction to this proposed by Le et al. (1996), a graphical inspection is represented yielding correlation coefficient boundaries in comparison with others using different values for exponents in multiplicative factor.

In chapter 4 a multivariate INGARCH(1,1) model is studied based on multivariate distribution proposed at the previous chapter. Two models have been also discussed by Fokianos et al. (2020) and Piancastelli et al. (2020). The pbjective of the new model is to allow flexibility on the way of study theoretical properties such as ergodicity and stationarity of the model and parameter estimation. In addition considering that in trivariate case number of parameters becomes large because multiplicative factor contains all the combinations of two out of three variables, other ways of estimations could be usefull for parameter estimation. For that reason decomposition of score function tends to be necessary and a likelihood by parts method briefly discussed.

In chapter 5 based on the work of Chou (2005);Chan et al. (2019);Miralles-Marcelo et al. (2013)} we study a CARR(1,1) model where variable is defined as the difference between daily maximum and minimum stock prices. Study for ergodicity does not differ from this of INGARCH(1,1) model but the main difference in bivariate model is that we do not consider interactions between the processes. The drawback is that no implementation has been provided due to lack of intradaily data that are not needed for model construction. The contribution is focused on the the structure of exponential CARR(1,1) model with strong correlation lying in the interval (-0.95,0.95).

In the last chapter some challenges for future work are summarized.

Chapter 2

Bayesian models selection for INGARCH models

This chapter focuses on the theory and model selection for a class of INGARCH processes. Assuming that observations Y_t, $t = 0, 1, \dots$ are Poisson distributed, we proposed two models where the difference obtained on the definition of the mean process. Based on Poisson distribution the first and second order conditional moments are equal. To overcome the problem of equidispersion and deal with observations with extreme values, further extensions with other conditional distributions of Y_t such as a negative binomial, a generalized Poisson, a COM-Poisson distribution and a zero-inflated Poisson by Zhu (2011, 2012a,b,c) have been studied respectively. We focus on the model selection among a few linear and log-linear Poisson INGARCH models. The reason to do this is that at log-linear INGARCH models parameter values lie on less restrictive intervals offering flexibility at volatility. More specifically we consider observations $\{X_t\} \in 0, 1, \dots$ and $\lambda_t, t \in \mathbb{Z}$ be a sequence of mean process. In the first Poisson heteroskedastic model, we consider that observations X_t given the past values are Poisson distributed and mean λ_t is linked linearly with its past values and past observations. Namely we assume that

$$X_t \mid \mathcal{F}_{t-1} \sim Poisson(\lambda_t) \tag{2.1}$$

where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{X_s : s \leq t-1\}$ and

$$\lambda_t = d_1 + \sum_{i=1}^{p} a_i \lambda_{t-i} + \sum_{j=1}^{q} b_j X_{t-j} \tag{2.2}$$

Assuming that λ_t is the mean of Poisson distribution, parameters d_1, a_i, b_j are positive for all $i = 1, \dots, p, j = 1, \dots, q$. p and q determine the order of the model and initial values X_0 and λ_0 are fixed. Furthermore considering Poisson process

$E(X_t \mid \mathcal{F}_{t-1}) = Var(X_t \mid \mathcal{F}_{t-1}) = \lambda_t$. This model proposed by Ferland et al. (2006) based on the idea of GARCH model proposed by Bollerslev (1986). Heinen (2003) studied unconditional moments of an INGARCH(1,1) model based on martingale properties and Ferland et al. (2006) study properties and moments of the same model based on theory of ARMA processes. Assuming that $\{X_t\}$ and $\{Y_t\}$ are zero mean stationary processes with autocovariance function $\gamma(\cdot)$ and $\{X_t\}$ is an ARMA(1,1) process defined as

$$Y_t - \phi Y_{t-1} = \epsilon_t + \theta\epsilon_{t-1} \tag{2.3}$$

where ϵ_t is an i.i.d. sequence $N(0, \sigma^2)$, $\phi_1 = a_1 + b_1$, $\theta_1 = -a_1$ and $\sigma^2 = E(Y_t)$ then $\{Y_t\}$ is also an ARMA(1,1) process. According to stationarity conditions

$$\begin{aligned} E(X_t) &= E(E(X_t \mid \mathcal{F}_{t-1})) = E(\lambda_t) = d_1 + a_1E(\lambda_{t-1}) + b_1E(X_{t-1}) \\ &= d_1 + a_1E(\lambda_t) + b_1E(E(X_{t-1} \mid \mathcal{F}_{t-1})) \\ &= d_1 + a_1E(\lambda_t) + b_1E(\lambda_t) \end{aligned} \tag{2.4}$$

where

$$\mu = E(\lambda_t) = d_1\,(1 - a_1 - b_1)^{-1} \tag{2.5}$$

Based on stationarity condition of ARMA(1,1) model we have that $a_1 + b_1 < 1$ and $d_1 > 0$ Considering martingale property where the deviation between dependent variable and conditional mean are independent from the information set at time t and $\mathcal{I}_t$ is the σ-algebra of $\{\lambda_s, s \leq t-1\}$ we have:

$$\begin{aligned} Cov(X_t - \lambda_t, \lambda_{t-k}) &= E\left[(X_t - \lambda_t)(\lambda_{t-k} - \mu)\right] \\ &= E\left[(\lambda_{t-k} - \mu)E(E(X_t - \lambda_t) \mid \mathcal{I}_{t-1})\right] \\ &= E\left[(\lambda_{t-k} - \mu)\left(E(E(X_t) \mid \mathcal{I}_{t-1}) - \lambda_t\right)\right] = 0 \end{aligned} \tag{2.6}$$

$$\begin{aligned} Cov(X_t, X_{t-k} - \lambda_{t-k}) &= E\left[(X_t - \mu)(Y_{t-k} - \lambda_{t-k})\right] \\ &= E([(X_t - \mu)E(X_{t-k} - \lambda_{t-k} \mid \mathcal{F}_{t-k+1})]) = 0 \end{aligned} \tag{2.7}$$

$$Cov(\lambda_t, \lambda_{t-k}) = a_1Cov(X_{t-1}, \lambda_{t-k}) + b_1Cov(\lambda_t, \lambda_{t-k}) \tag{2.8}$$

and by iterating (2.8) we obtain:

$$Cov(\lambda_t, \lambda_{t-k}) = (a_1 + b_1)^k\gamma_\lambda(0)$$

and

$$Cov(X_t, X_{t-k}) = Cov(X_t, \lambda_{t-k}) = Cov(\lambda_t, X_{t-k})$$

In order to prove the second order unconditional moment we could consider that $Var(X_t) = E(Var(X_t \mid \mathcal{F}_{t-1})) + Var(E(X_t \mid \mathcal{F}_{t-1}))$ or based on ARMA(1,1) model we have that $Var(X_t) = \frac{(1-(a_1+b_1)^2+a_1^2)E(X_t)}{1-(a_1+b_1)^2}$ where $E(X_t)$ given by (2.4). The drawback with this model is that parameters take only positive values according to the expression of volatility. To overcome this problem another model based on theory of GLM has been proposed by Fokianos and Tjøstheim (2011) assuming

$$X_t \mid \mathcal{F}_{t-1} \sim Poisson(v_t = \log(\lambda_t)) \tag{2.9}$$

where

$$v_t = d_1 + \sum_{i=1}^{p} a_i v_{t-i} + \sum_{j=1}^{q} b_j \log(X_{t-j} + 1) \tag{2.10}$$

In this model, parameters d_1, a_i, b_j for any $i = 1, \ldots, p$, $j = 1, \ldots, q$ take values in $\mathbb{R}$ and both negative and positive correlations can occur. Certainly other specifications are possible. In particular, one may consider a model for the log-mean process by introducing $\log(X_{t-j} + u)$, where u is a constant. Fokianos and Tjøstheim (2011) presented results of the data analysis that do not indicate any gross deviations in terms of the mean square error of residuals for values of u varying from 1 to 10 with a step equal to 0.5. We consider that u = 1 throughout the thesis and parameter λ_t is defined in (2.1), (2.2), (2.9), (2.10). Zeger and Qaqish (1988) studied parameters estimation approach without discussion for conditions of ergodicity and stationarity. Furthermore Shephard (1995) studied a similar model called Integer moving average OD model, where the main difference is that mean expressed log-linearly only from past values of observations. Fokianos and Kedem (2004) studied moments of the model given by (2.9), (2.10) and a different approach of estimation has been discussed.

2.1 Stationarity of linear and log-linear INGARCH models

Ferland et al. (2006) discussed first and second order stationarity based on backshift operator for the general INGARCH(p,q) process. Based on $B(d)\lambda_t = d_1 + A(d)X_t$ where $B(d) = 1 - B^d$ and $A(d) = 1 - A^d$ are lag-operators used to produce previous elements at volatilities and time series. Then $\lambda_t = B^{-1}(d)d_1 + B^{-1}dA(d)$. There are numerous criteria in order to prove process ergodicity. Fokianos et al. (2009) study conditions of ergodicity considering a pertubed model and based on criteria of ϕ-irreducibility, geometric ergodicity has been provided. Assuming that λ_t is a Markov chain, where $\lambda_t = d_1 + a_1\lambda_{t-1}$, Fokianos et al. (2009) tried to prove geometric ergodicity considering ϕ-irreducibility. This is equivalent to show that $\sum P^n(\lambda, A) > 0$ or $\phi(A) > 0$, where $P^n(\lambda, A) = P\{\lambda_n \in A \mid \lambda_0 = \lambda\}$. Streett (2000) proved that under certain conditions there exist a stationary initial distribution for λ_t by showing that a point λ^* is reachable. The problem is that some sets A of positive Lebesque measure are not open sets and for this reason Fokianos et al. (2009) used a perturbed model defined by

$$\lambda_t^m = d_1 + a_1\lambda_{t-1}^m + b_1 X_{t-1}^m + \epsilon_t^m$$

where $\epsilon_{t,m} = c_m \mathbb{1}(X_{t-1}^m = 1)U_t$, $\mathbb{1}(\cdot)$ is an indicator function, U_t is a sequence of iid uniform random variables appropriate to establish irreducibility. Furthermore $c_m \to 0$ when $m \to \infty$. The above indicate small differences between the initial and the perturbed model offering to provide infinite partial derivatives and estimator convergence for both of two models. Several specific cases and examples have been also proposed by Dedecker et al. (2007). Franke (2010) based on θ and τ weak dependence criteria proved stationarity of the univariate INGARCH process. The main idea for the ergodicity is that considering two convenient functions f,g, $Cov(f(past), g(future))$ tends to 0 as the distance between past and future becomes large, where 'past' and 'future' are elementary events given through finite dimensional marginals. For those criteria we study functions which satisfy Lipschitz condition. Considering Euclidean space $\mathbb{R}^m$ and $h : \mathbb{R}^m \to \mathbb{R}^n$

$$Liph = sup\left\{\frac{|h(\lambda_t, X_t) - h(\tilde{\lambda}_t, \tilde{X}_t)|}{|\lambda_t - \tilde{\lambda}_t| + |X_t - \tilde{X}_t|}\right\}$$

Let $\Lambda_1(\mathbb{R}^m)$ be the set of functions h where $Lip(h) \leq 1$

The above conditions indicate that model (2.1); (2.2) is ergodic and has a strictly stationary solution. Doukhan et al. (2012) studied weak dependence criteria for two INGARCH time series models studied also by Fokianos et al. (2009). In contrast for the model defined by (2.10) criteria of weak dependence cannot be proved easily because logarithmic function does not possess the contraction property. Fokianos and Tjøstheim (2011) provided conditions of geometric ergodicity based on the idea of Fokianos et al. (2009) and considering a perturbed model given by:

$$v_t^m = d_1 + a_1 v_{t-1}^m + b_1 \log(X_{t-1}^m + 1) + \epsilon_{t,m}$$

where similar to the linear model $\epsilon_{t,m} = c_m \mathbb{1}(X_{t-1}^m = 1)U_t$, U_t are iid uniformly distributed random variables, $\mathbb{1}()$ is the indicator function and $c_m > 0$, $c_m \to \infty$, $m \to \infty$. Considering the above study the chain v_t is ϕ-irreducible and geometric ergodic when $|a_1| < 1$ and $|a_1 + b_1| < 1$. Liboschik et al. (2015) proposed alternative and more flexible conditions for log-linear model's stationarity given by:

$$max\{|a_1|, |b_1|, |a_1 + b_1|\} < 1$$

Fokianos and Tjøstheim (2012) studied geometric ergodicity and moments for a class of non-linear models considering perturbed models

Similar to linear model considering conditions of stationarity,for the log-linear INGARCH(1,1) we obtain that:

$$\begin{aligned}
E(X_t) &= E(E(X_t \mid \mathcal{F}_{t-1})) = E(E(v_t)) = E(E(\log(\lambda_t))) \\
&= E(d_1 + a_1 E(\log(\lambda_{t-1})) + b_1 \log(Y_{t-1} + 1)) \\
&= d_1 + a_1 E(\log(\lambda_{t-1})) + b_1 E(\log(\lambda_{t-1}))
\end{aligned}$$

and

$$E(\log(\lambda_t)) = \frac{d_1}{1 - a_1 - b_1}.$$

2.2 Parameters estimation

Estimation of parameters is necessary before applying one of the methods for bayesian model selection. Problem of estimation via maximum likelihood is treated firstly by Ferland et al. (2006). The conditional likelihood function for the parameters of linear INGARCH model given by:

$$L(\mathbf{x} \mid \theta) = \prod_{t=1}^{n} \frac{e^{-\lambda_t} \lambda_t^{x_t}}{x_t!} \tag{2.11}$$

where λ_t is expressed via (2.2) and $\theta = (\theta_1, \theta_2, \ldots, \theta_{p+q+1}) = (\alpha_1, \ldots, \alpha_p, d_1, b_1, \ldots, b_q)$ is the vector of the unknown parameters and y_t the observed value at time t. The score function defined as :

$$S(\theta) = \frac{\partial l}{\partial \theta} = \sum_{t=1}^{n} \left(\frac{X_t}{\lambda_t} - 1 \right) \frac{\partial \lambda_t}{\partial \theta} \tag{2.12}$$

while the second derivatives are:

$$\frac{\partial^2 l}{\partial \theta_i \partial \theta_j} = \frac{\partial^2 \lambda_t}{\partial \theta_i \partial \theta_j} \left(\frac{X_t}{\lambda_t} - 1 \right) - \frac{X_t}{\lambda_t^2} \frac{\partial \lambda_t}{\partial \theta_i} \frac{\lambda_t}{\partial \theta_j}. \tag{2.13}$$

where ℓ is the log-likelihood given by:

$$l = \sum_{t=1}^{n} \lambda_t + x_t \log(\lambda_t - \log(x_t!)$$

First and second order derivatives are calculated and considering conditions of stationarity it is easily to prove that $E||\frac{\partial^2 l}{\partial\theta_i\theta_j}|| \leq \infty$.

$$\frac{\partial^2 l}{\partial a_i^2} = -\frac{X_t}{\lambda_t^2}\frac{\partial^2 \lambda_t}{\partial a_i^2} \tag{2.14}$$

$$E\left(\frac{\partial^2 l}{\partial\theta_i\partial\theta_j} \mid \mathcal{F}_{t-1}\right) = -\frac{1}{\lambda_t}\frac{\partial\lambda_t}{\partial\theta_i}\frac{\partial\lambda_t}{\partial\theta_j} \tag{2.15}$$

Considering Taylor series expansion for $\frac{\partial l}{\partial\theta}$ and considering Lemma 6 of Lee et al. (2016)

$$-\frac{1}{T}\frac{\partial^2 l(\theta_0)}{\theta_i\theta_j} \to I(\theta_0) \tag{2.16}$$

we have that

$$\sqrt{T}(\hat{\theta} - \theta_0) \to N(0, I(\theta_0)^{-1}). \tag{2.17}$$

Likelihood function defined similarly to (2.11) for the linear model and considering construction for the volatility

$$l(\theta) = \sum_{i=1}^{T}(X_t v_t - e^{v_t} - \log(X_t!))$$

On the other hand the log-likelihood function for the perturbed model given by:

$$l(\theta) = \sum_{t=1}^{n}\left(X_t^m v_t^m - e^{v_t^m} - \log(X_t^m!) + \log(f_u(U_t))\right)$$

where $f_u(\cdot)$ denotes the uniform density respectively, second order derivatives are given by:

$$\begin{aligned}
\frac{\partial^2 v_t(\theta)}{\partial d_1^2} &= a_1\frac{\partial^2 v_{t-1}}{\partial b_0^2} \\
\frac{\partial^2 v_t(\theta)}{\partial a_1^2} &= 2\frac{\partial v_{t-1}}{\partial a_1} + a_1\frac{\partial^2 v_{t-1}}{\partial a_1^2} \\
\frac{\partial^2 v_t(\theta)}{\partial b_1^2} &= a_1\frac{\partial^2 v_{t-1}}{\partial b_1^2}
\end{aligned}$$

Providing that difference between the perturbed and the unperturbed model can be eliminated as $m \to \infty$. Fokianos and Tjøstheim (2011) studied asymptotic normality and consistency of estimators by using the fact that the likelihood function of the perturbed model is three times differentiable. By using proposition 6.3.9 of Brockwell and Davis proved that the score function, information matrix and third derivatives tend to the quantities of the unperturbed model.

$$\begin{aligned} E\left(\frac{1}{\lambda_t^m}\frac{\partial \lambda_t^m}{\partial \theta}\frac{\partial \lambda_t^m}{\partial \theta}^T\right) &\to E\left(\frac{1}{\lambda_t}\frac{\partial \lambda_t}{\partial \theta}\frac{\partial \lambda_t}{\partial \theta}^T\right) \\ \frac{1}{\sqrt{n}}\frac{\partial^2 l^m(\theta_0)}{\partial \theta_i \partial \theta_j} &\to I^m(\theta_0)\ \ as\ \ T \to \infty \\ I^m(\theta_0) &\to I(\theta_0)\ \ as\ \ m \to \infty \end{aligned} \tag{2.18}$$

2.3 Parameters estimation based on Bayesian statistics methods

Using Bayes theorem posterior density $p(\theta \mid \mathbf{y})$ for the parameter vector $\theta = (\theta_1, ..., \theta_{p+q+1}) = (a_1, a_2, \ldots, a_p, d_1, b_1, \ldots, b_q)$ can be derived. We have as usual that

$$\begin{aligned} p(\theta \mid \mathbf{x}) &\propto L(\mathbf{x} \mid \theta)p(\theta) \\ &\propto L(\mathbf{x} \mid \theta)p(\alpha_1)\ldots p(\alpha_p)p(d_1)\ldots p(b_q) \end{aligned} \tag{2.19}$$

where $L(\mathbf{x} \mid \theta)$ is the likelihood function and $p(\theta)$ is the prior distribution. Here we assume independent priors for all the parameters. We use MCMC method to obtain samples from the posterior distribution. Based on Bayes theorem and under the assumption of stationarity we choose as prior distributions truncated normals for the parameters α_i and b_j where $0 < \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} b_j < 1$ or $-1 < \alpha_i, b_j < 1$ and $-1 \leq \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} b_j < 1$ in linear and log-linear case accordingly and a Gamma(1,2) prior distribution for the parameter d_1 and after Metropolis-Hasting steps we reject values that they do not satisfy stationarity conditions.

Given that $X_t \mid \mathcal{F}_{t-1} \sim Poisson(\lambda_t)$, where λ_t is expressed via equations (2.2) or (2.10), posterior density of parameter vector $\theta = (\alpha_1, ..., \alpha_p, d_1, b_1, ...b_q)$ can be expressed by equation

$$p(\theta \mid \mathbf{x}) \;\propto\; \prod_{t=1}^{n} e^{-\lambda_t}\lambda_t^{x_t} \times \frac{e^{\frac{-1}{2\sigma^2_{\alpha_1}}(\alpha_1-\mu_{\alpha_1})^2}}{\Phi\left(\frac{1-\mu_{a_1}}{\sigma^2_{a_1}}\right)-\Phi\left(\frac{0-\mu_{a_1}}{\sigma^2_{a_1}}\right)} \times \ldots \times \frac{e^{\frac{-1}{2\sigma^2_{\alpha_p}}(\alpha_p-\mu_{\alpha_p})^2}}{\Phi\left(\frac{1-\sum_{i=1}^{p-1}a_i-\mu_{\alpha_p}}{\sigma^2_{\alpha_p}}\right)-\Phi\left(\frac{u-\mu_{\alpha_p}}{\sigma^2_{\alpha_p}}\right)} \times$$
$$\frac{e^{\frac{-1}{2\sigma^2_{b_1}}(b_1-\mu_{b_1})^2}}{\Phi\left(\frac{\sum_{i=1}^{p}a_i-\mu_{b_1}}{\sigma^2_{b_1}}\right)-\Phi\left(\frac{u-\mu_{b_1}}{\sigma^2_{b_1}}\right)} \times \ldots \times \frac{e^{\frac{-1}{2\sigma^2_{b_q}}(b_q-\mu_{b_q})^2}}{\Phi\left(\frac{1-\sum_{i=1}^{p}a_i-\sum_{j=1}^{q-1}b_j-\mu_{b_q}}{\sigma^2_{b_q}}\right)-\Phi\left(\frac{u-\mu_{b_q}}{\sigma^2_{b_q}}\right)} \times$$
$$\frac{d_1^{\gamma-1}e^{-\frac{d_1}{\beta}}}{\gamma\left(\gamma_1,\frac{d_1}{\beta}\right)} \times I(\theta \in \mathbb{S}) \tag{2.20}$$

where

$$\mathbb{S} = \{\alpha_1,\ldots a_p, b_1,\ldots,b_q\} : = \begin{cases} \sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j \in (0,1) \\ |\alpha_i|, \quad |b_j| < 1, \quad |\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q} b_j| < 1, \end{cases}$$

,

$$u \;=\; \begin{cases} 0, \text{ if } \sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j \in (0,1) \\ -1, \text{ if } |\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q} b_j| < 1. \end{cases}$$

and $\gamma\left(\gamma_1,\frac{d_1}{\beta}\right)$ is the lower incomplete gamma function. We use $p+q+1$ independent Metropolis steps (Tierney (1994);Chib and Greenberg (1995)) for all parameters θ_i and as proposal distribution $q(\theta^{(can)} \mid \theta_1,\theta_2,\ldots,\theta_{p+q+1})$ we use $N(\theta,\sigma^2)$ distributions. The choice for this variance σ^2 is important for the convergence of the chain. Considering the conditions of stationarity, the interval for each parameter our work is $(0,1)$ and $(-1,1)$ at the linear and the log-linear cases respectively but for the parameter d_1 the positivity is the only restriction in both cases. In the work of Fokianos and Tjøstheim (2011) the restrictions of the parameters have been discussed, because they provide different correlation structure and they are deterministic for the calculation of the maximum likelihood estimators. Acceptance probability at the j-th iteration is calculated by

$$\alpha = min\left\{1, \frac{\prod_{t=1}^{T} e^{-\lambda_t^{(can)}}(\lambda_t^{(can)})^{y_t} p(\theta_i^{(can)})}{\prod_{t=1}^{T} e^{-\lambda_t^{(j)}}(\lambda_t^{(j)})^{y_t} p(\theta_j^{(can)})}\right\}, \tag{2.21}$$

where $\lambda^{(can)}$ implies that it has been calculated based on the candidate value for the parameter. Marginal posterior means and posterior variances are used on trans-dimensional MCMC method for the definition of pseudopriors.

Denote as $\theta = (\alpha_1 \ldots, \alpha_p, d_1, b_1, \ldots, b_{p+q})$ the vector with all the parameters. We describe the MCMC algorithm in a general form applicable to both linear and log-linear models. The key distinction is the conditions that need to be satisfied for each model to ensure stationarity.

We use $p+q+1$ independent Metropolis steps for all parameters θ_i, $i = 1, \ldots, p+q+1$. The algorithm runs as follows

- Initialize $\theta^{(0)} = (\alpha_1^{(0)}, \ldots, \alpha_p^{(0)}, b_1^{(0)}, \ldots, b_q^{(0)}, d_1^{(0)})$ from truncated normal distributions for α's and $b_1, \ldots, b_q$ and a Gamma for d_1 such as to satisfy the stationarity conditions of the particular model
- For the $r+1$-th iteration of the algorithm For each θ_i, $i = 1, \ldots, p+q+1$
  - Generate values from proposal densities for $\theta_i^{(can)} \sim N(\theta_i, \sigma_i^2)$ where the tuning parameter σ_i^2 was selected such as to achieve 25% acceptance rate.
  - Update $\theta_i^{(r)}$ to $\theta_i^{(r+1)}$ with probability given by (2.21) and if the candidate value does not violate the stationarity conditions of the model
  - Iterate this updating procedure.
- Go to the next iteration or stop if the chain has converged convergence

2.4 Methods for Bayesian model selection

While the literature on discrete valued time series is increasing very fast, there are very few papers that consider the problem of model selection, namely that of selecting the order of the model. In simple INAR(p) models this is merely the order of autoregressive terms, in our INGARCH setting this relates to the order of p and q to be used. An efficient reversible jump Markov chain Monte Carlo (RJMCMC) algorithm is constructed by Enciso-Mora et al. (2009) for moving between INARMA processes of different orders. See also work for model selection in INAR(p) models in Bu and McCabe (2008).

In addition, Alzahrani et al. (2018) studied the model selection problem among INAR and linear INGARCH models by using particle filtering MCMC method. Wang et al. (2020) used a penalized conditional maximum likelihood to estimate the parameters of an INGARCH model. Note that this approach while can help to identify useful terms in the model it does not select between competing models. Assume that we have a countable set of models denoted by $\mathcal{M}$ for a given set of data $\mathbf{x}$. Let consider a model indexed by $m \in \mathcal{M}$, we denote the vector of unknown parameters for model m as $\theta_m \in \Theta_m$. Two models, say m_1 and m_2, can have different dimensions. Let $p(m)$ the prior distribution for each model m. Then posterior probability for model $m \in \mathcal{M}$ is given by

$$p(m \mid \mathbf{x}) = \frac{p(m)L(\mathbf{x} \mid m)}{\sum_{m' \in \mathcal{M}} p(m')L(\mathbf{x} \mid m')}, \tag{2.22}$$

where $L(\mathbf{x} \mid m)$ is the marginal likelihood for model m calculated by the integral $\int L(\mathbf{x} \mid \theta_m)p(\theta_m \mid m)d\theta_m$ and $p(\theta_m \mid m)$ is the prior distribution of the parameter vector θ_m. There are many approaches for model selection problem using Bayes factor of model m_i against model m_j. Basically there are two interpretations for Bayes factor. Firstly Bayes factor is the ratio of two marginal likelihoods, representing how well the data are fitted under each model. More analytically

$$BF = \frac{p(m_i \mid \mathbf{x})p(m_j)}{p(m_j \mid \mathbf{x})p(m_i)} = \frac{\int_{\Theta_{m_i}} p(\mathbf{x} \mid m_i, \theta_{m_i})p(\theta_{m_i} \mid m_i)d\theta_{m_i}}{\int_{\Theta_{m_j}} p(\mathbf{x} \mid m_j, \theta_{m_j})p(\theta_{m_j} \mid m_j)d\theta_{m_j}}. \tag{2.23}$$

Secondly, we could say that posterior model odds can be represented as

$$\frac{p(m_i \mid \mathbf{x})}{p(m_j \mid \mathbf{x})} = B_{ij} \times \frac{p(m_i)}{p(m_j)}. \tag{2.24}$$

The integrals required for Bayes factor are analytically intractable. Consequently many methods have been proposed to approximate Bayes factor. Green (1995) proposed the so called reversible jump MCMC method allowing jumps between models with low computational cost. The idea of Dellaportas et al. (2002) based on the above methods. Reversible jump MCMC is more flexible because proposals only for a subset of parameters is required. Assuming that we want to make a jump from model m_i with parameter vector $\theta_{m_i} = (\theta_1, \theta_2)$ to model m_j with parameter vector $\theta_{m_j} = (\theta_1, \theta_2, \theta_3')$ where those two models are nested. For this reason we generate u from a specified proposal density $q(u \mid \theta_{m_i}, m_i, m_j)$. When we are in a model where those 2 models are nested, a proposal only for the third parameter is needed. In the general case $(\theta_{m_j}, u') = g_{m_i m_j}$ where $g_{m_i m_j}$ is a specified function where $g_{m_i m_j} = g^{-1}_{m_i m_j}$. Then the acceptance probability is calculated by:

$$\frac{p(x \mid m_j, \theta_{m_j})p(\theta_{m_j} \mid m_j)}{p(x \mid m_i, \theta_{m_i})p(\theta_{m_i} \mid m_i)q(u \mid \theta_{m_i}, m_i, m_j)}.$$

Methods, introduced by Green (1995), Carlin and Chib (1995), Dellaportas et al. (2002), generate observations from the joint posterior distribution from (m, θ_m), for estimating $p(\theta_m \mid m)$. A method introduced by Carlin and Chib (1995) proposes jumps between all models. The full conditional posterior density for each model is given by

$$p(m \mid \theta_m, \mathbf{x}) = \frac{A_m}{\sum_{m' \in \mathcal{M}} A_{m'}}, \tag{2.25}$$

where numerator is given by equation

$$A_m = p(\mathbf{x} \mid \theta_m, m) \prod_{m' \in \Updownarrow} p(\theta_{m'} \mid \mathbf{x}) p(m'). \tag{2.26}$$

There are alternative constructions proposed by Bayarri et al. (2012) in the case where the models are nested. The method proposed by Carlin and Chib (1995) requires the calculation of all $A_{m'}$ in the denominator at each iteration. An alternative method called as "Metropolised Carlin and Chib" proposed by Dellaportas et al. (2002). The main difference with the method proposed by Chib and Greenberg (1995) is that at this study for the model selection step only a proposal for the move to the new model is required. Assuming that we are in model m_i and we suggest a move to model m_j with probability $j(m_i, m_j)$. Only pseudopriors for parameters $\theta_{\mathbf{m_i}}$ and $\theta_{\mathbf{m_j}}$ are required. The acceptance probability calculated as:

$$\frac{p(x \mid m_j, \theta_{m_j}) p(\theta_{m_j} \mid m_j) p(\theta_{m_i} \mid m_j) p(m_j) j(m_j, m_i)}{p(x \mid m_i, \theta_{m_i}) p(\theta_{m_i} \mid m_i) p(m_j \mid \theta_{m_i}) p(m_i) j(m_i, m_j)}. \tag{2.27}$$

Among three above methods the less mathematically complicated is this of Carlin and Chib (1995) because only appropriate proposal distributions are needed for all models' parameters. For this reason the parametric space becomes large and computational methods more intensive than those of Green (1995) and Dellaportas et al. (2002). On the other hand the more complicated method for Bayesian model selection of Green (1995) the less computationally intensive.

Reversible jump MCMC is a more complicated method in our case because we have a certain number of models, so an appropriate transformation is necessary to define each time the parameter spaces for making jumps between models. Metropolised Carlin and Chib method proposes a move from one model to some other and examine the acceptance or rejection of this proposal. Lodewyckx et al. (2011) suggested the product space method both in the cases of nested and non-nested models.

For example let us consider three models, namely INGARCH(1,1), INGARCH(1,2) and INGARCH(2,1) denoted as m_1, m_2 and m_3. The parameter vectors are $\theta_{m_1} = (\alpha_1, d_1, b_1)$, $\theta_{m_2} = (\alpha_1, \alpha_2, d_1, b_1)$ and $\theta_{m_3} = (\alpha_1, d_1, b_1, b_2)$. The three parameter vectors are combined in one parameter vector $\theta = (\theta_{m_1}, \theta_{m_2}, \theta_{m_3})$ which takes any value of the Cartesian product of the three models' parameter spaces $\Theta_{m_1} \times \Theta_{m_2} \times \Theta_{m_3}$.

In this method and in the method introduced by Dellaportas et al. (2002) there is no linked densities and also the calculation of the Jacobian is not needed. Those methods require proposal densities of higher dimensions, called pseudopriors.

The crucial problem here is the matching of dimensions of the three models. For doing this consider that if model m_1 is visited then the corresponding parameter vector θ_{m_1} is connected to the model likelihood. So this parameter vector can be updated from posterior distributions. If m_1 is not visited then the parameter vector is disconnected from the likelihood and is generated from pseudopriors. A similar reasoning can be followed for the models m_2 and m_3. For example, for model m_1 the joint distribution is defined by

$$\begin{aligned} p(\mathbf{x}, \theta \mid m_1) &= p(\mathbf{x} \mid \theta, m_1) p(\theta \mid m_1) \\ &= p(\mathbf{x} \mid \theta, m_1) p(\theta_{m_1} \mid m_1) p(\theta_{m_2} \mid m_1) p(\theta_{m_3} \mid m_1), \end{aligned} \tag{2.28}$$

where $p(\theta_{m_2} \mid m_1)$, $p(\theta_{m_3} \mid m_1)$ are proper distributions integrating to 1. Then for model m_i, $i = 1, 2, 3$ we sample from

$$p(m_i \mid \theta, \mathbf{x}) \propto \begin{cases} L(\mathbf{x} \mid \theta, m_1) p(\theta_{m_1} \mid m_1) p(\theta_{m_2} \mid m_1) p(\theta_{m_3} \mid m_1) p(m_1), & \text{for } i = 1 \\ L(\mathbf{x} \mid \theta, m_2) p(\theta_{m_2} \mid m_2) p(\theta_{m_1} \mid m_2) p(\theta_{m_3} \mid m_2) p(m_2), & \text{for } i = 2 \\ L(\mathbf{x} \mid \theta, m_3) p(\theta_{m_3} \mid m_3) p(\theta_{m_2} \mid m_3) p(\theta_{m_1} \mid m_3) p(m_3), & \text{for } i = 3 \end{cases}$$

where

$$p(\theta_{m_1} \mid m_i) \propto \begin{cases} p(x \mid \theta_{m_1}, m_1) p(\theta_{m_1} \mid m_i) & \text{if } i = 1 \\ p(\theta_{m_1} \mid m_i) & \text{if } i = 2 \\ p(\theta_{m_1} \mid m_i) & \text{if } i = 3 \end{cases} \tag{2.29}$$

and $p(m_1), p(m_2), p(m_3)$ are prior model probabilities. Those methods are deterministic for the time that the algorithm visit and stay at each model. Finally posterior model probabilities for model $m \in \mathcal{M}$ are estimated by

$$\hat{p}(m \mid \mathbf{x}) = \frac{\sum_{i=1}^{B} I(m^{(i)} = m)}{B}, \tag{2.30}$$

where B is the total number of iterations and $m^{(i)}$ denotes the model we are in the i-th iteration. We will apply this trans-dimensional approach to jump between different INGARCH models. A detailed description of the trans-dimensional MCMC can be found in the Appendix B. Predictions in time varying volatility in linear and log-linear INGARCH models can be obtained from the output of the MCMC. McCabe and Martin (2005) studied methods for prediction of INAR(1) model. Raftery et al. (1997) discussed model averaging and the choice of models for Bayesian model averaging. A p-step ahead predictive pmf is defined as

$$p(x_{T+p} \mid \mathbf{x}) = \sum_{m \in \mathcal{M}} p(x_{T+p} \mid \mathbf{x}, m) p(m \mid \mathbf{x}), \tag{2.31}$$

where $\mathcal{M}$ denotes the countable set of candidate models and $P(x_{T+p} \mid \mathbf{x}, m)$ is the p-step ahead predictive pmf from model m and $P(m \mid \mathbf{x})$ is the posterior probability of model m calculated in trans-dimensional MCMC. So in our work

$$p(\lambda_{T+1} \mid \mathbf{x}) = \sum_{m \in \mathcal{M}} p(\lambda_{T+1} \mid \mathbf{x}, m) p(m \mid \mathbf{x}), \tag{2.32}$$

which is an average of the posterior predictive distribution under each model weighted by their posterior model probabilities. For each parameter vector we calculate at each iteration λ_{T+1}. Each sampled point should be taken with probability $p(m \mid \mathbf{x})$. Then we obtain the sample of $p(\lambda_{T+1} \mid \mathbf{x})$ by weighting all samples of $p(\lambda_{T+1} \mid m, \mathbf{x})$ by the corresponding $p(m \mid \mathbf{x})$.

Consider s candidate models, namely $m_1, \ldots, m_s$. We denote the parameter vector for model m as θ_m. Due to the different number of parameters for each model they may have different dimension. Also denote the set of indices not having j as $S_{-j} = \{1, 2, \ldots, j-1, j+1, \ldots, s\}$. We have obtained from the MCMC from all the models the mean $\tilde{\mu}_{\theta_{ij}}$ and the variance $\tilde{\sigma}^2_{\theta_{ij}}$ of parameter θ_{ij} which is the i-th parameters of the j-th model. Then the algorithm runs as follows.

For a number of iterations say B we work as follows

1. For model m_k we generate the parameter vector for this model, say θ_{m_k} from the conditional posterior distribution $L(\mathbf{x} \mid \theta_{m_k}, m_k) p(\theta_{m_k} \mid m_k)$ given by (2.20) and the other parameters θ_{m_j}, from proper pseudopriors $p(\theta_{m_j} \mid m_j) = \prod_{i=1}^{p_j} N(\tilde{\mu}_{\theta_{ij}}, \tilde{\sigma}^2_{\theta_{ij}})$ for $j \in S_{-k}$ where p_k denotes the number of parameters for model k.

2. We repeat this for each candidate model so for all $k = 1, \ldots, s$.

3. Calculate posterior probability for model k defined by

$$Pr_k = p(m_k \mid \theta, \mathbf{x}) = \frac{A}{B},$$

where

$$A = L(\mathbf{x} \mid \theta_{m_k}, m_k) p(\theta_{m_k} \mid m_k) p(m_k) \times \prod_{j \in S_{-k}} \prod_{i=1}^{p_j} \frac{1}{\sqrt{2\pi}\tilde{\sigma}_{\theta_{ij}}} \exp\left(-\frac{(\theta_{ij} - \tilde{\mu}_{\theta_{ij}})^2}{2\tilde{\sigma}^2_{\theta_{ij}}}\right)$$

and

$$B = \sum_{w=1}^{s} L(\mathbf{x} \mid \theta_{m_w}, m_w) p(\theta_{m_w} \mid m_w) p(m_w) \times \prod_{j \in S_{-w}} \prod_{i=1}^{p_j} \frac{1}{\tilde{\sigma}_{\theta_{ij}} \sqrt{2\pi}} \exp\left(-\frac{(\theta_{ij} - \tilde{\mu}_{\theta_{ij}})^2}{2\tilde{\sigma}^2_{\theta_{ij}}} \right).$$

4. We select the model with larger value Pr_k. this is the model selected at the i-th iteration,

5. Repeat steps 1-3. for B times.

6. At the end calculate s posterior model probabilities given by (2.30).

2.4.1 Simulations

In order to examine if our approach can identify the correct structure of the time series we study a small simulation study considering two criteria proposed by Liang et al. (2008) for model selection consistency. These are:

CRITERION 1: Each conditional prior must be proper (integrating to one) and cannot be arbitrarily vague in the sense of almost all of its mass being outside any believable compact set.

CRITERION 2: (Model selection consistency) If data $\mathbf{y}$ have been generated from model M_i then posterior model probability of model M_i should converge to 1 as the sample size n $\longrightarrow \infty$.

If criterion 2 is not accomplished may be a different choice of prior distributions is necessary. In our case for the accomplishment of criterion 1 we suggest as pseudo-priors normal densities by taking means and variances after "pilot runs" of MCMC for each model and considering that those pseudopriors are proper. More specifically we generate 100 datasets of size $n = 200$ with randomly chosen parameter values for each model, and we run a trans-dimensional Markov chain with length 10000. According to criterion 2 posterior model probability of model from which we generate the data, must be close to 1 while the probabilities must be small for the other models. Furthermore the choice of prior model probabilities is deterministic for the definition of the posterior model probabilities. Lodewyckx et al. (2011) proposed bisection algorithm to choose prior model probabilities which have a vital role at the algorithm. We present the results of averages of 100 replications for each model. The columns correspond to the frequency of selecting each one of 10 models, while the rows are the true models that generated the data. Note that in addition to the 10 INGARCH models we have two additional models different from INGARCH type. The first model was an INAR(2) model of the form $X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + R_t$, where $\circ$ is the binomial thinning operator and R_t are innovations terms that follow a Poisson distribution, while the second model was an INAR(1) model of the form $X_t = \alpha_1 \circ X_{t-1} + R_t$, $t = 1, \ldots, n$ where now the innovation term depends on some covariate, namely $R_t \sim Poisson(\exp(\gamma_1 z_t))$ with $Z_t \sim Ber(\zeta)$ where $\gamma_1 = 2$ and $\zeta = 0.2$. So, reading from table (2.1) 91.5% of the times we have generated the data from model M_1 we found this model as the chosen one.

Looking the results from Table (2.1) we can see that even for small sample size ($n = 100$) we can identify the correct structure with great success for most of the models. Furthermore based on the criteria mentioned above and considering results on Tables (2.2) and (2.3), when sample sizes are $n = 200$ and 500 respectively, we have increasing probability to find the correct model.

Overall, the experiment shows that the method can identify the underlying model with good success. Our approach correctly detects the model with high probability.

Model	M_1	M_2	M_3	M_4	M_5	LM_1	LM_2	LM_3	LM_4	LM_5
M_1 = INGARCH(1)	90.57	1.74	1.13	3.48	0.09	2.55	0.24	0.10	0.07	0.03
M_2 = INGARCH(1,1)	0.04	81.43	0.59	16.12	0.06	0.86	0.63	0.17	0.10	0.00
M_3 = INGARCH(1,2)	0.04	1.23	96.76	0.17	0.31	0.09	0.10	0.42	0.10	0.38
M_4 = INGARCH(2,1)	0.00	3.01	0.42	93.72	0.38	2.07	0.36	0.01	0.03	0.00
M_5 = INGARCH(2,2)	0.00	1.25	2.84	0.09	87.34	0.02	7.32	1.02	0.01	0.11
LM_1 = INGARCH(1)	0.14	0.18	0.16	0.09	0.07	97.38	0.94	0.43	0.31	0.28
LM_2 = INGARCH(1,1)	0.08	0.37	0.26	0.14	0.01	3.01	90.72	1.72	2.34	1.35
LM_3 = INGARCH(1,2)	0.02	0.10	0.08	0.02	0.00	0.00	2.35	87.76	9.32	0.37
LM_4 = INGARCH(2,1)	0.10	0.04	0.29	0.20	0.00	0.28	8.72	0.41	84.64	5.32
LM_5 = INGARCH(2,2)	0.00	0.00	0.04	0.00	0.12	0.05	0.42	1.07	8.28	89.02
IN2=INAR(2)	6.28	37.48	4.72	12.29	1.24	22.12	12.28	2.71	2.98	0.10
INexp=Exponential INAR(1)	0.35	40.16	3.82	11.96	2.53	7.60	24.72	2.91	5.21	0.44

Table 2.1: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size n=100, M_i is the class of linear and LM_i for log-linear INGARCH models accordingly.

Model	M_1	M_2	M_3	M_4	M_5	LM_1	LM_2	LM_3	LM_4	LM_5
M_1 = INGARCH(1)	91.50	4.18	0.19	2.16	0.03	1.75	0.10	0.06	0.01	0.01
M_2 = INGARCH(1,1)	0.10	78.73	0.39	18.91	0.15	0.79	0.45	0.21	0.21	0.07
M_3 = INGARCH(1,2)	0.01	0.56	98.27	0.09	0.22	0.01	0.08	0.36	0.05	0.36
M_4 = INGARCH(2,1)	0.00	2.82	0.60	92.80	0.47	2.50	0.42	0.07	0.26	0.05
M_5 = INGARCH(2,2)	0.00	1.19	3.22	0.27	82.98	0.26	10.24	1.08	0.18	0.56
LM_1 = INGARCH(1)	0.00	0.00	0.00	0.00	0.00	99.05	0.41	0.28	0.15	0.11
LM_2 = INGARCH(1,1)	0.00	0.01	0.01	0.00	0.00	2.81	91.23	2.19	2.77	0.97
LM_3 = INGARCH(1,2)	0.00	0.01	0.00	0.00	0.00	0.00	1.92	89.13	7.63	1.29
LM_4 = INGARCH(2,1)	0.03	0.09	0.01	0.04	0.01	0.05	9.46	8.17	72.88	9.27
LM_5 = INGARCH(2,2)	0.00	0.00	0.01	0.00	0.04	0.02	0.77	1.67	9.95	87.54
IN2=INAR(2)	4.67	31.96	4.20	13.19	1.42	25.24	14.30	1.71	3.12	0.19
INexp=Exponential INAR(1)	0.15	41.50	3.10	13.01	1.32	8.05	26.81	1.33	4.54	0.19

Table 2.2: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size is n=200, M_i is the class of linear and LM_i for log-linear INGARCH models accordingly.

Model	M_1	M_2	M_3	M_4	M_5	LM_1	LM_2	LM_3	LM_4	LM_5
M_1 = INGARCH(1)	95.83	1.87	0.21	1.12	0.03	0.80	0.08	0.04	0.01	0.01
M_2 = INGARCH(1,1)	0.07	90.79	0.32	5.03	0.09	2.72	0.55	0.24	0.13	0.06
M_3 = INGARCH(1,2)	0.01	0.32	99.03	0.08	0.11	0.06	0.05	0.18	0.03	0.13
M_4=INGARCH(2,1)	0.03	2.13	0.42	95.62	0.27	1.01	0.23	0.14	0.10	0.05
$M_5 = INGARCH(2,2)$	0.06	0.89	1.72	0.11	94.07	0.10	2.32	0.43	0.09	0.21
LM_1 = INGARCH(1)	0.01	0.04	0.02	0.01	0.02	99.36	0.25	0.12	0.10	0.07
LM_2 = INGARCH(1,1)	0.01	0.04	0.03	0.02	0.01	1.73	96.46	0.34	1.12	0.24
LM_3 = INGARCH(1,2)	0.00	0.13	0.02	0.09	0.01	0.00	0.43	95.72	3.42	0.18
LM_4 = INGARCH(2,1)	0.08	0.97	0.11	0.17	0.14	0.24	3.98	1.14	91.34	1.83
LM_5 = INGARCH(2,2)	0.02	0.04	0.11	0.08	0.28	0.13	0.52	1.07	4.32	93.43
IN2=INAR(2)	2.25	49,12	2.09	3.12	1.34	26.32	17.21	1.82	2.07	0.03
INexp=Exponential INAR(1)	0.01	56.13	1.98	9.52	0.08	6.04	23.06	0.16	3.01	0.01

Table 2.3: Averages of posterior model probabilities after 100 samples of data from each model (multiplied by 100) where sample size n=500, M_i is the class of linear and LM_i for log-linear INGARCH models accordingly.

2.4.2 Applications

Polio data

The data consist of monthly counts of poliomyelitis cases in the United States from 1970 to 1983 (168 observations) reported by the Centres for Disease Control and discussed in Zeger (1988) among others. Figure (2.1) presents the original series and the autocorrelation function of the series. In trans-dimensional method we consider jumps between five linear and five log-linear INGARCH model.

Table (2.4) presents estimators for each of five linear and log-linear INGARCH(p, q) models. The proposal densities are normal distributions where the value for the variance is very small and was selected to achieve acceptance rate between 25% and 30%,

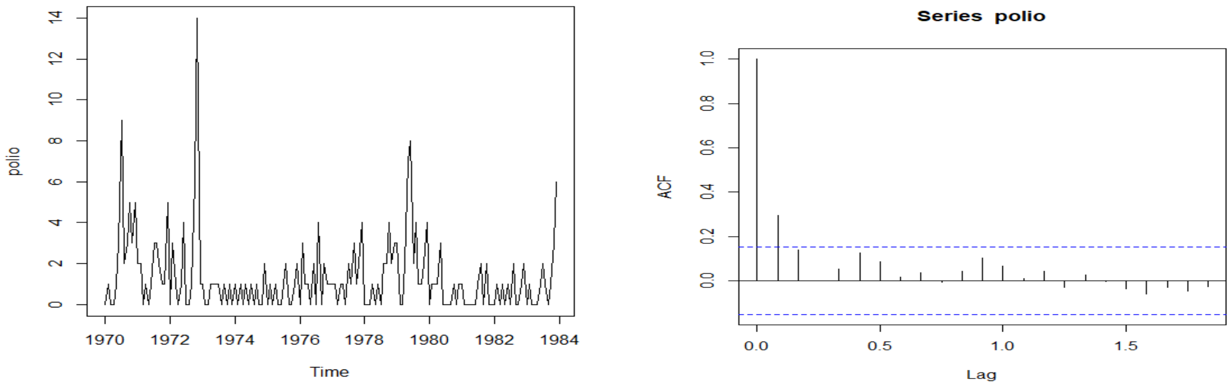


Figure 2.1: Series of polio dataset (left) and its ACF (right)

namely it has a value such as $0.01 < \sigma^2 < 0.025$ for the parameters $\alpha_1, ..\alpha_p, b_1, ..., b_q$ and 0.04 for the parameter d_1 in linear case and 0.01 in log-linear case accordingly. The chain was run for 10000 to 20000 iterations depending on the model. A burn-in part of 2500 iterations was discarded and the resulting samples were checked for convergence by using the test proposed by Geweke (1992). Z-scores are presented in Table (2.5) and indicate that the convergence has been achieved.

Model		linear					log-linear				
Parameters		d_1	α_1	b_1	α_2	b_2	d_1	α_1	b_1	α_2	b_2
INGARCH(0,1)	mean	0.86	-	0.37	-	-	0.04	-	0.47	-	-
	sd	0.09	-	0.07	-	-	0.03	-	0.07	-	-
INGARCH(1,1)	mean	0.62	0.20	0.35	-	-	0.04	0.04	0.44	-	-
	sd	0.15	0.12	0.07	-	-	0.03	0.15	0.08	-	-
INGARCH(1,2)	mean	0.67	0.11	0.34	-	0.08	0.05	-0.47	0.39	-	0.35
	sd	0.12	0.09	0.07	-	0.06	0.05	0.32	0.08	-	0.19
INGARCH(2,1)	mean	0.54	0.19	0.33	0.09	-	0.07	0.06	0.54	-0.52	-
	sd	0.15	0.11	0.07	0.08	-	0.07	0.14	0.08	0.11	-
INGARCH(2,2)	mean	0.56	0.11	0.31	0.09	0.08	0.08	-0.49	0.47	-0.54	0.42
	sd	0.14	0.09	0.07	0.05	0.08	0.08	0.22	0.09	0.12	0.16

Table 2.4: Estimated Posterior Means and standard deviations(sd) for the competing models.

For the models where parameters satisfy stationarity conditions we apply trans-dimensional MCMC method of Lodewyckx et al. (2011) and posterior probabilities are presented in Table (2.6). The INGARCH(1,1) model is the one mostly visited which indicates that this is the selected model. Note that the 4 best models are of the linear type while the log-linear models have much smaller posterior probabilities. Bayes factors are also reported in Table (2.6). They have been calculated with

Model	linear					log-linear				
Parameters	d_1	α_1	b_1	α_2	b_2	d_1	α_1	b_1	α_2	b_2
INGARCH(0,1)	0.454	-	0.128	-	-	-0.081	-	-1.444	-	-
INGARCH(1,1)	-0.677	0.737	-0.842	-	-	-0.455	-0.129	-0.481	-	-
INGARCH(1,2)	0.182	-0.117	-0.854	-	1.178	1.292	-0.109	-0.192	-	0.191
INGARCH(2,1)	1.644	0.642	-1.281	-0.848	-	-1.163	-0.394	0.057	0.478	-
INGARCH(2,2)	1.724	-0.849	-0.404	0.283	-0.733	0.644	0.708	0.857	0.870	-0.825

Table 2.5: Geweke's convergence z-scores for the parameters of five linear and five log-linear INGARCH models

respect the last (and more complicated) model. They also support simple models.

Model	Posterior probability	Bayes Factor
INGARCH(1,1)	0.4564	3245.066
INGARCH(0,1)	0.3741	2659.801
INGARCH(1,2)	0.0641	455.838
INGARCH(2,1)	0.0588	418.188
LINGARCH(0,1)	0.0341	242.644
INGARCH(2,2)	0.0085	60.462
LINGARCH(1,2)	0.0023	16.199
LINGARCH(2,1)	0.0016	11.527
LINGARCH(2,2)	0.0001	1.000

Table 2.6: Posterior probabilities and Bayes factor for the competing models

Finally, considering that we have posterior model probabilities for all models, we apply a Bayesian model-averaging (BMA) procedure. Based on posterior model probabilities, as estimated via trans-dimensional MCMC, we derive the density of predictive volatilities. Consequently we calculate $p(\lambda_{T+1} \mid m, \mathbf{y})$, i.e. the mean of the next unseen observation, weighted at its posterior model probability. We present densities for the four linear models with the higher posterior model probabilities in Figure (2.3). As expected the BMA estimate summarizes the individual densities.

Campylobacterosis data

The data (e.g. Liboschik et al., 2015) consist of number of campylobacterosis cases in the North of Québec in Canada. Figure (2.2) presents the original series and the autocorrelation function of the series. Similarly, as for the polio dataset, we present parameters estimation and Z-scores concerning and checking the convergence. Considering same conditions of stationarity as for the polio dataset we present in Table 7 results after the trans-dimensional MCMC algorithm. In this case we observe that the 4 best models are not only linear but also they have a complicated structure. Furthermore the simple model INGARCH(1,0) is the less preferable.

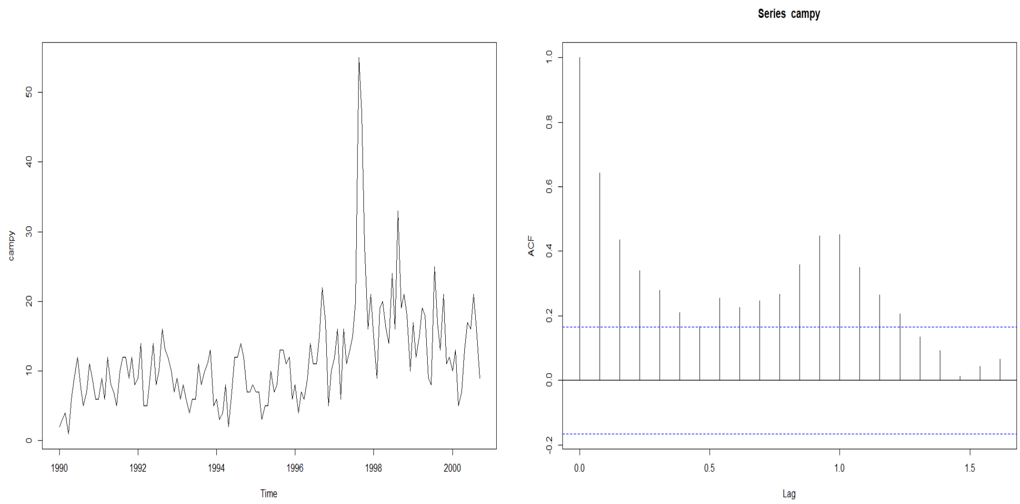


Figure 2.2: Series of campylobacterosis dataset (left) and its ACF (right)

Model		linear					log-linear				
Parameters		d_1	α_1	b_1	α_2	b_2	d_1	α_1	b_1	α_2	b_2
INGARCH(0,1)	mean	2.81	-	0.67	-	-	0.68	-	0.71	-	-
	sd	0.44	-	0.66	-	-	0.21	-	0.09	-	-
INGARCH(1,1)	mean	1.82	0.32	0.52	-	-	0.41	0.23	0.59	-	-
	sd	0.43	0.08	0.07	-	-	0.12	0.08	0.07	-	-
INGARCH(1,2)	mean	1.87	0.27	0.52	-	0.05	0.41	0.24	0.60	-	-0.02
	sd	0.45	0.10	0.05	-	0.07	0.22	0.21	0.07	-	0.30
INGARCH(2,1)	mean	1.64	0.13	0.51	0.23	-	0.37	0.15	0.58	0.11	-
	sd	0.44	0.08	0.07	0.08	-	0.12	0.11	0.07	0.09	-
INGARCH(2,2)	mean	1.69	0.10	0.51	0.20	0.05	0.42	-0.01	0.59	0.13	0.09
	sd	0.41	0.08	0.07	0.09	0.04	0.17	0.30	0.07	0.10	0.20

Table 2.7: Estimated Posterior Means and standard deviations(sd) for the competing models.

Finally, we have again calculated $p(\lambda_{T+1} \mid y)$ for the next unseen observation, weighted at its posterior model probability, depicted in Figure (2.4). As the sample size becomes large, the probability where we generate the data converge to 1. This is evident that our prior distributions are appropriate for parameters. As we can see from the first implementation the most visited model is the $INGARCH(1,1)$ model instead of the simplest $INGARCH(0,1)$ model. On the contrary at the second application the greatest probability is gotten in $INGARCH(2,1)$ model that is the expected considering the results presented at table 2.9. Larger data values

Model	linear					log-linear				
Parameters	d_1	α_1	b_1	α_2	b_2	d_1	α_1	b_1	α_2	b_2
INGARCH(0,1)	0.037	-	-0.099	-	-	0.856	-	-1.101	-	-
INGARCH(1,1)	0.592	-0.990	1.318	-	-	-0.241	-0.987	1.480	-	-
INGARCH(1,2)	-0.057	0.544	-0.335	-	-0.962	1.722	1.409	-0.765	-	-0.487
INGARCH(2,1)	-0.373	-1.256	-1.607	1.749	-	-0.909	0.592	-0.839	0.147	-
INGARCH(2,2)	-0.823	0.431	0.745	-0.719	-0.294	1.235	0.316	-1.543	-1.478	0.746

Table 2.8: Geweke's convergence z-scores for the parameters of five linear and five log-linear INGARCH models

Model	Posterior probability	Bayes Factor
INGARCH(2,1)	0.5595	576.0052
INGARCH(1,1)	0.2784	286.5797
LINGARCH(1,1)	0.0515	52.9751
INGARCH(2,2)	0.0339	34.8693
LINGARCH(1,2)	0.0222	22.8036
LINGARCH(0,1)	0.0177	18.2094
INGARCH(1,2)	0.0170	17.5422
LINGARCH(2,1)	0.0125	12.9080
LINGARCH(2,2)	0.0064	6.5652
INGARCH(0,1)	0.0009	1.0000

Table 2.9: Posterior probabilities and Bayes factor for the competing models

generally provide more information and reduce the uncertainty associated with parameter estimation, leading to greater precision. As a result, the posterior model probabilities become more precise and reliable, reflecting the increased confidence in the model's fit to the data.

In this section a bayesian model selection method has been discussed for the selection of the appropriate model among linear and log-linear INGARCH models. The more complicated method, the more flexible to use. Reversible Jump MCMC is difficult to construct assuming that appropriate proposal distributions must be defined for the variables of the new model. In contrary while the other two methods proposed by Carlin and Chib (1995) and Dellaportas et al. (2002) have a tolerant computational cost, they are less complicated at construction. Based on two criteria for model consistency, we observed that the average of posterior model probability is closed to one for the model that data are generated without any previous reference to this. Moreover to our simulation study two models different from INGARCH type have been also considered in order to examine algorithm's efficacy.

Two datasets are taken into account to examine which of the proposed five linear

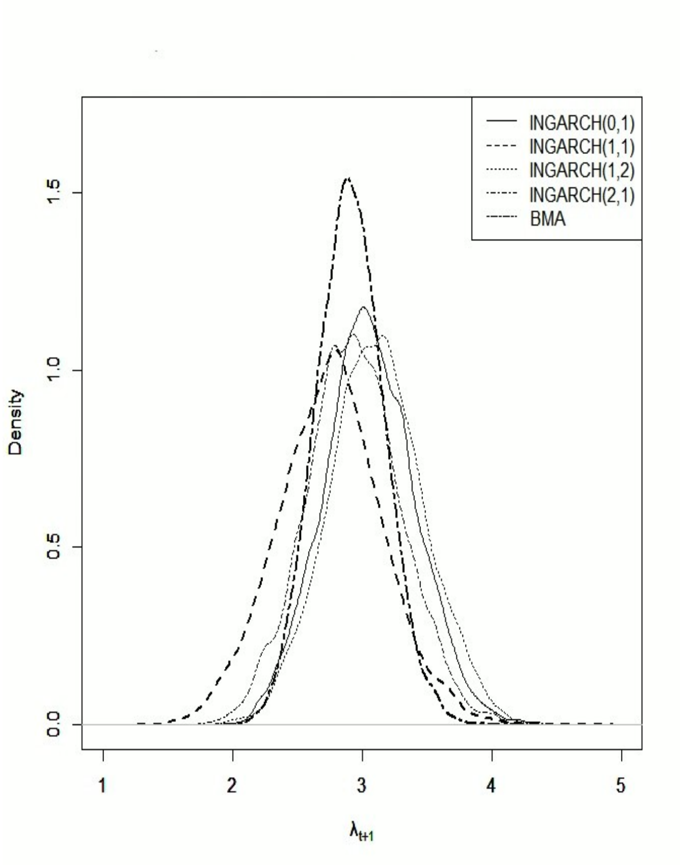


Figure 2.3: Posterior density for λ_{T+1} for polio data

and five log-linear INGARCH models better fits the data. In both cases linear INGARCH models have been selected, indicating that less complicated models are preferable for studies. Finally Bayesian Model Averaging procedure has been also applied to get densities of predictive volatilities. The method of Carlin and Chib (1995) provides flexibility with the construction of parameter space and the way that method decide to jump from one model to another. For those reasons the above method is very useful both in cases of nested and non-nested models and it can be generalized when we have processes of different types such as homoskedastic and heteroskedastic.

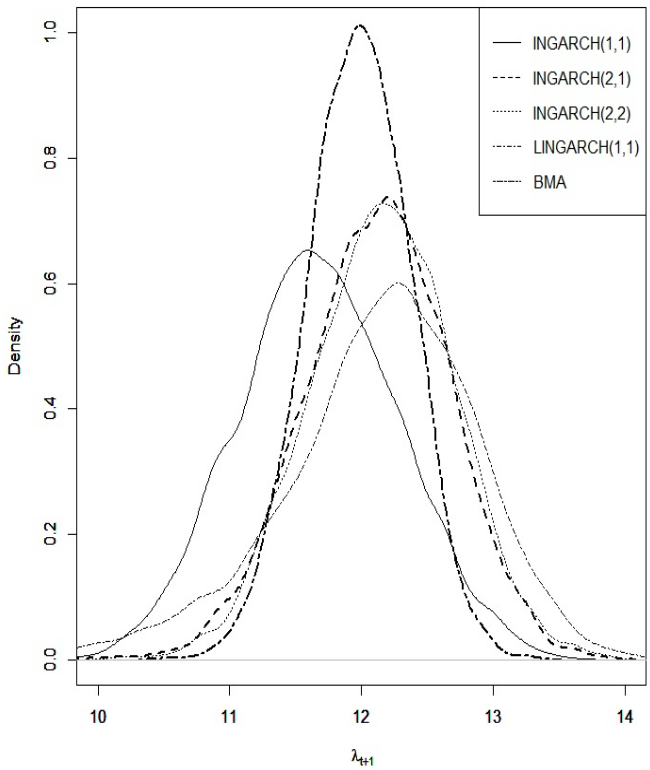


Figure 2.4: Posterior density for λ_{T+1} for Campylobacterosis data

Chapter 3

A novel multivariate Sarmanov distribution

3.1 Sarmanov distribution & FGM copula

The study of copulas in the domain of statistics plays an important role for the construction of multivariate distributions. Bivariate distributions could be a starting point for the generalization in more than three dimensions. Considering different structures, copula is a way for the calculation and modeling of the dependence. Sklar in 1959 proposed a theorem which plays an important role between multivariate distributions and their univariate margins. Considering inverse transformation method:

Theorem 1. *Let H be a joint distribution function with margins F,G. Then there exist a copula C such for all $/x_1, x_2$ in $\mathbb{R}$*

$$H(X_1, X_2) = C(F(X_1), G(X_2); \omega). \tag{3.1}$$

If F,G are continuous, then C is unique otherwise C is uniquely determined on $RangeF \times RangeG$. Conversely if C is a copula and F,G are distribution functions, then the function H defined by (3.1) is a joint distribution function with margins F,G.

The main assumption is that the marginals must be monotone functions. Some further properties for copulas proposed by Nelsen (2007):

1. For every $u, v \in [0,1]$ $C(u,0) = C(0,v) = 0$ and $C(u,1) = u, C(1,v) = v$

2. For every $u_1, u_2, v_1, v_2 \in [0,1]$ such that $u_1 \leq u_2$ $v_1 \leq v_2$

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0.$$

The first condition ensures that U, V are uniform marginals and the second referred as rectangular inequality $P(u_1 \leq U \leq u_2, v_1 \leq V \leq v_2) \geq 0$. Based on the above theorem there are many studies provided bivariate and multivariate parametric continuous and discrete distributions. Furthermore copulas is a way to study scale-free measures of dependence. There is flourishing literature about different named structures of copulas. Archimedean copulas firstly appeared in study of probabilistic metric spaces Schweizer (1991). Methods for constructions of copulas and tail dependence presented on a review of Nelsen (2007). There are numerous studies for bivariate distributions with some restrictive assumptions such as the range of dependence or the form of the cdf. At many studies alternative constructions of copulas have been proposed in order to provide a wide range at values of parameters of dependence. Furthermode those values depend on copulas function each time. The bounds for dependence based on Fréchet-Hoeffding inequality

$$max\{u + v - 1, 0\} \leq C(u, v) \leq min\{u, v\}.$$

The left side of inequality proves negative dependence and the right strong positive dependence. Cambanis (1977) studied bounds depending on the above inequality for a specified family of copula.

Since we have a form of bivariate distribution via a copula, it will be reasonable to measure dependence and association between the variables. Moreover the opportunity to capture tail dependence is offered via copulas. Two measures of dependence namely Spearman's ρ and Kendall's tau also used from many researchers for the above reason. In each case we examine the probability of concordance minus the probability of discordance for the pairs (x_i, y_i) and (x_j, y_j), where those two pairs are concordant when $(X_i - X_j)(Y_i - Y_j) \geq 0$ and discordant when $(X_i - X_j)(Y_i - Y_j), \leq 0$. Under the assumption that these are two vectors of independent variables (X_1, Y_1) and (X_2, Y_2) with common margins F and G, Kendall's tau defined as

$$\tau_C = 4 \int \int_{I^2} C(u, v) dC(u, v) - 1.$$

For Spearman's correlation examination of concordance and discordance between independent pairs of variables $(X_i, Y_i), i = 1, 2, 3$ defined as

$$\begin{aligned} \rho_C &= 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]) \\ &= 12 \int_0^1 \int_0^1 uv dC(u, v) - 3, \end{aligned} \tag{3.2}$$

where X_2, Y_3 are independent with marginals F and G respectively. Pearson correlation is often used as linear dependence measure and for this reason it is not appropriate for copula dependence measure. Although the above two measures of association take different values, depending on the family of copula each time, there is a relationship between them and measures of dependence.

Definition 7. *Let X and Y be two random variables. X and Y are positive quadrant dependent(PQD) if for all* $(x_1, x_2) \in \mathbb{R}^2$

$$P[X_1 \leq x_1, X_2 \leq x_2] \geq P[X_1 \leq x_1]P[X_2 \leq x_2],$$

or equivalently

$$P[X_1 > x_1, X_2 > x_2] \geq P[X_1 > x_1]P[X_2 > x_2].$$

One important consequence for the above measures of association for continuous positively quadrant dependent variables is the following theorem

Theorem 2. *Let X and Y are two continuous random variables with joint distribution function H and marginals F and G respectively and copula C. If X and Y are PQD, then*

$$3\tau_C \geq \rho_C \geq 0.$$

Furthermore with copulas the tail monotonicity is deterministic for positive quadrant dependence.

Definition 8. *Let X and Y be two random variables,*

1. *X_2 is left tail decreasing in X_1 if $P[X_2 \leq x_2, X_1 \leq x_1]$ is a non-increasing function of x_1 for all x_2.*

2. *X_1 is left tail decreasing in X_2 if $P[X_1 \leq x_1, X_2 \leq x_2]$ is a non-increasing function of x_2 for all x_1.*

3. *X_2 is right tail increasing in X_1 if $P[X_2 > x_2, X_1 > x_1]$ is a non-decreasing function of x_1 for all x_2.*

4. *X_1 is right tail increasing in X_2 if$P[X_1 > x_1, X_2 > x_2]$ is a non-decreasing function of x_2 for all x_1.*

Chapter 5 in Nelsen (2007) provides more details about tail monotonicity of continuous and discrete random variables. There are numerous families of copulas constructed with different methods. Some of those methods are geometric methods, algebraic methods and the method of inversion. Nelsen (2007) mentioned as geometric methods ordinal sums and singular copulas with prescribed support. We briefly discuss about those methods.

Definition 9. *Assuming a finite index set K and J_K a partition of $\mathbb{I}$ with $J_K \in [a_k, b_k]$ for $k \in K$ and $\{C_k\}_{k\in K}$ a collection of copulas the ordinal sum defined by:*

$$C(u,v) = \begin{cases} a_k + (b_k - a_k)C\left(\frac{u-a_k}{b_k-a_k}, \frac{v-a_k}{b_k-a_k}\right), & \text{if } (u,v) \in J_k^2. \\ M(u,v), & \text{if } (u,v) \notin J_k^2. \end{cases} \tag{3.3}$$

Furthermore three examples have been also proposed in order to construct singular copulas with support consists of line segments or arcs of circles. As algebraic methods Nelsen (2007) referred Plackett distributions and Ali-Mikhail-Haq distributions. The idea for the construction of the first family based on contigency tables. Considering two continuous random variables X, Y with marginals F, G respectively and their joint distribution H, column variables correspond to the events "$X \leq x$" and "$X > x$" similarly for the row variable Y. Thus we obtain

	$X_2 \leq x_2$	$X_2 > x_2$
$X_1 \leq x_1$	$H(x_1,x_2)$	$F(x_1) - H(x_1,x_2)$
$X_1 > x_1$	$G(x_2) - H(x_1,x_2)$	$1 - F(x_1) - G(x_2) + H(x_1,x_2)$

$$\theta = \frac{H(x_1,x_2)\left[1 - F(x_1) - G(x_2) + H(x_1,x_2)\right]}{\left[F(x_1) - H(x_1,x_2)\right]\left[G(x_2) - H(x_1,x_2)\right]} \tag{3.4}$$

and using probability transforms $F(x_1) = u$, $G(x_2) = v$ and Sklar's theorem

$$\theta = \frac{C(u,v)\left[1 - u - v + C(u,v)\right]}{\left[u - C(u,v)\right]\left[v - C(u,v)\right]}.$$

Second family of distributions constructed based on Poisson process. In the univariate case the odds for survival defined as: $\frac{\bar{F}(x)}{F(x_1)} = \frac{1-F(x_1)}{F(x_1)}$. Analogously in bivariate case survival odds ratio given by:

$$\frac{1 - H(x_1,x_2)}{H(x_1,x_2)}.$$

For example Gumbel's bivariate logistic distribution defined considering two exponential distributions and the above function. In this case the odds ratio given as the sum of the survivals odds ratio for variables X and Y. Furthermore considering two independent random variables X, Y and expressing each one as $F(x) = \left(1 + \left[\frac{(1-F(x_1))}{F(x_1)}\right]\right)^{-1}$ then

$$\frac{1 - H(x_1, x_2)}{H(x_1, x_2)} = \frac{1 - F(x_1)}{F(x_1)} + \frac{1 - G(x_2)}{G(x_2)} + \frac{1 - F(x_1)}{F(x_2)} \frac{1 - G(x_2)}{G(x_2)}. \tag{3.5}$$

Ali-Mikhail-Haq based on (3.5) proposed a bivariate distribution given by:

$$\frac{1 - H(x_1, x_2)}{H(x_1, x_2)} = \frac{1 - F(x_1)}{F(x_1)} + \frac{1 - G(x_2)}{G(x_2)} + (1 - \theta) \frac{1 - F(x_1)}{F(x_1)} \frac{1 - G(x_2)}{G(x_2)},$$

where θ lies in the interval $[-1, 1]$. The third and mostly used method is this of inversion assuming Sklar's theorem and considering two marginal distributions F,G then $C(u, v) = H(F^{(-1)}(u), G^{(-1)}(v))$.

3.2 Families of copulas

In this section we provide a brief overview of the most commonly used families of parametric copulas. Based on Sklar's theorem one can construct bivariate one parameter or two parameter families of copulas. The first prevalent family is the family of Archimedean copulas with a construction based on Laplace transform appeared firstly in study of probabilistic metric spaces Schweizer (1991). A CDF F then it could be expressed via the laplace transform of an other one CDF G as

$$F(X) = \int_0^\infty G^a(x) dM(a) = \phi(-\log G(x)).$$

Respectively in bivariate case

$$C(F(x_1), G(x_2)) = \phi(-\log G_1 - \log G_2) = \phi(\phi^{-1}(F_1) + \phi^{-1}(F_2)),$$

or after transformation

$$C(u, v) = \phi\left(\phi^{-1}(u) + \phi^{-1}(v)\right),$$

where $\phi : \mathbb{I} \to [0, 1]$ are continuous, strictly decreasing and convex functions with $\phi(1) = 0$ This family has an appeal at many different scientific fields because it does not have many difficulties. Obviously this copula has a closed form if both ϕ and ϕ^{-1} have closed forms and parameter of dependence appears in Laplace transform. Numerous copulas have been constructed with different Laplace transforms as generators. For example Gumbel (1960b) proposed a new copula using as generator

Laplace transform of Exponential distribution where parameter of dependence lies in the interval $[-1, 1]$, namely the cfd is given by:

$$C(u, v) = e^{\left(-\left((-\log u)^{\theta}+(-\log v)^{\theta}\right)^{1/\theta}\right)}, \quad \phi_{\theta}(x) = (-\log x)^{\theta}.$$

Furthermore another discussed family is this of Frank (1979) which first appeared in a non-statistical context. The reason is that this family is the only one which satisfies radial symmetry. Details and statistical properties have been also provided by Nelsen (1986);Genest (1987). The Frank copula has cdf given by:

$$C(u, v) = -\frac{1}{\theta}\log\left(1 + \frac{\left(e^{-\theta u} - 1\right)\left(e^{-\theta v} - 1\right)}{e^{-\theta} - 1}\right), \quad \phi_{\theta}(x) = -\log\frac{e^{-\theta x} - 1}{e^{-\theta} - 1}.$$

At Nelsen (2007) a variety of families of Archimedean copulas is concentrated with a little discussion about some of them and many scatterplots are also provided in order to highlight the importance of parameter θ.

3.3 Alternative construction of FGM copula

Another family of copulas with discernible effects is Sarmanov family of bivariate distributions proposed firstly by Sarmanov (1966). Assume two univariate probability mass functions (p.m.f.) or probability density functions (p.d.f.) with supports defined on $A_1 \subseteq \mathbb{R}$, $A_2 \subseteq \mathbb{R}$ respectively. Additionally assume $q_i(x_i), i = 1, 2$ two bounded non constant functions where

$$\textstyle\sum_{x_i=-\infty}^{\infty} q_i(x_i) f_i(x_i) = 0 \text{ or } \int_{x_i=-\infty}^{\infty} q_i(x_i) f_i(x_i) = 0.$$

The joint probability mass function (p.m.f.)(or probability density function (p.d.f.)) for the discrete and the continuous case is:

$$f(x_1, x_2) = f_1(x_1) f_2(x_2)[1 + \omega q_1(x_1) q_2(x_2)], \tag{3.6}$$

where the multiplicative factor $[1 + \omega q_1(x_1) q_2(x_2)]$ indicates the deviance of two variables X_1, X_2 from independence and parameter ω is a real number that satisfies the condition

$$1 + \omega q_1(x_1) q_2(x_2) \geq 0. \tag{3.7}$$

In the case where $\omega = 0$, the variables X_1, X_2 are independent.

Values of parameter ω are deterministic for values of correlation coefficient and depend on approximate values of functions $q_i(x_i)$, $i = 1, 2$. Ting Lee (1996) proposed countless constructions of mixing functions $q_i(x_i)$ considering each time support of $f_i(x_i)$. When the support of $f_i(x_i)$ is contained in [0,1], then $q_i(x_i) =$

$x_i - \mu_i,\ i = 1,2$ where $\mu_i = \sum_{i=1}^{2}\sum_{z_i=0}^{1} z_i f_i(z_i)\left(\int_{z_i=0}^{1} z_i f_i(z_i)dz_i\right)$. The same construction could be applied when variable's support is any finite interval. In the case where $x_i \in \mathbb{R}$ then the proposed mixing functions are $q_i(x_i) = f_i(x_i) - \sum_{z_i=0}^{\infty} f_i^2(z_i)\left(\int_0^{\infty} f_i^2(t_i)dt_i\right)$ and in case where $x_i \in \mathbb{R}^+$ then $q_i(x_i) = exp(-x_i) - L_i(t)$ or $q_i(x_i) = 1 - F_i(x_i)$. Ting Lee (1996) expressed $q_i(x_i)$ by the function below where $|q_i(x_i)| < 1$ and variable's support is $[0,1]$:

$$q_i(x_i) = 1 - 2F_i(x_i), \quad i = 1,2 \tag{3.8}$$

Based on (3.6), (3.7),considering Sklar's theorem and inversion method Farlie, Gumbel and Morgerstern at the same period constructed the FGM family given by:

$$c(u,v) = 1 + \omega(1-2u)(1-2v),\ \ u,v \in [0,1], \tag{3.9}$$

where $F(x_1) = u$ and $F(x_2) = v$ The overall constraint on ω is given by:

$$c(u,v) \geq 0, \tag{3.10}$$

or equaivalently:

$$1 + \omega\left(1-2u\right)\left(1-2v\right) \geq 0,$$

where finally $\omega \in [-0.33, 0.33]$ Product moment and correlation coefficient are given by:

$$E(X_1X_2) = \mu_1\mu_2 + \omega v_1 v_2, \tag{3.11a}$$

$$\rho = \frac{\omega v_1 v_2}{\sigma_1\sigma_2}, \tag{3.11b}$$

where

$$v_i = E\left(f_i(x_i)\phi_i(x_i)\right),\ \ i = 1,2. \tag{3.12}$$

Based on (3.11b), it is prevalent that values of correlation coefficient depend on values of ω. Considering (3.11b) and (3.12) values of ω lie in the interval $[-1,1]$. Considering FGM, Farlie (1960) extended the idea of Morgerstern and Gumbel by proposing the following

$$C(u,v) = uv\left(1 + \omega A(u)B(v)\right),$$

where $A(u), B(v) \to 0$ as $u,v \to 1$ According to the above general family, Farlie (1960) and then Huang and Kotz (1999) proposed alternative construction embedding mixtures of powers, namely

$$\begin{aligned}
C(u,v) &= uv\left(1+\omega\left(1-u\right)^{\gamma}\left(1-v\right)^{\gamma}\right), \;\; \gamma \geq 1 \\
C(u,v) &= uv\left(1+\omega\left(1-u^{\gamma}\right)\left(1-v^{\gamma}\right)\right), \;\; \gamma \geq 1/2 \\
C(u,v) &= uv\left(1+\omega(1-u)^{\gamma_1}\left(1-v\right)^{\gamma_2}\right), \;\; \gamma_1 \neq \gamma_2 \\
C(u,v) &= uv\left(1+\omega(1-u^{\gamma})(1-v^{\gamma})\right), \;\; \gamma \geq 0
\end{aligned}$$

The first two cases are focused in symmetry. For the third there are unclear properties because it is about a future work. The fourth has been studied by Huang and Kotz (1999) considering a division of the unit square into four quadrants to obtain lower and upper bound for parameter of association ω. In this case the inequality that provide bounds for parameter ω is given by:

$$1+\omega\left(1-\left(1+\gamma\right)u^{\gamma}\right)\left(1-\left(1+\gamma\right)v^{\gamma}\right),$$

where finally the range for ω is $[-(max\{1,p\}^{-2}), p^{-1}]$

Bekrizadeh et al. (2012) probed the workings of amendment of correlation values constructing an alternative model as

$$C(u,v) = uv\left[1+\theta(1-u^{a})(1-v^{a})\right]^{n},$$

where $a>0$ and $n \in \mathbb{N}$. Taking advantage of the fact that $c(u,v) \geq 0$, the interval for ω is defined as $-min\{1, \frac{1}{na^2}\} \leq \omega \leq \frac{1}{na}$. The above studies have been concentrated(focused) in the case where a bivariate distribution function can be interpreted as a bivariate copula or in other words when u,v are uniformly distributed.

3.4 A proposed alternative FGM copula

In the whole study about association parameter's interval, numerous applications have been mentioned. Gumbel (1960a) proposed a new bivariate Exponential distribution based on the first construction of FGM copula and Bekrizadeh et al. (2012) studied his proposed FGM copula considering as marginals Beta and Weibull distributions. Assuming the properties studied by Huang and Kotz (1999), Ting Lee (1996) proposed Gamma, Beta, Exponential and Normal bivariate distributions considering each time different construction for functions encompassed in multiplicative factor. The main problems are: (a) there is not always a closed form for the bivariate distribution leading to infeasibility of studying other properties of copulas e.g. parameters estimation and (b) restrictive correlation coefficients values are provided.

At the case of discrete variables there are many studies focusing not only on the interval of parameter ω but also on methods for parameters' estimation. Ting Lee (1996) proposed two distributions considering Binomial and Poisson marginals respectively. Lakshminarayana et al. (1999) and after many years Cui and Zhu (2018) proposed bivariate Poisson distributions. The main drawback is the violation of Frechet-Hoeffding boundaries in order to get better values for dependence parameter range. Piperigou (2009) studied a Poisson and a Geometric bivariate distributions considering probability generating functions in order to increase values for ω and correlation coefficient ρ. For the construction of boundaries for dependence parameter Piperigou (2009) used the theory introduced by Cambanis (1977). Miravete (2009a) based on Ting Lee (1996) built two multivariate count data regression models considering as marginals Double Poisson distributions. Moreover Ting Lee (1996) proposed an extended Sarmanov distribution and Miravete (2009a) based on this extension proposed trivariate Double Poisson and Gamma distributions without studying theoretical properties. Lakshminarayana et al. (1999) provided different general construction for the mixing functions

$$q_i(x_i) = g_i(x_i) - E\left(g_i(x_i)\right), \;\; i = 1, 2,$$

and discussed about the vital problem of the choice of $g_i(x_i)$. Assuming that mixing functions must be bounded functions, damped polynomials have been proposed where as damped polynomials Lakshminarayana et al. (1999) considered any polynomial multiplied by a dampening function $w(x)$ that tends to zero faster than any polynomial tends to infinity as x increases. A prevalent option is e^{-x} where $g_i(x_i) = e^{-x_i}P(x_i)$. The idea of Lakshminarayana et al. (1999) seems to be similar to this of Ting Lee (1996) where under the assumption that $f_i(x_i)$ are defined in the interval $[0, \infty)$. Ting Lee (1996) proposed as mixing functions

$$q_i(x_i) = e^{-x_i} - LT_i(1), \;\; i = 1, 2, \;\; \forall\, x_i \geq 0. \tag{3.13}$$

At this study $g_i(x_i) = e^{-x_i}$ and $LT_i(t) = \sum_{x_i=0}^{\infty} e^{-tx_i} f_i(x_i)$ is the Laplace transform of f_i for $i = 1, 2$. Integrating $x_1 x_2$ with respect to (3.6) the product moment and correlation coefficient are given by (3.11a; 3.11b) where

$$v_i = -LT_i^{'}(1) - LT_i(1)\mu_i, i = 1, 2. \tag{3.14}$$

Under the restriction (3.7) $\omega \in [b_1, b_2]$, where

$$b_1 = \frac{-1}{max\left\{LT_1(1)LT_2(1), \left(1 - LT_1(1)\right)\left(1 - LT_2(1)\right)\right\}} \tag{3.15}$$

and

$$b_2 = \frac{1}{max\{LT_1(1)(1 - LT_2(1)), (1 - LT_1(1))(LT_2(1))\}} \tag{3.16}$$

and based on (3.6) product moment and correlation coefficient are defined as:

$$E(X_1, X_2) = \mu_1\mu_2 + \omega v_1 v_2, \tag{3.17}$$

$$\rho = \frac{\omega v_1 v_2}{\sigma_1 \sigma_2}, \tag{3.18}$$

where $v_i,\ i = 1,2$ given by (3.14). The drawback with all of those constructions is that the range of correlation is too restrictive. Vernic (2020) and a few years later Piancastelli et al. (2023), based on the previous definition and on the idea that $q_i(x_i)$ could be expressed as $u_i(x_i) - E[u_i(X_i)]$, proposed a new function $u_i(x_i)$ defined as $u_i(x_i) = e^{-ax_i}$ where $a \neq 1$. The proposed function is called Exponential kernel yielding a greater range for correlation coefficient boundaries. In our study based on Sarmanov distribution given by (3.6), on the work of Ting Lee (1996) and generalized the idea of Vernic (2020), we focus firstly on bivariate case, defining new mixing functions as alternative approaches of Laplace transforms given by:

$$q_i(x_i) = c_i^{-k_i x_i} - \textstyle\sum_{x_i=0}^{\infty} c_i^{-k_i x_i} f(x_i)\ , \quad i = 1,2$$

where

$$c_i > 1, i = 1,2 \text{ and } k_i > 0, i = 1,2.$$

Based on the above structure, functions $q_i(x_i), i = 1,2$ are bounded, non constant functions. In discrete case the definition of bivariate distribution has a drawback at the use of finite sums. For the above reason many studies have been done based on Sarmanov distribution and it is alternative construction. Moreover the crucial problem is tracked for the variables with support $\mathbb{R}^+$ because there is a restricted number of bounded functions which allow flexibility at values of correlation coefficient. Considering now two discrete variables and by definition of joint distribution a new construction has been developed where $q_i(x_i) = f_i(x_i) + 2F_x(x_i) - 1$:

$$\begin{aligned}
p(x_1,x_2) &= P(X_1=x_1,X_2=x_2)=P(X_1\leq x_1X_2=x_2)-P(X_1<x_1,X_2=x_2)\\
&= P(X_1\leq x_1,X_2\leq x_2)-P(X_1\leq x_1,X_2\leq x_2-1)\\
&- P(X_1\leq x_1-1,X_2\leq x_2)+P(X_1\leq x_1-1,X_2\leq x_2-1)\\
&= P(x_1,x_2)-P(x_1,x_2-1)-P(x_1-1,x_2)+P(x_1-1,x_2-1)\\
&= P(x_1)P(x_2)[1+\omega\left(1-P(x_1)\right)\left(1-P(x_2)\right)]\\
&- P(x_1)P(x_2-1)[1+\omega\left(1-P(x_1)\right)\left(1-P(x_2-1)\right)]\\
&- P(x_1-1)P(x_2)[1+\omega\left(1-P(x_1-1)\right)\left(1-P(x_2)\right)]\\
&+ P(x_1-1)P(x_2-1)[1+\omega\left(1-P(x_1-1)\right)\left(1-P(x_2-1)\right)]\\
&= \left(P(x_1)-P(x_1-1)\right)\left(P(x_2)-P(x_2-1)\right)\\
&\quad [\left(P(x_1)+P(x_1-1)-1\right)\left(P(x_2)+P(x_2-1)-1\right)], \qquad (3.19)
\end{aligned}$$

and considering that

$$\begin{aligned}
P(x_1)+P(x_1-1)-1 &= P(x_1)-P(x_1-1)+2P(x_1-1)-1\\
&= -[P(x_1)-P(x_1-1)]+2P(x_1)-1\\
&= -p(x_1)+2P(x_1)-1,
\end{aligned}$$

we obtain finally another construction of multiplicative factor different from those that have been discussed previously. This formula provide an improvement at bounds of correlation coefficient. Barbiero (2019) studied values of correlation coefficient based on the above structure for a discrete Weibull distribution.

$$p(x_1,x_2)=\prod_{i=1}^{2}[P(x_i)-P(x_i-1)]\left(1+\omega\prod_{i=1}^{2}[p(x_i)-2P(x_i)+1]\right).$$

For the construction of bivariate Poisson distribution two general closed forms have been also proposed from Kocherlakota and Kocherlakota (2001) and Lakshminarayana et al. (1999) respectively. The focus is concentrated on the flexibility of the closed form of bivariate distribution and on the construction which allows high values for correlation. At the first general construction, researchers tried to define a bivariate model paying attention on methods for estimation caused of the non-closed form in conjunction with the second where different functions have been proposed to increase correlation coefficient bounds. Kocherlakota and Kocherlakota (2001) studied a bivariate Poisson distribution based on trivariate reduction method. Assuming that variables X_i follow independent Poisson distributions $P(\lambda_i)$ with parameters λ_i, $i = 1, 2, 3$ and the variables $X = X_1 + X_3$, $Y = X_2 + X_3$ follow jointly a bivariate distribution given by

$$\begin{aligned} p(x = r, y = s) &= \sum_{k=0}^{min(r,s)} f_1(r-k) f_2(s-k) f_3(k) \\ &= e^{-\sum_{i=1}^{3}\lambda_i} \sum_{k=0}^{min(r,s)} \frac{\lambda_1^r \lambda_2^s \lambda_3^k}{\left(\prod_{i=1}^{2}\lambda_i\right)^k (r-k)!(s-k)!k!}, \\ & r, s = 0, 1, \ldots, \quad \lambda_i > 0, i = 1, 2, 3. \end{aligned}$$

Marginally each one of those variables follows a Poisson distribution with mean $E(X) = \lambda_1 + \lambda_3$ and $E(Y) = \lambda_2 + \lambda_3$. Parameter λ_3 is appropriate for dependence between two variables X, Y while $cov(X, Y) = \lambda_3$. In case where $\lambda_3 = 0$ two variables are independent and the bivariate Poisson distribution is the product of two independent Poisson distributions. Vernic (1997) studied properties for the above bivariate Poisson distribution and parameters' estimation using method of moments while Kocherlakota and Kocherlakota (2001) proposed method of maximum likelihood for parameters' estimation. Karlis and Ntzoufras (2003) proposed an EM algorithm considering the bivariate Poisson model proposed by Kocherlakota and Kocherlakota (2001) and assuming log link regression functions for the mean assuming

$$\log \lambda_{ik}^{*} = w_{ik}\beta_k, \quad k = 1, 2, 3, \quad i = 1, ..., n$$

where w_{ik} denotes the vector of explanatory variables and β_k is the vector of regression coefficients.

Lakshminarayana et al. (1999) studied a bivariate Poisson model based on Sarmanov distribution as we mentioned at the beginning of this chapter. Correlation coefficient bounds in this study violated the Frechet-Hoeffding bounds because additional restriction is necessary. Similar to the above model another model was also proposed by Cui and Zhu (2018). The main difference with the model construct by Lakshminarayana et al. (1999) is the non deterministic and inappropriate way of correlation coefficient bounds' definition. While Lakshminarayana et al. (1999) proposed that:

$$|\rho| \leq \frac{\sqrt{\lambda_1\lambda_2} e^{-\left(1-e^{-1}\right)(\lambda_1+\lambda_2)} \left(1-e^{-1}\right)^2}{\left(1-e^{-(1-e^{-1})\lambda_1}\right)\left(1-e^{-(1-e^{-1})\lambda_2}\right)}.$$

Cui and Zhu (2018) restrict ρ as:

$$|\rho| \leq \min\left\{\frac{\sqrt{\lambda_1\lambda_2}e^{-\left(1-e^{-1}\right)(\lambda_1+\lambda_2)}\left(1-e^{-1}\right)^2}{\left(1-e^{-(1-e^{-1})\lambda_1}\right)\left(1-e^{-(1-e^{-1})\lambda_2}\right)}, 1\right\}.$$

In our study based on (3.4) we define a new bivariate Poisson distribution as:

$$f(x_1, x_2) = \prod_{i=1}^{2} \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!}\left[1 + \omega \prod_{i=1}^{2} q_i(x_i)\right],$$

where

$$q_i(x_i) = c_i^{-k_i x_i} - e^{\lambda_i(c_i^{-k_i x_i} - 1)}, \;\; c_i \geq 0, k_i \geq 0.$$

Moreover considering (3.19) ,we can conclude to the same bivariate Poisson distribution proposed above.

$$\begin{aligned}
F(x_1, x_2) &= \prod_{i=1}^{2}\left(e^{-\lambda_i}\sum_{t=0}^{X_i}\frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i(c_i^{-k_i}-1)}F(x_i)\right),\\
F(x_1 - 1, x_2) &= \left(e^{-\lambda_1}\sum_{t=0}^{X_1-1}\frac{(\lambda_1 c_1^{-k_1})^t}{t!} - e^{\lambda_1(c_1^{-k_1}-1)}F(x_1 - 1)\right)\\
&\quad \left(e^{-\lambda_2}\sum_{t=0}^{X_2}\frac{(\lambda_2 c_2^{-k_2})^t}{t!} - e^{\lambda_2(c_2^{-k_2}-1)}F(x_2)\right),\\
F(x_1, x_2 - 1) &= \left(e^{-\lambda_1}\sum_{t=0}^{X_1}\frac{(\lambda_1 c_1^{-k_1})^t}{t!} - e^{\lambda_1(c_1^{-k_1}-1)}F(x_1)\right)\\
&\quad \left(e^{-\lambda_2}\sum_{t=0}^{x_2-1}\frac{(\lambda_2 c_2^{-k_2})^t}{t!} - e^{\lambda_2(c_2^{-k_2}-1)}F(x_2 - 1)\right),\\
F(x_1 - 1, x_2 - 1) &= \prod_{i=1}^{2}\left(e^{-\lambda_i}\sum_{t=0}^{X_i-1}\frac{(\lambda_i c_i^{-k_i})^t}{t!} - e^{\lambda_i(c_i^{-k_i}-1)}F(x_i - 1)\right).
\end{aligned}$$

Assuming that $f(x_1, x_2) = F(x_1, x_2) - F(x_1 - 1, x_2) - F(x_1, x_2 - 1) + F(x_1 -$

$1, x_2 - 1)$ and substituting the equations above we have :

$$
\begin{aligned}
f(x_1, x_2) = \prod_{i=1}^{2} \left(F\left(x_i\right) - F\left(x_i - 1\right)\right) e^{\sum_{i=1}^{2} \lambda_i (c_i^{-k_i} - 1)} \left(F\left(x_2\right) - F\left(x_2 - 1\right)\right) \\
- e^{\lambda_1 \left(c_1^{-k_1} - 1\right) - \lambda_2} \left(F\left(x_1\right) - F\left(x_1 - 1\right)\right) \sum_{t=0}^{x_2} \frac{\left(\lambda_2 c_2^{-k_2}\right)^t}{t!} \\
+ e^{\lambda_1 \left(c_1^{-k_1} - 1\right) - \lambda_2} \left(F\left(x_1\right) - F\left(x_1 - 1\right)\right) \sum_{t=0}^{x_2 - 1} \frac{\left(\lambda_2 c_2^{-k_2}\right)^t}{t!} \\
- e^{-\lambda_1 + \lambda_2 \left(c_2^{-k_2} - 1\right)} \left(\sum_{t=0}^{X_1} \frac{\left(\lambda_1 c_1^{-k_1}\right)^t}{t!} - \sum_{t=0}^{X_1 - 1} \frac{\left(\lambda_1 c_1^{-k_1}\right)^t}{t!} \right) \left(F\left(x_2\right) - F\left(x_2 - 1\right)\right) \\
+ e^{-(\lambda_1 + \lambda_2)} \prod_{j=1}^{2} \left(\sum_{t=0}^{X_j} \frac{\left(\lambda_j c_j^{-k_j}\right)^t}{t!} - \sum_{t=0}^{X_j - 1} \frac{\left(\lambda_j c_j^{-k_j}\right)^t}{t!} \right).
\end{aligned}
$$

Taking into account that in discrete case $f(x_i) = F(x_i) - F(x_i - 1), \; i = 1, 2$ and

$$
f(x_i) = \left(\sum_{t=0}^{X_i} \frac{\left(\lambda_i c_i^{-k_i}\right)^t}{t!} - \sum_{t=0}^{X_i - 1} \frac{\lambda_i c_i^{-k_i}}{t!} \right), \quad i = 1, 2
$$

for variables $X_i, \; i = 1, 2$ respectively we take that

$$
p(x_1, x_2) = \prod_{i=1}^{2} f(x_i) \left(1 + \omega \prod_{i=1}^{2} \left(c_i^{-k_i x_i} - e^{\lambda_i (c_i^{-k_i} - 1)} \right) \right).
$$

Another way is the variable transformation to change the domain and to bypass the problem especially at the case of $[0, \infty)$. Miravete (2009b) based on the model defined by Lakshminarayana et al. (1999) constructed two bivariate models the first when marginals are distributed according to a double Poisson distribution and the second one when marginals follow a Gamma distribution. At those models there is a closed form for the bivariate cdf. On the contrary in case on Poisson distribution this is not efficient cause of cdf's construction. For this reason a closed form for a copula is not always possible and methods for parameters estimation are more extensive.

Our new model is flexible not only in case of a discrete variable but also for continuous variables. Calculating cdf of a bivariate distribution and considering $g_i(x_i) = c_i^{-k_i x_i}$ in order to construct a new family of copulas with flexible correlation bounds depend each time on the choice of mixing functions, we have:

$$
C(u, v) = uv \left[1 + \omega \left(\sum_{t_1=0}^{F^{-1}(u)} c_1^{-k_1 t_1} - uLT_1(f(x_1)) \right) \left(\sum_{t_2=0}^{F^{-1}(v)} c_2^{-k_2 t_2} - vLT_2(f(x_2)) \right) \right].
$$

In general this family has the properties of a copula.

Considering structure for correlation coefficient and assuming that $q_i(x_i),\ i = 1,2$ are defined by the mixing functions given above we have that:

$$\begin{aligned} E(X_1X_2) &= \prod_{i=1}^{2}\lambda_i + \omega\prod_{i=1}^{2}\left[-\frac{1}{\log c_i}q^{'}(k_i) - \lambda_i e^{\lambda_i(c_i^{-k_i}-1)}\right] \\ &= \prod_{i=1}^{2}\lambda_i + \omega\prod_{i=1}^{2}\lambda_i e^{\lambda_i(c_i^{-k_i}-1)}(c_i^{-k_i}-1),\ \ i=1,2 \end{aligned} \tag{3.20}$$

Under the assumption that $q_i^{'}(k_i)$ are the derivatives of $\exp(\lambda_i(c_i^{-k_i})-1)$, i=1,2 with respect to $k_i,\ \ i = 1,2$ or

$$\frac{\partial}{\partial k_i}\sum_{x_i=0}^{\infty} f_i(x_i)c_i^{-k_i x_i} = \sum_{x_i=0}^{\infty} f_i(x_i)\frac{\partial}{\partial k_i}c_i^{-k_i x_i} = -\log c_i \sum_{x_i=0}^{\infty} x_i f_i(x_i)c_i^{-k_i x_i}.$$

Combining (3.20) into (3.18) correlation coefficient given by:

$$\rho = \omega\prod_{i=1}^{2}\sqrt{\lambda_i}e^{\lambda_i(c_i^{-k_i}-1)}\left(c_i^{-k_i}-1\right)\ \ ,i=1,2. \tag{3.21}$$

The above structure of (3.20) has been calculated based on (3.14). According to the above mixing functions new boundaries are calculated considering restriction given by (3.6) and calculating the limits as $x_i,\ \ i = 1,2$ approaches to 0 and ∞ and given by:

$$b_1 = \frac{-1}{max\left\{\left(1-e^{\lambda_1(c_1^{k_1}-1)}\right)\left(1-e^{\lambda_2(c_2^{-k_2}-1)}\right), e^{\lambda_1(c_1^{-k_1-1})}e^{\lambda_2(c_2^{-k_2-1})}\right\}}, \tag{3.22}$$

$$b_2 = \frac{1}{max\left\{e^{\lambda_1(c_1^{-k_1-1})}\left(1-e^{\lambda_2(c_2^{-k_2-1})}\right), e^{\lambda_2(c_2^{-k_2-1})}\left(1-e^{\lambda_1(c_1^{-k_1-1})}\right)\right\}}. \tag{3.23}$$

Based on the Cambanis (1977) this construction does not violate Frechet-Hoeffding inequality and it offers flexible boundaries for correlation coefficient. Firstly we examine ρ's values taking 3 possible values for $\lambda_i,\ \ i = 1,2$ and numerous values for $k_i,\ \ i = 1,2$. From the graphs below we observe that the optimal value for ρ is obtained when $\lambda_i = 0.5,\ \ i = 1,2$.

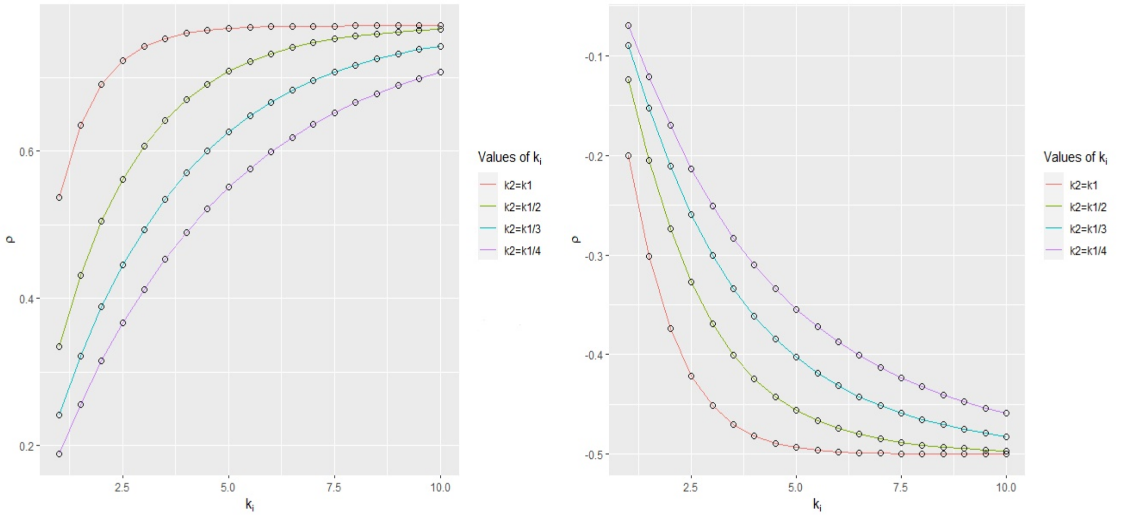


Figure 3.1: Values for correlation with equal values at both $\lambda_i, k_i, i = 1, 2$. Lower boundary(first graph) and upper boundary(second graph)

Then keeping the value for $\lambda_i,\ \ i = 1, 2$ and considering different combinations between two parameters k_1, k_2 boundaries for correlation coefficient are also presented. From those graphs and by using $c_i = exp(1),\ \ , i = 1, 2$ we observe that upper boundary for correlation is 0.77 and lower boundary is -0.5. Furthermore approximately when k_1, k_2 are equal to 10 correlation coefficient value converge to upper and lower boundaries respectively and it is closed to the same value when $k_1 = k_2$, $k_1 = k_2/2$ and $k_1, k_2 > 8$.

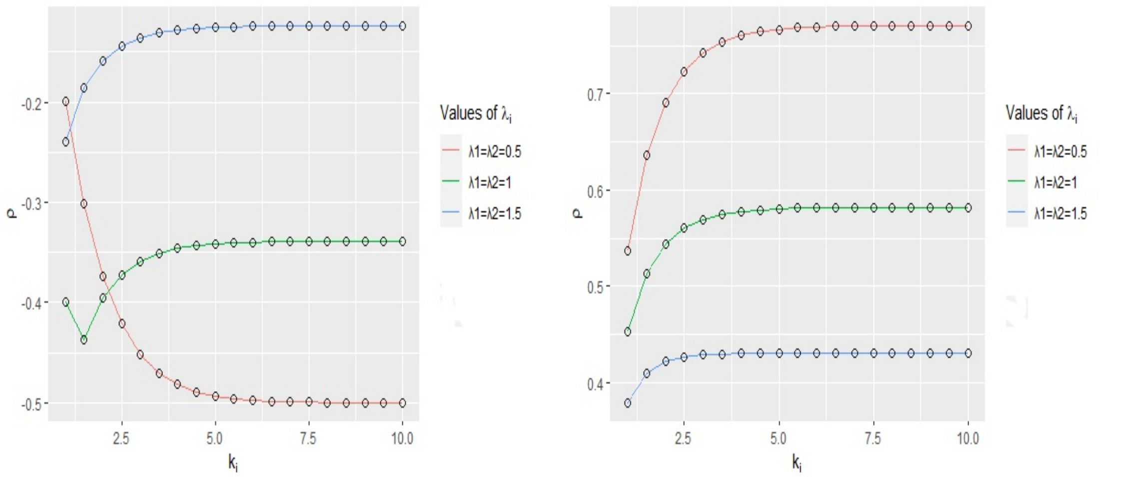


Figure 3.2: Values for correlation with equal values at both $\lambda_i, k_i, i = 1, 2$. Lower boundary(first graph) and upper boundary(second graph)

Finally a comparison among the models that have been also proposed with the new model is presented via graphs with the appropriate values for ρ lie in different interval each time based on values for parameters c_i, λ_i and $k_i,\ \ i = 1, 2,$

We obtain values for ρ for countless values of c_i, $i = 1, 2$. At the graphs below we present correlation coefficient values taking numerous values for parameters λ_i and for $k_1 = k_2 = 1$. Considering models proposed by Lakshminarayana et al. (1999), Ting Lee (1996) and only assuming conditions for inequality proposed by Cambanis (1977) the optimal values are given when $c_i = 3.5$.

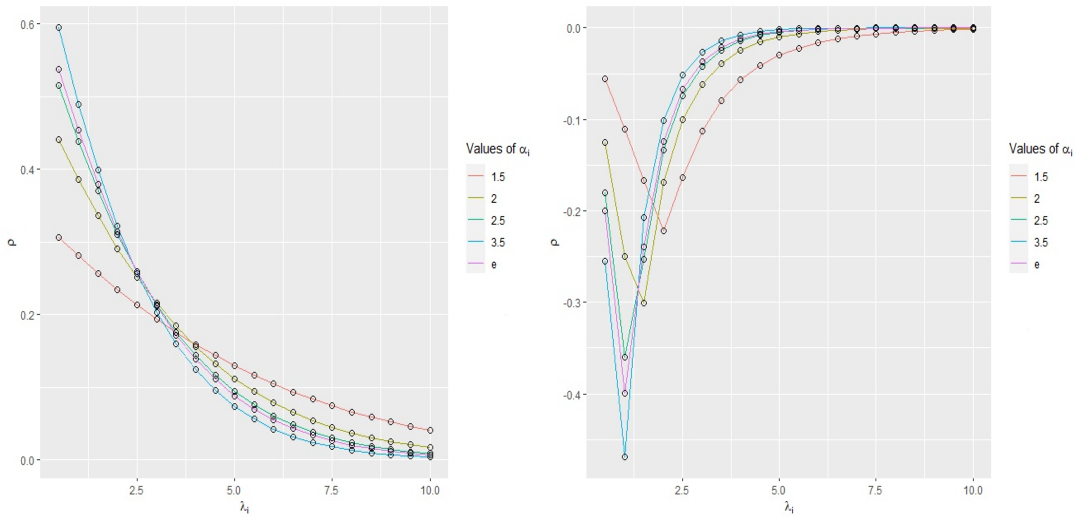


Figure 3.3: Correlation coefficient values considering numerous values for bases c_i and powers k_i, lower boundary(first graph) and upper boundary(second graph)

Under the assumption that correlation coefficient given by (3.18) and assuming Sarmanov distribution, the proposed bivariate Poisson distribution can be written equivalently as:

$$f(x_1, x_2) = \prod_{i=1}^{2} f_i(x_i) + \rho \prod_{i=1}^{2} \frac{[f_i(x_i^{'}) - f_i(x_i)]}{\sqrt{\lambda_i}(c_i^{-k_i} - 1)}, \tag{3.24}$$

where

$$f_i(x_i^{'}) \sim Poisson(\lambda_i \alpha^{-k_i}), \quad i = 1, 2.$$

$$
\begin{aligned}
f(x_1, x_2) &= \prod_{i=1}^{2} f_i(x_i)[1 + \rho \prod_{i=1}^{2} \frac{\sigma_i}{v_i} q_i(x_i)] \\
&= f_1(x_1) f_2(x_2) + \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} c_2^{-k_2 x_2} \\
&- \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} e^{\lambda_2 (c_2^{-k_2} - 1)} \\
&- \quad \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) e^{\lambda_1 (c_1^{-k_1} - 1)} (c_2^{-k_2} - 1 \\
&+ \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) e^{\lambda_1 (c_1^{-k_1} - 1)} e^{\lambda_2 (c_2^{-k_2} - 1)} \\
&= f_1(x_1) f_2(x_2) + \rho \frac{f_1(x_1^{'})}{\sqrt{\lambda_1}(c_1^{-k_1} - 1)} \frac{f_2(x_2^{'})}{\sqrt{\lambda_2}(c_2^{-k_2} - 1)} \\
&- \rho \frac{f_1(x_1)}{\sqrt{\lambda_1}(c_1^{-k_1} - 1)} \frac{f_2(x_2^{'})}{\sqrt{\lambda_2}(c_2^{-k_2} - 1)} \\
&- \rho \frac{f_1(x_1^{'})}{\sqrt{\lambda_1}(c_1^{-k_1} - 1)} \frac{f_2(x_2)}{\sqrt{\lambda_2}(c_2^{-k_2} - 1)} + \rho \frac{f_1(x_1)}{\sqrt{\lambda_1}(c_1^{-k_1} - 1)} \frac{f_2(x_2)}{\sqrt{\lambda_2}(c_2^{-k_2} - 1)} \\
&= f_1(x_1) f_2(x_2) + \rho \frac{[f_1(x_1^{'}) - f_1(x_1)][f_2(x_2^{'}) - f_2(x_2)]}{\sqrt{\lambda_1 \lambda_2}(c_1^{-k_1} - 1)(c_2^{-k_2} - 1)}. \qquad (3.25)
\end{aligned}
$$

Based on (3.6) and considering (3.7) the construction of continuous distributions is possible. Miravete (2009b) based on Sarmanov distribution studied not only a Double Poisson distribution in discrete case but also a bivariate Gamma distribution. Assuming that exponential distribution gain a dominant role because it holds for distances in time and more precisely the time between events, it has served as a first approach to a model for life testing. Many researchers developed bivariate exponential distributions without considering Sarmanov distribution for the construction. Gumbel (1960a) and Marshall and Olkin (1967) study constructions for new bivariate exponential distribution. On the work of Gumbel (1960a) bounds for correlation have been proved to be between -0.25 to +0.25 in one model and on the second model correlation takes only negative values between -0.4 to 0. Gumbel (1960a) based on work of Fréchet (1951) constructed a new family of copula named Gumbel copula. Farlie after his investigation to correlation coefficient studied a generalization of bivariate Farlie-Gumbel-Morgerstern distribution proposed by Morgernstern and Gumbel. For a review of various exponential distributions someone can study Balakrishnan (2014). Shubina and Lee (2004) study an alternative construction of bivariate distributions by using a new upper boundary for parameter ω. This idea could be correct in general case or in case where we have a variable in a bounded interval but not in case of $[0, \infty)$. Ting Lee (1996) proposed different construction of multiplicative factor depend on the variable's domain. Considering that X is continuous and taking values in $[0, \infty)$, we choose as marginal distributions exponential with parameters λ_1 and λ_2 accordingly. This choice of continuous variable provide that the construction of the proposed mixing function $q_i(x_i), i = 1, 2$ can be used either with discrete or with continuous variables. At exponential distribution the choice for the bounds for dependence parameter ω differ from those that have been already proposed at discrete case in previous chapter. Considering that $X_i \sim Exp(\lambda_i) = \lambda_i \exp(-\lambda_i x_i), \lambda_i, x_i > 0$ and based on (3.6), (3.7) and alternative definition for mixing functions given at section 3 we have:

$$q_i(x_i) = c_i^{-k_i x_i} - \frac{\lambda_i}{\log(e^{\lambda_i} c_i^{k_i})}, \quad i = 1, 2. \tag{3.26}$$

By using (3.18) correlation given by $\rho = \frac{\omega v_1 v_2}{\sigma_1 \sigma_2}$ where $v_i = E(q_i(x_i) f_i(x_i))$ or equivalently by

$$\rho = \omega \prod_{i=1}^{2} \frac{\lambda_i k_i \log c_i}{(\lambda_i + k_i \log c_i)^2}. \tag{3.27}$$

Bounds for ω are calculated assuming that $1 + \omega q_1(x_1) q_2(x_2)$ must be positive. The construction of boundaries based on Cambanis (1977) and the boundaries defined in section 4 could be given by:

$$b_1 = min \left\{ \frac{-1}{\prod_{i=1}^{2} \left(1 - E_{f_i} c_i^{-k_i x_i}\right), \prod_{i=1}^{2} E_{f_i} c_i^{-k_i x_i}} \right\}$$

and

$$b_2 = min \left\{ \frac{1}{\left(1 - E_{f_1} c_1^{-k_1 x_1}\right) E_{f_2} c_2^{-k_2 x_2}, \left(1 - E_{f_2} c_2^{-k_2 x_2}\right) E_{f_1} c_1^{-k_1 x_1}} \right\}, \tag{3.28}$$

where $E_{f_i} c_i^{-k_i x_i} = \int_{i=0}^{\infty} c_i^{-k_i x_i} \lambda_i \exp(-\lambda_i x_i)$.

In case of exponential distribution equations (3.18, 3.27) can be written as

$$\rho = \omega \prod_{i=1}^{2} (1 - E_{f_i} c_i^{-k_i x_i}) E_{f_i} c_i^{-k_i x_i}. \tag{3.29}$$

Taking all possible cases for ω based on the above boundaries give by (3.28) and correlation coefficient given by (3.29), we can see that ρ is always the other choice of the boundary and is always smaller than 1. For this reason we can reconstruct boundaries for ω as :

$$b_1 = \frac{-1}{min \left\{ \prod_{i=1}^{2} \left(1 - E_{f_i} c_i^{-k_i x_i}\right), \prod_{i=1}^{2} E_{f_i} c_i^{-k_i x_i} \right\}} \tag{3.30}$$

and

$$b_2 = \frac{1}{min\left\{\left(1 - E_{f_1}c_1^{-k_1x_1}\right)E_{f_1}c_1^{-k_1x_1}, E_{f_1}c_1^{-k_1x_1}\left(1 - E_{f_2}c_2^{-k_2x_2}\right)\right\}}. \tag{3.31}$$

Calculation of those boundaries in general case based on the study of Cambanis (1977). The main property that must be satisfied is the inequality of Frechet-Hoeffding bounds for joint distribution $F(x_1, x_2)$ where $F(x_1)$ and $G(x_2)$ are Exponential .

Based on the above equations we observe that $-0.95 \leq \rho \leq 0.95$ when $c_i = \exp(1), \;\; i = 1, 2$. As we can see at the first two graphs in case of $k_i = \lambda_i \in [0.5, 9]$ and $\alpha_i \in [0.5, 10]$ as values of parameters k_i, λ_i increase and α_i decrease ρ is in [-0.249, 0.506]. The color of line change when equation for the calculation of correlation change. As we can see for the lower bound in the case where parameters take the same values simultaneously equation does not change.

At the graphs below we observe the higher upper and lower boundaries for correlation of exponential distribution. As we can see when $\lambda_1 = \lambda_2$ and $k_1 = k_2$ we observe the higher value for the upper boundary and the opposite when $\lambda_i = k_j, \;\; , i, j = 1, 2, \;\; i \neq j$.

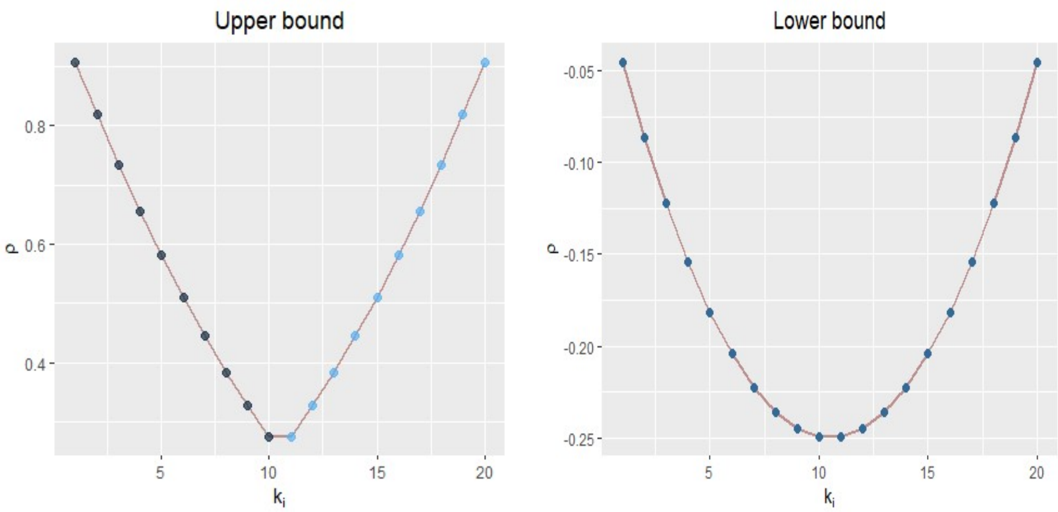


Figure 3.4: Correlation coefficient values for numerous equal values of λ_i and k_i

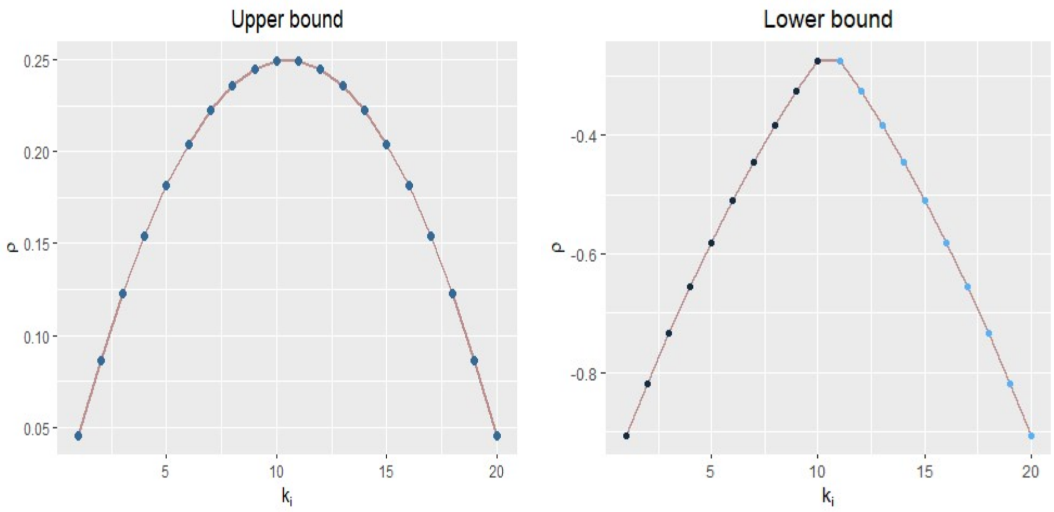


Figure 3.5: Correlation coefficient values when $\lambda_i = k_j,\ \ i \neq j,\ \ i,j = 1,2$

The fact that Exponential distribution has a cdf in a closed form provide a bivariate copula in a closed form. Bivariate cdf given by:

$$F(x_1,x_2) = \prod_{i=1}^{2}\left(1-e^{-\lambda_i x_i}\right)+\omega\prod_{i=1}^{2}\left[\frac{\lambda_i}{\lambda_i+k_i\log(c_i)}\left(1-e^{-(\lambda_i+k_i\log(c_i))x_i}\right)-\frac{\lambda_i}{\lambda_i+k_i\log(c_i)}F_i(x_i)\right]$$

and copula defined as

$$\begin{aligned} C(u,v) = uv \quad &+ \quad \omega\left(\frac{\lambda_1}{\lambda_1+k_1\log(c_1)}\left(1-e^{-(\lambda_1+k_1\log(c_1))F^{-1}(u)}-u\right)\right) \\ &\times \quad \left(\frac{\lambda_2}{\lambda_2+k_2\log(c_2)}\left(1-e^{-(\lambda_2+k_2\log(c_2))F^{-1}(v)}-v\right)\right). \end{aligned}$$

Bivariate distribution is given by:

$$f(x_1,x_2) \quad = \quad \prod_{i=1}^{2} f_i(x_i) + \prod_{i=1}^{2}\frac{\log(e^{\lambda_i}c_i^{k_i})}{k_i\log c_i}[f_i(x_i^{''})-f_i(x_i)], \tag{3.32}$$

where

$$f_i(x_i^{''}) = Exp(\lambda_i + k_i\log c_i),\ \ i=1,2.$$

$$
\begin{aligned}
f(x_1, x_2) &= \prod_{i=1}^{2} f_i(x_i)[1 + \rho \prod_{i=1}^{2} \frac{\sigma_i}{v_i} q_i(x_i)] \\
&= f_1(x_1) f_2(x_2) + \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} c_2^{-k_2 x_2} \\
&- \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) c_1^{-k_1 x_1} \frac{\lambda_2}{\log(e^{\lambda_2 c_2^{k_2}})} \\
&- \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) \frac{\lambda_1}{\log(e^{\lambda_1 c_1^{k_1}})} c_2^{-k_2 x_2} \\
&+ \rho \frac{\sigma_1}{v_1} \frac{\sigma_2}{v_2} f_1(x_1) f_2(x_2) \frac{\lambda_1}{\log(e^{\lambda_1 c_1^{k_1}})} \frac{\lambda_2}{\log(e^{\lambda_2 c_2^{k_2}})} \\
&= f_1(x_1) f_2(x_2) + \rho \frac{f_1(x_1^{''})}{\sqrt{\lambda_1}(c_1^{-k_1} - 1)} \frac{f_2(x_2^{''})}{\sqrt{\lambda_2}(c_2^{-k_2} - 1)} \\
&- \rho \frac{f_1(x_1)}{-\lambda_1 - 1} \frac{f_2(x_2^{''})}{-\lambda_2 - 1} \\
&- \rho \frac{f_1(x_1^{''})}{-\lambda_1 - 1} \frac{f_2(x_2)}{-\lambda_2 - 1)} + \rho \frac{f_1(x_1)}{-\lambda_1 - 1} \frac{f_2(x_2)}{-\lambda_2 - 1)} \\
&= f_1(x_1) f_2(x_2) + \rho \, \frac{[f_1(x_1^{''}) - f_1(x_1)][f_2(x_2^{''}) - f_2(x_2)]}{(\lambda_1 + 1)(\lambda_2 + 1)}. \qquad (3.33)
\end{aligned}
$$

3.4.1 Copulas in higher dimensions

Ting Lee (1996) proposed the multivariate Sarmanov distribution extended (3.6) as:

$$
f_{12\ldots n}(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} f_i(x_i)[1 + R_{q_1 q_2 \ldots q_n}(x_1, x_2, \ldots, x_n)],
$$

where

$$
\begin{aligned}
R_{12\ldots n} &= \sum_{i<}^{n-1} \sum_{j=1}^{n} \omega_{ij} q_i(x_i) q_j(x_j) + \\
& \sum_{i<}^{n-2} \sum_{j<}^{n-1} \sum_{k=3}^{n} \omega_{ijk} q_i(x_i) q_j(x_j) q_k(x_k) + \cdots + \omega_{ijk\ldots m} \prod_{i=1}^{m} q_i(x_i) \qquad (3.34)
\end{aligned}
$$

and $\Omega_n = \{\omega_{ijk}, \ldots, \omega_{ijk\ldots m}\}$ is the set of real numbers chosen such that

$$
1 + R_{q_1 q_2 \ldots q_n}(x_1 x_2 \ldots x_n) \geq 0.
$$

The main drawback is the high computational cost needed to calculate all the combinations mentioned to the above equation. Miravete (2009b) based on the given multivariate extension, proposed the extended Compound Poisson and Gamma distributions in three dimensions. Another multivariate distribution is proposed considering dependence only between two variables.

$$R_{q_1 q_2 \dots q_n}(x_1, x_2, \dots, x_n) = \sum_{i<}^{n-1} \sum_{j=2}^{n} \omega_{ij} q_i(x_i) q_j(x_j). \tag{3.35}$$

Correlation coefficient is given by:

$$\rho_{ij} = \frac{\omega_{ij} v_i v_j \mu_k}{\sigma_i \sigma_j}, \quad i, j, k = 1, 2, 3, i \neq j \neq k,$$

where v_i , $i = 1, 2, 3$ given by (3.14) and $\mu_i = E f_i(x_i)$, $i = 1, 2, 3$.

A trivariate Poisson distribution similar to this proposed previously given by:

$$\begin{aligned} p(x_1, x_2, x_3) &= P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3) - P(X_1 \leq x_1 - 1, X_2 \leq x_2, X_3 \leq x_3) \\ &- P(X_1 \leq x_1, X_2 \leq x_2 - 1, X_3 \leq x_3) - P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1, X_3 \leq x_3) \\ &- [P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3 - 1) - P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1, X_3 \leq x_3 - 1)] \\ &- [P(X_1 \leq x_1 - 1, X_2 \leq x_2, X_3 \leq x_3 - 1) - P(X_1 \leq x_1 - 1, X_2 \leq x_2 - 1, X_3 \leq x_3 - 1)], \end{aligned}$$

where trivariate Sarmanov distribution can be defined more simplier as:

$$\begin{aligned} F(x_1, x_2, x_3) = \prod_{i=1}^{3} F(x_i) & [1 + \omega_{12} F(x_1) F(x_2) (1 - F(x_1)) (1 - F(x_2)) F(x_3) \\ & + \omega_{13} F(x_1) F(x_3) (1 - F(x_1)) (1 - F(x_3)) F(x_2) \\ & + \omega_{23} F(x_2) F(x_3) (1 - F(x_2)) (1 - F(x_3)) F(x_1)]. \end{aligned}$$

At this chapter we discussed a way to construct multivariate distributions. Based on Sklar's theorem and Sarmanov distribution, many alternative approaches have been also studied. The novelty in our study is not only that we try to construct a less complicated multivariate distribution at the same period with Piancastelli et al. (2023) but we provide a wider range at correlation coefficient considering another function in the term of multiplicative factor. The importance is that we could imply this function in both discrete and continuous case taking each time different construction for ρ's boundaries. Furthermore when we have a continuous variable, calculations are less restrictive and complicated as proved in case of exponential distribution. The construction of our new copula has been done not only to be useful for construction of any bivariate distribution considering appropriate marginals each time, but also to construct a new Poisson INGARCH(1,1) and a new CARR(1,1) time series models with studies provided at the next two chapters.

Chapter 4

A multivariate INGARCH(1,1) time series model

4.1 Model

Fokianos et al. (2020) studied linear and log-linear models for multivariate count time series models with Poisson marginals. The model was structured taking into account fundamental properties of Poisson process via a vector of continuous distributions and considering a copula form. The dependence among marginal Poisson components based on the properties of the Poisson process. The construction of the multivariate Poisson INGARCH model based on the construction of Freeland and McCabe (2004) A brief discussion about this model was done in the introduction of the thesis. In what follows we assume that $X_t = (X_{i,t}),\ \ i = 1, ..., p,\ \ t = 1, 2...,$ denotes a p–dimensional count time series. Let $\lambda_t = (\lambda_{i,t}),\ \ i = 1, ..., p,\ \ t = 1, 2, \ldots, T$ be the corresponding p-dimensional intensity process and $\mathcal{F}_t^{X,\lambda}$ the σ–field generated by $X_0, , X_t, \lambda_0$ with λ_0 being a p-dimensional vector denoting the starting value of λ_t. With this notation, the intensity process is given by $\lambda_t = E[X_t|\mathcal{F}_t^{X,\lambda}]$. The new trivariate model constructed assuming that the linear dependence of λ_{it} on $\lambda_{j,t-1}$ and $X_{j,t-1}$ in the presence of $\lambda_{i,t-1}$ and $X_{j,t-1}$, $i \neq j,\ \ i, j = 1, 2, 3$

$$X_t \mid \mathcal{F}_{t-1}^{X,\lambda} \sim TP(\lambda_\mathbf{t})$$

where

$$\lambda_\mathbf{t} = \mathbf{d} + \mathbf{A}\lambda_\mathbf{t-1} + \mathbf{B}\mathbf{X}_\mathbf{t-1}, \tag{4.1}$$

assuming that for each i = 1, 2, 3

$$X_{i,t} \mid \mathcal{F}_{t-1}^{X,\lambda} \sim P(\lambda_{i,t}),\ \ i = 1, 2, 3, \tag{4.2}$$

where

$$\lambda_{i,t} = d_i + \sum_{j=1}^{3} a_{i,j}\lambda_{j,t-1} + \sum_{j=1}^{3} b_{i,j}X_{j,t-1}, \quad i,j = 1,2,3. \tag{4.3}$$

Each $\lambda_{it}, \ i = 1, \ldots, p$ depends not only on its own past but also to past observations of the other series and this indicates that **A** and **B** are matrices of dimension $p \times p$ (in multivariate case). Those matrices are responsible for heteroskedasticity because in case of constant conditional mean and variance($\mathbf{A} = \mathbf{B} = 0$) the model becomes an integer AR Poisson model. Furthermore if elements $a_{ij}, b_{ij}, \ i \neq j, \ i,j = 1,2,3$ are zero then matrices A, B are diagonal and conditional mean depends only on its own past and its past values. In the case where $\mathbf{A} = 0$ then the model reduced to an INARCH process. Fokianos et al. (2020) proved that consider model (1.14) and suppose that $||A||_1 + ||B||_1 < 1$, $||A||_1 = max_{1 \leq j \leq p} \sum_{i=1}^{q} |a_{ij}|$, then there exists a unique solution (Y_t, λ_t) to model (4.1) which is stationary and ergodic. The elements of **d,A,B** are assumed to be positive considering positivity of $\lambda_{i,t}$

4.2 Ergodic Properties

4.2.1 Weak dependence criteria

Fokianos et al. (2020) studied stationarity and ergodicity of the linear model by using Tweedie criteria. Based on the way with which Fokianos et al. (2009) studied ergodicity of univariate linear INGARCH model using a perturbed model, on a similar way Fokianos et al. (2020) study ergodic properties for a bivariate INGARCH model. At the proposed trivariate Poisson INGARCH model we study ergodicity of the model by using weak dependence criteria based on the work of Dedecker et al. (2007) without using a pertubation model. There are numerous criteria of weak dependence but we study ergodicity based on θ and τ weak dependence criteria and based on (Dedecker et al. (2007), pg. 115) $\tau_1(\mathcal{F}_{t-1}^X, X_t) = ||X_t - \tilde{X}_t||_1$. Under the assumption that $h : \mathbb{R}^m \to \mathbb{R}^n$ is a Lipschitz function, where m is and n is , . We prove that $\{X_t\}$ is ergodic. Proof is given in the Appendix.

Theorem 3. *If $h(\lambda_\mathbf{t}, \mathbf{X_t})$ satisfies the Lipschitz condition, then $\{X_t\}_{t\in\mathbb{Z}}$ is stationary and θ-weak dependent.*

Considering that $\tilde{\mathcal{F}}_{t-1}$ is a σ-algebra generated by $\lambda_i, X_i, \tilde{\lambda}_i, \tilde{X}_i$, i=0,...,t-1 where $\tilde{\lambda}_t, \tilde{X}_t$ are perturbed volatilities, then

$$||X - \tilde{X}||_1 = E|X - \tilde{X}| = E[E\left(|X - \tilde{X}| \mid \tilde{\mathcal{F}}_{t-1}\right)] = E|\lambda_t - \tilde{\lambda}_t| = ||\lambda - \tilde{\lambda}||_1.$$

Based on the above, under the assumption that multivariate function $h : \mathbb{R}^m \to \mathbb{R}^n$ satisfy the Lipschitz condition $|h(\mathbf{x}) - h(\mathbf{y})| \leq L|\mathbf{x} - \mathbf{y}|$ where L is constant less than 1 and by using Minkowski's inequality where A_i,B_j are squared, invertible matrices in general case that we have a multivariate INGARCH(p,q) model we have that volatility given by :

$$\lambda_t = \mathbf{d} + \mathbf{A}\lambda_{t-i} + \mathbf{B}\mathbf{X}_{t-i},$$

where

$$\lambda_t = \begin{bmatrix} \lambda_{1t} \\ \vdots \\ \lambda_{nt} \end{bmatrix}, \mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}, A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}, \text{ and } B = \begin{bmatrix} B_1 \\ \vdots \\ B_n \end{bmatrix}$$

and

$$\lambda_t = d + \textstyle\sum_{i=1}^p A_i\lambda_{t-i} + \sum_{j=1}^q B_j Y_{t-j}.$$

$$\begin{aligned}
||X_t - \tilde{X}_t||_1 &= ||\lambda_t - \tilde{\lambda}_t||_1 = ||h(\lambda_{t-i}, X_{t-i}) - h(\tilde{\lambda}_{t-i}, \tilde{X}_{t-i})||_1 \\
&= ||\sum_{i=1}^p A_i(\lambda_{t-i} - \tilde{\lambda}_{t-i}) + \sum_{j=1}^q B_j(X_{t-j} - \tilde{X}_{t-j})||_1 \\
&\leq ||\sum_{i=1}^p A_i||_1||\lambda_{t-i} - \tilde{\lambda}_{t-i}||_1 + ||\sum_{j=1}^q B_j||_1||X_{t-j} - \tilde{X}_{t-j}||_1 \\
&\leq \sum_{i=1}^p ||A_i||_1||\lambda_{i,t-i} - \tilde{\lambda}_{i,t-i}|| + \sum_{j=1}^q ||B_j||_1||X_{j,t-j} - \tilde{X}_{j,t-j}||_1 \\
&\leq \left(\sum_{i=1}^p ||A_i||_1 + \sum_{j=1}^q ||B_j||_1\right) max_{1\leq k_1 \leq m_1}||X_{k_1,t-k_1} - \tilde{X}_{k_1,t-k_1}||_1.
\end{aligned}$$

By iterating the procedure we have that:

$$||X_t - \tilde{X}_t||_1 \leq \left(\sum_{i=1}^p ||A_i||_1 + \sum_{j=1}^q ||B_j||_1\right)^{\sum_{i=1}^s k_i} \max ||X_{k_s,t-\sum_{i=1}^s k_i} - \tilde{X}_{k_j,t-\sum_{j=1}^s k_s}||_1,$$

where $k_i \in [1, max(p,q)]$.

In our study we consider a trivariate INGARCH time series model with Poisson marginals. More specifically by iterating (1.14) we have an equation which decrease computational cost :

$$
\begin{aligned}
\lambda_t &= \mathbf{d} + \mathbf{A}\lambda_{t-1} + \mathbf{B}X_{t-1} \\
&= \mathbf{d} + \mathbf{A}\mathbf{d} + \mathbf{A^2}\lambda_{t-2} + \mathbf{AB}X_{t-2} + \mathbf{B}X_{t-1} \\
&= \dots \\
&= \sum_{i=0}^{k-1} \mathbf{A^i}\mathbf{d} + \mathbf{A^k}\lambda_{t-k} + \mathbf{B}\sum_{i=0}^{k-1} \mathbf{A^i}X_{t-i-1}, \qquad (4.4)
\end{aligned}
$$

where d_i is the i-th element of $\mathbf{d}$ and a_{ij} is the (i,j)-element of $\mathbf{A}$ ($\mathbf{B}$ respectively). Assuming first order stationarity, taking conditional expectation and from passage to the limit as $k \to \infty$ the above equation can be written as:

$$
\lambda_t = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{d} + \sum_{i=0}^{\infty} \mathbf{A}^i \mathbf{B} \mathbf{X}_{t-i-1}, \qquad (4.5)
$$

where $\mathbf{I}$ is the identical matrix of 3-dimensions.

In trivariate INGARCH model according to construction of trivariate Poisson distribution given by (3.34) both negative and positive correlation can be calculated. In trivariate Poisson distribution from the restriction $\omega_{12}q_1(x_1)q_2(x_2) + \omega_{13}q_1(x_1)q_3(x_3) + \omega_{23}q_2(x_2)q_3(x_3)$ parameters $\omega_{ij}, \ , i, j = 1, 2, 3, i \neq j$ are bounded based on (3.22 , 3.23) but the main difference is that at the numerators $l_k, k = 1, 2, 3$ are chosen randomly such that $l_1 + l_2 + l_3 = 1$.

4.3 Estimation

The vital problem in multivariate distribution is the expression of the joint distribution. At those models that have been already proposed the multivariate pmf is quite complicated and maximum likelihood based inference is challenging. Considering that $\{\mathbf{X}_t, \ t = 1, 2, ..., n\}$ is a sample from the count time series model proposed above and the vector of the parameters denoted by $\boldsymbol{\theta}$ where $\boldsymbol{\theta} = \left(\boldsymbol{d}^T, \boldsymbol{A}^T, \boldsymbol{B}^T, \boldsymbol{\omega}^T\right)$. Using the general distribution defined by (3.36), (3.34) we can estimate parameters of each multivariate model. Fokianos et al. (2020) maximized the Quasi log-likelihood function. Based on lemmas (4.1) and (4.2), the examination of asymptotic normality has been provided based on the unperturbed model. The log-likelihood function can be written as:

$$
\log L(\theta) = \sum_{t=1}^{n}\sum_{i=1}^{3} \log p_i(x_{it}) + \sum_{t=1}^{n} \log[1 + \sum_{i<}^{2}\sum_{j=2}^{3} \omega_{ij} q_i(x_{it}) q_j(x_{jt})].
$$

Considering (4.3) log-likelihood function for trivariate Poisson INGARCH model is given by:

$$l(\theta) = \sum_{t=1}^{n} \left(\sum_{i=1}^{3} [-\lambda_i(\theta) + X_{i,t} \log(\lambda_i(\theta)) - \log(X_{i,t}!)] + \log \phi_t(\theta) \right), \quad (4.6)$$

where

$$\phi_t(\theta) = 1 + \sum_{i<}^{2} \sum_{j=2}^{3} \omega_{ij} \left(c_i^{(-k_i-1)} - e^{\lambda_i(c_i^{-k_i}-1)} \right) \left(c_j^{(-k_j-1)} - e^{\lambda_j(c_j^{-k_j}-1)} \right). \quad (4.7)$$

Score function is given by:

$$S_n(\theta) = \frac{\partial l(\theta)}{\partial \theta} = \sum_{t=2}^{n} \frac{\partial l_t(\theta)}{\partial \theta} = \sum_{t=1}^{n} \left(\sum_{i=1}^{3} \left(\frac{X_{i,t}}{\lambda_{i,t}} - 1 \right) \frac{\partial \lambda_{i,t}(\theta)}{\partial \theta} + \frac{1}{\phi_t(\theta)} \frac{\partial \phi_t(\theta)}{\partial \theta} \right),$$

where $\lambda_{i,t}$ defined by (1.14) and $\theta = (d_i, a_{ij}, b_{ij}, \omega_{ij})$, $i, j = 1, 2, 3$. The score function satisfies:

$$\frac{1}{\sqrt{n}} S_n(\theta) \to N(0, I(\theta_0)) \quad (4.8)$$

Theorem 4. *For the INGARCH model as $n \to \infty$*

$$\sqrt{n}(\hat{\theta} - \theta_0) \to N(0, I(\theta_0)^{-1})$$

where

$$I(\theta_0) = E\left(\frac{\partial l(\theta_0)}{\partial \theta} \frac{\partial l(\theta_0)}{\partial \theta^T} \right) = -E\left(\frac{\partial^2 l(\theta_0)}{\partial \theta \partial \theta^T} \right).$$

Based on the work of Lee et al. (2016) conditions for strong consistency of MLE estimator in INGARCH models are accomplished in our case. The true parameter value θ_0 is an interior poit on $\boldsymbol{\Theta}$, with $\boldsymbol{\Theta}$ being a compact set. Furthermore considering conditions of stationarity **d**,**A**,**B** do not have negative entries and $Cov(X_{it}, X_{jt}) < min(a_1, a_2)$ where $(a_1, a_2)^T = (I - \mathbf{A})^{-1}\mathbf{d}$ for all $\theta \in \Theta$. Considering (3.22, 3.23), $\phi_L \leq \phi_t(\theta) \leq \phi_U$ and based on condition of stationarity λ_t is bounded $\lambda_L \leq \lambda_t \leq \lambda_U$. Assuming parameter vectors $\theta_i = (\alpha_{ij}, b_{ij}, d_i)$, $i, j = 1, 2, 3$ and $\theta_4 = (\omega_{12}, \omega_{13}, \omega_{23})$ first and second order partial derivatives are presented in appendix B. Moreover based on equation (1.14) derivatives for $\lambda_{it}(\theta)$ are:

$$\begin{aligned}
\frac{\partial \lambda_i(\theta)}{\partial d_i} &= d_i + b_{ij} \frac{\partial \lambda_{i,t-1}}{\partial d_i}, \\
\frac{\partial \lambda_i(\theta)}{\partial b_{ij}} &= \lambda_{j,t-1} + b_{ij} \frac{\partial \lambda_{j,t-1}}{\partial b_{ij}}, \\
\frac{\partial \lambda_i(\theta)}{\partial a_{ij}} &= \lambda_{j,t-1} + a_{ij} \frac{\partial \lambda_{i,t-1}}{\partial a_{ij}}.
\end{aligned}$$

and they are $\mathcal{F}_{t-1}$-measurable. Assuming second order partial derivatives we have:

$$E\left(\frac{X_{it}}{\lambda_i^2(\theta)} \mid \mathcal{F}_{t-1}\right) = -\frac{1}{\lambda_{it}}E\left(\left(c_i^{-k_i x_{it}} - e^{\lambda_i(c_i^{k_i}-1)}\right) \mid \mathcal{F}_{t-1}\right) = 0$$

Under above conditions and considering Minkowski's inequality we have:

$$E\left(-\frac{\partial^2 l}{\partial \theta_{ij}^2} \mid \mathcal{F}_{t-1}\right) \leq \frac{1}{\lambda_U} + \frac{c_i^{-k_i}-1}{\phi_U^2} + \frac{(c_i^{-k_i}-1)^2}{\phi_U} + 2\frac{(c_i^{-k_i}-1)^2}{\phi_U^2}. \tag{4.9}$$

Based on the above $E\sup_{\theta\in\Theta}||\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta^T}||_1 < \infty$. Proofs are given in Appendix A. Using Taylor series expansion for $\frac{\partial l(\theta)}{\partial \theta}$ we have

$$0 = \frac{\partial l(\hat{\theta})}{\partial \theta} = \frac{\partial l(\theta_0)}{\partial \theta} + (\hat{\theta}-\theta_0)\frac{\partial^2 l(\tilde{\theta})}{\partial\theta\partial\theta^T} \qquad \text{or} \qquad \frac{1}{\sqrt{n}}\frac{\partial l(\theta_0)}{\partial \theta} = \sqrt{n}(\hat{\theta}-\theta_0)\left(-\frac{1}{n}\frac{\partial^2 l(\theta)}{\partial\hat{\theta}\partial\theta^T}\right).$$

Considering Lemma 6 of Lee et al. (2016) $-\frac{1}{n}\frac{\partial^2 l(\theta_n^*)}{\partial\theta\partial\theta^T} \to I(\theta_0)$ and by using martingale central limit theorem $\frac{1}{\sqrt{n}}\frac{\partial l(\theta_0)}{\partial \theta} \to N(0, I(\theta_0))$. Based on the above

$$\sqrt{n}(\hat{\theta}-\theta_0) \to N(0, I(\theta_0)^{-1}).$$

Considering that in trivariate case there are numerous parameter for estimations other ways for estimations could be provided. Here we briefly discussed the method of likelihood by parts based on score function's decomposition. Assuming that model given by (4.2), (4.3) likelihood by parts defined as:

$$l(\theta) = l(\theta_1) + l(\theta_1, \theta_2), \tag{4.10}$$

where

$$l(\theta_1) = \sum_{t=1}^{n}\left(\sum_{i=1}^{3}[-\lambda_i(\theta) + X_{i,t}\log(\lambda_i(\theta)) - \log(X_{i,t}!)]\right),$$

$$l(\theta_1, \theta_2) = \sum_{t=1}^{n}\log\phi_t(\theta)$$

and

$$\theta_1 = (d_i, a_{ij}, b_{ij}), \theta_2 = (\omega_{ij}), \ \ i, j = 1, 2, 3.$$

Song et al. (2005) proposed repetition of steps 1-k to solve the above equations under the assumption that as $k \to \infty$, θ_n^k will converge to θ.

- **Step 1:** $\frac{\partial l(\theta_1)}{\partial \theta_1} = 0$, for θ_1^1 and $\frac{\partial l(\theta_1^1, \theta_2)}{\partial \theta_2} = 0$ for θ_2^1.

- **Step k:**$\frac{\partial (l(\theta_1))}{\partial (\theta_1)} = -\frac{\partial l(\theta_1^{k-1}, \theta_2^{k-1})}{\partial \theta_1}$ for θ_1^k and $\frac{\partial l(\theta_1^{k-1}, \theta_2)}{\partial \theta_2} = 0$ for θ_2^k.

In our case this method is very useful in comparison with the maximum likelihood approach of the full bivariate distribution. This means that calculating the partial derivative of $l(\theta_1)$ and then substituting at $l(\theta_1^{k-1}, \theta_2)$ the number of parameters foe estimation of $l(\theta_1^{k-1}, \theta_2)$ is decreased and the algorithm becomes less intensive.

4.4 Simulations

We run a simulation study in order to check the performance of estimators. In our simulation study we use sample size of n=50, n=200 and n=500 with m=200 replications. Based on inverse transformation method for the data generation we have:

- **Step 1:** Choose initial values for the parameter vector $\theta = (d_i, a_{ij}, b_{ij}, \omega_{ij})$, $i, j = 1, 2$.

- **Step 2:** Generate an observation x_1 from a $P(\lambda_1)$ where $\lambda_{1t} = d_1 + a_{11}\lambda_{1,t-1} + a_{12}\lambda_{2,t-1} + b_{11}X_{1,t-1} + b_{12}X_{2,t-1}$.

- **Step 3:** Generate a random number $u_1 \sim U(0, 1)$.

- **Step 4:** If there exist a k_1 such that $P_1(k_1 - 1 \mid x_1) \leq u_1 \leq P_1(k_1 \mid x_1)$ set $x_2 = k_1$, where $P_1(k_1 - 1 \mid x_1)$ and $P(k_1 \mid x_1)$ are conditional distributions.

- **Step 5:** Generate a number $u_2 \sim U(0, 1)$.

- **Step 6:** If there exist a k_2 such that $P_2(k_2 - 1 \mid x_1, x_2) \leq u_2 \leq P_2(k_2 \mid x_1, x_2)$ then set $x_3 = k_2$.

where

$$P_1(x_2 \mid x_1) = P(\lambda_2)\left(1 + A(\lambda_1, \lambda_2)\right),$$

$$P_2(x_3 \mid x_1, x_2) = P(\lambda_3^\star)\frac{1 + \sum_{i<j=1}^{3} A(\lambda_i^\star, \lambda_j^\star)}{A(\lambda_1^\star, \lambda_2^\star)},$$

$$A(\lambda_i,\lambda_j)=\omega_{ij}\left(e^{-k_i x_i}-e^{\lambda_i^\star\left(e^{-k_i}-1\right)}\right)\left(e^{-k_j x_j}-e^{\lambda_j^\star\left(e^{-k_j}-1\right)}\right)$$

and λ_{it} defined in step 2 and $\lambda_{it}^\star$ defined as

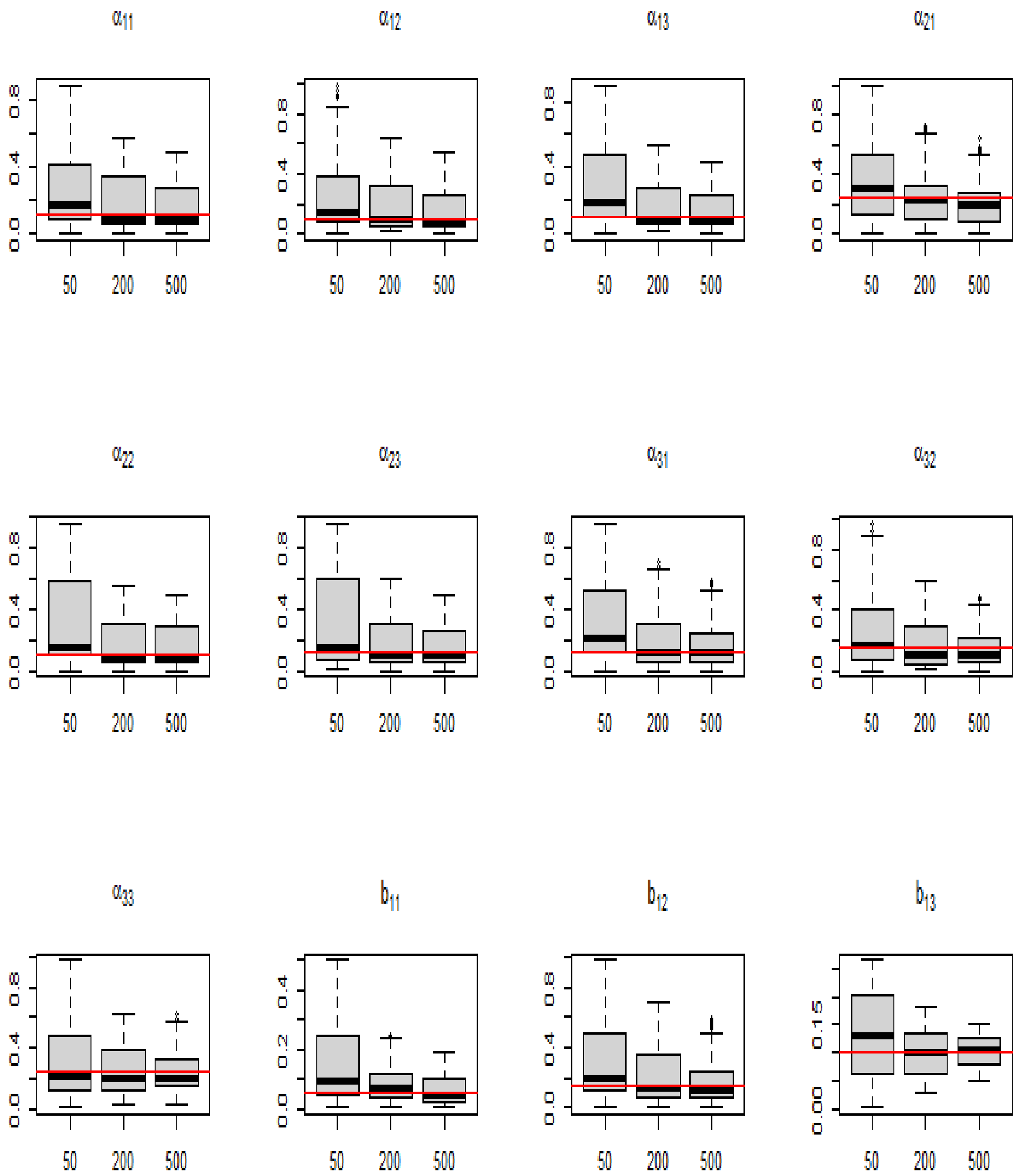
$$\lambda_{it}^\star=d_i+a_{i1}\lambda_{1,t-1}^\star+a_{i2}\lambda_{2,t-1}^\star+a_{i3}\lambda_{3,t-1}^\star+b_{i1}X_{1,t-1}+b_{i2}X_{2,t-1}+b_{i3}X_{3,t-1}.$$

Furthermore at the formulas above $P(\lambda_i)$ and $P(\lambda_i^\star)$ are Poisson distributions where λ_{it} and $\lambda_{it}^\star$ defined at steps above. For the trivariate Poisson model we estimate parameter vector $\theta=(d_1,d_2,d_3,a_{ij},b_{ij},\omega_{ij})$ with both negative and positive values for parameters $\omega_{ij},\ \ i,j=1,2,3$. In order to take optimal values for parameters $\omega_{ij},\ \ i,j=1,2,3$ we take into account (3.22; 3.23) and substituting optimal values of the other parameters by the way described above , we get corresponding ω_{ij}'s values. At the case in which the joint distribution does not have a closed form the choice for initial value of ω is demanded. Fokianos et al. (2020) at the proposed model overcome the problem considering the method studied by Tjøstheim and Hufthammer (2013) for local Gaussian correlation. Considering a two-dimensional random variable the bivariate density can be approximated by a Gaussian bivariate density in a neighbourhood at each point $x=(x_1,x_2)$ where correlation coefficient defined as $\rho=\frac{\sigma_{12}}{\sigma_1\sigma_2}$. Tjøstheim and Hufthammer (2013) provided that the parameter vector of Gaussian distribution is defined by using a likelihood related penalty function q given by :

$$q=\int K_b(v-x)\left[\psi(v,\theta(x))-\log\psi(v,\theta(x))f(v)\right]dv, \tag{4.11}$$

where $K_b(v-x)=\frac{Kb_1^{-1}(v_1-x_1)Kb_2^{-1}(v_2-x_2)}{b_1b_2}$ and $b=(b_1,b_2)$ is a bivariate bandwidth. An other one interpretation for the above equation is a locally weighted Kullback-Leibler distance from f to $\psi(.,\theta(x))$.

The biases and standard deviations are presented in (4.1; 4.2; 4.3) in Appendix. Simulations were carried out in R using constrOptim considering constraints defined at the previous sections. All estimators perform well as small biases are obtained.



α_{11}
α_{12}
α_{13}
α_{21}
α_{22}
α_{23}
α_{31}
α_{32}
α_{33}
b_{11}
b_{12}
b_{13}
0.0
0.2
0.4
0.8
0.00
0.15
50
200
500

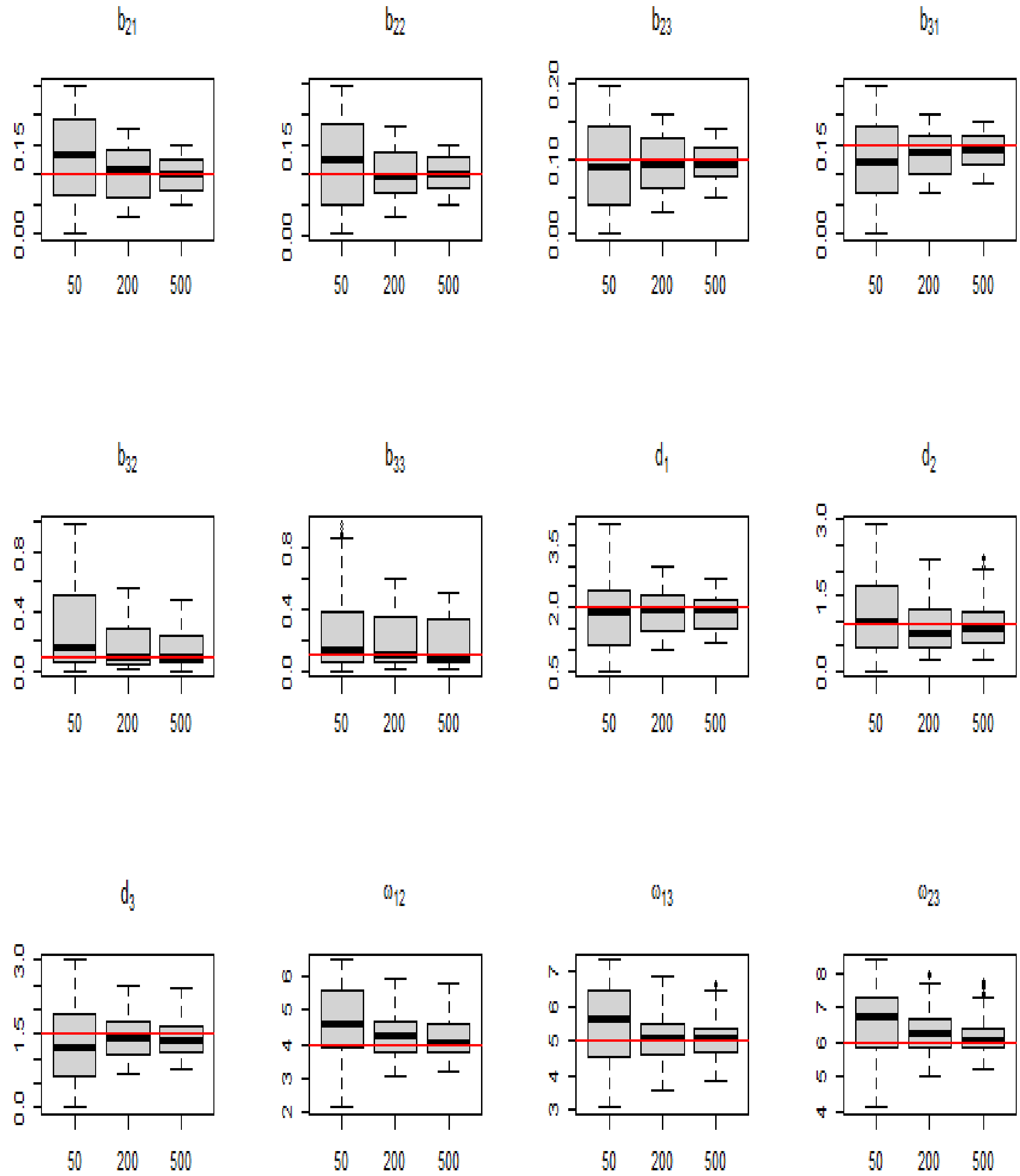
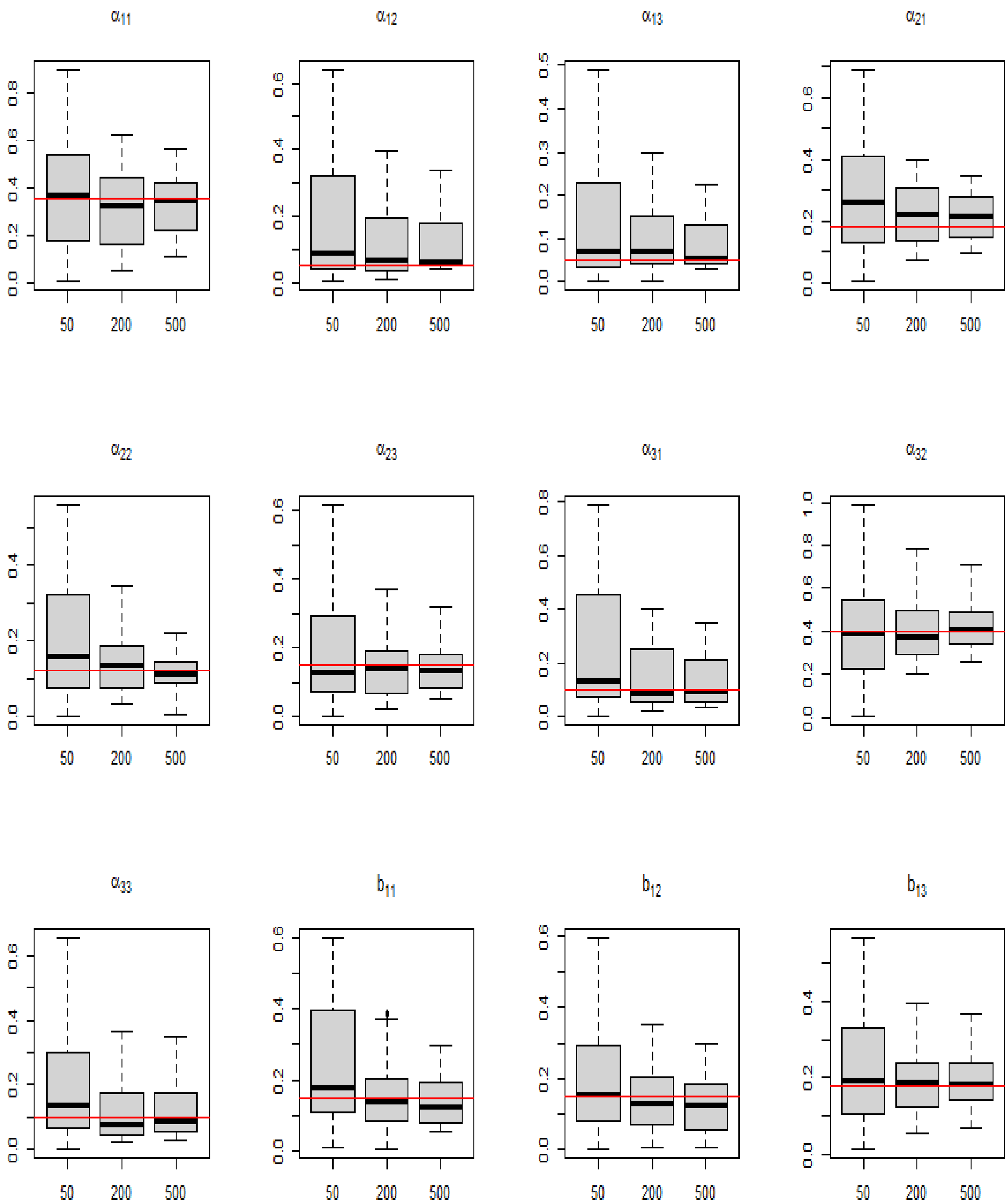


Figure 4.1: Boxplot with results from the simulation experiment with initial parameter values $(a_{ij},b_{ij},d_{ij},\omega_{ij})$=(0.12,0.1,0.1,0.22,0.15,0.18,0.15,0.18, 0.22,0.16,0.18,0.18,0.1,0.1,0.1,0.1,0.15,0.1,0.1,2,1,1.5,4,5,6)



α_{11}
α_{12}
α_{13}
α_{21}
α_{22}
α_{23}
α_{31}
α_{32}
α_{33}
b_{11}
b_{12}
b_{13}
50 200 500

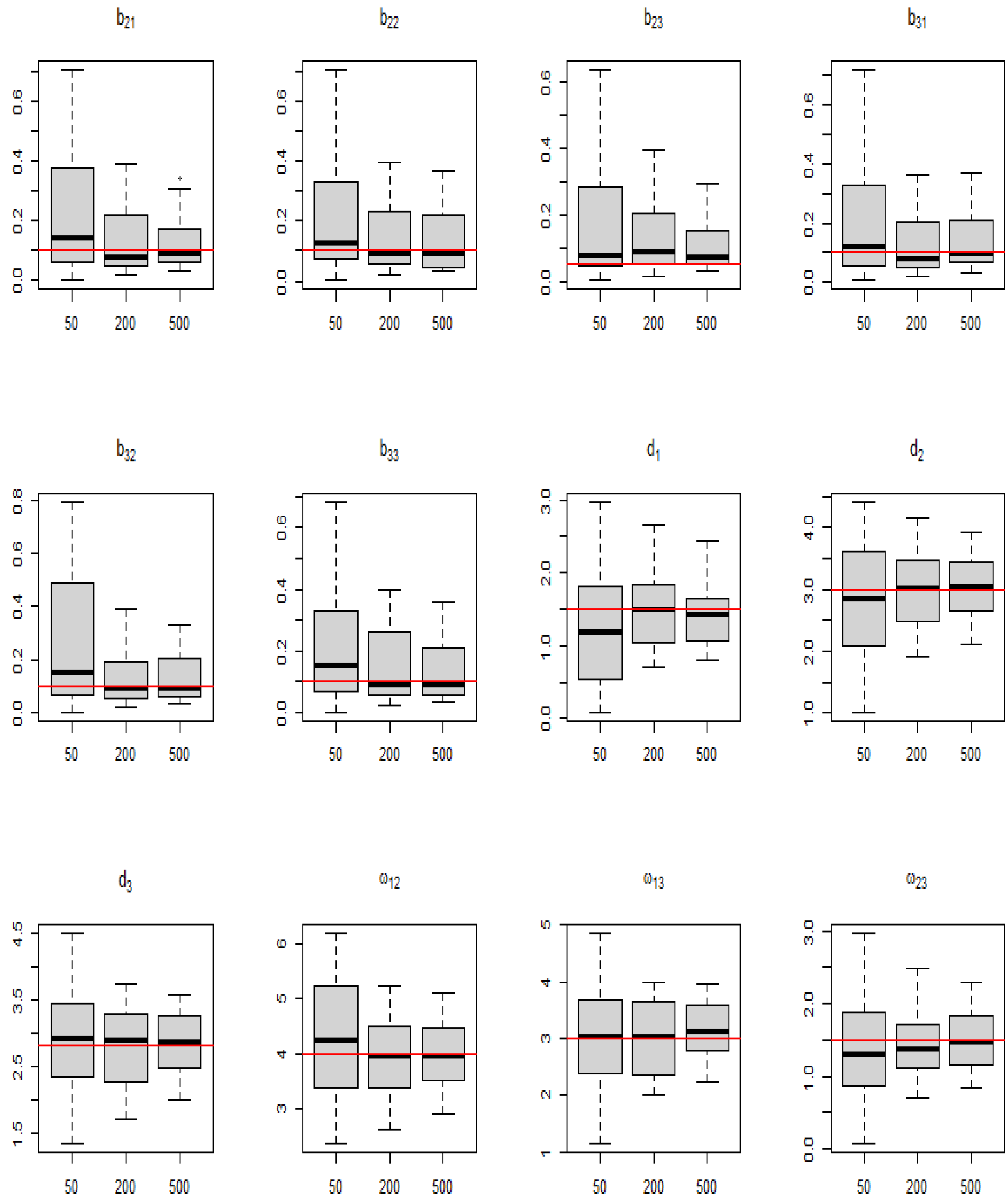
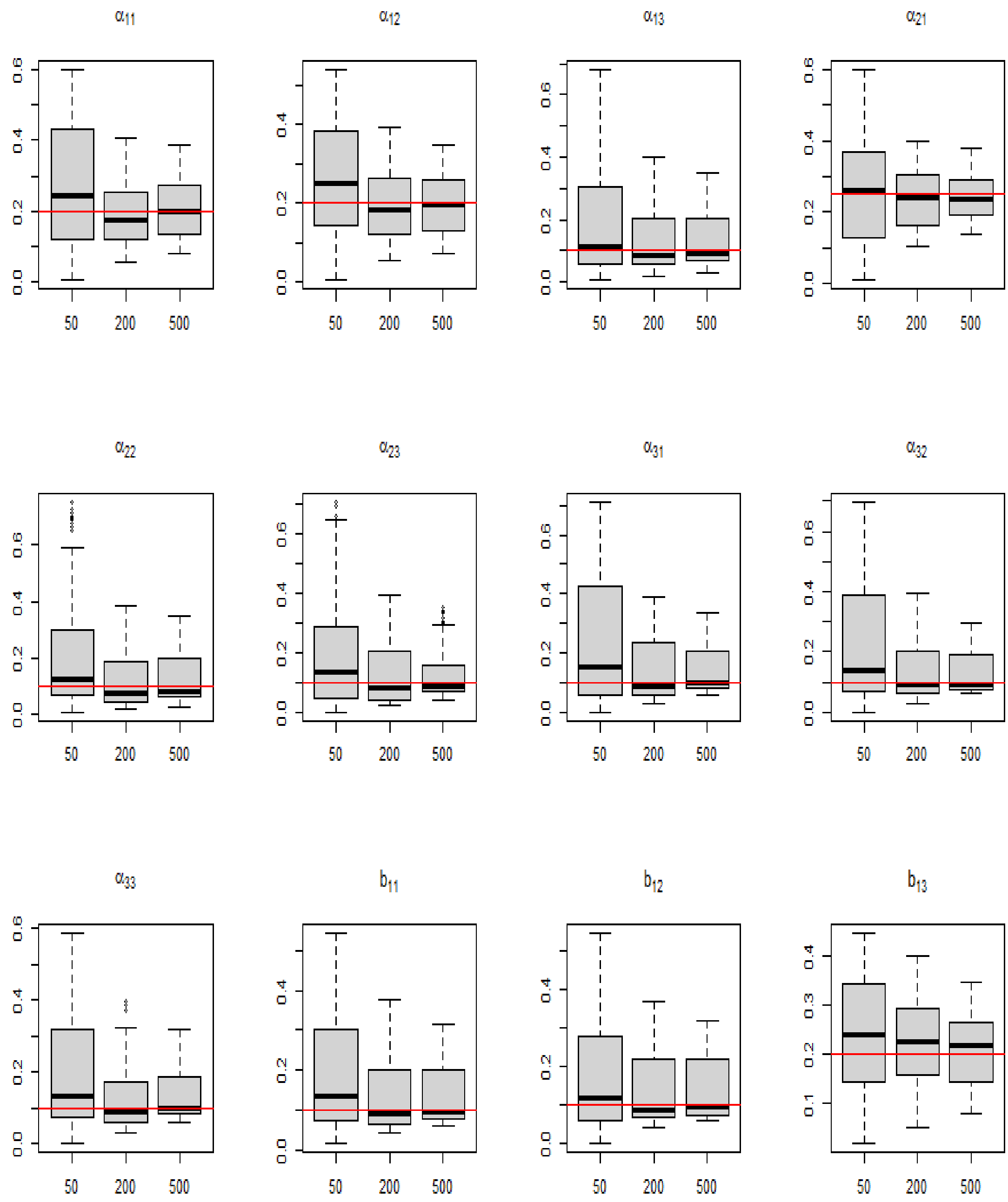


Figure 4.2: Boxplot with results from the simulation experiment with initial parameter values $(a_{ij},b_{ij},d_{ij},\omega_{ij})$=(0.35,0.12,0.05,0.18,0.12,0.15,0.1,0.4, 0.1,0.15,0.15,0.18,0.1,0.1,0.05,0.1,0.1,0.1,1.5,1.5,3,2.8,4,3,1.5)



α_{11}
α_{12}
α_{13}
α_{21}
α_{22}
α_{23}
α_{31}
α_{32}
α_{33}
b_{11}
b_{12}
b_{13}
50
200
500

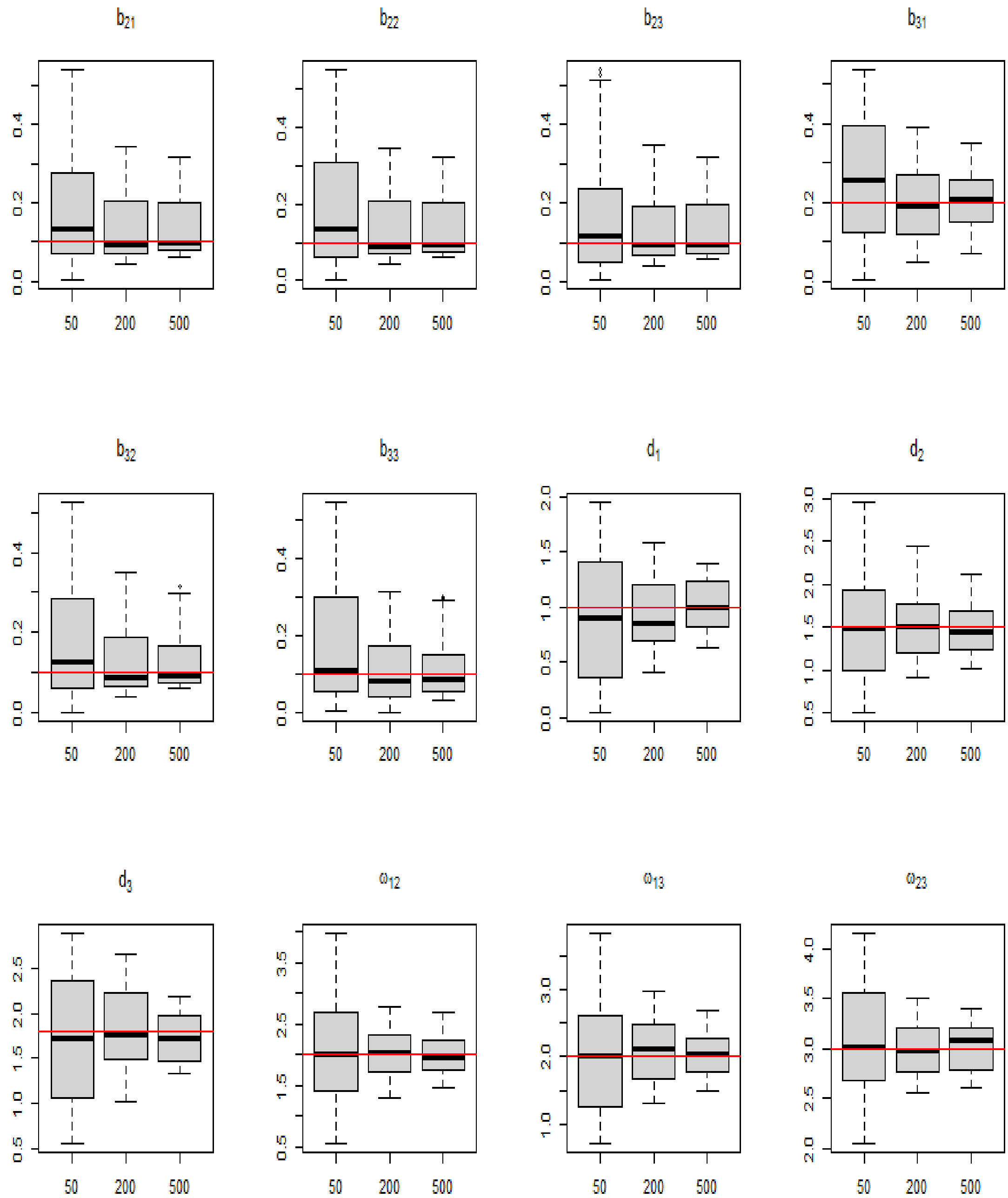
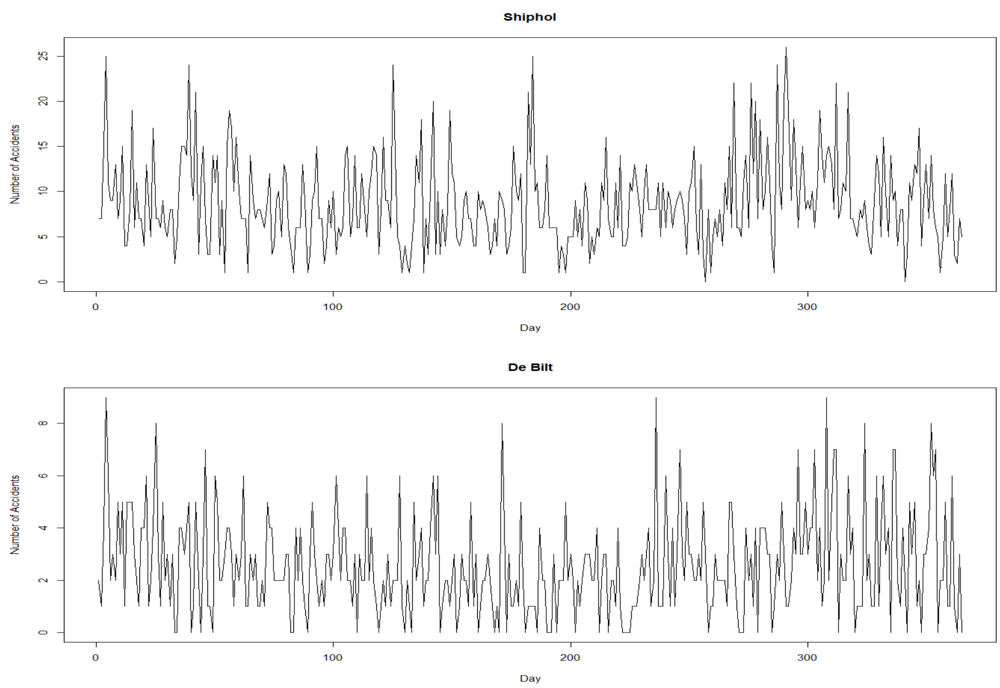


Figure 4.3: Boxplot with results from the simulation experiment with initial parameter values $(a_{ij},b_{ij},d_{ij},\omega_{ij}))$=(0.2,0.2,0.1,0.25,0.1,0.1,0.1,0.1, 0.1,0.1,0.1,0.2,0.1,0.1,0.1,0.2,0.1,0.1,1,1.5,1.8,2,2,3)

4.5 Application

Recent years there is a flourishing interest about data for road accident statistics. A number of approaches have been suggested to model time-series crash count data, most of them do not explicitly take into account the significant autocorrelation that traditionally is present in the data. The ignorance of autocorrelation lead to statistical inconsistency of the coefficient estimates producing standard estimates that are too optimistic. Brijs et al. (2008) proposed a Poisson INAR(1) regression model for modelling daily car crash data with time interdependencies and examine the risk impact of weather conditions on the observed counts. At this application, we use daily crash counts in three areas of Netherlands that is, Shiphol, De Bilt and Soesterberg. The data share some similar environment, especially with respect to weather conditions, road characteristics and traffic exposure. From time series plots presented in (4.4) we could see that the large number of accidents is obtained in weekdays.



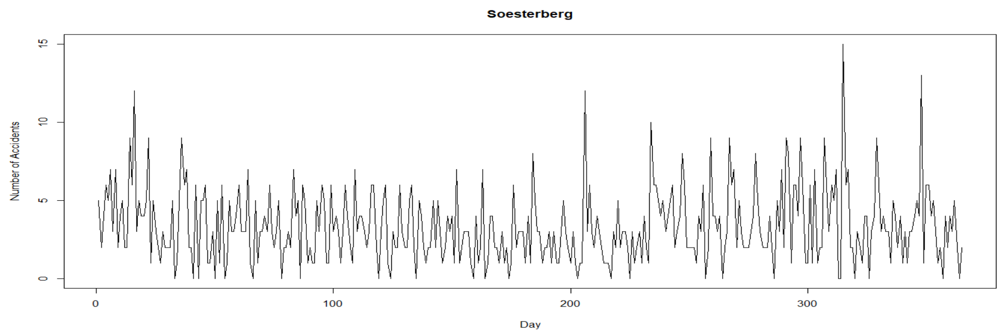


Figure 4.4: Time series plots of the series of accidents in Schiphol, De Bilt and Soesterberg

Pedeli and Karlis (2013) calculated cross-correlation between the three time series providing the emergency study of a multivariate time series process. Autocorrelation for the three series and cross-correlation also presented at (4.5, 4.5). The first two series are auto-correlated at lag=1. Some descriptive statistics are also presented at the table 5.3 . Calculating means and variances we take values for variances larger than those for means with an indication to overdispersion. This implies that Poisson distribution is not appropriate some other distribution such as a Negative binomial distribution is more appropriate for those data. Considering the study of Zhu (2011) about a univariate negative binomial integer-valued GARCH model, a multivariate model need to discussed for modeling those data. Fokianos et al. (2020) based on study of Latour (1997) approaches INGARCH(1,1) model with an ARMA(1,1) model as:

$$(X_t - \mu) - (B + D)(X_{t-1} - \mu) = e_t - De_{t-1}, \tag{4.12}$$

where $\mu = E(Y_t) = (I - B - A)^{-1}d$. Based on the general ARMA(1,1) model $X_t - \phi X_{t-1} = e_t + \theta e_{t-1}$ compared with (4.12) we get that $\phi = A + B$, $\theta = -A$ and $e_t \sim Normal(0, \Sigma)$.

	Mean	Variance
Shiphol Weekdays/Weekends	8.781 9.770/6.298	24.600 2.39/3.06
De Bilt Weekdays/Weekends	2.584 2.893/1.740	3.681 1.28/1.55
Soesterberg Weekdays/Weekends	3.370 3.667/2.625	5.673 1.53 /1.95

Table 4.1: Descriptive statistics for the time series of accidents

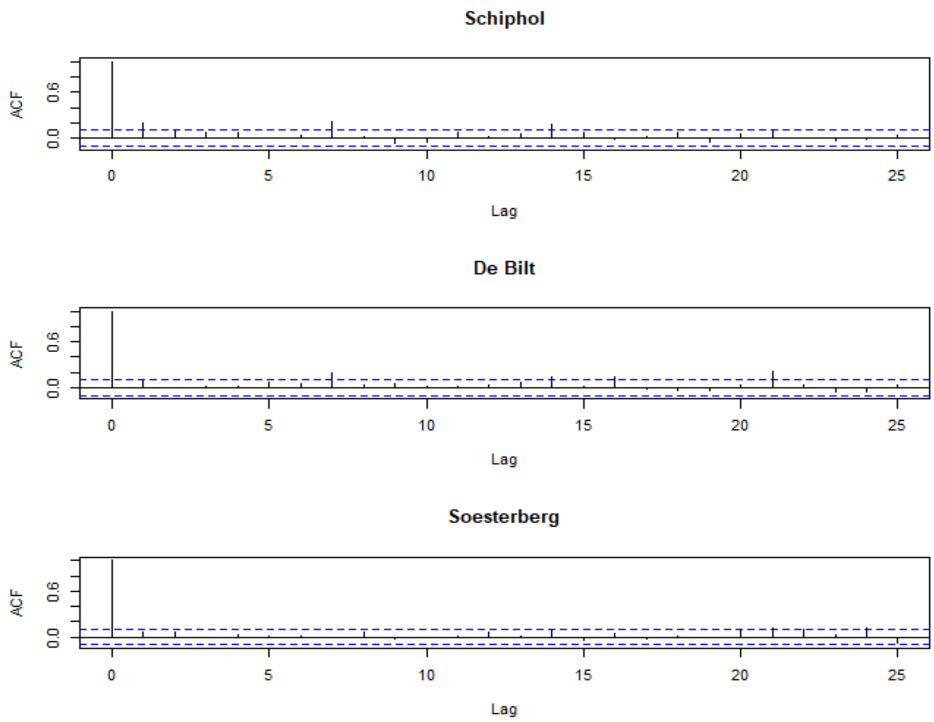


Figure 4.5: ACF for the three time series

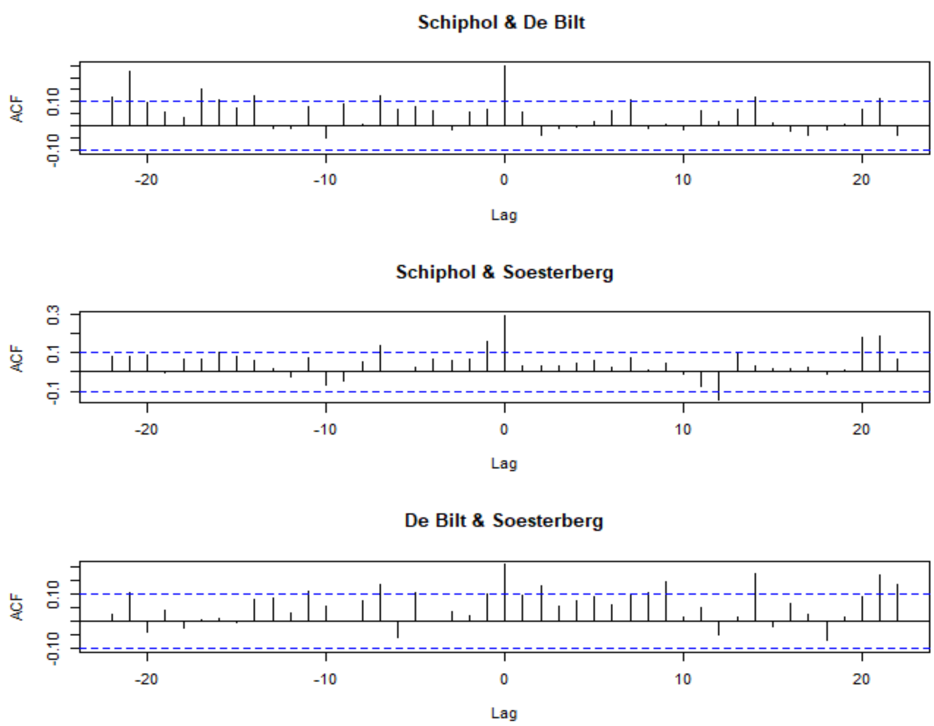


Figure 4.6: Cross-correlation function

We have indication of dependence between the number of crashes at three cities.

One can observe that cross correlation between Schiphol and De Bilt decline faster than the others. On the other hand positive cross correlation is observed between daily crashes at Soesterberg and De Bilt indicating that when there is an increase at number of daily crashes at Soesterberg, there is also an increase at those of De Bilt h days later, where h stands for the time lag.

In our study the parameters of the multivariate Poisson model are calculated by ML method. Furthermore the parameters of the trivariate INGARCH Poisson time series model are calculated by using constrOptim in R and considering parameter vector θ as presented in simulation. As constraints we consider conditions of stationarity for parameters $a_{ij}, b_{ij},\ \ i,j=1,2,3$ and under the assumption that $k_i = 1,\ \ i=1,2,3$, we restrict parameters $\omega_{12}, \omega_{13}, \omega_{23}$ according to (3.22), (3.23).

$\hat{d}_1$	1.002(0.016)	$\hat{b}_{11}$	0.124(0.024)
$\hat{d}_2$	0.810(0.010)	$\hat{b}_{12}$	0.123(0.023)
$\hat{d}_3$	0.814(0.019)	$\hat{b}_{13}$	0.113(0.013)
$\hat{a}_{11}$	0.215(0.015)	$\hat{b}_{21}$	0.084(0.016)
$\hat{a}_{12}$	0.192(0.007)	$\hat{b}_{22}$	0.0115(0.015)
$\hat{a}_{13}$	0.113(0.013)	$\hat{b}_{23}$	0.105(0.014)
$\hat{a}_{21}$	0.128(0.028)	$\hat{b}_{31}$	0.243(0.043)
$\hat{a}_{22}$	0.117(0.017)	$\hat{b}_{32}$	0.084(0.015)
$\hat{a}_{23}$	0.188(0.012)	$\hat{b}_{33}$	0.122(0.022)
$\hat{a}_{31}$	0.299(0.001)	$\hat{\omega}_{12}$	-2.465(0.035)
$\hat{a}_{32}$	0.134(0.034)	$\hat{\omega}_{13}$	-1.448(0.051)
$\hat{a}_{33}$	0.039(0.061)	$\hat{\omega}_{23}$	-1.961(0.039)

Table 4.2: ML estimates for trivariate Poisson model parameters based on volatility construction defined by (1.14; 4.3). Numbers in parentheses are standard errors.

4.5.1 Goodness of fit

We consider Pearson residuals defined by $e_{i,t} = \frac{Y_{i,t}-\lambda_{i,t}}{\sqrt{\lambda_{i,t}}}, i=1,2$ in order to examine model fit. Assuming that those residuals could be expressed as:

$$e_t = \sum_{j=1}^{n/2} \beta_1\,(j/n)\,cos(2\pi\omega_j t) + \beta_2\,(j/n)\,sin(2\pi\omega_j t),$$

and calculating periodograms observations with high-valued frequencies were not observed. This means that there are not extreme values at Pearson residuals and equivalently there are no divergences between estimators and real values. Based on theory under the correct model residuals come from a white noise with constant variance and cumulative periodograms presented at 4.5.1 are smooth. In our study for calculating $\hat{e}_{i,t}$, we consider $\lambda_{i,t}$ as $\lambda_{i,t}(\hat{\theta})$ where $\hat{\theta} = (\hat{\omega}_{ij}, \hat{a}_{ij}, \hat{b}_{ij})$. In graphs below we present cumulative periodograms for trivariate INGARCH model.

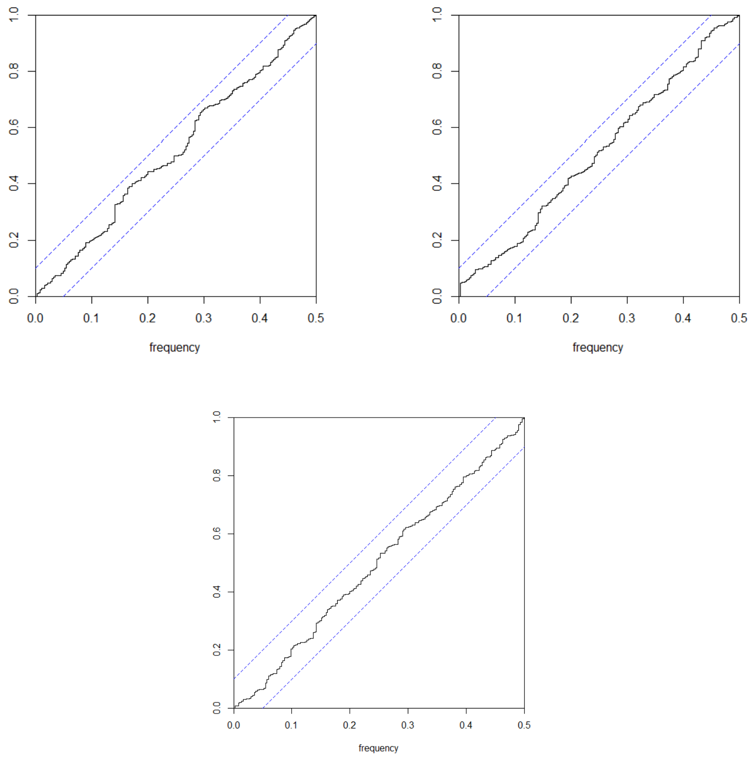


Figure 4.7: Cumulative periodogram plots of the Pearson residuals for daily crash counts for Shiphol(left upper), De Bilt(right upper) and Soesterberg.

4.5.2 Predictive marginals

We make prediction based on method of conditional expectation. Based on eqn (4.4) an one step ahead forecasting is:

$$\hat{\lambda}_{t+1} = E(\lambda_{t+1} \mid \mathcal{F}) = \mathbf{d} + \mathbf{A}\lambda_t + \mathbf{B}X_t \tag{4.13}$$

and using recursion we have:

$$\hat{\lambda}_{t+1} = \sum_{i=0}^{k-1} \left(\mathbf{A}^i \mathbf{d} + \mathbf{A}^i \mathbf{B} X_{t-i}\right) + \mathbf{A}^k \lambda_{t-k-1}.$$

Considering recursive formula a p-steps ahead forecasting can be calculated by:

$$\hat{\lambda}_{t+p} = E\left(\lambda_{t+p} \mid \mathcal{F}^X\right) = \sum_{i=0}^{k-1} A^i d + \sum_{i=0}^{k-1} \left(BA^i X_{t+p-i}\right) + A^k E\left(\lambda_{t+p-k} \mid \mathcal{F}\right).$$

Consecutively calculating $\lambda_{t+p}, \ ,p = 1, 2, ..., n$ we choose every time this λ_t with the greatest probability. Based on (3.4) we predict λ at time $t = 366$ and then we calculate bivariate marginal predictive pmfs given by:

$$p(x_i, x_j) = p(x_i)p(x_j)\left[1 + \omega_{ij}\left(c_i^{(-k_i-1)} - e^{\lambda_i(c_i^{-k_i}-1)}\right)\left(c_j^{(-k_j-1)} - e^{\lambda_j(c_j^{-k_j}-1)}\right)\right].$$

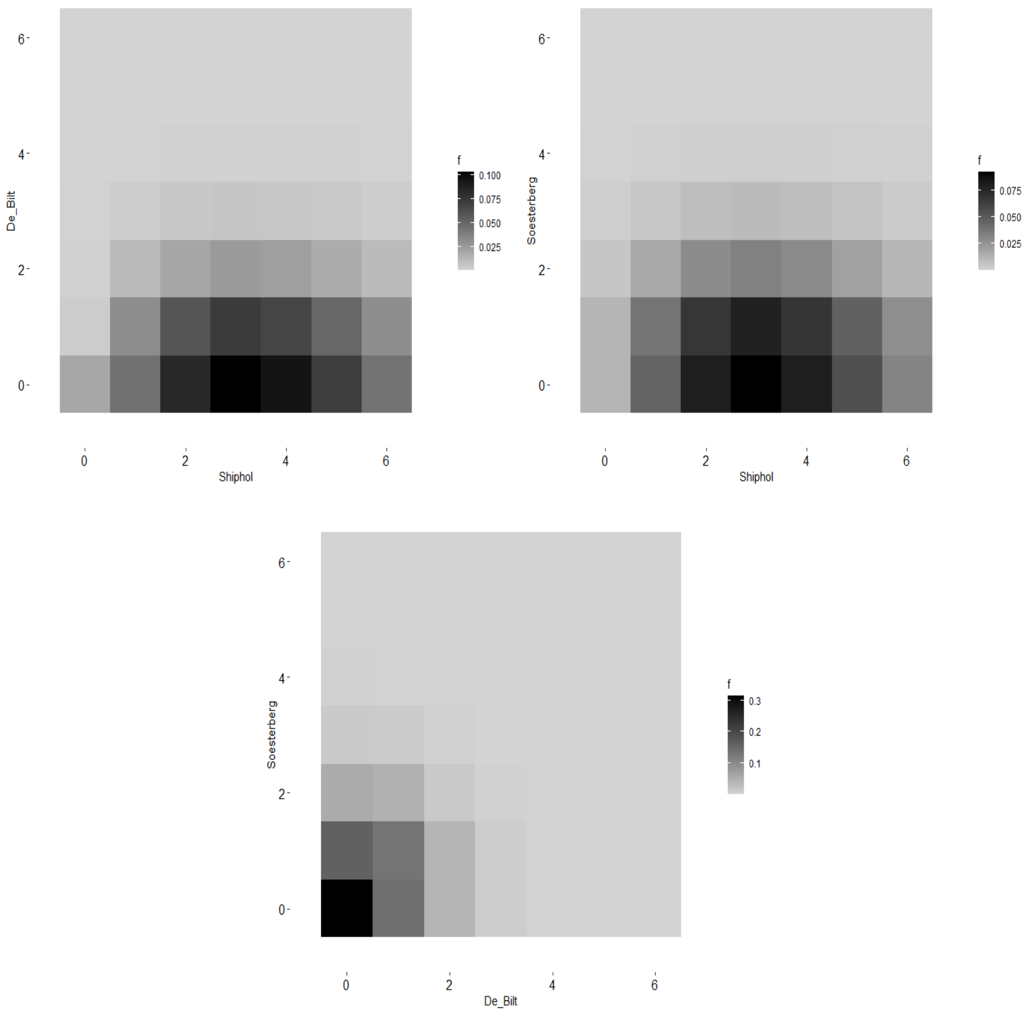


Figure 4.8: One step ahead bivariate predictive probability for each one of the three places in Netherlands.

Moreover we compare our model with this studied by Fokianos et al. (2020). At this illustration working with the same data consists of number of transactions per 15 seconds for the stocks Coca-Cola(KO) and IBM on September 19^{th} 2005, a likelihood ratio test is provided and graphical inspection for residuals for both models is discussed. On the same way we keep 1440 where transactions before 9:30 and after 16:30 are not included. Figure (4.9) shows time series plots and (4.10) depicts the autocorrelation and cross-correlation functions. Autocorrelation functions plot indicates high correlated transactions data within and between the individual transaction series.

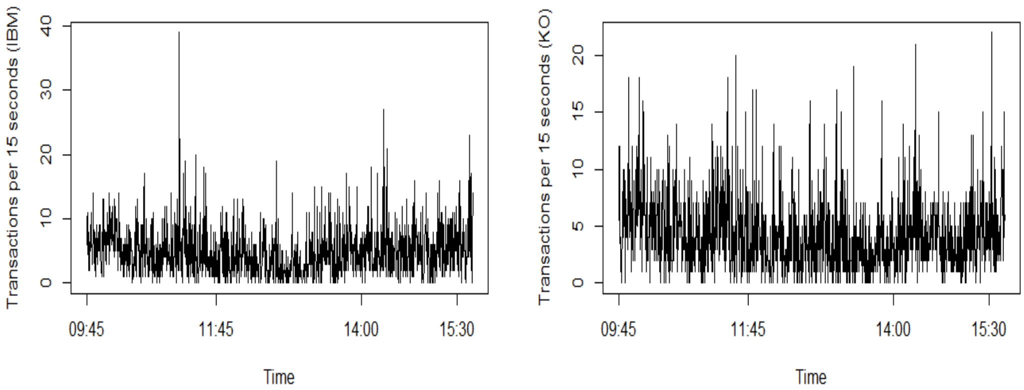


Figure 4.9: Number of transactions per 15 seconds for IBM (top) and Coca-Cola (bottom).

There is evidence that our proposed model fit better the data providing a likelihood ratio test where log-likelihoods are -2519.479 and -6884.62 respectively. Moreover we calculate Pearson residuals as $\hat{e}_{i,t} = \frac{X_{i,t}-\hat{X}_{i,t}}{\sqrt{\hat{X}_{i,t}}}$, where $\hat{X}_{i,t} = \lambda_{i,t}(\hat{\theta})$ and $\hat{\theta}$ is parameters estimations vector. Based on the graph below (4.11) in model 2, which is the INGARCH model constructed based on Sarmanov distribution discussed in chapter 3, we have a more consistent spread implying a less substantial dispersion.

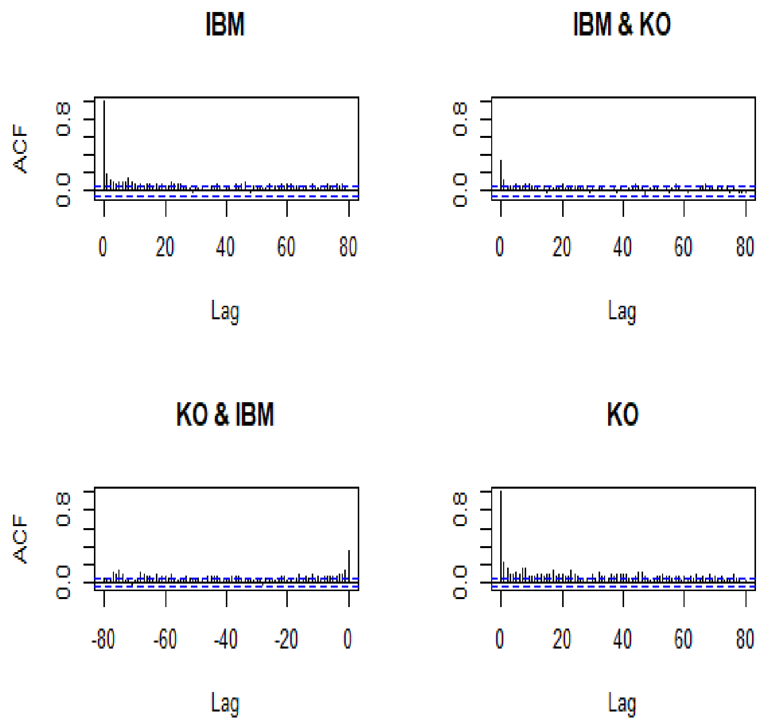


Figure 4.10: Auto- and cross-correlation function of the transaction data.

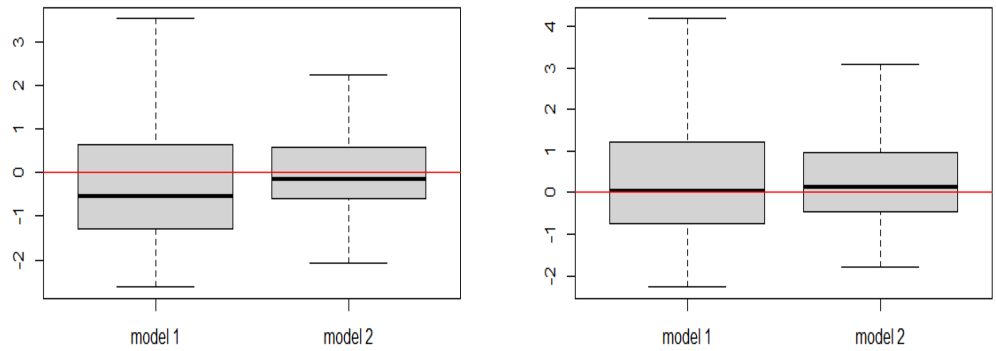


Figure 4.11: Boxplot of residuals $\hat{e}_{1,t}$ for IBM(left) and $\hat{e}_{2,t}$ for KO(right).

At this chapter, considering multivariate distribution discussed at chapter 3, a new multivariate INGARCH time series model has been provided. The keypoint is the construction of a flexible Poisson distribution which provide many ways for parameters estimation. The contribution to the almost existed models is that we provide a flexible INGARCH(1,1) model with contemporaneous correlation assuming that conditional expectation depends on past values. In our study volatilities are calculated considering interdependencies. Parameters estimation was provided by Maximum Likelihood approach but many other methods like likelihood by parts could be applied in order to simplify and to decrease computational cost. A brief discussion of likelihood by parts method is also provided. A dataset with daily crash counts has been studied without being totally appropriate to study with a Poisson INGARCH model. Moreover in order to provide the efficacy of the copula proposed in chapter 3 in case of continuous variable and using the new range of correlation coefficient, a CARR model is also studied in chapter 5.

Chapter 5

A bivariate CARR(1,1) model

5.1 The model

The study of interactions between two or more assets in finance seems to be of particular importance. Assuming that X_t is a variable for daily stock prices, Fernandes et al. (2005) studied a CARR model where R_t is Exponentially distributed range R_t between daily maximum and minimum stock prices. At the proposed trivariate CARR model volatilities expressed linearly with previous volatilities and ranges of days up to $t-1$. The above structure is capable to capture covariates and spill-over effects but many drawbacks lead to more extensively study of multivariate CARR models. The first one is the restrictive interval of correlation coefficient provided by the srtucture of multivariate exponential distribution. The second one is that no implementations have been provided yet considering the lack of intradaily data and for that reason any parameter estimation method does not have been discussed. Furthermore ergodic properties have been studied using criteria developed by Tweedie (1976) and β-mixing property but other methods assuming weak dependence criteria discussed by Doukhan et al. (2012) could be also yielded for simplicity. The main shortcoming is that while ergodic properties and estimation approaches are not difficult to study theoretically, it is hard to implement data under the assumption of Exponential distribution. In our study we contribute at CARR models by constructing a bivariate conditional autoregressive range process based on bivariate exponential distribution defined by (5.10). A maximum likelihood approach is discussed and convergence of maximum likelihood estimations is also provided. Numerous efforts for strengthening goodness of fit criteria at heteroskedastic models have been done by Meintanis et al. (2020) and Jiménez-Gamero et al. (2020). Based on those, a brief discussion presented in our study and results of simulation are . A review for range volatilities of univariate and multivariate conditional stochastic

processes given by Chou et al. (2015).
Assume that $R_t = \lambda_t \epsilon_t$ where ϵ_t is a sequence of positive independent and identical exponentially distributed random variables with mean equal to 1 and λ_t equal to conditional expectation of $R_t = max(X_t) - min(X_t)$ up to time t, where X_t is a continuous variable usually express the price. The proposed bivariate exponential process is defined as:

$$R_t \mid \mathcal{F}_{t-1}^{R,\lambda} \sim BExp(\lambda_t).$$

In our study we assume that each conditional expectation expressed only on its past ignoring the interactions between the two processes and volatility spill-over effects. In other words:

$$\begin{bmatrix} R_{1t} \\ R_{2t} \end{bmatrix} = \begin{bmatrix} \lambda_{1t} & 0 \\ 0 & \lambda_{2t} \end{bmatrix} \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}, \tag{5.1}$$

where

$$\lambda_{jt} = d_j + a_{j1}\lambda_{jt-1} + b_{j1}R_{jt-1}, \;\; j = 1, 2. \tag{5.2}$$

Furthermore based on (5.2) and under the assumption of exponential distribution on residuals ϵ_t conditional first and second order moments are given by:

$$E[R_{jt} \mid \mathcal{F}_{t-1}] = \lambda_{jt}, \;\; Var[R_{jt} \mid \mathcal{F}_{t-1}] = \lambda_{jt}^2.$$

Unconditional moments of our process are defined based on the idea of Engle and Russell (1998) for martingale difference. Considering a martingale difference $\eta_{jt} = R_{jt} - \lambda_{jt}, \;\; j = 1, 2$, where λ_{jt} defined by (5.2), a new expression for the calculation of unconditional moments is given by:

$$R_{jt} = d_j + (a_{j1} + b_{j1})R_{jt-1} - a_{j1}\eta_{jt-1} + \eta_{jt},$$

where

$$E(\eta_{jt}) = 0 \text{ and } Var(\eta_{jt}) = E(\eta_{jt}^2).$$

For the calculation of first and second order moments we assume conditions of weak stationarity. Assuming that the process is first order stationarity we have that

$$E(\lambda_{jt} \mid \mathcal{F}_{t-1}) = E(\lambda_{jt-1} \mid \mathcal{F}_{t-1}) \text{ and } E[R_{jt}] = \frac{d_j}{1-\sum_{i=0}^{max(p,q)}}(a_{ji} + b_{ji})$$

and

$$Var[R_{jt}] = \frac{E[R_{jt}](1-(a_j+b_j)^2+a_j^2)}{1-(a_j+b_j)^2},$$

in case of an ECARR(p,q) model. When we have a ECARR(1,1) model then $\mu_j = E(R_{jt}) = \frac{d_j}{1-(a_{1j}+b_{1j})}$. Based on the above equation for the range $R_{jt},\ j=1,2$ we have that:

$$\begin{aligned}
(R_{jt}-\mu_j)-(a_{j1}+b_{j1})(R_{jt-1}-\mu_j) &= \eta_{jt}-a_{j1}\eta_{jt-1}\\
(R_{jt}-\mu_j)R_{jt-1}-(a_{j1}+b_{j1})(R_{jt-1}-\mu_j)R_{jt-1} &= \eta_{jt}R_{jt-1}-a_{j1}\eta_{jt-1}R_{jt-1}\\
(R_{jt}-\mu_j)R_{jt}-(a_{j1}+b_{j1})(R_{jt-1}-\mu_j)R_{jt} &= \eta_{jt}R_{jt}-a_{j1}\eta_{jt-1}R_{jt}
\end{aligned} \tag{5.3}$$

It is easy to verify that

$$\begin{aligned}
E(\eta_{jt-1}R_{jt}) &= E(\eta_{jt-1}(d_j+(a_{j1}+b_{j1})R_{jt-1}+\eta_{jt}-a_{j1}\eta_{jt-1})) = b_{j1}\sigma^2_{\eta_j},\\
E(\eta_{jt}R_{jt}) &= E(\eta_{jt}(d_j+(a_{j1}+b_{j1})R_{jt-1}+\eta_{jt}-a_{j1}\eta_{jt-1})) = \sigma^2_{\eta_j},
\end{aligned} \tag{5.4}$$

Combining (5.3) and (5.4) we have:

$$\begin{aligned}
Var(R_{jt}) &= (a_{j1}+b_{j1})\gamma_{R_{jt}}(1)+\sigma^2_{\eta_j}-a_{j1}b_{j1}\sigma^2_{\eta_j},\\
\gamma_{R_{jt}}(1) &= (a_{j1}+b_{j1})Var(R_{jt})-a_{jt}\sigma^2_{jt},
\end{aligned} \tag{5.5}$$

or

$$Var(R_{jt}) = \frac{E[R_{jt}](1-(a_{j1}+b_{j1})^2+a^2_{j1})}{1-(a_{j1}+b_{j1})^2},$$

$$Cov(R_{jt}R_{jt-1}) = \frac{b_{j1}\left(1-a_{j1}(a_{j1}+b_{j1})\right)E(R_{jt})}{\left(1-(a_{j1}+b_{j1})\right)^2},$$

and

$$Cor(R_{jt}R_{jt-1}) = \frac{b_{j1}\left(1-a_{j1}(a_{j1}+b_{j1})\right)}{(1-2a_{j1}b_{j1}-a^2_{j1})}.$$

5.1.1 Ergodic properties

Fernandes et al. (2005) proposed a trivariate model where each expected value expressed not only on its previous values but on the previous volatilities of other models in order to capture the interactions between the variance. In other words the model given by:

$$R_t = diag(\lambda_t)\epsilon_t,$$

where $R_t = (R_{1t}, R_{2t}, R_{3t})$, $\lambda_t = (\lambda_{1t}, \lambda_{2t}, \lambda_{3t})$ and $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t})$.

Carrasco and Chen (2002) studied stationarity and properties of ergodicity based on Tweedie criteria. Fernandes et al. (2005) studied under which conditions the process is strictly stationary and β-mixing with exponential decay after expressing the model defined above as a trivariate autoregressive model of order 1(MAR(1)) given by:

$$X_{t+1} = C(\epsilon_t)X_t + D(\epsilon_t),$$

where $X_t = \left(R_t', \ \lambda_t'\right)$, $D(\epsilon_t) = (d_1\epsilon_{1t}, d_2\epsilon_{2t}, d_3\epsilon_{3t}, d_1, d_2, d_3)$ and

$$C(\epsilon_t)=\begin{bmatrix} b_{11}\epsilon_{1t} & b_{12}\epsilon_{2t} & b_{13}\epsilon_{3t} & a_{11}\epsilon_{1t} & a_{12}\epsilon_{2t} & a_{13}\epsilon_{3t} \\ b_{21}\epsilon_{1t} & b_{22}\epsilon_{2t} & b_{23}\epsilon_{3t} & a_{21}\epsilon_{1t} & a_{22}\epsilon_{2t} & a_{23}\epsilon_{3t} \\ b_{31}\epsilon_{1t} & b_{32}\epsilon_{2t} & b_{33}\epsilon_{3t} & a_{31}\epsilon_{1t} & a_{32}\epsilon_{2t} & a_{33}\epsilon_{3t} \\ b_{11} & b_{12} & b_{13} & a_{11} & a_{12} & a_{13} \\ b_{21} & b_{22} & b_{23} & a_{21} & a_{22} & a_{23} \\ b_{31} & b_{32} & b_{33} & a_{31} & a_{32} & a_{33}, \end{bmatrix}$$

Dedecker et al. (2007) proposed numerous weak dependence criteria appropriate for the examination processes ergodicity. For more information someone can study the book of Dedecker et al. (2007). The main difference between Tweedie criteria and weak dependence criteria is that in the first case we have strong conditions of stationarity for all moments in contrast with the second case of weak stationarity of the first two moments. On the same way to previous chapter based on the work that we have also provided for INGARCH(1,1) model and considering study of Franke (2010), ergodic properties of a CARR(1,1) model are also provided. Assuming a Lipschitz function $h : \mathbb{R}^m \to \mathbb{R}^n$, we prove that the process is ergodic. At those models where the conditional expectation expressed linearly, criteria of weak dependence for the proof of ergodicity are less complicated in the study. On the other hand when the expression is not linear other methods must be used for example see Fokianos et al. (2009).

Theorem 5. *If $||A||_1+||B||_1 \leq 1$ then the process $\{R_t\}$ defined by (5.1) is stationary and θ-weak dependent.*

Based on Lemma 5.2(p 115) Dedecker et al. (2007) we have that $||R - R^*||_1 = \tau_1(F^R_{t-1}, R)$, where τ-weak dependence criterion has been discussed in chapter 2. Furthermore if the process R_t is τ-weak dependent then satisfies the criterion of θ-weak dependence. The proof is similar to the proposed in trivariate INGARCH Poisson process. In addition considering that parameter λ_t is the conditional expectation of the Exponential process all parameters $d_j, a_{ij}, b_{ij}, j = 1, 2$ are positive.

5.2 Ergodic properties

Conditions of stationarity embedded on conditions of ergodicity. Carrasco and Chen (2002) studied stationarity and properties of ergodicity based on Tweedie criteria. Fernandes et al. (2005) studied appropriate conditions that the process is strictly stationary and β-mixing with exponential decay.

Assuming the model defined by (5.1) and considering that mixing functions defined as (3.26) for two exponentially distributed random variables with parameters $\frac{1}{\lambda_{it}}, i = 1, 2$ we have:

$$f(r_{1t}, r_{2t}) = \left[\prod_{i=1}^{2} \frac{1}{\lambda_{it}} e^{-\frac{R_{it}}{\lambda_{it}}}\right] q(r_{1t}, r_{2t}), \tag{5.6}$$

where

$$q(r_{1t}, r_{2t}) = 1 + \omega \prod_{i=1}^{2} \left(c_i^{-k_i r_{it}} - \frac{1}{1 + \lambda_{it} k_i \log c_i} \right) \tag{5.7}$$

and

$$\lambda_{jt} = d_j + a_j \lambda_{j,t-1} + b_j R_{j,t-1}, \;\; j = 1, 2. \tag{5.8}$$

Log-likelihood function is given by:

$$l(\theta; r) = -\sum_{i=1}^{2}\sum_{t=1}^{n} \log \lambda_{it} - \sum_{i=1}^{2}\sum_{t=1}^{n} \frac{R_{it}}{\lambda_{it}} + \sum_{t=1}^{n} \log(q(r_{1t}, r_{2t})). \tag{5.9}$$

Score function is

$$S_n(\theta) = \frac{\partial \ell}{\partial \theta},$$

$$\frac{1}{\sqrt{n}} S_n(\theta) \to N(0, I(\theta_0)).$$

Theorem 6. *For the CARR model as $n \to \infty$*

$$\sqrt{n}(\hat{\theta} - I(\theta_0)^{-1}),$$

where $I(\theta_0)^{-1}$=-$E\left(\frac{\partial^2 l(\theta_0)}{\partial\theta\partial\theta^T}\right)$.

First and second order partial derivatives are calculated in order to prove consistency of MLE. Assuming parameter vector $\theta = (d_j, b_j, \omega)$, $j = 1, 2$ and based on condition of stationarity and boundedness of parameter ω, $E_{sup_{\theta\in\Theta}}||\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta^T}||_1 \leq \infty$. Proof are given in Appendix. Moreover using Taylor series expansion for $\frac{\partial l(\theta)}{\partial\theta}$ we have:

$$0 = \frac{\partial l(\hat{\theta})}{\partial\theta} = \frac{\partial l(\theta_0)}{\partial\theta} + (\hat{\theta} - \theta_0)\frac{\partial^2 l(\tilde{\theta})}{\partial\theta\partial\theta^T} \qquad \text{or} \qquad \frac{1}{\sqrt{n}}\frac{\partial l(\theta_0)}{\partial\theta} = \sqrt{n}(\hat{\theta} - \theta_0)\left(-\frac{1}{n}\frac{\partial^2 l(\theta)}{\partial\hat{\theta}\partial\theta^T}\right).$$

Considering Lemma 6 of Ting Lee (1996) $-\frac{1}{n}\frac{\partial^2 l(\theta_n^*)}{\partial\theta\partial\theta^T} \to I(\theta_0)$ and by using martingale central limit theorem $\frac{1}{\sqrt{n}}\frac{\partial l(\theta_0)}{\partial\theta} \to N(0, I(\theta_0))$. Based on the above $\sqrt{n}(\hat{\theta} - \theta_0) \to N(0, I(\theta_0)^{-1})$

5.3 Estimation

Taking into account (3.7) and considering that $q_i(r_i)$ defined as (3.26) for two exponentially distributed random variables with parameters $\frac{1}{\lambda_{jt}}, j = 1, 2$, bivariate Exponential distribution defined as:

$$f(r_{1t}, r_{2t}) = \prod_{i=1}^{2}\frac{1}{\lambda_{it}}e^{-\frac{R_{it}}{\lambda_{it}}}q(r_{1t}, r_{2t}), \tag{5.10}$$

where

$$q(r_{1t}, r_{2t}) = 1 + \omega\prod_{i=1}^{2}\left(c_i^{-k_i r_{it}} - \frac{1}{1 + \lambda_{it}k_i\log c_i}\right). \tag{5.11}$$

The score function is given by:

$$l(\theta; r) = -\sum_{i=1}^{2}\sum_{t=1}^{n}\log\lambda_{it} - \sum_{i=1}^{2}\sum_{t=1}^{n}\frac{R_{it}}{\lambda_{it}} + \sum_{t=1}^{n}\log(q(r_{1t}, r_{2t})). \tag{5.12}$$

Partial derivatives are given in the Appendix. Considering conditions of stationarity given by Theorem 1 and bounds for parameter ω given by (3.30; 3.31) assumptions for estimators consistency are satisfied. The score function is given by

$$S_n(\theta) = \sum_{i=1}^{2}\sum_{t=1}^{n}\left(\frac{R_{it}}{\lambda_{it}^2} - \frac{1}{\lambda_{it}}\right)\frac{\partial\lambda_{it}}{\partial\theta} + \sum_{t=1}^{n}\frac{1}{q(r_{1t}, r_{2t})}\frac{\partial q(r_{1t}, r_{2t})}{\partial\theta}, \tag{5.13}$$

where $\hat{\theta} = \arg\max_\theta l(\theta)$.

5.4 Goodness of fit test

Consider observations $r_{it},\ i = 1, 2,\ t = 1, \ldots, n$ and under the assumption that residuals ϵ_{it} come from a unit exponential distribution we test the null hypothesis that bivariate distribution is a bivariate exponential distribution defined by (5.10) and (5.11).Jiménez-Gamero et al. (2020) proposed a GOF test for residuals considering the classical GARCH model. The idea is based on characterization that $\mathbb{E}(X \mid Y) = 0$ holds if and only if $\mathbb{E}(Xe^{iuy}) = 0$. Meintanis et al. (2020) proposed GOF tests for the class of Autoregressive Conditional Duration(ACD) models in univariate case. Considering that residuals $\epsilon_t = x_t/\mathbb{E}(x_t \mid \mathcal{F}_t)$ are unobservable a test for the null hypothesis that residuals come from a unit exponential distribution has been examined. At this study there is a brief discussion about a Goodness of fit test for a bivariate Stochastic Conditional duration(SCD) model proposed by Bauwens and Veredas (2004). Meintanis et al. (2020) for the bivariate SCD model considered the characteristic function of the bivariate distribution and defined the distance statistic measure as:

$$K(u_1, u_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\hat{K}_T(u_1, u_2) - K_F(u_1, u_2)|^2 w(u_1, u_2) du_1 du_2.$$

Furthermore in a similar way Hudecová et al. (2021) discussed goodness of fit tests for a class of bivariate time series of counts considering the L2-type distance between two estimators of the probability generating function one nonparametric and one semiparametric under the same null hypothesis as an extension of the study in univariate case proposed by Meintanis and Karlis (2014). At all of the above test statistics the importance is concentrated in the asymptotic behavior of the statistic. At the proposed model where bivariate distribution defined by (5.10) the statistic based on the moment generating function defined as :

$$S = \int_0^{\infty}\int_0^{\infty} |\hat{M}_T(u_1, u_2) - M_F(u_1, u_2)|^2 w(u_1, u_2) du_1 du_2,$$

where

- $M_F(u_1, u_2) = \prod_{i=1}^{2} \frac{1}{\lambda_i(u_i - 1/\lambda_i)} + \omega \left(\frac{\lambda_i}{1 - \lambda_i u_i + \lambda_i k_i \log c_i} - \frac{1}{\lambda_i + k_i \log c_i} \frac{1}{\lambda_i u_i - 1} \right).$
- $\hat{M}_t(u_1, u_2) = \frac{1}{n} \sum_{t=1}^{n} e^{u_1 \hat{\epsilon}_{1t} + u_2 \hat{\epsilon}_{2t}}.$
- $w(u_1, u_2)$ is a weight function satisfying $0 \leq \int \int w(u_1, u_2) du_1 du_2 \leq 1.$

The asymptotic distribution of the test statistic S depends on several unknown quantities. For this reason we use a parametric bootstrap approach.

1. Generate $\epsilon^{\star}_{it}$, i = 1,2, where ϵ_{it} are iid and follow Exp(1).
2. Compute pseudo-observations $r^{\star}_{it}, i = 1, 2, t = 1, \dots, n$ from (5.1) where ϵ_t replaced by $\epsilon^{\star}$ and λ_t by $\hat{\lambda}_t$ where all parameters a_{ii} and b_{ii} replaced by their estimators.
3. Fit the model (5.1) using $r^{\star}_{it}, i = 1, 2, t = 1, \dots, n$ and compute the bootstrap estimator $\lambda^{\star}_t$ of λ_t.
4. Compute the corresponding test statistic $S^{\star}$.
5. Repeat steps 1-4 several times, say B, and obtain the sequence of test statistics, $S^{\star}_1, S^{\star}_2, \dots, S^{\star}_B$.
6. Compute the p-value as $p = \frac{I(S^{\star}-S)}{B}$ where I=$\begin{cases} 1, S^{\star} > S \\ 0, Otherwise \end{cases}$.

Asymptotic behavior of the test statistic implies asymptotic behavior of the MLE estimator as $Y_t, t \geq 0$ is strictly stationary, estimators converge to a Normal distribution given by theorem 4 and parametric space Θ is a compact set. Partial derivatives with respect to all parameters are presented below.

- $\frac{\partial l_1(\theta_1;r)}{\partial d_j} = \sum_{i=1}^{n} \left(-\lambda_{jt}^{-1} + R_{jt}\lambda_{jt}^{-2}\right) \frac{\partial \lambda_{jt}}{\partial d_j}$
- $\frac{\partial l_2(\theta_1,\theta_2;r)}{\partial d_2} = \sum_{t=1}^{n} \frac{1}{q(r_{1t}r_{2t})}\omega \left(c_1^{-k_1 r_{1t}} - \frac{1}{1+\lambda_{1t}k_1 \log(c_1)}\right) \frac{k_2 \log c_2}{(1+\lambda_2 k_2 \log c_2)^2} \frac{\partial \lambda_{1t}}{\partial d_2}$
- $\frac{\partial l_2(\theta_1,\theta_2;r)}{\partial k_2} = -\sum_{t=1}^{n} \frac{1}{q(r_{1t},r_{2t})} \left(a_1^{-k_1 r_{1t}} - \frac{1}{1+\lambda_{1t}k_1 \log(c_1)}\right) (c_2^{-k_2 r_{2t}} \log c_2 r_{2t})$
- $\frac{\partial \lambda_{jt}}{\partial a_j} \sum_{t=1}^{n-1} t a_j^{t-1} d_j + T a_j^{n-1} \lambda_{jt-T} + b_j \sum_{t=1}^{n-1} t a_j^{t-1} R_{jt-t-1}$
- $\frac{\partial \lambda_{jt}}{\partial b_j} = \sum_{t=1}^{n-1} a_j^t R_{jt-t-1}$
- $\frac{\partial \lambda_{jt}}{\partial d_j} = \sum_{t=0}^{n-1} a_j^t$

5.5 Simulations

We obtain starting values of $l(\theta; r)$ by considering a bivariate ARMA(1,1) model similar to this defined in chapter 2 for the calculation of the unconditional moments see Ferland et al. (2006). Considering (3.27), we define initial value for parameter

ω after obtaining initial values of other parameters and calculating volatilities. For data generation:

1. Choose initial values for the parameter vector $\theta = c(a_{ij}, b_{ij}, \omega)$,

2. Generate a value for the univariate $Exp(\lambda_{1t})$, where λ_{1t} defined by (5.2),

3. Generate a $u \sim U(0, 1)$,

4. If this $u \leq f(r_2 \mid r_1)$ then let $r_2 = u$.

sample size(n)	50	200	500	50	200	500	50	200	500
initial value		mean			sd			MADE	
d_1=2.000	1.933	1.981	1.958	0.545	0.234	0.195	0.711	0.349	0.223
d_2=1.500	1.578	1.479	1.516	0.394	0.189	0.123	0.357	0.243	0.146
a_{11}=0.300	0.378	0.319	0.294	0.209	0.118	0.087	0.256	0.157	0.102
a_{22}=0.450	0.386	0.391	0.398	0.134	0.062	0.048	0.137	0.081	0.062
b_{11}=0.200	0.204	0.199	0.201	0.119	0.056	0.045	0.144	0.073	0.063
b_{22}=0.150	0.181	0.166	0.152	0.135	0.088	0.066	0.118	0.086	0.075
ω=1.800	1.992	1.767	1.827	0.919	0.553	0.317	1.054	0.537	0.402

Table 5.1: Mean, sd and MADE values after 200 replications of simulation experiment.

sample size(n)	50	200	500	50	200	500	50	200	500
initial value		mean			sd			MADE	
d_1=3.000	3.035	2.989	2.995	0.450	0.274	0.239	0.518	0.313	0.299
d_2=0.500	0.461	0.503	0.509	0.211	0.103	0.079	0.196	0.123	0.100
a_{11}=0.400	0.468	0.414	0.404	0.205	0.094	0.051	0.202	0.120	0.060
a_{22}=0.600	0.538	0.608	0.617	0.248	0.100	0.077	0.232	0.126	0.102
b_{11}=0.500	0.483	0.476	0.490	0.223	0.093	0.069	0.252	0.112	0.087
b_{22}=0.300	0.362	0.286	0.272	0.225	0.115	0.081	0.277	0.131	0.087
ω=-0.200	-0.100	-0.154	-0.172	0.518	0.313	0.257	0.655	0.316	0.260

Table 5.2: Mean, sd and MADE values after 200 replications of simulation experiment.

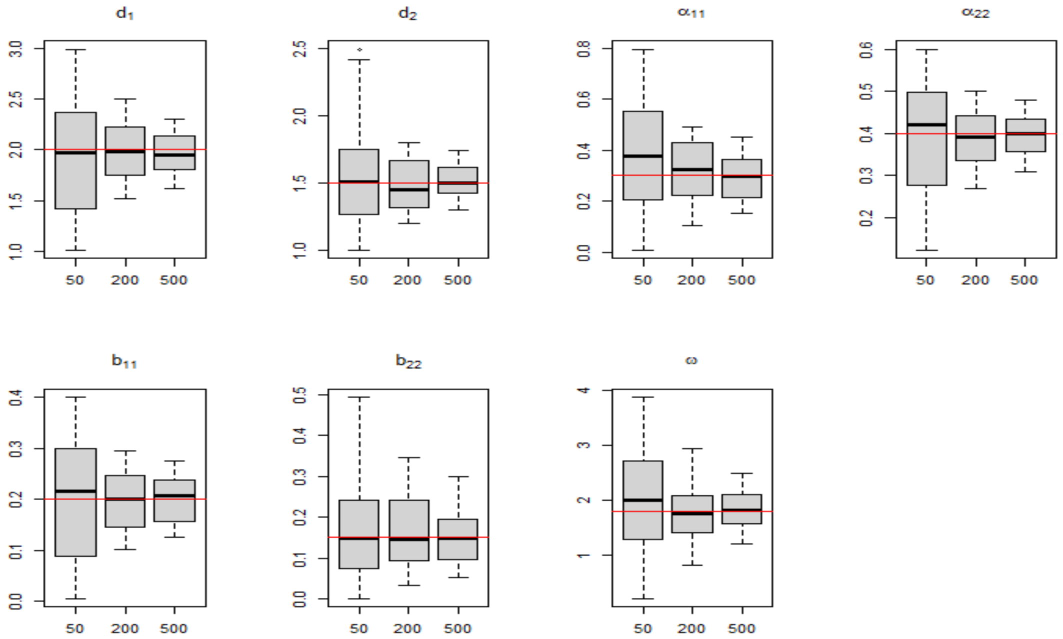


Figure 5.1: Boxplot with results from the simulation experiment with initial parameter values $(d_1,d_2,a_{11},a_{22},b_{11},b_{22},\omega)$=(2,1.5,0.3,0.45,0.2,0.15,1.8)

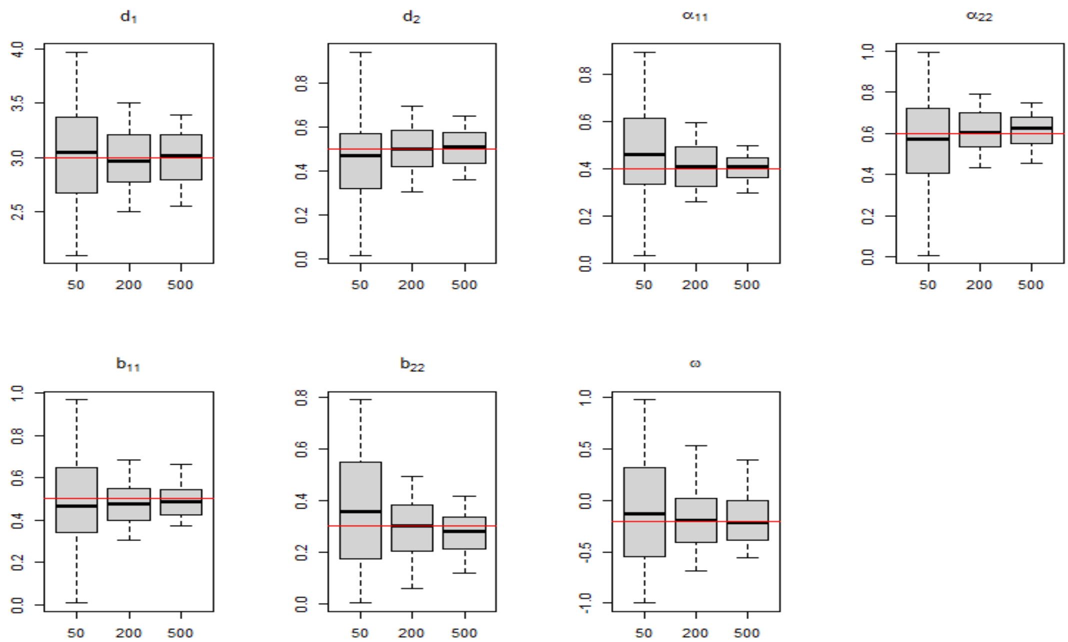


Figure 5.2: Boxplot with results from the simulation experiment with initial parameter values $(d_1,d_2,a_{11},a_{22},b_{11},b_{22},\omega)$=(3,0.5,0.4,0.6,0.5,0.3,0.2)

$$\hat{\lambda}_{jt+1} = E(\lambda_{jt+1} \mid \mathcal{F}_{jt}) = \sum_{i=0}^{T} \left(a_j^i d_j + b_j a_j^i R_{jt-i}\right) + a_j^T \lambda_1. \tag{5.14}$$

5.6 Implementation

Many efforts to protect the climate provide creation of a variety of regulations. One way for the environment protection is the introduction and usefulness of Renewable sources. With renewable sources we mean wind and photovoltaic electricity generators which have an impact on energy market. This impact depend on weather conditions because when weather is rainy or solar energy is not enough there is scarcity of renewable sources and an augmentation of prices of electricity market is obtained. Firstly German policy makers introduced regulations to promote renewable sources but the issue of how Germany's green energy policy has affected electricity market volatility (i.e., the question of whether such mechanisms are effective) has not been analysed so far. To avoid unstable increased market volatility Energy Exchange Law (Energiewirtschaftsgesetz, EnWG) and Renewable Energy Law (Erneuerbare-Energien-Gesetz, EEG), seek to promote renewables in a way that maintains electricity market stability. Auer (2016) are the first who proposed a pure electricity volatility model. GARCH models are inappropriate because they cannot capture the extraordinary characteristics of electricity prices and volatilities which have their origins in the non-storability of electricity requiring a constant balance between production and consumption and (ii) the dependence of electricity demand on the weather and the intensity of business and everyday activities. Moreover CARR models is designed to accommodate extreme values. It is worth mentioned that in Nordic European countries where climate is cold during winter hydroelectric power plants are activated for low prices in electricity . In contarst at Pennsylvania-New Jersey-Maryland(PJM) high and very votatile electricity prices are obtained during the summer because during the winter gas-fired plants are used to generate electricity. used weekly data about German electricity prices from Energy Exchange Market(EXX) for period of June 2000 to August 2015. In our study we use data from Hellenic Energy exchange group and Gestore dei Mercati Energetici SpA in Italy. Theoretically there is a correlation based on similarities of climate conditions(daytime maximum temperature and daily low temperature), humidity, yearly shiny hours, yearly rainy days and hours of daily light. At both countries high electricity values are observed at periods with maximum temperature at summer because electrical devices are very useful to decrease temperature. At

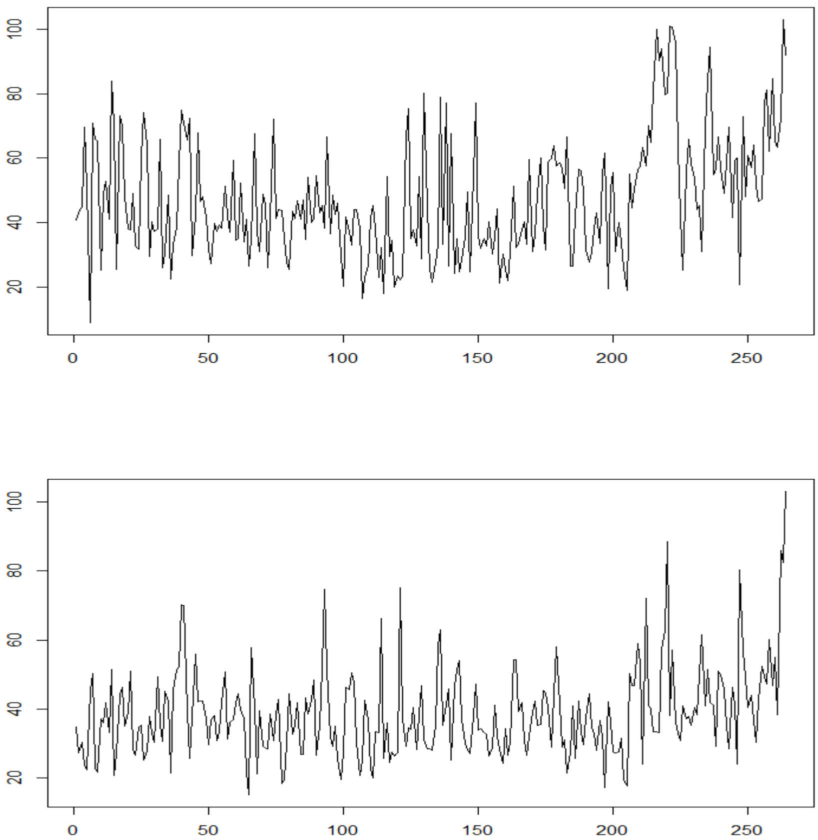


Figure 5.3: Time series plots of the series of daily ranges between maximum and minimum electricity prices for Greece (first plot) and for Italy(second plot)

countless studies with energy economics trends and periodicity both at heteroskedastic and homoskedastic time series models are frequently observed. To eliminate trend models considering exogenous variables have been proposed. Narayan and Liu (2015) studied trend-based GARCH unit root model that outperforms a GARCH model without trend. Sin (2013) proposed a conditional autoregressive range model (CARRX) with exogenous variables in volatilities. At this application only data from September to April with size 264 are used in order to avoid any trend. The drawback obtained at forecasting because intradaily data are not used only a range of maximum and minimum daily values. For that reason we cannot predict specific values of energy. At the table below descriptove statistics are presented considering skewness and kyrtosis. Base don the results the ranges at both countries are exponentially distributed. To estimate parameters $d_i, a_i, b_i, \quad i = 1, 2$, we choose initial

	Mean	Sd	Skewness	Kyrtosis
Greece	76.499	56.774	1.627	5.531
Italy	59.989	41.585	1.534	4.629

Table 5.3: Descriptive statistics for the time series of electricity price ranges

	d_1	d_2	a_{11}	a_{22}	b_{11}	b_{22}	ω
estimate	6.810	9.760	0.678	0.572	0.181	0.211	1.687
se	0.022	0.024	0.019	0.018	0.029	0.024	0.037
		Log-Lik=	-3702.26				

Table 5.4: ML estimates and se for model parameters.

values based on ARIMA model described at the beginning of the chapter and for dependence parameter we take into account boundaries given by (3.30; 3.31).

To conclude with this chapter, a new bivariate CARR time series model has been studied based on copulas defined by (5.10). While there are few studies were bivariate and trivariate CARR processes have been provided, difficulties with other properties like ergodicity or parameters estimation can be found. In our study the structure of bivariate exponential distribution offers the opportunity to study those properties for this model. The main drawback with all of those studies is the way of prediction. To predict a value at this model data observations during a time period are very useful. The fact that variable R_t is a range between maximum and minimum values without considering values in between this interval leads to appropriateness for forecasting. An implementation using daily electricity prices has been studied including data from Greece and Italy. This aforementioned model addressed to researchers who want to study financial data and the shortcomings are that financial data usually have periodicity and trends. For those reasons time series processes at finance could be used to model data for certain time periods dropping out in this way the above problems.

Chapter 6

Main contribution & Future challenges

6.1 Main contribution

We have developed a trans-dimensional MCMC approach for model selection between a family of Poisson INGARCH type models including linear and log-linear models. This allows to fit the models and select the one that best fits the data. Model selection for discrete valued time series is a topic of less research and we think that we contribute towards this. Our approach is based on satisfying the stationarity conditions, i.e. assuming that the series are stationary. Allowing for combinations of parameters beyond stationarity is possible but the interpretation of the models is more difficult afterwards.

The method that has been discussed by Carlin and Chib (1995) is the most flexible method in our case when we have a number of nested models. This is an important issue because since now there are limited studies about trans-dimensional methods for count data.

Furthermore we note that, although we only analyzed methodology by using linear and log-linear INGARCH models when all parameters are positive, our analysis can be extended when parameters are negative in case of log-linear INGARCH models. In addition the use of other functions for the mean process λ_t presented by Fokianos et al. (2009); Fokianos and Tjøstheim (2011) could be examined for computational efficacy.

Additionally we treated only Poisson INGARCH models. However the method can be easily extended to the case of other models of the same type with different conditional distributions that have appeared in the literature.

We contributed at the construction of multivariate discrete and continuous distributions that offer strong correlation based on the new structure of Sarmanov distribution. We only discussed this family assuming properties of p.m.f and p.d.f. and we proposed the closed formula of a copula without taking into account any study about copula's properties. We presented boundaries of correlation coefficient at the case of Poisson and exponential distributions providing the flexible way of use in both cases. While Miravete (2009b) proposed a multivariate Sarmanov distribution, in our study an alternative construction of this distribution is provided. The advantage is that this multiplicative factor consider only products between all combinations of appropriate functions each time. This decreases computational cost and theoretical properties are more easier to study.

While a multivariate Poisson and a multivariate Exponential time series models have been structured and theoretical properties have been discussed, the proposed Sarmanov could be used for the study of the above multivariate distributions. Moreover computational cost for parameters estimation could be decreased if maximization by parts method proposed by Song et al. (2005) implemented. Poisson INGARCH model dealt with the problem of equidispersion and for this reason less datasets can be studied with this model. In univariate case INGARCH models of Binomial and Negative Binomial distributions have been studied. Mimic those models and based on the new structure of copula, other multivariate INGARCH models where processes X_t are simulated for other discrete distibutions could be provided.

In case of exponential distribution homoskedastic time series processes are difficult to study. Assuming that innovations simulated from exponential distributions and using characteristic functions, one can proved that process X_t can not simulated from an exponential distribution and this leads to more complicated structures. In contrary in case of data with heteroskedasticity, models with exponential distribution can be constructed without obstacles. Based on the idea of Bollerslev et al. (1994), an exponential CARR model has been developed and an application with data of electricity energy sources has been provided.

The model becomes complicated when interdependencies can be considered at volatilities expressions. One can observe difficulties not only with the study of theoretical properties like ergodicity but also with the intensive computational way of parameters estimation. At both studies weak dependence criteria have been discussed for studying ergodic properties but other theoretical tools could be also used to provide stationarity at p^{th}-order moments where $p \geq 2$ as future work.

6.2 Future challenges

Several extensions of Bivariate INGARCH or Multivariate INGARCH models can enhance the research. Studies of multivariate INGARCH models where volatilities will be expressed by another nonlinear functions could be produced. Log-linear INGARCH model has been also constructed in one dimension but the construction in more dimensions is a topic of interest. One can study the model when

$$\mathbf{X_t} \sim BP(\lambda_\mathbf{t}),$$

where

$$\lambda_t = \mathbf{d} + A\lambda_{t-1} + B\mathbf{Y}_{t-1}.$$

BP is a bivariate Poisson distribution, A= $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and B= $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$. Assuming the constructions above, positive and negative correlations and cross-correlations can occur and interdependencies can be captured to the model. Furthermore INGARCH(p,q) models with incorporation of other discrete distributions can be account. For example Zhu (2011), Zhu (2012c);Zhu (2012b) studied univariate Negative Binomial and Compound Poisson INGARCH models and Miravete (2009b) studied bivariate Negative Binomial distribution without considering time series processes. Based on the above study one can develop bivariate Negative Binomial and Compound Poisson INGARCH processes to model data with over or under-dispersion. In our study we propose a trivariate INGARCH(1,1) model defined by (2.1), (5.2) but one may extend this model in n dimensions $n \geq 3$ where volatilities could be expressed by :

$$\lambda_t = \mathbf{d} + A\sum_{i=1}^{p} \lambda_{t-i} + B\sum_{j=1}^{q} \mathbf{Y}_{t-j},$$

where

$$\mathbf{d}= \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n, \end{bmatrix} \quad \text{A}= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \ddots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix} \text{ and B}= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \ddots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nq} \end{bmatrix}.$$

Stationarity could be provided using other theoretical approaches except weak dependence criteria but taking into account how many weak dependence criteria there are, stationarity of low order moments could be also provided in cases where more theoretical approaches are difficult to proceed.

Moreover the contribution to the new bivariate distribution given by (5.1) is the definition of another new where one variable will be discrete and the other will be continuous. At this case we will have:

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \omega q_1(x_1)q_2(x_2)],$$

where:

$$q_1(x_1) = c_1^{-k_1 x_1} - \sum_{x_1=0}^{\infty} c_1^{-k_1 x_1} f_1(x_1)$$

and

$$q_2(x_2) = c_2^{-k_2 x_2} - \int_{x_2=0}^{\infty} c_2^{-k_2 x_2} f_2(x_2).dx_2.$$

This is a challenging problem because there are not many studies with distributions of this form to account for. The problem becomes more intricated when a multivariate distribution is going to constructed and multivariate time series models could be defined without assuming interdependencies. The transformation of this bivariate distribution to a copula is:

$$C(u,v) = uv\left[1 + \omega\left(\sum_{t_1=0}^{F^{-1}(u)} c_1^{-k_1 t_1} - uLT_1(f(x_1))\right)\left(\sum_{t_2=0}^{F^{-1}(v)} c_2^{-k_2 t_2} - vLT_2(f(x_2))\right)\right].$$

Considering the appropriate marginals, the construction is easy to derived calculating the corresponding Laplace transforms. While the formula has been given and at chapter 3 a copula with exponential marginals was proposed, properties to be a copula is an aforementioned problem.

At CARR time series model, studies with data from financial and econometrics fields have been implied. The crucial problem with this family of models is to find a way to forecast values considering not only the maximum and minimum values of a period but also values in between those periods. The difficulty stems from volatility's expression. Moreover Chou and Wang (2006) compared the performance of an Exponential CARR model and a GARCH model. An exponential CARR model with the incorporation of exogenous variables has been also appeared at this study. A model similar to this could be studied in order to decrease model's order and in general to improve it's accuracy.

$$\mathbf{R}_t \sim BExp(\lambda_t),$$

where

$$\lambda_t = \mathbf{d}_1 + A\lambda_{t-1} + B\mathbf{R}_{t-1} + \Gamma\mathbf{Y}_{t-1}$$

and $\mathbf{Y}_t$ could be simulated from any continuous distribution.

Chapter 7

Appendix

7.1 Appendix A: Bivariate distribution based on copulas

Proof of bivariate Poisson distribution:

$$\begin{aligned}
f(x_1, x_2) &= \prod_{i=1}^{2} f_i(x_i)\left[1+\rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}q_i(x_i)\right] = \prod_{i=1}^{2} f_i(x_i) + \rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}f_i(x_i)c_i^{-k_i x_i} \\
&- \rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}f_i(x_i)\left[c_1^{-k_1x_1}e^{\lambda_2(c_2^{-k_2}-1)} + c_2^{-k_2x_2}e^{\lambda_1(c_1^{-k_1}-1)}\right] \\
&+ \rho\prod_{i=1}^{2}P_i(x_i) - \rho D_1(x_1)P_2(x_2) - \rho P_1(x_1)D_2(x_2) + \rho\prod_{i=1}^{2}D_i(x_i) \\
&= \prod_{i=1}^{2} f_i(x_i) + \rho\prod_{i=1}^{2}\frac{[f_i(x_i^{'}) - f_i(x_i)]}{\sqrt{\lambda_i}(c_i^{-k_i}-1)}, \qquad (7.1)
\end{aligned}$$

where

$$D_i(x_i) = \frac{f_i(x_i)}{\sqrt{\lambda_i}(c_i^{-k_i}-1)}, \quad P_i(x_i) = \frac{f_1(x_1^{'})}{\sqrt{\lambda_1}(c_i^{-k_1}-1)}, \quad f_i(x_i^{'}) \sim Poisson(\lambda_i c_i^{-k_i}).$$

Proof of bivariate Exponential distribution:

$$\begin{aligned}
f(x_1x_2) &= \prod_{i=1}^{2} f_i(x_i) + \rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}f_i(x_i)c_i^{-k_ix_i} + \rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}f_i(x_i)Q(\theta) + \rho\prod_{i=1}^{2}\frac{\sigma_i}{v_i}f_i(x_i)\frac{\lambda_i}{\log\left(e^{\lambda_i}c_i^{k_i}\right)} \\
&= \prod_{i=1}^{2} f_i(x_i) + \rho\prod_{i=1}^{2}P_i(x_{i,t}) - \rho D_1(x_{1,t})P_2(x_{2,t}) - \rho\prod_{i=1}^{2}P_1(x_{1,t}D_2(x_{2,t})) + \rho\prod_{i=1}^{2}D_i(x_{i,t}) \\
&= \prod_{i=1}^{2} f_i(x_i) + \prod_{i=1}^{2}\frac{\log\left(e^{\lambda_i}c_i^{k_i}\right)}{k_i\log c_i}[f_i(x_i^{''}) - f_i(x_i)],
\end{aligned}$$

where

$$Q(\theta) = \left[c_1^{-k_1 x_1} \frac{\lambda_2}{\lambda_2 + k_2 \log(c_2)} + \frac{\lambda_1}{\lambda_1 + k_1 \log c_1} c_2^{-k_2 x_2} \right],$$

$$P_i(x_{i,t}) = \frac{\log\left(e^{\lambda_i} c_i^{k_i}\right)}{k_i \log c_i} f_i(x_i^{''})],$$

$$D_i(x_{i,t}) = \frac{\log\left(e^{\lambda_i} c_i^{k_i}\right)}{k_i \log c_i} f_i(x_i),$$

$$f_i(x_i^{''}) \sim Exp(\lambda_i + k_i \log c_i).$$

7.2 Appendix B: Partial derivatives of INGARCH process

Assuming parameter vector $\theta_i = (\alpha_{ij}, b_{ij}, d_i, \omega_{ij})$, $i, j = 1, 2, 3$ first order partial derivatives with respect to parameter vector θ_i and second order partial derivatives with respect to b_{ij}, ω_{ij} are given accordingly by:

$$\begin{aligned}
\frac{\partial l(\theta_i)}{\partial d_i} &= \sum_{t=1}^{n} \left(Q_t(\theta)\right) \frac{\partial \lambda_i(\theta_i)}{\partial d_i}, \\
\frac{\partial l(\theta_i)}{\partial b_{ij}} &= \sum_{t=1}^{n} \left(Q_t(\theta)\right) \frac{\partial \lambda_i(\theta)}{\partial b_{ij}}, \\
\frac{\partial l(\theta_i)}{\partial a_{ij}} &= \sum_{t=1}^{n} \left(Q_t(\theta)\right) \frac{\partial \lambda_i(\theta)}{\partial a_{ij}}, \\
\frac{\partial l(\theta_i)}{\partial \omega_{ij}} &= \sum_{i=1}^{n} \left(\frac{1}{\phi_t(\theta)} R_i(x_{i,t}) R_j(x_{j,t}) \right) \frac{\partial \lambda_i(\theta)}{\partial \omega_{ij}}, \\
\frac{\partial^2 l(\theta)}{\partial \omega_{ij}^2} &= -\frac{\sum_{t=1}^{n} R_i^2(x_{i,t}) R_j^2(x_{j,t}}{\phi_t^2(\theta)} \frac{\partial \lambda_i^2(\theta)}{\partial \omega_{ij}^2}.
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 l(\theta)}{\partial b_{ij}^2} = & \sum_{t=1}^{n}\left(\sum_{i=1}^{3}\left(-\frac{X_{i,t}}{\lambda_i^2(\theta)}\left(\frac{\partial\lambda_i(\theta)}{\partial b_{ij}}\right)^2\right)+\sum_{i=1}^{3}\left(\frac{X_{i,t}}{\lambda_i(\theta)}-1\right)\frac{\partial^2\lambda_i(\theta)}{\partial b_{ij}^2}\right) \\
& + \sum_{t=1}^{n}\frac{1}{\phi_t(\theta)}\left(\sum_{i<j=1}^{3} -\omega_{ij}\left(c_i^{k_i}-1\right)D_i(x_{i,t})R_j(x_{j,t})\left(\frac{\partial\lambda_i(\theta)}{\partial b_{ij}}\right)^2\right) \\
& - \left(\omega_{ij}\left(c_j^{k_j}-1\right)D_j(x_{j,t})R_i(x_{i,t})\left(\frac{\partial\lambda_j(\theta)}{\partial b_{ij}}\right)^2\right) \\
& + 2D_i(x_{i,t})D_j(x_{j,t}))\frac{\partial\lambda_i(\theta)}{\partial b_{ij}}\frac{\partial\lambda_j(\theta)}{\partial b_{ij}} \\
& - \frac{1}{\phi_t^2(\theta)}\left(\sum_{i<j=1}^{3}\left(\omega_{ij}^2 D_i^2(x_{i,t})R_j^2(x_{j,t})\right)\left(\frac{\partial\lambda_i(\theta)}{\partial b_{ij}}\right)^2\right) \\
& + \left(\omega_{ij}^2 D_j^2(x_{j,t})R_i^2(x_{i,t})\left(\frac{\partial\lambda_j(\theta)}{\partial b_{ij}}\right)^2\right. \\
& + 2\omega_{ij}^2 D_i(x_{i,t}D_j(x_{j,t}))R_i(x_{i,t})R_j(x_{j,t})\Big),
\end{aligned}
$$

where

$$
\begin{aligned}
R_i(x_{i,t}) &= c_i^{-k_i x_{i,t}} - e^{\lambda_i(c_i^{-k_i}-1)}, \\
D_i(x_{i,t}) &= e^{\lambda_i(c_i^{-k_i}-1)}\left(c_i^{-k_i}-1\right), \\
Q_t(\theta) &= \sum_{i=1}^{3}\left(\frac{X_{i,t}}{\lambda_i(\theta)}-1\right)-\frac{1}{\phi_t(\theta)}\sum_{i<j=1}^{3}\left(\omega_{ij}e^{\lambda_i(c_i^{-k_i}-1)}R_i(x_{i,t})+\omega_{ij}R_j(x_{j,t})e^{\lambda_j(c_j^{-k_j}-1)}\right).
\end{aligned}
$$

Based on equation (1.14) derivatives of $\lambda_{it},\ ,i=1,2$ with respect to parameter vector $theta_i,\ i=1,2$ defined previously are given by:

$$
\begin{aligned}
\frac{\partial\lambda_i(\theta)}{\partial d_i} &= d_i + b_{ij}\frac{\partial\lambda_{i,t-1}}{\partial d_i}, \\
\frac{\partial\lambda_i(\theta)}{\partial b_{ij}} &= \lambda_{j,t-1} + b_{ij}\frac{\partial\lambda_{j,t-1}}{\partial b_{ij}}, \\
\frac{\partial\lambda_i(\theta)}{\partial a_{ij}} &= \lambda_{j,t-1} + a_{ij}\frac{\partial\lambda_{i,t-1}}{\partial a_{ij}}.
\end{aligned}
$$

7.3 Appendix C: Partial derivatives of CARR process

Considering bivariate CARR model defined by (5.1) and volatilities defined by (5.2) first and second order partial derivatives with respect to parameters a_i and ω are given by:

$$
\begin{aligned}
\frac{\partial l(\theta)}{\partial c_i} &= \sum_{t=1}^{n} \omega \left(c_j^{-k_j r_{jt}} - \frac{1}{1+\lambda_{jt} k_j \log c_j} \right) \frac{k_i \log c_i}{\left(1+\lambda_{it} k_i \log c_i\right)^2 q(r_{1t}, r_{2t})} \frac{\partial \lambda_{it}}{\partial c_i} \\
&+ \sum_{t=1}^{n} \left(\frac{\lambda_{it} + R_{it}}{\lambda_{it}^2} \right) \frac{\partial \lambda_{it}}{\partial c_i}, \\
\frac{\partial l(\theta)}{\partial \omega} &= \sum_{t=1}^{n} \prod_{i=1}^{2} \frac{\left(c_i^{-k_i r_{it}} - \frac{1}{1+\lambda_{it} k_i \log c_i} \right)}{q(r_{1t}, r_{2t})}, \\
\frac{\partial l^2(\theta)}{\partial c_i^2} &= \sum_{t=1}^{n} \left(\frac{\lambda_{it} - 2R_{it}}{\lambda_{it}^3} - \frac{2\omega k_i \log c_i c_j^{-k_j r_{jt}} q(r_{1t}, r_{2t})}{(1+\lambda_{it} k_i \log c_i)^3} \right) \frac{\partial \lambda_{it}}{\partial c_i} \\
&+ \sum_{t=1}^{n} \left(\frac{2\omega k_i \log c_i q(r_{1t}, r_{2t})}{\left(1+\lambda_{jt} k_j \log c_j\right)\left(1+\lambda_{it} k_i \log c_i\right)^3} - \frac{\omega c_j^{-k_j r_{jt}}}{\left(1+\lambda_{it} k_i \log c_i\right)^4} \right) \frac{\partial \lambda_{it}}{\partial c_i} \\
&+ \sum_{t=1}^{n} \frac{\omega}{(1+\lambda_{it} k_i \log c_i)^4 \left(1+\lambda_{jt} k_j \log c_j\right)} \frac{\partial \lambda_{it}}{\partial c_i} + \left(\frac{\lambda_{it} + r_{it}}{\lambda_{it}^2} \right) \frac{\partial^2 \lambda_{it}}{\partial c_i^2} \\
&+ \sum_{t=1}^{n} \omega k_i \log c_i \left(c_j^{-k_j r_{jt}} q(r_{1t}, r_{2t}) - \frac{1}{1+\lambda_{jt} k_j \log c_j} \right) \frac{\partial^2 \lambda_{it}}{\partial c_i^2}, \\
\frac{\partial l^2(\theta)}{\partial \omega^2} &= -\sum_{t=1}^{n} \prod_{i=1}^{2} \frac{\left(c_i^{-k_i r_{it}} - \frac{1}{1+\lambda_{it} k_i \log c_i} \right)^2}{q^2(r_{1t}, r_{2t})}.
\end{aligned}
$$

where $q(r_{1t}, r_{2t})$ given by (5.11) and partial derivatives with respect to d_i and b_i are calculated in a similar way.

Bibliography

Abraham, B. and N. Balakrishna (1999). Inverse Gaussian autoregressive models. *Journal of Time Series Analysis 20*(6), 605–618.

Al-Osh, M. and A. A. Alzaid (1988). Integer-valued Moving Average (INMA) process. *Statistical Papers 29*(1), 281–300.

Al-Osh, M. A. and A. A. Alzaid (1987). First-order Integer-valued Autoregressive (INAR (1)) process. *Journal of Time Series Analysis 8*(3), 261–275.

Alexander, C. and A. Chibumba (1997). Multivariate orthogonal factor GARCH. *University of Sussex, Mimeo*.

Alomani, G. A., A. A. Alzaid, and M. A. Omair (2018). A Skellam GARCH model. *Brazilian Journal of Probability and Statistics 32*(1), 200–214.

Alzahrani, N., P. Neal, S. E. Spencer, T. J. McKinley, and P. Touloupou (2018). Model selection for time series of count data. *Computational Statistics & Data Analysis 122*, 33–44.

Alzaid, A. and M. Al-Osh (1990). An Integer-valued pth-order Autoregressive structure (INAR (p)) process. *Journal of Applied Probability 27*(2), 314–324.

Anderson, H. M., K. Nam, and F. Vahid (1999). Asymmetric nonlinear smooth transition GARCH models. In *Nonlinear time series analysis of economic and financial data*, pp. 191–207. Springer.

Andreassen, C. M. (2013). *Models and inference for correlated count data.* Ph. D. thesis, Aarhus University, Department of Mathematics.

Auer, B. R. (2016). How does Germany's green energy policy affect electricity market volatility? An application of conditional autoregressive range models. *Energy Policy 98*, 621–628.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics 74*(1), 3–30.

Balakrishna, N. and C. S. Ranganath (2015). ARMA models with generalized error distributed innovations. *Journal of the Indian Statistical Association 53*(1), 11–34.

Balakrishna, N. and K. Shiji (2014). Extreme value Autoregressive model and its applications. *Journal of Statistical Theory and Practice 8*(3), 460–481.

Balakrishnan, N. (2014). Continuous multivariate distributions. *Wiley StatsRef: Statistics Reference Online*.

Barbiero, A. (2019). A bivariate count model with discrete Weibull margins. *Mathematics and Computers in Simulation 156*, 91–109.

Barreto-Souza, W., L. S. Piancastelli, K. Fokianos, and H. Ombao (2022). Time-Varying Dispersion Integer-Valued GARCH Models. *arXiv preprint arXiv:2208.02024*.

Bauwens, L. and P. Giot (2000). The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique*, 117–149.

Bauwens, L., P. Giot, et al. (2001). *Econometric modelling of stock market intraday activity*, Volume 38. Springer Science & Business Media.

Bauwens, L. and D. Veredas (2004). The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of Econometrics 119*(2), 381–412.

Bayarri, M. J., J. O. Berger, A. Forte, G. García-Donato, et al. (2012). Criteria for Bayesian model choice with application to variable selection. *The Annals of Statistics 40*(3), 1550–1577.

Bekrizadeh, H., G. A. Parham, and M. R. Zadkarmi (2012). The new generalization of Farlie–Gumbel–Morgenstern copulas. *Applied Mathematical Sciences 6*(71), 3527–3533.

Bentarzi, M. and W. Bentarzi (2017). Periodic integer-valued GARCH (1, 1) model. *Communications in Statistics-Simulation and Computation 46*(2), 1167–1188.

Berben, R.-P., W. J. Jansen, et al. (2005). Bond market and stock market integration in Europe. Technical report, Netherlands Central Bank, Research Department.

Berkhout, P. and E. Plug (2004). A bivariate poisson count data model using conditional probabilities. *Statistica Neerlandica 58*(3), 349–364.

Bibi, A. and A. Aknouche (2008). On periodic GARCH processes: Stationarity, existence of moments and geometric ergodicity. *Mathematical methods of Statistics 17*(4), 305–316.

Bickel, P. J. and P. Bühlmann (1999). A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli*, 413–446.

Billio, M., M. Caporin, and M. Gobbo (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters 2*(02), 123–130.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics*, 498–505.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. *Handbook of Econometrics 4*, 2959–3038.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy 96*(1), 116–131.

Boussama, F., F. Fuchs, and R. Stelzer (2011). Stationarity and geometric ergodicity of BEKK multivariate GARCH models. *Stochastic Processes and their Applications 121*(10), 2331–2360.

Box, G. E., G. M. Jenkins, G. C. Reinsel, and G. M. Ljung (2015). *Time Series Analysis: Forecasting and Control.* John Wiley & Sons.

Brijs, T., D. Karlis, and G. Wets (2008). Studying the effect of weather conditions on daily crash counts using a discrete time-series model. *Accident Analysis & Prevention 40*(3), 1180–1190.

Bu, R. and B. McCabe (2008). Model selection, estimation and forecasting in INAR (p) models: a likelihood-based Markov chain approach. *International Journal of Forecasting 24*(1), 151–162.

Cai, J. (1994). A markov model of switching-regime ARCH. *Journal of Business & Economic Statistics 12*(3), 309–316.

Cambanis, S. (1977). Some properties and generalizations of multivariate Eyraud-Gumbel-Morgenstern distributions. *Journal of Multivariate Analysis 7*(4), 551–559.

Cappiello, L., R. F. Engle, and K. Sheppard (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics 4*(4), 537–572.

Carlin, B. P. and S. Chib (1995). Bayesian model choice via Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(3), 473–484.

Carrasco, M. and X. Chen (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory 18*(1), 17–39.

Chan, J. S.-K., K.-H. Ng, and R. Ragell (2019). Bayesian return forecasts using realised range and asymmetric CARR model with various distribution assumptions. *International Review of Economics & Finance 61*, 188–212.

Chen, C. W. and S. Lee (2016). Generalized Poisson autoregressive models for time series of counts. *Computational Statistics & Data Analysis 99*, 51–67.

Chen, C. W., M. K. So, J. C. Li, and S. Sriboonchitta (2016). Autoregressive conditional Negative Binomial model applied to over-dispersed time series of counts. *Statistical Methodology 31*, 73–90.

Chiang, M.-H., R. Y. Chou, and L.-M. Wang (2016). Outlier detection in the lognormal logarithmic conditional autoregressive range model. *Oxford Bulletin of Economics and Statistics 78*(1), 126–144.

Chib, S. and E. Greenberg (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician 49*(4), 327–335.

Chou, H.-C. and D. Wang (2006). Using conditional autoregressive range model to forecast volatility of the stock indices. In *9th Joint International Conference on Information Sciences (JCIS-06)*, pp. 592–595. Atlantis Press.

Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: the Conditional Autoregressive Range (CARR) model. *Journal of Money, Credit and Banking*, 561–582.

Chou, R. Y. (2006). Modeling the asymmetry of stock movements using price ranges. In *Econometric Analysis of Financial and Economic Time Series*. Emerald Group Publishing Limited.

Chou, R. Y., H. Chou, and N. Liu (2010). Range volatility models and their applications in finance. In *Handbook of Quantitative Finance and Risk Management*, pp. 1273–1281. Springer.

Chou, R. Y., H. Chou, and N. Liu (2015). Range volatility: a review of models and empirical studies. *Handbook of Financial Econometrics and Statistics*, 2029–2050.

Clark, N. J. and P. M. Dixon (2021). A class of spatially correlated self-exciting statistical models. *Spatial Statistics 43*, 100493.

Consul, P. C. and G. C. Jain (1973). A generalization of the Poisson distribution. *Technometrics 15*(4), 791–799.

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

Cui, Y., Q. Li, and F. Zhu (2020). Flexible bivariate Poisson integer-valued GARCH model. *Annals of the Institute of Statistical Mathematics 72*(6), 1449–1477.

Cui, Y. and F. Zhu (2018). A new bivariate Integer-valued GARCH model allowing for negative cross-correlation. *Test 27*(2), 428–452.

Cushman, D. O. and T. Zha (1997). Identifying monetary policy in a small open economy under flexible exchange rates. *Journal of Monetary Economics 39*(3), 433–448.

Davis, R. A., K. Fokianos, S. H. Holan, H. Joe, J. Livsey, R. Lund, V. Pipiras, and N. Ravishanker (2021). Count time series: A methodological review. *Journal of the American Statistical Association 116*(535), 1533–1547.

Davis, R. A. and W. P. McCormick (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Processes and their Applications 31*(2), 237–250.

Dedecker, J., P. Doukhan, G. Lang, L. R. J. Rafael, S. Louhichi, and C. Prieur (2007). Weak dependence. In *Weak dependence: With examples and applications*, pp. 9–20. Springer.

Deistler, M. and E. Hannan (1988). The statistical theory of linear systems.

Del Castillo, J. and M. Pérez-Casany (1998). Weighted Poisson distributions for overdispersion and underdispersion situations. *Annals of the Institute of Statistical Mathematics 50*(3), 567–585.

del Castillo, J. and M. Pérez-Casany (2005). Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference 134*(2), 486–500.

Dellaportas, P., J. J. Forster, and I. Ntzoufras (2002). On Bayesian model and variable selection using MCMC. *Statistics and Computing 12*(1), 27–36.

Denuit, M. and P. Lambert (2005). Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis 93*(1), 40–57.

Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics 105*(1), 131–159.

Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *Journal of Applied Econometrics 4*(1), 1–21.

Ding, Z., C. W. Granger, and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*(1), 83–106.

Doukhan, P. and P. Doukhan (1994). Mixing. *Mixing: Properties and Examples*, 15–23.

Doukhan, P., K. Fokianos, and D. Tjøstheim (2012). On weak dependence conditions for Poisson autoregressions. *Statistics & Probability Letters 82*(5), 942–948.

Doukhan, P. and S. Louhichi (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and Their Applications 84*(2), 313–342.

Duan, J.-C. (1997). Augmented GARCH(p, q) process and its diffusion limit. *Journal of Econometrics 79*(1), 97–127.

Dufour, A. and R. F. Engle (2000). Time and the price impact of a trade. *The Journal of Finance 55*(6), 2467–2498.

Efron, B. (1986). Double Exponential families and their use in generalized linear regression. *Journal of the American Statistical Association 81*(395), 709–721.

Enciso-Mora, V., P. Neal, and T. Subba Rao (2009). Efficient order selection algorithms for integer-valued ARMA processes. *Journal of Time Series Analysis 30*(1), 1–18.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Eonometrics 131*(1-2), 3–27.

Engle, R. F., C. W. Granger, and D. Kraft (1984). Combining competing forecasts of inflation using a bivariate ARCH model. *Journal of Economic Dynamics and Control 8*(2), 151–165.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*(1), 122–150.

Engle, R. F. and G. Lee (1999). A long-run and short-run component model of stock return volatility. *Cointegration, causality, and forecasting: A Festschrift in honour of Clive WJ Granger*, 475–497.

Engle, R. F. and V. K. Ng (1993). Measuring and testing the impact of news on volatility. *The journal of Finance 48*(5), 1749–1778.

Engle, R. F., V. K. Ng, and M. Rothschild (1990). Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics 45*(1-2), 213–237.

Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, 1127–1162.

Engle III, R. F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH.

Farlie, D. J. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika 47*(3/4), 307–323.

Feigin, P. D. and S. I. Resnick (1992). Estimation for autoregressive processes with positive innovations. *Communications in Statistics. Stochastic Models 8*(3), 685–717.

Ferland, R., A. Latour, and D. Oraichi (2006). Integer-valued GARCH process. *Journal of Time Series Analysis 27*(6), 923–942.

Fernandes, M., B. de Sá Mota, and G. Rocha (2005). A multivariate conditional autoregressive range model. *Economics Letters 86*(3), 435–440.

Fokianos, K. (2012). Count time series models. In *Handbook of statistics*, Volume 30, pp. 315–347. Elsevier.

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

Fokianos, K., A. Rahbek, and D. Tjøstheim (2009). Poisson autoregression. *Journal of the American Statistical Association 104*(488), 1430–1439.

Fokianos, K., B. Støve, D. Tjøstheim, P. Doukhan, et al. (2020). Multivariate count autoregression. *Bernoulli 26*(1), 471–499.

Fokianos, K. and D. Tjøstheim (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis 102*(3), 563–578.

Fokianos, K. and D. Tjøstheim (2012). Nonlinear poisson autoregression. *Annals of the Institute of Statistical Mathematics 64*(6), 1205–1225.

Fonseca, G. and R. Tweedie (2002). Stationary measures for non-irreducible non-continuous Markov chains with time series applications. *Statistica Sinica*, 651–660.

Francq, C. and J.-M. Zakoïan (2006). Mixing properties of a general class of GARCH (1, 1) models without moment assumptions on the observed process. *Econometric Theory 22*(5), 815–834.

Frank, M. J. (1979). On the simultaneous associativity off (x, y) andx+ y- f (x, y). *Aequationes mathematicae 19*(1), 194–226.

Franke, J. (2010). Weak dependence of functional ingarch processes.

Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, 3ˆ e serie, Sciences, Sect. A 14*, 53–77.

Freeland, R. and B. P. McCabe (2004). Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis 25*(5), 701–722.

Gaver, D. P. and P. Lewis (1980). First-order autoregressive Gamma sequences and point processes. *Advances in Applied Probability 12*(3), 727–745.

Gelfand, A. E. and S. R. Dalal (1990). A note on overdispersed Exponential families. *Biometrika 77*(1), 55–64.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika 74*(3), 549–555.

George, D. and S. George (2013). Marshall-Olkin Esscher transformed Laplace distribution and processes. *Brazilian Journal of Probability and Statistics*, 162–184.

GEORGE, E. (1970). Box, Gwilym M Jenkins, and G Reinsel. *Time Series Analysis: forecasting and control holden-day san francisco. BoxTime Series Analysis: Forecasting and Control Holden Day1970*.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian Statistics 4*, 641–649.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of Finance 48*(5), 1779–1801.

Gonçalves, E., N. M. Lopes, and F. Silva (2015). A new approach to integer-valued time series modeling: The Neyman type-A INGARCH model. *Lithuanian Mathematical Journal 55*(2), 231–242.

Gonçalves, E., N. Mendes-Lopes, and F. Silva (2016). Zero-inflated compound Poisson distributions in integer-valued GARCH models. *Statistics 50*(3), 558–578.

González-Rivera, G. (1998). Smooth-transition GARCH models. *Studies in Nonlinear Dynamics & Econometrics 3*(2).

Gorgi, P. (2020). Beta–Negative Binomial auto-regressions for modelling integer-valued time series with extreme observations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 82*(5), 1325–1347.

Gouriéroux, C. and Y. Lu (2019). Negative Binomial autoregressive process with stochastic intensity. *Journal of Time Series Analysis 40*(2), 225–247.

Granger, C. W. and N. Hyung (2004). Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns. *Journal of Empirical Finance 11*(3), 399–421.

Granger, C. W. J. and Z. Ding (1995). Some properties of absolute return: An alternative measure of risk. *Annales d'Economie et de Statistique*, 67–91.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics 42*(1), 27–62.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Grunwald, G. K., R. J. Hyndman, L. Tedesco, and R. L. Tweedie (2000). Theory & methods: Non-gaussian conditional linear ar (1) models. *Australian & New Zealand Journal of Statistics 42*(4), 479–495.

Gumbel, E. J. (1960a). Bivariate Exponential distributions. *Journal of the American Statistical Association 55*(292), 698–707.

Gumbel, E. J. (1960b). Distributions des valeurs extremes en plusiers dimensions. *Publ. Inst. Statist. Univ. Paris 9*, 171–173.

Haas, M., S. Mittnik, and M. S. Paolella (2004). A new approach to markov-switching GARCH models. *Journal of financial Econometrics 2*(4), 493–530.

Hagerud, G. E. (1997). *A new non-linear GARCH model.* Economic Research Institute, Stockholm School of Economics.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 357–384.

Hamilton, J. D. and R. Susmel (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics 64*(1-2), 307–333.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

Heinen, A. (2003). Modelling time series count data: an autoregressive conditional Poisson model. *Available at SSRN 1117187*.

Heinen, A. and E. Rengifo (2007). Multivariate autoregressive modeling of time series count data using copulas. *Journal of Empirical Finance 14*(4), 564–583.

Hsiao, C. (1979). Autoregressive modeling of Canadian money and income data. *Journal of the American Statistical Association 74*(367), 553–560.

Hsiao, C. (1981). Autoregressive modelling and money-income causality detection. *Journal of Monetary Economics 7*(1), 85–106.

Hu, X. and B. Andrews (2021). Integer-valued asymmetric GARCH modeling. *Journal of Time Series Analysis 42*(5-6), 737–751.

Huang, J. S. and S. Kotz (1999). Modifications of the Farlie-Gumbel-Morgenstern distributions. a tough hill to climb. *Metrika 49*(2), 135–145.

Hudecová, Š., M. Hušková, and S. G. Meintanis (2021). Goodness–of–fit tests for bivariate time series of counts. *Econometrics 9*(1), 10.

Ing, C.-K. and C.-Y. Yang (2014). Predictor selection for positive autoregressive processes. *Journal of the American Statistical Association 109*(505), 243–253.

Jacobs, P. A. (1978). A cyclic queueing network with dependent Exponential service times. *Journal of Applied Probability 15*(3), 573–589.

Jacobs, P. A. (1980). Heavy traffic results for single-server queues with dependent (EARMA) service and interarrival times. *Advances in Applied Probability 12*(2), 517–529.

Jacobs, P. A. and P. A. Lewis (1978). Discrete time series generated by mixtures ii: asymptotic properties. *Journal of the Royal Statistical Society: Series B (Methodological) 40*(2), 222–228.

Jayakumar, K. and A. Kuttykrishnan (2007). A time-series model using asymmetric Laplace distribution. *Statistics & probability letters 77*(16), 1636–1640.

Jiménez-Gamero, M. D., S. Lee, and S. G. Meintanis (2020). Goodness-of-fit tests for parametric specifications of conditionally heteroscedastic models. *TEST 29*(3), 682–703.

Jose, K., L. Tomy, and J. Sreekumar (2008). Autoregressive processes with Normal-Laplace marginals. *Statistics & Probability Letters 78*(15), 2456–2462.

Karlis, D. and I. Ntzoufras (2003). Analysis of sports data by using bivariate Poisson models. *Journal of the Royal Statistical Society: Series D (The Statistician) 52*(3), 381–393.

Kawakatsu, H. (2006). Matrix Exponential GARCH. *Journal of Econometrics 134*(1), 95–128.

Keating, J. W. (2000). Macroeconomic modeling with asymmetric vector autoregressions. *Journal of Macroeconomics 22*(1), 1–28.

Kim, W.-K. and B. Lynne (1997). Existence condition for the stationary ergodic new Laplace autoregressive model of order p-NLAR (p). *Journal of the Korean Statistical Society 26*(4), 521–530.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. In *Advances in Markov-switching models*, pp. 223–254. Springer.

Kocherlakota, S. and K. Kocherlakota (2001). Regression in the bivariate Poisson distribution.

Kotz, S., N. Balakrishnan, and N. L. Johnson (2004). *Continuous multivariate distributions, Volume 1: Models and applications*, Volume 1. John Wiley & Sons.

Krishnan, B. and D. George (2017). A mixed asymmetric Laplace moving average process. *Global J. Pure Appl. Math 13*, 6919–6928.

Kwan, C., W. Li, and K. Ng (2005). A multivariate threshold GARCH model with time-varying correlations. *Econometric Reviews 1180.*

Lakshminarayana, J., S. Pandit, and K. Srinivasa Rao (1999). On a bivariate Poisson distribution. *Communications in Statistics-Theory and Methods 28*(2), 267–276.

Lamoureux, C. G. and W. D. Lastrapes (1990). Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics 8*(2), 225–234.

Lange, T. and A. Rahbek (2009). An introduction to regime switching time series models. In *Handbook of Financial Time Series*, pp. 871–887. Springer.

Lanne, M. and P. Saikkonen (2005). Non-linear GARCH models for highly persistent volatility. *The Econometrics Journal 8*(2), 251–276.

Lanne, M. and P. Saikkonen (2007). A multivariate generalized orthogonal factor GARCH model. *Journal of Business & Economic Statistics 25*(1), 61–75.

Latour, A. (1997). The multivariate GINAR (p) process. *Advances in Applied Probability 29*(1), 228–248.

Lawrance, A. (1980). The mixed Exponential solution to the first-order autoregressive model. *Journal of Applied Probability 17*(2), 546–552.

Lawrance, A. and P. Lewis (1980). The Exponential autoregressive-moving average EARMA (p, q) process. *Journal of the Royal Statistical Society: Series B (Methodological) 42*(2), 150–161.

Lawrance, A. and P. Lewis (1981). A new autoregressive time series model in Exponential variables (NEAR (1)). *Advances in Applied Probability 13*(4), 826–845.

Lawrance, A. and P. Lewis (1985). Modelling and residual analysis of nonlinear autoregressive time series in Exponential variables. *Journal of the Royal Statistical Society: Series B (Methodological) 47*(2), 165–183.

Le, N. D., R. D. Martin, and A. E. Raftery (1996). Modeling flat stretches, bursts outliers in time series using mixture transition distribution models. *Journal of the American Statistical Association 91*(436), 1504–1515.

Lee, J. and E. Hwang (2018). A generalized regime-switching integer-valued GARCH (1, 1) model and its volatility forecasting. *Communications for Statistical Applications and Methods 25*(1), 29–42.

Lee, S., Y. Lee, and C. W. Chen (2016). Parameter change test for zero-inflated generalized Poisson Autoregressive models. *Statistics 50*(3), 540–557.

Lescheb, I. (2015). On stationarity of the periodic AGARCH processes. *Random Operators and Stochastic Equations 23*(1), 31–37.

Lewis, P. A., E. McKenzie, and D. K. Hugus (1989). Gamma processes. *Stochastic Models 5*(1), 1–30.

Li, C. and W. K. Li (1996). On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Econometrics 11*(3), 253–274.

Liang, F., R. Paulo, G. Molina, M. A. Clyde, and J. O. Berger (2008). Mixtures of g-priors for Bayesian variable selection. *Journal of the American Statistical Association 103*(481), 410–423.

Liboschik, T., K. Fokianos, and R. Fried (2015). *tscount: An R package for analysis of count time series following generalized linear models.* Universitätsbibliothek Dortmund Dortmund, Germany.

Ling, S. and M. McAleer (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics 106*(1), 109–117.

Liu, H. (2012). *Some models for time series of counts.* Ph. D. thesis, Columbia University.

Liu, J.-C. (2006). Stationarity of a Markov-switching GARCH model. *Journal of Financial Econometrics 4*(4), 573–593.

Liu, T. and X. Yuan (2013). Random rounded integer-valued autoregressive conditional heteroskedastic process. *Statistical Papers 54*(3), 645–683.

Lodewyckx, T., W. Kim, M. D. Lee, F. Tuerlinckx, P. Kuppens, and E.-J. Wagenmakers (2011). A tutorial on Bayes factor estimation with the product space method. *Journal of Mathematical Psychology 55*(5), 331–347.

Lu, Z. (1996). A note on geometric ergodicity of autoregressive conditional heteroscedasticity (ARCH) model. *Statistics & probability letters 30*(4), 305–311.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer Science & Business Media.

MacDonald, I. L. and W. Zucchini (1997). *Hidden Markov and other models for discrete-valued time series*, Volume 110. CRC Press.

Mardia, K. V. (1970). *Families of bivariate distributions.* Number 27. Lubrecht & Cramer Limited.

Marshall, A. W. and I. Olkin (1967). A generalized bivariate Exponential distribution. *Journal of Applied Probability 4*(2), 291–302.

McCabe, B. and G. Martin (2005). Bayesian predictions of low count time series. *International Journal of Forecasting 21*(2), 315–330.

McDonald, J. B. (1991). Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. *Economics Letters 37*(3), 273–278.

McDonald, J. B. (2008). Some generalized functions for the size distribution of income. In *Modeling income distributions and Lorenz curves*, pp. 37–55. Springer.

McKenzie, E. (1985). Some simple models for discrete variate time series 1. *JAWRA Journal of the American Water Resources Association 21*(4), 645–650.

Medeiros, M. C. and A. Veiga (2004). Modeling multiple regimes in financial volatility with a flexible coefficient GARCH model. Technical report, Texto para discussão.

Meintanis, S. G. and D. Karlis (2014). Validation tests for the innovation distribution in INAR time series models. *Computational Statistics 29*(5), 1221–1241.

Meintanis, S. G., B. Milošević, and M. Obradović (2020). Goodness-of-fit tests in conditional duration models. *Statistical Papers 61*(1), 123–140.

Meitz, M. and T. Teräsvirta (2006). Evaluating models of autoregressive conditional duration. *Journal of Business & Economic Statistics 24*(1), 104–124.

Meyn, S. P. and R. Tweedie (1994). State-dependent criteria for convergence of Markov chains. *The Annals of Applied Probability*, 149–168.

Meyn, S. P. and R. L. Tweedie (1992). Stability of markovian processes i: Criteria for discrete-time chains. *Advances in Applied Probability 24*(3), 542–574.

Meyn, S. P. and R. L. Tweedie (1993). Stability of Markovian processes II: Continuous-time processes and sampled chains. *Advances in Applied Probability 25*(3), 487–517.

Meyn, S. P. and R. L. Tweedie (2012). *Markov chains and stochastic stability.* Springer Science & Business Media.

Miralles-Marcelo, J. L., J. L. Miralles-Quirós, and M. d. M. Miralles-Quirós (2013). Improving the carr model using extreme range estimators. *Applied Financial Economics 23*(21), 1635–1647.

Miravete, E. (2009a, 01). Multivariate Sarmanov count data models. *C.E.P.R. Discussion Papers, CEPR Discussion Papers*.

Miravete, E. J. (2009b). Multivariate Sarmanov count data models.

M'Raïhi, D., D. Naccache, and M. Tunstall (2000). Asymmetric currency rounding. In *International Conference on Financial Cryptography*, pp. 192–201. Springer.

Narayan, P. K. and R. Liu (2015). A unit root model for trending time-series energy variables. *Energy Economics 50*, 391–402.

Nelsen, R. B. (1986). Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics-Theory and Methods 15*(11), 3277–3285.

Nelsen, R. B. (2007). *An introduction to copulas.* Springer Science & Business Media.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Nguyen, H. D., G. J. McLachlan, J. F. Ullmann, and A. L. Janke (2016). Laplace mixture autoregressive models. *Statistics & Probability Letters 110*, 18–24.

Opschoor, A., A. Lucas, I. Barra, and D. Van Dijk (2021). Closed-form multi-factor copula models with observation-driven dynamic factor loadings. *Journal of Business & Economic Statistics 39*(4), 1066–1079.

Pedeli, X. and D. Karlis (2013). On composite likelihood estimation of a multivariate INAR (1) model. *Journal of Time Series Analysis 34*(2), 206–220.

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics 131*(1-2), 445–473.

Piancastelli, L. S., W. Barreto-Souza, and H. Ombao (2020). Flexible bivariate INGARCH process with a broad range of contemporaneous correlation. *Journal of Time Series Analysis*.

Piancastelli, L. S., W. Barreto-Souza, and H. Ombao (2023). Flexible bivariate ingarch process with a broad range of contemporaneous correlation. *Journal of Time Series Analysis 44*(2), 206–222.

Piancastelli, L. S., N. Friel, W. Barreto-Souza, and H. Ombao (2023). Multivariate conway-maxwell-poisson distribution: Sarmanov method and doubly intractable bayesian inference. *Journal of Computational and Graphical Statistics 32*(2), 483–500.

Piperigou, V. E. (2009). Discrete distributions in the extended FGM family: The pgf approach. *Journal of Statistical Planning and Inference 139*(11), 3891–3899.

Pitt, M. K., C. Chatfield, and S. G. Walker (2002). Constructing first order stationary autoregressive models via latent processes. *Scandinavian Journal of Statistics 29*(4), 657–663.

Pitt, M. K. and S. G. Walker (2005). Constructing stationary time series models using auxiliary variables with applications. *Journal of the American Statistical Association 100*(470), 554–564.

Pitt, M. K. and S. G. Walker (2006). Extended constructions of stationary autoregressive processes. *Statistics & Probability Letters 76*(12), 1219–1224.

Polasek, W. and J. Pai (2004). CPO plots for ARMA model selection. *Central European Journal of Operations Research 12*(2).

Rabemananjara, R. and J.-M. Zakoian (1993). Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics 8*(1), 31–49.

Raftery, A. E., D. Madigan, and J. A. Hoeting (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association 92*(437), 179–191.

Rahbek, A. (2003). Stochastic properties of multivariate time series equations with emphasis on ARCH. *IFAC Proceedings Volumes 36*(16), 227–232.

Rao, P. S. and D. H. Johnson (1988). A first-order AR model for non-Gaussian time series. In *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, pp. 1534–1535. IEEE Computer Society.

Reed, W. J. and M. Jorgensen (2004). The double Pareto-lognormal distribution—a new parametric model for size distributions. *Communications in Statistics-Theory and Methods 33*(8), 1733–1753.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences of the United States of America 42*(1), 43.

Rydberg, T. H. and N. Shephard (2000). Bin models for trade-by-trade data. *Modelling the number of trades in fixed interval of time. Paper 740*, 28.

Rydén, T., T. Teräsvirta, and S. Åsbrink (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics 13*(3), 217–244.

Saikkonen, P. (2007). Stability of mixtures of vector autoregressions with autoregressive conditional heteroskedasticity. *Statistica Sinica 17*(1), 221–239.

Sarmanov, O. V. (1966). Generalized normal correlation and two-dimensional frechet classes. In *Doklady Akademii Nauk*, Volume 168, pp. 32–35. Russian Academy of Sciences.

Schweer, S. and C. H. Weiß (2014). Compound Poisson INAR (1) processes: stochastic properties and testing for overdispersion. *Computational Statistics & Data Analysis 77*, 267–284.

Schweizer, B. (1991). Thirty years of copulas. In *Advances in probability distributions with given marginals*, pp. 13–50. Springer.

Sentana, E. (1995). Quadratic ARCH models. *The Review of Economic Studies 62*(4), 639–661.

Sentana, E. (1998). The relation between conditionally heteroskedastic factor models and factor GARCH models. *The Econometrics Journal 1*(2), 1–9.

Shephard, N. (1995). Generalized linear autoregressions. Economics Papers 8., Economics Group, Nuffield College, University of Oxford.

Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics 25*(2), 197–231.

Shmueli, G., T. P. Minka, J. B. Kadane, S. Borle, and P. Boatwright (2005). A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 54*(1), 127–142.

Shubina, M. and M.-L. T. Lee (2004). On maximum attainable correlation and other measures of dependence for the Sarmanov family of bivariate distributions. *Communications in Statistics-Theory and Methods 33*(5), 1031–1052.

Silva, F. A. C. d. (2016). *Compound Poisson integer-valued GARCH processes*. Ph. D. thesis, 00500:: Universidade de Coimbra.

Silvennoinen, A. and T. Teräsvirta (2005). Multivariate autoregressive conditional heteroskedasticity with smooth transitions in conditional correlations. Technical report, SSE/EFI Working Paper Series in Economics and Finance.

Silvennoinen, A. and T. Teräsvirta (2009a). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model. *Journal of Financial Econometrics 7*(4), 373–411.

Silvennoinen, A. and T. Teräsvirta (2009b). Multivariate GARCH models. In *Handbook of financial time series*, pp. 201–229. Springer.

Sim, C.-H. (1990). First-order autoregressive models for Gamma and Exponential processes. *Journal of Applied Probability 27*(2), 325–332.

Sim, C. H. (1993). First-order autoregressive logistic processes. *Journal of Applied Probability 30*(2), 467–470.

Sin, C.-Y. C. (2013). Using CARRX models to study factors affecting the volatilities of asian equity markets. *The North American Journal of Economics and Finance 26*, 552–564.

Song, P. X.-K., Y. Fan, and J. D. Kalbfleisch (2005). Maximization by parts in likelihood inference. *Journal of the American Statistical Association 100*(472), 1145–1158.

Straumann, D. (2005). Maximum likelihood estimation in conditionally heteroscedastic time series models. *Estimation in Conditionally Heteroscedastic Time Series Models*, 141–168.

Streett, S. B. (2000). *Some observation driven models for time series*. Colorado State University.

Tauchen, G. E. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica: Journal of the Econometric Society*, 485–505.

Taylor, S. J. (2008). *Modelling financial time series.* World Scientific.

Teräsvirta, T. (2009). An introduction to univariate GARCH models. In *Handbook of Financial Time Series*, pp. 17–42. Springer.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *The Annals of Statistics*, 1701–1728.

Ting Lee, M.-L. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Communications in Statistics-Theory and Methods 25*(6), 1207–1222.

Tjøstheim, D. (2012). Some recent theory for autoregressive count time series. *Test 21*(3), 413–438.

Tjøstheim, D. and K. O. Hufthammer (2013). Local Gussian correlation: A new measure of dependence. *Journal of Econometrics 172*(1), 33–48.

Tomy, L. and K. K. Jose (2009). Generalized Normal-Laplace AR process. *Statistics & probability letters 79*(14), 1615–1620.

Tong, H. (1990). *Non-linear time series: a dynamical system approach.* Oxford university press.

Toulemonde, G., A. Guillou, P. Naveau, M. Vrac, and F. Chevallier (2010). Autoregressive models for maxima and their applications to CH4 and N2O. *Environmetrics: The Official Journal of the International Environmetrics Society 21*(2), 189–207.

Trindade, A. A., Y. Zhu, and B. Andrews (2010). Time series models with asymmetric Laplace innovations. *Journal of Statistical Computation and Simulation 80*(12), 1317–1333.

Truong, B.-C., C. W. Chen, and S. Sriboonchitta (2017). Hysteretic Poisson INGARCH model for integer-valued time series. *Statistical Modelling 17*(6), 401–422.

Tse, Y. K. and A. K. C. Tsui (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics 20*(3), 351–362.

Tweedie, R. (1976). Criteria for classifying general Markov chains. *Advances in Applied Probability 8*(4), 737–771.

Van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics 17*(5), 549–564.

Vernic, R. (1997). On the bivariate generalized Poisson distribution. *ASTIN Bulletin: The Journal of the IAA 27*(1), 23–32.

Vernic, R. (2020). On a class of bivariate mixed sarmanov distributions. *Australian & New Zealand Journal of Statistics 62*(2), 186–211.

Vrontos, I. D., P. Dellaportas, and D. N. Politis (2003). A full-factor multivariate GARCH model. *The Econometrics Journal 6*(2), 312–334.

Wang, K.-L., C. Fawson, C. B. Barrett, and J. B. McDonald (2001). A flexible parametric GARCH model with an application to exchange rates. *Journal of Applied Econometrics 16*(4), 521–536.

Wang, X., D. Wang, and H. Zhang (2020). Poisson autoregressive process modeling via the penalized conditional maximum likelihood procedure. *Statistical Papers 61*(1), 245–260.

Wang, Z. (2011). One mixed Negative Binomial distribution with application. *Journal of Statistical Planning and Inference 141*(3), 1153–1160.

Wong, C. S., W.-S. Chan, and P. Kam (2009). A Student t-mixture autoregressive model with applications to heavy-tailed financial data. *Biometrika 96*(3), 751–760.

Wong, C. S. and W. K. Li (2000). On a mixture autoregressive model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 62*(1), 95–115.

Xie, H. and X. Wu (2017). A conditional autoregressive range model with Gamma distribution for financial volatility modelling. *Economic Modelling 64*, 349–356.

Xie, M., B. He, and T. Goh (2001). Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis 38*(2), 191–201.

Xu, H.-Y., M. Xie, T. N. Goh, and X. Fu (2012). A model for integer-valued time series with conditional overdispersion. *Computational Statistics & Data Analysis 56*(12), 4229–4242.

Yang, J., M. Xie, and T. N. Goh (2011). Outlier identification and robust parameter estimation in a zero-inflated Poisson model. *Journal of Applied Statistics 38*(2), 421–430.

Yatigammana, R. P., S. Choy, and J. S. Chan (2016). Autoregressive conditional duration model with an extended Weibull error distribution. In *Causal Inference in Econometrics*, pp. 83–107. Springer.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and control 18*(5), 931–955.

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika 75*(4), 621–629.

Zeger, S. L. and B. Qaqish (1988). Markov regression models for time series: a quasi-likelihood approach. *Biometrics*, 1019–1031.

Zhang, M. Y., J. R. Russell, and R. S. Tsay (2001). A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics 104* (1), 179–207.

Zhu, F. (2011). A Negative Binomial integer-valued GARCH model. *Journal of Time Series Analysis 32* (1), 54–67.

Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *Journal of Mathematical Analysis and Applications 389* (1), 58–71.

Zhu, F. (2012b). Modeling time series of counts with COM-Poisson INGARCH models. *Mathematical and Computer Modelling 56* (9-10), 191–203.

Zhu, F. (2012c). Zero-inflated Poisson and Negative Binomial integer-valued GARCH models. *Journal of Statistical Planning and Inference 142* (4), 826–839.

Zhu, F., Q. Li, and D. Wang (2010). A mixture integer-valued ARCH model. *Journal of Statistical Planning and Inference 140* (7), 2025–2036.