



Facultad de Ciencias

**Assessing normality in stationary
stochastic processes**

Evaluación de normalidad en procesos
estocásticos estacionarios

Trabajo de Fin de Máster
para acceder al

MÁSTER EN MATEMÁTICA Y COMPUTACIÓN

Autor: Izhar Asael Alonzo Matamoros

Directora: Alicia Nieto-Reyes

octubre - 2020

Acknowledgments

First of all, I would like to thank to *Fundación Carolina* for providing the scholarship opportunity and funding this work. Second, to my professor and advisor *Alicia Nieto-Reyes* for give me this opportunity, it was the chance that I need to work in a statistical programming field. To Professor *Rob Hyndman* for his implementation of the **forecast** package in which the **nortsTest** package's functionality is based on. Enrique Rivera Gomez for his help in the edition of this work. To my classmate and friend Angel Ornella for his unconditional help in all this year

And a special thank to my mother *Mireya Matamoros* for her help in every problem I have encountered so far. My life, successes and work are just a reflection of her guide.

Abstract

Normality (a set of observations being sampled from a Gaussian process) is an important assumption in a wide variety of statistical models. Therefore, developing procedures for testing this assumption is a topic that has gained popularity over several years. Extensive literature exists on goodness of fit tests for normality under the assumption of independent identical distributed random variables. However, this is not the case for the context of stationary stochastic process, case in which the independence assumption is violated. For this matter, several tests have been proposed over the years. The aim of this work is to present a discussion and references of the most common tests for normality in stationary processes, such as *Epps, Lobato and Velasco, the random projections, and Psaradakis and Vavra*. For assessing model adequacy in a Bayesian approach, we propose an alternative methodology for checking model's assumptions inspired by the random projection results with promising results, in all the designed case studies. Additionally we present our implemented **nortsTest** package, an R package that performs all the reviewed tests mentioned above.

Keywords: *Hypothesis test, stochastic process, stationarity, Gaussian process, Model's diagnostic.*

Resumen

Normalidad (un conjunto de observaciones son muestreadas de un Proceso Gaussiano) es un supuesto importante en una gran cantidad de modelos estadísticos. Debido a eso, desarrollar procedimientos para corroborar estos supuestos es un tema que ha ganado popularidad en los últimos años. Existe una gran cantidad de literatura para el caso de variables aleatorias independientes e idénticamente distribuidas, pero, este no es el caso en el contexto de procesos estocásticos estacionarios, donde el supuesto de independencia no se mantiene. Algunas pruebas de hipótesis han sido propuestas a través de los años para resolver esta problemática. El objetivo de este trabajo es presentar una discusión de las pruebas más utilizadas para probar normalidad en procesos estacionarios, tales como *Epps, Lobato y Velasco, las proyecciones aleatorias, y Psaradakis y Vavra*. Para diagnóstico de modelos en un enfoque Bayesiano, proponemos una metodología alternativa para la corroboración de supuestos inspirada en los resultados del método de las proyecciones aleatorias con prometedores resultados. Adicionalmente, presentamos nuestro implementado paquete **nortsTest**, un paquete en R que realiza las pruebas mencionadas anteriormente.

Palabras clave: *Pruebas de hipótesis, procesos estocásticos, estacionaridad, procesos Gaussianos, diagnóstico de modelos.*

Contents

Introduction	5
Preliminary Concepts	6
Stochastic process	6
Bayesian inference	9
Introduction to Markov Chain Monte Carlo methods	9
Normality tests for stationary processes	12
Software	12
Test for stationarity	13
Unit root tests	13
Seasonal unit root tests	13
The Epps' test	14
The Lobato and Velasco's test	16
The random projection test	17
The Psaradakis and Vavra's test	19
Simulations and data analysis	22
Numerical experiments	22
Real data application	24
Model assessment in Bayesian Methods	27
The random projections method for Bayesian model assessing	27
Case studies	28
Case 1: stationary Gamma AR(1) process	28
Case 2: stationary Gaussian MA(1) process	29
Conclusion	31
Future work and projects	31

List of Tables

1	Rejection rates estimates, part 1	22
2	Rejection rates estimates, part 2	23
3	Case 1: Estimated values of M_1 for a Gamma AR(1) process	29
4	Case 2: Estimated values of M_2 for a Gaussian MA(1) process	30

List of Figures

1	CO2 levels at Mauna Loa	24
2	Check residuals plot	26
3	Case 1: Diagnostic plots for projected errors in M_1 model for Gamma AR(1) process	30
4	Case 2: Diagnostic plots for projected errors in M_2 model for Gaussian AR(1) process	31

Introduction

Normality (*A set of observations are sampled from a Gaussian process*) is an important assumption in a wide variety of statistical models. Therefore, developing procedures for testing these assumptions is a problem that has gained popularity over several years. Most of the existing literature and implementations are for independent identically distributed data (D'Agostino and Stephens 1986), but there are no results showing that these tests are consistent in the context of stationary processes. Moreover, Gasser (1975) provides a simulation study where *Pearson's test* had an excessive rejection of the alternative hypothesis for dependent data. For this matter, a small number of tests have been proposed over the years, but as far of our knowledge, there are no consistent implementation of all these tests. The aim of this work is to present a review of the implemented tests for normality, providing an general framework of the four most important tests. And introduce our proposed methods for model assessing in time-series analysis.

The four presented tests are discussed in chronological order; (i) the *Epps test* (Epps 1987) based on the characteristic function, (ii) the corrected *Skewness-Kurtosis* (SK) test implemented by Lobato and Velasco (2004), (iii) the *random projection test* proposed by Nieto-Reyes, Cuesta-Albertos, and Gamboa (2014) and (iv) the *Psadarakis and Vávra test* (Psaradakis and Vávra 2017) that uses a bootstrap approximation of the Anderson and Darling (1952) test statistic for stationary linear processes. Additionally, this work introduces our implemented **nortsTest** package. The first R package for testing normality in stationary processes, that provides implementations of all the previously mentioned tests. For assessing model adequacy we present the `check_residual()` function for checking the assumptions in time-series models, it returns a report of tests for stationarity, seasonality, and normality, as diagnostic plots for visual check. And in a Bayesian approach, we propose an alternative method for checking the model's assumptions inspired in the random projections methods of Cuesta-Albertos, del Barrio, Fraiman, and Matrán (2007). The main procedure is, estimate the model's errors and projected in the real line, if the projected errors are normally distributed then, the model's normality assumptions are satisfied.

Section 1 provides the theoretical background, including preliminary concepts and results. Section 2 introduces the normality tests for stationary processes, each subsection introducing a test framework and including examples of the tests functions with simulated data. Section 3 provides numerical experiments with simulated data and a real data application. In Subsection 3.1 reports a simulation study for all the implemented tests, and Subsection 3.2 shows an application of the normality tests for model diagnostic in Time series analysis. Section 4 introduces our proposal method for model diagnostic in Bayesian methods, showing the method's procedure, and presenting two study cases with simulated data. Finally, section 5 discusses the package functionality and provides our conclusions. Furthermore, we include our future work on the package.

1. Preliminary concepts

This section provides some theoretical aspects of stochastic processes and Bayesian inference that are a necessary theoretical framework for the following sections. The presented definitions and results were obtained from [Shumway and Stoffer \(2010\)](#), [Tsay \(2010\)](#) and [Migon, Gamerman, and Louzada \(2014\)](#).

1.1. Stochastic process

For the purpose of this work, T is a set of real values denoted as time, $T \subseteq \mathbb{R}$, for instance $T = \mathbb{N}$ or $T = \mathbb{Z}$, the naturals or integer numbers respectively. We denote by $X := \{X_t\}_{t \in T}$ a *stochastic process* with X_t a real random variable for each $t \in T$. Following this notation, a *time-series* is just a finite collection of ordered observations of X ([Shumway and Stoffer 2010](#)). An important measure for a stochastic process is its mean function $\mu(t) := E[X_t]$ for each $t \in T$, where $E[\cdot]$ denotes the usual expected value of a random variable. A generalization of this measure is the k -th order centered moment function $\mu_k(t) := E[(X_t - \mu(t))^k]$ for each $t \in T$ and $k > 1$; with the process variance function the second order centered moment, $\sigma^2(t) := \mu_2(t)$. Other important indicators are the auto-covariance and auto-correlation functions, which measure the linear dependency between two different time points of a given process. For any $t, s \in T$, they are, respectively,

$$\gamma(t, s) := E[(X_t - \mu(t))(X_s - \mu(s))] \text{ and } \rho(t, s) := \frac{\gamma(t, s)}{\sqrt{\mu_2(t)}\sqrt{\mu_2(s)}}.$$

Other widely used indicator functions for the analysis of processes are the skewness and kurtosis functions, defined for each $t \in T$ as $s(t) := \mu_3(t)/[\mu_2(t)]^{3/2}$ and $k(t) := \mu_4(t)/[\mu_2(t)]^2$ respectively.

A generally used assumption for stochastic processes is stationarity. It has a key role in forecasting procedures of classic time-series modeling ([Tsay 2010](#)) or as a principal assumption in de-noising methods for signal theory ([Wasserman 2006](#)).

Definition 1 *A stochastic process X is said to be strictly stationary if, for every collection $\tau = \{t_1, t_2, \dots, t_k\} \subset T$ and $h > 0$, the joint distribution of $\{X_t\}_{t \in \tau}$ is identical to that of $\{X_{t+h}\}_{t \in \tau}$.*

The previous definition is strong for applications. A milder version of it, which makes use of the process first two moments, is weak stationarity.

Definition 2 *A stochastic process X is said to be weakly stationary if its mean function is constant in time, $\mu(t) = \mu$, its auto-covariance function only depends on the difference between times, $\gamma(s, t) = \sigma|t - s|$ for a $\sigma \in \mathbb{R}$, and it has a finite variance function, $\mu_2(t) = \mu_2 < \infty$.*

For the rest of this work, the term *stationary* will be used to specify a weakly stationary process. A direct consequence of the stationarity assumption is that the previous indicator functions get simplified. Thus, given a stationary stochastic process X , its mean function, k -th order centered moment, for $k > 1$, and auto-covariance function are respectively,

$$\mu = E[X_{t_1}], \mu_k = E[(X_{t_1} - \mu)^k] \text{ and } \gamma(h) = E[(X_{t_1+h} - \mu)(X_{t_1} - \mu)],$$

which are independent of $t_1 \in T$.

Given a sample x_1, \dots, x_n , $n \in \mathbb{N}$, of equally spaced observations of X , their corresponding estimators, sample mean, sample k -th order centered moment and sample auto-covariance, are respectively

$$\hat{\mu} := n^{-1} \sum_{i=1}^n x_i, \hat{\mu}_k := n^{-1} \sum_{i=1}^n (x_i - \hat{\mu})^k \text{ and } \hat{\gamma}(h) := n^{-1} \sum_{i=1}^{n-h} (x_{i+h} - \hat{\mu})(x_i - \hat{\mu}).$$

A particular case in which stationarity implies strictly stationarity are Gaussian processes.

Definition 3 *A stochastic process X is said to be a Gaussian process if for every finite collection $\tau = \{t_1, t_2, \dots, t_k\} \subset T$, the joint distribution of $\{X_t\}_{t \in \tau}$ has a multivariate normal distribution.*

A series of mean zero uncorrelated random variables with finite constant variance is known as *white noise*. If additionally, it is formed of independent and identically distributed (i.i.d) normal random variables, it is known as *Gaussian white noise*; which is a particular case of stationary Gaussian process. For the rest of the work, $X_t \sim N(\mu, \sigma^2)$ denotes that the random variable X_t is normally distributed with mean μ and variance σ^2 and $\chi^2(v)$ denotes the chi square distribution with v degree freedom.

Other classes of stochastic processes can be defined using collections of white noise, for instance, the linear process.

Definition 4 *Let X be a stochastic process. X is said to be linear if it can be written as*

$$X_t = \mu + \sum_{i \in \mathbb{Z}} \phi_i \epsilon_{t-i},$$

where $\{\epsilon_i\}_{i \in \mathbb{Z}}$ is a collection of white noise random variables and $\{\phi_i\}_{i \in \mathbb{Z}}$ is a set of real values such that $\sum_{i \in \mathbb{Z}} |\phi_j| < \infty$.

An important class of processes is the *auto-regressive moving average (ARMA)*. [Box and Jenkins \(1990\)](#) introduced it for time series analysis and forecast, becoming very well-known in the 90s and early 21st century.

Definition 5 *For any non-negative integers p, q , a stochastic process X is an $ARMA(p, q)$ process if it is a stationary process and*

$$X_t = \sum_{i=0}^p \phi_i X_{t-i} + \sum_{i=0}^q \theta_i \epsilon_{t-i}, \tag{1}$$

where $\{\phi_i\}_{i=1}^p$ and $\{\theta_i\}_{i=0}^q$ are sequences of real values with $\phi_0 = 0$, $\phi_p \neq 0$, $\theta_0 = 1$ and $\theta_q \neq 0$ and $\{\epsilon_i\}_{i \in \mathbb{Z}}$ is a collection of white noise random variables.

Particular cases of *ARMA* processes are the auto-regressive ($AR(p) := ARMA(p, 0)$) and the mean average ($MA(q) := ARMA(0, q)$) processes. Additionally, a *random-walk* is a non-stationary process satisfying (1) with $p = 1$, $\phi_1 = 1$ and $q = 0$. Several properties of an

ARMA process can be extracted from its structure. For that, the AR and MA polynomials are introduced

$$AR: \phi(z) = 1 - \sum_{i=0}^p \phi_i z^i \text{ and } MA: \theta(z) = \sum_{i=0}^q \theta_i z^i,$$

where z is a complex number and, as before, $\phi_0 = 0$, $\phi_p \neq 0$, $\theta_0 = 1$ and $\theta_q \neq 0$. Conditions for stationarity, order selection and process behavior are properties studied from these two polynomials.

For modeling volatility in financial data [Bollerslev \(1986\)](#) proposed the *generalized auto-regressive conditional heteroscedastic* (GARCH) class of processes as a generalization of the *auto-regressive conditional heteroscedastic* (ARCH) processes ([Engle 1982](#)).

Definition 6 For any $s, k \in \mathbb{N}$, a stochastic process X is a GARCH(s, k) process if it has the following structure

$$X_t = \mu + \sigma_t \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^s \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^k \beta_i \sigma_{t-i}^2,$$

where s and k are non-negative integers, μ is the process mean, $\{\alpha_i\}_{i=1}^s$ and $\{\beta_i\}_{i=1}^k$ are non-negative sequences of real values, σ_0 is a positive constant value, and $\{\epsilon_t\}_{t \in T}$ is a collection of white noise random variables.

A more general class of processes are the *state-space models* (SSMs), which have gained popularity over the years as techniques for smoothing ([West and Harrison 2006](#)) and forecasting ([Hyndman and Khandakar 2008](#)) in time series analysis.

Definition 7 Let X be an stochastic process and let be $\theta_t \in \Theta \subset \mathbb{R}^d$ $d > 0$, a vector of parameters defined in a parametric space Θ for every $t \in T$, then X is a SSM process if satisfies

$$X_t = f(\theta_{t-1}) + r(\theta_{t-1})\epsilon_t$$

$$\theta_t = g(\theta_{t-1}) + w(\theta_{t-1})\epsilon_t,$$

where f, g, r and w are functional over Θ , and $\{\epsilon_t\}_{t \in T}$ is a collection of white noise random variables denoted as errors.

A particular type of STMs that analyzes the level, trend and seasonal components of the process is known as *error, trend, and seasonal* (ETS) models. There are over 32 different variations of the ETS models ([Hyndman, Koehler, Ord, and Snyder 2008](#)). One of them is the *ETS(M, A, A)* model.

Definition 8 A SSM process X follows an *ETS(M, A, A)* model, if the process accepts

$$X_t = [L_{t-1} + T_{t-1} + S_{t-1}](1 + \epsilon_t)$$

as innovation equation and

$$\begin{aligned}
L_t &= L_{t-1} + T_{t-1} + \alpha(L_{t-1} + T_{t-1} + S_{t-m})\epsilon_t \\
T_t &= T_{t-1} + \beta(L_{t-1} + T_{t-1} + S_{t-m})\epsilon_t \\
S_t &= S_{t-m} + \gamma(L_{t-1} + T_{t-1} + S_{t-m})\epsilon_t,
\end{aligned}$$

as state equations. $\alpha, \beta, \gamma \in [0, 1]$, $m \in \mathbb{N}$ denotes the period of the series and $\{\epsilon_t\}$ is a collection of white noise random variables denoted as errors. For each $t \in \mathbb{Z}$, L_t , T_t and S_t represent respectively the level, trend and seasonal component.

1.2. Bayesian inference

Let $Y := \{Y_1, Y_2, \dots, Y_n\}$ a finite collection of random variables with common distribution P depending on a vector of unknowns parameters $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ assuming values in a set denoted by $\Theta \subseteq \mathbb{R}^d$, for $d > 0$. The likelihood function of θ :

$$l(\theta; Y) : \Theta \rightarrow \mathbb{R}^+,$$

is a function that associates the value $P(Y|\theta)$ to each θ . Therefore, for given values of Y , the likelihood is computed as follows:

$$L(\theta; Y) = \prod_{i=1}^n P(Y_i | Y_{i-1}, Y_{i-2}, \dots, Y_2, Y_1, \theta),$$

note that in a *frequentist approach*, the likelihood is sufficient to compute adequate estimates of the unknown parameters θ , but in a *Bayesian scheme*, additional information can be incorporated using a *prior distribution* $P(\theta)$, and the unknown quantities can be obtained using the Bayes theorem:

$$P(\theta|Y) = \frac{P(\theta, Y)}{P(Y)} = \frac{l(\theta; Y)P(\theta)}{P(Y)}, \quad (2)$$

Where $P(\theta|Y)$ is denoted as the *posterior distribution* of θ , $l(\theta; Y)$ is the likelihood function, and $P(Y)$ does not depend of θ , then, it can be removed from (2) as constant. Therefore, the Bayes theorem can be written as:

$$P(\theta|Y) \propto l(\theta; Y)P(\theta). \quad (3)$$

Then, a numeric estimator for θ is estimated using the expected value of the posterior distribution as

$$E[\theta|Y] = \int_{\Theta} \theta P(\theta|Y) d\theta,$$

denoted as the *posterior mean* of θ . A relevant aspect of a Bayesian scheme is that it allows to obtain automatic prediction for future observations. To predict a new value Y_{n+1} , whose probabilistic description is $P(Y_{n+1}|\theta)$, then

$$P(Y_{n+1}|Y) = \int_{\Theta} P(Y_{n+1}, \theta|Y) d\theta = \int_{\Theta} P(Y_{n+1}|\theta)P(\theta|Y) d\theta,$$

where $P(Y_{n+1}|Y)$ is denoted as the predictive distribution.

1.3. Introduction to Markov Chain Monte Carlo methods

A Bayesian approach for inference has become very popular over the last years, because of its advantage of incorporate prior knowledge about the parameters into a model. But, its practical use was limited because for most of the problems an analytical solution for the posterior can not be found, and an approximate solution might be a computational challenge. In the last decades, the Markov chain Monte Carlo (MCMC) algorithms (Brooks, Gelman, Jones, and Meng 2011) allow an approximation of the solution by drawing random samples from the posterior distribution, but for most of the problems, this approach was too time-consuming.

MCMC methods allow to simulate dependent realizations of a stochastic process, such that these simulations form an irreducible Markov chain that has the target distribution (i.e the posterior) as its stationary distribution. The MCMC has origin with Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953), they propose this method for simulating a liquid in equilibrium with its gas phase. Hastings (1970) generalize Metropolis algorithm using asymmetric distributions. In Gelfand and Smith (1990) introduce a particular case of the Metropolis-Hastings known as the Gibbs Sampler. Over the years, several adaptations of the Metropolis-Hastings algorithm such as Reversible Jump MCMC, Sequential Monte Carlo, and Hamiltonian Monte Carlo (HMC) Duane (1987), have been proposed for improving the speed of convergences.

Let's present the original Metropolis-Hastings algorithm for Bayesian Inference. For a finite collection of random variables X with unknown parameter θ , and $P(\theta)$ a selected prior, let's denote as:

- m the number of simulations
- $f(\theta) = P(X/\theta)P(\theta)$ is the proposal function for the posterior distribution.
- $J(\theta_t/\theta_{t-1})$ a predefined Markov Chain's jump distribution.
- The Metropolis-Hastings ratio

$$r = \frac{f(\theta_*)J(\theta_{t-1}/\theta_*)}{f(\theta_{t-1})J(\theta_*/\theta_{t-1})}$$

Then the Metropolis-Hastings algorithm is shown in Algorithm 1, its working principle is as follows: for an arbitrary value θ_0 a Markov Chain $\{\theta_i\}_{i=1}^m$ is generated using a transition function J with stationary distribution f , which converges in law to a random variable from f . Hastings (1970) provides further theoretical subjects such as convergences and Ergodic theorems for his proposed MCMC method, and Brooks *et al.* (2011) provides a practical overview of MCMC methods and its applications.

The major problem of the Metropolis Hastings methodology is its slow convergence in high dimensional problems, where the number of iterations m is too big to guarantee the Markov Chain convergence. However, in the last years this has changed with the development of new algorithms and the rapid increase of general computing power (Hoffman and Gelman 2014; Betancourt 2017).

Algorithm 1 Metropolis-Hasting algorithm

Require: m , f and J

Ensure: A sample $\theta_1, \theta_2, \dots, \theta_m$ of $P(\theta/X)$

Draw a value θ_0 such that $f(\theta_0) > 0$

for $t \in 1, 2, 3 \dots, m$ **do**

 Draw θ_* from J

 calculate r

 set $\alpha = \min(r, 1)$

 Draw a value U from a uniform $U(0, 1)$ distribution.

if $U < \alpha$ **then**

 Set $\theta_t = \theta_*$

else

 Set $\theta_t = \theta_{t-1}$

end if

end for

2. Normality tests for stationary processes

Extensive literature exists on goodness of fit tests for normality under the assumption of independent and identical distributed random variables, including Pearson's chi-squared test (Pearson and Henrici 1895), Kolmogorov-Smirnov test (Smirnov 1948), Anderson-Darling test (Anderson and Darling 1952), SK test (Jarque and Bera 1980) and Shapiro-Wilk test (Shapiro and Wilk 1965; Royston 1982) among others. These procedures have been widely used in many studies and applications, see D'Agostino and Stephens (1986) for further details. There are no results, however, showing that the above tests are consistent in the context of stationary processes, case in which the independence assumption is violated. For instance, Gasser (1975) provides a simulation study where Pearson's chi-squared test has an excessive rejection rate under the null hypothesis for dependent data. For this matter, several tests have been proposed over the years; a selection of which we reference here. Epps (1987) provides a test based on the characteristic function and a similar test is proposed by Hinich (1982) based on the process' spectral density function (Berg, Paparoditis, and Politis 2010, for further insight). Gasser (1975) gives a correction of the SK test, with several modifications made in Lobato and Velasco (2004); Bai and Ng (2005); Psaradakis (2017), which are popular in many financial applications. Bontemps and Meddahi (2005) constructs a test based on Stein's characterization of a Gaussian distribution. Using the random projection method (Cuesta-Albertos *et al.* 2007), Nieto-Reyes *et al.* (2014) build a test that upgrades the performance of Epps (1987) and Lobato and Velasco (2004) procedures. Furthermore, Psaradakis and Vávra (2017) proposed a bootstrap approximation of the Anderson and Darling (1952) test statistic for stationary linear processes.

Despite the existing literature, consistent implementations of goodness of fit test for normality of stationary processes in programming languages such as R or Python are limited. We present here the **nortsTest** package: it performs the tests proposed in Epps (1987), Lobato and Velasco (2004), Nieto-Reyes *et al.* (2014) and Psaradakis and Vávra (2017). To install the latest release version of **nortsTest** from CRAN, type `install.packages("nortsTest")` within R. The current development version can be installed from GitHub using the next code:

```
R> if (!requireNamespace("remotes")) install.packages("remotes")
R> remotes::install_github("asael697/nortsTest",dependencies = TRUE)
```

Additionally, the package offers visualization functions for descriptive time series analysis and several diagnostic methods for checking stationarity and normality assumptions for the most used time series models of several R packages. To elaborate on this, Subsection 2.1 introduces the package functionality and software and Subsection 2.2 provides an overview of the used methods for checking stationarity and seasonality. Finally, Subsections 2.3-2.6 present a general framework of each of the implemented test and their functionality by providing simulated data examples.

2.1. Software

The package works as an extension of the **nortest** package (Gross and Ligges 2015), which performs normality tests in random samples but for independent data. The building block functions of the **nortsTest** package are:

- `epps.test()`, function that implements the test of Epps,

- `lobato.test()`, function that implements the test of Lobato and Velasco,
- `rp.test()`, function that implements the random projection test of Nieto-Reyes, Cuesta-Albertos and Gamboa, and
- `vavra.test()`, function that implements the test of Psaradaki and Vavra.

Each of these functions accepts a `numeric` (*numeric*) or `ts` (*time series*) class object for storing data, and returns a `hctest` (*hypothesis test*) class object with the main results for the test. To guarantee the accuracy of the results, each test performs unit root tests for checking stationarity and seasonality (see Subsection 2.2) and displays a warning message if any of them not satisfied.

For visual diagnostic, the package offers different plot functions based on the `ggplot2` package (Wickham 2009): the `autoplot()` function plots `numeric`, `ts` and `mts` (*multivariate time series*) classes while the `gghist()` and `ggnorm()` functions are for plotting histogram and qq-plots respectively; and on the `forecast` package (Hyndman and Khandakar 2008): `ggacf()` and `ggPacf()` for the display of the auto-correlation and partial auto-correlations functions respectively.

Furthermore, inspired in the function `check.residuals()` of the `forecast` package, we provide the `check_residuals()` function for checking assumptions of the model using the estimated residuals. Thus this function checks stationarity, seasonality (see Subsection 2.2) and normality, presenting a report of the used tests and conclusions. If the `plot` option is `TRUE`, the function displays several plots for visual checking. An illustration of these functions is provided in Subsection 3.2, where we show the functions details and their utility for assumptions commonly checked in time series modeling.

2.2. Tests for stationary

For checking stationarity, the `nortsTest` package uses *unit root* and *seasonal unit-roots* tests. These tests work similarly, checking whether a specific process follows a random-walk model, which clearly is a non-stationary process.

Unit root tests

A stochastic process X is non stationary if it follows a random-walk model. This statement is equivalent to say that the AR(1) polynomial ($\phi(z) = 1 - \phi z$) of X has a unit root¹. The most commonly used tests for unit root testing are *Augmented Dickey Fuller* (Said and Dickey 1984), *Phillips-Perron* (Perron 1988), *kpps* (Kwiatkowski, Phillips, Schmidt, and Shin 1992) and *Ljung-Box* Box and Pierce (1970). The `urrot.test()` and `check_residual()` functions perform these tests, making use of the `tseries` package (Trapletti and Hornik 2019).

Seasonal unit root tests

Let X be a stationary process and m be its period². X follows a seasonal random walk if it can be written as

$$X_t = X_{t-m} + \epsilon_t,$$

¹If $\phi = 1$, then $\phi(z) = (1 - z)$ which its only root is one

²For observed data, m is the number of observations per unit of time

where ϵ_t is a collection of i.i.d random variables. In a similar way, the process X is non-stationary if it follows a seasonal random-walk. Or equivalently, X is non-stationary if the seasonal AR(1) polynomial ($\phi_m(z) = 1 - \phi z^m$) has a unit root. The `seasonal()` and `check_residuals()` functions perform the *OCSB test* (Osborn, Chui, Smith, and Birchenhall 1988) from the `forecast` package, and the *HEGY* (Beaulieu and Miron 1993) and *Ch* (Canova and Hansen 1995) tests from the `uroot` package (de Lacalle 2019).

2.3. Test of Epps

The χ^2 test for normality proposed by Epps (1987) compares the empirical characteristic function of the one-dimensional marginal of the process with the one of a normally distributed random variable evaluated at certain points on the real line. Several authors, such as Lobato and Velasco (2004) and Psaradakis and Vávra (2017), point out that the greatest challenge in this test is its implementation procedure.

Let X be a stationary stochastic process that satisfies

$$\sum_{t=-\infty}^{\infty} |t|^k |\gamma(t)| < \infty \text{ for some } k > 0. \quad (4)$$

The null hypothesis is that the one-dimensional marginal distribution of X is a Gaussian process. As we see in what follows, the procedure for constructing the test consists of defining a function g , estimating its inverse spectral matrix function, minimizing the generated quadratic function in terms of the unknown parameters of the random variable and, finally, obtaining the test statistic, which converges in distribution to a χ^2 .

Given $N \in \mathbb{N}$ with $N > 2$, let

$$\Lambda := \{\lambda := (\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N : \lambda_i \leq \lambda_{i+1} \text{ and } \lambda_i > 0, \text{ for } i = 1, 2, \dots, N\}$$

and $g : \mathbb{R} \times \Lambda \rightarrow \mathbb{R}^n$ be a measurable function, where

$$g(x, \lambda) := [\cos(\lambda_1 x), \sin(\lambda_1 x), \dots, \cos(\lambda_N x), \sin(\lambda_N x)].$$

Additionally, let $g_\theta : \Lambda \rightarrow \mathbb{R}^N$ be a function defined by

$$g_\theta(\lambda) := [\text{Re}(\Phi_\theta(\lambda_1)), \text{Im}(\Phi_\theta(\lambda_1)), \dots, \text{Re}(\Phi_\theta(\lambda_N)), \text{Im}(\Phi_\theta(\lambda_N))]^t,$$

where the $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$ are the real and imaginary components of a complex number and Φ_θ is the characteristic function of a normal random variable with parameters $\theta = (\mu, \sigma^2) \in \Theta$, an open bounded set contained in $\mathbb{R} \times \mathbb{R}^+$. For any $\lambda \in \Lambda$, let us also denote

$$\hat{g}(\lambda) := \frac{1}{n} \sum_{t=1}^n [\cos(\lambda_1 X_t), \sin(\lambda_1 X_t), \dots, \cos(\lambda_N X_t), \sin(\lambda_N X_t)]^t.$$

Let $f(v; \theta, \lambda)$ be the spectral density matrix of $\{g(X_t, \lambda)\}_{t \in \mathbb{Z}}$ at a frequency v . Then, for $v = 0$, it can be estimated by

$$\hat{f}(0; \theta, \lambda) := \frac{1}{2\pi n} \left(\sum_{t=1}^n \hat{G}(X_{t,0}, \lambda) + 2 \sum_{i=1}^{\lfloor n^{2/5} \rfloor} (1 - i/\lfloor n^{2/5} \rfloor) \sum_{t=1}^{n-i} \hat{G}(X_{t,i}, \lambda) \right),$$

where $\widehat{G}(X_{t,0}, \lambda) = (\widehat{g}(\lambda) - g_\theta(\lambda))(\widehat{g}(\lambda) - g_\theta(\lambda))^t$ and $\lfloor \cdot \rfloor$ denotes the floor function. The test statistic general form under H_0 is

$$Q_n(\lambda) := \min_{\theta \in \Theta} \{Q_n(\theta, \lambda)\},$$

with

$$Q_n(\theta, \lambda) := (\widehat{g}(\lambda) - g_\theta(\lambda))^t G_n^+(\lambda) (\widehat{g}(\lambda) - g_\theta(\lambda))$$

where G_n^+ is the generalized inverse of the spectral density matrix $2\pi\widehat{f}(0; \theta, \lambda)$. Let $\widehat{\theta} = \arg \min_{\theta \in \Theta} \{Q_n(\theta, \lambda)\}$ be the argument that minimizes $Q_n(\theta, \lambda)$ such that $\widehat{\theta}$ is in a neighborhood of $\theta_n = (\widehat{\mu}, \widehat{\gamma}(0))$. To guarantee its' existence and uniqueness, the following assumptions are required. We refer to them as assumption (A.).

(A.) Let θ_0 be the true value of $\theta = (\mu, \sigma^2)$ under H_0 , then for every $\lambda \in \Lambda$ the following conditions are satisfied.

- $f(0; \theta, \lambda)$ is positive definite.
- $\Phi_\theta(\lambda)$ is twice differentiable with respect to θ in a neighborhood of θ_0 .
- The matrix $D(\theta_0, \lambda) = \frac{\partial \Phi_\theta(\lambda)}{\partial \theta}|_{\theta=\theta_0} \in \mathbb{R}^{N \times 2}$, for $N > 2$, has rank 2.
- The set $\Theta_0(\lambda) := \{\theta \in \Theta : \Phi_\theta(\lambda_i) = \Phi_{\theta_0}(\lambda_i), i = 1, \dots, N\}$ is a finite bounded set in Θ . And θ is a bounded subset $\mathbb{R} \times \mathbb{R}^+$.
- $f(0; \theta, \lambda) = f(0; \theta_0, \lambda)$ and $D(\theta_0, \lambda) = D(\theta, \lambda)$ for all $\theta \in \Theta_0(\lambda)$.

Under these assumptions, the Epps's main result is presented as follows.

Theorem 1 (Epps (1987) Theorem 2.1) *Let X be a stationary Gaussian process such that (4) and (A.) are satisfied, then $nQ_n(\lambda) \rightarrow_d \chi^2(2N - 2)$ for every $\lambda \in \Lambda$.*

For the current **nortsTest** version, we define $\Lambda := \{(1.0, 1.0, 2.0, 2.0)/\widehat{\gamma}(0)\}$, where $\widehat{\gamma}(0)$ is the sample variance. Therefore, the implemented test statistics converges to a χ^2 distribution with two degree freedom. In the next version of the package, the user will set Λ as desired, with the current value as default.

Example 1 *A, stationary, AR(2) process is drawn using a beta distribution with `shape1 = 9` and `shape2 = 1` parameters, and the implementation of the test of Epps, `epps.test()`, is performed. At significance level $\alpha = 0.05$, the null hypothesis of normality is correctly rejected.*

```
R> set.seed(298)
R> # Simulating the AR(2) process
R> x = arima.sim(250, model = list(ar = c(0.5, 0.2)),
                rand.gen = rbeta, shape1 = 9, shape2 = 1)
R> epps.test(x)
```

Epps test

```
data: x
epps = 16.597, df = 2, p-value = 0.0002489
alternative hypothesis: x does not follow a Gaussian Process
```


2.4. Test of Lobato and Velasco

Lobato and Velasco (2004) provides a consistent estimator for the corrected SK test statistic³ for stationary processes (Lomnicki 1961; Gasser 1975, for further insight). On the contrary to the test of Epps, it does not require of additional parameters for the approximation of the test sample statistic. The general framework for the test is presented in what follows.

Let X be a stationary stochastic process that satisfies

$$\sum_{t=0}^{\infty} |\gamma(t)| < \infty. \quad (5)$$

The null hypothesis is that the one-dimensional marginal distribution of X is normally distributed, that is

$$H_0 : X_t \sim N(\mu, \sigma^2) \text{ for all } t \in \mathbb{R}.$$

Let $k_q(j_1, j_2, \dots, j_{q-1})$ be the q -th order cumulant of $X_1, X_{1+j_1}, \dots, X_{1+j_{q-1}}$. H_0 is fulfilled if all the marginal cumulants above the second order are zero. In practice, it is tested just for the third and fourth order marginal cumulants, equivalently, in terms of moments, the marginal distribution is normal by testing whether $\mu_3 = 0$ and $\mu_4 = 3\mu_2^2$. For non correlated data, the SK test compares the SK statistic against upper critical values from a $\chi^2(2)$ distribution (Bai and Ng 2005). For a Gaussian process X satisfying (5), it holds the limiting result

$$\sqrt{n} \begin{pmatrix} \hat{\mu}_3 \\ \hat{\mu}_4 - 3\hat{\mu}_2^2 \end{pmatrix} \rightarrow_d N[0_2, \Sigma_F],$$

where $0_2 := (0, 0)^t \in \mathbb{R}^2$ and $\Sigma_F := \text{diag}(6F^{(3)}, 24F^{(4)}) \in \mathbb{R}^{2 \times 2}$ is a diagonal matrix with $F^{(k)} := \sum_{j=-\infty}^{\infty} \gamma(j)^k$ for $k = 3, 4$ (Gasser 1975).

The following consistent estimator in terms of the auto-covariance function is proposed in Lobato and Velasco (2004)

$$\hat{F}^{(k)} = \sum_{t=1-n}^{n-1} \hat{\gamma}(t) [\hat{\gamma}(t) + \hat{\gamma}(n - |t|)]^{k-1},$$

to build a *generalized SK test* statistic

$$G := \frac{n\hat{\mu}_3^2}{6\hat{F}^{(3)}} + \frac{n(\hat{\mu}_4 - 3\hat{\mu}_2^2)^2}{24\hat{F}^{(4)}}.$$

Similar to the SK test for non-correlated data, the G statistic is compared against upper critical values from a $\chi^2(2)$ distribution. This is seen in the below result that establishes the asymptotic properties of the test statistics, so that the general test procedure can be constructed. The result requires the following assumptions, denoted by (B.), for the process X .

- (B.)
- $E[X_t^{16}] < \infty$ for $t \in T$.
 - $\sum_{j_1=-\infty}^{\infty} \dots \sum_{j_{q-1}=-\infty}^{\infty} |k_q(j_1, \dots, j_{q-1})| < \infty$ for $q = 2, 3, \dots, 16$.

³Also known as the Jarque-Bera test, Jarque and Bera (1980).

$$\begin{aligned}
& - \sum_{j=1}^{\infty} \left(E \left[E[(X_0 - \mu)^k | B_j] - \mu_k \right]^2 \right)^{1/2} < \infty \text{ for } k = 3, 4, \text{ where } B_j \text{ denotes the} \\
& \quad \sigma\text{-field generated by } X_t, t \leq -j. \\
& - E \left[(X_0 - \mu)^k - \mu_k \right]^2 + 2 \sum_{j=1}^{\infty} E \left(\left[(X_0 - \mu)^k - \mu_k \right] \left[(X_j - \mu)^k - \mu_k \right] \right) > 0 \text{ for } k = \\
& \quad 3, 4.
\end{aligned}$$

Note that these assumptions imply that the higher-order spectral densities up to order 16 are continuous and bounded.

Theorem 2 (Lobato and Velasco (2004), Theorem 1) *Let X be a stationary process. If X is Gaussian and satisfies (5) then $G \rightarrow_d \chi^2(2)$, and under assumption (B.), the test statistic G diverges whenever $\mu_3 \neq 0$ or $\mu_4 \neq 3\mu_2^2$.*

Example 2 *A, stationary, MA(3) process is drawn using a gamma distribution with rate = 3 and shape = 6 parameters and the test of Lobato and Velasco is performed using the function `lobato.test()` of the proposed **nortstTest** package. At significance level $\alpha = 0.05$, the null hypothesis of normality is correctly rejected.*

```

R> set.seed(298)
R> # Simulating the MA(3) process
R> x = arima.sim(250,model = list(ma =c(0.2,0.3,-0.4)),
                rand.gen = rgamma,rate = 3,shape = 6)
R> lobato.test(x)

Lobato and Velasco's test

data:  x
lobato = 62.294, df = 2, p-value = 2.972e-14
alternative hypothesis: x does not follow a Gaussian Process

```

2.5. The Random Projections test

The previous two proposals only test for the normality of the one-dimensional marginal distribution of the process, which is inconsistent against alternatives whose one-dimensional marginal is Gaussian. Nieto-Reyes *et al.* (2014) provides a procedure to fully test normality of a stationary process using a Crammér-Wold type result (Cuesta-Albertos *et al.* 2007) that uses random projections to differentiate among distributions. We show this result below. The result works for separable Hilbert spaces, however here, for its later application, we restrict it to l^2 , the space of square summable sequences over \mathbb{N} , with inner product $\langle \cdot, \cdot \rangle$.

Theorem 3 (Cuesta-Albertos *et al.* (2007), Theorem 3.6) *Let η be a dissipative distribution on l^2 and Z a l^2 -valued random element, then Z is Gaussian if and only if*

$$\eta\{h \in l^2 : \langle Z, h \rangle \text{ has a Gaussian distribution}\} > 0.$$

A dissipative distribution (Nieto-Reyes *et al.* 2014, Definition 2.1) is a generalization of the concept of absolutely continuous distribution to the infinite dimensional space. To construct

a dissipative distribution in l^2 , it is made use of the Dirichlet process (Gelman, Carlin, Stern, Dunson, Vehtari, and Rubin 2013). In practice, the $h \in l^2$ is drawn with a stick-breaking process that makes use of beta distributions.

Let $X = \{X_t\}_{t \in \mathbb{Z}}$ be a stationary process. As X is normally distributed if the process $X^{(t)} := \{X_k\}_{k \leq t}$ is Gaussian for each $t \in \mathbb{Z}$, using the result above, Nieto-Reyes *et al.* (2014) provides a procedure for testing that X is a Gaussian process by testing whether the process $Y^h = \{Y_t^h\}_{t \in \mathbb{Z}}$ is Gaussian.

$$Y_t^h := \sum_{i=0}^{\infty} h_i X_{t-i} = \langle X^{(t)}, h \rangle, \quad (6)$$

where $\langle X^{(t)}, h \rangle$ is a real random variable for each $t \in \mathbb{Z}$ and $h \in l^2$. Thus, Y^h is a stationary process constructed by the projection of $X^{(t)}$ on the space generated by h . Therefore, X is a Gaussian process if and only if the marginal distribution of Y^h is normally distributed. Additionally, the hypothesis of the tests *Lobato and Velasco* or *Epps*, such as (4), (5), (A) and (B), imposed on X are inherited by Y^h . Then, those tests can be applied to evaluate the normality of the marginal distribution of Y^h . Further conditions such as, a discussion on the specific beta parameters used to construct the distribution from which to draw h , select a proper amount of combinations to establish the number of projections required to improve the method performance, have to be considered. All of these details are discussed in Nieto-Reyes *et al.* (2014).

Next, we summarize the test of random projections in practice:

1. Select k , the number of independent projections to be used (*by default* $k = 64$).
2. Half of the random elements in which to project are drawn from a dissipative distribution that makes use of a particular beta distribution ($\beta(2, 7)$ *by default*). Then the test of *Lobato and Velasco* is applied to the odd number of projected processes, and the *Epps* test to the even.
3. The other half are drawn analogously but using another beta distribution ($\beta(100, 1)$ *by default*). Then again the test of *Lobato and Velasco* is applied to the odd number of projected process, and the *Epps test* to the even.
4. The obtained k p -values are combined using the false discover rate (Benjamini and Yekutieli 2001).

The `rp.test()` function implements the above procedure. The user might provide optional parameters such as the number of projections `k`, the parameters of the first beta distribution `pars1` and those of the second `pars2`. In the next example, the `rp.test` is applied to a stationary GARCH(1,1) process drawn using normal random variables.

Example 3 A stationary GARCH(1, 1) process ⁴ is drawn using standard normal distribution and the parameters $\alpha_0 = 0$, $\alpha_1 = 0.2$ and $\beta_1 = 0.3$.

⁴A GARCH(1, 1) process is stationary if the parameters α and β satisfy the inequality $\alpha + \beta < 1$ (Bollerslev 1986).

```

R> set.seed(3466)
R> # Simulating the GARCH(1,1) process
R> spec = garchSpec(model = list(alpha = 0.2, beta = 0.3))
R> x = ts( garchSim(spec, n = 300) )

R> rp.test(x,k=250)

      k random projections test

data:  x
k = 250, lobato = 1.2720, epps = 3.2891, p-value = 0.7946
alternative hypothesis: x does not follow a Gaussian Process

```

The *random projections* test is applied to the simulated data with $k = 250$ as the number of projections (*as recommended by the authors*). At significance level $\alpha = 0.05$, there is no evidence to reject null hypothesis of normality.

The `random.projection()` function upgrades the `lobato.test()` and `epps.test()` functions for fully testing normality. This function generates the projected process Y^h as in (6), the `shape1` and `shape2` function's arguments are the parameters of a beta distribution used to generate the stick-breaking process h . And then, the `lobato.test()` or `epps.test()` functions can be applied to the resulting Y^h process for fully testing.

Example 4 We use the *AR(2)* process simulated in Example 1, to fully check of normality using the `epps.test()` and `random.projection()` functions, where `shape1 = 100` and `shape2 = 1` are the arguments for generating the new projected process Y^h . At significance level $\alpha = 0.05$, the null hypothesis of normality is again correctly rejected.

```

R> set.seed(298)
R> x = arima.sim(250,model = list(ar =c(0.5,0.2)),rand.gen = rbeta,
      shape1 = 9,shape2 = 1)
R> y = random.projection(x,shape1 = 100, shape2 = 1,seed = 298)
R> epps.test(y)

```

```

      Epps test

data:  x
epps = 11.645, df = 2, p-value = 0.002961
alternative hypothesis: x does not follow a Gaussian Process

```

2.6. The Psaradakis and Vavra's test

Psaradakis and Vávra (2017) proposed a distance test for normality of the one-dimensional marginal distribution of a stationary process. The test is based on the Anderson and Darling (1952) test statistic and makes use of an auto-regressive sieve bootstrap approximation to the null distribution of the sample test statistic. Although the test is said to be applicable to a wider class of non-stationary processes, by transforming them into stationary by means

of a fractional difference operator, no theoretic result was apparently provided to sustain this transformation. Therefore, here we restrict the presentation and implementation of the procedure to stationary processes.

Let X be a stationary process satisfying

$$X_t = \sum_{i=0}^{\infty} \theta_i \epsilon_{t-i} + \mu_0, \quad t \in \mathbb{Z}, \quad (7)$$

where $\mu_0 \in \mathbb{R}$, $\{\theta_i\}_{i=0}^{\infty} \in l^2$ with $\theta_0 = 1$ and $\{\epsilon_t\}_{i=0}^{\infty}$ a collection of mean zero i.i.d random variables. The null hypothesis is that the one-dimensional marginal distribution of X is normally distributed,

$$H_0 : F(\mu_0 + \sqrt{\gamma(0)x}) - F_N(x) = 0, \quad \text{for all } x \in \mathbb{R},$$

where F is the cumulative distribution function of X_0 , and F_N denotes the standard normal cumulative distribution function. Note that if ϵ_0 is normally distributed, then the null hypothesis is satisfied. Conversely, if the null hypothesis is satisfied, then ϵ_0 is normally distributed and consequently X_0 .

The considered test for H_0 is based on the Anderson-Darling distance statistic

$$A_d = \int_{-\infty}^{\infty} \frac{[F_n(\hat{\mu} + \sqrt{\hat{\gamma}(0)x}) - F_N(x)]^2}{F_N(x)[1 - F_N(x)]} dF_N(x), \quad (8)$$

where $F_n(\cdot)$ is the empirical distribution function associated to F based on a simple random sample of size n . Psaradakis and Vávra (2017) propose an auto-regressive sieve bootstrap procedure to approximate the sampling properties of A_d arguing that making use of classical asymptotic inference for A_d is problematic and involved. This scheme is motivated by the fact that under some assumptions for X , including (7), ϵ_t admits the representation,

$$\epsilon_t = \sum_{i=1}^{\infty} \phi_i (X_{t-i} - \mu_0), \quad t \in \mathbb{Z}, \quad (9)$$

for certain type of $\{\phi_i\}_{i=1}^{\infty} \in l^2$. The main idea behind this approach is to generate a bootstrap sample ϵ_t^* to approximate ϵ_t with a finite-order auto-regressive model. This is because the distribution of the processes ϵ_t and ϵ_t^* coincide asymptotically if the order of the auto-regressive approximation grows simultaneously with n at an appropriate rate (Bühlmann 1997). The procedure makes use of the ϵ_t^{*s} to obtain the X_t^{*s} through the bootstrap analog of (9). Then, a bootstrap sample of the A_d statistic, A_d^* , is generated making use of the bootstrap analog of (7).

This test is implemented in the `vavra.test()` function. 1,000 sieve-bootstrap replications are used by default. The presented values are Monte-Carlo estimates of the A_d statistic and `p.value`.

Example 5 *A stationary ARMA(1,1) process is simulated using a standard normal distribution, and the implementation of the test of Psaradakis and Vávra is performed. At significance level $\alpha = 0.05$, there is no evidence to reject the null hypothesis of normality.*

```
R> set.seed(298)
R> # Simulating the ARMA(1,1) process
R> x = arima.sim(250,model = list(ar = 0.2, ma = 0.34))
R> vavra.test(x)
```

Psaradakis-Vavra test

```
data: x
bootstrap A = 1.5798, p-value = 0.796
alternative hypothesis: x does not follow a Gaussian Process
```

3. Simulations and data analysis

3.1. Numerical experiments

Inspired in [Psaradakis and Vávra \(2017\)](#) and [Nieto-Reyes *et al.* \(2014\)](#) simulation studies, this work proposes a similar procedure. This study involves drawing data from the $AR(1)$ process

$$X_t = \phi X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}, \quad \text{for } \phi \in \{0, \pm 0.25, \pm 0.4\} \quad (10)$$

where the $\{\epsilon_t\}_{t \in \mathbb{Z}}$ are i.i.d random variables. For the distribution of the ϵ_t we consider different scenarios: standard normal (N), standard log-normal ($\log N$), Student t with 3 degrees of freedom (t_3), chi-squared with 10 degrees of freedom ($\chi^2(10)$) and beta with parameters (7, 1). As in [Psaradakis and Vávra \(2017\)](#), $m = 1,000$ independent draws of the above process are generated for each pair of parameter ϕ and distribution.

distribution	n = 100					n = 250				
	ϕ -0.4	-0.25	0.0	0.25	0.4	-0.4	-0.25	0.0	0.25	0.4
Lobato and Velasco										
N	0.045	0.042	0.039	0.052	0.037	0.044	0.054	0.056	0.050	0.048
$\log N$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
t_3	0.806	0.872	0.873	0.882	0.796	0.980	0.992	0.999	0.993	0.983
$\chi^2(10)$	0.553	0.685	0.776	0.667	0.559	0.968	0.995	0.997	0.996	0.964
$\text{beta}(7, 1)$	0.962	0.995	0.999	0.996	0.958	1.000	1.000	1.000	1.000	1.000
Epps										
N	0.066	0.077	0.084	0.075	0.070	0.058	0.059	0.065	0.073	0.068
$\log N$	0.825	0.884	0.969	0.958	0.948	0.998	0.999	1.000	1.000	1.000
t_3	0.202	0.294	0.363	0.288	0.207	0.716	0.838	0.909	0.855	0.752
$\chi^2(10)$	0.319	0.465	0.548	0.461	0.361	0.631	0.836	0.917	0.841	0.729
$\text{beta}(7, 1)$	0.781	0.953	0.991	0.960	0.887	0.996	1.000	1.000	1.000	0.999
Random Projections, k = 10										
N	0.007	0.009	0.007	0.006	0.006	0.021	0.027	0.025	0.021	0.018
$\log N$	0.891	0.865	0.772	0.625	0.515	1.000	1.000	1.000	1.000	1.000
t_3	0.293	0.267	0.204	0.122	0.098	0.989	0.993	0.995	0.983	0.961
$\chi^2(10)$	0.223	0.231	0.201	0.131	0.086	0.954	0.993	0.992	0.949	0.840
$\text{beta}(7, 1)$	0.605	0.533	0.363	0.184	0.108	1.000	1.000	1.000	1.000	1.000
Psaradakis and Vavra										
N	0.056	0.046	0.038	0.052	0.046	0.050	0.061	0.044	0.055	0.047
$\log N$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
t_3	0.714	0.802	0.850	0.768	0.644	0.959	0.987	0.997	0.990	0.960
$\chi^2(10)$	0.500	0.692	0.800	0.660	0.542	0.911	0.985	0.995	0.985	0.922
$\text{beta}(7, 1)$	0.956	1.000	1.000	0.998	0.972	1.000	1.000	1.000	1.000	0.999

Table 1: Rejection rate estimates over $m = 1,000$ trials of the four studied goodness of fit test for the null hypothesis of normality. The data is drawn using the process defined in (10) for different values of ϕ and n displayed in the columns and different distributions for ϵ_t in the rows.

Each is taken of length $past + n$, with $past = 500$ and $n \in \{100, 250, 500, 1000\}$. The first 500 data points of each realization are then discarded in order to eliminate start-up effects. The n remaining data points are used to compute the value of the test statistic of interest. In each particular scenario, the rejection rate is obtained by computing the proportion of times that the test is rejected among the m trials.

distribution	n = 500					n = 1,000				
	ϕ -0.4	-0.25	0.0	0.25	0.4	-0.4	-0.25	0.0	0.25	0.4
Lobato and Velasco										
N	0.042	0.053	0.037	0.041	0.043	0.049	0.051	0.046	0.043	0.047
$\log N$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
t_3	0.968	0.987	0.989	0.983	0.962	1.000	1.000	1.000	1.000	1.000
$\chi^2(10)$	0.902	0.965	0.996	0.976	0.880	1.000	1.000	1.000	1.000	1.000
$\text{beta}(7,1)$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Epps										
N	0.063	0.077	0.078	0.072	0.073	0.051	0.048	0.052	0.056	0.062
$\log N$	0.989	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
t_3	0.569	0.705	0.781	0.694	0.587	0.999	1.000	1.000	1.000	0.999
$\text{chisq}10$	0.534	0.745	0.859	0.740	0.611	0.999	1.000	1.000	1.000	1.000
$\text{beta}(7,1)$	0.983	0.998	1.000	1.000	0.989	1.000	1.000	1.000	1.000	1.000
Random Projections k = 10										
N	0.016	0.015	0.012	0.019	0.017	0.015	0.016	0.019	0.018	0.018
$\log N$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
t_3	1.000	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000
$\text{chisq}10$	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000	1.000	1.000
$\text{beta}(7,1)$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Psaradakis and Vavra										
N	0.064	0.046	0.048	0.038	0.050	0.055	0.049	0.045	0.057	0.042
$\log N$	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
t_3	0.908	0.972	0.982	0.958	0.896	1.000	1.000	1.000	1.000	1.000
$\text{chisq}10$	0.824	0.954	0.988	0.958	0.856	1.000	1.000	1.000	1.000	1.000
$\text{beta}(7,1)$	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: Rejection rate estimates over $m = 1,000$ trials of the four studied goodness of fit test for the null hypothesis of normality. The data is drawn using the process defined in (10) for different values of ϕ and n displayed in the columns and different distributions for ϵ_t in the rows. $\phi \in \{0, \pm 0.25, \pm 0.4\}$, $n \in \{500, 1000\}$.

Tables 1 and 2 presents the rejection rate estimates. For every process of length n , the columns represent the used $AR(1)$ parameter, and the rows the proposed distribution where the processes were sampled. The obtained results are consistent with those obtained in the publications where the different tests were proposed. As expected, rejection rates are around 0.05 when the data is drawn making use of the standard normal distribution, as in this case the data is drawn from a Gaussian process. Conversely, high rejection rates are registered for the other distributions. Although low rejection rates for the $\chi^2(10)$ distribution for the *Epps* and *random projection* test are observed, they consistently tend to 1 when the length of the

process, n , increases.

The number of projections for this study is $k = 10$, this is by far a lower number than the recommended by Nieto-Reyes *et al.* (2014), but even in these conditions, the obtained results are satisfactory, having even better performance than the tests of Epps (1987), or Psaradakis and Vávra (2017).

3.2. Real data application

As an illustrative example, we analyze the monthly mean carbon dioxide, in parts per million (*ppm*), measured at the Mauna Loa Observatory, in Hawaii, from March 1958 to November 2018. The carbon dioxide data measured as the mole fraction in dry air on Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. This dataset is available in the `astsa` package (Stoffer 2020) under the name `cardox` data and it is displayed in the left panel of Figure 1.

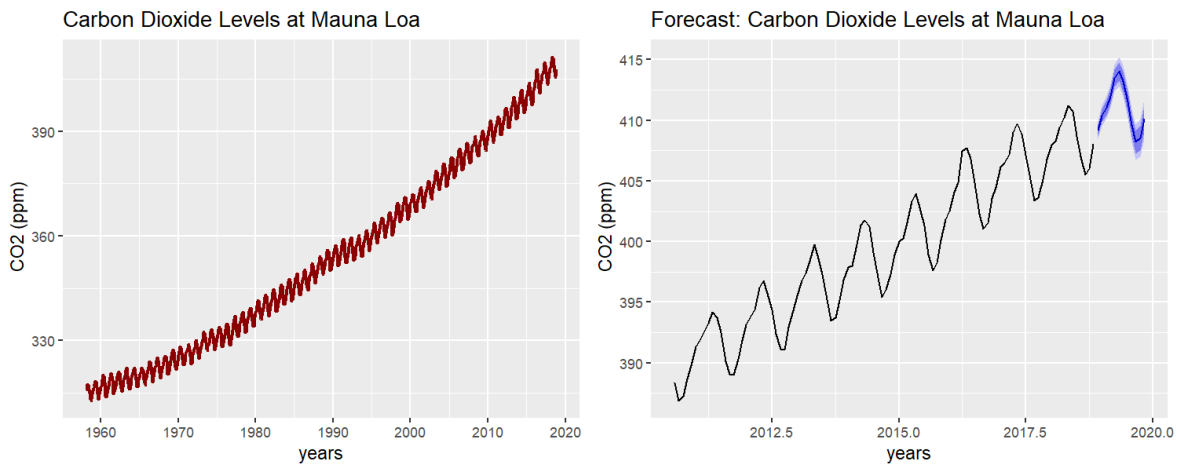


Figure 1: Left panel: CO_2 Levels at Mauna Loa, time-series plot. The `cardox` data show a positive tendency and strong seasonality. Right panel: forecast of the next 12 months for the CO_2 levels at Mauna Loa, the model’s predictions capture the time-series behaviour.

The time-series clearly has trend and seasonal components (see Figure 1), therefore, an adequate model that filters both components has to be selected. We propose a Gaussian linear state space model (STM), implemented by the `forecast` package (Hyndman and Khandakar 2008). This function fits 32 different STMs with different characteristics for trend and seasonality, and selects the best one according to the *Bayesian Information criteria* (BIC) (Chen and Chen 2008). For the `cardox` data, the best fitted model is a *Multiplicative error, additive trend-seasonality* (ETS(M,A,A)) model, (see Definition 8).

```
R> library(forecast)
R> summary(ets(astsa::cardox))
```

```
ETS(M,A,A)
```

```

Call:
ets(y = cardox)

Smoothing parameters:
  alpha = 0.5591
  beta  = 0.0072
  gamma = 0.1061

Initial states:
  l = 314.6899
  b = 0.0696
  s = 0.6611 0.0168 -0.8536 -1.9095 -3.0088 -2.7503
      -1.2155 0.6944 2.1365 2.7225 2.3051 1.2012

sigma: 9e-04

      AIC      AICc      BIC
3136.280 3137.140 3214.338

Training set error measures:
              ME  RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 0.02324 0.3120 0.24308 0.0063088 0.068840 0.1559102 0.07275949

The cardox data satisfies the models assumptions if the model's errors follow a stationary Gaussian process. Using the estimated model's errors (residuals), we propose a normality test to check Gaussianity ( Section 2) and a unit root test for stationarity (subsection 2.2). This can be performed automatically using our propose function check_residuals(), that displays the results of the Augmented Dickey-Fuller and random projection tests.

R> check_residuals(model,unit_root = "adf",normality = "rp",plot = TRUE)

*****

Unit root test for stationarity:

      Augmented Dickey-Fuller Test

data: y
Dickey-Fuller = -9.7249, Lag order = 8, p-value = 0.01
alternative hypothesis: stationary

Conclusion: y is stationary
*****

Goodness of fit test for Gaussian Distribution:

```

```

k random projections test

data: y
k = 64, lobato = 3.679, epps = 1.3818, p-value = 0.5916
alternative hypothesis: y does not follow a Gaussian Process

```

Conclusion: y follows a Gaussian Process

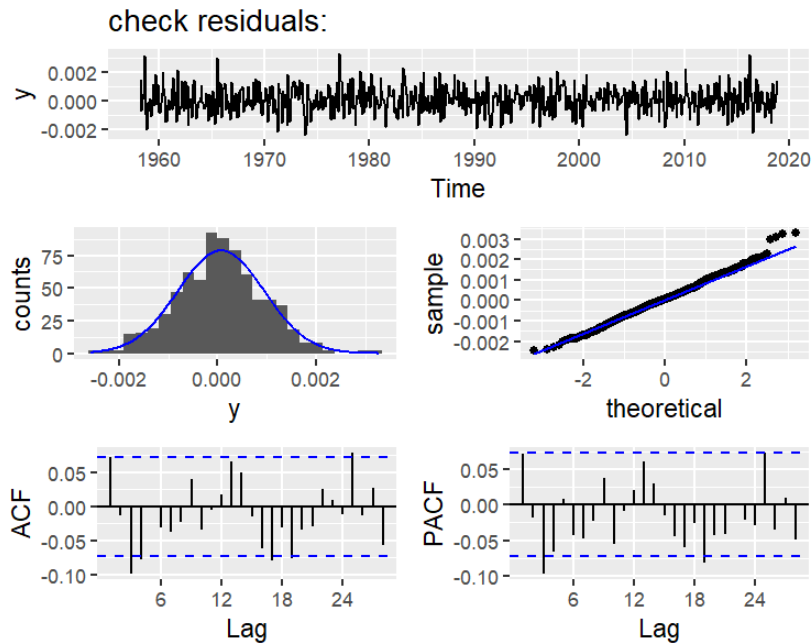


Figure 2: Diagnostic plots for the residuals of the ETS(M,A,A) model. The upper panel shows the residuals time-series plot, showing small oscillations around zero, which insinuates stationarity. The middle plots are the residuals histogram (middle-left) and quantile-quantile plot (middle-right), both plots suggest that the residuals have a normal distribution. The lower panel shows the autocorrelation functions, for both plots, the autocorrelations are close to zero giving the impression of stationarity.

The obtained results indicate that the null hypothesis of non-stationary is rejected at significance $\alpha = 0.01$. Additionally, there is no evidence to reject the null hypothesis of normality at significance $\alpha = 0.05$. Consequently, we conclude that the residuals follow a stationary Gaussian process. Additionally, Figure 2 performs different diagnostic plots for the residuals, indicating that the assumptions of the model are satisfied.

As the assumptions of the model have been checked, it can be used for instance to forecast. The result of applying the following function is displayed in the right panel of Figure 1. It presents the Carbon dioxide data for the last 8 years and a forecast of the next 12 months. It is observable from the plot that the model captures the process trend and periodicity.

4. Model assessment in Bayesian Methods

Model assessment in a Bayesian approach remains as computational challenge. Several solutions based on the predictive distribution have been considered. The Bayes factor (Kass and Raftery 1995) for model selection allows comparing alternative models when proper prior distributions are established. For model selection, the Watanabe information criteria (WAIC) of Watanabe (2010) as an improvement of the deviance information criteria (DIC) proposed by Spiegelhalter, Best, Carlin, and Van Der Linde (2002). And Vehtari, Gelman, and Gabry (2016) implement an *approximate leave one out cross-validation*.

For assessing model adequacy, Guttman (1967) and Gelman (1996) propose the posterior predictive checks (PPC) methods, that compares the simulated predictive distribution with the data empirical distribution. For this matter, Johnson (2004) implements a Bayesian version of Pearson's goodness of fit test, and Gabry, Simpson, Vehtari, Betancourt, and Gelman (2019) implement visual PPC methods for models diagnostic.

On the last years, the PPC methods are the most used for model assessment in Bayesian modeling, the main problem of this methods in stochastic process and time series analysis is that stationarity is required for an accurate computation of the process' empirical distribution, and this condition might not be satisfied in practice. As an alternative procedure, we propose a visual method for model adequacy using the random projection method of Cuesta-Albertos *et al.* (2007). The main idea of this procedure is to obtain a multivariate sample of the model's residuals using a MCMC algorithm and then, use the random projections method to assess normality. The advantage of this approach is that the process' empirical distribution is not need to check the models assumptions, giving a clear advantage to the PPC methods.

For the purpose of this work, the proposed method is applied to time series models with Gaussian assumptions, but it can be extended to any type of parametric model, with any distribution assumption. In the following subsection we present the methods procedure, and two case studies using simulated data.

4.1. The random projections method for Bayesian model assessing

Let $X = \{X_1, X_2, \dots, X_n\}$ be a finite collection of an stochastic process, and denote M a parametric Gaussian model imposed to X restricted to $\theta \in \Theta$, a vector of unknown parameters. Then, lets assume that the model M can be expressed State space model as in **Definition 7**, with equations:

$$\begin{aligned} X_t &= f(\theta_{t-1}) + r(\theta_{t-1})\epsilon_t \\ \theta_t &= g(\theta_{t-1}) + w(\theta_{t-1})\epsilon_t, \end{aligned}$$

where f, g, r and w are functional over Θ , and $\{\epsilon_t\}_{t=1}^n$ is a collection of white noise random variables denoted as errors. Then for assessing model adequacy, lets check if the models errors are normally distributed, where the errors can be estimated as

$$\epsilon_i := Y_i - Y_i^*,$$

where Y_i^* is an estimate of Y_i , based on the predictive distribution, see Pettit (1986) for further details. In practice, the sample obtained from the posterior distribution in a MCMC algorithm can be used to generate observations from the predictive distribution, which in turn can be used to compute the estimate errors. Let $\mathbb{E} := \{\epsilon^{(1)}, \epsilon^{(2)}, \dots, \epsilon^{(m)}\}$ be a multivariate

sample of the errors $\epsilon := [\epsilon_1, \epsilon_2, \dots, \epsilon_n]$, for an arbitrary model M , where m is the number of simulations obtained by a MCMC methodology. A straightforward procedure is to check if every sample of the error components (ϵ_i) , for $i := 1, 2, \dots, n$ are normally distributed, but this task is computational expensive. Another approach is to check if all the m sampled values

$$\epsilon^{(k)} := [\epsilon_1^{(k)}, \epsilon_2^{(k)}, \dots, \epsilon_n^{(k)}], \text{ for } k = 1, 2, \dots, m$$

are random elements of a multivariate normal distribution, this approach presents a computational challenge, but can be easily solved using the random projections methods of [Cuesta-Albertos et al. \(2007\)](#).

Let $\mathbb{H} = \mathbb{R}^n$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$, η a dissipative distribution on \mathbb{R}^n , and every vector of errors $\epsilon^{(k)}$ is a \mathbb{R}^n -valued random element for $k = 1, 2, \dots, m$. Then, applying **Theorem 3.6** of [Cuesta-Albertos et al. \(2007\)](#), every $\epsilon^{(k)}$ is Gaussian if and only if $\eta(E) > 0$, where

$$E = \{h \in \mathbb{R}^n : \langle \epsilon^{(k)}, h \rangle \text{ has a Gaussian distribution}\}.$$

Let $\{h_k\}_{k=1}^m$ be a collection of \mathbb{R}^n -valued random elements sampled from a stick-breaking process η , every h_k is then used to project every simulated error $\epsilon^{(k)}$ to the one-dimensional space \mathbb{R} , obtaining an one-dimensional sample of projected errors $\{\langle \epsilon^{(i)}, h_i \rangle\}_{i=1}^m$, where normality can be easily checked. Therefore, the model errors ϵ are multivariate Gaussian distributed if and only if the collection $\{\langle \epsilon^{(i)}, h_i \rangle\}_{i=1}^m$ has a one-dimensional Gaussian distribution. Finally, for assessing normality in the obtained one-dimensional sample, we propose using histograms, density plots, boxplots and quantile-quantile plots for visual methods.

Even so this new approach validates the model without any additional computations such as the process' empirical distribution, the major problem is that visual methods are not a proper methodology, then a more robust way is need for a proper validation, but unfortunately there are not much references for hypothesis testing in a Bayesian inference approach.

4.2. Case Studies

In this subsection we propose 2 simulations studies using simulated data. All cases are analyzed using *Bayesian ARMA* models, implemented by the **varstan** package ([Matamoros and Torres 2020](#)), an R package that implements Bayesian time series models using a Hamiltonian Monte Carlo (HMC) [Betancourt \(2017\)](#). For every parameter, this algorithm returns a simulated Markov chain, which stationary distribution is the parameter posterior distribution. Additionally, the package provides a multivariate sample of the model's errors as well.

Case 1: stationary Gamma AR(1) process

For this case, we draw 250 simulations from a stationary *AR(1)* process.

$$X_i = \mu_0 + \phi X_{i-1} + \epsilon_i,$$

with $\mu_0 = 0$, $\phi = 0.34$ and the white noise errors are simulated using a Gamma distribution with parameters $\alpha = 2$ and $\beta = 3$. For analyze this data set we propose the next model:

M_1 : Gaussian AR(1) process

$$X_i = \mu_0 + \phi X_{i+1} + \epsilon_i,$$

where:

$$\epsilon_i \sim N(0, 1),$$

and priors:

$$\begin{aligned} \mu_0 &\sim t_7, \\ \phi &\sim N(0, 0.5), \end{aligned}$$

where t_7 represent a t student distribution with 7 degrees of freedom and N represents a normal distribution. For the HMC we propose 1 chain of 4,000 iteration, and then, the first 2,000 iterations are discarded in order to eliminate start-up effects. As Table 3 shows, a bad selection of models assumption is hardly noted in the estimation process. Even so, that the posterior mean values are bad estimations for μ_0 and ϕ , the *potential scale reductions factors* \widehat{R} 's (Vehtari, Gelman, Simpson, Carpenter, and Bürkner 2020) are 1, indicating that the simulated chains are stationary. And the *effective sample size* (ess) are all around 2,000 (the number of defined iterations) indicating low correlation in the simulated chain, therefore, no illness is found in the HMC procedure. Figure 3, shows the sample distribution of the

parameters	mean	se	2.5%	97.5%	ess	\widehat{R}
μ_0	0.90	0.00	0.89	0.90	2019.03	1.00
ϕ	0.17	0.00	0.17	0.18	2007.52	1.00
loglik	-198.13	0.03	-198.18	-198.07	2179.49	1.00

Table 3: Case 1: Estimated values of M_1 , the mean column present the parameters posterior mean estimates, *se* is the posterior estimates standard error, columns 4 and 5 represents the credible intervals at a 95% credibility, the effective sample size shows the number of independent equivalent sample are generated from the simulated chain. And the *potential scale reductions factors* \widehat{R} 's (Vehtari *et al.* 2020) shows the chains convergences for values close to 1.

projected errors, as it is shown, the distribution is asymmetric, and the quantile-quantile plot show that the sample quantiles do not match the theoretical ones of a normal distribution. We conclude, the errors are not normally distributed, and therefore, the model's assumptions are not adequate for the simulated AR(1) gamma process.

Case 2: stationary Gaussian MA(1) process

In a similar way, we draw $n = 300$ simulations of a stationary MA(1) process.

$$X_i = \mu_0 + \theta\epsilon_{i-1} + \epsilon_i$$

with $\mu_0 = 0$, $\phi = 0.15$ and the white noise errors are simulated using a standard normal distribution. For analyze this dataset, we use the following model: M_2 : Gaussian MA(1) process

$$X_i = \mu_0 + \theta\epsilon_{i-1} + \epsilon_i,$$

where,

$$\epsilon_i \sim N(0, 1),$$

and priors,

$$\mu_0 \sim t_7,$$

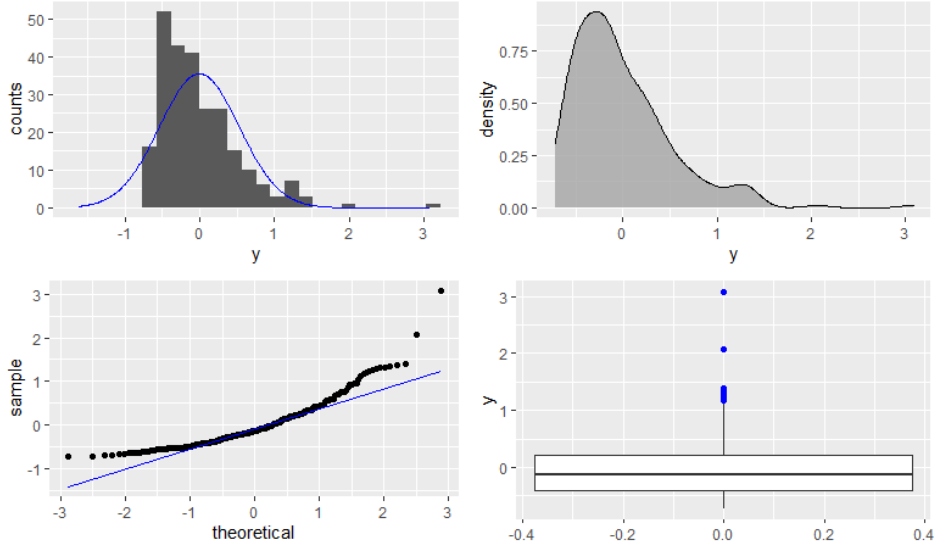


Figure 3: Diagnostic plot for projected errors, model M_1 : Gamma AR(1) process dataset. The upper left plot compares the projected errors histogram with the theoretic density plot of a normal distribution. The upper right plot show the density plot for the projected errors. The lower left plot is a qqplot it compares the projected errors sample quantiles with the ones of a standard normal distribution. And the lower right plot is the projected error's boxplot.

$$\theta \sim N(0, 0.5).$$

Following the same technical specifications for the HMC as in the previous study case, the indicators in Table 4 shows that the simulated chains are stationary, and the posterior mean estimates are accurate to the real specified parameters.

	mean	se	2.5%	97.5%	ess	\widehat{R}
μ_0	0.03	0.00	0.02	0.03	886.19	1.00
θ	0.18	0.00	0.018	0.19	1084.23	1.00
loglik	-423.29	0.04	-423.36	-423.22	964.35	1.00

Table 4: Case 2: Estimated values for Gaussian MA(1) process, the mean column present the parameters posterior mean estimates, se is the posterior estimates standard error, columns 4 and 5 represents the credible intervals at a 95% credibility, the effective sample size shows the number of independent equivalent sample are generated from the simulated chain. And the *potential scale reductions factors* \widehat{R} 's show the chains convergences for values close to 1.

Figure 4, shows the sample distribution of the projected errors, as it is shown, the distribution is symmetric, and the quantile-quantile plot shows that the sample quantiles match the theoretical ones of a normal distribution. We conclude, the errors are normally distributed, and therefore, the model's assumptions are adequate for the simulated Gaussian MA(1) process.

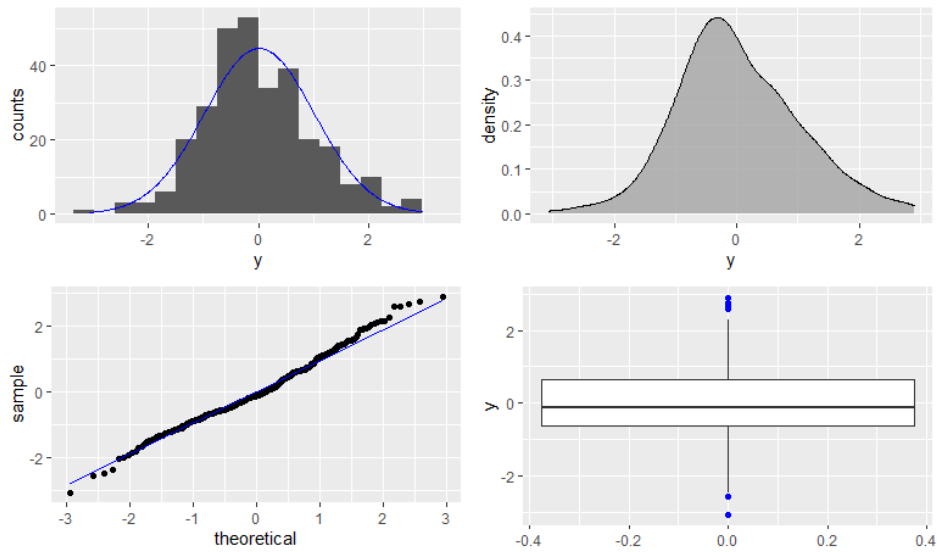


Figure 4: Diagnostic plot for projected errors, model M_2 : Gaussian MA(1) process. The upper left plot compares the projected errors histogram with the theoretic density plot of a normal distribution. The upper right plot show the density plot for the projected errors. The lower left plot is a qqplot it compares the projected errors sample quantiles with the ones of a standard normal distribution. And the lower right plot is the projected error's boxplot.

5. Conclusions

This work gives a general overview from most tests available for normality in stationary process, and provides examples that illustrate every test implementation. At the date of this work, no consistent application or package of all these tests was found. In the context of stationary processes, some authors have available the **Matlab** code used from their simulation studies, but in these cases, the code was restrictive and not suitable for use in real data applications. Therefore, the **nortsTest** is the first package that implements consistent test for normality in stationary process. Additionally the packages provides a too tool for automatic model checking and diagnostic, for the most common models implemented in several R packages, [Matamoros and Nieto-Reyes \(2020\)](#).

Model's assumption check is an important part in statistical analysis, this works focus on providing tools, for model diagnostic in time series model. On a Bayesian approach, our proposed method is an alternative for assessing model adequacy, providing a tool for model diagnostic in time series analysis where stationary might not be satisfied, therefore, a PPC method might not be reliable. Even so we obtained promising results in both case studies, a further study might be necessary, several situations such as the number of projections, the method's sensibility, and a general implementation for distribution different than a normal distribution need to be considerate.

Future work and projects

For the **nortsTest** package, additional tests such as Bispectral (Hinich 1982) and Stein’s characterization (Bontemps and Meddahi 2005) will be incorporated in its second version; as upgrades in the optimization and bootstrap procedures for the *Epps* and *Psaradaskis & Vavra’s* tests for faster performance. Also the creation of different implementations of the Skewness-Kurtosis test besides *Lobato & Velasco’s*.

References

- Anderson TW, Darling DA (1952). “Asymptotic Theory of Certain Goodness of Fit Criteria Based on Stochastic Processes.” *Annals of Mathematical Statistics*, **23**(2), 193–212. doi: [10.1214/aoms/1177729437](https://doi.org/10.1214/aoms/1177729437).
- Bai J, Ng S (2005). “Tests for Skewness, Kurtosis, and Normality for Time Series Data.” *Journal of Business & Economic Statistics*, **23**(1), 49–60. doi: [10.1198/073500104000000271](https://doi.org/10.1198/073500104000000271).
- Beaulieu J, Miron JA (1993). “Seasonal Unit Roots in Aggregate U.S. Data.” *Journal of Econometrics*, **55**(1), 305 – 328. ISSN 0304-4076. doi: [https://doi.org/10.1016/0304-4076\(93\)90018-Z](https://doi.org/10.1016/0304-4076(93)90018-Z).
- Benjamini Y, Yekutieli D (2001). “The Control of the False Discovery Rate in Multiple Testing under Dependency.” *The Annals of Statistics*, **29**(4), 1165–1188. ISSN 00905364. URL <http://www.jstor.org/stable/2674075>.
- Berg A, Paparoditis E, Politis DN (2010). “A Bootstrap Test for Time Series Linearity.” *Journal of Statistical Planning and Inference*, **140**(12), 3841 – 3857. ISSN 0378-3758. doi: <https://doi.org/10.1016/j.jspi.2010.04.047>.
- Betancourt M (2017). “A Conceptual Introduction to Hamiltonian Monte Carlo.” [1701.02434](https://arxiv.org/abs/1701.02434).
- Bollerslev T (1986). “Generalized Autoregressive Conditional Heteroskedasticity.” *Journal of Econometrics*, **31**(3), 307 – 327. ISSN 0304-4076. doi: [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1).
- Bontemps C, Meddahi N (2005). “Testing Normality: a GMM approach.” *Journal of Econometrics*, **124**(1), 149 – 186. ISSN 0304-4076. doi: <https://doi.org/10.1016/j.jeconom.2004.02.014>.
- Box G, Pierce DA (1970). “Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models.” *Journal of the American Statistical Association*, **65**(332), 1509–1526. doi: [10.1080/01621459.1970.10481180](https://doi.org/10.1080/01621459.1970.10481180).
- Box GEP, Jenkins G (1990). *Time Series Analysis, Forecasting and Control*. Holden-Day, Inc., USA. ISBN 0816211043. URL <https://www.wiley.com/en-us/Time+Series+Analysis>.
- Brooks S, Gelman A, Jones G, Meng XL (2011). *Handbook of Markov Chain Monte Carlo*. CRC press.
- Bühlmann P (1997). “Sieve Bootstrap for Time Series.” *Bernoulli*, **3**(2), 123–148. ISSN 13507265. URL <http://www.jstor.org/stable/3318584>.

- Canova F, Hansen BE (1995). “Are Seasonal Patterns Constant Over Time? A Test for Seasonal Stability.” *Journal of Business & Economic Statistics*, **13**(3), 237–252. doi:[10.1080/07350015.1995.10524598](https://doi.org/10.1080/07350015.1995.10524598).
- Chen J, Chen Z (2008). “Extended Bayesian Information Criteria for Model Selection With Large Model Spaces.” *Biometrika*, **95**(3), 759–771. ISSN 0006-3444. doi:[10.1093/biomet/asn034](https://doi.org/10.1093/biomet/asn034).
- Cuesta-Albertos J, del Barrio E, Fraiman R, Matrán C (2007). “The Random Projection Method in Goodness of Fit for Functional Data.” *Computational Statistics & Data Analysis*, **51**(10), 4814 – 4831. ISSN 0167-9473. doi:<https://doi.org/10.1016/j.csda.2006.09.007>.
- D’Agostino RB, Stephens MA (1986). “Goodness-of-fit techniques.” *Quality and Reliability Engineering International*, **3**(1), 71–71. doi:[10.1002/qre.4680030121](https://doi.org/10.1002/qre.4680030121).
- de Lacalle JL (2019). **uroot**: *Unit Root Tests for Seasonal Time Series*. R package version 2.1-0, URL <https://CRAN.R-project.org/package=uroot>.
- Duane S ea (1987). “Hybrid Monte Carlo.” *Physics Letters B*, **95**(2), 216 – 222. ISSN 0370-2693. doi:[https://doi.org/10.1016/0370-2693\(87\)91197-X](https://doi.org/10.1016/0370-2693(87)91197-X).
- Engle RF (1982). “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation.” *Econometrica*, **50**(4), 987–1007. ISSN 00129682, 14680262. URL <http://www.jstor.org/stable/1912773>.
- Epps TW (1987). “Testing That a Stationary Time Series is Gaussian.” *The Annals of Statistics*, **15**(4), 1683–1698. doi:[10.1214/aos/1176350618](https://doi.org/10.1214/aos/1176350618).
- Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A (2019). “Visualization in Bayesian workflow.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **182**(2), 389–402. doi:[10.1111/rssa.12378](https://doi.org/10.1111/rssa.12378).
- Gasser T (1975). “Goodness-of-Fit Tests for Correlated Data.” *Biometrika*, **62**(3), 563–570. ISSN 00063444. URL <http://www.jstor.org/stable/2335511>.
- Gelfand AE, Smith AFM (1990). “Sampling-Based Approaches to Calculating Marginal Densities.” *Journal of the American Statistical Association*, **85**(410), 398–409. ISSN 01621459. URL <http://www.jstor.org/stable/2289776>.
- Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D (2013). *Bayesian Data Analysis, Third Edition*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis. ISBN 9781439840955. URL <https://books.google.nl/books?id=ZXL6AQAAQBAJ>.
- Gross J, Ligges U (2015). **nortest**: *Tests for Normality*. R package version 1.0-4, URL <https://CRAN.R-project.org/package=nortest>.
- Guttman I (1967). “The Use of the Concept of a Future Observation in Goodness-of-Fit Problems.” *Journal of the Royal Statistical Society. Series B (Methodological)*, **29**(1), 83–100. ISSN 00359246. URL <http://www.jstor.org/stable/2984569>.

- Hastings WK (1970). “Monte Carlo Sampling Methods Using Markov Chains and Their Applications.” *Biometrika*, **57**(1), 97–109. ISSN 00063444. URL <http://www.jstor.org/stable/2334940>.
- Hinich MJ (1982). “Testing For Gaussianity and Linearity of a Stationary Time Series.” *Journal of Time Series Analysis*, **3**(3), 169–176. doi:10.1111/j.1467-9892.1982.tb00339.
- Hoffman MD, Gelman A (2014). “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo.” *Journal of Machine Learning Research*, **15**, 1593–1623. URL <http://jmlr.org/papers/v15/hoffman14a.html>.
- Hyndman R, Khandakar Y (2008). “Automatic Time Series Forecasting: The **forecast** Package for R.” *Journal of Statistical Software, Articles*, **27**(3), 1–22. ISSN 1548-7660. doi:10.18637/jss.v027.i03.
- Hyndman RJ, Koehler AB, Ord JK, Snyder RD (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer. ISBN 9783540719168. doi:https://doi.org/10.1111/j.1751-5823.2009.00085_17.
- Jarque CM, Bera AK (1980). “Efficient tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals.” *Economics Letters*, **6**(3), 255 – 259. ISSN 0165-1765. doi:[https://doi.org/10.1016/0165-1765\(80\)90024-5](https://doi.org/10.1016/0165-1765(80)90024-5).
- Johnson VE (2004). “A Bayesian Chi Test for Goodness-of-Fit.” *Annals of Statistics*, **32**(6), 2361–2384. doi:10.1214/009053604000000616.
- Kass RE, Raftery AE (1995). “Bayes Factors.” *Journal of the American Statistical Association*, **90**(430), 773–795. ISSN 01621459. URL <http://www.jstor.org/stable/2291091>.
- Kwiatkowski D, Phillips PC, Schmidt P, Shin Y (1992). “Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How sure Are We that Economic Time Series Have a Unit Root?” *Journal of Econometrics*, **54**(1), 159 – 178. ISSN 0304-4076. doi:[https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y).
- Lobato I, Velasco C (2004). “A simple Test of Normality for Time Series.” *Econometric Theory*, **20**, 671–689. doi:10.1017/S0266466604204030.
- Lomnicki Z (1961). “Tests for Departure from Normality in the Case of Linear Stochastic Processes.” *Metrika: International Journal for Theoretical and Applied Statistics*, **4**(1), 37–62. URL <https://EconPapers.repec.org/RePEc:spr:metrik:v:4:y:1961:i:1:p:37-62>.
- Matamoros IAA, Nieto-Reyes A (2020). “An R Package For Normality in Stationary Processes.” **2009.10481**.
- Matamoros IAA, Torres CAC (2020). “varstan: An R Package for Bayesian Analysis of Structured Time Series Models With Stan.” **2005.10361**.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953). “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics*, **21**(6), 1087–1092. doi:10.1063/1.1699114.

- Migon H, Gamerman D, Louzada F (2014). *Statistical inference. An integrated approach*. Chapman and Hall CRC Texts in Statistical Science. Chapman and Hall. ISBN 9781439878804.
- Nieto-Reyes A, Cuesta-Albertos JA, Gamboa F (2014). “A Random-Projection Based Test of Gaussianity for Stationary Processes.” *Computational Statistics & Data Analysis*, **75**, 124 – 141. ISSN 0167-9473. doi:<https://doi.org/10.1016/j.csda.2014.01.013>.
- Osborn DR, Chui APL, Smith JP, Birchenhall CR (1988). “Seasonality and the Order of Integration for Consumption.” *Oxford Bulletin of Economics and Statistics*, **50**(4), 361–377. doi:[10.1111/j.1468-0084.1988.mp50004002.x](https://doi.org/10.1111/j.1468-0084.1988.mp50004002.x).
- Pearson K, Henrici OMFE (1895). “X. Contributions to the Mathematical Theory of Evolution. Skew Variation in Homogeneous Material.” *Philosophical Transactions of the Royal Society of London. (A.)*, **186**, 343–414. doi:[10.1098/rsta.1895.0010](https://doi.org/10.1098/rsta.1895.0010).
- Perron P (1988). “Trends and Random Walks in Macroeconomic Time Series: Further Evidence From a New Approach.” *Journal of Economic Dynamics and Control*, **12**(2), 297 – 332. ISSN 0165-1889. doi:[https://doi.org/10.1016/0165-1889\(88\)90043-7](https://doi.org/10.1016/0165-1889(88)90043-7).
- Pettit LI (1986). “Diagnostics in Bayesian Model Choice.” *Journal of the Royal Statistical Society. Series D (The Statistician)*, **35**(2), 183–190. ISSN 00390526, 14679884. URL <http://www.jstor.org/stable/2987522>.
- Psaradakis Z (2017). “Normality Tests for Dependent Data.” *Working and Discussion Papers WP 12/2017*, Research Department, National Bank of Slovakia. URL <https://ideas.repec.org/p/svk/wpaper/1053.html>.
- Psaradakis Z, Vávra M (2017). “A Distance Test of Normality for a Wide Class of Stationary Processes.” *Econometrics and Statistics*, **2**, 50 – 60. ISSN 2452-3062. doi:<https://doi.org/10.1016/j.ecosta.2016.11.005>.
- Royston JP (1982). “An Extension of Shapiro and Wilk’s W Test for Normality to Large Samples.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **31**(2), 115–124. ISSN 00359254, 14679876. URL <http://www.jstor.org/stable/2347973>.
- Said SE, Dickey DA (1984). “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order.” *Biometrika*, **71**(3), 599–607. ISSN 0006-3444. doi:[10.1093/biomet/71.3.599](https://doi.org/10.1093/biomet/71.3.599).
- Shapiro SS, Wilk MB (1965). “An Analysis of Variance Test for Normality (Complete Samples).” *Biometrika*, **52**(3-4), 591–611. ISSN 0006-3444. doi:[10.1093/biomet/52.3-4.591](https://doi.org/10.1093/biomet/52.3-4.591).
- Shumway R, Stoffer D (2010). *Time Series Analysis and Its Applications: With R Examples*. Springer Texts in Statistics. Springer New York. ISBN 9781441978646. URL <https://books.google.es/books?id=dbS5IQ8P5gYC>.
- Smirnov N (1948). “Table for Estimating the Goodness of Fit of Empirical Distributions.” *Annals of Mathematical Statistics*, **19**(2), 279–281. doi:[10.1214/aoms/1177730256](https://doi.org/10.1214/aoms/1177730256).

- Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002). “Bayesian measures of model complexity and fit.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**(4), 583–639. doi:10.1111/1467-9868.00353.
- Stoffer D (2020). *astsa: Applied Statistical Time Series Analysis*. R package version 1.10, URL <https://CRAN.R-project.org/package=astsa>.
- Trapletti A, Hornik K (2019). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-47., URL <https://CRAN.R-project.org/package=tseries>.
- Tsay R (2010). *Analysis of Financial Time Series*. Second edi edition. Wiley-Interscience, Chicago. ISBN 978-0470414354. doi:10.1002/0471264105.
- Vehtari A, Gelman A, Gabry J (2016). “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC.” *Statistics and Computing*, **27**(5), 1413–1432. ISSN 1573-1375. doi:10.1007/s11222-016-9696-4.
- Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC (2020). “Rank-Normalization, Folding, and Localization: An Improved \hat{R} for Assessing Convergence of MCMC.” doi:10.1214/20-BA1221. Advance publication.
- Wasserman L (2006). *All of Nonparametric Statistics*. Springer, New York. ISBN 9780387251455. doi:10.1007/0-387-30623-4.
- Watanabe S (2010). “Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory.” *Journal of Machine Learning Research*, **11**. URL <http://www.jmlr.org/papers/volume11/watanabe10a/watanabe10a.pdf>.
- West M, Harrison J (2006). *Bayesian Forecasting and Dynamic Models*. Springer Series in Statistics. Springer New York. ISBN 9780387227771. URL <https://books.google.nl/books?id=0mPgBwAAQBAJ>.
- Wickham H (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-0-387-98140-6. URL <http://ggplot2.org>.