International Journal of Forecasting 31 (2015) 1056–1066

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting long memory series subject to structural change: A two-stage approach

Fotis Papailias [a,b,*], Gustavo Fruet Dias [c,d]

[a] Queen's University Management School, Queen's University Belfast, UK
[b] quantf Research[1], UK
[c] Department of Economics and Business, Aarhus University, Denmark
[d] CREATES, Aarhus University, Denmark

ARTICLE INFO

Keywords:
Time series forecasting
Spurious long memory
Fractional integration
Local Whittle

ABSTRACT

A two-stage forecasting approach for long memory time series is introduced. In the first step, we estimate the fractional exponent and, by applying the fractional differencing operator, obtain the underlying weakly dependent series. In the second step, we produce multi-step-ahead forecasts for the weakly dependent series and obtain their long memory counterparts by applying the fractional cumulation operator. The methodology applies to both stationary and nonstationary cases. Simulations and an application to seven time series provide evidence that the new methodology is more robust to structural change and yields good forecasting results.

1. Introduction

The issue of analysing economic and other series which possess hyperbolically decaying autocorrelations has long been of concern in the time series analysis literature. The work of Granger (1980), Granger and Joyeux (1980) and Hosking (1981), among others, has been influential in the study and modelling of such *long memory* series; see Beran (1994) and Baillie (1996) for an extensive survey of this field.

There has been a major debate on the estimation of long memory series in both full and semi-parametric setups, e.g., see, among others, Abadir, Distaso, and Giraitis (2007), Beran, Bhansali, and Ocker (1998), Fox and Taqqu (1986), Hualde and Robinson (2011), Robinson (1995, 2006), Shimotsu and Phillips (2006) and Sowell (1992) for more details.

However, the literature on the forecasting of long memory series is still growing. Baillie, Kongcharoen, and Kapetanios (2012), Bhansali and Kokoszka (2002), Bhardwaj and Swanson (2006), Chan and Palma (1998) and Diebold and Lindner (1996), among others, have been concerned with predictions from ARFIMA models. A well-known approach is to obtain predictions using a truncated version of the infinite autoregressive representation of the model. Peiris (1987) and Peiris and Perrera (1988) discuss computationally feasible ways of calculating these predictions, and Crato and Ray (1996) and Poskitt (2007) analyse information criteria in order to determine the lag of the autoregression.

In this paper, we suggest the use of a two-stage forecasting approach (TSF). The TSF is a simple and intuitive methodology. We begin by estimating the long memory parameter using any consistent estimator, then apply the

* Correspondence to: Queen's University Management School, Queen's University Belfast, Riddel Hall, 185 Stranmillis Road, BT9 5EE, Northern Ireland, UK. Tel.: +44 2890974667; fax: +44 2890974201.
E-mail addresses: f.papailias@qub.ac.uk, f.papailias@quantf.com (F. Papailias).

1 www.quantf.com.

http://dx.doi.org/10.1016/j.ijforecast.2015.01.006

fractional differencing operator, resulting in the underlying weakly dependent series. Finally, we compute multi-step-ahead forecasts for the latter and apply the fractional cumulation operator[2] in order to obtain the corresponding forecasts for the original long memory series. A similar approach was adopted by Papailias, Kapetanios, and Taylor (2013), who are interested in the bootstrapping of long memory series.

Our claim is that, when forecasts of the underlying weakly dependent series are translated to their long memory equivalents, they should provide smaller forecast errors *on average*, given that the weakly dependent series is less persistent, and hence, the models are able to provide better forecasts. Therefore, TSF avoids any "loss" of information that might occur when employing the truncation of the infinite AR representation of the model.

It should be noted that we are not concerned with the nature of the estimation of the series, i.e., full or semi-parametric methods, and hence, we do not discuss the advantages and/or disadvantages of such methods. We simply rely on the consistency of the estimators for carrying out our forecasting methodology. In our simulations and applications, we use the Fully Extended Local Whittle (FELW) of Abadir et al. (2007); however, other consistent estimators can be used equivalently.

A common issue that often arises when working with real time series that might exhibit long memory is the possibility of structural change. This is commonly referred to as *spurious long memory*. In such cases, the change(s) in the structure of the series might be mistaken for long memory, or strong dependence and structural change may even coexist. This poses threats to the analysis, and consequently the forecasting, of the series. Diebold and Inoue (2001) were among the first in the field to analyse the phenomena of long memory and structural change jointly and to prove that structural change can be misinterpreted as long memory. Berkes, Horváth, Kokoszka, and Shao (2006), Iacone, Leybourne, and Taylor (2013), Lazarovà (2005), Ohanissian, Russell, and Tsay (2008), Qu (2011) and Shao (2011), among others, develop tests to accommodate this spurious long memory effect.

However, the question is, how should the applied researcher forecast series which might exhibit spurious long memory? What happens if the tests fail to distinguish between pure long memory and structural change? Wang, Bauwens, and Hsiao (2013) suggest that a simple autoregressive (AR) model should be used in the forecasting, as it does a good job of approximating an ARFIMA process that is subject to a mean shift or a change in the long memory parameter.

In this paper, we show via simulations that a simple AR model used in the second step of the TSF methodology suggested here results in accurate and more robust forecasts when applied to long memory series with a break in the mean or a change in the long memory parameter. This result is useful for practitioners, who can employ the methodology even when there is a possibility of spurious long memory. An empirical exercise on seven real time series illustrates the applicability and advantages of the TSF methodology, and in some cases of the truncated version of the infinite AR representation of the model, in an applied setup.

The rest of the paper is organised as follows: Section 2 provides some basic definitions and the algorithm of the proposed forecasting methodology. Section 3 introduces the structural change in long memory series and discusses the simulation results. Section 4 relates to the empirical exercise, and Section 5 summarises the conclusions.

2. Long memory: concepts and forecasting

2.1. Existing framework

We start by considering the following general fractionally integrated model

$$(1-L)^d x_t = u_t, \quad t = 1, \dots, T, \tag{1}$$

where L denotes the lag operator, d is the degree of long memory, and u_t is a weakly dependent, or short-range dependent, process. Hence, in the above setup, x_t is $I(d)$ and u_t is $I(0)$. We define $I(0)$ processes such that their partial sums converge weakly to Brownian motion; for more information regarding the definition of $I(0)$ processes, see Davidson (2002, 2009), Davidson and DeJong (2000), Müller (2008) and Stock (1994), among others. We model u_t as

$$u_t = \psi(L)\varepsilon_t, \tag{2}$$

with $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ and $E(\varepsilon_t\varepsilon_s) = 0$ for all $t \neq s$. $\psi(\lambda)$ is given by $\psi(\lambda) = \sum_{i=0}^{\infty} \psi_i \lambda_i$, where ψ_i is a sequence of real numbers such that $\psi_0 = 1$, $\sum_{i=0}^{\infty} |\psi_i| < \infty$ and $\sum_{i=0}^{\infty} \psi_i \lambda_i \neq 0$. In the case where u_t follows a stationary and invertible ARMA(p, q) model, x_t becomes the widely-known ARFIMA(p, d, q) model. For $|d| < 0.5$, the process is stationary and invertible, whereas for $d > 0.5$ the process is nonstationary. The above-defined process belongs to the Type I fractionally integrated process; see Marinucci and Robinson (1999) and Robinson (2005) for definitions regarding nonstationary processes and Baillie (1996) and Beran (1994) for a more detailed introduction to long memory processes.

We can write x_t as an infinite autoregression process as follows:

$$x_t = \sum_{i=1}^{\infty} \beta_i x_{t-i} + u_t, \tag{3}$$

where $\beta_i = \frac{\Gamma(i-d)}{\Gamma(i+1)\Gamma(d)}$, with $\Gamma(\cdot)$ being the gamma function. The above results follow from the definition of the fractional differencing operator, $(1-L)^d$, that is valid formally for any real d; see Hosking (1981) for more details.

The standard forecasting method in the literature suggests that, given a knowledge of the parameters and using Eq. (3), the theoretical s-step-ahead forecast conditional on the information available at time T is given by

$$\widehat{x}_{T+s} = \sum_{i=1}^{\infty} \beta_i x_{T+s-i}. \tag{4}$$

[2] This is the inverse of the fractional differencing operator.