Chapter

Time Trends and Persistence in the Snowpack Percentages by Watershed in Colorado

Luis Alberiko Gil-Alana

Abstract

In this paper we investigate the time trend coefficients in snowpack percentages by watershed in Colorado, US, allowing for the possibility of long range dependence or long memory processes. Nine series corresponding to the following watersheds are examined: Arkansas, Colorado, Gunnison, North Platte, Rio Grande, South Platte, San Juan-Animas-Dolores-San Miguel, Yampa & White and Colorado State-wide, based on annual data over the last eighty years. The longest series start in 1937 and all end in 2019. The results indicate that most of the series display a significant decline over time, showing negative time trend coefficients, and thus supporting the hypothesis of climate change and global warming. Nevertheless, there is no evidence of a long memory pattern in the data.

Keywords: snowpack percentages, time trends, long memory, Colorado, watersheds

1. Introduction

It is a well known fact that temperatures have been increasing during the last 50 years not only at global level but also at specific locations all over the world. In this paper we examine the statistical properties of nine time series corresponding to the snowpack percentages in watersheds in Colorado, US. Using annual data dating back to 1937, we are interested in the long memory feature of the data along with the time trend coefficients to check if the snowpack percentages have been declining in the last eighty years as a consequence of the effects of global warming. In addition, we test this hypothesis under the assumption that the underlying series display a long memory property, a feature that is very common in climatological data. As far as we know there are no previous works dealing with the statistical modeling of snow packs with time series data. Our results, using fractional integration, show no evidence of long memory, and the time trend coefficients of the snow packs are statistically significantly negative in the majority of the series examined, supporting thus the global warming hypothesis.

The standard approach to test for significant trends in time series is to consider a linear regression model of the following form:

$$y_t = \alpha + \beta t + x_t, \tag{1}$$

where a significant slope coefficient for β implies the presence of a trend (positive or negative, depending on the sign of the coefficient). However, in order to get consistent estimates of the unknown coefficients in (1), this set-up implicitly assumes that the error term, x_t must be well behaved and more specifically, it must be integrated of order 0 or I(0). This is a standard regularity condition and indicates that the infinite sum of its autocovariance values should be finite. This condition, however, is not always satisfied. If that sum is infinite, the series is said to be long memory, a feature widely observed in time series in many different disciplines including geophysical and climatological series, e.g., Beran [1], Percival et al. [2], Gil-Alana [3, 4], Ercan et al. [5], Graves et al. [6], etc.

In this article, this long memory feature is incorporated in our set-up by means of using a fractional integration model, which is described in the following section, and that is used to describe the error term x(t) in (1). Based on that, we test for the presence of significant time trends in the snowpack percentages at Colorado's watersheds.

The rest of the paper is structured as follows. Section 2 briefly describes the main idea of long memory or long range dependence processes and also presents the series under examination. Section 3 is devoted to the empirical results, while Section 4 concludes the manuscript.

2. Methodology and data: long memory

Given a covariance stationary process $\{x(t), t = 0, \pm 1, \ldots\}$ we say that it is short memory or integrated of order 0 (and denoted as $x(t) \sim I(0)$) if the infinite sum of its autocovariances, defined as $\gamma(u) = Cov(x(t), x(t + u))$ is finite, i.e.,

$$\sum_{u=-\infty}^{\infty} \gamma(u) < \infty. \tag{2}$$

Within this category, we can include the white noise process but also the standard stationary AutoRegressive Moving Average (ARMA) type of models. This latter category allows for some type of dependence between the observations and is named "weak" (dependence) in the sense that the autocorrelations decay exponentially fast. However, many time series show higher degrees of dependence and belong to a category denominated as "long memory", characterized because the infinite sum of the autocovariances is infinite, i.e.,

$$\sum_{u=-\infty}^{\infty} \gamma(u) = \infty. \tag{3}$$

This long memory feature has been observed in time series data referring to many different disciplines, including economics and finance [7–9], energy [10–13], tourism [14, 15], environmental issues [16] and climatology [3, 17–19] among many others.

A very simple model, very popular among econometricians, and satisfying the above property (3) is the fractionally integrated or I(d, d > 0) model, which is expressed as:

$$(1 - L)^d x(t) = u(t), \quad t = 1, 2, \ldots, \tag{4}$$

where L is the lag-operator, ie., $L^k x(t) = x(t-k)$, d can be any real positive value, and u(t) is I(0) or short memory as defined above. In this context, x(t) displays the property of long memory if d > 0. Using a Binomial expansion, the polynomial in the left-hand side in (4) can be expressed as:

$$(1-L)^d = 1 - dL + \frac{d\,(d-1)}{2}L^2 - \frac{d\,(d-1)\,(d-2)}{6}L^3 + \ldots \quad (5)$$

and x(t) in (4) can then be expressed as:

$$x(t) = d\,x(t-1) - \frac{d\,(d-1)}{2}x(t-2) + \frac{d\,(d-1)\,(d-2)}{6}x(t-3) + \ldots + \varepsilon(t) \quad (6)$$

and higher the value of d is, the higher the level of association between the observations is. A wide range of possibilities can be examined depending on the value of d in the real range. Examples are.

i. anti-persistence, if $d < 0$,

ii. short memory, if $d = 0$,

iii. stationary long memory processes, if $0 < d < 0.5$,

iv. nonstationary long memory mean reverting patterns, if $0.5 \le d < 1$,

v. unit root processes, if $d = 1$, and

vi. explosive patterns with $d > 1$.

This specification is clearly more general than the standard methods used in the literature and that are based only on integer degrees of differentiation, i.e., d = 0 for stationarity and d = 1 for unit root or nonstationarity.

The series examined refer to the snowpack percentages in seven watersheds in Colorado, US, (see, **Figure 1**) namely, Arkansas, Colorado, Gunnison, North Platte, Rio Grande, San Juan-Animas-Dolores-and-San Miguel, South Platte, Yampa &

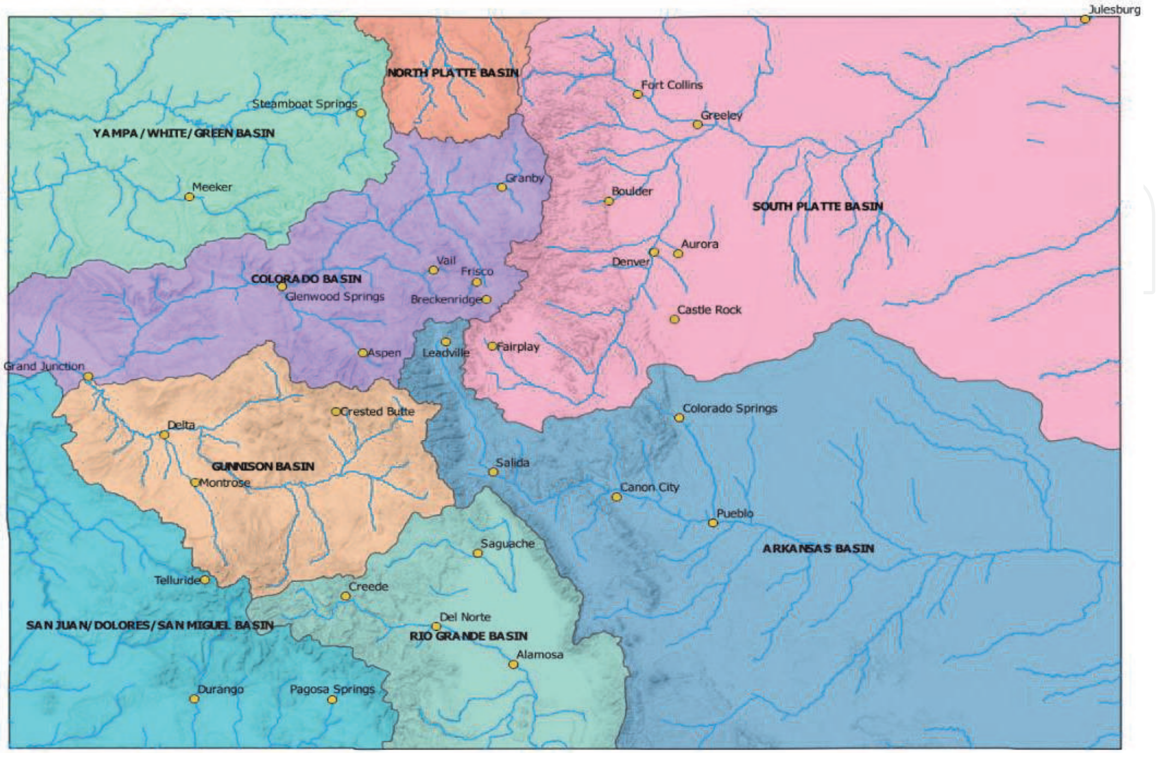

Figure 1.
Watersheds in Colorado, US. The different colors represent the different areas under study.

White and Colorado Statewide, using annual data. The starting year changes from one watershed to another and all of them end in 2019. The source of the data is the USDA (United States Department of Agriculture), Natural Resources Conservation Service Colorado (https://www.nrcs.usda.gov). All values use medians calculated from the most recent normal period between 1981 and 2010 For each basin we use data for the months of March, April and May, which are the months that present complete datasets for the whole period examined.

Table 1 displays the time series examined along with the acronym and the starting and ending years for each series. We observe that the long series are those referring to North Platte and Colorado Statewide (with data starting in 1937) and followed by South Platte and Colorado (starting in 1938), while the shortest one refers to San Juan, Animas, Dolores and San Miguel, starting in 1973.

The objective in this paper is twofold. First, to determine if long memory holds in the snowpack data examined, and second, to see if there is a decline in the time evolution of the data as a consequence of climate warming. Thus, in order to examine these two issues together, we consider the following model

$$y(t) = \alpha + \beta t + x(t); \quad (1 - L)^d x(t) = u(t), \tag{7}$$

where y(t) refers to the observed data, the snowpack percentages, α and β are unknown parameters referring respectively to a constant and a linear time trend, and the regressions errors x(t) are supposed to be I(d). Therefore, there are two main parameters in the above specification: β, related with the evolution over time of the series, and d, dealing with the short/memory feature of the data.

In the following section we display the estimated coefficients of d in Eq. (7) (along with their corresponding 95% confidence intervals) under three potential set-ups. The results presented in the second column in the table refer to the case where we impose α = β = 0 in (7); thus, no deterministic terms are included in the model. The results in column 3 refer to the model with an intercept, i.e. imposing β = 0 in (7). Finally, the last column reports the results where α and β are estimated from the data along with d.

We have marked the most appropriate case for each series in **Table 2** in bold. This selection has been made based on the significance of the estimated coefficients in the d-differenced process, noting that the two equations in (7) can be expressed in a single one as:

$$\tilde{y}(t) = \alpha \tilde{1}(t) + \beta \tilde{t}(t) + u(t), \tag{8}$$

Watershed	Acronym	Starting year	Ending year	n. of observ.
Arkansas	ARK	1950	2019	70
Colorado	COL	1938	2019	82
Gunnison	GUN	1941	2019	79
North Platte	NPB	1937	2019	83
Rio Grande	RIO	1950	2019	70
San Juan …	SJ	1973	2019	47
South Platte	SPB	1938	2019	82
Yampa & White	YAM	1952	2019	68
Colorado statewide	YAM	1937	2019	83

San Juan … refers to San Juan, Animas, Dolores and San Miguel.

Table 1.
Time series under examination.

where

$$\tilde{y}(t) = (1 - L)^d y(t), \tag{9}$$

$$\tilde{1}(t) = (1 - L)^d 1, \tag{10}$$

$$\tilde{t}(t) = (1 - L)^d t, \tag{11}$$

and noting that u(t) in (8) is I(0) by assumption, standard t-values hold here.

3. Results

We observe in **Table 2** that the time trend is required in 15 out of the 27 cases presented, while for the remaining 12, an intercept is sufficient to describe the deterministic part of the model. Of these 12 cases where only an intercept is required, five correspond to the month of March, another five occur in May and only two in April. Thus, the time trend is required in more than half of the series examined and as we will show below the coefficients are significantly negative in all cases implying that the snow packs are decreasing with time.

Watershed	Month	No terms	An intercept	A linear time trend
ARKANSAS BASIN	March	0.47 (0.20, 0.73)	0.02 (−0.13, 0.27)	**−0.16 (−0.40, 0.21)**
	April	0.47 (0.20, 0.74)	−0.02 (−0.15, 0.18)	**−0.28 (−0.47, 0.03)**
	May	−0.06 (−0.13, 0.48)	**−0.08 (−0.25, 0.18)**	−0.11 (−0.29, 0.17)
COLORADO BASIN	March	−0.05 (−0.08, 0.08)	**0.24 (−0.41, 0.01)**	−0.28 (−0.48, 0.00)
	April	0.08 (−0.08, 0.15)	−0.12 (−0.23, 0.06)	**−0.25 (−0.40, −0.02)**
	May	−0.07 (−0.11, 0.35)	**−0.19 (−0.36, 0.06)**	−0.19 (−0.37, 0.05)
GUNNISON BASIN	March	−0.06 (−0.11, 0.46)	**0.28 (−0.47, 0.00)**	−0.28 (−0.48, 0.00)
	April	−0.08 (−0.13, 0.41)	−0.11 (−0.23, 0.08)	**−0.20 (−0.35, 0.02)**
	May	−0.07 (−0.13, 0.29)	**−0.11 (−0.26, 0.13)**	−0.12 (−0.27, 0.12)
NORTH PLATTE BASIN	March	−0.03 (−0.06, 0.46)	**0.10 (−0.22, 0.09)**	−0.10 (−0.22, 0.09)
	April	0.09 (−0.08, 0.56)	**−0.05 (−0.17, 0.07)**	−0.11 (−0.21, 0.04)
	May	−0.06 (−0.10, 0.33)	**−0.09 (−0.20, 0.06)**	−0.13 (−0.25, 0.03)
RIO GRANDE BASIN	March	−0.13 (−0.21, 0.35)	−0.25 (−0.32, −0.06)	**−0.51 (−0.73, −0.22)**
	April	−0.08 (−0.12, 0.37)	**−0.12 (−0.24, 0.05)**	−0.17 (−0.29, 0.02)
	May	−0.02 (−0.10, 0.21)	**−0.03 (−0.15, 0.16)**	−0.04 (−0.16, 0.15)
SAN JUAN, ANIMAS, DOLORES AND SAN MIGUEL	March	−0.16 (−0.20, −0.07)	−0.27 (−0.44, 0.02)	**−0.53 (−0.82, −0.09)**
	April	0.00 (−0.23, 0.47)	0.00 (−0.13, 0.21)	**−0.44 (−0.73, −0.02)**
	May	−0.25 (−0.30, 0.51)	0.10 (−0.04, 0.31)	**−0.39 (−0.67, 0.07)**
SOUTH PLATTE BASIN	March	−0.06 (−0.10, 0.40)	**−0.09 (−0.22, 0.10)**	−0.15 (−0.31, 0.07)
	April	0.28 (0.15, 0.44)	0.04 (−0.06, 0.17)	**−0.16 (−0.31, 0.06)**
	May	0.27 (−0.12, 0.48)	0.03 (−0.09, 0.21)	**−0.16 (−0.35, 0.12)**
YAMPA & WHITE BASIN	March	0.23 (−0.04, 0.53)	0.18 (−0.31, 0.04)	**−0.37 (−0.63, −0.05)**
	April	0.04 (−0.21, 0.51)	−0.03 (−0.13, 0.12)	**−0.40 (−0.56, −0.14)**
	May	−0.06 (−0.33, 0.48)	−0.01 (−0.14, 0.18)	**−0.22 (−0.41, 0.08)**

Watershed	Month	No terms	An intercept	A linear time trend
COLORADO STATEWIDE BASIN	March	−0.04 (−0.09, 0.37)	**−0.19 (−0.33, 0.04)**	−0.21 (−0.37, 0.03)
	April	0.05 (−0.12, 0.45)	−0.04 (−0.14, 0.12)	**−0.18 (−0.32, 0.02)**
	May	−0.07 (−0.17, 0.36)	−0.06 (−0.19, 0.14)	**−0.15 (−0.30, 0.08)**

The values in parenthesis refer to the 95% confidence bands; in bold, the selected model for each series according to the deterministic terms.

Table 2.
Estimates of the differencing parameter.

Watershed	Month	No terms	An intercept	A linear time trend
ARKANSAS BASIN	March	−0.16 (−0.40, 0.21)	120.30 (27.37)	−0.338 (−2.96)
	April	−0.28 (−0.47, 0.03)	119.72 (43.53)	−0.348 (−4.65)
	May	−0.08 (−0.25, 0.18)	104.66 (29.77)	—
COLORADO BASIN	March	0.24 (−0.41, 0.01)	103.06 (102.35)	—
	April	−0.25 (−0.40, −0.02)	112.66 (48.55)	−0.153 (−2.86)
	May	−0.19 (−0.36, 0.06)	105.17 (57.61)	—
GUNNISON BASIN	March	0.28 (−0.47, 0.00)	106.76 (102.99)	—
	April	−0.20 (−0.35, 0.02)	114.57 (34.43)	−0.185 (−2.38)
	May	−0.11 (−0.26, 0.13)	108.27 (30.54)	—
NORTH PLATTE BASIN	March	0.10 (−0.22, 0.09)	104.10 (58.55)	—
	April	−0.05 (−0.17, 0.07)	106.25 (54.47)	—
	May	−0.09 (−0.20, 0.06)	106.45 (47.14)	—
RIO GRANDE BASIN	March	−0.51 (−0.73, −0.22)	112.80 (64.88)	−0.263 (−5.05)
	April	−0.12 (−0.24, 0.05)	99.769 (36.79)	—
	May	−0.03 (−0.15, 0.16)	92.624 (14.40)	—
SAN JUAN, ANIMAS, DOLORES AND SAN MIGUEL	March	−0.53 (−0.82, −0.09)	115.59 (44.98)	−0.461 (−4.03)
	April	−0.44 (−0.73, −0.02)	126.23 (38.08)	−1.005 (−7.06)
	May	−0.39 (−0.67, 0.07)	176.66 (26.53)	−2.556 (−9.12)
SOUTH PLATTE BASIN	March	−0.09 (−0.22, 0.10)	110.33 (49.03)	—
	April	−0.16 (−0.31, 0.06)	123.95 (34.76)	−0.337 (−4.26)
	May	−0.16 (−0.35, 0.12)	129.15 (26.99)	−0.391 (−3.68)
YAMPA & WHITE BASIN	March	−0.37 (−0.63, −0.05)	114.18 (57.22)	−0.225 (−3.88)
	April	−0.40 (−0.56, −0.14)	123.11 (71.75)	−0.441 (−8.73)
	May	−0.22 (−0.41, 0.08)	134.97 (22.75)	−0.663 (−4.09)
COLORADO STATEWIDE BASIN	March	−0.19 (−0.33, 0.04)	105.51 (83.54)	—
	April	−0.18 (−0.32, 0.02)	115.49 (38.53)	−0.222 (−3.35)
	May	−0.15 (−0.30, 0.08)	119.25 (22.40)	−0.267 (−2.30)

The values in parenthesis in the last two colums are their corresponding t-values.

Table 3.
Estimated coefficients of the selected models.

Table 3 displays the estimated coefficients of the selected models in **Table 2**. If we first focus on the values of d, we observe that there is no evidence of long memory behavior in any single case, since the confidence intervals include the value 0 in all cases examined. There are 21 series where the I(0) hypothesis of short memory cannot be rejected and in the remaining 6, anti-persistence ($d < 0$) is detected. The series showing anti-persistence are Colorado (April), Rio Grange (March) and San Juan … and Yampa & White (in March and April). Anti-persistence is not as frequent as persistence, though some examples can be found in the literature [20, 21]. Anti-persistent behavior exhibits prolonged damped oscillations with the spectral density function showing a zero value at the origin [22].

If we look now at the time trend coefficients the first noticeable feature is that all the significant coefficients are negative, supporting the hypothesis of a decline in the snowpack percentage in a number of cases. The highest coefficients refer to the cases of San Juan, Animas, Dolores and San Miguel in May (with an estimated time trend coefficient of −2.556) and in April (−1.005). Other high significant negative coefficients are those of Yampa & White (in May, −0.663, and April, −0.441) and San Juan … in March (−0.461). All these cases support the hypothesis of a decreasing trend in the snow packs in various Colorado's watersheds, which might be a consequence of the global warming climate hypothesis.

As a robustness method, we also employ alternative parametric and semiparametric methods of estimating the differencing parameter in the context of fractional integration, including among other Sowell's [23] maximum likelihood estimation method, the classical semiparametric Geweke and Porter-Hudak's [24] approach and the most recent developed method in Shimotsu [25] and the results support our conclusions in all cases, finding evidence of short memory and negative time trend coefficients in most of the series examined.

4. Conclusions

We have examined nine time series in this paper corresponding to snowpack percentages in Colorado, investigating if there has been a significant decline over time in the series in the context of long memory processes. For this purpose, we have tested for the significance of the time trend coefficient in a model where the regression errors are fractionally integrated or integrated of order d. Long memory occurs then if d is a positive value.

Our results indicate that there is no evidence of long memory behavior since all the orders of integration are close to zero or below it, implying short memory or anti-persistence behavior. Focusing on the time trend coefficients, these are significant in 15 of the 27 series examined, and in all these cases, they are found to be significantly negative, supporting thus the hypothesis of a decline in the amount of snow in Colorado watersheds probably as a consequence of global climate warming.

Future work should focus on alternative modeling approaches including for example non-linear structures which are clearly related with long memory and fractional integration models [26]. Thus, non-linear deterministic terms like those based on Chebyshev polynomials in time [27] or Fourier transforms [28], in both cases based on I(d) models, can also be implemented on these or on similar data.

Acknowledgements

The author gratefully acknowledges financial support from the MINEIC-AEI-FEDER ECO2017-85503-R project from 'Ministerio de Economía, Industria y

Competitividad' (MINEIC), 'Agencia Estatal de Investigación' (AEI) Spain and 'Fondo Europeo de Desarrollo Regional' (FEDER).

Comments from the Editor and an anonymous reviewer are also gratefully acknowledged.

Conflict of interest

The authors declare no conflict of interest.

Author details

Luis Alberiko Gil-Alana
University of Navarra, Pamplona, Spain

*Address all correspondence to: alana@unav.es

References

[1] Beran, J. (1994). Statistics for long-memory processes. Chapman & Hall: New York, NY, USA, 1–315.

[2] Percival, D.B., Overland, J.E. and Mofjeld, H.O. (2001) Interpretation of North Pacific variability as a short- and long-memory process. Journal of Climate, 14(24), 4545–4559.

[3] Gil-Alana, L.A, (2005), Statistical model of the temperatures in the northerm hemisphere using fractionally integrated techniques. Journal of Climate 18 (24), 5357–5369.

[4] Gil-Alana, L.A. (2006) Nonstationary, long memory and antipersistence in several climatological time series data. Environmental Modeling and Assessment, 11(1), 19–29.

[5] Ercan, A., Kavvas, M.L., Abbasov, R. K. (2013). Long-range dependence and sea level forecasting. Springer International Publishing, 1–51.

[6] Graves, T., Gramacy, R., Watkins, N., and Franzke, C. (2017). A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA, 1951–1980. Entropy, 19(9) 437, 1–21.

[7] Abbritti, M., L.A. Gil-Alana, Y. Lovcha and A. Moreno (2016). Term structure persistence, Journal of Financial Econometrics, 14, 2, 331–352.

[8] Gil-Alana, L.A. and A. Moreno (2012). Uncovering the U.S. term premium.An alternative route. Journal of Banking and Finance, 36, 1184–1193.

[9] Haug, A.. (2014), On Real Interest Rate Persistence: The Role of Breaks, Applied Economics, 46, 1058–1066

[10] Barros, C.P., L.A. Gil-Alana and P. Wanke, 2016, Energy production in Brazil. Empirical facts based on persistence, seasonality and breaks, Energy Economics 54, 88–95,

[11] Elder, J., Serletis, A.(2008). "Long memory in energy futures prices". Review of Financial Economics . *17* , 146–155

[12] Gil-Alana, L.A., H. Carcel and R. Mudida (2017), Shocks affecting electricity prices in Kenya. A fractional integration study, Energy 124, 1, 521–530.

[13] Solarin, S.A., L.A. Gil-Alana and C. Lafuente, 2020, An investigation of long range reliance on shale oil and shale gas production in the US market, Energy 195.

[14] Al-Shboul, M. and S. Anwar, 2017, Long memory behavior in Singapore's tourism market, International Journal of Tourism Research 19, 5, 524–534.

[15] Gil-Alana, L.A. and P. Huijbens (2018), Tourism in Iceland. Persistence, seasonality and long memory behaviour, Annals of Tourism Research 68, 20–29.

[16] Gil-Alana, L.A., S.A. Solarin and C. Lafuente (2019), Persistence in carbon footprint emissions. An overview over of 92 countries, Carbon Management 10, 4, 405–415.

[17] Gil-Alana, L.A. and M. Monge, (2020), Global CO_2 emissions and global temperatures. Are they related, International Journal of Climatology, forthcoming.

[18] Rea, W., M. Reale and J. Brown, 2011, Long memory in temperature reconstruction, Climatic Change 107, 247–265.

[19] Rodrigo-Caballero, S., J. Brix and A. Brix, 2020, Long memory in surface air temperatura, Climate Research 21, 127–140.

[20] Caporale, G. M., Gil-Alana, L. A., and Poza, C., (2020). Persistence, non-linearities and structural breaks in European stock market indices. The Quarterly Review of Economics and Finance 77, 50–61.

[21] Dimitrova, V., Fernández-Martínez M., Sánchez-Granero, M. A., and Trinidad Segovia, J. E., (2019). Some comments on Bitcoin market (in) efficiency. Plos One 14, e0219243.

[22] Dittmann, I., and Granger, C. W. J., (2002). Properties of nonlinear transformations of fractionally integrated processes. Journal of Econometrics. 110, 113–133.

[23] Sowell, F., (1992), Maximum likelihood estimation of stationary univariate fractionally integrated time series models, Journal of Econometrics 53, 1–3, 165–188.

[24] Geweke, J. and S. Porter-Hudak, (1982), The estimation and application of long memory time series models, Journal of Time Series Analysis 4, 4, 221–238.

[25] Shimotsu, K. (2010): Exact Local Whittle Estimation Of Fractional Integration with Unknown Mean and Time Trend. Econometric Theory, Vol. 26, pp. 501–540.

[26] Diebold, F.X. and A. Inoue (2001), Long memory and regime switching, Journal of Econometrics 105, 131–159.

[27] Cuestas J.C. and L.A. Gil-Alana (2016). "Testing for long memory in the presence of non-linear deterministic tre nds with Chebyshev polynomials," Stud ies in Nonlinear Dynamics and Econome trics, De Gruyter, vol. 20(1), pages 57–74, February.

[28] Gil-Alana, L.A. and O. Yaya, 2020, Testing fractional unit roots with non-linear smooth break approximations using Fourier functions, Journal of Applied Statistics, DOI: 10.1080/02664763.2020.1757047