

## Use of fresh water over the long run measuring persistence

Marta del Rio <sup>a</sup>, Juan Infante <sup>a</sup> and Luis A. Gil-Alana <sup>b,c,\*</sup>

<sup>a</sup> Universidad Villanueva, Madrid, Spain

<sup>b</sup> Faculty of Economics and ICS, University of Navarra, E-31080 Pamplona, Spain

<sup>c</sup> University Francisco de Vitoria, Facultad de Ciencias Empresariales y Jurídicas, Madrid, 28223, Spain

\*Corresponding author. E-mail: alana@unav.es

 MDR, 0000-0003-3233-7831; JI, 0000-0003-2094-6656; LAG, 0000-0002-5760-3123

### ABSTRACT

In this article, we carry out a study of the degree of persistence of a time series of data on freshwater use in the long term, using fractional integration or I(d) techniques. Using annual data from 1901 to 2014, we observe that the order of integration of the series is close to 1 if the errors are not correlated, and if they are correlated, then it is greater than 1. This shows that the series is highly persistent. On the other hand, we detect two structural breaks, one at 1951 and the other one at 1980. In these cases, we observe a reversion to the mean since the integration orders are much lower in the subsamples. This supports the hypothesis that when these breaks are not taken into account, there is an overestimation of the differentiation parameter, misspecifying the reversion to the mean of the data. The series also shows segmented trends with the higher time trend coefficient observed during the years 1951 and 1980.

**Key words:** Fractional integration, Long memory, Persistence, Time trends, Use of water

### HIGHLIGHTS

- The degree of persistence of freshwater use is examined with fractional integration.
- Annual data are examined from 1901 to 2014.
- The order of integration is close to 1 if the errors are not correlated, and if they are correlated, then it is greater than 1.
- This shows that the series is highly persistent.
- Two breaks were found in 1951 and 1980 finding mean reversion.

## 1. INTRODUCTION

The increase in water consumption has multiplied sixfold since the beginning of the 20th century. This has been due to the increase in the world population and an economy that needs more resources, in all sectors: agricultural, industrial and private consumption. We see that it is since the middle of the 20th century that the greatest increase occurs, but that in the 21st century, it has begun to slow down.

We can make a distribution by geographical areas regarding the consumption of fresh water in the world. Although consumption has increased since 1900, the distribution between these areas has not changed. The OECD nations consume 20–25%, the so-called BRICS (Brazil, Russia, India, China and South Africa) countries consume 45% and the rest represent between 30 and 33%. Regarding specific countries and taking 2014 as a

---

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

reference, it was India that had the most consumption (760 billion cubic metres per year), followed by China (around 600 billion cubic metres) and the United States (480–490 billion cubic metres).

The amount of fresh water that is extracted for agricultural, industrial or domestic consumption varies greatly from one place to another in the world. The average estimate of that amount per capita per year is estimated as the water that is extracted from underground or surface sources (lakes or rivers). The factors that determine this amount are very varied, from latitude, climate, the importance of the agricultural sector to how developed the industry is in a given country.

It is important that the levels of water resources are sustainable. To achieve this, the extraction of water must be below its replacement. Renewable internal flows (internal river flows and groundwater from rains) are the most important indicators to see if a country can be sustainable in terms of freshwater use in the long term. It is evident that if the extraction is greater than the renewable flows, the resources begin to diminish.

According to the United Nations, freshwater resources will decrease by 40% by 2030. [The UN World Water Development Report \(2020\)](#) forecasts an increase in water demand of 20–30% by 2050. In this scenario, the international body declared the decade 2018–2028 as the Decade of Action for Water: Water and Sustainable Development, which is an unequivocal recognition of water as a key factor.

The concept of sustainability for Urban Water Cycle Solutions (UWCS) was also defined as ‘those water resources systems designed and managed to contribute fully to society’s goals now and in the future while maintaining their ecological, environmental and hydrological integrity’. This definition emphasizes primarily the environmental dimension, although ‘societal goals’ also encompass the economic and social dimensions.

Urban water services have evolved significantly over time, as customers and society demand more. According to [Da Cruz & Marques \(2013\)](#), it is no longer enough to have an adequate drinking water service. Water services must be efficient, effective and responsive to customers’ needs.

Stakeholders comprising several areas with different objectives and interests make governance issues (e.g. participation and transparency) also very important in this area, so in addition to taking into account economic, social and environmental aspects, we need to take into account technical and governance aspects when assessing the sustainability of water services.

The total amount of renewable flows and the size of the population influence the renewable resources per capita. If there is a great variability of annual rainfall in a country – as is the case of the monsoon seasons – and there is a decrease in renewable resources, there will be a decrease in renewable extractions per capita. In the same way, although renewable sources remain constant, if the population increases, there will also be a decrease in renewable resources per capita, as is already happening in many countries.

In this study, we deal with the analysis of the time series of the use of fresh water in the long term and its statistical characteristics. Specifically, we want to determine the degree of integration it has and if the sudden changes that occur at some moments have a temporary or permanent effect. For this, we use fractional integration methods, since it seems more appropriate in this case than the standard methods based exclusively on integers that determine stationarity  $I(0)$  and non-stationarity  $I(1)$ . Being able to obtain a fractional integration order allows a greater degree of flexibility in the dynamic specification of the model, which is important from a perspective of the policy that must be implemented.

The main contributions of this work are: (1) the use of fractional integration or  $I(d)$  techniques to study the degree of persistence of a time series of long-term freshwater use data, and (2) the dataset used based on annual global freshwater use. In addition, the use of structural breaks in the context of fractional integration is another methodological contribution of the paper. The rest of the paper is structured in the following sections. A literature review is in Section 2. Data used in the paper are included in Section 3. In Section 4, we present the methodology and empirical results. We finish with conclusions in Section 5.

## 2. LITERATURE REVIEW

The increase in agricultural production in recent years has been due to the massive use of water through irrigation or controlled flooding (World Bank, 2008). This has led to an increase in the water used for this sector, especially in South and East Asia and the Middle East. If we look at use by country, we see that Pakistan, Bangladesh and South Korea irrigate more than half of their agricultural land, while India irrigates 35%. In the case of the African continent, the sub-Saharan part has increased the extraction of water for agriculture, but it is still less than in the north of the same continent. This may explain the lower crop yields in the sub-Saharan part, together with other causes such as the use of fewer fertilizers and less efficient crop varieties. Worldwide, the use of water for agriculture is at 70% of the total extraction.

The use of water in agriculture varies according to the country and the wealth of its inhabitants (Gleick & Heberger, 2014). If we take into account this last factor, the richest countries only consume 41% in agriculture, for the countries in the intermediate zone it is 79%, and for the poorest it reaches 90%. This figure may even be exceeded for countries in Latin America, Africa and South Asia. The record is held by Sudan, with 96%. On the contrary, the countries of the Northern Hemisphere are those that tend to use less water for agriculture. With less than 1%, we find Germany and the Netherlands.

As we mentioned above, the amount of fresh water available is related to the extractions that are made and the internal (renewable) resources that are available. These latter data should be treated with caution, since they are collected intermittently and do not take into account seasonal and annual changes, which can be of great importance for certain countries. On the other hand, sometimes there are some very local areas which are overlooked and not included in the national calculations. If there is a greater consumption of water than renewable water resources, we are faced with water stress. The World Resources Institute has drawn up a classification according to the following percentages: if more than 80% is extracted, the stress is extremely high; if it is between 40 and 80%, it is a high stress; for values between 20 and 40%, we would be talking about a medium to high stress; finally, if it is between 10 and 20%, the stress is low. According to Gassert *et al.* (2013), if we analyse these data for different countries, we see that the Middle East, North Africa and South Asia have extremely high levels. In some cases, it may even be above 100%, as is the case in Saudi Arabia, Egypt, the United Arab Emirates, Syria, Pakistan and Libya. With values between 40 and 80%, we find countries from South Asia and a little below are East Asia, the United States and much of southern and eastern Europe (20–40%). Finally, those that are in the lowest part of water stress (10–20%) are northern Europe, Canada, much of Latin America, sub-Saharan Africa and Oceania.

For the study of residential water demand, Danielson (1979) conducted a research that estimated consumption taking into account rainfall, house value, the price of water and the size of the household. The sample consisted of 261 family residences in Raleigh, North Carolina, between May 1969 and December 1974. The methodology used provided unbiased estimates of parameters and standard errors with data showing serially correlated residuals. The estimated demand was calculated taking into account the total for houses for the winter and the use in irrigation. By subtracting the use in winter (November–April) from the summer (May–October), it was possible to determine the hose and sprinkler use per customer for each year. The greatest variation in the data was given by the size of the household and the use of a greater number of sprinklers was given by the price of water and changes in the climate. However, total demand and winter use responded less to price changes.

Beck (1987) reviewed the role of uncertainty in the identification of mathematical models of water quality and in the application of these models to problems of prediction. More specifically, four areas were examined in detail: uncertainty about model structure, uncertainty in the estimated model parameter values, the propagation of prediction errors and the design of experiments in order to reduce the critical uncertainties associated with a model. Wada *et al.* (2016) assessed the state of the art for estimating and projecting water use regionally and

globally in a consistent manner. It provided an overview of different approaches, the uncertainty, strengths and weaknesses of the various estimation methods, types of management and policy decisions for which the current estimation methods are useful. They also discussed additional information most needed to be able to improve water use estimates and be able to assess a greater range of management options across the water–energy–climate nexus. Other recent papers dealing with the modelization of water consumption and prices include among others Reynaud (2015), Bi *et al.* (2019), García-Lopez & Montano (2020) and Zheng *et al.* (2022).

Within the multivariate AutoRegressive Moving Average (ARMA) models, Salas *et al.* (1985) proposed alternative approaches to model multiple series of water resource systems. The methodology was applied in an iterative process in three phases: identification of the model, estimation of parameters and diagnostic checks. If high-order vector models are used, even with statistical tools available to follow this process, it is not an easy task. However, for the calculations of water resources, simpler models can be used, such as contemporary models and transfer function models. Accordingly, two examples of series models for bivariate and trivariate river flows were proposed. The statistical characteristics of the analysed time series can be reliably reproduced by using low-order models, as well as contemporary ARMA models. We can deduce from these results that the same conclusions would apply to the time series of water resources.

Water prices were investigated by Monge & Gil-Alana (2020) in different regions around the world using fractional integration methods. Unlike the ARMA models, where category  $I(0)$  is given, the proposed model allows fractional degrees of differentiation. Series corresponding to global data and to the following regions were studied: Asia Pacific and Russia, Europe, the United States and Latin America. The values obtained were close to one, indicating a high degree of persistence, even higher under the assumption of uncorrelated errors. On the other hand, a small degree of reversion to the mean could be verified for all cases except for Latin America, if autocorrelation was allowed. It is true that several breaks were observed in the series analysed (three in Latin America and global data; four in Europe and the United States, and five in Asia Pacific and Russia). In all of them, however, a significant change in the degree of persistence was not demonstrated, but when autocorrelation was allowed, the reversion to the mean appeared once again.

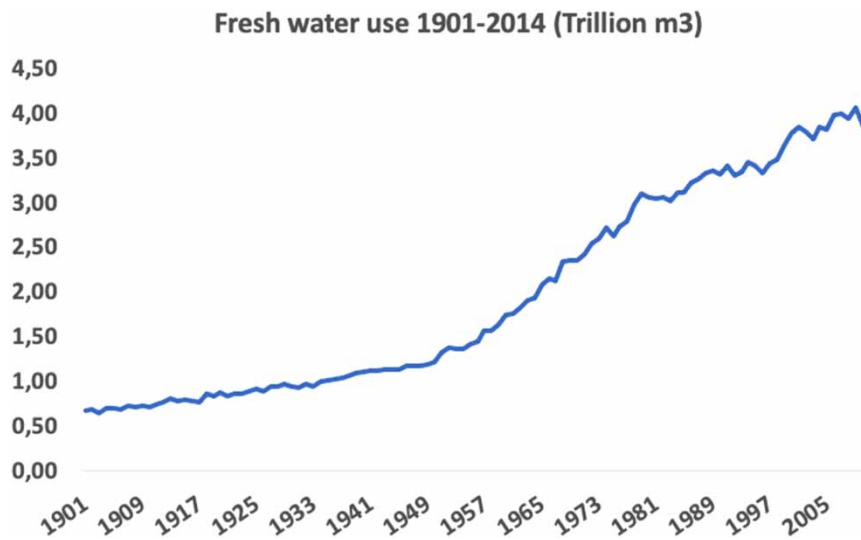
### 3. DATA

Data series used for our calculation goes from 1901 to 2014. Data are taken from the Global International Geosphere-Biosphere (IGB) Programme and World Bank. Data measures global freshwater use which is the sum of water withdrawals for agriculture, industrial and domestic uses.

Resources of data series can be divided into two periods. From 1900 to 2010, they are sourced from the IGB Database. IGB's data is estimated using the WaterGAP model from Flörke *et al.* (2013). Global data has been extended to 2014 as reported in the World Bank – World Development Indicators, under the variable 'Annual Freshwater Withdrawals · Total (billion cubic metres)'. Available at: <http://data.worldbank.org/data-catalog/world-development-indicators> (accessed 8 November 2017).

Data are available at aggregates in OECD, BRICS and the Rest of the World (ROW). OECD members are defined as countries who were members in 2010 and their membership was carried back in time. BRICS countries are Brazil, Russia, India, China and South Africa.

Figure 1 shows the time-series plot. We observe an increase in the use of freshwater over the long run, although in some periods it is higher than in others. At the beginning of the series, the increase is small, but after the 1950s, it is higher due to drought in that decade. After the 1980s, we see a decrease on the slope due to better implementation of irrigation techniques. In the 1990s, Emerging Market Countries increased rapidly agricultural water consumption – almost doubling between 1975 and 2010 – as its population and total food demand increased. Since 2000 appears to be plateauing, or at least slowing.



**Fig. 1** | Time-series plots.

#### 4. METHODOLOGY AND EMPIRICAL RESULTS

In order to allow for non-stochastic terms and the following standard practice, we consider a regression model of the following form,

$$y_t = \beta_0 + \beta_1 t + x_t, \quad t = 1, 2, \dots, \quad (1)$$

where  $y_t$  represents the series of interest, and  $\beta_0$  and  $\beta_1$  are unknown coefficients corresponding, respectively, to an intercept and a linear time trend; in addition, the error term, which is  $x_t$  in (1), is supposed to be integrated of order  $d$  or  $I(d)$  and represented as:

$$(1 - B)^d x_t = u_t, \quad t = 1, 2, \dots, \quad (2)$$

where  $B$  is the backshift operator, i.e.  $B^k x_t = x_{t-k}$ , and  $u_t$  is  $I(0)$  (see Granger & Joyeux, 1980; Hosking, 1981). The left-hand side of (2) can be expanded using the Binomial expansion,

$$(1 - B)^d = \sum_{j=0}^d \binom{d}{j} (-1)^j B^j = 1 - dB + \frac{d(d-1)}{2} B^2 - \dots \quad (3)$$

such that the higher the value of  $d$ , the higher the degree of association between observations distant in time. The differencing parameter  $d$  actually plays a very important role from many different viewpoints. Firstly, the estimation of the deterministic terms in (1) can be seriously affected if the differencing parameter  $d$  is positive (displaying long memory) instead of  $d = 0$  as in the standard short memory case (i.e. ARMA processes); secondly, this parameter can also be taken as a measure of the degree of persistence in the data, and it allows us to consider

non-stationary models albeit with mean-reverting behaviour if  $0.5 \leq d < 1$ ; the lack of mean reversion takes place if  $d \geq 1$ .

For inference, we use a version of a testing approach suggested in Robinson (1994) that is based on the Lagrange Multiplier (LM) principle and that uses a frequency domain version of the Whittle function. Using this method, we test the null hypothesis:

$$H_0: d = d_0 \quad (4)$$

in the model given by Equations (1) and (2), where  $d_0$  can be any real value and thus potentially include fractional numbers, even those belonging to the non-stationary range ( $d \geq 0.5$ ). This permits us to study a wide range of possibilities, including, for instance, the short memory case ( $d = 0$ ); stationary long memory processes ( $0 < d < 0.5$ ); non-stationary though mean reverting processes ( $0.5 \leq d < 1$ ), unit roots ( $d = 1$ ) or explosive patterns ( $d \geq 1$ ). The test is the most efficient one in the Pitman sense against local departures from the null (Robinson, 1994), and the limit distribution of his test is standard  $N(0, 1)$  independently of the inclusion of the non-stochastic terms and the way of modelling the error term, and the functional form of the test statistic can be found in any of the numerous empirical applications based on the test (see, e.g., Gil-Alana & Robinson, 1997).

Table 1 displays the estimated values of  $d$  in the model given by Equations (1) and (2) under different scenarios. Thus, we consider first the case where no deterministic terms are included, i.e., imposing  $\beta_0 = \beta_1 = 0$  *a priori* (in column 2); then, we include an intercept, i.e., with  $\beta_1 = 0$  (column 3), and finally, both  $\beta_0$  and  $\beta_1$  are freely estimated from the data. The first row displays the estimates under the assumption that  $u_t$  in (2) is a white noise process, while in the following row, we impose autocorrelation throughout the non-parametric method of Bloomfield (1973).<sup>1</sup> In all cases, we display the estimates of  $d$  along with the values where we cannot reject the null hypothesis (4) using the tests of Robinson (1994) at the 5% level. We have marked the selected specification in bold in the table in relation to the deterministic terms, and we observe here that the time trend is required in the two cases of uncorrelated and autocorrelated errors.

Table 2 displays the estimated coefficients of the selected models, and we see that the time trend is significantly positive; the estimates of  $d$  are 0.98 (with a confidence interval of 0.91, 1.08) with white noise errors and much higher, 1.17 (1.04, 1.34) under autocorrelation; thus, the unit root null hypothesis (i.e.,  $d = 1$ ) cannot be rejected in the former case, being rejected in favour of  $d > 1$  in the latter. Thus, according to these results, the series is highly persistent, with a large degree of dependence and showing a lack of mean-reverting behaviour.

**Table 1** | Estimates of the differencing parameter.

Type of errors	No terms	An intercept	An intercept with a linear time trend
White noise	0.94 (0.87, 1.08)	0.99 (0.93, 1.07)	<b>0.98 (0.91, 1.08)</b>
Autocorrelation	0.91 (0.82, 1.06)	1.14 (1.04, 1.29)	<b>1.17 (1.04, 1.34)</b>

In parenthesis, the 95% confidence intervals for the differencing parameter. In bold, the selected model.

<sup>1</sup> Bloomfield (1973) proposed a model that is implicitly described in terms of its spectral density function and that produces autocorrelations decaying exponentially as in the case of autoregressions.

**Table 2** | Estimated coefficients in the selected models.

Type of errors	$d$ (95% band)	Intercept (t-value)	Time trend (t-value)
White noise	0.98 (0.91, 1.08)	0.638 (10.15)	0.030 (5.53)
Autocorrelation	1.17 (1.04, 1.34)	0.648 (10.82)	0.028 (2.38)

These results, however, can be biased due to the presence of structural breaks in the data.<sup>2</sup> In fact, many authors argue that if breaks (or any other type of non-linear structures) are present in the data and they are not taken into account, the evidence of long memory can be overestimated (Diebold & Inoue, 2001; Granger & Hyung, 2004; Sibbertsen, 2004; Choi *et al.*, 2010; etc.).

Based on the above, we next conducted the approach developed in Gil-Alana (2008) that basically identifies breaks in the context of I(d) models. Using this method, and allowing for two breaks as maximum, we identify them at 1951 and 1980 (see Table 3). The same results are obtained when using Bai & Perron's (2003) approach. This is not surprising noting that Gil-Alana (2008) is basically an extension of Bai & Perron's (2003) method to the fractional case and the estimates of  $d$  are close to 1 in the two series.

We can explain the breaks in the following way. For the first, droughts come in cycles, and the 1950s saw a series of dry years that caused farmers to dig thousands of new irrigation wells, hastening the development of centre pivot irrigation systems and causing a crisis as underground aquifers fell. Hoerling *et al.* (2009) focused on the relationship between global sea surface temperatures and drought indices derived from both historical observations and a multi-model suite of climate simulations. Distinct factors were found to be responsible for the 1930s and the 1950s droughts.

On the other hand, after 1980 water use continued to increase, albeit at a lower rate, possibly due to the world making more use of water-conservation measures. Chambers (1980) studied the main system operation of large and medium-size canal irrigation in South and Southeast Asia. The conclusions were designed to complement and support the many initiatives that were being undertaken or contemplated to improve large and medium-size irrigation systems.

Jägermeyr *et al.* (2015) found pronounced regional patterns in beneficial irrigation efficiency (a refined irrigation efficiency indicator accounting for crop-productive water consumption only), due to differences in different

**Table 3** | Break dates.

Series	Number of breaks	Break dates
Fresh water	2	1951; 1980

<sup>2</sup> When climate change is greater, more droughts occur in the short term (less than 6 months), especially in the tropics and mid-latitudes. If we talk about mid and high latitudes, then droughts occur within 7–12 months. Finally, it is the northern latitudes and sub-Saharan Africa where the longest droughts occur, which can last up to 12 months. An extreme case is that of the Sahel, where drought situations occur over a long period of time and of extreme severity. We can also speak of drought limited to very specific areas, such as the one that occurred in northern Europe in 1996 (3 months) and 1975 (12 months), and if we take into account soil moisture, there are some anomalies during the winter that have led to droughts in northern Asia. The most extensive and severe known droughts were those of 1998 in the USA, 1982/83 in Australia, 1983/4 in the Sahel and 1965/66 in India (Sheffield & Wood, 2007).

features, with the lowest values (<30%) in South Asia and sub-Saharan Africa and the highest values (>60%) in Europe and North America. Cui *et al.* (2022) demonstrated that increasing agricultural machinery and urbanization rates and reducing the agricultural water rate are conducive to improving the resource utilization efficiency in pumping irrigation systems.

Tables 4 and 5 present the estimated coefficients for each subsample for the same cases as in Table 1, i.e. with no deterministic terms, with an intercept, and with an intercept and a linear time trend, using both uncorrelated and autocorrelated errors.

The first thing we observe in these tables is that the time trends are once more significant in all cases; however, the orders of integration are now much smaller, particularly if the errors are autocorrelated. Starting with the case of white noise errors, we observe that the estimated coefficients are 0.30, 0.54 and 0.42, respectively for the first, the second and the third subsamples, respectively, and though the confidence intervals are wide (due to the shorter sample sizes), the unit root null hypothesis is rejected in the three cases in favour of mean reversion. If autocorrelation is permitted, the estimated values of  $d$  are even smaller, even being negative in the second and third subsamples (−0.07 and −0.48), but the confidence intervals are now extremely large. Thus, for example, for the third subsample, we cannot reject the nulls of I(0) and I(1) behaviour.

**Table 4** | Estimates of the differencing parameter for each subsample.

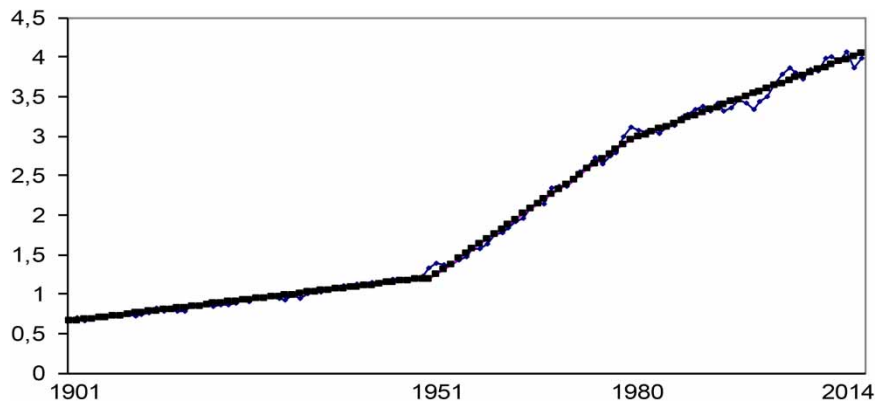
Type of errors	No terms	An intercept	An intercept with a linear time trend
(i) White noise errors			
1st period	0.92 (0.70, 1.20)	0.71 (0.64, 0.78)	<b>0.30 (0.15, 0.52)</b>
2nd period	0.86 (0.54, 1.24)	0.81 (0.70, 0.95)	<b>0.54 (0.21, 0.92)</b>
3rd period	0.84 (0.55, 1.22)	0.73 (0.62, 0.93)	<b>0.42 (0.04, 0.88)</b>
(ii) Autocorrelated errors			
1st period	0.67 (0.29, 1.21)	0.83 (0.73, 0.95)	<b>0.41 (0.10, 0.85)</b>
2nd period	0.45 (0.28, 1.30)	0.72 (0.05, 0.97)	−0.07 (−0.65, 0.83)
3rd period	0.15 (0.08, 1.12)	0.77 (0.54, 1.13)	−0.48 (−1.06 1.11)

In parenthesis, the 95% confidence intervals for the differencing parameter. In bold, the selected model.

**Table 5** | Estimates of the differencing parameter for each subsample.

Type of errors	No terms	An intercept	An intercept with a linear time trend
(i) White noise errors			
1st period	0.30 (0.15, 0.52)	0.637 (54.15)	0.011 (29.10)
2nd period	0.54 (0.21, 0.92)	1.201 (23.42)	0.062 (18.56)
3rd period	0.42 (0.04, 0.88)	2.988 (52.27)	0.032 (10.65)
(ii) Autocorrelated errors			
1st period	0.41 (0.10, 0.85)	0.641 (17.79)	0.011 (9.01)
2nd period	−0.07 (−0.65, 0.83)	1.126 (9.00)	0.063 (8.53)
3rd period	−0.48 (−1.06 1.11)	2.951 (61.48)	0.034 (11.33)





**Fig. 2** | Data and a segmented time trend.

Figure 2 displays the estimated trends (which are similar in the two cases of white noise and autocorrelated errors, see Table 5). We observe a sharp increase in the trend around 1951, followed by another, smoother one, in 1980.

## 5. CONCLUSION

In this paper, we have examined the time series behaviour of the use of fresh water over the period from 1901 to 2014 using annual data. Looking at the order of integration of the series from a fractional perspective, the first thing we observe is that the series is highly persistent with an order of integration equal to or higher than 1 depending on whether the disturbances are autocorrelated or not. However, if structural breaks are taken into account, we notice two breaks, one at 1951 and the other one at 1980, and looking at the degree of integration of the corresponding subsamples, we observe first the existence of segmented trends, along with lower orders of integration and show mean reversion in the three subsamples. Thus, our results validate those studies that argue that if non-linear components or structural breaks are present in the data and are not taken into account, the results overestimate the degree of dependence, finding, in our case, values of the differencing parameter close to or higher than 1 and showing a lack of mean reversion. However, once the breaks are detected, the hypothesis of mean reversion cannot be rejected in any of the subsamples, thus supporting the idea of transitory shocks and disappearing in the long run. This has important implications if exogenous shocks occur in the series since little actions should then be conducted since the shocks will have transitory effects disappearing by themselves in the long run. Nevertheless, though transitory, shocks may have long-lasting effects due to the long memory property and this should be taken into account by practitioners and academics.

The results presented in this work can be extended in several directions. Thus, it would be interesting to investigate if our results may be affected by the location of the study. In that respect, the use of alternative datasets may solve this question. From a methodological viewpoint, the model based on structural breaks can be substituted by others that use non-linear deterministic trends based, for example, on Chebyshev polynomials in time (Cuestas & Gil-Alana, 2016), Fourier functions (Gil-Alana & Yaya, 2021) or neural networks (Yaya *et al.*, 2021). In doing so, we avoid the abrupt changes produced by the breaks in the data. Work in these directions is now in progress.

## ACKNOWLEDGEMENTS

Luis A. Gil-Alana gratefully acknowledges financial support from the MINEIC-AEI-FEDER PID2020-113691RB-I00 project from ‘Ministerio de Economía, Industria y Competitividad’ (MINEIC), ‘Agencia Estatal de Investigación’ (AEI) Spain and ‘Fondo Europeo de Desarrollo Regional’ (FEDER). An internal Project from the Universidad Francisco de Vitoria is also acknowledged. Comments from the editor and two anonymous reviewers are gratefully acknowledged.

## DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

## CONFLICT OF INTEREST

The authors declare there is no conflict.

## REFERENCES

- Bai, J. & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18(1), 1–22.
- Beck, M. B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research* 23(8), 1393–1442.
- Bi, R., Chen, C., Tan, J., Jia, X. & Xiang, S. (2019). Two-level optimization model for water consumption based on water prices in eco-industrial parks. *Resources, Conservation and Recycling* 146, 308–315.
- Bloomfield, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* 60(2), 217–226.
- Chambers, R. (1980). *In Search of a Water Revolution: Questions for Managing Canal Irrigation in the 1980 s. Report of a Planning Workshop on Irrigation Water Management.*
- Choi, K., Yu, W. C. & Zivot, E. (2010). Long memory versus structural breaks in modeling and forecasting realized volatility. *Journal of International Money and Finance* 29, 857–875.
- Cuestas, J. C. & Gil-Alana, L. A. (2016). Testing for long memory in the presence of nonlinear deterministic trends based on Chebyshev polynomials. *Studies in Nonlinear Dynamics and Econometrics* 21(4), 57–74.
- Cui, S., Wu, M., Huang, X. & Cao, X. (2022). Unravelling resources use efficiency and its drivers for water transfer and grain production processes in pumping irrigation system. *Science of The Total Environment* 818, 151810.
- Da Cruz, N. F. & Marques, R. C. (2013). A multi-criteria model to determine the sustainability level of water services. *Water Asset Management International* 9(3), 16–20.
- Danielson, L. E. (1979). An analysis of residential demand for water using micro time-series data. *Water Resources Research* 15(4), 763–767.
- Diebold, F. X. & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics* 105(1), 131–159.
- Flörke, M., Kynast, E., Bärlund, I., Eisner, S., Wimmer, F. & Alcamo, J. (2013). Domestic and industrial water uses of the past 60 years as a mirror of socio-economic development: a global simulation study. *Global Environmental Change* 23(1), 144–156.
- García-Lopez, M. & Montano, B. (2020). Water price effects on consumption and territorial imbalances in Spain in the context of the water framework directive. *Water* 12(6), 1604.
- Gassert, F., Reig, P., Luo, T. & Maddocks, A. (2013). *A Weighted Aggregation of Spatially Distinct Hydrological Indicators.* World Resources Institute, Washington, DC.
- Gil-Alana, L. A. (2008). Fractional integration and structural breaks at unknown periods of time. *Journal of Time Series Analysis* 29(1), 163–185.
- Gil-Alana, L. A. & Robinson, P. M. (1997). Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *Journal of Econometrics* 80(2), 241–268.
- Gil-Alana, L. A. & Yaya, O. (2021). Testing fractional unit roots with nonlinear smooth break approximations using Fourier functions. *Journal of Applied Statistics* 48(13–15), 2542–2559.

- Gleick, P. H. & Heberger, M. (2014). Water and conflict. In *The World's Water* (Gleick, P., ed). Island Press, Washington, DC, pp. 159–171.
- Granger, C. W. & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis* 1(1), 15–29.
- Granger, C. W. J. & Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance* 11, 399–421.
- Hoerling, M., Quan, X. W. & Eischeid, J. (2009). Distinct causes for two principal US droughts of the 20th century. *Geophysical Research Letters* 36(19), 1–6.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* 68, 168–176.
- Jägermeyr, J., Gerten, D., Heinke, J., Schaphoff, S., Kummu, M. & Lucht, W. (2015). Water savings potentials of irrigation systems: global simulation of processes and linkages. *Hydrology and Earth System Sciences* 19, 3073–3091. doi:10.5194/hess-19-3073-2015.
- Monge, M. & Gil-Alana, L. A. (2020). Water prices: persistence, mean reversion and trends. *Water Policy* 22(6), 1200–1216.
- Reynaud, A. (2015). *Modelling Household Water Demand in Europe*. JRC Technical Reports, European Commission.
- Ritchie, H. & Roser, M. (2017). *Water Use and Stress*. Published online at OurWorldInData.org
- Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association* 89(428), 1420–1437.
- Salas, J. D., Tabios III., G. Q. & Bartolini, P. (1985). Approaches to multivariate modeling of water resources time series 1. *JAWRA Journal of the American Water Resources Association* 21(4), 683–708.
- Sheffield, J. & Wood, E. F. (2007). Characteristics of global and regional drought, 1950–2000: analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. *Journal of Geophysical Research: Atmospheres* 112(D17), 1–21.
- Sibbertsen, P. (2004). Long memory versus structural breaks: an overview. *Statistical Papers* 45(4), 465–515.
- The United Nations World Water Development Report (2020). *Water and Climate Change (2020)*. ONU.
- Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S. & Wiberg, D. (2016). Modeling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches. *Geoscientific Model Development* 9(1), 175–222.
- World Bank (2008). *World Development Report (2008): Agriculture for Development*. World Bank, Washington, DC.
- Yaya, O., Ogbonna, A. E., Furuoka, F. & Gil-Alana, L. A. (2021). A new unit root test for unemployment hysteresis based on the autoregressive neural network. *Oxford Bulletin of Economics and Statistics* 83(4), 960–981.
- Zheng, Y., Zhang, W., Xie, J. & Liu, Q. (2022). A water consumption forecasting model by using a nonlinear autoregressive network with exogenous inputs based on rough attributes. *Water* 14(3), 329.

First received 22 October 2021; accepted in revised form 1 August 2022. Available online 12 August 2022