⁶Long Memory in Average Monthly Temperatures and Precipitations in Guatemala

LUIS A. GIL-ALANA^{a,b,c,d} AND MARLON J. CASTILLO^a

^a Faculty of Economics, University of Navarra, Pamplona, Spain ^b Navarra Center for International Development, University of Navarra, Pamplona, Spain ^c Institute of Data Science and Artificial Intelligence, University of Navarra, Pamplona, Spain

^d Universidad Francisco de Vitoria, Madrid, Spain

(Manuscript received 1 February 2023, in final form 17 August 2023, accepted 24 October 2023)

ABSTRACT: In this paper, we perform a fractional integration analysis of the average monthly temperature and precipitation data in 17 departments of Guatemala. Two analyses are performed, the first with the original data and the second with the anomalies based on the period January 1994–December 1999. The results indicate that there is a significant positive time trend in temperatures in the departments of Guatemala ($0.0045^{\circ}C \mod^{-1}$), Quetzaltenango ($0.0040^{\circ}C \mod^{-1}$), Escuintla ($0.0034^{\circ}C \mod^{-1}$), and Huehuetenango ($0.0047^{\circ}C \mod^{-1}$), whereas in the case of precipitation no time trend was observed. An important relevant result is that the departments of El Progreso, Baja Verapaz, and Guatemala occupy the second, third and fourth highest levels of persistence for both temperatures and precipitation, with Sacatepéquez and Quiché displaying the first places for temperature and precipitation, respectively, thus making these five departments the ones that are most vulnerable to climate change since a shock would take a long time to disappear.

KEYWORDS: Filtering techniques; Numerical analysis/modeling; Stochastic models

1. Introduction

Climate change has been a controversial and much-studied topic in the last decade; it has been strongly addressed by governments, and policies have been created around the world to help mitigate the negative effects of this phenomenon. Climate change is defined as global variations in Earth's climate (De Lorenzo and Liaño 2017). Climate encompasses many variables, for example, humidity, temperature, rainfall, and wind, among others. Some of the most noticeable effects of climate change have been on global temperature and precipitation. According to Kumar et al. (2021), Earth's temperature increased 1°C since 1900, while precipitation has increased by 2% since the beginning of the twentieth century (Dore 2005). On the temperature side, several studies have shown a positive and statistically significant trend in global mean temperatures (Vogelsang and Franses 2005; Jenkins et al. 2022; Helbling and Meierrieks 2023). On the precipitation side, studies show that there is a global readjustment of precipitation in magnitude and timing, causing an increase in the severity and variability of rainfall that ultimately leads to a worsening of conditions in different areas, that is, wetter areas become wetter and arid areas become drier (Dore 2005; Sun et al. 2021; Dong et al. 2021).

Although it is important to analyze climate variations globally, it is also important to analyze them in a more segmented way, as this makes it possible to identify whether the region of interest is being affected to a greater extent by one or other factor, which in the end allows governments to design specialized et al. (2005) analyzed temperatures and precipitation in the region of Central America and northern South America and found that the annual maximum value of daily maximum temperatures has increased by 0.3°C decade⁻¹ during the years 1961-2003, while the annual maximum 1-day precipitation and the annual count of days when precipitation ≥ 20 mm have increased by 2.6 mm decade⁻¹ and 0.1 days decade⁻¹, respectively. These results suggest a general warming of the region and an increase in rainfall intensity. A simulation by Lyra et al. (2017) shows that by 2021-50, precipitation is projected to increase by 0.2-1 mm day⁻¹ in the western part from El Salvador to Costa Rica during the spring, while in the summer an increase of 1–4 mm day⁻¹ is expected in southeastern Nicaragua and along the eastern coast of Costa Rica and Panama. The same study projects that temperatures in the northern part of Central America (Guatemala, El Salvador, Honduras, and Nicaragua) will increase between 1.8° and 2.4°C by 2021-50, while in the southern part (Costa Rica and Panama) an increase between 1.6° and 2.0°C is expected. Some other important studies have shown a positive increase in temperatures in Europe (Gil-Alana and Sauci 2019a); an increase between 0.08° and 0.96°C in temperature trend in West Antarctica (Ludescher et al. 2016), and an increase in precipitation in the west and a decrease in the east in China between 1960 and 2012 (Wu et al. 2016). Deng et al. (2022) supports the results of Ludescher et al. (2016) indicating that a study conducted with data from 1990 to 2019 at 698 meteorological stations shows that daytime and nocturnal precipitation have increased in the western parts of China. In addition to analyzing changes in climate, it is also important to know the consequences of the changes. Shi et al. (2015) found that a 1°C increase in average summer temperatures in New England is associated with a 1% increase in mortality. Liu et al. (2020) indicates that higher interannual precipitation variability tends to reduce vegetation productivity, and Desbureaux and Rodella (2019) found that

policies to address the most worrying climate changes. Aguilar

© 2024 American Meteorological Society. This published article is licensed under the terms of the default AMS reuse license. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

^o Denotes content that is immediately available upon publication as open access.

Corresponding author: Luis A. Gil-Alana, alana@unav.es

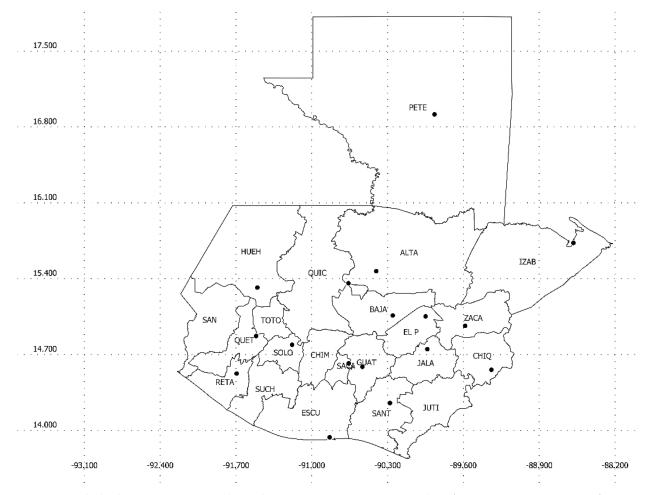


FIG. 1. Distribution of the 17 meteorological stations throughout the Guatemalan territory (based on the EPSG:4326 system).

droughts in cities can cause a reduction of up to 6.4% in the labor income of casual employees.

Several methods have been used to analyze climate trends, and one of the most used methods has been fractional integration due to its characteristic of being able to deal with stationary and nonstationary series since many of the climatic variables are nonstationary. According to Baillie (1996), long memory models have been of special interest in recent years in the analysis of climate change trends since they allow us to determine whether these changes are simply natural cycles (e.g., if the series are stationary with no deterministic trends) or whether they have been caused by human intervention (in case of nonstationarity and/or deterministic trends). Ventosa-Santaulària et al. (2014) and Gil-Alana (2005, 2012) have found that some climatological variables, such as sea level and temperatures, demonstrate characteristics of a long memory process. Aguilar et al. (2005) and Gil-Alana (2012) are the closest works to the one in this paper, to the best of our knowledge, a detailed analysis of climate in Guatemala using fractional integration has not yet been carried out and this would be the first study. In addition to this, it is known that Guatemala is an agricultural country. According to data from the Bank of Guatemala in 2020 the agricultural sector

represents 13.3% of the country's GDP. For this reason, it is important to conduct a comprehensive analysis of climate trends in order to create contingency plans for the future. This paper analyzes trends in average monthly temperatures and precipitation in 17 of the 22 departments of Guatemala (Fig. 1) between 1994 and 2021 using a fractional integration approach.

Thus, the main objectives of this paper are the following: First, we want to investigate if the two variables (temperatures and precipitation) display the property of long memory, and we do so by using fractional integration methods. Second, based on the estimated order of integration for each series, we check if stationarity holds. Third, we examine if time trends are present in the data, and we do that by taking into account the potential long memory feature previously tested. Note that the estimation of the time trend coefficients will be clearly affected by the assumption made in relation with the error term, the result being clearly biased (and inconsistent) if the series display long memory (or if it presents an order of integration above 0), which has not been taken into account in the analysis.

The structure of the present article is as follows. Section 2 details the model used in the study, section 3 describes the

2. The model

The classic model for determining the linear trend of a time series is as follows:

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

where y_t represents in this case the monthly average temperature or precipitation, x_t is the error term, β_0 is an unknown coefficient representing the intercept (a constant), and β_1 is the time trend; therefore, if β_1 is positive there is an upward trend in the series, but if it is negative there is a downward trend.

To correctly determine β_1 , it is necessary to define x_t , generally assumed to be stationary; for example, x_t has a normal distribution with zero mean and constant variance $N(0, \sigma^2)$. This definition allows the coefficients to be determined using simple methods such as ordinary least squares (OLS) or generalized least squares (GLS); however, several studies have shown that some climatological variables may not be stationary or may display at least a long memory pattern (Gil-Alana 2012; Ventosa-Santaulària et al. 2014; Gil-Alana and Sauci 2019b). Because of this, we use in this paper a long memory model that uses fractional integration and that is characterized by the following model:

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

where *d* is a value belonging to the set of real numbers and representing the number of differences required to obtain I(0), *L* is the lag operator, and u_t is a short memory process I(0). Note that the polynomial on the left-hand side in Eq. (2) can be expressed as

$$(1-L)^{d} = \sum_{j=0}^{\infty} {\binom{d}{j}} (-1)^{j} L^{j} = 1 - dL + \frac{d(d-1)}{2} L^{2} - \cdots,$$
(3)

implying that

$$(1-L)^{d}x_{t} = x_{t} - dx_{t-1} + \frac{d(d-1)}{2}x_{t-2} - \cdots$$
 (4)

From this definition it can be concluded that if *d* is an integer value, then x_t depends on a finite number of its past observations, but if *d* is not an integer value, then x_t will depend on an infinite number of its past observations, the higher the value of *d*, the higher the level of data dependence, that is, the higher the level of long memory (Gil-Alana 2012). Furthermore, in the case of temperatures, several studies have shown that the values of *d* range between 0 and 1 (Gil-Alana 2012). All processes are mean reverting if d < 1, are nonstationary if $0.5 \le d < 1$, and are stationary if 0 < d < 0.5; for example, a process with d = 0.8 is nonstationary mean reverting (Belbute and Pereira 2017). Baillie (1996) defines I(0) as a stationary and ergodic process with a bounded and positively valued spectrum at all frequencies, within this specification there are models such as stationary autoregressions (AR), moving averages (MA), and

stationary ARMA models. Since in this case we are dealing with climatological monthly data, and these are seasonally unadjusted, the simplest and best fitting definition for this type of data is a monthly AR(1) process defined as follows:

$$u_t = \rho u_{t-12} + \varepsilon_t, \tag{5}$$

where ρ is the seasonality indicator and ε_t is a white noise process. The model used in this study is composed of Eqs. (1), (2), and (5), and with these three equations it will be possible to determine the time trend β_1 , the level of persistence *d*, and seasonality ρ of the data, respectively.

The estimation is conducted by using a frequency domain version of the Whittle function and is implemented by means of a simple version of the tests of Robinson (1994), widely used in empirical applications in many different fields. This method is based on the Lagrange multiplier (LM) principle and tests the null hypothesis:

$$H_0: d = d_0 \tag{6}$$

in a model given by Eqs. (1), (2), and (5) for any real value d_0 , including stationary ($d_0 < 0.5$) and nonstationary ($d_0 \ge 0.5$) cases. This is one of the advantages of this approach since it does not require preliminary differentiation in the case of nonstationary data. In addition, the limit distribution is standard normal, independently of the inclusion of deterministic terms like those in Eq. (1), and it is the most efficient method in the Pitman sense against local departures from the null [see Gil-Alana and Robinson (1997) for the functional form of the test statistic employed in this work].

3. The data

The data used in this study were provided by the Instituto Nacional de Sismología, Vulcanología, Meteorología e Hidrología (INSIVUMEH) of Guatemala. The data provided are daily average temperatures (°C) and daily precipitation (mm) of 21 of the 22 departments into which the country is divided. The algorithm for the estimation of the results was programmed in Fortran, but there was a particular limitation: each series could not have more than 3000 or 3500 data entries because of the computational load this represented. Since there were more than 10000 data entries per series, it was decided to calculate the monthly average and use these data instead of the daily value.¹ The range of years varies between departments, and it was found that much of the older data were inconsistent or there were long periods of years without data. For this reason, we decided to only take into account the data from January 1994 (1994m1) through December 2021 (2021m12) because from 1994 onward the data were more consistent and reliable. After establishing the study period, we proceeded to convert the daily frequency series to a monthly frequency. This was done by simply averaging the daily temperatures of each month. Another

¹ The versions of the Fortran codes employed in this work in Python and R are being developed at present, which will allow us to work with longer time series data.

TABLE 1. Estimated values for temperatures. The material in parentheses gives the 95% confidence bands in column 2 and the *t* values of the estimated coefficients for the intercept and the time trend in columns 3 and 4, respectively. The last column reports the seasonal coefficient.

Series	d value	Intercept (t value)	Time trend (t value)	Seasonality
Alta Verapaz	0.27 (0.18, 0.39)	19.4943 (54.72)	_	0.785
Baja Verapaz	0.48 (0.39, 0.62)	21.6137 (28.74)	_	0.405
Chiquimula	0.24 (0.16, 0.36)	22.4374 (79.27)	—	0.729
El Progreso	0.66 (0.56, 0.79)	14.0248 (10.72)	—	0.516
Escuintla	0.38 (0.30, 0.49)	27.5772 (86.57)	_	0.772
Guatemala	0.43 (0.34, 0.56)	19.5569 (45.02)	—	0.611
Huehuetenango	0.35 (0.25, 0.48)	18.2302 (42.40)	_	0.723
Izabal	0.25 (0.16, 0.36)	26.6270 (80.31)	_	0.866
Jalapa	0.28 (0.19, 0.40)	17.0409 (52.19)	_	0.750
Petén	0.32(0.22, 0.46)	26.5885 (51.27)	_	0.819
Quetzaltenango	0.38 (0.29, 0.50)	14.4974 (32.60)	0.0041 (1.72)	0.719
Quiché	0.32(0.23, 0.45)	25.4385 (58.69)	_	0.708
Retalhuleu	0.27 (0.16, 0.40)	27.4543 (116.48)	_	0.321
Sacatepéquez	0.74 (0.63, 0.89)	15.5955 (15.30)	_	0.399
Santa Rosa	0.33 (0.23, 0.44)	24.6198 (111.37)	_	0.652
Sololá	0.34 (0.24, 0.47)	14.6478 (53.15)		0.615
Zacapa	0.36 (0.27, 0.48)	28.2216 (53.75)	_	0.764

important rule for selecting the data was that after obtaining the series for monthly frequency, it was decided that these should not have more than six consecutive missing data entries, otherwise the series was discarded from the analysis. After applying these rules, we obtain a database of average monthly temperatures and precipitations from 1994 to 2021 for 17 of the 22 departments in Guatemala (each series with 336 values; see Fig. 1).

4. The empirical results

Although not reported, we started this section by conducting standard stationary/nonstationary (unit root) methods. In particular, we performed Dickey and Fuller (1979; ADF), Phillips and Perron (1988; PP), Kwiatkowski et al. (1992; KPSS), and Elliott et al. (1996; ERS) tests in the two variables, temperatures and precipitation. The results generally support the integration of order 1 [I(1)] hypothesis for the temperatures and the I(0) one for precipitation. These results, however, should be taken with caution, noting that the abovementioned methods have very low power if the data-generating process is fractionally integrated (see, e.g., Diebold and Rudebush 1991; Hassler and Wolters 1994; Lee and Schmidt 1996). Because of this, we focus on the I(d) models described in section 2.

As earlier mentioned, we use a version of the tests of Robinson (1994) based on the Whittle function in the frequency domain. For this study two types of analysis were performed, the first analysis was using the data without any modification and the second was using the anomalies. The results of the first analysis are shown across Tables 1 and 2, and the results on the anomalies [with base period from January 1994 through December 1999 (1994m1–1999m12)] are reported across Tables 3 and 4. The tables are described in more detail in the rest of section 4.

TABLE 2. As in Table 1, but for precipitation.

Series	d value	Intercept (t value)	Time trend (t value)	Seasonality
Alta Verapaz	0.18 (0.07, 0.31)	6.2022 (10.89)	_	0.270
Baja Verapaz	0.22 (0.10, 0.38)	2.6688 (5.28)	_	0.521
Chiquimula	-0.08(-0.18, 0.05)	4.5892 (26.43)	_	0.762
El Progreso	0.26 (0.14, 0.41)	4.9766 (5.07)	_	0.580
Escuintla	0.15 (0.02, 0.30)	4.6093 (7.27)	_	0.470
Guatemala	0.20 (0.08, 0.36)	3.3647 (5.92)	_	0.624
Huehuetenango	0.13 (0.02, 0.26)	2.9293 (8.61)	_	0.646
Izabal	0.07(-0.04, 0.20)	9.1487 (21.32)	_	0.195
Jalapa	0.17 (0.06, 0.30)	3.3994 (7.18)	_	0.581
Petén	0.13 (0.04, 0.26)	5.0473 (11.89)	_	0.507
Quetzaltenango	0.10 (0.00, 0.24)	2.4505 (10.61)	_	0.674
Quiché	0.28 (0.18, 0.42)	4.1576 (4.82)	_	0.343
Retalhuleu	0.02(-0.08, 0.15)	8.5428 (19.37)	_	0.739
Sacatepéquez	0.18 (0.07, 0.33)	3.2156 (6.74)	_	0.645
Santa Rosa	0.08(-0.02, 0.23)	4.9856 (11.70)		0.700
Sololá	0.09(-0.01, 0.23)	3.8444 (10.79)	_	0.683
Zacapa	0.17 (0.07, 0.29)	2.2001 (6.05)	_	0.486

Series	d value	Intercept (t value)	Time trend (<i>t</i> value)	Seasonality
Alta Verapaz	0.27 (0.18, 0.39)		_	0.783
Baja Verapaz	0.48 (0.39, 0.62)			0.404
Chiquimula	0.24 (0.16, 0.36)	_	—	0.729
El Progreso	0.67 (0.56, 0.79)	-2.232(-1.68)		0.514
Escuintla	0.39 (0.31, 0.51)	-0.2584(-1.93)	0.0034 (1.78)	0.772
Guatemala	0.43 (0.34, 0.56)	-0.5797(-2.09)	0.0045 (1.72)	0.611
Huehuetenango	0.35 (0.25, 0.49)	-0.2576(-2.44)	0.0047 (1.68)	0.723
Izabal	0.25 (0.16, 0.37)			0.866
Jalapa	0.28 (0.18, 0.40)	_	—	0.750
Petén	0.32 (0.22, 0.46)	—	—	0.819
Quetzaltenango	0.38 (0.29, 0.50)	-0.4520(-2.01)	0.0040 (1.72)	0.719
Quiché	0.32 (0.23, 0.45)			0.709
Retalhuleu	0.27 (0.16, 0.40)	_	—	0.321
Sacatepéquez	0.74 (0.63, 0.89)	-1.6536 (-1.72)	—	0.271
Santa Rosa	0.33 (0.23, 0.44)		_	0.653
Sololá	0.34 (0.24, 0.47)	—	—	0.616
Zacapa	0.36 (0.27, 0.48)	_	_	0.764

TABLE 3. As in Table 1, but for temperature anomalies.

a. Results based on the original data

Table 1 shows the estimation results for the case of temperatures. We started with a model containing a linear time trend, that is, using Eq. (1) and looking at their corresponding *p* values. If both coefficients were statistically significant then the model with both, intercept and a time trend was selected; if only the first coefficient β_0 was significant then the model with an intercept is selected, and if both coefficients were insignificant then the model does not contain any deterministic term. Note that the model given by Eqs. (1) and (2) can be expressed in a single equation as

$$\tilde{y}_t = \beta_0 \tilde{1}_t + \beta_1 \tilde{t}_t + u_t, \quad t = 1, 2, \dots,$$
 (7)

where

$$\tilde{y}_t = (1-L)^d y_t, \quad \tilde{1}_t = (1-L)^d 1, \quad \tilde{t}_t = (1-L)^d t,$$

and, since u_t is I(0) by construction, standard t values hold in Eq. (7).

Starting with the *d* values of the selected model for each series, it can be observed that the values are statistically significantly positive in all cases with a 95% confidence level. The maximum value is 0.74 and belongs to Sacatepéquez, while the minimum value is 0.24 and belongs to Chiquimula. Something important to take into account in these results is that the null hypothesis d = 0 can be rejected, that is, the hypothesis of a short memory behavior I(0) in the temperatures is rejected in all cases. Since for all the intervals of the *d* values, the values are between 0 and 1, the results suggest that there is a long memory behavior in the monthly mean temperatures of all the departments in Guatemala, that is, there is a high dependence on past observations. In addition to this, it is observed that in Baja Verapaz and Guatemala the interval includes values less and greater than 0.5, this indicates that nothing can be concluded

TABLE 4. As in Table 1, but for precipitation an	TABLE 4. As	n Table 1	1, but i	for precip	pitation	anomalies.
--	-------------	-----------	----------	------------	----------	------------

Series	d value	Intercept (t value)	Time trend (t value)	Seasonality
Alta Verapaz	0.18 (0.07, 0.31)	_	_	0.270
Baja Verapaz	0.23 (0.10, 0.38)	_	_	0.519
Chiquimula	-0.08(-0.18, 0.05)	_	_	0.762
El Progreso	0.26 (0.14, 0.41)	_	_	0.579
Escuintla	0.15 (0.03, 0.31)	_	_	0.471
Guatemala	0.20 (0.08, 0.36)	_	_	0.625
Huehuetenango	0.13 (0.02, 0.26)	_		0.647
Izabal	0.07(-0.04, 0.20)	_	_	0.195
Jalapa	0.17 (0.07, 0.31)	_		0.581
Petén	0.13 (0.04, 0.26)	0.9053 (0.506)	_	0.506
Quetzaltenango	0.10 (0.00, 0.24)			0.674
Quiché	0.28 (0.18, 0.42)	_	_	0.343
Retalhuleu	0.02(-0.07, 0.14)	_	_	0.737
Sacatepéquez	0.18 (0.07, 0.33)	_	_	0.645
Santa Rosa	0.08(-0.02, 0.22)	_		0.700
Sololá	0.09(-0.02, 0.23)		_	0.683
Zacapa	0.17 (0.07, 0.29)	—	—	0.486

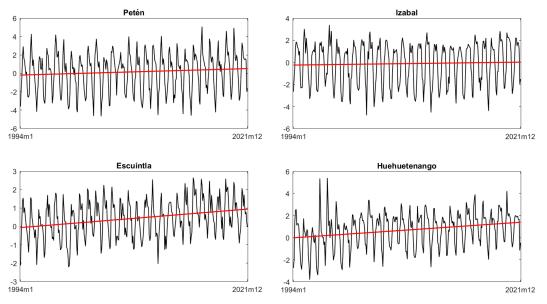


FIG. 2. Time series of monthly average temperature anomalies.

regarding whether it is stationary or not since values at both extremes are included, that is, these departments present a long memory behavior, but it cannot be concluded whether they are stationary or nonstationary. For the departments of El Progreso and Sacatepéquez, however, it is observed that they present a nonstationary behavior ($d \ge 0.5$). For the rest of the departments, it can be said that they are long memory processes with stationary characteristics because the intervals are between 0 and 0.5. In addition, column 5 shows that the seasonality coefficient is positive in all cases, with a minimum of 0.321 in Retalhuleu and a maximum of 0.866 in Izabal, these results suggest a seasonal monthly effect in the temperature data. On the other hand, only Quetzaltenango shows a positive linear trend with a coefficient of 0.0041, this indicates an increase of 0.0041° C month⁻¹ (or 0.0492° C yr⁻¹). For the rest of the departments there is no evidence of a time trend. According to the census conducted in 2018 by the Instituto Nacional de Estadística (INE) of Guatemala, Quetzaltenango is the department with the second highest percentage of urban population (Guatemala being the first and Escuintla the third), and, in addition, it is known that in recent years it has had strong economic growth. It is possible to argue that the observed trend might be attributed to this strong population and economic growth. (https://www.censopoblacion.gt/graficas).

In the case of precipitations, the results are shown in Table 2. For the departments of Chiquimula, Izabal, Retalhuleu, Santa Rosa, and Sololá, the short memory hypothesis d = 0 cannot be rejected since the intervals range from a negative to a positive value. For all the cases the selected model is the one with an intercept and no time trend, and based on this model, the minimum value of the differencing parameter is -0.08, which corresponds to Chiquimula (in which the intervals suggest a short memory behavior), while the maximum is found in Quiché with a value of d of about 0.28. Something important to note in this table is that a time trend is not observed in all cases. In addition, it is observed that the upper value of the intervals for all d values does not

exceed 0.5, which allows the null hypothesis of nonstationary $(d \ge 0.5)$, to be rejected, that is, it allows the null hypothesis that the series present a nonstationary behavior for monthly average precipitations to be rejected. All series present a stationary long memory process behavior (except for the departments where the null hypothesis d = 0 cannot be rejected). The seasonal coefficient ranges from 0.195 to 0.762, again suggesting a seasonal monthly effect in the data.

b. Results based on the anomalies

Some studies suggest that it is better to use anomalies instead of unmodified data because this allows us to observe in a better way if there is a time trend (Gil-Alana 2012; Gil-Alana and Sauci 2019b); for this reason, the previous analysis was performed again but this time using temperature and precipitation anomalies. In this context, the period 1994m1–1999m12 was taken as the base. This means that from our monthly frequency series we take all the monthly temperatures (or precipitation) values of these years and calculate the average value. Then, to calculate the anomalies of each of the series we use their average value, calculated from the base period and subtract it from each of the values of the series in monthly frequency. In this way we can obtain how much each monthly value deviates from the average value (anomalies) calculated between the period 1994m1– 1999m12. Therefore, the equation for the anomalies is

anomaly_{*s,t*} =
$$V_{s,t} - V_{avg,s}$$
 $s = 1, 2, ..., 34$ and
 $t = 1, 2, ...,$ (8)

where *s* represents each of the 34 temperature and precipitation series under the study, *t* is the time in months, *V* represents the value of the series *s* at time *t*, and V_{avg} is the average value of the series calculated in the base period. The model, amount of data and estimation process remain the same, the only thing that changes in this case is the data. Figure 2 shows the graphs of the monthly average temperature anomalies for four departments; these departments were selected to show the comparison of series with and without a time trend. The rest of the figures for each series of original and anomaly data are shown in the appendix. At first glance, a clear upward trend can be seen in Escuintla and Huehuetenango, whereas in Petén and Izabal, this trend is not very noticeable.²

Table 3 shows the estimates of the *d* values; in most cases, the selected model is the no terms model, except in El Progreso and Sacatépequez where the model with intercept was selected and in Escuintla, Guatemala, Huehuetenango, and Quetzaltenango where the model with a time trend was selected. This table shows the estimates of each of the coefficients according to the selected model, and something important to mention is that the *d* values are similar, or the same for most cases, as those in Table 1, but with the notable difference that in this case a time trend is not only observed in Quetzaltenango but also in Escuintla, Guatemala, and Huehuetenango.

The value of the time trend coefficient of these four departments is positively significant, Huehuetenango being the maximum with an increase of 0.0047°C month⁻¹ (or 0.0564°C yr⁻¹) and Escuintla the minimum with 0.0034° C month⁻¹ (or 0.0408° C yr⁻¹). As mentioned above, in the three departments with the largest urban population (Guatemala, Quetzaltenango, and Escuintla) a positive trend is observed; this urban growth could explain this trend. On the other hand, Huehuetenango is the sixth department with the largest urban population, but something important to mention is that the meteorological station of this department is located in the municipality of Huehuetenango, which is the municipality with the largest urban population in the entire department, again, the upward trend could be explained by this increase in the urban population rather than by climate change at the regional level. Again, it is observed that all series present monthly seasonality and characteristics of a long memory process; for Baja Verapaz, Escuintla, and Guatemala nothing can be concluded about stationarity; El Progreso and Sacatepéquez are nonstationary, while the rest of the series are stationary.

For the case of precipitation, the results are displayed in Table 4, and unlike the results with the original data, in this case the model selected for all cases is the no terms model (except for Petén where the model with the intercept was selected). We can observe that we have the same conclusions about seasonality, stationarity, and long memory as in Table 2, since there are no significant differences in the intervals and d values. These data confirm the results of Table 2, where we can clearly conclude that there is no time trend in the average monthly precipitation for the departments of Guatemala. If we order the d values from highest to lowest for Tables 3 and 4 in order to determine the departments with the highest levels of persistence, we will observe that in both scenarios the departments of El Progreso, Baja Verapaz, and Guatemala occupy the second, third, and fourth place, respectively, that is, these three departments present the highest levels of

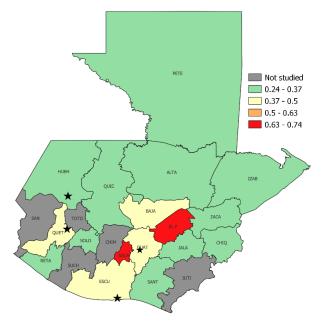


FIG. 3. Estimations of d value and time trend (the star symbol indicates departments with a time trend) for temperature anomalies, based on results of Table 3.

persistence in average monthly temperatures and precipitation, along with Sacatepéquez for temperatures and Quiché for precipitation. Therefore, of the 17 departments analyzed in this study, the five departments mentioned above would be the most affected by a shock in temperatures and precipitation, since it would take a long time to disappear.

Figure 3 shows the distribution of d values for temperatures throughout the Guatemalan territory. We observe from this figure that the great majority of departments have a low persistence (long memory) value, ranging between 0.24 and 0.37; four departments have a moderate persistence and two have a high persistence. This means that in the departments of El Progreso and Sacatepéquez, temperature shocks will take longer to disappear than in the rest of the departments. In addition, the departments with a time trend are marked in the figure with a black star. It is observed that in three of the four departments with a moderate persistence there is a significant time trend. Figure 4 shows the seasonality values for the temperatures. We see that in most of the departments this value is greater than 0.5, which indicates that there is a marked repetition of patterns or behavior at regular intervals in these departments.

Figure 5 shows the *d* values for precipitation. It can be seen that there is no dominant range since the departments are evenly distributed among the different ranges. Another point to consider is that although all the departments have a low persistence level (less than 0.28), there are some that have a much lower persistence. The same happens in the case of seasonality, in Fig. 6; the departments are evenly distributed among the different ranges, although it is observed that there is a slightly greater number of departments with a high seasonality in precipitation.

² Note that these trends were obtained in a model that incorporates both the long memory feature and the seasonality issue.

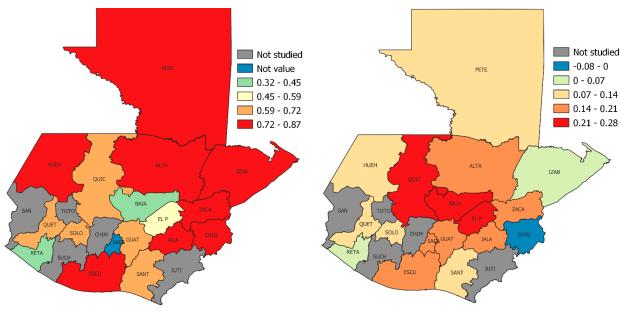


FIG. 4. Estimations of seasonality for temperature anomalies, based on results of Table 3.

FIG. 5. Estimations of *d* value for precipitation anomalies, based on results of Table 4.

5. Discussion

The dry corridor in Guatemala is an area made up of 11 departments of the country; this is an area strongly affected by very prolonged droughts, in addition to being an area with low food security and extreme poverty. Not all the territory of each of the departments is part of the dry corridor, but El Progreso is a special case in which a large part of its territory is within this area. This is especially worrisome since, as shown in Fig. 3, this is one of the departments with a very high temperature persistence, which means that a positive temperature shock will worsen the prolonged droughts in this territory and therefore, affecting issues such as food security and extreme poverty. In addition, Fig. 5 shows that persistence in precipitation is one of the highest in comparison with the rest of the departments, but note also that this persistence degree is only 0.26, which, at a general level, is a low level of persistence (persistence is more intense when it is close to 1 and lower when it is close to 0). In conclusion, this department will be strongly affected by changes in temperatures, but slightly by changes in precipitation. Then it is also observed that three of the four departments with a moderate persistence in temperature also present a positive trend; this gives indications not only that with time the temperatures in these places will increase, but also that they can also be affected if there is a shock in temperatures, which in the worst case, could intensify the increase in temperatures in these departments. It was previously mentioned that urbanization could be one of the possible causes of the time trend observed in the temperatures in Guatemala, Quetzaltenango, Escuintla and Huehuetenango, but in addition, another possible reason for the temperature increase in Huehuetenango is that this occupies the fourth place of the departments with the highest forest cover; this is of 250734 ha, but according to a report of the Instituto Nacional de Bosques (INAB), from 2016 to 2020, this has lost about 3178 ha. This loss

of forest cover may be one of the possible causes of the increase in temperatures in this department. The same occurs in Escuintla, which has lost about 3691 ha from 2016 to 2020.

In the case of precipitation, we see that no time trend was observed in any department, in addition to the fact that the highest degree of persistence in precipitation is 0.28, so that at a general level throughout the territory, the persistence of precipitation is low and therefore the shocks will not

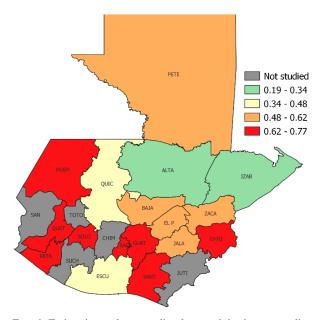


FIG. 6. Estimations of seasonality for precipitation anomalies, based on results of Table 4.

represent a great impact as in the case of temperatures. Although in this case a time trend was not observed, in future studies it would be worthwhile to analyze whether the distribution of precipitation has changed, as it has been shown in studies such as Dong et al. (2021), which indicates that a readjustment in precipitation has been observed globally. This readjustment in precipitation may have an impact especially on crop planning since it is possible that rainy seasons will be more intense and shorter while dry seasons will be longer.

As mentioned in section 1, Guatemala is an agricultural country, representing 13.3% of the country's GDP, which is why it is so important to know how vulnerable the country is to changes in temperature and precipitation. On the other hand, the seasonality values indicate how marked the cycles in temperature (or precipitation) are; a value close to 1 indicates more marked cycles while a value close to 0 indicates the opposite. To know how marked the cycles in climatic variables are can help us to better plan agricultural activities, water resource management, planting and harvesting seasons. In Fig. 4, we observe the seasonality values distributed in the Guatemalan territory. For the case of temperatures, we see that in most departments this seasonality is greater than 0.59, which is a good indication for the country since this indicates that the patterns in temperatures are well marked. The same happens in the case of precipitation, (see Fig. 6) where most of the departments have seasonality values higher than 0.48. According to a report from the Economic and Commercial Office of the Spanish Embassy in Guatemala, the three most exported products of the country are cardamom, bananas, and coffee, representing 10%, 7%, and 5%, respectively, of the country's total agricultural production. According to the Ministerio de Agricultura, Ganadería y Alimentación, 68% of the total production of cardamom is concentrated in Alta Verapaz and 14% in Quiché, and we see that the seasonality of these two departments in temperatures is high, while in precipitation is low. In addition, the persistence in temperatures is low whereas that in precipitation is high (relative to the rest of the departments). Therefore, we see that the production of this crop can be strongly affected by changes in rainfall (low seasonality and high persistence) but very little affected by changes in temperatures (high seasonality and low persistence). This gives an indication that contingency plans should be created to counteract changes in rainfall such as artificial irrigation systems, efficient drainage systems, more resistant varieties against droughts or high humidity. We see that the seasonality of temperatures is the lowest in Retalhuleu and Baja Verapaz, which indicates less marked temperature cycles and that it may affect the production of sesame and tomato. Retalhuleu yields 60% of total sesame production and Baja Verapaz generates 20% of tomato production. In the case of precipitation, the lowest seasonality is in Alta Verapaz and Izabal, endangering cocoa and banana crops. Alta Verapaz generates 31% of total cocoa production in the country and Izabal yields 33% of banana production. This shows that the sector most affected by changes in temperature and precipitation is agricultural production; persistence can affect crops but also seasonality, since it is necessary to resort to mitigation systems against changing cycles in temperature and precipitation.

6. Conclusions

In this paper, an analysis of monthly average temperatures and precipitation during the period 1994–2021 in 17 of the 22 departments of Guatemala was carried out. A fractional integration approach was used for this analysis due to its versatility when working with stationary and nonstationary series. The model proposed in this study allowed us to know the time trend, persistence, and seasonality of each series. Two analyses were carried out, the first one with the original data without modifications and the second one with the anomalies, in the latter case using as a base period 1994m1–1999m12.

For the first analysis, it was found that temperatures showed a long memory process behavior in every department of Guatemala, with Sacatepéquez being the department with the highest persistence level with an estimated value of d of 0.74 and Chiquimula the lowest with 0.24. Regarding stationarity, the results are diverse, with El Progreso and Sacatepéquez being nonstationary, Baja Verapaz and Guatemala being inconclusive, and the rest of the departments being stationary. The most relevant result in this case is that a time trend was only observed in Quetzaltenango, which was 0.0041°C month⁻¹. Since this time trend was not observed in the rest of the departments, a possible explanation is that this trend is mainly due to an increase in the urban population rather than to a climatic change per se in the region, since Quetzaltenango is the department with the second largest urban population in the country. In the case of precipitation, for the departments of Chiquimula, Izabal, Retalhuleu, Santa Rosa, and Sololá, the null hypothesis of short memory behavior cannot be rejected. The highest level of persistence corresponds to Quiché with an estimated d equal to 0.28 and the lowest to Chiquimula with d = -0.08. For all precipitation series, no time trend was observed.

For the second analysis, the anomalies were used, which allowed us to observe that not only was there a time trend in Quetzaltenango, but also in Guatemala, Escuintla, and Huehuetenango. The maximum trend was 0.0047°C month⁻¹ for Huehuetenango and the minimum was 0.0034°C month⁻¹ for Escuintla. These results agree with other studies in which using the anomalies instead of the original data was preferred, since the former allows the time trends that could not be observed with the latter to be visualized. Again, a possible explanation for this trend is due to the high percentage of urban population, since Guatemala, Quetzaltenango, and Escuintla occupy the first, second and third place among the departments with the highest urban population. In the case of Huehuetenango, the meteorological station is located in the municipality of the same name, which is the municipality with the largest urban population in the department. Since there is no time trend in temperatures in most of the Guatemalan territory, it is possible that the trend observed only in some areas is explained mainly by the high percentage of urban population. Another possible explanation is deforestation as Huehuetenango and Escuintla have lost more than 3000 ha of their forest cover from 2016 to 2020. In the case of precipitation, it is again observed that there is no trend in any of the series. This result supports that of the first analysis, that is, that there is no time trend in the average monthly precipitation in the Guatemalan territory.

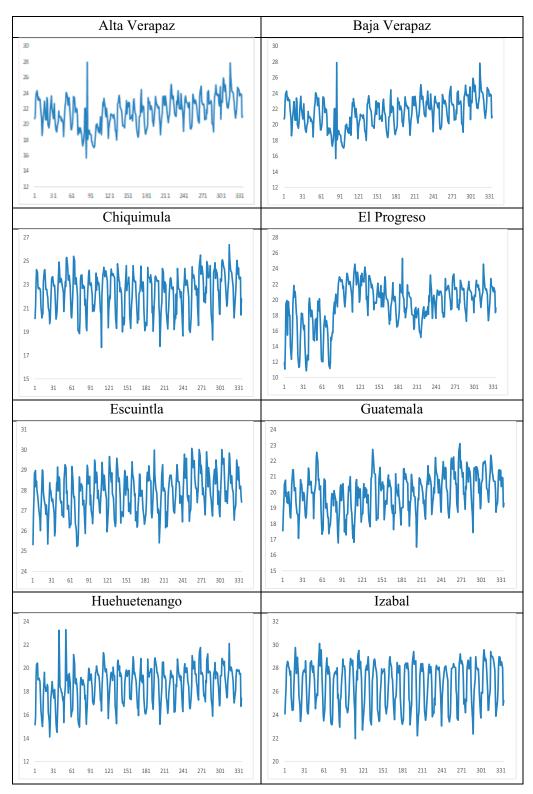


FIG. A1. Plots of each original data series: temperatures.

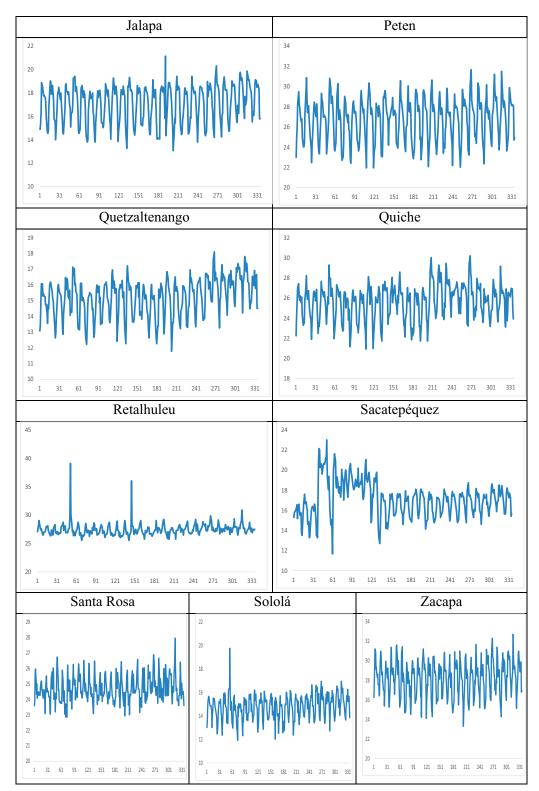


FIG. A1. (Continued)

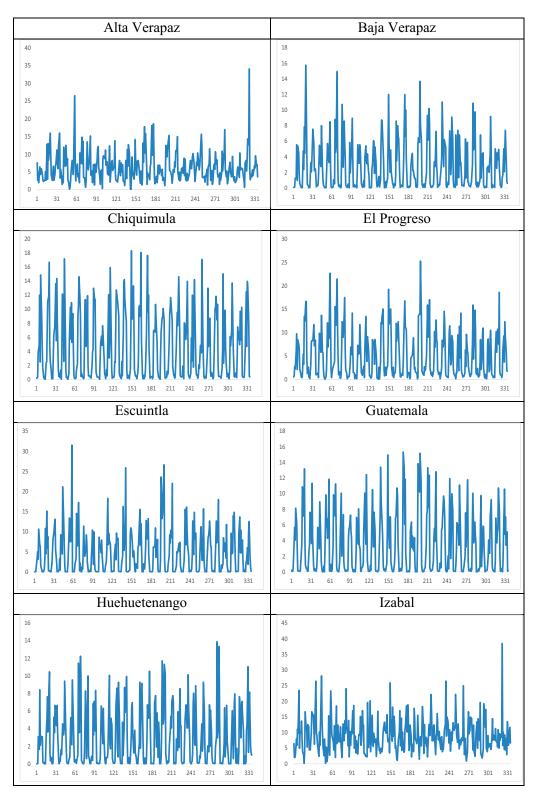


FIG. A2. Plots of each original data series: precipitation.

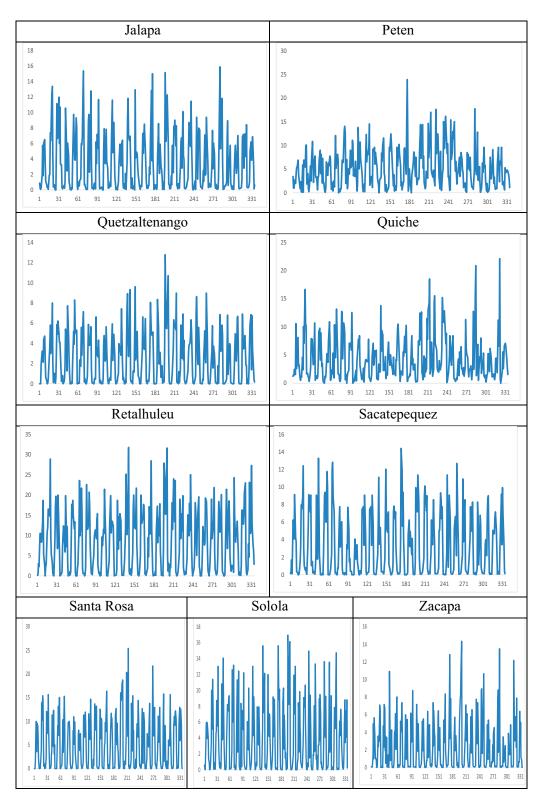


FIG. A2. (Continued)

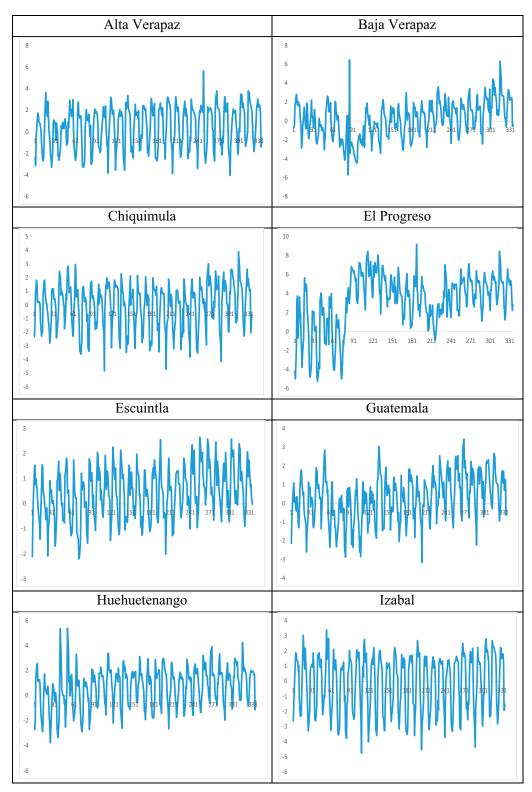


FIG. A3. Plot of each anomaly data series: temperatures.

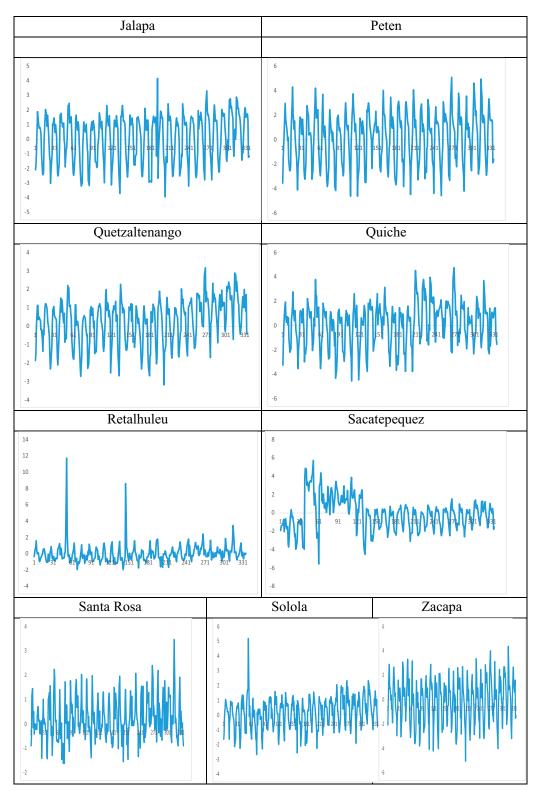


FIG. A3. (Continued)

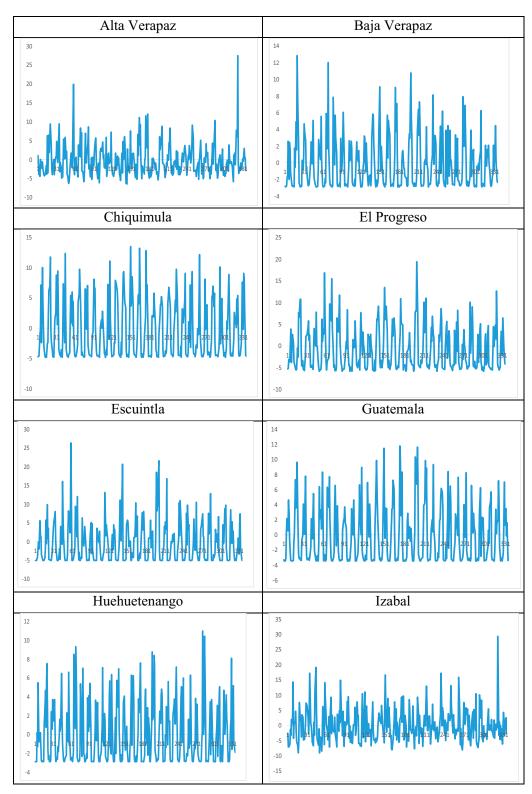


FIG. A4. Plots of each anomaly data series: precipitation.

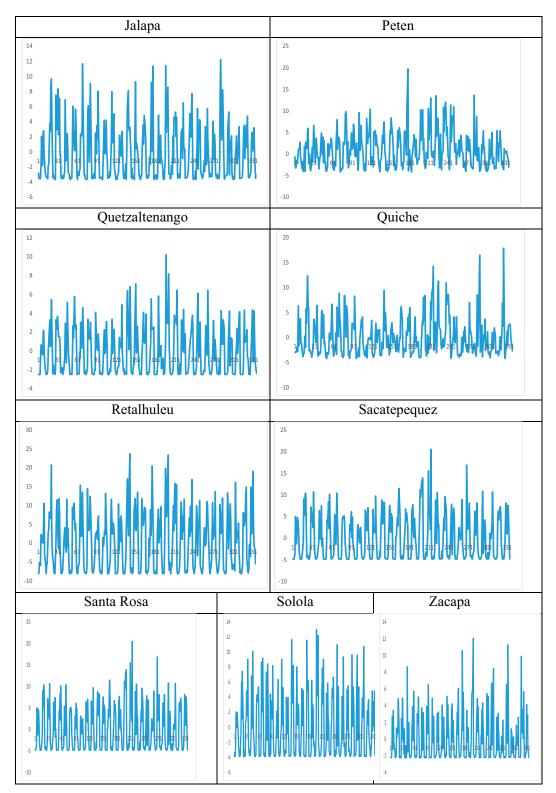


FIG. A4. (Continued)

When a series presents a high level of persistence, it means that, if it experiences a shock, it will take a long time to disappear. Further, as mentioned above, Guatemala is an agricultural country, and it is also known that temperature and precipitation shocks can significantly affect crops. Moreover, the results obtained show that El Progreso, Baja Verapaz, and Guatemala occupy the second, third, and fourth place among the departments with the highest levels of persistence in average monthly temperatures and precipitation. Therefore, it can be said that these three departments are the most vulnerable areas of the country to the presence of climate shocks, as they can greatly affect crops. The most important thing to take into account is that the meteorological stations are located in one municipality in each department. Of these three departments, two of them (San Jerónimo and San Agustín Acasaguastlán) are part of Guatemala's dry corridor. The dry corridor is an area made up of 82 municipalities that is affected by high levels of drought. Therefore, these two municipalities could be significantly affected if there is a shock that increases temperatures. It is important that the Guatemalan government is capable of managing contingency plans to mitigate the damage that a shock could cause not only in these municipalities but also in the departments with the highest levels of persistence.

Unfortunately, most of the data provided by the Guatemalan government entities are not completely reliable, they lack a standard format, there are too many missing data, there are data that do not match the rest, and so on. Therefore, it is important to mention that it is recommended that this study be conducted again with a different database (preferably of international origin) and over a longer period of years so as to be able to compare the results. For future studies, the use of satellite data could be of great help in performing a more detailed segmentation; in this way an analysis could be performed by municipality instead of by department. This could help to better visualize which municipalities are more vulnerable to shocks; that is, it could help to determine the municipalities most vulnerable to climate change.

This article can be extended in several directions. For example, alternative datasets can be employed, such as that proposed by the Climate Change Knowledge Portal belonging to the World Bank (2023). Climatic Research Unit (CRU) gridded Time Series (CRU TS) are used and are created from observational data, providing quality-controlled temperature and precipitation measurements from hundreds of weather stations worldwide. This dataset was produced by the CRU at the University of East Anglia (Harris et al. 2020). However, the data have an annual frequency and are not available for all departments in Guatemala. From a methodological viewpoint, nonlinear deterministic trends can also be examined even in the context of long memory models. In fact, various authors found that the two issues (nonlinearities and long memory) were intimately related (e.g., Diebold and Inoue 2001; Granger and Hyung 2004; Ohanissian et al. 2008). With respect to this, alternative modeling frameworks to the linear function in Eq. (1) may include Chebyshev polynomials in time (Cuestas and Gil-Alana 2016), Fourier functions in time (Gil-Alana and Yaya 2021) or even neural networks (Yaya et al. 2021). Other issues that might be investigated

include the presence of extreme observations and the study of the range between maximum and minimum temperatures over the region. Work in these directions is now under progress.

Acknowledgments. Author Luis A. Gil-Alana 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 four anonymous reviewers are gratefully acknowledged.

Data availability statement. Data are available from the authors upon request.

APPENDIX

Additional Data Series Figures

Figures A1 and A2 present the original data series of temperature and precipitation, respectively, for the 17 sites. Figures A3 and A4 present the anomaly data series of temperature and precipitation, respectively, for those sites.

REFERENCES

- Aguilar, E., and Coauthors, 2005: Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. J. Geophys. Res., 110, D23107, https://doi. org/10.1029/2005JD006119.
- Baillie, R. T., 1996: Long memory processes and fractional integration in econometrics. J. Econom., 73, 5–59, https://doi.org/ 10.1016/0304-4076(95)01732-1.
- Belbute, J. M., and A. M. Pereira, 2017: Do global CO₂ emissions from fossil-fuel consumption exhibit long memory? A fractional-integration analysis. *Appl. Econ.*, **49**, 4055–4070, https://doi.org/10.1080/00036846.2016.1273508.
- Cuestas, J. C., and L. A. Gil-Alana, 2016: A non-linear approach with long range dependence based on Chebyshev polynomials. *Stud. Nonlinear Dyn. Econom.*, 23, 445–468.
- De Lorenzo, A., and F. Liaño, 2017: High temperatures and nephrology: The climate change problem. *Nefrología*, **37**, 492– 500, https://doi.org/10.1016/j.nefro.2016.12.008.
- Deng, H., and Coauthors, 2022: Dynamics of diurnal precipitation differences and their spatial variations in China. J. Appl. Meteor. Climatol., 61, 1015–1027, https://doi.org/10.1175/JAMC-D-21-0232.1.
- Desbureaux, S., and A.-S. Rodella, 2019: Drought in the city: The economic impact of water scarcity in Latin American metropolitan areas. *World Dev.*, **114**, 13–27, https://doi.org/10.1016/ j.worlddev.2018.09.026.
- Dickey, D. A., and W. A. Fuller, 1979: Distribution of the estimators for autoregressive time series with a unit root. J. Amer. Stat. Assoc., 74, 427–431, https://doi.org/10.2307/2286348.
- Diebold, F. X., and G. D. Rudebush, 1991: On the power of Dickey-Fuller tests against fractional alternatives. *Econ. Lett.*, 35, 155–160, https://doi.org/10.1016/0165-1765(91)90163-F.
- —, and A. Inoue, 2001: Long memory and regime switching. J. Econom., 105, 131–159, https://doi.org/10.1016/S0304-4076 (01)00073-2.

103

- Dong, Q., W. Wang, K. E. Kunkel, Q. Shao, W. Xing, and J. Wei, 2021: Heterogeneous response of global precipitation concentration to global warming. *Int. J. Climatol.*, **41**, E2347–E2359, https://doi.org/10.1002/joc.6851.
- Dore, M. H. I., 2005: Climate change and changes in global precipitation patterns: What do we know? *Environ. Int.*, 31, 1167–1181, https://doi.org/10.1016/j.envint.2005.03.004.
- Elliott, G., T. J. Rothenberg, and J. H. Stock, 1996: Efficient tests for an autoregressive unit root. *Econometrica*, **64**, 813–836, https://doi.org/10.2307/2171846.
- Gil-Alana, L. A., 2005: Statistical model for the temperatures in the Northern Hemisphere using fractional integration techniques. J. Climate, 18, 5357–5369, https://doi.org/10.1175/JCLI3543.1.
- —, 2012: Long memory, seasonality and time trends in the average monthly temperatures in Alaska. *Theor. Appl. Climatol.*, **108**, 385–396, https://doi.org/10.1007/s00704-011-0539-0.
- —, and P. M. Robinson, 1997: Testing of unit root and other nonstationary hypotheses in macroeconomic time series. *J. Econom.*, 80, 241–268, https://doi.org/10.1016/S0304-4076(97)00038-9.
- —, and L. Sauci, 2019a: Temperatures across Europe: Evidence of time trends. *Climatic Change*, **157**, 355–364, https://doi.org/ 10.1007/s10584-019-02568-6.
- —, and —, 2019b: US temperatures: Time trends and persistence. *Int. J. Climatol.*, **39**, 5091–5103, https://doi.org/10.1002/joc.6128.
- —, and O. Yaya, 2021: Testing fractional unit roots with nonlinear smooth break approximations using Fourier functions. *J. Appl. Stat.*, **48**, 2542–2559, https://doi.org/10.1080/02664763. 2020.1757047.
- Granger, C. W. J., and N. Hyung, 2004: Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. J. Empirical Finance, 11, 399–421, https://doi.org/10.1016/j.jempfin.2003.03.001.
- Harris, I., T. J. Osborn, P. Jones, and D. Lister, 2020: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, 7, 109, https://doi.org/10.1038/s41597-020-0453-3.
- Hassler, U., and J. Wolters, 1994: On the power of unit root tests against fractional alternatives. *Econ. Lett.*, 45, 1–5, https://doi. org/10.1016/0165-1765(94)90049-3.
- Helbling, M., and D. Meierrieks, 2023: Global warming and urbanization. J. Popul. Econ., 36, 1187–1223, https://doi.org/10. 1007/s00148-022-00924-y.
- Jenkins, S., A. Povey, A. Gettelman, R. Grainger, P. Stier, and M. Allen, 2022: Is anthropogenic global warming accelerating? J. Climate, 35, 7873–7890, https://doi.org/10.1175/JCLI-D-22-0081.1.
- Kumar, V., D. Ranjan, and K. Verma, 2021: Global climate change: The loop between cause and impact. *Global Climate Change*, Elsevier, 187–211.
- Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, 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? J. Econom., 54, 159–178, https://doi. org/10.1016/0304-4076(92)90104-Y.

- Lee, D., and P. Schmidt, 1996: On the power of the KPSS test of stationarity against fractionally-integrated alternatives. J. Econom., 73, 285–302, https://doi.org/10.1016/0304-4076(95)01741-0.
- Liu, J., X. Ma, Z. Duan, J. Jiang, M. Reichstein, and M. Jung, 2020: Impact of temporal precipitation variability on ecosystem productivity. *Wiley Interdiscip. Rev.: Water*, 7, e1481, https://doi.org/ 10.1002/wat2.1481.
- Ludescher, J., A. Bunde, C. L. E. Franzke, and H. J. Schellnhuber, 2016: Long-term persistence enhances uncertainty about anthropogenic warming of Antarctica. *Climate Dyn.*, **46**, 263– 271, https://doi.org/10.1007/s00382-015-2582-5.
- Lyra, A., P. Imbach, D. Rodriguez, S. Chan, S. Geoirgiou, and L. Garofolo, 2017: Projections of climate change impacts on central America tropical rainforest. *Climatic Change*, **141**, 93–105, https://doi.org/10.1007/s10584-016-1790-2.
- Ohanissian, A., J. R. Russell, and R. S. Tsay, 2008: True or spurious long memory? A new test. J. Bus. Econ. Stat., 26, 161– 175, https://doi.org/10.1198/07350010700000340.
- Phillips, P. C. B., and P. Perron, 1988: Testing for a unit root in time series regression. *Biometrika*, **75**, 335–346, https://doi. org/10.1093/biomet/75.2.335.
- Robinson, P. M., 1994: Efficient tests of nonstationary hypotheses. J. Amer. Stat. Assoc., 89, 1420–1437, https://doi.org/10.1080/ 01621459.1994.10476881.
- Shi, L., I. Kloog, A. Zanobetti, P. Liu, and J. D. Schwartz, 2015: Impacts of temperature and its variability on mortality in New England. *Nat. Climate Change*, 5, 988–991, https://doi. org/10.1038/nclimate2704.
- Sun, Q., X. Zhang, F. Zwiers, S. Westra, and L. V. Alexander, 2021: A global, continental, and regional analysis of changes in extreme precipitation. J. Climate, 34, 243–258, https://doi. org/10.1175/JCLI-D-19-0892.1.
- Ventosa-Santaulària, D., D. R. Heres, and L. C. Martínez-Hernández, 2014: Long-memory and the sea level-temperature relationship: A fractional cointegration approach. *PLOS ONE*, 9, e113439, https://doi.org/10.1371/journal.pone.0113439.
- Vogelsang, T. J., and P. H. Franses, 2005: Are winters getting warmer? *Environ. Modell. Software*, 20, 1449–1455, https:// doi.org/10.1016/j.envsoft.2004.09.016.
- World Bank, 2023: Climate Change Knowledge Portal. World Bank Group, accessed 12 June 2023, https://climateknowledgeportal. worldbank.org/.
- Wu, Y., S.-Y. Wu, J. Wen, M. Xu, and J. Tan, 2016: Changing characteristics of precipitation in China during 1960–2012. *Int. J. Climatol.*, 36, 1387–1402, https://doi.org/10.1002/joc.4432.
- Yaya, O. S., A. E. Ogbonna, F. Furuoka, and L. A. Gil-Alana, 2021: A new unit root test for unemployment hysteresis based on the autoregressive neural network. *Oxford Bull. Econ. Stat.*, 83, 960–981, https://doi.org/10.1111/obes.12422.