



# Particulate matter 10 (PM<sub>10</sub>): persistence and trends in eight European capitals

Guglielmo Maria Caporale<sup>1</sup> · Luis A. Gil-Alana<sup>2,3</sup> · Nieves Carmona-González<sup>3</sup>

Received: 9 September 2020 / Accepted: 23 February 2021 / Published online: 11 March 2021

© The Author(s) 2021

## Abstract

This paper examines the statistical properties of daily PM<sub>10</sub> in eight European capitals (Amsterdam, Berlin, Brussels, Helsinki, London, Luxembourg, Madrid and Paris) over the period 2014–2020 by applying a fractional integration framework; this is more general than the standard approach based on the classical dichotomy between I(0) stationary and I(1) non-stationary series used in most other studies on air pollutants. All series are found to be characterised by long memory and fractional integration, with orders of integration in the range (0, 1), which implies that mean reversion occurs and shocks do not have permanent effects. Persistence is the highest in the case of Brussels, Amsterdam and London. The presence of negative trends in Brussels, Paris and Berlin indicates some degree of success in reducing pollution in these capitals.

**Keywords** Fractional integration · Long memory · Persistence · Trends · Air pollutants · PM<sub>10</sub>

**JEL classification** C22 · Q53

## Introduction

Particulate matter (PM) are **microscopic particles of solid or liquid matter suspended** in the **air**. Its sources can be natural or anthropogenic. It has a significant impact on both climate and precipitation and therefore has both health and social costs. In particular, it affects the amount of incoming solar radiation and outgoing terrestrial radiation. The coarse particles can have a diameter between 2.5 and 10  $\mu\text{m}$  (PM<sub>10</sub>) and are known to be a very harmful form of air pollution given their ability to penetrate into the lungs and blood streams and cause respiratory and heart diseases as well as premature death. Various countries have therefore set limits for particulates in the air, which are emitted during the combustion of vehicle engine fuels, braking and tyre wearing. In particular, the European Union has defined in a series of directives the

acceptable limits for exhaust emissions of new vehicles sold in the European Union and EEA member states.

Numerous studies have analysed the connection between pollution and harmful health effects (e.g. Schwartz and Marcus 1990; Anderson et al. 1996; Atkinson et al. 1999; Gardner and Dorling 1999). The present study contributes to another branch of the literature which focuses instead on modelling various pollutants such as sulphur dioxide (SO<sub>2</sub>), nitrogen dioxide (NO<sub>2</sub>), carbon monoxide (CO), ozone (O<sub>3</sub>), PM<sub>2.5</sub> and PM<sub>10</sub>. For instance, Zamri et al. (2009) applied the Box-Jenkins ARIMA approach to model CO and NO<sub>2</sub> in Malaysia and found an upward trend. Li et al. (2017) analysed air quality in Beijing from 2014 to 2016 using the spatio-temporal deep learning (STDL) model, the time delay neural network (TDNN) model, the ARMA model, the support vector regression (SVR) model and the long short-term memory neural network extended (LSTME) model and concluded that the LSTME model is the most suitable one for time series characterised by long-term dependence with optimal time delays. Naveen and Anu (2017) studied air quality in India using ARIMA, seasonal ARIMA (SARIMA) and other models. Pan and Chen (2008) is one of the few studies using long-memory autoregressive fractional integrated moving average (ARFIMA) models for air pollution data (in the case of Taiwan) and concluding that these are more accurate than autoregressive integrated moving average (ARIMA) models

---

✉ Guglielmo Maria Caporale  
Guglielmo-Maria.Caporale@brunel.ac.uk

<sup>1</sup> Department of Economics and Finance, Brunel University London, London UB8 3PH, UK

<sup>2</sup> University of Navarra, Pamplona, Spain

<sup>3</sup> Universidad Francisco de Vitoria, Madrid, Spain

for such series. We also apply the latter type of framework for our analysis; however, instead of imposing a specific ARMA structure for the differenced process, we use white noise errors or alternatively the non-parametric approach of Bloomfield (1973), thus avoiding the issue of misspecification that might occur when choosing the short-run components.

It is clearly important to investigate the dynamics of air pollution to develop suitable models for prediction purposes and design policies to manage air quality. This paper examines the statistical properties of daily PM<sub>10</sub> in eight European capitals (Amsterdam, Berlin, Brussels, Helsinki, London, Luxembourg, Madrid and Paris) over the period 2014–2020 by applying a fractional integration framework that is more general than the standard approach based on the classical dichotomy between I(0) stationary and I(1) non-stationary series used in the vast majority of previous studies on air pollutants, since it allows for fractional as well as integer degrees of differentiation and thus for a much wider set of stochastic behaviours. In particular, it enables the researcher to analyse the long-memory properties of the series of interest and the possible presence of trends, to test for mean reversion and to measure the degree of persistence and the speed of adjustment to the long-run equilibrium level. Therefore, it provides information about whether the effects of shocks are transitory or permanent, which is a crucial piece of information for adopting appropriate policy measures.

The remainder of the paper is structured as follows: the ‘Methodology’ section outlines the methodology used for the analysis; the ‘Data’ section describes the data; the ‘Empirical results’ section presents the empirical results; and the ‘Conclusions’ section offers some concluding remarks.

## Methodology

As mentioned above, we adopt a long-memory approach based on fractional integration. Long memory is a feature of time series that are characterised by a high degree of dependence between observations which are far apart in time. It has been found to be displayed by many time series in different fields such as climatology (Gil-Alana 2005, 2008, 2017; Vyushin and Kushner 2009; Franzke 2012; Ludescher et al. 2016; Bunde 2017; Yuan et al. 2019; Bruneau et al. 2020), environmental sciences (Barros et al. 2016; Gil-Alana et al. 2016; Tiwari et al. 2016; Gil-Alana and Solarin 2018, Gil-Alana and Trani 2019; Xayasouk et al. 2020) and economics and finance (Gil-Alana and Moreno 2012; Abritti et al. 2017; Kalemkerian and Sosa 2020; Murialdo et al. 2020; Qiu et al. 2020).

There exist a variety of statistical models that can describe this type of behaviour; a very popular one among time series analysts is based on the concept of fractional integration, which occurs when the number of differences required to

make a series stationary I(0) is a fractional value. More precisely, a time series is said to be integrated of order  $d$  or I( $d$ ) if it can be expressed as:

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

where  $B$  is the backshift operator ( $Bx_t = x_{t-1}$ ), the differencing parameter  $d$  can be any real value, and  $u_t$  is I(0) defined as a covariance stationary process with a spectral density function that is positive and bounded at all frequencies in the spectrum. This framework encompasses different cases such as short memory ( $d = 0$ ), stationary long memory ( $0 < d < 0.5$ ), non-stationary though mean-reverting processes ( $0.5 \leq d < 1$ ), unit roots ( $d = 1$ ) and explosive patterns ( $d \geq 1$ ).

## Data

The series analysed is the daily average air quality taken from the World Air Quality Index (WAQI) at <https://aqicn.org/map/>

**Table 1** Descriptive statistics

Amsterdam		Berlin	
Start date	End date	Start date	End date
01/01/2014	13/06/2020	20/08/2014	13/06/2020
No. of observations:	2145	No. of observations:	2114
Mean:	25.7	Mean:	28.0
Standard deviation:	11.1	Standard deviation:	11.9
Variance:	124.1	Variance:	140.6
Min/max:	8/263	Min/max:	8/95
Range:	255	Range:	87
Brussels		Helsinki	
Start date	End date	Start date	End date
31/12/2013	13/06/2020	02/05/2014	13/06/2020
No. of observations:	2334	No. of observations:	2214
Mean:	24.0	Mean:	18.1
Standard deviation:	11.8	Standard deviation:	10.0
Variance:	140.4	Variance:	99.9
Min/max:	1/100	Min/max:	3/90
Range:	99	Range:	87
London		Luxembourg	
Start date	End date	Start date	End date
31/12/2013	13/06/2020	19/06/2015	13/06/2020
No. of observations:	2353	No. of observations:	1583
Mean:	26.5	Mean:	19.5
Standard deviation:	9.9	Standard deviation:	6.8
Variance:	97.1	Variance:	46.2
Min/max:	5/89	Min/max:	2/52
Range:	84	Range:	50
Madrid		Paris	
Start date	End date	Start date	End date
31/12/2013	13/06/2020	31/12/2013	13/06/2020
No. of observations:	2323	No. of observations:	2227
Mean:	24.3	Mean:	39.3
Standard deviation:	11.7	Standard deviation:	14.5
Variance:	138.0	Variance:	210.5
Min/max.:	5/160	Min/max:	6/122
Range:	155	Range:	116

**Table 2** Estimates of *d*: white noise errors

Series	No deterministic terms	An intercept	An intercept and a time trend
Amsterdam	0.42 (0.39, 0.46)	<b>0.39 (0.35, 0.44)</b>	0.39 (0.35, 0.44)
Berlin	0.63 (0.58, 0.68)	<b>0.61 (0.56, 0.67)</b>	0.61 (0.56, 0.67)
Brussels	0.52 (0.48, 0.56)	<b>0.50 (0.46, 0.55)</b>	0.50 (0.46, 0.55)
Helsinki	0.57 (0.52, 0.61)	<b>0.54 (0.50, 0.59)</b>	0.54 (0.50, 0.59)
London	0.54 (0.50, 0.59)	<b>0.52 (0.47, 0.57)</b>	0.52 (0.47, 0.57)
Luxembourg	0.56 (0.51, 0.62)	<b>0.54 (0.48, 0.61)</b>	0.54 (0.48, 0.61)
Madrid	0.63 (0.58, 0.68)	<b>0.62 (0.57, 0.67)</b>	0.62 (0.57, 0.67)
Paris	0.55 (0.51, 0.59)	<b>0.53 (0.49, 0.58)</b>	0.53 (0.49, 0.58)

We report the estimates of *d* and its 95% confidence band (in parenthesis). In bold, the selected specification for each series

[world/es/](#). All data have been converted using the US EPA standard (United States Environmental Protection Agency). Specifically, we use daily data for the past 7 years (2014–2020) concerning eight European capitals: Amsterdam, Berlin, Brussels, Helsinki, London, Luxembourg, Madrid and Paris. The series represents the daily level of air quality (PM<sub>10</sub>) measured in micrograms per cubic metre of air (µg/m<sup>3</sup>). The WAQI data come from the following original sources: Madrid, <http://www.mambiente.madrid.es/opencms/opencms/cal aire/> (Ayuntamiento de Madrid); Paris, <http://www.airparif.asso.fr/> (AirParif—Association de surveillance de la qualité de l'air en Île-de-France); Amsterdam, <https://www.luchtmeetnet.nl/> (RIVM); Luxembourg, <https://environnement.public.lu/fr.html> (Portail de l'Environnement du Grand-duché de Luxembourg); London, <https://uk-air.defra.gov.uk/> (UK-AIR, air quality information resource, Defra, UK); Helsinki, <https://www.ilmatieteenlaitos.fi/ilmanlaatu> (Ilmanlaatu Suomessa); Brussels, <https://www.irceline.be/en/> (Belgian Interregional Environment Agency); and Berlin, <https://www.berlin.de/senuvk/umwelt/luftqualitaet/> (Luftqualität).

Table 1 reports the sample periods for each capital and provides some descriptive statistics for each series. It can be seen that Paris exhibits the highest mean value, while Helsinki

has the lowest. Paris also has the most volatile series, while Luxembourg has the least volatile.

### Empirical results

We estimate the following model:

$$y_t = \alpha + \beta t + x_t, \quad (1-B)^d X_t = u_t, \quad t = 1, 2, \dots, \quad (2)$$

where  $y_t$  stands for PM<sub>10</sub> in each European capital in turn, and  $x_t$  is an I(*d*) process such that the error term  $u_t$  is I(0); the disturbances are assumed to follow a white noise (see Tables 2 and 3) and an autocorrelated process (see Tables 4 and 5) in turn, where the latter is modelled using the exponential spectral framework of Bloomfield (1973). In all cases, we display the estimated values of *d* (and their associated 95% confidence bands) for three different specifications: (i) no deterministic terms in (2), i.e. we impose the restriction  $\alpha = \beta = 0$  (the results for this case are reported in the second column in Tables 2 and 4); (ii) an intercept only (see the third column in both tables); and (iii) an intercept and a linear time trend (see the fourth column in both tables). The estimated values reported in bold in these tables are those corresponding to our

**Table 3** Estimated coefficients in the selected model: white noise errors

Series	<i>d</i> (95% band)	Intercept ( <i>t</i> -value)	Time trend
Amsterdam	0.39 (0.35, 0.44)	26.0450 (8.77)	---
Berlin	0.61 (0.56, 0.67)	25.3864 (3.93)	---
Brussels	0.50 (0.46, 0.55)	20.9146 (4.17)	---
Helsinki	0.54 (0.50, 0.59)	20.4913 (4.57)	---
London	0.52 (0.47, 0.57)	23.0696 (5.23)	---
Luxembourg	0.54 (0.48, 0.61)	15.6394 (4.61)	---
Madrid	0.62 (0.57, 0.67)	16.5249 (2.59)	---
Paris	0.53 (0.49, 0.58)	32.9226 (5.16)	---

The values in parenthesis in column 3 are the corresponding *t*-values

**Table 4** Estimates of  $d$ : autocorrelated errors

Series	No deterministic terms	An intercept	An intercept and a time trend
Amsterdam	0.34 (0.29, 0.39)	<b>0.26 (0.21, 0.30)</b>	0.25 (0.20, 0.30)
Berlin	0.38 (0.33, 0.43)	0.30 (0.25, 0.36)	<b>0.29 (0.23, 0.36)</b>
Brussels	0.33 (0.28, 0.38)	0.26 (0.21, 0.32)	<b>0.25 (0.21, 0.31)</b>
Helsinki	0.40 (0.35, 0.45)	<b>0.33 (0.29, 0.39)</b>	0.33 (0.29, 0.38)
London	0.34 (0.30, 0.39)	<b>0.26 (0.21, 0.31)</b>	0.27 (0.22, 0.30)
Luxembourg	0.32 (0.28, 0.38)	<b>0.22 (0.16, 0.29)</b>	0.21 (0.15, 0.28)
Madrid	0.34 (0.30, 0.39)	<b>0.30 (0.25, 0.35)</b>	0.30 (0.25, 0.35)
Paris	0.41 (0.37, 0.46)	0.33 (0.28, 0.38)	<b>0.33 (0.28, 0.38)</b>

We report the estimates of  $d$  and its 95% confidence band (in parenthesis). In bold, the selected specification for each series

preferred specification, which has been selected on the basis of the statistical significance of the regressors.

When assuming that  $u_t$  is a white noise, the intercept is found to be the only significant deterministic term in all cases; the estimated values of  $d$  are in the interval (0, 1), which implies long memory and fractional integration. They range between 0.39 (Amsterdam) and 0.62 (Madrid). For Amsterdam, the values are all within the stationary range ( $d < 0.5$ ); for Brussels, London, Paris and London, they are around the stationary boundary ( $d = 0.5$ ), while non-stationarity ( $d \geq 0.5$ ) is found in the case of Helsinki, Berlin and Madrid.

When allowing for autocorrelation (Tables 4 and 5), the time trend appears to be negative and statistically significant in the case of Brussels, Berlin and Paris, which might reflect the anti-pollution policies adopted in these three capitals. In particular, a low emission zone (LEZ) was established in the Brussels region with the aim of meeting the European air quality standards and emission ceilings; in Berlin the German Climate Action Plan 2050 is being implemented to control air pollution by laying down environmental quality standards and emission reduction

requirements; a LEZ based on Euro norm vehicle classification has also been introduced in Paris.

The estimated values of  $d$  are once more in the interval (0, 1), though they are now significantly smaller than in the previous case. In fact, they are all within the stationary range, specifically between 0.22 (Luxembourg) and 0.33 (Helsinki and Paris). These lower estimates are likely to reflect the competition between the fractional integration and Bloomfield parameters in describing time dependence between the observations. Both sets of estimates, under the assumption of white noise and autocorrelated errors, respectively, indicate that the degree of persistence is highest in the case of Brussels, Amsterdam and London and lowest in the case of Helsinki, Berlin, and Madrid; thus, the effects of shocks are more long-lived in the former capitals.

## Conclusions

This paper has used fractional integration methods to obtain evidence on persistence and time trends in  $PM_{10}$  in eight European capitals (Amsterdam, Berlin, Brussels, Helsinki, London, Luxembourg, Madrid and Paris). This approach is more general than the standard ones used in most of the literature on air pollutants and thus is more informative about the time series properties of the series of interest. The results indicate that all of them display fractional integration with orders of integration in the range (0,1); this implies that mean reversion occurs and shocks do not have permanent effects. However, the degree of persistence is different in the eight capitals examined; in particular, the effects of shocks take longer to die away in the case of Brussels, Amsterdam and London. Such evidence should be taken into account by policymakers aiming to design effective measures to reduce pollution.

The estimated values of  $d$  are lower under the assumption of autocorrelated errors; in this case, three of the capitals examined (Brussels, Paris and Berlin) exhibit statistically significant

**Table 5** Estimated coefficients in the selected model: autocorrelated errors

Series	$d$ (95% band)	Intercept ( $t$ -value)	Time trend ( $t$ -value)
Amsterdam	0.26 (0.21, 0.30)	25.6832 (21.18)	---
Berlin	0.29 (0.23, 0.36)	32.6088 (16.38)	-0.0046 (-2.92)
Brussels	0.25 (0.21, 0.31)	26.3621 (15.22)	-0.0022 (1.97)
Helsinki	0.33 (0.29, 0.39)	19.6888 (12.88)	---
London	0.26 (0.21, 0.31)	25.9193 (24.21)	---
Luxembourg	0.22 (0.16, 0.29)	19.2309 (34.57)	---
Madrid	0.30 (0.25, 0.35)	23.5407 (16.78)	---
Paris	0.33 (0.28, 0.38)	42.2180 (13.79)	-0.0044 (-1.98)

The values in parenthesis in columns 3 and 4 are the corresponding  $t$ -values for the intercept and the time trend, respectively

negative time trends, which suggests that the policies they have adopted to reduce pollution (such as the establishment of LEZs) have been successful, at least to some extent.

Other issues that could be investigated in the case of PM<sub>10</sub> are the following: seasonality with fractional integration (Gil-Alana and Robinson 2001; Bisognin and Lopes 2009; del Barrio Castro and Rachinger 2021; etc.); non-linearities and structural breaks, which is of particular interest given the fact that these are both strongly related to long memory and fractional integration (see, e.g. Diebold and Inoue 2001; Granger and Hyung 2004; Ohanissian et al. 2008; Kongcharoen 2013; etc.); and the forecasting performance of alternative specifications; all these topics are left for future work.

**Acknowledgements** We are grateful to an anonymous referee for useful comments. Prof. Luis A. Gil-Alana also 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). He also acknowledges financial support from internal Projects of the Universidad Francisco de Vitoria.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abritti M, Gil-Alana LA, Lovcha Y, Moreno A (2017) Term structure persistence. *J Financ Econ* 14:331–352
- Anderson HR, Ponce de Leon A, Martin-Bland J, Bower JS, Strachan DP (1996) Air pollution and daily mortality in London: 1987–92. *BMJ* 312:665–669. <https://doi.org/10.1136/bmj.312.7032.665>
- Atkinson RW, Anderson HR, Strachan DP, Bland JM, Bremner SA, Ponce de Leon A (1999) Short-term associations between outdoor air pollution and visits to accident and emergency departments in London for respiratory complaints. *Eur Respir J* 13(2):257–265. <https://doi.org/10.1183/09031936.99.1322579>
- Barros CP, Gil-Alana LA, Perez de Gracia F (2016) Stationarity and long range dependence of carbon dioxide emissions. Evidence from disaggregated data. *Environ Resour Econ* 63(1):45–56
- Bisognin C, Lopes SRC (2009) Properties of seasonal long memory processes. *Math Comput Model* 49(9–10):1837–1851
- Bloomfield P (1973) An exponential model in the spectrum of a scalar time series. *Biometrika* 60(2):217–226
- Bruneau N, Wang S, Toumi R (2020) Long Memory Impact of Ocean Mesoscale Temperature Anomalies on Tropical Cyclone Size, *Geophysical Research Letters* 47. Vol. 6
- Bunde, A. (2017), Long-term memory in climate: detection, extreme events and significance of trends, Chapter 11 in *Nonlinear and Stochastic Climate Dynamics*, edited by Christian L.E. Franzke and Terence O’Kane, Cambridge University Press
- Del Barrio Castro, Rachinger (2021) Aggregation of Seasonal Long-Memory Processes. *Econometrics and Statistics*, forthcoming
- Diebold FX, Inoue A (2001) Long memory and regime switching. *J Econ* 105:131–159
- Franzke C (2012) Nonlinear trends, long-range dependence, and climate noise properties of surface temperature. *J Clim* 25(12):4172–4183
- Gardner MW, Dorling SR (1999) Neural network modelling and prediction of hourly NO<sub>x</sub> and NO<sub>2</sub> concentrations in urban air in London. *Atmos Environ* 33:709–719. [https://doi.org/10.1016/S1352-2310\(98\)00230-1](https://doi.org/10.1016/S1352-2310(98)00230-1)
- Gil-Alana LA (2005) Statistical modeling of the temperatures in the Northern hemisphere using fractional integration techniques. *J Clim* 27(10):3477–3491
- Gil-Alana LA (2008) Time trend estimation with breaks in temperature series. *Clim Chang* 89(3):325–337
- Gil-Alana LA (2017) Alternative modeling approaches for the ENSO time series. Persistence and seasonality. *Int J Climatol* 37(5):2354–2363
- Gil-Alana LA, Gupta R, Perez de Gracia F (2016) Modelling persistence of carbon emissions allowance prices. *Renew Sust Energ Rev* 55: 221–226
- Gil-Alana LA, Moreno A (2012) Uncovering the US term premium. An alternative route. *J Bank Financ* 36(4):1181–1193
- Gil-Alana LA, Robinson PM (2001) Testing seasonal fractional integration in the UK and Japanese consumption and income. *J Appl Econ* 16:95–114
- Gil-Alana LA, Solarin SA (2018) Have US environmental policies been effective in the reduction of US emissions? A new approach using fractional integration. *Atmospheric Pollution Research* 9(1):53–60
- Gil-Alana LA, Trani T (2019) Time trends and persistence in the global CO<sub>2</sub> emissions across Europe. *Environ Resour Econ* 73:213–328
- Granger CWJ, Hyung N (2004) Occasional structural breaks and long memory with an application to the S&P 500 absolute stock return. *J Empir Financ* 11:399–421
- Kalemkerian J, Sosa A (2020) Long-range dependence in the volatility of returns in Uruguayan sovereign debt indices. *American Institute of Mathematical Sciences* 7(3):225–237
- Kongcharoen C (2013) Forecasting using nonlinear long memory models with artificial neural network expansion. In: Huynh VN., Kreinovich V., Sriboonchitta S., Suriya K. (eds) *Uncertainty Analysis in Econometrics with Applications. Advances in Intelligent Systems and Computing*, vol 200. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-35443-4\\_17](https://doi.org/10.1007/978-3-642-35443-4_17)
- Li X, Peng L, Yao X, Cui S, Hi Y, You C, Chi T (2017) Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation. *Environ Pollut* 231:1997–1004. <https://doi.org/10.1016/j.envpol.2017.08.114>
- Ludescher J, Bunde A, Franzke CL, Schellnhuber HJ (2016) Long-term persistence enhances uncertainty about anthropogenic warming of Antarctica. *Clim Dyn* 46(1–2):263–271. <https://doi.org/10.1007/s00382-015-2582-5>
- Murialdo P, Ponta L, Carbone A (2020) Long-range dependence in financial markets: a moving average cluster entropy approach. *Entropy* 22(6):634
- Naveen V, Anu N (2017) Time series analysis to forecast air quality indices in Thiruvananthapuram District, Kerala, India. *Int J Eng Res Appl* 7(6):66–84. <https://doi.org/10.9790/9622-0706036684>
- Ohanissian A, Russell JR, Tsay RS (2008) True or spurious long memory? A new test, *Journal of Business and Economic Statistics* 26: 161–175

- Pan JN, Chen ST (2008) Monitoring long-memory air quality data using ARFIMA model. *Environmetrics* 19(2):209–219. <https://doi.org/10.1002/env.88>
- Qiu J, Wang B, Zhou C (2020) Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PLoS One* 15:1
- Schwartz O, Marcus A (1990) Mortality and air pollution in London: a time series analysis. *Am J Epidemiol* 131(1):185–194. <https://doi.org/10.1136/jech.54.10.750>
- Tiwari AK, Kyophilavong P, Albulescu CT (2016) Testing the stationarity of CO<sub>2</sub> emissions series in Sub-Saharan African countries by incorporating nonlinearity and smooth breaks. *Res Int Bus Financ* 37:527–540
- Vyushin DI, Kushner PJ (2009) Power-law and long-memory characteristics of the atmospheric general circulation. *J Clim* 22(11):2890–2904
- Xayasouk T, Lee HM, Lee G (2020) Air pollution prediction using long short-term memory (LSTM) and deep autoencoder (DAE) models. *Sustainability* 12:6
- Yuan N, Huang Y, Duan J, Zhu C, Xoplaki E, Luterbacher J (2019) On climate prediction; how much can we expect from climate memory? *Clim Dyn* 52(1-2):855–864
- Zamri IM, Roziah Z, Marzuki I, Muhd SL (2009) Forecasting and time series analysis of air pollutants in several area of Malaysia. *Am J Environ Sci* 5:625–632. <https://doi.org/10.3844/ajessp.2009.625.632>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.