

Multiscaled causality of infections on viral testing volumes: The case of COVID-19 in Tunisia

Foued Saâdaoui¹  | Hana Rabbouch² | Hayet Saadaoui³ | Frédéric Dutheil⁴

¹Department of Statistics, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

²Université de Tunis, Institut Supérieur de Gestion de Tunis, Tunis, Tunisia

³Montpellier SupAgro, UMR LAMETA, Montpellier, France

⁴Université Clermont Auvergne, CNRS, LaPSCo, Physiological and Psychosocial Stress, University Hospital of Clermont-Ferrand, CHU Clermont-Ferrand, Occupational and Environmental Medicine, WittyFit, Clermont-Ferrand, France

Correspondence

Foued Saâdaoui, Department of Statistics, Faculty of Sciences, King Abdulaziz University, P.O BOX 80203, Jeddah 21589, Saudi Arabia.
Email: foued.saadaoui@gmail.com

Abstract

Objectives: Coronavirus disease (COVID-19) is one of the most detrimental pandemics that affected the humanity throughout the ages. The irregular historical progression of the virus over the first year of the pandemic was accompanied with far-reaching health and social damages. To prepare logistically against this worsening disaster, many public authorities around the world had set up screening and forecasting studies. This article aims to analyse the time-frequency co-evolution of the number of confirmed cases (NCC) in Tunisia and the related number of performed polymerase chain reaction (PCR) tests over the COVID-19 first year. Accurately predicting such a relationship allows Tunisian authorities to set up an effective health prevention plan.

Study Design: In order to keep pace with the speed of evolution of the virus, we used uninterrupted daily time series from the Tunisian Ministry of Public Health (TMPH) recorded over the COVID-19 first year. The objective is to: (1) analyse the time-frequency progress of the NCC in relationship with the number of PCR tests, (2) identify a multi-scale two-factor stochastic model in order to develop a robust bivariate forecasting technique.

Methods: We assume a bivariate stochastic process which is projected onto a set of wavelet sub-spaces to investigate the scale-by-scale co-evolution of the NCC/PCR over the COVID-19 first year. A wavelet-based multiresolutional causality test is then performed.

Results: The main results recommend the rejection of the null hypothesis of no instantaneous causality in both directions, while the statistics of the Granger test suggest failing to reject the null hypothesis of non-causality. However, by proceeding scale-by-scale, the Granger causality is proven significant in both directions over varying frequency bands.

Conclusions: It is important to include the NCC and PCR variables in any time series model intended to predict one of these variables. Such a bivariate and multi-scale model is supposed to better predict the needs of the public health sector in screening tests. On this basis, testing campaigns with multiple periodicities can be planned by the Tunisian authorities.

KEYWORDS

causality, COVID-19, demand forecasting, health planning and logistics, PCR tests, virus dynamics

Highlights

- We test the causality of COVID-19 infections on viral testing volumes in Tunisia
- A wavelet-based approach is applied to examine the relationship scale-by-scale
- The results recommend the rejection of the noncausality at all frequency bands
- Screening campaigns with multiple periodicities can be conducted by Tunisian MOH

1 | INTRODUCTION

Over the few last decades, medical practices have known significant changes, which altered the demand for supplies/service and the ways they are employed. These changes therefore have, in fact, budgetary requirements, especially in terms of capital investment and revenue requirements. With the onset of Coronavirus disease 2019 (COVID-19), this equation has become even more complicated, hence the need for effective management plans to avoid stock-out of medical supplies. Thus, accurately planning health care needs is an important issue, yet existing forecasting models are often inadequate. On this basis, there is a need to methods whereby the optimum level of provision can be predicted and the likely effects of changes proposed can be evaluated.¹ An effective multiresolutional forecasting model,²⁻⁵ for instance, should enable health planners/managers to minimise costs and maximise the use of projected capital and investment. In all of this, the role of the authorities would therefore be to plan with certainty the necessary financial resources by the request for additional budgets intended for the public health sector over COVID-19.

COVID-19 is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Since identified in 2019 in Wuhan (China), COVID-19 has generated much fear due to its relatively high mortality rate. Today, the current pandemic caused by the highly contagious SARS-CoV-2 and its effects on the global economy is of major global concern.⁶⁻⁸ Preventive and organisational measures have been taken at the global and regional

levels to limit and control the spread of the pandemic and its repercussions.^{9,10} Tunisia, in particular, was not only among the first north-African countries to be affected by the virus, but also the among the firsts to know how to fight it effectively. The Tunisian Public Health Ministry (TPHM) had in fact opted for protectionist and proactive strategies based upon effective screening and forecasting plans. This had immediately succeeded in stopping the scourge in a very effective manner, despite the fact that the North African nation is situated in a Mediterranean zone of great liveliness, and in particular in the neighbourhoods of Italy, which was one of the most affected by the pandemic during the first year.

On the other hand, the COVID-19 pandemic is a powerful reminder of the importance of having health services that are effective and just-in-time. For the public authorities, predicting the instantaneous evolution of the number of cases affected by this disease is both a delicate and important mission in the sense that its success allows to accurately determine in advance the material and logistic needs in order to protect themselves from severe viral waves.¹¹ The most popular models used in a wide range are based on a single factor model developed in a continuous-time setting, which provides a rich framework for specifying the dynamic behaviour in the short-term.¹² However, the most known models does not capture important information about nonstationary processes. Therefore, wavelets are commonly used to decompose data into a sum of orthogonal sub-signals.^{3,4,13,14} Wavelets are known as very practical and accurate mathematical tools for dealing with nonlinearities and problems where phenomena of very different scales occur. In the wavelet-decomposed signals, the finest time-scales represent the short-run dynamics (corresponding to the very high- and high-frequency components), the coarsest time scales characterise cycles (epidemic waves), while the trend is associated to approximation components corresponding to the long-run elements.

In this paper, a wavelet-based bivariate time series methodology is proposed for analysing the relationship between the daily COVID-19 infections and the number of polymerase chain reaction (PCR) tests in Tunisia. The idea of such a relationship has in fact recently been spread in some research directions.¹⁵ The main interest of this conjoint modelling is to allow health authorities in Tunisia to prepare logistically against the arrival of new COVID-19 waves in order to optimise their needs in terms of PCR tests. To keep pace with the speed of evolution of the virus during its first year, uninterrupted daily time series are used in our study. The strength of the wavelet-based multiresolution analysis (MRA), in comparison with classical approaches, is that it offers a powerful way to decompose the dynamics of the signals over a variety of different time horizons. Such a framework could therefore represent the cornerstone of a multi-periodic screening plan. The main results recommend the rejection of the null hypothesis of no instantaneous causality in both directions, while the statistics of the Granger test suggest failing to reject the null hypothesis of non-causality also in both directions. By proceeding rather scale-by-scale, the Granger causality is proven significant in both directions over varying frequency bands. The implication of such results is that it is important to include both variables in any model intended to predict one of the variables. Such a bivariate and multi-scale model is supposed to better predict the needs of the public health sector in screening tests.

2 | METHODS

The wavelet transform^{3,4,13,14} decomposes a signal $f(t)$ into some regular functions called daughter wavelets $\psi_{s,\tau}(t)$, such as $\psi_{s,\tau}(t) = s^{-\frac{1}{2}}\psi[(t - \tau) / s]$, where $s > 0$. The discrete wavelet representation of $f(t)$ in $L^2(\mathbb{R})$ is as follows:

$$f(t) = \sum_{k \in \mathbb{Z}} c_{0,k} \phi_{0,k}(t) + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t), \quad (1)$$

where $c_{0,k}$ summarise the general form of the function and the $d_{j,k}$ represent the local details. These coefficients, known as scaling and wavelets coefficients, are calculated as $c_{0,k} = \int_{-\infty}^{+\infty} f(t) \phi_{0,k}^*(t) dt$ and $d_{j,k} = \int_{-\infty}^{+\infty} f(t) \psi_{j,k}^*(t) dt$. The filter coefficients of a compactly supported wavelet, commonly denoted $\{h_j\}_{j=0}^{L-1}$ and $\{g_j\}_{j=0}^{L-1}$, are defined via the following quadrature mirror relationship $h_m = (-1)^m g_{L-m-1}$, $m = 0, \dots, L - 1$. L is the width of the filters. The filters must also

satisfy $\sum_{l=0}^{L-1} h_l = 0$, $\sum_{l=0}^{L-1} h_l^2 = 0$ and $\sum_{l=0}^{L-1} h_l h_{l+2k} = 0$. The coefficients of a Maximal Overlap Discrete Wavelet Transform (MODWT) can be defined as:

$$\tilde{d}_{j,k}^{(f)} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} f(k - l_{\text{mod}}N) \text{ and } \tilde{c}_{j,k}^{(f)} = \sum_{l=0}^{L_j-1} \tilde{g}_{j,l} f(k - l_{\text{mod}}N), \tag{2}$$

where $\tilde{g}_{j,l} = g_{j,l} / 2^{j/2}$ and $\tilde{h}_{j,l} = h_{j,l} / 2^{j/2}$, so that $\sum_{l=0}^{L_j-1} \tilde{g}_{j,l}^2 = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l}^2 = 1/2^j$. $L_j = (2^j - 1)(L - 1) + 1$ is the length of the new wavelet filter for level j . Finally, the MODWT algorithm provides a MRA of f , which is represented as follows:

$$f(t) = \sum_{j=1}^J \tilde{D}_{j,k}^{(f)} + \tilde{S}_{j,k}^{(f)}, \tag{3}$$

where $\tilde{D}_{j,k}^{(f)}$ and $\tilde{S}_{j,k}^{(f)}$ are respectively the detail and smooth components. $\tilde{D}_{j,k}^{(f)} = \sum_{l=0}^{N-1} \tilde{h}_{j,l} d_{j,k+l_{\text{mod}}N}$ and $\tilde{S}_{j,k}^{(f)} = \sum_{l=0}^{N-1} \tilde{g}_{j,l} c_{j,k+l_{\text{mod}}N}$. The sequences $\{\tilde{h}_{j,l}\}$ and $\{\tilde{g}_{j,l}\}$ are respectively, $\{h_{j,l}\}$ and $\{g_{j,l}\}$ periodised to the length N . Let $\mathcal{F}_{j,t}^{(f_1)}$ and $\mathcal{F}_{j,t}^{(f_2)}$ denote the information contained in the past observations of $\tilde{D}_{j,s}^{(f_1)}$ and $\tilde{D}_{j,s}^{(f_2)}$, respectively, $\forall s \leq t$. Assuming that details are strictly stationary, $f_1(t)$ is a cause of $f_2(t)$ at the scale j if, for some $h > 0$,

$$\tilde{D}_{j,t+1}^{(f_2)}, \dots, \tilde{D}_{j,t+h}^{(f_2)} | (\mathcal{F}_{j,t}^{(f_1)}, \mathcal{F}_{j,t}^{(f_2)}) \sim \tilde{D}_{j,t+1}^{(f_2)}, \dots, \tilde{D}_{j,t+h}^{(f_2)} | \mathcal{F}_{j,t}^{(f_2)}, \tag{4}$$

where \sim denotes the equivalence in distribution. At each scale j , the two alternative hypotheses are expressed as follows:

$$\begin{cases} H_0 : \tilde{D}_{j,t+1}^{(f_1)} \not\rightarrow \tilde{D}_{j,t+1}^{(f_2)} \\ H_a : \tilde{D}_{j,t+1}^{(f_1)} \rightarrow \tilde{D}_{j,t+1}^{(f_2)} \end{cases} \tag{5}$$

where the crossed out arrow in the null hypothesis means $\tilde{D}_{j,t+1}^{(f_1)}$ does not cause $\tilde{D}_{j,t+1}^{(f_2)}$. Similarly, inverting the arrows' directions implies that the causality is tested in the opposite direction.

3 | RESULTS

The objective of this experiment is to test the uni- and multi-scaled causalities between the number of COVID-19 infections and the related number of PCR viral tests in Tunisia. A daily database from the TMPH (Data on the epidemic situation are provided by the Tunisian Ministry of Public Health [<http://www.santetunisie.rns.tn/>] and the Tunisian observatory for new and emerging diseases [<http://www.onmne.tn/>],) is used in this study. By quickly reading the statistics of Table 1, we can notice that the two indices show almost the same characteristics. Asymmetry and excess of kurtosis (Jarque–Bera JB test), nonlinearity (Brock–Dechert–Scheinkman BDS test), and slight long memory (Geweke–Porter–Hudak GPH test), are the common facts for the two variables. These statistics are also confirmed by the histograms of the two variables as well as the cross-correlation function between the same variables (Figure 1). The significance of these statistics is a sign that a multiresolutional decomposition method can lead to a better appreciation of the interplay between the two variables.

The causality analysis between these two variables will in fact allow us to measure whether one of these variables have a significant effect on the other or not. Furthermore, the main interest of such a conjoint empirical study is to allow health authorities in Tunisia to prepare logistically against the arrival of new COVID-19 waves in order to optimise their needs in terms of PCR tests. In the other direction, it would also be possible for academia to re-predict the evolution of the number of cases according to the number of tests, especially in relationship with the credibility of the results announced by the authorities. To examine the causality assumption, we use the daily data ranging between

TABLE 1 Descriptive statistics and goodness-of-fit tests

Statistics	NCC	PCR
Minimum*	28	0
Maximum*	15,132	5752
Mean*	2.97×10^3	685.26
Coef. variation*	89.275%	126.46%
Kurtosis	5.8680	6.9836
Skewness	1.3720	1.6645
GPH* (\hat{d})	-0.0502	0.1667
JB test	1.56×10^3	6.77×10^3
ADF test	-32.922	-30.152
BDS test	12.785	10.498

Note: Statistics with asterisks are performed on level time series, while the remaining are obtained from differenced data. Abbreviations: NCC, number of confirmed cases; PCR, polymerase chain reaction.

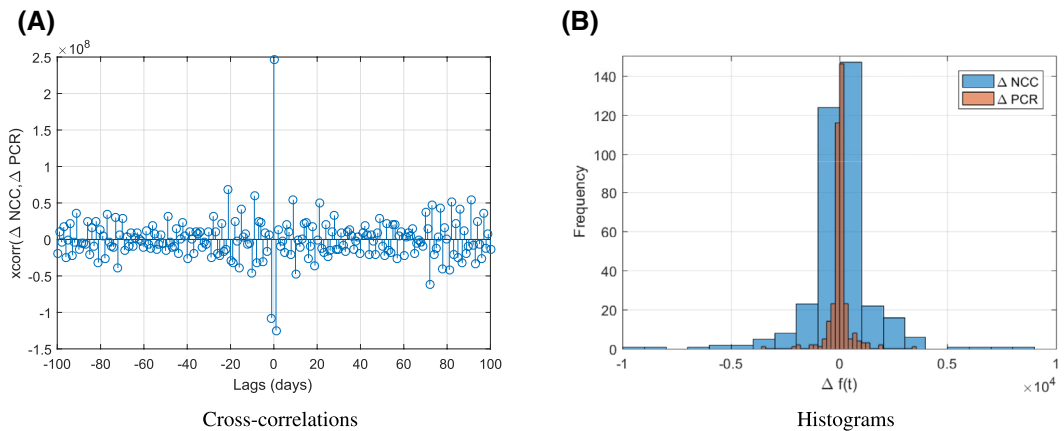


FIGURE 1 Cross-correlation and histogram plots of the daily changes ($\Delta X_t = X_t - X_{t-1}$) of COVID-19 confirmed cases and polymerase chain reaction (PCR) tests [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

March 2020 and March 2021 (see Figure 2). Three types of tests are conducted: the Granger test of causality,¹⁶ instantaneous test of causality,¹⁷ and multiscaled causality test.³ Since the simple principle of the multiscaled causality is to perform a scale-by-scale Granger test, we rely on the same rule as the classical test, which is to calculation an empirical information criterion to determine the order of autoregression lags.

The results reported in Table 2 recommend the rejection of the null hypothesis of no instantaneous causality in both directions, while the statistics of the Granger test suggest failing to reject the null hypothesis of non-causality also in both directions. The interpretation of these results is that, a forecasting model that uses current, past and future values of 'number of cases' (number of tests) and current and past values of 'number of tests' (number of cases) to predict the 'number of tests' (number of cases) has smaller forecast error than a model than only uses current and past values of 'number of cases' (number of tests) and current and past values of 'number of tests' (number of cases). The time series chart in Figure 2, however, shows signs of local stationarity and persistence, which could be interpreted as a type of long-range co-dependence.¹⁸ For this reason, the causality hypothesis is addressed in more depth by employing wavelets. As explained above, wavelets allow to restructure the causality test on ordered frequency

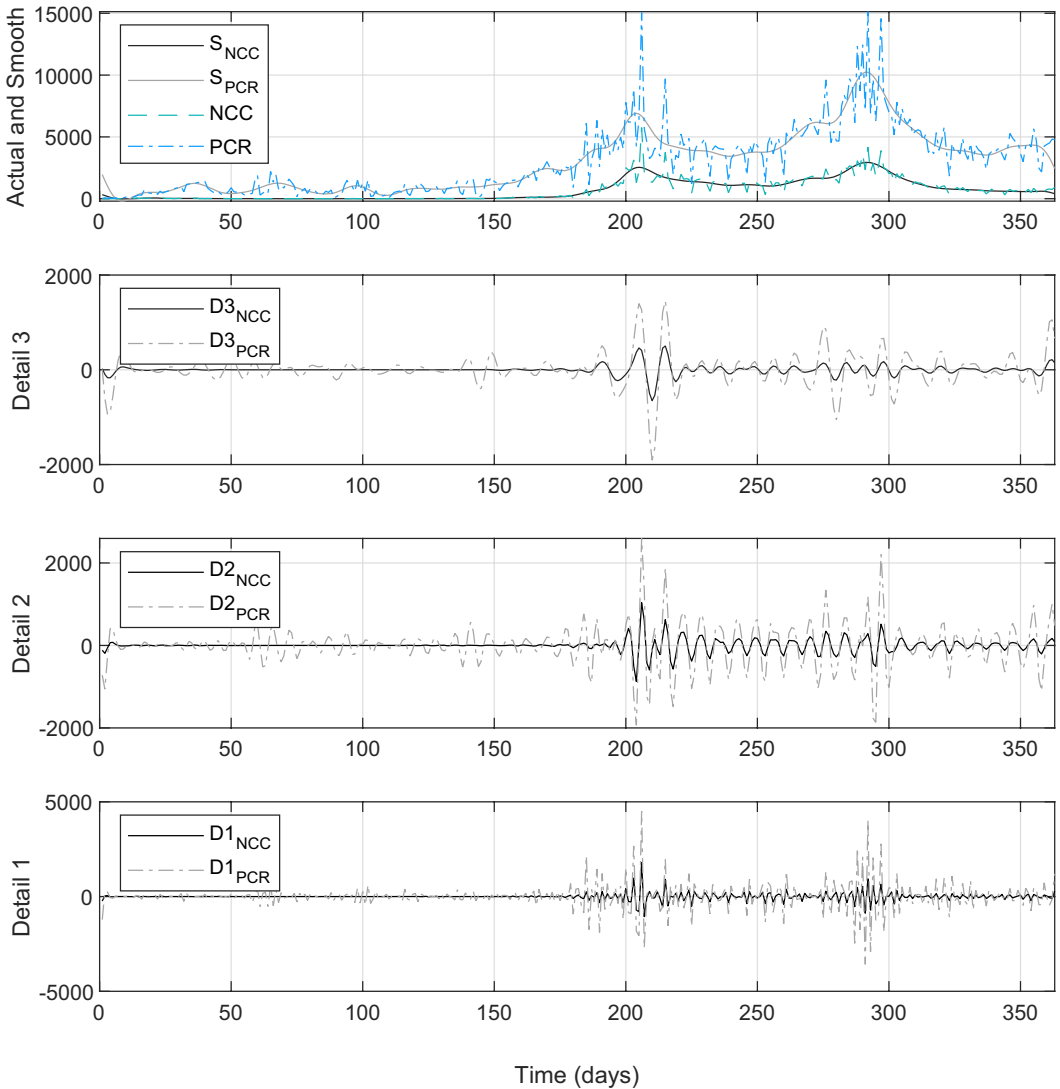


FIGURE 2 Daily number of COVID-19 confirmed cases and related polymerase chain reaction (PCR) tests in Tunisia from March 2020 to March 2021. Both actual (blues) and MODWT-decomposed (black and grey) time series are plotted. In the subplots, S and D_j correspond to the smooth components and details of level j , respectively. The subscripts NCC and PCR are abbreviations for the number of confirmed cases and performed PCR tests, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

sub-bands. As detailed in Table 3, on a decomposition level $J = 3$, the non-causal hypothesis is rejected in both directions at all scales. This implies that a bidirectional multiscaled forecasting model⁴ can be used with potential success.

Proving such a significant causal relationship between the number of infected cases and the number of performed PCR tests can in fact be one of the levers of a national COVID screening strategy, and therefore, a reliable informational advantage for an efficient public health logistics. This also means that it is possible to use data from the two factors to forecast the values of one of them. Forecasts that can be of great interest, especially if they are sufficiently accurate. Moreover, given the dynamics of the virus's progression and the NCC, showing several types of irregularities, it was crucial to use a sophisticated tool to better understand the relationship. It was for this reason that we have used wavelet techniques to also measure scale-by-scale causality. Recorded

TABLE 2 Causality analysis between the number of confirmed cases (NCC) and the number of polymerase chain reaction (PCR) tests

	F stat	p value		F stat	p value
Granger causality			Instant. causality		
NCC \rightarrow PCR	1.4492	0.2161	NCC \rightarrow PCR	189.231	0.0000
PCR \rightarrow NCC	1.6018	0.1721	PCR \rightarrow NCC	164.120	0.0000

Note: The notation $X \rightarrow Y$ corresponds to the causality null hypothesis H_0 : X does not cause Y . We reject the null hypothesis if the p value < 0.05

Abbreviations: NCC, number of confirmed cases; PCR, polymerase chain reaction.

TABLE 3 MODWT-based multiscaled causality. A Daubechies least asymmetric wavelet filter of length $L = 5$ is used for the MODWT with a decomposition level of $J = 3$

	F stat	Critical value		F stat	Critical value
$\tilde{D}_{1,t}^{(ncc)} \rightarrow \tilde{D}_{1,t}^{(pcr)}$	6.9094	3.0209	$\tilde{D}_{1,t}^{(pcr)} \rightarrow \tilde{D}_{1,t}^{(ncc)}$	4.3795	3.0211
$\tilde{D}_{2,t}^{(ncc)} \rightarrow \tilde{D}_{2,t}^{(pcr)}$	4.1604	3.8679	$\tilde{D}_{2,t}^{(pcr)} \rightarrow \tilde{D}_{2,t}^{(ncc)}$	4.9692	3.8676
$\tilde{D}_{3,t}^{(ncc)} \rightarrow \tilde{D}_{3,t}^{(pcr)}$	4.3765	3.8675	$\tilde{D}_{3,t}^{(pcr)} \rightarrow \tilde{D}_{3,t}^{(ncc)}$	4.6806	3.0210

Note: The notation $X \rightarrow Y$ corresponds to the causality null hypothesis H_0 : X does not cause Y . We reject the null hypothesis if the F -statistic is greater than the critical value.

statistics provided significant scale-by-scale causal relationships, showing the importance of building periodic short, medium, and long-term screening planning strategies. As a perspective for this work, it would be interesting, today, to introduce an additional variable representing the daily number of vaccinations. This would certainly allow to build a more up-to-date public health strategy with regard to COVID-19. It would be important, for instance, to make a similar reasoning to forecast the needs for vaccinations conditioned on a set of factors and constraints, such as the need and the will of the population, the storage conditions, the levels of spread of the virus, etc.

4 | CONCLUSION

A wavelet-based statistical analysis is conducted to measure the relationship between the daily COVID-19 infections in Tunisia and the viral tests allocated by the public health authorities. To continuously keep pace with the dynamics of the evolution of the virus during its first year, we used uninterrupted daily time series describing the two studied variables. The hypotheses claiming a relationship were therefore tested scale-by-scale. The question addressed was about the significance of the causal relationship between the NCC and the number of PCR tests conducted by the various national public and private health services. Most of the results recorded point to a significant two-way causal relationship between the number of infected cases and the number of tests. Such a time-scale co-evolution provides a clue to the Tunisian public health authorities in order to implement supply chain strategies and logistic plans to protecting themselves from future COVID-19 waves. In the future, the current analysis could be extended to also include vaccination data. A multivariate model would therefore allow to better understand the reaction of the spread of the virus in the population according to the achieved vaccination levels. Such a device could therefore represent the cornerstone of an adaptive vaccination plan.

ACKNOWLEDGMENT

We would like to thank the anonymous reviewers for their insightful and constructive comments that greatly contributed to improving the paper. Our many thanks go also to the editorial staff for their generous support and assistance during the review process. This work has not received any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ETHICS STATEMENT

Data were publicly available and did not include individual identities and thus did not require ethics approval.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available at the websites of the Tunisian Ministry of Public Health (<http://www.santetunisie.rns.tn/>) and the Tunisian Observatory for New and Emerging Diseases (<http://www.onmne.tn/>).

ORCID

Foued Saâdaoui  <https://orcid.org/0000-0002-0574-2922>

REFERENCES

1. Farmer RD, Emami J. Models for forecasting hospital bed requirements in the acute sector. *J Epidemiol Community Health*. 1990;44:307-312.
2. Rabbouch H, Sadaoui F, Mraïhi R. A vision-based statistical methodology for automatically modeling continuous urban traffic flows. *Adv Eng Inf*. 2018;38:392-403.
3. Saâdaoui F. The price and trading volume dynamics relationship in the EEX power market: a wavelet modeling. *Comput Econ*. 2013;42(1):47-69.
4. Saâdaoui F, Rabbouch H. A wavelet-based multiscale vector-ANN model to predict comovement of econophysical systems. *Expert Syst Appl*. 2014;41(13):6017-6028.
5. Saâdaoui F, Bertrand PR, Boudet G, Rouffiac K, Dutheil F, Chamoux A. A dimensionally reduced clustering methodology for heterogeneous occupational medicine data mining. *IEEE Trans NanoBioscience*. 2015;14(7):707-715.
6. Hesham F, Harizi R, Nasr KS. What have we learned about the effects of the COVID-19 pandemic on consumer behavior? *Sustainability*. 2021;13(8):4304.
7. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. *Nat Rev Microbiol*. 2021;19:141-154.
8. Ordu M, Akin HK, Demir E. Healthcare systems and Covid19: lessons to be learnt from efficient countries. *Int J Health Plan Manag*. 2021;36(5):1476-1485.
9. Gulec EY, Cesur NP, Fazlioglu GY, Kazezoglu C. Effect of different storage conditions on COVID-19 RT-PCR results. *J Med Virol*. 2021;93(12):6575-6581.
10. Musalam AMA, Abunamous RES, El-Ashgar NM, Alhana AS. Emergency management to avoid the spread of the COVID-19 pandemic in Gaza Strip. *Int J Health Plan Manag*. 2021;36(5):1742-1757.
11. Liu M, Ning J, Du Y, et al. Modelling the evolution trajectory of COVID-19 in Wuhan, China: experience and suggestions. *Public Health*. 2020;183:76-80.
12. Triacca M, Triacca U. Forecasting the number of confirmed new cases of COVID-19 in Italy for the period from 19 May to 2 June 2020. *Infect Dis Model*. 2021;6:362-369.
13. Arfaoui S, Mabrouk AB, Cattani C. *Wavelet Analysis: Basic Concepts and Applications*. 1st ed. Chapman and Hall/CRC; 2021.
14. Khalfa N, Bertrand PR, Boudet G, Chamoux A, Billat V. Heart rate regulation processed through wavelet analysis and change detection. Some case studies. *Acta Biotheor*. 2012;60(1-2):109-129.
15. Sonoo M, Kanbayashi T, Idogawa M, Shimohata T, Hayashi H. Correlation between polymerase chain reaction (PCR) examination rate among the population and the containment of the COVID-19 pandemic. *Public Health*. 2021;191:31-32.
16. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 1969;37(3):424-438.

17. Gianetto QG, Rassi H. Testing instantaneous causality in presence of nonconstant unconditional covariance. *J Bus Econ Stat.* 2015;33(1):46-53.
18. Omay T, Baleanu D. Fractional unit-root tests allowing for a fractional frequency flexible Fourier form trend: predictability of Covid-19. *Adv Differ Equ.* 2021. Article number: 167.

How to cite this article: Saâdaoui F, Rabbouch H, Saadaoui H, Dutheil F. Multiscaled causality of infections on viral testing volumes: the case of COVID-19 in Tunisia. *Int J Health Plann Mgmt.* 2022;37(3):1838-1846. <https://doi.org/10.1002/hpm.3427>