APPLICATION OF CROSS-WAVELET AND SINGULAR VALUE DECOMPOSITION ON COVID-19 AND BIO-PHYSICAL DATA

Iftikhar U. Sikder[1] and James J. Ribero[2]

[1]Department of Information Syatems, Cleveland State University, USA
[2]IBA, University of Dhaka, Bangladesh

ABSTRACT

The paper examines the bivariate relationship between COVID-19 and temperature time series using Singular Value Decomposition (SVD) and continuous cross-wavelet analysis. The COVID-19 incidence data and the temperature data of the corresponding period were transformed using SVD into significant eigen-state vectors for each spatial unit. Wavelet transformation was performed to analyze and compare the frequency structure of the single and the bivariate time series. The result provides coherency measures in the ranges of time period for the corresponding spatial units. Additionally, wavelet power spectrum and paired wavelet coherence statistics and phase difference were estimated. The result suggests statistically significant coherency at various frequencies. It also indicates complex conjugate dynamic relationships in terms phases and phase differences.

KEYWORDS

COVID-19, SVD, Wavelet analysis, Cross-wavelet power, Wavelet coherence.

1. INTRODUCTION

The outbreak of COVID-19 has significantly changed the landscape of the global health. However, the dynamics between the bio-physical or climatic variables and the diffusion of COVID-19 is poorly understood [1]-[3]. There are various claims with regards to the dependency between the incidence or prevalence and environmental variables. It has often been argued that lower (cold) temperature act as a catalyst in significantly increasing the spread of COVID-19 [4]-[5]. There also exist alternative claims that warm temperatures slow down the spread of COVID-19 [6]. In contrast to these claims, some scholars assert that temperature does not play any role in the spread of COVID-19 [7]. In this paper, we have examined some specific empirical relationships of such dependencies, namely wavelet coherence and its statistical significance, phases and phase differences using the dataset of the USA.

2. OBJECTIVES AND SCOPE

The primary objective of the paper is to characterize the dynamic relationship of COVID-19 and a time-variant bio-physical parameter namely the temperature. This paper aims to provide an empirical investigation that captures and analyzes the characteristic relationship of these variables. The study area was limited within the United States. The data for COVID-19 cases was collected from the fifty (50) states, and the corresponding data on temperature of the same period was collected from these states. The period covered was between Jan. 21, 2020, till date. Around

David C. Wyld et al. (Eds): EMSA, SEA, AIFU, NLCAI, BDML, BIoT, NCOM, CLOUD, CCSEA, SIPRO - 2022
pp. 121-126, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.120612

40,000 records (20000 COVID-19 data records and 20000 temporal temperature data records) have been collected and used for the research [8]-[9].

3. METHODOLOGY

The variables used in the model are featured as time series data, and thus expected to fluctuate with an associated noise. Employing conventional smoothing technique involving amplitude-based statistical analysis would not be appropriate to achieve the research objective. Therefore, we adopt a Wavelet Transform algorithm not only to capture the periodicities of the variables over the time, but also to establish coherence among the variables in the frequency domain.

3.1. Modeling Framework

The Figure 1 below details the process flow of the research:

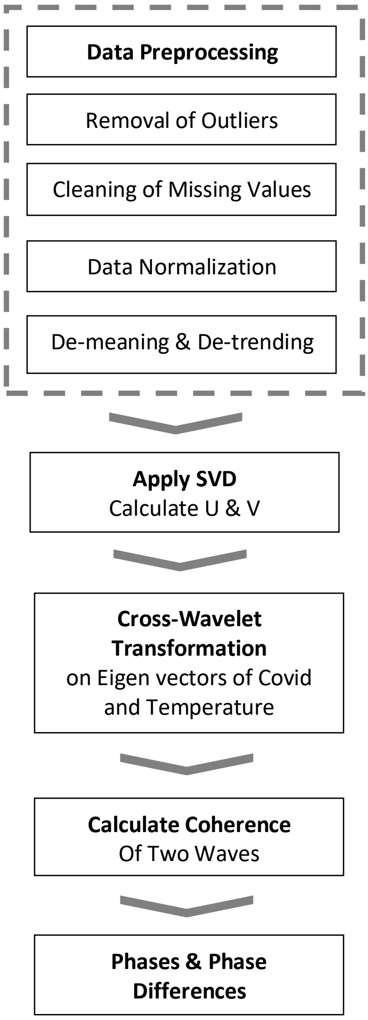


Figure 1. Methodology flow chart

3.1.1. Data Pre-Processing

The data sets on COVID-19 cases and temperature across the 50 states accessed in a format that was not readily available for analysis. The pre-processing steps involved the basic data cleaning functions such as removal of irrelevant attributes and missing values, removal of outliers, de-meaning, linear detrending, and data normalization. The data processing was done in R environment using *WaveletComp* package [10]. The outputs of the pre-processing are data frames that were transformed into rectangular matrices, where the rows represent either COVID-19 cases by states or temperature, and the columns represent the date.

3.1.2. Calculating Singular Value Decomposition (SVD)

A widely adopted matrix factorization or dimension reduction technique namely Singular Value Decomposition (SVD) was applied on both the COVID-19 and temperature data sets to compress the data into ortho normal eigen basis to rectangular matrices. Based on the top singular values, top three eigen states for both COVID-19 and temperature were selected, which in combination accounts for significant total variance of the original data. Plots of the transformed data set are displayed in Figure 2 below showing negative correlation (with *correlation coefficient* of -0.34) of cases and temperature.

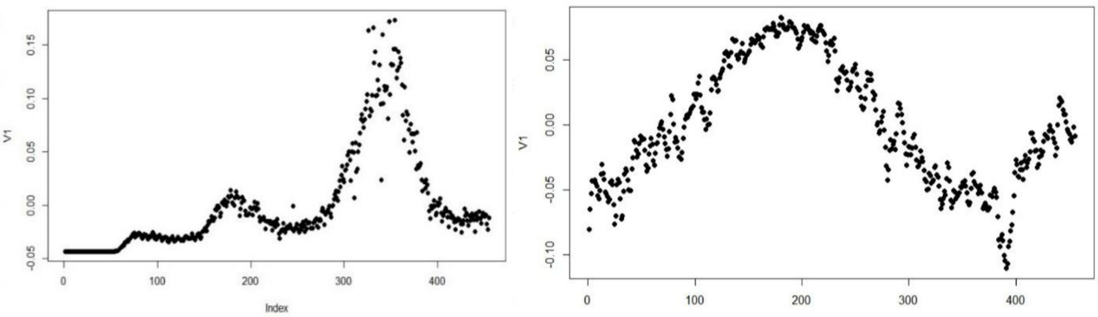


Figure 2. Significant eigenvectors of temperature and Covid-19 cases

3.1.3. Calculating Wavelet Transformation

A Wavelet transform decomposes a time series into a set off wavelets localized in time. The Wavelet transformation was performed on both COVID-19 and temperature time-series of *eigen* vectors. We applied *Morlet* continuous Wavelet transforms to the transformed data using the *WaveletComp* R package. Wavelet transformation leads to a continuous, complex-valued output of the time series that preserves both time and frequency resolution parameters. The transform is separable into its real part and imaginary part providing information on both local amplitude and instantaneous phase. This allows for the investigation of coherency between the two time series. Given two time series $X(t)$ and $Y(t)$, and corresponding *wavelet spectrums* $W_x(s,\tau)$ and $W_y(s,\tau)$ which could be considered as localized energy spectrum varying with scale s, and translation τ, and associated frequency ω and time t. The cross-wavelet transformation $W_{xy}(s,\tau)$ is associated with complex-valued wavelet coherency:

$$\Upsilon(s,\tau) = \frac{\langle W_{xy}(s,\tau)\rangle}{\sqrt{\langle W_x(s,\tau)\rangle * \langle W_y(s,\tau)\rangle}} \qquad (1)$$

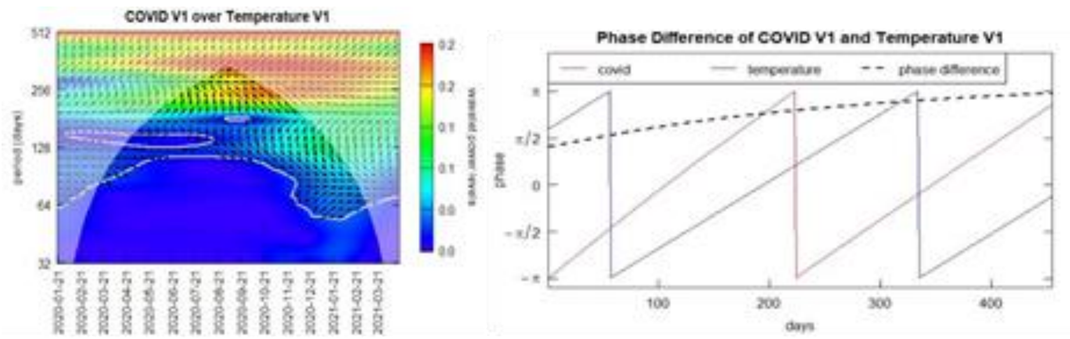
and the normalizing wavelet power spectra coherence is:

$$\Upsilon^2(s,\tau) = \frac{\left[\langle Re(W_{xy}(s,\tau))\rangle\right]^2 + \left[\langle Im(W_{xy}(s,\tau))\rangle\right]^2}{\langle W_x(s,\tau)\rangle * \langle W_x(s,\tau)\rangle} \qquad (2)$$

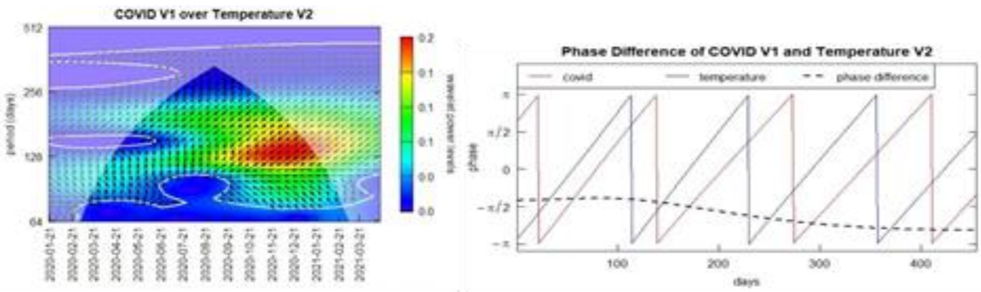
The angle brackets ⟨ ⟩ denotes the smoothing operation over time. The squaring of the amplitude component gives us the wavelet power spectrum $0 \leq \Upsilon^2(s,\tau) \leq 1$, which is some what analogous to conventional correlation coefficient. After computing the wavelet power spectrum for each of *eigenvector*, we analyze the coherence of the paired waves of COVID-19 and temperature using the coherence function. The phase lags between the variables were also computed. The cross-wavelet transformation provided cross-magnitude, phase differences, non-stationarity, and coherency between signals. Using these results of the cross-wavelet transformation, a series' synchronicity at certain periods and across certain ranges of time was analyzed.

4. RESULTS AND MODEL INTERPRETATION

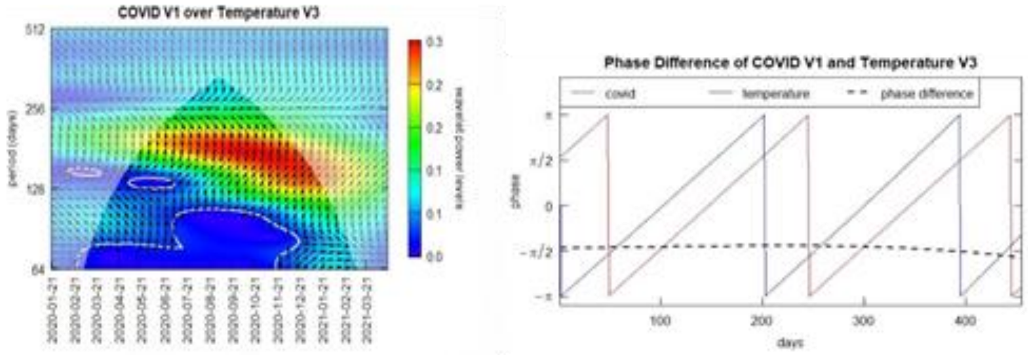
The cross-wavelet analysis generated coherence plot that shows that that there is a coherence (correlation) between COVID-19 and temperature, and these relationships are statistically significant (the region enveloped or bounded by the white line). The phases and phase differences show varied results. Figures 3a, 3b, 3c shows COVID-19 and temperature are out of phase with varying phase lags while Figure 3d and 3e shows that are in phase. Comparing the result of plots 3a and 3e, COVID-19 and temperature were both out of phase in 3a, with temperature leading and COVID-19 lagging by 96days, while from 3e both time series were in phase. Though temperature was leading, the lag period was much narrow (around 5 days) compared to 3a.



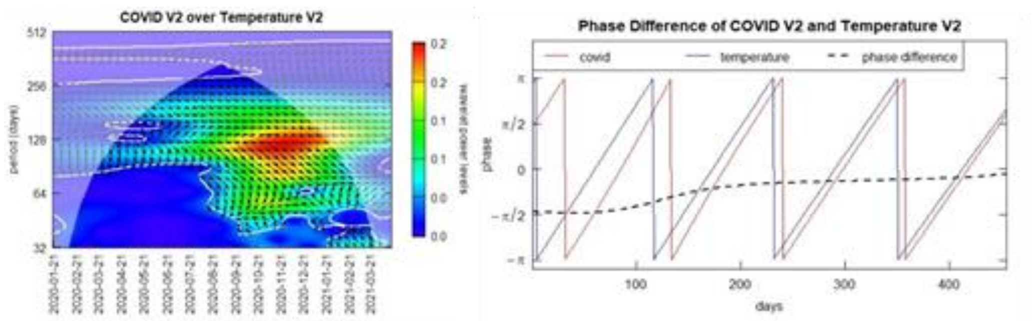
3a. COVID-19 V1 & Temperature V1



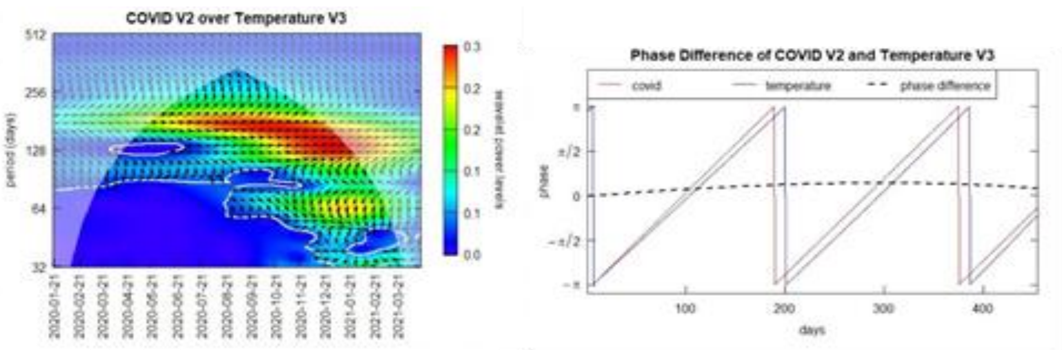
3b. COVID-19 V1 & Temperature V2



3c. COVID-19 V1 & Temperature V3



3d. COVID-19 V2 & Temperature V2



3e. COVID-19 V2 & Temperature V3

Figure 3. Cross-wavelet analysis generated coherence plots

5. CONCLUSION

This article studies the dynamics between two-time series variables – COVID-19 and temperature- using SVD and wavelet transform approach. The results from the continuous cross-wavelet transform shows power spectrum strengths and coherence corresponding at various frequencies (periods). The coherence statistics suggest statistically significant relationship. The results also show varying phases and phase lags with leading and lagging behavior showing complex conjugate dynamics. Future studies focusing on spatially explicit mapping of coherence could provide additional explanatory schemes and better understanding of the spatio-temporal dynamics of the disease.

REFERENCES

[1] Sahoo, P.K., Powell, M.A., Mittal, S., and Garg, V. (2020) “Is the transmission of novel coronavirus disease (COVID-19) weather dependent?”, *Journal of the Air & Waste Management Association*, , Vol. 70, (11), pp. 1061-1064

[2] Kroumpouzos, G., Gupta, M., Jafferany, M., Lotti, T., Sadoughifar, R., Sitkowska, Z., and Goldust, M. (2020) “COVID-19: A relationship to climate and environmental conditions?”, *Dermatologic therapy*

[3] Rosario, D.K., Mutz, Y.S., Bernardes, P.C., and Conte-Junior, C.A. (2020) “Relationship between COVID-19 and weather: Case study in a tropical country”, *International journal of hygiene and environmental health*, Vol. 229, pp. 113587

[4] Wang, M., Jiang, A., Gong, L., Luo, L., Guo, W., Li, C., Zheng, J., Li, C., Yang, B., and Zeng, J. (2020) “Temperature significant change COVID-19 Transmission in 429 cities”, *MedRxiv*

[5] Prata, D.N., Rodrigues, W., and Bermejo, P.H. (2020) “Temperature significantly changes COVID-19 transmission in (sub) tropical cities of Brazil”, *Science of the Total Environment*, , 729, pp. 138862

[6] Oliveiros, B., Caramelo, L., Ferreira, N.C., and Caramelo, F. (2020) “Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases”, *MedRxiv*

[7] Jamil, T., Alam, I., Gojobori, T., and Duarte, C.M. (2020) “No evidence for temperature-dependence of the COVID-19 epidemic”, *Frontiers in public health*, , 8, pp. 436

[8] COVID-19 Dataset. *https://www.kaggle.com/joelhanson/coronavirus-COVID-19-data-in-the-united-states?select=us-states.csv* [Accessed on 1st September 2021]

[9] Temporal Temperature Dataset. *https://catalog.data.gov/dataset* [Accessed on 1st September 2021]

[10] Roesch, A., Schmidbauer, H., and Roesch, M.A. (2014) “Package ‘WaveletComp”, *The Comprehensive R Archive Network*

AUTHORS

Iftikhar U. Sikder is an Associate Professor jointly appointed in the department of Information Systems and the Electrical Engineering and Computer Science at Cleveland State University, USA. He holds a PhD in Computer Information Systems from the University of Maryland, Baltimore. His research interests include soft computing, granular computing, spatial databases, spatio-temporal data mining and adaptive complex systems. His papers appeared in the journal of Knowledge-Based Systems, Risk Analysis, Expert Systems with Applications, Intl. journal of Digital Earth, International Journal of Mobile Communications, and Information Resources Management Journal, Intl. Journal of Management & Decision Making, and Intl. Journal of Aerospace Survey and Earth Sciences and many others journals. He has authored many book chapters and presented papers in many national and international conferences. Dr. Sikder is also currently serving in the editorial board of International Journal of Computational Models and Algorithms in Medicine and Intl. Journal of Computers in Clinical Practice.

James J. Ribero is an Adjunct Faculty at IBA, Dhaka University, Bangladesh. He holds MBBS, MS (Microbiology) and MBA from Dhaka University, Bangladesh. His research interests include applications of Machine Learning and Big Data technologies into medical and life sciences. He has authored book chapters, monographs, and presented papers in many national and international conferences. He is an ex-executive editor of *The Orion* medical journal.

© 2022 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.