

## Article

# Inferring a Causal Relationship between Environmental Factors and Respiratory Infections Using Convergent Cross-Mapping

Daipeng Chen <sup>1,2</sup> , Xiaodan Sun <sup>1,\*</sup> and Robert A. Cheke <sup>3</sup> 

<sup>1</sup> School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China; chendaipeng@stu.xjtu.edu.cn

<sup>2</sup> Mathematical Institute, Leiden University, 2333 CA Leiden, The Netherlands

<sup>3</sup> Natural Resources Institute, University of Greenwich at Medway, Central Avenue, Chatham Maritime, Chatham ME4 4TB, Kent, UK; robert.cheke@btinternet.com

\* Correspondence: xiaodansun@mail.xjtu.edu.cn

**Abstract:** The incidence of respiratory infections in the population is related to many factors, among which environmental factors such as air quality, temperature, and humidity have attracted much attention. In particular, air pollution has caused widespread discomfort and concern in developing countries. Although the correlation between respiratory infections and air pollution is well known, establishing causality between them remains elusive. In this study, by conducting theoretical analysis, we updated the procedure of performing the extended convergent cross-mapping (CCM, a method of causal inference) to infer the causality between periodic variables. Consistently, we validated this new procedure on the synthetic data generated by a mathematical model. For real data in Shaanxi province of China in the period of 1 January 2010 to 15 November 2016, we first confirmed that the refined method is applicable by investigating the periodicity of influenza-like illness cases, an air quality index, temperature, and humidity through wavelet analysis. We next illustrated that air quality (quantified by AQI), temperature, and humidity affect the daily influenza-like illness cases, and, in particular, the respiratory infection cases increased progressively with increased AQI with a time delay of 11 days.

**Keywords:** environmental factors; respiratory infection; nonlinear system; causality



**Citation:** Chen, D.; Sun, X.; Cheke, R.A. Inferring a Causal Relationship between Environmental Factors and Respiratory Infections Using Convergent Cross-Mapping. *Entropy* **2023**, *25*, 807. <https://doi.org/10.3390/e25050807>

Academic Editors: Hector Zenil, Jiang Zhang and Peng Cui

Received: 24 April 2023

Revised: 12 May 2023

Accepted: 13 May 2023

Published: 17 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

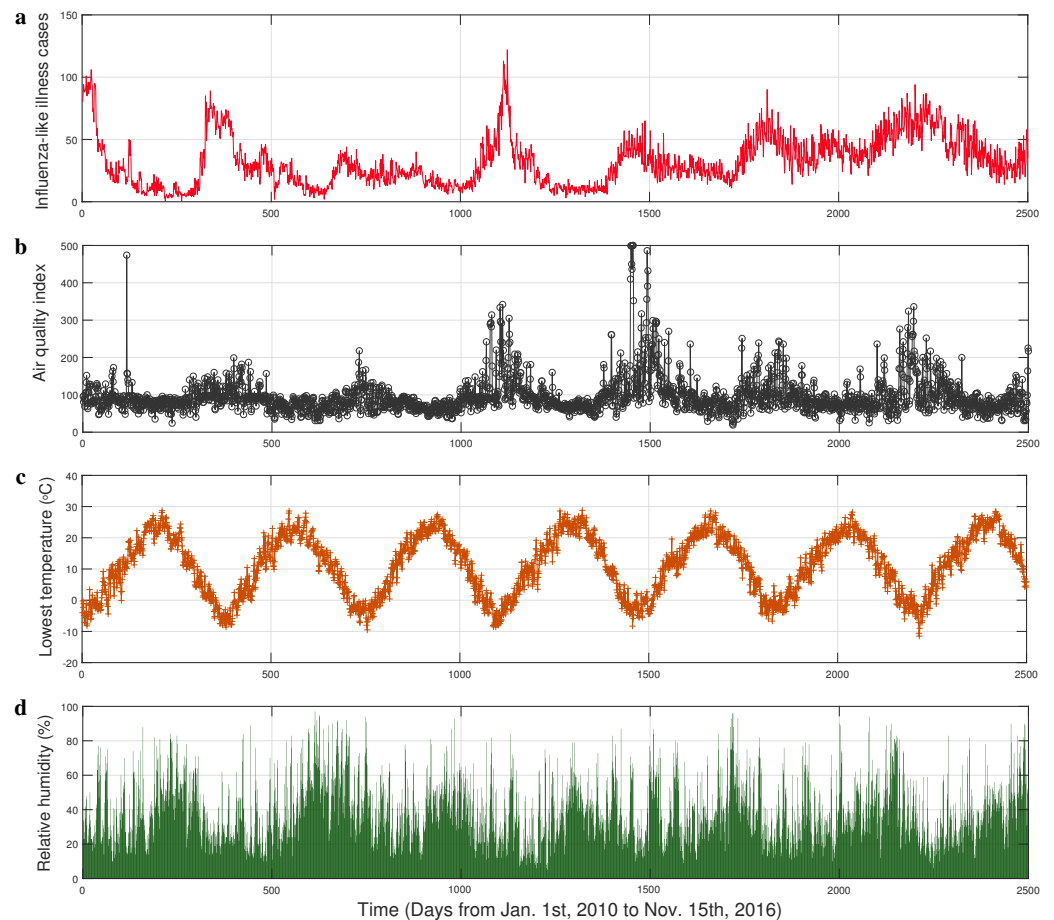
Understanding the driving force of infectious disease spread is critical to designing effective interventions curbing the threat of diseases transmission to public health around the world [1]. In tracking driving factors of infectious disease spread, examining the relationship between environmental changes and disease transmission is one of the most important topics [2–5]. Previous works have established that monsoon rains and temperature affect the epidemiology of cholera [6] and the life cycles of vectors such as mosquitoes, and the parasites that they transmit, so they are important environmental drivers of malaria [7], dengue [8] and Ross River fever [9]. Moreover, regional temperature and humidity are also related to influenza transmissibility [10,11]. Recently, there is an increasing recognition that poorer air quality is synchronized with a higher incidence of infectious diseases [12]. For example, air pollution is associated with an increased risk of tuberculosis [13,14], influenza [15–17], influenza-like illness [18,19] and COVID-19 [20,21]. Unlike the vector-transmitted diseases [7–9] with clear biological evidence, however, the driving effect of environments on respiratory infection is still controversial, at least partly because two unlinked variables in complex systems may have significant correlations [22]. To design interventions which can reduce infection risk effectively, it is of great importance to infer or falsify causal links between environmental factors and respiratory infection.

Recent studies have revealed a broad correlation between respiratory infections and environmental factors during climate change [23]. While correlation does not imply causal-

ity [24], correlated variables may potentially share information in a complex system and increase the complexity of this system. To address the issue that a system is too complex to be parameterized, researchers have developed a nonparametric framework called empirical dynamic modeling (EDM) that is designed to analyze complex systems using observed time series [25–29]. In the EDM methods, convergent cross-mapping (CCM) is specifically used to detect causal relationships in complex systems [28]. This approach is based on the mathematical theory of reconstructing attractor manifolds [30,31], that is, in the homeomorphic sense, the attractor of a dynamical system can be reconstructed from the time series of a single observed variable of this system. Therefore, the reconstructed manifolds of two bidirectionally coupled variables are homeomorphic [30]. For two unidirectionally coupled variables, the reconstructed manifold of the response variable is homeomorphic to the original attractor, while the reconstructed manifold of the driving variable is only a subset of the original attractor and is therefore a subset of the reconstructed manifold of the response variable [31]. Based on these consequences, Sugihara et al. [28] developed convergent cross-mapping (CCM) to predict the points in one reconstructed manifold using the points in another reconstructed manifold to then infer/falsify a causal relationship between two variables through the accuracy of predictions (CCM skill). To date, CCM has been used to analyze the causality involved in a prey–predator system [28], Earth system [32], locust abundance [33], and deep-sea biodiversity [34].

Although CCM has been widely applied to infer causality between variables with weak to moderate coupling strengths [28,32–34], Sugihara et al. [28,35] found that it is very easy to infer two variables with an unidirectionally strong coupling as having two-way causality because of the phenomenon of “generalized synchrony” [36]. To resolve this problem, Ye et al. [35] extended the CCM by considering the time delay between interacting variables. In contrast to the CCM where the predicted values of one of two variables is based on the values of another with the same time label, the extended CCM takes into account that the value of the driving variable should be more suitable for predicting the future values of the response variable and that the response variable is better at predicting the past values of the driving variable [35]. The time delay with optimal prediction suggests a causal relationship between two variables and gives an estimate of the interaction lag [35]. The biggest challenge for extended CCM is that the optimal time delay is not unique when the observation data present periodicity [37]. For respiratory infections such as influenza [38] and influenza-like illness (Figure 1a), an annual cycle is one of the most obvious characteristics. Thus, extended CCM appears to be limited by its potential inability to infer causality between respiratory infections and environmental factors.

Here, we theoretically demonstrated that while the time lags that make the CCM skill locally optimal are not unique, they occur periodically, and the period has a lower bound. This inspired the notion that the time delay with optimal prediction is unique in a narrow testing window whose width does not exceed the lower bound. The numerical analysis on a mathematical model is consistent with our theoretical analysis. For real data in Xi’an, our analysis shows that air quality, temperature, and humidity are driving factors of respiratory infection with different time delays, and suggests that the interventions such as improving air quality and appropriately increasing temperature as well as humidity could reduce the respiratory infection risk in the population.



**Figure 1.** The time series of influenza-like illness (ILI) cases and experimental factors in Xi’an. **(a)** The influenza-like illness (ILI) cases collected from seven hospitals. **(b)** The real time air quality index (AQI). **(c)** The lowest daily temperature. **(d)** The relative humidity.

## 2. Materials and Methods

### 2.1. The Data

The data of respiratory infections (Figure 1a) we used consist of reports of daily cases who seek medical attention with influenza-like illness (ILI) (symptoms commonly include fever, shivering, chills, malaise, dry cough, loss of appetite, body aches, and nausea) in Xi’an from 1 January 2010 to 15 November 2016 [19]. Air pollution is a mixture of multiple pollutants, so an air quality index (AQI) is used by government agencies to communicate to the public the pollution levels of the air. [39]. The AQI in Xi’an (Figure 1b) was collected from the website [40] which is an open platform for weather data. The time series of temperature and relative humidity (Figure 1c,d) were downloaded from the shared portal [41] of the China Meteorological Administration.

An alternative test for causality is the Granger test [42], but this method is inappropriate for nonlinear dynamic systems [43], so the convergent cross-mapping was proposed by Sugihara et al. [28]. One of the limitations of CCM is that CCM is sensitive to high levels of process noise in the data [44]. In order to reduce the noise level, we split the time series of collected data into low-frequency series and residuals (Figure A1a) using Kalman filtering [45]. Since the standardized residuals follow the normal distribution (Figure A1b), we assumed that the main information in real time series was included in the filtered low-frequency series which is used in the following analysis.

### 2.2. The Method

The general dynamical system actually corresponds to a complex causal network of interlocking variables. We apply extended convergent cross-mapping (CCM) [28,35] to examine the causality and interaction delay between two variables. We let  $x(t)$  and  $y(t)$  be the observed time series corresponding to the variables  $x$  and  $y$ , respectively, and begin by reconstructing the lagged-coordinate vectors  $X(t) = [x(t), x(t - s), \dots, x(t - (\Pi - 1)s)]$  and  $Y(t) = [y(t), y(t - s), \dots, y(t - (\Pi - 1)s)]$  with dimension  $\Pi$  based on the Takens embedding theorem [30,31]. We denote the reconstructed manifolds as  $M_x = \{X(t)\}$  and  $M_y = \{Y(t)\}$ . For any point  $Y(t_*)$  in  $M_y$ , we mark the time label of its  $\Pi + 1$  nearest neighbors in  $M_y$  as  $t_1, t_2, \dots, t_{\Pi+1}$ , and the estimated point  $\bar{X}(t_* + \tau)$  with some delay  $\tau$  on  $M_x$  is given by the simplex projection [25]:

$$\bar{X}(t_* + \tau) = \sum_{i=1}^{\Pi+1} \chi^i X(t_i + \tau), \tag{1}$$

where

$$\chi^i = \frac{\exp(-\|(Y(t_i) - Y(t_*))\| / \|Y(t_1) - Y(t_*)\|)}{\sum_{i=1}^{\Pi+1} \exp(-\|(Y(t_i) - Y(t_*))\| / \|Y(t_1) - Y(t_*)\|)}.$$

We denote  $\bar{x}(t | \tau)$ , the first coordinate of  $\bar{X}(t)$ , as the estimated value of time series  $x(t)$  using this method. We use the Pearson correlation coefficient  $\rho(\tau) = \text{Corr}(\bar{x}(t | \tau), x(t))$  between estimated values and observed values to quantify CCM skill. From the perspective of the extended CCM [35], the CCM skill  $\rho(\tau)$  reaching its maximum at a negative  $\tau$  means that there is a driving force from  $x$  to  $y$  with time delay  $|\tau|$ .

Given a  $T$ -periodic observed time series  $x(t)$ , the reconstructed manifold  $M_x$  is a closed orbit in the  $\Pi$ -dimensional space. According to the Whitney embedding theorem [46],  $\Pi = 3$  is sufficient to ensure that all information of the original complex system is represented in the periodic orbit  $M_x$ . During any small period  $[t, t + \Delta t]$ , the length  $\Delta l(t)$  of the corresponding small arc on  $M_x$  can be approximated by

$$\Delta l(t) = \sqrt{x'(t)^2 + x'(t - s)^2 + x'(t - 2s)^2} \Delta t. \tag{2}$$

The prediction using simplex projection (1) mainly depends on the local distance between points [25] on the reconstructed manifold  $M_x$ . Therefore, the characters of  $\Delta l(t)$  affect the evaluation accuracy of extended CCM. If  $s$  is very small (i.e., the data are collected densely), then  $\Delta l(t) \approx \sqrt{3} \|x'(t)\| \Delta t$ . To investigate the mathematical characters of  $\Delta l(t)$  (i.e.,  $\|x'(t)\|$ ), as the first step, we give the following proposition.

**Proposition 1.** *A continuous, periodic, and nonconstant function  $x(t)$  has a smallest (positive) period  $T$ . For any other period  $\tilde{T}$  of  $x(t)$ , there is an integer  $n$  such that  $\tilde{T} = nT$ .*

**Proof.** Suppose there is no smallest (positive) period, then there is a decreasing (positive) sequence  $T_1, T_2, \dots, T_n, \dots$  of periods such that  $\lim_{n \rightarrow \infty} T_n = 0$ . For any given  $t_0$  and  $t_1$  ( $t_0 < t_1$ ), we define a sequence of integers:

$$z_1 = \lceil \frac{t_1 - t_0}{T_1} \rceil, z_2 = \lceil \frac{t_1 - t_0}{T_2} \rceil, \dots, z_n = \lceil \frac{t_1 - t_0}{T_n} \rceil, \dots$$

then the sequence  $t_1, t_{n+1} = t_n - z_n T_n$  ( $n = 1, 2, \dots$ ) such that

$$x(t_n) = x(t_{n+1}) \text{ and } \lim_{n \rightarrow \infty} t_n = t_0.$$

By continuity of  $x(t)$ ,  $x(t_1) = \lim_{n \rightarrow \infty} x(t_n) = x(t_0)$ , which contradicts that  $x(t)$  is a non-constant function. Consequently, there is a smallest (positive) period  $T$  of  $x(t)$ .

For any other period  $\tilde{T}$  of  $x(t)$ , there always is an integer  $k$  such that  $kT < \tilde{T} \leq (k + 1)T$ . Consequently,

$$0 < \tilde{T} - kT \leq T.$$

If  $\tilde{T} \neq (k + 1)T$ , then  $\tilde{T} - kT$  is a new smallest (positive) period of  $x(t)$  because

$$x(t + \tilde{T} - kT) = x(t).$$

Therefore, there is an integer  $n = k + 1$  such that  $\tilde{T} = nT$  for any other period  $\tilde{T}$  of  $x(t)$ .  $\square$

Based on the result of Proposition 1, we further draw the following proposition about the derivative of a nonconstant periodic function.

**Proposition 2.** *For any smooth periodic function  $x(t)$ , the derivative  $x'(t)$  is a periodic function and minimum period of  $x'(t)$  is the minimum period  $T$  of the original function.*

**Proof.** On the one hand,

$$x'(t) = \lim_{\Delta t \rightarrow 0} \frac{x(t + \Delta t) - x(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{x(t + T + \Delta t) - x(t + T)}{\Delta t} = x'(t + T).$$

Therefore,  $x'(t)$  is a periodic function. Suppose the minimum period of derivative  $x'(t)$  is  $\tilde{T}$ , and then Proposition 1 implies that there is a positive integer  $k_1$  such that  $T = k_1\tilde{T}$ . On the other hand,

$$x'(t) = x'(t + \tilde{T})$$

gives

$$\int_t^{t+\tilde{T}} x'(s)ds = \int_t^{t+\tilde{T}} x'(s + \tilde{T})ds = \int_{t+\tilde{T}}^{t+2\tilde{T}} x'(s)ds.$$

Consequently,  $x(t + \tilde{T}) = x(t) + C$  for some constant  $C$ . Because function  $x(t)$  is periodic, we must have  $C = 0$ . Therefore, there is a positive integer  $k_2$  such that  $\tilde{T} = k_2T$ . Collectively,  $k_1 = k_2 = 1$ .  $\square$

Finally, we give a proposition which presents the mathematical characters of the length  $\Delta l(t)$  (i.e.,  $\|x'(t)\|$ ) of the corresponding small arc on the reconstructed manifold  $M_x$  from time series  $x(t)$ .

**Proposition 3.** *Given a smooth function  $x(t)$  with minimum period  $T$ , suppose that the function has at most  $k$  extreme points in a single period, then the function  $\|x'(t)\|$  is periodic and the minimum period of  $\|x'(t)\|$  is not less than  $T/k$ .*

**Proof.** The result of Proposition 2 yields that

$$\|x'(t + T)\| = \|x'(t)\|.$$

This shows that the function  $\|x'(t)\|$  is periodic. Suppose the minimum period of function  $\|x'(t)\|$  is  $\tilde{T}$ , and then by Proposition 1 that there is a positive integer  $n$  such that  $T = n\tilde{T}$ . For any period  $[t, t + T]$ , we have

$$\int_t^{t+T} x'(s)ds = x(t + T) - x(t) = 0.$$

Therefore, there is a  $t_0 \in [t, t + T]$  such that  $x'(t_0) = 0$ , and then  $\|x'(t_0)\| = 0$ . Without loss of generality, we assume that  $t_0 - t < \tilde{T}$ . Consequently, we have  $t_0, t_0 + \tilde{T}, t_0 + 2\tilde{T}, \dots, t_0 + (n - 1)\tilde{T} \in [t, t + T]$  such that

$$\|x'(t_0)\| = \|x'(t_0 + \tilde{T})\| = \|x'(t_0 + 2\tilde{T})\| = \dots = \|x'(t_0 + (n - 1)\tilde{T})\| = 0.$$

Thus,  $n \leq k$ , which means that  $\tilde{T} = T/n \geq T/k$ .  $\square$

Simplex projection (1) means that the higher-density data points (i.e., smaller arc length  $\Delta l(t)$ ) on manifold  $M_x$  corresponds to lower uncertainty of estimates [25], which is consistent with the tests on simulated data and real data [37]. The proposition 3 links the period of arc length  $\Delta l(t)$  (2) to the period of time series  $x(t)$  via the function  $\|x'(t)\|$ . For the time series  $x(t)$  of common environmental infectious disease with period  $T$ , based on Proposition 3, the minimum positive period of  $\|x'(t)\|$  is the heuristic  $T/2$ . In addition, an infectious disease such as hand-foot and mouth disease (HFMD) with multiple peaks [47], wavelet analysis [48] of corresponding time series estimates the periodicity of peaks. Collectively, for environmental infectious disease with estimated minimal period  $T$ , if the Pearson’s correlation coefficient  $\rho(\tau)$  reaches a local maximal at  $\tau_*$ ,  $\rho(\tau)$  will not peak again in  $(\tau_* - T/2, \tau_* + T/2)$ . Thus, we provide a boundary  $B = (-T/4, T/4)$  with width  $T/2$  as the empirical testing window so that the extended CCM can be used to infer causality between periodic variables such as seasonal infectious diseases and environmental factors.

We next present the procedure to infer causality between variables  $x$  and  $y$ . For observed time series  $x(t)$  and  $y(t)$  with significant period  $T$ , the response time delay of variable  $x$  to variable  $y$  can be estimated by the following formula:

$$\bar{\tau}_{xy} = \arg \max_{\tau \in B} \text{Corr}(\bar{y}(t | \tau), y(t)), \tag{3}$$

where  $\bar{y}(t | \tau)$  are the estimated values of time series  $y(t)$ . Similarly, we can also obtain the estimated response time delay of variable  $y$  to variable  $x$  using the formula

$$\bar{\tau}_{yx} = \arg \max_{\tau \in B} \text{Corr}(\bar{x}(t | \tau), x(t)). \tag{4}$$

We consider the sign of  $\bar{\tau}_{xy}$  and  $\bar{\tau}_{yx}$  comprehensively to infer the causality between the variables  $x$  and  $y$ . If  $\bar{\tau}_{xy} \geq 0$  and  $\bar{\tau}_{yx} < 0$ , then variable  $x$  affects future values of  $y$  unidirectionally with time delay  $-\bar{\tau}_{yx}$ . If  $\bar{\tau}_{xy} < 0$  and  $\bar{\tau}_{yx} \geq 0$ , then variable  $y$  affects future values of  $x$  unidirectionally with time delay  $-\bar{\tau}_{xy}$ . If  $\bar{\tau}_{xy} < 0$  and  $\bar{\tau}_{yx} < 0$ , then  $x$  and  $y$  have two-way cause and effect, and the action time delay is  $-\bar{\tau}_{yx}$  and  $-\bar{\tau}_{xy}$ , respectively. If  $\bar{\tau}_{xy} \geq 0$  and  $\bar{\tau}_{yx} \geq 0$ , then we conclude that there is no causal evidence between variables  $x$  and  $y$ . We choose the negative optimal cross-lag as the estimated time delay because CCM is a historical information-dominated method [28], which is also consistent with previous extension [35].

Collectively, we have made an adjustment on the basis of CCM, and this update makes up for the limitation of the extended CCM in inferring causality between periodic time series.

### 3. Results

#### 3.1. Testing on an Infectious Disease Model

As the first step, we test the causal inference methods on an air quality index (AQI)-embedded susceptible-infectious-susceptible (SIS) epidemic model (Figure 2a). In this model, the total population ( $N$ ) consists of classes of individuals that are susceptible ( $S$ ) and infectious ( $I$ ), yielding

$$\begin{aligned} S(t+1) &= f(N(t)) + \sigma \phi\left(\frac{I(t)}{N(t)}, \alpha F(t)\right) S(t) + \gamma \sigma I(t), \\ I(t+1) &= \sigma \left(1 - \phi\left(\frac{I(t)}{N(t)}, \alpha F(t)\right)\right) S(t) + (1 - \gamma) \sigma I(t). \end{aligned} \tag{5}$$

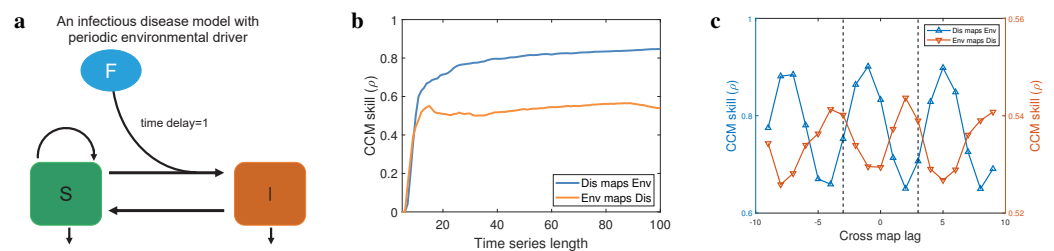
where  $t$  is time,  $f(N(t))$  is the density-dependent birth rate or recruitment according to the formula  $f(N(t)) = N(t) \exp(r - N(t))$ ,  $\sigma$  is the probability of survival,  $\alpha$  is a modified parameter for the effect of air pollution on incidence, and  $\gamma$  is the recovery rate. We assume that the proportion of susceptible individuals that do not become infected at time  $t$  is  $\phi(z(t), w(t)) = e^{-\beta z(t)} e^{-\beta w(t)}$  given the disease prevalence  $z(t) = I(t)/N(t)$  and air

pollution effect  $w(t) = \alpha F(t)$ , that is, encounters leading to infection are modeled via a Poisson process with the transmission constant  $\beta$ . The environmental driver  $F(t)$  (i.e., air quality index) varies according to the following formula, which is a discrete time form of a continuous time system [49].

$$F(t + 1) = \lambda(t)F(t) + C \tag{6}$$

where  $\lambda(t) = a - b * \sin(t/\omega)$  is the remaining proportion of mixed pollutants in the air from time step  $t$  to  $t + 1$ ;  $C$  is the constant rate of inflow of pollutants into the air, mainly depending on the persistent release of various air pollutants. We simulated the system with parameters setting in (Appendix A.1) and generated time series of the environmental factor and infectious individuals.

We first test CCM [28] using these simulated data, and we find that the CCM skill in both direction becomes better as the length of time series increases (Figure 2b), which implies a bidirectional causality between environmental factor and disease incidence [28] instead of the true unidirectional causality in our simulated system (Figure 2a). Thus, CCM is not suitable for inferring causality in the strongly coupled systems, because the strong coupling strength leads to a synchrony between response variable and driving variable, resulting in the dynamics of a response variable becoming dominated by those of the driving variable [35]. We next apply extended CCM [35] to identify the optimal cross-map lags between environmental factor and disease. In our numerical simulations,  $\lambda(t)$  is a  $4\pi$ -period function (Appendix A.1), so the period of time series  $F(t)$  and  $I(t)$  are approximately 12. The results of extended CCM show that the optimal time lag is not unique (Figure 2c), and the difference between two adjacent local optimal delays is around 6, which is consistent with Proposition 3 and the period of time series  $F(t)$  and  $I(t)$ .



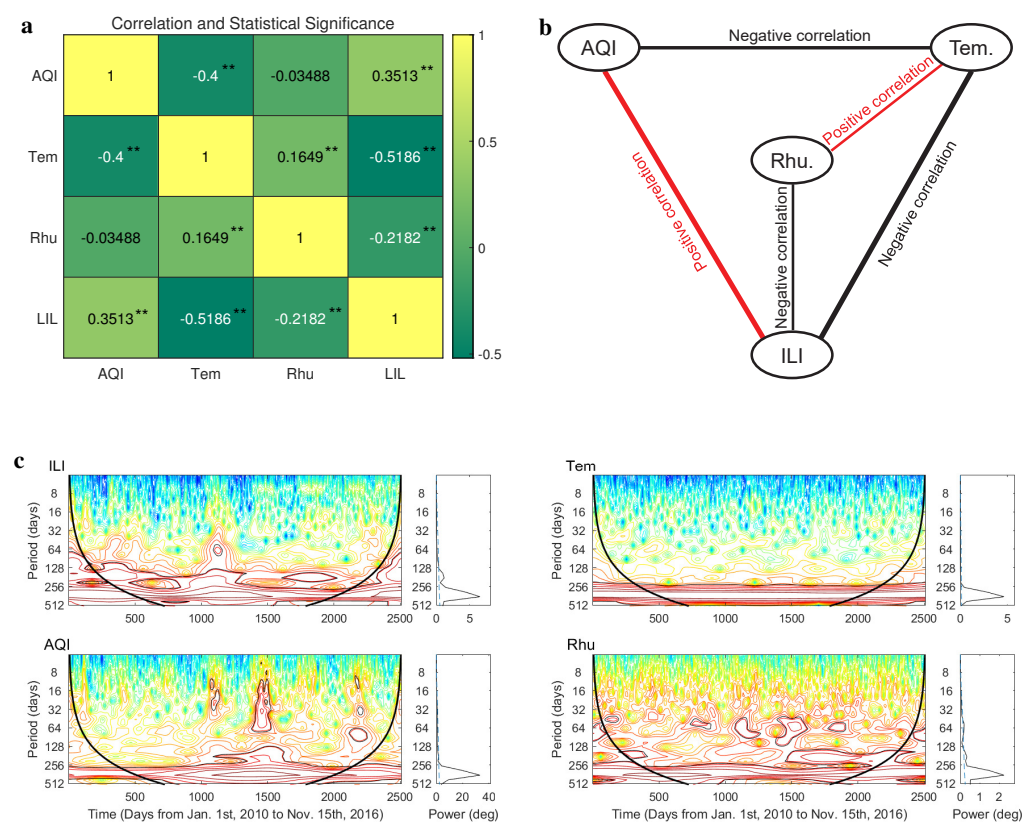
**Figure 2.** Numerical validation of theoretical results. (a) An environmental factor (F)–embedded susceptible–infectious–susceptible (SIS) epidemic model, in which the dynamics of environmental factor is periodic. The effect of environmental factor on disease incidence has a time delay 1. (b) The performance of CCM and the CCM skill as a function of the length of time series used to reconstruct the high–dimensional manifold. (c) The performance of extended CCM and the CCM skill as a function of the tested time delay. Here, the length of time series used to reconstruct the high–dimensional manifold is fixed.

In addition, by setting  $B = (-3, 3)$  as shown by the vertical dashed lines in Figure 2c, we estimated that  $\bar{\tau}_{Dis,Env} = -1$  and  $\bar{\tau}_{Env,Dis} = 2$  using the Formulas (3) and (4). Therefore, our inference is that the environmental factor affects disease incidence with a time delay 1, which is consistent with the modeling (Figure 2a). Collectively, we showed the limitations of CCM and extended CCM in inferring the causality between periodic time series (Figure 2b,c), which are overcome by adding an estimation interval for the extended CCM (Formulas (3) and (4)).

### 3.2. Correlation Analysis and Wavelet Analysis of Real Data

As a comparison, before inferring the causality among respiratory infection, air pollution, temperature, and humidity using the real data in Xi’an (Figure 1), we analyzed the correlation between these data and provided statistic significance of the correlation (Figure 3a). Based on the result, we constructed a correlation network which is undirected

(Figure 3b). Edges in this network simply indicate that the variations in two connected variables are positively or negatively correlated, but do not distinguish between a driving variable and a response variable. In addition, we found that there is no significant correlation between air quality index (AQI) and relative humidity (Rhu) in Xi’an. According to the theoretical analysis (Proposition 3) and numerical analysis (Figure 2), the periodicity of observed time series affects the performance of extended CCM. Moreover, the minimum positive period of the time series also determines the setting of the estimation interval for the optimal time delay between two variables. Here, we investigated the periodicity of influenza-like illness (ILI) cases, air quality index (AQI), daily temperature, and relative humidity in Xi’an using the wavelet analysis ([48]; Figure 3c). The results of wavelet analysis [48] show that all time series have only a significant annual cycle (see Figure 3c), so we can detect the optimal time lag in a narrow window  $B$  with the width of no more than 180 days using extended CCM (Formulas (3) and (4)). In the following analysis, we detect the optimal lag in the estimation interval  $(-50, 50)$ .



**Figure 3.** Correlation analysis and wavelet analysis of real data in Xi’an. (a) The correlation and statistical significance among the observed time series of influenza-like illness (ILI) cases, air quality index (AQI), lowest daily temperature, and relative humidity in Xi’an.  $p$ -values are  $< 0.01$  (\*\*). (b) The correlation network among the four variables we studied. (c) Wavelet analysis for time series of influenza-like illness (ILI) cases, air quality index (AQI), lowest daily temperature, and relative humidity. Wavelet power spectra are depicted on the left, and the right-hand panels show the mean spectra (vertical solid black line) with their significant threshold value of 0.05 (blue dashed line).

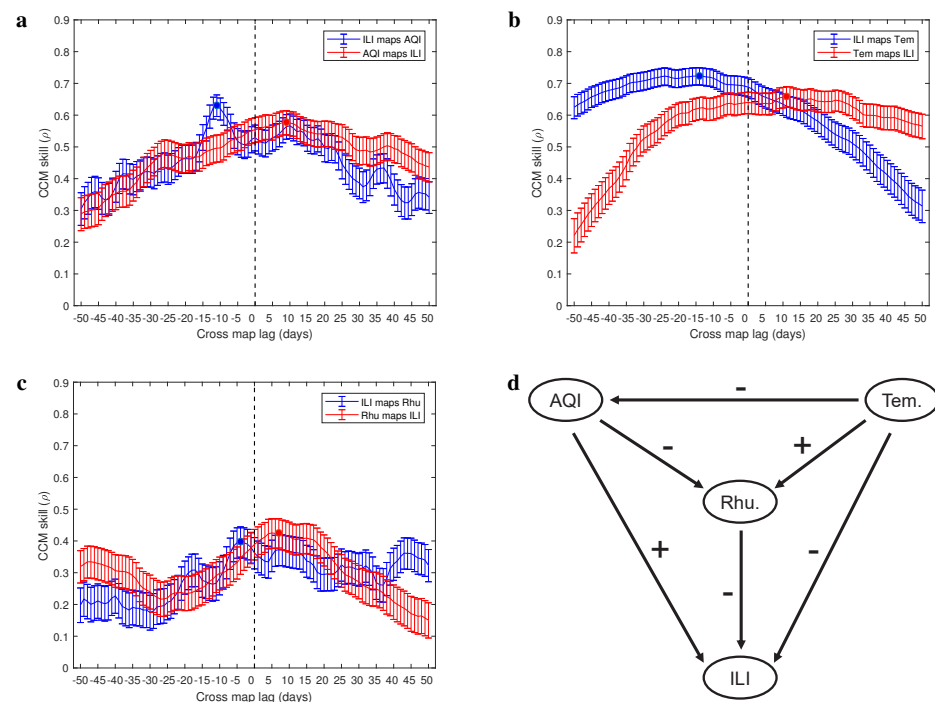
### 3.3. Causality Analysis of Real Data

In Figures 4a–c and 5a–c, we present the CCM skills using extended CCM in a narrow testing window. The cross-mapping skills between influenza-like illness and AQI time series indicate that there is a driving force from air pollution to respiratory infections with time delay of 11 days ( $\bar{\tau}_{ILI,AQI} = -11$  and  $\bar{\tau}_{AQI,ILI} = 9$ ; Figure 4a). The cross-mapping skills between influenza-like illness and temperature time series indicate that there is a driving force from temperature to respiratory infections with a delay of 14 days ( $\bar{\tau}_{ILI,Tem} = -14$

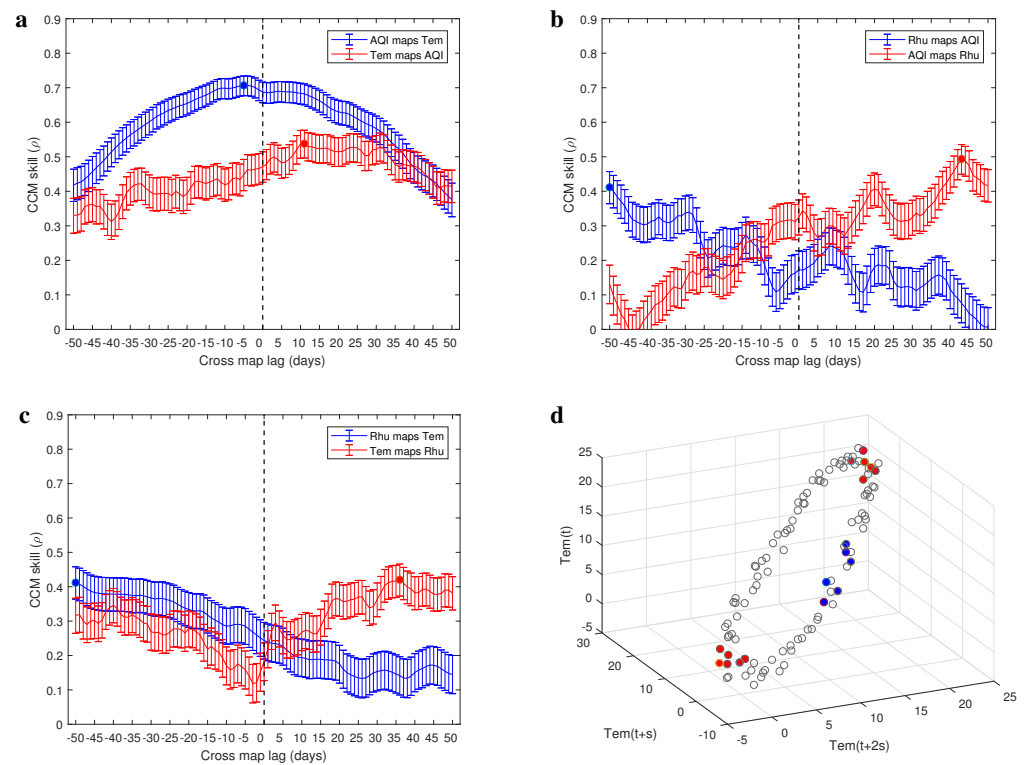


and  $\bar{\tau}_{Tem,ILI} = 11$ ; Figure 4b). Figure 4c shows the cross-mapping skills between influenza-like illness and relative humidity time series, which indicate that there is a driving force from relative humidity to respiratory infections with time delay of 4 days ( $\bar{\tau}_{ILI,Rhu} = -4$  and  $\bar{\tau}_{Rhu,ILI} = 7$ ). We further analyzed the causality between environmental factors (see Figure 5a–c) and obtained a causal network among the four variables studied (see Figure 4d). The sign on each side comes from the correlation analysis (Figure 3a). In contrast to the correlation network (Figure 3b), the causality network is a directed network where the edge distinguishes the driving variable and response variable (Figure 4d). From this directed network, we predict that increasing temperature increases relative humidity and decreases the air pollution degree as well as respiratory infections risk. More serious air pollution decreases relative humidity and increases respiratory infections risk, but higher relative humidity decreases respiratory infections risk. This indicates that in order to reduce the risk of respiratory infections, the indoor temperature and humidity can be improved by using air conditioners and air humidifiers.

Collectively, we inferred the causality between respiratory infections and environmental factors in Xi'an using the extended CCM by limiting the width of testing window for searching for the optimal time delay. The results enrich the studies on health effects of environmental factors.



**Figure 4.** Causal evidence among four variables (ILI—influenza like illness, AQI—air quality index, Tem.—temperature and Rhu.—relative humidity). (a–c) The CCM skills between involved variables as a function of tested cross-map lag. The negative optimal cross-map lag is the estimated interaction delay between them, e.g., the estimated delay for air pollution to drive influenza-like illness cases is 11 days. (d) Estimated causal network. The signs associated with arrows represent positive or negative correlation between two nodes. A negative sign means that increasing the drive variable would inhibit the response variable, and a positive sign means that higher drive variable would promote the response variable.



**Figure 5.** (a–c) Causal evidence between environmental factors. (d) Reconstructed manifold using the time series of temperature. Consistent with the theoretical analysis, the variations of points density in the reconstructed manifold is periodic (the blue points and red points are the 6 nearest neighbors of some points).

#### 4. Discussion

In this study, we presented evidence (Figures 4 and 5) for a causal relationship between respiratory infection and several environmental factors such as air quality, temperature, and humidity by adopting and refining a published method, CCM (convergent cross-mapping [28]). Related to this, we first performed theoretical analysis, centered chiefly on the extended CCM [35], which supports the result that the estimated optimal time delay between driving variable and response variable is not unique as long as these variables show synchrony and periodicity [37]. In addition, our theoretical result also suggests an idea to overcome the limitation of the extended CCM in inferring causality between synchronized periodic variables, which is estimating the optimal time delay in a bounded testing window (Formulas (3) and (4)). Consistently, we illustrated the limitations of the CCM as well as the extended CCM in inferring a causal relationship, and visualized the theoretical results (Proposition 3) using the data generated from an epidemic model (Figure 2) and the real data (Figure 5d). To illustrate that the characteristics of the real data (Figure 1) are consistent with the assumptions in our theoretical result and fall within the scope of our refined approach, we evaluated the periodicity of the data using wavelet analysis (Figure 3c). The minimal period of real data determines the width of testing window for estimating the time delay between the driving variable and response variable (Formulas (3) and (4)). By performing causal analysis for the real data, we suggested that all of air quality, temperature, and humidity have an effect on the incidence of respiratory infections. In particular, taking the reporting delay into account, air pollution promotes the respiratory infections risk with a time delay of 11 days.

In China, air pollution has been a public issue for a long time [50]. In recent years, the number of respiratory infections has highly synchronized with the variations of air quality index in China [19]. Through analyzing the time series of influenza-like illness and some environmental factors in Xi'an from 1 January 2010 to 15 November 2016, one

important result of our research is that air pollution fuels the risk of respiratory infections. Different from the significant correlation confirmed in previous studies [19], our study gives the causality between air pollution and respiratory infection, and identifies the detailed time delay. In addition, we found that lower temperature and humidity also fuel the respiratory infections risk. According to our estimated causal network (Figure 4d), temperature is located at the most upstream of the entire causal network and affects respiratory infections, air quality, and relative humidity. Because the temperature in the Northern Hemisphere is mainly affected by the relative position of the Earth and the sun, the variables we studied may come from a larger and more complex system.

In addition to respiratory diseases [51], air pollution also leads to 3.3 million premature deaths per year worldwide [52] and has a substantial role in many noncommunicable diseases [53] such as cancer [54], stroke [55], cardiovascular disease [56,57], and Alzheimer's disease [58,59]. Furthermore, available evidence suggests that air pollution can prevent the beneficial cardiopulmonary effects of walking in people with heart or chronic lung disease [60,61] and results in poor lung function in children [62,63]. Therefore, our results complement previous research on the health effects of air pollution, which strengthens the importance of improving air quality.

It is well known that the variables being correlated does not imply that they are causal. Comparing the causal analysis with correlation analysis (see Figures 3b and 4d), we find that the causality does not imply significant correlation either. The possible reason is that standard correlation analysis mainly captures the linear relationships between variables, while real data may arise from complex nonlinear systems in which ephemeral correlations are common [64,65]. By broadening the scope of application and improving the accuracy as we did in this study, CCM can help us understand the relationship between variables in complex systems such as molecular systems [66] and public health systems [49], in which the causality is useful for designing novel experiments and interventions, respectively.

It is worth noting that in order to ensure that the optimal cross-map lag of extended CCM is unique, we define a bounded testing window which depends on the period of time series (Formulas (3) and (4)). This means that the real time delay between driving variable and response variable should be in the testing window. Otherwise, we would give a wrong causality and a wrong time delay between two variables. Nonetheless, we are still confident that our work has a broad applicability in time series analysis. For seasonal diseases such as respiratory infections, the period of the time series is usually very long and the unit is years. Therefore, the testing window is wide enough to cover the real time delay between driving variable and response variable. In addition, due to the temporal decay of information in transfer process, when the extended CCM is used to infer causality, only a rough approximation of the action time delay can be estimated. In some cases, the estimates of action time delay may not be reliable (Figure 5b,c). Related to this, we inferred causality by following the rule that the values of the driving variable are better to estimate the future value of the response variable, whereas the values of the response variable are better to estimate the past value of the driving variable [35].

In summary, we refined a published method (CCM) for causal inference through theoretical analysis. Using this refined method to analyze the time series of influenza-like illness, air quality, temperature, and humidity in Xi'an, we established a causal network where the nodes are these variables. In this network, air pollution promotes respiratory infections risk while higher temperature and humidity limit the risk (Figure 4d). It will be of interest to test the robustness of this causal network using different datasets, and to determine how these results will impact the design of novel interventions against respiratory infections in populations.

**Author Contributions:** Conceptualization, D.C., X.S. and R.A.C.; methodology, D.C., X.S. and R.A.C.; software, D.C.; validation, D.C., X.S. and R.A.C.; formal analysis, D.C. and X.S.; investigation, D.C.; resources, X.S.; data curation, X.S.; writing—original draft preparation, D.C.; writing—review and editing, D.C., X.S. and R.A.C.; visualization, D.C.; supervision, X.S.; project administration, X.S.; funding acquisition, X.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key R&D Program of China (2022YFA1003704), by the National Natural Science Foundation of China (12071366) and by the China Scholarship Council.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sources are already included in the main text.

**Acknowledgments:** We would like to thank Yanni Xiao and Luonan Chen for discussions and advice.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

Tem	Temperature
AQI	Air quality index
Rhu	Relative humidity
ILI	Influenza-like illness

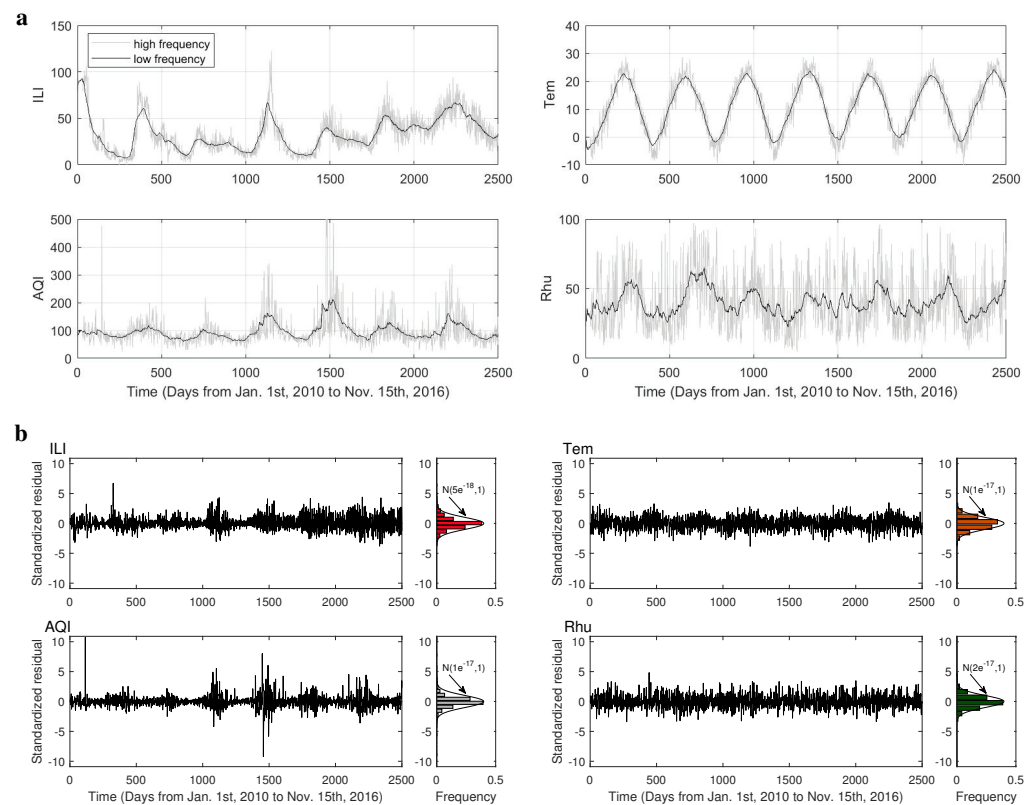
## Appendix A

### Appendix A.1. The Generation and Analysis of Simulated Data

The system (5) is initialized as  $S(1) = 5$ ,  $I(1) = 4$ , and  $F(1) = 5$ , and run for 3500 time steps with the parameters  $\lambda(t) = 0.6 - 0.3\sin(t/2)$ ,  $r = 6$ ,  $\sigma = 0.5$ ,  $\alpha = 8$ ,  $\beta = 0.01$ , and  $\gamma = 0.01$  to generate the time series of  $I(t)$  and  $F(t)$ . These time series are analyzed using  $s = 1$  and  $\Pi = 3$  which ensure the requirements of Takens' embedding theorem [30,31]. We first test CCM [28] using these simulated data, selecting 100 pairs of 100 vectors of random libraries over time points 101–1000, computing the cross-map skill for each pair of libraries using time series length 7–100, and averaging out the cross-map skill of 100 random libraries, respectively. We then apply extended CCM [35] to identify the optimal cross-map lags, selecting 100 pairs of random libraries of 2000 vectors over time points 101–3000, computing cross-map skill for each pair of libraries with maximum length using different cross-map lag and averaging out the cross-map skill of these random libraries, respectively.

### Appendix A.2. Kalman Filtering

Kalman filtering was achieved by using the function `kalman` in MATLAB (available from: <https://www.mathworks.com/help/control/ref/ss.kalman.html> accessed on 19 January 2021). The low-frequency component was derived first (Figure A1a) and this was subtracted from the original series (high-frequency component) to provide the residual. After standardization, we found that the residuals were fitted with a standard normal distribution (Figure A1b).



**Figure A1.** Results of Kalman filtering. (a) Data of the influenza–like illness (ILI) cases, air quality index (AQI), lowest daily temperature and relative humidity, low–frequency component, and original date (high–frequency component). (b) The time series and distribution of standardized residuals. The residuals were calculated by subtracting low–frequency component from the original series.

## References

1. Jones, K.E.; Patel, N.G.; Levy, M.A.; Storeygard, A.; Balk, D.; Gittleman, J.L.; Daszak, P. Global trends in emerging infectious diseases. *Nature* **2008**, *451*, 990–993. [[CrossRef](#)] [[PubMed](#)]
2. Weiss, R.A.; McMichael, A.J. Social and environmental risk factors in the emergence of infectious diseases. *Nat. Med.* **2004**, *10*, S70–S76. [[CrossRef](#)] [[PubMed](#)]
3. McMichael, A.J. Environmental and social influences on emerging infectious diseases: Past, present and future. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* **2004**, *359*, 1049–1058. [[CrossRef](#)] [[PubMed](#)]
4. Hay, S.; Tatem, A.; Graham, A.; Goetz, S.; Rogers, D. Global environmental data for mapping infectious disease distribution. *Adv. Parasitol.* **2006**, *62*, 37–77. [[PubMed](#)]
5. Eisenberg, J.N.; Desai, M.A.; Levy, K.; Bates, S.J.; Liang, S.; Naumoff, K.; Scott, J.C. Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. *Environ. Health Perspect.* **2007**, *115*, 1216–1223. [[CrossRef](#)]
6. Koelle, K.; Rodó, X.; Pascual, M.; Yunus, M.; Mostafa, G. Refractory periods and climate forcing in cholera dynamics. *Nature* **2005**, *436*, 696–700. [[CrossRef](#)]
7. Mordecai, E.A.; Caldwell, J.M.; Grossman, M.K.; Lippi, C.A.; Johnson, L.R.; Neira, M.; Rohr, J.R.; Ryan, S.J.; Savage, V.; Shocket, M.S.; et al. Thermal biology of mosquito-borne disease. *Ecol. Lett.* **2019**, *22*, 1690–1708. [[CrossRef](#)]
8. Johansson, M.A.; Apfeldorf, K.M.; Dobson, S.; Devita, J.; Buczak, A.L.; Baugher, B.; Moniz, L.J.; Bagley, T.; Babin, S.M.; Guven, E.; et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24268–24274. [[CrossRef](#)]
9. Shocket, M.S.; Ryan, S.J.; Mordecai, E.A. Temperature explains broad patterns of Ross River virus transmission. *eLife* **2018**, *7*, e37762. [[CrossRef](#)]
10. Deyle, E.R.; Maher, M.C.; Hernandez, R.D.; Basu, S.; Sugihara, G. Global environmental drivers of influenza. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 13081–13086. [[CrossRef](#)]
11. Ali, S.T.; Cowling, B.J.; Wong, J.Y.; Chen, D.; Shan, S.; Lau, E.H.; He, D.; Tian, L.; Li, Z.; Wu, P. Influenza seasonality and its environmental driving factors in mainland China and Hong Kong. *Sci. Total Environ.* **2022**, *818*, 151724. [[CrossRef](#)]
12. Sly, P.D.; Trottier, B.; Ikeda-Araki, A.; Vilcins, D. Environmental impacts on infectious disease: A literature view of epidemiological evidence. *Ann. Glob. Health* **2022**, *88*, 91. [[CrossRef](#)]

13. Lin, H.H.; Ezzati, M.; Murray, M. Tobacco smoke, indoor air pollution and tuberculosis: A systematic review and meta-analysis. *PLoS Med.* **2007**, *4*, e20. [[CrossRef](#)]
14. Xiang, K.; Xu, Z.; Hu, Y.Q.; He, Y.S.; Dan, Y.L.; Wu, Q.; Fang, X.H.; Pan, H.F. Association between ambient air pollution and tuberculosis risk: A systematic review and meta-analysis. *Chemosphere* **2021**, *277*, 130342. [[CrossRef](#)]
15. Wong, C.M.; Yang, L.; Thach, T.Q.; Chau, P.Y.K.; Chan, K.P.; Thomas, G.N.; Lam, T.H.; Wong, T.W.; Hedley, A.J.; Peiris, J.M. Modification by Influenza on Health Effects of Air Pollution in Hong Kong. *Environ. Health Perspect.* **2009**, *117*, 248–253. [[CrossRef](#)]
16. Liang, Y.; Fang, L.; Pan, H.; Zhang, K.; Kan, H.; Brook, J.R.; Sun, Q. PM 2.5 in Beijing—temporal pattern and its association with influenza. *Environ. Health* **2014**, *13*, 102. [[CrossRef](#)] [[PubMed](#)]
17. Chen, G.; Zhang, W.; Li, S.; Zhang, Y.; Williams, G.; Huxley, R.; Ren, H.; Cao, W.; Guo, Y. The impact of ambient fine particles on influenza transmission and the modification effects of temperature in China: A multi-city study. *Environ. Int.* **2017**, *98*, 82–88. [[CrossRef](#)]
18. Feng, C.; Li, J.; Sun, W.; Zhang, Y.; Wang, Q. Impact of ambient fine particulate matter (PM<sub>2.5</sub>) exposure on the risk of influenza-like-illness: A time-series analysis in Beijing, China. *Environ. Health* **2016**, *15*, 17. [[CrossRef](#)]
19. Tang, S.; Yan, Q.; Shi, W.; Wang, X.; Sun, X.; Yu, P.; Wu, J.; Xiao, Y. Measuring the impact of air pollution on respiratory infection risk in China. *Environ. Pollut.* **2018**, *232*, 477–486. [[CrossRef](#)]
20. Fattorini, D.; Regoli, F. Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. *Environ. Pollut.* **2020**, *264*, 114732. [[CrossRef](#)] [[PubMed](#)]
21. Travaglio, M.; Yu, Y.; Popovic, R.; Selley, L.; Leal, N.S.; Martins, L.M. Links between air pollution and COVID-19 in England. *Environ. Pollut.* **2021**, *268*, 115859. [[CrossRef](#)]
22. Berkeley, G. *A Treatise Concerning the Principles of Human Knowledge*; JB Lippincott and Company: Philadelphia, NY, USA, 1881.
23. Mirsaeidi, M.; Motahari, H.; Taghizadeh Khamesi, M.; Sharifi, A.; Campos, M.; Schraufnagel, D.E. Climate change and respiratory infections. *Ann. Am. Thorac. Soc.* **2016**, *13*, 1223–1230. [[CrossRef](#)] [[PubMed](#)]
24. Eichler, M., Causal Inference in Time Series Analysis. In *Causality*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 2012; Chapter 22, pp. 327–354.
25. Sugihara, G.; May, R.M. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **1990**, *344*, 734–741. [[CrossRef](#)]
26. Sugihara, G. Nonlinear forecasting for the classification of natural time series. *Philos. Trans. R. Soc. London. Ser. A Phys. Eng. Sci.* **1994**, *348*, 477–495.
27. Dixon, P.A.; Milicich, M.J.; Sugihara, G. Episodic fluctuations in larval supply. *Science* **1999**, *283*, 1528–1530. [[CrossRef](#)]
28. Sugihara, G.; May, R.; Ye, H.; Hsieh, C.H.; Deyle, E.; Fogarty, M.; Munch, S. Detecting Causality in Complex Ecosystems. *Science* **2012**, *338*, 496–500. [[CrossRef](#)]
29. Ye, H.; Sugihara, G. Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality. *Science* **2016**, *353*, 922–925. [[CrossRef](#)]
30. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381.
31. Stark, J. Delay Embeddings for Forced Systems. I. Deterministic Forcing. *J. Nonlinear Sci.* **1999**, *9*, 255–332. [[CrossRef](#)]
32. Runge, J.; Bathiany, S.; Bollt, E.; Camps-Valls, G.; Coumou, D.; Deyle, E.; Glymour, C.; Kretschmer, M.; Mahecha, M.D.; Muñoz-Mari, J.; et al. Inferring causation from time series in Earth system sciences. *Nat. Commun.* **2019**, *10*, 1–13. [[CrossRef](#)]
33. Cheke, R.A.; Young, S.; Wang, X.; Tratalos, J.A.; Tang, S.; Cressman, K. Evidence for a causal relationship between the solar cycle and locust abundance. *Agronomy* **2020**, *11*, 69. [[CrossRef](#)]
34. Doi, H.; Yasuhara, M.; Ushio, M. Causal analysis of the temperature impact on deep-sea biodiversity. *Biol. Lett.* **2021**, *17*, 20200666. [[CrossRef](#)] [[PubMed](#)]
35. Ye, H.; Deyle, E.R.; Gilarranz, L.J.; Sugihara, G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.* **2015**, *5*, 14750. [[CrossRef](#)]
36. Rulkov, N.F.; Sushchik, M.M.; Tsimring, L.S.; Abarbanel, H.D. Generalized synchronization of chaos in directionally coupled chaotic systems. *Phys. Rev. E* **1995**, *51*, 980. [[CrossRef](#)]
37. Sugihara, G.; Deyle, E.R.; Ye, H. Reply to Baskerville and Cobey: Misconceptions about causation with synchrony and seasonal drivers. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2272–E2274. [[CrossRef](#)]
38. Bjørnstad, O.N.; Viboud, C. Timing and periodicity of influenza epidemics. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12899–12901. [[CrossRef](#)]
39. wikipedia. Air Quality Index. 2022. Available online: [https://en.wikipedia.org/wiki/Air\\_quality\\_index](https://en.wikipedia.org/wiki/Air_quality_index) (accessed on 18 February 2020).
40. tianqihoubao. Air Quality Index. 2022. Available online: <http://www.tianqihoubao.com/aqi/xian.html> (accessed on 18 February 2020).
41. Administration, C.M. Meteorological Data. 2022. Available online: <http://data.cma.cn> (accessed on 18 February 2020).
42. Granger, C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econom. J. Econom. Soc.* **1969**, *37*, 424–438. [[CrossRef](#)]

43. Tsonis, A.A.; Deyle, E.R.; May, R.M.; Sugihara, G.; Swanson, K.; Verbeten, J.D.; Wang, G. Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3253–3256. [[CrossRef](#)]
44. Cobey, S.; Baskerville, E.B. Limits to causal inference with state-space reconstruction for infectious disease. *PLoS ONE* **2016**, *11*, e0169050. [[CrossRef](#)]
45. Southallzy, B.; Buxtony, B.; Marchant, J. Controllability and observability: Tools for Kalman filter design. In Proceedings of the British Machine Vision Conference, Southampton, UK, 14–17 September 1998; Volume 98, pp. 164–173.
46. Whitney, H. Differentiable manifolds. *Ann. Math.* **1936**, *37*, 645–680. [[CrossRef](#)]
47. Xia, F.; Deng, F.; Tian, H.; He, W.; Xiao, Y.; Sun, X. Estimation of the reproduction number and identification of periodicity for HFMD infections in northwest China. *J. Theor. Biol.* **2020**, *484*, 110027. [[CrossRef](#)]
48. Torrence, C.; Compo, G.P. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 61–78. [[CrossRef](#)]
49. Chen, D.; Xiao, Y.; Tang, S. Air quality index induced nonsmooth system for respiratory infection. *J. Theor. Biol.* **2019**, *460*, 160–169. [[CrossRef](#)]
50. Li, M.; Zhang, L. Haze in China: Current and future challenges. *Environ. Pollut.* **2014**, *189*, 85–86. [[CrossRef](#)] [[PubMed](#)]
51. Guarnieri, M.; Balmes, J.R. Outdoor air pollution and asthma. *Lancet* **2014**, *383*, 1581–1592. [[CrossRef](#)] [[PubMed](#)]
52. Lelieveld, J.; Evans, J.S.; Fnais, M.; Giannadaki, D.; Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **2015**, *525*, 367–373. [[CrossRef](#)]
53. Maria Neira, A.P.U.; Mudu, P. Reduce air pollution to beat NCDs: From recognition to action. *Lancet* **2018**, *392*, 1178–1179. [[CrossRef](#)]
54. Loomis, D.; Grosse, Y.; Lauby-Secretan, B.; El Ghissassi, F.; Bouvard, V.; Benbrahim-Tallaa, L.; Guha, N.; Baan, R.; Mattock, H.; Straif, K. The carcinogenicity of outdoor air pollution. *Lancet Oncol.* **2013**, *14*, 1262–1263. [[CrossRef](#)]
55. Pandian, J.D.; Gall, S.L.; Kate, M.P.; Silva, G.S.; Akinyemi, R.O.; Ovbiagele, B.I.; Lavados, P.M.; Gandhi, D.B.C.; Thrift, A.G. Prevention of stroke: A global perspective. *Lancet* **2018**, *392*, 1269–1278. [[CrossRef](#)]
56. Xie, W.; Li, G.; Zhao, D.; Xie, X.; Wei, Z.; Wang, W.; Wang, M.; Li, G.; Liu, W.; Sun, J.; et al. Relationship between fine particulate air pollution and ischaemic heart disease morbidity and mortality. *Heart* **2015**, *101*, 257–263. [[CrossRef](#)] [[PubMed](#)]
57. Franklin, B.A.; Brook, R.; Arden Pope, C. Air Pollution and Cardiovascular Disease. *Curr. Probl. Cardiol.* **2015**, *40*, 207–238. [[CrossRef](#)]
58. Hautot, D.; Pankhurst, Q.A.; Khan, N.; Dobson, J. Preliminary evaluation of nanoscale biogenic magnetite in Alzheimer’s disease brain tissue. *Proc. R. Soc. London. Ser. B Biol. Sci.* **2003**, *270*, S62–S64. [[CrossRef](#)] [[PubMed](#)]
59. Maher, B.A.; Ahmed, I.A.M.; Karloukovski, V.; MacLaren, D.A.; Foulds, P.G.; Allsop, D.; Mann, D.M.A.; Torres-Jardón, R.; Calderon-Garciduenas, L. Magnetite pollution nanoparticles in the human brain. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 10797–10801. [[CrossRef](#)]
60. Pope III, C.A.; Dockery, D.W. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manag. Assoc.* **2006**, *56*, 709–742. [[CrossRef](#)] [[PubMed](#)]
61. Sinharay, R.; Gong, J.; Barratt, B.; Ohman-Strickland, P.; Ernst, S.; Kelly, F.J.; Zhang, J.J.; Collins, P.; Cullinan, P.; Chung, K.F. Respiratory and cardiovascular responses to walking down a traffic-polluted road compared with walking in a traffic-free area in participants aged 60 years and older with chronic lung or heart disease and age-matched healthy controls: A randomised, crossover study. *Lancet* **2018**, *391*, 339–349.
62. Gauderman, W.J.; Urman, R.; Avol, E.; Berhane, K.; McConnell, R.; Rappaport, E.; Chang, R.; Lurmann, F.; Gilliland, F. Association of Improved Air Quality with Lung Development in Children. *N. Engl. J. Med.* **2015**, *372*, 905–913. [[CrossRef](#)] [[PubMed](#)]
63. Dockery, D.W.; Ware, J.H. Cleaner Air, Bigger Lungs. *N. Engl. J. Med.* **2015**, *372*, 970–972. [[CrossRef](#)]
64. Mysterud, A.; Stenseth, N.C.; Yoccoz, N.G.; Langvatn, R.; Steinheim, G. Nonlinear effects of large-scale climatic variability on wild and domestic herbivores. *Nature* **2001**, *410*, 1096–1099. [[CrossRef](#)]
65. Wagner, B.K.; Kitami, T.; Gilbert, T.J.; Peck, D.; Ramanathan, A.; Schreiber, S.L.; Golub, T.R.; Mootha, V.K. Large-scale chemical dissection of mitochondrial function. *Nat. Biotechnol.* **2008**, *26*, 343–351. [[CrossRef](#)]
66. Chen, D.; Forghany, Z.; Liu, X.; Wang, H.; Merks, R.M.; Baker, D.A. A new model of Notch signalling: Control of Notch receptor cis-inhibition via Notch ligand dimers. *PLoS Comput. Biol.* **2023**, *19*, e1010169. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.