Environmental Pollution 303 (2022) 119057

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Causation inference in complicated atmospheric environment☆

Ziyue Chen[a], Miaoqing Xu[a], Bingbo Gao[b], George Sugihara[c], Feixue Shen[d], Yanyan Cai[d], Anqi Li[d], Qi Wu[d], Lin Yang[d,*], Qi Yao[a], Xiao Chen[a], Jing Yang[a], Chenghu Zhou[d], Manchun Li[d]

[a] State Lab of Remote Sensing Sciences of China, College of Global and Earth System Sciences, Beijing Normal University, 19 Xinjiekou Street, Haidian, Beijing, 100875, China
[b] College of Land Science and Technology, China Agricultural University, Beijing, 100083, China
[c] Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, 92093, USA
[d] School of Geography and Ocean Science, Nanjing University, Nanjing, 210023, China

A R T I C L E I N F O

Keywords:
Causation inference
Atmospheric environment
Multi-model comparison
Indirect causation

A B S T R A C T

Reliable attribution is crucial for understanding various climate change issues. However, complicated inner-interactions between various factors make causation inference in atmospheric environment highly challenging. Taking $PM_{2.5}$-Meteorology causation, which involves a large number of non-Linear and uncertain interactions between many meteorological factors and $PM_{2.5}$, as a case, we examined the performance of a series of mainstream statistical models, including Correlation Analysis (CA), Partial Correlation Analysis (PCA), Structural Equation Model (SEM), Convergent Cross Mapping (CCM), Partial Cross Mapping (PCM) and Geographical Detector (GD). From a coarse perspective, the Top 3 major meteorological factors for $PM_{2.5}$ in 190 cities across China extracted using different models were generally consistent. From a strict perspective, the extracted dominant meteorological factor for $PM_{2.5}$ demonstrated large model variations and shared a limited consistence. Such models as SEM and PCM, which are capable of further separating direct and indirect causation in simple systems, performed poorly to identify the direct and indirect $PM_{2.5}$-Meteorology causation. The notable inconsistence denied the feasibility of employing multiple models for better causation inference in atmospheric environment. Instead, the sole use of CCM, which is advantageous in dealing with non-linear causation and removing disturbing factors, is a preferable strategy for causation inference in complicated ecosystems. Meanwhile, given the multi-direction, uncertain interactions between many variables, we should be more cautious and less ambitious on the separation of direct and indirect causation. For better causation inference in the complicated atmospheric environment, the combination of statistical models and atmospheric models, and the further exploration of Deep Neural Network can be promising strategies.

1. Introduction

Climate change has become one of most crucial social-economical-ecological issues that received growing emphasis. Such environmental events as global warming, air pollution, heat waves and drought have been exerting a severe threat to public health (Bryan et al., 2020; El-Sayed and Kamel, 2020; Bhat et al., 2021), crop yields (Trnka et al., 2014; Wang et al., 2016a, 2016b; Pagani et al., 2017), biodiversity (Lee et al., 2017; Zhang et al., 2018) and international trade (Azam et al., 2016). Against a variety of side effects induced by climate change, scholars have been working substantially towards a better understanding of the cause and effects of climate change. Despite a rapid progress, major challenges remain in the reliable attribution of climate change induced environmental issues. Climate change mainly occurs in the atmospheric environment, where a large number of meteorological factors and airborne (gaseous, liquid and particulate) components interact intensely. For instance, solar radiations influence the near-ground temperature (Zhong et al., 2017), which further affecting the vertical

☆ This paper has been recommended for acceptance by Maria Cristina Fossi.

* Corresponding author.
E-mail addresses: zychen@bnu.edu.cn (Z. Chen), xumiaoqing@mail.bnu.edu.cn (M. Xu), gaobingbo@cau.edu.cn (B. Gao), gsugihara@ucsd.edu (G. Sugihara), mg1827017@smail.nju.edu.cn (F. Shen), mg1927050@smail.nju.edu.cn (Y. Cai), lianqidlmu@163.com (A. Li), MG20270027@smail.nju.edu.cn (Q. Wu), yanglin@nju.edu.cn (L. Yang), yaoqi@mail.bnu.edu.cn (Q. Yao), xchen@mail.bnu.edu.cn (X. Chen), jYang@mail.bnu.edu.cn (J. Yang), zhouch@lreis.ac.cn (C. Zhou), limanchun@nju.edu.cn (M. Li).

https://doi.org/10.1016/j.envpol.2022.119057
Received 17 December 2021; Received in revised form 21 February 2022; Accepted 23 February 2022
Available online 26 February 2022
0269-7491/© 2022 Elsevier Ltd. All rights reserved.

and horizontal air motion (wind speed and wind direction), humidity (Bhardwaj et al., 2013), boundary layer height (Bianco et al., 2011). Successively, the variation of solar radiation, temperature, wind, humidity and boundary layer height can all individually impact Particulate Matters (PM), ozone and other airborne components through enhanced or restricted photolysis (Luo et al.; Xing et al., 2017), chemical reaction (Tie et al., 2019), diffusion (Luo et al., 2017), and hygroscopic increase (Tie et al., 2017). Meanwhile, airborne components (e.g. PM, and CO_2) can have a strong feedback effect on such meteorological factors as temperature, wind and humidity and solar radiation (Tai et al., 2010; Yang et al., 2016; Zhong et al., 2017; Zhou et al., 2018). Therefore, meteorology-component interactions are bidirectional and asymmetric, presenting uncertain patterns under different circumstances (Chen et al., 2020). On the other hand, the physical and chemical reactions between different airborne components are even more complicated. There are dozens of reactions between SO_x, NO_x, CO_x, O_x, H_xO_x and VOCs, subject to varying meteorological and composition conditions (Baklanov et al., 2014). Consequently, the reaction types and intensities between various airborne components are highly unpredictable (Chen et al., 2020). Specifically, due to the disturbance of other influencing factors, the relative contribution of individual factors (e.g. CO_2 concentrations) to a specific environmental process (e.g. warming) can hardly be isolated and quantified, which has become one major difficulty in properly attributing climate change events (Fang et al., 2017).

Scholars have increasingly employed global or regional atmospheric models, such as GCAM (Global Change Assessment Model) (Calvin et al., 2013), GEOS (Goddard Earth Observing System) (Liu et al., 2009; Molod et al., 2012; Colarco et al., 2014) and WRF (Weather Research and Forecasting Model) (Chotamonsak et al., 2011; Srinivas et al., 2013) to simulate the climate dynamics, and such Chemical Transport Models (CTMs) as WRF-CAM_x (Skjøth et al., 2013; Shahbazi et al., 2017) and GEOS-Chem (Sherwen et al., 2016) to simulate the interactions between different airborne components and estimate their variations based on the combination of ground-level observation and remote sensing data. Despite their wide implementations, some major limitations exist (Morice et al., 2012; Mizuta et al., 2017). For atmospheric models, the reaction mechanisms (physically or chemically) between atmospheric factors should be clearly explained for reliable prediction. However, till now, many reaction mechanisms remain unclear. Furthermore, the reaction mechanisms between the same variables may vary significantly under different conditions, while atmospheric models can hardly explain the actual types, proportion and orders of reactions in a highly complicated atmospheric environment (Saikawa et al., 2014).

Different from climate model simulation, statistical models attempt to extract the relationship between variables by examining their simultaneous temporal variations. On one hand, the precise description of interaction mechanisms between different variables is not required for statistical models, which avoids one major source of uncertainty (VanderWeele et al., 2014). On the other hand, disturbed by multi-direction interactions between a diversity of meteorological factors and airborne components, commonly employed correlation models can be affected significantly by other factors (Sugihara et al., 2012; Chen et al., 2020). In this regard, inferring causation in atmospheric environment based on proper statistical models remains a promising, yet challenging strategy.

Given the great significance of the attribution of climate change issues, effective causation inference based on statistical models should be further explored. Compared with the mature and massive use of climate models, statistical models have been limitedly employed for understanding the drivers and causes in atmospheric environment, and their reliability has yet been evaluated. To this end, we take $PM_{2.5}$-Meteorology interactions, which involve various atmospheric factors and is one of major and challenging issues in climate change research, as a case research. We employ a diversity of mainstream statistical models to quantify the influence of individual meteorological factors on $PM_{2.5}$ concentrations. Due to the lack of observable reference data, we considered those $PM_{2.5}$-Meteorology prior knowledge from previous studies as reference to roughly judge the performance of these models. By comparing model differences, we aim to examine the reliability, limitations and underlying connections between these models. Furthermore, we attempt to check whether a cross-verification of multiple models can lead to a shared, more reliable output, or this multi-model comparison can confirm the reliability of specific models in the complicated atmospheric environment. This research provides a comprehensive picture how the complicated atmospheric environment pose a real challenge on reliable causation inference, and useful decision support for selecting and interpreting statistical models accordingly.

2. Materials and methods

2.1. Data sources

$PM_{2.5}$ data in 190 monitoring cities were acquired from the website www.PM25.in, which collected data from China National Environmental Monitoring Center (CNEMC). The daily $PM_{2.5}$ concentrations were calculated using the averaged hourly $PM_{2.5}$ concentrations. For a consecutive division of different seasons (Spring: March to May, Summer: June to August, Autumn: September to November, Winter: December to February) and multiple-year analysis, we employed $PM_{2.5}$ data from March 1st, 2014 to February 28th, 2017.

The meteorological data in 190 cities during the same period were obtained from the "China Meteorological Data Sharing Service System", and a set of meteorological factors, including humidity, temperature, wind speed, wind direction, solar radiation, evaporation, precipitation, and air pressure, which were closely related to $PM_{2.5}$ concentrations (Chen et al., 2018; Chen et al., 2020) were selected for this research. For better analysis of meteorology-$PM_{2.5}$ interactions, the meteorological data included some sub-factors for each meteorological factor, such as temperature (daily max temperature, mean temperature, minimum temperature, and largest temperature difference for the day), precipitation (total precipitation from 8am to 8pm, total precipitation from 8pm to 8am and total precipitation for the day) and wind speed (mean wind speed, max wind speed and extreme wind speed). In this research, the strongest calculated interactions between one sub-factor (e.g. mean temperature) and $PM_{2.5}$ concentrations were regarded as the influence of the meteorological factor (e.g. temperature) on $PM_{2.5}$ concentrations.

2.2. Model descriptions

In recent years, causation inference in complicated ecosystems has become a hot, yet challenging topic. Against this background, growing statistic models have been proposed and implemented. However, these models have been limited implemented in atmospheric environment and their reliability has rarely been verified. The major difficulty lies in the lack of observable reference data. To address this issue from a feasible perspective, we employed a series of mainstream statistical models, including Correlation Analysis (CA), Partial Correlation Analysis (P-CA), Convergent Cross Mapping (CCM), Structural Equation Modelling (SEM), Partial Cross Mapping (P-CM) and Geographical Detector (GD) to respectively quantify the influence of individual meteorological factors on $PM_{2.5}$ concentrations. Following this, we attempted to compare and interpret the differences between these model outputs according to their mechanisms. Through cross-verification, we can better understand the reliability and suitability of mainstream statistical models and their underlying connections in causation inference in atmospheric research. Except for the well-known Correlation Model and Partial Correlation Model, the general principle and setting of other statistical models employed are briefly introduced as follows.

2.3. Convergent Cross Mapping (CCM)

Convergent Cross Mapping (CCM) (Sugihara, G. et al., 2012) aims to identify mirage correlations and extract the robust quantitative interactions between two variables in complicated ecosystems by effectively removing the influence of other factors. Since CCM is advantageous of detecting weak coupling and calculating the asymmetric, bidirectional causality between two variables, this model has been massively implemented. CCM calculates the predictive skill of variable A on B (defined as ρ, ranging from 0 to 1), which provides quantitative reference for comparing the magnitude of different variables on one variable.

The principle and establishment of CCM is briefly explained as follows. For two variables X and Y and their time series observations [X(1), …, X(T)] and [Y(1), …, Y(T)], it firstly builds the shadow manifolds for X and Y as Eq. (1) and (2).

$$M_{X,t} = \left[X_t, X_{t-1}, X_{t-2}, \ldots, X_{t-(E-1)}\right] \tag{1}$$

$$M_{Y,t} = \left[Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_{t-(E-1)}\right], \tag{2}$$

where $M_{X,t}$ and $M_{Y,t}$ are the shadow manifold of X and Y at time t respectively, X_t and Y_t are observations of X and Y at time t respectively, E is a parameter defining the number of dimensions. Next, $\widehat{Y}_t$ is estimated as a linear sum of $E+1$ neighbors, which are defined through $M_{X,t}$ in following steps:

(1) Find $E+1$ nearest neighbors of $M_{X,t}$ from the shadow manifold of X, (2) Obtain the contemporaneous Y_i using the time indices of the $E+1$ shadow manifold of X.

$$\widehat{Y}_t \middle| M_{X,t} = \sum_{i=1}^{E+1} w_i Y_i, \tag{3}$$

where Y_i are contemporaneous values of Y and w_i is its weight calculated according to the distance between $M_{X,t}$ and its ith nearest neighbor on $M_{X,i}$ (as shown in Eq. (4)).

$$w_i = u_i \Big/ \sum_{j=1}^{E+1} u_j \tag{4}$$

where

$$u_i = \exp\left(-\frac{d\left(M_{x,k_i}, M_{x,k_0}\right)}{d\left(M_{x,k_1}, M_{x,k_0}\right)}\right) \tag{5}$$

where d is the Euclidean distance between two vectors and exp represents the exponential function.

The CCM use the correlation between the $\widehat{Y}_t|M_{X,t}$ and Y_t after convergence to measure the causation effect of Y on X:

$$\rho_{x\to y} = \lim_{L\to+\infty} \mathrm{cor}\left(Y_t, \widehat{Y}_t \middle| M_{X,t}\right), \tag{6}$$

where cor is Pearson correlation, and L is the size of the library.

Only a limited number of variables are required for running CCM: E (number of dimensions), (time lag) and b (number of nearest neighbors). According to previous studies (Chen et al., 2018; 2020), E, and b was set as 2 days, 3 and 4 in this research.

2.4. Partial Cross Mapping (PCM)

Partial Cross Mapping (PCM) (Leng et al., 2020) was proposed to specifically distinguish the direct causation of one variable from the indirect causation of other variables based on the partial correlation of the mutual cross mapping outputs. PCM is advantageous of identifying the direct and indirect interactions between three variables.

Taking X, Y and Z as example, the direct causation of X on Y without the indirect causation from Z is calculate in Eq. (7)

$$\rho_D = \left|Pcc\left(X, \widehat{X}^Y \middle| \widehat{X}^{\widehat{Z}^Y}\right)\right| \tag{7}$$

where $\widehat{X}^Y$ is cross mapping result of X from Y, and $\widehat{X}^{\widehat{Z}^Y}$ is the cross mapping result of X from $\widehat{Z}^Y$, which is in turn the cross mapping of Z from Y. PCC is a function that explains the partial correlation coefficient.

The PCM model is advantageous of detecting the direct and indirect interactions between three variables. However, PCM cannot directly quantify the combined indirect effects of multiple factors on variable C through variable B. To address this issue, we respectively calculated the indirect influence of individual meteorological factors on $PM_{2.5}$ concentrations through one specific factor (e.g. humidity) and the summed up indirect influence of all other factors was regarded as the overall indirect influence of this specific factor on $PM_{2.5}$ concentrations.

2.5. Structural Equation Modelling (SEM)

Structural Equation Modelling (SEM) is a composition of mechanistic and data-driven method to model the cause-effect relationships between multiple variables in complex systems (Grace et al., 2012). SEM includes three progressively refined models, conceptual model, graphical model and mathematical model. In the conceptual model, concepts and links among variables are built to explain the target system. Following this, those concepts are refined to variables (observed, latent or composite variables) to build relations among them in the graphical model. The graphical model is finally transformed into mathematical model consisted by mathematical equations. SEM is established following the equations as follow.

$$x = \Lambda\xi + \sigma \tag{8}$$

$$y = K\eta + \varepsilon \tag{9}$$

$$\eta = B\eta + \Gamma\xi + \zeta \tag{10}$$

where Eqs. (8) and (9) are measurement models and (10) is structural model; x is observed exogenous variable, ξ is its latent variable and Λ is corresponding coefficient; y is observed endogenous variable, η is its latent variable and K is corresponding coefficient; B and Γ are coefficients of latent variables; σ, ε and ζ are errors after fitting.

When setting SEM for quantifying $PM_{2.5}$-meteorology relationship, a two-layer structure was established. For the first layer, two most important meteorological factors are selected as the primary influencing factors. For the second layer, four meteorological factors, which were strongly related to the two factors in the first layer, are selected as secondary influencing factors. To acquire the primary and secondary factors for each city, we employed an automatic strategy to respectively calculate ρ (using CCM) of all meteorological factors and selected the two factors with the largest ρ as the primary factors in the first layer. Similarly, we calculated ρ of all other parameters on the two primary factors and selected four factors with the largest ρ as secondary factors in the second layer.

2.5.1. Geographical Detector (GD)

Geographical Detector (GD) measures the relationship between the target variable and the explanatory variables based on the consistency of their spatial distribution pattern (Wang et al., 2010, 2016a, 2016b). The consistency is measured using the factor detector, q as shown in (11).

$$q = 1 - \frac{\sum_{h=1}^{H} n_h \delta_h^2}{N\delta^2} \tag{11}$$

where N is the total size the study area, which is stratified into H strata using the explanatory variable. n_h is the size of stratum h. δ_h^2 is the variance of the target variable in stratum h. δ^2 Is total variance of target variable in the whole study area. The significance of q can be tested using the non-central F distribution, as shown in 12.

$$F \sim F(H-1, n-H, \lambda) \tag{12}$$

where n is the sample size; λ is noncentrality.

Different from other temporal statistical models, GD is a spatial model. Therefore, GD is not capable of detecting the $PM_{2.5}$-Meteorology relationship in each country. Instead, GD can calculate the general influence on $PM_{2.5}$ concentrations across China, presenting as one q for each meteorological factor.

3. Results

Based on CA (Correlation Analysis), P-CA (Partial Correlation Analysis), CCM (Convergent Cross Mapping), SEM (Structural Equation Modelling), PCM (Partial Cross Mapping) and GD (Geographical Detector), $PM_{2.5}$-Meteorology interactions across China were examined respectively. Due to the difference in algorithms, the absolute measurement of $PM_{2.5}$-Meteorology causation from different models cannot be directly compared. To address this issue, we presented the multi-model comparison from different perspectives.

The lack of reference data is always the major challenge for evaluating model performance in causation inference. This is because the causation between two variables were very hard to observe or measure and this is the major reason why the comparison of different statistical models for causation inference was rarely conducted. To address this issue, we employed the strategy of referring to some well-accepted prior knowledge concerning $PM_{2.5}$-Meteorology association, which were extracted using atmospheric models (e.g. WRF-CAM_x and GEOS-Chemistry) and regression models. For instance, Yang et al. (2016) suggested that wind was the dominant meteorological driver for $PM_{2.5}$ pollution in Eastern China (including Beijing-Tianjin-Hebei Region, BTH), while wind and precipitation were major meteorological driver in Southern China (including Pearl River Delta). Similarly, Wang et al. (2021) revealed that wind contributed most to the decay of haze episodes in BTH. Gui et al. (2019) also suggested a major influence of wind on $PM_{2.5}$ in Pearl River Delta. These findings were similar and consistent with our prior-knowledge, and can be considered as reference for understanding the reliability of these statistical models. Only if the model outputs were consistent with these $PM_{2.5}$-Meteorology prior-knowledge, we can regard the model is at least qualitatively reliable for causation inference in atmospheric environment.

3.1. The dominant meteorological factor for $PM_{2.5}$ across China

The relative strength (the ranking) of the calculated influence of individual meteorological factors on $PM_{2.5}$ concentrations can be compared. Here we take the meteorological factor, which has the largest

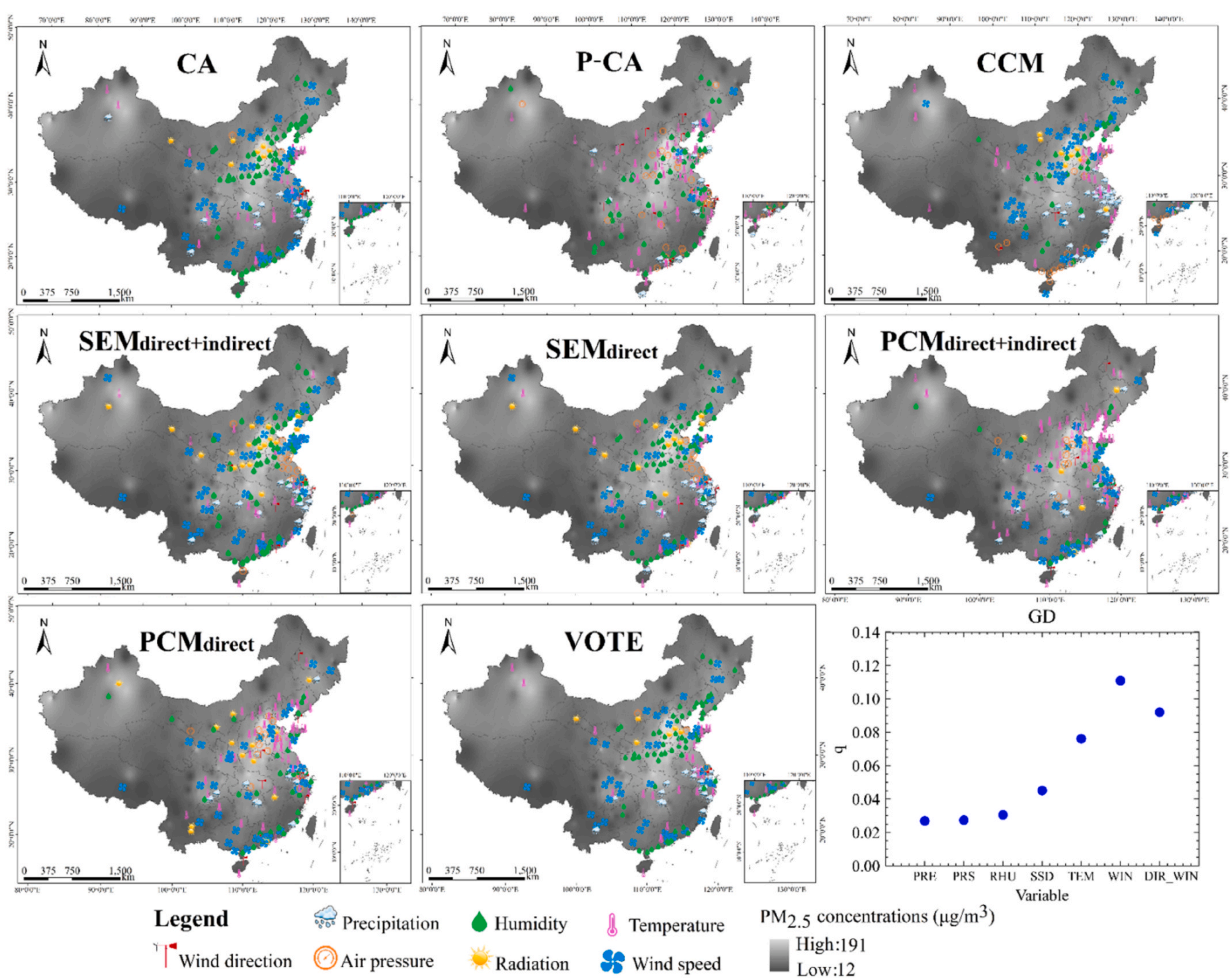

Fig. 1. The dominant meteorological factor for $PM_{2.5}$ concentrations across China extracted by different models.

association with $PM_{2.5}$ concentrations, as the dominant factor. $PM_{2.5}$-Meteorology interactions present notable seasonal patterns and thus we respectively calculated the dominant meteorological factor for $PM_{2.5}$ concentrations across China in winter and other seasons, as shown in Fig. 1 and Fig. S1. Outputs from multiple models presented notable differences in all four seasons. Since $PM_{2.5}$ concentrations were highest and $PM_{2.5}$-Meteorology interactions were strongest and presented the most stable patterns in winter (Chen et al., 2020), for a clearer explanation, this manuscript mainly discussed the strongest $PM_{2.5}$-Meteorology causation in winter. Some major characteristics of model performances are as follows:

a. P-CA and PCM, which specifically aims to separate the underlying direct and indirect interactions between multiple influencing factors, led to outputs remarkably different from CA, CCM, SEM, and was not fully consistent with previous findings and prior-knowledge (e.g. the dominant role of wind and humidity). The relatively poor performance of P-CA and PCM revealed the extreme difficulty in isolating multi-direction $PM_{2.5}$-Meterology interactions in the complicated atmospheric environment. The imprecise extraction of indirect causation can meantime lead to large uncertainties in the extracted direct causation.
b. CA, SEM and CCM suggested that humidity, wind speed and temperature were three major dominant factors for most Chinese cities. Despite this rough similarity, low consistence existed between these models. Amongst 190 cities, CA-SEM, CA-CCM and SEM-CCM respectively presented 114, 65 and 60 cities with a similar dominant factor. The low consistence may be attributed to the complicated and strong $PM_{2.5}$-Meteorology interactions. For a specific city, there may be several meteorological factors strongly associated with $PM_{2.5}$. How to clearly isolate disturbing factors and effectively identify the slight difference between major influencing factors poses a real challenge. A slight underestimation or overestimation could cause a notable variation in extracted dominant factors. Compared with other models, CCM specifically identified the dominant role of wind in Northern China, and the dominant role of precipitation and wind in Southern China, which agreed well with previous findings.
c. To better understand the consistence between models, we conducted a vote process as follows. For each city, we compared the dominant meteorological factor extracted using CA, CCM, SEM and PCM, and regarded a similar output from two or more models as a voted output. Among 190 cities, there were 178 cities with a voted factor. Specifically, voted output was the same with CA, CCM, SEM and PCM output in 160, 96, 132 and 66 cities respectively (Table S1). The largest similarity with other models does not mean CA achieved a most reliable output. Instead, it revealed that many advanced statistical models were established based on CA.

Since there can be two or more factors that exert a strong influence on $PM_{2.5}$ concentrations, the sole consideration of one dominant factor adds difficulties in reaching a consistence between multiple models. Therefore, we also compared two, instead of one, dominant factors (two factors with largest causation) from multiple models. We set that if two models shared at least one of two dominant factors, they had a similar output. The output of two-factor comparison is shown in Table S2. From a two-factor perspective, CA-CCM (170), CA-SEM (172), CA-PCM (141), CCM-SEM (147), PCM-CA (141), PCM-CCM (124) and PCM-SEM (127) indicated a significantly enhanced consistence between models outputs.

The one-factor and two-factor comparison highlighted the challenge in causation inference in atmospheric environment. When aiming for some precise conclusions (e.g. the sole dominant meteorological factor), it is necessary to clearly isolate the multi-direction interactions between many factors and reliably compare the slight difference between them, which is highly challenging and makes outputs from different models varied notably. As a comparison, when aiming for some rough conclusions (e.g. two or more major influencing factors), the underestimation or overestimation of causation had less a influence on the extraction of multiple major factors. In this regard, outputs from different models presented a high consistence.

The statistics of the causation of individual meteorological factors on $PM_{2.5}$ across China.

In addition to the dominant meteorological factor, we also presented the statistics of the causation of individual meteorological factors on wintertime $PM_{2.5}$ based on different models (Fig. 2). As shown in Fig. 2, the large difference of $PM_{2.5}$-Meteorology relationship across China is observed in all models. Meanwhile, the rain cloud plots of these models presented limited consistence, indicating the capability of reducing the disturbing factors and quantifying specific $PM_{2.5}$-Meteorology causation varied significantly. Due to the large model variations, a close output from multiple models was not available for cross-verification.

In terms of the mean $PM_{2.5}$-Meteorology causation across China, we summarized Top 3 meteorological factors from different models as follows: CA (Humidity, Wind, Temperature), P-CA (Temperature, Humidity, Precipitation), CCM (Temperature, Humidity, Wind), SEM (Humidity, SSD, Wind), SEM_{direct} (Humidity, Wind, Temperature), $SEM_{indirect}$ (Radiation, Humidity, Temperature), PCM (Temperature, Wind, Humidity), PCM_{direct} (Temperature, Wind, Humidity), $PCM_{indirect}$ (Wind direction, Precipitation, Temperature), and GD (Wind, Wind direction and Temperature). Without quantitive reference, we cannot easily judge which model performs the best according to the ranking of TOP 3 meteorological factors. However, despite a difference in ranking, most models (except for GD and P-CA) revealed that Temperature, Humidity, Wind exerted a major influence on $PM_{2.5}$ concentrations, which were consistent with our $PM_{2.5}$-Meteorology prior-knowledge and thus regarded as roughly reliable. On one hand, due to complicated interactions between multiple atmospheric factors, the exact $PM_{2.5}$-Meteorology causation in a specific city varied significantly between multiple models. On the other hand, some rough patterns of $PM_{2.5}$-Meteorology at the national scale could be reliably extracted by most models. Generally, this $PM_{2.5}$-Meteorology research proved the feasibility of most statistical models in inferring rough patterns, yet their uncertainty in reliably quantifying extract causation in the atmospheric environment.

Different from other models, PCM and SEM aim to further isolate the direct and indirect causation. For this $PM_{2.5}$-Meteorology research, the TOP 3 factors extracted by PCM_{direct} and SEM_{direct} were generally similar to other models, presenting a reasonable output. Meanwhile, $PCM_{indirect}$ and $SEM_{indirect}$ outputs varied significantly and came to no shared findings. Specifically, the two-layer SEM was established based on sufficient prior-knowledge and hypothetical setting, while $PCM_{indirect}$ led to nearly similar outputs for all meteorological factors except for wind direction, which deviated largely from previous studies. In this regard, both SEM and PCM failed to reveal unknown and reliable indirect $PM_{2.5}$-Meteology causation in the atmospheric environment.

3.2. The inner-interactions between models and model evaluation

Despite a large and hard-to-interpret model differences, we attempted to extract some inner-connections between these modelling through model clustering. Firstly, the causation between each sub-factor and wintertime $PM_{2.5}$ in 190 cities was averaged and the 17 mean causation value formed a vector for each model. Based on these vectors, the model clustering was conducted according to the means of square Euclidean distance for Between-groups linkage method (Yuan et al., 2013; Murtagh and Legendre, 2014).

The result of model clustering (as shown in Fig. 3) suggested CCM, CA and PCM, which aim to extract the complete causation between two variables, were categorized to the same group. Meanwhile, $PCM_{indirect}$, PCM_{direct} and PCA, which aim to separate the direct and indirect causation, were categorized to another group. SEM uniquely included a two-layer structure and thus categorized into one individual group. The clustering outputs based on $PM_{2.5}$-Meteorology research were highly

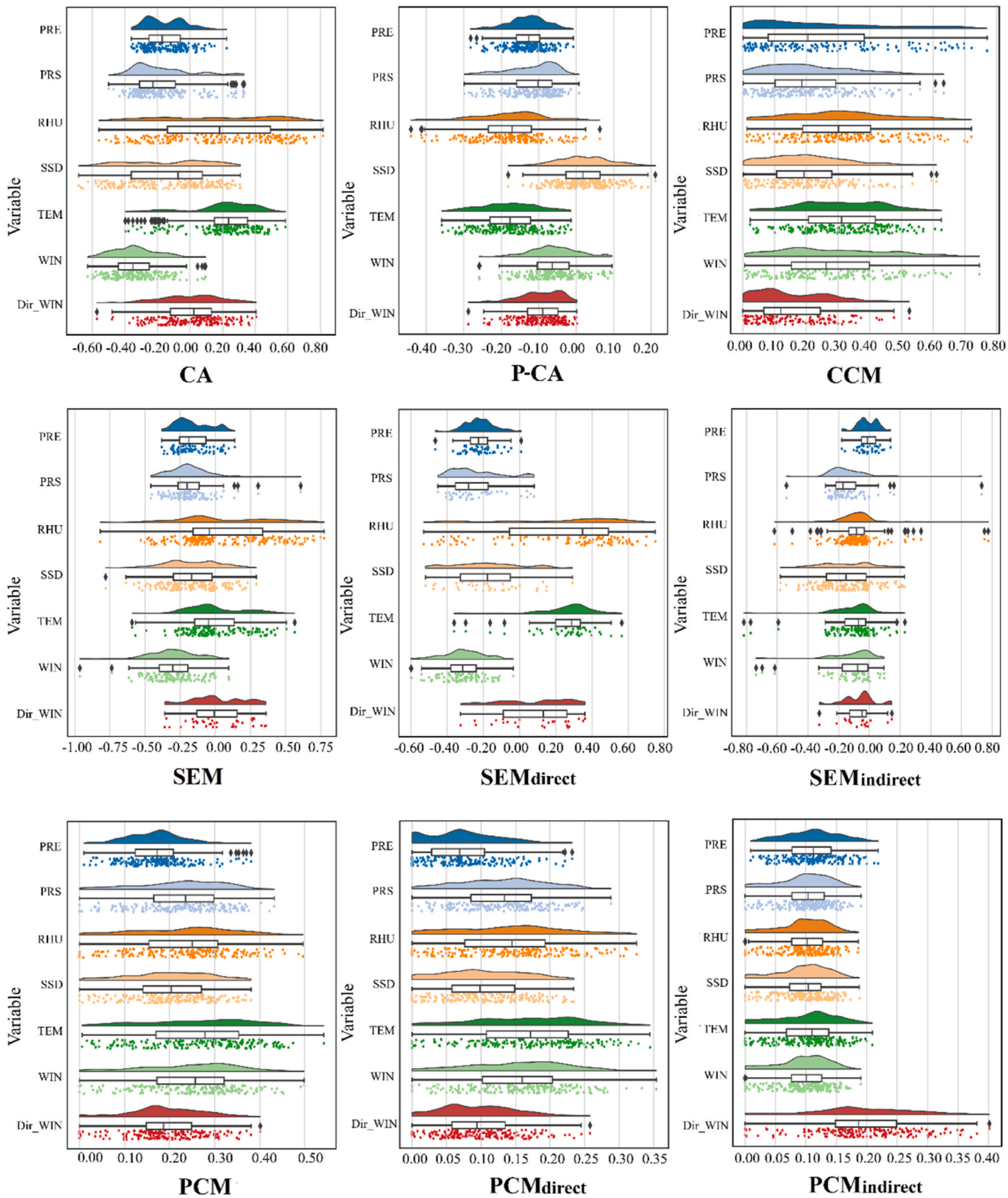

Fig. 2. Raincloud plots of the influence of individual meteorological factors on wintertime $PM_{2.5}$ concentrations across China calculated using different models. Note: the extracted causation from all models were normalized to [-1, 1]. PRE stands for precipitation, PRS stands for air pressure, RHU stands for humidity, SSD stands for radiation, TEM stands for temperature, WIN stands for wind speed, Dir_WIN stands for wind direction.

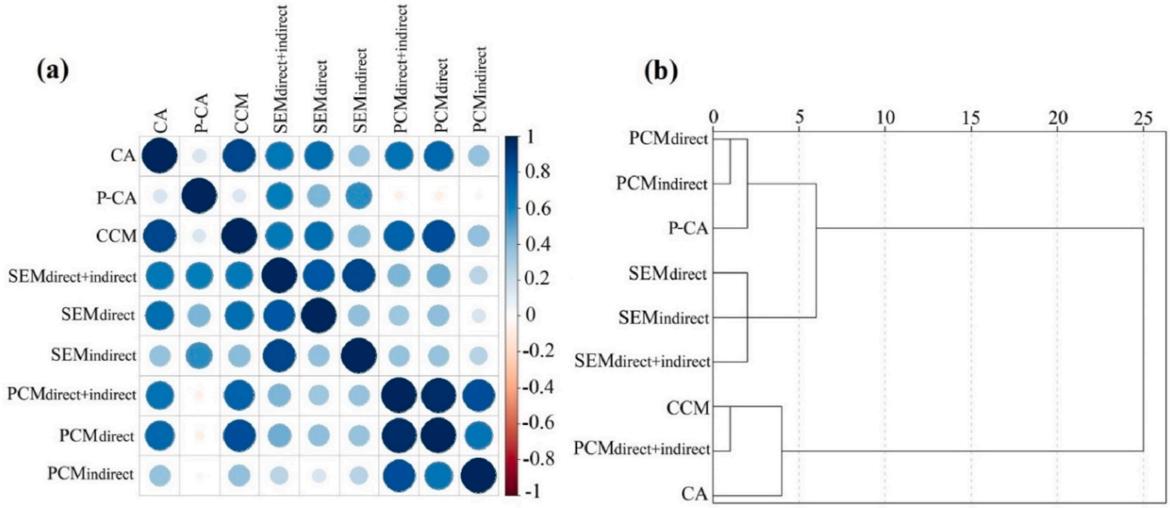

Fig. 3. Clustering analysis based on the $PM_{2.5}$-Meteorology outputs from different models. a. The correlation between multiple models b. The clustering of multiple models.

consistent with the fundamental mechanisms of these statistical models. This indicated that although inner-interactions were highly complicated in atmospheric environment, these statistical models effectively conducted causation inferences following their strategies. In this regard, it is crucial to better interpret the advantages and limitations of different statistical models in atmospheric research, and select proper models accordingly.

Based on the model algorithms and their performances in $PM_{2.5}$-Meteorology research, a brief evaluation of these statistical models were as follows:

GD (Geographical Detector): In this research, GD suggested Wind, Wind direction and Temperature were three major influencing factors for $PM_{2.5}$ across China, which were not consistent with other temporal models. One major reason was that there were many confounding variables across regions, leading to large uncertainties in extracted causation. Therefore, GD or other spatial models may not be suitable for inferring causation in atmospheric environment.

CA (Correlation Analysis): $PM_{2.5}$-Meteorology relationship is nonlinear and affected by other disturbing factors. So although CA acquired some reasonable patterns at the national scale (e.g. the dominant role of temperature and humidity), quantitative $PM_{2.5}$-Meteorology causation in specific cities can be largely underestimated or overestimated by CA (Sugihara et al., 2012; Chen et al., 2017).

P-CA (Partial Correlation Analysis): P-CA aims to isolate the inner-interactions between multiple variables and focus on the direct influence of one variable on the target variable. In this research, P-CA led to model outputs largely different from other models. One potential reason was that P-CA failed to separate the non-linear and uncertain interactions between massive meteorological factors, significantly affecting the extraction of causation between $PM_{2.5}$ and individual factors.

PCM (Partial Cross Mapping): PCM aims to further separate the direct causation of variable A on B and the indirect causation of variable A on B through C. In this research, PCM failed to effectively identify and quantify the strong indirect influence from other meteorological factors, and simply highlight the role of wind direction. This suggested that PCM, which is advantageous of separating the direct and indirect interactions in simple systems, may not be suitable for direct-indirect causation inference in atmospheric environment with massive atmospheric factors.

SEM (Structural Equation Modelling): In this research, SEM aimed to examine $PM_{2.5}$-Meteorology causation by establishing a two-layer structure and presenting both direct and indirect causation. While SEM_{direct} outputs were generally reasonable, $SEM_{indirect}$ outputs, characterized with the dominant indirect causation of precipitation, were not consistent with our prior-knowledge of $PM_{2.5}$-Meteorology (the strong indirect causation of wind and atmospheric pressure). The unsatisfactory outputs were mainly attributed to large uncertainties involved in the establishment of SEM. Firstly, SEM requires the setting of a specific function (e.g. Linear and exponential) between two variables. However, $PM_{2.5}$-Meteotogloy causation is non-Linear and uncertain. Secondly, SEM requires a pre-description of the interaction between two variables (e.g. wind and humidity) using a coefficient (e.g. 0.3), which required reliable prior-knowledge or hypothesis. Without prior-knowledge, we could only employ other models (e.g. CCM) to extract causation between meteorological factors, which significantly reduced the feasibility and transferability of the sole use of SEM.

CCM: CCM aims to improve the causation inference between two variables by removing the influences from other factors in complicated ecosystems. CCM is designed by repeatedly experimenting all possible time lags and interaction forms. Therefore, CCM is highly automatic, and little prior-knowledge or parameter setting is required. CCM is advantageous for extracting non-Linear causation. These characteristics making CCM especially suitable for $PM_{2.5}$-Meteorology causation inference. Similar to other models, CCM also revealed the dominant role of temperature and humidity for $PM_{2.5}$ concentrations across China. Furthermore, CCM uniquely identified the dominant influence of wind on wintertime $PM_{2.5}$ in BTH, and the dominant role of wind and precipitation in Pearl River Delta, which was consistent with previous studies, yet limitedly detected by other models.

4. Discussion

Given the complexities of multiple models, uncertainties exist in this multi-model comparison. Firstly, for such models as SEM, model setting has a major influence on outputs. To effectively reduce the uncertainty, we repeatedly tuned these model parameters to extract the potential largest causation. Secondly, different models generate outputs with various dimensions, which cannot be compared directly. To reduce the disturbance of the dimension issue, we considered the relative contribution (the ranking) of individual meteorological factors to $PM_{2.5}$ concentrations for model comparison. With this strategy, the $PM_{2.5}$-Meteorology research provides a reliable reference for cross-verifying multiple models.

Integrating outputs from multiple atmospheric models has been a commonly employed strategy for estimating a more reliable range. For instance, we could integrated the output from model A (70%), model B (85%) and model C (90%) and came to a more reliable conclusion that emission-cut contributed 70%~90% to the decrease of $PM_{2.5}$ in Beijing. Nevertheless, due to notable dimension differences, the outputs of causation inference from multiple statistical models cannot be directly compared or computed. To address this issue, we considered the integration of some shared relative results (e.g. the ranking of causation between $PM_{2.5}$ and individual factors) from different models. However, even from a qualitative perspective, large model variations remain, making the cross-verification and an integrated outputs from multiple statistical models not feasible. In addition to the multi-model integration, the other strategy for reliable causation inference is to confirm the robustness of one model. The $PM_{2.5}$-Meteorology research, which is a typical and challenging atmospheric research with a diversity of strongly interacted factors, provides valuable references. GD is a spatial model, which can extract the general causation between two variables at large-scales, yet fails to extract causation at a specific site. Compared with other models, the problematic outputs from GD highlighted the large disturbance of the spatial variation of multiple confounding variables on causation inference. PCM and PCA performed poorly when separating multi-direction interactions between many meteorological variables. CA can be affected by the existence of agent variables, which is a major issue in atmospheric environment, and cause largely underestimated or overestimated outputs. SEM requires sufficient prior-knowledge and hypothesis, adding extra difficulties to model setting and reducing model reliability. As a comparison, CCM might be the most suitable model for causation inference in atmospheric environment. Theoretically, firstly, CCM is designed specifically to process non-linear relationship between two variables, which fully suits the non-linear relationship between atmospheric factors. Secondly, CCM automatically considers all possible interaction forms and lagging effects between the time series of two variables, which effectively reduces the disturbing influence and avoids mirage correlation. Thirdly, CCM requires little parameter setting and prior-knowledge, removing the uncertainties caused by improper parameter setting. Practically, for the $PM_{2.5}$-Meteology research, CCM generated a reasonable output, including some general findings consistent with other models, and some unique findings consistent with previous studies. Given the inconsistence of multiple statistical models, the sole use of CCM can be a preferable strategy for causation inference in such complicated ecosystems as atmospheric environment.

In complicated ecosystems, the key for causation inference between two variables is to effectively separate the influence of disturbing factors and isolate the influence of the target factor. Furthermore, the separation of direct and indirect causation has been hotly debated and some advanced models (e.g. Partial Cross Mapping) have been intendedly proposed. Despite its effectiveness in simple systems, PCM performed poorly to clearly separate the direct and indirect $PM_{2.5}$-Meteorology interactions. The major reason is that massive atmospheric factors led to a series of underlying, or even unknown interactions. Meanwhile, PCM, or other statistical models, can only simultaneously consider the multi-direction interactions between limited factors. With the increase of total variables, the robustness of extracted indirect-causation decreases significantly. For other atmospheric research, which also involves many interacting variables, scholars should be more cautious and less ambitious. Admittedly, a comprehensive understanding of clearly explained direct and indirect causation between all variables is ideal for fully attributing and predicting atmospheric processes. However, as revealed in this $PM_{2.5}$-Meterology research, it is very challenging for current statistical models to accurately separate, quantify and interpret direct and indirect causation between various atmospheric factors. As one of the most complicated and dynamic ecosystems, the inference of direct and indirect causation in the atmospheric environment remains a major gap to fill.

Causation inference in atmospheric environment holds great importance for attributing and managing natural disasters, and should be further explored. Firstly, due to the lack of observation reference data, we considered the prior-knowledge of $PM_{2.5}$-Meteorology association as qualitative reference and the result proved that even at the qualitative level, it remained challenging to extract reliable causation patterns. For a better, quantitative evaluation of statistical models, specific experiments based on lab-experiment (e.g. smog chamber) should be designed to explore the possibility of extracting measurable causation between atmospheric variables by carefully controlling other variables. With reliable quantitative evidences, we may further verify some advanced hypothesis for causation inference in atmospheric environment. For instance, is it possible for largely varied model performance at different spatial or temporal scales? Is it possible for a good performance for causation inference in a specific region using a model, which performs poorly at a large scale? Is it possible for a much improved performance for direct and indirect causation inference in an atmospheric issue, which involves many less confounding variables? Secondly, given the low efficiency of integrating multiple statistical models, the combination of both statistical models (e.g. CCM) and atmospheric models (e.g. Climate Models or Chemical Transport Models) for causation inferences is preferred. For instance, Chen et al. (2019) employed both a statistical model, Kolmogorov–Zurbenko (KZ) filter and a climate model, WRF-CMAQ, and respectively quantified the relative contribution of meteorological variations to $PM_{2.5}$ reduction was close to 20%. Similar outputs could cross-verified the reliability of both models. Meanwhile, causation coefficients extracted from statistical models can be used to optimize parameter setting in atmospheric models. Thirdly, existing causation models, which are capable of investigating causation in simple systems, fail to fully clarify all direct and indirect interactions between massive atmospheric variables. Deep Neural Network (DNN), which has been limitedly designed for extracting, is a promising tool for revealing underlying multi-direction interactions in a complicated ecosystem. DNN is capable of simulating even thousands of variables and improving the simulation of their inner interactions through repetitive training based on long-term and large-scale observation data, the accumulation of which has become one major advantage in recent atmospheric research. Currently, DNN has been commonly employed to estimate a target variable based on a series of influencing factors, yet limitedly designed for extracting relationship or causation between variables. With the future development of relationship-oriented DNNs, the multi-direction causation between a large number of atmospheric factors may be better extracted. In conclusion, for improved causation inference in atmospheric environment, a comprehensive strategy, which considers and potentially makes full use of different approaches, is encouraged.

5. Conclusion

We employed a series of mainstream statistical models to examine $PM_{2.5}$-meteorological association. From a coarse perspective, the Top 3 major meteorological factors for $PM_{2.5}$ extracted using different models were generally consistent. From a strict perspective, the extracted dominant meteorological factor for $PM_{2.5}$ demonstrated large model variations and shared a limited consistence. Specifically, SEM and PCM, which are capable of further separating direct and indirect causation in simple systems, performed poorly to identify the direct and indirect $PM_{2.5}$-Meteorology causation. The notable model variations denied the feasibility of employing multiple models for better causation inference. Instead, the sole use of CCM, which is advantageous of dealing with non-linear causation and removing disturbing factors, is a preferable strategy. Meanwhile, given the multi-direction, uncertain interactions between massive variables, we should be more cautious and less ambitious on the separation of direct and indirect causation. For better causation inference in complicated atmospheric environment, the combination of statistical models and atmospheric models, and further exploration of Deep Neural Network can be promising strategies.

Author statement

Ziyue Chen: Conceptualization, Methodology, Writing - review & editing, Supervision. **Miaoqing Xu:** Writing - original draft, Conceptualization, Data curation, Formal analysis, Methodology. **Bingbo Gao:** Writing - original draft, Formal analysis. **George Sugihara:** Conceptualization. **Feixue Shen:** Data curation & Formal analysis. **Yanyan Cai:** Data curation& Formal analysis. **Anqi Li:** Data curation& Formal analysis. **Qi Wu:** Formal analysis& Formal analysis. **Lin Yang:** Conceptualization. **Qi Yao:** Data curation& Formal analysis. **Xiao Chen:** Data curation& Formal analysis. **Jing Yang:** Data curation& Formal analysis. **Chenghu Zhou:** Conceptualization & Writing - review & editing. **Manchun Li:** Conceptualization & Writing - review & editing.

Funding sources

This work was supported by the National Natural Science Foundation of China (Grant No. 42171399, 41971054) and Beijing Municipal Natural Science Foundation, China (Grant No. 8202031).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2022.119057.

References

Azam, M., 2016. Does environmental degradation shackle economic growth? A panel data investigation on 11 Asian countries. Renew. Sustain. Energy Rev. 65, 175–182.
Baklanov, A., et al., 2014. Online coupled regional meteorology chemistry models in Europe: current status and prospects. Atmos. Chem. Phys. 14 (1), 317–398.
Bhardwaj, S., et al., 2013. Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model. Sol. Energy 93, 43–54.
Bhat, T.H., Guo, J., Farzaneh, H., 2021. Air pollution health risk assessment (ap-hra), principles and applications. Int. J. Environ. Res. Publ. Health 18 (4), 1935.
Bianco, L., et al., 2011. Diurnal evolution and annual variability of boundary-layer height and its correlation to other meteorological variables in California's central valley. Boundary-Layer Meteorol. 140 (3), 491–511.
Bryan, K., et al., 2020. The health and well-being effects of drought: assessing multi-stakeholder perspectives through narratives from the UK. Climatic Change (1), 1–23.
Calvin, K., et al., 2013. Implications of simultaneously mitigating and adapting to climate change: initial experiments using GCAM. Climatic Change 117 (3), 545–560.
Chen, Z, et al., 2017. Detecting the causality influence of individual meteorological factors on local $PM_{2.5}$ concentration in the Jing-Jin-Ji region. Sci. Rep. 7, 407352.
Chen, Z, et al., 2018. Understanding meteorological influences on PM2.5 concentrations across China: a temporal and spatial perspective. Atmos. Chem. Phys. 18 (8), 5343–5358.
Chen, Z., et al., 2019. The control of anthropogenic emissions contributed to 80% of the decrease in PM2.5 concentrations in Beijing from 2013 to 2017. Atmos. Chem. Phys. 19 (21), 13519–13533.
Chen, Z., et al., 2020. Influence of meteorological conditions on PM2. 5 concentrations across China: a review of methodology and mechanism. Environ. Int. 139, 105558.
Chotamonsak, C., et al., 2011. Projected climate change over Southeast Asia simulated using a WRF regional climate model. Atmos. Sci. Lett. 12 (2), 213–219.
Colarco, P.R., et al., 2014. Impact of radiatively interactive dust aerosols in the NASA GEOS-5 climate model: sensitivity to dust particle shape and refractive index. J. Geophys. Res. Atmos. 119 (2), 753–786.
El-Sayed, A., Kamel, M., 2020. Climatic changes and their role in emergence and re-emergence of diseases. Environ. Sci. Pollut. Control Ser. 27 (18), 22336–22352.
Fang, X., et al., 2017. Isolating and quantifying the effects of climate and CO2 changes (1980–2014) on the net primary productivity in arid and semiarid China. Forests 8 (3), 60.
Grace, J.B., et al., 2012. Guidelines for a graph-theoretic implementation of structural equation modeling. Ecosphere 3 (8).
Gui, K., et al., 2019. Satellite-derived pm2.5 Concentration Trends over Eastern china from 1998 to 2016: Relationships to Emissions and Meteorological Parameters. Environmental Pollution.
Lee, B.P.Y., Davies, Z.G., Struebig, M.J., 2017. Smoke pollution disrupted biodiversity during the 2015 El Niño fires in Southeast Asia. Environ. Res. Lett. 12 (9), 094022.
Leng, S., et al., 2020. Partial cross mapping eliminates indirect causal influences. Nat. Commun. 11, 2632.
Liu, Y., Paciorek, C.J., Koutrakis, P., 2009. Estimating regional spatial and temporal variability of PM2.5 concentrations using satellite data, meteorology, and land use information. Environ. Health Perspect. 117 (6), 886–892.
Luo, X.S., et al., 2017. Effects of emission control and meteorological parameters on urban air quality showed by the 2014 Youth Olympic Games in China. Fresenius Environ. Bull. 26, 4798–4807.
Mizuta, R., et al., 2017. Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. Bull. Am. Meteorol. Soc. 98 (7), 1383–1398.
Molod, A., et al., 2012. The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna.
Morice, C.P., et al., 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. J. Geophys. Res. Atmos. 117 (D8).
Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J. Classif. 31 (3), 274–295.
Pagani, V., et al., 2017. Improving cereal yield forecasts in europe - the impact of weather extremes. Eur. J. Agron. 89, 97–106.
Saikawa, E., et al., 2014. Global and regional emissions estimates for N 2 O. Atmos. Chem. Phys. 14 (9), 4617–4641.
Shahbazi, H., et al., 2017. Investigating the influence of traffic emission reduction plans on Tehran air quality using WRF/CAMx modeling tools. Transport. Res. Transport Environ. 57, 484–495.
Sherwen, T., et al., 2016. Global impacts of tropospheric halogens (Cl, Br, I) on oxidants and composition in GEOS-Chem. Atmos. Chem. Phys. 16 (18), 12239–12271.
Sugihara, G., et al., 2012. Detecting causality in complex ecosystems. Science 338 (6106), 496–500.
Skjøth, C.A., Geels, C., 2013. The effect of climate and climate change on ammonia emissions in Europe. Atmos. Chem. Phys. 13 (1), 117–128.
Srinivas, C.V., et al., 2013. Simulation of the Indian summer monsoon regional climate using advanced research WRF model. Int. J. Climatol. 33 (5), 1195–1210.
Tai, A.P.K., Mickley, L.J., Jacob, D.J., 2010. Correlations between fine particulate matter (pm2.5) and meteorological variables in the United States: implications for the sensitivity of pm2.5 to climate change. Atmos. Environ. 44 (32), 3976–3984.
Tie, X., et al., 2017. Severe pollution in China amplified by atmospheric moisture. Sci. Rep. 7 (1), 1–8.
Tie, X., et al., 2019. Ozone enhancement due to the photodissociation of nitrous acid in eastern China. Atmos. Chem. Phys. 19 (17), 11267–11278.
Trnka, M., et al., 2014. Adverse weather conditions for european wheat production will become more frequent with climate change. Nat. Clim. Change 4 (7), 637–643.
VanderWeele, T.J., Knol, M.J., 2014. A tutorial on interaction. Epidemiol. Methods 3 (1), 33–72.
Wang, L., et al., 2016a. Increasing flash droughts over China during the recent global warming hiatus. For. Rep. 6, 30571.
Wang, J.F., et al., 2016b. A measure of spatial stratified heterogeneity. Ecol. Indicat. 67, 250–256.
Wang, J., et al., 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. Int. J. Geogr. Inf. Sci. 24, 107–127.
Wang, X., et al., 2021. Dominant synoptic patterns associated with the decay process of PM2.5 pollution episodes around Beijing. Atmos. Chem. Phys. 21, 2491–2508.
Xing, J., et al., 2017. Impacts of aerosol direct effects on tropospheric ozone through changes in atmospheric dynamics and photolysis rates. Atmos. Chem. Phys. 17 (16), 9869–9883.
Yang, Y., et al., 2016. Increase in winter haze over eastern China in recent decades: roles of variations in meteorological parameters and anthropogenic emissions. J. Geophys. Res. Atmos. 121 (21), 13, 050.
Yuan, F., Zhan, Y., Wang, Y., 2013. Data density correlation degree clustering method for data aggregation in WSN. IEEE Sensor. J. 14 (4), 1089–1098.
Zhang, Y., et al., 2018. Climate variability decreases species richness and community stability in a temperate grassland. Oecologia 188 (1), 1–10.
Zhong, J., et al., 2017. Relative contributions of boundary-layer meteorological factors to the explosive growth of PM 2.5 during the red-alert heavy pollution episodes in Beijing in December 2016. J. Meteorol. Res. 31 (5), 809–819.
Zhou, S., et al., 2018. Shift of daily rainfall peaks over the Beijing–Tianjin–Hebei region: an indication of pollutant effects? Int. J. Climatol. 38 (13), 5010–5019.