Electroinc Supplementary Material to: Understanding the Soil Temperature Variability at Different Depths: Effects of Surface Air Temperature, Snow Cover, and the Soil Memory

Haoxin ZHANG[1,2], Naiming YUAN[1], Zhuguo MA[1], and Yu HUANG[3]

[1]*Key Laboratory of Regional Climate–Environment Research for Temperate East Asia*, *Institute of Atmospheric Physics*, *Chinese Academy of Sciences*, *Beijing* 100029, *China*

[2]*University of Chinese Academy of Sciences*, *Beijing* 100049, *China*

[3]*Laboratory for Climate and Ocean–Atmosphere Studies*, *Department of Atmospheric and Oceanic Sciences*, *School of Physics*, *Peking University*, *Beijing* 100871, *China*

ESM to : Zhang, H. X., N. M. Yuan, Z. G. Ma, and Y. Huang, 2021: Understanding the soil temperature variability at different depths: Effects of surface air temperature, snow cover, and the soil memory. *Adv. Atmos. Sci.*, https://doi.org/10.1007/s00376-020-0074-y. (in press)

1. Introduction

Figures S1 to S12 are provided in this file, as a supplement to detect the effects of surface air temperature (AT), snow depth (SD), and the soil memory on the soil temperature (ST). To demonstrate the advantage of the time-lagged convergent cross mapping (CCM) under the strong coupling conditions and the presence of noise, we present two numerical models in Figs. S1–S3. From the linear correlation aspect, the time-delayed relationships between the AT and ST, and their differences for different layers, are shown in Fig. S4. The optimal embedding dimension (E) and the nonlinearity of the AT and ST time series are calculated by the simplex projection and S-map methods, respectively, in Figs. S5 and S6. Results of the pairwise asymmetric inference (PAI) to confirm the AT effects identified by the CCM are shown in Fig. S7. Figure S8 shows that the linear trends of the AT and STs in each season from 1998 to 2013 and the AT are always consistent with the trends of the STs. For the wintertime trends, however, the AT is cooling while the STs are warming, in northeastern China. In this region, the wintertime SD and its increasing trends are shown in Fig. S9. This may be the reason why the trends of the STs diverged from the trends of the AT in winter during the recent global warming hiatus period. Since the winter AT effects are blocked, the effects of SD and the soil memory on winter STs become more important in northeastern China. In Fig. S10, the optimal embedding dimension (E) and the nonlinearity of the SD time series are calculated by the simplex projection and S-map methods, respectively. In Fig. S11, a nonlinear causality analysis method, CCM, is used to identify the

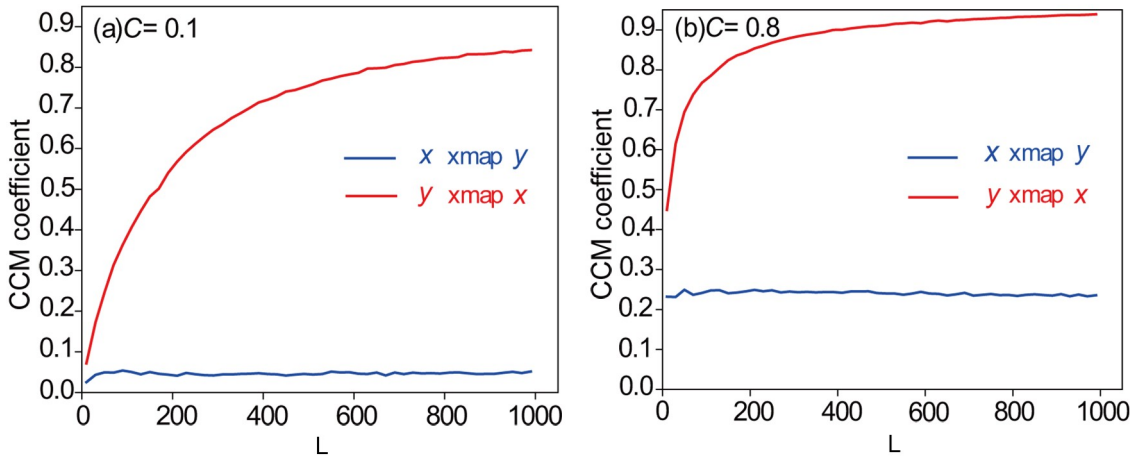

Fig. S1. CCM analyses of the coupled logistic systems [Eq. (S1)] in the cases of moderate coupling strength (a) and strong coupling strength (b).

*The online version of this article can be found at https://doi.org/10.1007/s00376-020-0074-y.

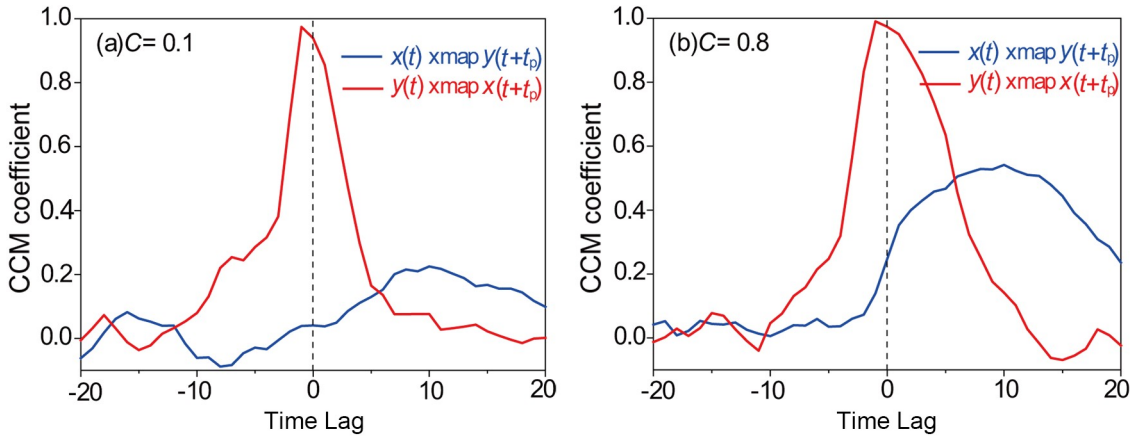

Fig. S2. Time-lagged CCM analyses of the coupled logistic systems [Eq. (S1)] in the cases of moderate coupling strength (a) and strong coupling strength (b).

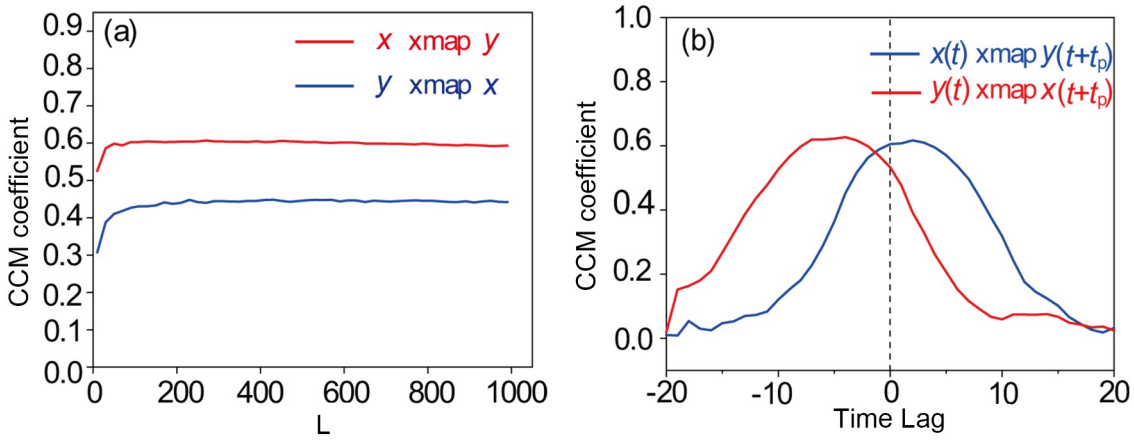

Fig. S3. (a) CCM and (b) time-lagged CCM analyses of the stochastic processes [Eq. (S2)].

causal effects of the SD and the soil memory on the winter STs. Results of the PAI to confirm the SD effects identified by the CCM are shown in Fig. S12.

2. Data information

The data used in these figures were obtained from the China Meteorological Administration, including the monthly mean surface AT records (recorded at a height of 1.5 m), the monthly mean ST records at the nine layers of 0 cm, 5 cm, 10 cm, 15 cm, 20 cm, 40 cm, 80 cm, 160 cm and 320 cm (ST0, ST5, ST10, ST15, ST20, ST40, ST80, ST160 and ST320, respectively), and the SD. All the data cover the period 1960–2013. For each element, only those stations with no more than 5.5% missing data during 1960–2013 were selected, and the gaps caused by the missing data in the time series were filled using a simple linear interpolation algorithm. The data records of ST0 are inhomogeneous owing to the change in observational infrastructure in the year 2005.

3. Advantages of the time-lagged CCM method

To demonstrate the advantage of the time-lagged CCM method under strong coupling conditions and the presence of noise, we present two numerical models as follows:

3.1. *Testing the strong coupling effect*

$$\begin{cases} x(t+1) = x(t)[3.8 - 3.8x(t)] \\ y(t+1) = y(t)[3.1 - 3.1y(t) - Cx(t)] \end{cases}. \tag{S1}$$

Equation (S1) shows the coupled logistic systems, and the parameter C controls the coupling strength between x and y.

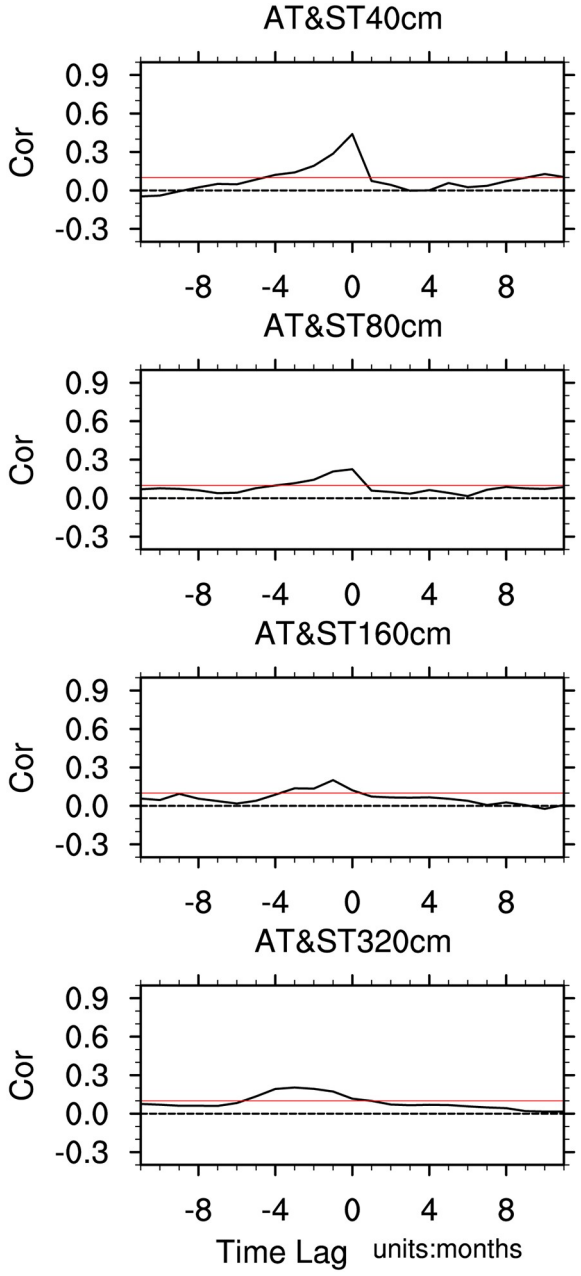

Fig. S4. Time-delayed correlation coefficients between the AT and the STs (ST40, ST80, ST160 and ST320, from the top panel to bottom panel) averaged over all stations. Above the red line indicates the correlation coefficients are statistically significant at the 95% confidence level. The black dashed line indicates where correlation coefficients are equal to 0. In each panel, the value on the horizontal axis, where the correlation coefficient is maximum, represents the time taken to propagate the signals. Negative (positive) values on the horizontal axis indicate the variations in the AT are ahead (behind) those of the ST. From the top panel to bottom panel, the aforementioned horizontal axis values are 0, 0, −1 and −3. This implies that the AT has impacts on the STs, and it takes 0 months, 0 months, 1 month and 3 months for the effects of AT to reach the soil at the 40 cm, 80 cm, 160 cm and 320 cm layers, respectively. The effects of the ST on the AT cannot be exhibited by this method.

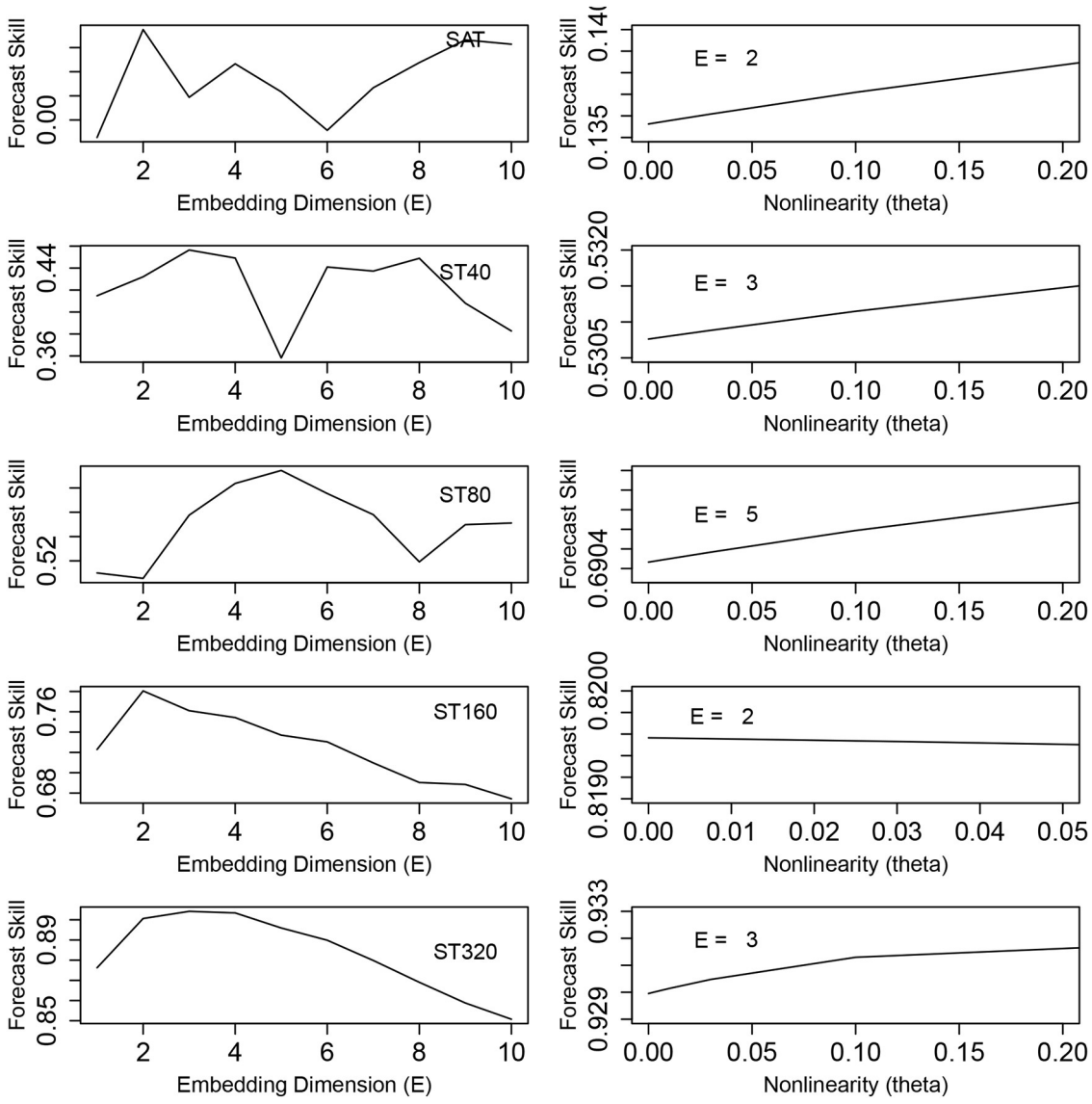

Fig. S5. Left-hand column: Forecast skills calculated by simplex projection at different embedding dimensions (E) for the monthly records of the AT, ST40, ST80, ST160 and ST320 at Lanzhou station. The best E is chosen when the forecast skill is optimal. According to the best E, the nonlinearities of the time series were detected by the S-map method (right-hand column).

When $C = 0.1$, the coupled systems have a moderate coupling strength. When C is set to 0.8, it is acting as a strong coupling.

According to Sugihara et al. (2012) and Ye et al. (2015), when $C = 0.1$, both the CCM and time-lagged CCM methods can accurately infer the causal direction between x and y. However, when $C = 0.8$, Sugihara et al. (2012) found that CCM cannot infer the correct causal direction, and this is because the strong coupling induces a general synchrony effect for x and y. Hence, Ye et al. (2015) demonstrated that time-lagged CCM can accurately infer the causal direction, in the case of synchrony induced by strong coupling.

In the following, we briefly report the results from experiments to test the strong coupling impact on the performances of CCM and time-lagged CCM, by analyzing the system shown by Eq. (S1).

3.1.1. *Inferring causality through CCM*

Figure S1a shows that when the coupling strength is moderate (C = 0.1), CCM can infer causal direction for x and y. The cross-mapping skill from y to x is of high magnitude, and this is due to the fact that the information content of x is shared with y. However, the cross-mapping skill from x to y is nearly zero, and this is due to the fact that the information content of y is not shared with x. This means that x causes y, but y does not cause x, which is consistent with the real dynamics of Eq. (S1).

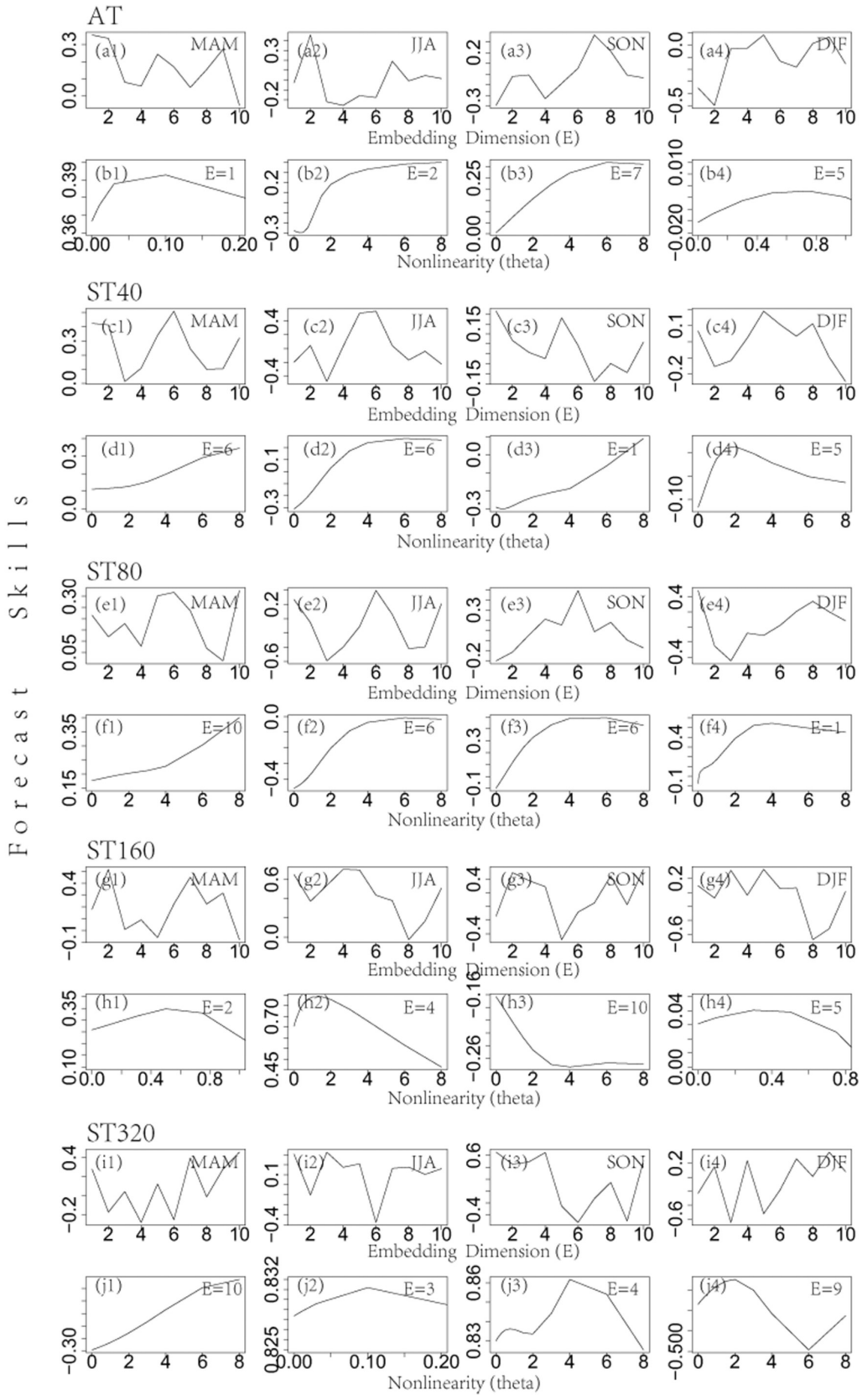

Fig. S6. (a), (c), (e), (g) and (i): Forecast skills calculated by simplex projection at different embedding dimensions (*E*) for the AT, ST40, ST80, ST160 and ST320 in spring (MAM), summer (JJA), autumn (SON) and winter (DJF). The best *E* is chosen when the forecast skill is optimal. According to the best *E*, the nonlinearities of the time series were detected by the S-map method ((b), (d), (f), (h) and (j)).

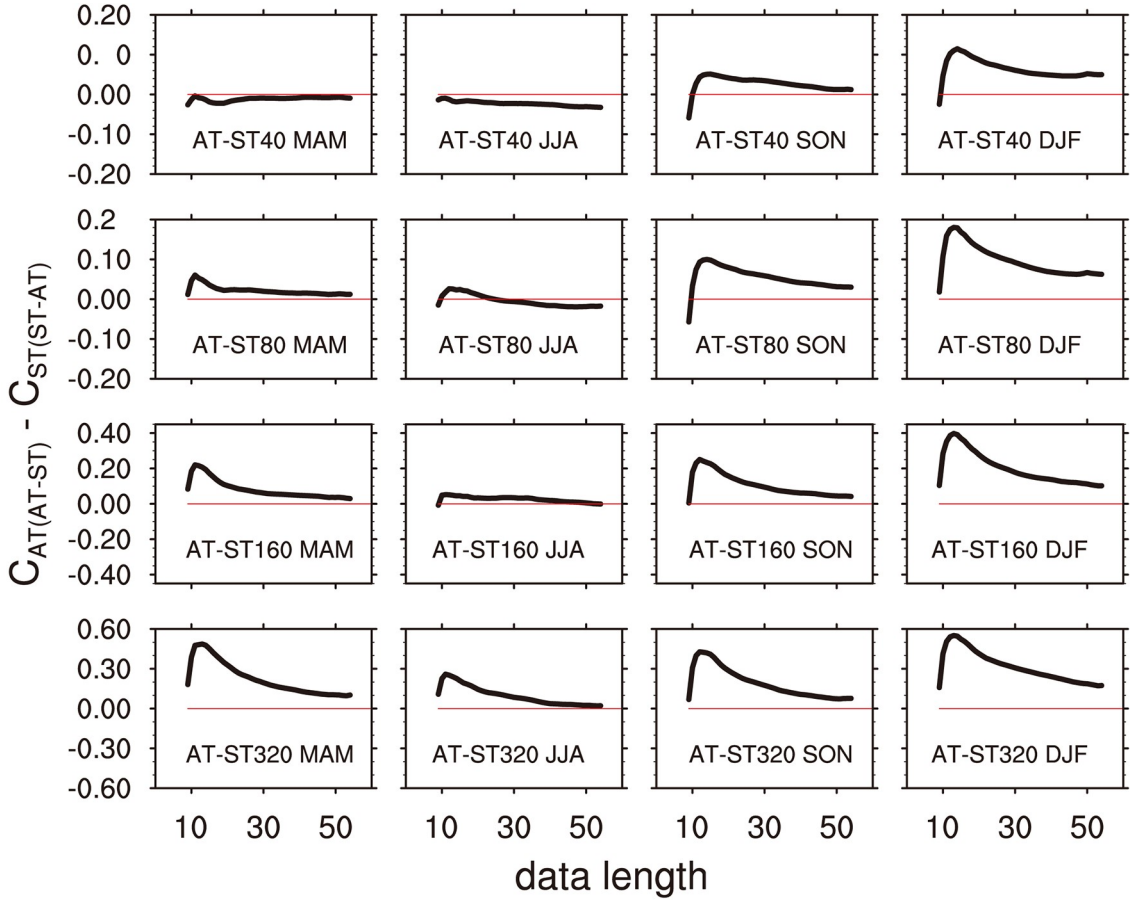

Fig. S7. Results of the PAI analysis to confirm the effects of the AT on ST40 (first row), ST80 (second row), ST160 (third row) and ST320 (last row) in spring (first column), summer (second column), autumn (third column) and winter (last column). The *y*-axis shows the "$\Delta = C_{AT(AT\text{-}ST)} - C_{ST(ST\text{-}AT)}$", where $C_{AT(AT\text{-}ST)}$ is the correlation between the estimated and original AT and $C_{ST(ST\text{-}AT)}$ is the correlation between the estimated and original ST. When Δ is always greater than 0 as the data length increases, it means that the effects of the AT on the STs are very significant and the causality from the AT to the STs is robust. It can be seen that Δ for all soil depths (ST40, ST80, ST160 and ST320) is always greater than 0 in autumn and winter, implying robust causality from the AT to the STs. In spring and summer, the AT and the ST have stronger coupling, and the springtime Δ for the ST40 and summertime Δ for the ST40 and ST80 are not always greater than 0. This may be because of the low temporal resolution of the seasonal records, which eliminates some differences between the AT and the STs from the monthly time scale and makes them more synchronous at the seasonal time scale.

Then, Fig. S1b shows that when the coupling strength is strong ($C = 0.8$), CCM cannot correctly infer causal direction for x and y. For the cross-mapping skill both from y to x and from x to y, they are not zero. This means that x and y share common information content. x causes y, and y also causes x. This is not consistent with the real dynamics of Eq. (S1). This is due to the fact that the strong coupling makes x and y become synchronized (Sugihara et al., 2012; Ye et al., 2015), and Ye et al. (2015) suggested to look at the performance of time-lagged CCM.

3.1.2. *Inferring causality through time-lagged CCM*

Figure S2a shows that when the coupling strength is moderate ($C = 0.1$), time-lagged CCM can infer causal direction for x and y. When using $x(t)$ to cross-map $y(t + t_p)$, the optimal cross-mapping skill occurs at a positive lag; when using $y(t)$ to cross-map $x(t + t_p)$, the optimal cross-mapping skill occurs at a negative lag. This reveals that the information content of x at the past time is encoded in the state of y at the latter time; and this means that x causes y, but y does not cause x, which is consistent with the real dynamics of Eq. (S1).

Then, Fig. S2b shows that when the coupling strength is strong ($C = 0.8$), time-lagged CCM can infer causal direction for x and y. Similar to the case of Fig. S2a, when using $x(t)$ to cross-map $y(t + t_p)$, the optimal cross-mapping skill occurs at a positive lag; when using $y(t)$ to cross-map $x(t + t_p)$, the optimal cross-mapping skill occurs at a negative lag. The conclusion for this causal inference is still that x causes y, but y does not cause x, which is consistent with the real dynamics of Eq. (S1). The performance of time-lagged CCM is better than that of CCM.

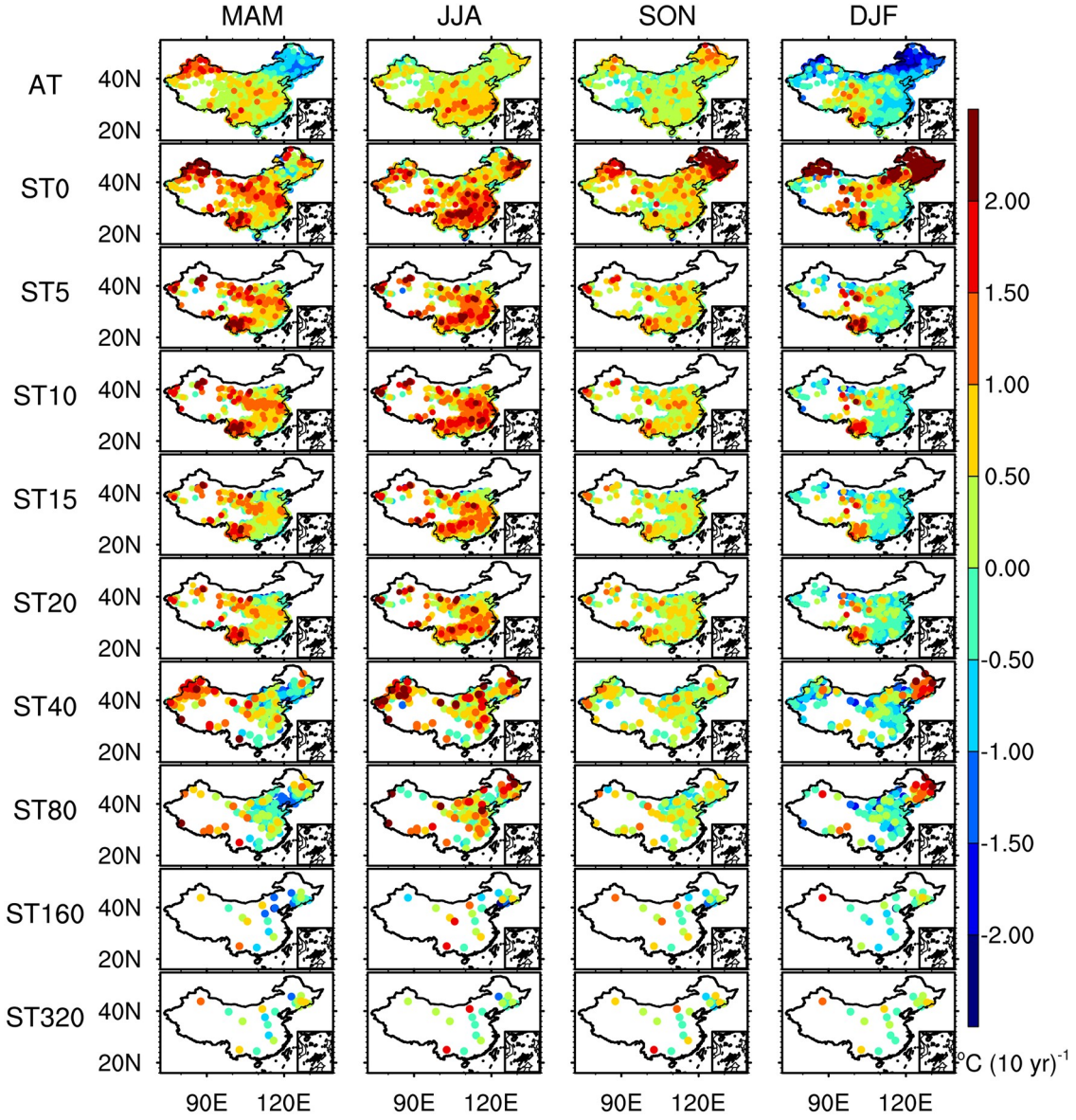

Fig. S8. Linear trends of the AT and STs (ST0, ST5, ST10, ST15, ST20, ST40, ST80, ST160, ST320) in spring (MAM, first column), summer (JJA, second column), autumn (SON, third column) and winter (DJF, last column) from 1998 to 2013. Despite these trends being region-dependent, the spatial distributions of the trends of the AT are similar to those of the STs in spring, summer and autumn. This phenomenon indicates that the AT and the STs are always coupled in terms of their long-term trends. However, for the cooling trends of the wintertime AT over northeastern China, they diverge from the warming trends of the STs (ST0, ST40, ST80, ST160 and ST320). This suggests that the AT effects on the STs are not dominant.

3.2. *Testing the presence of noise*

$$\begin{cases} x(t+1) = 0.85x(t) + \varepsilon_1(t) \\ y(t+1) = 0.8y(t) + 0.2x(t) + \varepsilon_2(t) \end{cases}. \tag{S2}$$

Equation (S2) shows two stochastic processes, and they are coupled through a unidirectional causal influence from x to y. $\varepsilon_1(t)$ and $\varepsilon_2(t)$ denote the Gaussian-distributed white noise with zero mean and unit variance. This coupled system mimics the dynamical processes with random noise. Previous studies (Ye et al., 2015; Mønster et al., 2017; Huang et al., 2020) have suggested that CCM is highly sensitive to the presence of noise, but time-lagged CCM can perform better under the presence of noise. The associated technical discussion is presented in the papers of Ye et al. (2015) and Huang et al. (2020).

In the following, we briefly report the results of experiments to test the impact of noise on the performances of CCM and time-lagged CCM, by analyzing the system shown by Eq. (S2).

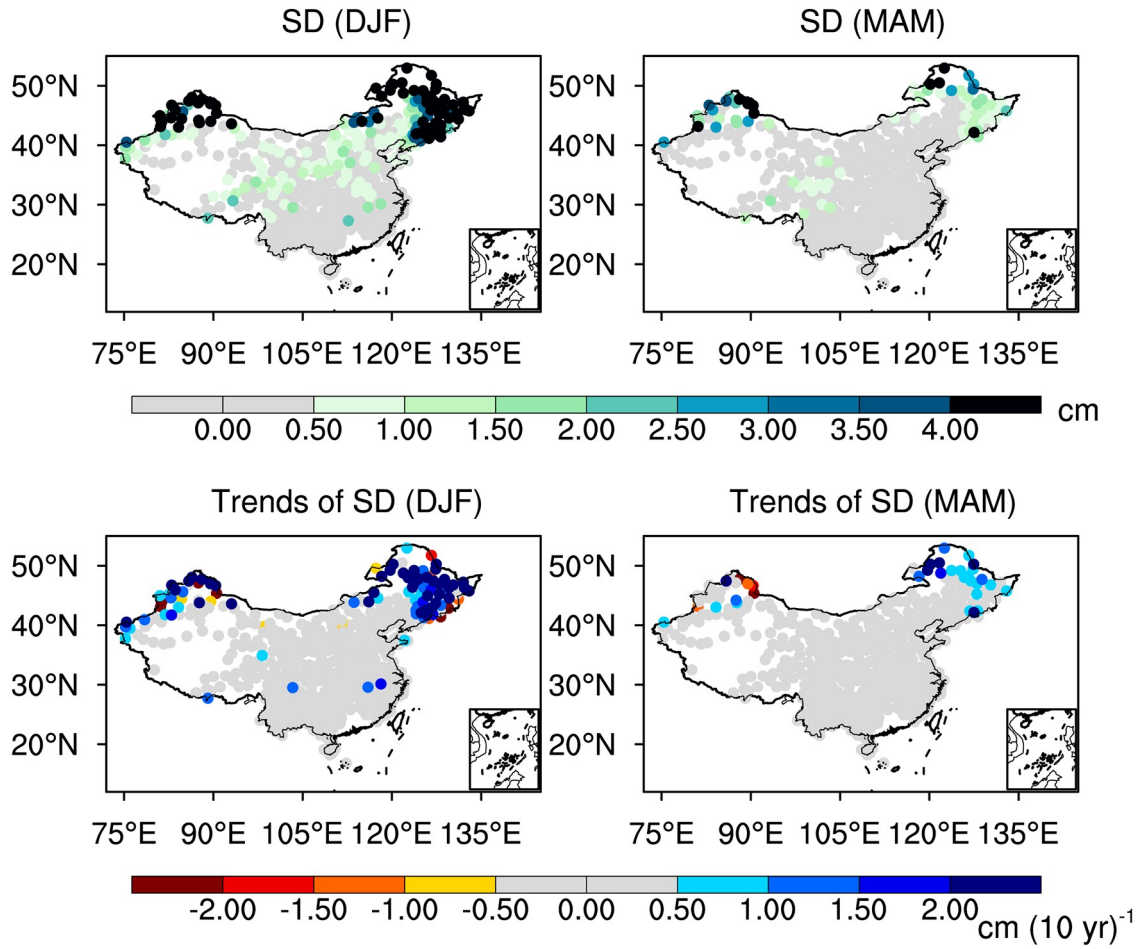

Fig. S9. In our paper, the divergence between the AT trends and the STs trends only occurred in winter in northeastern China. This is intriguing, and the insulation effect of the thicker snow may explain this phenomenon. The two panels in the top row show the climatological SD in winter (DJF) and spring (MAM). The wintertime SD in most of China (except northeastern China and northern Xinjiang) is less than 2.5 cm. In winter, the SD in northeastern China is more than 4 cm and is the thickest over China. In spring, both the thickness and area of coverage of the snow are decreased. The SD is only 0.5--1.5 cm at most stations in northeastern China. The two panels in the bottom row show the trends of the SD in winter and spring from 1998 to 2013. The wintertime SD exhibits remarkable increasing trends in northeastern China. The springtime SD is too thin in northeastern China, though it also has an increasing trend.

Figure S3a shows that the CCM cannot correctly infer causal direction for x and y of Eq. (S2). For both cross mapping skills from y to x and from x to y, they are not zero. This means that x and y share the common information content. x causes y, and y also causes x. This is not consistent with the real dynamics of Eq. (S2).

Figure S3b shows that time-lagged CCM can infer causal direction for x and y of Eq. (S2). When using $x(t)$ to cross-map $y(t + t_p)$, the optimal cross-mapping skill occurs at a positive lag; when using $y(t)$ to cross-map $x(t + t_p)$, the optimal cross mapping skill occurs at a negative lag. This reveals that the information content of x at the past time is encoded in the state of y at the latter time; and this means that x causes y, but y does not cause x, which is consistent with the real dynamics of Eq. (S2).

4. Simplex projection method and S-map test

Firstly, the optimal embedding dimension (E) should be determined. The simplex projection method was used to produce forecasts and examine the prediction skill, computed as the correlation between observed and predicted values. When implementing the prediction, an exploratory series of E was used to evaluate the prediction. When the prediction skill was optimal, the corresponding E was chosen. Secondly, based on the best E, the S-map test was used to test the nonlinearity in the time series. In the S-map test, the local linear maps were fitted to the forecast from the reconstructed state space, and the nonlinear parameter θ determined the distance weight of the vectors when fitting the local linear map. If θ was equal to 0,

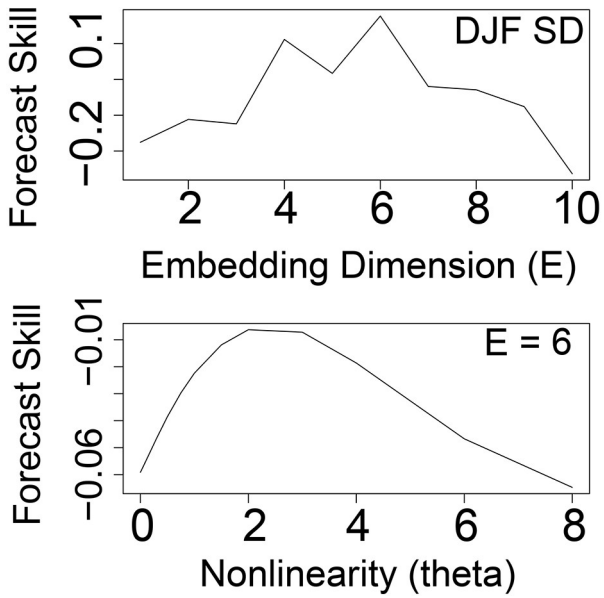

Fig. S10. As in Fig. S6 but for the winter SD in northeastern China. Forecast skills are calculated by simplex projection at different embedding dimensions (E) (top panel). The best E is chosen when the forecast skill is optimal. According to the best E, the nonlinearity of the time series was calculated by the S-map method (bottom panel).

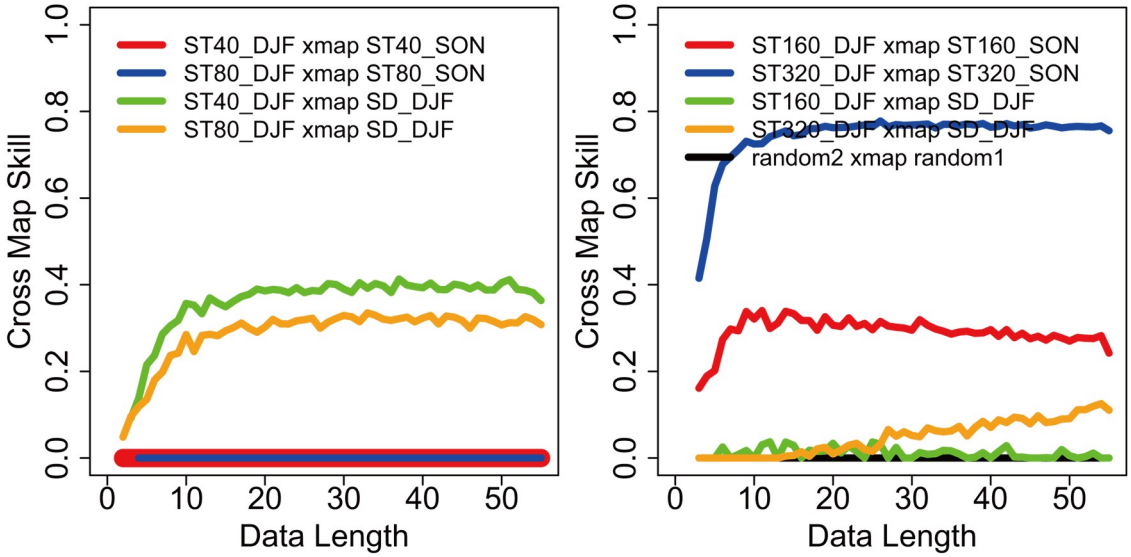

Fig. S11. Nonlinear causality analysis (the mentioned CCM) was implemented to identify the effect of the SD and the soil memory on the STs. The green (orange) line in the left-hand panel represents the cross-mapping skill of the CCM between the SD and the winter ST40 (ST80), and the green (orange) line in the right-hand panel represents the cross-mapping skill of the CCM between the SD and the winter ST160 (ST320). If the cross-mapping skills converge with the increase in data length conducted in the CCM, and are not equal to 0, there is a causality between the two variables. In the left-hand panel, the green line and the orange line approximately converge to 0.4 and 0.3, respectively, indicating that the SD has impacts on ST40 and ST80 (more effect on ST40) in winter. For the deeper soil, the SD seems to have no impacts on ST160 and ST320 in winter, with their cross-mapping skills almost vanishing (green line and orange line in the right-hand panel). The blue lines and red lines in the two panels represent the effect of the autumn ST on the winter ST. The cross-mapping skills between the autumn ST320 and the winter ST320 converge (blue line in the right-hand panel), and are close to 0.8 when the data length exceeds 20. This means that the autumn ST320 has a strong impact on the winter ST320. With the decrease in depth, the soil memory effects are weakened. The effect of the autumn ST160 on the winter ST160 is doubtful because the red line (in the right-hand panel) is not convergent. The autumn ST40 (ST80) has no effect on the winter ST40 (ST80) because the red line and blue line (in the left-hand panel) are always equal to 0.

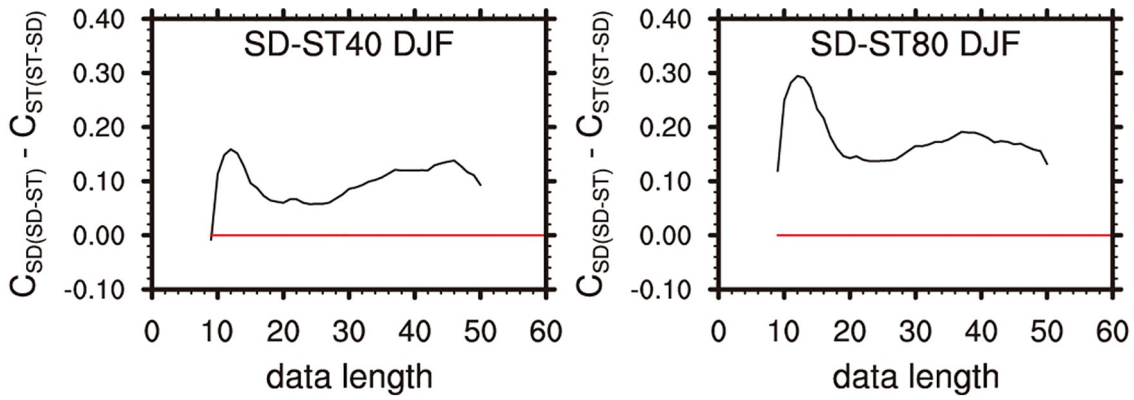

Fig. S12. Results of the PAI analysis to confirm the effects of the SD on the ST40 (left-hand panel) and ST80 (right-hand panel) in winter. The *y*-axis shows the "$\Delta = C_{SD(SD-ST)} - C_{ST(ST-SD)}$", where $C_{SD(SD-ST)}$ is the correlation between the estimated and original SD and $C_{ST(ST-SD)}$ is the correlation between the estimated and original ST. When Δ is always greater than 0 as the data length increases, it means that the effects of the SD on the STs are highly significant and the causality from the SD to the STs is robust.

the weights were equal and the fitting model was a single linear model. However, for $\theta > 0$, the weights relied on the local information (i.e., distance) and the fitting model was nonlinear. If the optimal prediction skill occurred at $\theta = 0$, the implication was that the time series could be predicted by a single linear model and exhibited linearity. Otherwise, the time series exhibited nonlinearity when the optimal prediction skill occurred at $\theta > 0$. For the AT, ST40, ST80, ST160 and ST320, their time series exhibited nonlinearity in each season (except ST160 in autumn), and the wintertime SD in northeastern China also reflected a nonlinear process (Figs. S5, S6 and S12). Therefore, the CCM method is suitable for detecting the causality between these variables.

REFERENCES

Huang Y., C. L. E. Franzke, N. M. Yuan, and Z. T. Fu, 2020: Systematic identification of causal relations in high-dimensional chaotic systems: Application to stratosphere-troposphere coupling. *Climate Dyn.*, **55**, 2469–2481, https://doi.org/10.1007/s00382-020-05394-0.

Mønster D., R. Fusaroli, K. Tylén, A. Roepstorff, and J. F. Sherson, 2017: Causal inference from noisy time-series data—Testing the Convergent Cross-Mapping algorithm in the presence of noise and external influence. *Future Generation Computer Systems*, **73**, 52–62, https://doi.org/10.1016/j.future.2016.12.009.

Sugihara G., R. May, H. Ye, C. H. Hsieh, E. Deyle, M. Fogarty, and S. Munch, 2012: Detecting causality in complex ecosystems. *Science*, **338**(6106), 496–500, https://doi.org/10.1126/science.1227079.

Ye H., E. R. Deyle, L. J. Gilarranz, and G. Sugihara, 2015: Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.*, **5**, 14750, https://doi.org/10.1038/srep14750.