S1. The choosing of research object

The reasons for choosing Harbin-Changchun urban agglomeration as research object lie in following aspects:

On one hand, the old industrial areas in Northeast China are limited by their own development characteristics and have long been plagued with serious industrial emissions. This kind of development at the expense of the environment is harmful, so it is necessary to make timely and accurate air quality prediction for it. Especially, the model construction is credible with the Harbin-Changchun urban agglomeration as a typical example.

On the other hand, the northeastern region is dominated by agricultural development, which plays a key role in the food supply of the whole of China, and the development of traditional agriculture can cause huge air pollution. To balance the dual goals of agricultural development and pollution control in the northeast region, and to strive to comply with the concept of green development, it is particularly important to make timely and effective air pollution forecasts for the region. Among them, the Harbin-Changsha urban agglomeration is in an important position in the development of Northeast China, and this paper takes the air pollution in Harbin-Changchun urban agglomeration as the re-search object.

S2. The Dynamic analysis of social network

Setting the year of comparison here as 2015, the spatial network map of air pollution in that year is plotted in the same way as in Section 3 as shown in Figure 1. And the difference between these two years was showed in Figure 2, evaluated by the indicators introduced above.

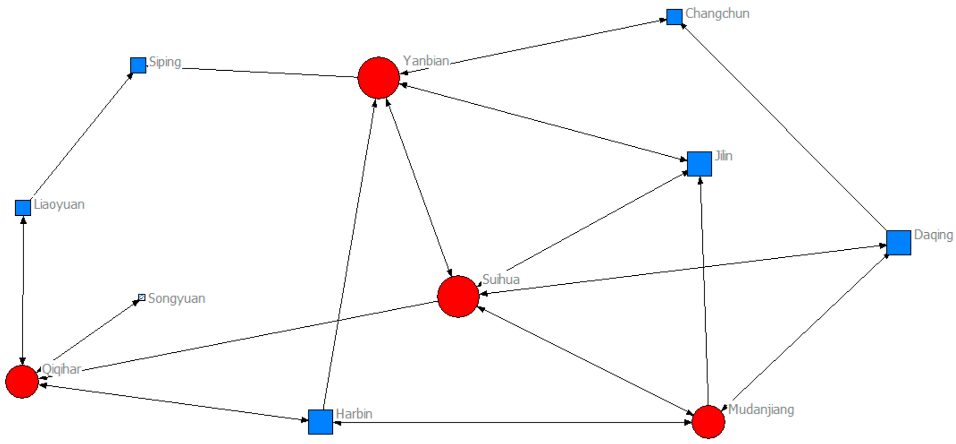

Figure S1. The comparative correlation network of 2015.

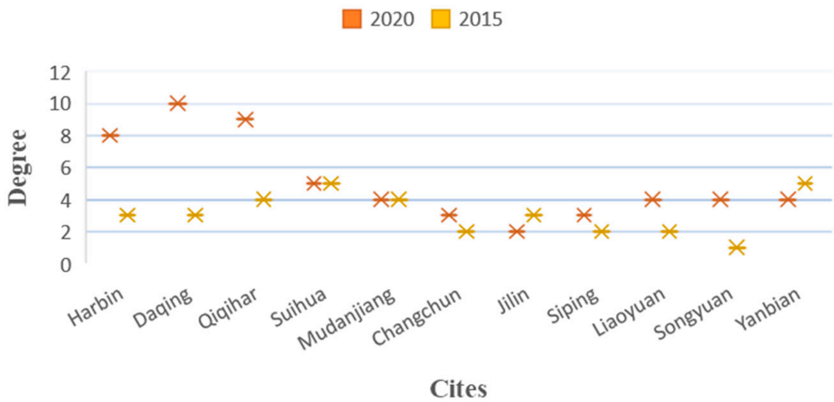

Figure S2. The comparison of 2015 and 2020 in Degree.

It is evident that the structure of social network in 2015 is simpler in terms of the connected direction, when comparing with that of 2020. The comparative graph in Figure 2 clearly shows that the temporal correlation between the nodes of the air pollution network in 2015 is not strong, and it is difficult to determine the main creators of pollution in urban agglomerations. In addition, the linkages represented to be undirected while the directed one appeared more and more in 2020. By 2020, due to a number of factors such as economic interaction and population mobility, the correlation between urban agglomerations has become increasingly close, as evidenced by a significant increase in the degree of the emergence of key cities for joint defense and joint control.

S3. The Application of multistep forecasting

The results of multistep forecasting could be seen as follows:

Table S1. The accuracy of multistep forecasting.

	ε_{MAE}	ε_{RMSE}	ε_{SMAPE}	ε_{U1}	ε_{r}
1 day	7.8692	9.9289	11.5215	0.0744	0.9816
3 days	10.1762	11.9826	15.0847	0.1281	0.9104
5 days	13.5700	15.9853	19.5894	0.1620	0.8518
7 days	19.6287	20.1934	21.9382	0.2682	0.7910

Based on the experiment on air quality forecasting of different steps ahead, it is observed that the prediction accuracy of the proposed model remains good in the short term, like three days ahead, keeping the mean absolute error close to ten percent, which is sufficient for daily pollution monitoring. When the step is up to seven days, the model accuracy decreases significantly, which is due to the constant accumulation of errors and is a problem encountered by almost all forecasting models.

But overall, from the above results, the predictive performance of the model and its stability are recognizable, at least to a high degree of accuracy in the short term. As far as practical applications are concerned, multi-step prediction of air quality facilitates people's daily travel planning, reduces exposure risks, facilitates transportation planning, and is helpful for government departments to formulate appropriate preventive measures in advance.