Supplemental materials

Case study conducted by a new domain expert and the interview

To evaluate the effectiveness of *AirLens* for domain experts not involved in the system design, we invited a new domain expert E_C to conduct another case study, who focused on the air pollution control technology. Under the background that data analysis techniques are being widely used in the environmental field, E_C is curious to know what our systems can do. We introduced our system through video meeting and deployed the system in advance. E_C was interested in the urban agglomeration of Fen-Wei plain, which had attracted the attention of many researchers in recent years. Actually, E_A had already performed a case study of this region. Thus, to guide the exploration for E_C who did not have clear analysis targets, we specified the following tasks according to E_A's findings:

Task 1. Explore the pollutant patterns and their spatiotemporal distributions.

Task 2. Explore the stages with different characteristics of evolution.

Task 3. Explore the frequent pattens starting from different pollutant patterns.

Exploring the pollutant patterns and their spatiotemporal distributions. E_C selected the data in urban agglomeration of Fen-Wei plain (Figure 1-A) from March 1, 2015 to January 3, 2016. The selected pollutants (Figure 1-B) are $PM_{2.5}$, PM_{10}, NO_2 and O_3, because their concentrations have high associations. E_C preselected the pollutant patterns summarized using 6 as the cluster number and inspect their difference in pollutant IAQIs and spatiotemporal distribution. Then, E_C classified these pollutant patterns that can summarized as Figure 2.

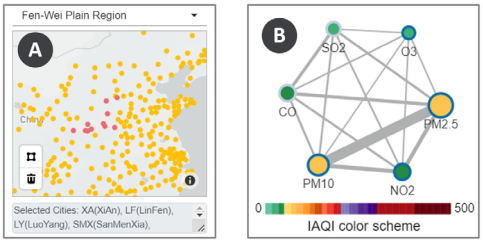

Figure 1. *The selected urban agglomeration and pollutants.*

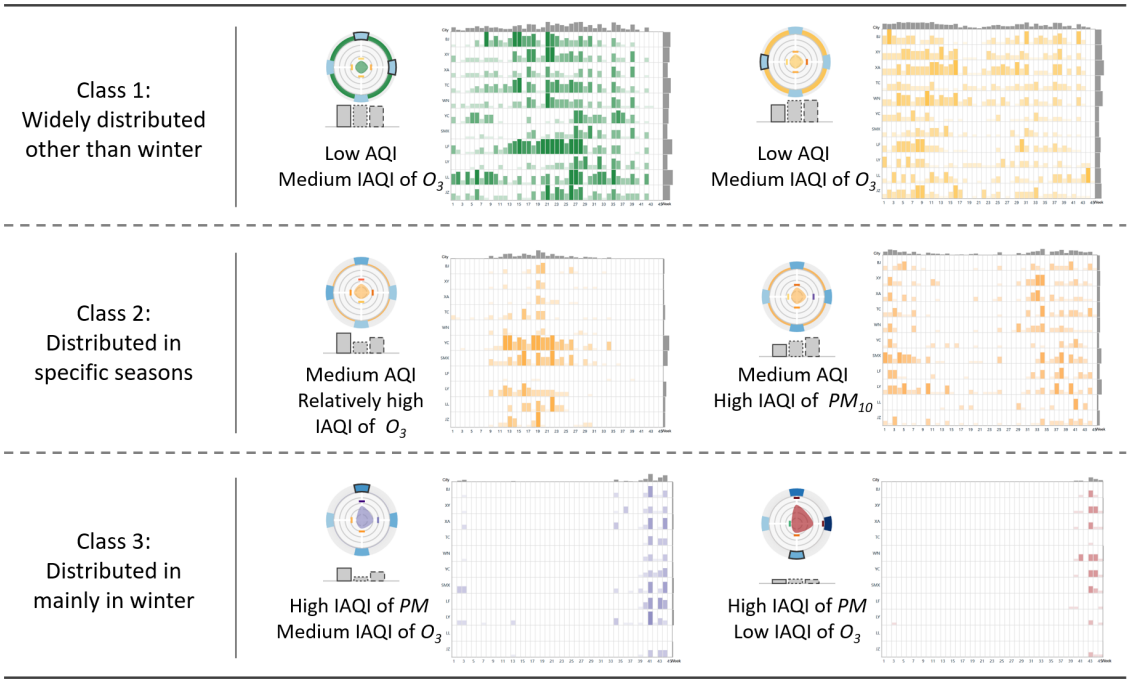

Figure 2. *The pollutant patterns classified by their pollutant IAQIs and spatiotemporal distribution.*

Exploring the stages with different characteristics of evolution. E_C thought the above group with 6 pollutant patters is fine enough and selected it to explore the *evolution analysis module*. After observing the overall-level view (Figure 3), E_C noticed four stages (Figure 4) that attracted her attention. Among them, the stage A and stage D showed different evolution trends in the city communities, while stage C only had one city community.

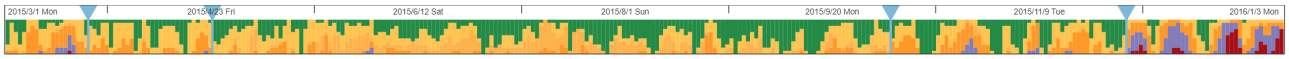

Figure 3. *Overall-level view.*

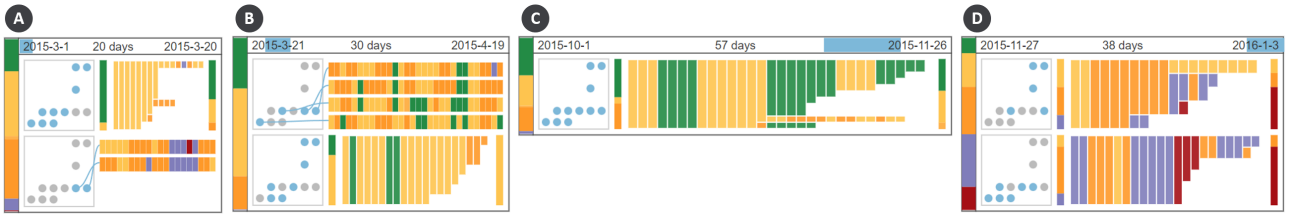

Figure 4. *Stages noticed by* E_C.

E_C first selected stage C to observe the detailed evolution and explore why the cites were not divided to communities. After setting the city number filter as shown in Figure 5-A1, it could be observed that more than 5 cities changed together on most of the time steps as shown in the top row of the *time controller* (Figure 5-A2). This indicated that the air quality in the cities tended to change together in stage C. Then, E_C selected stage D. From the corresponding *spatiotemporal flow chart* (Figure 5-B) with setting the change level filter (Figure 5-B1), E_C found that the flows changed drastically, which meant the air quality are unstable in stage D.

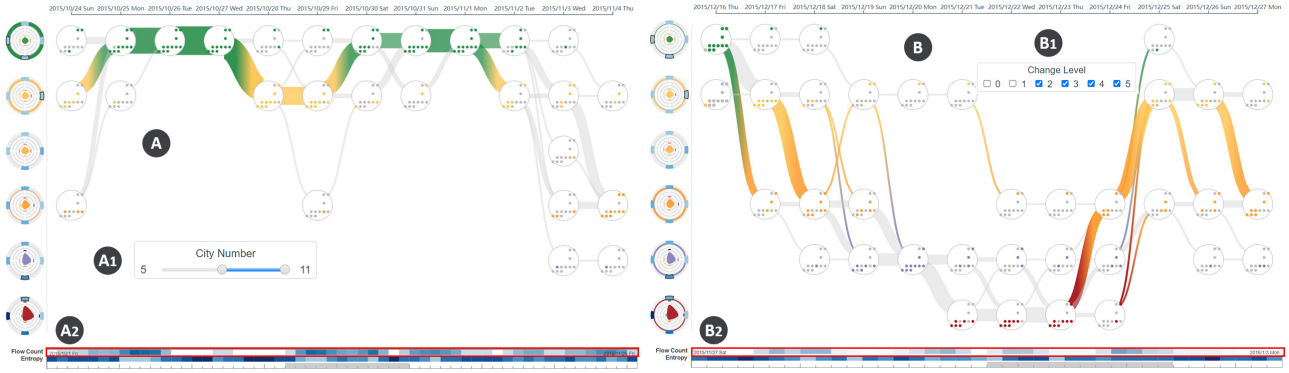

Figure 5. *Spatiotemporal flow charts of* (A) *stage C and* (B) *stage D.*

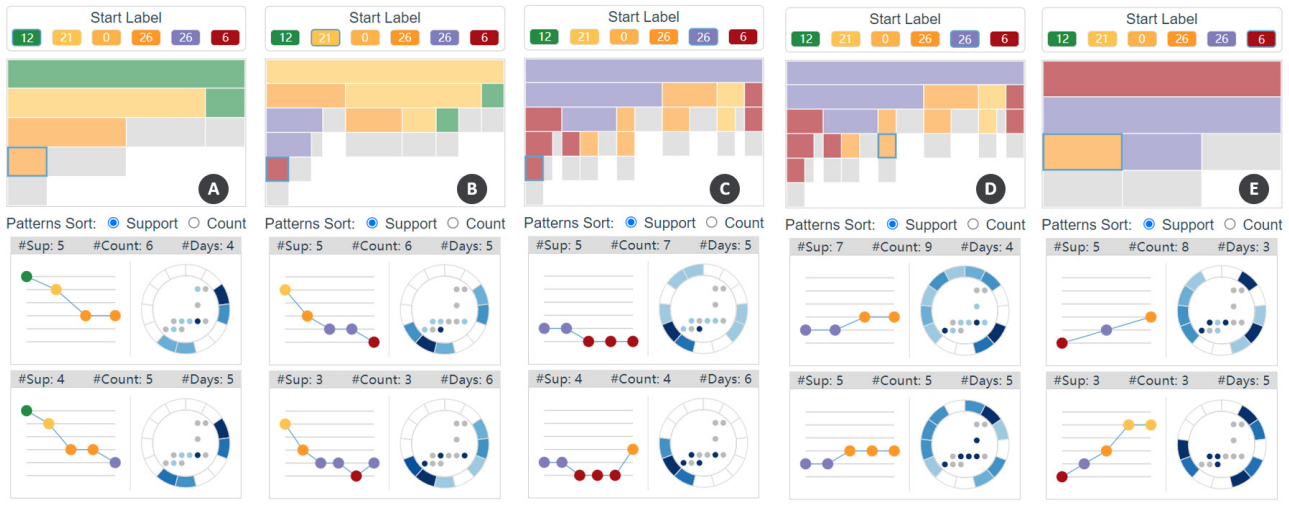

Figure 6. *Frequent patterns* starting from different pollutant patterns *in stage D.*

Explore the frequent pattens starting from different pollutant patterns. E_C further explored the frequent patterns in stage D. E_C first clicked the button of the green pollutant pattern, to investigate how air quality change from a good condition. As shown in Figure 6-A, after selecting the marked node in the icicle plot, E_C found that the two frequent patterns both presented that air quality became poor within nearly the same cities and during nearly the same period. The similar condition could be found as the two frequent patterns in Figure 6-B. Then, E_C clicked the button of the purple pollutant pattern. From the icicle plot (Figure 6-C and D), E_C noticed that the corresponding frequent patterns can be divided into two classes. One class was that air quality became worse, as shown in Figure 6-C. While, the other class was that air quality became better, as shown in Figure 6-D. Finally, E_C clicked the button of the red pollutant pattern, and the corresponding frequent patterns showed up trends as shown in Figure 6-E.

Interview

System. After exploring our system, E_C thought *AirLens* was effective to find the significant evolution patterns and useful information, which could provide clues or evidence for her research.

Visual designs and interactions. E_C considered the design of the cluster glyph and frequent pattern card are easy to understand. Although she did not well understand all the information shown by the stage cards, she could distinguish the different evolution trends of different city communities. For the spatiotemporal flow view, E_C was initially confused with complex flows. While, after using the various filter, E_C could freely get the flows of interest.

Suggestions. E_C commented that the evolution analysis was based on the pollutant patterns fully summarized by computes, and she wanted to specify a pollutant pattern by herself. Thus, E_C suggested us to add the function that user could specify the patterns first and incorporate them with the clustering algorithm.