Visual Analytics for Regional Economic Environment Factors Based on a Dashboard Design

Conference Paper

Author(s):
Zuo, Chenyu; Ding, Linfang; Meng, Liqiu

Visual Analytics for Regional Economic Environment Factors Based on a Dashboard Design

Chenyu Zuo[a] , Linfang Ding[a,b],*, Liqiu Meng[a]

[a] *Chair of Cartography, Technical University of Munich, Germany*
[b] *KRDB Research Centre, Faculty of Computer Science, Free University of Bozen-Bolzano, Italy*

* Corresponding Author

Abstract: Economic environment is vital for commercial investment, city planning and company strategy planning in urban areas. Mastering the economical trend may help the entrepreneurs, government officers and individuals in their decision-making process. In this study, we explore multiple geo-economic datasets using visual analytics methods for understanding the economic environment. More specifically, we user time-series Gross Domestic Product (GDP) data as an economic indicator of economic development and land use data to support the spatial analysis at a refined geographic scale. The spatiotemporal patterns of the regional economic environment are revealed both qualitatively and quantitatively. The work has a three-fold contributions: (1) we apply a grid-based spatial interpolation model to derive GDP values at a file granularity based on land use data; (2) we design a novel interactive dashboard for the GDP data exploration, which serves as a visual analytical tool between data and users; (3) we combine quantitative analysis with visualizations to strengthen the qualitative analysis. The feasibility of visual analytics methods and the dashboard design are tested in one of the most developed regions, Jiangsu Province, China. Both expected and unexpected economical patterns were extracted.

Keywords: visual analytics, spatiotemporal analysis, choropleth maps, small multiples

1. Introduction

Economic environment is unquestionably important for entrepreneurial activities and innovations. It is the result of the influence of many factors, such as business, education, logistic, politics, natural resources. Understanding the trend of economic environment is vital for commercial investment, city planning and company strategy planning. However, the analysis of economic environment is challenging due to its numerous impact factors, and large amount of heterogeneous datasets. Therefore, mastering the information related with economic environment is hard for the public.s

Economic environment is an intensively observed research domain, where an enormous number of studies related with spatial economics were conducted. Spatial autoregression model and its extensions for spatial econometrics were concluded (LeSage and Pace, 2009), to aid spatial economic analysis, such as national commuting time at county-level analysis. Additionally, the relationship between economic environment with infrastructure (Shenggen and Zhang, 2004), air quality (Lin et al., 2014) and land use change in urban area (He et al., 2014) were explored. Especially, the economic environment in Yangtze River Delta was analyzed systematically by Wu et al. (Wu et al., 2016). In their research, many economical topics such as industry clustering, urban agglomeration economic integration, division of functions of urban agglomeration were discussed, economic models (such as spatial lag model and linear regression) and economic indexes (such as Theil index and Moran's I) were applied, and finally thought-provoking conclusions and insights of the regional economic development status were drawn. However, the researches were mainly intended to solve the defined questions, therefore, the information flows unidirectionally from experts to non-experts. Besides, visualizations applied in those research were only served to present the results of the analysis.

Visualizations could be used as an interface for communication between information and humans. Visual analysis involves human perception as part of the analysis process, combining the strong data computing capability of the machines, to benefit the interpretation and dissemination of information. Particularly, knowledge presenting different perspectives of interests can be obtained from the visualizations by the visualization recipients. For instance, Minard designed a flow map to illustrate the Napoleon's ill-fated march on Moscow (Corbett, 2012) visualizing the information such as size and location of the marches, and the temperature; (Bostock, 2016) applied choropleth mapping to show US unemployment rates of different regions from 2008; (Wu et al., 2016) developed a visual analysis platform to visually explore the urban human mobility. Furthermore, visualizations were increasingly applied and designed for the understanding of environments. For example, the choropleth mapping was applied to ecosystem services by Le Clec'h et al. to monitor the environment (Le Clec'h et al., 2018). Interestingly, choropleth maps were designed to illustrate the characteristics of regional economic, for instance, population and Gross Domestic Product (GDP) in China (Zilong, 2014). In their research, the visualization unit of the choropleth maps is as Chinese province, which might lead to misunderstanding of data density since the area of each unit varies. Similarly, heat maps and charts were used to present the urban agglomeration in Yangtze River Delta (Zheng and Bohong,

2012). The heat maps were grid-based, showing the population density. Moreover, urbanization is illustrated in line charts. However, graphics-enhanced maps could deliver more information and benefit the correlation analysis. Another study was dedicated for visualizing economic model outputs and designed visualizations like treemap and sunburst (Webbera et al., 2006). In short, the visualizations can help the data owners to publish them to the public in a more easily understandable way.

Dashboard is a visual communication tool, that displays important information on a single screen so it can be monitored at a glance (Few, 2006). With the organic visualization organization of a dashboard, the attention of the users are leaded to those important hidden patterns. However, the geo-locateded, user-centered, intelligent dashboard has not been used for economic environment analyzing. Therefore, we designed a dashboard-integrated fraemework for spatiotemporal macro econnomic environment factors visual analyzing. To achieve this, we propose an interactive dashboard for visual analysis: visualization methods such as choropleth mapping, small multiples, spatial interpolation are applied. The experiment results show that our proposed methods are helpful in examining the spatiotemporal inequity of regional economic environment, in understanding the correlations between economic environment and the related factors, and in providing the possibilities for general public to obtain insights from the raw data.

2. Visual analysis methodology

The aim of this research is to develop a framework for economic environment visual analytics with heterogeneous geo-data, which can allow public to access and analysis the economic spatiotemporal patterns. The visual analytics methods are applied in this study, to broaden the human understandability of the large amount of numerical geo-data, to get access to the hidden economic environment knowledge.
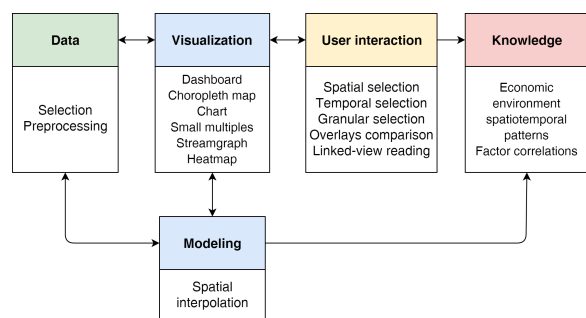

Figure 1. The workflow of visual analytics of regional economic environment factors.

As Figure 1 shows, visual analytics of regional economic environment factors is composed of five components, including data, visualization, model, user interaction with visualizations and knowledge obtain. In each component, customized methods are designed. Firstly, the appropriate economic environment related factors are selected as a preliminary data and be preprocessed. Secondly, the data is visualized with different methods. Meanwhile, the results of the visualizing invoke the data modeling. Then the results of the modeling contributes to the visualizing. Here we investigate the visualization methods such

as dashboard, choropleth map, small multiples, streamrgaph and heatmap. The spatial interpolation is applied in the model component. Thirdly, the interactive visualizations are organized by the dashboard. Users can select by location, time, granularity and overlays. Lastly, the knowledges, such as spatiotemporal economic environment patterns and factor correlations, can be interpreted from the model and visualizations. The following sections describe each component in detail.

2.1 Data selection and preprocessing

Census data is wildly used in social-economic research (He et al., 2014, Shenggen and Zhang, 2004), due to its high accessibility and reliability. As our focus is on the regional economic environment, GDP data (part of the census data) is selected as the main factor to reflect the general living cost and commercial activities. The spatial variation in economic development can be explained by differences in local policy and investment. But within the small administrative unit, that influence can be minimized.

However, the census data is highly aggregated from the geographic perspective. Thus the supplementary data needs to be integrated to the model. The other dimensions which correlated to spatial economic are societal and geographic factors. Land use area and types, and the transportation are vital geographic and societal factors for the local economics. As different land use types produce economic activities. In the land such as urban land, the economic development level is higher than other land types. While barren land hardly benefit the local economic. Given the condition that the local transportation is the fundamental high economic development level condition, canals and ports increase the local commercial activities. Therefore those factors are included in our study.

2.2 Spatial interpolation modeling

The most research which used census data kept the original geographic character: county was used as the analysis unit. It is convenient for the large scale (global or national scale) analysis, while not precise enough for regional economics analysis. A regional geographic poverty indicator identifying study (Xu et al., 2018) was conducted and the topographic indicators were successfully detected. Based on their finding, a further step, a geographical social-economic factors analysis can be made. In this study, a multi-granular geographic economic environment analysis framework is introduced.

The county-level and fine-level economic environment analysis and cross level analysis are conducted. To get the fine-level economic data, spatial interpolation methods are investigated and applied. Geographically weighted and regression (Brunsdon et al., 1998) includes the geographical variables in spatial modeling. Moreover, the value of the weights can reveal the importance of the variables.

The land use type and area is integrated in the spatial interpolation model, as previous studies show that: there is strong positive correlation between land use subcategories and GDP values. A few studies has been done to gain the fine-level GDP by using land use data (Gaffin et al., 2004, Grübler et al., 2007). (Huang et al., 2014) proposed a model to calculate gridded GDP with the census county-wise GDP data and land use data . In their model GDP can be calculated as:

$$GDP_j = \sum_{i=1}^{n} \left(g_i \times L_{ij} \right) + B_j \qquad (1)$$

where GDP_j is the GDP value of the j-th county, g_i is the GDP density of the i-th land use subcategory (100 million Chinese Yuan/km^2), L_{ij} is the area of the i-th land use subcategory of the j-th county and B_j is the intercept. The values of g_i can be calculated by the least square method by applying the method to all the counties. Then the gridded GDP values can be calculated by multiplying the area of land use types and the GDP density g_i.

In this study, we apply Huang et al.'s model to interpolate GDP values to continuous gridded values.

2.3 The dashboard design

Unlike read-only maps which show certain amount of geographic information, the dashboard with interactive maps and charts offer customized solutions. They serve as an information platform, providing users with multi-dimensional, multi-variable visualizations revealing different aspects of data distribution, and with the possibility to interactively explore the relations among different visual patterns. The visualization influences the how humans obtain insights out of the data (Kohlhammer et al., 2011). Thus, the design is vital for visual analytics of economic environment.

A dashboard is designed to the governmental officers and entrepreneurs for decision making. Firstly, the spatiotemporal patterns of GDP is vital for economic environment understanding. So that the textual GDP data is transfered into maps. Secondly, the patterns on large scale and small scale reflect different aspects of the spatiotemporal patterns. Thus, a spatial interpolation model is performed to gain the fine scale data for the large scale visualizations. Thirdly, both of the general patterns and data details are interested by different user groups. Therefore, the overviews and detail views need to be displayed. Last but not least, the dashboard provides adaptive contents for different users, so that the content of interest can be show to the users in one screen at a glance.

Based on the above requirements, a dashboard with six linked-views is designed, 1) a map view showing the spatiotemporal GDP distribution, 2) a time view focusing on the temporal changing at an overall scale, 3) a data view displaying the textual data of the GDP details, 4) a text description view emphasizing the name, time and GDP value of the under showing area, 5) a pie chart view showing the proportion of the under showing area's GDP contribution to higher levels administration, 6) a bar chart view presenting the temporal pattern of one area.

3. Case study

3.1 Study area

The study area of Jiangsu Province shown in Figure 2 shows is one of the most developed regions in China. The Chinese longest river, Yangtze River, passes through the province. The area of Jiangsu is 102,600 km^2, with a plain area of 70,600 km^2, and water surface 17,300 km^2. It is one of the economically most developed regions in China,
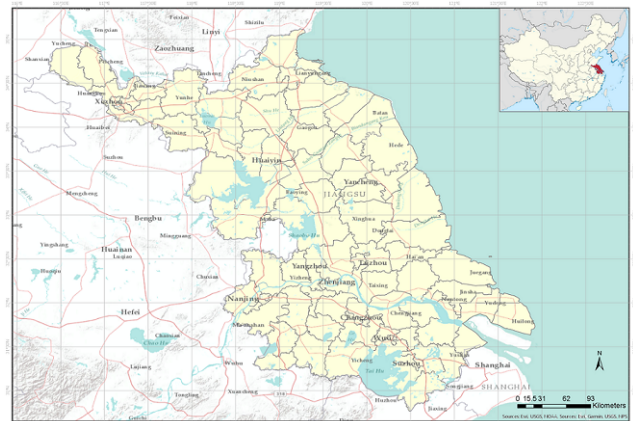

Figure 2. Jiangsu Province and the location in China (the overview image by Uwe Dedering, in Wikimedia Commons undera CC-BY-SA-3.0-PL license).

with the GDP of 137.87 billion of US Dollars for 80.3 million population in 2018 (Bureau, n.d.).

The administrative division hierarchy in China is 5 levels, namely, provincial level, prefectural level, county-level, township level and village level. In 2018, Jiangsu Province includes 1 sub-provincial-level city, 12 prefectural-level cities, 44 districts, 22 county-level city and 19 counties.

3.2 Data

To visually analyze the GDP patterns, we acquired the following types of data: (1) administrative boundaries, (2) the GDP data, (3) land use map.

3.2.1 Administrative map

The administrative map contains the province, city, county boundaries in the study area at the scale of 1:1,000,000 in 2017. It is stored in the format of vector data. The dataset was downloaded from the National Catalogue Service for Geographic Information of China[1]. The administrative map is used as base map to spatially project GDP data.

3.2.2 GDP data

The GDP data was extracted from the census data, which is normally stored as tables. The values in Jiangsu from 1980 to 2010 were selected to use in our experiment. The GDP data contains two levels of aggregated data: city-level and county-level [2] data.

However, the data quality is not ideal. Figure 3 shows the statistics of the GDP value and availability from 1980 to 2010. The light green line shows the number of county-level GDP available counties. Obviously, the county-level data in the early 1980s was in lack. The best data available year is 1989, while the data availability decreased in the ten years from 2000 to 2010. The dark green line shows the total county-level administration. Jiangsu Province included 11 cities from 1980 to 1996, and 13 cities from 1996 onwards. In 1996, Taizhou County and Suqian County upgraded to cities, so that the affiliate of counties changed. It reduced the consistency of city-level GDP data. Moreover,

---

[1] www.webmap.cn
[2] Sub-provincial-level city, prefectural-level city affiliate to city-level. District, county-level city and county affiliate to county-level.
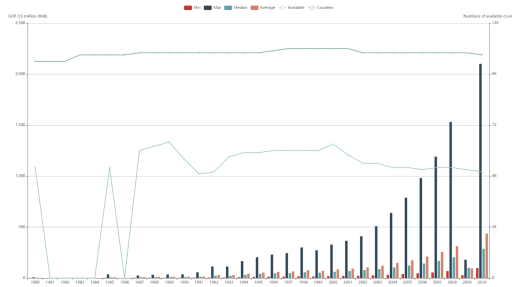
Figure 3. The GDP statistics in Jiangsu Province.

the names of some administrations changed. For instance, Huaiyin City changed to Huaian City in 2001. Therefore, the changing of the administration caused uncertainty for the spatial data projection. Under the condition of unavailable historical administrative maps, the changing names brought difficulties in projecting historical data to the up-to-date maps. The bars present the yearly statistics of GDP, respectively minimum GDP value among all the counties in study area in red, maximum in dark blue, median in light blue, and average in light red.
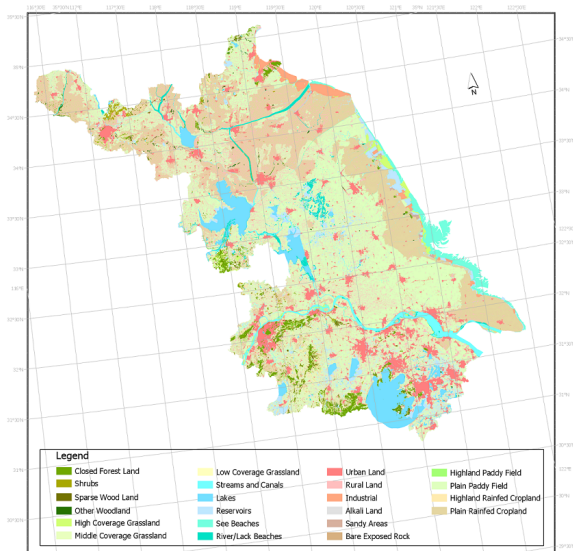
3.2.3   Land use map

Figure 4. The land use types in Jiangsu Province in 2010.

We used the land use map at the scale of 1:100,000 in 2010. The land use types and area are obtained by fusion enhancement and interpreted by remote sensing data (Landsat imagery and aerophotos) and auxiliary historical documents. The land use types fell into six categories (grass land, forest land, crops land, residential and build-up land, water and unused land) and 27 subcategories. Specifically, the agricultural land is classified by local topographical information and irrigation methods. Figure 4 shows the land use types and areas in Jiangsu 2010 in detail.

3.3   Spatial interpolation

As introduced in Section 2.2, the spatial interpolation was applied on the GDP data and land use data of Jiangsu in 2010. Previous study, such as (Huang et al., 2014), spatially interpolated GDP into 1*1 km km^2 grids. However, the grid size is defined by the smallest land use area size.

To guarantee the interpolation precision, we take the smallest land use polygon size, 5*5 km^2 is chosen as our interpolated unit size.

Firstly, the study area was devided into 5*5 km^2 grids. Secondly, the area of every land use type in each grid is calculated. Thirdly, the coefficients between each land use type and GDP was calculated by the regression model. Finally, the fine-level GDP was calculated per grid. Table 1 presents the calculated coefficients of each land use type. The intercept of the linear model is 160.43932 (100 million Yuan/square kilometer/year), which is the expected mean value of the Jiangsu GDP in 2010.

Table 1. The Jiangsu GDP interpolation model coefficients of different land use types with the intercept of 160.43932 (100 million Yuan/square kilometer/year).

Landuse Category	Subcategory	Coefficient	Sum
Agricultural Land	Highland Paddy Field	-0.07902	
	Plain Paddy Field	-0.13615	
	Highland Rainfed Cropland	1.28506	1.03603
	Plain Rainfed Cropland	-0.03386	
Forest Land	Closed Forest Land	0.33067	
	Shrubs	-1.33001	
	Sparse Wood Land	-2.90484	-12.6523
	Other Woodland	-8.74807	
Grassland	High Coverage Grassland	0.48797	
	Middle Coverage Grassland	50.24663	50.7346
	Low Coverage Grassland	NA	
Water	Streams and Canals	0.81486	
	Lakes	-0.40637	
	Reservoirs	0.56104	2.26795
	See Beaches	2.42068	
	River/Lack Beaches	-1.12226	
Build-up Land	Urban Land	6.74171	
	Rural Land	-0.82187	5.06545
	Industrial	-0.85439	
Barren Land	Alkali Land	19.55501	
	Sandy Areas	-6.86577	-1.91442
	Bare Exposed Rock	-14.60366	

3.4   The dashboard

The aforementioned interactive dashboard is implemented and shown in Figure 5. The section A, B, C, D, E, F are map view, time view, data view, text description view, pie chart view and bar chart view respectively. The innovation of dashboard design is in four aspects: 1) unlike conventional at-a-glance dashboard, our interactive dashboard allows user to obtain both of the at-a-glance and customized information, 2) the map-based dashboard gives users multi-dimensional map services, 3) the various visualizations offer a overview + details multi-aspects data interpretation, 4) the dashboard provide a multi-granular spatiotemporal view.

Map view presents the spatial distribution of GDP and the background geoinformation. By default, the foreground shows a choropleth map of regional GDP distribution of a specific year. Three granularity levels of maps are offered: city-level, county-level and grid-level. In addition, three overlays are offered as optionally layers: rivers, lakes and ports.

Time view serves as a time controller. Users can select the time by clicking or dragging the time controller. Accordingly, the GDP information of the selected time period would be shown in the views. For each selection, only one year can be activated.
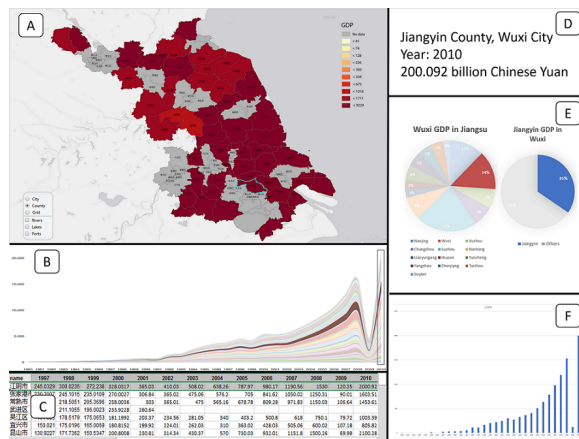
Figure 5. The interface design of dashboard.

Data view displays the detailed information of each data unit. The years, the county names, and the precise GDP values.

Text description view shows the information according to the user selection. Firstly, the name of the selected county and the name of the city that county belongs to are shown. Secondly, the selected year is shown. Thirdly, the GDP value of the selected county and the selected year is shown.

Pie chart view displays two pie charts showing GDP proportion of the selected area of the selected year. One of the pie chart depicts the proportion of the cities GDP in the province. Similarly, the other pie chart depicts the proportion of the counties GDP in the city.

Bar chart view serves as an additional chart to the time view streamgraph. It shows the GDP values in a bar chart of a selected county, as the accurate values of a sub-area is hardly available from the streamgraph. The temporal trend of the GDP values of the selected county can be easily perceived.

Interactivity The map and streamgraph are used as the spatial and temporal filters. The users can make selections by clicking the map and dragging the streamgraph. Other views will adjust the content in the same time.

4. Results and discussion

The map-based dashboard provide the users an innovational tool to explore and analysis with the regional economic environment from a geographic perspective. To be specific, 1) the original economic factors, 2) the spatiotemporal distribution patterns of the economical factors, 3) the multi-granular patterns are presented to the user from the tool.

4.1 Interactive analysis

We take one of the most economic developed county in Jiangsu Province, Jiangyin County, as an example to visually analyze the underlaying patterns by using the dashboard, as the Figure 5 shows. Firstly, the county Jiangyin is selected in the map view, shown as highlighted in the the choropleth map. We can read that Jiangyin lies on the south of Jiangsu and had a high GDP value revealed by the dark red color of the county in the map. What is more, the neighboring counties were as well with high GDP values,

as the red color shows. Secondly, we select 2010 in the time view. We can see Jiangyin is highlighted as the color of the other counties are faded in the streamgraph. The temporal pattern of Jiangyin follows the temporal pattern of Jiangsu, as the shape of Jiangyin stream is similar with the overall stream. The GDP kept increasing, while with a few fluctuations from 1980 to 2010. More specifically, the GDP increase of Jiangsu from 1996 to 1990 is relatively slow and smooth, whereas the increase from 1992 to 2008 is exponential. However, it is followed by a dramatic drop in the year of 2009. In terms of spatial comparison, among all the counties in Jiangsu Province, Jiangyin holds one of the widest bars in the streamgraph, meaning Jiangyin is one of the most economically developed counties in Jiangsu. After the selection of county and year, more details can be found in the other views of the dashboard. The numerical GDP values can be seen from the data view and text description view. Unlike text description view which gives only the information of the selected item, the data view showed all the available GDP values with Jiangyin being highlighted. The GDP of Jiangyin in 2010 was 200.902 billion Chinese Yuan, which is higher than earlier years of the area. Comparing with the other counties, the Jiangyin GDPs were always in a leading position among the counties during those years. From the pie chart view, we notice that the GDP of Wuxi City took 14% of Jiangsu in 2010 (the second biggest proportion), whereas Jiangyin took 35% of Wuxi. Last but not the least, the bar chart shows the changing of Jiangyin GDP along those 31 years. It is clear that the pattern of the bar chart is similar with the pattern of the streamgraph. The GDP of Jiangsu had a positive correlation with the GDP in Jiangyin; and the increasing speed of the GDP of Jiangsu and Jiangyin were also positively related. Therefore, we guess that the economic development of Jiangyin was synchronous with the regional development.

4.2 Temporal analysis of GDP data at the city- and county-level

From the dashboard, we pick the yearly choropleth maps at city-level and county-level and then compose them into small multiples.

Figure 6 shows the city-level choropleth maps of the GDP from 1980 to 2010. Generally, the GDP of the cities kept increasing in the 31 years: the GDP of the cities in south increased faster than the cities in north, the GDP of the cities in middle is the slowest. Temporarily, 1992 is the turning point of economic booming, the GDP increasing was much faster than before. However, there is a dramatic decrease in 2009. Lastly, the city-level GDP data from 1981 and 1983 is missing.

We can see form the small multiple that the Jiangsu GDP kept increasing, but with unexpected patterns, i.e., increasing one rapid increasing turning point (in 1992) and a outlier (in 2009). It is hard to say what is the main reason for the fluctuations of the Jiangsu GDP. But some influential historical events can be listed here as a temporal GDP pattern reasoning aid. Deng Xiaoping had his southern tour speech in 1992 (Vogel, 2011). In 1998, the China experienced the severe flooding, including Yangtze River. In 2001, the Shanghai Cooperation Organization (SCO) was founded. In the same year, China joint the World Trade Organization (WTO). From 2002 to 2003, a severe acute
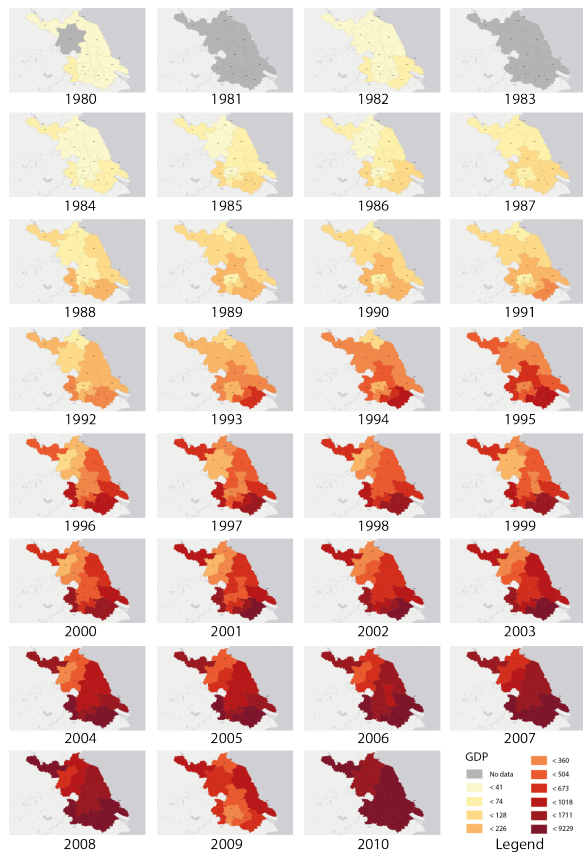
Figure 6. Small multiples of GDP distribution at city-level.

Figure 7. Small multiples of GDP distribution at county-level.

respiratory syndrome (SARS) exploded in most regions in China, with in total 4698 people effected. Similarly, in 2004, a bird flu effected a lot of people in China. In 2008, a great earthquake with measuring 8.0 surface wave magnitude happened in Sichaun, China. Plus, a global financial crisis occurred in 2008. We assumed that the Jiangsu GDP decreasing in 2009 was related to the financial crisis.

The availability of GDP data at county-level and its spatiotemporal distribution are shown in Figure 7. Obviously, the data availability was significantly increasing along the time. However, in the years from 1981 to 1984 and 1986, no county-level GDP data is available. Oddly, the data availability in urban area is lower than in rural area. Moreover, in the southern part of Jiangsu, the GDP increased faster than northern part of Jiangsu, which is the same as we observed from the city-level small multiples. Commonly, people would assume that Shanghai, lying on the south of the Jiangsu Province, leading the regional economical development. Therefore, the closer an area to Shanghai is, the higher GDP of the area would have. In Figure 6, we can see the spatial GDP change matches with this assumption at the city-level: the GDP of the cities close to Shanghai increased earlier and faster than the cities in the north. However, that assumption does not match at the county-level. For instance, the GDP of Jiangyin is higher than the GDP of Kunshan, where as Kunshan is closer to Shanghai than Jiangyin. Another interesting phenomena is, that Guanyun County lies close to the sea, but the GDP increased slowly, not as rapid increase the GDP of other coastal counties.

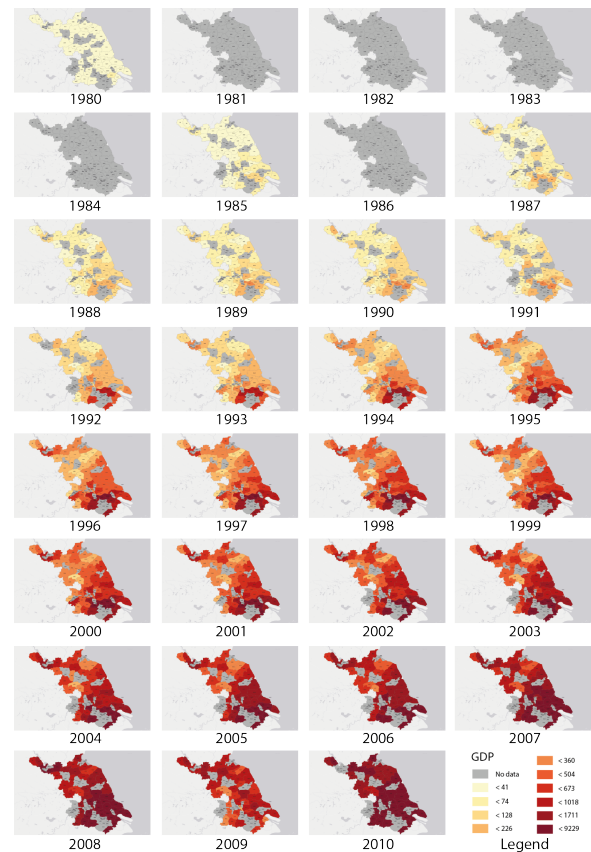From the small multiple visualizations, we analyze the spa-

tiotemporal GDP distribution. The area of different colors help us to identify patterns from the maps. The temporal trend is easily perceived by noticing the color change, which enables the automated analysis. Moreover, the contrast of gray and red color emphasized the area where the data is missing. However, the small multiples with choropleth maps have drawbacks. The visualization unit, county in this study, varies from 22.43 km^2 (Nanchang District) to 1869 km^2 (Dongtai County). Due to single color spreads over an entire county, it might mislead that the larger counties owns a higher GDP than smaller counties in the case of those counties sharing similar GDP density. Moreover, some details are not discernible because of the small size of the maps.

4.3 Analysis of GDP at the grid-level

From the dashboard, we obtain the gridded GDP heatmap of Jiangsu in 2010. As shown in Figure 8, the GDP is visualized as a heat map based on the $5km \times 5km$ grids. Besides, two overlays are added: rivers, canals and lakes were in blue, the ports were demoted by white circles with boats inside. The darker red the area is, the higher GDP it owns. According to the model introduced in Section 2.2, the input data should be county-level GDP data. However, spatial completeness of the county-level data is poor. Thus, we obtain the non-data areas' GDP value subtraction: in every city, the unavailable are GDP equals to the city-level GDP value minus all the available counties' GDP values.

We analysis the correlation of the GDP and land use types by comparing the model coefficients from Table 1. The water category owns the highest coefficient, followed by
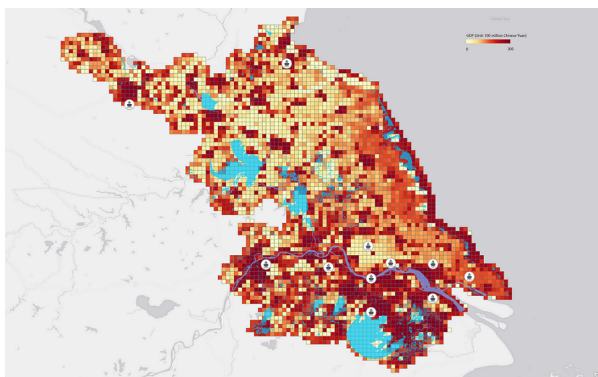
Figure 8. The GDP of Jiangsu Province in 5km * 5km grid, with the unit value of 100 million Chinese Yuan.

build-up land, water, agricultural land, barren land and forest respectively. Additionally, middle coverage grassland, urban land, alkali land and sea beaches show a strong positive correlation with the GDP. Conversely, other woodland, sandy areas and base exposed rock land use types show a strong negative correlation. In the subcategories, highland agricultural land owns a more positive correlation with GDP than plain agricultural land. Surprisingly, sea beaches are much more positively related with GDP than river/lake beaches.

With the help of visual analysis from Figure 8, we realize that the urban areas and coastal areas have relatively high GDP values, whereas lakes and industrial areas have relatively low GDP values. Although the middle coverage grassland has the highest coefficient, it does not influence the GDP a lot in the $5km \times 5km$ grid because of its small occupation area. There are two large high GDP regions. One is at south Jiangsu and the other lies along the coast. The south part of Jiangsu with high GDP is located close to Yangtze River with a lot of ports. Therefore, we may reason that water transport is a very important factor contributing to the local economic.

5. Conclusion

In this study, designed the visual analytics framework improved the understanding of the economic environment. Specifically, we visually examined the GDP data in Jiangsu Province. Visualization methods such as choropleth map, heat map, small multiples were designed to show the regional spatiotemporal data distribution. A dashboard as an interactive user interface was designed for users to explore the data freely. Plus, a spatial interpolation model was applied to analyze the GDP data at a fine scale. The process of the work was done iteratively: the results of visualizations, user interactions and the modeling inspired and stimulated each other, and they formed a complementary analysis. Moreover, the design of the dashboard could help the data distributors to publish their data in an easy-understandable way, which accelerates the flow of the shared information, and increases the accessibility of public to the information of economic environment.

The design of the dashboard is suitable for the economic environment factors illustration, as the spatiotemporal patterns are shown, both of the overviews and detailed views are provided and the interactivity of the dashboard meet different needs of users. However, the data grouping and

color design for a series of choropleth maps is a a trade-off between too many groups of indistinguishable colors in some maps and pattern suppression in the data tails. In the small multiple visualization, we can see the colors of choropleth maps in the early years and late years are similar. Although we have 10 GDP groups, the patterns in the data tail is suppressed.

With the aid of the dashboard, we visually analyzed the economic environment qualitatively and quantitatively. Both of the general patterns and data details were shown in the visualizations. The expected and unexpected spatiotemporal GDP patterns are revealed, such as the GDP growth differences between south and north of Jiangsu, the correlation between GDP and land use. For this reason, we assume that the economic environment can be reflected by the GDP patterns. Moreover, the visualizations are informative: e.g., the overlays give an intuitive view of the correlation between the GDP and other factors, the 31 years' temporal trend of GDP in a whole province. The interactive mapping allows users to customize the shown data and explore it freely. Additionally, the GDP interpolation model breaks the administrative boundary limitations for the analysis, so that the local economic development was analyzed on a higher resolution.

However, this research is still at its preliminary stage. Due to the complexity of the economic environment, the reasons of regional inequality are hard to explain. Many other factors, e.g., population, education, industry structure, logistic, policy need to be integrated in our visual analysis in the next step. Meanwhile, in the future, the multivariable visualization methods could be evaluated in a usability study. The user test of the dashboard and the visualization will be performed and analyzed.

6. Acknowledgements

References

Bostock, M., 2016. Merged choropleth. https://bl.ocks.org /mbostock/670b9fe0577a29c39a1803f59c628769.

Brunsdon, C., Fotheringham, S. and Charlton, M., 1998. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)* 47(3), pp. 431–443.

Bureau, J. P. S., n.d. Statistical communique of the national economic and social development of jiangsu province in 2018. http://tj.jiangsu.gov.cn/art/2019/3s/8/ art_4031_8257205.html.

Corbett, J., 2012. Charles joseph minard: Mapping napoleon's march, 1861.

Few, S., 2006. Information dashboard design.

Gaffin, S. R., Rosenzweig, C., Xing, X. and Yetman, G., 2004. Downscaling and geo-spatial gridding of socio-economic projections from the ipcc special report on emissions scenarios (sres). *Global Environmental Change* 14(2), pp. 105–123.

Grübler, A., O'Neill, B., Riahi, K., Chirkov, V., Goujon, A., Kolp, P., Prommer, I., Scherbov, S. and Slentoe, E., 2007. Regional, national, and spatially explicit scenarios of demographic and economic change based on sres. *Technological forecasting and social change* 74(7), pp. 980–1029.

He, C., Huang, Z. and Wang, R., 2014. Land use change and economic growth in urban china: A structural equation analysis. *Urban Studies* 51(13), pp. 2880–2898.

Huang, Y., Jiang, D., Zhuang, D., Zhu, Y. and Fu, J., 2014. An improved approach for modeling spatial distribution of water use profit—a case study in tuhai majia basin, china. *Ecological indicators* 36, pp. 94–99.

Kohlhammer, J., Keim, D., Pohl, M., Santucci, G. and Andrienko, G., 2011. Solving problems with visual analytics. *Procedia Computer Science* 7, pp. 117–120.

Le Clec'h, S., Sloan, S., Gond, V., Cornu, G., Decaens, T., Dufour, S., Grimaldi, M. and Oszwald, J., 2018. Mapping ecosystem services at the regional scale: the validity of an upscaling approach. *International Journal of Geographical Information Science* 32(8), pp. 1593–1610.

LeSage, J. and Pace, R. K., 2009. *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Lin, G., Fu, J., Jiang, D., Hu, W., Dong, D., Huang, Y. and Zhao, M., 2014. Spatio-temporal variation of pm2. 5 concentrations and their relationship with geographic and socioeconomic factors in china. *International journal of environmental research and public health* 11(1), pp. 173–186.

Shenggen, F. and Zhang, X., 2004. Infrastructure and regional economic development in rural china. *China economic review* 15(2), pp. 203–214.

Vogel, E. F., 2011. *Deng Xiaoping and the transformation of China*. Vol. 10, Belknap Press of Harvard University Press Cambridge, MA.

Webbera, R., Herbertb, R. D. and Jiangbc, W., 2006. Space-filling techniques in visualizing output from computer based economic models. In: *Computing in Economics and Finance*, Vol. 67.

Wu, W., Xu, J., Zeng, H., Zheng, Y., Qu, H., Ni, B., Yuan, M. and Ni, L. M., 2016. Telcovis: Visual exploration of co-occurrence in urban human mobility based on telco data. *IEEE transactions on visualization and computer graphics* 22(1), pp. 935–944.

Xu, Z., Cai, Z., Wu, S., Huang, X., Liu, J., Sun, J., Su, S. and Weng, M., 2018. Identifying the geographic indicators of poverty using geographically weighted regression: A case study from qiandongnan miao and dong autonomous prefecture, guizhou, china. *Social Indicators Research* pp. 1–24.

Zheng, Z. and Bohong, Z., 2012. Study on spatial structure of yangtze river delta urban agglomeration and its effects on urban and rural regions. *Journal of Urban Planning and Development* 138(1), pp. 78–89.

Zilong, J., 2014. Analysis of agglomeration and equilibrium characteristics of regional economic in china with spatial visualization in science computing technique. In: *2014 Sixth International Conference on Measuring Technology and Mechatronics Automation*, IEEE, pp. 91–94.