Construction of regional economic development model based on remote sensing data

Hailing Gu¹, Chao Chen¹*, Ying Lu¹, Yanli Chu², and Yuxiang Ma¹

¹Marine Science and Technology College, Zhejiang Ocean University, Zhoushan, PR China
²School of Port and Transportation Engineering, Zhejiang Ocean University, Zhoushan, PR China

* Corresponding author: chenchao@zjou.edu

Abstract. In order to break through the time-consuming and laborious limitations of traditional regional economic development surveys, this study will build some regional economic development models by virtue of the advantages of remote sensing technology. First, based on multi-source and multi-temporal satellite remote sensing data, obtain surface morphological changes and land use information; second, analyze the correlation between land use types and regional economic indicators, and optimize sensitive factors; third, combine the social survey data to build a regional economic development model; finally, an accuracy evaluation is performed to verify the validity and applicability of the model. Zhoushan City, China, is selected as the research area to carry out verification experiments. The experimental results show that the construction land area is the most sensitive factor relating to various economic indicators, and the correlation coefficients with GDP, PPI, PSI and PTI are respectively 0.9591, 0.9390, 0.9546 and 0.9573. The coefficient of determination R² of the regional economic development model built with the survey data is 0.9884. This study provides a new way of thinking for regional economic development prediction and economic data correction as well as a possibility for humans to observe economic activities and their impact. The model built in this study is simple and clear yet with high precision. It is of great significance for understanding regional economic development, adjusting and correcting statistical data.

1. Introduction
Regional economic development refers to various economic activities such as economic growth, population growth, upgrading of industrial organizations and structures within a certain time-space. Assessing regional economic development is conducive to objectively understanding environmental and social issues such as resource consumption and environmental pollution resulted by regional economic development, and is of great significance for maintaining social and human sustainable and healthy development [1].

Usually, the regional economic development is mainly assessed based on the statistics such as indexes of gross domestic product (GDP), proportion of primary industry (PPI), proportion of secondary industry (PSI) and proportion of tertiary industry (PTI) etc. However, defects exist in statistical data such as data missing or lack of spatial information [1, 2]. On the other hand, statistical data is often obtained by means of “field investigation, layer-by-layer reporting”, which consumes high-cost and is limited to subjective factors and exist plenty of uncertainty. Therefore, how to obtain
the real information reflecting regional economic development and how to objectively understand regional economic development, now become hot topics attracting widespread concern of researchers. Economic activities related to regional economic development are often reflected in the transformation of the earth's surface and the changes in land use types. Remote sensing technology has the advantages of wide observation range, long time series and low acquisition cost [3]. It can objectively reflect the intensity of land use and economic activity, and facilitate to understand regional land use and cover changes from a macroscopic view and assess regional economic development [4]. In view of the problems existing in traditional regional economic development assessment, this paper intends to use the multi-source and multi-temporal satellite remote sensing data to establish the relationships between land use types and various economic indicators, and to construct a regional economic development model based on remote sensing data [5]. This research is of great significance for understanding and assessing regional economic development, verifying social survey data, and so on.

2. Research methods and materials

2.1. Research area
Zhoushan City is located in northeast of Zhejiang Province, China with a geographical area from 121°30'E to 123°25'E, and 29°32'N to 31°04'N, right facing the Pacific Ocean. It is the first prefecture-level city in China composed of group of islands. As shown in Figure 1, Zhoushan has 1,390 large and small islands. The hills on the island are undulating, with high hills accounting for 9%, low hills accounting for 61%, and plains 30%, forming different soil types and crop patterns. According to statistics from the Zhoushan Municipal Bureau of Statistics, there were 504,000 permanent residents in 1982, and the resident population increased to 1.168 million to 2017.

Fig.1. Schematic diagram of the study area.

2.2 Research data
According to the research needs, this paper collects two aspects of data. One is the Landsat satellite remote sensing data from 1984 to 2017, and another is the economic indicator and population data released by the Zhoushan Municipal Bureau of Statistics including GDP, PPI, PSI and PTI etc. This paper divides the statistical data into two groups, one totaling 18 years for model construction dated 1884, 1987, 1990, 1991, 1993, 1997, 1998, 1999, 2001, 2003, 2005, 2006, 2008, 2010, 2013, 2014, 2016, 2017, another totaling 9 years for model validation dated 1988, 1992, 1996, 2000, 2004, 2007, 2009, 2011 and 2015.

2.3 Research method
Based on long-term sequence satellite remote sensing data, this study selects the most sensitive factors from the perspective of correlation, and constructs a regional economic development model to analyze the regional economic development situation. This research mainly includes data preprocessing, image classification, sensitivity factor selection, model construction, and accuracy evaluation. The technical roadmap is shown in Fig. 2.
Firstly, pre-processing operations such as radiometric calibration, atmospheric correction, image splicing and cropping are performed on the collected long-term sequence remote sensing images. The gradation values are converted into radiance data with clear physical meaning by radiometric calibration, and atmospheric correction is performed, the effects of atmospheric absorption and scattering are removed, and reflectance data that reflects the real situation on the surface is obtained. Image stitching and cropping are used to obtain Landsat satellite data covering the entire study area. Secondly, the training samples can be selected in considering of the distinctions of the target features in terms of spectrum, shape, texture, etc. and the supervised classification method is used to divide the research into construction land (including urban and rural buildings, port terminals, roads, etc.), vegetation (including forest land and farmland), water bodies (river lakes, farmed paddy fields), and bare land (including tidal flats, sandy land etc.), then, the total area of each type of land use is available. Thirdly, we introduce economic data and analyze the correlations between the indexes of GDP, PPI, PSI, PTI, household GDP per capita and various types of land areas. Among the correlations, the most relevant feature categories are selected as the sensitive factors. Then, based on the remote sensing classification data, we consider comprehensively the impact of resident population on regional economic development and construct a regional economic development model by the regression statistical analysis method. Finally, under the support of statistical data, indicators such as percent deviation, root mean square error, and correlation coefficient will be counted to evaluate the accuracy of the model.

3. Experimental results and analysis

3.1 Classification result
In this paper, the classical maximum likelihood classifier in the supervised classification is used to divide the research into four categories (i.e. construction land, vegetation, water body, bare land, etc.) (Fig. 3), and the area of each category is counted on this basis. The results are shown in Table 1. Show. From the classification results, the vegetation coverage of Zhoushan City is maintained at a high level, and the land area is increasing. From 1984 to 2017, the vegetation coverage area of Zhoushan was relatively stable; the construction land area increased from 50.8 km² to 220.0 km², an increase of approximately 3.3 times; the total water area showed an increasing trend.

3.2 Sensitive factor selection
Correlation analysis was carried out between each economic indicator and each land use type, and the most relevant feature category was selected as the economic sensitivity factor. In this study, in order to
remove the data redundancy, the economic indicators and land use type data are converted logarithmically and exponentially in order to certain the data correlations more accurately. The correlation between the construction land area and indexes of GDP, PPI, PSI, PTI is the strongest after logarithmic transformation, and the correlation coefficients are respectively 0.9591, 0.9390, 0.9546 and 0.9573.

Table 1. Areas of land use type.

| Time /yr | 1984 | 1987 | 1988 | 1990 | 1991 | 1992 | 1993 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2003 |
|----------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Vegetation /km² | 972.51 | 950.54 | 994.48 | 990.62 | 967.19 | 974.73 | 1043.44 | 956.23 | 955.41 | 1020.90 | 1039.48 | 998.41 | 1025.00 | 988.57 |
| Water /km² | 28.44 | 60.18 | 50.45 | 71.42 | 49.45 | 81.43 | 45.96 | 42.40 | 34.89 | 57.74 | 40.67 | 48.43 | 51.56 | 47.31 |
| Construction Land /km² | 50.84 | 39.17 | 62.55 | 56.23 | 56.03 | 53.86 | 47.76 | 86.13 | 78.55 | 107.19 | 73.07 | 107.83 | 111.08 | 126.06 |
| Bare Land /km² | 255.55 | 247.57 | 188.59 | 157.36 | 221.10 | 170.19 | 211.54 | 188.53 | 114.52 | 83.03 | 126.23 | 130.59 | 118.22 |
| Total Area /km² | 1307.34 | 1297.45 | 1296.07 | 1275.76 | 1280.21 | 1268.28 | 1296.31 | 1259.38 | 1300.36 | 1236.24 | 1280.90 | 1318.24 | 1280.16 |

Fig.3. Classification results graph.
3.3 Regional economic development Model Construction
According to the selected sensitive factors, the regional economic development model is constructed based on the comprehensive consideration of single-factor (construction land area) and two-factor (resident population and construction land area). The results are shown in Table 3. It can be seen from Table 3 that the two-factor model involving the influence of resident population is significantly better than the single-factor model, and the average correlation coefficient $R^2$ of the fitting function is increased from 0.9101 to 0.9795.

4. Accuracy evaluation
In order to verify the validity and applicability of the model, the percent deviation, root mean square error, and correlation coefficient between the measured data and the model calculation data are calculated. The results of the accuracy evaluation are shown in Table 3. It can be seen from the results that the regional economic development model combined with remote sensing data and social survey data has higher accuracy, and the average correlation coefficient with the measured economic indicators is 0.9594, the average percent deviation is -7.29%, and the average root mean square error is 4.099 billion yuan.

| Economic index | Single factor | Double factor |
|----------------|---------------|---------------|
|                | D | RMSE | $R^2$ | D | RMSE | $R^2$ |
| GDP            | 8.44% | 107.66 | 0.9465 | -8.50% | 78.11 | 0.9670 |
| PPI            | 6.79% | 9.44 | 0.9525 | -3.06% | 9.87 | 0.9890 |
| PSI            | 6.50% | 123.75 | 0.9013 | -11.59% | 46.67 | 0.9083 |
| PTI            | 16.56% | 64.93 | 0.9453 | -6.00% | 29.33 | 0.9735 |
| Average value  | 9.57% | 76.44 | 0.9364 | -7.29% | 40.99 | 0.9594 |

The unit of RMSE : 100 million RMB

5. Conclusion and Outlook
This paper combines land use/cover change information and statistical data to construct a regional economic development model combining remote sensing data and social survey data. The study found that the construction land can well represent the basic situation of regional economic development. With higher goodness of fit and lower prediction error, the regional economic development model, established by using the construction land area and the resident population, are more applicable. The correlation coefficients between GDP, PPI, PSI, PTI model prediction results and real results are 0.9670, 0.9890, 0.9083, 0.9735, respectively. The average percentage deviations are -8.50%, -3.06%, -11.59%, and -6.00%, respectively.

This study uses the medium-resolution satellite remote sensing data to classify the land use. In the future, higher spatial resolution remote sensing data will be used to obtain more detailed classification data, eliminate the comprehensive effects of multiple interference factors in the coarse category, and to
dig deep into the relationship between local object categories and various economic indicators. In addition, more detailed statistical data will be collected to reduce the impact of statistical data in the modeling process, which is also the next step of this work.

Acknowledgment
We are very grateful to the anonymous reviewers for their comments and suggestions. This work was supported by the National Natural Science Foundation of China (41701447), the Project supported by the Open Foundation from Fishery Sciences in the First-Class Subjects of Zhejiang Ocean University.

References
[1] G.C. Nelson, R.D. Robertson, *International Journal of Remote Sensing* **28**, 19 (2007).
[2] C. Li, J. Wang, L. Wang, L. Hu, P. Gong, *Remote Sensing* **6**, 2 (2014).
[3] G. Grekousis, G. Mountrakis, M. Kavouras, *GIScience & Remote Sensing* **53**, 1 (2016).
[4] P.E. Osgouei, S. Kaya, *Environmental monitoring and assessment* **189**, 4 (2017).
[5] M.Z. Al-hamdan, P. Oduor, A.I. Flores, S.M. Motikot, R. Mugo, J. Ababu, H. Farah, *International Journal of Applied Earth Observation and Geoinformation* **62**, (2017).