Double Vagueness: Uncertainty in Multi-scale Fuzzy Assignment of Duneness

CHENG Tao   Pete Fisher   LI Zhilin

ABSTRACT   In the automation of identification of landscape features the vagueness arises from the fact that the attributes and parameters that make up a landscape vary over space and scale. In most of existing studies, these two kinds of vagueness are studied separately. This paper investigates their combination in identification of coast landscape units. Fuzzy set theory is used to describe the vagueness of geomorphic features due to the continuity in space. The vagueness resulted from the scale of measurement is evaluated by statistic indicators. The differences of fuzzy objects derived from data at differing resolutions (in size from 3×3 cells to 25×25 cells) are studied in order to examine these higher-order uncertainties.

KEY WORDS   vagueness; uncertainty; multiscale; fuzzy spatial objects

CLC NUMBER   P237

Introduction

The surface parameters of a landscape can vary in a spatial nature as well as with scale. The continuous distribution of conceptual features and landform parameters over space means that it is not possible to draw a crisp boundary between the foreshore and the beach[1]. Further, it is widely accepted that a single scale of analysis is insufficient for accurate description or characterization of a landscape, as this would rely too heavily on the resolution of the original DEM. For example, a channel at one scale would be a ridge at a different scale[2]. When classifying a landscape in terms of surface parameters, the scale of analysis must be considered. Therefore, in the automation of identification of landscape features the vagueness arises from the fact that the attributes and parameters that make up a landscape vary over space and scale.

Fuzzy set theory can be used to model vagueness,

imprecision and in-accuracy. It has been successfully used, for example, to describe the accuracy of land cover classifications, imprecision in geographical boundaries and vagueness in geographic objects[1,3,4]. The assignment of fuzzy membership values is not clearly understood, however. The assignment of fuzzy membership values is always subjective, subject to algorithm, spatial scales, and individual preference, etc., giving rise to a higher level of uncertainty, and often making this the weakest part in the application of fuzzy set theory.

Scale is a major unsolved issue in geographical information related sciences although some attempts have been made. For example, in cartography, "how to derive small scale maps from large scale maps" is a key issue for automated map generalization[5]. In human geography, "how to aggregate data from small enumeration units to larger units for processing" on spatial analysis and modelling is called "the modifiable areal unit problem"[6]. In physical geography, "how to extrapolate informa-

tion across scales" is often being asked to improve the cost-befit ratio of sampling[7].

When vagueness is involved in the definition of the geographical objects, the multiscale problem makes the modeling and representation of the geographical objects more complicated and uncertain; there is not just uncertainty about the extent of the object, but about the estimate of the uncertainty of the extent of the object. However, very little research has looked at this higher order vagueness: vagueness from both space and scale, at the same time. "Will scale affect the result of modeling fuzzy spatial objects" is an un-tackled question. The research reported here aims to answer this question.

1   Modeling of fuzzy spatial objects

The inherent characteristics of geographical entities, i.e. continuity, heterogeneity, dynamics and scale-dependence, imply that most of them are naturally indeterminate or fuzzy[8]. If we want to model geographical entities in GIS, we have to first define the objects and then measure them in term of the parameters involved in the definition. Since the definition of objects in GIS is related to category theory[9], the indeterminacy of geographical entities is revealed in the uncertainties in category theory, as the definition of the category might be fuzzy, multi-criteria, spatially incomplete, and/or time-incoherent. If one of the reasons is valid, the object class should be defined as an uncertain category. If we use set theory to define the class of the geographical entities, we may use fuzzy set theory or rough set theory to model the vagueness[9-11].

A number of researchers have also introduced the idea that terrain objects are fundamentally vague, and may be appropriate for analysis by fuzzy sets. Vague landscape features have been defined by two methods: either elevation has been as the basis of a semantic import model where some a priori knowledge is used to assign a certain value of fuzzy membership of a landscape feature to a particular height above datum[1, 12], or they use a number of surface derivatives, such as slope and curvature, in a multi-

variate fuzzy classification[13, 14].

2   Modeling of scale

Scale can refer both to the level of detail of a description, and to the scope or extent of the area covered. As mentioned above, scales are inherent in the ways process operates. To deal with scale in modeling human and physical systems, and to model the effect of scale on description is a challenging issue in geographical information science. In cartography, maps are produced at certain scales with different application, e.g. 1:10 000 and 1:100 000. Small-scale maps provide better overview while large-scale maps provide more detailed and precise information. It is intuitive that the same number of map symbols cannot be represented when the map scale is smaller. It means that the representation of the same features on the ground will be different on maps of different scales. The issue arising is "how to derive small scale maps from large scale maps" through operations such as simplification, aggregation and selective omission[15]. This issue is on the representation of spatial data and is called "map generalization". As map generalization is not directly relevant to current research, it will not be discussed further here.

In geography, there is a similar issue. Normally, geographical data are sampled in small enumeration units (also called small scale), and in some applications, these data need to be aggregated to a larger enumeration unit. However, the statistical results will be different when the analysis is carried out on the basis of different size of enumeration units (specially on the zones used to produce aggregate statistics, i.e. different scales), and different aggregations of the same size. Therefore, there is an issue of "how to aggregate data from small enumeration units to larger units for processing". This issue is called "the modifiable areal unit problem"[6].

There is a similar issue in all geographical information related sciences, such as geomorphology, oceanography, soil science, biology, biophysics, social sciences, hydrology, environmental sciences

and landscape ecology. In general, there are two related but distinctive goals for conducting a multiscale analysis in these studies. The first is to characterize the multiscale structure of a landscape. The second is to detect or identify "scale breaks" or "hierarchical levels" in the landscape, which often can be studied as a spatially nested hierarchy.

Two approaches to multiscale analyses are possible: ① the direct multiscale approach that use inherently multiscale methods, and ② the indirect multiscale approach that uses single methods repeatedly at different scales. Frequently used multiscale methods include semivariance analysis, wavelet analysis, fractal analysis, lacunarity analysis, and blocking quadrate variance analysis. All these methods contain multiscale components in their mathematical formulation or procedures, and thus are either hierarchical or multiscaled[16]. On the other hand, the indirect approach to multiscale analysis can use methods redesigned from single scale analysis, such as a wide variety of landscape metrics (e.g. diversity, contagion, perimeter-area ratios, spatial autocorrelation indices) as well as statistical measures (e.g. mean, variance, correlation or regression coefficients). The scale multiplicity in the indirect approach is realized by resampling the data at different scales, albeit grain or extent, and then repeatedly computing the metrics or statistical measures using sampled data at different scales[17].

Many studies have dealt with numerical aggregation such as zoning or modifiable area unit problems[18, 19]. Some studies have used categorical aggregation based either on a majority or a random rule[20]. The statistical approach has been broadly applied in multiscale analysis, just as it has been widely used to model spatial uncertainty and its propagation. However, Wood (1996) had used fuzzy sets to calibrate the vagueness resulting from multiscale analysis[21].

Recent research on scales in GIS can be found in Reference [22]. Five key issues of scales such as, "changing the scale of measurement", "non-stationary modeling", "dynamic modeling", "conditional simulation" and "constrained optimization",

are put forward as recommended for further research for GI Science. It is argued that while regularization provides an important tool for modeling change of scale, it does not solve the problem of changing the scale of measurement for an actual data layer. When changing the scale of measurement is facilitated by interpolation, the inherent smoothing which results in the predicted values may alter the bivariate distribution between that variable and any other. Solutions based on simulation are inadequate. Therefore, the issue of "Changing the scale of measurement" is the most important and should be given the highest priority by researchers among these five problems.

In summary, although fuzzy set theory has been widely used in GIS, the scale issue has not been investigated; while statistical approaches has been applied in multiscale analysis, the fuzzy aspect of the geographical features is ignored. This current research attempts to combine these two relevant and inherent issues by studying the effect of scale on modeling fuzzy spatial objects, as an approach to modeling the higher order fuzziness in spatial objects: double fuzziness.

3   Case and methodology

3.1   Case study area

A barrier island, Ameland, in the north of the Netherlands is adopted as a case study here (Fig. 1). The process of coast change involves the erosion and accumulation of sediments along the coast. It can be monitored through the observation of changes of landscape units such as foreshore, beach and foredune. The process of coastal change is scale-dependent in space and time.

The landscape units are defined on the basis of water lines. The foreshore is the area above the closure depth and beneath the low water line, beach is the area above the low water line and beneath the dune foot, the foredune is the first row of the dunes inland from dune foot. These
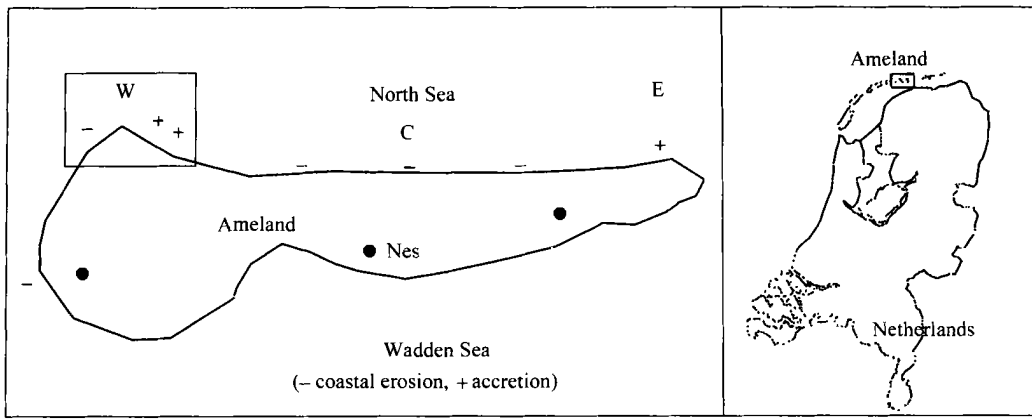
Fig. 1   Test site— Ameland, The Netherlands

definitions are normally different from surveyor to surveyor, from case to case and from time to time. For example, the low water line was set to be −6 m in 1965 to 1984 and 1989, and −8 m in 1985 to 1988 and in 1990 to 1993. Therefore, the extent of these landscape units are a fuzzy concept, but on the basis of height observation, it is possible to derive a measure of foreshore, beach and duneness.

Height observations have been made by laser scanning of the beach and dune area and by echo sounding on the foreshore. These data have been interpolated to form a full height raster of the test area. Experiments show that the uncertainty of the interpolated heights of the raster can be expressed by standard deviation ($\sigma = 0.15$ m). However, in the following analysis, the error of the height raster, which was used as the original fine resolution DEM, is ignored.

3.2   Multiscale analysis of DEM

Since the hierarchical analysis does not have to assume the existence of a hierarchical structure in the landscape under study, we adopted the indirect approach to multi-scale analysis. The multiplicity of scales is realized by resampling the data at different resolutions, resolution acting as a surrogate for scale, and then repeatedly computing the statistical measures using sampled data at different resolutions. One way of resampling data is to systematically aggregate the original fine-resolution data set and produce a hierarchically nested data set, which

leads to a hierarchical analysis using single-resolution methods. We have used the software Landsurf developed by Wood (2003) to aggregate the original fine-resolution DEM (60 m × 60 m) to coarse data sets, using a moving window ranging in size from 3 × 3 cells to 25 × 25 cells. In Wood's software the surface is modeled as a quadratic surface using the central point and the outer points of an expanding window, and calculating a generalized value of the elevation for the centre point of the surface[2]. The characterization of scale-based uncertainty so far has been described independently of the model of the surface and any operational definition of scale itself. In this way, a series of DEMs with cell size from 60 m × 60 m (1 × 1 cells) to 1 500 m × 1 500 m (25 × 25 cells) are created.

3.3   Fuzzy classification

As described in Section 3.1, the extent of the coastal landscape units are a fuzzy concept, but based on height observation, it is possible to derive a measure of foreshore, beach and duneness. We use the fuzzy set to represent the vagueness in the definition of these landscape units. The fuzzy membership function is built to modify the crisp classification criteria and a trapezodial membership function was adopted[23] for fuzzy classification. According to the definitions given by geomorphologists for the situation of Ameland, the height values of the closure depth, low water line and dune foot are suggested to be about −6.0 m, −1.1 m, and 2

m, respectively. After fuzzy classification, each grid cell has a membership vector, containing a value for each of the three classes. As multiresolution DEMs were created in Section 3. 2, a series of fuzzy membership vectors were created at 13 resolutions.

3. 4　Identification of multiscale fuzzy geomorphologic objects

The estimation of the spatial extent of objects from the fuzzy classifications is related to the interpretation of fuzziness of the objects and their topological relationships, hence a pre-defined fuzzy object model is required. For example, if foreshore, beach and foredune are considered to be spatially disjointing objects, the conceptual model suggests that a specific location should either belong to beach or foredune, but not to both and a boundary has to be set to define explicitly the spatial extent of any object by assigning each grid cell to exactly one object. In such cases criteria（conditions）have to be applied to assign a cell to a specific class. After segmentation, the spatial extents of objects are identified and the boundaries between them are apparent automatically. These boundaries are called conditional-boundaries since they are based upon conditions（or criteria）. In this case, the concept of objects with fuzzy spatial extent is applied, which means the objects are represented as fields with varying fuzziness and conditional boundaries[8] .

4　Effect of scale on fuzzy spatial objects

In the beginning of the multi-scale analysis, we actually proposed two opposing hypotheses: fuzzy spatial object will vary smoothly with changing grain size as pixels are aggregated reflecting a decrease in variability; or fuzzy spatial objects will show discrete changes as grain changes. Additionally, we wanted to determine if these changes could be modeled and, if so, could these models predict scale change effects on fuzzy spatial objects at either finer or coarser scales. In order to test the hypothe-

ses, we use statistical analysis. We calculate the total cells belonging to three landscape units based on the effective image window created at grain size 25 $\times$ 25 cells. Then we calculate by scale the mean, minimum and standard deviation of fuzziness for those cells belonging to each landscape units, respectively. Please notice the regression equations, y_1, y_2 and y_3, in the following figures represents the relationship of foreshore, beach and foredune with scale, respectively.

4. 1　Change in The Fuzzy Area

There is obvious change of the area of the three fuzzy objects with scale（Fig. 2）. With the increase of scale, the area of beach decreased till to scale at 15, but the area of foreshore and foredune increased till to scale at 9 and 15, respectively; then they change in opposite ways.

In order to test the hypotheses set in the beginning of this section, linear regression analyses and diagnoses were implemented. The results are also shown in Fig. 2. We can see that the area of all the landscape units change significantly（at significant level 95%）with scale, while foreshore and foredune are in positive relationships and beach in negative relationship. However, the coefficient of determination for foreshore and foredune（$R^2=0.35$）shows only about 35% of the variance of area is explained by its common variance with scale, suggesting low levels of explanation, indicating that other factors could be involved; but the coefficient of determination for the area change of beach indicates 67% variance is coming from the effect of scale.

4. 2　Change in mean of fuzziness

The change in mean of the fuzziness of three landscape units with scales is illustrated in Fig. 3. It can be seen from Fig. 3 that the changes in the mean of fuzziness of foreshore and beach have similar tune, mostly down with scale; but the change of foredune is unstable.

Regression analyses and diagnoses have also applied to the data above. The results are shown in
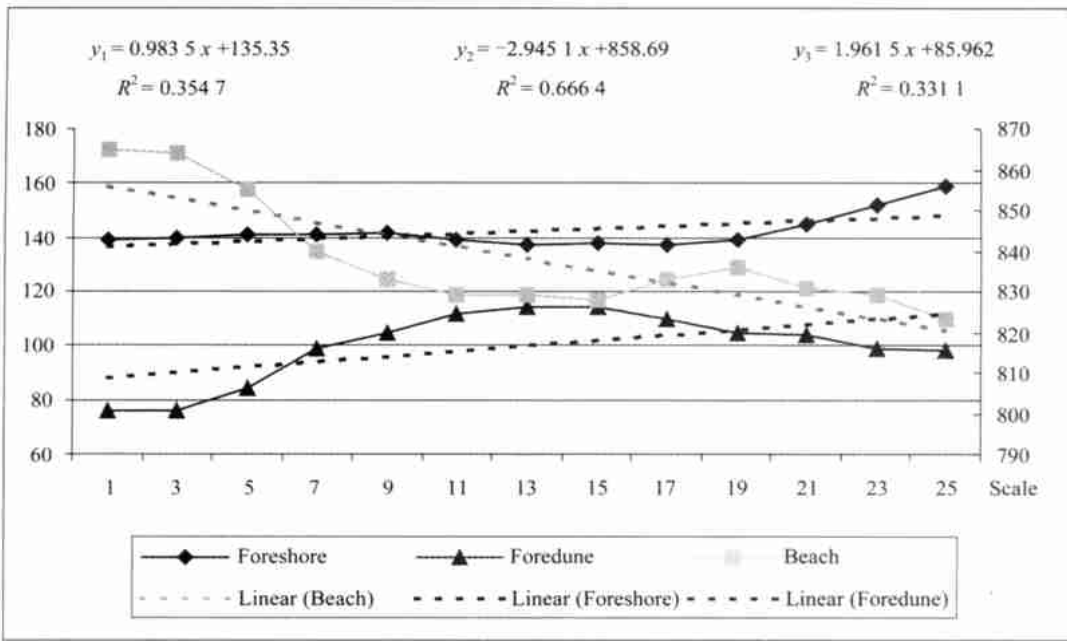
$y_1 = 0.983\ 5\ x + 135.35$ $y_2 = -2.945\ 1\ x + 858.69$ $y_3 = 1.961\ 5\ x + 85.962$

$R^2 = 0.354\ 7$ $R^2 = 0.666\ 4$ $R^2 = 0.331\ 1$

Fig. 2　Change in the fuzzy area of the three objects with scales

$y_1 = -0.002\ 2\ x + 0.946\ 7$ $y_2 = -0.000\ 4\ x + 0.966\ 1$ $y_3 = -0.008\ 2\ x + 0.939\ 8$

$R^2 = 0.795\ 7$ $R^2 = 0.042\ 3$ $R^2 = 0.649$
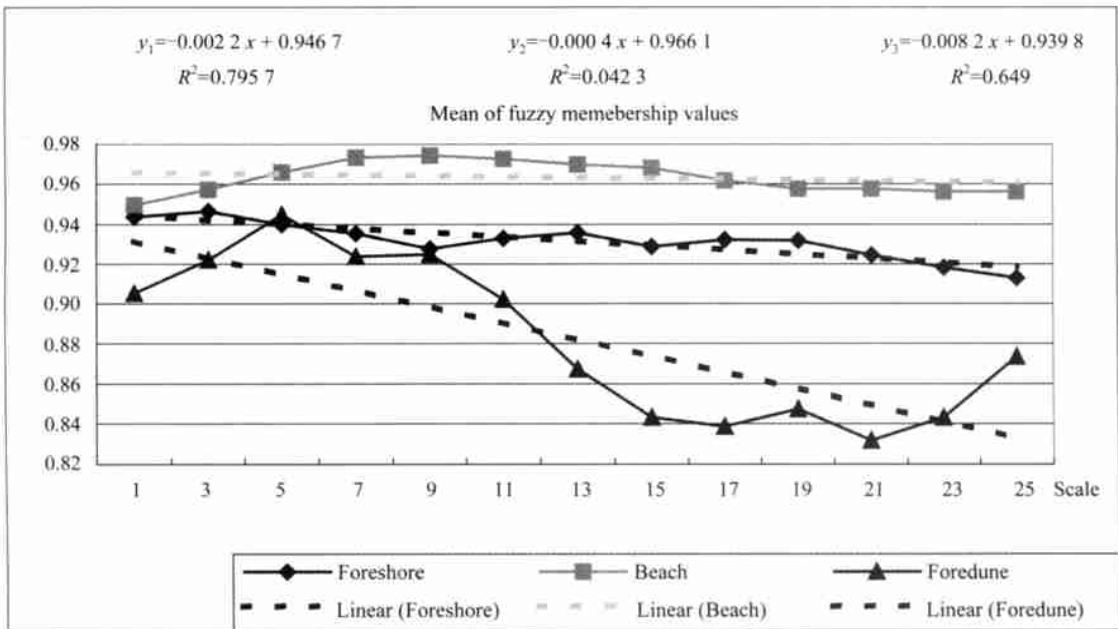
Mean of fuzzy memebership values

Fig. 3　Change in the mean fuzziness of three landscape units with the scale

Fig. 3. From the results we can only say that the mean of fuzziness of foreshore and beach change significantly and linearly (at level 95%) with scale. But for beach, the linear regression test is not passed.

4. 3　Change in minimum of fuzziness

The change in minimum of fuzziness of the three landscape units with scales is illustrated in Fig. 4, which shows cyclic patterns. But the amplitude and ranges are different for these three units. The change exhibits cyclic fluctuations, indicative of the periodic pattern in the landscape. Regression analyses reveal that there is no obvious linear correlation of the minimum fuzzy membership function value of the landscape units with scale. It is intuitively showed in Fig. 4.
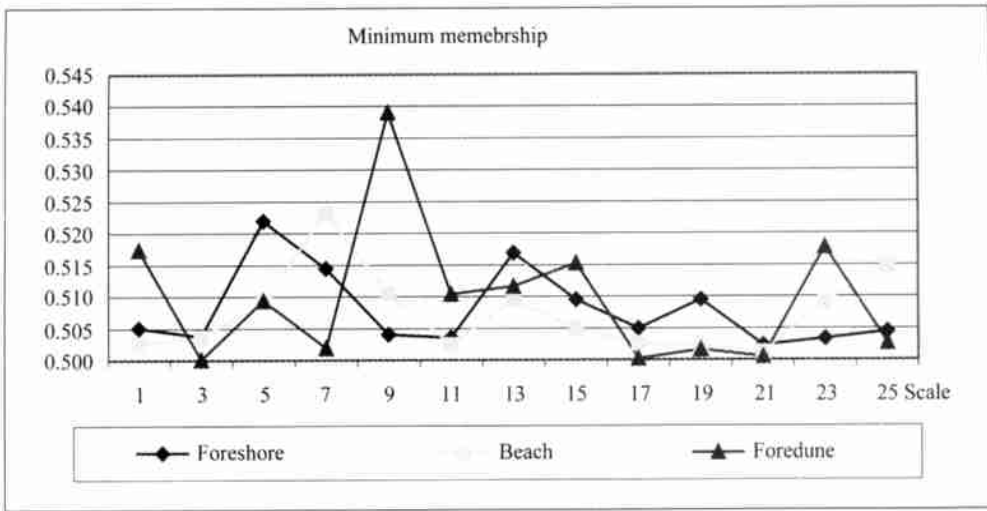
Fig. 4   Change of the minimum of fuzziness of three landscape units

4. 4   Change in standard deviation (σ) of fuzziness

We also calculated the standard deviation of the fuzziness of each landscape unit at different scales. There is no systematically change with the scale, but relevant (Fig. 5). It can be seen that a linear relationship exists between σ of the fuzziness of foreshore and of foredune with the scale. We can see that the standard deviation of the fuzziness of all the landscape units change significantly (at level 95%) with scale positively. The coefficient of determination for foreshore and foredune ($R^2 = 0.6$) suggests high levels of explanation, i. e. around 60% of the variance of STDEV is explained by its common variance with scale, but the coefficient of determination of foredune is relatively low as 0. 36, indicating that other factors might be involved in the variance.

4. 5   Discussion

In summary, we found that the area of three landscape units, the mean and the standard deviation of the foreshore and beach change significantly with scale; but the minimum of the fuzziness of the landscape units doesn't change significantly with scale. It implies that the scale has effect on fuzzy classification, i. e. the fuzzy membership values changed so that the class of the cells changed which

resulted in that the areas of fuzzy objects are different with scale. The change of STDEV is obvious with scale, if implies that the fuzziness of the landscape units increase with scale (becoming more uncertainty with scale). This is because the aggregation enlarges (and smoothes) the transition zone between the landscape units.

Further, we would like to say although the linear regressions have been applied, the coefficients of determination (R^2) are generally low, specially the area of the three landscape units, indicating low explanation of the variance of the statistical indicators with scale. It also suggests that the liner regression lines are not the best-fit lines. Therefore, polynomial trend lines are tried out. We may say that fouth-order polynomial line fits the trend of area change very well. The results are illustrated in Fig. 6.

5   Conclusions

In this paper, we evaluate the effect of scale on modeling fuzzy spatial objects, i.e., the subjectivity of the assignment of the fuzzy membership values to the scale of measurement. The work is illustrated by a coastal geomorphologic case. Multi-scale analysis of the landscape is carried out using a moving window, ranging in size from 3×3 cells to 25×25 cells. The differences of fuzzy memberships derived from data at different resolutions are studied
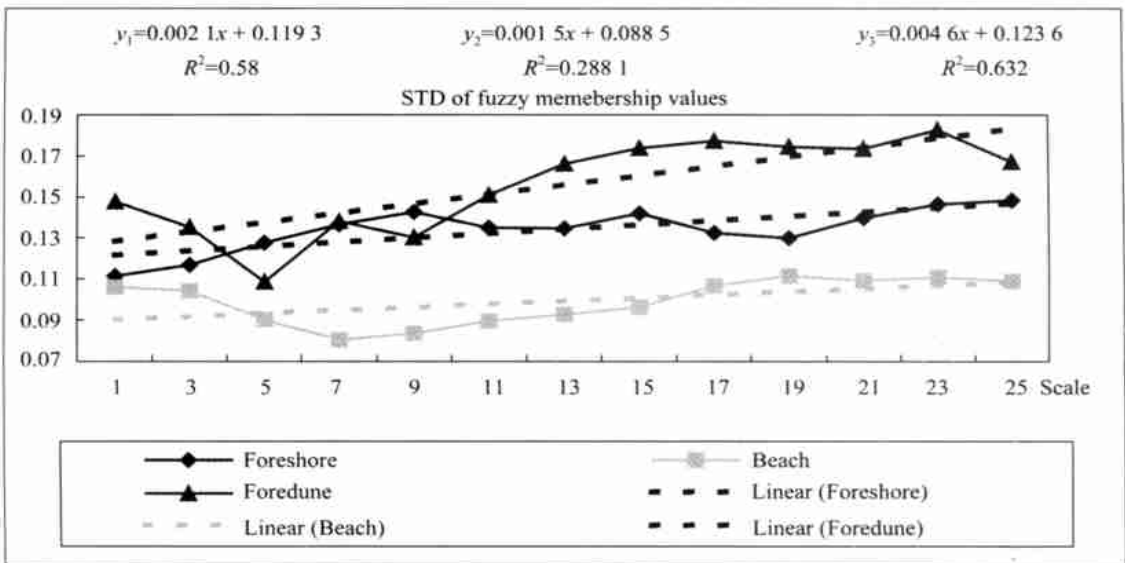
$y_1 = 0.002\,1x + 0.119\,3$ $y_2 = 0.001\,5x + 0.088\,5$ $y_3 = 0.004\,6x + 0.123\,6$
$R^2 = 0.58$ $R^2 = 0.288\,1$ $R^2 = 0.632$

STD of fuzzy memebership values

Fig. 5   Change in standard deviation of fuzziness of three landscape units

$y_1 = -0.000\,4\,x^4 + 0.083\,x^3 - 1.311\,6\,x^2 + 5.801\,9\,x + 133.8$ $y_3 = 0.025\,9\,x^4 - 0.714\,9x^3 + 5.712\,7\,x^2 - 8.623\,4\,x + 78.007$
$R^2 = 0.977\,4$ $R^2 = 0.981\,1$

$y_2 = -0.025\,5\,x^4 + 0.631\,9x^3 - 4.401\,2\,x^2 + 2.821\,5\,x + 868.19$
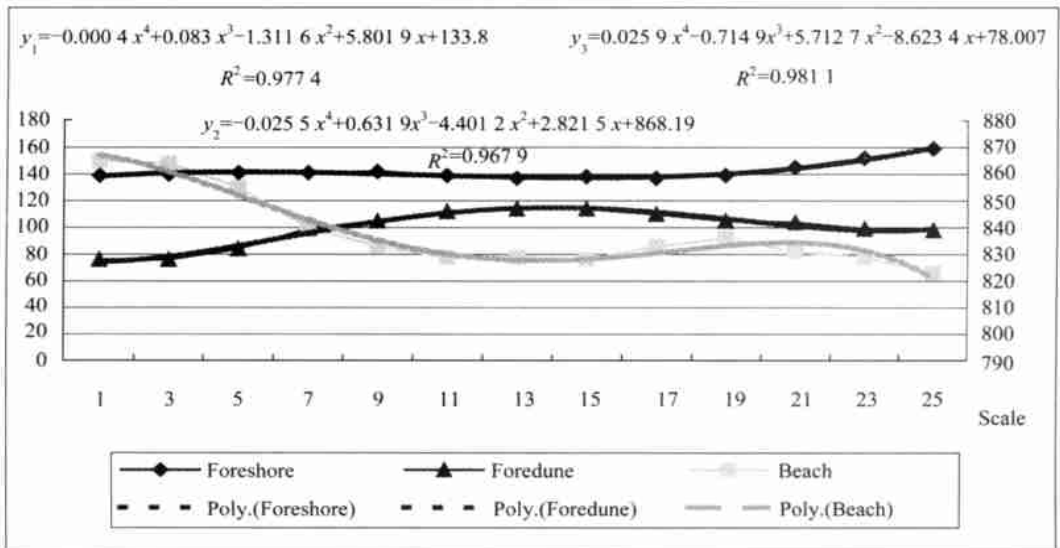$R^2 = 0.967\,9$

Fig 6   Forth-order polynomial trendlines

in order to examine these higher-order uncertain-ties. The statistics of the fuzziness of the fuzzy landscape units are calculated, and the variability of them with scale is assessed.

Results showed that it is difficult to accurately predict the effect of scale on fuzzy spatial objects, although the change with the scale exhibits linear relationship in some statistical indicators. In other words, the change of the fuzzy spatial objects with scale exhibits cyclic fluctuations, indicative of the periodic pattern in the landscape, which more suit-able to be polynomial than linear. In conclusion the identification of geomorphologic landscape units are

dependent upon the scale of the measurement, par-ticularly the area of the landscape units. A fine res-olution affords more detail in original data does not necessarily results in the landscape appearing to be more highly fragmented and complex than the same landscape examined with a coarser resolution. In view of this, caution must be exercised in compar-ing landscapes at different scales and in choosing the resolution of the data that best describes the process under study.

For the effect of scale on modeling of spatial data, it is still in its infancy stage. However, fieldwork data and satellite remote sensing data have been

more and more widely used without understanding the problems associated with the outcome. In this way， misleading decision may be made based on the uncertain modeling results. This topic has attracted increasing attention from GIS community. The results from this study will enable us to be aware of the level of uncertainty associated with the modeling outcome and thus make precautions if necessary. Further research using additional data of landscape and a greater range of resolution is necessary to determine whether general scaling laws be determined. Moreover, the effect of scale on dynamic processes should also be investigated.

REFERENCES

1   Cheng T, Molenaar M（1999）Objects with fuzzy spatial extent. *Photogrammetric Engineering and Remote Sensing*, 65: 797-801

2   Wood J（1996）The geomorphological characterisation of digital elevation models. http://www.soi.city.ac.uk/~jwo/phd

3   Foody G M（2002）Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80: 185-201

4   Wang F, Hall G B（1996）Fuzzy representation of geographical boundaries in GIS. *International Journal of Geographical Information Systems*, 10: 573-590

5   Li Z L, Openshaw S（1993）A natural principle for objective generalisation of digital map data. *Cartography and Geographic Information System*, 20: 19-29

6   Marceau D J（1999）The scale issue in the social and natural sciences. *Canadian Journal of Remote Sensing*, 25: 347-356

7   Wu J, Qi Y（2000）Dealing with scale in landscape analysis: an overview. *Geographic Information Sciences*, 6: 1-5

8   Cheng T（2002）Fuzzy spatial objects, their change and uncertainties. *Photogrammetric Engineering and Remote Sensing*, 68: 41-49

9   Ahkqvist O, Keukelaar J, Oukbir K（2000）Rough classification and accuracy assessment. *International Journal of Geographical Information Science*, 14: 475-496

10 Ahkqvist O, Keukelaar J, Oukbir K（2003）Rough and fuzzy geographical data integration. *International Journal of Geographical Information Science*, 17: 223-234

11 Zadeh L A（1965）Fuzzy sets. *Information and Control*, 8: 338-353

12 Usery E L（1996）A conceptual framework and fuzzy set implementation for geographic feature. In: Burrough P A, Frank A U, eds. *Geographic Objects with Indeterminate Boundaries*. London: Taylor & Francis. 71-85

13 Burrough P A, van Gaans P F M, MacMillan R A（2000）High-resolution landform classification using fuzzy k-means. *Fuzzy Sets and Systems*, 113: 37-52

14 Mackay D S, Samanta S, Ahl D E, et al.（2003）Automated parameterization of land surface process models using fuzzy logic. *Transactions in GIS*, 7: 139-153

15 Brassel K E, Weibel R（1988）A review and conceptual framework for automated map generalisation. *International Journal of Geographical Information Systems*, 2: 229-244

16 Wu J, Jelinski D E, Luck M（2000）Multiscale analysis of landscape heterogeneity: scale variance and pattern metrics. *Geographic Information Sciences*, 6: 6-19

17 Francis J M, Kopatek J M（2000）Multiscale effects of grain size on landscape pattern analysis. *Geographic Information Sciences*, 6: 27-37

18 Fotheringham A S, Wong D W S（1991）The modifiable areal unit problem in statistical analysis. *Environment and Planning*, 23: 1 025-1 044

19 Jelinski D E, Wu J（1996）The modifiable area unit problem and implications for landscape. *Landscape Ecology*, 11: 129-140

20 He H S, Venturta S J, Mladenoff D J（2002）Effects of spatial aggregation approaches on classified satellite imagery. *International Journal of Geographical Information Science*, 16: 93-109

21 Wood J（2003）Landserf: visualisation and analysis of terrain models. http://www.soi.city.ac.uk/~jwo/landserf/

22 Tate N, Atkinson P（2001）Modelling scale in geographical information science. Chichester: Jone Wiley & Sons.

23 Cheng T, Molenaar M, Bouloucos T（1997）Identification of fuzzy objects from field observation data. In: Hirtle S C, Frank A U, eds. *Spatial Information Theory: A Theoretical Basis for GIS*, *Lecture Notes in Computer Sciences*. Berlin: Spring-Verlag. 241-259

24 Hay G J, Blaschke T, Marceau D J, et al.（2003）A comparision of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry & Remote Sensing*, 57: 327-345

25 Robinson V B（2003）A perspective on the fundamentals of fuzzy sets and their use in geographic information systems. *Transactions in GIS*, 7: 3-30