Deciding the Way of Remote Sensing Image Wetland Classification of Information Based on the Vague Soft Set Similarity Measures-Topsis Method

Yuan-yuan MENG, Bo WEI, Yao ZOU and Zhi-liang WU

Guilin University of Technology, Guilin, China
*Corresponding author

Keywords: Remote sensing image, Wetland classification, Vague soft set similarity Measures-topsis method.

Abstract. As wetland is one of the richest ecosystems in bio-diversity and the most important survival environment of humans, quite a large amount of multi-temporal data need to be processed in the wetland dynamic monitoring based on satellite remote sensing, which brings five contradictory factors such as cost, data accuracy, technology, time-consumption and pollution. Using the Minjiang river estuary wetland as example, this paper proposes wetland remote sensing image classification method based on the theory of vague soft set similarity measures-Topsis decision-making. According to the proximity methods and the positive ideal solution (PIS) to select the most appropriate way, the entropy value method of vague weight value was put forward to every evaluation. Thus, the best multilevel utility model was established. Building the satellite data which is suitable for large quantities of long-term decision model about information extraction, it is of great significance to wetland long-term protection and sustainable application.

Introduction

On the one hand, in the wetland dynamic monitoring based on satellite remote sensing, wetland situation of historical continuous time sequence needs to be obtained; on the other hand, the satellite image data which has to be processed has multiple time periods, consuming long time, high cost and there are a lot of repetitive work in a wetland information constantly updated in the future. As the research object, the Minjiang river wetland of Fujian province, in which the dominating factors are the cost, data accuracy, technology, time-consumption and pollution. To apply the theory of vague soft set similarity measures - Topsis, and beginning study on optimize model of wetland remote sensing image information extraction. how to extract remote sensing information, and how to apply Vague soft set similarity measures - Topsis theory to realize industrialization of remote sensing study conducted to explore innovative.

General Situation of Test Area

Due to the structure, lithology and dynamic conditions are complex in the study area, which formed the complex geomorphic types. In order to facilitate the progress of the research work, this study only selected the most common, low-cost single source of data of US Landsat TM images for the study. The image’s forming time is June 6, 2009, high tide time of irrigation in paddy field.

According to image characteristics and needs of the research, the Minjiang river estuary will be divided into two categories, one is the study of wetland types, divided into five types of wetlands of Hohai, floodplain wetlands, lake, paddy fields, reservoirs and ponds; the other type two small class are vegetation and other land type.
Extracting the Way of Wetland Information Based on the Vague Soft Set Similarity Measures -Topsis Method

Classification Based on the Vague Soft Set Similarity Measures -Topsis Method

Remote sensed wetland information extracting procedure based on the Vague soft set similarity measures -Topsis theory, which is based on classification, adding cost, data accuracy, technology, time-consumption and pollution as degree indicators between the layers in classification. Based on the theory of Vague soft set similarity measures - Topsis multi-objective decision-making model of remote sensing information extraction controlling every layers of these five indicators, from simple to complex, achieving the best synthetic utility between the whole remote sensing information extraction process.

Construction of Classification Scheme Based on Vague soft set similarity measures - Topsis

Theoretical Optimum Utility Model

Definition 2.1 [1]. Let \( U = \{u_1, u_2, \ldots, u_n\} \) be an initial universe set. A vague set over \( U \) is characterized by a truth-membership function \( t_A \) and a false-membership function \( f_A \),

\[
t_A : U \rightarrow [0, 1], \quad f_A : U \rightarrow [0, 1],
\]

where for any \( u_i \in U \), \( t_A (u_i) \) is a lower bound on the grade of membership of \( u_i \) derived from the evidence for \( u_i \), \( f_A (u_i) \) is a lower bound on the negation of \( u_i \) derived from the evidence against \( u_i \), and \( 0 \leq t_A (u_i) + f_A (u_i) \leq 1 \). The grade of membership of \( u_i \) in the vague set is bounded to a subinterval \([t_A (u_i), 1 - f_A (u_i)]\) of \([0, 1]\).

Definition 2.2 [2]. Let \( U \) be an initial universe set, \( P(U) \) the power set of \( U \), \( E \) a set of parameters, and \( A \subseteq E \). A pair \((F, A)\) is called a soft set over \( U \), where \( F \) is a mapping given by \( F: A \rightarrow P(U) \).

A soft set \((F, E)\) describes the attractiveness of the cost, data accuracy, technology, time-consumption and pollution.

references [3] Vague soft set similarity measures, denoted by \( M \)

\[
M((G, E), (F, E)) = \sum_{i=1}^{m} \alpha_i \left[ 1 - \frac{1}{2n} \sum_{j=1}^{n} \left[ |t_{G(e_j)}(u_j) - t_{F(e_j)}(u_j)| + |f_{G(e_j)}(u_j) - f_{F(e_j)}(u_j)| \right] \right] \sum_{i=1}^{m} \alpha_i = 1. \tag{1}
\]

The greater the similarity measure with the PIS, the higher the effectiveness of the process; on the contrary, the lower the utility.

Theorem 2.1[4]. Let \( U=\{x_1,x_2,\ldots,x_n\} \) be the universal set of elements and \( E=\{e_1,e_2,\ldots,e_m\} \) be the universal set of parameters.

Hence \((F, E)=\{F(e_i)|i=1,2,\ldots,m\}\) is a family of vague soft sets. Define \( H(F, E) \) as Fangyan Peng [5]:

\[
H(F, E) = \frac{1}{n} \sum_{j=1}^{n} \left[ \pi_{F(e_j)}(x_j) + \frac{\sqrt{2}}{2} (1 - \pi_{F(e_j)}(x_j)) - \frac{\sqrt{2}}{2} f_{F(e_j)}(x_j) \right]. \tag{2}
\]

then \(H(F, E)\) is an entropy of vague soft sets.

The weight of each index, can be allocated according to the formula (2) [6]:

\[
\omega_i = \frac{1 - H_i}{\sum_{i=1}^{m} (1 - H_i)} \; i=1,2,\ldots,m. \tag{3}
\]

In the wetland remote sensing dynamic monitoring, similar level degree of attention are lay on the cost, data accuracy, technology, time-consumption and pollution. Relatively speaking, remote Sensing put emphasis on long-term and continuity of monitoring, cost and pollution appear most prominent, following by time, thus data accuracy and technology with the staff’s accumulation of experience and increase of technology will not be a big problem in the long-term monitor. Above the
study, comparing the five indicators refer to Table 2 and then rating into Vague sets by the formula (2) -(3) to calculate the weights shown below in Table 1.

<table>
<thead>
<tr>
<th>Index</th>
<th>technology</th>
<th>data accuracy</th>
<th>time-consumption</th>
<th>cost</th>
<th>pollution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vague Value</td>
<td>[0.60,0.80]</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
</tr>
<tr>
<td>Weight</td>
<td>0.16</td>
<td>0.16</td>
<td>0.20</td>
<td>0.24</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Here, as it shown in Table 2, division level were Vague value simple satisfaction of conversion, thereby constructing at different levels of classification, the Vague value of various classification schemes quantify list in multi-objective case.

<table>
<thead>
<tr>
<th>Satisfaction</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td>lower</td>
<td>low</td>
<td>Middle High</td>
<td>High</td>
<td>Higher</td>
<td>Absolute High</td>
</tr>
<tr>
<td>Vague Value</td>
<td>[0.20,0.30]</td>
<td>[0.30,0.45]</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.80,0.90]</td>
<td>[1.00,1.00]</td>
</tr>
<tr>
<td>Abstaining degree</td>
<td>0.10</td>
<td>0.15</td>
<td>0.15</td>
<td>0.20</td>
<td>0.10</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Optimization of Extracting Wetland Classification

Separation of Land and Water

As shown in Table 3, artificial interpretation due to the large map sheet, the time it takes too long, low integrated utility, the similar measure is 0.512. Unsupervised classification will leak to extract smaller rivers, somewhat the accuracy is a bit poor, but not high technical requirement, short time-consuming, low cost, which is a good program, similarity measure up to 0.760. Supervised classification Because of the image paddy field in the irrigation period, some part type of paddy are easily extracted into the waters, its overall effectiveness is not ideal, similarity measure is only 0.664.

<table>
<thead>
<tr>
<th>Separation of Land and Water</th>
<th>technology</th>
<th>data accuracy</th>
<th>time-consumption</th>
<th>cost</th>
<th>pollution</th>
<th>similarity measure(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial Interpretation</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
<td>[0.20,0.30]</td>
<td>[0.30,0.45]</td>
<td>[0.60,0.80]</td>
<td>0.512</td>
</tr>
<tr>
<td>Unsupervised classification</td>
<td>[0.80,0.90]</td>
<td>[0.70,0.85]</td>
<td>[0.80,0.90]</td>
<td>[0.70,0.85]</td>
<td>0.760</td>
<td></td>
</tr>
<tr>
<td>Supervised classification</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>0.664</td>
<td></td>
</tr>
<tr>
<td>Classification based on knowledge</td>
<td>[0.60,0.80]</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
<td>[0.70,0.85]</td>
<td>0.744</td>
<td></td>
</tr>
</tbody>
</table>

Applying remote sensing ENVI software, unsupervised classification method is used for image, results are obtained as shown in Figure 1.

Remote sensing classification based on knowledge whose technical difficulty is higher, and the similarity measure is 0.744. By comparison, max M = 0.760, so choosing unsupervised classification program.

Separation of Terrestrial Vegetation and Non-vegetation

As shown in Table 4, due to the large map sheet, artificial interpretation is more time-consuming and low utility. Because of the influence of the mountain shadow, to obtain the appropriate precision, unsupervised classification method will result in reducing the degree of technology and time, which similarity measure is 0.708. In that the underlying surface soil and shadow, when supervised classification attains a high classification accuracy, the satisfaction of technology and time will reduce, therefore its similarity measure is 0.664. Vegetation thematic information extraction method based on
the knowledge of commonly used which effect is good and the level of precision is very high with less difficult technology, its similarity measure is 0.760. By comparison, \( \max M = 0.760 \), as a result electing the classification scheme based on knowledge. Results are obtained as shown in Figure 2.

<table>
<thead>
<tr>
<th>Separation of Terrestrial Vegetation and Non-vegetation</th>
<th>technology</th>
<th>data accuracy</th>
<th>time-consumption</th>
<th>cost</th>
<th>pollution</th>
<th>similarity measure(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial Interpretation</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
<td>[0.20,0.30]</td>
<td>[0.30,0.45]</td>
<td>[0.60,0.80]</td>
<td>0.512</td>
</tr>
<tr>
<td>Unsupervised classification</td>
<td>[0.70,0.85]</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.80,0.90]</td>
<td>[0.70,0.85]</td>
<td>0.708</td>
</tr>
<tr>
<td>Supervised classification</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.60,0.80]</td>
<td>[0.70,0.85]</td>
<td>[0.70,0.85]</td>
<td>0.664</td>
</tr>
<tr>
<td>Classification based on knowledge</td>
<td>[0.70,0.85]</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
<td>[0.80,0.90]</td>
<td>[0.70,0.85]</td>
<td>0.760</td>
</tr>
</tbody>
</table>

The Extraction of Paddy Field

Paddy field area is too much big, wide distribution and some relatively fragmented, the amount of time is too long by artificial extraction, so not taking to consider. Unsupervised classification separated paddy precision not up to standard, also eliminated directly. Supervised classification of corresponding Vague set for every index are \((0.7,0.85),(0.7,0.85),(0.6,0.8),(0.7,0.85),(0.7,0.85))\), the similar measure is 0.680. Vegetation extraction based on knowledge classification method is conducive to the use of probing the underlying surface soil moisture in TM5 taking into account the TM image is in paddy field irrigation period, the band takes the gray value threshold from 16 to 81 extract paddy fields, the effect is good, its satisfaction indicators are \((0.7,0.85),(0.7,0.85),(0.8, 0.9),(0.8, 0.9),(0.7,0.85))\), the similar measure is 0.744. By comparison, \( \max M = 0.744 \), therefore selection is based on classification knowledge method.
Summary

Through analysing the field survey of the Minjiang river estuary wetland, the result indicates that the various classification methods are obtained satisfaction with the quantitative in cost, data accuracy, technology, time-consumption and pollution on the five indexes, which finally illustrated based on the theory of Vague soft set similarity measures-Topsis decision-making wetland information of remote sensing image classification method is more significant and more effective due to the data explanation. Unexpectedly, there are many uncertain and influence factors such as a diverse root of remote sensing information and a flexible dynamics of information extraction methods in the actual extraction work, but further investigations should be continued to screen better method of remote sensing image wetland classification of information.

Acknowledgements

This research was jointly supported by the National Natural Science Foundation of China (41461085), the Natural Science Foundation of Guangxi Province (2016GXNSFAA380035, 2014GXNSFDA118032), the Science and Technology Development Program of Guangxi Province (1598019-4), the "BaGui Scholars" Special Funds of Guangxi Province (2013-3), and the Doctoral Foundation of the Guilin University of Technology (GUT1996015).

References