

Article

Spatio-Temporal Analysis of Wetland Changes Using a Kernel Extreme Learning Machine Approach

Yi Lin ^{1,2} , Jie Yu ^{1,3,*}, Jianqing Cai ⁴, Nico Sneeuw ⁴ and Fengting Li ³

¹ College of Surveying, Mapping and Geo-information, Tongji University, Shanghai 200092, China; linyi@tongji.edu.cn

² Research Center of Remote Sensing & Spatial Information Technology, Shanghai 200092, China

³ College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China; fengting@tongji.edu.cn

⁴ Institute of Geodesy, University of Stuttgart, 70174 Stuttgart, Germany; cai@gis.uni-stuttgart.de (J.C.); nico.sneeuw@gis.uni-stuttgart.de (N.S.)

* Correspondence: 2011_jieyu@tongji.edu.cn; Tel.: +86-188-1736-6254

Received: 19 May 2018; Accepted: 11 July 2018; Published: 18 July 2018



Abstract: Natural wetland ecosystems provide not only important habitats for many wildlife species, but also food for migratory and resident animals. In Shanghai, the Chongming Dongtan International Wetland, located at the mouth of the Yangtze River, plays an important role in maintaining both ecosystem health and ecological security of the island. Meanwhile it provides an especially important stopover and overwintering site for migratory birds, being located in the middle of the East Asian-Australasian Flyway. However, with the increase in development intensity and human activities, this wetland suffers from increasing environmental pressure. On the other hand, biological succession in the mudflat wetland makes Chongming Dongtan a rapidly developing and rare ecosystem in the world. Therefore, studying the wetland spatio-temporal change is an important precondition for analyzing the relationship between wetland evolution processes and human activities. This paper presents a novel method for analyzing land-use/cover changes (LUCC) on Chongming Dongtan wetland using multispectral satellite images. Our method mainly takes advantages of a machine learning algorithm, named the Kernel Extreme Learning Machine (K-ELM), which is applied to distinguish between different objects and extract their information from images. In the K-ELM, the kernel trick makes it more stable and accurate. The comparison between K-ELM and three other conventional classification methods indicates that the proposed K-ELM has the highest overall accuracy, especially for distinguishing between *Spartina alterniflora*, *Scirpus mariqueter*, and *Phragmites australis*. Meanwhile, its efficiency is remarkable as well. Then a total of eight Landsat TM series images acquired from 1986 to 2013 were used for the LUCC analysis with K-ELM. According to the classification result, the change detection and spatio-temporal quantitative analysis were performed. The specific analysis of different objects are significant for learning about the historical changes to Chongming Dongtan and obtaining the evaluation rules. Generally, the rapid speed of Chongming Dongtan's urbanization brought about great influence with respect to natural resources and the environment. Integrating the results into the ecological analysis and ecological regional planning of Dongtan could provide a reliable scientific basis for rational planning, development, and the ecological balance and regional sustainability of the wetland area.

Keywords: Dongtan wetland; land-use/cover change (LUCC); kernel extreme learning machine (K-ELM); spatio-temporal change analysis; Landsat TM imagery

1. Introduction

Wetlands play an important role in ecosystem management, ranging from the local to the global scale [1]. They are defined as the boundary area of terrestrial and aquatic ecosystems, having features of both [2]. Wetlands generate a variety of benefits to society and nature, such as providing fertile soils for agriculture, food, and habitat for shorebirds, generating oxygen, adjusting climate, improving water quality, etc. They have been regarded as one of the most valuable resources in the world [3]. In recent decades, however, the extensive loss of wetlands has occurred in many countries throughout the world due to the high speed of economic development and population increase [4]. Therefore, the study of wetland conservation has generated considerable interest among researchers and achieved significant progress [5].

As a prelude to wetland conservation, the land-use/cover change (LUCC) of wetlands needs to be assessed to obtain thematic information. The conventional method applied to LUCC by ground survey is time-consuming and laborious, and is prone to errors for inaccessible areas as well. Satellite multispectral imagery, therefore, has recently been widely used in the characterization of wetland ecosystems, since it is able to cover large geographic areas with different spatial and spectral resolutions according to different requirements. Satellite multispectral imagery is mainly applied in five major implementations: (1) the survey and recognition of wetland resources [6,7]; (2) the generation of landscape-type thematic maps for wetlands [8,9]; (3) the dynamic change analysis of wetlands [10,11]; (4) the evaluation and inversion of biomass in wetland areas [12,13]; and (5) the evolution of wetland landscape patterns [14].

The quality of the remotely-sensed data and the ability of classifiers affects the dynamic spatio-temporal analysis directly. A remote sensing image with a higher spatial resolution, or a classifier with higher accuracy, leads to a more reasonable analysis which fits the reality better. However, using data with higher spatial resolution brings increasing cost. Thus, many researchers have attempted to generate a more accurate spatio-temporal analysis of wetlands by enhancing the classification accuracy. Based on previous studies, a wide array of classification approaches have been developed and applied for monitoring and analyzing the LUCC of wetland resources. Decision trees (DTs), due to their intuitive simplicity and computational efficiency, are also widely used. They can easily accommodate data from all measurements scales [15]. Many applications have been explored for various wetlands [16–18]. However, DT has the disadvantage of overfitting, leading to a result with low classification accuracy. Resampling statistical methods, like bagging and boosting, can avoid this issue and enhance accuracy, but they also make the interpretation of the results more complex [19]. Maximum likelihood classification (MLC), according to Ozesmi and Bauer's research [4], leads to higher accuracy than DT and is the most commonly used supervised classification algorithm. Moreover, a large amount of classification algorithms based on machine learning have been proposed recently and implemented successfully in various fields with higher accuracy, such as support vector machines (SVMs) [20] and artificial neural networks (ANNs) [21]. However, some of the machine learning approaches have high complexity and large computation. Thus, how to generate the LUCC with high accuracy and efficiency is still the main issue in spatio-temporal analysis of wetlands.

The main purpose of this study is to analyze the evolution of Chongming Dongtan with high accuracy and efficiency by remote sensing data, which is mainly on the basis of the classification results. Therefore, we took a first step to perform the LUCC for the Chongming Dongtan wetland using an optimized extreme learning machine (ELM) [22], namely the Kernel-ELM (K-ELM). Then, in order to demonstrate the performance and benefits of K-ELM, we conducted a comparative study of K-ELM, MLC, SVM, and conventional ELM with multispectral remotely-sensed data from 2009. The spatio-temporal dynamic changes were analyzed in terms of area change between different years. Finally, the experimental results allow a clear conclusion to be made about the dynamic change of Chongming Dongtan from 1986 to 2013. The spatio-temporal analysis of the Dongtan wetland will help researchers learn about its evolution more completely. It is, moreover, more intuitive to analyze the relationship between different objects. Generally, the novelty of our research is divided into two

parts: higher classification accuracy and efficiency generated by K-ELM for the Chongming Dongtan wetland, and a long periodic analysis of spatio-temporal changes of the Dongtan wetland. The results could provide a reliable scientific basis for urban planning, ecological resource management, and sustainable development.

2. Study Area and Data Used

The Dongtan wetland is at the mouth of the Yangtze River, located in the far east of Chongming Island, in Shanghai, China, between $121^{\circ}50'$ and $122^{\circ}05'$ east longitude, and ranging from $31^{\circ}25'$ to $31^{\circ}38'$ north latitude. Due to the influence of subtropical oceanic monsoons, it has a mild and moist climate. Since 1949, it has undergone several reclamations and seven dikes (shown in Figure 1) have been built to protect the bank and to promote silt. Nowadays, the wetland consists of freshwater, salty marshes (mainly dominated by *Spartina alterniflora*, *Phragmites australis*, and *Scirpus mariqueter*) [23], tidal creeks, and intertidal mudflats. Due to the unique climate and ecosystem it has been regarded as an important habitat for many wildlife species and a migratory stopover site for shorebirds along the East Asian–Australasian Flyway. However, with the increase in development intensity and human activities, the Dongtan wetland has to face many common ecological problems as well, such as sea level rise, inundation, alien species invasion, and fresh water shortage. In short, human interference exceeds the wetland's ability to maintain its biodiversity. On the other hand, biological succession in the mudflat wetland, behaving as the eastern beach of Chongming Island, extends further each year, which has made Chongming Island a rapidly developing and worldwide rare ecosystem.

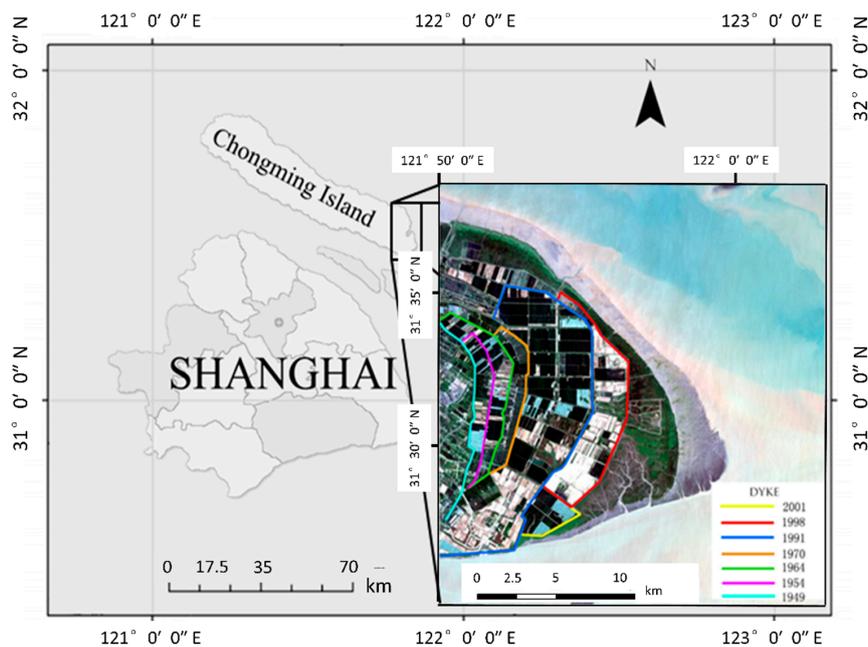


Figure 1. Study area: different line colors correspond to different years of dike construction.

Eight cloudless and terrain-corrected Landsat images of Chongming Island from 1986 to 2013 were utilized in this study (Table 1). For a number of reasons, growing season (July and August) images were selected. At this time of year, different kinds of vegetation are all at a stable stage. Meanwhile, it is easier to achieve cloud-free images due to the climatic conditions. The experimental data is comprised of six original spectral bands (band 1–5, 7 for TM and ETM+, band 2–7 for OLI) with 30 m spatial resolution.

A reference dataset was constructed based on the ground survey in 2009 which was used to train the classifier and evaluate the classification results. It contains 437 samples. Considering that the distribution of the bare flat is along the coastline, it is difficult and dangerous to obtain enough

ground survey points. We only obtained 32 samples for the bare flat, and another nine classes contain 45 samples, respectively.

Table 1. Acquisition dates of Landsat images.

Date	Sensor
19 August 1986	Landsat-5 TM
14 August 1990	Landsat-5 TM
14 August 1995	Landsat-5 TM
1 August 2000	Landsat-7 ETM+
2 August 2003	Landsat-5 TM
28 July 2007	Landsat-5 TM
17 July 2009	Landsat-5 TM
29 August 2013	Landsat-8 OLI

3. Methodology

3.1. Extreme Learning Machine

The extreme learning machine (ELM) is a simple learning method for single-hidden layer feed-forward networks (SLFNs), randomly assigning the input weights and hidden layer biases [24].

The output of a conventional SLFN with p hidden nodes, whose output nodes are linear, can be obtained by:

$$o_j = \sum_{i=1}^p \beta_i g(w_i x_j + b_i), \quad j = 1, 2, \dots, t \tag{1}$$

where $x_j = [x_{j1}, x_{j2}, \dots, x_{jm}]^T$ is the input vector with m nodes; $o_j = [o_{j1}, o_{j2}, \dots, o_{jn}]^T$ is the output vector with n nodes; $w_i = [w_{i1}, w_{i2}, \dots, w_{im}]$ is the weight vector connecting the input nodes and the i th hidden node; $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{in}]^T$ is the weight vector connecting the output nodes and the i th hidden node; b_i is the bias of the i th hidden node; and $g(x)$ is the activation function of hidden layer. Then Equation (1) can be written as:

$$\mathbf{O} = \mathbf{H}\beta \tag{2}$$

where \mathbf{H} is the hidden layer output matrix of the neural network, which can be expressed as [25]:

$$\mathbf{H} = \begin{bmatrix} g(w_1 x_1 + b_1) & \dots & g(w_n x_1 + b_n) \\ \vdots & \dots & \vdots \\ g(w_1 x_t + b_1) & \dots & g(w_n x_t + b_n) \end{bmatrix}_{t \times n} \tag{3}$$

According to Huang et al. [24], the input weights and hidden layer biases can be randomly assigned if only the activation function is infinitely differentiable. The least-squares solution of weight vector β is expressed as:

$$\beta = \mathbf{H}^+ \mathbf{O} \tag{4}$$

where \mathbf{H}^+ is the Moore-Penrose generalized inverse of \mathbf{H} [26]. \mathbf{H}^+ can be achieved, e.g., by orthogonal projection, iterative method, singular value decomposition. When $\mathbf{H}^T \mathbf{H}$ is nonsingular:

$$\mathbf{H}^+ = \left(\mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \tag{5}$$

Using ridge regression theory [27], the resultant solution is more stable and tends to have better generalization performance by adding a positive value to the diagonal of $\mathbf{H}^T \mathbf{H}$ [28]. Then the solution of β becomes:

$$\beta = \left(\mathbf{H}^T \mathbf{H} + \frac{\mathbf{I}}{c} \right)^{-1} \mathbf{H}^T \mathbf{O} \tag{6}$$

where \mathbf{I} is the identity matrix and c is the regularization parameter. In our study we set $c = 1$.

Hence the learning procedure of SLFNs is simplified as follows:

- Step 1 Define the training set, activation function and the number of hidden node.
- Step 2 Randomly generate the input weight and bias.
- Step 3 Calculate the hidden layer output matrix according to Equation (3).
- Step 4 Obtain the output weight based on Equation (6).

ELM is a faster learning algorithm, in which the hidden layer does not need to be tuned. It has been demonstrated that ELM has higher scalability and less computational complexity than the conventional feedforward network learning algorithms [28].

3.2. Kernel-Extreme Learning Machine

In this study an optimized ELM on the basis of kernel theory was used, namely the kernel-extreme learning machine (K-ELM) proposed by Yao et al. [29]. In the K-ELM, the activation function is replaced by a 'kernel trick' [30]. Theoretically, since K-ELM combines a kernel trick with ELM, it gives a better general performance and less computational complexity than the conventional popular learning algorithms for feedforward neural networks. Based on the kernel theory the kernel function for a SLFN [29] can be defined as:

$$\mathbf{K}(x_i, x_j) = \left[g(w_1 x_i + b_1) \quad \dots \quad g(w_n x_i + b_n) \right] \cdot \left[g(w_1 x_j + b_1) \quad \dots \quad g(w_n x_j + b_n) \right]^T \quad (7)$$

Solution (6) can now be rewritten as:

$$\beta = \left(\begin{bmatrix} \mathbf{K}(x_1, x_1) & \dots & \mathbf{K}(x_1, x_t) \\ \vdots & \dots & \vdots \\ \mathbf{K}(x_t, x_1) & \dots & \mathbf{K}(x_t, x_t) \end{bmatrix} + \frac{\mathbf{I}}{c} \right)^{-1} \begin{bmatrix} \mathbf{K}(x_1, x) \\ \vdots \\ \mathbf{K}(x_t, x) \end{bmatrix} \mathbf{O} \quad (8)$$

In K-ELM we only need to define the type of kernel function instead of the activation function and node number of the hidden layer, which simplifies the random feature mapping and makes the machine much more stable. In our experiment, a third-order polynomial function was used as the kernel function.

3.3. Support Vector Machine

Support vector machines (SVMs) are based on the statistical learning theory developed by Vapnik [20], aiming to find a separating hyperplane with the maximum geometric margin between classes. SVMs have been widely introduced for solving classification, regression, and other learning problems. The v -support vector classification (v -SVM) [31] was selected to recognize the different objects of Dongtan wetland in this study. The primal optimization problem can be represented by:

$$\begin{aligned} \min_{Q, b, \zeta, \rho} & \frac{1}{2} \mathbf{Q}^T \mathbf{Q} - v\rho + \frac{1}{l} \sum_{i=1}^l \zeta_i, \\ \text{subject to} & y_i (\mathbf{Q}^T \varnothing(x_i) + c) \geq \rho - \zeta_i, \\ & \zeta_i \geq 0, i = 1, \dots, l, \rho \geq 0. \end{aligned} \quad (9)$$

where x_i are the training vectors; $\varnothing(x_i)$ maps x_i into a higher-dimensional space by some kernel tricks; $v \in (0, 1]$ is a parameter giving an upper bound on the fraction of training errors and a lower bound of the fraction of support vector respectively [32]; y_i is the class label; ζ_i are slack variables; ρ is an additional variable; \mathbf{Q} and c are the components of the separating hyperplane:

$$(\mathbf{Q} \cdot \mathbf{X}) + c = 0, \mathbf{Q} \in \mathbb{R}^N, c \in \mathbf{R} \quad (10)$$

where \mathbf{X} is a matrix of reflectance of pure endmembers.

This optimization problem can be solved by obtaining its Lagrangian and transforming it to its dual problem [31].

3.4. Maximum Likelihood Classification

Maximum likelihood classification is one of the most commonly used parametric supervised classifiers [33]. This method distinguishes objects by computing the weighted distance or likelihood based on the Bayesian theory [34]. The discriminant function $L(x_i)$ for each pixel x_i is:

$$L(x_i) = \ln P_j - \ln |\Sigma_j| / 2 - (x_i - m_j)^T \Sigma_j^{-1} (x_i - m_j) / 2 \quad (11)$$

where P_j , Σ_j , and m_j are the probability, covariance, and mean value of j th class, respectively.

When conducting a hard classification objects would belong to the class in which it has the highest probability.

4. Experiment

Figure 2 presents a flowchart of the main process in this study. Aiming to validate the effectiveness and efficiency of K-ELM, a comprehensive comparison between K-ELM and three other conventional methods was made with the image in 2009. In the comparison experiment the classification results were evaluated by three accuracy indices (overall, producer's, and user's accuracy) and computation time. Then the spatio-temporal analysis of the Dongtan wetland from 1986 to 2013 was generated with the classification results of the K-ELM.

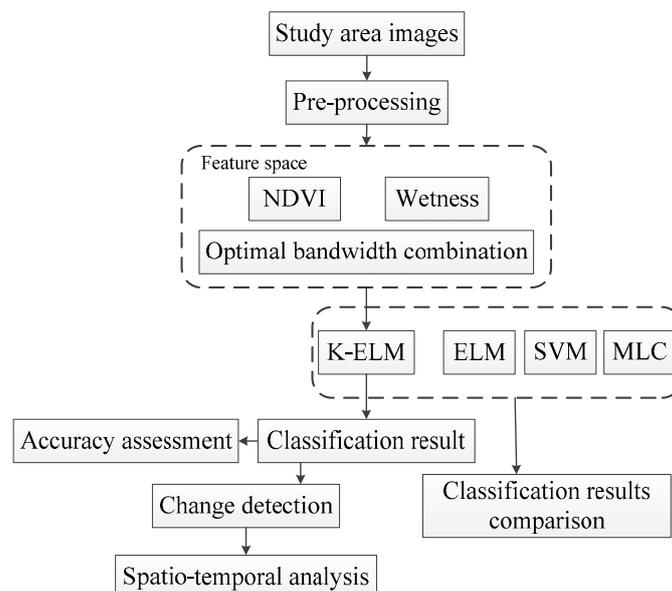


Figure 2. The main process in this study.

4.1. Pre-Processing

In order to eliminate the atmospheric influence on radiance in the remote sensing imagery and to achieve the real surface reflectance, the atmospheric interference was removed by FLAASH [35] in this study. Ten ground control points were selected to make the geometric correction. The root mean square errors of the images were all less than one pixel.

4.2. Band Analysis and Feature Space Construction

The feature space consists of one optimal bandwidth combination, normalized difference vegetation index (NDVI), and the wetness dimension of the Kanth–Thomas (K-T) transformation.

The optimum index factor (OIF) [36], which weighs the variance of each individual band, was utilized to select the optimal bandwidth combination. According to the results of the OIF, bands 3–5 (red, near-infrared, and short-wave infrared) for TM/ETM+ data and bands 3, 5, and 6 (green, near-infrared, and short-wave infrared) for OLI data were chosen.

4.3. Classifier Training and Accuracy Evaluation

According to the ground survey in 2009, to the west of the dike, the classification system consists of green land, building land, farmland, water area, reclaimed land, and bare land, while in the east, it is comprised of *Spartina alterniflora*, *Phragmites australis*, *Scirpus mariqueter*, and bare flat. Therefore, there are ten totally different classes in our study area.

For the image of 2009, since a total of 437 reference samples were obtained based on the ground survey, the training and testing samples were selected from the reference dataset. Then a three-fold cross-validation was applied for the classification of accuracy evaluation. Thus, 292 training samples, which were separately and randomly selected from each class sample in proportion, were utilized to train the classifier. In addition, the remaining 145 samples were regarded as testing data. For images of other years the training samples were defined by a priori knowledge. The final selected training samples for each class have separability greater than 1.9 in ENVI software (Harris Geospatial Solutions, Broomfield, CO, USA) (Deng 2010). The number of training samples for each class in other years is equal to that in the 2009 image.

In this study we applied a K-ELM with a third-order polynomial kernel function to distinguish between the objects. The node number in the input layer and output layer are equal to the dimension of the feature space and the number of classes, respectively. To examine the success and effectiveness of the K-ELM a comparison between K-ELM and other three methods: MLC, SVM, and standard ELM were generated by the 2009 image. The assessment is divided into two parts: accuracy and efficiency. The classification accuracy was based on the three-fold cross-validation result and evaluated by the average values of overall accuracy, producer's accuracy, and user's accuracy. On the other hand, computation time was recorded and applied for a comparison of the efficiency.

4.4. Land-Use Cover/Change

Based on the classification results, a quantitative change detection was generated by calculating the area difference between images. An index called LUCC dynamic degree was developed to measure the rate of changes, which led to a better description of the quantitative change of area year by year. The expression is given by:

$$D = \frac{U_2 - U_1}{U_1} \times \frac{1}{T_2 - T_1} \times 100\% \quad (12)$$

where D is the LUCC dynamic degree; U_1 and U_2 are the area of a given class in the year of T_1 and T_2 , respectively. After achieving each year's LUCC dynamic degree, we analyzed the dynamic spatio-temporal change.

5. Results

5.1. Classification Result and Evaluation

To examine the performance and benefits of the K-ELM method the classification accuracy and efficiency were evaluated by comparing them with the other three commonly utilized methods: MLC, ELM, and SVM. Figure 3 shows the classification results in 2009 of different methods. Different colors indicate different objects. The result of the K-ELM, according to ground survey data, is more consistent with the actual distribution. Especially for the east part (marked by a black rectangle), which is mainly

occupied by *Phragmites australis*, *Scirpus mariqueter*, and bare flat in practice, K-ELM could distinguish them significantly, while MLC and SVM misclassify some areas as grassland and *Spartina alterniflora*, respectively. The ELM could not effectively recognize *Scirpus mariqueter* as well.

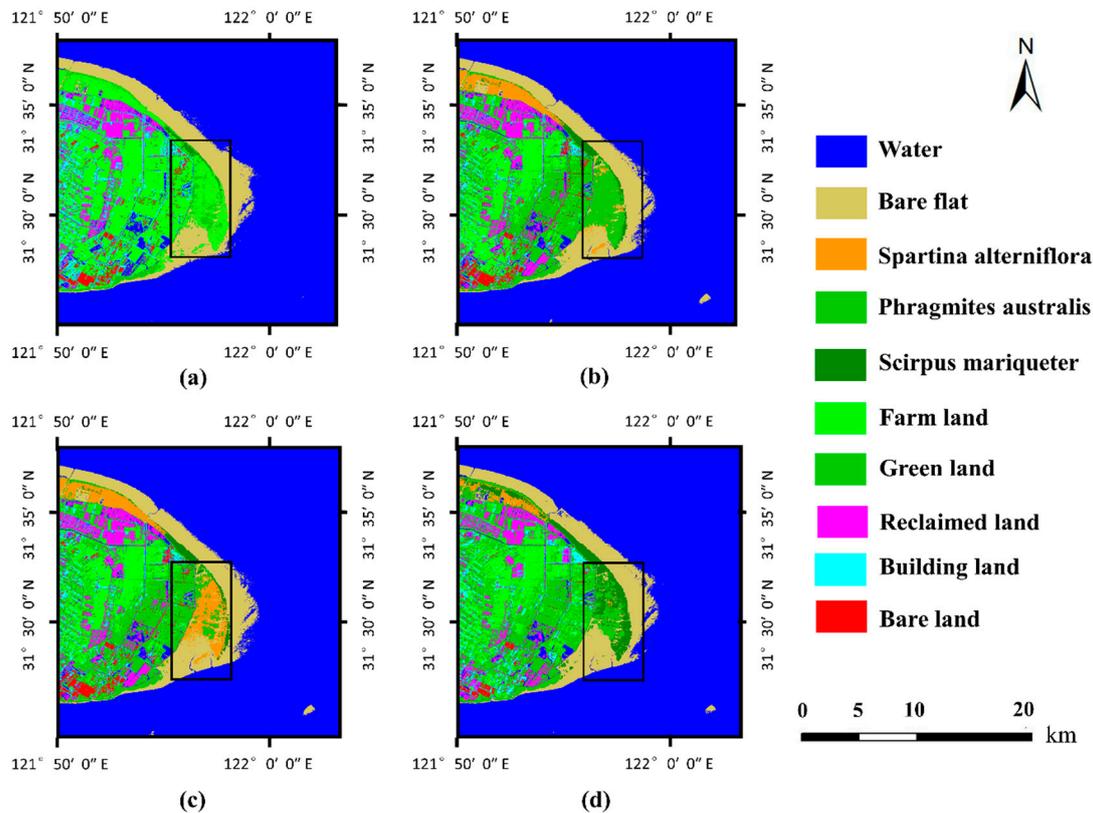


Figure 3. Classification results in 2009 of the four methods: (a) maximum likelihood classification (MLC); (b) extreme learning machine (ELM); (c) support vector machine (SVM); and (d) kernel extreme learning machine (K-ELM).

Then, based on the classification results, the accuracy of different methods was assessed with the producer's and user's accuracy for each class (given in Table 2) and the overall accuracy for the entire research area (given in Table 3). Meanwhile, the computation time was calculated to assess different methods' efficiencies, as shown in Table 3, as well. It can be seen that the K-ELM performs best with the highest overall accuracy (86.07%), followed by the SVM (85.35%). The conventional ELM, whose accuracy is 82.59%, performs worse than the SVM, while the MLC has the lowest accuracy, which is 79.31%. However, the classification ability of four classifiers varies slightly with classes. K-ELM generates both the highest producer's and user's accuracy for *Spartina alterniflora*, *Scirpus mariqueter*, *Phragmites australis*, reclamation land, and water, while SVM and ELM work better for farmland and green land. MLC is more suitable for building land and bare land recognition. As for the efficiency, SVM is the most time consuming, with a computation time more than twice as long as that of ELM. The computation time of K-ELM is 1.38 s, which is much less than MLC and SVM.

Generally, compared with the conventional ELM, the 'kernel trick' increases the classification accuracy of K-ELM significantly, especially for the area mixed by *Spartina alterniflora*, *Scirpus mariqueter*, and *Phragmites australis*. Even though there is, to some extent, a limitation to the efficiency of the 'kernel trick' its efficiency is still remarkable and much higher than that of SVM and MLC. Comprehensively, the K-ELM is advantageous when applied to the LUCC of the Dongtan wetland on Chongming Island.

Table 2. Producer's and user's accuracy of the four methods.

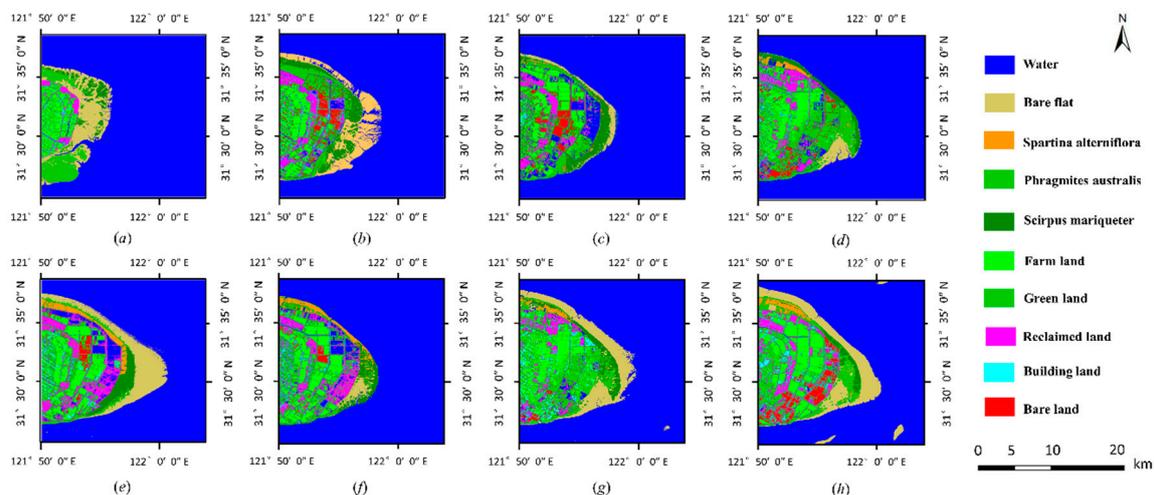
Class	Accuracy Index							
	Producer's Accuracy (%)				User's Accuracy (%)			
	MLC	ELM	SVM	K-ELM	MLC	ELM	SVM	K-ELM
Farm land	94.67	78.81	88.00	83.70	52.94	88.52	90.00	83.58
Building land	85.04	74.52	78.07	70.67	87.50	78.72	80.83	73.38
Reclamation land	87.85	93.93	92.74	95.85	90.40	92.02	91.79	92.56
Water	98.22	97.33	94.37	98.52	78.10	73.97	84.48	86.01
Green land	73.63	86.67	87.26	84.74	93.60	82.05	87.13	81.71
Bare land	84.89	78.96	91.78	72.74	85.01	77.13	78.37	75.54
Bare flat	95.33	90.44	81.04	94.89	91.08	94.21	96.72	93.23
<i>Scirpus mariqueter</i>	56.00	63.56	70.96	86.52	78.10	73.97	84.48	86.01
<i>Phragmites australis</i>	72.74	77.33	84.15	87.26	72.63	70.92	87.52	87.65
<i>Spartina alterniflora</i>	50.07	86.96	87.26	88.74	84.71	77.44	71.14	88.61

Table 3. Overall accuracy and efficiency assessment of the four methods.

Methods	Accuracy Index	Efficiency
	Overall Accuracy (%)	Time (s)
MLC	79.31	2.01
ELM	82.59	1.06
SVM	85.35	2.23
K-ELM	86.07	1.38

5.2. Dynamic Spatio-Temporal Analysis

Eight images, from 1986 to 2013, were classified by the K-ELM, shown in Figure 4, after which the change detection and spatio-temporal analysis were performed.

**Figure 4.** Classification results of K-ELM: (a) 1986; (b) 1990; (c) 1995; (d) 2000; (e) 2003; (f) 2007; and (g) 2009; (h) 2013.

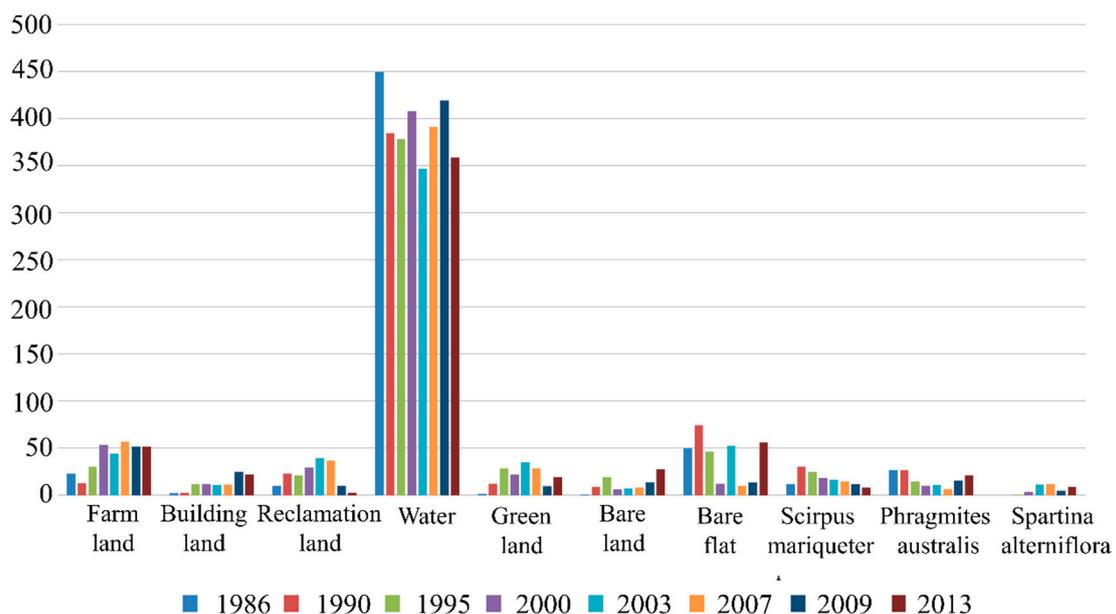
The area statistics of the ten classes and the corresponding LUCC dynamic degree are given in Tables 4 and 5, respectively. In order to make the results more intuitive, the corresponding bar graph of area statistics is drawn and shown in Figure 5. For each class the farm land increased quickly during 1990–1995 and 1995–2000, with growth rates of 27.41% and 15.73%, respectively. Its changing trend has a strong relationship with the two dikes built in 1991 and 1998. The average annual growth is 4.07 km², whereas the total area changed only slightly after the dike was built in 1998.

Table 4. Area statistics of ten classes (*Spartina alterniflora* was introduced into Chongming Island after 1995).

Classes	Area (km ²)							
	1986	1990	1995	2000	2003	2007	2009	2013
Farm land	22.94	12.60	29.87	53.37	43.98	56.46	51.29	51.31
Building land	1.84	2.34	11.83	11.65	10.91	11.16	24.89	21.81
Reclamation land	9.65	22.60	20.97	29.08	39.09	36.47	10.05	2.47
Water	449.3	384.12	378.13	407.6	346.82	390.91	419.17	358.4
Green land	1.08	12.10	28.21	21.89	34.84	28.23	9.27	19.06
Bare land	0.51	8.46	18.72	6.28	7.30	7.69	13.65	27.17
Bare flat	49.89	74.16	46.17	12.07	52.02	9.59	13.18	55.66
<i>Scirpus mariqueter</i>	11.39	30.03	24.33	18.02	16.38	14.28	11.90	7.66
<i>Phragmites australis</i>	26.38	26.57	14.60	9.57	10.50	6.35	14.95	21.07
<i>Spartina alterniflora</i>	0.00	0.00	0.15	3.45	11.14	11.84	4.63	8.37
Total	572.98							

Table 5. Land-use/cover changes (LUCC) dynamic degree of different objects from 1986 to 2013 (*Spartina alterniflora* was introduced into Chongming Island after 1995).

Classes	LUCC Dynamic Degree (%)						
	1986–1990	1990–1995	1995–2000	2000–2003	2003–2007	2007–2009	2009–2013
Farm land	−11.27	27.41	15.73	−5.86	7.09	−4.58	0.01
Building land	6.79	81.11	−0.30	−2.12	0.57	61.51	−3.09
Reclamation land	33.55	−1.44	7.73	11.47	−1.68	−36.22	−18.86
Water	−3.63	−0.31	1.56	−4.97	3.18	3.61	−3.62
Green land	255.09	26.63	−4.48	19.72	−4.74	−33.58	26.40
Bare land	389.71	24.26	−13.29	5.41	1.34	38.75	24.76
Bare flat	12.16	−7.55	−14.77	110.33	−20.39	18.72	80.58
<i>Scirpus mariqueter</i>	40.91	−3.80	−5.19	−3.03	−3.21	−8.33	−8.91
<i>Phragmites australis</i>	0.18	−9.01	−6.89	3.24	−9.88	67.72	10.23
<i>Spartina alterniflora</i>	0.00	0.00	440.00	74.30	1.57	−30.45	20.19

**Figure 5.** The bar graph of area statistics for different objects from 1986 to 2013.

The increase in building land reached its peak between 2007 and 2009, with 13.73 km², due to the development of the tourist industry indicating the quick economic development and increasing population on Chongming Island during that period.

The reclamation tended to increase when a new dike was built. After 2003 it began to degrade. It fell at the highest rate, −36.23%, from 2007 to 2009.

The water area is comprised of an east part (outside the dike) and a west part (inside the dike). The east water area depends on the sea tide, while the west part, which mainly consists of lakes and rivers, is strongly affected by rainfall. It generally has a stable trend.

The change in green land is greatly influenced by an increase in farmland. Before 2003 the green land area kept increasing, but after that it had an opposite trend. By analyzing the classification results in different years, we found that the decreased area of green land transferred to farmland.

The three kinds of vegetation (*Spartina alterniflora*, *Phragmites australis*, *Scirpus mariqueter*) differed distinctly from each other. *Spartina alterniflora*, after being introduced into Chongming Island in the middle of the 1990s, has an obvious trend of southward expansion, while dominating the growth area of *Phragmites australis*. This is consistent with the area statistics of *Spartina alterniflora* and *Phragmites australis*. *Scirpus mariqueter* is located in the east of the wetland. Therefore, its area is greatly affected by the tide and cannot be extracted accurately.

The bare flat extends eastward and its expansion in a northerly direction is larger than that towards the south. Though it did not change much from 2003 to 2007, its total area still kept increasing.

To make a more visual change analysis result the change detection in different periods was generated. Since we focused our study on the change of the wetland, which is separated by dikes and consisted of *Spartina alterniflora*, *Phragmites australis*, *Scirpus mariqueter*, and bare flat, the change detection only contained these four objects, as shown in Figure 6, where yellow indicates the increasing area of the wetland and blue shows the decreasing area.

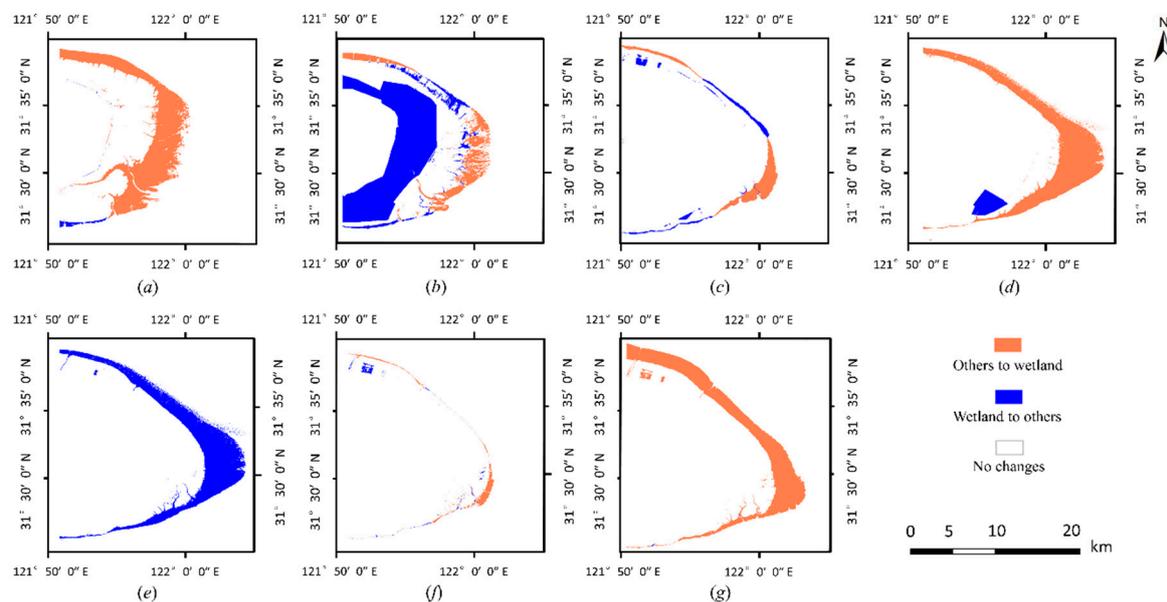


Figure 6. Change detection result of wetland from 1986 to 2013: (a) 1986–1990; (b) 1990–1995; (c) 1995–2000; (d) 2000–2003; (e) 2003–2007; (f) 2007–2009; and (g) 2009–2013.

It can be seen that the wetland extended quickly towards the east and its area kept increasing, generally, except for the period between 2003 and 2007. After comparing the classification results in 2003 and 2007 it was easy to ascertain that the large area of the wetland that disappeared was covered by sea water. This might have been caused by tidal influence. During two other periods, 1990 to 1995 and 2000 to 2003, the main reason which led to the decrease in wetland area was the fact that two

new dikes were finished in 1991 and 2001, respectively. Therefore, the region of wetland would be redefined at a time when a new dike was built and the part to the west of the new dike was not part of the wetland anymore. Dikes built at different times are shown in Figure 1. Overall, the change in wetland has a strong relationship to dike building.

6. Discussion

Compared with the conventional ANNs, the K-ELM combines ELM theory with a kernel trick. The users, therefore, only need to select a kernel function when applying it, instead of defining the active function, the number of hidden nodes, and input weights. The nonlinear mapping of an ELM is replaced by the kernel function, which gives the original inputs map a higher dimension space. This enables the K-ELM to be a more stable and accurate classifier. It has advantages for LUCC, especially for a region that contains similar and complex kinds of objects.

The classification results reveal the distribution characteristics of objects and provide their corresponding area size. Since the spatio-temporal analysis produced is based on classification results, and the accuracy of the classification results has a significant effect on the change detection.

With the rapid development of Chongming Dongtan the building land (west of the dike) grew quickly. Meanwhile, there was a sustained transformation between green land and farm land, which increased the landscape diversity. During the periods of new dike building, the area of reclamation would increase as well. Generally, human activities to the west of the dike have a great influence on the ecosystem. In the wetland ecosystem the bare flat decreased rapidly, since a large amount of vegetation was cultivated while it held the highest proportion before 1995. Additionally, the area of bare flat is also affected by the tide. For instance, the rapid increase in the bare flat in 2003 was mainly due to the low tide. Then, after *Spartina alterniflora* was introduced into Dongtan in 1995, the existence of local vegetation (*Phragmites australis* and *Scirpus mariqueter*) became seriously threatened. Thus, some measures were taken to manage the invasion of *Spartina alterniflora*. Meanwhile, the Chongming Dongtan Nature Reserve was established in 1998 and work began on it one year later, which was beneficial for improving the health of Chongming Dongtan. All the efforts are significant in keeping the ecological balance and raising the overall health level.

The spatio-temporal analysis, therefore, is advantageous for learning about the historical changes to the Chongming Dongtan and obtaining the evolution rules. To some extent, it could provide useful information for constructing and managing Chongming Dongtan.

7. Conclusions

In this study a novel method, which mainly takes advantages of the kernel extreme learning machine (K-ELM) algorithm, was applied to distinguish between different objects and extract their distribution information over the Dongtan wetland located on Chongming Island, Shanghai, China. In the K-ELM the active function is replaced by a kernel function, implying the random feature mapping of active functions and input connection weights are not necessary. Thus, the K-ELM is more stable and effective than a conventional ELM. On the other hand, the experimental results demonstrate that the K-ELM, whose overall accuracy is 86.07%, has a higher classification accuracy than the other three classification methods (MLC, SVM, and ELM).

Additionally, based on the classification results, another novelty of this study is a dynamic spatio-temporal analysis of the Dongtan wetland made from 1986 to 2013. The statistical information of the ten classes in the Dongtan wetland, in terms of area, LUCC dynamic degree, and change detection visualization, was analyzed specifically. Generally, the dike building has a great influence on changes in the wetland region. These quantitative results did not only help in assessing the subsequent ecosystems and supporting their management effectively, but it also led to a more reliable and scientific decision regarding the harmonious development of the ecological environment, regional management, urban planning, etc.

In the long-term, dynamic monitoring and analysis can play a very important role in the Dongtan wetland's protection and rational use. It could provide a reliable scientific basis for rational planning, development, ecological balance, and regional sustainability of the wetland area. According to the results of the remote sensing monitoring and spatio-temporal quantitative analysis the government could obtain a correct evaluation in order to execute ecological regional planning for the Dongtan zone.

Despite its success our study leads to some issues that need to be improved in future work. Firstly, as the spatial resolution of TM images is 30 m, there is a large number of mixed pixels in the image which can cause classification errors and cannot be reclassified, except when adding extra information. Therefore, some high spatial resolution satellite images, such as WorldView-2 or QuickBird, should be utilized. In addition, since the training dataset was not very large in our experiment, we did not consider the rare class issue which is, however, very important for classifier training. Finally, this study is mainly based on remotely-sensed images. More multi-source data, like environmental survey data and economics data, should be applied jointly to improve the assessment.

In addition, a monitoring network system should be established for future studies, which involves monitoring environmental qualities including land use, water quality, atmospheric environment, acoustic environment, and solid quality. It would, furthermore, involve aquatic and wetland ecosystems in order to realize comprehensive monitoring and evaluation.

Author Contributions: Y.L. conceived and designed the experiments and wrote the paper; J.Y. performed the experiments and analyzed the experimental results; J.C., N.S. and F.L. provided theoretical advises and contributed to the manuscript preparation.

Funding: This research were funded by the National Natural Science Foundation (NSFC) Project (No. 41771449) and the DAAD Thematic Network Project (No. 57173947).

Acknowledgments: The authors would like to thank the editors and three anonymous reviewers for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Henderson, F.M.; Lewis, A.J. Radar detection of wetland ecosystems: A review. *Int. J. Remote Sens.* **2008**, *29*, 5809–5835. [[CrossRef](#)]
- Keddy, P.A. *Wetland Ecology: Principles and Conservation*; Cambridge University Press: Cambridge, UK, 2010.
- Barbier, E.B.; Hacker, S.D.; Kennedy, C.; Koch, E.W.; Stier, A.C.; Silliman, B.R. The value of estuarine and coastal ecosystem services. *Ecol. Monogr.* **2011**, *81*, 169–193. [[CrossRef](#)]
- Ozesmi, S.L.; Bauer, M.E. Satellite remote sensing of wetlands. *Wetl. Ecol. Manag.* **2002**, *10*, 381–402. [[CrossRef](#)]
- Weinstein, M.P.; Kreeger, D.A. *Concepts and Controversies in Tidal Marsh Ecology*; Springer Science & Business Media: Berlin, Germany, 2000.
- La Cecilia, D.; Toffolon, M.; Woodcock, C.E.; Fagherazzi, S. Interactions between river stage and wetland vegetation detected with a Seasonality Index derived from LANDSAT images in the Apalachicola delta, Florida. *Adv. Water Resour.* **2016**, *89*, 10–23. [[CrossRef](#)]
- Torbick, N.; Salas, W. Mapping agricultural wetlands in the Sacramento Valley, USA with satellite remote sensing. *Wetl. Ecol. Manag.* **2015**, *23*, 79–94. [[CrossRef](#)]
- Crombé, P.; Verhegge, J.; Deforce, K.; Meylemans, E.; Robinson, E. Wetland landscape dynamics, Swifterbant land use systems, and the Mesolithic–Neolithic transition in the southern North Sea basin. *Quat. Int.* **2015**, *378*, 119–133. [[CrossRef](#)]
- Lu, X.; Zhang, J.; Hong, J.; Wang, L. Analysis of wetland landscape evaluation and its driving factors in Yellow River Delta based on remote sensing image. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 214–223.
- Ghobadi, Y.; Pradhan, B.; Shafri, H.Z.M.; Bin Ahmad, N.; Kabiri, K. Spatio-temporal remotely sensed data for analysis of the shrinkage and shifting in the Al Hawizeh wetland. *Environ. Monit. Assess.* **2015**, *187*, 4156. [[CrossRef](#)] [[PubMed](#)]
- Sivakumar, R.; Ghosh, S. Wetland spatial dynamics and mitigation study: An integrated remote sensing and GIS approach. *Nat. Hazards* **2016**, *80*, 975–995. [[CrossRef](#)]

12. Aslan, A.; Rahman, A.F.; Warren, M.W.; Robeson, S.M. Mapping spatial distribution and biomass of coastal wetland vegetation in Indonesian Papua by combining active and passive remotely sensed data. *Remote Sens. Environ.* **2016**, *183*, 65–81. [[CrossRef](#)]
13. Moulin, S.; Kergoat, L.; Viovy, N.; Dedieu, G. Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements. *J. Clim.* **1997**, *10*, 1154–1170. [[CrossRef](#)]
14. Li, S.; Wang, G.; Deng, W.; Hu, Y.; Hu, W.-W. Influence of hydrology process on wetland landscape pattern: A case study in the Yellow River Delta. *Ecol. Eng.* **2009**, *35*, 1719–1726. [[CrossRef](#)]
15. Wright, C.; Gallant, A. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sens. Environ.* **2007**, *107*, 582–605. [[CrossRef](#)]
16. Davranche, A.; Lefebvre, G.; Poulin, B. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sens. Environ.* **2010**, *114*, 552–562. [[CrossRef](#)]
17. Yuan, H.; Zhang, R.; Li, X. Classification of Wetland from TM imageries based on Decision Tree. *WSEAS Trans. Inf. Sci. Appl.* **2009**, *6*, 1155–1164.
18. Tian, B.; Zhou, Y.; Thom, R.M.; Diefenderfer, H.L.; Yuan, Q. Detecting wetland changes in Shanghai, China using FORMOSAT and Landsat TM imagery. *J. Hydrol.* **2015**, *529*, 1–10. [[CrossRef](#)]
19. Lawrence, R.; Bunn, A.; Powell, S.; Zambon, M. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sens. Environ.* **2004**, *90*, 331–336. [[CrossRef](#)]
20. Vapnik, V.; Golowich, S.E.; Smola, A. Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing. *Adv. Neural Inf. Process. Syst.* **1996**, *9*, 281–287.
21. Civco, D.L. Artificial neural networks for land-cover classification and mapping. *Int. J. Geogr. Inf. Syst.* **1993**, *7*, 173–186. [[CrossRef](#)]
22. Huang, G.; Ding, X.; Zhou, H. Optimization method based extreme learning machine for classification. *Neurocomputing* **2010**, *74*, 155–163. [[CrossRef](#)]
23. Yu, J.; Lin, Y.; Hu, C.; Zhang, Y. The spatio-temporal dynamic analysis of salt marsh vegetation in Chongming Dongtan based on remote sensing data. In Proceedings of the Third International Workshop on Earth Observation and Remote Sensing Applications (EORSA), Changsha, China, 11–14 June 2014; pp. 52–56.
24. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [[CrossRef](#)]
25. Huang, G.B. Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Trans. Neural Netw.* **2003**, *14*, 274–281. [[CrossRef](#)] [[PubMed](#)]
26. Rao, C.R.; Mitra, S.K. *Generalized Inverse of Matrices and Its Applications*; Wiley: New York, NY, USA, 1971.
27. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
28. Huang, G.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2012**, *42*, 513–529. [[CrossRef](#)] [[PubMed](#)]
29. Yao, W.; Zeng, Z.; Lian, C.; Tang, H. A Kernel ELM Classifier for High-Resolution Remotely Sensed Imagery Based on Multiple Features. In *Advances in Neural Networks—ISNN 2014*; Springer: Cham, Switzerland, 2014; pp. 270–277.
30. Mariéthoz, J.; Bengio, S. A kernel trick for sequences applied to text-independent speaker verification systems. *Pattern Recognit.* **2007**, *40*, 2315–2324. [[CrossRef](#)]
31. Schölkopf, B.; Smola, A.J.; Williamson, R.C.; Bartlett, P.L. New Support Vector Algorithms. *Neural Comput.* **2000**, *12*, 1207–1245. [[CrossRef](#)] [[PubMed](#)]
32. Chen, P.; Lin, C.; Schölkopf, B. A tutorial on ν -support vector machines. *Appl. Stoch. Models Bus. Ind.* **2005**, *21*, 111–136. [[CrossRef](#)]
33. Erbek, F.S.; Özkan, C.; Taberner, M. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *Int. J. Remote Sens.* **2004**, *25*, 1733–1748. [[CrossRef](#)]
34. Otukei, J.R.; Blaschke, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12* (Suppl. 1), 27–31. [[CrossRef](#)]

35. Cooley, T.; Anderson, G.P.; Felde, G.W.; Hoke, M.L.; Ratkowski, A.J.; Chetwynd, J.H.; Lewis, P. FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Toronto, ON, Canada, 24–28 June 2002; Volume 3, pp. 1414–1418.
36. Dwivedi, R.S.; Rao, B.R.M. The selection of the best possible Landsat TM band combination for delineating salt-affected soils. *Int. J. Remote Sens.* **1992**, *13*, 2051–2058. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).