A novel optimization parameters of support vector machines model for the land use/cover classification

Ying Liu 1,2,*, Bai Zhang 1*, Lihua Huang 1 and Limin Wang 2

1. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, 130012 Changchun, Jilin province, China.
2. Faculty of Management Science and Information Engineering, Jilin University of Finance and Economics, 130117 Changchun, Jilin province, China.

Received 1 January 2012, accepted 20 April 2012.

Abstract

Nowadays, support vector machines (SVM) are receiving increasing attention in land cover/use classification although one of the major drawbacks of the technique is the kernel function selection and its parameters setting. In this paper, a novel SVM parameters optimization method based on self-adaptive mutation particle swarm optimizer (SAMPSO-SVM) is proposed to improve the generalization performance of the SVM classifier. The SAMPSO algorithm, which is based on the variance of the population’s fitness, can break away the local optimum by the operation of self-adaptive mutation. Accordingly, very high classification accuracy will be achieved with the best value of the parameters of SVM, which have been searched using SAMPSO. In order to verify the validity of this SAMPSO-SVM method, a remote sensing land use/cover classification model is constructed using multi-spectral Landsat-5 TM data. In particular, they are organized so as to test the sensitivity of the SAMPSO-SVM model and that of the other reference classifiers used for comparison, i.e. maximum likelihood classifier (MLC), SVM classifier and standard PSO algorithm for SVM parameters optimization model (PSO-SVM). On an average, the SAMPSO–SVM model yielded an overall accuracy of 93.59% against 83.92% for maximum likelihood classier and outperformed PSO-SVM classifier in terms of overall accuracy (by about 2%). The obtained results clearly confirm the effectiveness and robustness of the SAMPSO-SVM approach to the remote sensing land use/cover classification.

Key words: Support vector machines, self-adaptive mutation, particle swarm optimization, land use/cover, classification.

Introduction

Land cover/use information about the Earth’s surface has been identified as one of the crucial data components for many aspects of global change studies and environmental applications. Remote sensing technologies have the capability of monitoring the Earth’s surface with different spatial, spectral and temporal resolutions. How to extract accurately and timely knowledge about land use/cover from remote sensing imagery relies upon not only the data quality and resolution, but also the classification techniques used. Therefore, improvement of remote sensing classification accuracy is always a concern. A bundle of approaches, e.g. the maximum likelihood classifier (MLC), artificial neural network classifier (ANN) 6,7 have been developed to build classification and prediction models. However, the MLC is a parametric classifier based on statistical theory. The major limitation of the method is that its assumption of normal distribution of class signature 8. In contrast to MLC, ANN is a more recent non-parametric classification technique which does not depend upon an assumption of normally distributed data, but accurately describing processes that translate input data into output classes can be difficult due to the combined use of multiple nonlinear activation functions at different layers. 9

Support vector machines (SVM), which were suggested by Vapnik 10, are one of the latest additions to the existing catalog of image classification techniques 11,12 and have been widely applied to land use/cover classification 13-16, but the most crucial problem is how to select the kernel function and its parameter values for SVM classification method 17. Many scholars use the intelligent optimization algorithms for SVM parameters optimization, including genetic algorithm (GA) 18,19, particle swarm optimization (PSO) 20,21. Although PSO has fast converging characteristics and more global searching ability at the beginning of the run than GA, it is more likely that PSO will explore local optima at the end of the run. The phenomenon of premature convergence can produce inaccurate parameters, which thereby affect on the generalization ability of SVM.

In this study, based on the lack of PSO-SVM model, a novel SVM parameters optimization method based on self-adaptive mutation particle swarm optimizer (SAMPSO-SVM) is proposed. In order to test its validity, SAMPSO-SVM is used for land use/cover classification using Landsat-5 TM image acquired in 2006. Also, performances of SAMPSO-SVM, PSO-SVM and MLC are compared and statistically analyzed.

Materials and Methods

Study area and data sets: The study area chosen for this research located in the Yanji district of Jilin province, the boundary between
China, North Korea and Russia (Fig. 1). Because of its geographic advantages and regional conditions, the region possesses remarkable position in the economic development belt along the boundary. With grouping and internationalizing of the world economy the region will be a hot point area of economic development in Northeast Asia in the future. However, the coordinative economic development depends on reasonable utilization of various resources. Land is a basic condition and site for production and living of human being. Land resource inventory and effective utilization are necessary key link for the economic development of the region. It is undoubtedly important to research the best method and model of land use/cover classification.

The newest Landsat-5 TM image is selected in the survey. The cloud-free scene image was acquired on September 30, 2006 (orbit number 115/30, 30 m spatial resolution, UTM projection). As it was just the growing season in our study area, each land use/cover type is relatively spectrally distinct. However, it is difficult to visually reflect the performance of various classification algorithms because of the limited format and smaller scale. So the typical region is chosen for our experiment from original TM images (5, 4, 2 band synthetic) so that classification algorithms are compared through the microscopic point of view (Fig. 2).

Support vector machines: Support vector machines (SVM) are supervised learning algorithms based on statistical learning theory, which are considered as heuristic algorithms. The aim of the SVM for classification is to determine a hyper plane that optimally separates two classes. An optimum hyper plane is determined using train data sets and its generalization ability is verified using test data sets.

Given a training set of instance-label pairs \((x_i, y_i), i = 1, 2, ..., m\) where \(x_i \in \mathbb{R}^n\) and \(y_i \in \{+1, -1\}\), the generalized linear SVM find an optimal separating hyper plane \(f(x) = <w \cdot x> + b\) by solving the following optimization problem:

\[
\begin{align*}
\text{Minimize} & \quad \frac{1}{2} <w \cdot w> + c \sum_{i=1}^{m} \xi_i \\
\text{Subject to} & \quad y_i(<w \cdot x_i> + b) - 1 \geq 0 \\
& \quad \xi_i \geq 0
\end{align*}
\]

where \(c\) is a penalty parameter on the training error, and \(\xi_i\) is the non-negative slack variables. SVM find the hyper plane that provides the minimum number of training errors (i.e. to keep the constraint violation as small as possible). This optimization model can be solved by introducing the Lagrange multipliers \(a_i\) for its dual optimization model. After the optimal solution \(a_i^*\) is obtained, the optimal hyperplane parameters \(w^*\) and \(b^*\) can be determined, and the indicator function (classifier) can be written as:

\[
\text{sign}(<w^* \cdot x> + b^*) \quad \text{or} \quad \text{sign}(\sum_{i=1}^{m} a_i^* y_i <x_i \cdot x> + b^*)
\]

When it is not possible to define the hyper plane by linear equations, the data may be mapped into a higher dimensional space \((H)\) through some nonlinear mapping functions \((\Phi)\), as
shown in Fig. 3. An input data point \( x \) can be represented as \( \Phi(x) \) in the high-dimensional space. The expensive computation of \( < \Phi(x), \Phi(y)> \) is reduced by using a kernel function \( ^{25} \). Kernel functions commonly used in SVM can be generally aggregated into four groups: namely, linear, polynomial, radial basis function and sigmoid kernels. Equations of kernel functions are following:

Linear kernel: \( K(x, x) = (x \cdot x) \)  

(4)

Polynomial kernel: \( K(x, x) = (x \cdot x + 1)^d \)  

(5)

where \( d \) is the natural number. 

Radial basis function kernel:

\[
K(x, x) = \exp(-\frac{1}{\sigma} |x - x|^2) \quad \text{or} \quad k(x, x) = \exp(-y|x - x|^2) 
\]

(6)

Sigmoid kernel: \( K(x, x) = \tanh(kx \cdot x - \delta) \)  

(7)

By introducing the kernel function, the nonlinear SVM classifier has the following forms:

\[
\text{sign} \left( \sum_{i=1}^{m} y_i a_i^* \Phi(x_i) \Phi(x) + b^* \right) \quad \text{or} \quad \text{sign} \left( \sum_{i=1}^{m} y_i a_i^* k(x_i, x) + b^* \right) 
\]

(8)

Several strategies including one-against-all, one-against-one and all-together \( ^{26} \) have been developed to apply SVM for multiclass problems. In one-against-all strategy, a set of binary SVM classifiers, each trained to separate one class from the rest is applied. On the other hand, in one-against-one strategy \( k(k-1)/2 \) SVM are constructed for each pair of classes. It can be considered as more symmetric than the one-against-all approach in terms of class sizes \( ^{27} \). The disadvantage of the method is the increase in the number of classifiers as the number of classes increase.

Common particle swarm optimization algorithm (PSO): PSO is an evolutionary computation technique proposed by Kennedy and Eberhart \( ^{28} \) that originated from the simulation of the behavior of a group of a flock of birds or school of fish or the social behavior of a group of people. Particles (individuals) representing a potential problem solution move through an n-dimensional search space. Each particle \( i \) maintains a record of the position of its previous best performance in a vector called pbest. The nbest, is another “best” value that is tracked by the particle swarm optimizer. This is the best value obtained so far by any particle in that particle’s neighbourhood. When a particle takes the entire population as its topological neighbours, the best value is a global best and is called gbest. All particles can share information about the search space \( ^{29} \).

Representing a possible solution to the optimization problem, each particle moves in the direction of its best solution and the global best position discovered by any particles in the swarm. Each particle calculates its own velocity and updates its position at the end of each iteration. Let \( P_{best} \) denote the best previous position encountered by the \( i^{th} \) particle. \( P_{best} \) denotes the global best position thus far, and \( k \) denotes the iteration counter. The current velocity of the \( d^{th} \) dimension of the \( i^{th} \) particle at time \( k \) is the following:

\[
V_{i,d}^{k+1} = \omega V_{i,d}^{k} + c_1 r_1 (P_{best,i}^{d} - X_{i,d}^{k}) + c_2 r_2 (P_{best}^{d} - X_{i,d}^{k}) 
\]

(9)

In the above formula, \( r_1, r_2 \) is a random function in the range \([0, 1]\), positive constant \( c_1 \) and \( c_2 \) are personal and social learning factors, and \( \omega \) is the inertia weight. Inertia weight was first introduced by Shi and Eberhart \( ^{30} \). Inertia weight balances the global exploration and local exploitation. The new position of a particle is calculated using the following formula:

\[
X_{i,d}^{k+1} = X_{i,d}^{k} + V_{i,d}^{k+1} 
\]

(10)

Self-adaptive mutation particle swarm optimization (SAMPSON): Standard PSO is distinctly different from other evolutionary-type methods in a way that it does not use the filtering operation (such as crossover and mutation) and the members of the entire population are maintained through the search procedure so that information is socially shared among individuals to direct the search towards the best position in the search space. However, if “the best position” is only local optimum, rather than the global optimum, the particles will not be able to search in the solution space again. The problem with more local optima is the phenomenon of premature convergence. In order to overcome the drawback of the PSO algorithm, it is necessary to provide a mechanism to make the algorithm jump into other areas of the solution space to continue the search until the global optimum is found in the final.

Lv and Hou \( ^{31} \) proposed the phenomenon of premature convergence is usually determined by two factors:

- The variance of the population’s fitness \( \sigma^2 \):

\[
\sigma^2 = \sum_{i=1}^{m} \left( \frac{f_i - f_{avg}}{f} \right)^2 
\]

(11)

Where \( f_i \) is the fitness of the \( i^{th} \) particle, \( f_{avg} \) is the average of the population’s fitness. The equation of \( f \) is the following:

\[
f = \begin{cases} 
\max \left( \left| f_i - f_{avg} \right| \right), & \text{if } \left| f_i - f_{avg} \right| > 1 \\
1, & \text{otherwise}
\end{cases}
\]

(12)

\( \sigma^2 \) reflects the convergence level of all the particles. The smaller the value of \( \sigma^2 \) is, the greater probability of particle swarm tends to converge, otherwise, particles will be in the random search phase. If PSO algorithm is driven to premature convergence stage or achieve global convergence, some particles will gather one or more specific locations in the search space. Here, the value of \( \sigma^2 \) is zero.

The comparison of the global optimal solution and theoretical optimal solution: It is known that premature convergence or global convergence cannot be distinguished only by virtue of the
variance of the population’s fitness. The optimal solution at this time must be judged whether it can belong to the theoretical global optimal solution \( f_d \) or not. If currently PSO algorithm is premature convergence rather than the global convergence, the onward direction of the particles can be changed by altering the global optima \( p_{gd} \) (mutation operation), so that particles can continue search into other areas. In the subsequent search process, the algorithm may find a new individual extreme \( p_\alpha \) and global optimal solution \( p_{gd} \). Such the cycle is executed; the algorithm can find the global optimum.

Mutation operation is designed as a random operators, namely \( p_\alpha, p_{gd} \) that meet the aforementioned conditions can complete the mutation operation with the probability \( p_m \). Equation is the following:

\[
p_\alpha = \begin{cases} k, & \text{if } \sigma^2 < \sigma_j^2 \text{ and } f(p_{gd}) < f_d, \\ 0, & \text{otherwise} \\ k \in [0,1] \end{cases}
\]  

The mutation operation of \( p_{gd} \):

\[
p_{gd} = p_{gd} * (1 + 0.5 * \eta)
\]

where \( \eta \) is a random variable subject to Gauss [0, 1] distribution. The SAMPSO algorithm is described in steps in Table 1.

The self-adaptive and normal mutation operators can restore the diversity loss of the population and improve the capacity of the global search of the algorithm.

**The architecture of SAMPSO-SVM classification modeling:** The parameters that should be optimized include the penalty parameter \( c \) and the kernel function parameters such as the value of gamma \( g \) for the radial basis function (RBF) kernel. Fig. 4 shows the flowchart of the SAMPSO-based parameters determination approach for the SVM classifier.

The fitness function is used to evaluate the quality of every particle which must be designed before searching for the optimal values of both the SVM parameters. The fitness function is based on the classification accuracy of a SVM classifier, which is as follows:

\[
\text{Fitness function} = \frac{y_i}{y_i + y_j} \times 100
\]

**Table 1. SAMPSO algorithm.**

<table>
<thead>
<tr>
<th>Step #</th>
<th>Action taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data preparation.</td>
</tr>
<tr>
<td>2</td>
<td>Initialize an array of particles with random positions and their associated velocities inside the problem space.</td>
</tr>
<tr>
<td>3</td>
<td>( p_b ) will be defined as the current particle position, ( p_g ) will be set to the best position of the initial particle population.</td>
</tr>
<tr>
<td>4</td>
<td>Check whether the algorithm meets the convergence condition or not. If not, go to step 5, otherwise step 11.</td>
</tr>
<tr>
<td>5</td>
<td>Update the particle position and velocity by Eqs. (9) and (10).</td>
</tr>
<tr>
<td>6</td>
<td>Evaluate the fitness value of each particle. Compare the current fitness value with the particles’ previous best value ( p_{gb} ). If the current fitness value is better, then assign the current fitness value to ( p_{gb} ) and assign the current coordinates to ( p_{gb} ) coordinates.</td>
</tr>
<tr>
<td>7</td>
<td>Determine the current global maximum among particle’s best position. If the current global maximum is better than ( p_g ), then assign the current global maximum to ( p_g ) and assign the current coordinates to ( p_g ) coordinates.</td>
</tr>
<tr>
<td>8</td>
<td>Count the variance of the population’s fitness ( \sigma^2 ) and ( f(p_{gd}) ) by Eqs. (11) and (12).</td>
</tr>
<tr>
<td>9</td>
<td>Calculate the mutation probability ( p_m ) by Eqs. (13).</td>
</tr>
<tr>
<td>10</td>
<td>Generate ( r ) and ( r \in [0,1] ), if ( r &lt; p_m ), carry out mutation operation by Eqs.(14). otherwise go to step 4.</td>
</tr>
<tr>
<td>11</td>
<td>End the algorithm, output ( p_g ).</td>
</tr>
</tbody>
</table>

Results and Discussion

**Experiment settings:** For comparison purposes, we implement two experiments, namely the SVM classifier with different kernels, and SVM parameters optimization model with two intelligent optimization algorithms (PSO and SAMPSO). In addition, there are several coefficients whose values can be adjusted to produce a better rate of convergence. Table 2 shows the coefficient values.

**Sample and feature selection:** Feature selection is very important in pattern classification. At this paper, based on spectral features and spatial distribution of vegetation, the features of the decision table include spectral bands, K-L and K-T principal component analysis (PCA) and normalized differential vegetation index (NDVI). Spectral bands include blue (Band 1), green (Band 2), red
(Band 3), near-infrared (Band 4), and two mid infrared (Band 5 and 7). Thermal band TM6 is excluded because it is less informative for vegetation classification and has a larger pixel size than the other bands.

According to land use/cover classification system and the map of vegetation distribution in Northeast China, the land use categories in the study area include Urban, Farmland, Water, Bare Land, and Forestland. Based on field experience and high-resolution aerial photographs, two data sets are formed using random pixel selection strategy, which guarantees the maximum variation and representativeness available for each class in the experiment. The first set is used for training purpose with each class comprising independent randomly selected samples. The second set is used for testing with each class comprising independent samples randomly selected. The detailed description of these land use/cover types and the number of samples are listed in Table 3.

### Table 3. The number of types and samples.

<table>
<thead>
<tr>
<th>Class code</th>
<th>Class name</th>
<th>The number of train samples</th>
<th>The number of test samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>o₁</td>
<td>Urban</td>
<td>335</td>
<td>418</td>
</tr>
<tr>
<td>o₂</td>
<td>Farmland</td>
<td>550</td>
<td>371</td>
</tr>
<tr>
<td>o₃</td>
<td>Water</td>
<td>125</td>
<td>237</td>
</tr>
<tr>
<td>o₄</td>
<td>Bare Land</td>
<td>367</td>
<td>389</td>
</tr>
<tr>
<td>o₅</td>
<td>Forestland</td>
<td>682</td>
<td>364</td>
</tr>
<tr>
<td>Total number</td>
<td></td>
<td>2059</td>
<td>1779</td>
</tr>
</tbody>
</table>

### Classification accuracy assessment: As widespread used overall accuracy (OA) can give misleading results, kappa coefficient is also used to assess classification results. Kappa coefficient expresses whether correctly assigned pixels may have been assigned by chance or not based on the classification decision rule.

### Experiment 1-performance evaluation for SVM classifier with different kernel functions: To evaluate the classification accuracy of the proposed classification model, we implement, in the first experiment, the SVM classifier with two kernels, namely sigmoid and radial basis function kernels, leading therefore to two SVM classifiers, termed as SVM-Sigmoid and SVM-RBF, respectively. The related parameters $c$ and $g$ for this kernel are determined by a grid search method using cross validation approach. For example, the best $(c, g)$ is $(100, 0.143)$ resulting the smallest generalization error. In addition, classification results for the maximum likelihood classification produce from the images are given at the same time in Table 5 that includes overall accuracies (OA) with Kappa coefficients. It can be seen from the table that SVM with radial basis function produced an overall accuracy of 87.07%, Kappa coefficients of 0.8372 whilst sigmoid kernel produced an overall accuracy of 84.60%, Kappa coefficients of 0.8058. It is also observed that the maximum likelihood classifier produced an overall accuracy of 83.92%, Kappa coefficients of 0.7974.

When the classification results presented in Table 4 are analyzed, several important conclusions can be drawn. Firstly, it is found that the lowest classification accuracies are produced for forestland class. The reason can be related to complex or close class boundaries (e.g. farmland class) resulting from high spectral resemblance to other classes and mixed pixels in the training and test samples. Secondly, it is observed that the water class is generally classified with 100% accuracy. This indicates high spectral discrimination capabilities for this particular class. Thirdly, SVM with RBF kernel, for the most cases, outperform SVM with sigmoid kernel with about 2.47% improvement. SVM with RBF kernel is found more powerful (approximately 3.15% higher) than the maximum likelihood classifier for the data sets used in this study. So the radial basis function kernel will be chosen in the second experiment.

### Experiment 2-remote sensing classification with SAMPSO-SVM: After testing the performances of the trained SVM-RBF model and verifying its effectiveness, at the second experiment, model parameters (the penalty parameter $c$ & kernel function parameter $g$) will be optimized using two different approaches respectively, including traditional PSO and our proposed SAMPSO. Classification results for SVM classifier, PSO-SVM classifier and SAMPSO-SVM classification model produced from the image are given in Table 5 that includes the value $c$ and $g$ and overall accuracies together with Kappa coefficients. As can be seen from the table, overall accuracy of SAMPSO-SVM model is 93.59%. Kappa coefficients of 0.9175 whilst PSO-SVM classification produce an overall accuracy of 91.50%, Kappa coefficients of 0.8903. It can be seen from the table that original SVM-RBF classifier only produced an overall accuracy of 87.07%, Kappa coefficients of 0.8372. SAMPSO-SVM classification model

### Table 4. Classification accuracy by different classifiers.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>$o₁$</th>
<th>$o₂$</th>
<th>$o₃$</th>
<th>$o₄$</th>
<th>$o₅$</th>
<th>OA</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM-Sigmoid</td>
<td>93.68</td>
<td>95.85</td>
<td>100.00</td>
<td>81.64</td>
<td>63.10</td>
<td>84.60</td>
<td>0.8058</td>
</tr>
<tr>
<td>SVM-RBF</td>
<td>98.80</td>
<td>94.16</td>
<td>99.53</td>
<td>92.80</td>
<td>64.68</td>
<td>87.07</td>
<td>0.8372</td>
</tr>
<tr>
<td>Maximum likelihood</td>
<td>98.56</td>
<td>78.72</td>
<td>99.53</td>
<td>92.48</td>
<td>61.69</td>
<td>83.92</td>
<td>0.7974</td>
</tr>
</tbody>
</table>

### Table 5. Comparison of classification parameters and classification accuracy, Kappa coefficient.

<table>
<thead>
<tr>
<th>Classification model</th>
<th>Penalty parameter $c$</th>
<th>Kernel function parameter $g$</th>
<th>OA (%)</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>100</td>
<td>0.125</td>
<td>87.07</td>
<td>0.8372</td>
</tr>
<tr>
<td>PSO-SVM</td>
<td>166.9423</td>
<td>1.366</td>
<td>91.50</td>
<td>0.8903</td>
</tr>
<tr>
<td>SAMPSO-SVM</td>
<td>246.7891</td>
<td>0.134</td>
<td>93.59</td>
<td>0.9175</td>
</tr>
</tbody>
</table>
is found more powerful (overall accuracy approximately 6.52% and 2.09% higher) than the SVM classifier and SVM-PSO classifier respectively for the data sets used in this study.

In addition, the results of study area image classification using four different classification models are presented in Fig. 5. Different class identifications by the four classification approaches are highlighted using white circles on three particular regions for Landsat-5 TM image. The experiment shows that SAMPSO-SVM model can improve the common SVM and PSO-SVM in classification performance for remote sensing data, which contains uncertainty or vagueness among farmland and forestland, caused by spectral confusion between-class and spectral variation with in-class. The result also reveals that the thematic maps produced using the maximum likelihood classifier have more noisy output compared to the other maps.

### Conclusions

Land cover/use classification using remotely sensed images is one of the most common applications in remote sensing, and many algorithms have been developed and applied for this purpose in the literature. Support vector machines (SVM) are a group of supervised classification algorithms that have been recently used in the remote sensing field. Although it is reported that they produce more accurate classification results than the conventional methods, the selection of optimum kernel function and its parameters are the major issues that largely affect their performances. In this study, a new version of PSO, named SAMPSO, is proposed to optimize the parameters of support vector machine (SAMPSO-SVM). During the running time for SAMPSO algorithm, the mutation probability for the current best particle is determined by the variance of the population’s fitness and the current optimal solution. In other words, SAMPSO can be helpful to escape more easily from local minima than can be done through the traditional PSO, so that the ability of global convergence will be effectively improved.

The performance of the SAMPSO-SVM is evaluated by means of classification the data of land cover/use, also, compared to MLC, SVM, PSO-SVM classifiers. Results noticeably indicate that the SVM with the radial basis function kernel method, for almost all cases, outperformed maximum likelihood classifier in terms of overall accuracy (by about 3%) and individual class accuracies. It was found that the SAMPSO-SVM produced more accurate results than PSO-SVM by about 2% overall accuracy. The simulation results demonstrate that the SAMPSO presented here is available for the SVM to seek optimized parameters for the remote sensing land use/cover classification. But SAMPSO-SVM method might be improved when taking these two reasons into account: first, how to reduce the computational time; second, the decision of theoretical global optimal solution is more or less subjective.

### Acknowledgements

This research was supported by the Knowledge Innovation Program of Chinese Academy of Sciences (KZCX2-YW-Q06-4).

### References