Using ensemble systems to study natural processes
Stanislav Yamashkin, Milan Radovanovic, Anatoliy Yamashkin and Darko Vukovic

ABSTRACT
Increasing accuracy of the data analysis of remote sensing of the Earth significantly affects the quality of decisions taken in the field of environmental management. The article describes the methodology for decoding multispectral space images based on the ensemble learning concept, which can effectively solve important problems of geosystems mapping, including diagnostics of the structure and condition of catchment basins, inventory of water bodies and assessment of their ecological state, study of channel processes; monitoring and forecasting of functioning, dynamics and development of geotechnical systems. The developed methodology is based on an algorithm for analyzing the structure of geosystems using ensemble systems based on a fundamentally new organization of the metaclassifier that allows for a weighted decision based on the efficiency matrix, which is characterized by an increase in accuracy of the decoding of space images and resistance to errors. The metaclassification training algorithm based on the method of weighted voting of monoclassifiers is proposed, in which the weights are calculated on the basis of error matrix metrics. The methodology was tested at the test site ‘Inerka’. The performed experiments confirmed that the use of ensemble systems increases the final accuracy, objectivity, and reliability of the analysis.

INTRODUCTION
The problems of effective water resources management, rational water use, and water quality assessment are priorities for mankind at the beginning of the 21st century. The development of projects to optimize water use should be based on a comprehensive study of geosystems, since all components of the landscape have a certain influence on the water balance-moisture permeability and moisture saturation of rocks and soils, their granulometric composition, infiltration capacity, morphology and morphometry of the relief, vegetation cover and land use structure. For example, on saturated sandy soils, quickly percolating moisture is almost not redistributed over the surface, and the natural complexes folded by sands are moistened more or less evenly throughout the area. A feature of most natural and artificial water bodies (rivers, lakes, canals, reservoirs, etc.) is their actual area and length, and uneven distribution.

Obviously, the quality of decisions in the field of water management is significantly affected by the availability of integrated and frequently updated information on the structure of the catchment area and the dynamics of changes in conditions of spontaneous or technogenic development. In this regard, the need to improve methods of remote sensing (RS) of the Earth from space and geoinformation technologies for decoding of space images is obvious. With the help of remote sensing data and software complexes for their processing, many important tasks can be solved, including: inventory of water bodies, assessment of their ecological state, study of channel processes and mapping.

Key words | data analysis, ensemble learning, geosystems, remote sensing
of geosystems, and forecasting and operational monitoring of floods.

Identifying land or water surface plots that are relatively homogeneous in terms of geophysical and geochemical properties (geological environment, geological and geomorphological as well as hydrological processes, the development of vegetation phenological phases, etc.) allows for the evaluation of the conditions of lands and prediction of their resistance to technogenous load and natural or technogeneous emergencies. Data collected by Earth remote sensing (ERS) as well as the analysis of such data using automated classification methods are becoming more and more important for research and practice activities related to assessing the spatio-temporal structure of the Earth’s surface and natural hazards research. This approach is important because it allows for the extraction of the necessary objective data to make science-based management decisions.

Instrumental deciphering of multispectral space images is based on the methods and algorithms of data signal analysis (Woodcock et al. 1984; Kruse et al. 1993; Haykin 1999; Duda et al. 2001; Landgrebe 2003), diagnosis of the types and properties of objects on the basis of systemic linkages in their spectral properties and structural characteristics (Schowengerdt 2007), statistical classification (Breiman 1996), neural networks (Rosenblatt 1964; Rumelhart et al. 1999; Kohonen 2001; Galushkin 2007), and the support vector machine (Vapnik 1995; Foody & Mathur 2004). Advanced search for new information acquisition methods is important because ERS data interpretation requires quality improvements to solve some research and practice problems related to hazards prediction and nature management optimization. We have analyzed the current situation in the field of pattern recognition and found it is promising to carry out experiments aimed at improving the efficiency of multispectral space image interpretation based on the ensemble learning concept (Woods et al. 1997; Polikar 2006) combined with the use of contextual information through various synthetic descriptors. Developments in this direction are the core of scientific and technical research.

Ensembles are sets of learning classifiers that combine in some way their decisions to obtain more reliable and more accurate predictions in supervised and unsupervised learning problems. Nowadays, ensemble methods represent one of the main current research lines in machine learning (Anvari et al. 2014), and the interest of the research community in ensemble methods is witnessed by conferences and workshops specifically devoted to ensemble. Several ideas have been proposed to explain the characteristics and the successful application of ensembles to different application domains, including hydroinformatics (Boucher et al. 2009; Kim & Seo 2015). Empirical studies showed that in classification problems ensembles improve on single learning machines and, moreover, large experimental studies compared the effectiveness of different ensemble methods on benchmark data sets (Evora & Coulibaly 2009). The interest in this research area is motivated by the availability of fast computers that allow the implementation and the experimentation of complex ensemble methods. Nevertheless, the theoretical problems behind ensemble methods need to be reviewed and discussed more in detail, and the discussion of the application of ensemble methods to real-world problems (such as analysis of remote sensing data) has been limited. Another current problem is the relationships between ensemble methods and remote sensing data complexity.

ERS-based analysis of the geophysical shell is focused on identifying the objectively existing linkages between the areal properties and the field of an area’s reflected and own radiation. The developed models of geophysical shells allow can addressing many challenges of land use.

The problem consists in automating the ERS data analysis to monitor the conditions of lands and predict emergent natural processes; the solution to this problem should be based on the presence of objective linkages between the spectral properties of the areal model and the characteristics thereof. Physical surfaces can reflect and radiate electromagnetic waves in various ways. The measured value of the spectral brightness coefficient is affected by the conditions of the atmosphere as well as the physical-chemical and geometrical parameters of the area. The dependency of the spectral brightness coefficient on the wavelength is an important objective characteristic reflecting the properties of the area under research. The analysis of this dependency can provide information not only on the type of surface, but also on its properties (Yamashkin & Yamashkin 2013).

Available multispectral and hyperspectral ERS data allow the study of the optical properties of the geophysical shell in different spectral bands. This significantly expands
the analysis capabilities: if in a particular spectral band different land types have similar reflectivity, one can identify their regular objective differences taking into account the imagery of different channels. In contrast, geophysical shells of the same class feature similar properties of spectral curves.

Approbation of the developed methodology for the classification of remote sensing data on the basis of ensemble systems was carried out at the test site ‘Inerka’ when solving the problem of mapping geosystems for the purpose of assessing their stability to recreational loads, organizing monitoring of a natural landmark. The study of the territory was carried out using widely applied neural networks and ensemble systems. The developed ensemble system-based ERS data classification methodology was tested at the Inerka test site while solving the lake condition monitoring problem.

METHODOLOGY FOR ENSEMBLE SYSTEM TERRITORY ANALYSIS

Analysis and classification of ERS data should be based on extensive use of automatic and automated methods as a key research tool (Yamashkin 2015). To solve the problems in this area more efficiently, one has to determine the key characteristics of input data.

Given the importance of the spectral characteristics of a geophysical surface, initial approximation reduces the problem of remote sensing data interpretation to finding a functional dependency type:

\[ X = f(B(\lambda)) \]

(1)

where \( X \) is the desired characteristics of the object displayed; and \( B(\lambda) \) is the spectral characteristics of the object displayed.

Spectral brightness is the most important parameter obtained from remote sensing data and characterizing the areal objects under research; however, it is not the only such parameter. Additional data such as the characteristics of the neighborhood of an object, its heterogeneity and components as well as whether it belongs to the boundary areas are obtainable from ERS data by a synthetic method using numeric algorithms and mathematical modeling methods. Taking such parameters into account allows greater accuracy and efficiency to be achieved when determining the characteristics of the objects displayed:

\[ X = f(B(\lambda), D) \]

(2)

where \( D \) is additional spatial information obtained from ERS data.

Various factors affect the value of the spectral brightness coefficient, which on the one hand allows determination of a number of characteristics of the area under analysis; while on the other hand, it makes it difficult to identify the surface type. One needs \textit{a priori} information about such areas to take into account how such factors as noise exposure, atmospheric impact, and light affect the spectral properties of an object.

\textit{A priori} information about natural and anthropogenic objects is obtained by various measurements that supplement satellite imagery. This information is found in statistical data, cartographic materials, as well as field research output. \textit{A priori} information may affect the results of determining the properties of an object on the basis of its ERS-based characteristics, which makes the functional dependency suggested above considerably more complex and yet more objective:

\[ X = f(B(\lambda), D, I) \]

(3)

where \( I \) is \textit{a priori} information about the object displayed.

Remote sensing allows the collection of objective data in dangerous and inaccessible areas without costly and lengthy field research. However, analysis of remote sensing data always has to rely on \textit{a priori} information about the research object.

Therefore, the method for solving the inverse ERS problem is reduced to analyzing the spectral characteristics of the objects displayed, algorithmically extracted spatial data, and \textit{a priori} information about the research object. The functional dependency \( f(B(\lambda), D, I) \) is non-linear and rather complex due to the complexity and ambiguity of the object; the solution of the inverse RS problem is difficult to formalize. It is virtually impossible to identify a dependency that would provide exceptionally true and optimal results. At the same time, it is pivotal to enhance the objectivity of results.

That is why research and development of algorithms and methods for ERS data analysis are relevant and
multidirectional. This area of research employs statistical, neuronetic, and fuzzy algorithms, while the creation of new and integrated approaches to solving the inverse ERS problem remains a relevant and open question.

Ensemble systems (hybrid systems) are essentially about integrating a number of classifiers into a single system in order to make the final classification more accurate. The key structural components of ensemble systems are:

- a set of monoclassifiers generated from individual algorithmic units of a controlled classification; classifiers may be of a similar nature (yet differing parameter-wise) or very different (neural networks, support vector machines, decision trees);
- a metaclassifier, which is the component that analyzes the monoclassifier-provided classification results and makes the final decision on whether an object belongs to this or that class.

Figure 1 shows a generalized block diagram of the ensemble system. The input is an object $X$ described with a vector of parameters $\{x_1, \ldots, x_i, \ldots, x_k\}$. Each of the system classifiers makes a specific hypothesis $y_i$ on whether the object belongs to a specific class; the hypothesis is then transferred to the metaclassifier, which makes the final decision $Y$ to determine the class that the object belongs to.

When developing an ensemble system of this kind, four fundamental problems must be addressed. One needs to:

- choose which classifiers (monoclassifiers) it makes sense to include in the system;
- determine the metaclassifier algorithm as a structural unit responsible for merging all the monoclassifiers into a single entity;
- train the monomodels and the metaclassifier;
- assess the classification quality and the performance of both the ensemble system and the individual monomodels it consists of.

When working with an ensemble system, you have to solve problems that are inherent in the use of monomodels; one very special problem concerns the preparation of data for training and testing.

We further detail the process of training an ensemble system, as shown in Figure 2.
The first step in training an ensemble system consists of training the individual monoclassifiers of which it is formed. This phase has a specific structure that depends on the nature of such monomodels, and is based on a training data set prepared by a specialist.

The second step is to assess the efficiency of the classifier N. The analysis of ERS data involves the processing of large volumes of data, in which multiple classes are identified based on the spectral parameters of the geophysical shell(s). For such analysis, researchers need a tool that helps assess the classifications obtained. A matrix of errors and error-based metrics does well in such cases. Error matrices are often used to describe the quality of classification models; such a description is based on the classification of test data for which the true class values are known. First of all, this phase provides information on how well individual monoclassifiers function, which enables a further comparison of their efficiency against the general system performance; second, this phase enables the researcher to configure the metaclassifier with due account of how well specific classes are isolated.

To build an error matrix, we have to classify the test data using the classifier under analysis and to compose a contingency table based on its output. The number of rows and columns in this table depends on the number of classes isolated in the classification. The rows of the error matrix are formed of the true classes to which the test sample objects belong; the columns are formed of the classifier-predicted classes. In other words, the Mij element of the matrix specifies the number of class i objects classified as class j objects.

The ratio of the error matrix values is used to calculate various metrics that describe the success rate, accuracy, and error rate of the classifier. These metrics are valuable because they help assess the efficiency of the algorithm in general as well as its ability to correctly classify the objects of a certain class.

Error matrix-based metrics are naturally objective; however, their reliability directly depends on the quality of the test sample, the composition of which is subjective in nature. This fact suggests the representativeness of the test sample that has to be maximized, while the number of errors has to be minimized. Only in this case can the error matrix and the metrics based thereupon give a high-quality assessment of the classifier efficiency.

To assess the classification quality, it makes sense to introduce metrics that integrate the properties of multiple classifier correctness indicators at once. The accuracy and sensitivity harmonic mean was experimentally proven to be informative:

\[
F_i = 2 \frac{\text{Sensitivity}_i \cdot \text{Precision}_i}{\text{Sensitivity}_i + \text{Precision}_i}
\]  

(4)

This indicator, calculated based on the sensitivity and accuracy parameters, describes the ability of the classifier to correctly classify the objects of a certain class while avoiding type I and type II errors. The value of this metric falls within a range of 0 to 1.

The error rate metrics are complementary to the corresponding correctness metrics and fall within a range of 0 to 1. The smaller the values of these parameters the greater the ability of the classifier to perform properly.

It would be logical to use the classifier correctness and error rate metrics to assess the functioning of individual classifiers. However, such metrics are no less valuable if included in an ensemble system to develop a metaclassifier.

Step three is about training the metaclassifier as a structural unit that analyzes the hypotheses accepted by individual classifiers. In the simplest case, the final decision can be made by equal voting, i.e., the object is assigned the class that gets the most classifier votes:

\[
H = \arg\max_{c \in C} \left( \sum_{i=1}^N V(i, c) \right).
\]

(5)

where C is the set of classes; N is the number of monoclassifiers in the system; \(V(i, c) = 1\), if the classifier i selects the class c; \(V(i, c) = 0\), if the classifier i rejects the class c.

Use of this decision-making method makes classification more accurate thanks to the system choosing the class that gets the most classifier votes. However, in the case that most classifiers make a wrong decision regarding the class of an object, while the efficient classifiers become a minority, the system makes a wrong decision as well. It is, therefore, necessary to introduce some class-specific classifier efficiency factors:

\[
\begin{align*}
H &= \arg\max_{c \in C} \left( \sum_{i=1}^N w(i, c) V(i, c) \right), \\
V(i, c) &= \begin{cases} 
1, & \text{if the classifier } i \text{ selects the class } c \\
0, & \text{if the classifier } i \text{ rejects the class } c 
\end{cases}
\end{align*}
\]

(6)
where \( w(i, c) \) is the weight factor specifying the efficiency of the \( i \)-th classifier in identifying class \( c \) objects.

In Yamashkin (2015), the algorithm for analyzing the structure of geosystems using ensemble systems based on the fundamentally new organization of the metaclassifier is conceptually justified. It allows for the making of a weighted decision based on the efficiency matrix, which is characterized by increased accuracy of decoding of multizone space images and error tolerance. The metrics calculated based on the error matrix help assess how well a classifier identifies a specific class; that being said, they are potentially good as efficiency factors. The accuracy and sensitivity harmonic mean (\( F \)), as discussed above, is an objective assessment of the classifier’s ability to correctly classify the objects of a specific class while avoiding type I and type II errors (Yamashkin 2015). Judging from the advantages of the accuracy and sensitivity harmonic mean, its value can be adopted as the classifier efficiency metric. However, it is advisable to refine it slightly; a classifier described as \( F_\epsilon \leq \epsilon \) (failing in more than \( \epsilon \) of all classification cases) will be considered inefficient. The coefficient \( \epsilon \) is the inefficiency boundary. Then, the efficiency of the \( i \)-th classifier in identifying class \( c \) objects is calculable as follows:

\[
w(i, c) = \begin{cases} 
0, & \text{if } F_c(i) - \epsilon \leq 0, \\
\frac{1}{1 - \epsilon} (F_C(i) - \epsilon), & \text{if } F_C(i) - \epsilon > 0,
\end{cases}
\]

(7)

where \( F_C(i) \) is the accuracy and sensitivity harmonic mean of the classifier \( i \) when classifying class \( c \) objects; and \( \epsilon \) is the inefficiency threshold.

The value of the efficiency metric will equal 1 if the classifier functions perfectly, or 0 if the classification quality is beyond the inefficiency threshold \( \epsilon \). In many cases, the value of the parameter \( \epsilon \) might be assumed to equal 0.5 (i.e., a classifier is deemed inefficient if it fails in 50% or more cases).

Calculating the efficiency factors for each monomodel of the ensemble system allows for the building of a metaclassifier that makes a weighted decision on whether to accept this or that classifier-suggested hypothesis.

Based on the efficiency factors, we build an ensemble system efficiency matrix, the rows of which are formed of the isolated classes, while the columns are formed of the system monomodels. The value at a row–column intersection stands for the efficiency of the classifier \( j \) in classifying class \( i \) objects – \( w(j, i) \).

It is also useful to introduce metrics to assess the efficiency of the ensemble system. When identifying class \( i \) objects, that efficiency is assumed to equal the maximum efficiency of the monomodels in identifying that class:

\[
\alpha_i = \max_{j=1,...,M} w(j, i) 
\]

(8)

If \( \alpha_i \) takes a value of zero, the system is inefficient in identifying class \( i \) objects. If \( \alpha_i = 0 \) for all values of \( i \), the system is deemed completely inefficient, and using it for classification purposes is not advisable.

Maximum efficiency of the classifier \( j \), the parameter \( \beta_j \) is calculated by finding the maximum value of its efficiency in identifying a class. If the value of this metric equals 0, the classifier is inefficient and must be removed from the ensemble system:

\[
\beta_j = \max_{i=1,...,N} w(j, i).
\]

(9)

The fourth training step is to assess the efficiency of the metaclassifier and the ensemble system, in general; it might as well be done by using the error matrix apparatus. Upon completion of this phase, we can answer two important questions: whether the ensemble system is more efficient than monomodels when it comes to classifications, as well as to which extent it is more efficient. If the first question is answered positively, and the second one is answered convincingly, then we can use the trained ensemble system for data classification.

To classify data by means of the ensemble system, they first must be monomodel-analyzed; the results obtained shall be assessed using an efficiency matrix-based metaclassifier.

Training time is a sum of the following components:

\[
T_{ST} = \sum_{i=1}^{N} T_{ST_i} + \sum_{i=1}^{N} T_T + T_w,
\]

(10)

where \( T_{ST} \) is time to train the ensemble system; \( N \) is the number of monomodels in the system; \( T_{ST_i} \) is time to
train the \(i\)th monoclassifier; \(T_{Ti}\) is time to test the \(i\)th monoclassifier; and \(T_{W}\) is time to build error matrices, to calculate the metrics and to generate a system efficiency matrix.

The ensemble system classification time is a sum of the following parameters:

\[
T_C = \sum_{i=1}^{N} T_{Ci} + T_H, \tag{11}
\]

where \(T_C\) is the time it takes the ensemble system to classify the data; \(T_{Ci}\) is the time it takes the \(i\)th monoclassifier to classify the data; and \(T_H\) is the time it takes the metaclassifier to accept the resulting hypothesis.

Increased time consumption is the price paid for the better classification accuracy ensemble that systems provide. It should be noted that the modern possibilities of parallel computing can be used to solve this problem too: monomodels can be trained and used in parallel, thereby reducing computational time depending on the purpose of the simulation.

The next section presents an experimental study of the capabilities of such an ensemble system.

**ENSEMBLE SYSTEM TESTING OF A SATELLITE IMAGERY ANALYSIS METHODOLOGY**

Lake Inerka is a natural landmark located in the Sura River valley 12 km south-west of Bolshiye Berezniki, a settlement in the Republic of Mordovia, Russian Federation. With an area of 56.3 ha, it is Mordovia’s largest lake. Inerka connects with the 12 ha Lake Ishak. The area of the lake gave it its name, which translates from Mordvin as Great Lake (from ine, ‘great’, and erke, ‘lake’). Landsat 7 satellite imagery was used as the input data for this research.

**Choosing the plots for training**

To sample training plots from the input data, we had to determine the set of classes and form a test and training data bank based on field research data and expert information. The goal of the study set out the need to distinguish between separate classes of geosystems and water bodies, as well as their anthropogenic modifications. This is necessary in order to determine the features of the natural differentiation of the Inerka site, which includes the lake and nearby territories. Extraction of this information from the space survey materials will allow us to approach the solving of inventory problems in a study area, assessing ecological status of water bodies, development of schemes for integrated use, and protection of water resources. The results of classification of geosystems of the study site should be an important starting point for studying the biological diversity of ecosystems and for assessing the sustainability of the functioning of natural territorial complexes in conditions of economic development within the territory.

The study of the territory was carried out using widely applied neural networks and ensemble systems. Carrying out this experiment reveals the advantages of the offered methodology over existing methods as efficiency criteria, efficiency and error metrics are calculated based on error matrices. 115 training plots were divided into six geophysical shell classes. 9,882 elementary plots with a total area of 8.89 km² were extracted by a specialist for supervised classification. The total area of the training plots made up 1.4% of the total area of the territory. To study the efficiency of the ensemble system, it made sense to divide the training sample into two sets so as to research the effect of the input data selection errors on the classification accuracy.

**Training the classifiers**

To classify the presented data, we used four feedforward neural networks with one hidden layer (of varying capacity). Transfer function is sigmoid:

\[
\sigma(x) = \frac{1}{1 + e^{-x}} \tag{12}
\]

We used a backpropagation training method with cross-entropy measure of error:

\[
E = -\sum_{q=1}^{Q} \sum_{i=1}^{n} \ell_i^q \ln y_i^q \tag{13}
\]

where \(\ell_i^q\) is the target value of the output of the network \(i\) when applying the example \(q\); \(y_i^q\) is the real value of the output of the network \(i\) when applying the example \(q\); \(Q\) is the number of examples in the training set; and \(n\) is the number of network outputs.
The networks were trained on various data sets independently generated from the common data bank. Therefore, each classifier, trained on various samples, will have a unique ability to classify the geosystems of specific classes.

**Classifier efficiency assessments**

Based on the data on how the models classify the test sample, error matrices can be generated that help assess the efficiency of a classifier and calculate the relative efficiency and error rate metrics described above.

**Constructing an ensemble system**

We calculated the efficiency factors for an inefficiency threshold \( \varepsilon = 0.8 \); the values of those factors were summarized in an efficiency matrix which is necessary to develop the system metaclassifier. The classifiers were integrated into an ensemble system.

Based on the experiments, we calculated the following efficiency metric values for individual neural networks and for the ensemble system in general (see Table 1):

- Neural Network 1 (NN1) has a high-capacity (10 neurons) internal neural layer and is trained on the first data set.
- Neural Network 2 (NN2) has a medium-capacity (5 neurons) internal neural layer and is trained on the first data set.
- Neural Network 3 (NN3) has a low-capacity (3 neurons) internal neural layer and is trained on the first data set.
- Neural Network 4 (NN4) has a medium-capacity (5 neurons) internal neural layer and is trained on the second data set.

**Table 1 | Efficiency and error rate metrics for classifiers**

<table>
<thead>
<tr>
<th>Efficiency metric</th>
<th>Classifier</th>
<th>Objects</th>
<th>Vegetation</th>
<th>Anthropogenic territories</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Softwood</td>
<td>Mixed and broad-leaf</td>
<td>Secondary</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>NN1</td>
<td>0.9532</td>
<td>0.9876</td>
<td>0.9485</td>
</tr>
<tr>
<td></td>
<td>NN2</td>
<td>0.9987</td>
<td>0.9021</td>
<td>0.9743</td>
</tr>
<tr>
<td></td>
<td>NN3</td>
<td>1.0000</td>
<td>0.9989</td>
<td>0.4715</td>
</tr>
<tr>
<td></td>
<td>NN4</td>
<td>0.9987</td>
<td>0.9848</td>
<td>0.9065</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>0.9993</td>
<td>0.9898</td>
<td>0.9262</td>
</tr>
<tr>
<td>Selectivity</td>
<td>NN1</td>
<td>0.9955</td>
<td>0.9661</td>
<td>0.9895</td>
</tr>
<tr>
<td></td>
<td>NN2</td>
<td>0.9947</td>
<td>0.9925</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>NN3</td>
<td>0.9907</td>
<td>0.8283</td>
<td>0.9983</td>
</tr>
<tr>
<td></td>
<td>NN4</td>
<td>0.9916</td>
<td>0.9806</td>
<td>0.9891</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>0.9933</td>
<td>0.9776</td>
<td>0.9925</td>
</tr>
<tr>
<td>Accuracy</td>
<td>NN1</td>
<td>0.9873</td>
<td>0.9757</td>
<td>0.9819</td>
</tr>
<tr>
<td></td>
<td>NN2</td>
<td>0.9955</td>
<td>0.9523</td>
<td>0.9496</td>
</tr>
<tr>
<td></td>
<td>NN3</td>
<td>0.9925</td>
<td>0.9042</td>
<td>0.9008</td>
</tr>
<tr>
<td></td>
<td>NN4</td>
<td>0.9930</td>
<td>0.9824</td>
<td>0.9738</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>0.9945</td>
<td>0.9831</td>
<td>0.9802</td>
</tr>
<tr>
<td>NPV</td>
<td>NN1</td>
<td>0.9889</td>
<td>0.9898</td>
<td>0.9883</td>
</tr>
<tr>
<td></td>
<td>NN2</td>
<td>0.9997</td>
<td>0.9268</td>
<td>0.9938</td>
</tr>
<tr>
<td></td>
<td>NN3</td>
<td>1.0000</td>
<td>0.9989</td>
<td>0.8926</td>
</tr>
<tr>
<td></td>
<td>NN4</td>
<td>0.9997</td>
<td>0.9877</td>
<td>0.9790</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>0.9998</td>
<td>0.9917</td>
<td>0.9834</td>
</tr>
<tr>
<td>Awareness</td>
<td>NN1</td>
<td>0.9487</td>
<td>0.9537</td>
<td>0.9380</td>
</tr>
<tr>
<td></td>
<td>NN2</td>
<td>0.9934</td>
<td>0.8947</td>
<td>0.9182</td>
</tr>
<tr>
<td></td>
<td>NN3</td>
<td>0.9907</td>
<td>0.8272</td>
<td>0.4699</td>
</tr>
<tr>
<td></td>
<td>NN4</td>
<td>0.9903</td>
<td>0.9653</td>
<td>0.8956</td>
</tr>
<tr>
<td></td>
<td>ES</td>
<td>0.9927</td>
<td>0.9675</td>
<td>0.9186</td>
</tr>
</tbody>
</table>
• Ensemble System (AS) merges all the four neural networks via the trained metaclassifier.

Remote sensing data analysis showed that a pronounced slope-like alteration of ecosystem types is what characterizes the natural differentiation of the Volga Uplands at the Inerka test site. Forest-type landscapes are located on the raised butte-watershed landmasses (Vukovic et al. 2017); whereas, meadow-steppe geocomplexes are located at the lower close-to-valley slopes. Intrazonal landscapes of softwood and mixed forest are located in the area of the Inerka site, in its ancient alluvial sands that were strongly affected by Eolian processes.

The research into the characteristics of work effectiveness of separate monomodels and ensemble systems in general on test grounds allowed for the conclusion that efficiency of the decision-made ensemble system in a certain area is always close to the value of efficiency of the most effective monomodels in the field, and often exceeds it on sensitivity, selectivity, accuracy, and awareness. This property is based on the features of creation of metaclassifier proposed in work. In this case, use of an ensemble system always allows for the avoidance of gross blunders within the separate monomodels which systematically make an incorrect hypothesis about the belonging of objects to a certain class.

By now, the ensembles using combinations of models are developed and successfully applied (bootstrap aggregating, boosting, stacked generalization, voting); they can improve results and increase classification accuracy in comparison with the use of separate classifiers. The ensemble system proposed in this work allows, similar to the stack generalization approach, for the uniting several classifiers, which are varied in nature. In this case, in contrast to already proposed ways of adopting the resultant decision, the metaclassifier considers indicators of accuracy of separate classifiers in determination of objects of concrete classes on the basis of a matrix of weight coefficients. In addition, the concept of a lower threshold of accuracy defines the possibility of the classifier system to make a hypothesis about the belonging of any object to a specific class.

It should be noted that the modern possibilities of parallel computing can be used to solve this problem too: monomodels can be trained and used in parallel, thereby reducing computational time depending on the purpose of the simulation.

In addition to the neural networks, algorithms with different features (decision trees, support vector machines (SVM)) can be used as classifiers to form ensembles.

Combining classifiers with different features into ensembles is an interesting direction for further research.

The following results were obtained for the artificial neural networks, support vector machines, decision trees, and ensemble systems.

When analyzing the test site Inerka, the developed ensemble system showed the best result in comparison with individual monomodels (2.29% more accurate than the result shown by artificial neural networks) and other ensemble systems (1.65% more accurate than the result of the begging-system, constructed on neural networks, which was the second result) (Table 2). From the point of view of time, the projected ensemble system shows the average result, in some cases ahead of the systems created on the basis of begging and boosting. In all cases, the ensemble system trained and classified objects within the final predictable time.

Lake Inerka stretches from the south-west to the north-east for 4.1 km. The coastline of the lake is slightly sinuous. The greatest depth is 11.2 m. To the north and the south in cross section the basin has a trough-shaped form. On the sides, near the shore, the depth is up to 4 m. The bed of the basin is flat, with a gradual decrease of depth towards the ends of the lake where shallows with a depth of up to 2 m were created. The largest of the nearby lakes is Lake Tatar which is a crescent extending along Lake Inerka (to the south-west). The length of the lake is about 1,000 m.

Two ecological dangers have become aggravated for the development of Lake Inerka. The first is the threat of its disapperance in connection with degradation of a crossing point between the lake and the Sura River, and the second is siltation.

Table 2 | Experimental accuracy and operating time of classification models

<table>
<thead>
<tr>
<th>Classification model</th>
<th>Accuracy (% of correctly classified objects)</th>
<th>Time, min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>94.03</td>
<td>1.1</td>
</tr>
<tr>
<td>SVM</td>
<td>86.30</td>
<td>0.1</td>
</tr>
<tr>
<td>DTs</td>
<td>94.24</td>
<td>0.01</td>
</tr>
<tr>
<td>Ensemble model</td>
<td>96.55</td>
<td>1.35</td>
</tr>
</tbody>
</table>

S. Yamashkin et al. | Using ensemble systems to study natural processes | Journal of Hydroinformatics | 20.4 | 2018

Downloaded from https://iwaponline.com/jh/article-pdf/20/4/753/244853/jh0200753.pdf by guest
Distinctive features of the regime of the Sura River are the existence of spring high water with flooding of floodplains, low summer-autumn stability of low water (mezhena) broken in rainy years by two to three floods, and steady winter low water (mezhena). The spring rise in water level begins under ice in the third week of March. The maximum of spring high water usually comes in the middle of April, and lasts 1–2 days. The height of high water varies from 2–3 m to 5–6 m. Downturn is rather slow, and is frequently with secondary peaks from the early rains which fall at this time, prolonging high water duration. Duration of a flood is about 1.5–2 months.

Water travels from the Sura River to Lake Inerka during spring flood along the intermountain floodplain and through channels between the lakes. A part of the rainfall flows down into the lake in temporary streams. Over the last 15–20 years high floods have not been observed. The drain of rain floods passes in courses, without coming to the floodplain.

Lake Inerka is mainly fed by underground waters, but there are no strong indications of considerable participation of underground waters in the process of lake water supply. Big springs are not found. Feeding by underground waters is carried out via the thickness of sandy-silt sediments of beaches and the bottom of the lake. Close hydraulic connection with surface water of the Sura River and Lake Inerka is proved by the chemical composition of waters. The underground and surface water is poorly mineralized (200–500 mg/dm³), and of hydrocarbonate, calcium magnesium composition, neutral and slightly alkaline. Water flowing down from boggy sites contains increased amounts of organic substances.

The water entering the lake partially evaporates in summer; a certain quantity during summer and winter periods of the year filters into the bottom and beaches and flows down to the Sura River. In general, a relative balance between water entering the lake and its exit is observed so water volume in the lake and its level remain approximately the same from year to year. Over 4 years of observations, water level at the end of June to the beginning of July fluctuated within 25–30 cm. The spring rise of the water level does not exceed one and a half meters.

Due to the river meandering towards Lake Inerka a crossing point between the water bodies was reduced to 50–60 m and there was a threat of complete destruction of the unique lake as low-flow water level in Lake Inerka is 4–5 m higher than low-flow level in the Sura River (Figure 3). To save Lake Inerka the ‘Antierosion measures on the Sura River near Lake Inerka’ project was developed, with a flattening of the riverbed of the Sura River. Its axis was shifted 250 m to the right, and the crossing point between the Inerka and the Sura River was increased to 300 m.

In developing the project to save the lake, a detailed interpretation of multizonal space images with the purpose of landscape indication of streams of underground waters, an assessment of the stability of geosystems for construction, recreational loadings, and accumulation of products of technogenesis was carried out.

When classifying the natural units on the basis of satellite imagery, the following geosystems were identified (Figure 4):
(1) Butte-watered landmasses of the right bedrock side of the Sura River valley, with elevations of up to 265 m. They consist of carbonate and silicon-carbonate rocks of Paleogene and Upper Cretaceous periods, closed with shallow diluvial formations. On the steep slopes, chalk, gaize, and marl deposits often protrude through the ground’s surface. The specific features of the lithogenic landscape foundation are determined by the distribution of underdeveloped gray forest gravelly soils under broadleaf forests.

(2) A complex of large and small ridges with underdeveloped sod sandy soils and sod-meadow shallow light-loam soils, bluegrass-fescue meadows. Oak forests are sometimes encountered, too.

(3) Large kettle-and-basic complexes.

(4) Medium-level floodplains covered in loams with sand layers, meander scars, and small watercourses with excessively moisturized floodplain sod-granular, gray forest podzolized, and meadow-swampy heavy-loam soils. Natural vegetation is dominated by the floodplain oak forests with elms and alders. There are also small-area grass-for-foxtail as well as horsetail-tallgrass-sedge meadows with willow and alder thickets.

(5) Duna-like, buttes of intra-floodplain, large and often endorheic and dry erosion basins. The maximum elevation is 128.8 m. It is mostly dry-grass as well as lily-of-the-valley pine forests that grow on the podzolic underdeveloped and sod-cryptopodzol soils.

(6) Large meander scars.

(7) Riverbed sand shoals (beaches).

(8) Natural aquatic complexes of the lakes.

(9) Natural aquatic complexes of the Sura River.

The results of mapping of geosystems of the natural landmark Lake Inerka are the basis for assessing stability of geocomplexes for the purposes of construction, formation of recreational zones, and accumulation of products of technogenesis.

Assessment of sustainability of natural complexes for construction was carried out by analysis of the lithogenic basis of landscapes. The main indicators used were: soil type, groundwater depth from the Earth’s surface, surface slope, horizontal dismemberment, and manifestation of modern exogenous geological processes.

In accordance with accepted indicators on the degree of stability, natural complexes are divided into the following groups:

(1) Sustainable natural complexes (conditionally favorable for engineering development) – the above floodplain terraces of the Sura River. They are composed of alluvial sandy rock varieties, which include interlayers and lenses of loam, more rarely than clay. The depth of groundwater occurrence depends on the ruggedness of the relief and seasonal dynamics of the landscapes. Water has weak carbonate aggressiveness in relation to concrete. When designing structures, it is necessary to foresee structural measures that exclude the possibility of deformation of buildings, which can occur due to uneven sedimentation of soils, owing to their differing compressibility.

(2) The floodplain of the Sura River and the hollow-beam complexes are unstable natural complexes. Landscapes are periodically flooded with flood waters. The soils are represented by different sized uncompacted watersaturated sands and very soft loam. Construction deals with the use of expensive measures to protect against flooding, as well as to compact the ground base.

(3) Quite unstable natural complexes – swamps and peat lands. It should also be taken into consideration that groundwaters in biogenic soils and mud, as a rule, are highly aggressive to the materials of underground structures.

Evaluation of the stability of PTK for recreational loads is based on the analysis of soil types, the depth of groundwaters, the thickness of the humus horizon and the mechanical composition of soils and the type of vegetation. Based on analysis of the characteristics of natural complexes, the following groups were identified according to their degree of resistance to recreational loads:

(1) Resistant to recreational loads – geocomplexes of floodplains with normal moisture. Development of destructive geoeconomic processes occurs with the destruction of vegetation and soil compaction. It is necessary to preserve relatively weakly modified geocomplexes between intergranular depressions and old channels, which perform the stabilizing functions of geosystems.

(2) Relatively stable – natural territorial complexes of the western part of the territory of the natural landmark,
represented by mixed forests on the terrace above the floodplain. In this area, degradation of the landscape is weakened by relatively well-developed grass cover that holds the soil together.

(3) Unstable – the fragments of terraces with aeolian relief forms under pine forests. Landscapes have the highest esthetic value, especially along the Inerka coastline.

Assessment of geocomplexes to the accumulation of products of technogenesis is of increasing relevance in connection with the active use of motor transport by recreational vehicles. The lithogenic basis of the neo-eluvial landscapes of the territory of the natural monument is characterized by the predominance of sediments of light mechanical composition, good water permeability, which does not contribute to the accumulation of pollutants in the soils. The least resistant to technogenic pollution are the super-elementary landscapes of floodplain complexes, where an increase in the content of clay and silt particles is noted, and the concentration of heavy metals is increasing. However, the average content of most heavy metals does not exceed their background concentrations in the soils of the Republic of Mordovia.

CONCLUSIONS

Studying the efficiency of ensemble systems at the Inerka test site led us to the following conclusions:

- The efficiency of the decisions made by the ensemble system in a specific area is always close to that of the most accurate classifier for this area, often being more sensitive, selective, accurate, and aware. This feature is due to the specific functioning of the metaclassifier, as proposed above and verified experimentally.
- The system error rate is, in fact, approximately equal to that of the most efficient classifier, often not exceeding it. At the same time, ensemble systems always avoid the gross errors made by individual monoclifiers that systematically formulate an incorrect object-class hypothesis.
- If each monoclifier in an ensemble system has a high error rate (i.e., false detection) for a specific class, the ensemble system inevitably inherits this behavior. However, it has been empirically determined that the system error rate never exceeds that of the weakest classifier. Theoretically, this fact is validated by the weighted approach to solving the final hypothesis problem.

The time to train an ensemble system equals the time to train and test its monoclifiers plus the time to build error matrices, to calculate the metrics, and to generate a system efficiency matrix. The time it takes the ensemble system to classify data is a sum of the time it takes a monoclifier to classify the same data plus the time it takes the metaclassifier to accept the resulting hypothesis. Increased time consumption is the price inescapably paid for the better classification accuracy ensemble systems provide. As all of the training and classification time components are finite provided that monoclifiers are trainable, these parameters are also finite.

Use of ensemble systems to analyze ERS data has a number of advantages: lower error rates and improved accuracy of individual classifiers thanks to their integration in a single system; ensemble analysis performed by a system of classifiers trained on independent data groups reduces the effect of errors occurring due to the generation of test and training samples. Ensemble systems help in solving the classifier parameterization problem by combined usage of the same-class models with different parameters.

When developing an ensemble system, one solves the problems of identifying and training a set of classifiers in such a system as well as developing a metaclassifier algorithm that merges monoclifiers into a single entity. Quality and performance assessment problems are solved as well. The ensemble system training process consists of the following steps: 1) training individual monoclifiers; 2) assessing the efficiency of monoclifiers; 3) training the metaclassifier; 4) assessing the efficiency of the ensemble system. It is proposed for the metaclassifier to be trained, and its efficiency be assessed, using an error matrix and metrics based thereupon. It is also proposed that the metaclassifier algorithm will use a weighted voting of the system monomodels on whether this or that object belongs to a particular class. The weight factor is the classifier efficiency calculated on the basis of the error matrix.

The use of ensemble systems enables the rapid automated analysis of big ERS data for the thematic mapping of land-use systems as well as for the analysis of emergency situations and natural processes. A comparative analysis of the monoclifier-specific and ensemble system
performance led to the following conclusions: an ensemble system, compared against monomodels, had a better or very close result; and the empirically assessed efficiency of the ensemble system significantly exceeds that of the lowest efficiency of individual monomodels while helping avoid gross classification errors.

The use of ensemble systems for decoding the materials of space imagery of the test site Inerka made it possible to assess the stability of geocomplexes of the territory for the purposes of construction, formation of recreational zones, and accumulation of products of technogenesis. Studies have shown that the diversity of the PTK of the natural landmark Lake Inerka and, in general, their weak resistance to anthropogenic impact, should pre-suppose special circumstance when planning and locating economic and cultural objects. When organizing nature management on the territory of the natural landmark Lake Inerka, special attention should be paid to hydromorphic complexes that are not only easily vulnerable to all types of economic activities, but also perform important functions to maintain the water balance. Recreational load should also be minimized for the intra-flood terraced complexes with aeolian relief forms, where at a low level of adherence of the rules of nature management, areas of intensive destruction of soil cover are formed and the probability of occurrence of forest fires is high.

REFERENCES


First received 22 May 2017; accepted in revised form 17 January 2018. Available online 5 March 2018.