DESIGNING FORECAST THEMATIC MAPS USING TIME SERIES REMOTELY SENSED IMAGES

A.V. Zamyatin, N.G. Markov

Tomsk Polytechnic University
84, Sovetskaya street, Tomsk, 634034, Russia,
e-mail: ZamyatinAV@prim.ce.cctpu.edu.ru

KEY WORDS: Remote Sensing, Land Use Mapping, Prediction, GIS, Neural Networks

ABSTRACT:

In the paper a number of original algorithms implemented as the program system are considered. The algorithms allow to perform the automated forecasting of land use/cover change using time series remotely sensed (RS) images. In the framework of the proposed approach the forecasting performs in two stages. In the first stage time series RS images are classified by the original classification algorithm, which is based upon both statistical nonparametric and artificial neural network (ANN) classification and separate processing of spectral and spatial features of RS images. The complex classification allows significantly decreasing the "noisiness" of the final thematic maps and to increase the classification accuracy in comparison with classification methods of traditional processing, RS and mapping software. To make the interpretation of RS images more flexible and effective it is proposed to perform classification using ANN with original way of forming feature space. In the second stage, the designed raster thematic maps are processed by the enhanced algorithm of time series analysis, which is based upon Markov chains for transition matrix calculation and cellular automata application allowing to take into account not only probability of transition from one class to another but also spatial interclass correlation.
The results of effectiveness investigation of the proposed algorithms, obtained with use of model RS images and real time series RS images for Uymon steppe area (Altay Region, Russia) obtained from satellite RESURS-O1 are discussed.

1. INTRODUCTION

Every year high accuracy and operative forecasting necessity of land cover change has been growing. At the same time the operative and accurate forecasting is not possible without joint application of the comprehensive technologies for time series (TS) RS data interpretation, geoinformation systems (GIS) performing complex spatial analysis of interpreted data, and comprehensive methods of spatial forecasting.

At present a lot of researchers are solving the task of designing forecast maps with use of simple classification methods and highly specialized models of land use/cover change. Application of simple methods and algorithms of classification leads to inadequate thematic maps. Saying about highly specialized models of land use/cover models with their advantages and disadvantages, we may say that these models often require a large amount of additional information (digital elevation model, dependability of different land types change, features of migration etc.) that makes their application field very limited. The all mentioned shortcomings restricts the application of existing program facilities, that does not allow to obtain forecast maps with the appropriate accuracy using TS RS images.

All mentioned above declare about the imperfections for using existing approaches and algorithms for RS images implementation in the task of forecasting land use/cover change. Also it says about urgent need for new methods and algorithms allowing to solve all mentioned problems and restrictions more effectively.

Let's consider the possibilities and restrictions of methods and algorithms applied for tasks of RS images interpretation and designing forecast maps more detailed.

Now the task of interpretation of TS aerospace images, which in their turn can be used for forecasting, traditionally is solving with image processing, RS and mapping software such as ER Mapper (Earth Resource Mapping), ERDAS Imagine (ERDAS), Idrisi 32 (Clark University). This software is based upon either parametric statistical methods using assumption about normal distribution of features (traditional maximum likelihood algorithm) or nonparametric methods that can produce acceptable results in few cases only. Besides the actual image processing RS, and mapping software do not use in full measure spatial (texture) information about classes on aerospace images. Application of such simple methods and approaches allow to use RS information to some extent only. Therefore for obtaining more accurate interpretation results it is required new and significantly more complicated scheme of RS images classification.

While forecasting the behaviour of complex systems such as mapping nature territorial complexes to which influence a lot of stochastic processes, basically have been modelled by stochastic forecasting methods (Baker W. L, 1989). The widespread among them are methods using Markov chains. Markov chains have been used in a variety of fields and have modelled changes on a variety of spatial scales (Baker W. L, 1989). In order to Markov model considers spatial interaction between classes on thematic map cellular automata (CA) are often applied (e.g. Park, S. and Wagner, 1997). A parameter of CA is the distance of the neighbourhood from the central grid-cell. In majority cases of CA application for land use/cover change modelling this parameter is taken equal for all types of classes. But such approach does not take into account features of spatial interaction of classes on thematic map (Verburg P.H. et al., 2003).

In the paper the approach, with combined application of original algorithm of RS images interpretation, geoinformation system (GIS) and also algorithm of enhanced designing of forecast maps is developing. The features of the proposed algorithms are considered and investigation results of these algorithms using modeled aerospace images and real TS RS images are discussed.

2. PROCESS DESCRIPTION

Processing and interpretation of TS RS images using the proposed methods and approaches are based upon two sequential stages. In the first stage TS RS images are classified by the original advanced interpretation algorithm with separate processing of spectral and spatial features of RS images. It allows to obtain the less noisiness final thematic map and to increase its classification accuracy compared to traditional interpretation, RS and mapping software. The first stage of RS images interpretation is a key stage, because on the basis of interpretation results the forecast maps will be designed. In that case more accurate and adequate of interpretation results will lead to more accurate and adequate forecast results.

In the second stage the obtained raster thematic maps are processed by the original TS analysis algorithm, which is based upon Markov chains and CA processing algorithm, includes the optimal neighborhood size determination. This TS analysis algorithm allows to consider not only the probability of transition one class to the another, but the spatial interclass correlation.

Let's consider the algorithm applied in the stages of the advanced interpretation and the algorithm of the enhanced TS analysis in details.

2.1 First stage – advanced interpretation

Advanced interpretation in the framework of the proposed approaches is based upon Bayes decision rule of empirical risk minimization:

$$p(\omega_i | X) = \frac{p(\omega_i) p(X | \omega_i)}{\sum_{k=1}^{M} p(\omega_k) p(X | \omega_k)}, \qquad (1)$$

where $p(\omega_i)$ – prior probability of class *i*, *M* – the number of classes, $p(X | \omega_i)$ – conditional probability density of class *i*.

At the same time according to (1) interpretation in two steps is performed. In the first step the posterior probability maps for each classes are designed, at that the feature space is considering spatial characteristics. In the second stage according to (1) designed maps are to be used as prior probabilities of classes and the feature space consist only of spectral features.

Besides, it should be noted that in the first step either statistical or artificial neural network (ANN) classification could be used.

Statistical classification is based upon combined application of parametric and nonparametric algorithm of density estimation depends on the agreement of the data with the normal distribution according to the chi-square criterion, and also statistical classification includes standard parallelepiped classification algorithm. The simple and fast parallelepiped classification algorithm is applied in case the sample data range is not intersected by data range of any other samples.

ANN classification is implemented together with the approach to the storage and the search of the ANNs in a database. The general purpose of the approach is to make the process of ANNs topology and parameters definition easier and also to make the learning process of ANN faster. The search might be done with test of sign-rank correlation between the investigated data sample and the ANN train data sample stored in the database. The possibility of ANN search makes the ANN learning more predictable and robust. That is why in case of successful search of the appropriate stored ANN for investigated data sample the designing of prior probability maps is performed by ANN classification.

2.2 Second stage – forecast maps designing

TS thematic maps, designed in the first stage, and other additional data, obtained by including GIS are carried out by enhanced TS analysis algorithm. The algorithm includes iterative performing of three operations.

The first operation of the considering process is the analysis of the neighborhood characteristics of the raster interpreted thematic maps. At that the optimal scale of classes representation is being determined. The optimal scale needs further for the effective CA application.

The second operation is the constructing of transition matrix (TM). The operation includes the analysis of two and more thematic maps using first-order or high-order Markov chains respectively. The use of suitability maps, which show the probability of change one class to another, allows to range all image elements from high disposed to change till low disposed to change.

The final operation is processing of primary forecast map by CA with the optimal neighborhood size, defined in the first step.

The final result of the enhanced TS analysis algorithm is a forecast map for the further time step. To perform longer forecast it is needed to pass the obtained forecast map to the input of the algorithm as the TS input thematic map.

3. FEATURES OF APPLIED ALGORITHMS

One of the original algorithms applied for interpretation of RS images is the density estimation nonparametric algorithm, which is based upon Rosenblatt-Parzen (RP) algorithm and k-nearest neighbour (k-NN) algorithm. The original algorithm provides a computational cost over dozens times compared to existing nonparametric algorithms. Moreover the ways of forming feature space for both statistical and neural network approaches are original. Also the approach to processing of thematic maps by CA with optimal neighbourhood size defined by the enrichment factor is original. It is proposed to consider these features more detailed.

3.1 Nonparametric algorithm of density estimation

The widespread approaches to nonparametric density estimation are approaches using RP and k-NN methods, defined by (2) and (3) respectively:

$$\hat{p}(X \mid \omega_i) = \left(n_i \prod_{\nu=1}^{P} c_\nu^i \right)^{-1} \sum_{s=1}^{n_i} \prod_{\nu=1}^{P} \Phi\left(\frac{x_\nu - x_\nu^s}{c_\nu^i} \right), \; i = \overline{1, M} \qquad (2)$$

where n_i – the number of elements in the sample V_i of the class ω_i; P – the number of image feature bands; $\beta_\nu^i = c_\nu^i$ – smoothing parameters of the class ω_i; $\Phi(u)$ – kernel function of density distribution function.

$$\hat{p}(X \mid \omega_i) = \frac{1}{N} \frac{k_n - 1}{V(k_n, N, X)}, \; i = \overline{1, M} \qquad (3)$$

where k_n – distance parameter, N – the sample size. At that in case of Euclidean distance:

$$V(k_n, N, X) = \frac{\pi^{n/2} R_k^n}{|A|^{1/2} \Gamma[(n+2)/2]} \qquad (4)$$

where $V(k_n, N, x)$ – in common case is the amount of all points for that the distance to the point x less or equal R_k, A – unit matrix, Γ – gamma-function.

Direct using of algorithm on the basis of (2) и (3) leads to significantly low computational performance of density estimation.

The proposed original density estimation algorithm is based upon modifications of the mentioned RP and k-NN algorithms. Let's consider the issues of these modifications.

To increase the computational performance of conditional density estimation algorithms the advance kernel function calculation algorithm is applied. The idea of the algorithm is based upon the exclusion of the periodic low computational operations of kernel functions $\Phi(u)$ during density estimation in each component ν-th of multidimensional feature vector $X = \{x_\nu, \nu = \overline{1, P}\}$ at (2) by buffering (caching) once calculated function values.

To increase the computational performance of conditional density estimation of k-NN algorithm it is proposed and developed the modification of the algorithm. The main issue of the modification is in the following. According to (3) the parameter defined computational performance of density estimation is $V(k_n, N, x)$, which represents the distance in a current metrics. Therefore the acceleration of calculation of this parameter will lead to increase the computational efficiency of density estimation as a whole.

To perform a faster search of nearest neighbors the application of the algorithm of a spatial indexing is proposed. The spatial indexing allows to find easier and faster a necessary point in multidimensional space according to (3) и (4) and to calculate the density probability value.

The spatial indexing of data is a process of reflection a multidimensional space to the one-dimensional space by the indexes structure where each index corresponds to the point of multidimensional space. There are some approaches to the spatial indexing with different curves, such as Zet, Hilbert etc. The developed algorithm of density estimation is based upon Zet-curves, because the investigation results of the proposed density estimation algorithm with two different types of the index curves (Zet and Hilbert) demonstrated that the density estimation by algorithm with Zet-curves gives more robust results and index structure creation in this case is performed faster.

Some conducted research shows that more effective is a combination of mentioned algorithms of density estimation defined by (2) и (3). At that what specific algorithm should be applied in each specific case is defined upon data dimension P. In case $P \leq 3$ the modification of RP algorithm is applied, in case $P \geq 4$ the modification of k-NN algorithm is applied. The obtained algorithm, combined possibilities of couple nonparametric density estimation algorithms, allows the computational performance increase in dozens times compared to traditional nonparametric algorithms.

3.2 Feature space forming

Statistical: Recently in the tasks of interpretation the additional (texture) information in some way usually is used (Haralick R.M. & Joo H. A, 1986). More widespread method for considering the texture information is statistical approach to forming Haralick texture characteristics. The texture features for each pixel are computed over a moving box of a defined size. In this study, first moment textures have been used, which are defined by first-order histograms representing the rate of occurrence of each grey level within the moving box. Further descriptions of the textures used can be found in (Haralick R.M. & Joo H. A, 1986). The following texture characteristics have been computed: variance, entropy, energy, skewness, kurtosis, coefficient of variation. However to apply this hopeful approach to RS images interpretation, difficulties of informative feature selection should be overcome.

The algorithm of forming feature space by the authors used might be represented in some sequential steps. In the first step the texture characteristics for all bands of RS image with neighborhood size 3x3, 7x7, 10x10 are calculated. In the second step the selection of 5 more informative features are carried out. The selection is performed according to algorithm of informative feature selection based upon criterion of pairwise separability Jeffries-Matusita (*JM*-distance). In common case *JM*-distance is:

$$JM_{ij} = \int_x \left\{ \sqrt{p(x \mid \omega_i)} - \sqrt{p(x \mid \omega_j)} \right\}^2 dx \cdot \qquad (5)$$

In case the data in the processed pairs of samples fit the normal distribution, the following expression is applied:

$$JM_{ij} = 2(1 - e^{-B}),$$

where

$$B = 1/8(\mu_i - \mu_j)^t \left\{\frac{\Sigma_i + \Sigma_j}{2}\right\}^{-1} (\mu_i - \mu_j) + 1/2 \ln\left\{\frac{\left|(\Sigma_i + \Sigma_j)/2\right|}{\left|\Sigma_i\right|^{1/2}\left|\Sigma_j\right|^{1/2}}\right\},$$

at that μ_i, μ_j and $\sum_i, \sum_j$ – parameters of normal distribution (means vectors and covariance matrices) of i and j samples.

The selection of more informative texture features is carried out in the following way. For all obtained textures band the average *JM*-distance is calculated. Then these bands are ranged by the average *JM*-distance from the best to the worst separability and the samples agreements with the normal distribution are defined (to apply parametric density estimation if possible). After that, in the band with the best separability, the worst separating (target) pair of samples is defined. Next, according to the ranged order, another band to this current band is added. First the second band is added, then the third etc. The separability of each two bands is calculated but only for target pairs samples. It significantly increases the computational performance of the best features selection algorithm. Two features with the best separability are taken as the intermediate complete combination and for them again the worst separating (target) pair is defined. Then the addition of band from the ranged bands to the complete bands combination is again performed and the separability of target pair samples is calculated. This time the best combination of three features is defined. This procedure is repeated up to the moment the five best bands to be selected and we obtained final complete combination of features.

Iterative increasing of feature set and calculation of separability only for target pair of samples is needed for the maximum increase of computational performance of the procedure that is very important for nonparametric computation of *JM*-distance according to (5).

ANN: The way of forming feature vector for ANN classification so called context-spectral is differed from the way of forming feature vector one for statistical classification by the significant simplicity.

Each component of the feature vector contains the focal and its neighbor elements from all bands of RS image. Such way of forming feature space allows to consider the interband and pixels correlation (texture information) without special calculation of texture features and feature space optimization.

3.3 Enhanced forecast maps designing by TS analysis algorithm

The procedure of obtaining forecast maps with use of interpreted RS images described in (A.V. Zamyatin & N.G. Markov, 2004), therefore only the main stages are to be discussed:

- determination of CA optimal neighborhood for each classes;
- constructing transition matrix with use of first-order and high-order Markov chains (if we have more then two TS RS images);
- making primary forecast map with the use of obtained transition matrix;
- processing of primary forecast map with use of CA with optimal neighborhood size.

One of the key moments is the determination of CA optimal neighborhood size for each class. It is determined by so called the enrichment factor (Verburg P.H. et al., 2003), which is defined by the occurrence of a land use type in the neighborhood of a location, relative to the occurrence of this land use type in the study area as a whole:

$$F_{i,k,d} = \frac{n_{k,d,i} / n_{d,i}}{N_k / N},$$

$F_{i,k,d}$, characterizes the enrichment of neighborhood d of location i with land use type k. The shape of the neighbourhood and the distance of the neighbourhood from the central grid-cell i is identified by d (for instance d = 1 means grid-cell 3x3). $n_{k,d,i}$ is the number of cells of land use type k in the neighbourhood d of cell i, $n_{d,i}$ the total number of cells in the neighbourhood while N_k is the number of cells with land use type k in the whole raster and N all cells in the raster. The algorithm of enrichment factor calculation is repeated for different neighbourhoods located at different distances (in this case d = 1,2,...,10) from the grid cell to study the influence of distance on the relation between land use types. The average neighbourhood characteristic for a particular land use type l ($\overline{F}_{i,k,d}$) is calculated by taking the average enrichment factors for all grid cells belonging to a certain land use type l, following:

$$\overline{F}_{i,k,d} = \frac{1}{N_l} \sum_{i \in L} F_{i,k,d},$$

where L – the set of all locations with land use type l and N_l the total number of grid-cells belonging to this set. The grid-cell of size d for each class type is fixed in case of maximum of the average enrichment factor. These values are to be used for every class in CA. In most land cover and land use change model first-order Markov chains and only two classified images are used.

4. RESULTS AND DISCUSSION

To investigate the efficiency of proposed algorithms the set of experiment with model and real RS images are carried out.

4.1 Aerospace image models used

In the conducting research model RS images of two types are applied. The multispectral images of the first type are images with implicit texture of classes and arbitrary distribution in classes (Figure 1). The multispectral images of the second type are images with evident texture of classes, designed with use of Brodatz textures (Figure 2).

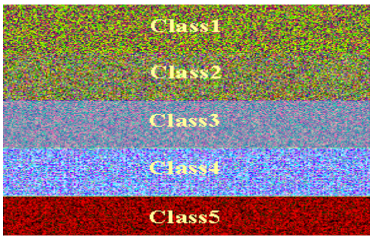


Figure 1. Model image with implicit textures of classes

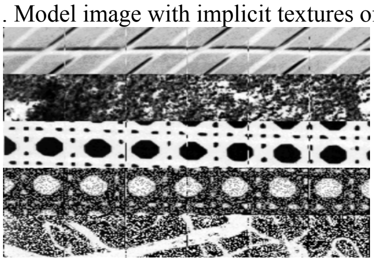
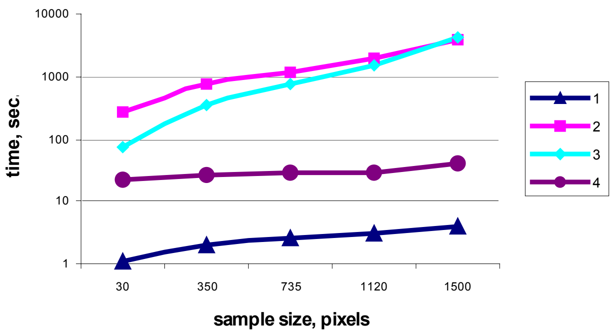


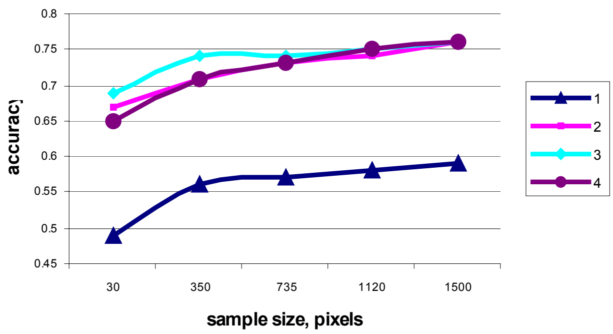
Figure 2. Model image with evident textures of classes

4.2 Nonparametric density estimation

The efficiency investigation of density estimation algorithms is performed for the determination of the computational cost (Figure 3 a)) and the accuracy (Figure 3 b)) of classification, on the example of seven bands image of type as shown in Figure 1.



a)



b)

Figure 3. Computational cost and accuracy of classification algorithms used

In Figure 3 the following symbols are defined: 1 – using of traditional maximum likelihood classification, 2 – using of ordinary RP algorithm based upon (2), 3 – using of ordinary k-NN algorithm based upon (3) and (4), 4 – using the proposed original statistical nonparametric algorithm.

Figure 3 a) shows the original algorithm provides computational performance increase in dozens times compared to traditional nonparametric density estimation algorithms. At that the performance of original algorithm is increasing together with the increasing of sample size.

At the same time Figure 3 b) shows the accuracy of the proposed algorithm is almost the same as the accuracy of traditional nonparametric algorithms, and also it should be noted the accuracy with use of parametric algorithm of density estimation is inappropriate low. It proofs necessity of developing nonparametric algorithms that are invariant to the distribution in a sample.

4.3 Classification with spatial features

The important element of the advanced interpretation is the classification of RS images with use of texture features of the classes that is needed for high accuracy interpretation. In the framework of the developing approach it is proposed to define prior probabilities of classes by either statistical or ANN ways and each way takes into account the spatial features of classes.

The efficiency investigation of some algorithms with use of different types of model images is conducted. The purpose of the research is to define how the proposed ways of forming spatial feature space consider texture information about classes.

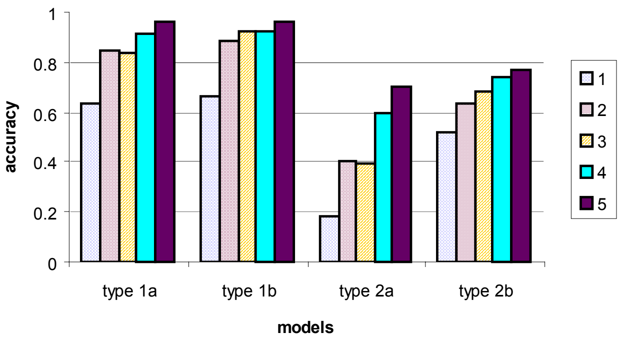


Figure 4. The classification accuracy of statistical and ANN algorithms with different types of model images used

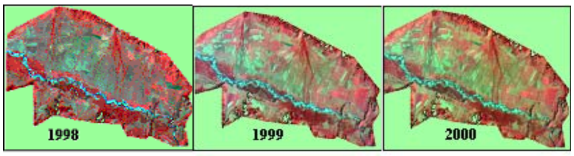
The Figure 4 shows the investigation results of accuracy classification for different ways of forming feature space.

In Figure 4 the following symbols are defined: 1 – traditional maximum likelihood classification, 2 – ordinary RP algorithm based upon (2), 3 – ordinary k-NN algorithm based upon (3) and (4), 4 – the proposed ANN original algorithm using context-spectral way of forming feature space, 5 – the proposed nonparametric algorithm using Haralick texture characteristics (Haralick R.M. & Joo H. A, 1986). The following types of model images on the abscissa axis are scaled: type1a and type1b – the first type model images (Figure 1) with three and seven bands consequently; type2a and type2b – the second type model images (Figure 2) with only one and six bands consequently.

Figure 4 presents the property to get more accurate results of the ANN classification with context-spectral forming feature space due to considering texture features in multispectral images. The noticeable effect of that in case of ANN classification with type2a image when spectral information almost is absent (the one band only), and classes textures are strongly marked.

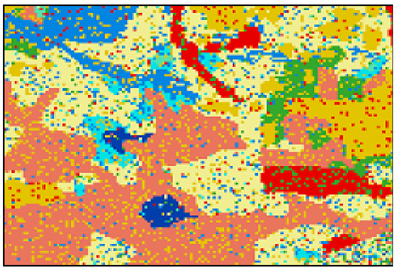
At the same time Figure 4 presents sufficient advantages of proposed complex nonparametric algorithm, which is based upon first-order Haralick texture characteristics and nonparametric optimization of feature space.

Besides, the investigation on the basis of time series mutlispectral RS images for the Uymon steppe area are conducted. The spatial resolution of RS data is 30 meters. The major classes of the RS images are the agricultural land, bushes, water, and vegetation.

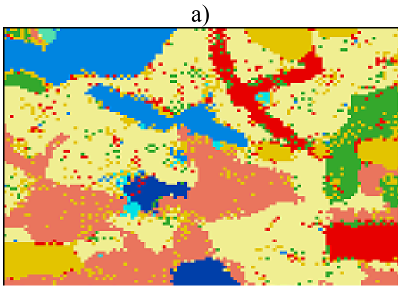


a) b) c)

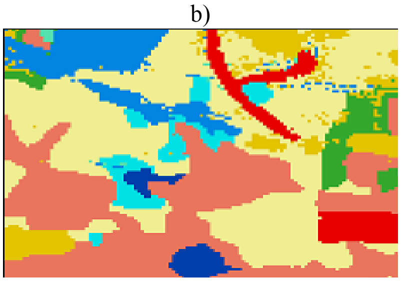
Figure 5. Time series RS images from RESURS-O1



a)



b)



c)

Figure 6. The fragment of forecast map designed for 2001 by various techniques

Figure 5 shows the time series RS mages for 1998, 1999 and 2000. Figure 6 shows the fragments of forecast maps for 2001 obtained by three various techniques: a) by the traditional maximum likelihood classification without CA; b) by the traditional maximum likelihood classification and ordinary CA; c) by the proposed algorithms of advanced interpretation and enhanced TS analysis.

Figure 6 shows that the best forecast map is in c) fragment. It is a less noisiness and it has more legible classes that could say about availability of the proposed algorithms for designing forecast maps.

CONCLUSIONS

Permanent improving of comprehensive satellite systems allow to obtain more qualitative RS images of high resolution, which might be effectively used for designing of thematic and forecast maps. Increasing amount of information requires the specialized processing, RS and mapping software, forecasting land use/cover change models.

The approach to the designing algorithms for advanced interpretation of RS images and on their basement including GIS designing forecast maps is developed in the paper. The proposed approach is based upon nonparametric statistical and ANN classification using spectral and spatial features. To make the forecasting more adequate it is proposed the approach with use of high-order Markov chains and CA with optimal neighborhood size.

The preliminary investigation results conducted with use of model RS images and TS RS images for Uymon steppe area, show high efficiency of the proposed approach and availability of further research in that field.

References

A.V. Zamyatin, N.G. Markov, 2004. Multi-stage Processing of Time Series Aerospace Images for Obtaining Enhanced Forecast Land Cover Maps. In: *Proceedings of the 7th AGILE Conference on Geographic Information Science*, Heraklion, Greece, AGILE, in press.

Baker W. L., 1989. A review of models of landscape change. *Landscape Ecology 2*, pp. 111–133.

Haralick R.M., Joo H. A, 1986. Context Classifier, *IEEE Transactions on Geoscience and Remote Sensing*, N24, pp. 997-1007.

J. R. Eastman, 2001. Idrisi32 release 2 tutorial, http://www.clarklabs.org, (accessed 1 Nov. 2003).

N.G. Markov, A.A. Napryushkin, A.V. Zamyatin, 2003a. Advanced Thematic Mapping Approach for Forecasting Landscape Change Using GIS. In: *Proceedings of the 6th AGILE Conference on Geographic Information Science*, Lyon, France, AGILE, pp. 687-693.

N.G. Markov, A.A. Napryushkin, A.V. Zamyatin, E. V. Vertinskaya, 2003b. Adaptive Procedure of RS Images Classification with Use of Extended Feature Space. In: *Proceedings of SPIE*, Vol. 4885, pp. 489-500.

Park, S. and Wagner, 1997. D. F., Incorporating Cellular Automata simulators as analytical engines in GIS, *Transaction in GIS*, vol.2, no. 3, pp. 213-231.

Verburg P.H. et al., 2003. A method to analyse neighbourhood characteristics of land use patterns, *Computers, Environment and Urban Systems*, in press.

Acknowledgements

The authors are grateful for the support this research by grant №03-07-90124 from Russian Foundation for Basic Research.