10
The MLRP-method for Analysis of Some Problems in Climate and Seismology

10.1 Introduction

The problem of detection correlations and regularities by data, which is presented by time series or data table, is used in different scientific domains. We have so problem in sociology, medicine, geology and seismology, economics and climate, etc. [Box and Dgenkins, 1974]. The problem of detection correlations and regularities by data is the inverse and ill-defined problem; therefore, we have to consider every possible information about studied object or process in order to raise solution stability. This information first of all is system of characteristic of studied object: multidimensional and heterogeneity characteristic's space; "noisy" and "duplicate" characteristics; insufficient or excess and missing sample data.

We have the most difficult problem, when the priory information about process or object is absent. In addition to that several attendant problems are appeared. Therefore developing new method, we have to consider all above. Firstly, it is necessary to define a class of decision functions (models). Secondary, we must pick out a method of plotting optimal decision function, in other words to define optimality criterion by sample. In the third place, we must to test our model on adequacy and effectiveness (capacity for general conclusion or statistical stability).

It is very important to lean against practical and scientific experience of researchers when solving out practical problems. Experts in the certain area own the big share of the aprioristic information and have wide experience in interpretation of results of decisions of practical problems. It is very important in working out and testing of methods. Here it is important to underline an urgency of different countries scientists association at the decision of the important common to all humanity problems, for example, the risk estimation of occurrence the extreme situations due to man-caused and natural accidents. Unfortunately, availability of data occasionally is the exclusive property of the country or separate business. Therefore, association of scientists and experts of the different countries in work on the general projects is simply common to all humankind necessity.

Method presented here depends on statistical information about studied object and can be used in many domains of data mining also. That approach is relied to intelligent analysis to a greater extent, since it uses expert and apriory information generally.

At present time, there is joint initiative of the European Commission and European Space Agency, which aims at achieving an autonomous and operational Earth observation capacity. It is named Global Monitoring for Environment and Security (GMES) (http://www.gmes-bg.org/). The main objective of GMES is to monitor and better understand our environment and to contribute to the security of every citizen. Land, sea, and atmosphere – each Earth component is observed through

GMES, helping to make our lives safer. We think that our investigations are included to GMES-project properly.

There are many well-known scientific schools, what make researches to that line of investigation [Lukashin, 2003], [Bezruchko and Smirnov, 2003], [Lbov and Starceva, 1999]. However, universal method is not exists. Several suppositions and apriory information are used by every method. It says that problems are actually. Than method is preferred if it uses lame suppositions to respect with decision function class and has capability for change-over model of correlation during learning on sample data. Current methods use neural-network technological, pyramidal-network, wavelet analysis, logical structures and others approaches. Such methods we can name so as adaptive methods. A conception of adaptive has a more comprehensive sense [Lukashin, 2003], [Lbov and Starceva, 1999].

We will interpret concept of adaptive so as consecutive selection of model's structure in during process of learning on sample data in order to take effective prediction by time series. At the same time, it is appeared additional problem – detection time moment of changing model's structure (criterion of adaptive).

We suggest two ways to joint analysis of several univariate time series by using MLRP-method. When the information about event is kept safe in the process, and when it is kept safe in depending process. That method was applied to prediction of multivariable heterogeneous time series [Stupina and Lbov, 2006]. The solving of practice actual problems from hydrological and seismology domains are presented here by MLRP-method.

In a hydrological problem, the most informative signs influencing volume of a reservoir of water in the rivers is investigated. Probable catastrophic consequences from occurrence of high waters or a drought are obvious; therefore, the problem is actual for many regions of a considerable quantity of the countries on the Earth. The second problem is not less actual in global measurement. It consists in research of probable regional migration of active seismic zones. The problem of a short-term prediction of earthquakes remains important on present time. Therefore, the majority of problems of modern seismology are focused on studying of process of earthquake in order to come little nearer to the decision of problems of the forecast.

Model from the logical decision function class we will name as a logical-and-probabilistic correlation [Lbov and Starceva, 1999], [Stupina, 2005].

Let us note below priority properties of the logical decision function class at the solving of inverse problems:

- easily interpreted decision is constructed. It is especially important at the decision of difficultly formalizable problems, at finding-out of cause-and-effect interrelations in object, for the organization of a dialogue mode of forecasting etc.;
- the given class allows to process heterogeneous experimental data (without reduction all signs to one scale) and is invariant to admissible transformations of signs scales;
- it has a small measure of complexity. At the same time, this class is "rich enough" for the effective decision of the applied problems characterized by high level of aprioristic uncertainty about a functional kind of distribution. Mathematical properties of the given class are proved and brought in this chapter;
- gives the chance to realize simple optimizing procedures of searching the selective decision function;

– allows to work in the presence of admissions in empirical tables and to receive adequate enough decision at small volumes of sample.

10.2 MLRP-method of creating logical-and-probabilistic model

10.2.1 Basic notations

From the beginning, we consider a commonly probabilistic statement problem. Let the value (x,y) is a realization of a multidimensional random variable (X,Y) on a probability space $<\Omega,\mathrm{B},\mathrm{P}>$, where $\Omega = D_X \times D_Y$ is μ-measurable set (by Lebesgue), B is the borel σ-algebra of subsets of Ω, P is the probability measure (we will define such as c, the strategy of nature) on B, D_X is heterogeneous domain of under review variable, $\dim D_x = n$, D_y is heterogeneous domain of objective variable, $\dim D_y = m$. The given variables can be of arbitrary types (quantitative, ordinal, nominal). For the pattern recognition problem, for example, the variable Y is nominal. Let us put $\Phi_\circ$ is a given class of decision functions. Class $\Phi_\circ$ is μ-measurable functions that puts some subset of the objective variable $E_y \subseteq D_Y$ to each value of the under review variable $x \in D_X$, i.e. $\Phi_\circ = \{f : D_x \to 2^{D_y}\}$. For example, the domain E_y can contain the several patterns $\{\omega_1,...,\omega_k\}$ for pattern recognition problem.

The quality $F(c,f)$ of a decision function $f \in \Phi_\circ$ under a fixed strategy of nature c is determined as $F(c,f) = \int\limits_{D_x}(P(E_y(x)/x) - \mu(E_y(x)))dP(x)$, where $E_y(x) = f(x)$ is a value of decision functions in x, $P(y \in E_y(x)/x)$ is a conditional probability of event $\{y \in E_y\}$ under a fixed x, $\mu(E_y(x))$ is measurable of subset E_y. Note that if $\mu(E_y(x))$ is probability measure, than criterion $F(c,f)$ is distance between distributions. If the specified probability coincides with equal distribution than such prediction does not give no information on predicted variable (entropy is maximum). On the nominal-real space $\Omega = D_н \times D_в$ a measure μ is defined so as any $E \in \mathrm{B}$, $E = \bigcup_{j=1}^{|E_н|} E_в^j \times \{z^j\}$, $\mu(E) = \sum_{j=1}^{|E_н|} \frac{\mu(E_в^j)}{|D_н|\mu(D_в)}$, were $E_н$ is projection of set E on nominal space $D_н$, z^j - item of $E_н$, $E_в^j$ - set in $D_в$ corresponding to z^j, $\mu(E_â^j)$ – Lebesgue measure of set $E_в^j$. For any subset of domains D_X or D_Y the measure μ is assigned similarly. Clearly, the prediction quality is higher for those E_y whose measure is smaller (accuracy is higher) and the conditional probability $P(y \in E_y(x)/x)$ (certainty) is larger. For a fixed strategy of nature c, we define an optimal decision function $f_\circ(x)$ as such as $F(c,f_\circ) = \sup_{f\in\Phi_\circ} F(c,f)$, where $\Phi_\circ$ is represented above class of decision functions.

In commonly when we solve this problem in practice the size of sample is small and type of variables is may be different type. In this case is used class of logical decision function Φ_M complexity M [Lbov and Starceva, 1999]. For the prediction problem of the heterogeneous system variables class Φ_M is defined as $\Phi_M = \{f \in \Phi_\circ \mid f \sim <\alpha,r(a)>,\alpha \in \Psi_M, r(\alpha) \in R_M\}$ (the mark '~' denotes the correspondence of pair $<\alpha,r(a)>$ to symbol f), were Ψ_M is set of all possible partitioning $\alpha = \{E_X^1,...,E_X^M \mid E_X^t = \prod_{j=1}^n E_{X_j}^t,\ E_{X_j}^t \subseteq D_{X_j}, t = \overline{1,M}, \bigcup E_X^t = D_X\}$ of domain D_X on M noncrossing subsets, R_M is set all possible decisions $r(\alpha)=\{E_y^1,...,E_y^M \mid E_y^t \in \Im_{D_Y}, t=\overline{1,M}\}$, $\Im_{D_Y}$ – set of all possible m-measuring intervals. For that class the measure $\mu(E_y(x)) = \frac{\mu(E_y)}{\mu(D_Y)} = \prod_{j=1}^{m}\frac{\mu(E_{y_j})}{\mu(D_{y_j})}$ is the normalized measure of subset E_y and it is introduced with taking into account the type of the variable. The measure $\mu(E_y(x))$ is measure of interval, if we have a variable with ordered set of values and it is quantum of set, if we have a nominal variable (it is variable with finite non-ordering set of values and we have the pattern recognition problem). A complexity of Φ_M class is assigned as M if we have invariante prediction (decision is presented by form: if $x \in E_X^t$, than $y \in E_y^t$), $M_\Phi = M$, and it is assembly $(k_1,...,k_M)$ if we have multivariate, i.e. $E_y^t = \bigcup_{i=1}^{k_t} E_y^i$, $t=1,...,M$ and $E_y^i \cap E_y^j = \varnothing$ for $i \neq j$ (decision is presented by form: if $x \in E_X^t$, than $y \in E_y^1 \vee E_y^2 \vee ... \vee E_y^{k_t}$). The class of logical decision function has universal property.

10.2.2 Properties of the method

Statement. For any function $f \in \Phi^\circ$ and $\varepsilon > 0$ there is M and several logical decision function $f_M \in \Phi_M$ so that $|F(c,f) - F(c,f_M)| \le \varepsilon$.

Others good properties of the logical decision function class are presented in work [Lbov and Stupina, 2002], [Stupina, 2006] for prediction system heterogeneous variables problem.

Definition 1. The strategy of nature is c={p(x,y)=p(x)p(y/x))}, where a conditional probability p(y/x) is specified for any elements on B.

Theorem. For a fixed type of the predicate, the class Φ_M of logic decision functions is a universal class in the problem of prediction multivariate heterogeneous value by criterion F(c,f), i.e. for any strategy of nature c and any $\varepsilon > 0$ there exists a number M (M=1,2,3,…) and for some logical decision function f $\in \Phi_M$ (it is represented in the form of decision tree on M vertices) such that $\left|F(c,f) - F(c,f_\circ)\right| \le \varepsilon$, where $f_\circ$ is optimal function in class $\Phi_\circ$.

The proof of this theorem readily follows from the property of μ-measurability and P-measurability of space D and its projections on the space D_X, D_Y correspondingly. This theorem for pattern recognition of two patterns was proofed in work [Lbov and Starceva, 1994].

We can introduce a complexity of distribution (strategy of nature c) using the class logical decision function. It is necessary for solving statistical stability problem of decision function [Lbov and Stupina, 2003].

Statement 1. For any nature strategy c the quality criterion $F(c,f)$ (risk function) of logical decision function f belonging to Φ_M is presented by following expression:

$$F(c,f) = \int_{D_X}\int_{D_Y}(1 - L(y,f(x)))p(x,y)dxdy = \sum_{t=1}^{M} p_x^t(p_{y/x}^t - \mu^t),$$

where the loss function $L(y,f)$ such as $L(y,f) = \begin{cases} p_o, & y \in \beta \\ 1+p_o, & y \notin \beta \end{cases}$, $p_o = \mu(E_Y^t)$, $\beta = f(\alpha), \alpha \in \Psi_M$.

Definition 2. To each subclass Φ_M we put in correspondence the subset $L_\varepsilon(M)=\{c:\exists f\in\Phi_M, \left|F(c,f)-F(c,f_\circ)\right|\le\varepsilon\}$ of nature strategies; ε is an arbitrarily small number determining an admissible error level of this subset of strategies, where $f_\circ$ is optimal function in class $\Phi_\circ$.

The complexity measure of each subset $L_\varepsilon(M)$ is defined as the complexity measure of the corresponding subclass of decision functions: $\nu(L_\varepsilon(M))=\nu(\Phi_M)=M$. Accordingly, the nature strategies c belonging to $L_\varepsilon(M)$ has complexity measure M. The important statement follows from this theorem and definition.

Statement 2. The set of all possible strategies can be ordered according to complexity, i.e. $L_\varepsilon(1)\subset L_\varepsilon(2)\subset\ldots\subset L_\varepsilon(M)\subset\ldots\subset L_\circ$, and $\varepsilon^{M+1}\le\varepsilon^M$, where $\nu(L_\varepsilon(M))$=M is the complexity and ε^M is the admissible error level of the strategy class $\nu(L_\varepsilon(M))$.

We can suppose that the true (optimal) decision function belongs to Φ_M it is followed from this statement 1.

Definition 3. Define a nature strategy c_M (generated by logical decision function $f\in\Phi_M$) such as set of parameters satisfying the following conditions:

1) $\sum_{t=1}^{M} p_x^t = 1$,

2) $P(E_Y^t/E_X^t)=p_{y/x}^t$ (conditional distribution is same for any $x\in E_X^t$ and $y\in E_Y^t$),

3) $P(\bar{E}_Y^t/E_X^t)=1-p_{y/x}^t$,

where $E_X^t\in\alpha$, $E_Y^t\in\beta$, $<\alpha,\beta>\sim f\in\Phi_M$. The complexity of this strategy is M, i.e. $\nu(c_M)=M$. Note that c_M generated by logical decision function belongs to class $L_\varepsilon(M)$. Clearly, the decision function that generated this strategy is optimal function in class Φ_M.

Statement 3. For a fixed nature strategy $c_M\in L_\varepsilon(M)$ of complexity M, the quality criterion $F(c_M,\tilde{f})$ (risk function) of logical decision function $\tilde{f}\in\Phi_{M'}$ of complexity M' is presented in following form:

$$F(c_M,\tilde{f})=F(\tilde{\alpha})=\sum_{t'=1}^{M'}\tilde{p}_x^{t'}\rho^{t'}=\sum_{t'=1}^{M'}\tilde{p}_x^{t'}(\tilde{p}_{y/x}^{t'}-\mu_y^{t'}),$$

where $\tilde{p}_x^{t'}=P(x\in\tilde{E}_X^{t'})=\sum_{t=1}^{M}p_x^t\frac{\mu(\tilde{E}_X^{t'}\cap E_X^t)}{\mu(E_X^t)}$,

$$\tilde{p}_{y/x}^{t'}=\frac{1}{\tilde{p}_x^{t'}}\sum_{t=1}^{M}p_x^t\frac{\mu(\tilde{E}_X^{t'}\cap E_X^t)}{\mu(E_X^t)}\left(p_{y/x}^t\frac{\mu(\tilde{E}_Y^{t'}\cap E_Y^t)}{\mu(E_Y^t)}+(1-p_{y/x}^t)\frac{\mu(\tilde{E}_Y^{t'})-\mu(\tilde{E}_Y^{t'}\cap E_Y^t)}{1-\mu(E_Y^t)}\right).$$

Remark. If the nature strategy c_M such that some subset E_Y^t coincides with the space D_Y, than

$$\tilde{p}_{y/x}^{t'}=\frac{1}{\tilde{p}_x^{t'}}\sum_{t=1}^{M}p_x^t\frac{\mu(\tilde{E}_X^{t'}\cap E_X^t)}{\mu(E_X^t)}p_{y/x}^t\frac{\mu(\tilde{E}_Y^{t'}\cap E_Y^t)}{\mu(E_Y^t)}.$$

It is followed from that $p_{y/x}^t=P(D_Y/E_X^t)=1$, $\mu(D_Y)=1$.

Consequence 1. If the decision function $\widetilde{f}$ belonging to Φ_M coincides with the function f belonging to Φ_M, than $F(c,\widetilde{f}) = F(c,f)$.

Consequence 2. For the decision function $\widetilde{f}$ belonging to $\Phi_{M'}$ we have the expression $P(\overline{\widetilde{E}}{}_Y^{t'} / \widetilde{E}_X^{t'}) = 1 - \widetilde{p}_{y/x}^{t'}$.

Really, it is follows from the statement 3, where $\dfrac{\mu(\overline{\widetilde{E}}{}_Y^{t'} E_Y^t)}{\mu(E_Y^t)} = \dfrac{\mu(E_Y^t) - \mu(E_Y^t \widetilde{E}_Y^{t'})}{\mu(E_Y^t)}$,

$\dfrac{\mu(\overline{\widetilde{E}}{}_Y^{t'} \overline{E}_Y^t)}{\mu(\overline{E}_Y^t)} = \dfrac{1 - \mu(E_Y^t) - \mu(\widetilde{E}_Y^{t'}) + \mu(E_Y^t \widetilde{E}_Y^{t'})}{1 - \mu(E_Y^t)}$.

Consequence 3. If we have M=1 and the optimal function f generating c_1 such that $E_Y^1 = D_Y$, than $F(c_1,f) = 0$.

Really, for the express of criterion we have $F(c,f) = \sum_{t=1}^{M} \left(P(E_X^t E_Y^t) - P_o(E_Y^t) \right) = P_o(D_X D_Y) - P_o(D_Y) = 0$. It means that we have the event distribution in D for the nature strategy of the complexity M=1. It is case when the entropy is maximum.

Consequence 4. If we have M=1 and the optimal function f generating c_1 such that $E_Y^1 = D_Y$, than for any decision function $\widetilde{f} \in \Phi_{M'}$ the criterion $F(c_1,\widetilde{f}) = 0$.

Really, $\widetilde{p}_{y/x}^{t'} = \dfrac{\mu(\widetilde{E}_Y^{t'} D_Y)}{\mu(D_Y)} P_o(D_Y / D_X) = \mu(\widetilde{E}_Y^{t'})$, $\widetilde{p}_x^{t'} = \dfrac{\mu(\widetilde{E}_Y^{t'} D_X)}{\mu(D_X)} P_o(D_X) = \mu(\widetilde{E}_X^{t'})$,

$F(c_1,\widetilde{f}) = \sum_{t'=1}^{M'} \mu(\widetilde{E}_X^{t'})(\mu(\widetilde{E}_Y^{t'}) - \mu(\widetilde{E}_Y^{t'})) = 0.$

Consequence 5. If the decision function $\widetilde{f}$ belongs to Φ_1 and $\widetilde{E}_Y^1 = D_Y$, than we have $F(c_M,\widetilde{f}) = 0$ for any complexity $M \geq 1$.

Really, we have $\widetilde{p}_x = \sum_{t=1}^{M} p_x^t \dfrac{\mu(D_X E_X^t)}{\mu(E_X^t)} = 1$, $\widetilde{p}_{y/x} = \sum_{t=1}^{M} p_x^t \left(p_{y/x}^t \dfrac{\mu(D_Y E_Y^t)}{\mu(E_Y^t)} + (1 - p_{y/x}^t) \dfrac{1 - \mu(D_Y E_Y^t)}{1 - \mu(E_Y^t)} \right) = 1$.

If the strategy of nature is unknown the sampling criterion $F(\bar{f})$ is used by method $Q(v_N)$ of constructing sample decision function $\bar{f}$, $\overline{F}(\bar{f}) = \sum_{t=1}^{M'} \bar{p}_x^t (\bar{p}_{y/x}^t - \bar{\mu}_y^t)$, were $\bar{p}_x^t = \frac{N(\tilde{E}_X^t)}{N(D_X)} = \frac{N^t}{N}$, $\bar{p}_{y/x} = \frac{N(\hat{E}_y^t)}{N(\tilde{E}_X^t)} = \frac{\hat{N}^t}{N^t}$, $\bar{\mu}_y = \mu(\hat{E}_y)$, $N('*')$ is number of sample points, generating the set "*", $\bar{f} \sim <\alpha, r(\alpha)>$, $\alpha = \{\tilde{E}_X^1, ..., \tilde{E}_X^{M'}\} \in \Psi_{M'}$, $r(\alpha) = \{\hat{E}_y^1, ..., \hat{E}_y^{M'}\} \in R_{M'}$. The optimal sample decision function is $\bar{f}^* = \arg\max_{\alpha \in \Psi_{M'}} \max_{r(\alpha) \in R_{M'}} \overline{F}(\bar{f})$. In order to solve this extreme problem we apply the algorithm MLRP of step-by-step increase attachments of decision trees. It do the branching of top point on that value criterion $\overline{F}(\bar{f})$ is maximum and the top point is divisible or $\overline{F}(\bar{f}) \geq F^*$. The top point is indivisible if 1) number of final top point is $M' = M^*$ or 2) $\hat{N}^t \leq N^*$. That criterion and parameters F^*, M^*, N^* assign method of constructing sample decision function.

In order to estimate the MLRP – method quality we did statistical modeling. The average of the criterion of sample decision function on samples of fixed size $m_F(c) = \mathrm{E}_{V_N} F(c,\bar{f})$ is estimated for fixed nature strategy. Moreover we researched the averaging-out empirical functional quality $\varepsilon_N(c) = \mathrm{E}_{V_N} F(c,\bar{f}) - \mathrm{E}_{V_N} \bar{F}(\bar{f})$ for given strategy of nature with the purpose of estimating decision quality, and maximal removal of empirical functional quality average of distribution $\varepsilon_N^*(c) = \sup\limits_{c:\bar{F}(\bar{f})=F_0} \varepsilon_N(c)$ for a given empirical quality value $F_\circ$. It was taken for some parametric nature strategy class, for given nature strategy complexity M, decision function complexity M'. The decision function is builder by MLRP-method on sample of size N. Parameters n, m (dimensions of domains D_X and D_Y) and a quantity of fixed type variables was considered in problem statement overall. It is defined the complexity of nature strategy and decision function in addition. The GenMLRP-algorithm was developed for modeling nature strategy parameters. Generation nature strategies were realized in accordance with definition, were parameters is established by random in the given interval. The properties of functional quality are presented in work [Stupina, 2006] for uniform distribution on set D_Y. Such approach to research of statistical stability was used in earlier works, but for simple and one-dimensionality cases [Raudis, 1976], [Startseva, 1995], [Berikov, 2002], [Lbov and Stupina, 1999]. The MLRP-method was applied for prediction multivariate time series. Three random processes were simultaneously considered instead of one. The feature systems (under review and predicted) was established. Procedure of building data table is offered in work [Stupina and Lbov, 2006]. The example of solving practical problem is presented in next paragraph.

10.2.3 Algorithm description

The MLRP- algorithm of construction of logical-and-probabilistic model algorithm carries out consecutive construction of fragmentation space D_X on two subspaces and representation of the decision in the form of a tree. Each fragmentation space let's compare with own node b. To initial fragmentation $\alpha = \{D_X\}$ there corresponds node b°, a set of corresponding decisions $r(\alpha) = \{D_Y\}$ and sample v of volume N.

Step 1. At the given stage on initial sample v of volume N for fixed fragmentation α search of the best set of decisions $r(\alpha) = \{E_y^1\}$ on empirical criterion of quality is carried out.

Procedure of consecutive truncation of space D_Y to cover $\hat{E}_y^1$ makes consecutive removal of "extreme" points of sample until value of criterion of quality will not start to decrease. Thus, for fragmentation α the maximum value of empirical criterion $\bar{F}(\bar{f}) = \bar{F}(\alpha)$ equal to $\bar{F}^1$ is received. Corresponding to the given value of criterion decision function is $\bar{f} \sim <\alpha, r(\alpha)>$, $\alpha \in \Psi_1(\nu)$, $r(\alpha) \in R_1(\nu)$, where $\Psi_1(\nu)$ – a class of fragmentations of complexity one, constructed on sample v, $R_1(\nu)$ – corresponding set of decisions.

Step 2. On the given step construction of fragmentation of initial area on two subspaces is carried out, i.e. branching of initial node (step 1) of which initial fragmentation consists, on two final nodes. If the node is indivisible, the algorithm finishes work and result is the fragmentation constructed on the previous step. If node is divisible then search on all variables $X_1,...,X_j,...,X_n$ is consistently made.

For each variable X_j various fragmentations of set D_{X_j} into two subsets E_{X_j} and $\bar{E}_{X_j}$ such, that $E_{X_j} \neq \varnothing, E_{X_j} \neq D_{X_j}$, where E_{X_j} – any subset for nominal variables, $E_{X_j} = \left\{x_j \mid x_j \leq \gamma\right\}$ – for serial and real variables, $\gamma \in \{\gamma_1,...,\gamma_N\}$, $\gamma_i \in R$ are considered. Let's designate set of every possible fragmentations α of space D_X on variable X_j on two subspaces through $\Psi_2^j(v)$, i.e. $\alpha \in \Psi_2^j(v)$, $\alpha = \left\{E_X^1, E_X^2\right\}$, $E_X^1 = D_{X_1} \times \cdots \times E_{X_j}^1 \times \cdots \times D_{X_n}$, $E_X^2 = D_{X_1} \times \cdots \times \bar{E}_{X_j}^1 \times \cdots \times D_{X_n}$. Fragmentation $\alpha \in \Psi_2^j(v)$ is compared with final nodes b^1, b^2 and decisions $r(\alpha) = \left\{E_Y^1, E_Y^2\right\}$, $r(\alpha) \in R_2(\nu)$ corresponded to them. Thus, object a_i belongs to node b^t, $t = 1,2$, if point $x^i \in E_X^t$ corresponds to its description $X(a_i)$, where $E_X^1 = \left\{x \mid x_j \in E_{X_j}\right\}$, $E_X^2 = \left\{x \mid x_j \in \bar{E}_{X_j}\right\}$. Further fragmentation α is fixed and search of the best decision $r(\alpha) = \{\hat{E}_y^1, \hat{E}_y^2\}$ such that criterion $\bar{F}(\bar{f})$ accepted the maximum value is made. After full search on all variables $j = 1,...,n$ (i.e. on all fragmentations $\alpha \in \Psi_2^j(v)$) and on probable decisions $r(\alpha) \in R_2(v)$ it is found the best fragmentation $\alpha^* \in \Psi_2^{j_1}(v)$ of space D_X and the best decision $r^*(\alpha^*) \in R_2(v)$ corresponding to it.

Let for fragmentation α^* on variable X_{j_1} the empirical criterion of quality has maximum value $\bar{F}^2$ then the decision on branching of initial node b^o is made. If $\bar{F}^2 > \bar{F}^1$ we fix branching and the initial table (sample) v breaks on two tables v^1 and v^2. Table v^1 includes realizations $x^i \in E_X^1$, table v^2 – realizations $x^i \in E_X^2$. These tables are put in conformity to nodes b^1 and b^2, as a result we receive a tree consisting of two nodes. To nodes b^1 and b^2 we give the status of initial nodes. If condition $\bar{F}^2 > \bar{F}^1$ is not carried out, initial node b^o is left former, we appropriate to it the status of final node and exit from a step 2.

Step 3. The step 2 repeats for each initial node. The stop-condition of algorithm is indivisibility of final node or achievement on the given step the specified number of final nodes M^* (entrance parameter of algorithm).

Node b^t is considered final and not subject to division, if number of selective points N^* in the given node is less than some entrance parameter N^* defining the minimum admissible number of objects in node.

The choice of the best values of parameters N^*, M^* is connected with the question of estimation of quality of a method of construction of logical-and-probabilistic model on the limited sample.

Algorithm MLRP is software implemented in programming language Microsoft Visual C ++. In an interface window, it is necessary to specify a full path to file with entrance data and the file name where found solving rule in the form of a set of laws will be written. In a data file, it is necessary to specify:

- dimension of space of entrance variables;
- dimension of space of target variables;
- volume of sample of training;
- describe types of all sets of variables and a range of their change (binary, nominal, continuous, discrete, serial);

– enter the table of data according to order of the description of variables.

In an interface window parameters N^* – the minimum number of points in tree node, M^* – the maximum number of dangling nodes of a tree of decisions, Contin Grid Step – a step of sampling of an interval for a continuous variable, Infimum of Pyt – is minimum admissible estimation of number of the points which belongs to given space at fixed fragmentation are assigned.

Algorithm MLRP carries out the directed search of the best decision, therefore is locally optimum.

10.3 Application MLRP-method to prediction multivariate time series problem

10.3.1 The MLRP-method to prediction multivariate time series problem

Let us consider terminal time series $\{x(t), t \in T\}$, that it is realization of sometime-dependent random process η(t). One is supposed that simultaneous distribution $p(\eta_1)$, $p(\eta_1, \eta_2)$, $p(\eta_1, \eta_2, \eta_3)$,…,$p(\eta_1,\ldots, \eta_T)$ is exist. The value set $D_{\eta(t)}$ of variables may be quantitative, nominal, and ordinal type in a more case. Let the values of random process η(t) are measured at consequent moments of the time with the gap $\Delta t = t_k - t_{k-1}$. Denote this set of moments as $T = \{t_1,\ldots,t_k,\ldots,t_N\}$, N << ∞.

Classical problem of prediction time series is consist in that we must to take predict at time moment $t = t_R$ on time period $t_{R+\tau}$ by analyzing prehistory $\{x(t_k)\}$, $k = 1,\ldots,d$, with length d. As a rule, the value τ is named as forestalling. The set of every possible all prehistory, that have length d denote as D_X, and the set of every possible all sets forestalling denote as D_Y. Let us understand a prediction decision function as a f mapping of the D_X set on the D_Y set, i. e. $f: D_X \to D_Y$, $\dim D_X = d$, $\dim D_Y = \tau$. Model's construction f of prediction is defined by decision function class Φ.

If a simultaneous distribution is known than optimal decision function, constructing predict to time $t + \tau$, is conditional average of distribution $E(\eta_{t+\tau} \mid \eta_{t-d},\ldots,\eta_t)$. In order to solve this problem it is necessary to restore conditional distribution. However, that way is not practical because we have not enough size of sample in applied tasks. Therefore, it is possible to offer a different depending on specified suggestions targets setting (concerning properties of random process) and the different methods (concerning decision function class) of their decision accordingly.

At present time it was developed many method for prediction depended of time random process (probabilistic characteristics of process are not changed on time). Its methods are based on constructing several models, which usually use some suggestion. For example, if we want to do long-time prediction than the best offer (concerning error variance value) is global model, if we want to do short-term prediction, than it is local model [Bezruchko and Smirnov, 2003]. Note that most of models accomplish solitary prediction. It is next time $t + \Delta t$ or at time moment $t + \tau\ \Delta t$, $\tau = 2,\ldots,N-d$.

We propose model, that accomplishes prediction on all forestalling term τ, in other words, to time moments $t + k\ \Delta t$, $k = 1,\ldots,\tau$. That prediction allows to take one decision function (structure of model) and to do simultaneously several predictions on future by one prehistory.

For that problem statement, it is important to researcher several steps:

– detection time moment of changing model's structure (criterion of adaptive);

– optimization of prehistory length d;
– optimization of forestalling term τ .

In order to solve these items we will use class of logical decision function. We will consider two ways:

a) when the information about event is kept safe in the process, and
b) when it is kept safe in depending process.

We will perform the primary ideas of these ways in following paragraphs.

✓ Analysis unvariate time series problem

Let us we have univariate time series { $x(t)$ } of any random process $\eta(t)$. It is necessary to problem of constructing predict function f by empirical data, which is presented as terminal points N for given prehistory length d and forestalling term τ. We will construct decision function from the logical decision function Φ_M by sample data, which is made from points of discrete time series. Procedure of building data table $v = \{v_x, v_y\}$ depends of problem statement and of data generally.

For example, it may be

a) Shift of prehistory window step-by-step on time series,
b) Shift of prehistory window to some position on time series,
c) Building prehistory window from the series points, that is positioned on some distance.

In addition, we can consider some combination of items indicated above. Visual illustration of univariate time series and principle of building sample table are presented on Figure 163.

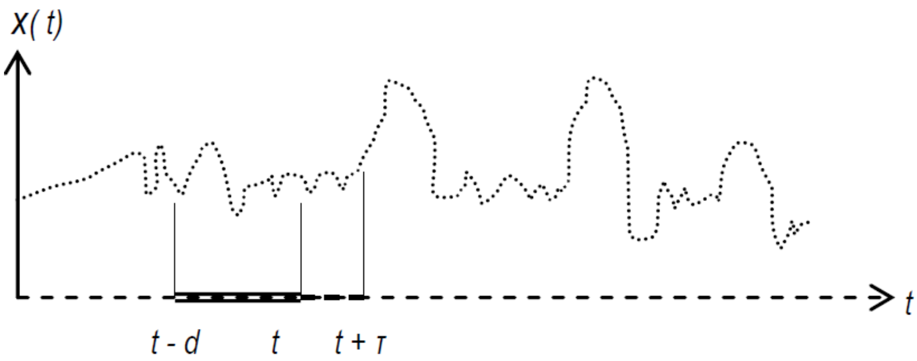


Figure 163. Analysis univariate time series

Not lose commonality let us consider, that $\Delta t = 1$, then prehistory table is built as $v_x = \{x_{kj}\} = \{x(j + k - 1)\}$, were $j = R - d + 1,...,R$, $d \le R \le N - \tau$, $k = 1,...,N - R - \tau$, and forestalling term table (future predictions) is builder as $v_y = \{x_{kj}\} = \{x(j + k)\}$, were $j = R + 1,...,R + \tau$, $k = 1,...,N - R - \tau$, for case (a) as above. With the help of data table $v = \{v_x, v_y\}$ of the size $N - R - \tau$ we will construct sample decision function $\bar{f}$ from the class Φ_M by the MLRP-method. So we have that choose of optimal length d^* of prehistory corresponds to choose of informative characteristic subset. A choose of optimal forestalling term length τ^* will be correspond to definition of likely problem size (complexity) for a given sample size.

We define a time moment of changing model's structure (adaptive) as a time moment $t^* = t_{R-d}$ for which the condition $|F(\bar{f}) - F^*| \geq h$ is carry out, were the value F^* is threshold value of model quality, h is admissible value of deviation for established quality.

Below we will consider logical decision function class Φ_M and its properties. We will define criterion of quality F(f) for decision function f.

- ✓ **The prediction with respect to other time series**

This paragraph is devoted to detection of correlation between two univariate time series. That problem statement is well known and is commonly applied for solving practice problem [Bezruchko and Smirnov, 2003]. However for the most part methods indicate some power of correlation for the given time point.

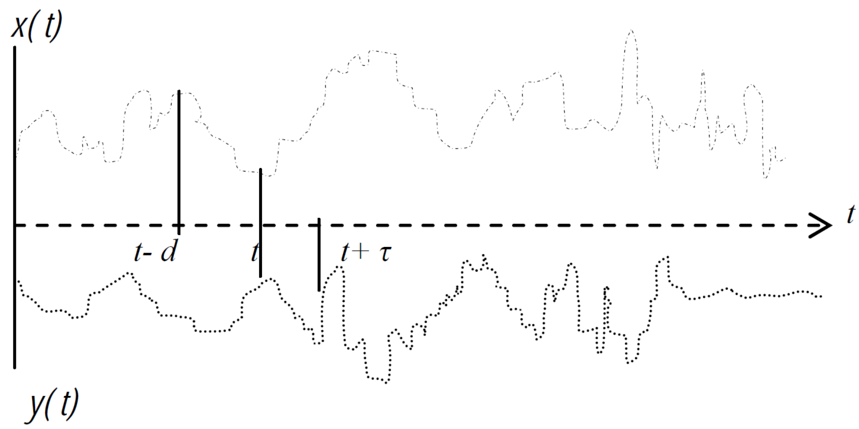


Figure 164. Two dependent time series

The method to attack is founded on constructing so function for that definitional domain is assigned in domain of realization one time series $\{x(t)\}$, and value domain (domain of prediction point $t + \tau\ \Delta t$, $\tau = 1,\dots,N-d$.) is assigned in domain of realization other time series $\{y(t)\}$. It is supposed that with respect to other process. Visual illustration of two dependent time series and principle of building sample table are presented on Figure 164.

The data table is constructed by principle like above. The power of correlation f is defined by quality value $F(f)$ a) on the learning sample and b) on the control sample. We will construct sample decision function $\bar{f}$ (logical-and-probabilistic correlation) from the logical decision function class Φ_M by the MLRP-method like above.

10.4 Application MLRP-method to prediction multivariate time series problem

10.4.1 Application for analysis some of problems in climate

This paragraph is devoted to some of practical problem from hydrological area. It consists in the prediction of the columbine $(k = 1)$, transmitting across Oby riverbed, the average monthly temperature $(k=2)$, and the atmospheric precipitates $(k=3)$ by like hydrometeor data (in the course of the 86 years) in control post of the city Kolpashevo of Novosibirsk region in Russia. In order to construct decision function of prediction variable system (y_1, y_2, y_3) in April by variable system in the course November, December, and January the average monthly data was worked up in the course of November $(i=1)$, December $(i=2)$, January $(i=3)$, and April $(j=1)$ in control post. So target setting provides potential for detecting possible shallowness or hydramnios in Oby lade in order to alarm about extremal hydrological situation.

We have considered several versions of this target setting and provided resolutions by MLRP-method. The sample decision functions were constructed by learning data {$x_k(t_{k_i+12})$, $y_k(t_{k_j})$} of the size 76.

First version – prediction of the columbine, the temperature and the atmospheric precipitates in April (three variables) by columbine data, temperature data and the atmospheric precipitates data for November, December and January (nine variables). We have constructed resolution so as logical probabilistic function, presented below.

Variable declaration:

X1 – the average monthly temperature for November,

X2 – the average monthly temperature for December,

X3 – the average monthly temperature for January,

X4 – the average monthly atmospheric precipitates for November,

X5 – the average monthly atmospheric precipitates for December,

X6 – the average monthly atmospheric precipitates for January,

X7 – the average monthly columbine for November,

X8 – the average monthly columbine for December,

X9 – the average monthly columbine for January,

Y1 – the average monthly temperature for April,

Y2 – the average monthly atmospheric precipitates for April,

Y3 – the average monthly columbine for April.

Logical regularity:

1. IF (X4=(45.000, 182.700]),
   THEN (Y1=[-4.200, 1.700] & Y2=[17.000, 60.000] & Y3=[873.000, 7390.000])

   The estimate of probability is 0.90000.

   Number of sample points (N) is 10.

2. IF (X7=[553.000, 1320.000] & X3=(-21.600, 22.300] & X4=[2.000, 31.200]),
   THEN (Y1=[-4.500, 2.500] & Y2=[4.100, 18.000] & Y3=[1010.000, 4890.000])

   The estimate of probability is 1.00000.

   Number of sample points (N) is 5.

3. IF (X4=(31.200, 36.000] & X7=[553.000, 2140.000]),
   THEN (Y1=[-2.100, 3.000] & Y2=[7.800, 46.000] & Y3=[1400.000, 5420.000])

   The estimate of probability is 1.00000.

Number of sample points (N) is 9.

4. IF (X4=(36.000, 45.000] & X7=[553.000, 2140.000]),
THEN (Y1=[-3.700, 1.000] & Y2=[14.000, 34.100] & Y3=[834.000, 3950.000])
The estimate of probability is 0.90000.
Number of sample points (N) is 10.

5. IF (X1=[-31.000, -11.600] & X7=(1320.000, 20700.000] & X3=(-21.600, 22.300] & X4=[2.000, 31.200]),
THEN (Y1=[-4.200, 4.000] & Y2=[7.800, 33.600] & Y3=[1870.000, 5110.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 7.

6. IF (X1=(-11.600, 22.300] & X7=(1320.000, 20700.000] & X3=(-21.600, 22.300] & X4=[2.000, 31.200]),
THEN (Y1=[-6.300, 1.800] & Y2=[13.000, 34.000] & Y3=[1120.000, 4850.000])
The estimate of probability is 0.90909.
Number of sample points (N) is 11.

7. IF (X1=[-31.000, -8.800] & X7=(2140.000, 20700.000] & X4=(31.200, 45.000]),
THEN (Y1=[-4.700, 0.400] & Y2=[18.000, 34.000] & Y3=[856.000, 3410.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 6.

8. IF (X1=(-8.800, 22.300] & X7=(2140.000, 20700.000] & X4=(31.200, 45.000]),
THEN (Y1=[-4.000, 1.900] & Y2=[2.000, 27.000] & Y3=[1530.000, 3270.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 6.

9. IF (X1=[-31.000, -13.100] & X3=[-31.000, -21.600] & X4=[2.000, 31.200]),
THEN (Y1=[-2.600, 1.300] & Y2=[12.500, 33.600] & Y3=[894.000, 6560.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 5.

10. IF (X1=(-13.100, 22.300] & X3=[-31.000, -21.600] & X4=[2.000, 31.200]),
THEN (Y1=[1.400, 3.800] & Y2=[5.400, 35.000] & Y3=[2110.000, 7060.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 7.

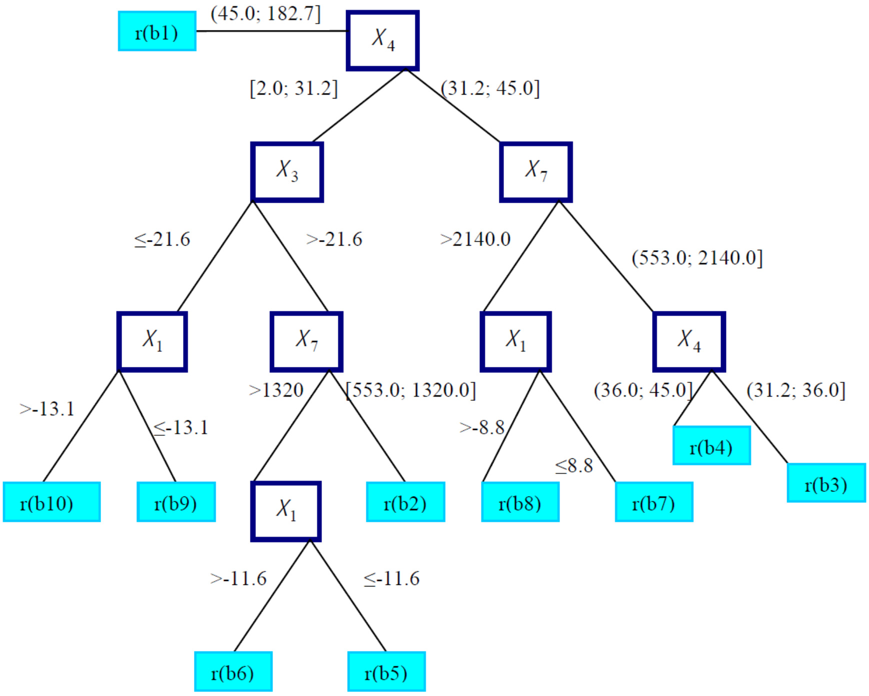
Value of decision function quality $F(\overline{f_{10}})$ is 0.9957355517.

Visual illustration of decision tree for that logical regularity and decision example on top $r(b\ 6)$ are presented on Figure 165 with complexity M = 10, $N^* = 3$.

Estimation of quality criterion (probability estimation of veritable decision by rule $\bar{f}$) was taken by control sample of the size 10 (it is last years of time series) and it was equal 0.8 that is satisfactory result.

For consecutive constructing decision three (the decision function from logical decision function may be presented by dichotomous count) four important features was choose from initial nine features. In compliance with construction decision function the average monthly temperature in November (x_1) and in January (x_2), the atmospheric precipitates in November (x_4), the columbine in November had the most influence on prediction quality.

At consecutive construction of a tree of decisions four significant signs from nine on which fragmentation has been done have been allocated. According to the constructed rule it has appeared, that the greatest influence on quality of the forecast is rendered by monthly average temperature in November (X1) and January (X3), by deposits in November (X4) and by a reservoir in November (X7).



$$r(b6) = "Y_1 \in [-6.3; 1.8]" \wedge "Y_2 \in [13.0; 34.0]" \wedge "Y_3 \in [1120.0; 4850.0]"$$

Figure 165. The decision tree of hydrological situations

Second version – prediction of meteorological values (the temperature and the atmospheric precipitates) in April (three variables) by corresponding data for November, December and January.

The minimum number of laws is 10; the minimum number of points in node is four.

First logical regularity:

1. IF (X5=(25.100, 56.000] & X4=[10.000, 33.000]),
   THEN (Y1=[-6.300, 3.000] & Y2=[13.000, 19.300])

   The estimate of probability is 1.00000.

   Number of sample points (N) is 6.

2. IF (X4=[10.000, 16.700] & X5=[8.700, 25.100]),
   THEN (Y1=[-5.800, 4.000] & Y2=[7.800, 33.600])

   The estimate of probability is 1.00000.

Number of sample points (N) is 6.

3. IF (X5=(18.300, 25.100] & X4=(16.700, 33.000]),
THEN (Y1=[-0.600, 2.500] & Y2=[4.100, 31.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 6.

4. IF (X5=[8.700, 23.000] & X4=(33.000, 68.000]),
THEN (Y1=[-4.200, 0.100] & Y2=[17.000, 32.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 7.

5. IF (X5=(23.000, 27.000] & X4=(33.000, 68.000]),
THEN (Y1=[-1.700, 1.900] & Y2=[2.000, 42.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 6.

6. IF (X5=[8.700, 12.400] & X4=(16.700, 33.000]),
THEN (Y1=[-3.000, 3.000] & Y2=[7.900, 33.600])
The estimate of probability is 1.00000.
Number of sample points (N) is 5.

7. IF (X5=(12.400, 18.300] & X4=(16.700, 33.000]),
THEN (Y1=[-2.000, 3.700] & Y2=[5.400, 34.500])
The estimate of probability is 1.00000.
Number of sample points (N) is 9.

8. IF (X3=(-18.400, -11.000] & X5=(27.000, 56.000] & X4=(33.000, 68.000]),
THEN (Y1=[-3.100, 2.900] & Y2=[23.000, 60.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 4.

9. IF (X6=[3.800, 21.000] & X3=[-31.000, -18.400] & X5=(27.000, 56.000] & X4=(33.000, 68.000]),
THEN (Y1=[-3.800, 1.900] & Y2=[7.700, 52.300])
The estimate of probability is 1.00000.
Number of sample points (N) is 7.

10. IF (X6=(21.000, 47.000] & X3=[-31.000, -18.400] & X5=(27.000, 56.000] & X4=(33.000, 68.000]),
THEN (Y1=[-2.100, -0.300] & Y2=[7.800, 46.000])
The estimate of probability is 1.00000.
Number of sample points (N) is 4.

Value of decision function quality $F(\bar{f}_{10})$ is 0.8372715792.

Visual illustration of decision tree for that logical regularity is presented on Figure 166 with complexity M = 10, $N^* = 4$.

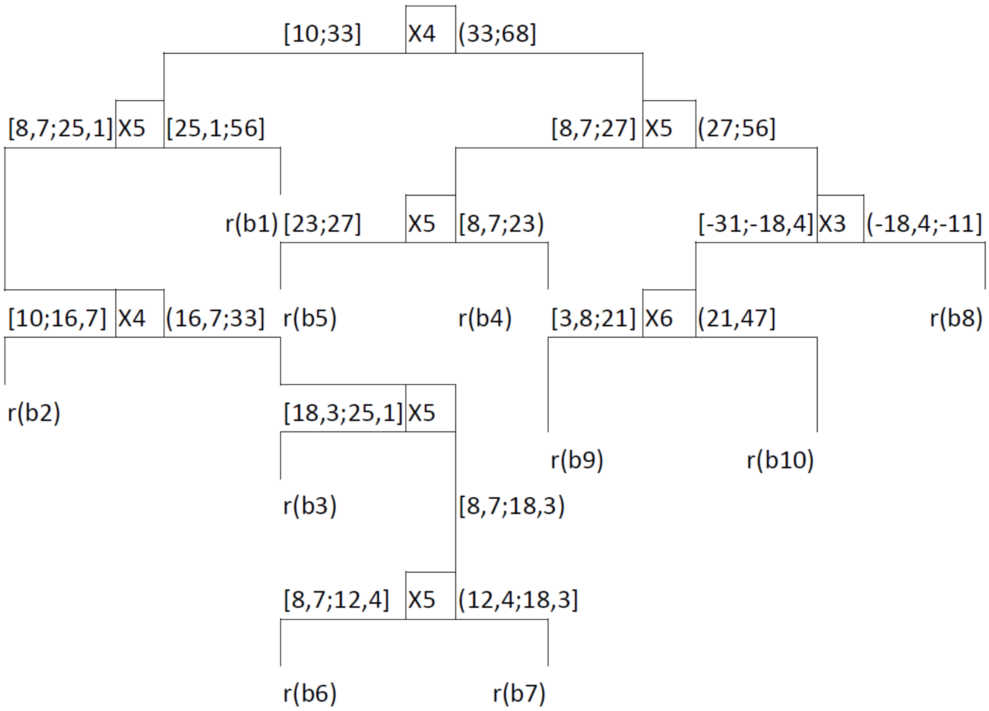


Figure 166. First decision tree of meteorological situations

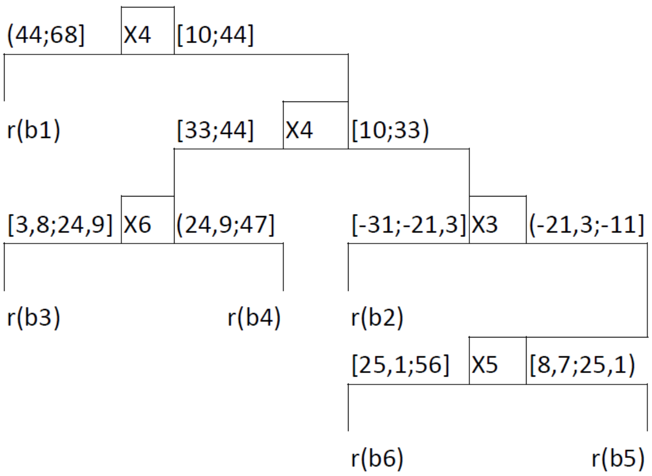


Figure 167. Second decision tree of meteorological situations

Forecast decision function has been constructed on training sample of volume 60. The estimation of criterion of quality is received on control sample of volume 26 (last twenty six years) and is equal 0.53, thus the error on the control at forecasting only monthly average temperature is equal 0.73, at forecasting only deposits – 0.69. Result visualization of the decision is presented in the form of a tree of decisions on Figure 167.

In addition the solving tree of complexity 6 (the maximum number of nodes of a tree of decisions) with the minimum number of points in node equal 4 has been received.

Second logical regularity:

1. IF (X4=(44.000, 68.000]),
   THEN (Y1=[-8.800, 1.700] & Y2=[15.000, 60.000])

The estimate of probability is 1.00000.

Number of sample points (N) is 11.

2. IF (X3=[-31.000, -21.300] & X4=[10.000, 33.000]),
THEN (Y1=[-3.000, 3.900] & Y2=[5.400, 35.000])

The estimate of probability is 1.00000.

Number of sample points (N) is 16.

3. IF (X6=[3.800, 24.900] & X4=(33.000, 44.000]),
THEN (Y1=[-3.800, 1.900] & Y2=[2.000, 37.400])

The estimate of probability is 1.00000.

Number of sample points (N) is 16.

4. IF (X6=(24.900, 47.000] & X4=(33.000, 44.000]),
THEN (Y1=[-4.700, 2.900] & Y2=[15.000, 46.000])

The estimate of probability is 1.00000.

Number of sample points (N) is 9.

5. IF (X5=[8.700, 25.100] & X3=(-21.300, -11.000] & X4=[10.000, 33.000]),
THEN (Y1=[-5.800, 4.000] & Y2=[4.100, 34.000])

The estimate of probability is 1.00000.

Number of sample points (N) is 19.

6. IF (X5=(25.100, 56.000] & X3=(-21.300, -11.000] & X4=[10.000, 33.000]),
THEN (Y1=[-6.300, -2.800] & Y2=[13.000, 25.000])

The estimate of probability is 1.00000.

Number of sample points (N) is 5.

Value of decision function quality $F(\bar{f}_6)$ is 0.7169073023.

Forecast decision function has been constructed on training sample of volume 76. The estimation of criterion of quality is received on control sample of volume 10 and is equal 0.5, thus the error on the control at forecasting only monthly average temperature is equal 0.6, at forecasting only deposits – 0.8. Result visualization of the decision is presented in the form of a tree of decisions on Figure 167.

According to the constructed rules it has appeared, that the greatest influence on quality of the forecast of deposits and temperature in April is rendered by monthly average temperature in January (X3), by deposits in November, December, January (X4, X5, X6).

Demonstration of solving problems from the considered applied areas shows efficiency of the offered methods of construction of logical-and-probabilistic models and can be applied in many natural-science areas.

10.4.2 Application for analysis some of problems in seismology

In this section, we consider the problem of detection of statistical regularities in the analysis of seismological data presented time series. The task of study of the earthquake as a physical process is quite complicated at least because of the earthquake originated in the deep bowels of the earth, not accessible to direct observation and measurement. Analysis of spatial-temporal course of seismicity is one of the promising directions in the modern seismology. Study of the regularities that occur before, during or after the earthquake, provides an opportunity to get closer to the mystery of earthquakes and, possibly, give some predictions of their occurrence or migration [Nikonov, 2006]. We were interested in testing of the algorithm MLRP on seismic data to test the hypothesis of logical-probabilistic regularities between seismic events occurring in two geographical areas.

The approach proposed to solve this problem, based on the construction of predictive function f: $D_x \rightarrow D_y$ from the class Φ_M of logical decision functions. The area of definition D_x is given in the space of realizations of a time series (x (t)), and the area of values D_y (forecasts) – in the space of realizations of another time series (y (t)), the complexity of the class of decision functions defined by the parameter M. A series {x (t)} is formed from the values of maximum magnitude of the earthquakes for every half of day in zone I, a series {y (t)} – the values of maximum magnitude earthquakes for every half of day in zone II, geographically distinguished from the Zone I.

In the data table v = $\{v_x,v_y\}$ of volume N – d – τ, d – length of the background, τ – the period of advance, we will build a sample decision function $\bar{f}$ of the class Φ_M by algorithm MLRP. Selecting the optimal length d* of prehistory corresponds to choosing an informative subset of features, the choice of the optimal length of the period of advance τ* corresponds to definition of appropriate dimension (complexity) of the problem for a fixed amount of training.

Two zones of equal area, located on a small geographically remote from each other (the first zone presented by the territory between -10° and 0° the latitude and between 150° and 160° the longitude, the second – from -20° to -10° the latitude and from 160° to 170° the longitude) were arbitrarily selected from the data presented in the directory ftp://www.ncedc.org/pub/catalogs/cnss//

For d = 10 we have a set of variables X_i, i = 1,..., 10, whose values are defined as the maximum magnitude of events for the corresponding half of day rounded to the integer value to transfer to the nominal space of variables). Let τ = 2, ie, the variables Y_i, i = 1, 2, refer to the following day for X_{10}. A decision tree for M = 3 and time series of 10000 samples per period of 5000 days, starting from 1970, was constructed.

1. IF ($X_8 \in$ {4, 5, 6, 10}),

THEN ($Y_1 \in$ {0, 5} & $Y_2 \in$ {0, 4, 5, 6})

The estimate of probability is 0,902.

Number of sample points N_1 is 1748.

2. IF ($X_6 \in$ {0, 4, 5, 6, 8, 10} & $X_8 \in$ {0, 1, 2, 3, 7, 8, 9}),

THEN ($Y_1 \in$ {0, 4, 5, 6} & $Y_2 \in$ {0, 5})

The estimate of probability is 0,906.

Number of sample points N_2 is 8092.

3. IF ($X_6 \in$ {1, 2, 3, 7, 9} & $X_8 \in$ {0, 1, 2, 3, 7, 8, 9}),

THEN ($Y_1 \in$ {0, 2, 5, 9} & $Y_2 \in$ {0, 5, 6})

The estimate of probability is 0,900.

Number of sample points N_3 is 160.

Value of decision function quality is 0,839.

Unfortunately, the accuracy of seismic event with a given magnitude in a given class of decision functions is poor. The reason may be either too rough rounding of data and the presence of a large number of events with a magnitude of zero. Nevertheless, it is interesting result, which consists in the fact that relatively short time events (variables X_6, X_8) were separated with the choice of informative subspace of variables.

10.5 Discussion

In this chapter, two ways to solving analysis univariate time series problem were considered. It was founded on the MLRP-method constructing logical-and-probabilistic model for prediction heterogeneous variables system. The idea's approach and ways of realizations was formulated here.

The decision was constructed by MLRP-method from the logical decision function class. It allows taking optimization parameters as such prehistory length and forestalling term for univariate time series. Our method has very priority properties for investigation of very complexity processes. It is: a) easily interpreted decision is constructed; b) the given class allows to process heterogeneous experimental data; c) it has a small measure of complexity; d) it allows to work in the presence of admissions in empirical tables and to receive adequate decision at small volumes of sample.

Very actual practice problems from hydrological and seismology domains are presented here by MLRP-method. We have received quite good results. Results of seismology problem allow us to understand process of geographic migration of seismic zones and perhaps even to detect earthquake precursors. Results of hydrological problem show us the most likely signs, which have most influence on the forecast of meteorological and hydrological features. Demonstration of solving problems from the considered applied areas shows efficiency of the offered methods of construction of logical-and-probabilistic models and can be applied in many natural-science areas.

We want to note that proposed approach to joint analyses of some time series can be have more than enough applications. We can to solve problem of statistically important correlation detection between processes arising on the most distant region. Method presented here depends on statistical information about studied object and can be used in many domains of data mining. We think that many actual problems of monitoring and better understanding our environment can be solved by our method.

We think that our investigations are included to GMES-project properly and we look forward to collaboration with colleagues from other countries. We have science contact with colleagues from Bulgaria GMES and we hope to join efforts to solve general problems. Our knowledge and science experience will be beneficial to GMES project.