PROBLEMS OF STATISTICAL DECISIONS IN OCEAN MONITORING

F. A. Mkrtchyan*

Department of Informatics, V.A. Kotelnikov's Institute of Radioengineering and Electronics, Russian Academy of Science, Fryazino, 141190, Russia
E-mail: ferd47@mail.ru

KEY WORDS: Geoinformation monitoring, statistical decisions, spottiness, recognition, small samples

ABSTRACT:

Application of means of geoinformation monitoring in many cases is connected to acceptance of the statistical decision on presence on a surveyed part Terrestrial surface of this or that phenomenon. One of features of a condition of gathering of the information for such decision is the impossibility of reception the big statistical samples.
Therefore development and research of optimum algorithms of distinction of the casual signals characterized by samples of limited volume, in conditions of parametrical aprioristic uncertainty are necessary.
At present time there are many methods of recognition which are caused appreciably by variety of statements of concrete tasks.
The feature of remote measurements is information acquisition, when the data of measurements, acquired during tracing of flying system along routes of survey, are directed to input of the processing system. As result the two dimensional image of investigated object is registered. Statistical model of spottiness for investigated space is one of models for this image.
In real conditions, the study of spots, the acquiring of their statistical characteristics and their using in a problem of detection is enough a complex problem. It is necessary to develop the criteria allowing the distinguishing the spots from other phenomena. For example, it is necessary to determine such threshold the exceeding of which is the spot indicator. Also it is necessary to develop model presentation of processes of spots detection.
Statistical characteristics "spottiness" microwave temperatures can be used at recognitions and classifications of the phenomena on a surface of the ocean, distinguished by a degree of excitement.
The analysis of empirical histograms for "spottiness" shows, that in most cases (l+, l-) - characteristics will be coordinated with exponential distribution, and amplitude characteristics will be coordinated with normal distribution. Therefore for detection and classification of the phenomena on a surface of ocean it is necessary to apply optimal algorithms for the Computer training to taking statistical decisions for the aforesaid distributions
In the present work the generalized adaptive algorithm of training to acceptance of statistical decisions for exponential classes of distributions is developed at aprioristic parametrical uncertainty of conditions small samples. Numerical examples are shown. Efficiency of the developed optimum procedure for small samples is shown.

1. INDRUTACTION

Development of systems of geoinformation monitoring demands the decision of some problems of the organisation of data flows of measurements. Among these problems of one of important the problem of acceptance of the statistical decision on presence on a surveyed part of a terrestrial surface of this or that phenomenon is. One of features of conditions of gathering of the information for such decision is the impossibility of reception statistical samples small volumes. Therefore working out and research of optimum algorithms of acceptance of statistical decisions for sample small volume are necessary at informational restrictions.
For a case when the number of supervision is great enough, the problem dares a method of an estimation of parameters of likelihood distributions which is effective at unlimited growth of volumes sample on which basis the estimation of parameters is made. At the limited volumes sample, received by a method of an estimation of the parameters, the solving rule does not satisfy to necessary conditions of an optimality: to a constancy of average probability of an error of first kind and unbiasedness.

In the present work the generalised adaptive algorithm of training to acceptance of statistical decisions for exponential groups of distributions is developed at aprioristic parametrical uncertainty of conditions samples small volume (Mkrtchyan, 1982; Armand et al.., 1987).

2. PROBLEM DEFINITION

Very often there is a problem: to what of two classes to carry a measured random variable, and the full likelihood description of these classes is not known that does not allow to use classical results of the theory of statistical decisions for the decision of this problem. The decision can be received only by means of training sampling.

Let ξ, η, ζ independent random variables
$f_\xi(x/\omega_0)$, $f_\eta(y/\omega_1)$, $f_\zeta(z/\omega)$ distribution of probabilities
$\omega_0, \omega_1, \omega \in \Omega$. For parameter ω two alternatives: H_0: $\omega = \omega_0$ and H_1: $\omega = \omega_1$. The problem consists in construction of a solving rule on n_0, n_1, n to supervision of random variables

$$\xi: \quad x^* = (x_1, x_2, \ldots \ldots \ldots x_{no})$$

η: $y^* = (y_1, y_2, \dots \dots \dots y_{n1})$

ζ: $z^* = (z_1, z_2, \dots \dots \dots z_n)$

to specify, what alternative H_0 or H_1 is accepted

The solving rule can be set by means of function $\varphi(x^*, y^*, z^*)$.

If $\varphi = 0$: $H=H_0$, $\varphi = 1$: $H=H_1$

Errors of first and second kind:

$\alpha(\varphi, \omega_0, x^*, y^*) = \int \varphi(x^*, y^*, z^*)\, f\omega_0(z^*)\, dz^*$

$\beta(\varphi, \omega_0, x^*, y^*) = \int [1- \varphi(x^*, y^*, z^*)]\, f\omega_1(z^*)\, dz^*$ An optimality condition: 1. $\acute{\alpha} = \alpha_0$, (a constancy of an error of first kind) ;2.$1-\beta > \acute{\alpha}$ (unbiasedness).

In the present work it is developed a mathematical apparatus and the generalised adaptive procedure for the decision of a problem of training to distinction of random variables from exponential group distributions with unknown parameters for sampling small volume is offered at information restrictions.

It is shown that procedures of training applied now to distinction in which the estimation of parameters at first is resulted, and then a choice between hypotheses, do not meet the above-stated requirements shown to optimum procedures.

For a case when the number of supervision is great enough, the problem dares a method of an estimation of parameters of likelihood distributions which is effective at unlimited growth of volumes sampling on which basis the parameter estimation is made. At the limited volumes sampling, received by a method of an estimation of the parameters, the solving rule does not satisfy to necessary conditions of an optimality: to a constancy of average probability of an error of the first kind and unibasendness.

3. THE SOLVING RULE BASED ON ESTIMATION OF UNKNOWN PARAMETERS OF DISTRIBUTIONS

The classic method of the decision of a considered problem is based on enough developed theory of dot estimations of unknown parameters of likelihood distributions. In a considered problem for unknown parameters ω_0, ω_1, starting with n_0 sequence

$x^* = (x_1, x_2, \dots \dots \dots x_{no})$ and n_1

sequence $y^* = (y_1, y_2, \dots \dots \dots y_{n1})$ random variables ξ and η estimations of parameters ω_0, ω_1, θ_1, θ_0 accordingly turn out.

Further the way of construction of a solving rule is based on fundamental lemma Neyman–Pearson: the credibility relation is under construction:

$L(z^*/\theta_1, \theta_0) = [f(z_1, z_2, \dots z_n/\theta_1) / f(z_1, z_2, \dots z_n/\theta_0)] > C(\theta_1, \theta_0)$

Also threshold $C(\theta_1, \theta_0)$ gets out.

1. $\theta_1 > \theta_0$ $\{s < t\,(n_0/n_1),\ s < n_0/G^{-1}_n(1-\alpha_0)$
2. $\theta_1 < \theta_0$ $\{s > t\,(n_0/n_1),\ s > n_0/G^{-1}_n(1-\alpha_0)$

$s = x/z$, $t = y/z$, $G^{-1}_n = [1/(n-1)!] \int \exp(-z)\, z^{n-1} dz$,

θ_1 and θ_0 – dot estimations for ω_1 and ω_0

Areas of acceptance of hypothesis H_1 it is resulted on fig.1.

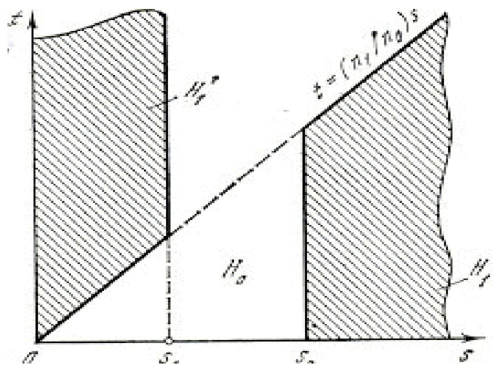

Figure 1. Areas of acceptance of hypotheses for a classical solving rule.

Apparently from Figure 2. The probability of an error of the first sort two from half of time exceeds admissible volume α_0.

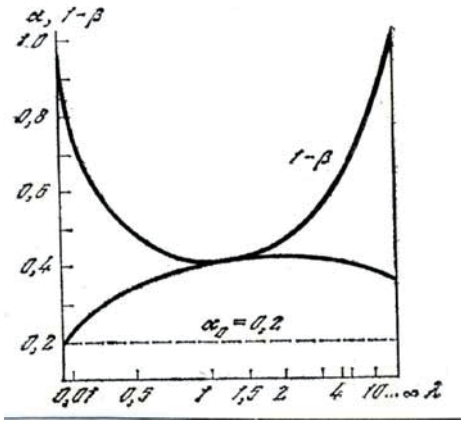

Figure 2. Probabilities of acceptance of correct solution and error of first kind.

4. THE SOLVING RULE SATISFYING NESSARY CONDITIONS OF AN OPTIMALITY

Optimum solving rule in a class of solving rules shown on Figure 3.

1)max $\alpha(s_1, s_2/\lambda) = \alpha_0$,
$0 < \lambda < \infty$
2)max min $D(s_1, s_2/\lambda)$
(s_1, s_2) $0 < \lambda < \infty$

$D(s_1, s_2/\lambda) = 1-\beta(s_1, s_2/\lambda)$

$\alpha(s_1, s_2/\lambda) = P_0(s_1) + [P_0(\infty) - P_0(s_1)]$

$D(s_1, s_2/\lambda) = P_1(s_1) + [P_1(\infty) - P_1(s_2)]$

$P_i(k) = \iint f_i(s, t)\, ds dt$, where $i = 0,1$.

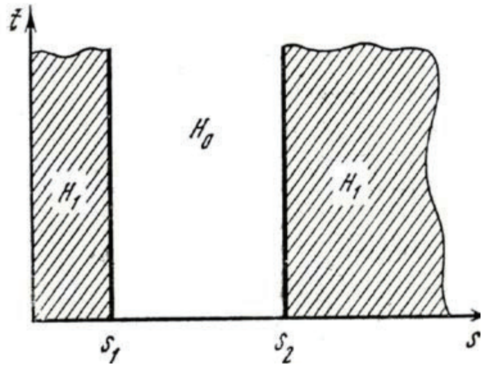

Figure 3. Areas of acceptance of hypotheses for the optimal inference engine

Applying a method of uncertain multipliers of Lagrange to optimum thresholds s_1 and s_2 we receive system of the equations:

1) $(s_1/s_2)^{n_0} = [(s_1+1)/(s_2+1)]^{(n+n_0)}$

2) $1 - \sum_{j=0}^{n_0-1} [(n_0+j)/n!] \, [[(n_0+n-1)! \, (s_1+1)^{(n_0-1)} - n]/(s_1+1)^{(n_0+n-1)}] -$

$- [[(n_0+n-1)! \, (s_2+1)^{(n_0-1)} - n]/(s_2+1)^{(n_0+n-1)})]] - \alpha_0 = 0$

The offered procedure satisfies following imperative conditions: 1) to a constancy of average probability of an error of the first kind α and 2) unbiasedness (1-β) <α).

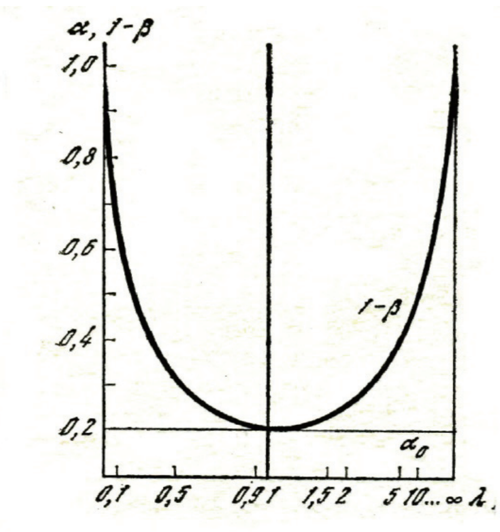

Figure 4. Probabilities of acceptance of correct solution and error of first kind.

5. APPLICATION

Analysis of statistical characteristics of "spottiness" for three types of areas of Atlantic and Pacific oceans was conducted. These statistical characteristics were determined for the most informative thresholds. At that time statistical characteristics of "spottiness" for the same areas, selected using criteria of minimal value of coefficient of correlation for joint sample of positive and negative spots. Analysis of these characteristics showed, that the statistical characteristics of "spottiness" coincide for areas with temperate sea roughness and storm zones. Minimum for the coefficient of correlation ρ_{min} is run down for a case of most informative thresholds. But for quiet area the situation is different[2].

. On Figure 5. the example of work of the automated system in a mode of monitoring of temperature of a surface of Northern Atlantic on data Space Satellite « Cosmos- 1151 » (8 - April, 14, 1980) is given. The system allows to receive maps of temperatures on enough rarefied grid of trajectory SS. Points on a map designate areas of realization of ship measurements. The analysis of satellite and contact measurements shows, that there is an appreciable regular understating satellite estimations of temperature of ocean concerning ship which on the average makes 1,6 K. The root-mean-square deviation(rejection) of satellite estimations T from ship measuring on all given sample makes *3,3K*. The dotted line on a map designates areas where the difference between ship and satellite measurements exceeds *4 K*. It is typical, that high overcast is registered in all these points (on the data weather forecasters). The root-mean-square deviation of satellite measurements of temperature from ship, designed without taking into account the allocated points, makes *1,4 K*.

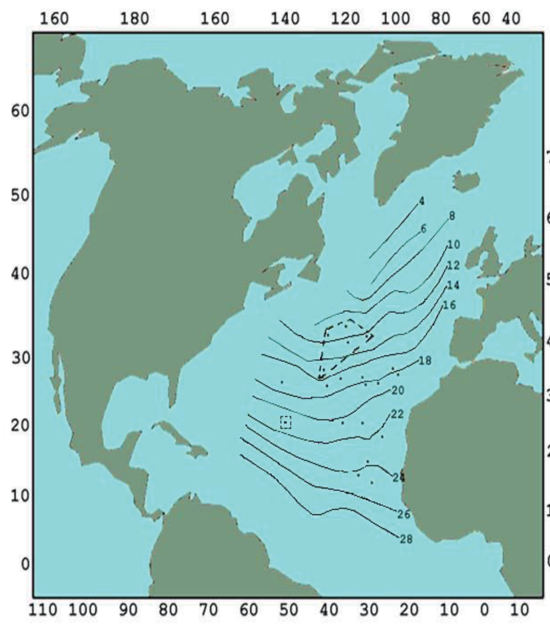

Figure 5. Map of temperature of Northern Atlantic.

From the aforesaid follows, that statistical characteristics for "spottiness" of brightness temperatures in microwaves can be used for detection and classification of the phenomena on a surface of the ocean, that was caused by a degree of sea roughness.

On Figure 2 the exemple of the automated system in a mode of monitoring of brightness temperature of a surface of Pacific Oceans on date Space Satellite "Intercosmos-21"(20-22, February 1979) is given.The system allows to receive maps of temperatures on enough rarefied grid of trajectory SS.From the aforesaid follows, that statisticalcharacteristics for "spottiness" of brightness temperatures in microwaves can be used for detection and classification of the phenomena on a surface of the ocean, that was caused by a degree of sea roughness.

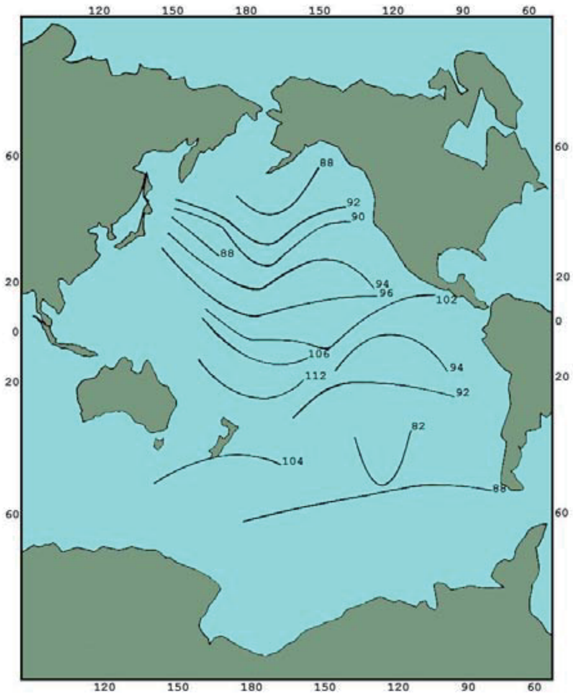

Figure 2. Map of brightness temperature of Pacific Ocean.

The analysis of empirical histograms for "spottiness of "brightness temperatures in microwaves" shows, that in most cases (l+, l-) - characteristics will be coordinated with exponential distribution, and amplitude characteristics will be coordinated with normal distribution. Therefore for detection and classification of the phenomena on a surface of ocean it is necessary to apply optimal algorithms for the COMPUTER training to taking statistical decisions for the aforesaid distributions((Mkrtchyan, 1982; Armand et al.., 1987).

.

5. CONCLUSION

The analysis of empirical histograms for "spottiness of "brightness temperatures in microwaves" shows, that in most cases *(l^+, l-)* - characteristics will be coordinated with exponential distribution, and amplitude characteristics will be coordinated with normal distribution. Therefore for detection and classification of the phenomena on a surface of ocean it is necessary to apply optimal algorithms for the COMPUTER training to taking statistical decisions for the aforesaid distributions.

REFERENCES

Mkrtchyan, F., A., 1982. *Optimal distinguish of signals and monitoring problems.* Nauka, Moscow, pp. 182(in Russian).

Armand, N., A., Krapivin, V., F., and Mkrtchyan, F.,A., 1987. *Methods of data processing of radiophysical research of an environment,* Nauka, Moscow, pp. 270 (in Russian).