DETECTION OF THERMAL ANOMALIES (FIRES) UNDER THE ATMOSPHERIC INFLUENCE FROM DATA OF AVHRR DEVICES AND METEOROLOGICAL SERVICES

K.T. Protasov

Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk
Received July 2, 1998

A two-stage procedure for detection of thermal anomalies (such as fires) on the territory of some region from satellite data of the AVHRR device is considered. At the first stage, the field of thermodynamic temperature of the underlying surface is reconstructed using a nonlinear non-parametric regression adapted to particular observation conditions found from the coordinated data of meteorological services and data of the AVHRR device from onboard the NOAA satellite. At the second stage, the reconstructed temperature field is used for constructing the Bayes adaptive rule to detect thermal anomalies. The rule is based on the principle of the component identification in a mixed distribution and approximation of the conditional density functions by the Johnson curves. An example of detection of thermal anomalies from satellite video data at the territory of the Tomsk Region is presented.

The problem of early detection of fires with initially small size is very urgent, especially for hard-to-reach and sparsely populated regions. In this connection, most promising way is to use satellite information for early detection of sources of thermal anomalies (possible fires). Most suitable for such purposes are satellites of the NOAA series, since they are most often orbiting over our region and they allow monitoring of the Earth's underlying surface. Such satellites carry an AVHRR device, which records radio brightness, in the form of images, in five spectral regions including thermal ones. Unfortunately, low resolution of the AVHRR device and comparatively narrow range of radio brightness that can be recorded with it does not permit efficient solution to the problem of early detection of small-size fires.

In this paper, we consider an approach that is based on the methods of predicting non-observed parameters (values of thermodynamic temperature) from the information contained in indirect measurements. In this case, the data observed in the five channels of the AVHRR device play the part of such indirect measurements. Note that efficiently solving the above-stated problem requires highly accurate referencing of the scanner pictures to the map of a territory surveyed. The referencing of video data by use orbital data does not provide a desired accuracy, first of all, due to the errors in determination of orbital elements and drifts in the orientation angles of a satellite. So, the orbital-data-based referencing can be considered only as a preliminary one and, thus, it should be corrected using reference points specified by an operator.

As known, self-radiation of an object is a function of its temperature, physical properties, and structure characteristics of the emitting surface. One may consider the spectral power brightness as a characteristic informative of incoherent self-radiation of a heated body. The power brightness, which can be determined by the Planck formula, depends on temperature, wavelength, and the factor $\varepsilon_\lambda(T)$ (emissivity). This factor is also called the spectral degree of blackness of a radiating surface at a given temperature and direction of viewing. For instance, the black body has $\varepsilon_\lambda = 1$ in the whole wavelength range. For a gray body, $0 < \varepsilon_\lambda < 1$ in a certain wavelength range. For some types of the underlying surface, the values of $\varepsilon_\lambda(T)$ are known. However, real devices record combined radiation brightness distorted by the atmosphere. Thus, it is problematic to estimate the temperature of the underlying surface from the radiation brightness measured under the influence from the atmosphere.

We suppose that images received from a satellite are pre-processed: geometrical distortions of the data are eliminated; data are referenced to geographical coordinates; a fragment of video data corresponding to Tomsk Region (TR) and the adjacent territories is "cut out"; data of the AVHRR device are corrected and calibrated by switching to the albedo values for the first and second channels and to thermodynamic temperatures in the third, fourth, and fifth channels (using inverted Planck formula).

0235-6880/98/12 1129-05 $02.00

Thus, we have a field of 1024 × 1024 five-dimensional vectors with geographic coordinates for every pixel of the image (the field of video data). Besides, we have the data of meteorological services on the temperature of the underlying surface (this may be temperature of soil or of the atmospheric surface layer) at some sites with known coordinates, which are spaced sufficiently uniformly over the territory. The data of meteorological services are given for the time of satellite observations. In the ideal case, the temperature at reference points for every class (type) of the underlying surface (water, field, plowed field, forest, etc.) should be known.

RECONSTRUCTION OF THE THERMODYNAMIC TEMPERATURE FIELD

The thermodynamic temperature of the underlying surface (US) on the field to be predicted is described by a random value $Y \in R^1$. The data on radiation brightness, which form the source of information used for prediction, are described by a random vector $\mathbf{X} \in R^k$ where R^k is the 5-dimensional Euclidean space; $\mathbf{X} = (X^1, \ldots, X^k)^\mathrm{T}$, X^i is the radiation brightness measured in the ith channel of the AVHRR device, $i = 1, \ldots, k$, T is the transposition sign.

The relation between the variable Y to be predicted and the vector $\mathbf{X}$ is described by the regression functional of the following form:

$$m(\mathbf{x}) = E(Y/\mathbf{X} = \mathbf{x}), \tag{1}$$

where $E(\ )$ is the operator of mathematical expectation, $E(|Y|) < \infty$.

If the following probability densities of the random variables $\mathbf{X}$ and Y exist, then, taking into account Eq. (1), we have

$$y = m(\mathbf{x}) = \int_{R^1} y \frac{f(\mathbf{x}, y)}{f(\mathbf{x})\, f(y)} \, \mathrm{d}F(y), \tag{2}$$

where $\mathbf{x} \in R^k$, $y \in R^1$, $f(\mathbf{x}, y)$ is the joint probability density of the random vector $\mathbf{X}$ and the random variable Y; $f(\mathbf{x})$ is the probability density of the random vector $\mathbf{X}$; $f(y)$ is the probability density of the random variable Y, and $F(y)$ is the integral distribution function of Y.

Let we have a sample of pairwise independent, similarly distributed random variables $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$, where n is the number of reference temperature values taken from meteorological services at the time of satellite observations. In Eq. (2), it is natural to use non-parametric estimates for unknown distributions using sample data.[1–3] Let us replace the unknown distributions with their non-parametric estimates of the kernel type and $F(y)$ with the empirical function $F_n(y)$. Then the estimate for the regression equation (2) takes the form[3]

$$\hat{m}(\mathbf{x}) = \sum_{l=1}^{n} \frac{Y_l \sum\limits_{j=1}^{n} K_h(Y_l - Y_j) \prod\limits_{i=1}^{k} K_h(x_i - X_j^i)}{\sum\limits_{j=1}^{n} \prod\limits_{i=1}^{k} K_h(x_i - X_j^i) \sum\limits_{i=1}^{n} K_h(Y_l - Y_i)}, \tag{3}$$

where h is the window width (smoothing parameter) described by the function $K_h(u) = h^{-1} K(u/h)$; the Epanechnikov kernel[1–3] or Gaussian kernel can be taken as $K(.)$.

The experience of using such estimates shows that accuracy characteristics of the regression equation $\hat{m}_h(\mathbf{x})$ depend basically on the scale parameter h, rather than the kernel form. Because the parameter h is important, it is natural to turn to the vector parameter $\mathbf{h} = (h^1, \ldots, h^k)^\mathrm{T}$ and to use the modified kernel $K'_h(u) = (h^i)^{-1} K(u^i/h^i)$, $i = 1, \ldots, k$ in Eq. (3). This brings up new problem on estimating $\mathbf{h}$ with the allowance for the observations $\{(\mathbf{X}_i, Y_i)\}_{i=1}^n$. To estimate $\mathbf{h}$, let us use the method of moving control. This method consists in the following. A modified estimate of the regression $\hat{m}_{h,j}(\mathbf{X}_j)$ is constructed with the jth observation omitted (excluded). The quality criterion of $\mathbf{h}$ estimate depends on the ability to predict a set of values $\{Y_j\}_{j=1}^n$ from the sets of subsamples $\{(\mathbf{X}_i, Y_i)\}_{i \neq j}$:

$$J(\mathbf{h}) = n^{-1} \sum_{j=1}^{n} [Y_j - \hat{m}_{h,j}(\mathbf{X}_j)]^2 \, w(\mathbf{X}_j), \tag{4}$$

where $w(.)$ is the weighting function, which may be omitted in the simplest cases. It is convenient to solve the optimization problem (4) by the search method,[4] using a two-stage estimation procedure for the global extremum of the functional (4). The search domain in this case is a multidimensional square $\prod\limits_{i=1}^{k} [h^i_{\min}, h^i_{\max}]$, where $h^i_{\min}$ and $h^i_{\max}$ are the lower and the upper estimated limits of the smoothing parameter, respectively. At the first stage, a point with uniform distribution is randomly selected in the search domain. Then the gradient descent is performed with the use of the search adaptation methods.[4] For this purpose, the quality functional (4) is varied by the smoothing parameters in the following way. The increments to the functional (4) are calculated:

$$J_+(\mathbf{h}, a) = (J(\mathbf{h} + a\mathbf{e}_1), \ldots, J(\mathbf{h} + a\mathbf{e}_k)),$$

$$J_-(\mathbf{h}, a) = (J(\mathbf{h} - a\mathbf{e}_1), \ldots, J(\mathbf{h} - a\mathbf{e}_k)),$$

where k is the number of the parameters h corresponding to the number of the components forming the vector $\mathbf{h} = (h^1, \ldots, h^k)^\mathrm{T}$; a is the scalar parameter determining the value of the search step;

$\mathbf{e}_i = \Big(\underbrace{0, \ldots, 1}_{i}, \ldots, 0\Big)^{\mathrm{T}}$, $i = 1, \ldots, k$ are basis vectors of the searching directions. The estimated value of the gradient is calculated in the following way:

$$J_{+}(\mathbf{h}, a)\ J_{-}(\mathbf{h}, a)/(2a) = \nabla_{h\,\pm}\, J(\mathbf{h}, a),$$

where $\nabla_{h\,\pm}$ denotes gradient. The search algorithm of adaptation in the recursion form is as follows:

$$\mathbf{h}[j] = \mathbf{h}[j-1] - \gamma\,[j]\ \nabla_{h\,\pm}\, J(\mathbf{h}[j-1], a[j]). \qquad (5)$$

Selection of the search step $a[.]$ and the working step $\gamma[.]$ is considered in Ref. 4; and $\gamma[.] < a[.]$.

Once the parameter $\mathbf{h}$ in the expression (3) for $\hat{m}_h(\mathbf{x})$ is found, the regression equation can be used to reconstruct the values of Y from observed $\mathbf{X}$ all over the field of video data. It should be noted that the regression model for prediction of non-observed values is valid only with statistically homogeneous data, from which the dependence is reconstructed, as well as the data to be reconstructed. Therefore, it is first necessary to analyze the entire image using a segmentation algorithm to separate statistically homogeneous areas and exclude the cloud fields. Thus, it is necessary to reconstruct the local expression (3) with its own values of $\mathbf{h}$ from the sample data at every such area. The reconstructed function (3) is then used to estimate the thermodynamic temperature for every pixel of the 1024×1024 field, including the image of the Tomsk Region. Figure 1 shows the sample distribution of the thermodynamic temperature at the TR territory. This distribution was reconstructed using the data of weather reports from 16 reference observation sites. Influence of the atmospheric inhomogeneity is taken into account in the model (3) by using some measured values of the US temperature for such US areas, radiation brightness of which has simultaneously been recorded with the AVHRR device.

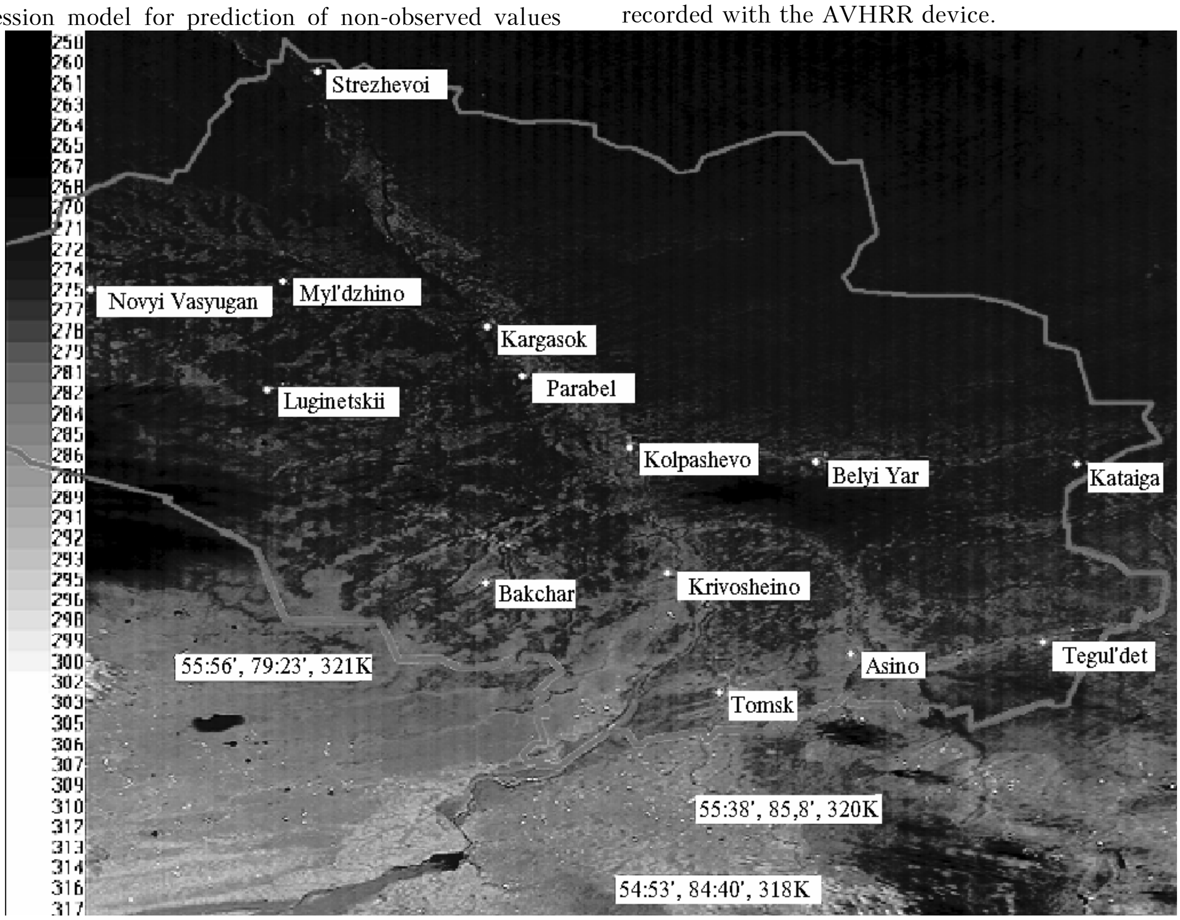

FIG. 1. Fragment of the temperature map of the Tomsk Region. The temperature (in K*) is demonstrated by grades of the gray color according to the color palette presented to the left.*

DECISION RULE FOR DETECTION OF THERMAL ANOMALIES

The AVHRR device allows temperature measurements only up to 45°C. So, we can have only the statistics describing observations in the situation with no thermal anomalies ("background" or "B" class). As to the observations in the situation with some thermal anomalies ("thermal anomaly" or "TA" class), they will be "cut off" at the level of 45°C. So, the background distribution should be reconstructed carefully and the threshold t of the decision rule should be defined by the level of false alarm $\alpha = \int_t^\infty f(y)\,dy$, where $f(y)$ is the probability density function of the background distribution. Let us denote the temperatures obtained at the first stage by Y_i, $i = 1, \ldots, n$ and reconstruct their probability distribution using the Johnson parametric family[5] with the chosen approximation S_B.

To construct the decision rule, let us use the following expression[5] for $f(y)$:

$$f_b(y) = \frac{\eta}{\sqrt{2\pi}} \frac{\lambda}{(y-\varepsilon)(\lambda - y + \varepsilon)} \times$$

$$\times \exp\left\{-\frac{1}{2}\left[\gamma + \eta \ln\left(\frac{y-\varepsilon}{\lambda - y + \varepsilon}\right)\right]^2\right\}, \quad (6)$$

$$\varepsilon \le y \le \varepsilon + \lambda, \quad \eta > 0, \quad -\infty < \gamma < \infty,$$

$$\lambda > 0, \quad -\infty < \varepsilon < \infty,$$

where ε (the lower limit of the value), as well as y and λ (sample range) can be estimated from observations $\{Y_j\}_{j=1}^n$, while the shape parameters η and γ are to be determined using, for instance, the method of maximum likelihood[5]:

$$\hat{\delta} = \left\{\sqrt{\frac{1}{n}\sum_{i=1}^{n} \tau^2(Y_i; \varepsilon, \lambda) - \left(\frac{1}{n}\sum_{j=1}^{n} \tau^2(Y_j; \varepsilon, \lambda)\right)^2}\right\}^{-1},$$

$$\hat{\gamma} = -\left[\frac{1}{n}\sum_{i=1}^{n} \tau(Y_i;\, \varepsilon,\, \lambda)\right]\hat{\delta},$$

where $\tau(y; \varepsilon, \lambda) = \ln[(y - \varepsilon)/(\lambda + \varepsilon - y)]$. Using Eq. (6) and having fixed the error level α, we can find the threshold t of the decision rule for TA detection.

Another approach to construction of the Bayes decision rule for detection of thermal anomalies is based on the fact that the histogram of reconstructed thermodynamic temperatures $\hat{f}(y)$ is a mixed distribution. It incorporates the temperature distribution in the situation with no thermal anomalies and that in the situation with some thermal anomalies observed at the territory. In this case, it is natural to use the approach associated with decomposition of the mixed distribution into mixture components each having its own weight. Suppose that the state of the nature in the situation B is described by the probability density function $f_0(y)$ and this, in the situation TA, is described by the probability density function $f_1(y)$ with unknown parameters. Then the next problem is to reconstruct the parameters of the mixture distributions $P f_0(y) + Q f_1(y)$ from the mixed histogram $\hat{f}(y)$ of the thermodynamic temperature distribution. Here, P and Q are *a priori* probabilities of the states B and TA of the nature, respectively, such that $P + Q = 1$. It should be noted that the TR territory always has thermal anomalies caused by the presence of about ten flares at ten gas deposits with known coordinates. This means that the TA class is not empty.

As $f_0(y)$ and $f_1(y)$, we have again used the distributions from the Johnson S_B family. The problem to be solved here is to optimize the quadratic quality criterion of identifying the following mixture:

$$J(\theta) = \frac{1}{m}\sum_{j=1}^{m} \{\hat{f}(y_j) - [P f_0(y_j) + Q f_1(y_j)]\}^2, \quad (7)$$

where m is the number of steps along the y axis, and $\theta = (P, \varepsilon_0, \lambda_0, \gamma_0, \eta_0, \varepsilon_1, \lambda_1, \gamma_1, \eta_1)^T$ is the vector of unknown parameters of the probability density functions $f_0(y)$ and $f_1(y)$ from the S_B family, respectively. The minimum in Eq. (7) was sought using the same adaptive search procedure (5). Once the mixture is identified and the distributions $f_0(y)$ and $f_1(y)$ reconstructed, the Bayes decision rule for detecting thermal anomalies takes the following form[6]:

$$u(y) = \arg\max \{P f_0(y),\ Q f_1(y)\}, \quad (8)$$

where $u(y)$ is the decision made about the presence (hypothesis H1) or absence (hypothesis H0) of TA.

The choice of the threshold with the use of the Johnson approximations for the reconstructed temperature field presented in Fig. 1 is illustrated in Fig. 2. The optimal criterion of quality of the approximation (7) was equal to 0.065 in this case.

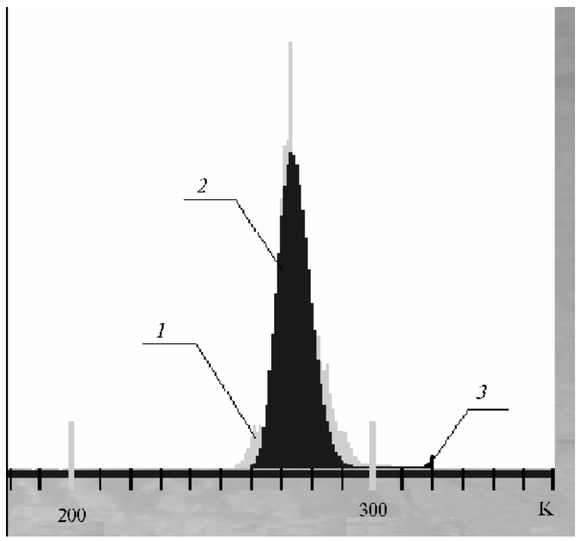

FIG. 2. The histogram of reconstructed temperatures (1) and Johnson approximations S_B for the observed B (2) and TA (3) class.

Figure 3 shows the results of detecting thermal anomalies during the NOAA measurement session on May 26, 1998. The fragments of images with thermal anomalies revealed by the algorithm have been referenced to the TR map via the "hot line" (the ArcView package) software.

When analyzing the video data, artifacts, such as flashes on water and clouds at low position of the Sun, should be considered. Those can easily be excluded using the threshold limits in the first and second channels, where the albedo of water and clouds is measured.

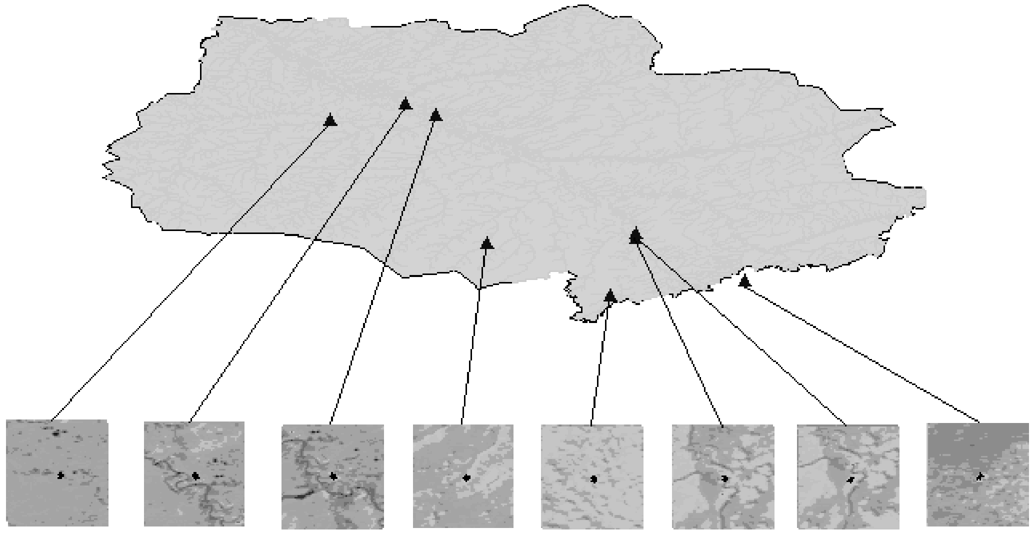

FIG. 3. Fragments of the video data measured with the AVHRR device from onboard the NOAA-14 satellite during the flight on May 26, 1998. Thermal anomalies detected by the program and assigned to particular places in Tomsk Region are marked on the image.

REFERENCES

1. V. Hardle, *Applied Non-parametric Regression* [Russian translation] (Mir, Moscow, 1993), 349 pp.
2. S.A. Aivazyan et al., *Applied Statistics* (Finansy i Statistika, Moscow, 1983), 471 pp.
3. F.P. Tarasenko, *Non-parametric Statistics* (State University, Tomsk, 1976), 294 pp.
4. Ya.Z. Tsypkin, *Adaptation and Learning in Automated Systems* (Nauka, Moscow, 1968), 400 pp.
5. G.J. Hahn and S.S. Shapiro, *Statistical Models in Engineering* (John Wiley & Sons, New York – London – Sydney, 1967).
6. H.L. Van Trees, *Detection, Estimation, and Modulation Theory*. Vol. 1. *Detection, Estimation, and Linear Modulation Theory* (John Wiley & Sons, New York – London – Sydney, 1968).