ISSN 0001-4370, Oceanology, 2016, Vol. 56, No. 3, pp. 315–319. © Pleiades Publishing, Inc., 2016.
Original Russian Text © M.A. Krinitskiy, A.V. Sinitsyn, 2016, published in Okeanologiya, 2016, Vol. 56, No. 3, pp. 341–345.

MARINE PHYSICS

Adaptive Algorithm for Cloud Cover Estimation from All-Sky Images over the Sea

M. A. Krinitskiy and A. V. Sinitsyn

Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia

e-mail: krinitsky@sail.msk.ru

Received June 5, 2015; in final form, July 15, 2015

Abstract—A new algorithm for cloud cover estimation has been formulated and developed based on the synthetic control index, called the grayness rate index, and an additional algorithm step of adaptive filtering of the Mie scattering contribution. A setup for automated cloud cover estimation has been designed, assembled, and tested under field conditions. The results shows a significant advantage of the new algorithm over currently commonly used procedures.

DOI: 10.1134/S0001437016020132

INTRODUCTION

Cloud cover (CC) is a key parameter in short-wavelength radiation parameterization [3]. In sea- and ground-based observations, CC is commonly estimated visually following the technique [1], which results in large uncertainties due to observer subjectivity. This prevents the use of these observations for both short-wave radiation estimation and validation of cloud parameters in reanalysis [5]. To automate observations and increase their accuracy, modern estimation techniques should be used, which are based on, e.g., data from wide-angle optical images of the visible sky hemisphere. Common sequentially produced packages for CC estimation can be of two types: packages with spherical mirrors equipped with standard-angle digital cameras and packages with fisheye lenses with viewing angles from 150° to 180° in the vertical plane and 360° in the horizontal plane. The peculiarities of digital shooting and algorithms used force one to use a sun-shading device (sun tracker) in the packages. Different techniques for CC estimation are similar except for the way of calculating the control index. The algorithms described in [7, 8] calculate the index by the equation $R_i = \frac{R}{B}$ (R and B are the red and blue components of pixel color in the RGB model). Useful pixels of an image are classified by discrimination with an empirically selected threshold value of this index $R_{i_{\text{thresh}}} = 0.8$. The algorithm [10] uses the so-called *SkyIndex*, calculated as $SI = \frac{B-R}{B+R}$. The threshold for classifying useful image pixels as CLEAR SKY and CLOUD has been also selected empirically: $SI_{\text{thresh}} = 0.23$. Again, CC is defined as

$$\text{CloudCover} = 10\,\text{int}\left[\frac{N_{\text{cloud}}}{N_{\text{total}}}\right], \tag{1}$$

where N_{cloud} is the number of pixels classified as CLOUD and N_{total} is the total number of useful image pixels.

More complex schemes similar to *SkyIndex* have been published; however, their use allows just an insignificant increase in CC estimation accuracy while solving only one of the problems typical of the observation method:

the need of sun-shading devices use;

classification of the solar disk and the region around it as cloudy area;

low sensitivity to thin clouds;

effect of water drops on the lens in the resultant all-aky images.

According to studies, though the accuracy of these algorithms is acceptable in some special cases, in general they result in a statistically biased distribution as compared to human-observed values (Fig. 1).

It should be noted that CC estimation for each image can be manually corrected adjusting the thresholds $R_{i_{\text{thresh}}}$ and SI_{thresh}, trying to achieve a statistically unbiased result as compared to human-observed values. However, the width of the distribution remains quite significant and there is no sense in automating observations in that case.

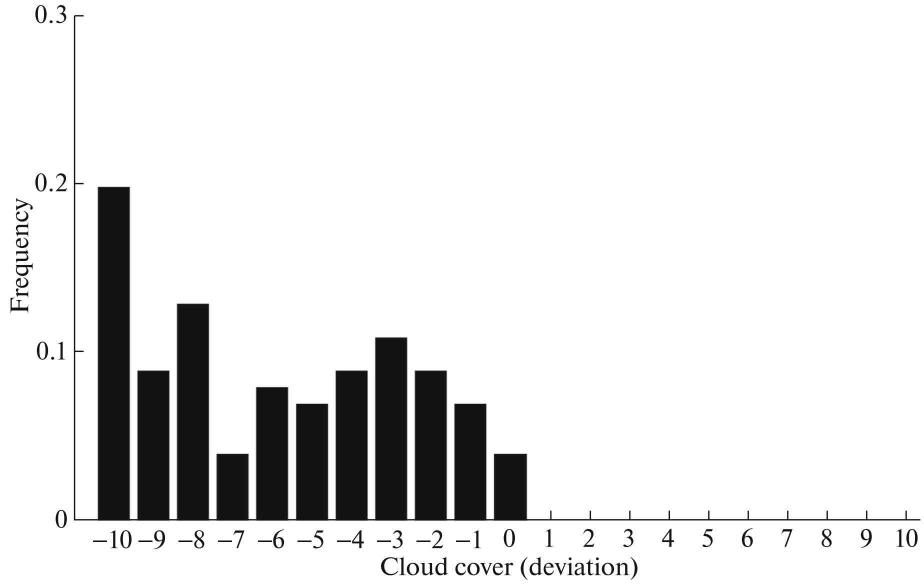

Fig. 1. Frequency of deviations of CC estimates by *SkyIndex* algorithm from observer evidence (based on data from 45th cruise of R/V *Akademik Ioffe*, 60° N, Atlantic Ocean, September 17–25, 2014).

PACKAGE FOR CC ESTIMATION

A novel device has been designed for complete automation of CC estimation. The basic principle of wide-angle imaging of the visible sky hemisphere remains unchanged: the key recording element is a digital optical camera with a wide-angle fisheye lens having viewing angles of 180° in the vertical plane and 360° in the horizontal plane. Ship rolling is an important factor, which introduces errors in all radiation measurements [2]. A principal difference is the use of an externally-mounted unit positioning system, which allows shooting under rolling conditions at the moment when the package is leveled. In addition, the externally-mounted unit is equipped with a GPS detector, which data allow calculation of the solar elevation at the shooting moment. Further positioning of the solar disk in an image is carried out with a special block of the image processing algorithm. This allows the external unit to be designed without a sun tracker and mechanical stabilization and orientation systems.

CC ESTIMATION ALGORITHM

To exclude the disadvantages of common algorithms, we have developed a new technique for CC estimation based on a new synthetic index, which we call the grayness rate index (*GrIx*). This index meaning is opposite to saturation in the HSV color model [6]:

$$GrIx = 1 - \frac{StdDev(R,G,B)}{Y}, \qquad (2)$$

where R, G, and B are the color components of a pixel in the RGB model and Y is the color brightness. The function $StdDev(R, G, B)$ is interpreted as the degree of spread of (R, G, B) values, which is analogous to the pixel saturation. Normalization to the intensity is used to eliminate the brightness dependence of the index. The resulting *GrIx* values observed in the images are within the range [0, 1]. Comparison of the indices R_i, *SI*, and *GrIx* has shown that the last index is the most sensitive to saturation gradients. Thus, this index allows a significant increase in the reliability of discrimination of thin clouds without adjustment of the threshold $GrIx_{thresh}$.

The block that estimates and, if necessary, filters the effects of sunlight scattering in the atmosphere in the part described by the Mie scattering is an individual branch of the *GrIx* algorithm (Fig. 2). The model used to estimate the Mie scattering contribution to image pixel saturation is empirical and adaptive. The model parameters are estimated for each image using statistical parameters of both the image (the distribution of *GrIx* values of pixels of the useful area of an image: median (m) and percentile 5 ($p5$)) and the optical camera used (density of the number of images over the space (m, $p5$) (Fig. 3) and image clustering results by the K-means algorithm [4] on the features space (m, $p5$) with a Euclidian metric of distance (Fig. 3)).

Image clustering on the space (m, $p5$) allows the following classification of images (Fig. 3).

Class 1: images without the solar disk and cloud gaps (or sparse gaps); there is no sense in estimating and filtering the Mie scattering contribution to distinguish thin clouds;

Class 2: images where the solar disk is highly probable and the fraction of clear sky is quite high for the Mie scattering contribution to be estimated and filtered in order to distinguish thin clouds;

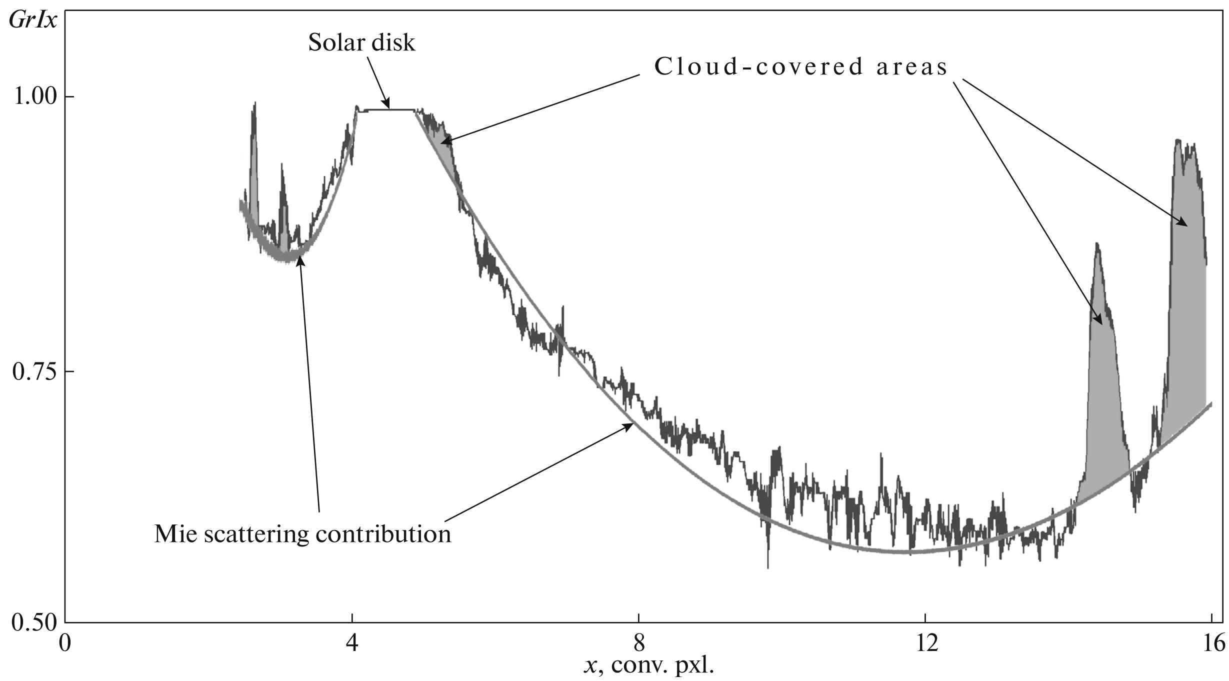

Fig. 2. Contribution of Mie scattering as background of useful signal. *GrIx* values along cross section passing through solar disk are shown.

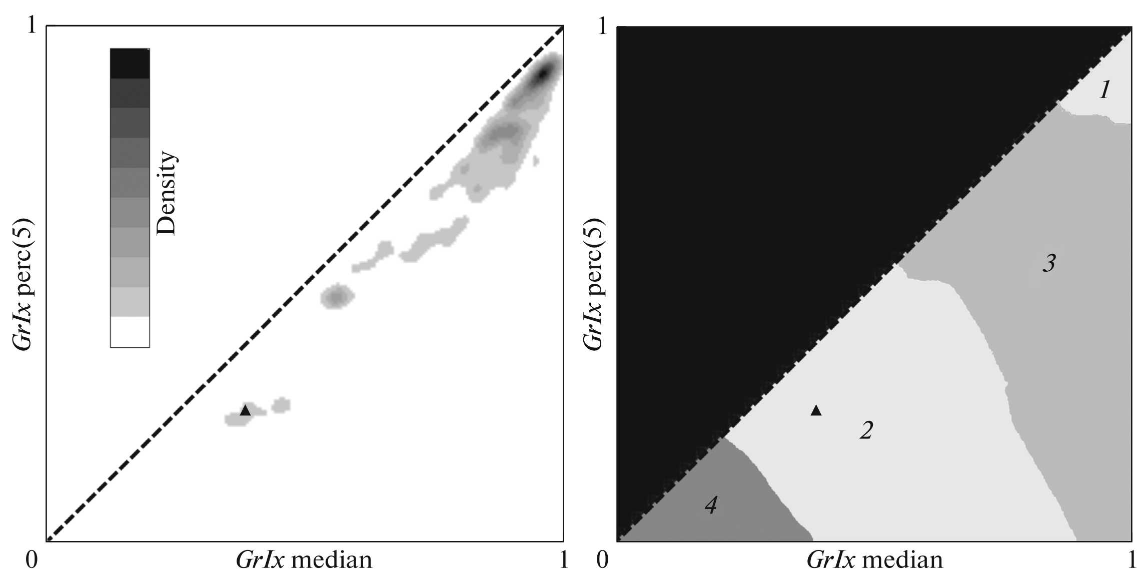

Fig. 3. Density of number of images over space (m, $p5$) (left) and space (m, $p5$) clustering for camera used on 45th cruise (right). Figures 1–4 show clustering classes *1–4*, respectively.

Class 3: images where it is difficult to detect the solar disk by statistical parameters; an individual branch of the *GrIx* algorithm has been developed for this class, which estimates the probability of the presence of the solar disk in an image and the need to filter the Mie scattering contribution;

Class 4: usually defective images with foreign objects or images taken at nighttime.

For example, the event (image) marked by the triangle in Fig. 3 relates to class 2, which entails estimation and filtering of the Mie scattering contribution and compensation of distortions introduced by it to the useful signal for CC signal estimation. Then, the image pixels are discriminated using $GrIx_{\text{thresh}}$.

Observations were carried out on the 45th cruise the R/V *Akademik Ioffe* under the MORE observation

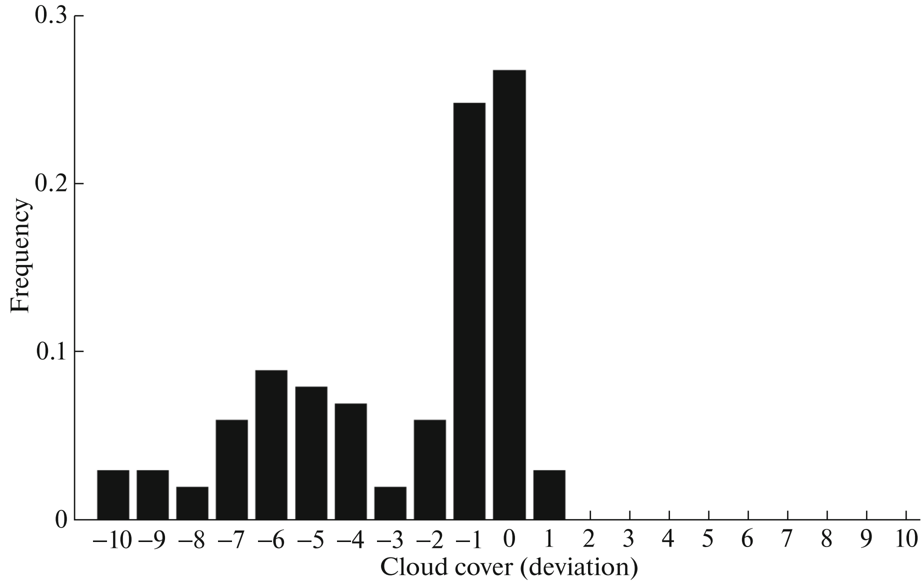

Fig. 4. Frequency of deviations of CC estimate by *GrIx* algorithm from observer evidence (based on data from 45th cruise).

program [9] in the Atlantic Ocean along 60° N from September 17 to 25, 2014. The cloud observation technique corresponded to that described in [1]. Digital recording of clouds was carried out and additional data of the above-described CC estimation package were recorded concurrently. According to the aforementioned technique, the statistical parameters of each image and the optical camera were calculated from the resulting set of digital all-sky images. Space (m, $p5$) was clustered using the K-means algorithm [4]. The results of image classification by this clustering were used for CC estimation by the *GrIx* technique.

To estimate the reliability of the results, deviations of the values calculated by the *GrIx* algorithm from human-observed values were analyzed (Fig. 4). The frequency of deviations shows that the technique used for CC estimation yields a more reliable result compared with common algorithms (Fig. 1). Accuracy is increased due to the use of the synthetic control index *GrIx* and the method suggested for filtering the Mie scattering contribution to the field of the useful signal. The study has demonstrated that the advantages of the *GrIx* algorithm are independent of image brightness, which allows this algorithm to be used in a wide range of weather and illumination conditions, as compared to common algorithms.

Detailed analysis has shown that cases of CC underestimation by the *GrIx* algorithm with deviations from human-observed values by –10 to –4 classes (using 10-classes CC scale) are caused by factors related to the observation procedure, in particular, with observations and shooting at a solar elevation less than 5°, which results in high errors in any case. No parameterizations of solar radiation are usually developed for these values of solar elevation. Another factor that affects the accuracy is the use of images from class 4 of clustering of the space (m, $p5$) (such images are inapplicable as source data). The developed algorithm allows automated long-term monitoring of CC. The data obtained can be used both to validate visual observations and directly calculate short-wave radiation fluxes. In addition, this package, once further improved, should allow measurement of many extremely important cloud parameters, in particular, optical depth and cloud base height, which cannot be measured in conventional ways. In particular, additional study of the data obtained on the 45th cruise of the R/V *Akademik Ioffe* makes it possible to talk about the potential of estimating the cloud base height with two synchronous images from two optical cameras of the described package and the optical stereoscopic effect. These data will make it possible to determine cloud type using the cloud base height.

ACKNOWLEDGMENTS

The author thanks Aleksei Sinitsyn of the Laboratory of Oceanic–Atmospheric Interaction and Climate Monitoring for the initial research data described in [2, 3, 9].

The work was supported by the Russian Science Foundation (analysis of data from the 45th cruise of R/V *Akademik Ioffe*, project no. 14-50-00095; development of the CC estimation algorithm, project no. 14-17-00697).

REFERENCES

1. *Recommendations for Hydrometeorological Stations and Offices,* No. 3, Part 1, Chap. 16: *Observation of the Clouds* (Gidrometeoizdat, Leningrad, 1985) [in Russian].
2. A. V. Sinitsyn and M. P. Alexandrova, "Assessment of direct determination uncertainties of incoming radiation fluxes associated with the ship motion," Oceanology (Engl. Transl.) **49** (4), 453–458 (2009).
3. A. V. Sinitsyn, M. P. Aleksandrova, and S. K. Gulev, "An improvement of parametrization of short-wave radiation at the sea surface on the basis of direct measurements in the Atlantic," Russ. Meteorol. Hydrol. **32** (4), 245–251 (2007).
4. J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles* (Addison-Wesley, New York, 1974).
5. E. Bedacht, S. K. Gulev, and A. Macke, "Intercomparison of global cloud cover fields over oceans from the VOS observations and NCEP/NCAR reanalysis," Int. J. Clim. **27** (13), 1707–1719 (2007). doi 10.1002/joc.1490
6. H. J. George and D. Greenberg, "Color spaces for computer graphics," in *Proceedings of the 5th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'78)* (Association for Computing Machinery, New York, NY, 1978), pp. 20–25. doi 10.1145/800248.807362
7. J. Kalisch and A. Macke, "Estimation of the total cloud cover with high temporal resolution and parameterization of short-term fluctuations of sea surface isolation," Meteorol. Z. **17** (5), 603–611 (2008).
8. C. N. Long and J. J. Deluisi, "Development of an automated hemispheric sky imager for cloud fraction retrievals," in *Proceedings of the. 10th Symposium on Meteorological Observations and Instrumentation, January 11–16, 1998* (Phoenix, AZ, 1998), pp. 171–174.
9. A. V. Sinitsyn, S. K. Gulev, A. Macke, et al., "2006: MORE cruises launched," Flux News, No. 1, 11–13 (2006).
10. M. Yamashita, M. Yoshimura, and T. Nakashizuka, "Cloud cover estimation using multi-temporal hemisphere imageries," Int. Arch. Photogr. Remote Sens. Spatial Inform. Sci. **35** (7), 826–829 (2004).

Translated by O. Ponomareva