Monitoring cumulus clouds by using global horizontal irradiance data

S V Zuev and N P Krasnenko

1 Institute of Monitoring of Climatic and Ecological Systems SB RAS, Tomsk, Russia
2 Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russia

zuev@imces.ru, krasnenko@imces.ru

Abstract. A previously proposed method for the detection of cumulus clouds by using time series of global horizontal irradiance from a pyranometer is developed. A characteristic feature of this method is the acceptance of short 21-minute moving analysis windows as stationary. This approach makes it possible not to construct a clear sky model for transforming a non-stationary daily time series of global horizontal irradiance to stationary time series.

1. Introduction
Cumulus clouds (Cu) are a basis for the formation of cumulonimbus clouds (Cb) – sources of various dangerous meteorological phenomena such as showers, thunderstorms, squalls, tornadoes, and hail. Monitoring of medium (4-7) Cu amount provides nowcasting information about the possibility of Cb formation. Satellite technologies cannot always provide continuous monitoring of clouds for various reasons. In turn, ground-based identification of cloud types is difficult to automate and it is most often performed by humans. Amplitude and variational analysis of the current actinometric data can be used to supplement or replace human observations of the sky. In the simplest case this can be global horizontal irradiance (GHI)

\[ Q = S \cdot \sin h + D, \]

the sum of direct normal \( S \) and diffuse horizontal \( D \) irradiance measured by an unshaded pyranometer (\( h \) is the height of the Sun).

Usually, before amplitude and variational analysis, a non-stationary GHI time series is converted to a stationary time series by using any clear-sky model. Many different methods for constructing such models are known [1-3]. In most cases, this requires some set of both: directly measured data (humidity, temperature, pressure, etc.) and reference data (dispersion coefficients of gases and aerosols, local average climatic characteristics, etc.). Measuring and processing such data usually has certain difficulties. Previously, the authors proposed a method for detecting Cu without using a clear sky model [4]. This is possible when short 21-minute samples of the GHI time series (± 10 minutes near the analysis point) are considered as stationary, and only variational analysis is applied to them.

2. Method
Variational analysis of global irradiance shows the opening and closing frequency of the solar disk by clouds with different densities, and it also shows the speed of these clouds over the observation point. Under stratiform clouds (Cs, As) and clear skies, the coefficient of variation (CV) is close to zero.
Waveform clouds ($Cc$, $Ac$, $Sc$, $St$) are more dynamic and under them the CV is middle. A feature of the cumulus clouds ($Ac$ cuf., $Sc$ cuf., $Cu$) is their high optical density and areas of clear skies between separate clouds or cloud ridges. There is a high coefficient of variation under these clouds. When dense opaque $Cu$ are present in the sky, the solar disk almost always has one of two stable states: either completely open and direct irradiance $S = S_{\text{max}}$, or completely closed and $S = 0$. As a result, the global irradiance has two values

$$
\begin{align*}
Q_{\text{max}} &= S_{\text{max}} \cdot \sin h + D, \\
Q_{\text{min}} &= D.
\end{align*}
$$

Figure 1. Example of GHI daily time series under $Cu$ at Ogurtsovo (WMO 29638) for August 6, 2015.

This leads to the formation of very heterogeneous time series of GHI (Figure 1). Using this feature, the detection of cumulus clouds can be reduced only to the search for heterogeneous intervals on the time series of global irradiance using a 21-min moving window.

Figure 2. Example of stable states $Q_{\text{min}}$ and $Q_{\text{max}}$ (left and right pictures, respectively) and unstable state (center picture) of the Sun under $Cu$ at Ogurtsovo (WMO 29638) for August 6, 2015.
Figure 3. Fragment of GHI daily time series for the example in Figure 2.

Studies have shown that the variation coefficient of a 21-minute sample of global irradiance $V_{21} \geq 0.33$ means that there are cumulus clouds $Cu\, hum.$ or $Cu\, med.$ in the sky with a probability of 70-80%. Such a coefficient value, $V = 0.33$, in statistical analysis is a threshold between homogeneous and heterogeneous series. It also turns out that stratus $St$ and nimbostratus $Ns$ clouds have similar CVs. This is due to high sensitivity of the variation coefficient to small changes in the low mean values of the analyzed sample.

To reduce or completely remove such errors, the use of some additional criterion is required. Another characteristic feature of the cumulus clouds is their distinct and unblurred edges and horizontal bases. Therefore, the transition between the states $Q_{\text{max}}$ and $Q_{\text{min}}$, as a rule, takes a short time of about 1-2 min (Figure 2). Thus, if within the analyzed 21-minute sample there is at least one sharp difference between the states $Q_{\text{max}}$ and $Q_{\text{min}}$, this is a sign of the presence of $Cu\, hum.$ and/or $Cu\, med.$ clouds (Figure 3). For instance, such differences can be found by analyzing the CV for a 3-min window ($\pm 1$ min near the analyzed point), which moves within a 21-minute sample, according to the condition $V_3 \geq 0.33$.

3. Results
The comparison of the $V_{21}$ and $V_{21} + V_3$ cumulus detector versions was made according to the GHI data for the warm period (May - August) 2018 ($h > 30^\circ$) from a CM-11 pyranometer (Kipp & Zonen, Netherlands) of the IMCES SB RAS geophysical observatory (Tomsk). The detection results were checked using a binary classifier and visual control by color panoramic images from an All Sky camera MVK-1653c (Byterg, Russia) installed near the pyranometer. In addition, one-hour visual observations of the sky were used.

Using a binary classifier, the state of an object can be attributed to one of two classes, "True" or "False" [5]. When monitoring cumulus clouds, there are four combinations for the result obtained (Figure 4):

1) **True Positive (TP) – $Cu$ is predicted, $Cu$ is present;**
2) **True Negative (TN) – $Cu$ is not predicted, $Cu$ is not present;**
3) **False Positive (FP) – $Cu$ is predicted, $Cu$ is not present;**
4) **False Negative (FN) – $Cu$ is not predicted, $Cu$ is present.**
The efficiency of both detector versions was compared by using three parameters - precision (Positive Predictive Value) 
\[ PPV = \frac{TP}{(TP + FP)} \], sensitivity (True Positive Rate) 
\[ TPR = \frac{TP}{(TP + FN)} \], and harmonic mean of precision and sensitivity (F–measure) 
\[ F_1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} \].

The results of cloud detection are presented in Table 1. Considering the difficulty of separating cumulus and cumulonimbus clouds above the observation point (by panoramic images), the determination of the sky state as Cu under the real state of Cb was taken as a positive result of the method (TP). It was also not considered an error to classify a low Cu amount (1-3) as its absence (TN).

**Table 1.** Comparison results of the Cu detector versions.

| Month | Version | Total analysis points | TP  | FP  | FN  | PPV | TPR | F1  |
|-------|---------|-----------------------|-----|-----|-----|-----|-----|-----|
| May   | V_{2j}  | 14806                 | 2821| 440 | 316 | 0.86| 0.90| 0.88|
|       | V_{2j} + V_3 | 14745             | 2051| 191 | 390 | 0.91| 0.84| 0.88|
| June  | V_{2j}  | 13255                 | 1885| 723 | 677 | 0.72| 0.74| 0.73|
|       | V_{2j} + V_3 | 13201             | 1556| 344 | 722 | 0.82| 0.68| 0.74|
| July  | V_{2j}  | 16913                 | 2628| 448 | 588 | 0.85| 0.82| 0.84|
|       | V_{2j} + V_3 | 16853             | 2621| 230 | 662 | 0.92| 0.80| 0.86|
| August| V_{2j}  | 14784                 | 2562| 592 | 913 | 0.81| 0.74| 0.77|
|       | V_{2j} + V_3 | 14732             | 1933| 306 | 1037| 0.86| 0.65| 0.74|
| Total (mean) | V_{2j}  | 59758                 | 9907| 2203| 2494| (0.81)| (0.80)| (0.81)|
|        | V_{2j} + V_3 | 59531             | 8188| 1071| 2817| (0.88)| (0.74)| (0.81)|

Table 1 shows that the use of the additional criterion V_3 increased the detection precision of cumulus clouds PPV by an average of 7%, with a decrease in the completeness of the sample TPR by 6%. At the same time, the number of false points FP decreased by more than 50%, and the harmonic mean F_1 did not change.

Table 2 shows that the errors with stratus St and nimbostratus Ns clouds are completely removed. The errors with cirrus clouds Ci decreased almost 7 times. Incorrect identification of cumuliform clouds Ac cuf. and Sc cuf. as Cu decreased by 30% and 50%, respectively. A high number of such errors is caused by the morphological similarity of such clouds and the complexity of their identification from panoramic All Sky images, which leads to subjective errors. Also, one of the detector V_{2j} + V_3 imperfections is incomplete filtering of the combined sky states Ci + Cu (Cb), which requires additional research.

**Figure 4.** Confusion matrix for Cu classification.
Table 2. Structure of the \textit{Cu} detector versions errors.

| Month | Version | Total analysis points | Total points $FP$ |
|-------|---------|-----------------------|------------------|
|       |         |                       | $Ci$ fib. | $Ac$ cuf. | $Sc$ cuf. | $St$ | $Ns$ |
| May   | $V_{21}$ | 14806                | 0       | 191      | 244      | 0    | 5    |
|       | $V_{21} + V_{3}$ | 14745            | 0       | 79       | 112      | 0    | 0    |
| June  | $V_{21}$ | 13255                | 125     | 365      | 223      | 0    | 18   |
|       | $V_{21} + V_{3}$ | 13201            | 17      | 233      | 94       | 0    | 0    |
| July  | $V_{21}$ | 16913                | 0       | 145      | 147      | 52   | 104  |
|       | $V_{21} + V_{3}$ | 16853            | 0       | 128      | 178      | 0    | 0    |
| August| $V_{21}$ | 14784                | 24      | 206      | 455      | 0    | 73   |
|       | $V_{21} + V_{3}$ | 14732            | 5       | 167      | 139      | 0    | 0    |
| Total (mean) | $V_{21}$ | 59758                | 149     | 907      | 895      | 52   | 200  |
|       | $V_{21} + V_{3}$ | 59531            | 22      | 607      | 442      | 0    | 0    |

4. Conclusions
The above-proposed and improved method allows one to monitor $Cu$ with a medium (4-7) cloud amount in near real time, with a delay of 10 min. The realization of this method does not require considerable computing power, which is often important when creating a monitoring network based on autonomous meters of integral and/or PAR global irradiance. In addition, the existing meteorological observing network can be used where continuous measurements of global irradiance are carried out by an unshaded pyranometer.

Acknowledgments
This research was supported by SB RAS under project of fundamental research no. IX.138.2.

References
[1] Duchon C E and O’Malley M S 1999 Estimating cloud type from pyranometer observations \textit{J. Appl. Meteorol.} \textbf{38} \textit{132-41}
[2] Badescu V, Guzyard C F, Cheval S, Oprea C, Baciu M, Dumitresku A, Iacobescu F, Milos I and Rada C 2012 Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models \textit{Renew. Sust. Energ. Rev.} \textbf{16} \textit{1636-56}
[3] Reno M J, Hansen C W and Stein J S 2012 \textit{Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis} (Albuquerque: Sandia National Laboratories) p 68
[4] Zuev S V and Krasnenko N P 2019 Simplified method for monitoring of Cumulus clouds using global horizontal irradiance \textit{IOP Conf. Ser.: Earth Environ. Sci.} \textbf{386} \textit{012044}
[5] Fawcett T 2006 An introduction to ROC analysis \textit{Pattern Recognit Lett} \textbf{27(8)} \textit{861-74}