

УДК [551.576+551.558.1]:551.507.362.2

Алгоритм распознавания и мониторинга облачности глубокой конвекции по данным МИСЗ на основе целочисленного программирования

А.Е. Шишов, И.А. Горлач

*Гидрометеорологический научно-исследовательский центр
Российской Федерации, г. Москва, Россия
ia_gorlach@mail.ru*

Измерения геостационарных метеорологических искусственных спутников Земли (МИСЗ) имеют большую ценность для мониторинга развития облачности глубокой конвекции (ОГК). Однако методы их обработки и возможности прогнозирования нельзя считать исчерпывающими. Цель данной статьи – описание нового алгоритма автоматизированного распознавания облачности глубокой конвекции и слежения за ее развитием на основе данных МИСЗ. На первом шаге алгоритма происходит распознавание объектов ОГК по пороговым значениям температуры; на втором шаге осуществляется поиск оптимального соответствия между распознанными объектами ОГК в последовательные моменты времени с помощью целочисленного программирования для оценки их смещения. Обосновывается преимущество предлагаемого алгоритма. Представлены результаты проверки качества его работы по данным за 11 и 12 июня 2019 года, когда над Европой и европейской частью России наблюдалось развитие мощной мезомасштабной конвективной системы. Характеристики распознанных объектов и их смещения соответствуют фактическим наблюдениям, что позволяет судить о корректности алгоритма.

Ключевые слова: искусственный спутник, облачность, распознавание, оптимизация, ассоциация, целочисленное программирование, трекинг, слежение

DOI: <https://doi.org/10.37162/2618-9631-2020-2-39-59>

An algorithm for the detection and tracking of deep convection using satellite data and integer programming

A.E. Shishov, I.A. Gorlach

*Hydrometeorological Research Center of Russian Federation,
Moscow, Russia
ia_gorlach@mail.ru*

Data from geostationary satellites are of great value for monitoring the development of deep convection. However, contemporary methods for their processing and forecasting capabilities are still far from perfect. The objective of the present article is to propose a new convective cloud detection and tracking algorithm based on satellite data. At the first stage of the algorithm, individual deep convective cells or systems are detected by a temperature thresholding technique. At the second stage, clouds detected at two consecutive time steps are matched to estimate their displacement using integer programming. Furthermore, the authors explain the strengths of the proposed method as they compare it to existing alternatives. Its performance is evaluated in a case-study for June 11-12, 2019.

During this period, a strong mesoscale convective system was observed over Europe and the European part of Russia. The correspondence of the features of the detected clouds and their displacement to observations suggests that the proposed algorithm is correct.

Keywords: satellite, deep convection, clouds, object detection, optimization, association, integer programming, object tracking

Введение

Ввиду большого потенциального ущерба как благосостоянию общества, так и здоровью людей, опасные явления, возникающие в виде гроз, града, ураганов и ливней, сильных порывов ветра, привлекают особое внимание метеорологов.

Известно, что одним из основных источников опасных явлений погоды считается облачность определенного типа, а именно, мощная кучево-дождевая облачность, и скопления конвективных облаков. Облака достигают больших высот и характеризуются большой вертикальной протяженностью, их верхняя граница может достигать уровня тропопаузы и превышать его, что осложняет оперативные наблюдения за ее развитием с поверхности Земли.

В настоящее время краткосрочный прогноз локальных конвективных явлений осуществляется с помощью негидростатических мезомасштабных моделей атмосферы, способных непосредственно воспроизводить глубокую конвекцию без использования каких-либо процедур ее параметризации. Пока численный прогноз нельзя назвать идеальным ввиду ошибок, возникающих из-за зависимости результатов моделирования от недостаточно высокого пространственного и временного разрешений моделей [1, 2] и недостаточного количества наблюдений.

Весьма полезными в этом случае оказываются большие массивы данных, полученные с геостационарных спутников. Однако методы их обработки по-прежнему далеки от идеала и поэтому привлекают внимание исследователей всего мира. Кроме того, рост интереса связан с повышением доступности вычислительных мощностей и развитием методов компьютерного зрения.

Цель данной статьи – представить новый алгоритм распознавания опасной кучево-дождевой облачности и слежения за нею (мониторинга, трекинга) на основе данных метеорологических искусственных спутников Земли (МИСЗ).

Достижение данной цели потребует выполнения нескольких задач. Во-первых, необходимо определить объект исследования и перечислить возникающие опасные конвективные явления, связанные с облачностью глубокой конвекции (ОГК), что требуется для формирования предупреждений при своевременном обнаружении и мониторинге облачности. Во-вторых, следует привести базовые принципы, использованные в задаче распознавания глубокой конвекции по данным МИСЗ. В-третьих,

безусловно, требуется представить описание разработанного алгоритма и выделить отличие от аналогов, предложенных другими исследователями. Наконец, качество работы алгоритма важно проверить, что и будет сделано на основе цифровых спутниковых данных наблюдений за период с 11 по 12 июня 2019 года, когда было зафиксировано большое число опасных явлений, связанных с ОГК в Европе.

Алгоритмы распознавания ОГК по данным МИСЗ

В последние годы было предложено множество инструментов для анализа глубокой конвекции по данным МИСЗ. Каждый из них можно разделить на три последовательных процедуры: чтение исходных данных, распознавание ОГК, трекинг ОГК, сохранение результатов (рис. 1).



Рис. 1. Краткая схема автоматизированного анализа глубокой конвекции по данным МИСЗ.

Fig. 1. Process flow diagram for the automated analysis of deep convection by using satellite data.

Реализация каждой процедуры зависит как от постановки задачи, так и от методов, используемых для решения. В частности, некоторые исследователи концентрируются на слежении (трекинге, мониторинге) уже образовавшихся конвективных штормов, например метод «RDT» [16], метод «MASCOTTE» [18]; либо при их выделении на этапе зарождения (инициации), например метод «SATCAST» [14]. Существуют и более универсальные методы и алгоритмы с учетом распознавания и мониторинга глубокой конвекции как на этапе возникновения, так и на более поздних стадиях развития конвективных образований. В качестве примера одного из наиболее успешных алгоритмов можно назвать метод «Cb-TRAM» [15, 23]. Авторы [18] представили исследование качества работы алгоритмов «SATCAST» и «Cb-TRAM» и установили довольно высокую долю ложных срабатываний, связанную с особенностями процесса возникновения конвекции.

В то же время несовершенство перечисленных алгоритмов мотивирует новые исследования, ведущие к их постепенному улучшению. В данной статье, воспользовавшись уже проверенным пороговым методом распознавания облачности глубокой конвекции, сосредоточимся на улучшении алгоритма слежения (мониторинга) или трекинга ОГК. Предлагаем способ реализации алгоритма, который позволяет эту задачу свести к задаче целочисленного линейного программирования.

Исходные данные

Основой для построения алгоритма послужили архивные материалы в виде ежедневных файлов, сформированных с дискретностью каждые 15 минут за несколько лет. Многоканальные спутниковые ежедневные измерения представлялись в виде бинарных записей. Каждая запись включала архив измерений в восьми каналах электромагнитного спектра, каждый канал составлял матрицу из 2312×1000 элементов (пикселей) ~ для $\frac{1}{4}$ части полного диска сканирования Земли с помощью прибора SEVIRI МИСЗ МЕТЕОСАТ-11 [19].

Остановимся кратко на особенностях выбранных каналов. Значения измерений инфракрасного канала в полосе прозрачности атмосферы (IR 10,8 мкм) позволяют оценить температуру поверхности Земли или верхней границы облачного покрова. Радиационная (измеряемая с помощью приборов МИСЗ) температура соответствует при минимальных отрицательных значениях большим высотам облачности над поверхностью Земли при доступности в любое время суток или зимним температурам заснеженных или оледеневших поверхностей Земли. Определение типа излучающей поверхности, верхней границы мощной облачности или подстилающей заснеженной поверхности не входило в наши задачи. Анализировались данные ИК-диапазона в умеренных широтах летом, при процессах, характерных для образования глубокой конвекции.

Излучение водяного пара (WV) в двух диапазонах (6,3 и 7,3 мкм) использовано для оценки наличия облаков в верхней и средней трети слоя тропосферы от ~2 км до 10–17 км. Поглощенное водяным паром излучение в нижних слоях оказывается незамеченным геостационарным спутником. Облачный покров, насыщенный водяным паром и достигающий больших высот в верхних слоях тропосферы (перистые и кучево-дождевые облака), напротив, удобнее анализировать с помощью измерений, сделанных в этом диапазоне.

Анализ данных видимого канала (радиация с длиной волны в диапазоне 0,6 мкм) дает возможность оценить более детальную, характерную для разных типов текстуру облачности [10, 16]. Однако отсутствие возможности измерений видимого канала в круглосуточном режиме не позволяет использовать преимущества таких измерений при мониторинге облачности.

Дополнительно использованы данные о разностях значений радиационной температуры в каналах электромагнитного излучения (WV 6,2 – IR 10,8) мкм и (WV 6,2 – WV 7,3) мкм, что позволило более точно выявить высокие, насыщенные влагой облака определенной текстуры, в том числе и относящиеся к глубокой конвекции. Выбор сделан на основе опубликованных в [12, 13] материалов.

Далее подробно опишем принцип работы алгоритма, в частности двух этапов: выделения и трекинга объектов.

Распознавание объектов с помощью метода «машинного зрения»

На первом этапе необходимо было получить маску объекта опасной кучевой облачности. Для получения начальной маски всех областей, относящихся к опасной кучевой облачности (ОГК), алгоритмом проводится сравнение измерений температуры каналов с длинами волн в диапазонах, описанных ранее, с заранее установленными пороговыми значениями в соответствии с литературными источниками.

Представим эту маску в виде матрицы C . Ее элементы можно вычислить с помощью функции f от трех переменных – измерений в трех диапазонах:

$$C_{ij} = f(T_{ij}^{10,8}, T_{ij}^{6,2}, T_{ij}^{7,3}); \quad (1)$$
$$j = 1, \dots, 2312; i = 1, \dots, 1000.$$

Функция f возвращает единицу при выполнении некоторого условия, нуль – в противном случае. Элемент матрицы с индексами (i, j) , (пиксель) можно считать частью конвективной ячейки ($C_{ij} = 1$), если выполняется ряд условий, каждое из которых опирается либо на попадание в заданные пределы температуры в определенном диапазоне длины волны, либо на заданные пределы значений разностей измерений температуры в двух диапазонах.

В литературе нет единого мнения относительно того, какие пороговые значения лучше всего использовать. В [20] проанализированы предельные значения для канала IR 10,8, использованные многими исследователями, и при построении собственного алгоритма выбрано значение 233 К. Мы остановимся на этом же значении.

Пороговые значения для разностей (WV 6,2 – IR 10,8) мкм и (WV 6,2 – WV 7,3) мкм были заимствованы из работ [12, 13]. Они провели статистический анализ измерений температуры во всех диапазонах, доступных для измерений с помощью прибора SEVIRI (МИСЗ MSG METEOSAT), для двух выборок: в случаях, когда имелись облака, вблизи которых за некоторый промежуток времени была зафиксирована гроза, и в случаях, когда имелись облака, которые не привели к возникновению опасных явлений. На основе полученных распределений температуры в [2] нами выбраны те же оптимальные пороговые значения. В таблице показаны результаты выбора порогов.

В итоге мы получили разбиение массива данных (исходного изображения) условно на два класса: «конвекция» (наличие ОГК) и «не конвекция» (отсутствие ОГК). Для анализа важно в первую очередь иметь маску каждого объекта, относящегося к классу «конвекция». Геометрическая форма этих объектов, как правило, такова, что индивидуальные пиксели группируются в кластеры выпуклой формы с четко выраженными центрами (для облаков более кратко назовем термином «верхушка»).

Таблица. Условия отнесения пикселя к классу глубокой конвекции (объекту)
Table. Threshold temperature values for assigning a pixel to the deep convection object

Величина	Значения температуры (К)	Описание
$T_{ij}^{10,8}$	< 233	Пороговое значение температуры верхней границы облака
$T_{ij}^{6,2} - T_{ij}^{10,8}$	> -10	Пороговое значение разности для верхней границы облака большой мощности.
$T_{ij}^{6,2} - T_{ij}^{10,8}$	> -4	Пороговое значение разности оценки вертикальной мощности в верхнем слое ВГО

Учет формы ОГК – используется алгоритм оценки эллипсообразной округлости формы объекта (наковальни). Подбор делается с помощью метода наименьших квадратов [8]. В результате процедуры получаем координаты центра эллипса, длины меньшей и большей его осей и углы их наклона. На основе этих данных получим матрицу E , каждый элемент которой, по аналогии с матрицей R , определяет принадлежность пикселя к эллипсам.

Далее рассчитаем два важных показателя:

1. Отношение меньшей оси к большей:

$K_1 = a/b$, где a – длина меньшей оси; b – длина большей оси.

2. Площадь пересечения объекта и его эллипса:

$K_2 = |I|/|U|$,

где $|I| = \{(i, j) : R_{ij} = n\} \cap \{(i, j) : E_{ij} = n\}$;

$|U| = \{(i, j) : R_{ij} = n\} \cup \{(i, j) : E_{ij} = n\}$.

Предполагаем, что мезомасштабные конвективные системы (МКС) не могут иметь слишком вытянутую форму, поэтому накладываем ограничение на приемлемые значения показателя K_1 . Далее исключаем случаи невыпуклой и угловатой формы по сравнению с идеальной. Учет таких случаев делается сравнением площади объекта и описывающего его эллипса с помощью показателя K_2 .

Проверку формы каждого объекта в алгоритме предлагается делать путем выделения области ОГК (ВГО ОГК) методом разрастания площади областей с учетом минимальных отклонений температуры ВГО ОГК для оценки соответствия следующему критерию:

$$\begin{cases} K_1 > 0,2 \\ K_2 > 0,7 \end{cases} \quad (2)$$

Только объекты, успешно прошедшие этот тест, признаются объектами ОГК. На следующем этапе распознавания и принадлежности

к одному и тому же объекту мы предлагаем применить простейший морфологический фильтр со структурным элементом

$$s = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (3)$$

для разметки связанных областей [8, 24] к маске, т. е. мы будем считать, что пиксель принадлежит объекту (ОГК), если

$$C_{ij} = C_{i+1,j} = C_{i,j+1} = C_{i-1,j} = C_{i,j-1}. \quad (4)$$

Таким образом мы получаем множество конвективных ячеек (когда ОГК в виде одного ядра – назовем ядро КЯ) или систем (когда ОГК в виде скопления – МКС или МКК). Для простоты мы будем называть их объектами. Очевидно, что каждый объект имеет свою маску, а значит, число объектов соответствует числу масок, которые мы получили на основе первоначальной маски C .

Так как маски характеризуют объекты в определенный момент времени, будем во всех дальнейших обозначениях применять верхний индекс t для привязки переменных к этому моменту времени.

Пусть Z_n^t – n -я маска в момент времени t . Размерность Z_n^t соответствует размерности исходной матрицы температуры, а ее элементы принимают значения из множества $\{0,1\}$.

Важно добавить, что маски объектов в один момент времени t не пересекаются, т. е. не имеют общих пикселей. Отсюда

$$\prod_{n=1}^{N^t} Z_n^t = 0, \quad (5)$$

где N^t – число объектов в момент времени t .

Трекинг выделенных объектов

Для установления связи между объектами, идентифицированными на шаге $(t-1)$ и шаге t необходимо сначала установить соответствие между объектами и масками. Мы предполагаем, что процедуру установления соответствия между известными объектами и новыми масками можно проводить последовательно, шаг за шагом.

Введем в рассмотрение случайную величину r_m^t ,

$$\text{где } r_m^t = \begin{cases} n, & \text{если } n\text{-я маска назначается } m\text{-му объекту;} \\ 0, & \text{если объект исчез} \end{cases} \quad (6)$$

$$m = 1, \dots, N^{t-1}, \quad n = 1, \dots, N^t, \quad t = 2, 3, \dots$$

Примечательно, что множество объектов непостоянно: на каждом шаге некоторые из старых объектов могут исчезнуть, а также могут

появиться новые. Это существенно усложняет задачу. С другой стороны, мы исключаем возможность «возвращения» объекта – ситуации, когда, исчезнув на каком-то предыдущем шаге, он появляется вновь. Отсюда следует, что если объект был выделен на шаге $(t - 1)$, на шаге t он отсутствовал и на шаге $(t + 1)$ появился вновь, то наш алгоритм зафиксирует его распад на шаге t , а затем рождение нового объекта.

Для наглядности приведем пример процесса установления соответствия между индексами распознанных ранее объектов с индексами новых масок: на первом шаге мы выделяем три маски и назначаем им объекты с индексами 1, 2, 3. На следующем шаге $(t = 2)$ мы выделяем пять масок с индексами 1, 2, 3, 4, 5 и устанавливаем взаимно однозначное соответствие между множествами $\{1, 2, 3\}$ и $\{1, 2, 3, 4, 5\}$. Третий шаг аналогичен второму.

Далее, так как в каждый момент времени мы знаем, к какому объекту какие пиксели относятся, можно проследить, куда смещается объект, его деформацию и рассчитать различные показатели, полезные для анализа.

Заметим, что поставленная задача решалась многими исследователями в рамках других алгоритмов трекинга конвективных штормов. Например, в пределах алгоритма «Сb-TRAM» и его многочисленных модификаций (например, «MASCOTTE») авторы опирались на сравнение фактических и прогнозируемых масок в момент времени [5]. Прогноз смещения и деформации объекта формировался на основе модели оптического потока, которая строилась на истории изменения матрицы наблюдений [23].

В качестве критерия схожести эти алгоритмы используют нормированную площадь пересечения двух объектов (их масок). То есть при сравнении n -й маски Z_n^t , сформированной в момент времени t , и прогнозируемого состояния k -й маски в момент времени $t \cdot \hat{Z}_k^t$, но выделенной в момент времени $(t - 1)$, рассчитывается показатель S_{kn}^t по формуле

$$S_{kn}^t = \frac{\sum_{i,j} Z_{nij}^t \hat{Z}_{kij}^t}{\sum_{i,j} \hat{Z}_{kij}^t}. \quad (7)$$

Далее назначение определяется по формуле:

$$\hat{r}_k^t = \arg \max_n S_{kn}^t. \quad (8)$$

Критерий схожести, используемый в нашем алгоритме, отличается тем, что прогнозируемые значения \hat{Z}_k^t заменяются фактическими значениями k -й маски Z_k^{t-1} в момент времени $(t - 1)$. Предполагаем, что смещение и деформация объекта за один временной шаг (15 минут) незначительны, а значит вычислительно сложной процедурой расчета оптического потока на данном этапе можно пренебречь.

Далее заметим, что формула определения назначения \hat{r}_k^t , используемая в методе «Сб-TRAM», может привести к некорректному назначению.

Чтобы исключить ситуацию некорректного решения при использовании формулы назначения метода «Сб-TRAM», мы предлагаем формулировать задачу установления соответствия между объектами в разные моменты времени (рис. 2) как оптимизационную с единой функцией ошибки. Похожий подход предлагался, например, в [9].

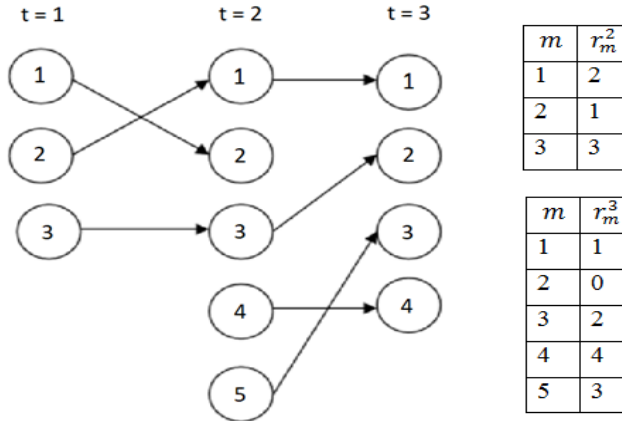


Рис. 2. Схема процесса установления соответствия между индексами распознанных ранее объектов с индексами новых масок.
Fig. 2. The Schematic illustration of the process of assigning mask indices to object indices.

Данная задача хорошо известна в области компьютерного зрения и называется задачей ассоциации. Если бы не возможность исчезновения старых объектов и возникновения новых, ее можно было бы переформулировать в задачу о назначениях во взвешенном двудольном графе, эффективные алгоритмы решения которых известны и активно применяются в контексте трекинга нескольких объектов [4, 22].

Таким образом, исходную задачу можно свести к задаче оптимизации с ограничениями. Предполагая, что оптимальное назначение на шаге t (в момент времени t) зависит лишь от предыдущего назначения, предыдущей маски и всех новых масок (выделенных в момент времени t), мы будем рассматривать этот процесс как марковский. Для краткости формул введем следующие обозначения:

$Z^t = \{Z_l^t, l = 1, \dots, N^t\}$ – множество масок, выделенных в момент времени t ;

$Z_{-n}^t = \{Z_l^t, l = 1, \dots, n-1, n+1, \dots, N^t\}$ – множество масок, выделенных в момент времени t , за исключением n -й.

Запишем формулу вероятности назначения n -й маски, выделенной на шаге t , m -му объекту, выделенному на шаге $(t-1)$.

По формуле Байеса для $n = 1, \dots, N^t$ имеем

$$P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}^t) = \frac{P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t) P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}_{-n}^t)}{\sum_{n=1}^{N^t} P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t) P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}_{-n}^t)}. \quad (9)$$

Условную вероятность исчезновения объекта ($n=0$) тогда можно задать формулой

$$P(r_m^t = 0 | Z_m^{t-1}, \mathbf{Z}^t) = 1 - \sum_{n=1}^{N^t} P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}^t). \quad (10)$$

Допустим, что априорные вероятности назначения на шаге t равны между собой (на шаге $(t-1)$ мы можем лишь догадываться о назначении на следующем шаге – "подбрасывать монету"), то есть

$$P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}_{-n}^t) = \frac{1}{N^t + 1}. \quad (11)$$

$$\forall n = 0, \dots, N^t$$

В таком случае выражение (9) можно значительно упростить за счет сокращения величин $P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}_{-n}^t)$:

$$P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}_{-n}^t) = \frac{P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t)}{\sum_{n=1}^{N^t} P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t)}. \quad (12)$$

Таким образом, определив функцию распределения $P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t)$, мы сможем давать оценку тому или иному назначению и использовать заданную функцию при принятии решений. Мы остановили свой выбор на формуле нормированной площади пересечения масок:

$$P(Z_n^t | r_m^t = n, Z_m^{t-1}, \mathbf{Z}_{-n}^t) = \frac{\sum_{i,j} z_{nij}^t z_{mij}^{t-1}}{\sum_{i,j} z_{mij}^{t-1}}, \quad (13)$$

где z_{nij}^t – элемент матрицы Z_n^t на пересечении i -й строки и j -го столбца.

Важно отметить, что $\sum_{i,j} z_{mij}^{t-1} \geq 1$, а значит, деление на нуль исключается. Кроме того, из неравенства $0 \leq z_{nij}^t \leq 1$ следует:

$$0 \leq \sum_{i,j} z_{nij}^t z_{mij}^{t-1} \leq \sum_{i,j} z_{mij}^{t-1}. \quad (14)$$

Тогда, очевидно,

$$0 \leq \frac{\sum_{i,j} z_{nij}^t z_{mij}^{t-1}}{\sum_{i,j} z_{mij}^{t-1}} \leq 1, \quad (15)$$

что соответствует определению функции распределения.

Составив из величин $P(r_m^t = n | Z_m^{t-1}, \mathbf{Z}^t)$ матрицу W размерности $N^{t-1} \times (N^t + 1)$, мы сможем найти оптимальные значения с помощью метода Global Nearest Neighbor Standard Filter (GNNSF), часто используемого для решения задачи ассоциации. Маски, оставшиеся без назначения (без соответствующего объекта), присваиваются новым объектам (наблюдается рождение нового объекта ОГК).

Метод GNNSF подразумевает поиск решения линейной задачи о назначениях с максимальной суммой (Linear Sum Assignment Problem) [3]. В конечном итоге эта задача сводится к задаче целочисленного линейного программирования с переменными, принимающими всего два значения: нуль и единица. Для этого случая есть алгоритмы полиномиальной сложности [17], что намного лучше экспоненциальной сложности, характерной для метода простого перебора.

Оценка качества работы алгоритма

Так как реализованный алгоритм состоит из двух этапов (распознавание объектов, трекинг) и ошибки возможны на каждом из них, полноценный анализ качества его работы также должен состоять из двух частей. Однако валидация каждого этапа в отдельности оказывается невозможной ввиду их тесной взаимосвязи, поэтому, как правило, более предпочтительной является оценка способности алгоритма выполнять поставленную задачу, например прогнозировать возникновение глубокой конвекции [6, 16].

На практике чаще всего используются три метода валидации:

1. Статистический анализ соответствия распознанных объектов наблюдаемым опасным явлениям (по географическим координатам и по времени).

2. Статистический анализ соответствия параметров полученных объектов параметрам, характерным для облачности глубокой конвекции (например, размеру, продолжительности жизни и траектории).

3. Детальное исследование отдельных случаев (case study) возникновения глубокой конвекции.

Для реализации первого метода валидации необходимо иметь доступ к обширной базе синоптических данных и данных сети грозоотметчиков.

Однако полученные оценки будут неточными: во-первых, наблюдения за грозами и измерения геостационарного спутника (распознанные объекты) будут чаще всего не согласованы по времени и географическим координатам; во-вторых, они не учитывают качества построенной траектории смещения объекта – неестественные «прыжки», слияния и распады не будут замечены [11]. Иными словами, данный метод не позволяет в полной мере проверить алгоритм трекинга.

Второй метод валидации, напротив, дает возможность оценить способность алгоритма следить за перемещениями объектов и их деформацией. С помощью анализа распределений продолжительности жизни и размера объектов можно сделать вывод о том, насколько они соответствуют процессу возникновения и развития облачности глубокой конвекции. Недостаток этого метода состоит в том, что на его основе нельзя судить о качестве распознавания самих объектов: т. е. можно ошибочно распознать скопления перистых облаков как кучево-дождевые, и тогда данный метод завысит оценку качества работы алгоритма.

Третий метод валидации опирается на экспертизу исследователя и небольшой объем значительно детализированных наблюдений за явлениями погоды в выбранные дни в выбранных регионах. Его преимущество – в более аккуратной оценке качества работы алгоритма. Более того, есть возможность проследить за каждой итерацией, сравнивая результаты с реальной синоптической ситуацией. Его недостаток в том, что сложно обеспечить репрезентативность выборки [10]. Этот метод использовался, например, для валидации алгоритма «ТООСАН» [6].

Учитывая перечисленные выше преимущества и недостатки, мы остановили свой выбор на третьем методе валидации. В качестве исходных данных для запуска алгоритма были выбраны 192 изображения (поля наблюдений МИСЗ) 11 и 12 июня 2019 года (с интервалом каждые 15 мин в течение 48 ч). Для проверки результатов его работы были использованы данные наблюдений наземной сети синоптических станций за рассматриваемый период, полученные по каналам связи Росгидромета, и комплексные данные сообщений из European Severe Storms Laboratory.

Алгоритму удалось выявить 856 объектов. Более 560 конвективных ячеек (КЯ) характеризовались малым сроком жизни, не превышающим 1 час: некоторые находились на этапе распада в начале рассматриваемого периода. Около 100 ячеек имели время жизни 1,5 ч. Всего 3–6 КЯ имели долгую продолжительность жизни – около 10 ч и лишь 1 или 2 КЯ наблюдались в течение 14–16 ч. Это согласуется с физическими законами развития конвекции разной интенсивности и мощности облачных образований. Возможно, некоторые не могли развиваться в мощные конвективные ячейки (ОГК) (рис. 3), а некоторые, скорее всего, были выявлены ошибочно.

Последнее замечание соответствует установленному свойству пороговых алгоритмов демонстрировать высокую долю «ложных

срабатываний» [14]. С одной стороны, это характеризует алгоритм как не идеальный. С другой стороны, данный недостаток нельзя считать критически значимым, так как основная задача алгоритма состоит в выявлении действительно опасных кучево-дождевых облаков (без пропусков) – важнее их не «упустить».

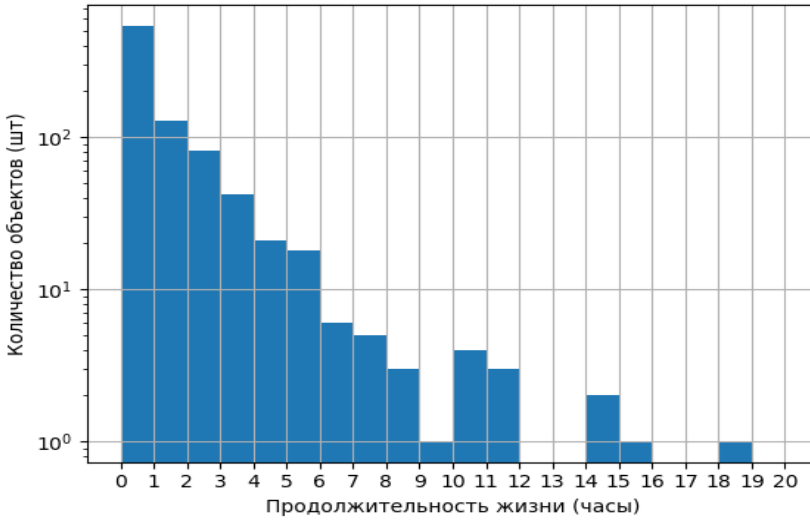


Рис. 3. Гистограмма распределения числа объектов (ОГК) в зависимости от продолжительности жизни выявленных объектов (лог. масштаб).

Fig. 3. The object lifetime histogram with log-scale for the y-axis.

Тем не менее хорошо видно, что алгоритму удалось также выделить объекты, продолжительность жизни которых была менее или около одного часа. Это значит, что процедура распознавания объектов и их перемещения работает и способна работать длительно.

Особого внимания заслуживает объект, продолжительность жизни которого превысила 18 ч. На самом деле он относится к МКС, которая существовала гораздо дольше. Алгоритму удалось проследить ее развитие как одного объекта с 00 ч ВСВ по 14.30 ч ВСВ 11 июня (в течение 14 ч 30 мин) (рис. 4). Затем система перестала выделяться алгоритмом. Рядом с ней был распознан новый объект, трекинг которого осуществлялся с 14.45 ч ВСВ 11 июня по 09.00 ч ВСВ 12 июня (18 ч 15 мин).

Около 8.00 ч ВСВ началось уменьшение площади объекта (рис. 5) и рост минимальной температуры верхней границы облака (ВГО) (рис. 6), а значит, и уменьшение высоты ВГО.

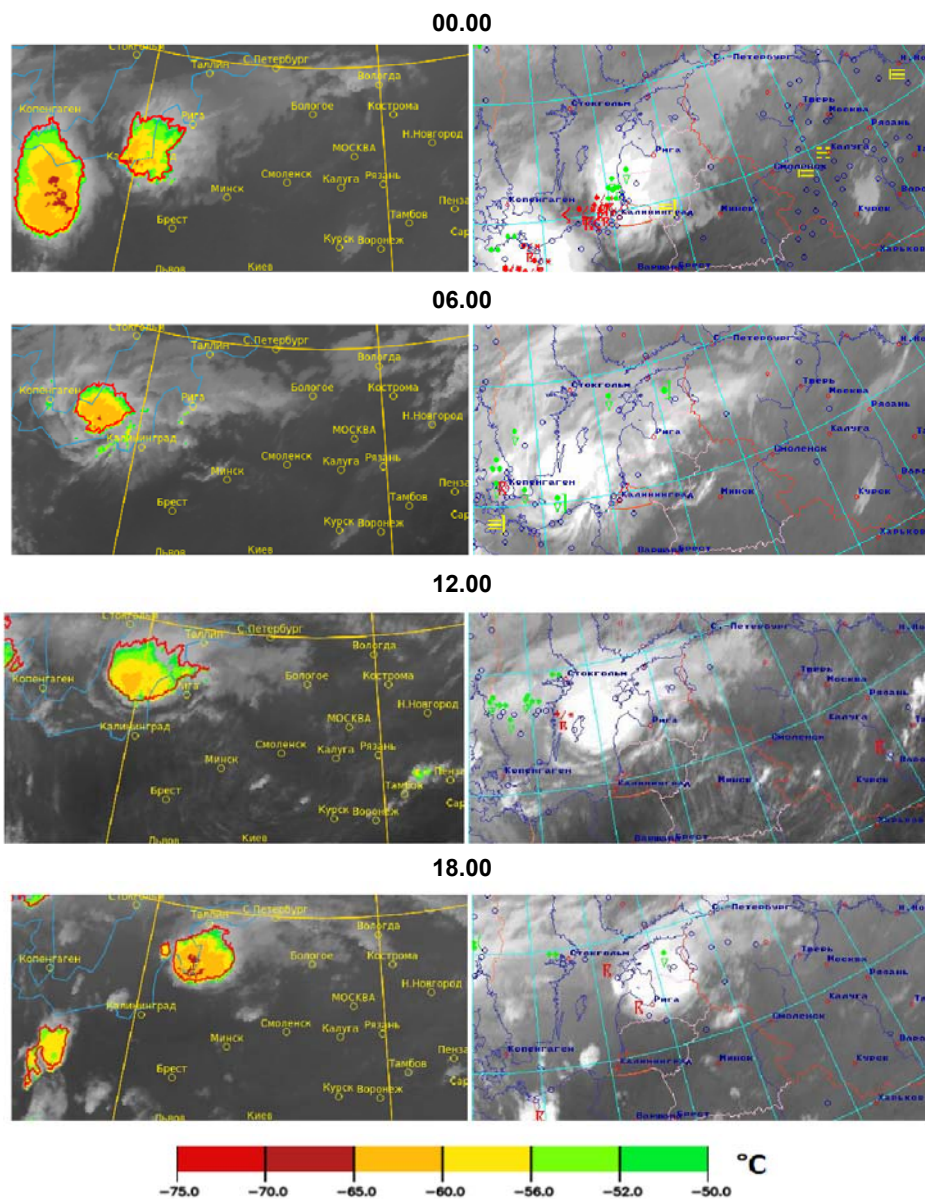


Рис. 4. Смещение МКС над восточной Европой 11 июня 2019 г. с 00 ч по 18.00 ч ВСВ. Визуализация результатов работы алгоритма — слева, объекты выделены красными контурами. Справа представлены данные синоптической сети наблюдений за опасными явлениями на фоне данных МИСЗ IR 10,8 мкм.

Fig. 4. The MCS movement over Eastern Europe on June 11, 2019, 12.00 - 06.00 UTC. The left subfigure illustrates detection and tracking results with the contours of detected objects colored red. The right subfigure illustrates severe weather observations reported by the synoptic network. IR 10.8 image is used as the background.

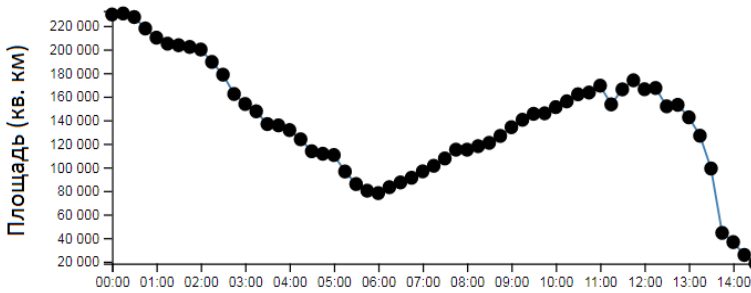


Рис. 5. Динамика площади МКС в период 00–14.30 ч ВСВ 11 июня 2019 г.
Fig. 5. Change in MCS area on June 11, 2019, 12 a.m. - 2.30 p.m.

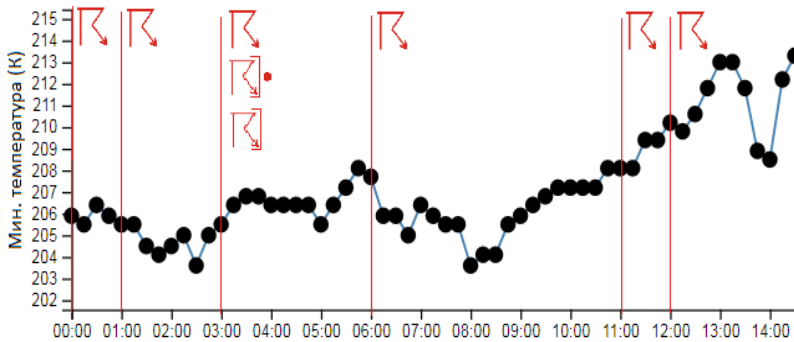


Рис. 6. Динамика распределения минимальной температуры облака ($T_{10.8 \text{ мкм}}$, К) верхней границы облаков в пределах МКС в период 00.00 – 14.30 ч ВСВ 11 июня 2019 г. Красными вертикальными линиями отмечены моменты времени, когда в районе объекта наблюдались грозы.
Fig. 6. MCS minimum cloud shield temperature dynamics (IR 10.8, K) on June 11, 2019, 12 a.m. – 2.30 p.m. UTC. Lightning occurrences within the object's area are indicated by red vertical lines.

Важно отметить, что за период с 12.00 по 15.00 ч ВСВ на территории, над которой находился рассматриваемый объект, не было зафиксировано ни одного случая грозы – это также свидетельствует об ослабевании МКС. Приблизительно для момента 14.15 ч ВСВ алгоритм выявил снижение интенсивности МКС, что привело к разделению на четыре более мелких и менее высоких ядра КЯ (рис. 7).

Отсутствие гроз и ливней на карте также подтверждает вывод алгоритма, что МКС не исчезла, но значительно понизилась степень ее опасности. Ядра КЯ в принципе могут регенерировать и усилиться позже и на некотором расстоянии при дополнительных условиях вынуждения.

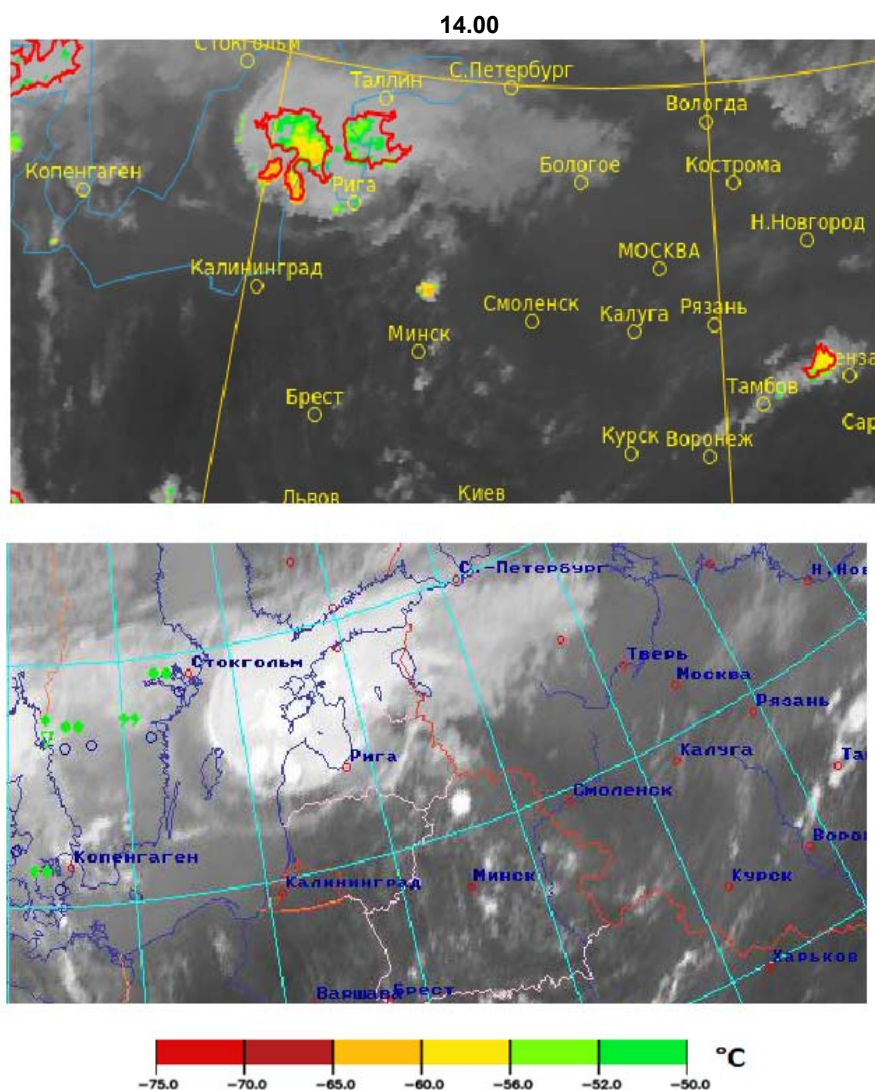


Рис. 7. МКС над восточной Европой в 14.00 ч ВСВ 11 июня 2019 г. Визуализация результатов работы алгоритма – слева, объекты ОГК выделены красными контурами. Справа представлена карта синоптических наблюдений на фоне спутниковых данных (T10,8 мкм).
Fig. 7. MCS over Eastern Europe on June 11, 2019, at 14.00 UTC. The left subfigure illustrates detection and tracking results with the contours of detected objects colored red. The right subfigure illustrates severe weather observations reported by the synoptic network. IR 10.8 image is used for the background.

Действительно, уже в 14.30 ч ВСВ алгоритм зафиксировал образование нового объекта в этой же области (рис. 8). Учитывая его географическое положение и отсутствие других объектов рядом, мы сделали вывод о том, что он является результатом регенерации ослабевшей, рассматриваемой ранее МКС.

Объект МКС перестал активно развиваться в ночь с 11 на 12 июня 2019 г.: наблюдался рост его площади (рис. 8), сопровождающийся небольшим понижением температуры верхней границы облачности в пределах контура (рис. 9), вплоть до полуночи, после чего площадь начала уменьшаться, а температура – расти. В 12.00 система окончательно распалась (алгоритм перестал ее выделять около 9.00 ч ВСВ).

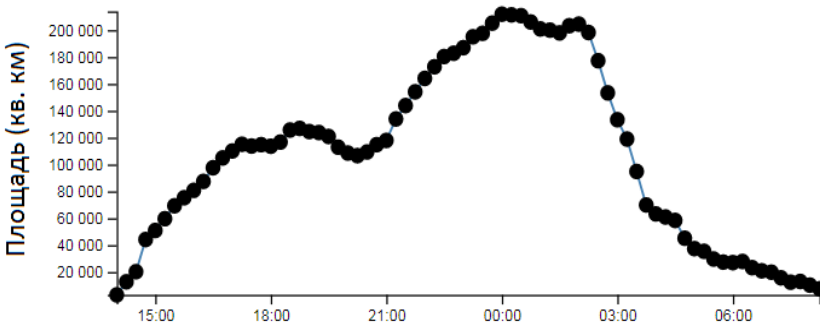


Рис. 8. Динамика площади МКС в период с 14.45 ВСВ 11 июня по 09.00 ВСВ 12 июня 2019 г.

Fig. 8. MCS area dynamics from June 11, 2019, 14.45 to June 12, 2019, 09.00.

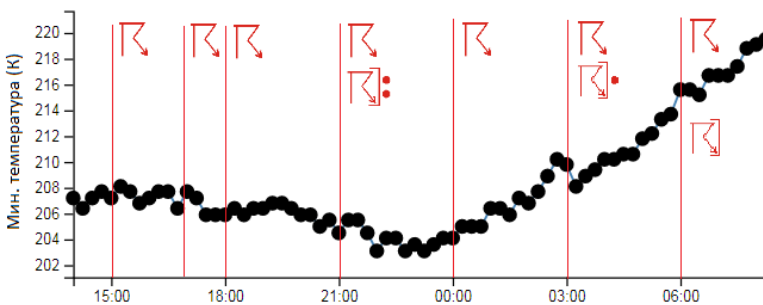


Рис. 9. Динамика минимальной температуры ВГО ($T_{10,8}$ мкм), регенерировавшей МКС с 14.45 ВСВ 11 июня по 09.00 ВСВ 12 июня 2019 г. Красными вертикальными линиями отмечены моменты времени, когда в районе объекта наблюдались грозы.

Fig. 9. MCS minimum cloud shield temperature dynamics (IR 10.8, T K) from June 11, 2019, 14.45 to June 12, 2019, 09.00. Lightning occurrences in the object's area are indicated by red vertical lines.

Таким образом, существенных отклонений от наблюдений в результатах работы алгоритма выявлено не было. Кроме того, контуры, выделяемые алгоритмом, соответствуют местам, где были действительно зафиксированы опасные явления, связанные с глубокой конвекцией. Следовательно, можно сделать вывод о корректности алгоритма и его практической ценности.

Выводы

В данной статье была описана реализация нового алгоритма распознавания опасной кучево-дождевой облачности и слежения за нею на основе данных метеорологических искусственных спутников Земли.

Первый этап работы алгоритма, связанный с выделением объектов (их маски), основан на рекомендациях из литературы и не имеет новизны.

Второй же этап – трекинг выделенных объектов – сводится к решению задачи ассоциации, которую, в свою очередь, мы свели к задаче целочисленного программирования с выбранными ограничениями. Была сформулирована функция ошибки и предложен алгоритм оптимизации. Кроме того, проведено сравнение предложенного подхода с аналогами других исследователей.

Корректность работы алгоритма была проверена на основе информации, наблюдаемой МИСЗ Метеосат-11 за период с 11 по 12 июня 2019 г. (48 часов). В этот день над Европой наблюдались несколько МКС, самая мощная из которых спровоцировала комплекс опасных явлений. Сведения о распознавании и динамике температуры верхней границы ОГК этой системы, сформированные в результате работы алгоритма, соответствуют выводам визуального анализа и данным European Severe Storms Laboratory. Поэтому можно сделать заключение, что качество работы предлагаемого алгоритма оптимизации для решения задачи распознавания и трекинга (мониторинга) ОГК удовлетворительно и его можно использовать в практической деятельности.

Список литературы

1. Васильев А.А., Вильфанд Р.М., Голубев А.Д. Совместное использование численных мезомасштабных и концептуальных моделей при оперативном прогнозе опасных явлений погоды // Труды Гидрометцентра России. 2016. Вып. 359. С. 48-57.
2. Калинин Н.А. Мониторинг, моделирование и прогноз состояния атмосферы в умеренных широтах. Пермь: Перм. гос. нац. исслед. ун-т, 2015. 308 с.
3. Blackman S.S., Popoli R. Design and Analysis of Modern Tracking Systems. Boston; London: Artech House, 1999. 1230 p.
4. Breitenstein M.D., Reichlin F., Leibe B., Koller-Meier E., Van Gool L. Robust tracking-by-detection using a detector confidence particle filter // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2009. P. 1515-1522.
5. Carvalho L.M., Jones C. A Satellite Method to Identify Structural Properties of Mesoscale Convective Systems Based on the Maximum Spatial Correlation Tracking Technique (MASCOTTE) // J. Appl. Meteorol. 2001. Vol. 40. P. 1683-1701.

6. *Carvalho L.M., Jones C.* A Satellite Method to Identify Structural Properties of Mesoscale Convective Systems Based on the Maximum Spatial Correlation Tracking Technique (MASCOTTE) // *J. Appl. Meteorol.* 2001. Vol. 40. P. 1683-1701.
7. *Fiolleau T., Roca R.* An Algorithm for the Detection and Tracking of Tropical Mesoscale Convective Systems Using Infrared Images From Geostationary Satellite // *Geoscience and Remote Sensing, IEEE Transactions.* 2013. Vol. 51. P. 4302-4315.
8. *Fiorio C., Gustedt J.* Two linear time Union-Find strategies for image processing // *Theoret. Comp. Sci.* 1996. Vol. 154. P. 165-181.
9. *Fitzgibbon A.* A Buyer's Guide to Conic Fitting. // *Proc. 5th British Machine Vision Conference, 1995.* P. 51.1-51.10.
10. *Han L., Fu S., Zhao L., Zheng Y., Wang H., Lin Y.* 3D Convective storm identification, tracking and forecasting – an enhanced TITAN algorithm // *J. Atm. Ocean. Tech.* 2009. Vol. 26. P. 719-732.
11. *Johnson J.T., Mackeen P.L., Witt A., Mitchell E.D., Stumpf G.J., Eilts D., Thomas K.W.* The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm // *Wea. Forecast.* June 1998. Vol. 13, no. 2. P. 263-276.
12. *Kuo K.S., Welch R.M., Weger R.C.* The three-dimensional structure of cumulus clouds over the ocean. 1. Structural analysis. // *J. Geophys. Res.* 1993. Vol. 98. P. 20685-20711.
13. *Lakshmanan V., Smith T.* An objective method of evaluating and devising storm-tracking algorithms // *Wea. Forecast.* April 2010. Vol. 25, no. 2. P. 701-709.
14. *Matthee R., Mecikalski J.R.* Geostationary infrared methods for detecting lightning-producing cumulonimbus clouds // *J. Geophys. Res.* 2013. Vol. 118. P. 6580-6592.
15. *Mecikalski J.R., MacKenzie W.M., Koenig M., Muller S.* Cloud-Top Properties of Growing Cumulus prior to Convective Initiation as Measured by Meteosat Second Generation. Part I: Infrared Fields // *J. Appl. Meteorol. Climatol.* 2010. Vol. 49. P. 521-534.
16. *Mecikalski J.R., Bedka K.M.* Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Daytime GOES Imagery // *Mon. Wea. Rev.* 2006. Vol. 134. P. 49-78.
17. *Merk D., Zinner T.* Detection of convective initiation using Meteosat SEVIRI: implementation in and verification with the tracking and nowcasting algorithm Cb-TRAM // *Atmos. Meas. Tech.* 2013. Vol. 6, no. 8. P. 1903-1918.
18. *Morel C., Senesi S.* A climatology of mesoscale convective systems over Europe using satellite infrared imagery. II: Characteristics of European mesoscale convective systems // *Q.J.R. Meteorol. Soc.* 2002. Vol. 128. P. 1973-1995.
19. *Nair U.S., Rushing J.A., Ramachadran R.K., Kuo S., Welch R.M., Graves S.J.* Detection of cumulus cloud fields in satellite imagery // *Proc. of SPIE Conference on Earth Observing Systems IV.* Denver, CO, SPIE, 2019. P. 345-355.
20. *Papadimitriou C.H.* On the complexity of integer programming // *J. ACM.* October 1981. Vol. 28, no. 4. P. 765-768.
21. *Reinhardt B., Dotzek N.* Validation of the MSG Convection Initiation Product (Final Report) // *EUMETSAT Technical Reports.* 2010.
22. *Schmetz J., Pili P.* An Introduction to METEOSAT Second Generation (MSG) // *BAMS.* 2002. P. 977-992.
23. *Silva Neto C.P., Alves Barbosa H., Assis Beneti C.A.* A method for convective storm detection using satellite data // *Atmósfera.* 2016. Vol. 29, is. 4. P. 343-358.
24. *Wu K., Otoo E., Shoshani A.* Optimizing connected component labeling algorithms // *Paper LBNL-56864 Lawrence Berkeley National Laboratory (University of California),* 2005.
25. *Xing J., Ai H., Lao. S.* Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses // *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* 2009. P. 1200-1207.
26. *Zinner T., Mannstein H., Tafferner A.* Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data // *Meteorol. Atmos. Phys.* 2008. Vol. 101. P. 191-210.

References

1. Vasiliev A.A., Vil'fand R.M., Golubev A.D. The joint use of mesoscale numerical and conceptual models in operational forecast of hazard weather phenomena. *Trudy Gidromettsentra Rossii [Proceedings of the Hydrometcentre of Russia]*, 2016, vol. 359, pp. 48-57 [in Russ.].
2. Kalinin N.A. Monitoring, modelirovanie i prognoz sostoyaniya atmosfery v umerennykh shirotah. Perm: Perm State University, 2015, 308 p. [in Russ.].
3. Blackman S.S., Popoli R. Design and Analysis of Modern Tracking Systems. Boston; London: Artech House, 1999, 1230 p.
4. Breitenstein M.D., Reichlin F., Leibe B., Koller-Meier E., Van Gool L. Robust tracking-by-detection using a detector confidence particle filter. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2009, pp. 1515-1522.
5. Carvalho L.M., Jones C. A Satellite Method to Identify Structural Properties of Mesoscale Convective Systems Based on the Maximum Spatial Correlation Tracking Technique (MASCOTTE). *J. Appl. Meteorol.*, 2001, vol. 40, pp. 1683-1701.
6. Carvalho L.M., Jones C.A. Satellite Method to Identify Structural Properties of Mesoscale Convective Systems Based on the Maximum Spatial Correlation Tracking Technique (MASCOTTE). *J. Appl. Meteorol.*, 2001, vol. 40, pp. 1683-1701.
7. Fiolleau T., Roca R. An Algorithm for the Detection and Tracking of Tropical Mesoscale Convective Systems Using Infrared Images From Geostationary Satellite. *Geoscience and Remote Sensing, IEEE Transactions*, 2013, vol. 51, pp. 4302-4315.
8. Fiorio C., Gustedt J. Two linear time Union-Find strategies for image processing. *Theoret. Comp. Sci.*, 1996, vol. 154, pp. 165-181.
9. Fitzgibbon A. A Buyer's Guide to Conic Fitting. Proc. 5th British Machine Vision Conference, 1995, pp. 51.1-51.10.
10. Han L., Fu S., Zhao L., Zheng Y., Wang H., Lin Y. 3D Convective storm identification, tracking and forecasting – an enhanced TITAN algorithm. *J. Atm. Ocean. Tech.*, 2009, vol. 26, pp. 719-732.
11. Johnson J.T., Mackeen P.L., Witt A., Mitchell E.D., Stumpf G.J., Eilts D., Thomas K.W. The storm cell identification and tracking algorithm: An enhanced WSR-88D algorithm. *Wea. Forecast.*, 1998, vol. 13, no. 2, pp. 263-276.
12. Kuo K.S., Welch R.M., Weger R.C. The three-dimensional structure of cumulus clouds over the ocean. 1. Structural analysis. *J. Geophys. Res.*, 1993, vol. 98, pp. 20685-20711.
13. Lakshmanan V., Smith T. An objective method of evaluating and devising storm-tracking algorithms. *Wea. Forecast.*, April 2010, Vol. 25, no. 2, pp. 701-709.
14. Matthee R., Mecikalski J.R. Geostationary infrared methods for detecting lightning-producing cumulonimbus clouds. *J. Geophys. Res.*, 2013, Vol. 118, pp. 6580-6592.
15. Mecikalski J.R., MacKenzie W.M., Koenig M., Muller S. Cloud-Top Properties of Growing Cumulus prior to Convective Initiation as Measured by Meteosat Second Generation. Part I: Infrared Fields. *J. Appl. Meteorol. Climatol.*, 2010, vol. 49, pp. 521-534.
16. Mecikalski J.R., Bedka K.M. Forecasting Convective Initiation by Monitoring the Evolution of Moving Cumulus in Daytime GOES Imagery. *Mon. Wea. Rev.*, 2006, vol. 134, pp. 49-78.
17. Merk D., Zinner T. Detection of convective initiation using Meteosat SEVIRI: implementation in and verification with the tracking and nowcasting algorithm Cb-TRAM. *Atmos. Meas. Tech.*, 2013, vol. 6, no. 8, pp. 1903-1918.
18. Morel C., Senesi S. A climatology of mesoscale convective systems over Europe using satellite infrared imagery. II: Characteristics of European mesoscale convective systems. *Q.J.R. Meteorol. Soc.*, 2002, vol. 128, pp. 1973-1995.
19. Nair U.S., Rushing J.A., Ramachandran R.K., Kuo S., Welch R.M., Graves S.J. Detection of cumulus cloud fields in satellite imagery. Proc. of SPIE Conference on Earth Observing Systems IV. Denver, CO, SPIE, 2019, pp. 345-355.

20. Papadimitriou C.H. On the complexity of integer programming. *J. ACM.*, October 1981, vol. 28, no. 4, pp. 765-768.
21. Reinhardt B., Dotzek N. Validation of the MSG Convection Initiation Product (Final Report). *EUMETSAT Technical Reports*, 2010.
22. Schmetz J., Pili P. An Introduction to METEOSAT Second Generation (MSG). *BAMS*, 2002, pp. 977-992.
23. Silva Neto C.P., Alves Barbosa H., Assis Beneti C.A. A method for convective storm detection using satellite data. *Atmósfera*, 2016, vol. 29, no. 4, pp. 343-358.
24. Wu K., Otoo E., Shoshani A. Optimizing connected component labeling algorithms. Paper LBNL-56864 Lawrence Berkeley National Laboratory (University of California), 2005.
25. Xing J., Ai H., Lao. S. Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses. *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1200-1207.
26. Zinner T., Mannstein H., Tafferner A. Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorol. Atmos. Phys.*, 2008, vol. 101, pp. 191-210.

Поступила в редакцию 20.05.2020 г.

Received by the editor 20.05.2020.