

## Использование данных лидара CALIOP для восстановления высоты нижней границы многослойной облачности по спутниковым данным MODIS на основе методов нечёткой логики

А. В. Скороходов

*Институт оптики атмосферы им. В. Е. Зуева СО РАН, Томск, 634055, Россия  
E-mail: vazime@yandex.ru*

Представлен алгоритм восстановления высоты нижней границы отдельных ярусов многослойной облачности по данным пассивного спутникового зондирования на основе методов нечёткой логики. Процедура оценки высоты основания облачности рассмотрена в качестве частного случая решения задачи классификации, где классами являются узкие диапазоны значений искомого параметра, а классификационными признаками — характеристики облаков, восстанавливаемые по данным пассивного спутникового зондирования. Методы нечёткой логики позволяют относить один объект к нескольким классам одновременно, но с различной степенью принадлежности, что обеспечивает возможность оценки высоты нижней границы для нескольких ярусов облаков, наблюдаемых одновременно. Обучение классификатора осуществляется на основе синхронных данных лидара CALIOP (*англ.* Cloud-Aerosol Lidar with Orthogonal Polarization) и спектрорадиометра MODIS (*англ.* Moderate Resolution Imaging Spectroradiometer), полученных летом над территорией Западной Сибири в период 2013–2018 гг. А его функционирование выполняется уже только по данным пассивного спутникового зондирования. Рассмотрено два метода нечёткой самоорганизации: Fuzzy C-means и Густафсона–Кесселя. Установлено, что второй подход является более эффективным и обеспечивает отклонение восстановленных значений высоты нижней границы облаков с оптической толщиной меньше 10 от эталонных в  $-0,5$  км при среднеквадратическом отклонении  $1,5$  км для вышележащего слоя облачности и  $-0,1$  км при  $2,1$  км для расположенного под ним.

**Ключевые слова:** высота нижней границы облаков, методы нечёткой логики, многослойная облачность, нейронная сеть, спутниковые данные, CALIOP, MODIS

Одобрена к печати: 16.01.2024  
DOI: 10.21046/2070-7401-2024-21-1-122-134

### Введение

Высота нижней границы облаков (ВНГО) — одна из наиболее важных их характеристик, используемая при решении большого числа задач климатологии и метеорологии. Информация о данном параметре необходима для обеспечения безопасности, контроля и планирования полётов летательных аппаратов различного типа и назначения, для оценки степени воздействия облачности на радиационный перенос и транспорт атмосферного и антропогенного аэрозоля, а также для радиолокации, например при выборе рабочего диапазона частот (Кхьюнг, 2020; Gebremariam et al., 2018; Mecikalski et al., 2007). Наземные и спутниковые средства активного зондирования, хотя и позволяют определять ВНГО «напрямую» с хорошей точностью, но не обеспечивают ежедневного глобального покрытия измерениями (особенно в высоких широтах и над поверхностью Мирового океана) из-за их недостаточного количества или низкой периодичности (Борейшо и др., 2019; Толмачева, Крючкова, 2013; Winker et al., 2009). Поэтому с начала 2000-х гг. интенсивно стали развиваться методы восстановления искомого параметра по данным пассивного спутникового зондирования (Forsythe et al., 2000; Hutchison, 2002; Wilheit, Hutchison, 2000). Современные геостационарные системы метеорологического назначения позволяют осуществлять глобальный мониторинг поверхности Земли с пространственным разрешением до 1000 м и с периодичностью до 10 мин (Bessho et al., 2016).

Основной недостаток пассивных спутниковых инструментов применительно к задаче восстановления ВНГО состоит в том, что «напрямую» им доступна информация исключительно о верхней границе облачности. Поэтому её решение может осуществляться только косвенно, на основе нахождения взаимосвязей между искомым параметром и другими характеристиками облаков, которые могут быть определены по данным таких сенсоров, как, например, MODIS (*англ.* Moderate Resolution Imaging Spectroradiometer) (спутник Aqua) (Platnick et al., 2017). В настоящее время существует три основных подхода к восстановлению ВНГО по данным пассивного спутникового зондирования: использование эмпирических зависимостей, применение концепции «донор—реципиент» и машинное обучение (Chen et al., 2020; Minnis et al., 2019; Noh et al., 2017). Их объединяет общая идея использования синхронных активных и пассивных наблюдений за облаками и сопоставления полученных данных. При этом чаще всего применяются результаты измерений искомого параметра лидаром CALIOP (*англ.* Cloud-Aerosol Lidar with Orthogonal Polarization) (спутник CALIPSO — *англ.* Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) и радара CPR (*англ.* Cloud Profiling Radar) (спутник CloudSat — *англ.* Cloud Satellite) в качестве эталонных. Наиболее эффективные алгоритмы восстановления ВНГО по данным оптических радиометров обеспечивают среднюю оценку смещения до  $-0,2$  км при среднеквадратическом отклонении  $1,2$  км (Minnis et al., 2021; Skorokhodov et al., 2023; Tan et al., 2021). Однако такие высокие показатели достигаются только для однослойной облачности. Наличие облаков в нескольких ярусах одновременно существенно искажает спектральные сигнатуры, регистрируемые пассивными сенсорами, что оказывает влияние на точность восстановления физических параметров облачности традиционными методами (Chang, Li, 2005; Teng et al., 2020).

Проблема восстановления высоты основания отдельных ярусов многослойной облачности затрагивалась только в исследованиях (Minnis et al., 2019; Noh et al., 2022; Tan et al., 2022). В работе (Minnis et al., 2019) многослойный перцептрон обучался на двух наборах данных. В первом из них облака различных ярусов рассматривались как единое целое, если расстояние между ними было менее  $3$  км, и ВНГО соответствовала самому нижнему из них. Второй набор характеризовал облачность с большим расстоянием между слоями, и высота основания восстанавливалась только для самого верхнего яруса. В исследовании (Noh et al., 2017) использовалась эмпирическая регрессионная модель для однослойных облаков. А уже на этапе тестирования было установлено, что в случаях наблюдения многослойной облачности ВНГО соответствовала либо самому нижнему слою из всех наблюдаемых, либо самому верхнему. Наилучшие результаты в данной области были достигнуты в работе (Tan et al., 2022), где высота основания определялась для каждого из наблюдаемых одновременно облачных ярусов. Здесь применялся алгоритм из исследования (Tan et al., 2021) для однослойной облачности совместно с методом экстраполяции результатов его использования, согласно которому ВНГО отдельных ярусов многослойных облаков можно оценить по высоте основания соседней однослойной облачности, исходя из предположения, что они расположены на одном уровне. Впервые такой подход был применён в исследовании (Mülmenstädt et al., 2018) для восстановления ВНГО облаков глубокой конвекции на основе значений этого параметра на их краях по данным пассивного спутникового зондирования.

Упомянутое выше предположение верно далеко не всегда, например не для облачных систем атмосферных фронтов, когда ВНГО может различаться на несколько километров в пределах одного поля облачности (Дашко, 2005; Облака..., 1989; Шакина, 1985). Кроме этого, соседние однослойные облака могут располагаться на значительном удалении от исследуемого поля многослойной облачности (с другой синоптической обстановкой) или вообще отсутствовать в пределах области снимка, что не позволит применить данный подход. Поэтому решение задачи восстановления ВНГО отдельных ярусов многослойной облачности остаётся по-прежнему актуальным. Одним из возможных вариантов представляется использование методов нечёткой логики, когда исследуемый образец может быть отнесён к нескольким классам одновременно, но с различной степенью принадлежности. В частности, такой подход был успешно применён в исследовании (Астафуров, Скороходов, 2015) для распознавания разновидностей облаков отдельных ярусов многослойной облачности.

Цель настоящей работы заключается в разработке алгоритма восстановления ВНГО многослойной облачности по данным пассивного спутникового зондирования на основе сопоставления результатов синхронных наблюдений за облаками лидаром CALIOP и спектрорадиометром MODIS и применения методов нечёткой логики.

## Исходные данные

В работе используются результаты тематической обработки синхронных (разница 2 мин) наблюдений за облаками активным сенсором CALIOP и пассивным инструментом MODIS, проведённых летом (с мая по сентябрь) в светлое время суток над территорией Западной Сибири в период с 2013 по 2018 г. В указанные месяцы над рассматриваемым регионом наблюдаются практически все типы облачности согласно международной классификации Всемирной метеорологической организации (ВМО) (Код..., 2013). При этом для обучения алгоритма восстановления ВНГО используются результаты наблюдения однослойных облаков лидаром и спектрорадиометром, взятые из публикации (Скороходов, Курьянович, 2022). Соответственно, и ограничения его применения по зенитному углу Солнца ( $\theta$ ) и оптической толщине облачности ( $\tau$ ) будут аналогичными:  $20 < \theta < 60^\circ$  и  $\tau < 10$ . Кроме этого, рассматриваются те же характеристики облаков и вспомогательные параметры окружающей среды:  $\tau$ , эффективный радиус частиц ( $r_{эф}$ ), водозапас ( $P$ ), фазовый состав ( $Q$ ), ВНГО ( $h_{НГ}$ ), высота ( $h_{ВГ}$ ), температура ( $T_{ВГ}$ ) и давление ( $p_{ВГ}$ ) на верхней границе облачности, её эффективная излучательная способность ( $\epsilon$ ), температура ( $T_{ПП}$ ) и тип ( $C$ ) подстилающей поверхности, а также  $\theta$ . Отметим, что в работе использовались значения  $\tau$ ,  $r_{эф}$  и  $P$ , рассчитанные по данным спектрального канала 2,2 мкм, а  $h_{НГ}$ ,  $T_{ВГ}$ ,  $p_{ВГ}$  и  $\epsilon$  — по данным спектрального канала 11 мкм. Указанные параметры извлекаются из тематических продуктов CALIOP и MODIS с пространственным разрешением 1000 м: MYD06\_L2, CLDPROP\_L2 и CAL\_LID\_L2\_01kmCLay-Standard (<https://ladsweb.modaps.eosdis.nasa.gov>, <https://www-calipso.larc.nasa.gov>). Объём обучающей выборки составляет 40 275 образцов однослойной облачности.

Тестирование алгоритма восстановления ВНГО осуществляется на основе результатов наблюдения многослойной облачности лидаром CALIOP и спектрорадиометром MODIS за рассматриваемый в работе период времени. При этом в качестве тестовых образцов используются только те облачные пиксели с  $\tau < 10$ , для которых наличие облаков в нескольких ярусах одновременно подтверждается значениями флагов Cloud\_Multi\_Layer\_Flag и Number\_Layers\_Found в продуктах MYD06\_L2 и CAL\_LID\_L2\_01kmCLay-Standard соответственно, а ВНГО являются истинными (т. е. лазерный импульс полностью не затухает при прохождении сквозь слои облачности) согласно флагу Opacity\_Flag в последнем из них. Объём тестовой выборки составляет 2622 образца многослойных облаков.

На *рис. 1* (см. с. 125) показан один из рассматриваемых в работе эпизодов наблюдения многоярусных облаков над Западной Сибирью, а также некоторые статистические характеристики такой облачности в этом регионе за период 2013–2018 гг. Зелёной линией здесь обозначено местоположение трассы лидара CALIOP на снимке MODIS, а красным выделена многослойная облачность с известными значениями ВНГО в каждом из её ярусов. Сопоставляя результаты активного и пассивного зондирования на *рис. 1а* и *б*, можно установить, что здесь наблюдаются перистые нитевидные (*Ci fib*) облака над кучевыми плоскими (*Cu hum*), а расстояние между слоями составляет приблизительно 7 км. В целом по Западной Сибири наиболее повторяющимися сочетаниями типов облачности в её многослойных образованиях с  $\tau < 10$  согласно *рис. 1в* являются: *Ci fib* над кучевыми средними/мощными (*Cu med/cong*), перистые плотные (*Ci sp*) над *Cu med/cong*, *Ci fib* над *Cu hum*, высокослоистые (*As*) над *Cu med/cong* и *As* над слоисто-кучевыми кучевообразными (*Sc cuf*). При этом наиболее характерным расстоянием между наблюдаемыми ярусами многослойной облачности в исследуемом регионе является  $\Delta h = 2\text{--}5$  км, что существенно затрудняет определение ВНГО хотя бы для одного из них, например, по методике из (Minnis et al., 2019).

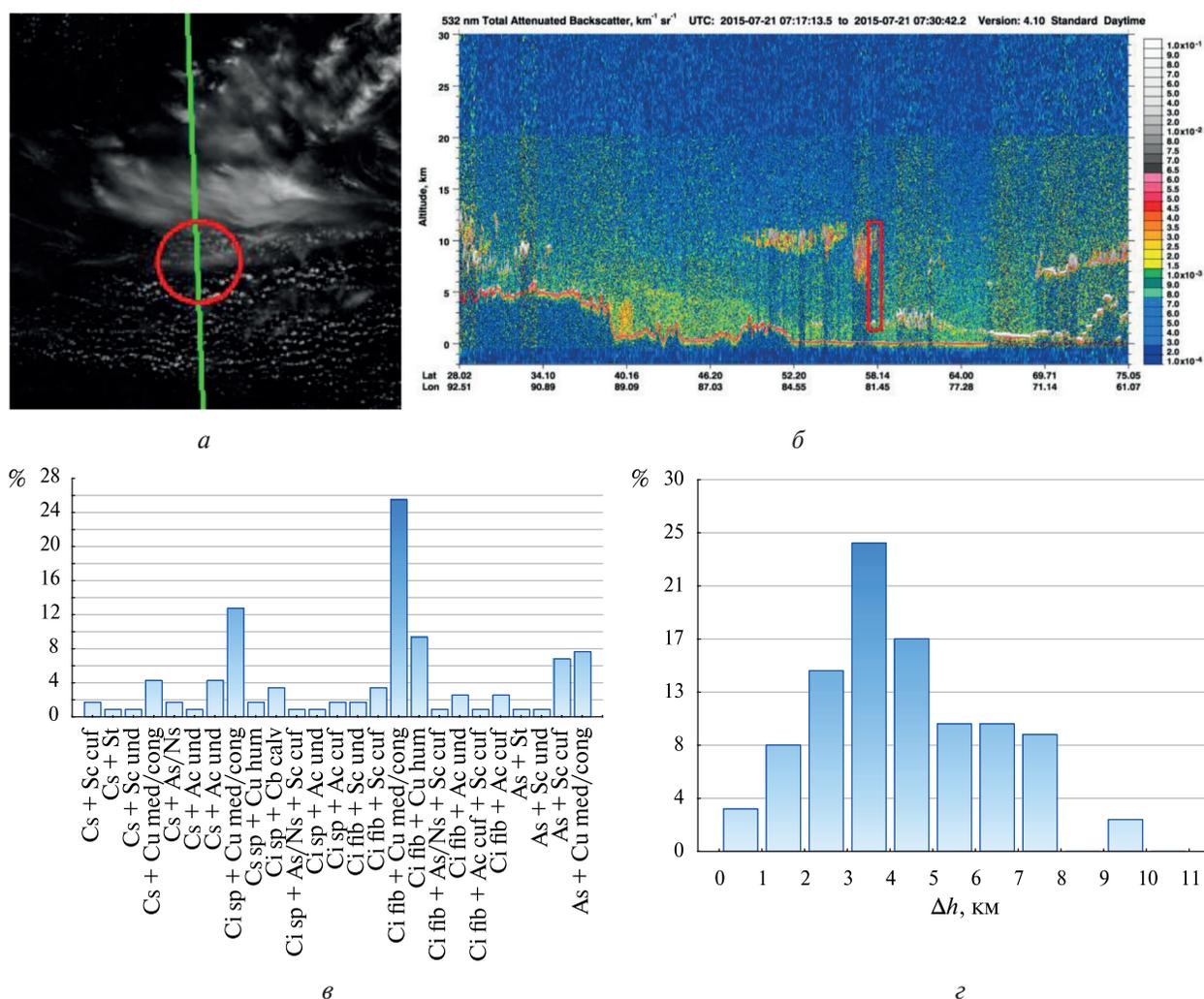


Рис. 1. Эпизод наблюдения двухслойной облачности над Западной Сибирью от 21.07.2015 по данным инструментов MODIS (а) и CALIOP (б); повторяемость различных сочетаний облаков (в) и характерных расстояний между ярусами (г) в этом регионе за период 2013–2018 гг.

### Методика восстановления высоты нижней границы многослойной облачности

Применение классических подходов, перечисленных в разд. «Введение», для восстановления ВНГО отдельных ярусов многослойной облачности представляется неэффективным способом решения поставленной задачи. Это можно продемонстрировать на примере использования алгоритма, предложенного в работе (Skorokhodov et al., 2023), который основан на применении самоорганизующихся нейронных сетей Кохонена и демонстрирует одни из лучших результатов определения высоты основания однослойных облаков по данным пассивного спутникового зондирования. На рис. 2 (см. с. 126) показаны результаты восстановления ВНГО многослойной облачности, состоящей из облаков верхнего яруса и расположенных под ними облаков среднего, нижнего и вертикального развития, указанным методом и лидаром CALIOP в 57 случайно выбранных эпизодах её наблюдения над Западной Сибирью. Из рисунка видно, что в целом алгоритм завышает оценки высоты основания облачности относительно облаков нижележащего слоя на 2 км, а самого верхнего — занижает на 3,5 км. В эпизодах, когда восстановленные значения ВНГО низкой облачности совпадали с измерениями лидара CALIOP (например, в 11, 16, 36 и 47), спектрометристом MODIS облака верхнего яруса фактически не были зарегистрированы ( $h_{\text{ВГ}}$  были низкими), несмотря на положительное значение флага многослойности. В остальных ситуациях оптическая толщина,

эффективный радиус частиц и водозапас соответствовали оптически более плотной нижележащей облачности, но значения  $h_{\text{ВГ}}$  были более высокими. Поэтому и восстанавливаемые алгоритмом оценки ВНГО получались завышенными.

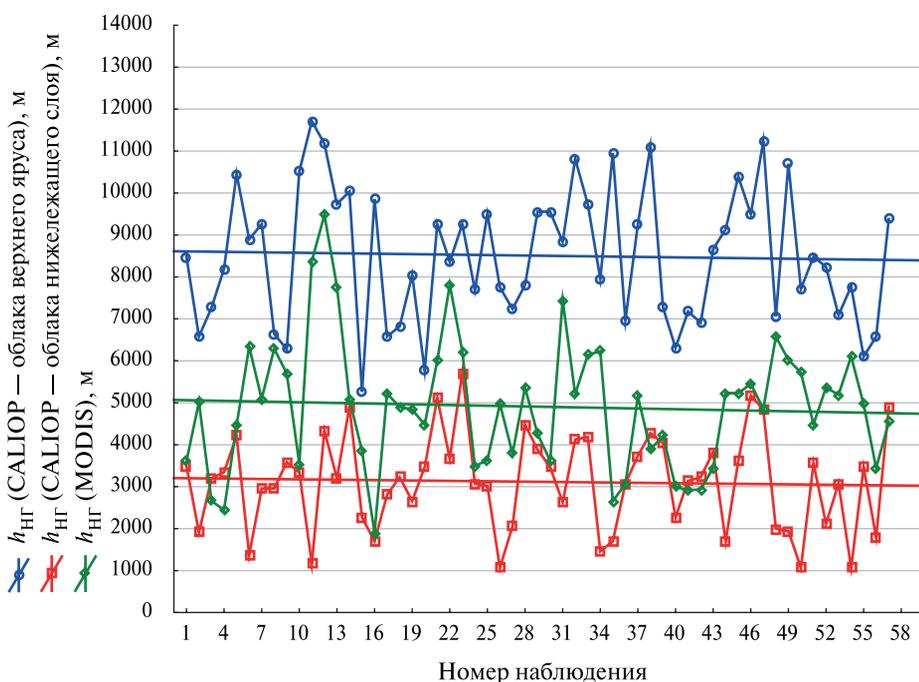


Рис. 2. Результаты измерений лидаром CALIOP ВНГО многослойной облачности, состоящей из облаков верхнего яруса и расположенных под ними облаков среднего, нижнего и вертикального развития, а также её значения, восстановленные по данным MODIS алгоритмом из работы (Skorokhodov et al., 2023)

Одно из возможных решений задачи восстановления ВНГО отдельных ярусов многослойной облачности — применение аппарата нечёткой логики, согласно которому один и тот же объект может относиться к разным классам одновременно, но с различной степенью принадлежности (Zadeh, 1965). Здесь, как и в случае с однослойными облаками в работе (Скороходов, Курьянович, 2022), оценка высоты основания облачности рассматривается в качестве частного случая решения задачи классификации. Классами выступают узкие интервалы ВНГО — 0,05 км в диапазоне от 0 до 20 км, а классификационными признаками являются характеристики облаков, восстановленные по данным пассивного спутникового зондирования. Поскольку обучающая выборка может охватывать не все возможные значения высоты основания облачности из-за сезонно-широтных особенностей её режима в исследуемом регионе, то целесообразным видится применение методов нечёткой самоорганизации (кластеризации), которые имеют сходство с сетями Кохонена (Осовский, 2002).

На рис. 3 показана структура нечёткого классификатора, представленного в виде нейронной сети, которая предлагается для восстановления ВНГО многослойной облачности по данным пассивного спутникового зондирования. На её вход предъявляется вектор  $x = (\tau, P, h_{\text{НГ}}, \dots, \theta)$  характеристик обучающих или исследуемых образцов однослойных или многослойных облаков соответственно. На этапе настройки классификатора элемент  $h_{\text{НГ}}$  используется и определяется по данным CALIOP, а при непосредственном функционировании (восстановлении ВНГО) исключается из его структуры вместе со всеми связями. Поэтому оценка ВНГО осуществляется только по данным пассивного спутникового зондирования. В сети существует  $K$  нечётких нейронов с центрами  $\beta_j = (\tau^{(j)}, P^{(j)}, h_{\text{НГ}}^{(j)}, \dots, \theta^{(j)})$  ( $j = 1, 2, \dots, K$ ), каждый из которых соответствует определённому диапазону (точнее, его центру) высот основания облаков, а  $w = (w_1, w_2, \dots, w_K)$  — вектор степени принадлежности обучающего или ис-

следуемого образца к кластерам  $\beta_j$ . Причём должно выполняться условие  $\sum_{j=1}^K w_j = 1$ . Выходом сети является вектор  $\mathbf{L}_j = (l_1, l_2, l_3)$ , где  $l_1, l_2$ , и  $l_3$  — восстановленные значения ВНГО нижнего, среднего и верхнего ярусов соответственно. При этом значения  $l$  соответствуют  $h_{\text{НГ}}^{(j)}$  кластеров  $\beta_j$  с тремя самыми высокими показателями степени принадлежности  $w_j$  при условии  $w_j > 0,3$ , заданного экспериментальным путём. Если для образца  $x$  данное неравенство выполняется только для одного или двух нейронов, то значения  $l_2$  и/или  $l_3$  заполняются нулями.

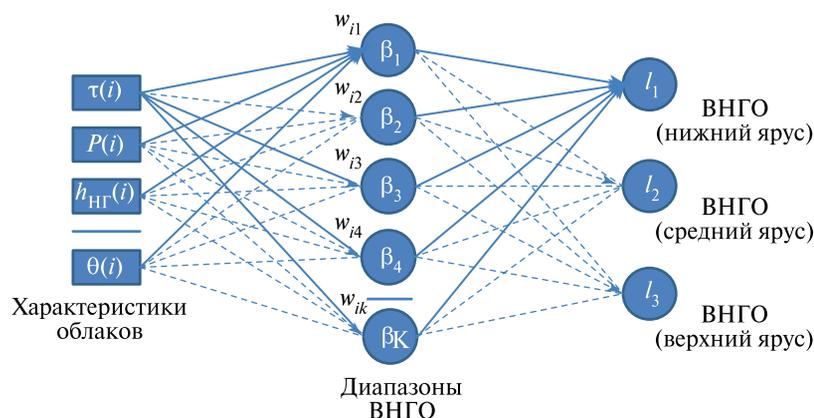


Рис. 3. Структура нечёткого классификатора

Основная суть обучения данного типа классификаторов заключается в минимизации функции

$$E = \sum_{j=1}^K \sum_{i=1}^N w_{ij}^m d^2(x_i, \beta_j),$$

где  $N$  — объём обучающей выборки;  $m$  — весовой коэффициент из интервала  $[1, \infty)$ ;  $d$  — расстояние между центром кластера  $\beta_j$  и вектором  $x_i$ . На начальном этапе  $\beta_j$  инициализируются случайными значениями. В классическом алгоритме, имеющем название Fuzzy C-means,  $d(x_i, \beta_j) = \|\beta_j - x_i\|$  — стандартное евклидово расстояние (Dunn, 1973). Однако его применение позволяет формировать кластеры только в виде идеальных гиперсфер, в результате чего не всегда удаётся правильно описать структуру входных данных. Более эффективным представляется алгоритм нечёткой самоорганизации Густафсона–Кесселя, согласно которому  $d(x_i, \beta_j) = \sqrt{(\beta_j - x_i)^T A_j (\beta_j - x_i)}$ , где  $A_j$  — масштабирующая (симметричная положительная определённая) матрица, которая на начальном этапе обучения имеет элементарную форму (Gustafson, Kessel, 1979).

Оба алгоритма обучаются по схожей схеме, согласно которой после предъявления очередного обучающего образца  $x_i$  происходит перерасчёт значений степени принадлежности  $w_{ij}$  и центров кластеров  $\beta_j$  по следующим формулам:

$$w_{ij} = \left( \frac{\sum_{k=1}^K \left( \frac{d^2(x_i, \beta_j)}{d^2(x_i, \beta_k)} \right)^{1/(m-1)}}{\sum_{k=1}^K \left( \frac{d^2(x_i, \beta_j)}{d^2(x_i, \beta_k)} \right)^{1/(m-1)}} \right)^{-1}, \quad \beta_j = \frac{\sum_{i=1}^N w_{ij}^m x_i}{\sum_{i=1}^N w_{ij}^m}.$$

Для алгоритма Густафсона–Кесселя необходимо ещё корректировать масштабирующую матрицу по формуле  $A_j = R \sqrt{\det(S_j)} \cdot S_j^{-1}$ , где  $R$  — размерность вектора  $x$ ;  $S_j$  — матрица ковариации, сформированная по формуле:

$$S_j = \sum_{i=1}^N w_{ij}^m (\beta_j - x_i) \cdot (\beta_j - x_i)^T.$$

В работе были реализованы оба алгоритма нечёткой самоорганизации для сравнения их эффективности по восстановлению значений ВНГО отдельных ярусов многослойной облачности. При этом были выбраны значения  $m = 2$  и  $K = 121$ . Большие значения  $m$  уменьшают скорость обучения классификаторов, но существенным образом не влияют на их эффективность. Число кластеров  $K$  было взято из работы (Скороходов, Курьянович, 2022), на эту величину сети Кохонена удалось разделить входные данные при решении задачи оценки высоты основания однослойных облаков. Результатом работы алгоритмов являются 2/3 изображения (по числу ярусов) в псевдоцветах, каждый из которых соответствует определённому значению ВНГО.

На рис. 4 показан результат восстановления ВНГО отдельных ярусов многослойной облачности над Западной Сибирью по спутниковому снимку MODIS от 21.07.2013 (06:45 UTC (англ. Coordinated Universal Time, всемирное координированное время)) разработанным алгоритмом на основе использования метода Густафсона–Кесселя. Отметим, что обрабатывались только пиксели многослойной облачности в соответствии со значениями флага из продукта MYD06\_L2. Из-за ограниченных возможностей лидара CALIOP ситуации с восстановлением истинных значений ВНГО трёхслойной облачности с  $\tau < 10$  практически не встречались. Поэтому постобработка, заключающаяся в сопоставлении  $l_1$ ,  $l_2$  и  $l_3$  со справочными значениями из работы (Беспалов и др., 2011), на данном этапе исследований не выполнялась. Фактически ВНГО оценивалась для двух слоёв облачности: самого высшего и расположенного под ним. К первому из них относятся облака верхнего и среднего ярусов, а ко второму — среднего и нижнего, а также конвективная облачность. Такой подход позволяет уменьшить число изображений, получаемых в результате применения разработанных алгоритмов, с трёх до двух без потери достоверности информации о высоте основания многослойной облачности с  $\tau < 10$ .

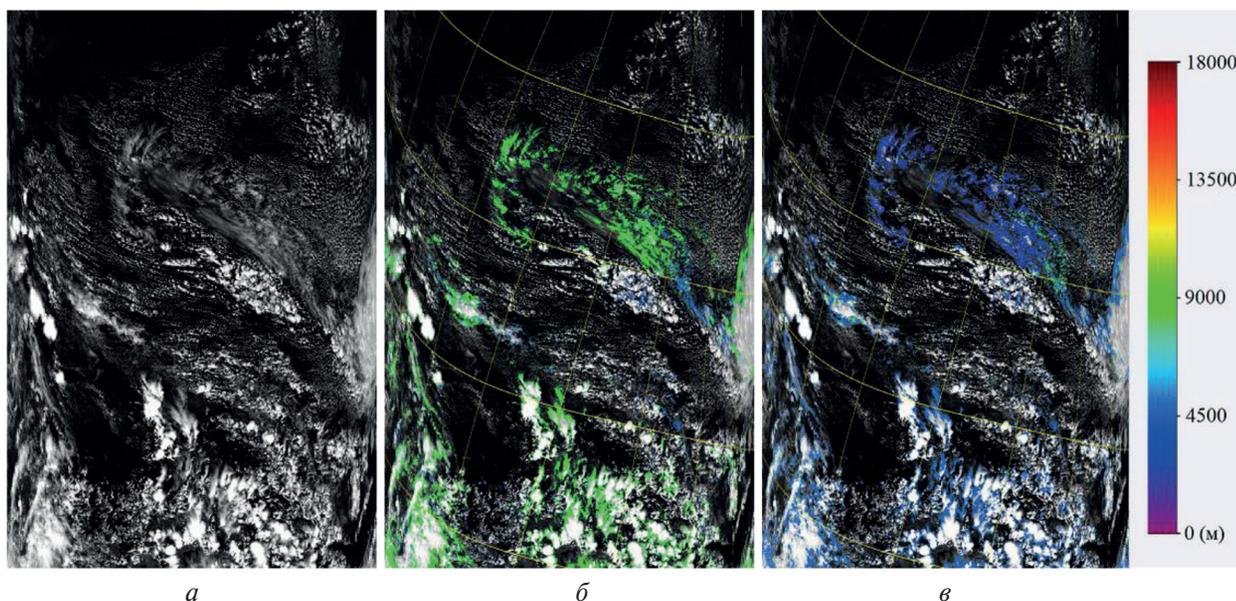


Рис. 4. Результаты восстановления ВНГО для самого верхнего слоя облаков (б) и расположенного под ним (в) по спутниковому снимку MODIS (а) территории Западной Сибири от 21.07.2013 (06:45 UTC)

### Обсуждение результатов

Эффективность восстановления ВНГО разработанными алгоритмами нечёткой самоорганизации оценивалась на основе результатов обработки тестовой выборки, состоящей из

2622 эпизодов синхронных наблюдений двухслойной облачности с  $\tau < 10$  лидаром CALIOP и спектрорадиометром MODIS над Западной Сибирью за исследуемый в работе период времени. Ещё раз отметим, что ситуации с регистрацией трёхъярусной облачности с указанной оптической толщиной и истинными значениями высоты основания в каждом слое по данным активного зондирования практически не встречались, а поэтому были исключены из анализа.

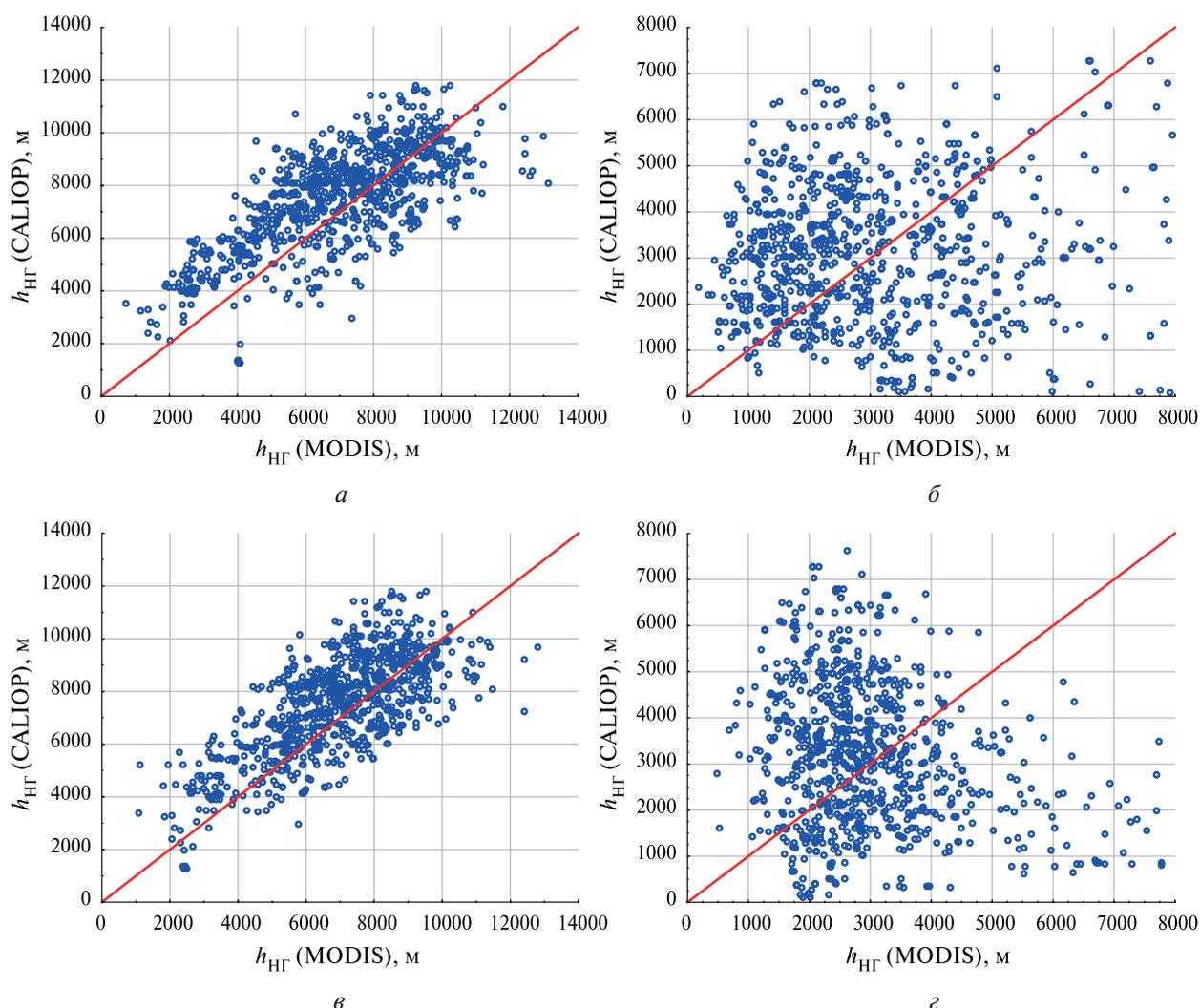


Рис. 5. Результаты сопоставления восстановленных по данным MODIS значений ВНГО многослойной облачности алгоритмами на основе метода Fuzzy C-means (а, б) и на основе метода Густафсона – Кесселя (в, г) с измерениями CALIOP для самого верхнего слоя облаков и расположенного под ним слоя соответственно

На рис. 5 показаны диаграммы рассеяния значений ВНГО для каждого из двух слоёв облаков, определённых по данным MODIS разработанными алгоритмами и измеренных лидаром CALIOP. Из рисунка видно, что применение обоих методов нечёткой самоорганизации даёт одинаково хорошие результаты определения высоты основания у самого верхнего слоя облачности. При этом наблюдается недооценка ВНГО разработанными алгоритмами по сравнению с активными наблюдениями, как и для однослойных облаков (Skorokhodov et al., 2023). Для нижележащего слоя результаты использования указанных методов различаются. Большой разброс значений даёт алгоритм, основанный на Fuzzy C-means, по сравнению с алгоритмом Густафсона – Кесселя. На рис. 5б и г наблюдается значительное количество завышенных оценок ВНГО, определённых с помощью обоих алгоритмов.

Предположительной причиной этого является сильное искажение восстанавливаемых по данным MODIS характеристик облаков нижележащего слоя более высоким. Оценка степени влияния облаков разных ярусов, наблюдаемых одновременно, на характеристики друг друга по данным пассивного спутникового зондирования — трудоёмкая задача, она выступает перспективным направлением развития этой работы.

Оценки эффективности восстановления ВНГО многослойной облачности разработанными алгоритмами на основе синхронных данных CALIOP и MODIS

Слой	$\langle \Delta \rangle$ , км	$\sigma_\Delta$ , км	$\delta_\sigma$	$R_\Delta^2$
<i>Метод Fuzzy C-means</i>				
Верхний	-0,6	1,7	0,23	0,50
Нижний	0,1	2,5	0,79	0,16
<i>Метод Густафсона – Кесселя</i>				
Верхний	-0,5	1,5	0,20	0,53
Нижний	-0,1	2,1	0,65	0,28

В таблице приведены стандартные для этой области исследований (Minnis et al., 2021; Noh et al., 2017; Tan et al., 2022) оценки эффективности восстановления ВНГО по данным пассивного спутникового зондирования, найденные с помощью их сопоставления с результатами активных измерений искомого параметра: средняя ошибка смещения ( $\langle \Delta \rangle$ ), среднеквадратичная ошибка ( $\sigma_\Delta$ ), относительная среднеквадратичная ошибка ( $\delta_\sigma$ ) и коэффициент детерминации ( $R_\Delta^2$ ). Из таблицы видно, что алгоритм на основе метода нечёткой самоорганизации Густафсона – Кесселя имеет лучшие показатели эффективности по сравнению с Fuzzy C-means, особенно при оценке высоты основания нижележащего слоя облачности. Оба алгоритма занижают оценки ВНГО самых верхних облаков на 0,5–0,6 км при среднеквадратичном отклонении 1,5–1,7 км, что подтверждает предположения, сделанные ранее на основании рис. 5. Для нижележащей облачности восстановленные значения высоты основания компенсируют друг друга, о чём свидетельствует показатель  $-0,1 < \langle \Delta \rangle < 0,1$ . Однако среднеквадратичное отклонение (а соответственно, и количество аномальных выбросов) меньше при использовании метода Густафсона – Кесселя. Таким образом, использование алгоритма восстановления ВНГО многослойной облачности на основе применения именно этого метода нечёткой самоорганизации является предпочтительным. В целом же оба алгоритма реализуют хорошую регрессионную модель для облаков самого верхнего уровня и менее эффективную для нижележащих слоёв, исходя из значений  $R_\Delta^2$ .

### Заключение

В результате проведённых исследований разработано два алгоритма восстановления ВНГО отдельных ярусов многослойной облачности по данным пассивного спутникового зондирования на основе методов нечёткой самоорганизации Fuzzy C-means и Густафсона – Кесселя. Одним из основных достоинств этих алгоритмов является их способность автоматически разбивать пространство входных данных на заранее неизвестное число кластеров (по аналогии с нейронными сетями Кохонена). Эта особенность позволяет выявлять область допустимых значений ВНГО, тем самым облегчая адаптацию указанных алгоритмов к работе с данными различных приборов, полученными над произвольно выбранными регионами при любых условиях съёмки. На текущем этапе исследований обучение алгоритмов осуществлялось на основе результатов синхронных наблюдений за облаками лидаром CALIOP и спектрорадиометром MODIS над Западной Сибирью в летнее время. По итогам сравнительного анализа

этих двух алгоритмов было установлено, что оба они более эффективно реализуют регрессионную модель для самого верхнего яруса облаков по сравнению с нижележащими слоями. При этом, исходя из *таблицы*, лучшие результаты восстановления ВНГО многослойной облачности в целом достигнуты алгоритмом, который основан на применении метода Густафсона – Кесселя. Основным ограничением разработанных алгоритмов на текущем этапе исследований является их способность обрабатывать спутниковые наблюдения многослойной (преимущественно двухъярусной) облачности с  $\tau < 10$ , сделанных только в светлое время суток при отсутствии снежного покрова. В перспективе планируется привлечение данных радара CPR (спутник CloudSat) для обеспечения работы произведённых алгоритмов с оптически более плотными облаками. Также целесообразным представляется проведение анализа влияния облачности различных ярусов, наблюдаемой одновременно, на восстанавливаемые пассивными сенсорами характеристики друг друга.

Работа в части создания алгоритма восстановления ВНГО многослойной облачности по данным пассивного спутникового зондирования выполнена при финансовой поддержке Российского научного фонда (грант № 21-71-10076, <https://rscf.ru/project/21-71-10076/>), а в части изучения статистических характеристик облачности над Западной Сибирью — в рамках государственного задания Института оптики атмосферы им. В. Е. Зуева СО РАН.

## Литература

1. Астафуров В. Г., Скороходов А. В. Классификация многоярусной облачности по данным MODIS с использованием технологии нейронных сетей и методов нечеткой логики // Современные проблемы дистанционного зондирования Земли из космоса. 2015. Т. 12. № 6. С. 162–173.
2. Беспалов Д. П., Девяткин А. М., Довгалюк Ю. А., Кондратьев В. И., Кулешов Ю. В., Светлова Т. П., Суворов С. С., Тимофеев В. И. Атлас облаков. СПб.: Изд-во Д'АРТ, 2011. 248 с.
3. Борейшо А. С., Ким А. А., Коняев М. А. и др. Современные лидарные средства дистанционного зондирования атмосферы // Фотоника. 2019. Т. 13. № 7. С. 648–657. DOI: 10.22184/1992-7296.FRos.2019.13.7.648.657.
4. Дашко Н. А. Курс лекций по синоптической метеорологии. Владивосток: Изд-во ДВГУ, 2005. 523 с.
5. Код для оперативной передачи данных приземных метеорологических наблюдений с сети станций Росгидромета (КН-01 SYNOP) / отв. ред. Фахрутдинова Н. П. М., 2013. 79 с.
6. Кхьюнг Н. В. Оценка влияния метеобразования на распространение радиоволн в X-диапазоне // Тр. Московского физико-техн. ин-та. 2020. Т. 12. № 3. С. 94–103.
7. Облака и облачная атмосфера: справ. / под ред. Мазина И. П., Хргиана А. Х. Л.: Гидрометеиздат, 1989. 647 с.
8. Осовский С. Нейронные сети для обработки информации / пер. с польск. Рудинского И. Д. М.: Финансы и статистика, 2002. 344 с.
9. Скороходов А. В., Курьянович К. В. Использование данных CALIOP для оценки высоты нижней границы облаков на спутниковых снимках MODIS // Современные проблемы дистанционного зондирования Земли из космоса. 2022. Т. 19. № 2. С. 43–56. DOI: 10.21046/2070-7401-2022-19-2-43-56.
10. Толмачева Н. И., Крючкова А. Д. Методы и средства метеорологических измерений: учеб. пособие. Пермь: Изд-во ПГНИУ, 2013. 253 с.
11. Шакина Н. П. Динамика атмосферных фронтов. Л.: Гидрометеиздат, 1985. 264 с.
12. Bessho K., Date K., Hayashi M. et al. An introduction to Himawari-8/9 — Japan's new-generation geostationary meteorological satellites // J. Meteorological Society of Japan. Ser. II. 2016. V. 94. No. 2. P. 151–183. DOI: 10.2151/jmsj.2016-009.
13. Chang F. L., Li Z. A near-global climatology of single-layer and overlapped clouds and their optical properties retrieved from Terra/MODIS data using a new algorithm // J. Climate. 2005. V. 18. No. 22. P. 4752–4771. DOI: 10.1175/JCLI3553.1.
14. Chen S., Cheng C., Zhang X. et al. Construction of nighttime cloud layer height and classification of cloud types // Remote Sensing. 2020. V. 12. Article 668. DOI: 10.3390/rs12040668.
15. Dunn J. C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters // J. Cybernetics. 1973. V. 3. No. 3. P. 32–57. DOI: 10.1080/01969727308546046.
16. Forsythe J. M., Vonder Haar T. H., Reinke D. L. Cloud base height estimates using a combination of meteorological satellite imagery and surface reports // J. Applied Meteorology. 2000. V. 39. P. 2336–2347. DOI: 10.1175/1520-0450(2000)039<2336:CBHEUA>2.0.CO;2.

17. *Gebremariam S., Li S., Weldegaber M.* Observed correlation between aerosol and cloud base height for low clouds at Baltimore and New York, United States // *Atmosphere*. 2018. V. 9. No. 4. Article 143. DOI: 10.3390/atmos9040143.
18. *Gustafson D. E., Kessel W. C.* Fuzzy clustering with a fuzzy covariance matrix // *Proc. IEEE Conf. Decision and Control including the 17<sup>th</sup> Symp. Adaptive Processes*. 1979. P. 761–766. DOI: 10.1109/CDC.1978.268028.
19. *Hutchison K. D.* The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission // *Intern. J. Remote Sensing*. 2002. V. 23. P. 5249–5265. DOI: 10.1080/01431160110117391.
20. *Mecikalski J. R., Feltz W. F., Murray J. J. et al.* Aviation applications for satellite-based observations of cloud properties, convection initiation, in-flight icing, turbulence, and volcanic ash // *Bull. American Meteorological Society*. 2007. V. 88. P. 1589–1607. DOI: 10.1175/BAMS-88-10-1589.
21. *Minnis P., Sun-Mack S., Smith W. L. J. et al.* Advances in neural network detection and retrieval of multilayer clouds for CERES using multispectral satellite data // *Proc. SPIE*. 2019. V. 11152. DOI: 10.1117/12.2532931.
22. *Minnis P., Sun-Mack S., Chen Y. et al.* CERES MODIS cloud product retrievals for Edition 4 — Part I: Algorithm changes // *IEEE Trans. Geoscience and Remote Sensing*. 2021. V. 59. P. 2744–2780. DOI: 10.1109/TGRS.2020.3008866.
23. *Mülmstädt J., Sourdeval O., Henderson D. S. et al.* Using CALIOP to estimate cloud-field base height and its uncertainty: the Cloud Base Altitude Spatial Extrapolator (CBASE) algorithm and dataset // *Earth System Science Data*. 2018. V. 10. P. 2279–2293. DOI: 10.5194/essd-10-2279-2018.
24. *Noh Y.-J., Forsythe J. M., Miller S. D. et al.* Cloud-base height estimation from VIIRS. Part II: A statistical algorithm based on A-Train satellite data // *J. Atmospheric and Oceanic Technology*. 2017. V. 34(3). P. 585–598. DOI: 10.1175/JTECH-D-16-0110.1.
25. *Noh Y.-J., Haynes J. M., Miller S. D. et al.* A framework for satellite-based 3D cloud data: An overview of the VIIRS cloud base height retrieval and user engagement for aviation applications // *Remote Sensing*. 2022. V. 14. Article 5524. DOI: 10.3390/rs14215524.
26. *Platnick S. K., Meyer G., King M. D. et al.* The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua // *IEEE Trans. Geoscience and Remote Sensing*. 2017. V. 55. P. 502–525. DOI: 10.1109/TGRS.2016.2610522.
27. *Skorokhodov A. V., Pustovalov K. N., Khatyutkina E. V., Astafurov V. G.* Cloud-base height retrieval from MODIS satellite data based on self-organizing neural networks // *Atmospheric and Oceanic Optics*. 2023. V. 36. No. 6. P. 723–734. DOI: 10.1134/S1024856023060209.
28. *Tan Z., Huo J., Ma S. et al.* Estimating cloud base height from Himawari-8 based on a random forest algorithm // *Intern. J. Remote Sensing*. 2021. V. 42. No. 7. P. 2485–2501. DOI: 10.1080/01431161.2020.1854891.
29. *Tan Z., Ma S., Wang X. et al.* Estimating layered cloud cover from geostationary satellite radiometric measurements: a novel method and its application // *Remote Sensing*. 2022. V. 14. Article 5693. DOI: 10.3390/rs14225693.
30. *Teng S., Liu C., Zhang Z. et al.* Retrieval of ice-over-water cloud microphysical and optical properties using passive radiometers // *Geophysical Research Letters*. 2020. V. 47. No. 16. Article e2020GL088941. DOI: 10.1029/2020GL088941.
31. *Wilheit T. T., Hutchison K. D.* Retrieval of cloud base heights from passive microwave and cloud top temperature data // *IEEE Trans. Geoscience and Remote Sensing*. 2000. V. 38. P. 1253–1259. DOI: 10.1109/36.843017.
32. *Winker D. M., Vaughan M. A., Omar A. et al.* Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms // *J. Atmospheric and Oceanic Technology*. 2009. V. 26. P. 2310–2323. DOI: 10.1175/2009JTECHA1281.1.
33. *Zadeh L. A.* Fuzzy sets // *Information and Control*. 1965. V. 8. No. 3. P. 338–353. DOI: 0.1016/S0019-9958(65)90241-X.

## Using CALIOP data for multilayer cloud base height estimation from MODIS imagery based on fuzzy logic methods

A. V. Skorokhodov

V. E. Zuev Institute of Atmospheric Optics SB RAS, Tomsk 634055, Russia  
E-mail: vazime@yandex.ru

We present an algorithm for base height estimation of separate levels in multilayer clouds from passive satellite sensors based on fuzzy logic methods. The procedure for retrieving the cloud-base height is considered as a special case of solving the classification problem. The classes are narrow value ranges of target parameter. The classification features are cloud parameters recovered from passive satellite sensors. One object can belong to several classes at the same time, but with different degrees of membership according to fuzzy set theory. This feature provides the ability to estimate the base height for multiple cloud levels at once. The classifier is trained based on synchronous data from the CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) and MODIS (Moderate Resolution Imaging Spectroradiometer) obtained in summer over the territory of Western Siberia in the period 2013–2018. Whereas cloud-base height estimation is performed using only data from passive satellite sensors. Two fuzzy self-organizing methods (Fuzzy C-means and Gustafson – Kessel) are considered. It has been found that the second approach is more efficient and provides a bias of the retrieved values of the base height for clouds with an optical thickness less than 10 compared to the reference ones of  $-0.5$  km at a standard deviation of 1.5 km for the overlying cloud layer and  $-0.1$  at 2.1 km for the underlying one.

**Keywords:** CALIOP, cloud base height, fuzzy logic methods, MODIS, multilayer clouds, neural network, satellite data

Accepted: 16.01.2024

DOI: 10.21046/2070-7401-2024-21-1-122-134

### References

1. Astafurov V.G., Skorokhodov A.V., Multi-layer cloud classification from MODIS data using neural network technology and fuzzy logic approach, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 6, pp. 162–173 (in Russian).
1. Bespalov D.P., Devyatkin A.M., Dovgalyuk Yu.A., Kondratyuk V.I., Kuleshov Yu.V., Svetlova T.P., Suvorov S.S., Timofeev V.I., *Atlas oblakov* (Cloud Atlas), Saint Petersburg: Publ. House of D'ART, 2011, 248 p.
2. Boreisho A.S., Kim A.A., Konyaev M.A. et al., Modern lidar systems for atmosphere remote sensing, *Fotonika*, 2019, Vol. 13, No. 7, pp. 648–657 (in Russian), DOI: 10.22184/1992-7296.FRos.2019.13.7.648.657.
3. Dashko N.A., *Kurs lektii po sinopticheskoi meteorologii* (Course of lectures on synoptic meteorology), Vladivostok: Publ. House of DVGU, 2005, 523 p (in Russian).
4. *Kod dlya operativnoi peredachi dannykh prizemnykh meteorologicheskikh nablyudenii s seti stantsii Rosgidrometa (KN-01 SYNOP)* (Code for live data transfer surface meteorological observations from the network of Roshydromet stations (KN-01 SYNOP)), Fakhrutdinov N.P. (ed.), Moscow, 2013, 79 p. (in Russian).
5. Kkhyong N.V., Evaluation of the influence of meteorology on the propagation of radio waves in X-bands, *Trudy Moskovskogo fiziko-tekhnicheskogo instituta*, 2020, Vol. 12, No. 3, pp. 94–103 (in Russian).
6. *Oblaka i oblachnaya atmosfera: spravochnik* (Clouds and cloudy atmosphere), Mazin I.P., Khrgian A.Kh. (eds.), Leningrad: Gidrometeoizdat, 1989, 647 p. (in Russian).
7. Osovskii S., *Neironnye seti dlya obrabotki informatsii* (Neural networks for information processing), Moscow: Finansy i statistika, 2002, 344 p. (in Russian).
8. Skorokhodov A.V., Kuryanovich K.V., Using CALIOP data to estimate the cloud base height on MODIS images, *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2022, Vol. 19, No. 2, pp. 43–56 (in Russian), DOI: 10.21046/2070-7401-2022-19-2-43-56.
9. Tolmacheva N.I., Kryuchkova A.D., *Metody i sredstva meteorologicheskikh izmerenii: uchebnoe posobie* (Methods and instruments of meteorological measurements), Perm: Publ. House of PGNIU, 2013, 253 p. (in Russian).
10. Shakina N.P., *Dinamika atmosferykh frontov* (Dynamics of weather fronts), Leningrad: Gidrometeoizdat, 1985, 264 p. (in Russian).

11. Bessho K., Date K., Hayashi M. et al., An introduction to Himawari-8/9 — Japan's new-generation geostationary meteorological satellites, *J. Meteorological Society of Japan, Ser. II*, 2016, Vol. 94, No. 2. pp. 151–183, DOI: 10.2151/jmsj.2016-009.
12. Chang F.L., Li Z., A near-global climatology of single-layer and overlapped clouds and their optical properties retrieved from Terra/MODIS data using a new algorithm, *J. Climate*, 2005, Vol. 18, No. 22, pp. 4752–4771, DOI: 10.1175/JCLI3553.1.
13. Chen S., Cheng C., Zhang X. et al., Construction of nighttime cloud layer height and classification of cloud types, *Remote Sensing*, 2020, Vol. 12, Article 668, DOI: 10.3390/rs12040668.
14. Dunn J. C., A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *J. Cybernetics*, 1973, Vol. 3, No. 3, pp. 32–57, DOI: 10.1080/01969727308546046.
15. Forsythe J. M., Vonder Haar T. H., Reinke D. L., Cloud base height estimates using a combination of meteorological satellite imagery and surface reports, *J. Applied Meteorology*, 2000, Vol. 39, pp. 2336–2347, DOI: 10.1175/1520-0450(2000)039<2336:CBHEUA>2.0.CO;2.
16. Gebremariam S., Li S., Weldegaber M., Observed correlation between aerosol and cloud base height for low clouds at Baltimore and New York, United States, *Atmosphere*, 2018, Vol. 9, No. 4, Article 143, DOI: 10.3390/atmos9040143.
17. Gustafson D. E., Kessel W. C., Fuzzy clustering with a fuzzy covariance matrix, *Proc. IEEE Conf. Decision and Control including the 17<sup>th</sup> Symp. Adaptive Processes*, 1979, pp. 761–766, DOI: 10.1109/CDC.1978.268028.
18. Hutchison K. D., The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission, *Intern. J. Remote Sensing*, 2002, Vol. 23, pp. 5249–5265, DOI: 10.1080/01431160110117391.
19. Mecikalski J. R., Feltz W. F., Murray J. J. et al., Aviation applications for satellite-based observations of cloud properties, convection initiation, in-flight icing, turbulence, and volcanic ash, *Bull. American Meteorological Society*, 2007, Vol. 88, pp. 1589–1607, DOI: 10.1175/BAMS-88-10-1589.
20. Minnis P., Sun-Mack S., Smith W. L. J. et al., Advances in neural network detection and retrieval of multilayer clouds for CERES using multispectral satellite data, *Proc. SPIE*, 2019, Vol. 11152, DOI: 10.1117/12.2532931.
21. Minnis P., Sun-Mack S., Chen Y. et al., CERES MODIS cloud product retrievals for Edition 4 — Part I: Algorithm changes, *IEEE Trans. Geoscience and Remote Sensing*, 2021, Vol. 59, pp. 2744–2780, DOI: 10.1109/TGRS.2020.3008866.
22. Mülmenstädt J., Sourdeval O., Henderson D. S., et al., Using CALIOP to estimate cloud-field base height and its uncertainty: the Cloud Base Altitude Spatial Extrapolator (CBASE) algorithm and dataset, *Earth System Science Data*, 2018, Vol. 10, pp. 2279–2293, DOI: 10.5194/essd-10-2279-2018.
23. Noh Y.-J., Forsythe J. M., Miller S. D. et al., Cloud-base height estimation from VIIRS. Part II: A statistical algorithm based on A-Train satellite data, *J. Atmospheric and Oceanic Technology*, 2017, Vol. 34, pp. 585–598, DOI: 10.1175/JTECH-D-16-0110.1.
24. Noh Y.-J., Haynes J. M., Miller S. D. et al., A framework for satellite-based 3D cloud data: An overview of the VIIRS cloud base height retrieval and user engagement for aviation applications, *Remote Sensing*, 2022, Vol. 14, Article 5524, DOI: 10.3390/rs14215524.
25. Platnick S. K., Meyer G., King M. D. et al., The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua, *IEEE Trans. Geoscience and Remote Sensing*, 2017, Vol. 55, pp. 502–525, DOI: 10.1109/TGRS.2016.2610522.
26. Skorokhodov A. V., Pustovalov K. N., Khatyutkina E. V., Astafurov V. G., Cloud-base height retrieval from MODIS satellite data based on self-organizing neural networks, *Atmospheric and Oceanic Optics*, 2023, Vol. 36, No. 6, pp. 723–734, DOI: 10.1134/S1024856023060209.
27. Tan Z., Huo J., Ma S. et al., Estimating cloud base height from Himawari-8 based on a random forest algorithm, *Intern. J. Remote Sensing*, 2021, Vol. 42, No. 7, pp. 2485–2501, DOI: 10.1080/01431161.2020.1854891.
28. Tan Z., Ma S., Wang X. et al., Estimating layered cloud cover from geostationary satellite radiometric measurements: a novel method and its application, *Remote Sensing*, 2022, Vol. 14, Article 5693, DOI: 10.3390/rs14225693.
29. Teng S., Liu C., Zhang Z. et al., Retrieval of ice-over-water cloud microphysical and optical properties using passive radiometers, *Geophysical Research Letters*, 2020, Vol. 47, No. 16, Article e2020GL088941, DOI: 10.1029/2020GL088941.
30. Wilheit T. T., Hutchison K. D., Retrieval of cloud base heights from passive microwave and cloud top temperature data, *IEEE Trans. Geoscience and Remote Sensing*, 2000, Vol. 38, pp. 1253–1259, DOI: 10.1109/36.843017.
31. Winker D. M., Vaughan M. A., Omar A. et al., Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, *J. Atmospheric and Oceanic Technology*, 2009, Vol. 26, pp. 2310–2323, DOI: 10.1175/2009JTECHA1281.1.
32. Zadeh L. A., Fuzzy sets, *Information and Control*, 1965, Vol. 8, No. 3, pp. 338–353, DOI: 10.1016/S0019-9958(65)90241-X.