

## Кластерный анализ в автоматическом выявлении и сопровождении грозовых очагов по данным грозопеленгационной сети

*А.А. Аджиева<sup>1</sup>, В.А. Шаповалов<sup>2</sup>*

<sup>1</sup>*ФГБОУ ВО «Кабардино-Балкарский государственный аграрный университет  
имени В.М. Кокова»*

<sup>2</sup>*ФГБУ "Высокогорный геофизический институт"*

**Аннотация:** Прогноз перемещения грозовых очагов требует внедрения современных методов оперативной обработки данных поступающих на автоматизированные рабочие места в центры приема. Для текущего прогноза быстроразвивающихся метеоявлений необходимо выделять и анализировать движение грозовых очагов по данным грозопеленгации. Для этих целей разрабатываются и совершенствуются методы выделения и сопровождения, основанные на математическом аппарате компьютерного зрения.

**Ключевые слова:** контурный анализ, кластерный анализ, опасные явления погоды, текущий прогноз, метеорадиолокация, грозопеленгация, грозовые очаги.

В то время как средства оповещения достигли высокого уровня массовости и оперативности, эффективность превентивных мер в настоящее время в значительной степени зависит от качества и заблаговременности прогноза катастрофических природных явлений, который в свою очередь основывается на сумме знаний о природных опасностях, причинах их возникновения, масштабах распространения, особенностях проявления и распространения в пределах определенной территории. Это сказывается на текущих прогнозах быстроразвивающихся метеоявлений, таких как, например, опасные явления погоды (ОЯП) связанные с атмосферной конвекцией: ливни, грозы, град, шквалы, смерчи, и вторичные, порождаемые ими: оползни, сели, наводнения, нагоны [1].

Для преодоления этих трудностей используют данные численного моделирования. Тем не менее, создание автоматизированной системы текущего прогнозирования опасных явлений, связанных с конвективной облачностью, на основе геоинформационного и гидродинамического моделирования сталкивается с рядом трудностей. Прогноз перемещения

---

грозовых очагов по какой-либо из современных мезомасштабных численных моделей так и не стал к настоящему времени простым, оперативным и обладающим высокой степенью оправдываемости методом, позволяющим предсказать поведение мощных конвективных ячеек. Необходимость же в автоматической обработке данных дистанционных метеорологических наблюдений в реальном масштабе времени с целью осуществления текущего прогноза опасных метеорологических явлений требует наличия алгоритмов в составе автоматизированного рабочего места (далее АРМ).

Метеорологическая информация на АРМ в центр приема поступает по каналам связи от радиолокаторов МРЛ-5, новых доплеровских станций ДМРЛ-С, грозорегистратора LS8000 и других источников [2-5]. Для целей выделения и сопровождения ОЯП внедряются новые усовершенствованные алгоритмы распознавания, основанные на комплексном анализе поступающей информации с использованием математического аппарата компьютерного зрения (Computer Vision) [6].

Кластерный анализ, или численная таксономия – это один из новых математических методов классификации получивших распространение благодаря развитию компьютерных технологий. Цель метода – типологическая группировка совокупностей массовых явлений на основе множества признаков. Он позволяет производить разбиение объектов не по одному параметру, а одновременно по целому их набору и при этом не накладывает никаких ограничений на вид рассматриваемых объектов, что позволяет рассматривать множество исходных данных практически произвольной природы [7, 8].

Одно из главных достоинств кластерного анализа в том, что его можно использовать циклически до тех пор, пока не будут достигнуты необходимые результаты. При этом каждый цикл дает информацию, которая влияет на применяемые в дальнейшем подход и направление. Этот процесс

---

можно представить как систему с обратной связью. На чем основывается еще одно применение кластерного анализа - распознавание с самообучением.

Каждая единица совокупности в кластерном анализе считается «точкой в признаковом пространстве». Значение каждого из признаков у данной единицы совокупности служит ее координатой в этом «пространстве». Таким образом, «признаковое пространство» – это область варьирования всех признаков совокупности изучаемых явлений, а «вектор признаков» - объект (точка) в данном «пространстве».

Для решения задачи классификации массива большого объема многомерных наблюдений, таких как данные грозопеленгации, воспользуемся хорошо зарекомендовавшим себя методом кластерного анализа -  $k$ -средних ( $k$ -means). Суть данного метода кластеризации заключается в разбиении массива данных на кластеры таким способом, чтобы минимизировать сумму расстояний от объектов до соответствующих им центров кластеров. За меру близости примем евклидово расстояние, т.к. рассматриваемое явление, а именно грозовой очаг является локализуемым в Евклидовой метрике [9, 10]. Тогда геометрическое расстояние в  $n$  - мерном признаковом пространстве между точками  $p$  и  $q$  с  $n$  - координатами (признаками) определяется так:

$$r_{p,q} = \sqrt{\sum_{i=1}^n (X_{ip} - X_{iq})^2}$$

где  $X$  – вектор признаков:  $X = (X_1, X_2, \dots, X_n)$  и квадрат евклидова расстояния между  $X_p$  и  $X_q$  можно так же представить в виде:

$$r_{pq}^2 = (X_p - X_q)^T (X_p - X_q)$$

В таком случае алгоритм стремится минимизировать суммарное квадратичное отклонение точек кластеров от центров этих кластеров.

Алгоритм работы  $k$ -средних имеет следующий вид:

1. Принимаем входные данные и предполагаемое число кластеров –  $k$ ;

2. Выбираем центры кластеров;
3. Рассчитываем расстояния от центров кластеров до всех объектов и ассоциируем каждый объект с ближайшим центром кластера;
4. Перемещаем центр кластера в центроид его точек данных, т.е. вычисленные средние становятся координатами нового центра каждого кластера;
5. Возвращаемся на шаг 3 пока не достигнута сходимость, например, центр кластера остается неподвижен, заняв устойчивое положение, т.е. пока состав кластеров не перестанет меняться.

Таким образом, алгоритм  $k$ -средних принимает в качестве входных данных вектор признаков  $X$ , содержащий  $n$  точек, а также параметр  $k$ , задающий требуемое количество кластеров. На выходе - набор из  $k$  центроидов кластеров, кроме того, всем точкам множества  $X$  присваиваются метки, относящие их к определенному кластеру, причем все точки в пределах данного кластера расположены ближе к своему центроиду, чем к любому другому центроиду.

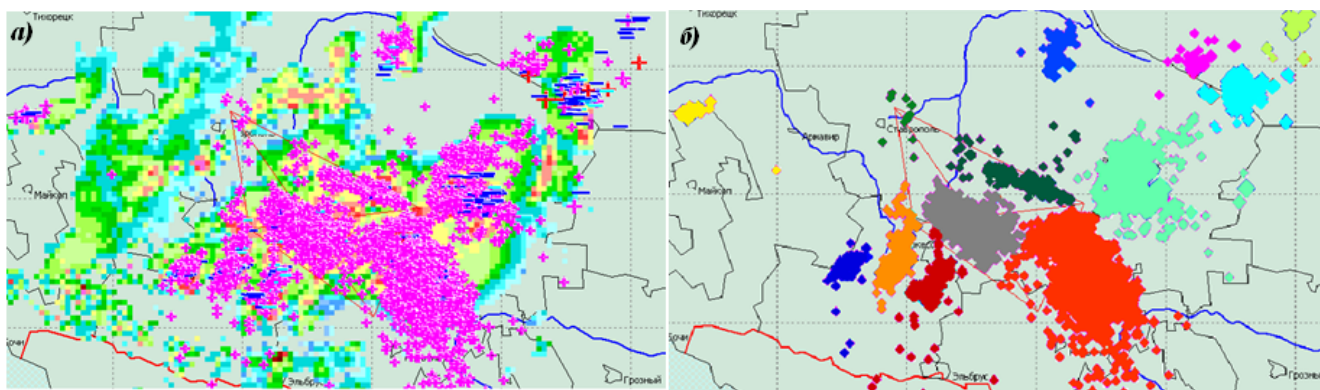


Рис. 1 – Совмещенные данные радиолокатора МРЛ-5 и грозопеленгатора LS8000 (а). Изображены отражаемость облака и разряды внутриоблачные и между облаком и землей. Плюс обозначает положительную полярность молний, минус – отрицательную. Полученные в результате обработки кластеры (б)

Для определения числа кластеров используется «разведочный» алгоритм: сначала совокупность делится на два кластера, затем на три и так до тех пор, пока не будет найдено оптимальное число кластеров.

После получения результатов следует проверить правильность кластеризации. Для этого рассчитываются средние значения для каждого кластера. При хорошей кластеризации должны быть получены сильно отличающиеся средние для всех измерений или хотя бы большей их части.

Среди достоинств алгоритма  $k$ -средних можно выделить: скорость, простоту использования, понятность и прозрачность алгоритма. В результате каждый объект назначен определенному кластеру рис. 1.

Для определения параметров движения грозового очага и предсказания их положения мало выделить и проанализировать его положение, необходимо определить параметры его движения. Эффективным средством в данном случае является контурный анализ. С целью сопровождения необходимо выделить и сопоставить контур каждого кластера на серии синтезированных изображений. Когда на последовательных изображениях выделены кластеры, с учетом некоторых ограничений возможно установление соответствий между их контурами, найденными на предыдущем и последующем кадрах, такой метод носит название детерминированное слежение [11]. Во время выполнения алгоритма для грозовых очагов возможны следующие варианты: возникновение нового, сопоставление прослеживаемому, разделение, слияние, затухание. Простейшая стратегия сопровождения в данном случае не срабатывает, поэтому вводятся различные ограничения: близость, максимальная скорость, общее движение, малое изменение вектора скорости, «жесткость», трендовая инертность и др. Применение методов кластеризации на этом этапе позволяет выполнять все эти условия [6].

---

Таким образом, развитие средств дистанционного зондирования сделало возможным быстрый обзор и регистрацию метеорологических параметров на больших площадях. Применение современных методов автоматического обнаружения и распознавания ОЯП позволяет оперативно и эффективно использовать дистанционные методы метеорологических наблюдений, и применять их данные наряду с синоптическими прогнозами в штормооповещении.

### Литература

1. Сергеева Г.А., Волобуева Л.Л., Кривошеева Е.А. Чрезвычайные ситуации, связанные с селевыми потоками на Северном Кавказе // Инженерный вестник Дона. 2012. №4. URL: [ivdon.ru/ru/magazine/archive/n4p1y2012/1154](http://ivdon.ru/ru/magazine/archive/n4p1y2012/1154).
2. Аджиев А.Х., Тапасханов В.О., Стасенко В.Н. Система грозопеленгации на Северном Кавказе // Метеорология и гидрология. 2013. № 1. С. 5-11.
3. Аджиев А.Х., Стасенко В.Н., Шаповалов А.В., Шаповалов В.А. Напряженность электрического поля атмосферы и грозовые явления на Северном Кавказе // Метеорология и гидрология. 2016. №3. С. 46-54.
4. Аджиева А.А., Шаповалов В.А., Машуков И.Х., Скорбеж Н.Н., Шаповалов М.А. Обнаружение и распознавание опасных конвективных процессов радиотехническими средствами // Известия Вузов. Северо-Кавказский регион. Естественные науки. 2014. №1 (179). С. 59-62.
5. Ашабоков Б.А., Шаповалов А.В., Кулиев Д.Д., Продан К.А., Шаповалов В.А. Численное моделирование термодинамических, микроструктурных и электрических характеристик конвективных облаков на стадии роста и максимального развития // Известия высших учебных заведений. Радиофизика. 2013. Т. 56. № 11-12. С. 900-907.

6. Шаповалов В.А., Машуков И.Х., Продан К.А., Колычев А.Г., Шаповалов М.А. АРМ обнаружения и предупреждения об опасных быстроразвивающихся конвективных явлениях на территории Южного и Северо-Кавказского федеральных округов // Материалы Международного симпозиума «Устойчивое развитие: проблемы, концепции, модели». - Нальчик, 2013. Том 2. С. 213-216.

7. Факторный, дискриминантный и кластерный анализ: Пер. с англ. / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка и др.; Под ред. И. С. Енюкова. - М.: Финансы и статистика, 1989. 215 с.

8. Колыхан Н.В., Тюрязев В.С. Информационные технологии статистической обработки данных выборок ограниченного объема // Инженерный вестник Дона. 2007. № 2. URL: ivdon.ru/ru/magazine/archive/n2y2007/21.

9. Maccorman D. R., Rust W. D. The Electrical Nature of storms. Oxford University Press, 1998. 422 p.

10. Hans Volland. Handbook on Atmospheric Electrodynamics. CRC Press, 1995. Vol. II. 528 p.

11. Гонсалес Р., Вудс Р. Цифровая обработка изображений: Пер. с англ./ Под ред. Чочиа П.А. - М.: Техносфера, 2005. 1072 с.

### References

1. Sergeeva G.A., Volobueva L.L., Krivosheeva E.A. Inženernyj vestnik Dona (Rus), 2012, №4. URL: ivdon.ru/ru/magazine/archive/n4p1y2012/1154.

2. Adzhiev A.H., Tapashanov V.O., Stasenko V.N. Meteorologija i gidrologija. 2013. № 1. pp. 5-11.

3. Adzhiev A.H., Stasenko V.N., Shapovalov A.V., Shapovalov V.A. Meteorologija i gidrologija. 2016. №3. pp. 46-54.



4. Adzhieva A.A., Shapovalov V.A., Mashukov I.H., Skorbezh N.N., Shapovalov M.A. Izvestija Vuzov. Severo-Kavkazskij region. Estestvennye nauki. 2014. №1 (179). pp. 59-62

5. Ashabokov B.A., Shapovalov A.V., Kuliev D.D., Prodan K.A., Shapovalov V.A. Izvestija vysshih uchebnyh zavedenij. Radiofizika. 2013. T. 56. № 11-12. pp. 900-907.

6. Shapovalov V.A., Mashukov I.H., Prodan K.A., Kolychev A.G., Shapovalov M.A. Ustojchivoe razvitie: problemy, koncepcii, modeli, Nal'chik, 2013, Vol. 2, pp. 213-216.

7. Dzh.-O. Kim, Ch. U. M'juller Faktornyj, diskriminantnyj i klasternyj analiz [Factor, discriminant and cluster analysis], Moscow. 1989. 215 p.

8. Kolyhan N.V., Tjurjaev V.S. Inženernyj vestnik Dona (Rus), 2007, № 2. URL: [ivdon.ru/ru/magazine/archive/n2y2007/21](http://ivdon.ru/ru/magazine/archive/n2y2007/21).

9. Maccorman D. R., Rust W. D. The Electrical Nature of storms. Oxford University Press, 1998. 422 p.

10. Hans Volland. Handbook on Atmospheric Electrodynamics. CRC Press. 1995. Vol. II. 528 p.

11. Gonsales R., Vuds R. Cifrovaja obrabotka izobrazhenij [Digital Image Processing], Moscow. 2005. 1072 p.