

АЛГОРИТМИЧЕСКОЕ ВЫДЕЛЕНИЕ КРЫЛЬЕВ  
БАТТЕРФЛЯЙ-ДИАГРАММAlgorithmic Identification for Wings  
in Butterfly Diagrams

**Abstract.** We investigate to what extent the wings of solar butterfly diagrams can be separated without an explicit usage of Hale's polarity law as well as the location of the solar equator. Two algorithms of cluster analysis, namely DBSCAN and C-means, have demonstrated their ability to separate the wings of contemporary butterfly diagrams based on the sunspot group density in the diagram alone. Here we generalize the method for continuous tracers, give results concerning the migration velocities and present clusters for cycles 12–20.

Волны солнечной активности выделяются с помощью широтно-временных (баттерфляй) диаграмм Маундера, на которые наносятся широта и время появления групп солнечных пятен. На этих диаграммах в каждом цикле активности видно по волне противоположной полярности в каждом полушарии, распространяющейся от средних широт к экватору. Кажется естественным дополнить это зрительное впечатление некоторым компьютерным алгоритмом, не апеллирующим непосредственно к полярности пятен. Это важно, например, для анализа исторических баттерфляй-диаграмм, для которых полярность групп неизвестна. Подобный алгоритм был предложен недавно (Illarionov E., Sokoloff D., Arlt R., Khlystova A., *Astron. Nachr.*, **332**, 590–596, 2011). Он основан на последовательном применении двух известных алгоритмов кластерного анализа. Один из них, DBSCAN, объединяет в кластеры точки на плоскости, лежащие в областях их повышенной плотности: две точки принадлежат одному кластеру, если они могут быть соединены последовательностью кружков заданного радиуса, в каждом из которых содержится достаточное число точек. При этом значительное число точек остается вне кластеров. Применение алгоритма C-mean позволяет отнести подавляющее большинство из этих точек к одному из кластеров, а именно к тому, расстояние до которого, измеренное с помощью аппроксимирующего этот кластер эллипса, оказывается существенно меньше, чем расстояние до других кластеров. Этот алгоритм оставляет не отнесёнными к кластерам лишь небольшое количество точек, для которых расстояния до двух ближайших кластеров оказались сравнимыми.

На примере нескольких современных и исторических кластеров авторы проверили, что описанный алгоритм действительно позволяет разбить баттерфляй-диаграммы на кластеры, представляющие собой крылья – индивидуальные волны активности. Это разбиение хорошо согласуется со зрительным впечатлением от диаграммы. Задача этой работы – провести дополнительное исследование предложенного алгоритма.

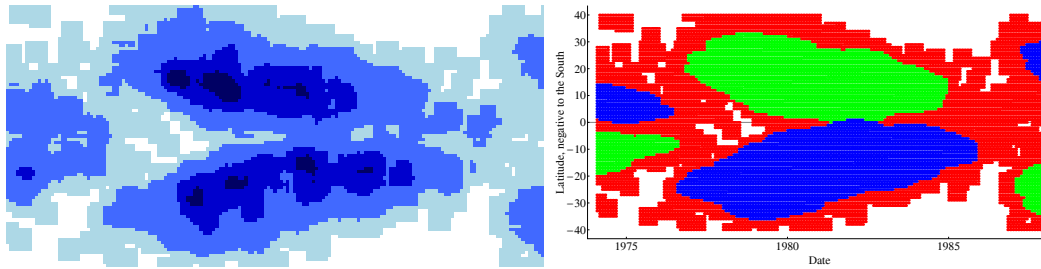


Рис. 1: Алгоритм DBSCAN выделяет крылья бабочки-диаграммы для цикла 21: слева – диаграмма, сглаженная гауссовским фильтром, справа – кластеры, выделенные алгоритмом. Координаты на левой панели соответствуют координатам на правой панели.

Таблица 1: Скорость миграции волны активности в  $^{\circ}$  год $^{-1}$  ветвей бабочки-диаграмм.

Cycle	Северное полушарие	Южное полушарие
12	3.6	5.4
13	4.4	6.5
14	6.3	7.6
15	4.9	4.5
16	4.7	3.5
17	5.6	4.2
18	4.2	4.3
19	3.3	2.3
20	5.7	6.0
21	3.8	3.4
22	4.7	4.0

Во-первых, далеко не все трасеры солнечной активности носят дискретный характер, как группы солнечных пятен, а алгоритм DBSCAN непосредственно применим только к дискретным сигналам. Мы показываем здесь, как воспользоваться им для непрерывного трасера. Для того, чтобы показать работу алгоритма на сопоставимом непрерывном трасере, мы сначала размываем бабочки-диаграмму по группам пятен для 21 цикла с помощью гауссовского фильтра до непрерывного распределения (рис. 1, левая панель). В качестве параметров гауссовского фильтра мы используем квадратичное отклонение  $\sigma = 1$  и радиус  $r = 3\sigma$ , где обе величины измерены в пикселях, а размер пикселя соответствует 1 месяцу по времени и одному градусу по широте. Интенсивность цвета на рисунке соответствует числу точек в области сглаживания, причем белый цвет соответствует нулевой плотности, светло-голубой – до 5, голубой – от 5 до 70, синий – от 70 до 140, тёмно-синий – более 140.

Небольшая модификация алгоритма DBSCAN позволяет воспользоваться им и в этом случае (рис. 1, правая панель). А именно, мы считаем, что два пикселя принадлежат одному кластеру, если их можно соединить последовательностью кружков данного радиуса  $r$ , каждый из которых лежит в области, в которой каждый пиксель имеет яркость, не меньшую  $m$ . Варьируя эти параметры, мы убедились, что при  $r = 0.08$  и  $m = 5$  алгоритм дает кластеры, соответствующие зрительному впечатлению от распределения

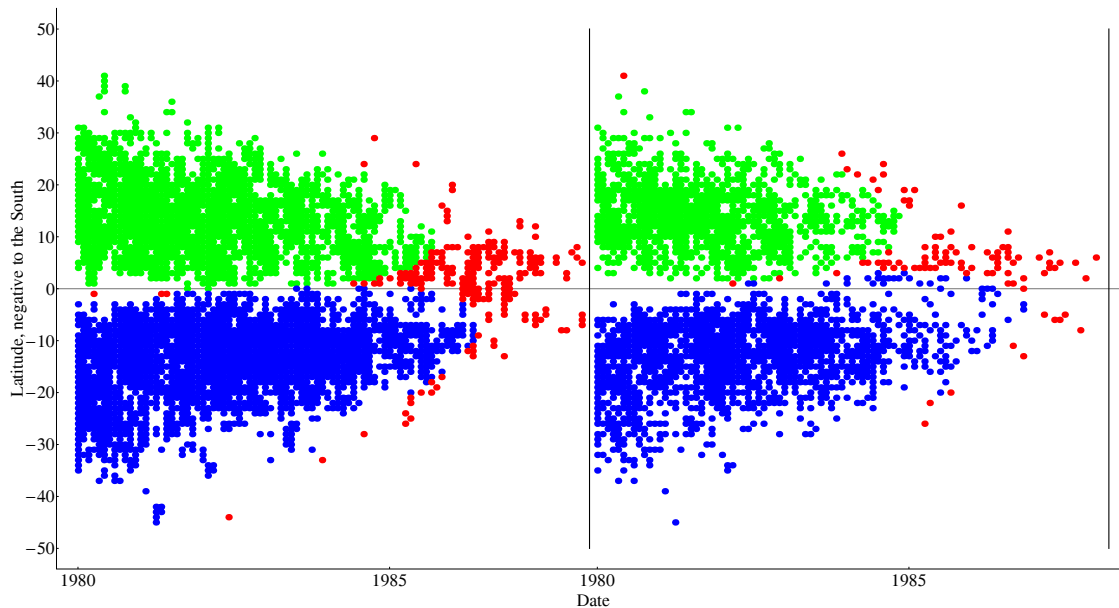


Рис. 2: Сравнение кластеров, выделенных для цикла 21 по группам солнечных пятен (слева) и по тем же данным, в которых временная привязка группы соответствовала первому её появлению (справа).

яркости.

Во-вторых, наш алгоритм на одном из своих этапов аппроксимирует получаемый кластер эллипсом, что даёт возможность вычислить по наклону этого эллипса скорость распространения волны активности. Мы приводим соответствующие результаты для циклов (табл. 1).

Далее, мы показываем, что наш алгоритм устойчив к небольшому изменению способа построения базы данных. В частности, в качестве временной привязки группы пятен можно в принципе использовать момент её первого появления. Мы попробовали эту возможность и убедились, что получающиеся таким образом кластеры мало отличаются от тех, которые мы получали ранее (рис. 2).

Наконец, мы показываем на рис. 3 кластеры, выделенные для циклов 12–20. Эти результаты дополняют примеры применения алгоритма, опубликованные в исходной работе.

Работа поддержана РФФИ, коды проектов 12-02-00170-а и 12-02-31128.

Механико-математический факультет МГУ  
 Москва, 119991, Россия  
 Физический факультет МГУ  
 Москва, 119991, Россия  
*sokoloff.dd@gmail.com*

Е.А. Илларионов  
 E.A. Illarionov  
 Д.Д. Соколов  
 D.D. Sokolov

Получено 25 декабря 2012 г.

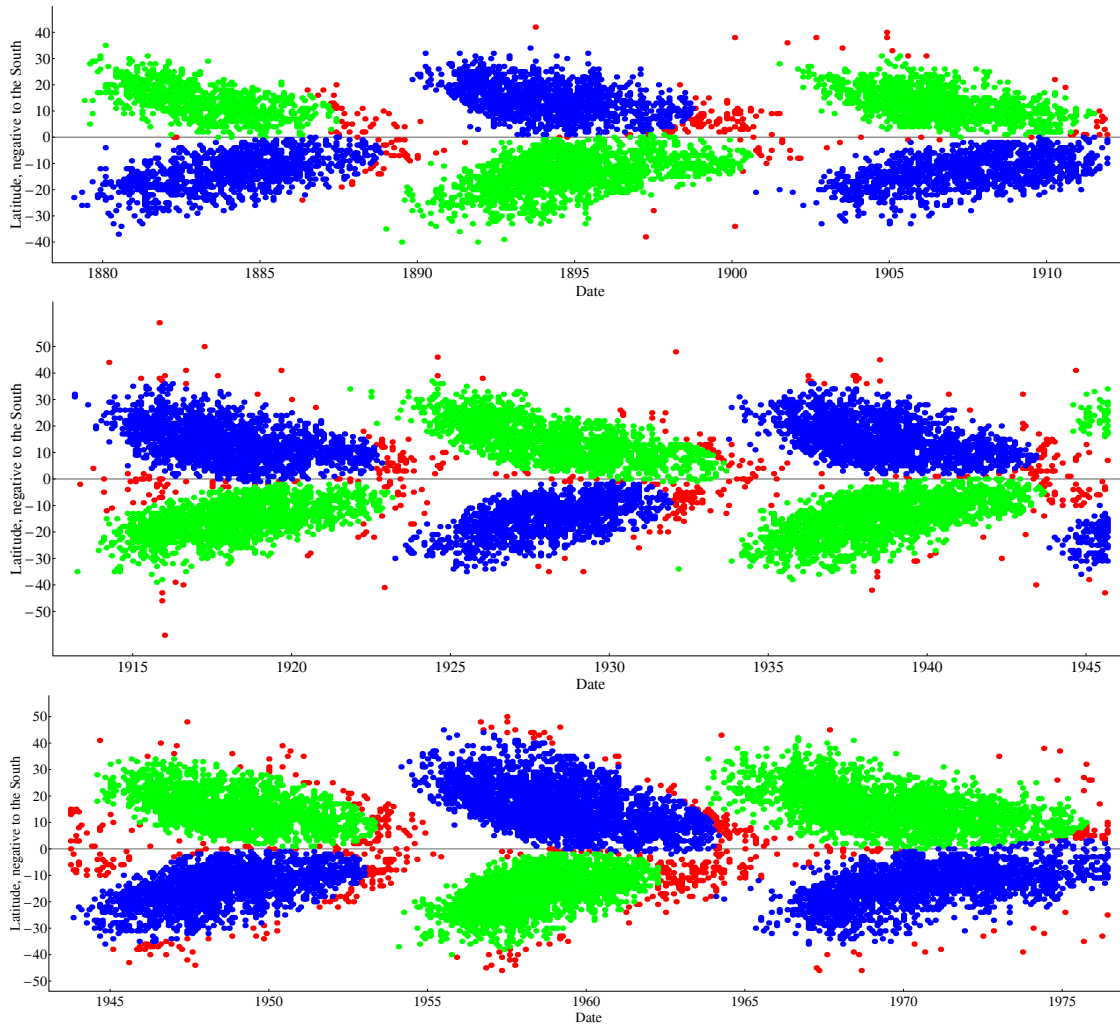


Рис. 3: Кластеры (зеленый и синий цвета), построенные с помощью нашего алгоритма для циклов 12–14 (сверху), 15–17 (в середине) и 18–20 (внизу). Красным показаны группы пятен, не попавшие в кластеры.