

Определение параметров межзвездного поглощения света по данным каталога Hipparcos

11 января 2015 г.

1 Абстракт

Основная задача исследования – автоматический поиск пылевых облаков в окрестности Солнца на основе массовых каталогов звезд. Первый этап этой задачи – построение двумерной панорамы распределения пылевых облаков на небесной сфере на основе данных каталога Hipparcos. Метод исследования основан на сравнении эталонного показателя цвета звезды данного спектрального класса с наблюдаемым показателем цвета. Так как полная двумерная спектральная классификация известна не для всех звезд каталога Hipparcos, то пришлось решать вспомогательную задачу: используя параллаксы звезд каталога Hipparcos дополнить информацию о звездах классом светимости на основе данных о видимой звездной величине и параллаксе звезды. Для этого использовался метод опорных векторов. В результате работы была построена карта распределения пылевой материи, вызывающей покраснения света звезд, по небесной сфере. Для построения трехмерного распределения пылевых облаков необходим более точный и массовый каталог параллаксов звезд, которым может стать каталог миссии GAIA, завершение которой планируется в ближайшие годы.

2 Введение

3 Исходные данные

4 Покраснение

4.1 Определение

Межзвездное поглощение может быть описано избытком цвета. Избыток цвета мы будем называть «покраснением». *Покраснение* звезды есть

$$E = E_{B-V} = (B - V)_{obs} - (B - V)_{int}$$

Где $(B - V)_{obs}$ — ее видимый нами показатель цвета, а $(B - V)_{int}$ — реальный показатель цвета звезды. $(B - V)_{obs}$ мы можем получить на основе данных фотометрии звезды. Их мы получим из каталога Hipparcos 2007. $(B - V)_{int}$ можно оценить статистически. Это проделано во многих работах, к примеру, в [4]. В результате, мы воспользуемся полученной в [4] двумерной таблицей «спектральный класс, класс светимости — показатель цвета». То есть для получения $(B - V)_{int}$ звезды нам потребуются ее спектральный класс и класс светимости. Их мы можем посмотреть в каталоге. Приведем пример расчета покраснения на звезде HIP 44800,

- У нее в каталоге $(B - V)_{obs} = 0.535^m$
- Класс F7V, поэтому (по [4]) $(B - V)_{int} = 0.493^m$

- Покраснение $0.535^m - 0.493^m = 0.042^m$

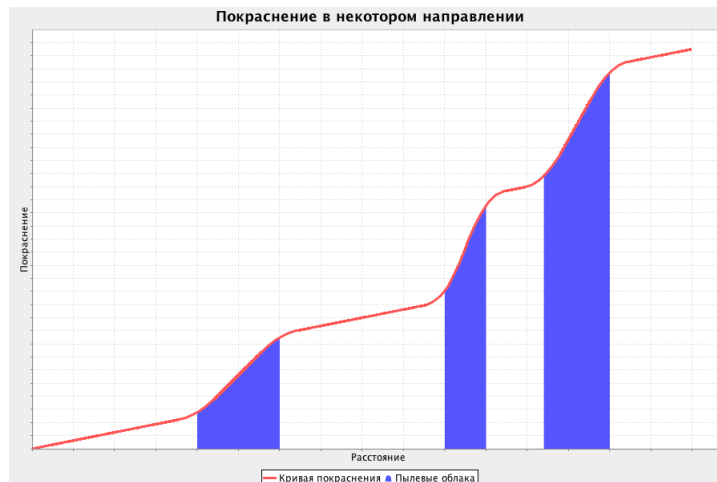
Покраснение — это количественное измерение межзвездного поглощения, поэтому мы можем сказать, что «между нами и звездой HIP 44800 пыли на 0.042 зв. вел».

4.2 Идеальная кривая покраснения

Предположим, что на некотором луче зрения бесконечно много звезд и они расположены на нем всюду плотно. Пусть для каждой звезды мы можем идеально измерить ее покраснение. Тогда, ход покраснения на этом луче зрения должен быть таким,



Где по оси x отложено расстояние по лучу зрения, а по y — покраснение у соответствующих звезд. То есть покраснение должно всегда монотонно расти, т.к. везде есть пыль. Очевидно, что там, где покраснение растет быстрее — пыли больше, там где медленнее — меньше. Поэтому, можно было бы сказать, что облака на этом луче зрения находятся там, где покраснение растет «очень быстро» (синие области),



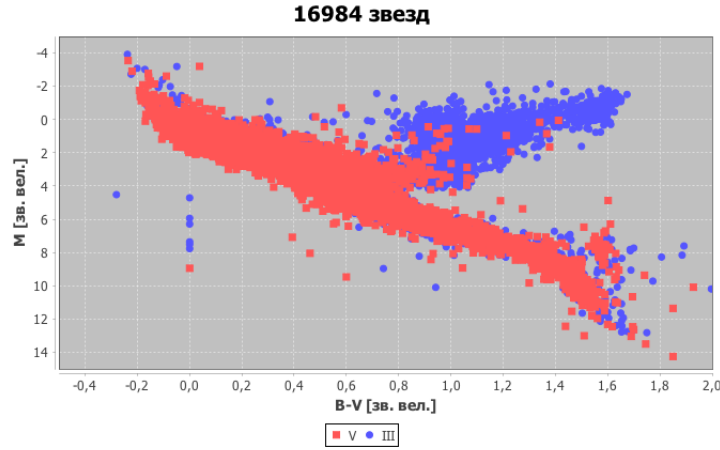
Тем самым, построение кривых покраснения в разных направлениях на небе может позволить находить области повышенного межзвездного поглощения, то есть находить пылевые облака.

5 Способ получения классов светимости

Практически у всех звезд северного полушария, представленных в каталоге Hipparcos, отсутствует класс светимости. Для нас его наличие чрезвычайно важно, ввиду того, что мы на основе класса светимости и спектрального класса мы считаем истинное значение $B-V$ для звезд $((B-V)_{int})$. Тем самым, отсутствие классов светимости у половины звезд делает невозможным проведение наших расчетов для всего северного полушария.

Исправим это. Рассчитаем неизвестные классы светимости. Сделаем это с помощью методов машинного обучения. Натренируем классификатор, который будет выдавать класс светимости для звезды по двум факторам - показателю цвета и абсолютной звездной величине. Этих факторов будет достаточно, т.к. классы светимости теоретически разделимы на диаграмме Гершпрунга-Рассела. Тренировать классификатор мы будем с помощью метода опорных векторов (Support Vector Machines, SVM) [ссылка на какую-нибудь статью]

Доля звезд, которые не относятся ни к III, ни к V классам светимости очень мала (3.6%, 3354 из 93230). Эти звезды никак на классификацию не влияют. Поэтому, в качестве обучающего множества возьмем звезды только III и V классов с приемлемыми ошибками в расстоянии (не более 10%). Диаграмма Гершпрунга-Рассела с нанесенным на нее обучающим множеством:



Факторами для обучения у нас будут показатель цвета и абсолютная звездная величина.

Для оценки качества классификатора проведем 10-fold кросс-валидацию. Ниже приведены ее результаты,

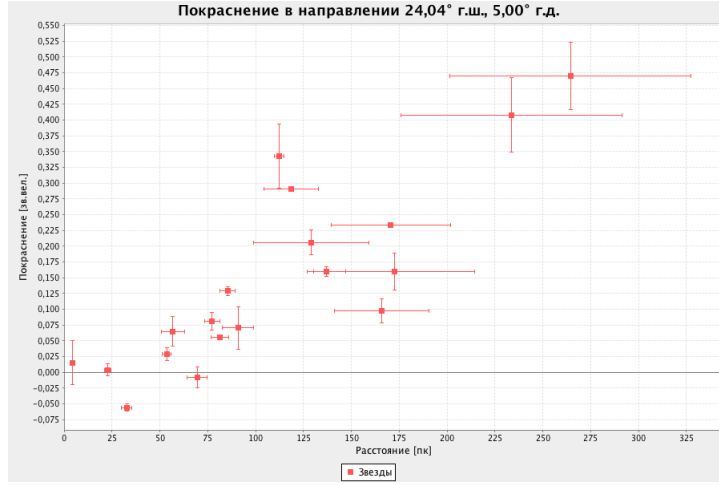
Класс	Точность	Полнота	F1-мера
III	98%	79%	88%
V	91%	99%	95%

Классификатор, естественно, получился неплохим даже несмотря на то, что многие звезды III класса из обучающего множества ошибочно находятся на главной последовательности (отсюда, низкая полнота у III класса). К подобным каталожным ошибкам мы вернемся в следующей статье.

Результат работы классификатора — наличие класса светимости у всех рассматриваемых звезд.

6 Результаты

Построение кривых покраснения в разных направлениях позволит нам понять пространственное распределение межзвездного поглощения (пыли). Звезд в каталоге не бесконечное число, поэтому реальные кривые покраснения будут не непрерывными кривыми, а будут наборами точек, описывающими ход покраснения. Аналогично, вместо звезд на луче зрения мы должны использовать звезды в малых конусах. Поэтому, ход покраснения у нас будет выглядеть, к примеру, так,



Следующий метод позволит построить «кривые» покраснения во всех направлениях на небесной сфере. Он состоит из трех этапов,

6.1 Картирование небесной сферы

Для картирования небесной сферы мы воспользуемся стандартным алгоритмом Healpix [ссылка], с помощью которого мы разобьем сферу на достаточно маленькие равновеликие части. Назовем их $\{P_i\}_{i=1}^{2n^2}$ (у нас $n = 18$). Обозначим конусы, высекаемые соответствующими частями через $\{C_i\}_{i=1}^{2n^2}$. Такое разбиение позволит нам,

1. Рассмотреть ход покраснения в каждом конусе как одномерную функцию $E(r)$. Это корректно, ввиду того, что конусы достаточно узкие;
2. Сделать наши результаты «независимыми», т.к. конусы не налагаются;
3. Поместить в каждый конус примерно одинаковое число звезд, чтобы избежать недостатка звезд в некоторых конусах.

6.2 Тренд

Тем самым, мы ищем $E_i(r)$, соответствующую каждому C_i . Ввиду того, что практически все $E_i(r)$ очень сильно зашумлены разного рода ошибками, не удастся проследить истинный ход этих функций. Но тренд вида kr все же можно вычислить ($E_i(r) \approx k_i r$). Он находится с помощью метода наименьших квадратов.

6.3 Критерии выбора звезд

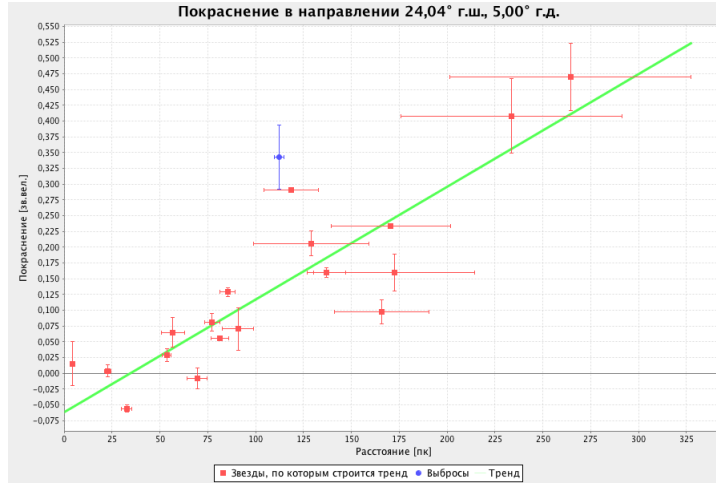
6.3.1 По параллаксу

В каталоге Hipparcos, далекие звезды имеют очень большие ошибки в параллаксе. На расстояниях, скажем, в 400 пк эти ошибки могут достигать 100%. Такие ошибки могут очень сильно испортить наши результаты. Тем самым, мы не будем рассматривать звезды, у которых относительная ошибка параллакса не превосходит 25%. Меньшее значение порога оставит нам очень малое число звезд, которых не хватит для того, чтобы вычисленные тренды были достоверными.

6.3.2 По тренду

Покраснения некоторых звезд очень сильно портят тренды. Иногда они даже бывают отрицательными, что вообще противоречит здравому смыслу. Но об этом мы поговорим позже.

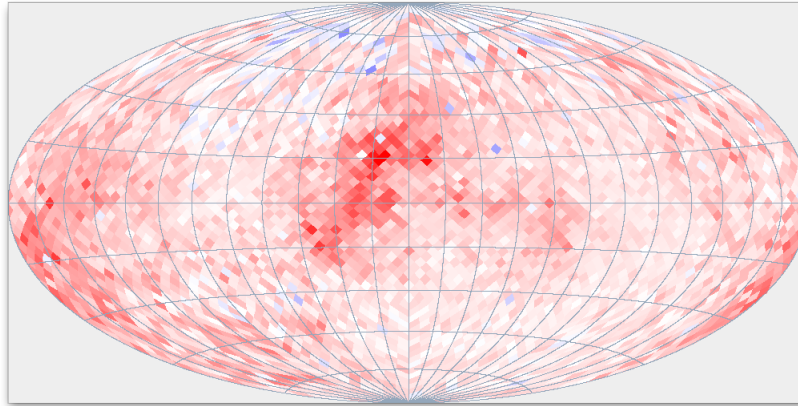
Как известно, метод наименьших квадратов не устойчив к выбросам, т.е. совсем неверные покраснение/параллакс могут очень сильно испортить тренд. Для более устойчивого построения тренда, мы сделаем следующее. После построения тренда по всем звездам в конусе, мы выбрасываем те, у которых отклонение от тренда самое большое. Затем, мы строим тренд заново, но уже только по оставшимся звездам. После выброса 10% самых плохих звезд тренд, к примеру, может быть таким



В данном случае, мы выбросили одну звезду (синюю). 10% звезд — это обычно 0, 1 или 2 звезды в каждом конусе. Опять же, нельзя выкидывать слишком много звезд из-за опасности «подгонки» модели.

6.4 Распределение коэффициента k

Тем самым, ход покраснения в конусе C_i мы описываем одним числом k_i — скоростью роста покраснения в этом конусе. Она, как мы ранее выясняли, коррелирует с наличием пыли. Поэтому, составив карту распределения коэффициента k , мы составим панораму пыли в окрестности Солнца.



Это небо в галактической системе координат с центром в центре галактики. Сфера разбита на 3888 кусочков алгоритмом Healpix. В каждом кусочке построен тренд покраснения kr . На этом рисунке изображено распределение значения коэффициента k по кусочкам. Синий цвет означает отрицательное

значение k , красный - положительное. Чем насыщеннее цвет, тем больше значение коэффициента по модулю.

7 Отрицательное поглощение

Примеры отрицательного поглощения. 1) Существуют звезды с аномальным значением поглощения: значимое отрицательное приращение к показателю цвета. Случайные отклонения, связанные, например, с необычными спектральными характеристиками. 2) Существуют «картинки», в которых целый ряд звезд имеют неправильное поглощение – систематическая ошибка – попытаться выяснить природу. Таблица наиболее значительных отклонений от нормы. Диаграмма.

8 Заключение

Получены определенные результаты. Найдены области с большим значением K (красные пиксели). Вопросы: отриц. поглощение – в дальнейшей работе будет проведено исследование.

9 Библиография

хотя бы 5-6 ссылок. 1. Каталог Hip 2007 2. Tycho 2 Sp T 3. Страйджест – книжка 4. Неточности в спектральной классификации звезд каталога Tycho-2 Spectral Type А.А.Сминов, А.С.Цветков, А.В.Попов 2006 5. Гончаров

10 Обзор литературы

Этот раздел еще не создан

11 Недостатки используемых методов

Этот раздел еще не создан

12 Метод нахождения параметров межзвездного поглощения

Нижеописанный метод состоит из трех этапов

12.1 Картирование небесной сферы

Для картирования небесной сферы мы воспользуемся стандартным алгоритмом Healpix, с помощью которого мы разобьем сферу на достаточно маленькие равновеликие части. Назовем их $\{P_i\}_{i=1}^{2n^2}$ (у нас $n = 18$). Обозначим конусы, отсекаемые соответствующими частями через $\{C_i\}_{i=1}^{2n^2}$

Такое разбиение позволит нам,

1. Рассмотреть ход покраснения в каждом конусе как одномерную функцию $E(r)$. Это корректно, ввиду того, что конусы достаточно узкие;
2. Сделать наши результаты «независимыми», т.к. конусы не налагаются;
3. Поместить в каждый конус примерно одинаковое число звезд, чтобы избежать недостатка звезд в некоторых конусах.

12.2 Тренд

Тем самым, мы ищем $E_i(r)$, соответствующую каждому C_i . Ввиду того, что практически все $E_i(r)$ очень сильно зашумлены разного рода ошибками, не удастся проследить истинный ход этих функций. Но тренд все же можно вычислить. Он находится с помощью метода наименьших квадратов.

12.3 Критерии выбора звезд

12.3.1 По параллаксу

В каталоге Hipparcos, далекие звезды имеют очень большие ошибки в параллаксе. На расстояниях, скажем, в 400 пк эти ошибки могут достигать 100%. Такие ошибки могут очень сильно испортить наши результаты. Тем самым, мы не будем рассматривать звезды, у которых параллакс (его среднее значение) превосходит 300 пк. Меньшее значение порога оставит нам очень малое число звезд, которых не хватит для того, чтобы вычисленные тренды были достоверными.

12.3.2 По тренду

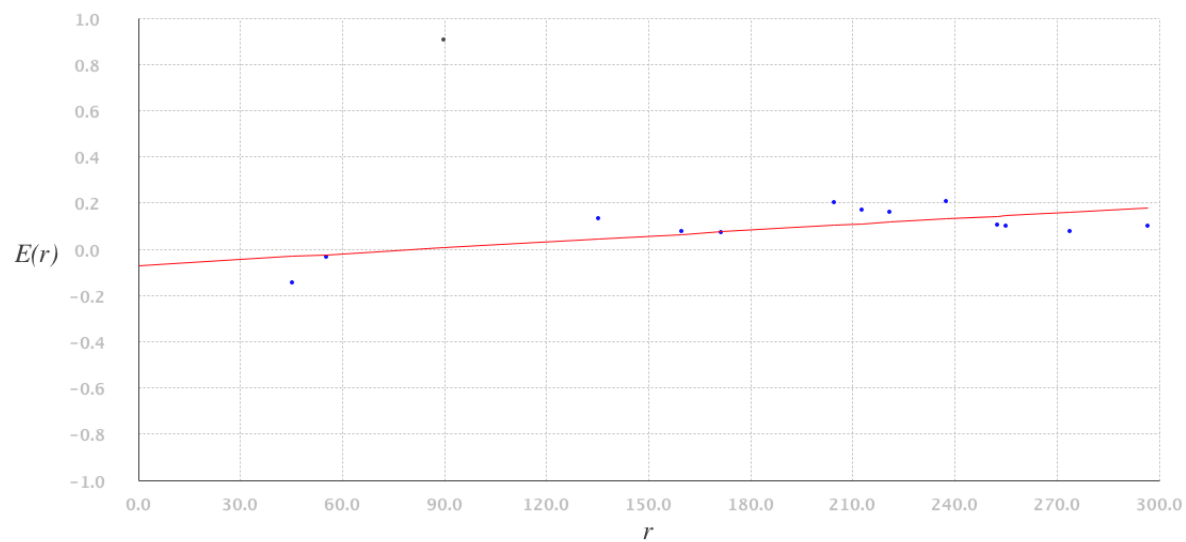
Покраснения некоторые звезды очень сильно портят тренды. Иногда они даже бывают отрицательными, что вообще противоречит здравому смыслу. Но об этом мы поговорим позже.

Как известно, метод наименьших квадратов не устойчив к выбросам, т.е. совсем неверные покраснение/параллакс могут очень сильно испортить тренд. Это пример $E(r)$ и тренда в одном из конусов,



На этом графике видно, как аномальное значение покраснения звезды (той, что сверху) «приподнимает» начало тренда.

Для устойчивого построения тренда, мы воспользуемся первой итерацией метода RANSAC. Она заключается в том, что после построения тренда по всем звездам, мы выбрасываем те, у которых отклонение от тренда самое большое. Затем, мы строим тренд заново, но уже только по оставшимся звездам. В методе RANSAC эта процедура выполняется много раз, но нам это не подойдет, т.к. тогда мы будем заниматься «подгонкой». После выброса 10% самых плохих звезд тренд в том же конусе становится таким,



Что чуть больше похоже на правду. В данном случае, мы выбросили одну звезду. 10% звезд — это 0, 1, 2 звезды в каждом конусе. Опять же, нельзя выкидывать слишком много звезд из-за опасности «подгонки» модели.

12.4

Подытожим все вышесказанное. Составим алгоритм расчета