

# Поиск аномалий в Открытом Каталоге Сверхновых Anomaly Detection in the Open Supernova Catalog

Maria Pruzhinskaya et al., Moscow University

9 октября 2017 г.

## Аннотация

Грядущее поколение крупных астрономических обзоров произведет революцию в нашем понимании Вселенной, однако при этом нам придется столкнуться с огромными объемами данных и связанными с ними проблемами. Одна из них — это невозможность использования человеческих ресурсов для поиска новых необычных астрофизических объектов. Кроме того, большинство данных, полученных этими обзорами, будут фотометрическими, и мы не сможем рассчитывать на спектроскопическую поддержку всех наблюдений. Целью проекта является поиск сверхновых звезд с уникальными свойствами в Открытом Каталоге Сверхновых (<http://sne.space/>) методами машинного обучения. В данном проекте мы объединим усилия экспертного анализа и методы машинного обучения. Проект состоит из двух частей: разработка программного обеспечения, которое позволит обнаружить аномалии в данных, и тщательный астрофизический анализ всех объектов с необычными свойствами. Для первичной классификации сверхновых по известным типам и подтипам требуется начальная выборка объектов со спектрами в качестве учебного образца, однако потом полученный классификатор может быть применен к большому набору данных, для которых доступны только фотометрические наблюдения. Если мы не хотим пропускать новую науку, скрытую в больших объемах данных, получение которых сопряжено с огромными умственными и финансовыми затратами, подобные инструменты поиска и анализа аномалий просто необходимы.

## Abstract

The next generation of astronomical surveys will revolutionize our understanding of the Universe, raising unprecedented data challenges in the process. One of them is the impossibility to rely on human scanning for the identification of unusual/unpredicted astrophysical objects. Moreover, given that most of the available data will be in the form of photometric observations, such characterization cannot rely on the existence of high resolution spectroscopic observations. The goal of this project is to detect the anomalies in the Open Supernova Catalog (<http://sne.space/>) with use of machine learning. We will develop a pipeline where human expertise and modern machine learning techniques can complement each other. Using supernovae as a case study, our proposal is divided in two parts: a first developing a strategy and pipeline — where anomalous objects are identified, and a second phase where such anomalous objects submitted to careful individual analysis. The strategy requires an initial data set for which spectroscopic is available for training purposes, but can be applied to a much larger data set for which we only have photometric observations. This project represents an effective strategy to guarantee we shall not overlook exciting new science hidden in the data we fought so hard to acquire.

## Ключевые слова

Сверхновые, транзиенты, машинное обучение в астрофизике

## 1 Описание фундаментальной научной задачи

Сверхновые звезды (СН) являются одними из самых мощных, а потому и интересных взрывов во Вселенной. Благодаря им происходит обогащение межзвездной среды химическими элементами; волны плотности, возникающие во время взрыва, служат толчком к звездообразованию; СН являются источником космических лучей высоких энергий; кроме того, с помощью сверхновых мы изучаем состав Вселенной и измеряем расстояния до далеких галактик.

Появление высокоточных и крупных обзоров по поиску сверхновых звезд в последние годы (Panoramic Survey Telescope & Rapid Response System (Pan-STARRS), Intermediate Palomar Transient Factory (iPTF), SDSS Supernova Survey, Supernova Legacy Survey (SNLS), All-Sky Automated Survey for Supernovae (ASAS-SN), MASTER supernovae, Nearby Supernova Factory и т.п.) привело к необходимости развития инструментов автоматического анализа большого количества данных. Подобные задачи открывают огромные возможности и являются вызовом как для астрономов, так и для специалистов по машинному обучению.

Недостаток спектроскопической поддержки в крупных обзорах по поиску СН приводит к острой необходимости в фотометрической классификации сверхновых. В представленном проекте мы предлагаем решать эту

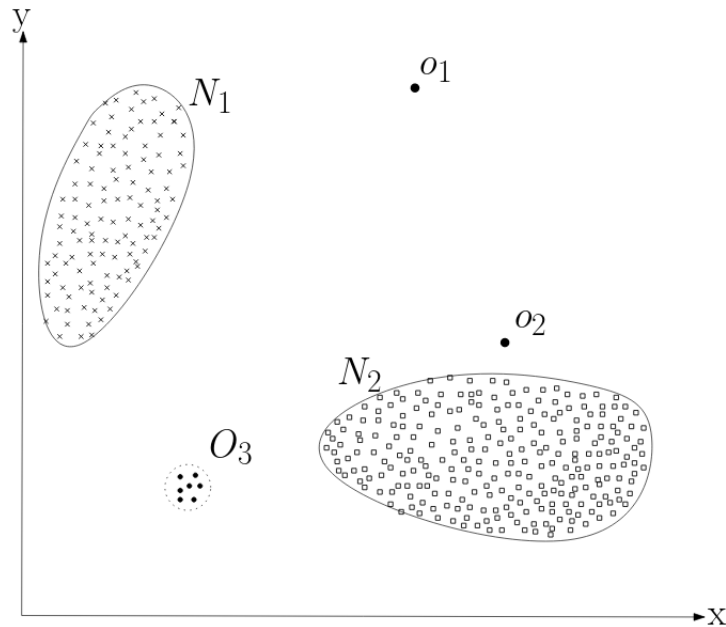


Рис. 1: Аномалии представляют собой данные с необычным поведением. Здесь представлены аномалии на примере выборки двумерных отсчетов. В основном наблюдения лежат в областях  $N_1$  и  $N_2$ . Данные, которые находятся на значительном удалении от  $N_1$  и  $N_2$  — аномалии ( $o_1$  и  $o_2$  — индивидуальные,  $O_3$  — групповая). Рисунок взят из Chandola et al., Anomaly detection: A survey, 2009, ACM Comput. Surv. 41, 3, Article 15, DOI=10.1145/1541880.1541882.

проблему путем анализа большой выборки сверхновых методами машинного обучения. Машинное обучение также поможет "очистить" выборку от вклада объектов, не являющихся СН — с подобной проблемой сталкиваются все базы данных/каталоги сверхновых звезд, которые формируются не только на результатах спектральной классификации, но и на основе одних фотометрических данных по присутствию "вторичных индикаторов": близость к галактике, переменность, поведение кривой блеска (наклон), абсолютная звездная величина. Объекты, не являющиеся СН, будут классифицированы. Кроме того, подобный анализ, возможно, позволит обнаружить новые, ранее не классифицированные объекты, а также сверхновые, отличные по своему поведению от остальных. В качестве примера таких уникальных объектов можно привести СН 2006jc — СН с очень мощными, но относительно узкими линиями He I в ранних спектрах (сейчас таких СН около 30; Pastorello et al. 2016, MNRAS), СН 2005bf — представляющую из себя СН Ib, но кривая блеска которой имеет два широких, ярких максимума, СН 2010mb — уникальная сверхновая Ic с очень медленным спаданием блеска после максимума, что не согласуется с радиоактивным распадом  $^{56}\text{Ni}$ , ASASSN-15lh — некоторое время считалась самой яркой из когда-либо наблюдавшихся СН (в два раза ярче, чем подкласс сверхъярких СН, к которому ее причислили), позднее принадлежность данного объекта к СН была поставлена под сомнение, сейчас считается, что наблюдалось приливное разрушение звезды главной последовательности сверхмассивной черной дырой.

Поиск подобных объектов с последующим подробным изучением является основной задачей данного проекта. Поскольку ожидается, что такие объекты довольно редки, эту задачу можно сформулировать как проблему поиска аномалий в крупных каталогах сверхновых звезд.

Мы будем рассматривать два типа аномалий: индивидуальные и групповые (см. рисунок в приложении). Индивидуальные аномалии включают единичные объекты с уникальными характеристиками. Групповые аномалии представляют собой аномальную группу объектов, находящихся близко друг к другу в некотором пространстве параметров. И те, и другие представляют огромный интерес для астрофизики, и наша цель — научиться находить их автоматически с помощью машинного обучения.

Все обнаруженные аномалии будут тщательно исследованы с привлечением всей доступной информации, в том числе в гамма, рентгеновском, ультрафиолетовом, оптическом, инфракрасном, микроволновом и радио диапазонах. Кроме того, физическая природа аномальных сверхновых будет изучена с помощью одномерного сферически-симметричного радиационно-гидродинамического кода Стелла, который позволяет рассчитывать процессы переноса излучения в сверхновых с учетом химического состава и внутренней структуры предсверхновых (Blinnikov et al. 1998, ApJ; Blinnikov et al. 2006, A&A).

## 2 Актуальность исследования

Развитие наблюдательной базы, появление крупных обзоров неба привело к открытию огромного числа сверхновых и кандидатов в сверхновые. Только 10% СН, открываемых каждый год, получает спектральное подтверждение. Количество астрономических данных продолжает расти и уже превышает человеческие возможности по их обработке. В то время как сейчас в нашем распоряжении имеются десятки тысяч СН и кандидатов в СН,

в течение своей десятилетней программы Большой обзорный телескоп (LSST) должен открыть 10 миллионов сверхновых, и только малая их часть будет иметь спектры (LSST Science Book, 2009). Ожидается, что LSST позволит получить кривые блеска  $>100$  тысяч сверхновых, но перед тем, как эти сверхновые можно будет использовать для астрофизических целей, их нужно классифицировать. Именно для подобных задач и необходимо машинное обучение. Оно позволит не только определить тип сверхновой, но и обнаружить другие транзитные объекты (новые, послесвечения гамма-всплесков, килоновые, оптические компаньоны гравитационно-волновых алертов), которые по ошибке были классифицированы как СН, а также выявить аномальные астрофизические объекты с необычными свойствами. Поиск новых объектов с последующим подробным изучением является одной из приоритетных астрономических задач.

Кроме того, в последние годы проблема классификации сверхновых обращает на себя все больше и больше внимания. Как мы уже отмечали, это происходит, во-первых, из-за недостатка спектроскопической поддержки. Во-вторых, существующая классификация СН уже устарела и нуждается в пересмотре. Например, детальные исследования показали, что сверхновые IP и IIL не представляют два подкласса СН, но существует континуум кривых блеска от IP до IIL (Anderson et al. 2014, ApJ; Sanders et al. 2015, ApJ). Появилось много новых типов и подтипов СН: СН Iaх, чьи спектры в максимуме блеска схожи со спектрами нормальных СН Ia, однако имеют меньшую светимость (проблема для космологии в случае ложной классификации); СН Ibp, сверхновые без водорода в спектрах с признаками взаимодействия с окружающей средой; СН Ib с двумя максимумами на кривой блеска; СН Ib с преобладанием линий Ca II в поздних спектрах (Ca-rich СН Ib); новый класс сверхъярких сверхновых, представленный тремя подклассами (SLSN-I, SLSN-II, SLSN-R), и т.д. Таким образом, классификация, которая изначально основывалась в основном на спектрах в максимуме блеска, становится все более сложной и временами сбивающей с толку, к тому же она иногда не отражает физические различия между СН. Машинное обучение позволит рассмотреть существующие типы и подтипы СН на многомерной сетке параметров и, возможно, поможет выявить какие-то новые особенности, которые в будущем смогут стать основой для более удачной классификации. Наличие таких особенностей/общих свойств (в многомерной выборке параметров) иногда не очевидно при ручном анализе, но может быть обнаружено с помощью методов машинного обучения.

### 3 Анализ современного состояния исследований в данной области

О проблеме выявления аномалий в данных было опубликовано много литературы (Chandola et al., Anomaly detection: A survey, 2009, ACM Comput. Surv. 41, 3, Article 15, DOI=10.1145/1541880.1541882), однако в основном ее можно разделить на два типа: обучение "с учителем" и обучение "без учителя". В обучении "без учителя" образцы не размечены и, следовательно, отсутствует понятие обучающей выборки. Задача алгоритма объединить данные в группы по степени схожести или расстояния между их наблюдательными свойствами в некоторой метрике. При таком рассмотрении за аномалии принимаются объекты, чьи группы наиболее малочисленны (по сравнению с объемом данных) или группы объектов, которые находятся далеко от остальных в рассматриваемом пространстве параметров. С другой стороны, в обучении "с учителем" поиск аномалий может рассматриваться как задача классификации, где обучающая выборка, содержащая заранее известные классы объектов, используется для создания классификатора, с помощью которого можно впоследствии разделить нормальные и аномальные данные (Goernitz et al., Toward Supervised Anomaly Detection, 2013, J. Artificial Intelligence Research, vol 46, pp. 235-262, arXiv1401.6424).

Автоматическое обнаружение аномалий в астрономии еще не было полностью реализовано по отношению к крупным каталогам. По существу, за некоторыми исключениями, большинство предыдущих исследований можно разделить на два различных направления: кластеризация (Rebbapragada et al., Finding anomalous periodic time series, 2009, Machine Learning, vol 74, Issue 3, pp. 281-313) и методы анализа подпространств (Henrion et al., CASOS: a Subspace Method for Anom. Detec. in HD Astron. Data, 2013, Stat. Anal. and Data Mining, vol 6, pp. 53-72). Недавно, стали применяться алгоритмы случайного леса как сами по себе (Baron & Poznanski 2017, MNRAS), так и в составе гибридного статистического анализа (Nun et al. 2014, ApJ). Также в недавней работе Nun et al. 2016, AJ были предложены новые алгоритмы для поиска уникальных объектов на примере каталога MACSO (Massive compact halo object).

### 4 Цель и задачи проекта

В этом проекте мы обратились к проблеме поиска аномалий/новых объектов в больших выборках сверхновых. Когда имеешь дело с большим и в основном фотометрическим набором данных, велика вероятность "пропустить" что-то новое. Именно в таких ситуациях машинные алгоритмы поиска аномалий могут сыграть значительную роль. Поиск аномалий нацелен на выявление объектов, чьи свойства отличны от свойств большинства. Однако определенные методом машинного обучения аномалии могут оказаться артефактами самого метода, и потому каждая из них должна быть изучена исследователем. Таким образом, наша цель не просто отождествление странных объектов, но и их критический анализ с привлечением всей доступной информации и радиационно-гидродинамического моделирования.

Одна из основных проблем, возникающих на пути внедрения алгоритмов машинного обучения в астрономию — это отсутствие благоприятного для астрономов окружения/интерфейса/удобных программ, которые позволяли бы использовать методы машинного обучения для повседневных проблем астрономических исследований. Поэтому одна из целей проекта состоит в создании программного обеспечения, которое облегчит применение

инструментов машинного обучения в астрономических задачах и обеспечит набор инструментов для работы с фотометрическими обзорами. Принимая во внимание опыт руководителя проекта в области СН (в частности в контексте больших данных LSST) и опыт исполнителя в современных машинных алгоритмах, мы намерены объединить эти два аспекта и разработать специальный инструмент для обнаружения аномалий. Кроме того, программы будут разработаны таким образом, что в дальнейшем смогут быть использованы не только применительно к сверхновым, но и другим данным.

## 5 Научная новизна

Астрономы уже использовали машинное обучение для решения некоторых задач (Ball & Brunner, *Data Mining and Machine Learning in Astronomy*, 2010, *International Journal of Modern Physics D*, Volume 19, Issue 07, pp. 1049-1106), в частности при поиске экзопланет (McCauliff et al. 2015, *ApJ*; Thompson et al. 2015, *ApJ*, Pearson et al. 2017, arXiv170604319), однако синергия еще далека от достигнутой в других областях, таких как генетика (Libbrecht & Noble, *ML applications in genetics and genomics*, 2015, *Nature Reviews Genetics* 16, pp. 321-332), экология (Crisci et al., *A review of ML and applications to ecology*, 2012, *Eco. Modelling*, v. 240, pp. 113-122), или медицина (Mathukumalli Vidyasagar, *Identifying Predictive Features in Drug Response Using Machine Learning: Opportunities and Challenges*, 2015, *Ann. Rev. of Pharm. and Tox.*, vol. 55, pp. 15-34), где новые научные задачи провоцировали развитие новых алгоритмов машинного обучения. Кроме того, учитывая относительно недавнее появление больших объемов данных в астрономии, в основном машинное обучение здесь применяется к задачам классификации (Kessler et al. 2010, *PASP*; Ishida & de Souza 2013, *MNRAS*; Lochner et al. 2016, *ApJS*) и регрессии (Hildebrandt et al. 2010, *A&A*; Cavuoti et al. 2015, *MNRAS*).

Однако машинное обучение практически не применялось для анализа транзиентов, а еще меньше для сверхновых. Поэтому предложенный проект является в своем роде уникальным.

## 6 Предлагаемые подходы и методы

При поиске интересных, необычных объектов в астрономии методы автоматического определения аномалий только первый шаг, за которым следует детальное экспертное исследование. Разные методы эффективны для обнаружения разных типов аномалий в различных обстоятельствах, и было бы наивно ожидать, что один единственный метод будет оптимальным. Поэтому мы будем применять несколько подходов, подобно тому, как предложено в работе Nun et al. 2016, *AJ*, комбинируя различные методы поиска аномалий: метод случайного леса (который относится к алгоритмам кластеризации), метод опорных векторов с одним классом (который можно классифицировать как алгоритм на основе расстояния) и смешанные гауссовы модели (которые являются частью статистических алгоритмов).

В качестве источника данных будет использоваться Открытый Каталог Сверхновых ("Open Supernova Catalog; <http://sne.space/>; Guillochon et al. 2017, *ApJ*). Каталог объединяет в себе данные многих других каталогов (Asiago Supernova Catalog, Carnegie Supernova Project, Gaia Photometric Science Alerts, Nearby Supernova Factory, Pan-STARRS, SDSS Supernova Survey, Sternberg Astronomical Institute Supernova Light Curve Catalogue, SNLS, ASAS-SN, iPTF, и т.п.) и отдельных публикаций. Наш выбор обусловлен тем, что каталог включает данные для 41989 СН/кандидатов в СН (16905 из которых имеют свободно доступные кривые блеска в простом формате, а 5516 имеют спектр). Для сравнения, каталог сверхновых SDSS содержит только 3911 СН с кривыми блеска и 588 СН со спектрами.

Исходный каталог содержит данные в разных фотометрических полосах. Большинство измерений доступно в полосах системы SDSS: *ugriz*. Для достижения однородности данных на первом этапе будут использоваться данные только из этих фотометрических полос, также будут отобраны объекты с хорошим отношением сигнал/шум. Важная задача состоит в том, чтобы привести кривые блеска в пространство одинаковой размерности. Для этого предполагается использовать аппроксимацию каждой индивидуальной кривой блеска Гауссовыми процессами с помощью уже существующего кода (Leget 2016, PhD thesis). Если потребуется, можно пополнить обучающий набор данных синтетическими кривыми блеска, полученными с помощью параметров аппроксимации реальных кривых Гауссовыми процессами. Эта идея недавно была успешно использована для фотометрической классификации сверхновых (Revsbech et al. 2017, arXiv:astro-ph/1706.03811). При необходимости, могут быть применены методы понижения размерности входных данных, такие как анализ главных компонент или метод основанный на вейвлетах (Lochner et al. 2016, *ApJS*). Целесообразность такого шага предстоит определить в ходе работы.

Для выявления аномалий мы хотели бы опробовать каждый из представленных ниже методов, сверяя затем их результаты между собой:

1. Случайный лес. Основная идея состоит в том, чтобы получить статистический ансамбль слабых классификаторов на основе деревьев принятия решений. Такой ансамбль способен давать хорошую классификацию на основе консенсуса.

2. Метод опорных векторов. Классический метод опорных векторов с одним классом позволит нам линейно разделить множество наших данных в пространстве повышенной размерности. В таком пространстве мы ожидаем, что наши нормальные данные будут разделены на две области повышенной плотности, а аномалии окажутся представлены отдельными точками или точками в областях меньшей плотности.

3. Смешанные Гауссовы модели. Можно моделировать статистическое распределение с помощью множества отдельных нормальных распределений. Чтобы определить число требуемых распределений, мы собираемся ис-

пользовать процедуру, описанную в de Souza et al. 2017, MNRAS. После оценки параметров нормальных распределений, например, с помощью EM-алгоритма, каждая аномалия должна будет иметь чрезвычайно низкую вероятность реализации.

Последующий анализ найденных аномалий будет включать сбор всей доступной информации по этим объектам в опубликованной литературе и каталогах. Особое внимание будет уделено данным, полученным в других диапазонах длин волн (гамма, рентген, ИК, радио).

Основным инструментом для анализа аномалий в наборе сверхновых будет код Стелла (Blinnikov et al. 1998, ApJ; Blinnikov et al. 2006, A&A). Стелла — это одномерный сферически-симметричный радиационно-гидродинамический код, который позволяет рассчитывать процессы переноса излучения в сверхновых с учетом химического состава и внутренней структуры предсверхновых. Теоретические спектры, рассчитываемые в определенные моменты времени после взрыва, позволяют моделировать многоцветные кривые блеска и фотосферные скорости расширяющейся оболочки СН. Расчеты учитывают распределение химических элементов в предсверхновой, взаимодействие между внутренними слоями оболочки и компактного ядра. Еще одним преимуществом подобного моделирования является исследование поглощения на луче зрения между наблюдателем и объектом путем сравнения модельных и теоретических кривых блеска в разных фильтрах. Результаты Стеллы хорошо согласуются с результатами других широко известных гидродинамических кодов (см. Woosley et al. 2007, ApJ; Kromer & Sim 2009, MNRAS; Sim et al. 2010, ASP Conf. Ser.; Kozuyeva et al. 2017, MNRAS). Код Стелла был многократно использован для моделирования различных типов сверхновых (Ia, Blinnikov et al. 2006, A&A; Ib/Ic, Folatelli et al. 2006, ApJ, Tauris et al. 2013, ApJ; Ib, Blinnikov et al. 1998, ApJ; IIn, Chugai et al. 2004, MNRAS; IP, Baklanov, Blinnikov & Pavlyuk 2005, Astron. Lett., Tominaga et al. 2009, ApJ). С помощью Стеллы можно определить физические свойства аномальных СН, такие как массу и радиус предсверхновой, ее химический состав, энергию взрыва и т.д.

## 7 Ожидаемые результаты научного исследования и их научная и прикладная значимость

В качестве основных результатов проекта можно выделить:

- создание классификатора сверхновых звезд
- обнаружение случаев ложной классификации в Открытом Каталоге Сверхновых
- обнаружение аномальных объектов
- создание пакета программ и соответствующего пользовательского интерфейса, которое облегчит применение инструментов машинного обучения в астрономических задачах и в будущем может быть использовано не только применительно к сверхновым, но и к другим данным.

Открытие интересных аномальных объектов может спровоцировать дальнейшие более тщательные наблюдения, возможно, с привлечением крупнейших телескопов мира.

## 8 Общий план работ на весь срок реализации Проекта

1 год:

- отбор сверхновых для обучающего набора
- предварительный ручной анализ данных
- реализация метода смешанных Гауссовых моделей
- тщательное исследование найденных аномалий/групп аномалий
- подготовка публикации

2 год:

- реализация метода случайного леса
- реализация метода опорных векторов
- анализ и сравнение результатов, получаемых разными методами
- создание пакета программ по поиску аномалий
- тщательное исследование найденных аномалий/групп аномалий
- подготовка публикации

## 9 План работ на первый год реализации Проекта

- отбор сверхновых для обучающего набора в соответствии с разработанными критериями: наличие спектра, отношение сигнал/шум, необходимое количество фотометрических данных (отв. исполнитель: Пружинская М.В.)
- предварительный ручной анализ данных: аппроксимация Гауссовыми процессами, уменьшение размерности (отв. исполнители: Корнилов М.В., Пружинская М.В.)
- реализация метода смешанных Гауссовых моделей (отв. исполнитель: Корнилов М.В.)
- тщательное исследование найденных аномалий/групп аномалий в помощью различных каталогов и кода Стелла (отв. исполнитель: Пружинская М.В.)
- подготовка публикации
- выступления участников проекта на конференциях

## 10 Ожидаемые результаты за первый год реализации проекта

- создание классификатора сверхновых на основе метода смешанных Гауссовых моделей
  - выявление в Открытом Каталоге Сверхновых объектов с необычными свойствами
  - выявление физической природы аномалий (в случае если аномалия окажется не СН, а, например, каким-нибудь перманентным объектом с признаками активности, можно будет подать заявку на наблюдения на крупных международных телескопах)

## 11 Имеющийся у коллектива научный задел по Проекту

Руководитель проекта Пружинская М.В. имеет опыт работы с транзиентами различной природы, включая их последующее изучение. Во время работы в проекте МАСТЕР (<http://observ.pereplet.ru/>, Lipunov et al. 2010, *Advances in Astronomy*), Пружинская М.В. была вовлечена в анализ различных транзиентных источников (СН, новых, квазаров и т.д.). В течение двух лет Пружинская М.В. была также членом рабочей группы LSST, где в основном занимаются обработкой большого количества данных с привлечением современных алгоритмов и компьютерных языков. Пружинская М.В. умеет моделировать кривые блеска коллапсирующих сверхновых с помощью радиационно-гидродинамического кода Стелла для различных параметров предсверхновых. В частности, используя код Стелла, в работе Moriya, Pruzhinskaya, et al. 2016, *MNRAS* авторы изучали природу аномалии (бампа) на кривых блеска быстро ослабевающих СН II. Недавно этот код был также использован руководителем проекта для изучения СН 2013dx, принадлежащей к редкому подклассу СН, ассоциированных с гамма-всплесками (~40 объектов; Volnova, Pruzhinskaya, et al. 2017, *MNRAS*).

Для проекта очень важно, что его исполнитель Корнилов М.В. имеет богатый опыт в реализации различных алгоритмов нелинейной, ограниченной и дискретной оптимизации в рамках прикладного ПО для решения задач экспериментальной астрономии. Корнилов М.В. умеет решать ряд обратных задач в области экспериментальной астрономии, в частности восстановление вертикальных профилей оптической турбулентности. Также для предполагаемого исследования значим опыт, приобретенный Корниловым М.В. в процессе статистического анализа рядов данных, полученных в автоматическом режиме астроклиматическим монитором КГО (Kornilov, V.G., Kornilov, M.V., et al. 2016, *Astronomy Letters*; Kornilov, M.V. 2016, *Exp Astron*), выявления аномалий, апостериорного определения параметров работы аппаратуры, и коррекции измерений с помощью разнообразных статистических методов (Kornilov, V.G., Kornilov, M.V., et al. 2016, *MNRAS*). Для описания долговременного временного поведения мощности оптической турбулентности автор использовал Гауссовы линейные модели авто-регрессии скользящего среднего, схожие с теми, которые мы собираемся использовать при выполнении проекта. Данные измерений получаются в автоматическом режиме в массовом объеме, поэтому важной задачей является не только обработка данных в рамках физической модели, а определение степени их соответствия этой модели и апостериорного определения инструментальных причин возникновения несоответствий. Такая задача близка к задаче, поставленной в проекте.

Данные, которые планируется использовать для реализации целей проекта, находятся в общем доступе (<http://sne.space/>). Кроме того, в нашем распоряжении имеется вся необходимая инфраструктура для проведения предлагаемого исследования, включая наличие рабочих мест с электронным доступом к основным международным периодическим изданиям в области астрофизики и архивам наблюдательных данных, а также необходимые вычислительные мощности, в том числе компьютерный кластер ГАИШ МГУ и доступ к суперкомпьютерам Московского университета.

Band name	Number of SNe with at least 1 spectrum	Total Number of SNe
B	808	1171
B'	1	5
Bj	1	1
C	210	1532
F625W	0	1
F775W	0	15
F850LP	0	24
G	332	2010
H	200	245
I	772	1562
Ic	3	3
J	200	253
Js	3	4
K	46	50
KS	3	3
Ks	79	109
LSQGR	1	1
M2	181	253
NUV	2	3
R	726	944
R1pg	1	2
RM	5	33
ROTSE	1	1
Rc	3	3
U	352	439
V	859	1268
V'	1	5
VM	5	34
W1	189	265
W2	182	255
Y	90	107
Z	7	31
blue	1	1
g	426	695
g'	228	4854
g+r	1	2
gr	4	7
i	628	3029
i'	364	5021
pg	18	47
pv	2	9
r	532	1978
r'	366	5026
u	103	129
u'	224	4762
w	172	4723
y	2	12
z	260	430
z'	218	4835

Таблица 1: Statistics from sne.space. When the number of points in each filter is more than 5.