

# Machine Learning Analysis of Supernova Light Curves

---

**Maria Pruzhinskaya\*†**

*Lomonosov Moscow State University, Sternberg Astronomical Institute, Universitetsky pr. 13,  
Moscow 119234, Russia*

*E-mail: [pruzhinskaya@gmail.com](mailto:pruzhinskaya@gmail.com)*

**Konstantin Malanchev**

*Lomonosov Moscow State University, Sternberg Astronomical Institute, Universitetsky pr. 13,  
Moscow 119234, Russia*

*National Research University Higher School of Economics, 21/4 Staraya Basmannaya Ulitsa,  
Moscow 105066, Russia*

*E-mail: [malanchev@sai.msu.ru](mailto:malanchev@sai.msu.ru)*

**Matwey Kornilov**

*Lomonosov Moscow State University, Sternberg Astronomical Institute, Universitetsky pr. 13,  
Moscow 119234, Russia*

*National Research University Higher School of Economics 21/4 Staraya Basmannaya Ulitsa,  
Moscow 105066, Russia*

*E-mail: [matwey.kornilov@gmail.com](mailto:matwey.kornilov@gmail.com)*

**Emille E.O. Ishida**

*Université Clermont Auvergne, CNRS/IN2P3, LPC, F-63000 Clermont-Ferrand, France*

*E-mail: [emille.ishida@clermont.in2p3.fr](mailto:emille.ishida@clermont.in2p3.fr)*

**Florian Mondon**

*Université Clermont Auvergne, CNRS/IN2P3, LPC, F-63000 Clermont-Ferrand, France*

*E-mail: [florian.mondon@clermont.in2p3.fr](mailto:florian.mondon@clermont.in2p3.fr)*

**Alina Volnova**

*Space Research Institute of the Russian Academy of Sciences (IKI), 84/32 Profsoyuznaya Str,  
Moscow, Russia, 117997*

*E-mail: [alinuss@gmail.com](mailto:alinuss@gmail.com)*

**Vladimir Korolev**

*Central Aerohydrodynamic Institute, 1 Zhukovsky st, Zhukovsky, Moscow Region, 140180,  
Russia*

*Moscow Institute of Physics and Technology, 9 Institutskiy per., Dolgoprudny, Moscow Region,  
141701, Russia*

*E-mail: [balodja@gmail.com](mailto:balodja@gmail.com)*

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. We introduce an analysis of anomaly detection in the Open Supernova Catalog (<http://sne.space/>) with use of machine learning. We developed a strategy and pipeline — where anomalous objects are identified and then submitted to careful individual analysis. 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.

*Accretion Processes in Cosmic Sources — II - APCS2018*  
*3–8 September 2018*  
*Saint Petersburg, Russian Federation*

---

\*Speaker.

†<https://pruzhinskaya.com/snad/>

## 1. Introduction

Supernova stars (SNe) are ones of the most beautiful and interesting objects in the Universe. They are responsible for chemical enrichment of interstellar medium; density waves induced by their energetic explosions causes the star formation; SNe are origin of high energy cosmic rays; moreover, thanks to SNe we are studying the composition and distance scale of the Universe which defines its following destiny.

The generation of precise, large, and complete supernova surveys in the last years has increased the need of developing automated analysis tools to process this large amount of data. These scientific observations present both great opportunities and challenges for astronomers and machine learning (ML) researchers.

The lack of spectroscopic support makes the photometrical supernova typing top required. The analysis of big supernova dataset with ML methods is needed to distinguish the supernova by types on base of N-parameter grid. Such study allows us to purify the considered SN sample from non-supernova contamination as well — the problem, which is relevant for all large supernova database that collect SN candidates without careful analysis of each candidate and basing on the secondary indicators (proximity to the galaxies, transient behavior, arise/decline rate on light curves (LCs), absolute magnitude). It is also expected that during such analysis the unknown variable objects or SNe with unusual properties can be detected. As an example of unique objects one can refer to SN 2006jc — SN with very strong but relatively narrow He I lines in early spectra ( $\sim 30$  similar objects are known, [1]), SN 2005bf — supernova attributed to SN Ib but with two broad maxima on LCs, SN 2010mb — unusual SN Ic with very low decline rate after the maximum brightness that is not consistent with radioactive decay of  $^{56}\text{Ni}$ , ASASSN-15lh — for some time it was considered as the most luminous supernova ever observed (two times brighter than super-luminous SNe), later the origin of this object was challenged and now it is considered as a tidal disruption of a main-sequence star by a black hole. Finding such objects (and then studying them more closely) is one of the main aims of the current project. As such sources are typically rare, the task of finding them can be framed as an anomaly detection problem.

Astronomers have already benefited from developments in machine learning [2], in particular for exoplanet search [3, 4, 5], but the synergy is far from that achieved by other endeavors in genetics [6], ecology [7] or medicine [8], where scientific questions drive the development of new algorithms. Moreover, given the relatively recent advent of large data sets, most of the ML efforts in astronomy are concentrated in classification [9, 10, 11, 12, 13, 14] and regression [15, 16] tasks.

Astronomical anomaly detection has not been yet fully implemented in the enormous amount of data that has been gathered. As a matter of fact, barring a few exceptions, most of the previous studies can be divided into only two different trends: clustering [17] and subspace analysis [18] methods. More recently, random forest algorithms have been extensively used by themselves [19] or in hybrid statistical analysis [20]. Although all of this has been done to periodic variables there is not much done for transients and even less for supernova.

In this study we search the anomalies in photometrical data of the Open Supernova Catalog<sup>1</sup> [21]. We use the Isolation Forest as an outlier detection algorithm that identifies anomalies

---

<sup>1</sup><https://sne.space/>

instead of normal observations [22]. This technique is based on the fact that anomalies are data points that are few and different. Similarly to Random Forest it is built on an ensemble of binary (isolation) trees.

## 2. Data

### 2.1 The Open Supernova Catalog

The data are drawn from the Open Supernova Catalog [21]. The catalog is constructed by combining many publicly available data sources (such as Asiago Supernova Catalog, Carnegie Supernova Project, Gaia Photometric Science Alerts, Nearby Supernova Factory, Panoramic Survey Telescope & Rapid Response System (Pan-STARRS), SDSS Supernova Survey, Sternberg Astronomical Institute Supernova Light Curve Catalogue, Supernova Legacy Survey (SNLS), MASTER, All-Sky Automated Survey for Supernovae (ASAS-SN), iPTF, etc.) and from individual publications. It represents an open repository for supernova metadata, light curves, and spectra in an easily downloadable format. This catalog also includes some contamination from non-SN objects.

Our choice is justified by the fact that the catalog incorporates the data for more than  $5 \times 10^4$  SNe/SNe candidates ( $\sim 1.2 \times 10^4$  of SNe have  $> 10$  photometrical observations and  $\sim 5 \times 10^3$  of SNe have spectra). For comparison, SDSS supernova catalog contains only  $\sim 4 \times 10^3$  of SNe LCs and  $\sim 600$  SNe with spectra.

The catalog contains the data in different photometrical passbands. To have a more homogeneous data sample, we chose only those SNe that have LCs in  $g'r'i'$ ,  $gri$  or  $BRI$  filters. We assume that  $g'r'i'$  filters are close enough to  $gri$  and transform  $BRI$  to  $gri$  (see Sect. 2.2). We require  $\geq 3$  photometrical points in each filter with a 3-day binning. This threshold is justified by the fact that our approach aims to deal with big data where the human expertise will not be possible. However, the cadence of future transient surveys does not allow to get a well-sampled multicolour light curve for all the transients with a nice coverage. Considering the sample we work with, the "3-points" threshold does not mean that we have a little number of points on multicolour light curve in total. We need them to better reconstruct a "poor" light curve basing on the others (see Sect. 2.3). After this cut, our sample contains 3197 objects (2026 objects in  $g'r'i'$ , 767 objects in  $gri$ , and 404 objects in  $BRI$ ).

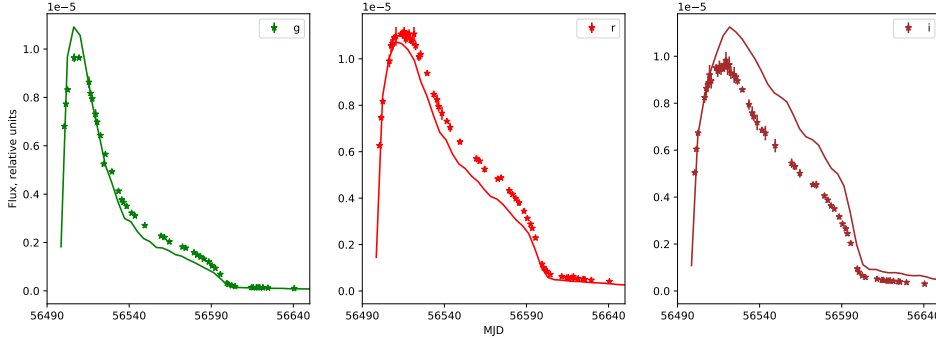
### 2.2 Transformation between $BRI$ and $gri$

To increase the sample we convert the Bessel's  $BRI$  into  $gri$  filters using the Lupton's (2005) transformation equations<sup>2</sup>. These equations are derived by matching SDSS DR4 photometry to Peter Stetson's published photometry for stars:

<sup>2</sup><http://www.sdss3.org/dr8/algorithms/sdssUBVRITransform.php>

$$\begin{cases} B = u - 0.8116 (u - g) + 0.1313 \\ B = g + 0.3130 (g - r) + 0.2271 \\ V = g - 0.2906 (u - g) + 0.0885 \\ V = g - 0.5784 (g - r) - 0.0038 \\ R = r - 0.1837 (g - r) - 0.0971 \\ R = r - 0.2936 (r - i) - 0.1439 \\ I = r - 1.2444 (r - i) - 0.3820 \\ I = i - 0.3780 (i - z) - 0.3974 \end{cases} \quad (2.1)$$

Before to apply the filter transformation, we fitted the LCs with Gaussian processes (see Sect. 2.3). First, we tried to use two Bessel’s filters to get *gri*. Obviously, in the Open Supernova Catalog more objects have photometry only in two filters than in three, and if such a transformation is enough we will have a larger sample. To check the quality of transformation with two filters only, we made a test: we chose few objects with LCs available in both, SDSS and Bessel’s filters, and compared the transformed *gri* with the original ones. As it is seen from the Fig. 1, the results of comparison are unsatisfactory. This means that at least one more filter has to be added in the analysis. The same test showed that three filters (*BRI*) are enough (Fig. 2).

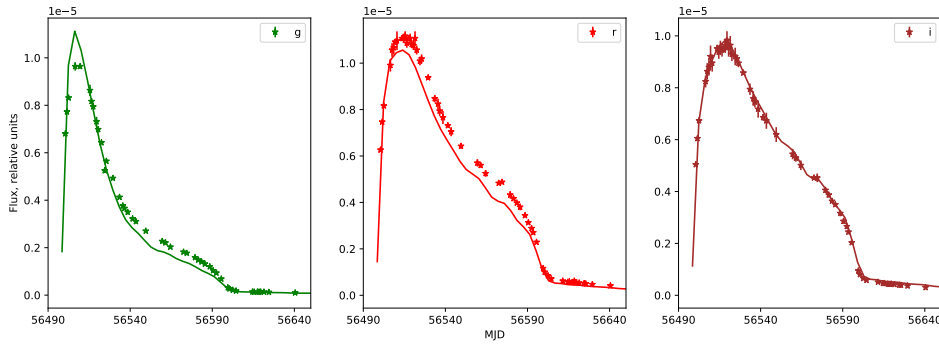


**Figure 1:** Light curves of SN2013ej. Points are the observations in *gri* filters [23, 24, 25]. Solid lines are the approximation of results of the transformation from *BR* to *gri* filters with use of Lupton’s (2005) transformation equations.

### 2.3 LCs fit

It is more convenient to implement the ML algorithm to the data with uniform time grid which is unfortunately not the case with supernovae. Commonly used technique to transform unevenly distributed data onto uniform grid is to fit them with Gaussian processes (GP). Usually, each light curve is fitted by GP independently. However, in this study we use the MULTIVARIATE GAUSSIAN PROCESS<sup>3</sup> interpolation (RBF kernel). For each object it allows to correlate its multicolour LCs and approximates the data by GP in all filters in a one global fit (for details see Kornilov et. 2019,

<sup>3</sup><https://github.com/matwey/gp-multistate-kernel>



**Figure 2:** Light curves of SN2013ej. Points are the observations in *gri* filters [23, 24, 25]. Solid lines are approximation of the results of the transformation from *BRI* to *gri* filters with use of Lupton’s (2005) transformation equations.

in prep.). With this technique we can reconstruct the missing parts of LC basing on LC behaviour in other filters. For example, in Fig. 4 the maximum in *g* filter is reproduced from the *r, i* light curves. This correlation does not assume any physical assumptions about LC shape.

When the fit by MULTIVARIATE GAUSSIAN PROCESS was done, we checked the results of approximation by eye. Those SNe with unsatisfactory fit were removed from the further consideration (mainly the objects with bad photometrical quality). We also extrapolated the fit to have a bigger temporal coverage. In the end we got a sample that consists on 1999 objects.

Based on the results of approximation we extracted photometry (in flux) in the range of  $[-20, 100]$  days with 1-day bin relative to the LC maximum in *r* filter and the kernel parameters. These values were used as features for our ML algorithm (see Sect. 3).

### 3. Anomaly detection

After the approximation procedure, each object has 373 features:  $121 \times 3$  fluxes in three bands, 9 fitted parameters of Gaussian Process kernel, and logarithm of likelihood of the fit. We examine two cases of outliers search: with all features and with smaller number of features obtained by dimensionality reduction.

#### 3.1 Dimensionality Reduction

Each object has its own flux scale due to the different origin and different distance. So, before the dimensionality reduction procedure we normalized each vector of 363 photometrical points by its maximum value and used the maximum value as one more feature. Then, we applied t-SNE [26], a variation of stochastic neighbor embedding method [27], for dimensionality reduction of the data with 374 features. Finally, we obtained 7 feature reduced data sets: from 2 to 8 features.

#### 3.2 Methods: Isolation Forest

Isolation forest is an ensemble of random isolation trees. Each isolation tree is a space partitioning tree similar to a widely-know Kd-tree. However, in contrast to Kd-tree, space coordinate (a

Name	Coordinates	Object type	Ref.
SN2016bfn	13 34 45.49 +13 51 14.3	SN Ia-91T	[28]
SN2013cv	16 22 43.16 +18 57 35.6	SN Ia-pec	[29, 30]
SN1000+0216	10 00 05.87 +02 16 23.6	SLSN	[31]
SN2006kg	01 04 16.98 +00 46 08.9	AGN	[32, 33, 34]
Gaia16aye	19 40 01.13 +30 07 53.4	Binary microlensing event	[35, 36]

**Table 1:** List of found anomalies.

feature) and a split value are selected at random for every node of the isolation tree. This algorithm leads to an unbalanced tree unusable for spatial search, but the tree has the following important property. A path distance between the root and a leaf is shorter on average for points distanced in space from "normal" data. This allows us to construct enough random trees to estimate average root-leaf path distance for every data sample that we have, and then rank the data samples based on the path length.

We run the Isolation Forest algorithm on each data set (see Sect. 3.1) and obtained a list of anomalies.

## 4. Results

We visually inspected  $\sim 100$  outliers among a total 1999 objects. Using the publicly available sources we checked what kind of astrophysical objects they are. The most prominent outliers are listed in Table 1 and described below, the rest are still being studied.

### 4.1 Peculiar SNe Ia

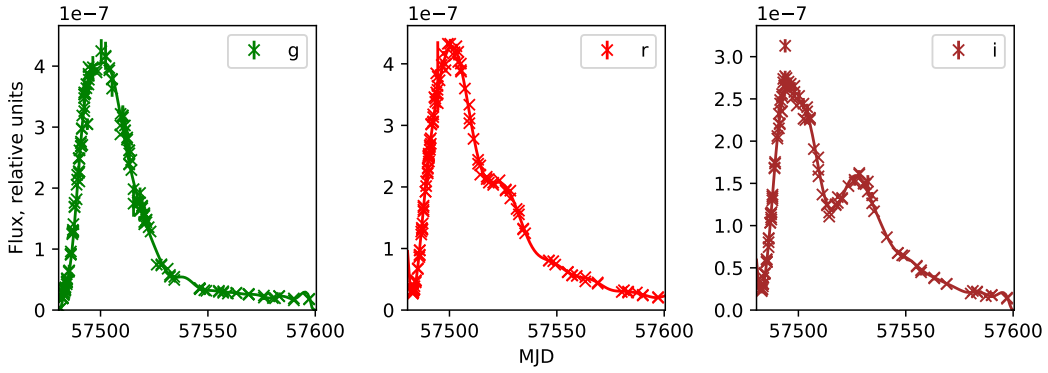
Type Ia supernova phenomenon is an explosion of a carbon-oxygen white dwarf that exceed the Chandrasekhar limit either by matter accretion from a companion star or by merging with another white dwarf [37, 38, 39]. SNe Ia are used as universal distance ladder since their luminosity at maximum light is approximately the same. However, SNe Ia can be divided by subtypes and not all of them are suitable for cosmology.

SN2016bfn [28], classified by our code as anomaly, belongs to the so-called 1991T-like-supernovae subtype (see Fig. 3). SNe Ia-91T are characterized by higher peak luminosity and broader LCs than "normal" SN Ia, and different early spectrum evolution.

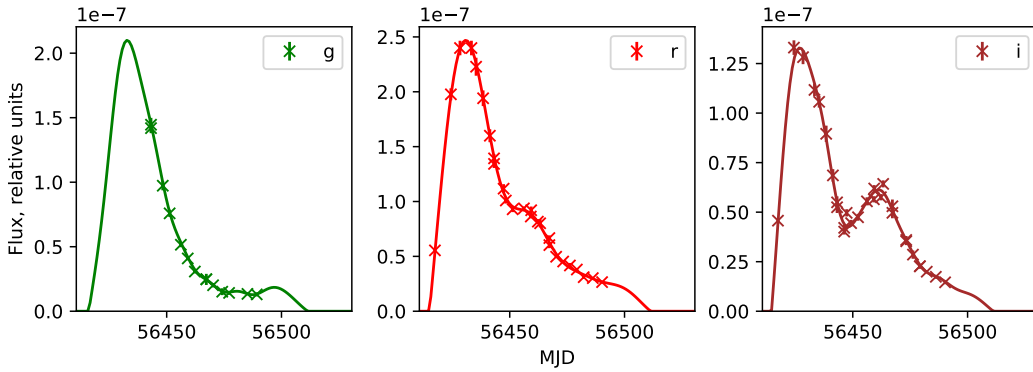
Another novelty is SN2013cv ([29], see Fig. 4). This peculiar supernova has large peak optical and UV luminosity and show an absence of iron absorption lines in the early spectra. [30] suggests that SN2013cv is an intermediate case between the normal and super-Chandrasekhar events.

### 4.2 Superluminous SNe

Superluminous SNe (SLSN) are supernovae with an absolute peak magnitude  $M < -21$  mag in any band. According to [41] SLSN can be divided into three broad classes: SLSN-I without hydrogen in their spectra, hydrogen-rich SLSN-II that often show signs of interaction with circum-stellar material (CSM), and finally, SLSN-R, a rare class of hydrogen-poor events with slowly evolving LCs, powered by the radioactive decay of  $^{56}\text{Ni}$ .



**Figure 3:** Light curves in *gri* filters of SN Ia-91T 2016bln [40]. Solid lines are the results of our approximation by MULTIVARIATE GAUSSIAN PROCESS.



**Figure 4:** Light curves in *gri* filters of peculiar SN2013cv [30, 33]. Solid lines are the results of our approximation by MULTIVARIATE GAUSSIAN PROCESS.

SN 1000+0216 (Fig. 5) was discovered in the framework of the Canada-France-Hawaii Telescope Legacy Survey Deep Fields and has a redshift  $z = 3.9$ . It may be an example of a pulsational pair-instability SN or a SLSN-II which extreme optical emission is explained by the strong interaction between the expanding ejecta and massive CSM [31].

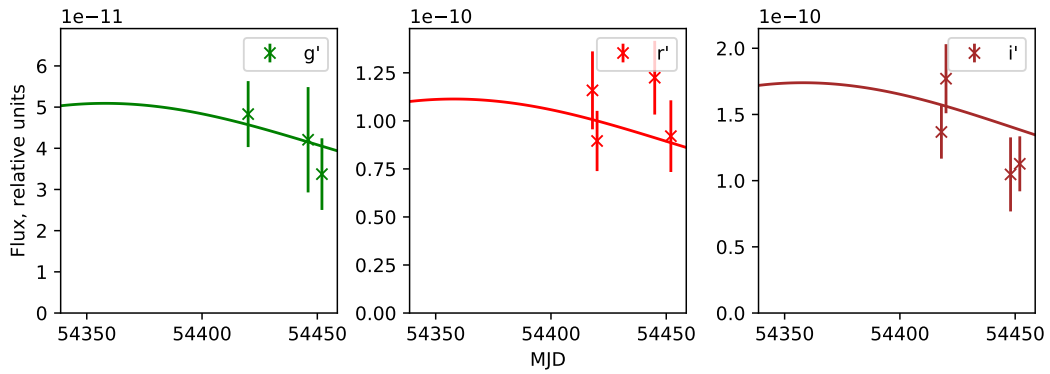
#### 4.3 AGN

SN2006kg was erroneously classified as Type II supernovae ([32], see Fig. 6). The following studies identified it as an active galactic nucleus (AGN, [33, 34]).

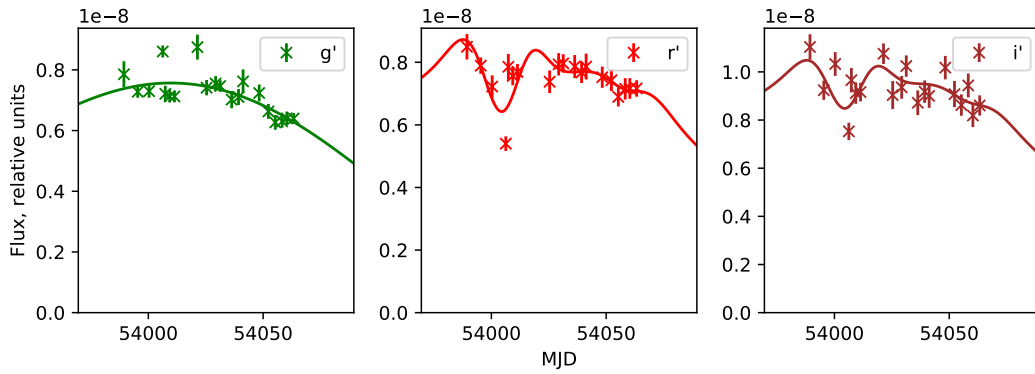
#### 4.4 Binary microlensing event

Gaia16aye [35] is an object with the most non-SN behavior in our set of outliers (Fig. 7). In [36] it was reported that Gaia16aye is a binary microlensing event — gravitational microlensing by binary systems — the first ever discovered towards the Galactic Plane.

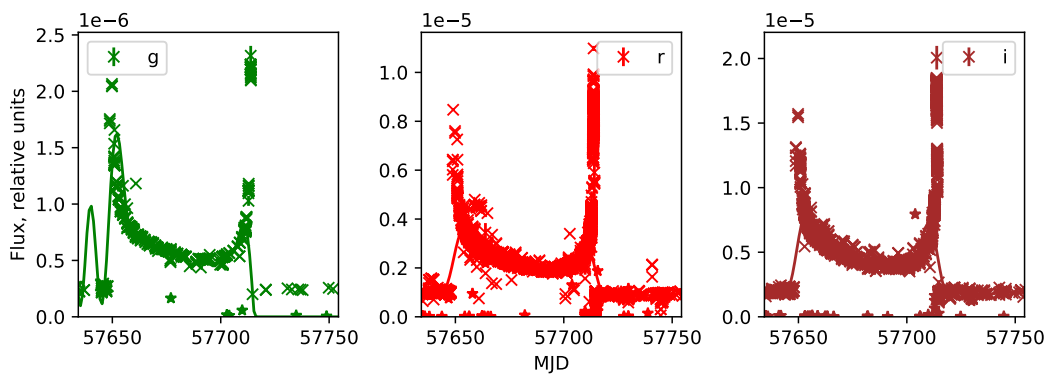




**Figure 5:** Light curves in *gri* filters of superluminous SN1000+0216 [31]. Solid lines are the results of our approximation by MULTIVARIATE GAUSSIAN PROCESS.



**Figure 6:** Light curves in *gri* filters of SN2006kg [34]. Solid lines are the results of our approximation by MULTIVARIATE GAUSSIAN PROCESS.



**Figure 7:** Light curves in *gri* filters of binary microlensing event Gaia16aye (<http://gsaweb.ast.cam.ac.uk/alerts/alert/Gaia16aye/followup>). Solid lines are the results of our approximation by MULTIVARIATE GAUSSIAN PROCESS.

## 5. Conclusions

The development of large synoptic sky surveys has led to a discovery of huge number of supernovae and supernova candidates. Among the SN discovered every year, only 10% have spectroscopical confirmation. The amount of astronomical data increases dramatically with time and already beyond human capabilities. While now community has dozens of thousands SN candidates, during ten-year survey Large Synoptic Sky Telescope (LSST, [42]) will discover over ten million supernovae (and only a small fraction of them will receive a spectroscopical confirmation). The LSST cadence will allow to receive the light curves for  $\sim 10^5$  SNe, but before these SNe will be used in any physical analysis, they must be classified by types. In order to process this information and to extract all possible knowledge, machine learning techniques become necessary. Such approach will allow not only to classify supernova candidates by known types, but to reveal other variable objects (novae, counterparts of GW alerts, kilonovae, GRB afterglows) that were mistakenly classified as SN and what is even more important to detect astronomical objects with strange physical properties — anomalies. Finding such objects (and then studying them more closely) is of high priority and one of the main aims of the current study.

We used the Isolation Forest algorithm to search the anomalies in the Open Supernova Catalog. During the data pre-processing we fitted the supernova LCs in three (*gri*) filters by Gaussian processes. The GP-MULTISTATE-KERNEL (Kornilov et al. 2019, in prep.) was specially developed to introduce the correlation between the filters. As a result, we found  $\sim 100$  anomalies, among which peculiar Type Ia SNe, SLSN, AGN, binary microlensing event.

## Acknowledgements

M. Pruzhinskaya and M. Kornilov are supported by RFBR grant according to the research project 18-32-00426\18 for anomaly analysis and LCs interpolation. K. Malanchev is supported by RBFR grant 18-32-00553 for preparing Open Supernova Catalog data. E. E. O. Ishida acknowledges support from CNRS 2017 MOMENTUM grant and Foundation for the advancement of theoretical physics and Mathematics "BASIS". A. Volnova acknowledges support from RSF grant 18-12-00522 for analysis of interpolated LCs. We used the equipment funded by the Lomonosov Moscow State University Program of Development. The authors (K. Malanchev and M. Pruzhinskaya) acknowledge the support from the Program of development of M.V. Lomonosov Moscow State University (Leading Scientific School "Physics of stars, relativistic objects and galaxies").

This research has made use of NASA's Astrophysics Data System Bibliographic Services. SCIKIT-LEARN machine learning software was used [43].

## References

- [1] A. Pastorello, X.-F. Wang, F. Ciabattari, D. Bersier, P. A. Mazzali, X. Gao et al., *Massive stars exploding in a He-rich circumstellar medium - IX. SN 2014av, and characterization of Type Ibn SNe*, *MNRAS* **456** (Feb., 2016) 853–869, [1509.09069].
- [2] N. M. Ball and R. J. Brunner, *Data Mining and Machine Learning in Astronomy*, *International Journal of Modern Physics D* **19** (2010) 1049–1106, [0906.2173].

- [3] S. D. McCauliff, J. M. Jenkins, J. Catanzarite, C. J. Burke, J. L. Coughlin, J. D. Twicken et al., *Automatic Classification of Kepler Planetary Transit Candidates*, *ApJ* **806** (June, 2015) 6, [1408.1496].
- [4] S. E. Thompson, F. Mullally, J. Coughlin, J. L. Christiansen, C. E. Henze, M. R. Haas et al., *A Machine Learning Technique to Identify Transit Shaped Signals*, *ApJ* **812** (Oct., 2015) 46, [1509.00041].
- [5] K. A. Pearson, L. Palafox and C. A. Griffith, *Searching for exoplanets using artificial intelligence*, *MNRAS* **474** (Feb., 2018) 478–491, [1706.04319].
- [6] M. W. Libbrecht and W. S. Noble, *Machine learning applications in genetics and genomics*, *Nature Reviews Genetics* **16** (2015) 321–332.
- [7] C. Criscia, B. Ghattasb and G. Pererac, *A review of supervised machine learning algorithms and their applications to ecological data*, *Ecological Modelling* **240** (2012) 113–122.
- [8] M. Vidyasagar, *Identifying predictive features in drug response using machine learning: Opportunities and challenges*, *Annual Review of Pharmacology and Toxicology* **55** (2015) 15–34, [<https://doi.org/10.1146/annurev-pharmtox-010814-124502>].
- [9] R. Kessler, B. Bassett, P. Belov, V. Bhatnagar, H. Campbell, A. Conley et al., *Results from the Supernova Photometric Classification Challenge*, *PASP* **122** (Dec., 2010) 1415, [1008.1024].
- [10] A. A. Mahabal, S. G. Djorgovski, A. J. Drake, C. Donalek, M. J. Graham, R. D. Williams et al., *Discovery, classification, and scientific exploration of transient events from the Catalina Real-time Transient Survey*, *Bulletin of the Astronomical Society of India* **39** (Sept., 2011) 387–408, [1111.0313].
- [11] E. E. O. Ishida and R. S. de Souza, *Kernel PCA for Type Ia supernovae photometric classification*, *MNRAS* **430** (Mar., 2013) 509–532, [1201.6676].
- [12] M. Lochner, J. D. McEwen, H. V. Peiris, O. Lahav and M. K. Winter, *Photometric Supernova Classification with Machine Learning*, *ApJS* **225** (Aug., 2016) 31, [1603.00882].
- [13] A. Mahabal, K. Sheth, F. Gieseke, A. Pai, S. G. Djorgovski, A. Drake et al., *Deep-Learnt Classification of Light Curves*, *arXiv e-prints* (Sept., 2017) , [1709.06257].
- [14] M. Graham, A. Drake, S. G. Djorgovski, A. Mahabal and C. Donalek, *Challenges in the automated classification of variable stars in large databases*, in *European Physical Journal Web of Conferences*, vol. 152 of *European Physical Journal Web of Conferences*, p. 03001, Sept., 2017, DOI.
- [15] H. Hildebrandt, S. Arnouts, P. Capak, L. A. Moustakas, C. Wolf, F. B. Abdalla et al., *PHAT: PHoto-z Accuracy Testing*, *A&A* **523** (Nov., 2010) A31, [1008.0658].
- [16] S. Cavuoti, M. Brescia, C. Tortora, G. Longo, N. R. Napolitano, M. Radovich et al., *Machine-learning-based photometric redshifts for galaxies of the ESO Kilo-Degree Survey data release 2*, *MNRAS* **452** (Sept., 2015) 3100–3105, [1507.00754].
- [17] U. Rebbapragada, P. Protopapas, C. E. Brodley and C. Alcock, *Finding anomalous periodic time series*, *Machine Learning* **74** (Mar, 2009) 281–313.
- [18] M. Henrion, D. J. Hand, A. Gandy and D. J. Mortlock, *Casos: a subspace method for anomaly detection in high dimensional astronomical databases*, *Statistical Analysis and Data Mining: The ASA Data Science Journal* **6** 53–72, [<https://onlinelibrary.wiley.com/doi/pdf/10.1002/sam.11167>].

- [19] D. Baron and D. Poznanski, *The weirdest SDSS galaxies: results from an outlier detection algorithm*, *MNRAS* **465** (Mar., 2017) 4530–4555, [1611.07526].
- [20] I. Nun, K. Pichara, P. Protopapas and D.-W. Kim, *Supervised Detection of Anomalous Light Curves in Massive Astronomical Catalogs*, *ApJ* **793** (Sept., 2014) 23, [1404.4888].
- [21] J. Guillochon, J. Parrent, L. Z. Kelley and R. Margutti, *An Open Catalog for Supernova Data*, *ApJ* **835** (Jan., 2017) 64, [1605.01054].
- [22] F. T. Liu, K. M. Ting and Z.-H. Zhou, *Isolation-based anomaly detection*, *ACM Trans. Knowl. Discov. Data* **6** (Mar., 2012) 3:1–3:39.
- [23] P. J. Brown, A. A. Breeveld, S. Holland, P. Kuin and T. Pritchard, *SOUSA: the Swift Optical/Ultraviolet Supernova Archive*, *Astrophysics and Space Science* **354** (Nov., 2014) 89–96, [1407.3808].
- [24] F. Huang, X. Wang, J. Zhang, P. J. Brown, L. Zampieri, M. L. Pumo et al., *SN 2013ej in M74: A Luminous and Fast-declining Type II-P Supernova*, *ApJ* **807** (July, 2015) 59, [1504.00446].
- [25] F. Yuan, A. Jerkstrand, S. Valenti, J. Sollerman, I. R. Seitenzahl, A. Pastorello et al., *450 d of Type II SN 2013ej in optical and near-infrared*, *MNRAS* **461** (Sept., 2016) 2003–2018, [1605.06117].
- [26] L. v. d. Maaten and G. Hinton, *Visualizing data using t-sne*, *Journal of machine learning research* **9** (2008) 2579–2605.
- [27] G. E. Hinton and S. T. Roweis, *Stochastic neighbor embedding*, in *Advances in neural information processing systems*, pp. 857–864, 2003.
- [28] S. B. Cenko, Y. Cao, M. Kasliwal, A. A. Miller, C. Fremling, M. West et al., *DCT and Gemini Spectroscopic Classification of AT 2016bln (=iPTF 16abc)*, *The Astronomer’s Telegram* **8909** (Apr., 2016) .
- [29] L. Zhou, X. Wang, K. Zhang, J. Chen, J. Liang, T. Zhang et al., *Supernova 2013cv = Psn J16224316+1857356*, *Central Bureau Electronic Telegrams* **3543** (June, 2013) .
- [30] Y. Cao, J. Johansson, P. E. Nugent, A. Goobar, J. Nordin, S. R. Kulkarni et al., *Absence of fast-moving iron in an intermediate type ia supernova between normal and super-chandrasekhar*, *The Astrophysical Journal* **823** (2016) 147.
- [31] J. Cooke, M. Sullivan, A. Gal-Yam, E. J. Barton, R. G. Carlberg, E. V. Ryan-Weber et al., *Superluminous supernovae at redshifts of 2.05 and 3.90*, *Nature* **491** (Nov., 2012) 228–231, [1211.2003].
- [32] B. Bassett, A. Becker, H. Brewington, C. Choi, D. Cinabro, F. Dejongh et al., *Supernovae 2006kg-2006lc*, *Central Bureau Electronic Telegrams* **688** (Oct., 2006) .
- [33] O. Yaron and A. Gal-Yam, *WISeREP — An Interactive Supernova Data Repository*, *PASP* **124** (July, 2012) 668, [1204.1891].
- [34] M. Sako, B. Bassett, A. C. Becker, P. J. Brown, H. Campbell, R. Wolf et al., *The Data Release of the Sloan Digital Sky Survey-II Supernova Survey*, *PASP* **130** (June, 2018) 064002, [1401.3317].
- [35] V. Bakis, U. Burgaz, T. Butterley, J. M. Carrasco, V. S. Dhillon, M. Dominik et al., *Gaia16aye: a flaring object of uncertain nature in Cygnus*, *The Astronomer’s Telegram* **9376** (Aug., 2016) .
- [36] L. Wyrzykowski, G. Leto, G. Altavilla, V. Bakis, N. Britavskiy, U. Burgaz et al., *Gaia16aye is a binary microlensing event and is crossing the caustic again*, *The Astronomer’s Telegram* **9507** (Sept., 2016) .

- [37] J. Whelan and I. Iben, Jr., *Binaries and Supernovae of Type I*, *ApJ* **186** (Dec., 1973) 1007–1014.
- [38] I. Iben, Jr. and A. V. Tutukov, *Supernovae of type I as end products of the evolution of binaries with components of moderate initial mass ( $M$  not greater than about 9 solar masses)*, *ApJS* **54** (Feb., 1984) 335–372.
- [39] R. F. Webbink, *Double white dwarfs as progenitors of R Coronae Borealis stars and Type I supernovae*, *ApJ* **277** (Feb., 1984) 355–360.
- [40] A. A. Miller, Y. Cao, A. L. Piro, N. Blagorodnova, B. D. Bue, S. B. Cenko et al., *Early Observations of the Type Ia Supernova iPTF 16abc: A Case of Interaction with Nearby, Unbound Material and/or Strong Ejecta Mixing*, *ApJ* **852** (Jan., 2018) 100, [1708.07124].
- [41] A. Gal-Yam, *Luminous Supernovae*, *Science* **337** (Aug., 2012) 927, [1208.3217].
- [42] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel et al., *LSST Science Book, Version 2.0*, *ArXiv e-prints* (Dec., 2009), [0912.0201].
- [43] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel et al., *Scikit-learn: Machine learning in Python*, *Journal of Machine Learning Research* **12** (2011) 2825–2830.