PHOTOMETRIC ANALYSIS OF PI OF THE SKY DATA

Rafał Opiela[a,*], Katarzyna Małek[a], Lech Mankiewicz[a],
Małgorzata Siudek[a], Marcin Sokołowski[b], Aleksander Filip Żarnecki[c]

[a] *Center for Theoretical Physics Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw.*

[b] *Sołtan Institute for Nuclear Studies, Hoża 69, 00-681 Warsaw, Poland.*

[c] *Faculty of Physics, University of Warsaw, Hoża 69, 00-681 Warsaw, Poland.*

[*] corresponding author: `opiela@cft.edu.pl`

Abstract. Two fully automatic Pi of the Sky detectors with a large field of view, located in Spain (INTA) and in Chile (SPDA) observe the sky in search of rare optical phenomena, and also collect observations which include many kinds of variable stars. To be able to draw proper conclusions from the data that is received, adequate quality of the detectors is very important. Pi of the Sky data are subject to systematic errors caused by various factors, e.g. cloud cover seen as significant fluctuations in the number of stars observed by the detector, problems with conducting mounting, a strong background of the moon or the passage of a bright object, e.g. a planet, near the observed star. Some of these adverse effects are already detected during cataloging of the individual measurements, but this is not sufficient to make the quality of the data satisfactory for us. In order to improve the quality of our data, we developed two new procedures based on two different approaches. In this paper we will say some words about these procedures, give some examples, and show how these procedures improve the quality of our data.

Keywords: Gamma Ray Burst (GRB), variable stars, robotic telescopes, photometry, astrometry, data quality, photometric corrections.

1. Introduction

The Pi of the Sky experiment has been designed for continuous observations of a large part of the sky, in search of astrophysical phenomena varying in scale from seconds to months, especially for prompt optical counterparts of Gamma Ray Bursts (GRBs). Other scientific goals include searching for nova stars and flare stars. The large amount of data obtained in the project also allows the identification and cataloging of many different types of variable stars. The project involves leading Polish academic and research units such as the Andrzej Sołtan Institute for Nuclear Studies, the Center for Theoretical Physics (Polish Academy of Science), the Institute of Experimental Physics (University of Warsaw), and many others.

A new detector belonging to the Pi of the Sky project was installed at the end of 2010 in the INTA El Arenosillo Observatory in south-western Spain. Unlike the prototype detector, which has been working since 2004 in Chile, it is equipped with four CCD cameras, and it has a much larger field of view of 40° × 40°, which significantly increases the flow of data collected during the night. By the end of the year, two more detectors will be installed in Spain, increasing the field of view of the whole system up to 4800 square degrees. This will dramatically increase the amount of data that needs to be analysed. The collected data are subject to systematic errors arising from the conditions under which the photograph was taken and from the design of the detector and the CCD
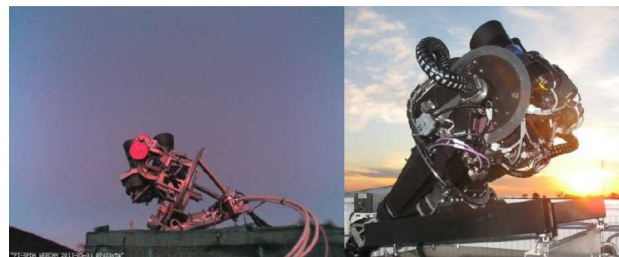

Figure 1. Two currently working Pi of the Sky detectors. The prototype detector (on the left) and the new detector (on the right).

chip. As has was already been mentioned, some of these adverse effects have been detected and properly marked, but this did not make the quality of our data satisfactory for us. We needed to take into account some other effects, the most important of which seems to be the dependence of the sensitivity of the detector on the spectral type of the observed object. We also looked for new, completely independent methods that can help us improve the quality of our data [1, 2].

2. Data processing

2.1. On-line data reduction

Data analysis in the Pi of the Sky experiment consist of two parts. The first part is called on-line analysis, and the second part is called off-line analysis. On-line data analysis is required to control the performance of the detector, and it is also responsible for finding

optical flashes in time scales from 10 to 22 seconds in real time. Fast algorithms optimised for transient search are included in the on-line data analysis. Frame-by-frame real time analysis provides an opportunity to distribute alerts in the community for follow-up observations.

On-line data reduction consists of several parts:

- Dark image subtraction. Dark frame subtraction is a way to minimize image noise for pictures taken with long exposure times. It takes advantage of the fact that a component of the image noise, known as fixed-pattern noise, is the same from shot to shot: noise from the sensor, dead or hot pixels. A dark frame is an image captured with the sensor in the dark, essentially just an image of noise in an image sensor. It works by taking a picture with the shutter closed.

- Image transformation, using a special method called Laplace Transformation. The value of each pixel is calculated as a simple function of several surrounding pixels. Values in the pixels just around transformed pixels are assumed, and values in other pixels far away from it are subtracted. The idea of this transformation is to calculate the simple aperture brightness for each pixel. In this step, a simple and fast aperture photometry algorithm is used to calculate star brightness. This photometry is performed every 300 seconds and it covers all night images.

- Comparison with a reference image (a series of previous images). Two images are compared in order to find differences in brightenes between existing objects, or to find entirely new objects.

2.2. Off-line data reduction

Off-line data analysis acts on the reduced data and catalogs it to the database. It consists of algorithms optimised for data reduction. The reduction pipeline is divided to three main stages: photometry, astrometry and cataloging to the database. The final reduced data give us information about star brightness measurements which is stored in the our database, which provides easy and effective public access. The off-line analysis consists of several algorithms developed for different purposes [3, 4].

Every image collected during one night is processed in the same way. Algorithms optimised for off-line data reduction have the following steps:

- Adding 20 subsequent frames (which takes $\sim 20 \cdot 10 = 200$ seconds exposure). The image coordinates are controlled and, if they change, the average chain is stopped in order not to allow for averaging images from different positions. In the case of single image reduction, the image averaging step is skipped.

- Dark frame subtraction and flat corrections. In order to reduce fluctuations, the dark image is calculated as the median of several dark images. This

step allows the signal offset produced by dark current and electronics to be subtracted. It also reduces the effect of hot pixels. Flat correction allows corrections to be made for non-uniformity of the optics and differences between pixel amplifications. The standard way of finding this correction is by taking images of a uniformly illuminated field. This is usually the sky just after dusk or just before dawn, when the sky is bright and stars are not visible. An alternative way is to use a uniformly illuminated screen. In the case of the Pi of the Sky prototype, due to the large field of view the detector flat image is obtained by taking images of the evening sky with the mount tracking switched off. After taking many images and calculating the median image, stars are eliminated. Finally, the image is normalized to one.

- Multiple aperture photometry. Photometry is a procedure which finds stars in the image and determines their chip coordinates (x, y) and their brightness. In the Pi of the Sky experiment, aperture photometry is adopted from the ASAS experiment. It is rather slow, so it cannot be used for reducing each individual image from a night. This consideration led to the development of our own fast photometry algorithm, as discussed above. This kind of photometry is used in a photometry experiment involving 20 averaged images, and in reducing scan images (3 images averaged). Both photometry procedures write the resulting list of stars with (x, y) coordinates and magnitudes to output mag files. The mag files are then available as input files for astrometry procedures.

- Astrometry and catalog star selection. This involves transforming chip coordinates (x, y) to celestial coordinates (λ, δ). It is an iterative minimization procedure for comparing stars.

- Identification in the image by photometry with catalog stars from external star catalogs. The astrometry procedure was adopted from the ASAS experiment. The star catalog currently used in the procedure is based on the TYCHO catalog, but it could be replaced by any star catalog. All night images are processed in the same way.

- Normalization to V magnitudes from the TYCHO catalog. The star magnitudes are normalized by comparing with the catalog of catalog stars. For each star, the corresponding star in the reference catalog is found. The correction for each star is calculated as the difference between the magnitudo of the catalog star and the magnitudo measured in the experiment. This calculated value must be added to the instrumental magnitude to obtain the normalized value.

- Cataloging of lightcurves to the PostgreSQL[1] database (see Section 3).

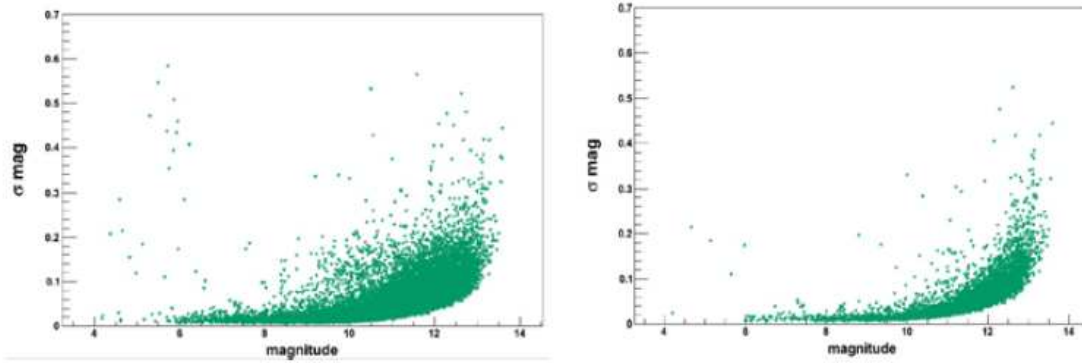[1] More information about PostgreSQL database are avalible at http://www.postgresql.org/.

FIGURE 2. Precision dispersion of star brightness measurements from standard photometry for 200 s exposures (20 co-added frames) from the Pi of the Sky prototype at the Las Campanas Observatory in Chile. The large dispersion (left) is mainly caused by false measurements. After the application of quality cuts (right), the photometry accuracy improves significantly.

2.3. MULTILEVEL SELECTION SYSTEM

Despite using the methods described above, we still did not obtain satisfactory data quality. Data quality varies with brightness, as shown in Fig. 2. The mean measured brightness dispersion in the range of magnitudo between 6^m–9^m is about 0.03^m. In other areas of magnitudo, the situation is even worse. The dispersion increases to 0.1–0.2 magnitudo. This is due to the sensitivity of the detector, which is at its most sensitive precisely in the range of magnitudo from 6 to 9. This is quite a lot if we want to detect variable stars, as one of the most popular methods for finding variable stars is to select stars with the largest variance. However, such an approach is impossible in the presence of large errors, because most of the stars with large variance measurements are simply stars with incorrect measurements. Better elimination of bad measurements will in the future lead to better identification of variable stars.

The value of brightness measurements may be distorted by the influence of various factors associated with the measurements. Precise determination of the causes of erroneous measurements of brightness is the key to eliminating these bugs and improving the quality of the data. Dedicated filters have been developed within our project to remove bad measurements or frames. The main causes of the errors that occur include:

- Read with an open shutter. With an open shutter, the image is irradiated all the time, and the light from bright stars forms strong streaks stretching down from the star images. If we measure the starlight from a star that has fallen into such a streak, its brightness will, of course, be many times greater than normal. Strips may therefore result in time-varying brightening of stars. This phenomenon is easy to detect, because it occurs on only one camera.

- Planet or planetoid passage. If a planet or a planetoid passes in front of stars, it will add its light to the star's light. This will cause momentary brighten-

ing of the stars. Such cases also need to be detected and removed. For this purpose, we use a specially prepared database in which the trajectories of every bright object in our Solar System are written. In a similar manner, incorrect measurements caused by brightening of stars due to the glow from Jupiter are detected and removed.

- Measurements near the CCD edge. Stars appearing at the edges of the detector have significantly distorted shapes. This is due to the fact that the sensitivity of the detector is much lower at the edges. This blurring is strong enough to change the brightness of the stars affected by it.

- Other causes include: hotpixels (discussed above), which can affect the brightness of nearby stars, frames with a very high background level caused by the strong background that occurs when observations are conducted during the full Moon, or frames with too few matched stars, mainly due to bad weather.

The accuracy of photometry improves significantly after bad quality data is removed. For stars from 7^m to 10^m, the photometry error is less than ~ 0.015 magnitudo. The dispersion is still larger for stars of greater or lower brightness, but the accuracy of the photometry is greatly improved [5, 6].

3. DATABASE

The data acquired during the Pi of the Sky observations are reduced, and only the light curves of stars are stored in the project databases. These databases contain all measurements taken by the Pi of the Sky detector. The first database covers VII.2004–VI.2005, and contains about 790 mln measurements for about 4.5 mln objects. The second database covers V.2006–XI.2007, and includes about 1002 mln measurements for about 10.8 mln objects. The third database covers V.2006–IV.2009, and includes about 2.16 billion measurements for about 16.7 mln objects. A dedicated web interface has been developed to facilitate
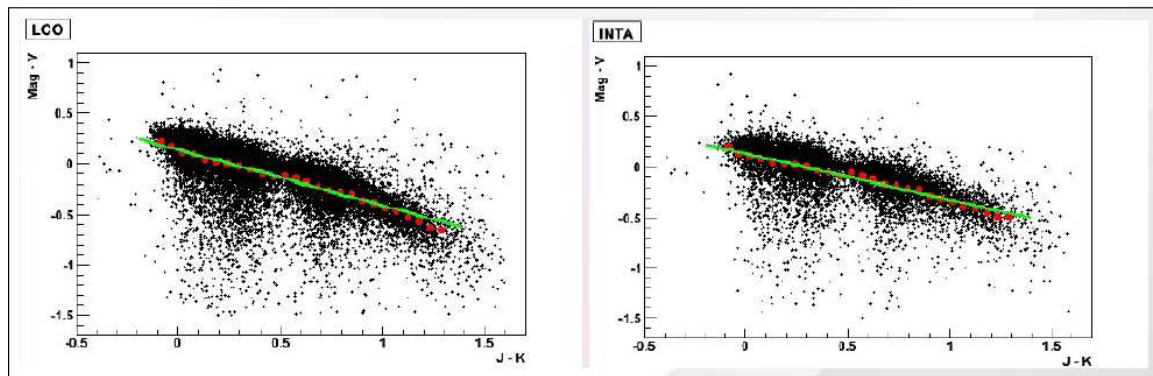
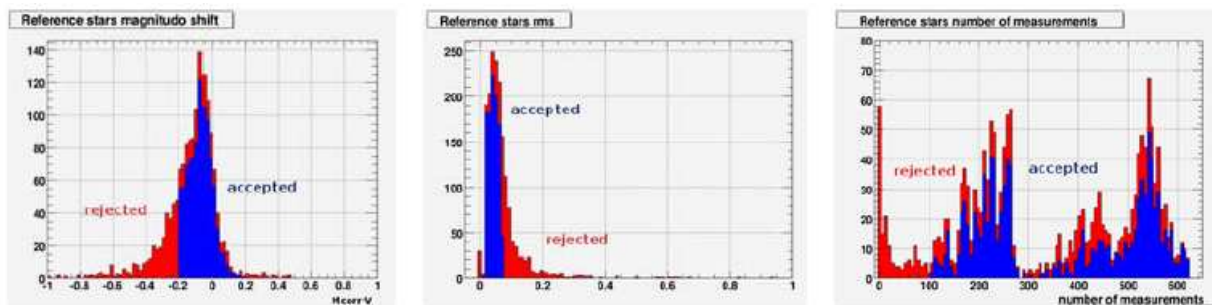FIGURE 3. The detector response is correlated with the spectral type(B-V or J-K) of catalog stars.

FIGURE 4. Only the best catalog stars were used (blue) for calculating photometry corrections, after rejecting stars with a magnitudo shift ($M_{corr} - M$) bigger than 0.2, RMS_{corr} bigger than 0.07 and with number of measurements smaller than 100.

public access to the databases of the Pi of the Sky project. The interface allows the user to search for stars by magnitude, coordinates and other parameters and to view the light curve of a selected star. Public databases are available on the Pi of the Sky web page http://grb.fuw.edu.pl/pi/databases [7].

4. METHODS FOR IMPROVING DATA QUALITY

4.1. COLOR CORRECTION

A CCD detector is not equally sensitive in all wavelengths. According to the manufacturer, the CCD detectors used in the Pi of the Sky project are at their most sensitive in the near infrared, while the average wavelength is, in their case $\langle \lambda \rangle \approx 585$ nm, which corresponds approximately to a wavelength characteristic of the V filter. The sensitivity of CCD detectors varies depending on the wavelength λ, and affects the quality of the results of measurements taken in white light. Two objects with the same luminosities, where the first object shines in near infra-red and the second object shines e.g. in blue, will have different brightnesses in our detectors. An object shining in the near infra-red will be brighter than a blue object. Since neither of the INTA cameras present in the new detector has a filter installed that would take care of this effect, we must take the effect into account while cataloging our data.

We have already determined that, in the case of

data collected by the prototype detector in Chile, the following formula may be used for approximating the correction of the standard photometry used in the Pi of the Sky project, which rests on taking into account the dependence of the sensitivity of the CCD chip on the observed star type:

$$M_{corr} = M - 0.2725 + 0.5258(J - K) \qquad (1)$$

In this formula, the J value and the K value correspond to the brightnesses of the object that is tested in the filter J and in the K filter. M is the brightness of the analysed object, measured by the detector and normalized to the brightness of catalog stars in V. M_{corr} represents the corrected magnitudo, taking into account color correction. The cameras of the prototype, however, have chips made by a different manufacturer (Fairchild) than the cameras in Spain (STA), so the colour calibration also had to be repeated for the data collected by the new detector (see Fig. 3).

Equation (1) provides information about the corrected magnitudo only for catalog stars which have J and K brighteness values. To calculate the photometry corrections for any stars visible in the resulting picture, we use a special procedure that requires the use of only the best catalog stars. We are interested in catalog stars in the magnitudo range from 6^m to 10^m, and we reject stars with a magnitudo shift ($M_{corr} - M$) bigger than 0.2. The number of measurements of catalog stars should also be more than 100, and we accept
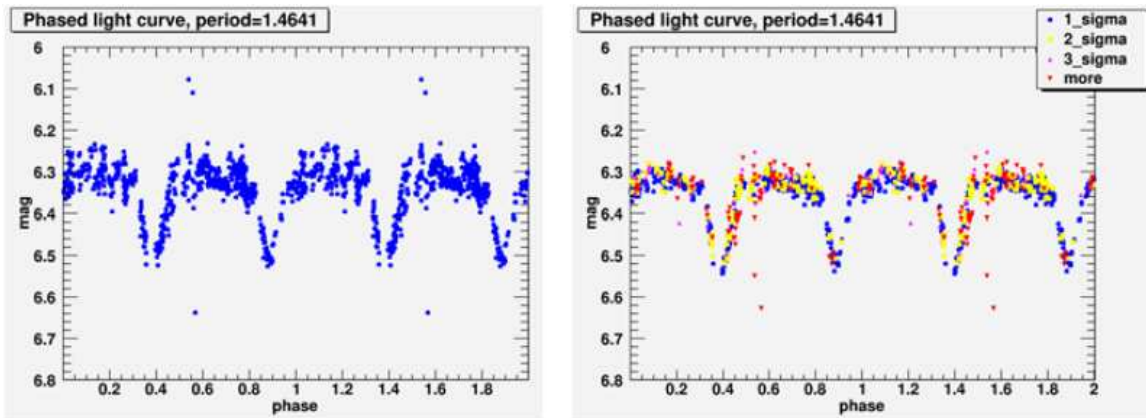
FIGURE 5. Uncorrected lightcurve for BGInd variable (left) and after spectral corrections with correction quality cut (right).
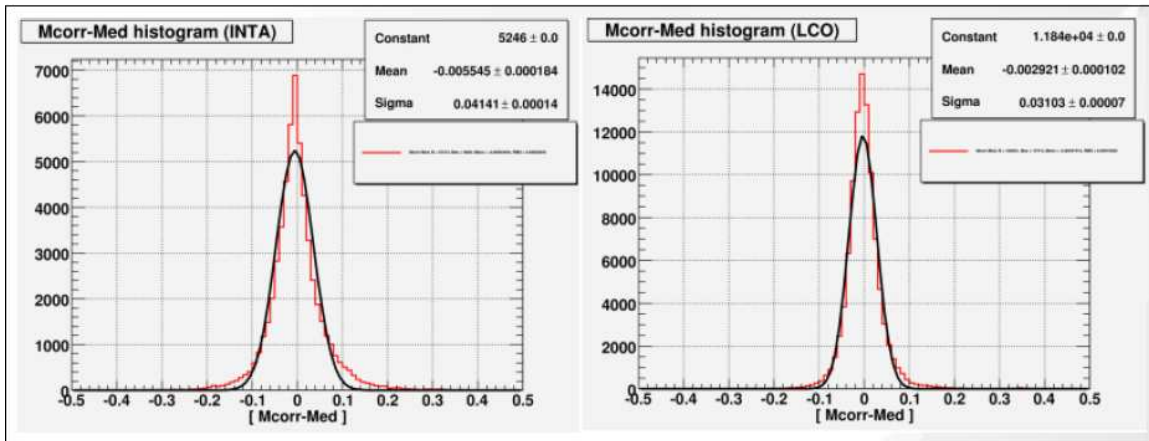

FIGURE 6. Gauss function fitted to the $|M_{\mathrm{corr}} - Med|$ histogram. Thanks to this fit, we can obtain the value of σ, which is then used to calculate the quality of the analysed frames.
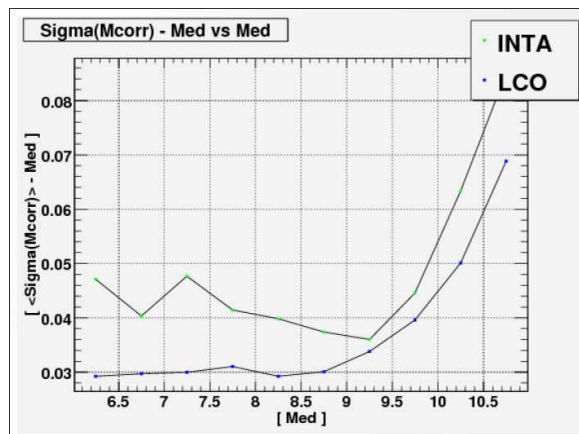

FIGURE 7. Values of σ obtained from the fitted Gauss function for $|M_{\mathrm{corr}} - Med|$ histograms.

only catalog stars which have $RMS_{M_{\mathrm{corr}}} < 0.07$, see Fig. 4.

For such selected catalog stars we calculate the quadratic surface correction, and we try to interpolate the value of the correction for the point where our analysed star exists. The average square distance of the reference stars from the fitted correction surface (χ^2) provides additional, independent information about the quality of the analysed measurement. This information can be used to select the measurements with most precise photometry.

The effect of photometry correction with a distribution of χ^2 on the reconstructed BG Ind light curve is shown in Fig. 5. In this case, applying the new algorithm improved the photometry quality, and uncertainty sigma of the order of 0.013^{m} was obtained. We also applied photometry correction to other stars, with as good results as in the case of BGInd variable. Spectral correction and additional χ^2 distribution enable selection of only the measurements with the highest precision [8, 9].

4.2. Statistical methods

Another, independent way to improve the data quality, which we also considered, is to study the statistical properties of a group of frames. In this method, based on the statistical properties of a group of frames, we calculate the quality of a single frame and, on the basis of this quality, we calculate the quality of the single measurement on this frame.

At the begining of this correction procedure we divide all analysed frames into "*good*" and "*bad*". In this case, we create $|M_{\mathrm{corr}} - Med|$ histograms based on all analysed frames. M_{corr} represents the corrected catalog star magnitudo taken from the analysed frame. The correction takes into account the dependence of the observed magnitudo on the brightness of the catalog star. *Med* is the median, which is calculated on the basis of good measurements (with quality $= 0$) of this catalog star. In the near future, we will calculate the median on the basis only of good measurements taken from the same field. For each frame, we found catalog stars in the given magnitudo range, and for these stars we calculate the values of $|M_{\mathrm{corr}} - Med|$. The results are used to plot $|M_{\mathrm{corr}} - Med|$ histograms for all ranges of magnitudo. To these histograms we fit the Gauss function, which gives us the value of σ, which is later used to obtain the quality of the analysed frames (see Fig. 6).

As was shown in Fig. 7, the smallest value of σ obtained from the Gauss function fitted to $|M_{\mathrm{corr}} - Med|$ histograms was calculated for the brightness range from 8^{m} to 8.5^{m} (for LCO data). This σ value was later used to obtain the quality of the whole frame.

In order to determine the quality of a given frame, we check how many catalog stars (which are visible in this frame) have $|M_{\mathrm{corr}} - Med| > 2\sigma$. On the basis of many tests, we assumed that "*good*" frames have

$\leq 10\%$ bad catalog stars, and "*bad*" frames have $> 10\%$ bad catalog stars (see Fig. 8).

If we know which frames are good and which are bad, we can calculate $\langle M \rangle$ and $\sigma \langle M \rangle$ values based on each group of frames. We can calculate these values on the basis of measurements taken only from good frames, from bad frames or from all frames. We also take into account the quality of the data calculated using the previous methods. The results are given in Fig. 9.

5. Conclusions

Much information about the Pi of the Sky project can be found on the project webpage, which is available on `http://grb.fuw.edu.pl`. We have created a system of dedicated filters to mark bad measurements and frames. This system is applied with a cataloging procedure for new data. To improve the quality of the data, we have created an approximate color calibration algorithm based on the spectral type of catalog stars. We have also developed another statistical method, which analyses all stars in the frame and rejects bad-quality exposures. After the new frame selection is applied, photometry accuracy of 0.01^{m}–0.03^{m} can be obtained. Further improvement can be made in dedicated analysis of selected objects[10].

References

[1] Sokołowski, M.: Investigation of astrophysical phenomena in short time scales with Pi of the Sky, July, 2008, Warsaw.

[2] Majczyna, A.: Search for optical flashes of astronomical origin with Pi of the Sky prototype, *The 2009 Europhysics Conference on High Energy Physics*, July, 2009.

[3] Małek, K.: General overview of the Pi of the Sky system, *Proceedings of SPIE Volume: 7502*, May, 2009.

[4] Majczyna, A.: Pi of the Sky catalogue of the variable stars from 2006–2007 data, *Proceedings of SPIE Volume: 7745*, May, 2010.

[5] Sokołowski, M.: Detection of short optical transient of astrophysical origin in real time, *Proceedings of SPIE Volume: 7502*, May, 2009.

[6] Małek, K.: Pi of the Sky detector, *Advances in Astronomy, Vol. 2010, Article ID 194946*, June, 2009.

[7] Sokołowski, M.: Automated detection of short optical transients of astrophysical origin in real time, *Advances in Astronomy, Vol. 2010, Article ID 463496*, June, 2009.
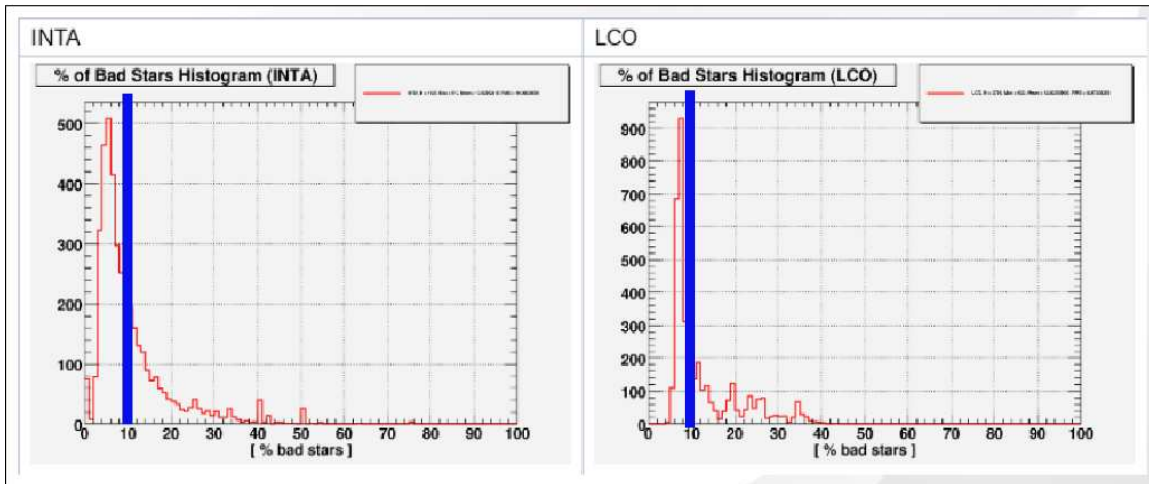
FIGURE 8. Histograms of the percentage of bad catalog stars in a single frame. These histograms were created for a given range of frames. For each frame, we calculated what percentage of the catalog stars from the analysed frame have $|M_{\mathrm{corr}} - Med| > 2\sigma$. We assumed that bad frames have more than 10 % bad catalog stars, and good frames have fewer than 10 % bad catalog stars.
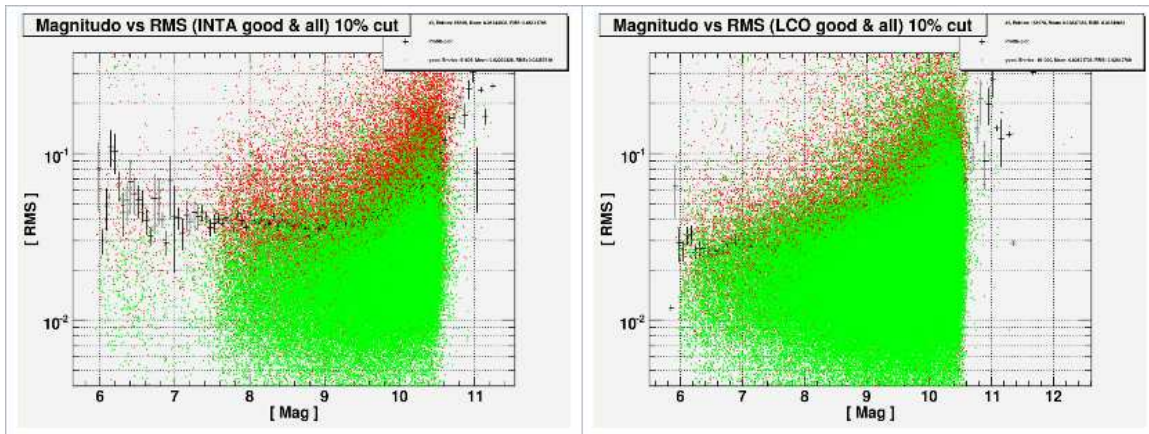

FIGURE 9. $\sigma\langle M \rangle$ vs $\langle M \rangle$ plot. On the red we can see points with positions corresponding to $\langle M \rangle$ and $\sigma\langle M \rangle$ calculated on the basis of all measurements. On the green, we can see points with positions corresponding to $\langle M \rangle$ and $\sigma\langle M \rangle$ calculated on the basis of good measurements only.

[8] Siudek, M.: Photometric analysis of the Pi of the Sky data, *Acta Polytechnica 2011/6 – Proceedings IBWS 2011*, June, 2011.

[9] Opiela, R.: Improving photometry of the Pi of the Sky, *Proceedings of SPIE Volume: 7745*, June, 2010.

[10] Żarnecki, F.: Improving the Photometry of the Pi of the Sky System, *Acta Polytechnica 2011/2 – Proceedings IBWS 2010*, June, 2010.