August 28, 2014

A pilot study of the Pan-STARRS potential for finding RR Lyrae

The MPIA Summer 2014 RR Lyrae Tiger Team

ABSTRACT

We used variable Pan-STARRS sources from the SDSS Stripe 82 region, together with WISE photometry, to select a sample of candidate RR Lyrae. Aided by a sample of known RR Lyrae with SDSS-based template light curves, we used template-fitting method and estimated its sample completeness and purity. Our results motivate the application of this method to the whole sky region covered by Pan-STARRS.

1. Introduction

Pan-STARRS data have the potential to extend the faint limit for finding RR Lyrae stars by about a magnitude, compared to the deepest available datasets. For motivation why to search for RR Lyrae, relevant references, and an attempt to find RR Lyrae in Pan-STARRS, see Abbas et al. (2014, AJ 148, 8).

Abbas et al. used SDSS colors (because they are nearly simultaneously measured) and Pan-STARRS variabiltiy estimators to select candidate RR Lyrae. They achieved completeness of 52% and purity (100% minus "contamination") of 77% (measured using a known sample of RR Lyrae from SDSS Stripe 82). Most of their contaminants are stars rather than quasars because the latter were efficiently rejected using the SDSS $u - g$ color. Therefore, this performance characteristics **cannot** be obtained over the entire Pan-STARRS footprint (about twice as large as covered by SDSS) due to the lack of SDSS u band measurements.

The only variability information that Abbas et al. used is a simple cut on the root-mean-square variability in the Pan-STARRS i_P, z_P and y_P bands, with at least three detections in each band. In particular, they did not attempt to fit RR Lyrae light curve templates to Pan-STARRS data. Light-curve template fitting can improve the selection of RR Lyrae in Pan-STARRS because of more information being used in the process. Furthermore, if the improvement is significant, it might be possible to abandon selection cuts based on SDSS data and thus extend the search to the full Pan-STARRS 3π area. Investigating this possibility was the main aim of this pilot study.

2. Pilot Study

We first discuss the selection of candidate RR Lyrae using Pan-STARRS and WISE data, and then discuss the results of applying the template fitting method to selected candidates.

2.1. Selection of candidate RR Lyrae using Pan-STARRS and WISE data

The final step of the selection algorithm operates in a 4-dimensional space spanned by WISE $W1 - W2$ color (3.4 μm $-$ 4.6 μm), Pan-STARRS $g - r$ color and Pan-STARRS based variability parameters q (statistical significance of the observed variability) and τ (time scale for a damped random walk fit, in days). Prior to this step, a number of quality cuts are performed by Nina's code, as explained in her document. In addition to quality cuts, Nina also places a cut on q, the statistical significance of the observed variability ($q = \chi^2_{dof}\sqrt{N_{dof}/2}$), and fits damped random walk model to all candidate sources. Nina's quality cuts are by and large driven by the quantity and quality of available Pan-STARRS data, and together with the $q > 5$ cut, result in a **starting selection completeness of about 75%.** Only 7.5% of all selected sources are confirmed RR Lyrae (purity=7.5%). The following cuts were used to significantly increase sample purity, while maintaining relatively high completeness.

Based on the behavior of known RR Lyrae and quasars (361 and 2373, respectively, out of 4831 selected sources), illustrated in Fig. 1, the following selection cuts were adopted (in addition to the general $14 < g < 21$ requirement):

1. $-1.0 < g - r < 0.4$ to reject red sources.

2. $q > 30$ to reject small-amplitude variables.

3. $\tau < 10$ days to reject most quasars and long-period variables.

4. $W1 - W2 < 0.5$ (when detected by WISE) to reject essentially all quasars.

These selection criteria result in a sample completeness of 75% (overall 50%, since Nina's selection had a completeness of, coincidentally, 75%) and purity of 50%. Selection by colors only results in completeness of 85% and purity of 24%, and selection by variability results in completeness of 89% and purity of 35%. Most of contaminants are expected to be eclipsing binary stars from the blue end of the main stellar locus. Quasars are not expected to represent a major contaminant: just like Abbas et al. rejected them using SDSS $u - g$ colors, here we rejected them using WISE $W1 - W2$ colors.

Therefore, instead of attempting light curve fitting for 4831 sources that include 361 RR Lyrae now we can fit 539 light curves that include 269 RR Lyrae (an incremental loss of 25% of RR Lyrae), while decreasing the initial sample by a factor of 9.

2.2. RR Lyrae template fitting results

Sandra used Branimir's code to fit RR Lyrae templates to selected 539 candidates. As a result of the above selection procedures, the starting sample from Stripe 82 includes equal numbers of known RR Lyrae and "Other" candidates. Again, the starting completness is 75%.

Unfortunately, the set of 214 templates only included ab-type stars; **the 59 c-type templates were not included** and this omission resulted in some problems discussed below.

The fitting code was sped up using two simple tricks: first, the period grid step was increased by a factor of two, and only every second template was fit. After the period corresponding to the mimimum χ^2 was found, its neighborhood (multiplicative tolerance of 0.015%, or typically ± 1 minute around the initial best-fit period) was explored with the original period step and the complete template set. As a result, up to about 500 stars can be fit in a 24 hour period on the available cluster. However, the throughput can be much lower if other users are loading the cluster.

Originally, we planned to allow for the "amplitude correction" factor but after some bugs in implementation, we decided to first try fits without template modifications (except for the period change, of course).

We have visually inspected all the best-fit phased light curves and concluded that on average better fits are obtained for known RR Lyrae subsample. However, there are some clearly bad fits for RR Lyrae where the best-fit period is probably wrong (judging from the scatter of data points; for examples of good and bad fits for RR Lyrae see Figs. 2–5). On the other hand, there are some formally excellent fits for the Other subsample. Most often they can be traced to a very small number of data points, or to large photometric uncertainties at the faint end (for illustration, see Figs. 6 and 7).

2.2.1. The χ^2_{dof} distributions and selection performance

The χ^2_{dof} distributions for the two subsamples (RR Lyrae and Other) are compared in Fig. 8. As evident, the distribution for known RR Lyrae is shifted towards smaller values. The corresponding "significance" factor (defined as $(\chi^2_{dof} - 1)\sqrt{N_{dof}/2}$) distributions are shown in Fig. 9. Even for RR Lyrae, a large fraction of fits have statistically unlikely large values. This implausibly large fraction could be due to (with decreasing likelihood) wrong best-fit period, inadequate template (e.g. requiring amplitude correction factor), or problems with Pan-STARRS data.

We measured how well known RR Lyrae can be separated from Other subsample using a χ^2_{dof} threshold. The results (the so-called ROC, Receiver Operating Characteristic, curve) are shown in Fig. 10. Using a selection cut for RR Lyrae as $\chi^2_{dof} < 5$, the sample purity can be boosted to 81%, with a (total) completeness of 38% (relative completeness of 50%). For comparison, at the sample completeness level obtained by Abbas et al. (52%), the template fitting method, using $\chi^2_{dof} < 10$, can deliver purity of 74%, compared to 77% for the Abbas et al. method. Therefore, the performance of the template fitting method is comparable to, and just slightly worse than for the Abbas et al. method. **However, this is with c type problem still unresolved!**

The number of Pan-STARRS observations per star varies (the median is 23, with interquartile-based scatter of about 20%) and the statistical significance factor might be a better parameter for sample selection. Fig. 10 also shows the ROC curve for selection based on the significance factor; unfortunately, no significant improvement is seen compared to the chi^2_{dof} based selection.

2.3. Problems with c-type RR Lyrae

Fig. 11 shows the period "correction" factor vs. the true (SDSS-based) period for known RR Lyrae. For c type RR Lyrae with $P_{true} < 0.42$ days, the lack of proper c-type templates resulted in grossly incorrect periods (by about a factor of 2) for a large fraction of stars (a few were able to recover the correct period because the template periods are allowed to vary by $\pm 20\%$). In addition, their χ^2_{dof} distribution is shifted towards the larger values than for the full RR Lyrae subsample. Therefore, it is likely that the performance (purity and completeness after a χ^2_{dof} cut) of the template fitting method could be improved by including c-type templates in the fitting template set.

Note added in proof:

When known RR Lyrae are separated, the $\chi^2_{dof} < 5$ cut selects 85% of ab type and only 36% of c type. The median χ^2_{dof} for ab type is 1.9, while for c type it is 5.9. We definitely need to use proper templates for c type stars!

2.4. Further analysis and other caveats

We asked a few other questions about the best-fit parameters:

- The best-fit period distributions are similar for RR Lyrae and Other samples.
- As shown in Fig. 12, the best-fit period and the period corresponding to the best-fit template can significantly differ (20% and more). Encouragingly, the distribution for RR Lyrae is centered on 1, with a scatter of about 10%.
- As shown in Fig. 13, the ratio of the best-fit period and the period corresponding to the best-fit template is not correlated with the best-fit χ^2_{dof}.

Of course, there are many more interesting questions to ask. Some of them are

- Could bad fits for known RR Lyrae be improved by sampling with even finer period step than originally used by Branimir?
- Can "amplitude correction" factors improve the fits for RR Lyrae? Would they perhaps allow too much freedom for Other subsample?
- Can we improve period estimates by combining best-fit periods for many templates in a Bayesian way (as per H-W's suggestion)?
- How do completeness and purity depend on apparent brightness? That is, it is plausible that a cut on χ^2_{dof} or the significance factor should be a function of brightness.

3. The Workflow for the Whole-Sky Extension

To summarize, our results motivate the application of this method to the whole sky region covered by Pan-STARRS. In order to do so, the following workflow is required:

1. A summary table from Pan-STARRS LSD, as suggested by Eddie, is used to extract the variance and weighted measurement uncertainty and to compute Nina's q parameter for all sources.

2. For a subset of sources selected by a q cut (originally chosen by Nina to be $q > 5$ but in the context of RR Lyrae search, it can at least in principle be changed to $q > 30$ – modulo its own scatter), Nina runs her DRW code to obtain τ.

3. For the same sources, WISE photometry needs to be extracted.

4. For a subset of sources that pass the selection cuts based on q, $g - r$, $W1 - W2$ and τ, light curve data have to be made available for template fitting. Based on this analysis, we expect about 2 candidates per sq. deg. (it could be many more if too close to the Galactic plane).

5. Using χ^2_{dof} from template fitting, a final sample of RR Lyrae can be constructed, as described here. Once all these steps are done for a large sky area, it would be prudent to redo the Stripe 82 analysis presented here to make sure that all the conclusions are still valid.

4. The Team

The MPIA Summer 2014 RR Lyrae Tiger Team was a cheerful bunch that included: Nina Hernitschek, Željko Ivezić, Sandra Mitrović, Hans-Walter Rix, Eddie Schlafly, and Branimir Sesar.

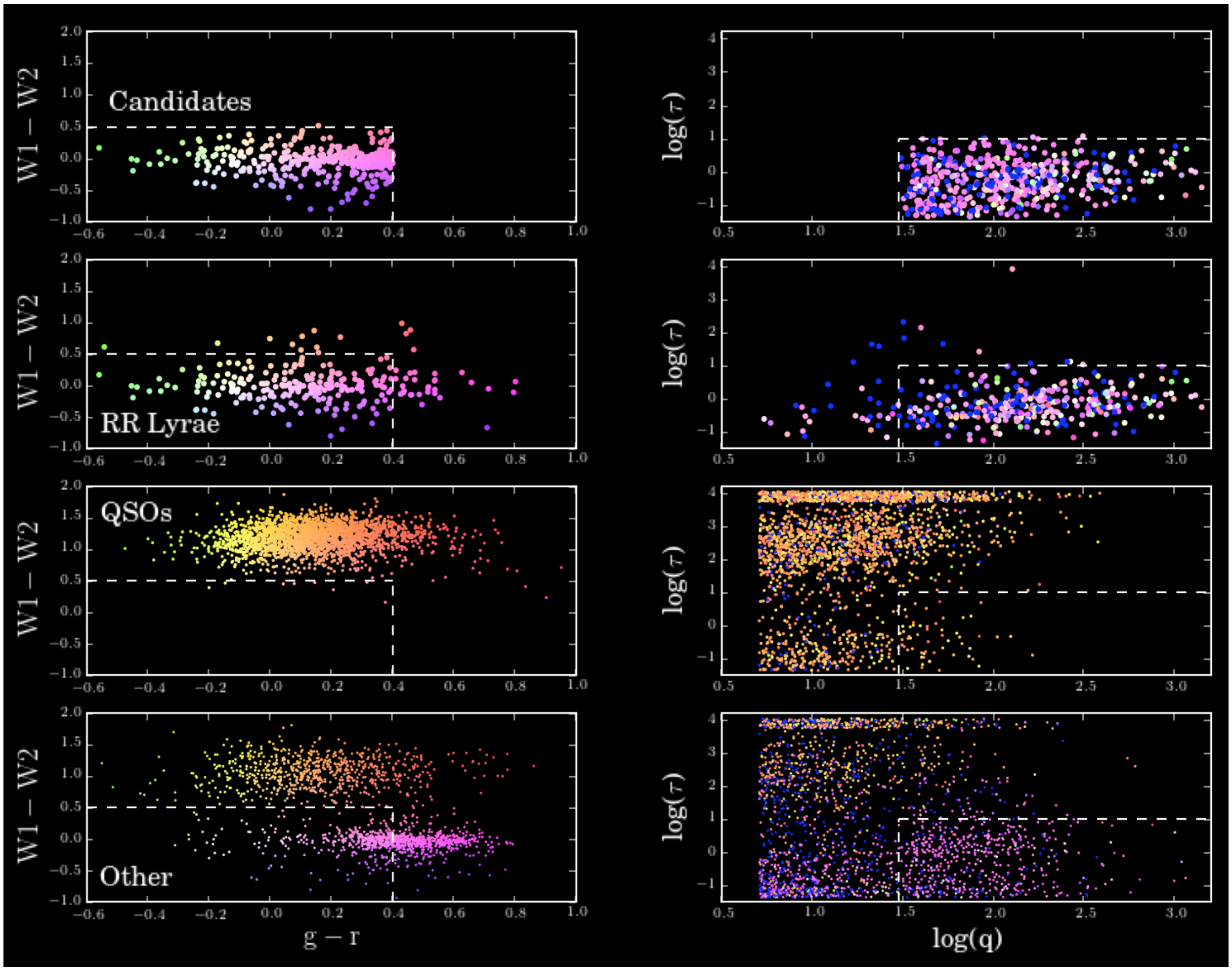

Fig. 1.— Summary of the selection procedure based on WISE $W1-W2$ color (3.4 μm − 4.6 μm) Pan-STARRS $g-r$ color and Pan-STARRS based variabilty parameters q (statistical significance of the observed variability) and τ (time scale for a damped random walk fit, in days). The left column shows the $W1-W2$ vs. $g-r$ color-color diagram and the right column shows the $\log(\tau)$ vs. q "variability" diagram for sources from the SDSS Stripe 82 region. The symbols are color-coded in 2 dimensions, according to their position in the color-color diagram (the blue symbols in the variability diagram for RR Lyrae are objects without WISE detections; practically all quasars are detected by WISE). The known RR Lyrae and quasars are shown in the second and third row, and unclassified sources are shown in the bottom row. The white dashed lines show the adopted selection boundaries for RR Lyrae based on the behavior of known RR Lyrae and quasars. The top row shows the selected candidate RR Lyrae.

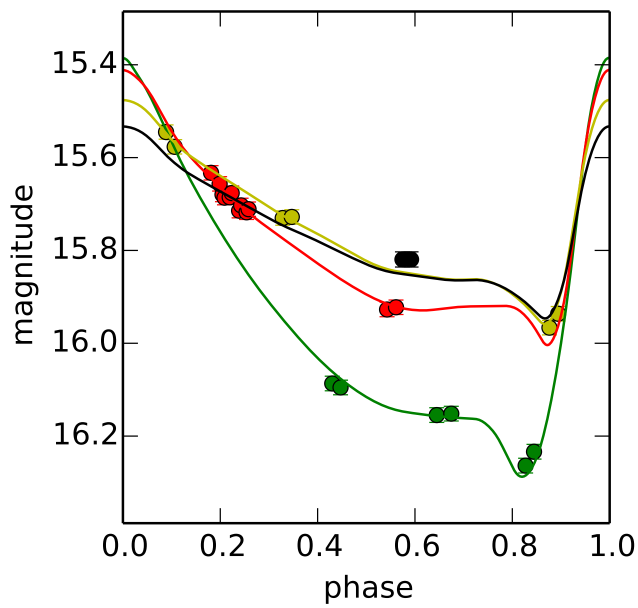

Fig. 2.— An example of a good fit for a known (bright) RR Lyrae.

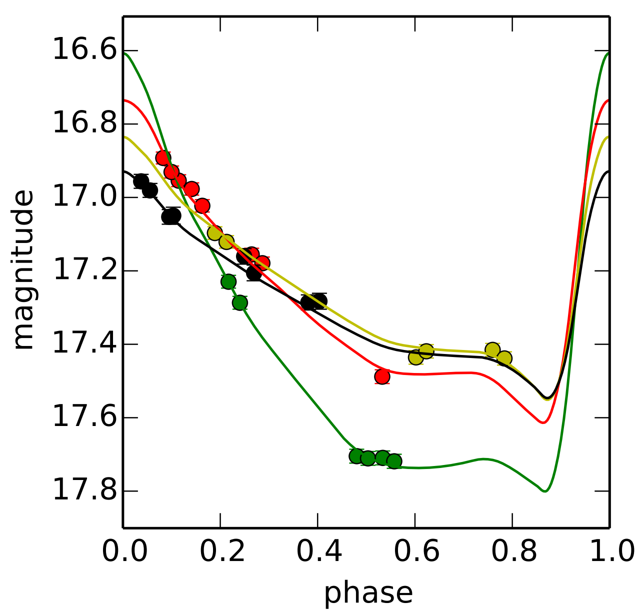

Fig. 3.— An example of a good fit for a known (intermediate brightness) RR Lyrae.

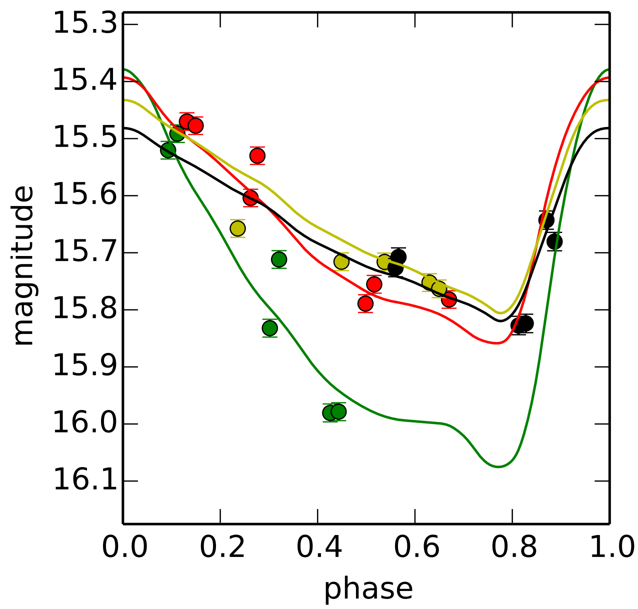

Fig. 4.— An example of a bad fit for a known RR Lyrae.

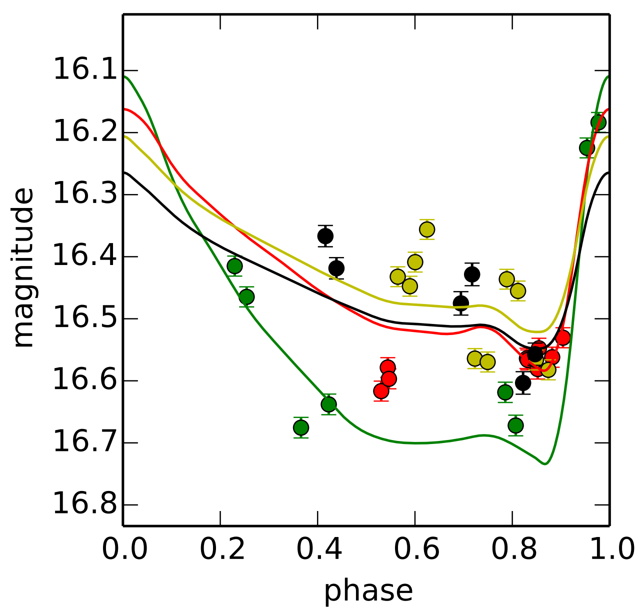

Fig. 5.— An example of a bad fit for a known RR Lyrae.

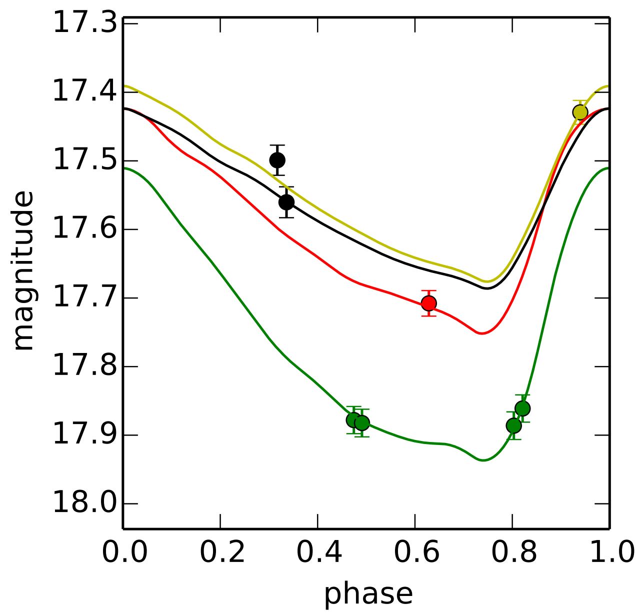

Fig. 6.— An example of a good fit for Other subsample.

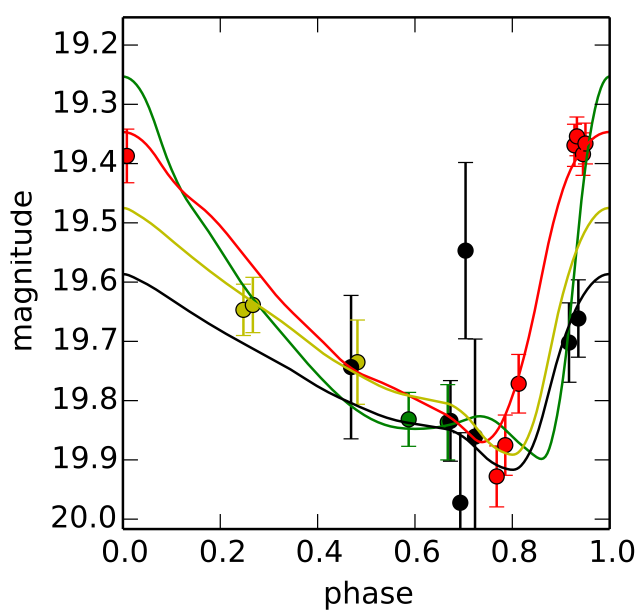

Fig. 7.— An example of a good fit for Other subsample.

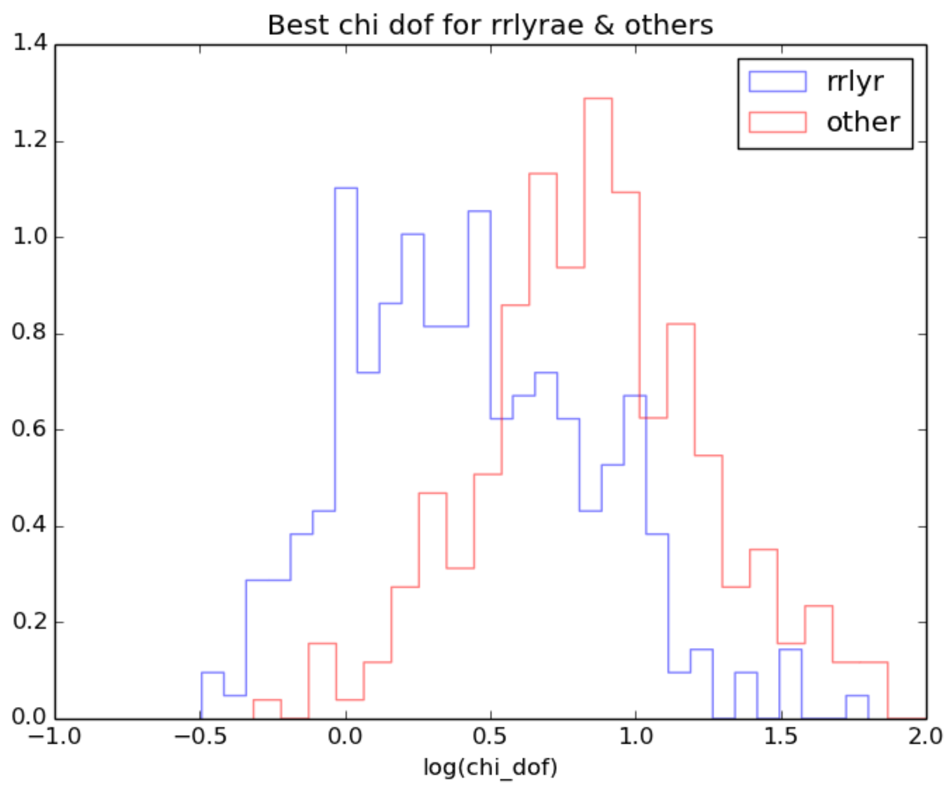

Fig. 8.— The distribution of χ^2_{dof} for known RR Lyrae (blue) and "Other" candidates (red).

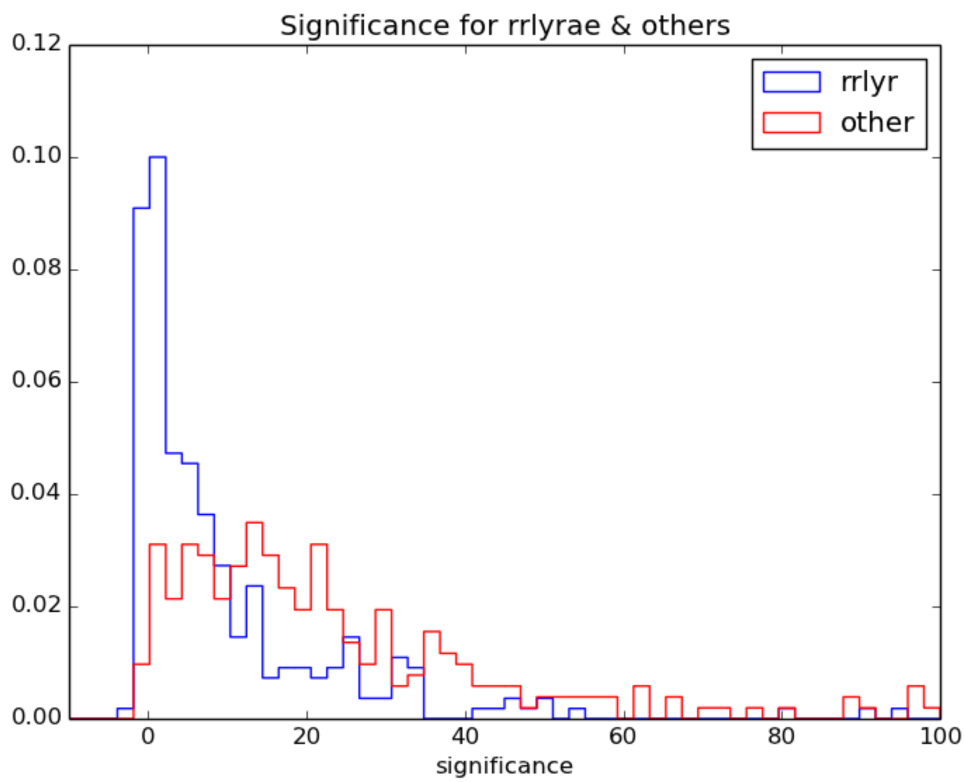

Fig. 9.— The distribution of the significance factor, $(\chi^2_{dof} - 1)\sqrt{N_{dof}/2}$, for known RR Lyrae (blue) and "Other" candidates (red).

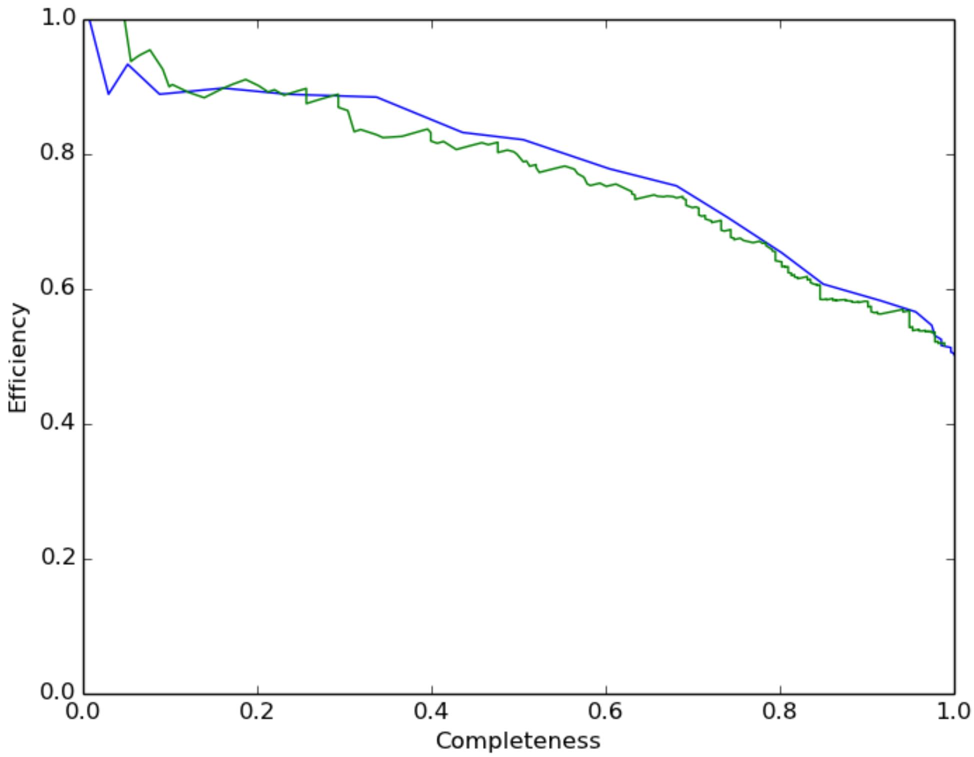

Fig. 10.— The blue curve shows the efficiency vs. completeness (ROC) curve for a classification based on χ^2_{dof}. The completeness is relative to the template fitting method (the starting completeness is 75%). The green curve shows performance based on the significance factor, $(\chi^2_{dof} - 1)\sqrt{N_{dof}/2}$.

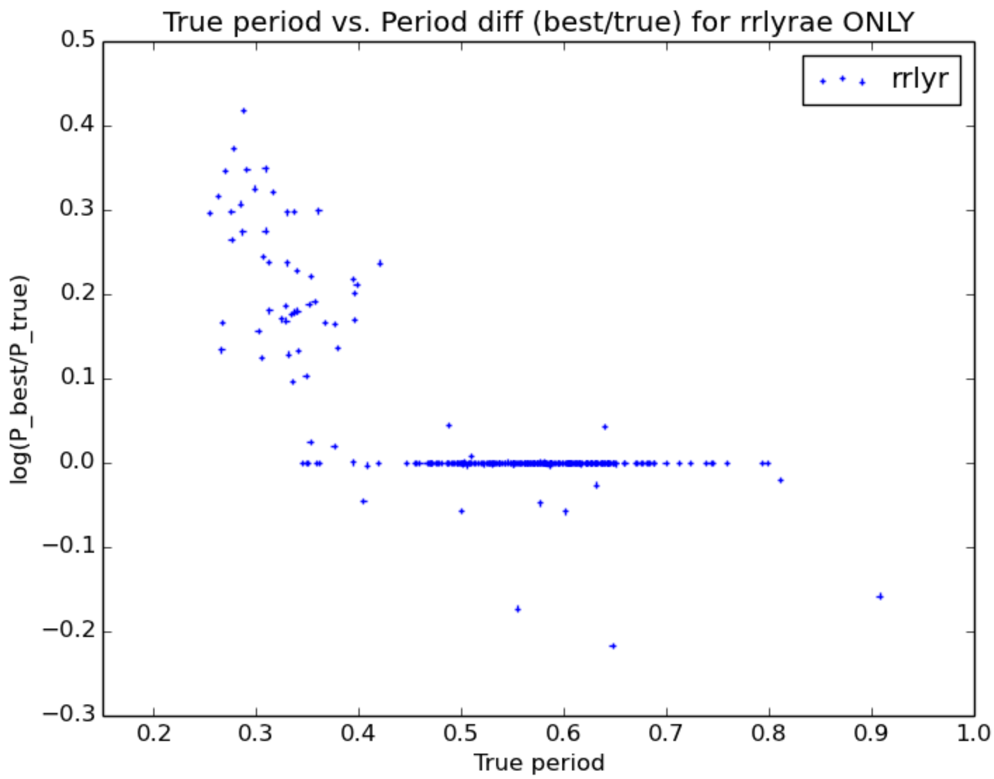

Fig. 11.— The period "correction" factor vs. the true (SDSS-based) period for known RR Lyrae. For the majority of c type RR Lyrae with $P_{true} < 0.42$ days, the lack of proper c-type templates resulted in grossly incorrect periods (by about a factor of 2). A few were able to recover the correct period because the template periods are allowed to vary by ±20%.

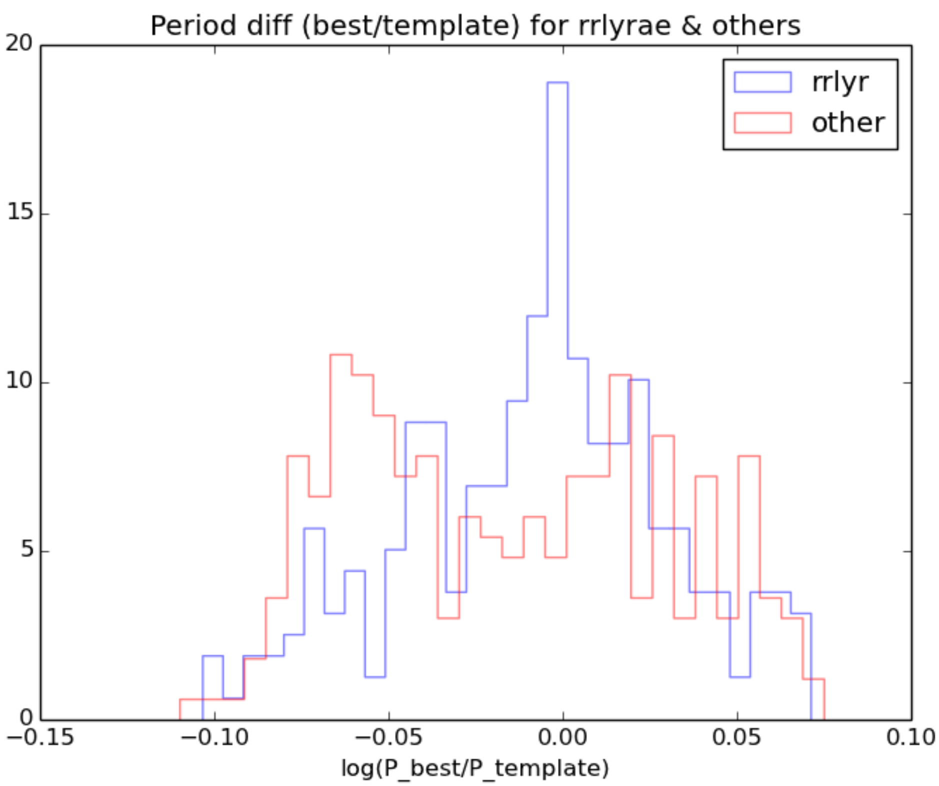

Fig. 12.— The distribution of the period "correction" factor for known RR Lyrae (blue) and "Other" candidates (red).

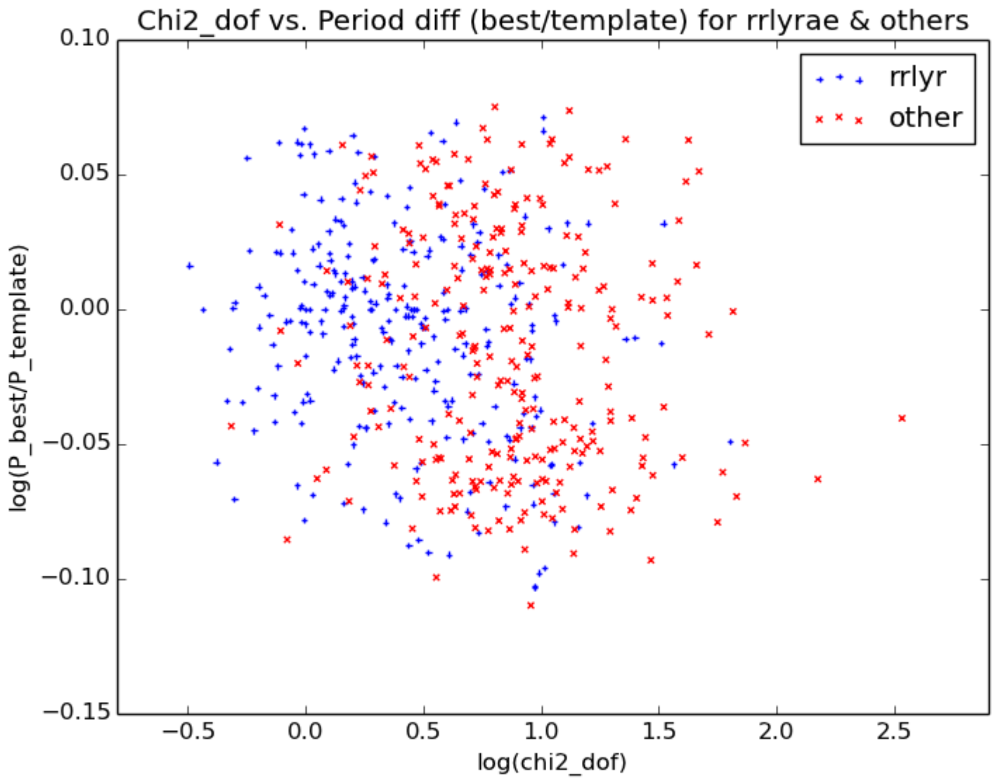

Fig. 13.— The period "correction" factor vs. χ^2_{dof} for known RR Lyrae (blue) and "Other" candidates (red).