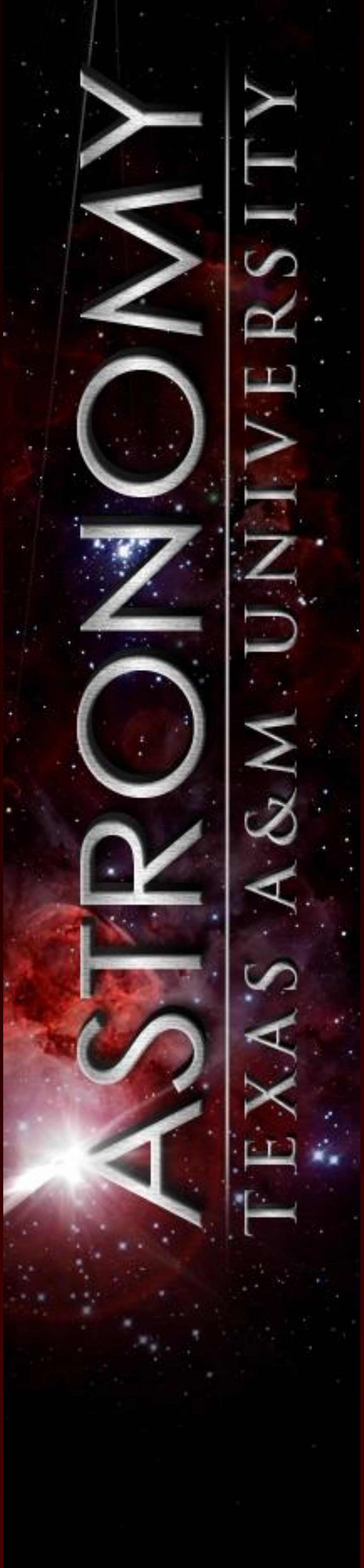
A Multi-band Extension of the Analysis of Variance Period Finding Algorithm

Nicholas Mondrik[1], Jennifer L. Marshall[1], James P. Long[2]

[1]*Texas A&M University Department of Physics and Astronomy*
[2]*Texas A&M University Department of Statistics*



Abstract

One of the largest challenges facing modern astronomical surveys is the automated classification of sources. In the case of transient surveys searching for variable stars, periods are among the most useful features for classification algorithms. In multi-band surveys such as the Dark Energy Survey, which cover a large area of the sky with relatively few visits, single band period finding algorithms can struggle due to poor phase coverage in any one band. We present here an extension of a single band algorithm to utilize an arbitrary number of additional bands. We test the multi-band algorithm on simulated RR Lyrae light curves and analyze performance as a function of photometric error and number of observations per band.

Introduction

Variable stars such as Cepheids and RR Lyraes have proven to be among the most useful astrophysical objects for measuring distance both within our Galaxy and in the local universe. Both of these variable star types lie on the instability strip on a Hertzsprung-Russell diagram and pulsate radially. These variables have a period-luminosity relation that allow astronomers to derive a distance to the star based on the observed period and magnitude. The distances derived from their period-luminosity relations allow us to not only trace substructures in our own Galaxy but can also be used to measure intergalactic distances, thereby calibrating the next step in the astronomical distance ladder. In order to find these variable stars, astronomers use transient surveys which image the same area of the sky across multiple epochs in order to detect changes in the brightness of objects with time.

In order for automated algorithms to classify the thousands of variable sources discovered by these surveys, certain features must be extracted from the data. The period of variable sources (when one exists) can be an important feature in classification, and much work has been done developing methods of period determination. Most modern period-finding algorithms use only one band of data to fit a function or template to determine the best fit period to a single band of data, but throw away data from all other bands. In this work, we show that the inclusion of additional bands can improve the performance of the Analysis of Variance (AoV) algorithm relative to its single-band implementation.

Simulation Procedure

In order to demonstrate the improvement of the AoV algorithm with the inclusion of multiple bands, we generate light curves in multiple bands (*u,g,r,i,z*) following the RR Lyrae templates of Sesar et. al. (2010). We then feed the synthetic light curves into the AoV algorithm (Schwarzenberg-Czerny 1989) and compare the output period to the generated period. To incorporate multiple bands, we sum the output periodogram (which is a measure of probability associated with each period) across all bands. This method depends on the fact that if a period is indeed the true period, it should have a peak in the AoV periodogram in each band which will grow significantly with each sum, while the lower, random peaks will grow more slowly since they are unique to each band. The specific steps of the process are as follows:

1. Generate a light curve with a period and amplitude appropriate for an RR Lyrae-type variable star for each of the 5 bands.
2. For each light curve, randomly generate *n_obs* [date, magnitude] pairs of simulated observations.
3. Adjust each magnitude by the output of a random normal distribution with standard deviation chosen to represent photometric error.
4. Run the AoV algorithm on the generated light curve. Compare the output best fit period with the generated period.
5. For the multi-band algorithm, add the output periodograms of the single band AoV method and select the best period.

Figure 1 shows the individual templates used in the simulation, as well as an example of simulated observations plotted on top of their template.

Algorithm Performance

We define a successfully recovered period as satisfying

$$\frac{|P_{Alg} - P_{Gen}|}{P_{Gen}} \le 0.01 * P_{Gen}$$

Where P_{Alg} and P_{Gen} are the period returned by the algorithm and the generated period, respectively. We generate 100 objects with randomly chosen periods and amplitudes, which corresponds to 500 runs of the single band algorithm and 100 runs of the multi-band implementation. To facilitate a more appropriate comparison, we normalize the number of recovered periods by the number of total runs for each method. In Figure 2, we present results from the simulation for photometric errors of 0.01 mag, 0.05 mag, and 0.1 mag. We also present a set of simulations with the number of observations in the *g*-band set to twice that of the other bands to demonstrate that the multi-band method performs better even when one band is dominant. Figure 3 shows how the algorithm performs when only 3 bands are used for both the single band and multi-band method.

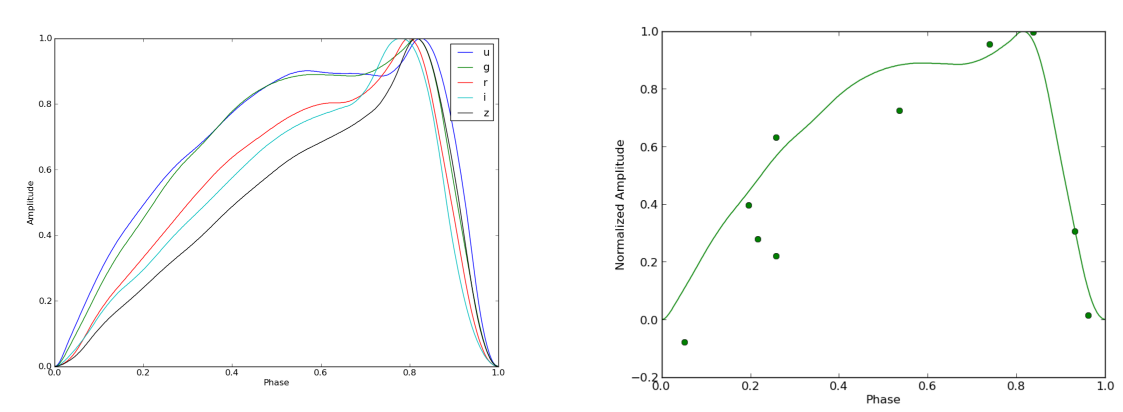


Fig. 1: Left: The RRab templates used to generate the simulated observations. Right: The *g*-band template with simulated observations phased and plotted. The amplitude of the simulated observations has been scaled to match the normalization of the template.

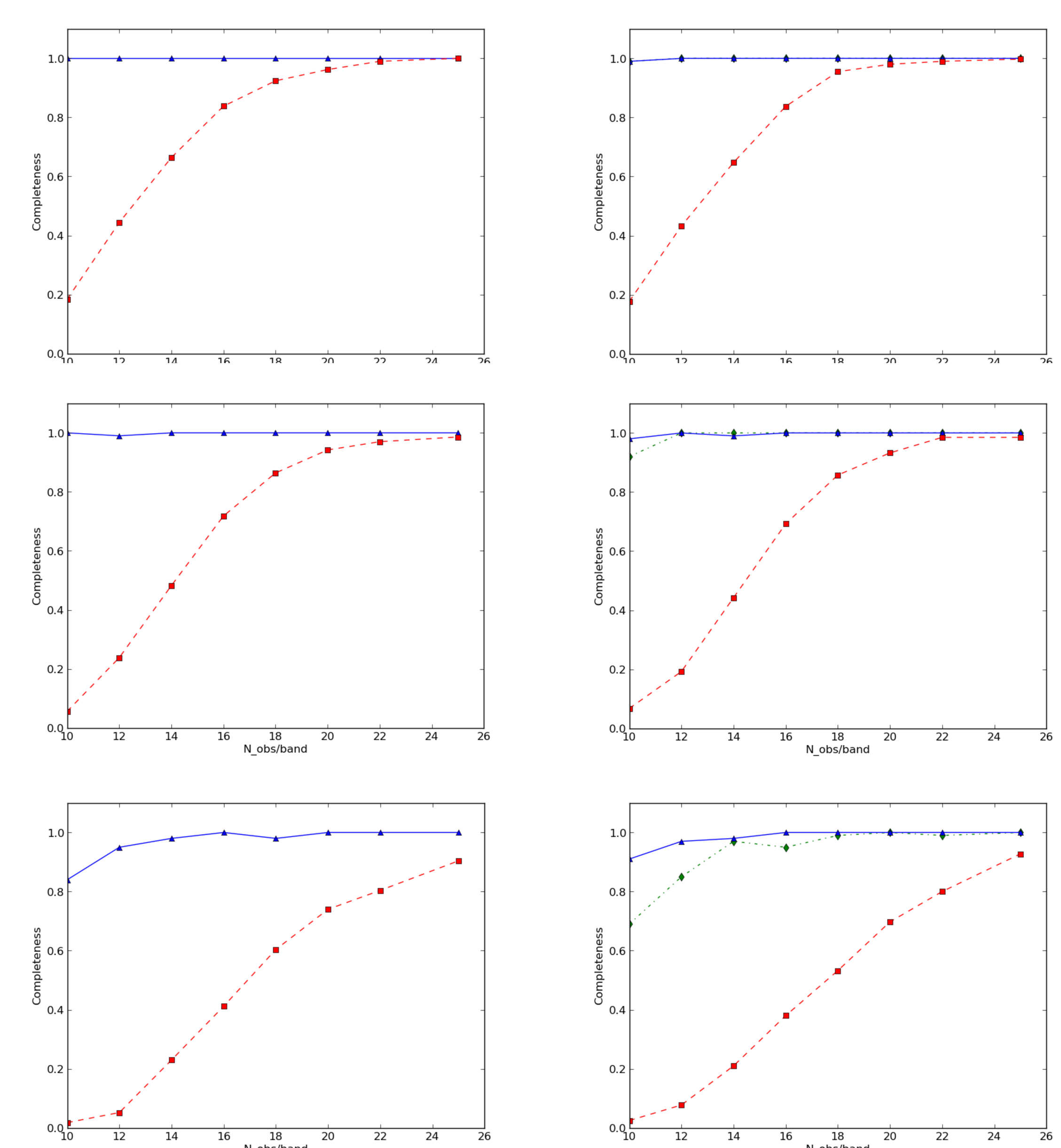


Fig. 2: Left column: A simple simulation with all bands having equal number of observations. Multi-band results are plotted as blue triangles, while single band results are represented as red squares. From top to bottom represents errors of 0.01, 0.05, and 0.1 mag., respectively. Right column: Same as left, but with N_obs for *g*-band doubled (i.e., a *g*-band point at N_obs/band of 12 uses 24 observations). The *g*-band is represented as green diamonds. The single band performance does not include the *g*-band.

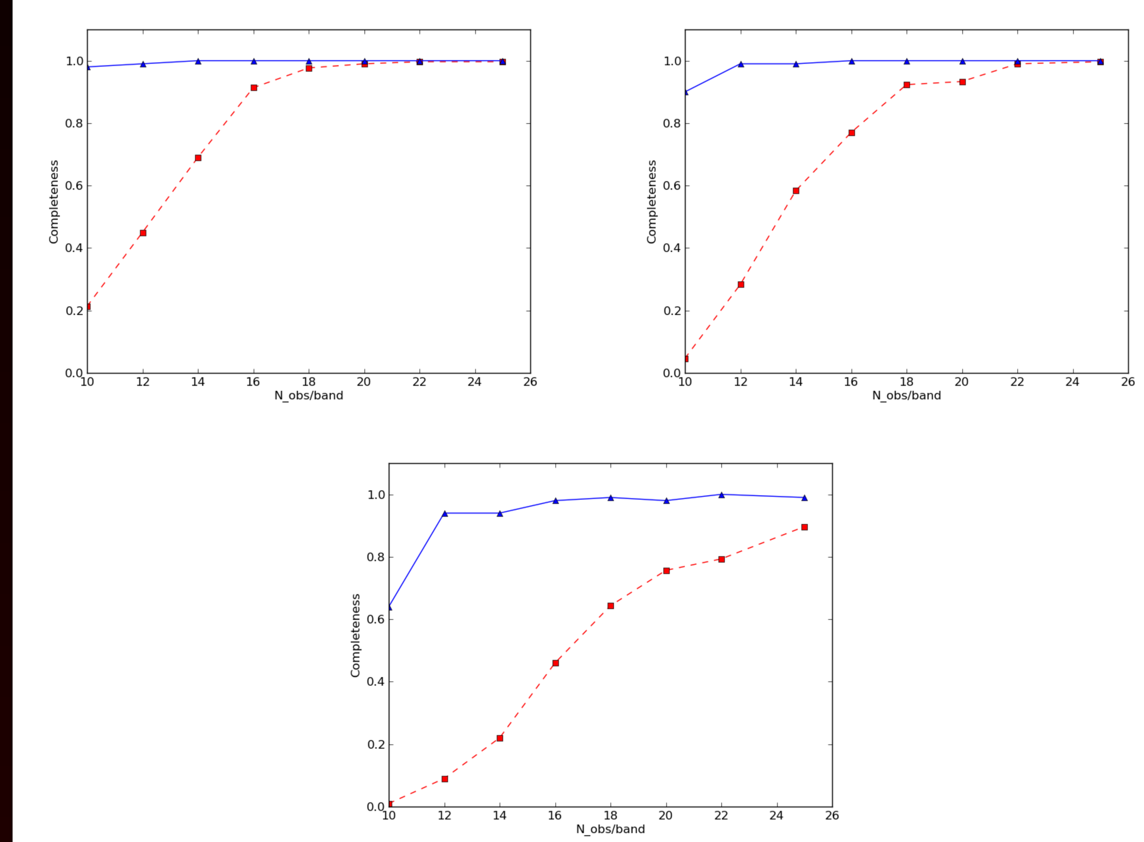


Fig. 3: A similar simulation as Fig. 2, but using only 3 bands for both the single band and multi-band method. Errors are given as 0.01 (top left), 0.05 (top right), and 0.1 mag (bottom).

Analysis of Results

In Fig 2., we show that for n_obs < 20, inclusion of additional bands in the AoV algorithm can improve period recovery rates by up to 80%. For Figure 3, we show that even when reduced to only three available bands, the multi-band method vastly outperforms the single band method out to n_obs = 20. In addition, both figures show that the multi-band AoV method retains the expected performance for large numbers of observations per band. Thus, for surveys involving less than ~20 observations per band, use of the multi-band algorithm can improve period recovery rates considerably. This allows surveys which may not be specifically designed for transient period identification to determine accurate periods of variable stars.

Conclusions

We have presented a method of incorporating multiple bands of data into the AoV period finding algorithm. Results for simulated RR Lyrae light curves show that period recovery rates can increase by up to 80% in the case of poorly sampled, but precise, light curves. It should be noted that this method of extension to multi-band data is not unique to the AoV algorithm. Indeed, any algorithm that outputs some measure of probability could in principle be extended in this manner. This method of modifying period finding algorithms could be particularly useful for surveys imaging fields a small to moderate number of times, which may produce too few observations in any one band to reliably recover periods. By combining data across multiple bands, these surveys can extended into the transient domain. This method could also impact the transient survey domain by allowing surveys to incorporate more than just a few filters, thereby opening up the ability to explore transient phenomena in a higher dimensional color-space than previously possible.

References

Sesar, B., et al. 2010, ApJ, 708,717

Schwarzenberg-Czerny, A., 1989, MNRAS, 241,153

Texas A&M University Department of Physics and Astronomy is an institutional member of:

Munnerlyn Astronomical Instrumentation Lab, Texas A&M University · GMT · HETDEX Hobby-Eberly Telescope Dark Energy Experiment, Illuminating the Darkness · LSST Large Synoptic Survey Telescope · DARK ENERGY SURVEY

Acknowledgments

Texas A&M University thanks Charles R. '62 and Judith G. Munnerlyn, George P. '40 and Cynthia Woods Mitchell, and their families for support of astronomical instrumentation activities in the Department of Physics and Astronomy. This work was supported by NSF grant Ast-1263034.