

# Automatyzacja analizy cefeid

Bartłomiej Kamiński

11 marca 2011

## 1 Wstęp

Ćwiczenie polegało na rozwijaniu programu do automatycznej analizy danych z pomiarów fotometrycznych cefeid wykonanych w eksperymencie Pi of the Sky. Analiza polega na znalezieniu okresu zmienności gwiazdy i dopasowaniu szeregu Fouriera do jej krzywej blasku, co pozwala określić jej klasę zmienności. Stworzono program, który umożliwia szybkie wykonanie tych czynności oraz zastosowano go do przykładowych danych zebranych dla 4 gwiazd zmiennych.

### 1.1 Cefeidy

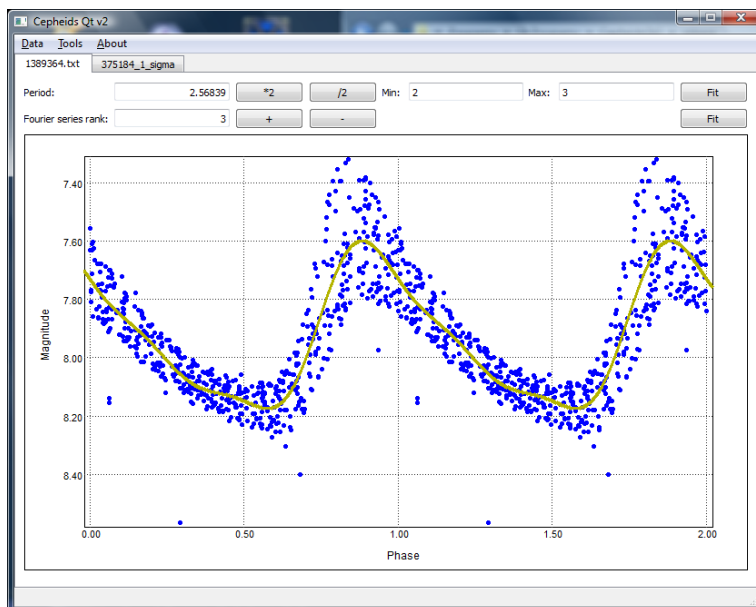
Cefeidy są to gwiazdy zmienne, których jasność ulega zmianom na skutek pulsacji, które polegają na zmianach gęstości materii gwiazdy i powodują wahania ciśnienia i temperatury [1]. Cefeidy dzielą się na kilka podtypów, wśród których można wyróżnić:

- Cefeidy II populacji (typ W Virginis) - oznaczane CW
- Typ  $\delta$  Cephei - oznaczane DCEP
- Cefeidy owertonowe - oznaczane DCEPS

Typ zmienności gwiazdy można określić na podstawie współczynników szeregu Fouriera.

## 2 Program

Program CepheidsQt2 został napisany w języku C++ z użyciem biblioteki Qt. Jego interfejs wygląda następująco:



Rys. 1: Interfejs programu

### 2.1 Funkcje

Program posiada następujące funkcje:

- Konwertowanie plików z danymi z katalogów Pi i ASASa na format czytelny dla programu.
- Wczytywanie danych z pliku lub bezpośrednio z bazy danych.
- Znajdowanie najlepszego dopasowania okresu w zadanym przedziale.
- Dopasowywanie szeregu Fouriera do danych.
- Zapisywanie parametrów szeregu Fouriera do pliku lub do bazy danych.
- Rysowanie wykresów krzywych blasku, sfazowanych krzywych blasku oraz parametrów szeregu Fouriera w zależności od logarytmu okresu zmienności.
- Ocenianie prawdopodobieństwa przynależności gwiazdy do różnych typów zmienności za pomocą sieci neuronowej na podstawie współczynników szeregu Fouriera.

Poniżej opisane są szczegóły niektórych funkcji.

### 2.1.1 Konwertowanie danych

Dane w katalogach Pi oraz ASAS zawierają dużo danych, które są zbędne przy wykreślaniu krzywych blasku oraz znajdowaniu parametrów Fouriera. Funkcja konwersji wybiera z nich wyłącznie potrzebne dane i zapisuje je do oddzielnego pliku.

### 2.1.2 Dopasowywanie okresu

Okres jest dopasowywany metodą tzw. "długości sznurka"[2]. Jest to szybka i prosta metoda, dająca jednocześnie bardzo dobre wyniki. Dla każdego okresu z zadanego przedziału obliczana jest "długość sznurka", wyrażająca się wzorem:

$$L = \sum_{i=1}^{N-1} \sqrt{(\phi_{i+1} - \phi_i)^2 + (m_{i+1} - m_i)^2}$$

gdzie  $N$  - liczba punktów pomiarowych,  $(\phi_i, m_i)$  - współrzędne  $i$ -tego punktu, przy czym  $\phi_i$  jest fazą obliczoną przy aktualnie założonym okresie oraz punkty są posortowane tak, że  $\phi_i < \phi_{i+1}$ , natomiast  $m_i$  oznacza jasność gwiazdy. Okres, dla którego  $L$  jest minimalne jest uznawany za najlepsze dopasowanie.

### 2.1.3 Dopasowywanie szeregu Fouriera

Szereg Fouriera rzędu wybranego przez użytkownika dopasowywany jest metodą najmniejszych kwadratów. Algorytm przebiega następująco:

1. Do danych dopasowywana jest funkcja:

$$f(x) = \sum_{i=0}^n C_i f_i(x)$$

gdzie  $C_i$  to pewne stałe rzeczywiste, natomiast  $f_i(x) = \cos(\frac{i}{2}x)$  dla parzystych  $i$  oraz  $f_i(x) = \sin(\frac{i+1}{2}x)$  dla nieparzystych  $i$ .

2. Generowana jest macierz  $Z_{ij}$  dana wzorem:

$$Z_{ij} = f_i(x_j)$$

gdzie  $x_j$  to współrzędne  $x$  punktów z danych wejściowych.

3. Szukany wektor współczynników  $C$  wyraża się równaniem:

$$C = (ZZ^T)^{-1}Zy$$

gdzie  $y$  - wektor współrzędnych  $y$  punktów z danych wejściowych. Macierz  $ZZ^T$  odwracana jest metodą Gaussa-Jordana.

### 2.1.4 Wykresy parametrów Fouriera

Program umożliwia naniesienie na wykresy parametrów więcej niż jednej gwiazdy naraz. Aby tego dokonać, należy otworzyć dane gwiazd, które mają być naniesione na wykresy, w zakładkach, dopasować do nich szeregi Fouriera a następnie uruchomić opcję rysowania wykresów. Na każdym wykresie znajdują się dane z wszystkich otwartych zakładek na tle gwiazd wzorcowych (gwiazdy wzorcowe mogą być wczytane z plików CW.txt, DCEP.txt, DCEPS.txt lub z bazy danych - źródło gwiazd wzorcowych można wybrać w menu Tools - Configure).

### 2.1.5 Rozpoznawanie typu zmienności gwiazdy

Rozpoznawanie typu zmienności gwiazdy jest realizowane przy pomocy sieci neuronowej.

Sieć przyjmuje na wejściu 9 parametrów szeregu Fouriera: logarytm okresu  $\log P$ , amplitudę  $A_1$ , amplitudy względne  $R_{21}$ ,  $R_{31}$ ,  $R_{41}$ , fazy względne  $\phi_{21}$ ,  $\phi_{31}$ ,  $\phi_{41}$  oraz różnicę faz między maksimum i minimum  $M - m$ . Na wyjściu pojawiają się 3 wartości odpowiadające typom CW, DCEP i DCEPS, które po znormalizowaniu mogą być interpretowane jako oszacowanie prawdopodobieństwa przynależności do danej klasy.

Sieć składa się z 3 warstw neuronów. Każdy neuron posiada  $k$  wejść o wagach  $w_i$ ,  $i = 1, \dots, k$  oraz jedno wyjście. Gdy na wejścia zostaną podane wartości  $x_i$ ,  $i = 1, \dots, k$ , wartość wyjścia będzie wynosić:

$$y = \frac{1}{1 + \exp(-a\bar{y})}$$

gdzie  $a$  - pewna dodatnia stała dobierana zależnie od potrzeb, natomiast:

$$\bar{y} = \sum_{i=1}^k w_i x_i$$

Każdy z neuronów  $n$ -tej warstwy posiada tyle wejść, ile jest neuronów w warstwie  $n-1$ , dzięki czemu każdy z nich może otrzymać na wejście wartości wszystkich wyjść poprzedniej warstwy. W pierwszej warstwie każdy neuron posiada tyle wejść, ile wejść ma cała sieć. Ostatnia warstwa składa się z liczby neuronów odpowiadającej liczbie wyjść sieci.

Uczenie sieci polega na dopasowywaniu wag wejść poszczególnych neuronów tak, aby sieć dawała poprawne wyniki dla pewnego wzorcowego zestawu danych. Jest to wykonywane metodą wstecznej propagacji błędów [3]. W przypadku programu opisywanego w niniejszym sprawozdaniu, zestawem wzorcowym były 3 pliki tekstowe zawierające dane dla 42 gwiazd wzorcowych typu CW, 178 typu DCEP oraz 18 typu DCEPS z katalogu ASAS. Dzięki zdolności sieci neuronowej do uogólniania, sieć po dopasowaniu się do danych wzorcowych potrafi podawać dobre wyniki także dla innych danych.

### 3 Analiza

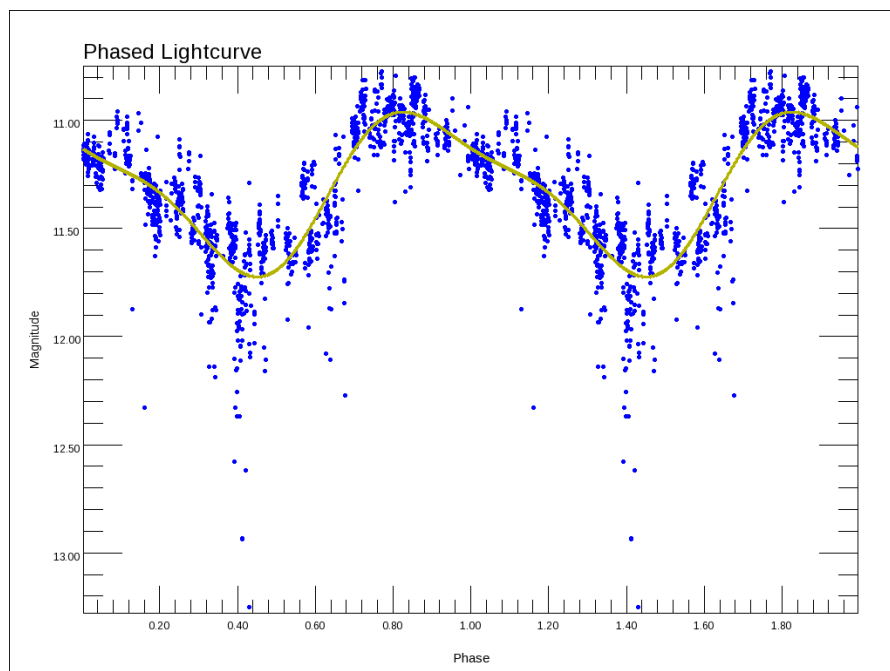
W celu przetestowania programu, zastosowano go do analizy 4 gwiazd. Ich numery identyfikacyjne w bazie danych Pi to 456849, 4936054, 5912203 oraz 9066932.

#### 3.1 Gwiazda 456849

Gwiazda niezidentyfikowana ani w katalogu ASAS, ani w katalogu GCVS.

Okres dopasowany przez program: 7.28671 dnia.

Sfazowana krzywa blasku:

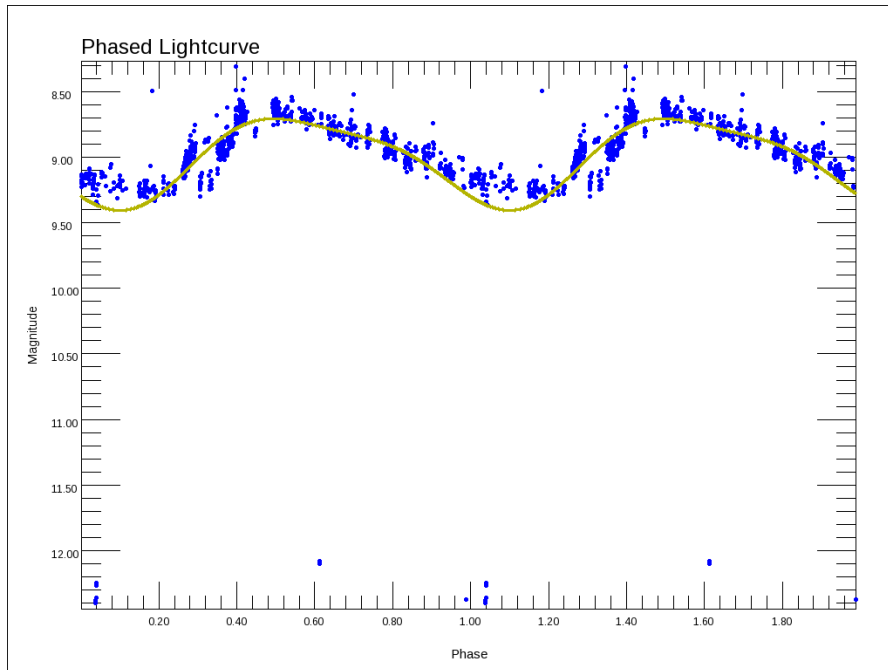


Prawdopodobieństwa przynależności do poszczególnych klas zmienności:

- CW - 4.24%
- DCEP - 93.83%
- DCEPS - 1.93%

### 3.2 Gwiazda 4936054

Gwiazda niezidentyfikowana ani w katalogu ASAS, ani w katalogu GCVS.  
Okres dopasowany przez program: 7.00654 dnia.  
Sfzowana krzywa blasku:



Prawdopodobieństwa przynależności do poszczególnych klas zmienności:

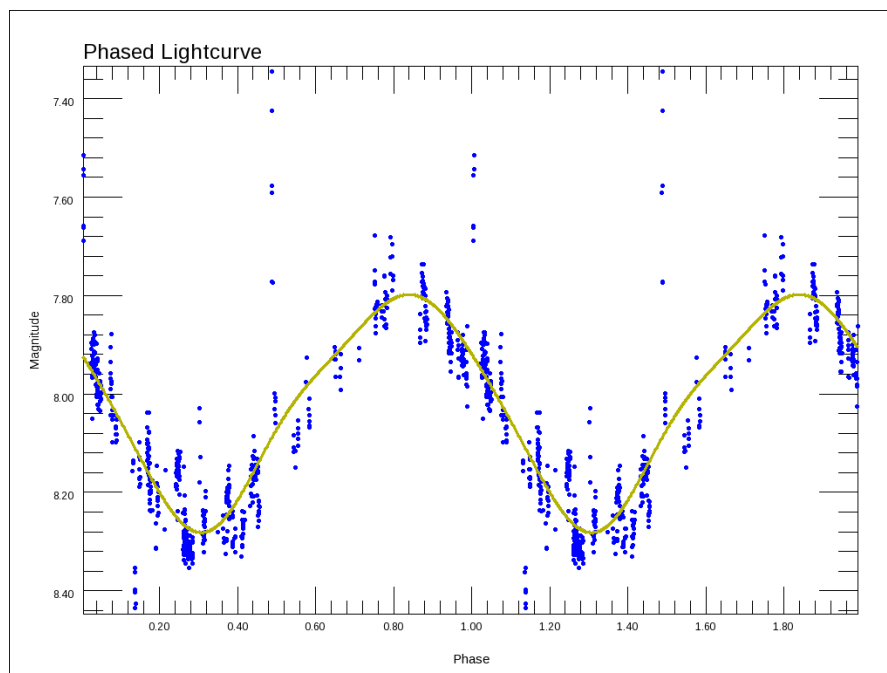
- CW - 1.15%
- DCEP - 93.1%
- DCEPS - 5.75%

### 3.3 Gwiazda 5912203

Gwiazda niezidentyfikowana ani w katalogu ASAS, ani w katalogu GCVS.

Okres dopasowany przez program: 9.84154 dnia.

Sfazowana krzywa blasku:



Prawdopodobieństwa przynależności do poszczególnych klas zmienności:

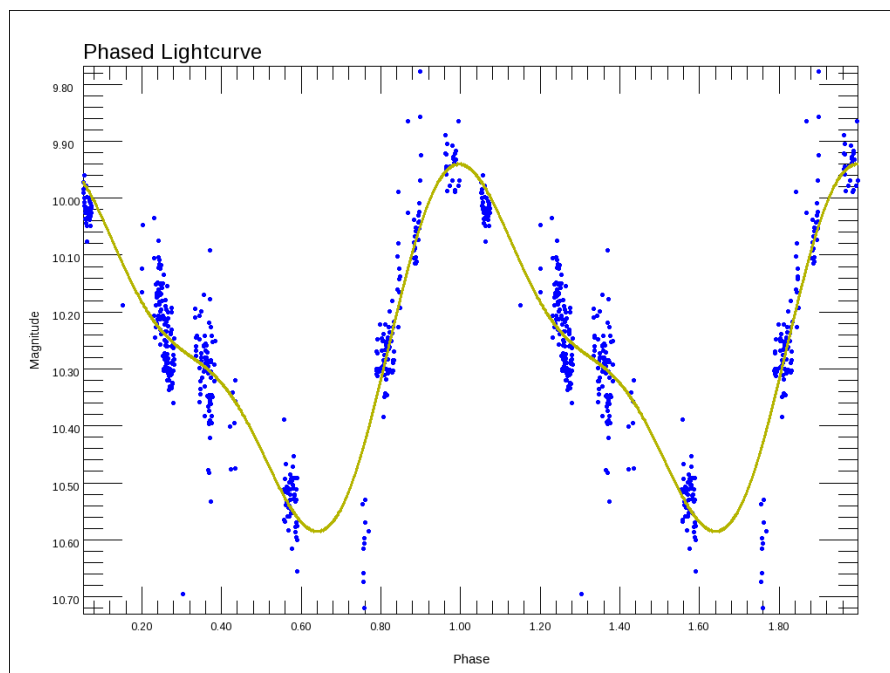
- CW - 0.41%
- DCEP - 94.21%
- DCEPS - 5.38%

### 3.4 Gwiazda 9066932

Gwiazda niezidentyfikowana ani w katalogu ASAS, ani w katalogu GCVS.

Okres dopasowany przez program: 5.72849 dnia.

Sfazowana krzywa blasku:



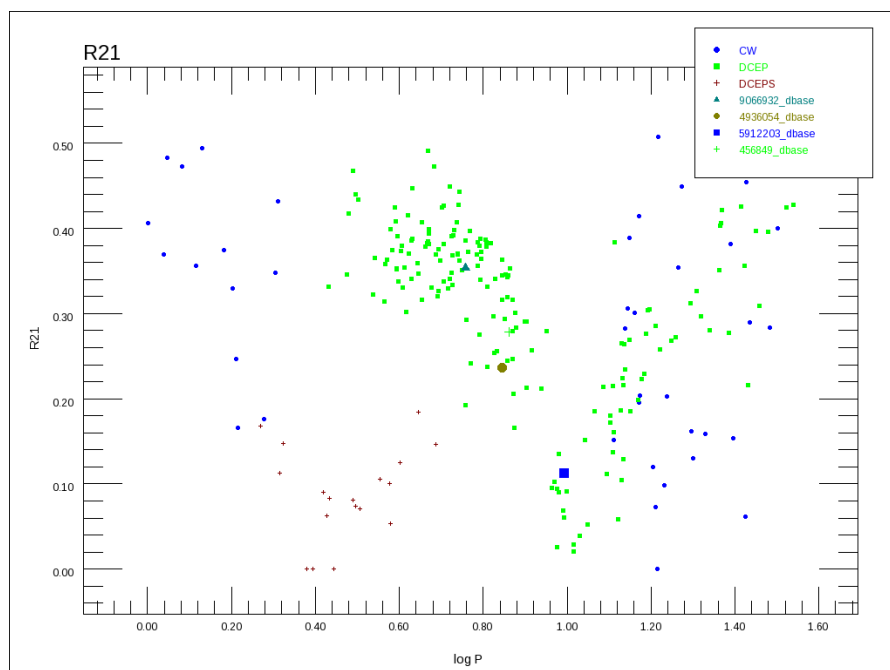
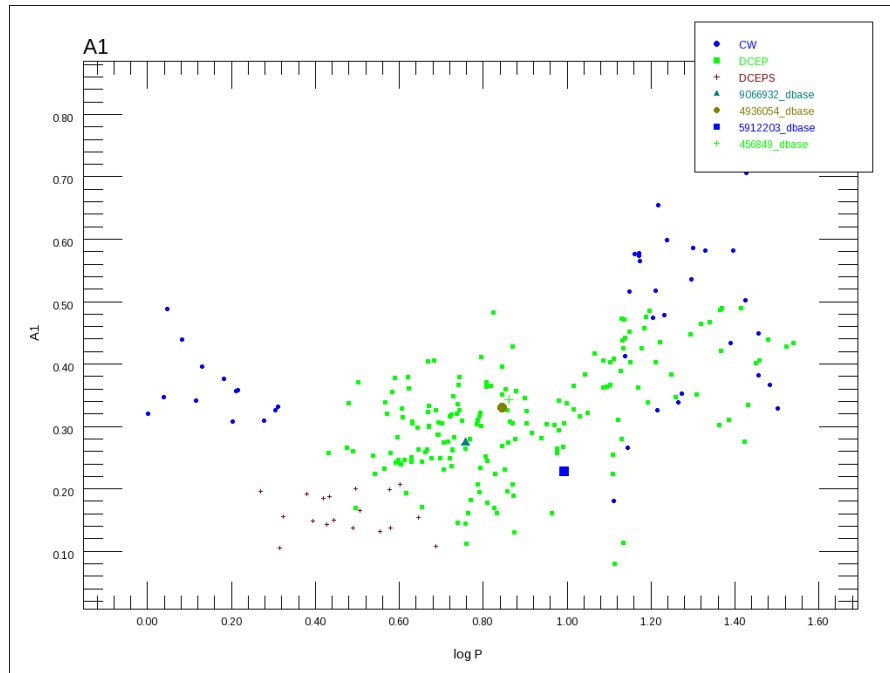
Prawdopodobieństwa przynależności do poszczególnych klas zmienności:

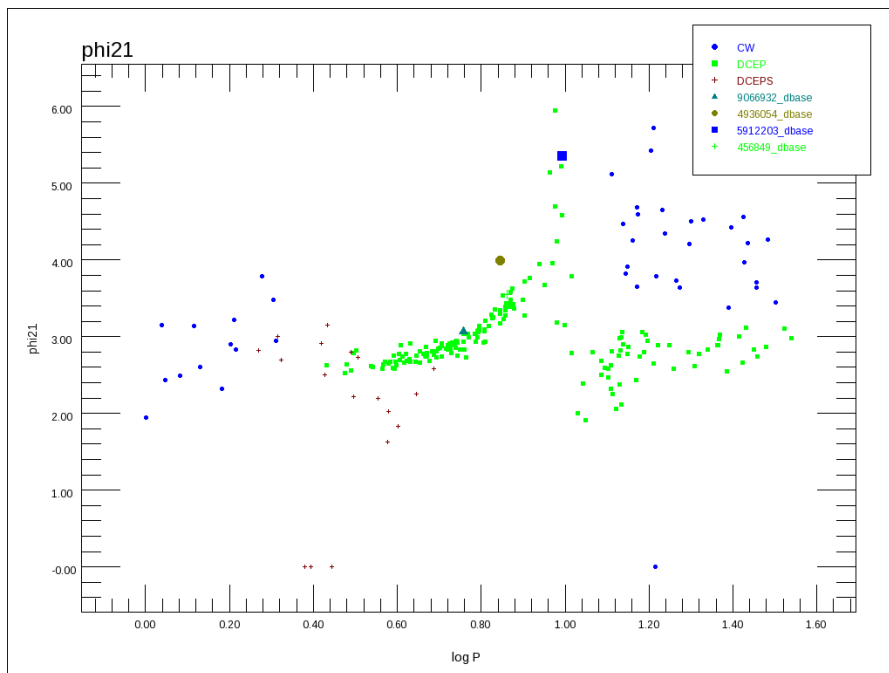
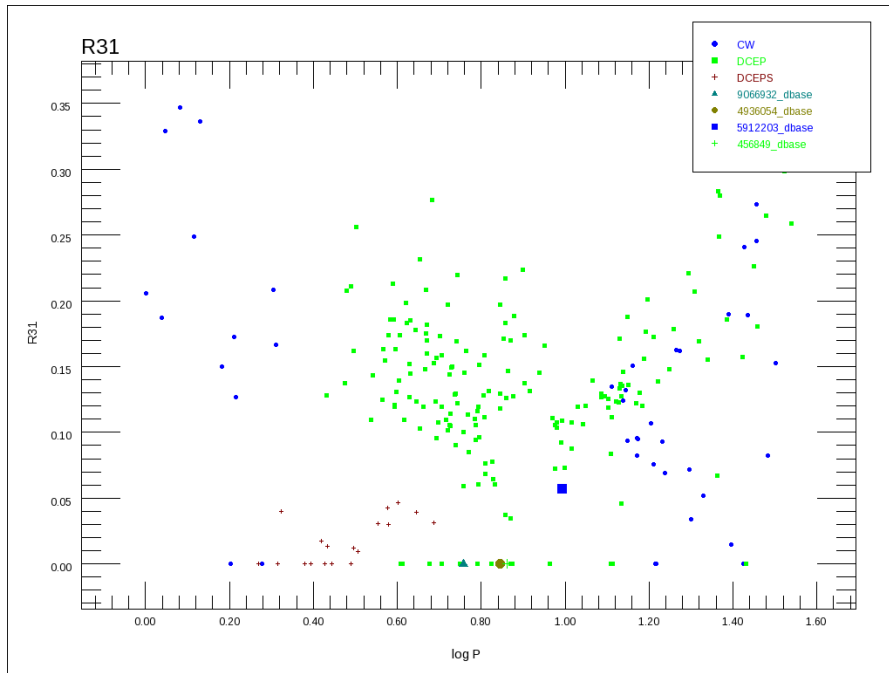
- CW - 4.36%
- DCEP - 92.08%
- DCEPS - 3.56%

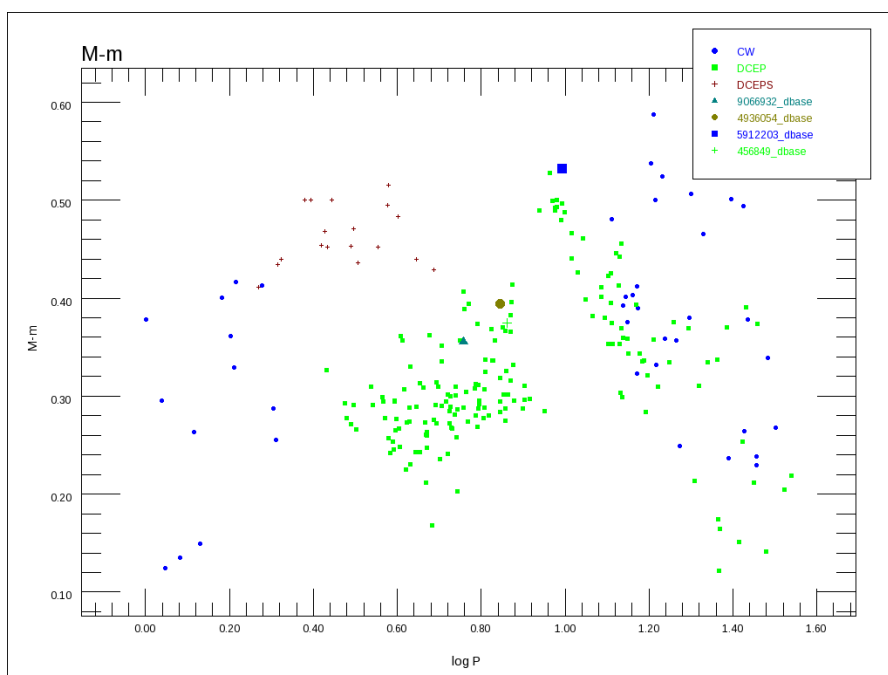
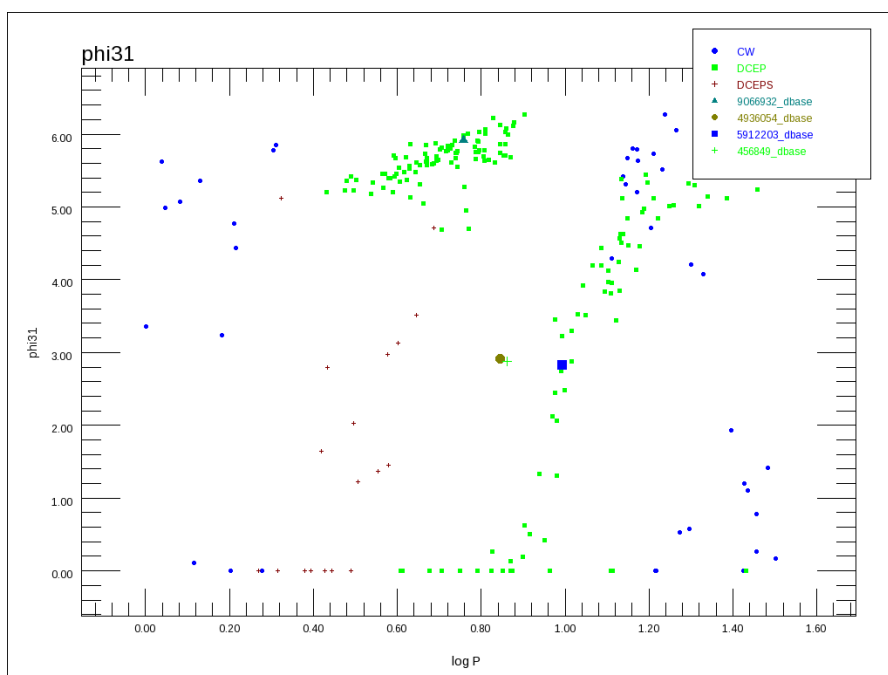


### 3.5 Parametry Fouriera

Wykresy parametrów szeregów Fouriera poszczególnych gwiazd w zależności od logarytmu okresu zmienności przedstawiają się następująco:







Do wszystkich gwiazd dopasowywany był szereg obcięty na maksymalnie trzeciej harmonicznej, gdyż szeregi z uwzględnionymi wyższymi harmonicznymi gorzej oddawały zmienność blasku, stąd amplitudy względne  $R_{41}$  wyniosły dla nich 0.

### 3.6 Wnioski

Dane dobrze fazują się z wyznaczonym okresem, co wskazuje na dobre działanie zastosowanej metody. Sfazowane krzywe blasku wykazują charakterystyczną dla cefeid asymetrię, z wyjątkiem gwiazdy 5912203. Znane są jednak przypadki, w których brak asymetrii, co można zobaczyć na wykresie parametru  $M - m$  (istnieją gwiazdy referencyjne z  $M - m > 0.5$ ). Na wykresach parametrów szeregów Fouriera analizowane gwiazdy leżą wyraźnie w grupie DCEP. Wynik podany przez sieć neuronową jest zgodny z tą obserwacją. Na podstawie danych z Pi możemy zatem sklasyfikować te gwiazdy jako gwiazdy typu DCEP.

## 4 Podsumowanie

Program CepheidsQt2 udostępnia kilka podstawowych funkcji przydatnych przy analizowaniu gwiazd zmiennych. Pozwala szybko wczytać krzywą blasku, dopasować okres i szereg Fouriera, co jest niezbędne do określenia typu zmienności, a także stworzyć wykresy i automatycznie określić typ zmienności. Zastosowane metody pozwalają uzyskać poprawne wyniki. Dzięki temu program może znacznie usprawnić analizę danych.

## Literatura

- [1] M. Siudek, *Klasyfikacja gwiazd zmiennych na podstawie analizy danych fotometrycznych w eksperymencie „Pi of the Sky”*, marzec 2010.
- [2] M. M. Dworetzky, *A period-finding method for sparse randomly spaced observations of ‘How long is a piece of string?’* Royal Astronomical Society, Monthly Notices (ISSN 0035-8711), vol. 203, p. 917-924, June 1983.
- [3] Opis metody wstecznej propagacji błędów na stronie AGH: [http://galaxy.agh.edu.pl/~vlsi/AI/backp\\_t\\_en/backprop.html](http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html)