Parametrization of binary stars with Gaia observations

Paraskevi Tsalmantza
Coryn A. L. Bailer-Jones

MPIA Heidelberg

Workshop on Astrostatistics and Data Mining of Large Astronomical Databases

30 May 2011

Outline

- The goal

Estimate astrophysical parameters for both stars
(e.g. T_{eff}, logg, A_0, brightness ratio)

- The sources

Unresolved binaries
Classified as binaries spectroscopically by DSC
Both stars at the same distance with the same extinction

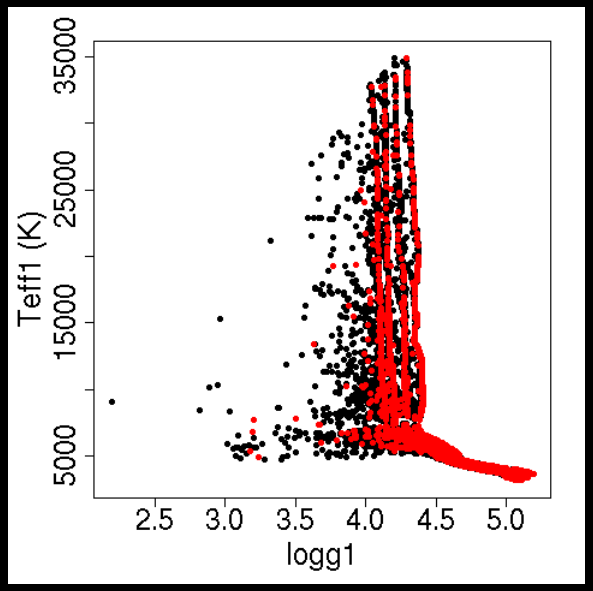
- The data

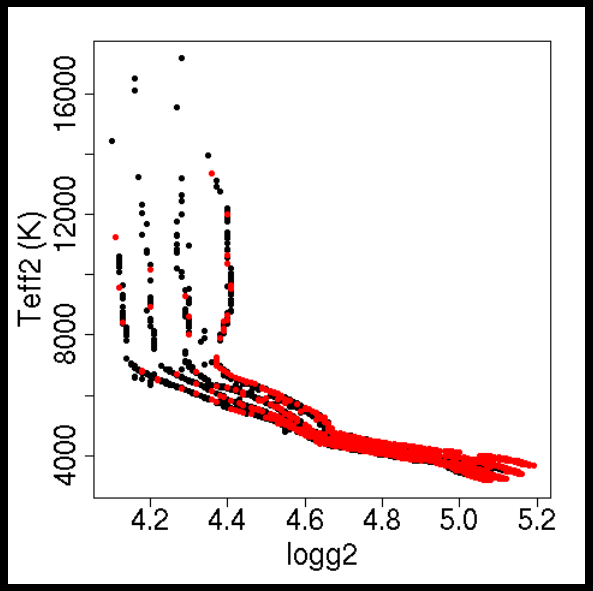
BP/RP spectra, magnitudes, parallaxes, proper motions

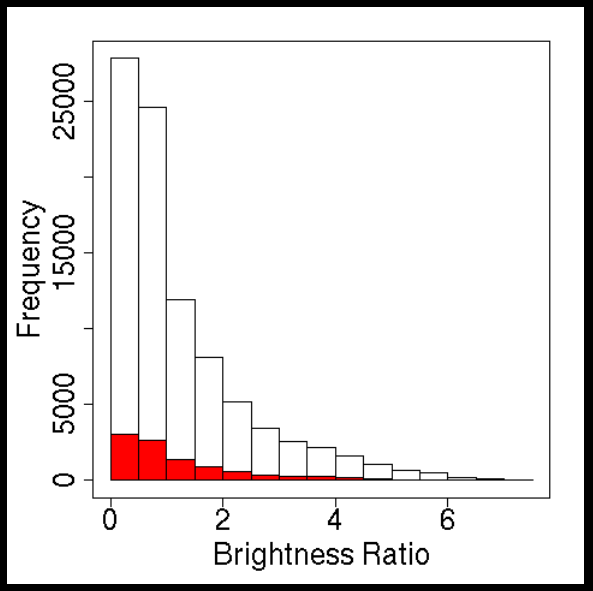
- The methods

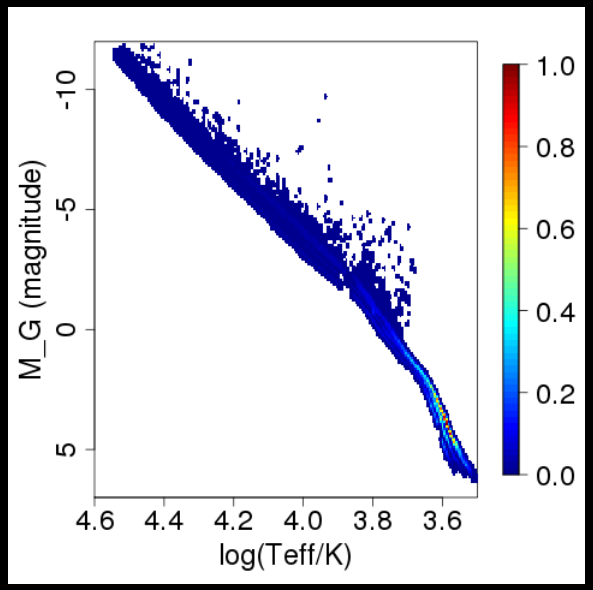
Support Vector Machines
Combination of SVM with H-R diagram prior information
q-method

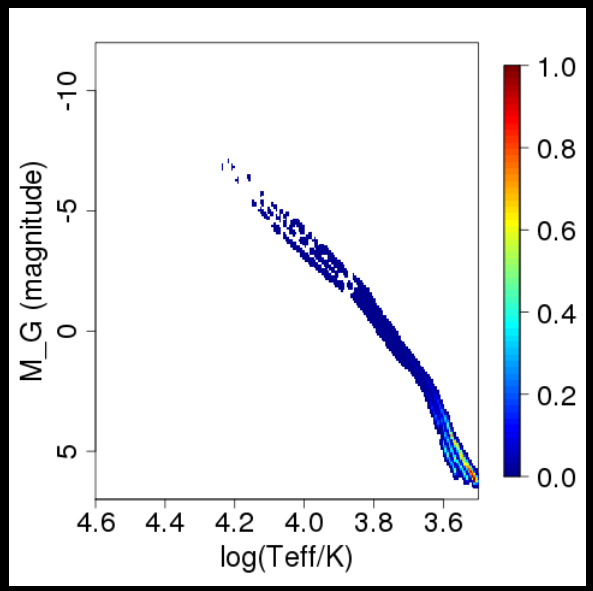
The synthetic library of spectra of binary stars









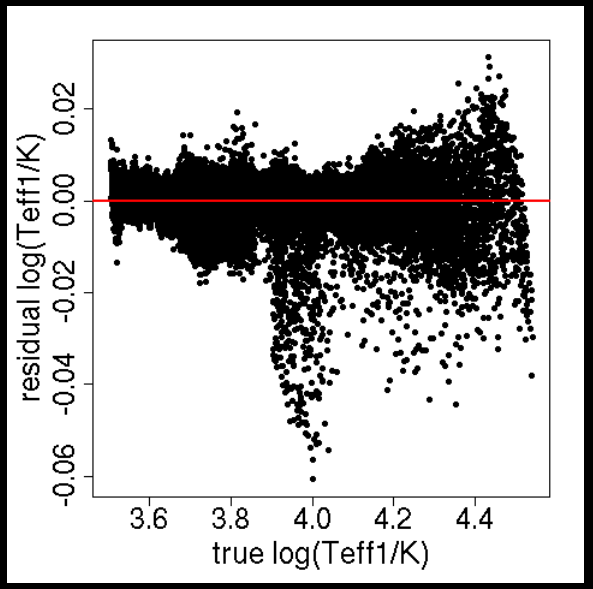


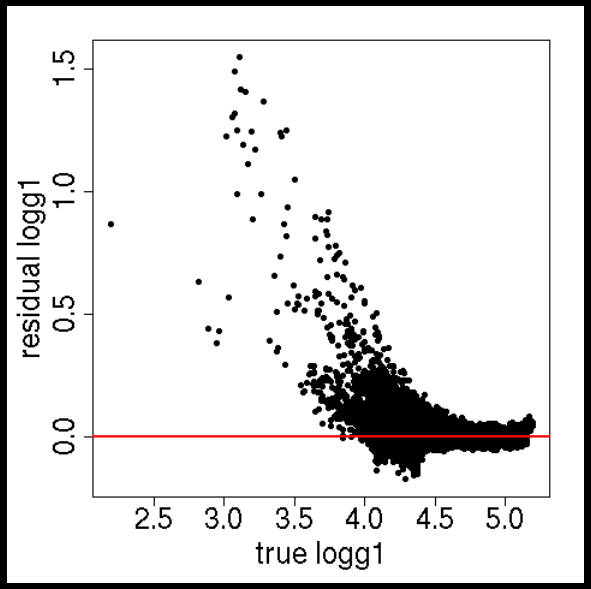
10^5 binary stars

red: training set
black: testing set

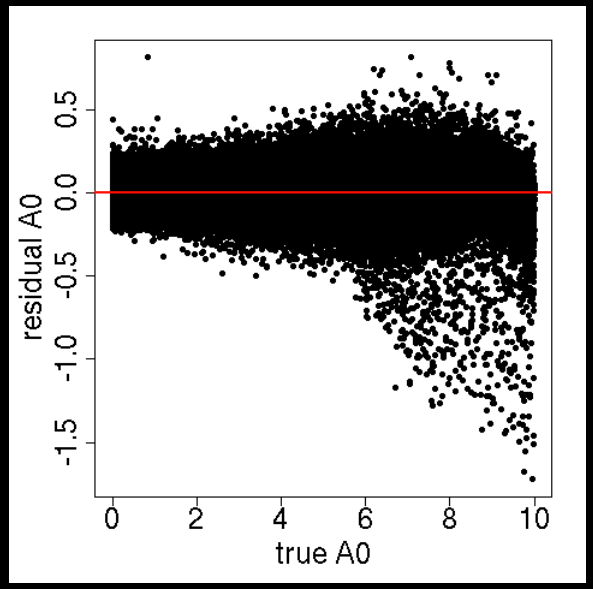
Method 1: Support Vector Machines

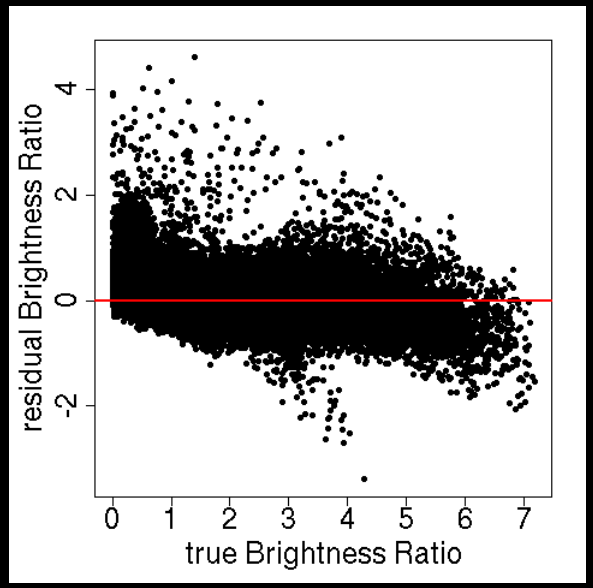
Results for G=15 and A_0=0

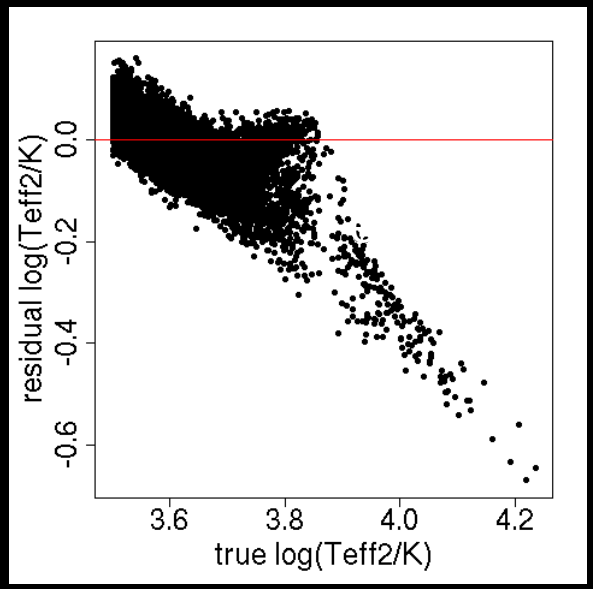


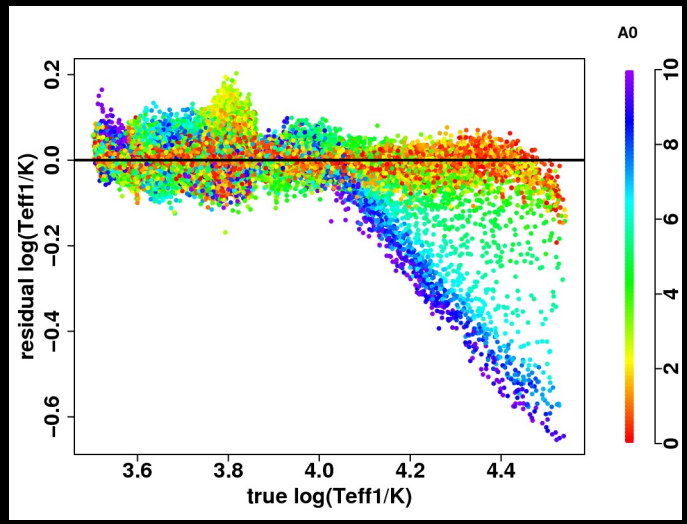


Results for G=15 and A_0>0



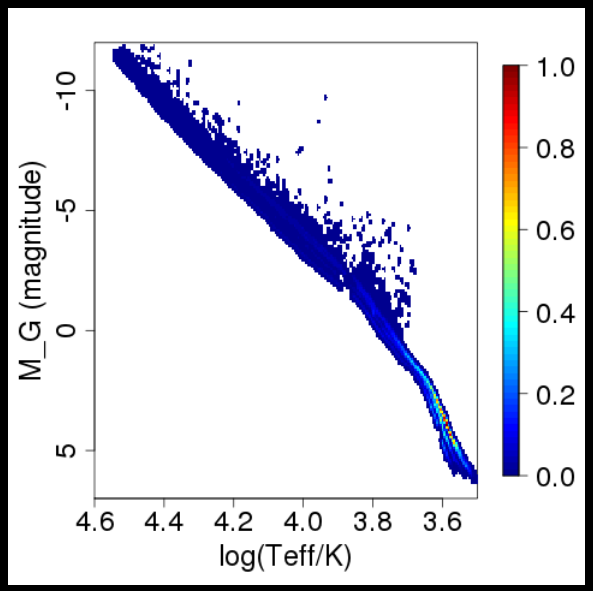




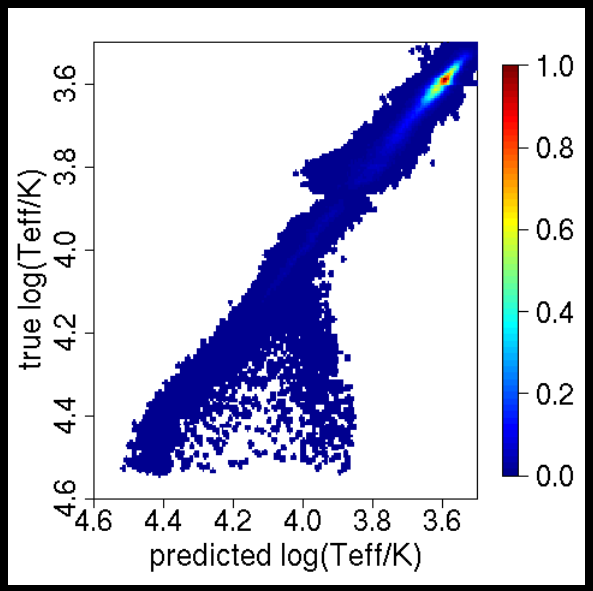


Method 2: SVM + HR prior

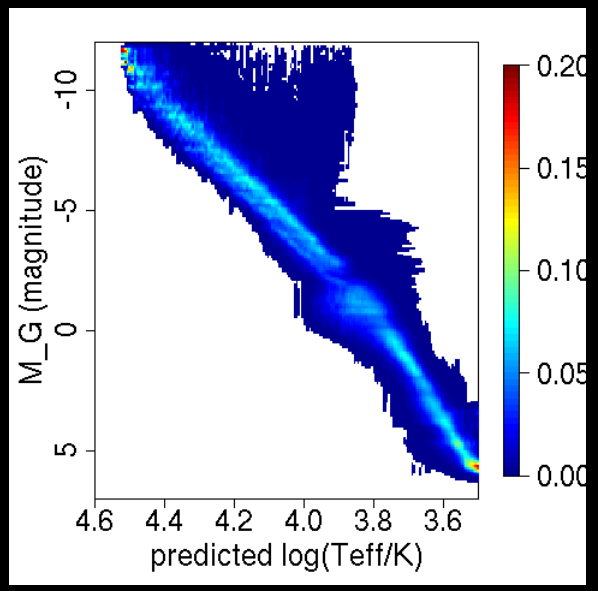
$$M_b = m_b + 5logp_b - A_b + 5$$



+



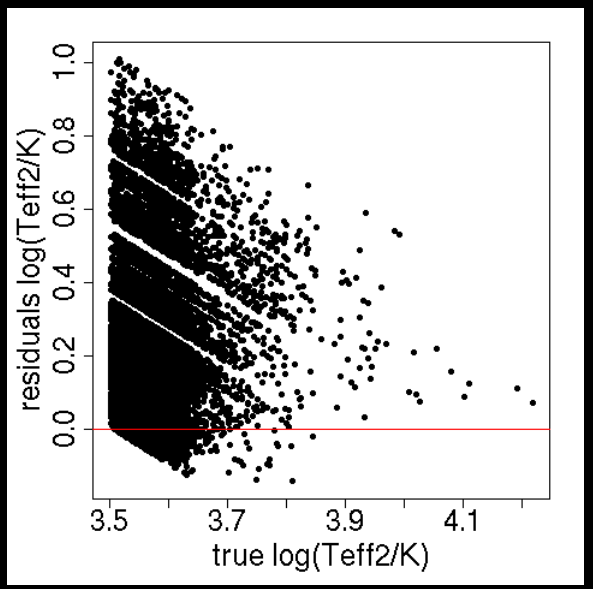
→



M_{b1}

$$M_b - M_{b1} = -2.5log(f_b/f_{b1})$$

$$M_{b2} = -2.5log(10^{-(M_b - M_{b1})/2.5} - 1) + M_{b1}$$

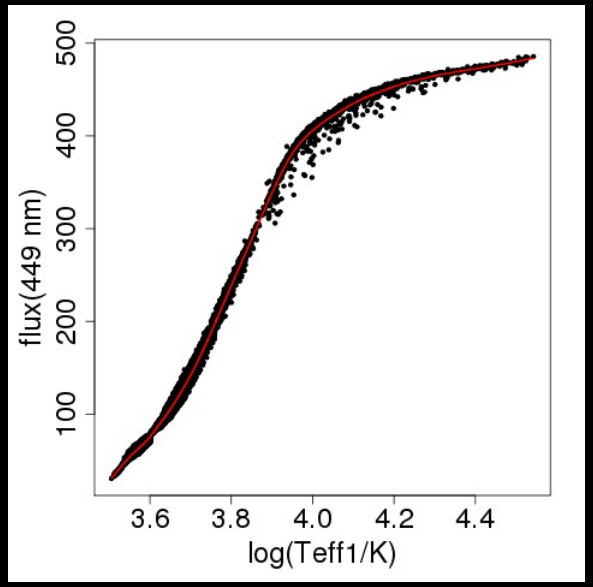


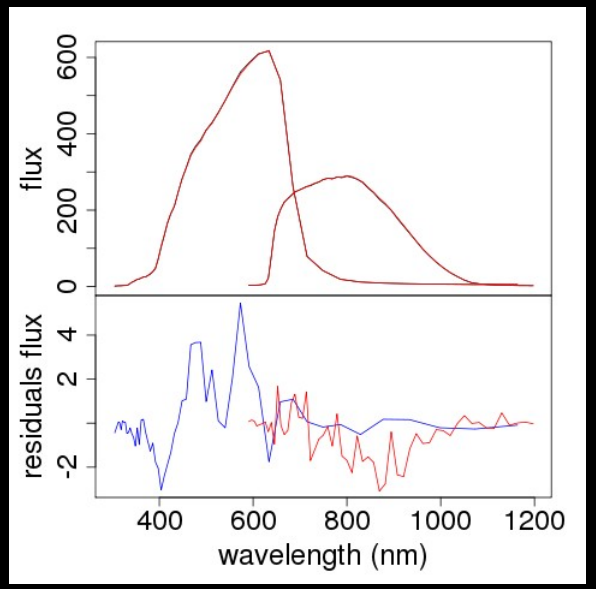
Method 3: q-method

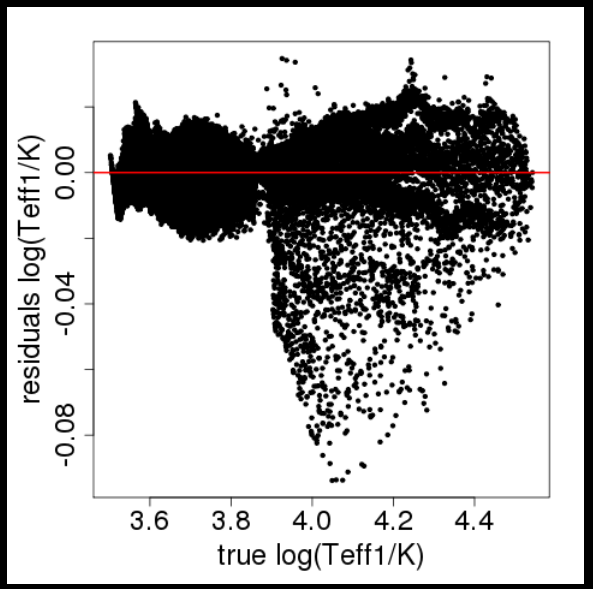
$$P(T_1, T_2|p) = \frac{P(p|T_1,T_2)}{P(p)} \int P(T_1, T_2, M_1, M_2) dM_1 dM_2$$

$$P(T_1, T_2|p) = \frac{P(p|T_1)}{P(p)} P(T_1, T_2)$$

$$P(p|T_1) = \frac{1}{\sqrt{2\pi|C_p|}} e^{-\frac{1}{2}\sum_{i=1}^{n} \frac{(p_{i,predicted} - p_{i,real})^2}{\sigma_i^2}}$$

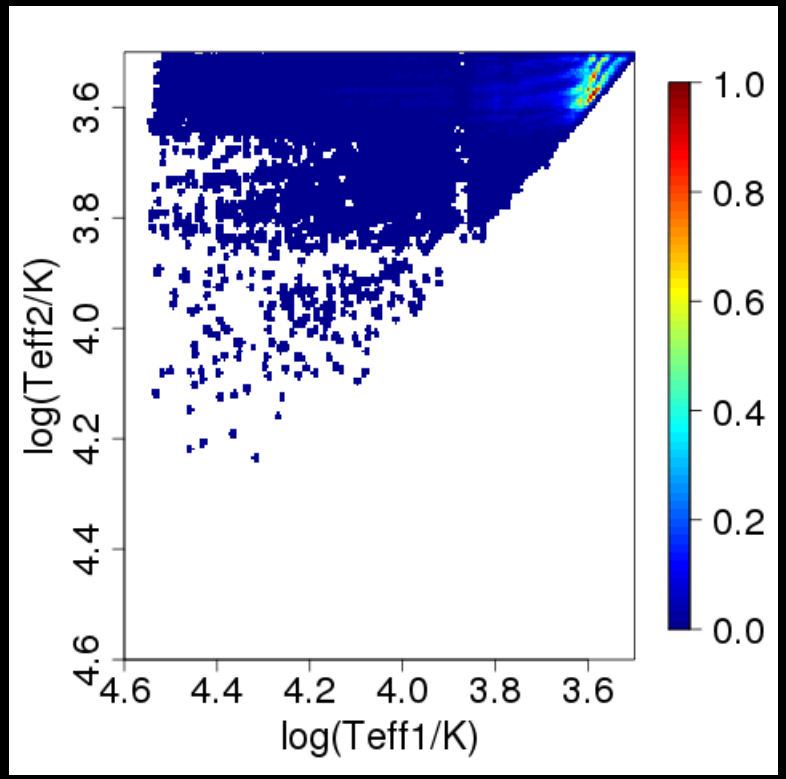


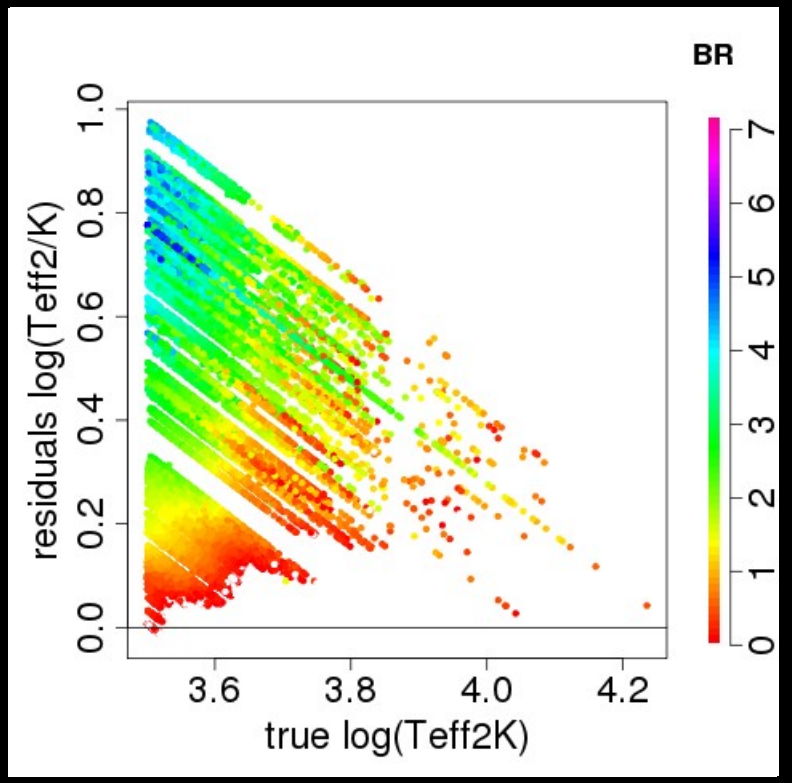




Method 3: q-method

$$P(T_1, T_2|p) = \frac{P(p|T_1)}{P(p)} P(T_1, T_2)$$





Conclusions

- Good performance of SVM in estimating parameters for the primary star (e.g. T_{eff1}, $logg_1$) and parameters for the binary system (e.g. brightness ratio and extinction).

- Good performance of the forward model in estimating $Teff_1$.

- Bad performance of all methods in estimating parameters for the secondary star. Problems in the methods or not enough information in the data?