Stellar Parameters in an Instant with Machine Learning

Earl P. Bellinger[1,2,3], George C. Angelou[1,2], Saskia Hekker[1,2], Sarbani Basu[4], Warrick H. Ball[5], and Elisabeth Guggenberger[1,2]
bellinger@mps.mpg.de

[1]Max-Planck-Institut für Sonnensystemforschung, Göttingen, Germany; [2]Stellar Astrophysics Centre, Aarhus University, Denmark;
[3]Institut für Informatik Göttingen, Germany; [4]Department of Astronomy, Yale University, USA; [5]Institut für Astrophysik Göttingen, Germany

Introduction

We use machine learning to build a **constrained multiple regression model** for rapidly estimating the fundamental stellar parameters of main-sequence solar-like stars [1]. We train a **random forest** of decision trees [2] with *scikit-learn* [3] on a matrix of stellar models generated with *MESA* [4] that we varied quasi-randomly in mass, initial helium and metallicity abundances, mixing length, the strength of convective overshooting, and the efficiency of gravitational settling. We additionally compute frequencies of each stellar model using *ADIPLS* [5] and summarize them to obtain averaged large and small frequency separations and frequency ratios [6].

We supply the global asteroseismic properties and other observable quantities of our stellar models to the random forest algorithm and produce a statistical model **relating observable quantities to fundamental stellar parameters** (Fig. 1). We validate this technique on a hare-and-hound exercise (Fig. 2) and the Sun (Fig. 3), and then apply it to 16 Cyg A & B (Fig. 4) and finally 34 planet-hosting candidates (Fig. 5).

Random forest regression

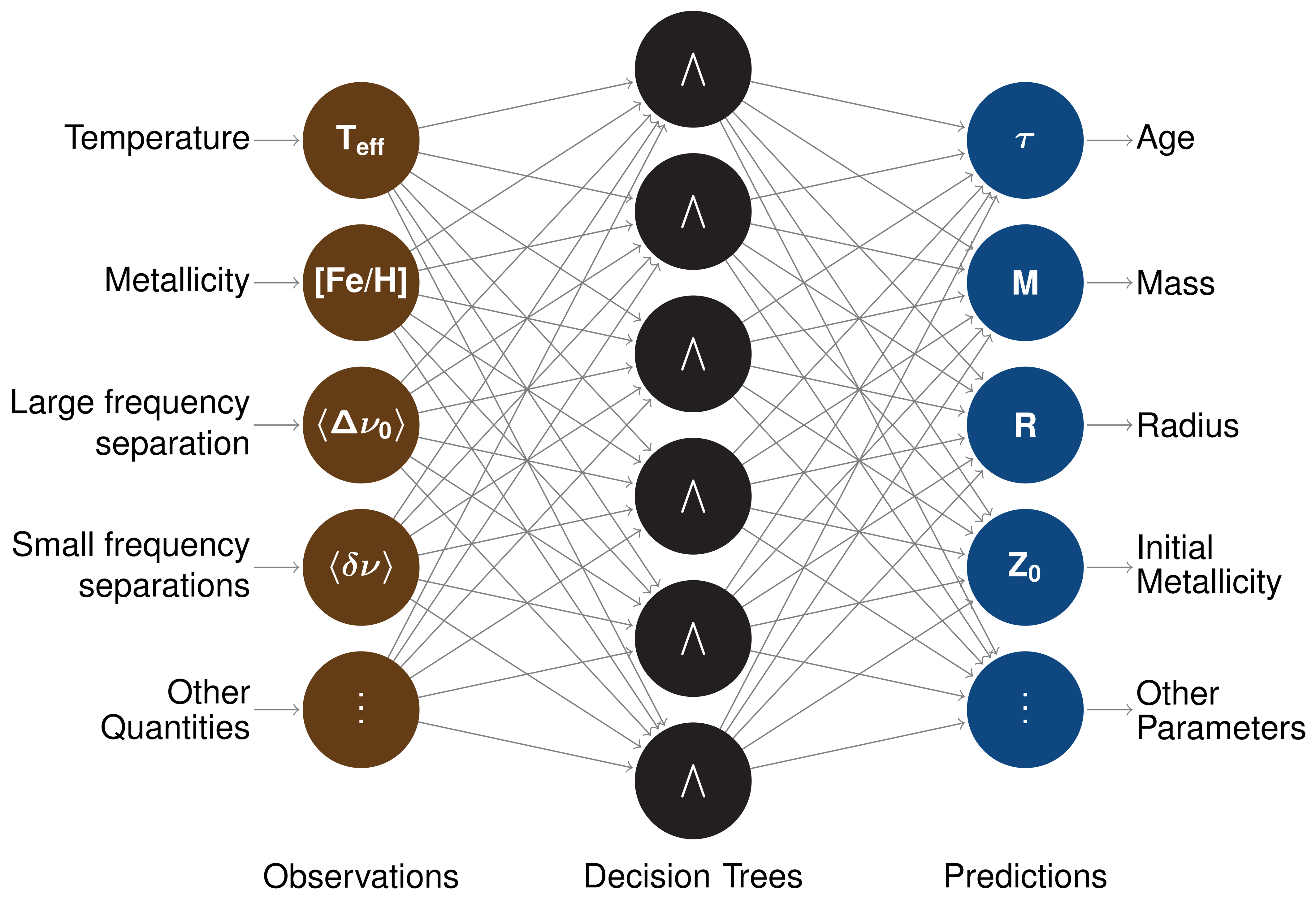

Fig. 1: A schematic representation of a random forest for inferring fundamental stellar parameters.

Hare-and-Hound

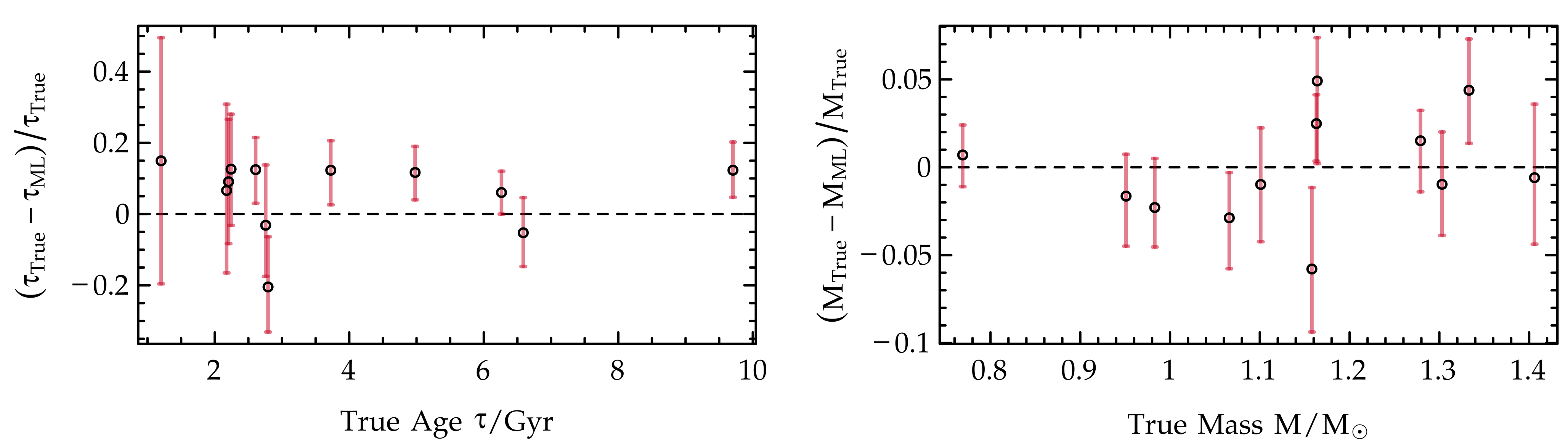

Fig. 2: Relative differences between predicted and true ages (left) and masses (right) for a blind hare-and-hound test.

The Sun

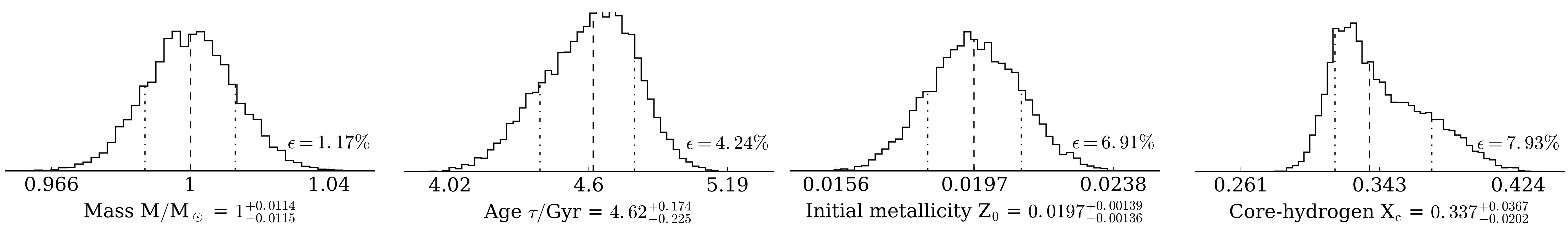

Fig. 3: Predictions from machine learning of stellar parameters for degraded BiSON [7] data of the Sun.

16 Cygni

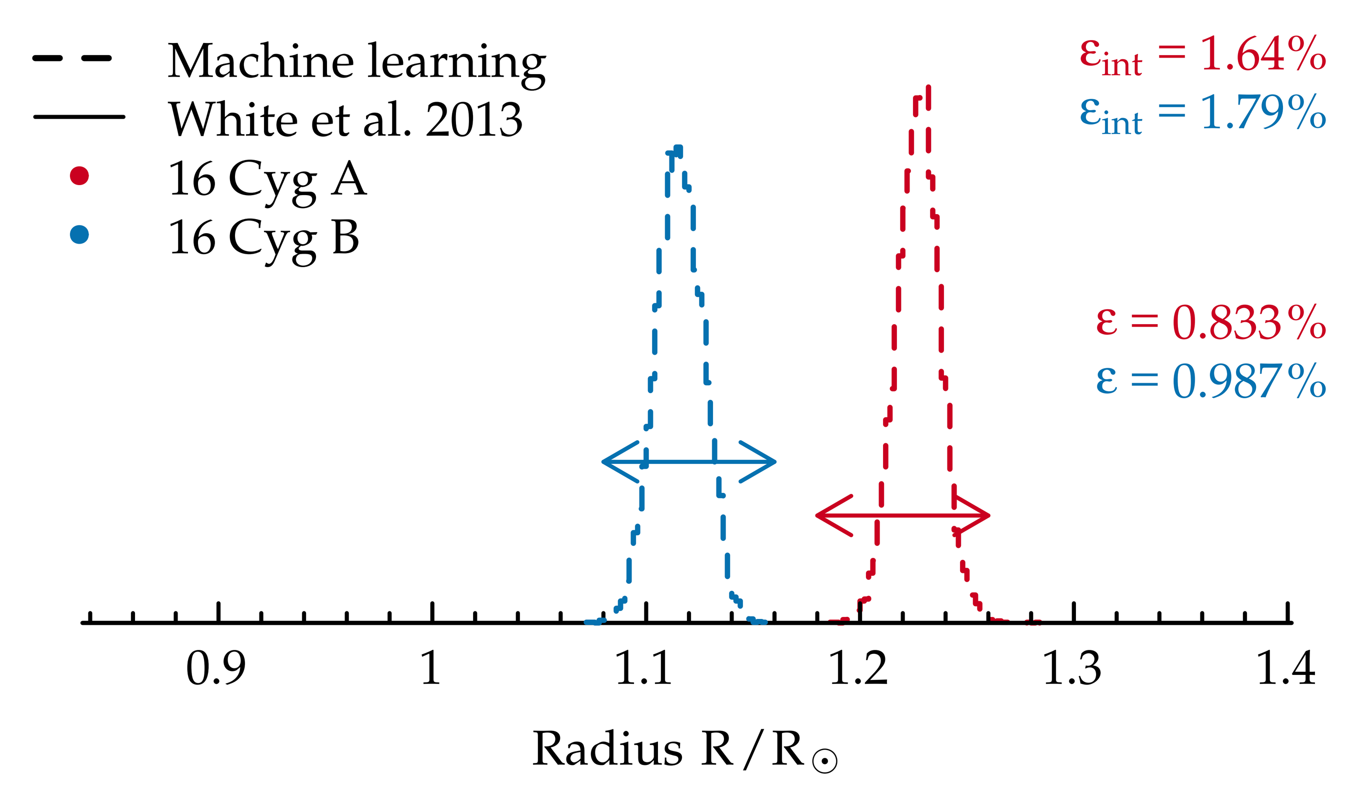

Fig. 4: Predictions from machine learning (- - -) on 16 Cyg A and B are in agreement with interferometric measurements (↔, 2σ) from White et al. [8].

Kepler objects-of-interest

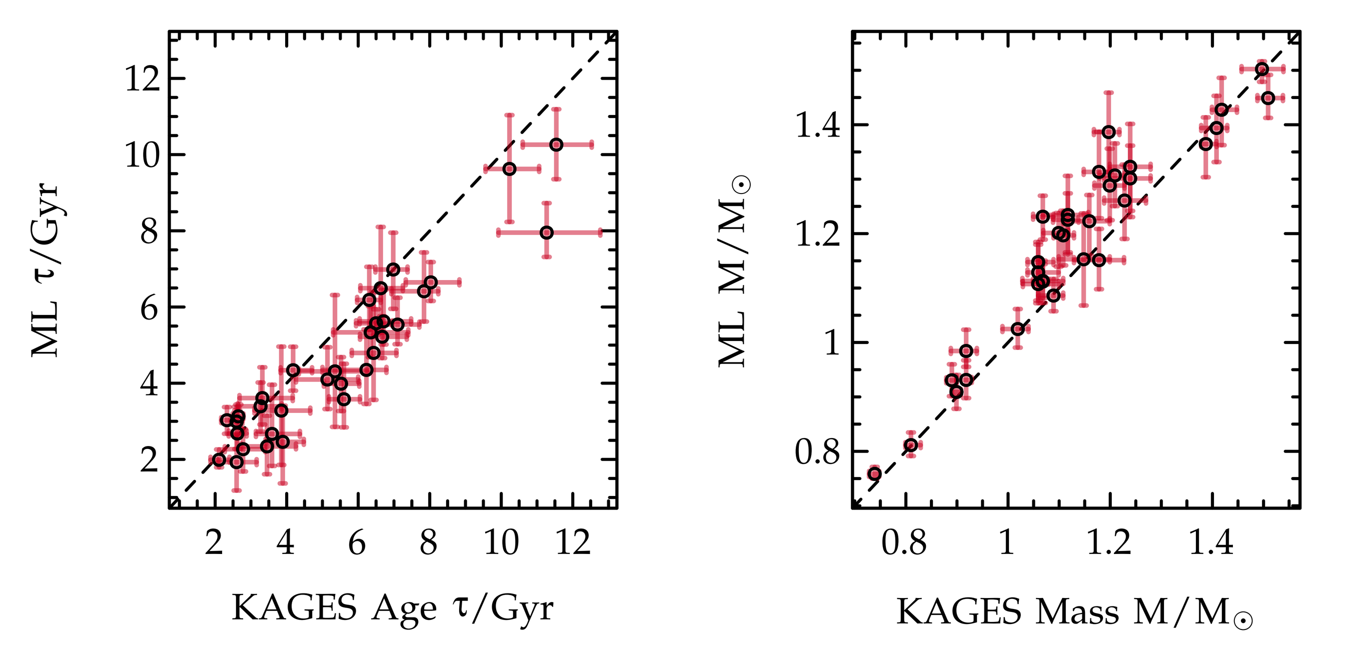

Fig. 5: Predictions from machine learning plotted against predictions from the *Kepler Ages* (KAGES, [9]) project on 34 planet-hosting candidates observed by *Kepler* [10].

Acknowledgements

The research leading to the presented results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no 338251 (StellarAges). This research was undertaken in the context of the International Max Planck Research School for Solar System Research. S.B. acknowledges partial support from NSF grant AST-1514676 and NASA grant NNX13AE70G. W.H.B. acknowledges research funding by Deutsche Forschungsgemeinschaft (DFG) under grant SFB 963/1 "Astrophysical flow instabilities and turbulence" (Project A18).

References

[1] Bellinger, E. P. & Angelou, G. C. et al.: *Fundamental Parameters of Main-Sequence Stars in an Instant with Machine Learning*, ApJ, accepted, arXiv:1607.02137.
[2] Breiman, L. 2001, Machine learning, 45, 5
[3] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
[4] Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3
[5] Christensen-Dalsgaard, J. 2008, APSS, 316, 113
[6] Roxburgh, I. W., & Vorontsov, S. V. 2003, AAP, 411, 215
[7] Davies, G. R. et al. 2014, MNRAS, 439, 2025
[8] White, T. R., Huber, D., Maestro, V., et al. 2013, MNRAS, 433, 1262
[9] Silva Aguirre, V., Davies, G. R., Basu, S., et al. 2015, MNRAS, 452, 2127
[10] Davies, G. R., Aguirre, V. S., Bedding, T. R., et al. 2016, MNRAS, 456, 2183