ODUSSEAS: a machine learning tool to derive effective temperature and metallicity for M dwarf stars

A. Antoniadis-Karnavas[1,2], S. G. Sousa[1], E. Delgado-Mena[1], N. C. Santos[1,2], G. D. C. Teixeira[1,2], V. Neves[3,4]

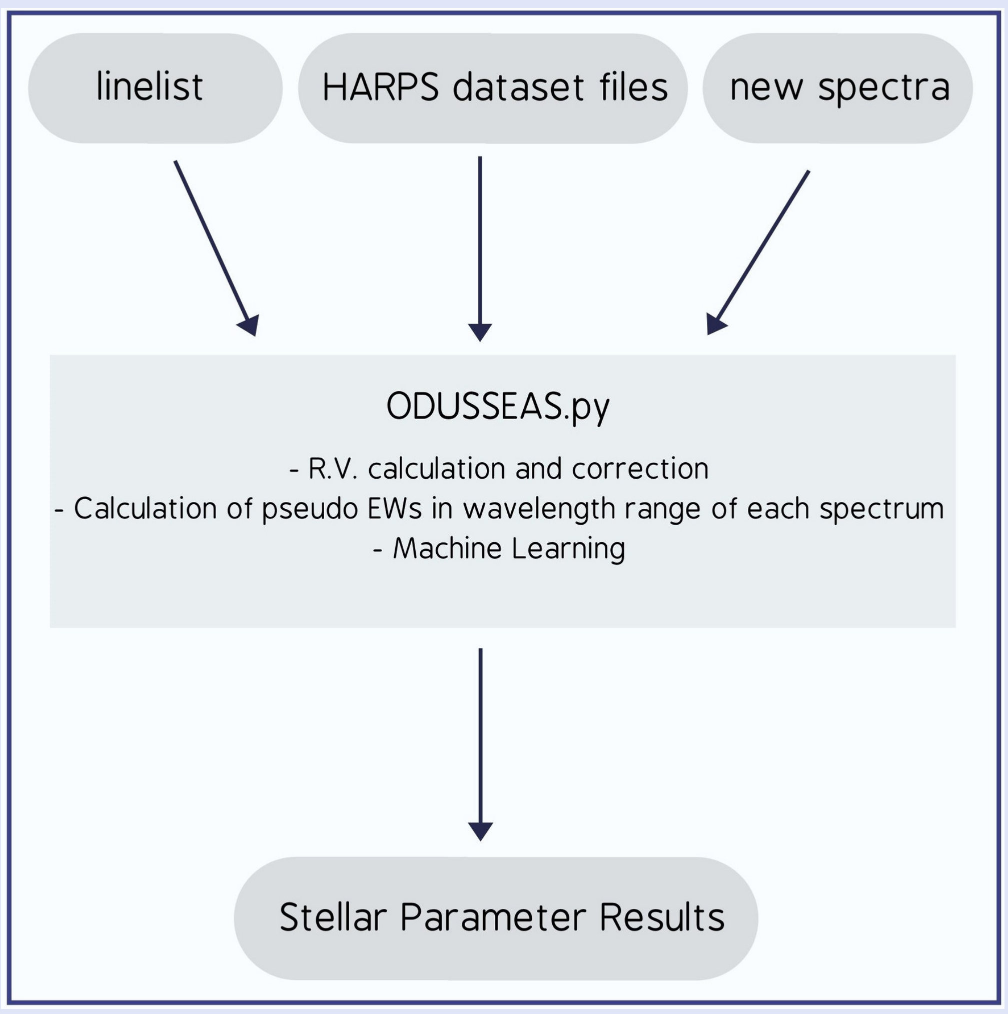
[1] Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal
[2] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 4169-007 Porto, Portugal
[3] Instituto Federal do Paraná, Campus Foz do Iguaçu, 85860000 Foz do Iguaçu-PR, Brazil
[4] Casimiro Montenegro Filho Astronomy Center, Itaipu Technological Park, 85867-900 Foz do Iguaçu-PR, Brazil
email: alexandros.antoniadis@astro.up.pt

Introduction

The derivation of spectroscopic parameters for M dwarf stars is very important in the fields of stellar and exoplanet characterization. Spectra can be used to reveal the chemical composition of the stars, as well as important stellar atmospheric parameters, such as effective temperature (Teff) and [Fe/H]. One of the most popular methods to derive stellar atmospheric parameters for FGK stars is by measuring the equivalent widths (EWs) of many metal lines of the spectrum. In our work, we follow the pseudo-EW approach. Since the identification of the continuum is very difficult in the spectra of M dwarfs, we set a pseudo continuum in each absorption feature. We measure the pseudo EWs of absorption lines and blended lines in the range between 530 and 690 nm.

ODUSSEAS Tool

- **Purpose**: Creation of an automatic, easy-to-use, computational tool being able to derive Teff and [Fe/H] reliably for M dwarfs with optical spectra of different resolutions.
- **Method**: Based on pseudo-EW measurements of more than 4000 stellar absorption lines and on the machine learning Python package “scikit-learn”.
- **Input to train & test**: The training input to the machine learning algorithm are the pseudo-EW measurements of 65 HARPS spectra, as convolved to the respective resolutions of the new stars, and their reference Teff and [Fe/H] values as expected output, for training and testing the models by applying ridge regression.
- **Input to analyze**: New 1D spectra and their resolutions. For each new star, the tool makes 100 determinations by randomly shuffling and splitting the training groups (~70% of the sample: 45 stars) and the testing groups (~30% of the sample: 20 stars).
- **Output**: The average values of Teff and [Fe/H], their dispersion, the average values of the model errors, the total error budget taking into consideration the intrinsic uncertainties of the reference parameters during the process, and the average "r2" and "explained variance" machine learning scores.



Results

- **Successful application on several spectrographs**: HARPS (R=115000), UVES (R=110000), CARMENES (R=94600), SOPHIE (R=75000), FEROS (R=48000).
- **High machine learning efficiency**: "r2" and "explained variance" scores are up to ~0.94 (with best possible value being 1.0).
- **Accurate determinations**: Differences are up to ~50 K for Teff and up to ~0.03 dex for [Fe/H] from the expected values.
- **Precise determinations**: The mean absolute errors of machine learning models are ~30 K for Teff and ~0.04 dex for [Fe/H].
- **Total uncertainties**: We perturbed the reference parameters by injecting their intrinsic uncertainties (100 K for Teff and 0.17 dex for [Fe/H]) into the training datasets with Gaussian distributions. In this case, our models have maximum uncertainties of ~80 K for Teff and ~0.13 dex for [Fe/H]. Still, our parameter determinations have consistent values with differences up to ~50 K and ~0.03 dex from the expected ones.
- **Validation of [Fe/H] determinations in binaries**: We determined [Fe/H] in eight FGK+M systems and in five M+M systems. Regarding the FGK stars, their [Fe/H] and respective uncertainties were derived using the methodology described in Sousa et al. (2008) and Santos et al. (2013). All binaries show differences within the uncertainties of the two methods.
- **Computational time**: The derivation of Teff and [Fe/H] by ODUSSEAS requires a few seconds per star.

Discussion

- **Machine learning VS traditional approach**: We compared our results with the ones by Neves et al. (2014), a work applying a least-squares weighted fit to determine parameters. Method by Neves et al. (2014) has precision errors of 91 K for Teff and 0.08 dex for [Fe/H], while the regression of our tool reduces these errors to 27 K and 0.04 dex respectively, both using the same HARPS dataset as reference scale.
- **Reference parameters**: Since supervised machine learning is trained to determine parameters based on given reference values, their systematic errors will apply to the results of new stars too. In this work, we have used the reference Teff and [Fe/H] photometric scales by Casagrande et al. (2008) and Neves et al. (2012) respectively, as they are derived in a homogeneous way for a sufficiently large number of spectra available to us.
- **Application limits**: Spectra should have S/N > 20 for optimal results. The tool is not applicable to very active or young stars.
- **Article**: The detailed description of method and analysis of results can be found in Antoniadis-Karnavas et al. (2020).
- **Available code**: https://github.com/AlexandrosAntoniadis/ODUSSEAS

References

Antoniadis-Karnavas, A., Sousa, S. G., Delgado-Mena, E., et al. 2020, A&A, 636, A9
Casagrande, L., Flynn, C., & Bessell, M. 2008, MNRAS, 389, 585
Neves, V., Bonfils, X., Santos, N. C., et al. 2012, A&A, 538, A25
Neves, V., Bonfils, X., Santos, N. C., et al. 2014, A&A, 568, A121
Santos, N. C., Sousa, S. G., Mortier, A., et al. 2013, A&A, 556, A150
Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373