Boletim da Sociedade Astronômica Brasileira, **37**, no. 1, 217-218

Multilevel machine learning models as a tool to understand the relationship between planetary systems and life-supporting planets

L. Bernardes[1]

[1] Centro Universitário Estácio de Ribeirão Preto e-mail:luander.uspicio@gmail.com

Abstract. Since the discovery of the first exoplanet to orbit a Main Sequence star, a huge effort has been made by the scientific community to try to understand some basic questions related to the existence of these new worlds. However, there are still gaps to be investigated, and one of them concerns the contribution of the physical-chemical conditions of planetary niches to the formation of Earth-like planets. Therefore, the research determined the influence of these niches on the formation of planets capable of maintaining the minimum conditions for the emergence and evolution of life, considering Earth as a model. To this end, several planetary systems were studied using unsupervised and supervised Machine Learning techniques with the aim of identifying the formation of clusters and investigating the multilevel relationship between planets and extrasolar niches. The work demonstrates that, probably, a wide variety of types of exoplanets can harbor atmospheres capable of being studied remotely, although these results do not consider the actual internal constitutions of the objects of study, since they are unknown, which prevents a historical reconstruction of the evolution process of these planets. The two-level multilevel approach quantitatively demonstrates that approximately 59% of the variation in the degree of similarity of exoplanets to Earth is due to the effect of the conditions of the planetary niches where they are located.

Resumo. Desde a descoberta do primeiro exoplaneta a orbitar uma estrela da Sequência Principal, um enorme esforço vem sendo realizado pela comunidade científica para tentar entender algumas questões básicas relacionadas à existência destes novos mundos. Entretanto, ainda há lacunas a serem investigadas e uma delas diz respeito à contribuição das condições físico-químicas de nichos planetários para a formação de planetas semelhantes à Terra. Sendo assim, a pesquisa determinou a influência desses nichos na formação de planetas aptos de manter as condições mínimas para o surgimento e a evolução da vida, considerando a Terra como modelo. Para isso, vários sistemas planetários foram estudados por meio de técnicas não supervisionadas e supervisionadas de Machine Learning com o objetivo de identificar a formação de clusters e investigar a relação multinível entre planetas e nichos extrassolares. O trabalho demonstra que, provavelmente, uma ampla variedade de tipos de exoplanetas pode abrigar atmosferas aptas a serem estudadas remotamente, apesar de que esses resultados não consideram as constituições internas reais dos objetos de estudo, já que elas são desconhecidas, o que impede uma reconstrução histórica do processo de evolução desses planetas. A abordagem multinível de dois níveis demonstra quantitativamente que aproximadamente 59% da variação do grau de semelhança de exoplanetas com a Terra se deve ao efeito das condições dos nichos planetários onde estão localizados.

Keywords. Astrobiology – Planetary Systems – Exoplanet

1. Introduction

There are still doubts about the role played by the physical and chemical conditions of planetary systems in the formation of planets capable of harboring life. Thus, the objective of this research was to quantitatively determine, through Multilevel Modeling, the influence of the characteristics of certain planetary clusters on the possibility of the existence of planets that could sustain the minimum conditions for the emergence and evolution of life. In this approach, the terrestrial environment was considered as a model and the Earth Similarity Index (Schulze-Makuch et al., 2011) as a parameter to estimate the degree of similarity of an exoplanet to Earth.

Multilevel modeling allowed the construction of a basic structure in which exoplanets were nested within planetary clusters, leading to the creation of a two-level model. This approach is advantageous compared to those using classical linear regression models, as it allows for the analysis of data from a hierarchical perspective, considering dependencies between observations belonging to the same group. This makes it possible to capture the behavior of variables studied at each of the proposed levels (Fávero and Belfiore, 2017). The hierarchy proposed here considered exoplanets as level 1 and planetary clusters as level 2 (see Figure 1).

2. Methodology

The data used in the multilevel modeling were linked to attributes of more than a hundred exoplanets and their host stars such as radius, masses, orbital period, surface temperature, luminosity, metallicity, etc. The chosen targets were Sub-Earth, Earth, Super-Earth, and Mini-Neptune planets orbiting stars of spectral types M, K, and G, as these stars can provide environments where at least extremophile life could develop. The structuring of the data in levels was used, as it allowed to discuss certain correlations based on a hierarchical perspective. Following the Step-up Strategy methodology (Raudenbush and Bryck, 2002), several models were created from clustering (Euclidean distance/complete chaining), as well as with the gradual increase in complexity, achieved from the insertion of random effects of intercepts, slope, explanatory and contextual variables, comparing their efficiencies through, for example, the Log-Likelihood gain (see Figure 2).

3. Results

Modeling through Machine Learning techniques was able to estimate the intraclass correlation index (approximately 59%), indicating the contribution of planetary clusters in the formation of Earth-like planets.

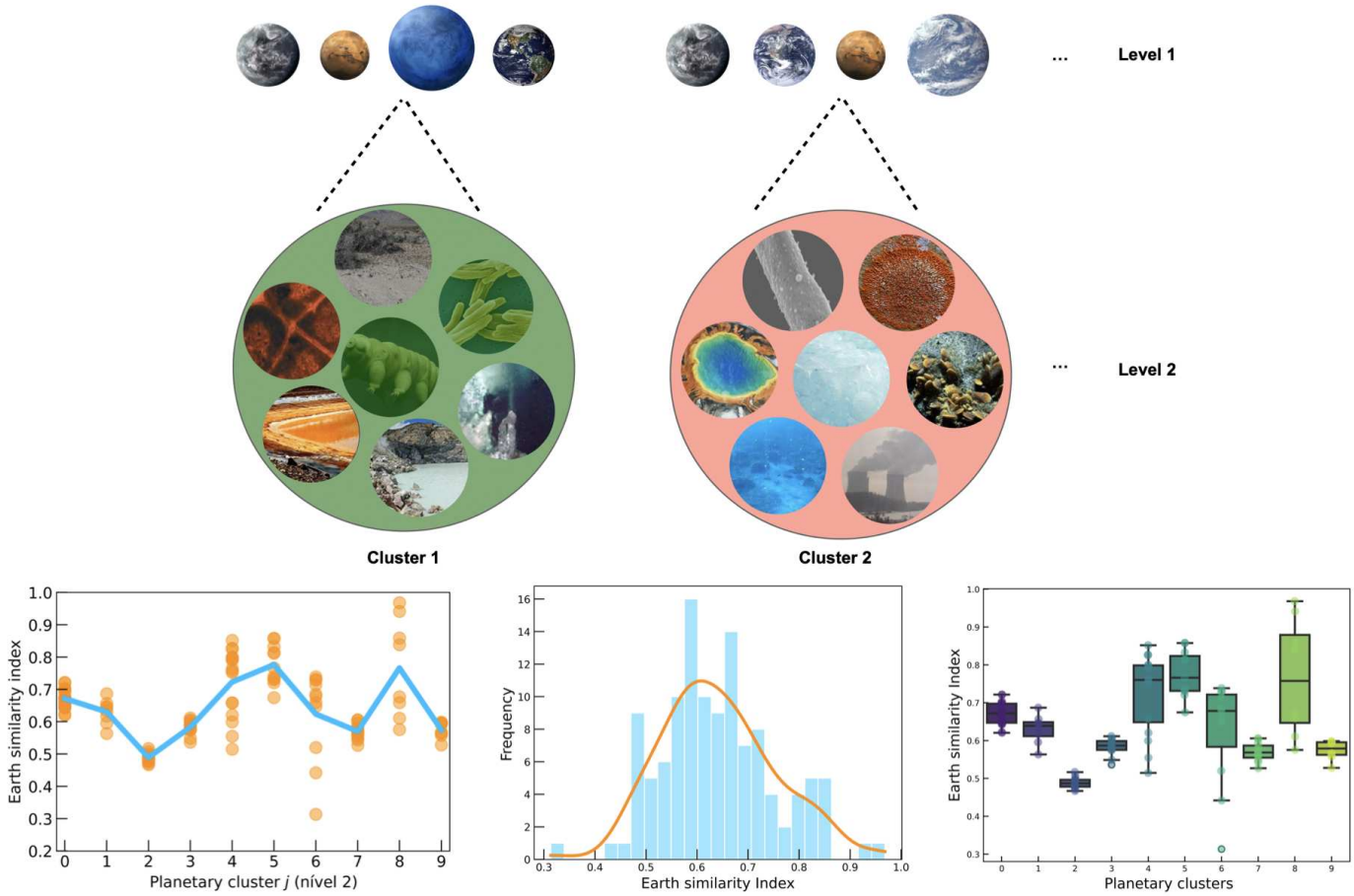

Figure 1. Two-level hierarchical model (HLM2) with planets nested within clusters containing planetary systems. The distribution of the degree of similarity of exoplanets to Earth is shown for each cluster.

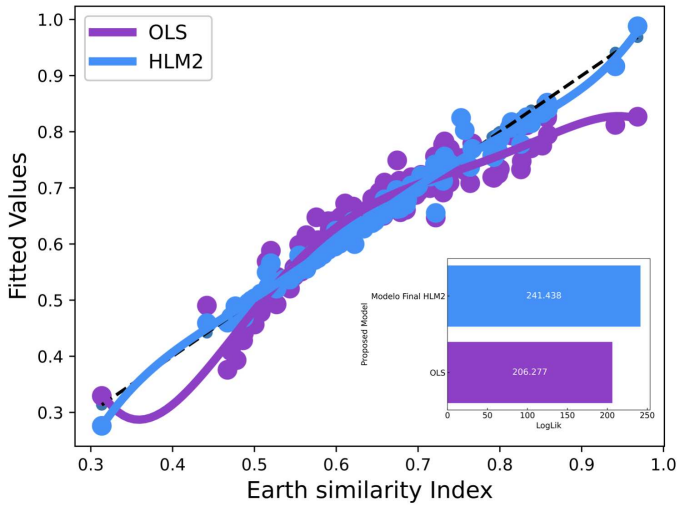

Figure 2. The final HLM2 model and its comparison with traditional OLS-type models.

Furthermore, several characteristics of planetary clusters most likely to provide optimized conditions for the existence of planets that could support life were highlighted. It was possible to create a quantitative and observational basis for presenting a mathematical equation capable of predicting the degree of similarity of an exoplanet to Earth. The model proves superior when compared to traditional Ordinary Least Squares - OLS type models (see Figure 2).

4. Conclusion

Even considering the limitations of the model, stemming from the small amount of data available, the work was able to quantitatively estimate the influence of planetary niches on the formation of Earth-like planets and offer a predictive mathematical equation for this similarity, serving as a tool for choosing promising targets in the search for signs of extraterrestrial life.

The next steps of the research aim to analyze a more robust database, as well as consider other hierarchical levels, such as the temporal dimension not addressed in the research.

References

Fávero, L. P., & Belfiore, P., 2017, Elsevier.
Raudenbush, S. W., & Bryk, A. S., 2002, CA: Sage Publications.
Schulze-Makuch, D. et al., 2011, Astrobiology, 11, 1041.