ExoPSI: A Planet Similarity Python toolkit

Vaibhav Garg  
Delhi Technological University

Aditya Rai  
Delhi Technological University

Divya Srinivasan  
Delhi Technological University

A. S. Rao (✉ drsrallam@dce.ac.in)  
Delhi Technological University

Research Article

Keywords: Astronomy software (1855), Exoplanets (498), Habitable planets (695), Publicly available software (1864), Astronomy data analysis (1858)

Posted Date: April 12th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2797683/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
ExoPSI: A Planet Similarity Python toolkit

Vaibhav Garg††, Aditya Rai††, Divya Srinivasan††, A. S. Rao†*

1†Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Rd, 110042, Delhi, India.

*Corresponding author(s). E-mail(s): drsrallam@dee.ac.in; Contributing authors: vaibhavgarg2022@gmail.com; rai.aditya01@gmail.com; divyasrini.2401@gmail.com; †These authors contributed equally to this work.

Abstract
The Earth Similarity Index (ESI) has proven to be an essential tool in locating Earth-like planets throughout the universe, yet its limited scope restricts its full potential. ExoPSI is a groundbreaking, open-source Python package that elevates the concept of similarity indexing. It provides the capability to calculate similarity indices of different planetary candidates with respect to any reference values, based on any numerical parameter, yielding a more accurate evaluation of the survivability of different species. ExoPSI offers unparalleled versatility, with the ability to use any planet as a reference point and visualisation functions that allow for easy comparison of the similarity between the interior and surface parameters of different planets. This makes it a critical tool for planetary science researchers, offering a comprehensive solution for evaluating planetary similarities.

Keywords: Astronomy software (1855) — Exoplanets (498) — Habitable planets (695) — Publicly available software (1864) — Astronomy data analysis (1858)

1 Introduction
A two-tiered classification system was created by Schulze-Makuch et al [18] to evaluate the habitability of exoplanets. The ESI was used to measure how similar exoplanets were to Earth, the only known habitable planet, in the first tier of the categorisation system. Based on currently available knowledge of the mass, radius, and temperature of exoplanets, the ESI was determined. They suggested the Planetary Habitability Index (PHI) for the next tier, which took into account elements such as stable substrate,
readily obtainable energy, acceptable chemical composition, as well as the ability to hold a liquid solution. This index was designed to account for the possibility of life in unusual environments and required more detailed information than was present for exoplanets at that time.

The ESI was developed to show the degree of resemblance between an exoplanet and Earth, and it is a critical aspect in determining whether or not an exoplanet is habitable [18]. The value ranges from 0 (no similarity) to 1 (the similarity of Earth). Typically, a planetary body is regarded as Earth-like if its ESI is more than 0.8. The temperature of the surface, the planet density, the planet’s escape velocity and its radius are used as the four variables in the ESI calculation. It is split into two parts, interior ESI$_i$ (calculated from the planet’s radius and its density) and surface ESI$_s$ (calculated from the planet’s escape velocity and the surface temperature). The ESI formula is as follows:

$$ESI_x = \left(1 - \frac{x - x_0}{x + x_0}\right)^w,$$

where ESI$_x$ is the calculated value of the ESI for that planet for the property $x$, $x_0$ is the value of that particular property with regards to Earth, and $w$ is the weighting component that adjusts the sensitivity of the scale.

As a potential future calculator for the possibility of life, Kashyap Jagadeesh et al [11] proposed a ‘Life Information Score’. They advocated for the creation of a multiparameter calculator that is based on the existing ESI scale and takes more input parameters into account, including those related to orbital characteristics, temperature, escape velocity, radius, density, activation energy, and others. They suggested incorporating other planet-like objects into this scale, such as the Mars Similarity Index (MSI). In order to determine the chance of life on exoplanets, subsequent studies modified the ESI model to produce various iterations of the same similarity scale.

After analysing these previous studies, we concluded that the biological survivability limits of organisms on Earth (and possibly on other planets in the future) could be used to modify the traditional ESI formula and create a suitable similarity scale for analysis. Therefore, we introduce ExoPSI, an open-source Python library that automates this process of modification. It provides a universal calculator that can potentially use any numerical parameter to determine a planet’s similarity to any other planet, with the ability to use any planet as a reference point. Additionally, ExoPSI includes visualisation functions that facilitate easy comparison of the interior and surface parameters of different planets.

Python’s strength lies in being a highly adaptable high-level language that is simple to program and integrates both conventional programming and capabilities for data reduction and analysis. Astronomy and astrophysics are only two of the many scientific and technological fields where the Python programming language has found many uses. Python’s user-friendliness was leveraged to develop IRAF Python command line, which is now commonly referred to as PyRAF [7] and the PyFITS module for manipulating, writing and reading FITS files [4] at the Space Telescope Science Institution (STScI) as tools for astronomical study in the early 2000s. Python’s usefulness as an application language has been aided by significant advancements in the tools that make it practical to use it to process astronomical data successfully. A wide array of packages and libraries have since been developed such as astroquery, SunPy, RadFil,
PySAP, ChiantiPy [8–10, 22] with the most popular being Astropy, a community developed and open-source Python library with core astronomy related functionality [21]. Based on this scalability of the Python language, we created ExoPSI.

2 Motivation

ExoPSI offers a more straightforward method for calculating similarity indices between exoplanets. It reduces the programming required for the user to compute similarity. The ESI is widely used to evaluate the habitability of exoplanets as the likelihood of the existence of other Earth-like planets [3, 17, 18]. ExoPSI’s main advantage is that it can calculate the similarity index between any two planets and not just provide the planet’s similarity with Earth. This proves helpful in identifying similarities and dissimilarities between planets that have their index values in a close range with respect to Earth. An example of this is the case of Mars in relation to Gliese 581 c, both of which have the same dissimilarity from Earth but do not have any shared characteristics [18]. The similarity of these two planets can be realised using ExoPSI. It also can convert the units used in terms of the specified planet. This allows us to use relative planetary properties to calculate the index, such as evaluating the similarity of exoplanets with respect to Jupiter, i.e., in terms of Jupiter Units. The addition of other parameters for assessing habitability and the study of finding the essential features for the existence of life is another application of similarity indices, which can be facilitated by ExoPSI as done by Chandra et al [6]. Other planetary properties such as revolution, surface gravity or eccentricity can also be included in studying the similarity of exoplanets. It can also evaluate flexible threshold values and their weights based on the higher and lower limits of sustaining different life forms. A study on rock-dependent extremophiles Acarospora and Chroococcidiopsis found that the weight of temperature parameter in the similarity index shifted to 2.26 for a 0.8 threshold similarity to support their existence in the temperature range 258 K to 395 K [12]. The concept of similarity index can also be used to identify planets for further study, as was done by [15] to infer the CME Magnetic Field Magnitude in Sun and Geospace where they studied the hypothesised stellar CMEs on Kepler 438b and Proxima b with high ESI. Comparative analysis of different methods of characterising exoplanets and their habitability and using this comparison to build a new index presents another field where ExoPSI can be of use, as seen in the comparison of Cobb-Douglas Habitability Score (CDHS) with ESI by Agrawal et al [2] and the development of the Statistical-likelihood Exo-Planetary Habitability Index (SEPHI) by Rodríguez-Mozos and Moya [16]. Considering the versatile use of ESI and understanding its limitations, ExoPSI was created.

3 Usage and Functionality

The class objects and their methods serve as the foundation for the ExoPSI library architecture. All methods and properties in ExoPSI accessible to the user are available in the exopsi class. The user can access functions such as calculating the similarity index, unit conversion and plotting similarity index graphs by invoking the exopsi class and creating an object. The exopsi class contains multiple methods which require
certain arguments to be passed. The readme.md file clearly defines which arguments are necessary for the function to work and which are optional. Each method in the exopsi class returns a certain output. This output can be either in the form of a Pandas Dataframe, figure object or a variable. The user can store these outputs in a variable or as a file for subsequent use. The library can easily be installed using pip.

ExoPSI can only handle datasets of the type pandas.core.frame.DataFrame. As a result, any dataset to be used has to be imported as a Pandas Dataframe. The dataset used in Appendix A is the PHL’s Exoplanet Catalog (PHL-EC) of the Planetary Habitability Laboratory @ UPR Arecibo (Version 20191205)\(^1\). The first 10 rows of this dataset, for the properties planet name, radius, density, escape velocity and surface temperature are shown in Table 1.

<table>
<thead>
<tr>
<th>Name</th>
<th>Radius (EU)</th>
<th>Density (EU)</th>
<th>Escape Velocity (EU)</th>
<th>Surface Temperature (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 Cnc e</td>
<td>1.87207</td>
<td>1.2178426</td>
<td>2.0659399</td>
<td>2066.9461</td>
</tr>
<tr>
<td>61 Vir b</td>
<td>2.10748</td>
<td>0.54327754</td>
<td>1.5533679</td>
<td>1214.7088</td>
</tr>
<tr>
<td>BD-06 1339 b</td>
<td>2.84734</td>
<td>0.37173883</td>
<td>1.7360343</td>
<td>772.23385</td>
</tr>
<tr>
<td>CD Cet b</td>
<td>1.81602</td>
<td>0.65963152</td>
<td>1.4749297</td>
<td>492.19235</td>
</tr>
<tr>
<td>CoRoT-7 b</td>
<td>1.6815</td>
<td>0.85768575</td>
<td>1.5572294</td>
<td>1867.6349</td>
</tr>
<tr>
<td>CoRoT-7 c</td>
<td>2.83613</td>
<td>0.368223</td>
<td>1.7210029</td>
<td>1135.9045</td>
</tr>
<tr>
<td>DMPP-1 c</td>
<td>3.06033</td>
<td>0.33488394</td>
<td>1.7709872</td>
<td>1277.4104</td>
</tr>
<tr>
<td>DMPP-1 d</td>
<td>1.64787</td>
<td>0.74862431</td>
<td>1.4257879</td>
<td>1683.3767</td>
</tr>
<tr>
<td>DMPP-1 e</td>
<td>1.86086</td>
<td>0.64070808</td>
<td>1.4895113</td>
<td>1335.3362</td>
</tr>
<tr>
<td>DMPP-3 A b</td>
<td>1.41246</td>
<td>0.91584022</td>
<td>1.3517177</td>
<td>943.53264</td>
</tr>
</tbody>
</table>

The PHL-EC dataset has been cleaned to drop rows where the data is missing for any of the desired columns.

### 3.1 Weight Calculation

The weights of individual parameters can be computed efficiently by providing the reference value, the lower and upper limits for any parameter where \(x_{\text{lower}} \leq x_{\text{ref}} \leq x_{\text{upper}}\). Optionally, a threshold value of the similarity index (default = 0.8) can also be provided. The calc_weight function takes in \ref_val, upper_lim, lower_lim and threshold (optional) for the reference values, upper limits, lower limits and threshold value, respectively. The code in Listing 1 uses the traditional values for planetary radius, density and escape velocity in Earth Units (EU) and temperature in Kelvin (K).

The calc_weight function gives the output: The calculated weight(s) is(are): [0.57, 1.07, 0.7, 5.58]. It can be observed that these results are in coherence with the results obtained by Schulze-Makuch et al [18] and serve as a validation test for this function.

---

3.2 PSI Calculation

The PSI (or planet similarity index) is the modified equivalent of the traditional ESI and can be calculated for single or multiple planetary properties. It can also calculate the weight (if it is unknown) and use it to compute the similarity. The \texttt{calc.psi} function takes in the following inputs:

- \texttt{params} - The dataset containing the values of the different parameters for which PSI is to be calculated.
- \texttt{upper_lim} - The list of upper limits for the given parameters.
- \texttt{lower_lim} - The list of lower limits for the given parameters.
- \texttt{ref_val} - The list of reference values for the given parameters.
- \texttt{threshold} (optional) - The threshold value to be considered for very high similarity (default = 0.8).
- \texttt{int_param} (optional) - List of column names that contribute to interior PSI.
- \texttt{surf_param} (optional) - List of column names that contribute to surface PSI.
- \texttt{p_index} (optional) - A column (passed as a pandas data frame) that is to be used as the index for the table.

The interior and surface parameter arguments are optional by nature. However, they are used to compute the interior PSI, surface PSI and global PSI. Thus, if they are not defined, then only the calculation of the individual PSI values for each parameter is possible.

The code in Listing 1 calculates the PSI values for the planetary radii, densities, escape velocities and surface temperatures on the PHL-EC dataset. The resulting dataset is shown in Table 2.

3.3 Unit Conversion

The similarity index calculations are simplified by converting the property values into a single unit scale. Unit conversion provides the ability to find the value of planetary properties relative to a single planet. This has the added advantage of providing a method for comparing planetary features. \texttt{ExoPSI} offers unit conversion through the function \texttt{unit_conv} requiring the data to be converted, the transformation reference value (the values of the given parameters in the same units as the data for the planet taken as the reference), and the name that should be assigned to the unit. These are taken as function arguments \texttt{data, ref_index} and \texttt{unit_name}. The \texttt{p_index} argument serves the same purpose as in the \texttt{calc.psi} function. It returns a pandas data frame containing the data values in the desired unit scale.

The code in Listing 1 converts the values for the planetary radii, densities and escape velocities of the PHL-EC dataset from EU to MU (Mars Units). The values for \texttt{ref_index} are the values of radius, density and escape velocity of Mars in EU. The resulting dataset is shown in Table 3.
Table 2  First 10 rows of PSI data

<table>
<thead>
<tr>
<th>Name</th>
<th>PSI&lt;sub&gt;r&lt;/sub&gt;</th>
<th>PSI&lt;sub&gt;d&lt;/sub&gt;</th>
<th>PSI&lt;sub&gt;e&lt;/sub&gt;</th>
<th>PSI&lt;sub&gt;t&lt;/sub&gt;</th>
<th>Interior PSI</th>
<th>Surface PSI</th>
<th>Global PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 Cnc e</td>
<td>0.813609</td>
<td>0.805275</td>
<td>0.741526</td>
<td>0.000387</td>
<td>0.8546357564747399</td>
<td>0.016940205488718253</td>
<td>0.120240946820440056</td>
</tr>
<tr>
<td>61 Vir b</td>
<td>0.777883</td>
<td>0.686974</td>
<td>0.842833</td>
<td>0.004745</td>
<td>0.7310166865687814</td>
<td>0.06323956502854838</td>
<td>0.215097140197864</td>
</tr>
<tr>
<td>BD-06 1339 b</td>
<td>0.688725</td>
<td>0.51925</td>
<td>0.803038</td>
<td>0.033221</td>
<td>0.5980137592480068</td>
<td>0.16333317298699612</td>
<td>0.3125307741453115</td>
</tr>
<tr>
<td>CD Cet b</td>
<td>0.822801</td>
<td>0.782244</td>
<td>0.861444</td>
<td>0.183938</td>
<td>0.8022662559549666</td>
<td>0.2980665677481867</td>
<td>0.5651116849139411</td>
</tr>
<tr>
<td>CoRoT-7 b</td>
<td>0.846681</td>
<td>0.518254</td>
<td>0.841935</td>
<td>0.006634</td>
<td>0.8814291042492635</td>
<td>0.02310382630648006</td>
<td>0.14270386326076857</td>
</tr>
<tr>
<td>CoRoT-7 c</td>
<td>0.689871</td>
<td>0.51541</td>
<td>0.80614</td>
<td>0.006408</td>
<td>0.5962938974281055</td>
<td>0.07187311820145276</td>
<td>0.2070205346642095</td>
</tr>
<tr>
<td>DMPP-1 c</td>
<td>0.667893</td>
<td>0.478094</td>
<td>0.795934</td>
<td>0.003777</td>
<td>0.5650807940772717</td>
<td>0.05482921409248905</td>
<td>0.17601970298241174</td>
</tr>
<tr>
<td>DMPP-1 d</td>
<td>0.852189</td>
<td>0.846992</td>
<td>0.873623</td>
<td>0.001043</td>
<td>0.8495865261925944</td>
<td>0.03018590381287258</td>
<td>0.16014224040009642</td>
</tr>
<tr>
<td>DMPP-1 e</td>
<td>0.815425</td>
<td>0.767618</td>
<td>0.857909</td>
<td>0.002888</td>
<td>0.7911604816028162</td>
<td>0.04970692024255778</td>
<td>0.19830822211419977</td>
</tr>
<tr>
<td>DMPP-3 A b</td>
<td>0.898658</td>
<td>0.95397</td>
<td>0.892794</td>
<td>0.014404</td>
<td>0.9254538987221532</td>
<td>0.1134010792541235</td>
<td>0.3239597906265763</td>
</tr>
</tbody>
</table>

Note: This table shows the PSI values for different planet parameters using equation 1. PSI<sub>r</sub> corresponds to the PSI calculation for radius. PSI values for density, escape velocity, and surface temperature are represented by PSI<sub>d</sub>, PSI<sub>e</sub>, and PSI<sub>t</sub>, respectively. The geometric mean of interior parameters (in this case, the planetary radius and its density) is used to calculate the interior PSI, whereas the geometric mean of surface parameters is used to get the surface PSI (the planet’s escape velocity and surface temperature in this case). The last column displays the global PSI, which is the geometric mean of interior and surface PSI.
Table 3 First 10 rows of mars data.

<table>
<thead>
<tr>
<th>Name</th>
<th>Radius (EU)</th>
<th>Density (EU)</th>
<th>Escape Velocity (EU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 Cnc e</td>
<td>3.532207547169811</td>
<td>1.7152712676056339</td>
<td>4.5999775555555556</td>
</tr>
<tr>
<td>61 Vir b</td>
<td>3.976373584905655</td>
<td>0.7651796338028168</td>
<td>3.4519286666666664</td>
</tr>
<tr>
<td>BD-06 1339 b</td>
<td>5.372339622641509</td>
<td>0.5235758169014085</td>
<td>3.857854</td>
</tr>
<tr>
<td>CD Cet b</td>
<td>3.426452830188679</td>
<td>0.9280584788732896</td>
<td>3.2776215555555552</td>
</tr>
<tr>
<td>CoRoT-7 b</td>
<td>3.172641509433962</td>
<td>1.208808983915493</td>
<td>3.4605764444444445</td>
</tr>
<tr>
<td>CoRoT-7 c</td>
<td>5.351188679245283</td>
<td>0.5186239436619718</td>
<td>3.824450888888889</td>
</tr>
<tr>
<td>DMPP-1 c</td>
<td>5.774207547169811</td>
<td>0.4716675211267606</td>
<td>3.935271111111111</td>
</tr>
<tr>
<td>DMPP-1 d</td>
<td>3.109188679245287</td>
<td>1.054400366197184</td>
<td>3.1684175555555556</td>
</tr>
<tr>
<td>DMPP-1 e</td>
<td>3.511056603773585</td>
<td>0.9024057464788733</td>
<td>3.310025111111111</td>
</tr>
<tr>
<td>DMPP-3 A b</td>
<td>2.665018867924528</td>
<td>1.2899158028169305</td>
<td>3.003817111111111</td>
</tr>
</tbody>
</table>

3.4 PSI Scale

The PSI Scale outputs a Surface PSI vs Interior PSI plot. This allows us to see the similarity or dissimilarities of other exoplanets regarding the planet under consideration and gives an understanding of whether the surface or interior properties are responsible for the value of the index. The psi_scale function takes the PSI data as input. It returns a scatter plot of Surface PSI vs Interior PSI with axes limits of 0.0 to 1.0. Hovering over the points of the scatter plot displays the planet’s name (or index of the PSI dataset in case of the absence of planet name), to make it easier for the user to identify the planets with higher PSI values.

The code in Listing 1 displays the PSI Scale for the PSI data. The resulting plot is shown in Figure 1.

Fig. 1 PSI Scale for the planets in PHL-EC dataset. Teegarden’s Star b shows the highest similarity.
3.5 PSI Distribution

The PSI Distribution plot is a histogram plot that outputs a distribution of the PSI values showcasing the count of exoplanets whose index values fall under specific ranges. This plot can be generated by passing the PSI data to the `psi.dist` function. It returns a count plot with bins ranging from ‘Very Low Similarity’ (0.0 - 0.2) to ‘Very High Similarity’ (0.8 - 1.0).

The code in Listing 1 displays the PSI Scale for the `PSI.data`. The resulting plot is shown in Figure 2.

![PSI Distribution](image)

**Fig. 2** PSI Distribution of the planets in PHL-EC dataset. 39 planets lie in the region of very high similarity.

4 Example

This example illustrates the applicability of ExoPSI in exoplanetary research. Here, we demonstrate how ExoPSI can be used to identify temperate, sub-Neptune-sized planets that are potential hosts for life as defined by Seager et al [20]. With radii between 1.5 and 4 Earth radii ($R_{Earth}$), sub Neptunes are planets that are larger in size than Earth but lesser in size than Neptune. Sub Neptunes have thick H$_2$ or H$_2$-He envelopes that make up 1 to 10% of their mass, which sets them apart from "super-Earths," planets that are likewise larger than Earth but less than sub Neptunes [1, 13]. In a previous study done by Seager et al [19], they demonstrated the feasibility of life on exoplanets with H$_2$-dominated atmospheres. Additionally we use the word "temperate" to emphasise the sub-Neptunes that get energy comparable to that of Earth from their host star. As a consequence, these planets have the unique potential to have liquid water, either in the form of clouds [5] or in the form of an ocean underneath the atmosphere [14]. We have used the characteristics defined for sub Neptunes to
calculate the weights of the PSI parameters. The values of planetary properties of the sub Neptune TOI-1266 c are used as reference values for this example. The planetary attributes of radius, mass, escape velocity, and equilibrium temperature were taken into account using the PHL-EC dataset (Table 4). Table 5 shows the final data of exoplanets similar to the sub Neptune TOI-1266 c, which comprises the result obtained using ExoPSI. As visualised in Figure 3, L 98-59e shows the highest similarity to TOI-1266c with a PSI of 0.926. Figure 4 depicts the distribution of exoplanets that are similar to TOI-1266 c among those described in the PHL-EC dataset, and it reveals that 51 of them lie in the region of very high similarity. Appendix B shows the code used for this example.

**Table 4** First 10 rows of the PHL-EC dataset for the desired columns

<table>
<thead>
<tr>
<th>Name</th>
<th>Radius (EU)</th>
<th>Density (EU)</th>
<th>Escape Velocity (EU)</th>
<th>Equilibrium Temperature (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Com b</td>
<td>12.1068</td>
<td>6165.8633</td>
<td>22.567438</td>
<td>813.07498</td>
</tr>
<tr>
<td>11 UMi b</td>
<td>12.2189</td>
<td>4684.7848</td>
<td>19.580725</td>
<td>833.92148</td>
</tr>
<tr>
<td>14 And b</td>
<td>12.8915</td>
<td>1525.5744</td>
<td>10.878399</td>
<td>771.53541</td>
</tr>
<tr>
<td>14 Her b</td>
<td>12.5552</td>
<td>2586.7417</td>
<td>14.353728</td>
<td>138.83404</td>
</tr>
<tr>
<td>17 Sco b</td>
<td>12.8915</td>
<td>1373.017</td>
<td>10.320155</td>
<td>774.21767</td>
</tr>
<tr>
<td>18 Del b</td>
<td>12.4431</td>
<td>3273.6284</td>
<td>16.219983</td>
<td>397.1771</td>
</tr>
<tr>
<td>24 Boo b</td>
<td>13.9004</td>
<td>289.22348</td>
<td>4.5614522</td>
<td>1622.6325</td>
</tr>
<tr>
<td>24 Sex b</td>
<td>13.3399</td>
<td>632.47773</td>
<td>6.885672</td>
<td>431.03198</td>
</tr>
<tr>
<td>24 Sex c</td>
<td>13.9004</td>
<td>273.33208</td>
<td>4.4343672</td>
<td>343.38196</td>
</tr>
</tbody>
</table>

Note: The PHL-EC dataset has been cleaned to drop rows where the data is missing for any of the desired columns.

## 5 Conclusion

In this paper, we have presented ExoPSI, a python package that provides the basic building block to calculate the similarity index between any two planets. It provides an intuitive, object-oriented user interface that simplifies the calculation and visualisation of similarity index calculations. This package’s adaptability allows for a wide range of potential applications and development in the future. It can be used to find similarities between any two planets, not just with earth (ESI). It can also be used to find relative planetary properties with unit conversion. Another important application is to compute the weight associated with each planetary parameter and visualise the extent of similarity with the surface vs interior PSI plots and the similarity distribution. The numerous advantages and uses of ExoPSI have also been discussed. The use of ExoPSI to locate planets like the exoplanet TOI-1266 c for research on the viability of life on sub-Neptunes has been illustrated with an example.
Table 5  Top 10 candidates with highest PSI with respect to TOI-1266 c

<table>
<thead>
<tr>
<th>Name</th>
<th>PSI&lt;sub&gt;r&lt;/sub&gt;</th>
<th>PSI&lt;sub&gt;m&lt;/sub&gt;</th>
<th>PSI&lt;sub&gt;e&lt;/sub&gt;</th>
<th>Interior PSI</th>
<th>Surface PSI</th>
<th>Global PSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>L 98-59 e</td>
<td>0.998787</td>
<td>0.836393</td>
<td>0.921602</td>
<td>0.9139904021875722</td>
<td>0.938000605000871</td>
<td>0.9259176800433893</td>
</tr>
<tr>
<td>GJ 625 b</td>
<td>0.953268</td>
<td>0.876645</td>
<td>0.930275</td>
<td>0.9141540493046016</td>
<td>0.9261297645713585</td>
<td>0.9201224236287391</td>
</tr>
<tr>
<td>Kepler-395 c</td>
<td>0.83571</td>
<td>0.977544</td>
<td>0.949279</td>
<td>0.903849156891559</td>
<td>0.931389941325329</td>
<td>0.9175162189761396</td>
</tr>
<tr>
<td>K2-3 c</td>
<td>0.998787</td>
<td>0.985037</td>
<td>0.996764</td>
<td>0.992341206054168</td>
<td>0.8423191788271237</td>
<td>0.9142581830152119</td>
</tr>
<tr>
<td>GJ 3223 b</td>
<td>0.762745</td>
<td>0.957688</td>
<td>0.960736</td>
<td>0.8546763911329246</td>
<td>0.9685194701375105</td>
<td>0.9076240902259464</td>
</tr>
<tr>
<td>GJ 1132 c</td>
<td>0.91467</td>
<td>0.908875</td>
<td>0.936403</td>
<td>0.9117678960404342</td>
<td>0.9026667850530449</td>
<td>0.9072059277888659</td>
</tr>
<tr>
<td>GJ 411 b</td>
<td>0.922437</td>
<td>0.900012</td>
<td>0.9341</td>
<td>0.826438</td>
<td>0.911155513052073</td>
<td>0.8786214974606529</td>
</tr>
<tr>
<td>Kepler-54 d</td>
<td>0.976137</td>
<td>0.853408</td>
<td>0.924573</td>
<td>0.9127119616264487</td>
<td>0.8739334874969001</td>
<td>0.8926522878136119</td>
</tr>
<tr>
<td>Kepler-560 b</td>
<td>0.903557</td>
<td>0.75974</td>
<td>0.907052</td>
<td>0.8285338829402211</td>
<td>0.9509802825884457</td>
<td>0.887648233329281</td>
</tr>
<tr>
<td>HD 85512 b</td>
<td>0.903557</td>
<td>0.772457</td>
<td>0.913171</td>
<td>0.965769</td>
<td>0.83543936317904</td>
<td>0.939101828077765</td>
</tr>
</tbody>
</table>

Note: This table shows the PSI values with respect to TOI 1266c. Here, PSI<sub>m</sub> corresponds to the PSI calculation for mass and PSI<sub>e</sub> corresponds to the PSI calculation for equilibrium temperature. Other definitions are the same as in Table 2.
Fig. 3 PSI Scale for the planets in PHL-EC dataset with respect to TOI-1266 c. L 98-59 e shows the highest similarity.

Fig. 4 PSI Distribution of the planets in PHL-EC dataset with respect to TOI-1266 c. 51 planets lie in the region of very high similarity.

6 Discussions

Possible future developments include the adaptability and compatibility with any data type (not just with Pandas Dataframes). Another essential inclusion would be to develop a formulation that allows the use of categorical data for similarity measures,
not just numerical calculations and perform life information calculations. We leave this endeavour to the future release of **ExoPSI**.

**Appendix A**

Listing 1 shows the code used for the examples in section 3.

Listing 1  Code to demonstrate the usage of **ExoPSI**

```python
#Importing the libraries
from ExoPSI import exopsi
import pandas as pd
#Instantiating the class
exopsi = exopsi()

#Importing the dataset
df = pd.read_csv(r"phl_exoplanet_catalog.csv")

#Weight Calculation
upper_lims=[1.9, 1.5,1.4,323]
lower_lims = [0.5, 0.7,0.4,273]
ref_val = [1,1,1,288]
weights = exopsi.calc_weight(ref_val,upper_lims,lower_lims)

#PSI Calculation
PSI_data = exopsi.calc_psi(df[['P_RADIUS','P_DENSITY','P_ESCAPE','P_TEMP_SURF']],upper_lims,
lower_lims,ref_val,0.8,
surf_param=['P_ESCAPE','P_TEMP_SURF'],
int_param=['P_RADIUS','P_DENSITY'],
p_index=df.loc[:,'P_NAME'])
print(PSI_data)

#Unit Conversion
mars_data = exopsi.unit_conv(df[['P_RADIUS','P_DENSITY','P_ESCAPE']],
[0.53,0.71,0.45],'MU',

df.loc[:,'P_NAME'])
print(mars_data)

#PSI Scale
exopsi.psi_scale(PSI_data)

#PSI Distribution
exopsi.psi_dist(PSI_data)
```

**Appendix B**

Listing 2 shows the code used to study the planetary candidates similar to TOI-1266 c.
Listing 2  Code to study the Planet Similarity with respect to TOI-1266 c

```python
# Importing the libraries
from ExoPSI import exopsi
import pandas as pd

# Instantiating the class
exopsi = exopsi()

# Loading and cleaning the dataset
df = pd.read_csv(r"phl_exoplanet_catalog.csv")
new_df = df.loc[df["P_TEMP_EQUIL"].isna()==False]
new_df = new_df.loc[new_df["P_RADIUS"].isna()==False]
new_df = new_df.loc[new_df["P_MASS"].isna()==False]
new_df = new_df.loc[new_df["P_ESCAPE"].isna()==False]
new_df = new_df.loc[new_df["P_NAME"].isna()==False]
new_df = new_df.loc[new_df["P_NAME"] != 'TOI-1266␣c']
df = new_df
print(df)

# Weight Calculation
upper_lims=[4,10,1.4,320]
lower_lims = [1.5,2,0.4,200]
ref_val = [1.56,2.2,1.18,315.3]
weights = exopsi.calc_weight(ref_val,upper_lims,lower_lims)

# PSI Calculation
PSI_data = exopsi.calc_psi(df[['P_RADIUS','P_MASS','P_ESCAPE','P TEMP_EQUIL']], upper_lims,
                          lower_lims,ref_val,0.8,
surf_param=['P_ESCAPE','P TEMP_EQUIL'],
int_param=['P_RADIUS','P_MASS'],
p_index=df.loc[:,"P_NAME"])
print(PSI_data)

# PSI Scale
exopsi.psi_scale(PSI_data)

# PSI Distribution
exopsi.psi_dist(PSI_data)

toi1266c_sim_candidates = PSI_data.loc[PSI_data["PSI_Global"]>0.8]
toi1266c_sim_candidates = toi1266c_sim_candidates.sort_values(by=['PSI_Global'],ascending = False)
print(toi1266c_sim_candidates)
```
References


**Declarations**

To the best of their knowledge, all authors declare that the work submitted is original and true. This manuscript accurately and completely represents the work, and none
of the authors withholds any information regarding the content, licensing, support, or interests. The authors sincerely believe that this work is critical to the relevant areas of research.

All authors agree that proper acknowledgement has been given for any kind of derivation or referencing done to the work of others, and all authors will be held equally responsible for any future issues that arise. All of the authors have agreed to the content and have given their permission to submit and publish this manuscript. Furthermore, all authors have agreed on the authorship order and have designated A. S. Rao as the manuscript’s corresponding author.

**Funding**

None of the authors received any funding to conduct this research.

**Competing Interests**

The authors confirm that they have no affiliation or active involvement with any organisation or institution that has a financial or non-financial interest in the topics or materials presented in this document.

**Author Contributions**

The study’s ideation and design involved input from all authors. Vaibhav Garg and Divya Srinivasan prepared the initial draft of the manuscript, and all authors provided feedback on subsequent versions. Data collection and analysis were performed by Aditya Rai and Divya Srinivasan. Vaibhav Garg, Aditya Rai and Divya Srinivasan wrote the source code for the library. Aditya Rai and Vaibhav Garg are in charge of the development and maintenance of the library. A. S. Rao had a guiding spirit over this research. All authors read and approved the final manuscript.