py_tps2tnt 4.2 User Manual

by: Jonathan Liria

1. Introduction

py_tps2tnt is a Python-based graphical tool that facilitates the conversion of TPS files containing morphometric landmark data into TNT (Goloboff & Morales, 2023) file format for subsequent phylogenetic analysis*. The program allows users to process Generalized Procrustes Analysis (GPA), calculate distances between landmarks (EDMA), and export results with options for centroid size calculations and interval estimations.

This manual will guide you through loading TPS files, running analyses, and exporting results using the program's graphical interface.

(*) For detailed theory and instructions to perform parsimony analysis with continuous data, please see: Catalano, S., Goloboff, P.A. (2018). A guide for the analysis of continuous and landmark characters in TNT (Tree Analysis using New Technologies). 10.13140/RG.2.2.23797.27360

A brief tutorial for Spanish speakers is available at: https://youtu.be/ubR3w-yRhx4

---

2. System requirements

Before using **py_tps2tnt**, make sure you have the following installed:

- Python 3.x
- And the following libraries:
    - `tkinter`
    - `numpy`
    - `scipy`
    - `matplotlib`
    - `morphops` (Patel, 2021)

Note: Make sure you have pip installed for your version of Python.

```
py -m ensurepip --default-pip
```

If you installed Python from source, with an installer from python.org, you should already have pip. If you're on Linux and installed using your OS package manager, you may have to install pip separately, see https://packaging.python.org/en/latest/tutorials/installing-packages/

---

3. Getting started

3.1 Launching the program

1. Run the script `py_tps2tnt.py` to open the graphical interface.
2. The main window of **py_tps2tnt** will appear, and you can start loading TPS files and performing analyses.

---

4. User interface overview

4.1 Main interface components

The graphical interface consists of buttons and input fields that allow you to load TPS files, set reference species, run analyses, and export results.
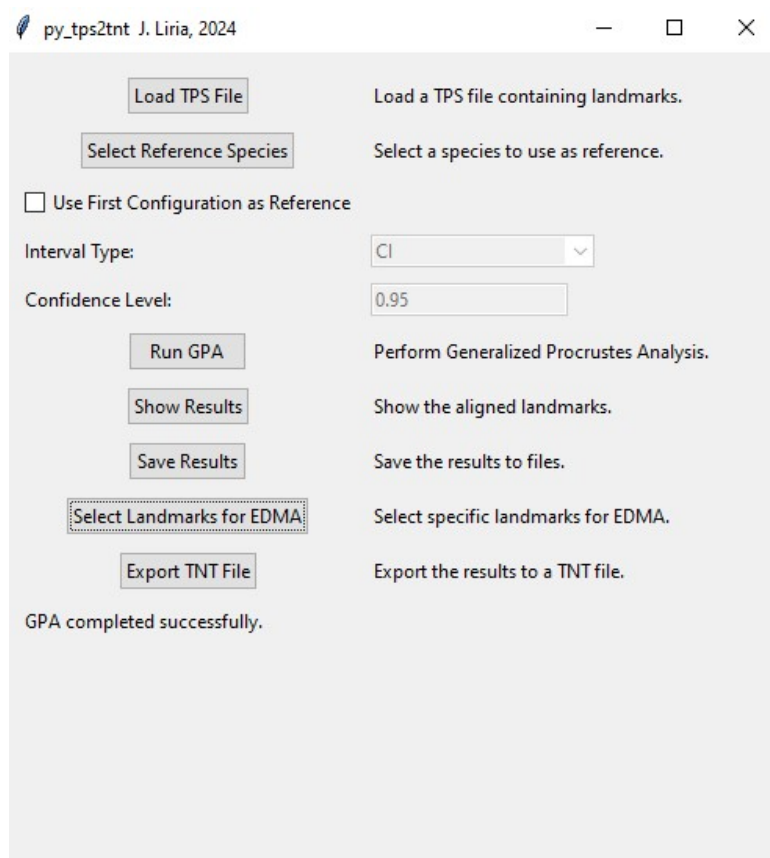
Here are the key components:

1. **Load TPS file**: Load a TPS file containing landmarks for one or more species. After loading, the program will prompt if the file contains multiple specimens per species. For multiple specimens, you must load a text file with the following format:

SpeciesOne 5

SpeciesTwo 3

SpeciesThree 6

.

.

2. **Select reference species**: Choose a specific species as a reference for GPA alignment. The reference species will be used as the starting point for the analysis.
3. **Use first configuration as reference**: Optionally, use the first configuration in the TPS file as the reference for alignment instead of selecting a species.
4. **Interval type**: Choose whether to export centroid size intervals as Confidence Intervals (CI) or as Mean ± Standard Error (SE).
5. **Confidence Level**: Input the confidence level for the interval calculation (e.g., 95%).
6. **Run GPA**: Perform Generalized Procrustes Analysis on the loaded landmarks, aligning them based on the reference configuration.
7. **Show results**: Display the aligned landmarks and centroid sizes graphically.
8. **Save results**: Save the results of the GPA and other analyses in TPS format.
9. **Select Landmarks for EDMA**: Choose specific landmarks for EDMA (Euclidean Distance Matrix Analysis).
10. **Export TNT File**: Export the results, including centroid sizes, EDMA distances, and landmark configurations in TNT format for phylogenetic analysis.

---

5. Features and Workflow

5.1 Loading a TPS file

1. Click **Load TPS File** and select a TPS file from your directory.
2. The program will ask whether the file contains multiple specimens per species.

TPS file format

The TPS file should include:

- **LM**: The number of landmarks per specimen.
- **Coordinates**: X and Y coordinates for each landmark.
- **ID**: Unique identifier for each specimen.
- **SCALE**: Optional scale factor to adjust landmark coordinates.

Note: If the TPS file contains a `SCALE` value, the program automatically multiplies each landmark coordinate by this scale factor. This ensures all subsequent calculations use the scaled coordinates.

5.2 Reference species selection

After loading the TPS file, click **Select Reference Species** to choose a species for the alignment in GPA. Alternatively, you can check **Use First Configuration as Reference** to use the first specimen.

5.3 Interval type and Confidence Level

Choose the type of interval to be calculated for centroid sizes: **CI (Confidence Interval)** or **Mean ± SE**. You can also specify the desired confidence level for the intervals.

5.4 Running Generalized Procrustes Analysis (GPA)

Once the file is loaded and the reference configuration is set, click **Run GPA**. This step aligns all landmarks based on the reference, removing effects of translation, rotation, and scaling.

5.5 Viewing results

Click **Show Results** to visualize the aligned landmarks in a scatter plot. The plot will display each specimen's landmark configuration after the GPA.
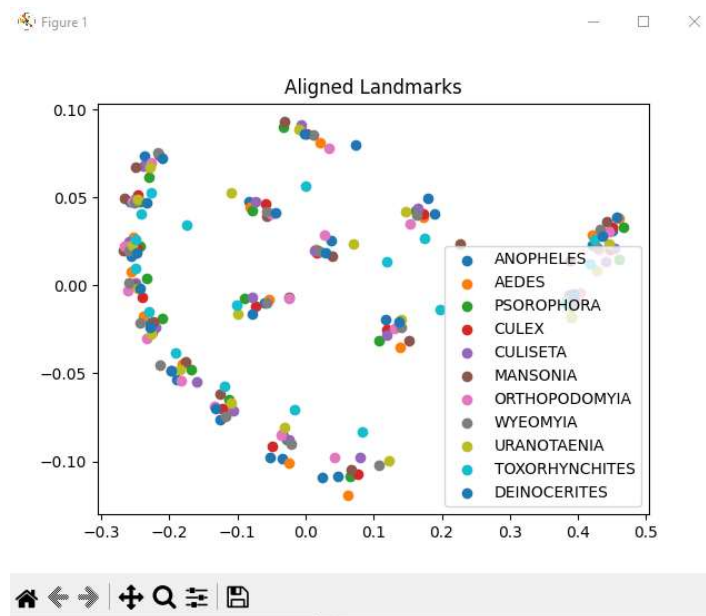

5.6 Saving results

After running GPA, you can save the aligned configurations and centroid sizes in TPS format by clicking **Save Results**. You can also save species means or centroid sizes.

5.7 EDMA: Selecting landmarks and calculating Euclidean distances

- Click **Select Landmarks for EDMA** to open a window where you can choose specific landmarks for calculating distances.
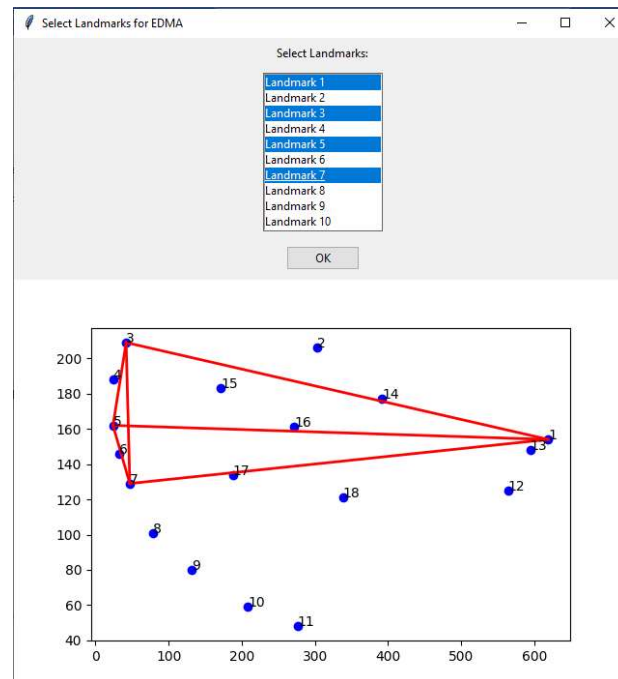
- The selected landmarks will be used to compute distances between points for further analysis.

5.8 Exporting TNT files

Click **Export TNT file** to generate a TNT format file. This file can include:

- Centroid sizes (with intervals if multiple specimens per species),
- Landmark configurations (mean or GPA-aligned),
- EDMA distances (with intervals).

The TNT export process lets you choose which data to include based on your analysis requirements.

6. Export Options

When exporting results to TNT format, the program provides the following options:

- **Include Centroid Size**: Choose to include centroid size data (either as a single value or as an interval if multiple specimens are available).
- **Include Mean Coordinates**: Export mean landmark coordinates.
- **Include GPA Coordinates**: Export GPA-aligned landmark coordinates.
- **Include EDMA Distances**: Export Euclidean distance matrices between selected landmarks, either as a single value or with intervals.

Interval calculations

- **Confidence Interval (CI)**: If selected, the program calculates a confidence interval using bootstrapping.
- **Mean ± SE**: Calculates the mean and standard error for the intervals.

---

7. Example Workflow

Step 1: Load TPS file

- Click **Load TPS File** and select a TPS file.

Step 2: Select Reference species

- Choose a reference species or use the first configuration.

Step 3: Set interval and Confidence Level

- Choose **CI** or **Mean ± SE** and input a confidence level (e.g., 95%).

Step 4: Run GPA

- Click **Run GPA** to align the landmarks.

Step 5: View results

- Click **Show Results** to visualize the aligned landmarks.

Step 6: Export TNT file

- Click **Export TNT File** and choose which data to include in the output.

---

8. TPS File Structure and data examples

A sample TPS file:

```
LM=14
995.00000 866.00000
1018.00000 796.00000
...
SCALE=0.003609
ID=SpecieOne1
```

In this file:

- `LM=14` indicates there are 14 landmarks.
- The subsequent lines provide X/Y coordinates.
- `SCALE` adjusts the coordinates based on the given factor.
- `ID` is the unique identifier for the specimen.

It is important to note that TNT software support only recognizes continuous characters with 65.000 states. However, the program scales all continuous values. To display the original states, it is necessary to utilize the command `nstates [`

In the code repository was included different data for testing:

File Name	Description	Data detail
mole_spp.tps	Mole cranial data from Rohlf, J. et al. (1996). Published in Systematic Zoology 45(3): 344-362.	13 landmarks: 10 species with 113 specimens.
mosqH.tps	Mosquito wing landmarks from Rohlf, J., Slice, D. (1990). Published in Systematic Biology, 39(1): 40–59.	18 landmarks: 11 species.
mosqH_various.tps	Mosquito wing landmarks from Rohlf and Slice (1990), with additional specimens hypothetical generate only for testing purposes.	18 landmarks: 11 species with 55 specimens.
pentagonoids.tps	Specimens hypothetical generate by De Luna, E. (2020). Published in Acta botánica mexicana 127: e164.	5 landmarks: 6 species.
triangles.tps	Specimens hypothetical generate by Liria, J, only for testing purposes.	3 landmarks: 5 species.
triangles_various.tps	Specimens hypothetical generate by Liria, J, only for testing purposes.	3 landmarks: 5 species with 16 specimens.
triangles_various_w_scale.tps	Specimens hypothetical generate by Liria, J, only for testing purposes.	3 landmarks: 5 species with 16 specimens, all scaled.

9. References

Patel, V. (2021). morphops 0.1.13. common operations and algorithms for Geometric Morphometrics, in Python 3 https://pypi.org/project/morphops/

Catalano, S., Goloboff, P.A. (2018). A guide for the analysis of continuous and landmark characters in TNT (Tree Analysis using New Technologies). 10.13140/RG.2.2.23797.27360

Goloboff, P.A., Morales, M.E. (2023). TNT version 1.6, with a graphical interface for MacOS and Linux, including new routines in parallel. Cladistics, 39: 144-153. https://doi.org/10.1111/cla.12524

Rohlf, F. J. (2015). The tps series of software. Hystrix, the Italian Journal of Mammalogy, 26(1), 9-12. https://doi.org/10.4404/hystrix-26.1-11264