

## ***What is IMP?***

IMP is a set of compiled software tools with graphic user interfaces that carry out a wide variety of tasks related to the display and analysis of 2-D landmark-based geometric morphometric data in the Windows environment. All of the IMP programs use the same file format, and are meant to work in conjunction with one another. Each program in the IMP series is meant to carry out a limited number of specific tasks, so that a complete analysis of your data will require the use of more than one IMP program. Since they share a common file format, progressing through a series of different programs to carry out the analysis is not as daunting as it might seem initially. Each program in the series is meant to be easy to use, and the controls of the various programs operate in similar ways. This is meant to ease the learning curve for the software, and allows for a lot of flexibility in how the software operates.

This is a brief introduction to the IMP Suite (Integrated Morphometrics Package) for analysis of landmark based geometric morphometric data. This document is a listing of the programs that make up IMP and what they do, it is not intended as an introduction to geometric morphometrics. Some user manuals are also available and included with the compressed program files.

If you have not already seen it, the morphometrics website at SUNY Stony Brook, <http://life.bio.sunysb.edu/morph/> maintained by F.J. Rohlf is a source of a wealth of information on morphometrics. You should also be aware of the excellent TPS series of software by F. James Rohlf.

### ***IMP Programming by:***

H. David Sheets, Dept. of Physics, Canisius College, Buffalo, NY 14208,

Dept. of Geology, SUNY at Buffalo, Buffalo NY 14260

sheets@canisius.edu

***Graphic User Interface Design:*** H. David Sheets, Miriam L Zelditch and Donald

Swiderski (Univ. of Michigan, Ann Arbor)

### ***Referencing IMP:***

The IMP series is completely free, but I would appreciate being referenced if you use it in published work. Please reference my name and address, and ideally the website address,

<http://www.canisius.edu/~sheets/morphsoft.html>

### ***Background***

All of this software was written in MATLAB, where the core components were all used as functions in a command line interface environment. MATLAB also allows programming graphic users interfaces to control the subroutines involved. I wrote the subroutines and functions used throughout IMP for my own research work, to allow myself greater flexibility and control, particularly when carrying out bootstrapping. Eventually, I found it convenient and productive to compile pieces of software to send to my collaborators (rather than running all the analyses myself!). Since these tools are now available, it makes sense to distribute them freely.

The graphic user interface forms of this software was converted into C/C++ code using the Matlab compiler from Mathworks, and compiled using Borland C++ compiler.

It will now run under Windows without MATLAB installed on the user's computer (a big savings, given what Matlab costs!). The downside of this is that the user does need to install a set of dynamic link libraries to support the mathematical and graphical routines used, these files are in the compressed self-extracting file mglinstaller.

The file mglinstaller.exe will install a variety of files used by IMP, you will need to download a copy of it and allow it to expand. It will install a number of files, and make several subdirectories. The subdirectory \bin\win32 created by mglinstaller must be in the PATH variable of your computer so that the IMP software can load the .DLL files it needs.

If you have trouble setting the PATH variable, simply expand all the IMP files into the \bin\win32 directory, which should take care of the problem.

IMP is not currently available for the MAC, although the MATLAB code should run fine on a MAC with MATLAB loaded on it. At least one person has run IMP within a Windows emulation package running on a MAC, and reported that it worked fine.

## **Two Categories of Programs in IMP: General Releases and Undocumented Software**

There are two categories of software in IMP, *General Release* and *Undocumented Software*. The difference between the two is that I have had the time to write a decent users manuals for the *General Release* Software, and haven't yet had the time to do so for the *Undocumented Software*. You are welcome to download software from both categories, but don't expect good documentation for software in the second category!

The reason for even posting the *Undocumented Software* on the website is that it makes it easier for my direct collaborators (and their collaborators and students) to get hold of this software, without having to wait for me to write a complete manual for it. Knowledge of how to use this software pretty much relies on passing information by word of mouth. I will try to write manuals for all the software eventually, but writing software manuals is pretty boring, and takes time away from other things. So when I can, I will. In the meantime, feel free to download the *Undocumented Software*, but don't expect much of a manual for now.

If you are in desperate need of a particular piece of *Undocumented Software*, contact me and plead for the manual, if I know someone is desperate I may be able to reallocate some time, by putting off some task that is less appealing at the time.

### **Programs available in IMP**

IMP consists of the following "core" programs:

- **CoordGen6** –This program generates a variety of different types of shape coordinates, and translates files back and forth between the IMP file format (called X1Y1, see below) and the TPS file format used by the programs produced by F. James Rohlf. This program is the entry point into IMP. It also produces reference forms for thin plate spline analysis. CoordGen is the “entry point” into the IMP system.
- **PCAGen6**-This is a Principal Components Analysis for geometric morphometric data. The program computes Partial Warp plus Uniform component scores, based on the Thin Plate Spline decomposition of the deformation, and then displays the deformations implied by various loadings of PCA Axes.

- **CVAGen6**- This is a Canonical Variates Analysis program for landmark data. The program computes the Partial Warp plus uniform component scores and then does the CVA on these scores. CVA determines the set of axes that best discriminates among groups, along with the statistics related to those axes. This is a useful tool for classifying groups. CVAGen operates in much the same way as PCAGen does, users of CVAGen need to be familiar with PCAGen.
- **Regress6**- This is a multipurpose program, producing thin plate spline deformation (partial warp) scores, and regressing them on a dependent variable, typically centroid size. It also computes Procrustes distances from a reference, and calculates the associated statistics.
- **VecDisplay6\***- Displays partial warp scores, and sums and differences of partial warp scores. This is useful for displaying the results of statistical analysis carried out using other software. Obtain partial warp scores and the reference form from one of the IMP programs, carry out your analysis in SAS, SPSS etc, and then load the results into VecDisplay6 to display them. The current version of VecDisplay does not have as good a deformation drawing function in it, and is currently in need of an update. Look for this in the near future.
- **Tmorphgen6**-Generates a set of traditional morphometric variables (lengths) from a geometric morphometrics data set.
- **TwoGroup6**- This program carries out pairwise tests of significant differences in shapes between groups, using Goodall's F-test and Hotelling's  $T^2$  test, in addition to bootstrap versions of these tests. The program also computes the Partial Procrustes distances between the means of the groups.

All of the above core programs are also included in the compressed data file ImpBasics6.exe, which is a self-extracting Zip file, as are all of the individual distribution files listed above.

If you are going to be working with IMP, you will probably want to have all of these files available, so loading IMPBASICS6 makes sense.

The asterisk after the name of the software indicates that it is currently (01/16/02) *Undocumented Software*.

For further examples of use of these programs, see the references below.

**IMP also contains the following "auxiliaries"- All of Which are currently Undocumented**

-----

These programs are somewhat more specialized, not all users may need them. They are available only as individual compressed files.

- **BigFix6\***- handles data that has a bilateral axis of symmetry, where points on only one side of the axis are to be used in the analysis. In the file BigFixpac6.exe
- **GrowProf6\***- Generates "Growth Profiles". In the file growp6.exe. Used in Zelditch, Sheets and Fink (2001).
- **ShuffleAllometry6\***- Tests for the range of vector correlations that may be generated by a growth vector, either in traditional morphometrics form, or in geometric morphometrics. In the file shufflepack6.exe.
- **Vecland6\***- Regresses shape coordinates against a dependent variable and displays the deformation, allows comparison of the patterns of deformation of two different groups.
- **VecCompare6\***- Computes the between and among growth growth vector angles for two groups. In veccompac6.exe

- **Standard6\***- This program fits a regression model to the data and computes the residual. This information is then used to “standardize” the data at a specific value of the independent variable, by forming a set of landmark configuration based on the regression model evaluated at a specified value of the independent variable, plus the residuals. This program must be used with caution, to avoid statistically invalid results.
- **DisparityBox\***- calculates the morphological disparity of a number of groups of specimens, and confidence intervals around those disparity measures.

The asterisk after the name of the software indicates that it is currently (01/16/02) *Undocumented Software*.

## **File Formats**

The IMP Suite uses a file format in which each specimen lies along a single row of the data matrix. The columns of the data matrix represent the coordinates (or partial warp scores) in the following order, x1y1x2y2....CS, where X1 is the x-coordinate of the first landmark, Y1 is the y-coordinate of the first landmark, X2 is the x-coordinate of the 2nd landmark and so on. CS refers to the centroid size of the specimen. More details about file formats are included in the manual for CoordGen.

No specimen labels are used, the order of samples is preserved throughout all subsequent analyses. Note that IMP programs will allow the user to add labels to the end of each line of data as long as they are prefixed with a %. IMP software will simply ignore such labels, but they are often tremendously helpful to the user. Cut and paste such labels using Excel, or a similar program.

The centroid size is also helpful in identifying specimens when uncertain. All files are in ASCII format text, and load easily into spreadsheets (like EXCEL) or statistical software (SPSS), which allow analysis of columns of data.

When partial warp scores are calculated they are stored starting with the X component of the Principal Warp of the highest bending energy, followed by the Y component along

this Principal Warp, finally the Uniform X and Uniform Y scores, and the centroid size. Partial warp scores may be obtained from the Regress6 program, or from PCAGen6.

It is assumed that most users will be starting with data files in the TPS file format, as produced by the TPS series programs produced by F.James Rohlf.

CoordGen6 will load and write TPS and X1Y1...CS files.

Note that the IMP system does not keep track of what content any given text file contains, the user must do this. We often use abbreviations for different file types.

An example set of files might be:

Fish.tps -> TPS format file, with comments, of original digitized  
data. Produced by TPSDig, or user.

Fishbc.txt -> Bookstein Coordinate data file generated from  
Fish.tps, from CoordGen6

Fishps.txt -> Procrustes superposition data file from Fish.tps, from  
CoordGen6

Fishsbr.txt -> Sliding Baseline Registration data file from Fish.tps,  
from CoordGen6

Fishpw.txt -> Partial Warp scores generated from Fish data file,  
from Regress6

Fishref.txt-> reference form generated from Fish data file, from  
CoordGen6 or Regress6

## **Users Manuals**

A few users manuals are currently available, they are intended as introductions to the software, not to the discipline of geometric morphometrics. You will need to do some other reading.....see the textbooks and papers discussed on the Stony Brook website.

These users manuals are often included with the compressed program files and will be expanded when the pac6.exe file is expanded.

More coming as time allows.....check the Website

.....A matter of time, not money or difficulty....

## References for IMP

-----

The listing below is the set of sources I consulted during the programming of the functions within IMP. The primary sources were the Blue Book (Rohlf and Bookstein 1990), the Orange Book (Bookstein 1991), the White Book (Marcus et al. 1996) and Dryden and Mardia (1998). The referencing throughout the manuals could certainly be more complete than it is at present, but the listing below is reasonably comprehensive.

Bookstein. F.L., 1991, Morphometric tools for landmark data. Geometry and Biology. Cambridge University Press, Cambridge.

Bookstein, F.L.1996. Standard formula for the uniform shape component in landmark data.. In Marcus, L.F., M. Corti, A. Loy, G.J.P. Naylor and D.E. Slice,(eds) 1996. Advances in Morphometrics. Nato ASI Series, Series A: Life Science Vol. 284, Plenum, N.Y. pg. 153-169.

Bookstein, F.L.1996. Combining the tools of geometric morphometrics. In Marcus, L.F., M. Corti, A. Loy, G.J.P. Naylor and D.E. Slice,(eds) 1996. Advances in Morphometrics. Nato ASI Series, Series A: Life Science Vol. 284, Plenum, N.Y. Pg. 131-152

Dryden, I.L. and K.V. Mardia, 1998. Statistical Shape Analysis. Wiley, Chichester



- Efron B. and R. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, NY.
- Freund, J.E. and R.E. Walpole.1980. Mathematical Statistics. Prentice-Hall, Englewood Cliffs, NJ.
- Goodall, C.R. 1991. Procrustes methods in the statistical analysis of shape (with discussion) Journal of the Royal Statistical Society, Series B. 53:285-339.
- Krzanowski, W. J. Principles of Multivariate Analysis. Oxford University Press, 1988.
- Marcus, L.F., M. Corti, A. Loy, G.J.P. Naylor and D.E. Slice,(eds) 1996. Advances in Morphometrics. Nato ASI Series, Series A: Life Science Vol. 284, Plenum, N.Y.
- Morrison D. F., 1967, 1990. *Multivariate Statistical Methods*, McGraw-Hill, N.Y.
- Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling.1988. Numerical Recipes in C, The art of scientific computing. Cambridge University Press, Cambridge.
- Rohlf F.J. and F.L. Bookstein, (eds.) 1990. Proceedings of the Michigan morphometrics workshop. University of Michigan Museum of Zoology. Special Publication 2.
- Rohlf, F.J., 1998, TPSRegress 1.18, Ecology and Evolution, SUNY at Stony Brook, <http://life.bio.sunysb.edu/morph/>
- Rohlf, F.J. 2000-2001 Personal Communication, in a number of instances.
- Stevens, J. 1996. Applied Multivariate Statistics for the Social Sciences. Lawrence

Erlbaum Associates, Mahwah, New Jersey.

Webster, M., H. D. Sheets, N. Hughes, 2001, Testing of Allometric Heterochrony in Trilobites, In "The Evolution of Development", ed. by Miriam L. Zelditch, Wiley, NY.

Zelditch, M. L., H. D. Sheets and W. L. Fink, 2000, "Spatiotemporal Reorganization of Growth Rates in the Evolution of Ontogeny", *Evolution* 54(4):1363-1371

Zelditch, M. L., H. D. Sheets and W. L. Fink, 2001, Heterochrony and heterotopy in the evolution of piranha body form. In "The Evolution of Development", edited by Miriam L. Zelditch, Wiley, NY.

This Document Last Updated

-----

1/16/02

-H. David Sheets