Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 6, No. 1, January 2002 (© 2002)

Data Library and Analysis Project

Stephen J. Guastello,[1,4] Kevin J. Dooley,[2] and Robert A. M. Gregson[3]

GENERAL INFORMATION ABOUT THE PROJECT

The Society for Chaos Theory in Psychology & Life Sciences is pleased to announce its new Data Library and Analysis Project. The goal of the project is to facilitate the testing of new software, algorithms, and computational procedures for nonlinear dynamical phenomena. The materials for the project will be housed on its web site, with exceptions as noted below www.societyforchaostheory.org.

All interested parties are invited to contribute data sets to the cause. Interested parties should read the next section, *How to Contribute a Data Set*, which includes a preview of the format of the data set descriptions that the project team wishes to acquire.

Descriptions of the data sets will be housed on the SCTPLS web site. The data sets themselves will be housed either on this web site or on the contributor's site, depending on which method represents the cybernetic path of least resistance. Downloading instructions are given at the point of access.

The Society will not provide analytic software or licenses for software. Fortunately, many options are available at no charge or involve facilities that analysts may own already. The Society will provide as complete information as possible about available software. The node of the web site designated *Other Resources/Software* will contain that information. The Project Team invites the use of computational techniques that are not yet known to the web site.

[1]Marquette University.
[2]Arizona State University.
[3]Australian National University.
[4]Correspondence should be directed to Stephen J. Guastello, Ph.D., Dept. of Psychology, Marquette University, P.O. Box 1881, Milwaukee, WI 53201-1881, USA.

1090-0578/02/0100-0071$21.50/0

The analytic strategy that a researcher might choose to pursue depends, of course, on the type of data being considered, and the programs that could be applicable to such data.

People who have downloaded data from the project are welcome to contribute their final reports to a section of this web site that is dedicated to housing such reports. Potential reports are not limited to those that use the available data from this site. Please refer to the section below, *Formatting Your Research Report for this Web Site*, for preparation and submission of these documents.

After a collection of reports has started to grow, some of the reports will be selected for publication in *Nonlinear Dynamics*, *Psychology*, *and Life Sciences*. The choice of papers will be based on the reviews of the Editors. The usual considerations of depth, clarity, originality, and topic coverage will apply. The authors whose work has been selected in this fashion will be contacted sufficiently in advance of publication to ascertain their interest in such publication.

HOW TO CONTRIBUTE A DATA SET

Those who would like to make a data set available for this project should submit the following materials to Kevin Dooley, kevin.dooley@asu.edu, SCTPLS President and Webmaster. E-mail submission is required for all materials that are intended for the web site; if there is a problem using e-mail for this purposes, please describe the problem to Dr. Dooley, and alternative arrangements will be worked out if possible. Please do not send "go to my website" messages.

Cover Letter/Message

Please identify yourself and your institutional affiliation. Give the name of the data set that you are enclosing as an attached file. Your message should clearly state an assurance from your that indeed you have the right to post this data set. The best case is where the person submitting the data is the one who made it. There are other acceptable conditions also, e.g., the data are already in the public domain.

Data Files

The data files should be in formatted in ASCII or MS EXCEL spreadsheet. Dr. Dooley will be screening incoming material for format and webworthiness. Large files may be ZIPPED or TAR COMPRESSED.

Provide a description of your data set in 750 words or less giving: (a) a descriptive title of the data; (b) the name(s) and affiliation(s) of the people submitting the data; (c) a description of the variables, how they were measured, and the conditions under which they were generated; (d) the format of the variables if not already given in (c) above; (e) the file type and any compression facility that may have been used; (f) a summary of the published research involving that data, if any; if none, so state; (g) bibliographic citations to (f) if there are any.

You may also include up to two figures if they are important for understanding your data set, its peculiarities, or particular analytic challenges that it might present. Figures should be prepared in .jpg, .tif, .gif, or postscript. The Webmaster may convert them to another format. BITMAP files are usually too large for the site; they may be submitted to the Webmaster, who will probably convert them to a different format.

The selection criteria for project data sets are simple. The general goal is to have good material available, to minimize redundancy, and to maximize variety. We encourage data sets for EEGs, EKGs, genomes, other biological phenomena, cognitive processes, clinical or other transcripts of human communication, other social and organizational behaviors, and economic phenomena.

SUGGESTED RESEARCH QUESTIONS

There are many possible questions that one might pose concerning new computational techniques and different types of data. We can make some suggestions, nonetheless:

If two computational techniques are meant to produce the same theoretical indicator (e.g., fractal dimension) by different means, is one method more suitable to one type of data than it is to another?

Various filtering strategies have been suggested over the years. Do they actually improve the data by removing noise, or do they strip out true measurement, or true dynamics, as well?

Sometimes we are interested in comparing entire modeling strategies. Does one computational system identify nonlinear systems more accurately than another?

Are the results of any of the new computational strategies more robust under different conditions of sampling, stationarity, cross-validation, etc.?

How are any of the questions given above modified when the researcher is faced with different types of data, such as EEGs, performance of work groups, hierarchically-organized phenomena, etc.

Do any of the new analytic techniques convey any new *meaning* concerning the phenomena from which the data were taken?

At this point we wish to convey some idea of the type of analyses that we *do not* wish to actively encourage. Researchers who are familiar with the nonlinear dynamics literature are probably aware of the large supply of journal articles where, "We took data generated from the logistic map equation [Lorenz attractor, etc.], added X% IID noise, ran it through [our favorite] computational procedure, and looked at what came out." Although it is probable that "pure attractor" data sets will be part of the data library, the goal here is to facilitate the solution of *substantive* problems that are encountered in the life and social sciences, which can only be solved by resolving problems that are encountered with real data.

More generally, we can anticipate four strategic approaches to any of the data analysis objectives listed above:

1. One data set is used in conjunction with one method of analysis, with interpretation relative to something that is known about the data.
2. Two or more data sets are used in conjunction with one method of analysis, with interpretations that address (in part) some aspect of the data that are different.
3. One data set is used in conjunction with two or more methods of analysis, with interpretations that address some aspect of the analyses that is different.
4. Two or more data sets are used in conjunction with two or more methods of analysis.

This last type of strategy lends itself to comprehensive thinking, which is good. From a practical standpoint, however it may be unwieldy. Analysts should be advised to break up their material into smaller chunks, as defined by 1, 2, 3 above. The next section, *Formatting Your Research Report for the SCTPLS Web Site* should provide additional useful information for framing your research questions and organizing your results.

FORMATTING YOUR RESEARCH REPORT FOR THE SCTPLS WEB SITE

When you are ready to write up and post your research report, please prepare your document according to the instructions below. For the most part, you should follow the *NDPLS Instructions for Authors* with the following exceptions:

1. Text should be single-spaced instead of double-spaced.
2. Place the tables and figures in the text where they are meant to appear, rather than at the end of the manuscript.

DO adhere to the following style points of manuscript preparation:

1. Begin your article with a title, list of authors, their institutional affiliations, and the address and e-mail of the corresponding author. In the case of multiple-author works, there will be only one "corresponding author."
2. The author/title information is followed by an abstract of approximately 150–250 words summarizing the nature of the research question, the data used and programs used, and the highlights of results.
3. References in the text and reference list should conform to the *NDPLS*/APA style.
4. If you are working with a data set that was downloaded from the SCTPLS web site, cite the contributing authors, their data set, and URL as it appears on the web site. Include other bibliographic citations to publications involving the data as given in the data description if such publications exist.
5. Research reports are accepted in English only. Both the British and United States spellings are acceptable.

Electronic formats for your reports are limited to the following preparations: MS WORD, Word Perfect, .pdf (for Adobe Acrobat Readers), and postscript. TEX users should prepare the postscript version of their document.

Figures take up a lot of space if wrong format is chosen. For those using MS WORD, convert your graphics to .jpg, .gif, or .tif before inserting them into your manuscript. BITMAP files tend to be overly large. Scanner documents tend to be large and incompatible with other software.

If you cannot convert your figures to a small format that is compatible with your word processor, put a MARKER in your text where figures should go (see ***NDPLS Instructions for Authors***), and send the figures to the Webmaster in separate files. The Webmaster will make the necessary conversions and post the document in a suitable format.

As we mentioned in the project introduction, the Society intends to collect a set of manuscripts culled from the web site postings for publication in *Nonlinear Dynamics*, *Psychology*, *and Life Sciences.* It is assumed that a posting to the SCTPLS web site constitutes submission for journal review, *unless you specify otherwise.* To specify otherwise, your title page must acknowledge that the paper is under review by another journal, or has been accepted for publication in another journal.

If you do in fact wish to have your manuscript considered for a special collection in *NDPLS*, it is required that your manuscript has not been accepted for publication elsewhere, and is not under review by another journal. Publication on your own website, or access through the SCTPLS website

does *not* constitute "publication." Distribution of your work through an "electronic journal" *does* constitute publication, however. Presentations at conferences, or publication of abstracts in conference documents do *not* constitute publication. Publication of the full text in a conference proceedings book *usually does* constitute publication. *NDPLS* cannot accept previously published material even if the original language of publication was not English. Contact the *NDPLS* Editor stephen.guastello@marquette.edu for any clarifications of the foregoing policy.

The provisions of the United States Copyright Law will apply to all documents loaded to this site. The US Code is generous to authors, and most of its provisions have been adopted as an international standard. The following points are particularly pertinent:

1. All rights to the paper belong to the authors until such time as a transfer of copyright has been signed between an author and another entity, such as a publisher. SCTPLS intends to honor the nature of agreements that exist between authors and other entities as a result of a transfer of copyright.
2. If the copyright has not been transferred, there is no problem with posting your paper to the SCTPLS site. If the paper has been accepted for publication, publishers typically allow authors to supply preprints through sites such as this one. When the paper is published, however, it is necessary for the authors to remove their paper from the site *or* request permission from the publisher to leave it on the site.
3. It is the author's responsibility to request permission from their publisher to continue to make their paper available on the SCTPLS site. To obtain such permission, contact the Permissions Office of the Publisher of the journal or book. Such permission is not usually authorized from the office of the Journal Editor.
4. If your document contains someone else's copyrighted materials other than the data provided on this site (e.g., figures, tables, or text excerpts greater than 99 words), it is the author's responsibility to secure such permission from the original publisher and/or authors. Journals typically give permission information on one of the inside covers of the journal.

Send your completed document to Dr. Kevin Dooley kevin.dooley@asu.edu. And last, but not least, analytic papers are welcomed by the *NDPLS* editors throughout the year, whether the authors prefer to participate in the web-based project or not.