

ESS (Emacs Speaks Statistics) as a User Interface to SAS

Richard M. Heiberger
Temple University

1. Emacs is a mature, powerful, and easily extensible text editing system that runs identically on all current computers (Unix, Windows, Macintosh, VMS).

Emacs knows the syntax of each programming language (Figure 1).

- directory editor
- FORTRAN
- C
- L^AT_EX
- SAS
- SAS log
- `*Buffer List*`

The screenshot shows an Emacs window with the title 'philasug_sas'. The menu bar includes 'File', 'Edit', 'Options', 'Buffers', 'Tools', 'ESS', and 'Help'. The main area displays several buffers, each with a header line indicating the file name, date, time, and language mode. The buffers and their contents are:

- philasug<rmh/ESS>**: A file directory listing showing files like 'sas-bash.el', 'philasug.sas', 'philasug-sas.log', 'philasug.sas~', 'regr', and 'software'.
- philasug.tex**: A LaTeX file containing a list item describing Emacs: '\begin{enumerate} \item Emacs is a mature, powerful, and easily extensible text editing system which is freely available under the GNU General Public License for a large number of platforms.'
- mgs.c**: A C file containing a function 'normcol' that normalizes a column of a matrix. The code uses 'for' loops and 'colprod'.
- mgs.f**: A Fortran file containing a 'DO' loop that calls 'NRMCOL' and 'COLPRD' to perform matrix operations on 'A(I,J)'. It includes comments like '/* NORMALIZE COLUMN K */'.
- philasug.sas**: A SAS file containing a 'proc glm' procedure with a 'model' statement and an 'output' statement.

Figure 1: Emacs frame with buffers showing a file directory, L^AT_EX file, C file, FORTRAN file, and SAS file. Each file is syntactically highlighted for its language.

2. Emacs can interact with and control other programs either as sub-processes or as cooperating processes.

- shell (Unix shell, on Unix or Windows machines)
- msdos (on Windows machines)
- telnet
- SAS

One major advantage of running other processes under emacs is that the user has complete search and editing capability on the interactive session. The interactive session is just another file that the computer is writing to at the same time you are reading it.

```

telnet-sbm*
File Edit Options Buffers Tools Complete In/Out Signals Help
c:/rmh/ESS/philasug$ ls -alF ph*
-rw-r--r--  1 rmh      544      1362 Nov 10 21:18 philasug-sas.log
-rw-r--r--  1 rmh      544      5192 Nov 11 14:47 philasug.dvi
-rw-r--r--  1 rmh      544        347 Nov 11 20:49 philasug.el
-rw-r--r--  1 rmh      544      3019 Nov 11 19:33 philasug.log
-rw-r--r--  1 rmh      544      1553 Nov 11 19:33 philasug.lst
-rw-r--r--  1 rmh      544        765 Nov 10 23:34 philasug.sas
-rw-r--r--  1 rmh      544      5298 Nov 11 14:47 philasug.tex
-rw-r--r--  1 rmh      544      1516 Sep 17 17:33 philsug.abstract
c:/rmh/ESS/philasug$ date
Wed Nov 14 18:36:14 2001
c:/rmh/ESS/philasug$

-t(Unix)** *shell* Wed Nov 14 6:45PM (Shell:run)--L117--C0--Bot-----
C:\rmh\ESS\philasug>dir ph*
PHILSU~1 ABS          1,516  09-17-01  4:33p  philsug.abstract
PHILASUG TEX          5,298  11-11-01  2:47p  philasug.tex
PHILASUG LOG          3,019  11-11-01  7:33p  philasug.log
PHILASUG DVI          5,192  11-11-01  2:47p  philasug.dvi
PHILASUG SAS           765  11-10-01 11:34p  philasug.sas
PHILAS~1 LOG          1,362  11-10-01  9:18p  philasug-sas.log
PHILASUG LST          1,553  11-11-01  7:33p  philasug.lst
PHILASUG EL           347  11-11-01  8:49p  philasug.el
                8 file(s)          19,052 bytes
                0 dir(s)          3,125.35 MB free

C:\rmh\ESS\philasug>

-t(Unix)** *msdos* Wed Nov 14 6:45PM (Shell:run msdos)--L49--C0--Bot-----
login: rmh
Password:
IRIX Release 6.5 IP19 surfer
Copyright 1987-2000 Silicon Graphics, Inc. All Rights Reserved.
You have mail.
sbm:/userdata/usrl/rmh> mail
Mail version SGI.950426. Type ? for help.
"/usr/mail/rmh": 62 messages 19 new 42 unread
>N 44 <Russell.Lavery. Fri Nov 9 16:52 PhilaSUG paper on emacs
-t(Unix)** *telnet-sbm* Wed Nov 14 6:45PM (Telnet:no process)--L42--C65-- 1%---

```

Figure 2: Emacs frame with buffers showing a Cygwin Unix-like shell, an MS-DOS shell, and a telnet session to a remote Unix computer.

3. ESS extends Emacs to provide a functional and uniform interface for multiple statistical languages including SAS.

The other languages ESS works with are

- S (S-Plus and R)
- XLispStat (including ARC and ViSta)
- Stata

The programs can be running simultaneously on the same or different computers.

Emacs provides:

- viewing two or more files at once
- editing formatted text
- visual comparison of two similar files
- syntax highlighting
- syntactically appropriate indentation of text
- enforcement of programming standards
- navigation in units of characters, words, lines, sentences, paragraphs, and pages.
- all Unix utilities (grep, awk, etc, are builtin).

```

iml.sas
File Edit Options Buffers Tools ESS Help

do j = 1 to n;
v = a[ ,j];
call ORTVEC(w, u, rho, lindep, v, q);
aro = abs(rho);
if aro > dmax then dmax = aro;
if aro <= 1.e-10 * dmax then lindep = 1;
if lindep = 0 then do;
nind = nind + 1;
q = q || w;
if nind = n then r = r || (u // rho);
else r = r || (u // rho // j(n-nind, 1, 1));
end;
else do;
print "Column " j " is linearly dependent";
ndep = ndep + 1; ind[ndep] = j;
end;
end;

do j = 1 to n;
v = a[ ,j];
call ORTVEC(w, u, rho, lindep, v, q);
aro = abs(rho);
if aro > dmax then dmax = aro;
if aro <= 1.e-10 * dmax then lindep = 1;
if lindep = 0 then do;
nind = nind + 1;
q = q || w;
if nind = n then r = r || (u // rho);
else r = r || (u // rho // j(n-nind, 1, 1));
end;
else do;
print "Column " j " is linearly dependent";
ndep = ndep + 1; ind[ndep] = j;
end;
end;

```

-1(Unix)-- **iml-nqr.sas** Wed Nov 14 8:59 -1(Unix)-- **iml.sas** Wed Nov 14 8:59

Figure 3: Syntactically appropriate indentation of text.

Left: SAS file as entered with all lines on the left margin.

Right: Same SAS file indented by ESS to display nesting structure of statements.

```
other.sas
File Edit Options Buffers Tools ESS Help
  gout=regr.gseg);
run;

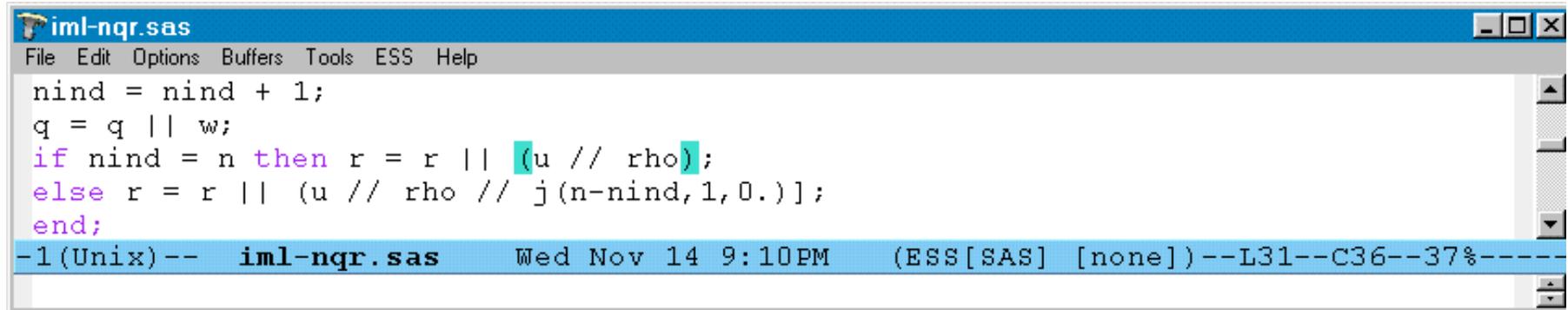
proc glm data=regr.fat;
  model bodyfat = abdomin biceps;
  output out=regr.fat2 p=bodyfathat ;
run;

A: --(Unix)-- other.sas      Wed Nov 14 9:07PM    (ESS[SAS] [none])--L20--C23--63%--
  gout=regr.gseg);
run;

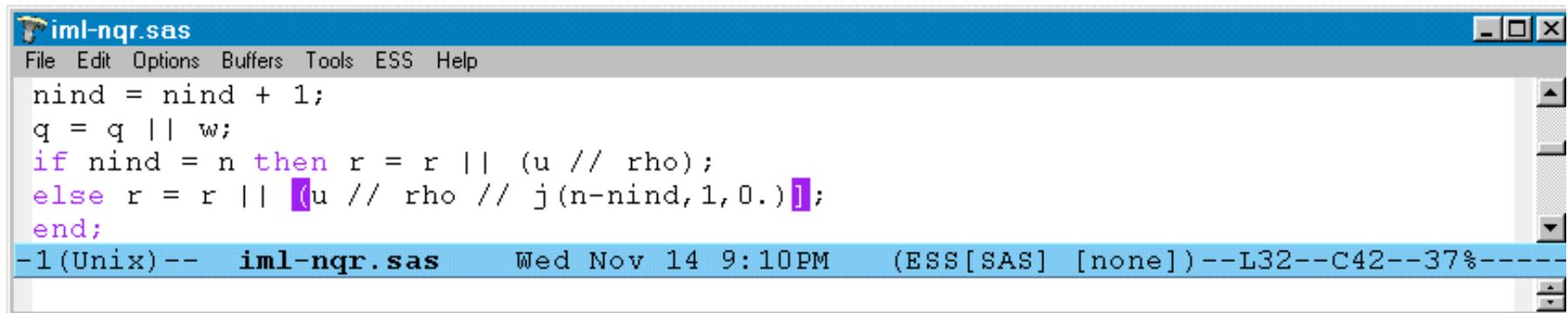
proc glm data=regr.fat;
  model bodyfat = abdomin;
  output out=regr.fat2 p=bodyfathat ;
run;

B: --(Unix)-- philasug.sas   Wed Nov 14 9:07PM    (ESS[SAS] [none])--L21--C0--64%--
```

Figure 4: Ediff session comparing two similar files.
The lines that differ are marked.
The words that differ on those lines are marked.



```
iml-nqr.sas
File Edit Options Buffers Tools ESS Help
nind = nind + 1;
q = q || w;
if nind = n then r = r || (u // rho);
else r = r || (u // rho // j(n-nind,1,0.));
end;
-1(Unix)-- iml-nqr.sas Wed Nov 14 9:10PM (ESS[SAS] [none])--L31--C36--37%-----
```



```
iml-nqr.sas
File Edit Options Buffers Tools ESS Help
nind = nind + 1;
q = q || w;
if nind = n then r = r || (u // rho);
else r = r || (u // rho // j(n-nind,1,0.));
end;
-1(Unix)-- iml-nqr.sas Wed Nov 14 9:10PM (ESS[SAS] [none])--L32--C42--37%-----
```

Figure 5: Programming made easier.

Matching parentheses are highlighted in a gentle color.

Mismatching parentheses are highlighted in a disturbing color.

Keywords are in the keyword color.

```
hilock.sas
File Edit Options Buffers Tools ESS Help
proc print data = abc ;
proc print data=fdssfa ;
procprint data=fdssfa ;
proc print other options;
proc print data b;
proc print ;
proc print
  data=fdssfa ;
-- (Unix)-- hilock.sas Wed Nov 14 9:12PM (ESS[SAS] [none] H)--L17--C0--69%--
```

Figure 6: Enforcement of programming standards.

The standard says all **PROC** statements must have a **DATA=datasetname** option.

Lines that satisfy the standard are green.

Lines that don't are red.

Lines that are unclear are yellow.

Learning emacs

- TUTORIAL `C-h t`
- online manual `C-h i`
- reference card `refcard.ps` and `survival.ps`
- online Help system for
 - functions `C-h f`
 - variables `C-h v`
 - keys `C-h k`

Learning ESS

- online manual `C-h i m ESS <RET>`

- Conference paper

Richard M. Heiberger.

Emacs speaks statistics: One interface — many programs. In Kurt Hornik and Friedrich Leisch, editors, *Proceedings of the 2nd International Workshop on Distributed Statistical Computing (DSC 2001)*. Technische Universität Wien, Vienna, Austria, 2001.

<http://www.ci.tuwien.ac.at/Conferences/DSC.html>,
ISSN 1609-395X.

Emacs and ESS are freely available under the GNU Public License.

- Emacs

<http://www.gnu.org/software/emacs/>

- ESS

A.J. Rossini, Martin Mächler, Kurt Hornik, Richard M. Heiberger, and Rodney Sparapani. ESS (emacs speaks statistics), 2001.

<http://www.analytics.washington.edu/Zope/wikis/ess/FrontPage>.

Conclusion:

Emacs and ESS enable major improvements in productivity.

ESS (Emacs Speaks Statistics) as a User Interface to SAS

Richard M. Heiberger

Temple University

`rmh@temple.edu`