

Tutorial 21: Using new gensspp script features to compare cycle times in runs

This tutorial describes a new way to use the gensspp script: polar plots comparing averages of runs with circular statistical comparison of large numbers of cycles.

In *Tutorial 13* there are details presented about how to create polar plots (circular plots) that express the phase shifts between the onset (or offset) of two waveforms within a run. We have further developed this ability of the analysis software and created a new way of comparing not only two waveforms within a run, but across two or more runs. This script essentially calculates the average burst difference between two waveforms over a specified number of cycles and converts that to a position on the polar plot for each runfile, and so the display on the plot is of these averages (not of each cycle as in the original version).

When can this be useful for science? With the original version of gensspp, we selected randomly a number of cycles (5, 10 but no more than 15 was optimal) and created a polar plot showing the phase shifts between two waveforms in each of those cycles. What if you have a preparation in which you record locomotor ENG activity for long periods at a time, and in several run files. Then you plan to use a drug and evaluate its effects on the coordination, as well as expecting a recovery of locomotor activity after the drug washes out. Because you have longer periods of locomotor activity with greater number of cycles (well over 15) in each run; the original version of the script is not the best choice for evaluating the changes caused by the drug on all of the cycles. However, with the new script, you can create the averages in those “control” runs that you collected before the drug application and then create the averages for those runs which were during the drug application as well as create the averages for those runs that were collected after “washout”. By comparing the averages from runs, you can basically see a trend (for example, if the drug had no effect on extensor-extensor coupling but changed flexor-extensor coupling) in a faster, simpler way that you could have done so with the original version of the gensspp.

Note that this gensspp option (with the -a option argument) requires specifying the runfile and X and Y waveform numbers in a different order than without the -a option. It also requires all the steps you would need to take for polar plotting (set range, filter if necessary & set cycles and specify the waveform number for cycle based analysis – as detailed in Tutorial 13) for all the runfiles you use.

Before the change (and now when you use the new version without -a) the usage is:

```
gensspp [various options] runfile xwfnum ywfnum [outfile]
```

In the new version, you need to add -a option to it, to tell it to show averaged angles for each run:

```
gensspp [various options] -a xwfnum ywfnum runfile1 runfile2 ...
```

The one or more run file arguments appear after the waveform numbers, rather than a single runfile name before the numbers. Also, you can't use the optional outfile argument with -a, so if you want to save the output of gensspp -a, you must use the file redirect, i.e. “> outfile”. With -a, it's assumed that

all the arguments after the two waveform numbers are runfile names, and you can specify as many as you wish.

Note also that you can use wildcards for file name matching (*, ?, [abc]) but when you do the filename extension (.prm or .frm) must be used so the shell can match actual file names. When you explicitly give run names as in the last example above, the .frm or .prm suffix is assumed automatically if none is given.

Also note that gensspp will use the same two waveform numbers for all the runs you specify, so it's assumed that you're consistent in how you named and numbered all your waveforms. That means that if you manually rectify and filter ENG/EMGs after capture, you should try to do it in the same order for all runs, or explicitly specify the resulting waveform numbers when filtering so they're consistent.

You can use "gensspp --help" for a description of all the options.

Beyond Averaging – Circular Statistics on two or more runs: This new gensspp script also add a **-an** option, used much like the **-a** option. With **-an**, gensspp uses a different statistical evaluation of significance levels than the previous one. It will run the statistical analysis on 2 or more runs, which includes the Watson-Williams test, the Shapiro-Wilk test, and Q-Q plots on all sets of angles (each run is a set, or group).

The reason for including the Q-Q plots is that this was strongly recommended when we read up on the Shapiro-Wilk test. (For more info see: <http://stackoverflow.com/questions/15427692/perform-a-shapiro-wilk-normality-test> and <http://stackoverflow.com/questions/7781798/seeing-if-data-is-normally-distributed-in-r/7788452#7788452>.) The reasoning is that all normality tests (and Shapiro-Wilk does seem to be held as the best of the lot) suffer from a couple of common problems: for small samples they can be inaccurate and not reject a sample that doesn't really fit a normal distribution, and for large samples they can be far too strict and reject a sample that deviates only marginally from a perfect normal distribution. So it's best to take these tests with a grain of salt and use other methods to get a feel for how the data really look. A Q-Q plot is recommended because it gives you a visual look at how well a sample fits the normal distribution, and you can see if the sample is bi-modal, skewed, or has just a few or many outlying values. Right now, the script does all 3 analyses when you use the **-an** option, and the Q-Q plots are just displayed on-screen in a pop-up window. Future possibilities would be options to select which tests to do or not, and output options for saving the Q-Q plots in a variety of formats (e.g. .eps, .pdf, .svg, .jpg or .png, but not .plt as R doesn't have an HPGL output device driver).

We have the updated version of **gensspp** installed on the SCRC systems, and it's available in the *May 20, 2015* release of the neuro software package. (*Note: At this time, we have only done minimal testing on these new analysis methods in gensspp, and we may consider refinements to these or trying other methods as the need arises after more thorough testing.*)

Examples: Here is a series of locomotor trials with 60 seconds of brain stimulation in the adult rat which resulted in well coordinated activity during nearly all of the stimulation period. We have collected 3 data files before giving any drugs, then 2 with the drug in effect and there are 3 files as "recovery" after the washout of the drug. The figure below shows the control runs, with rectified and filtered ENGs.

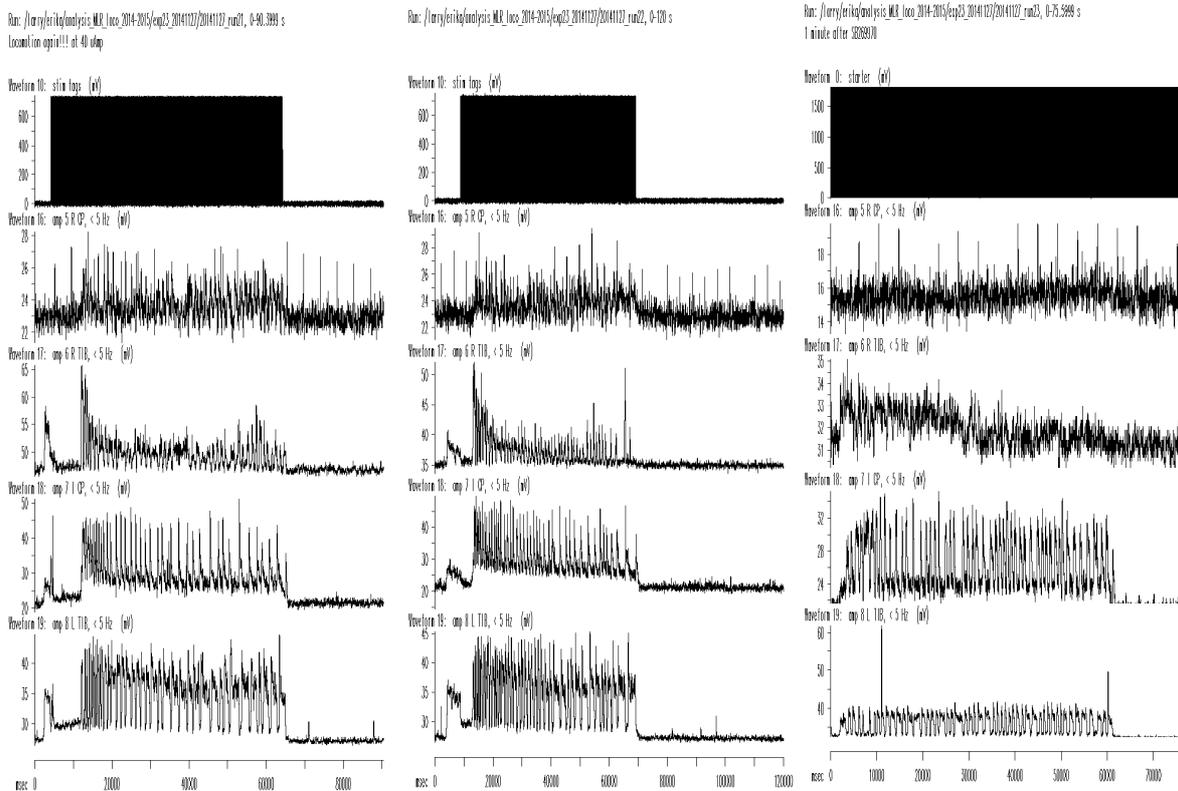
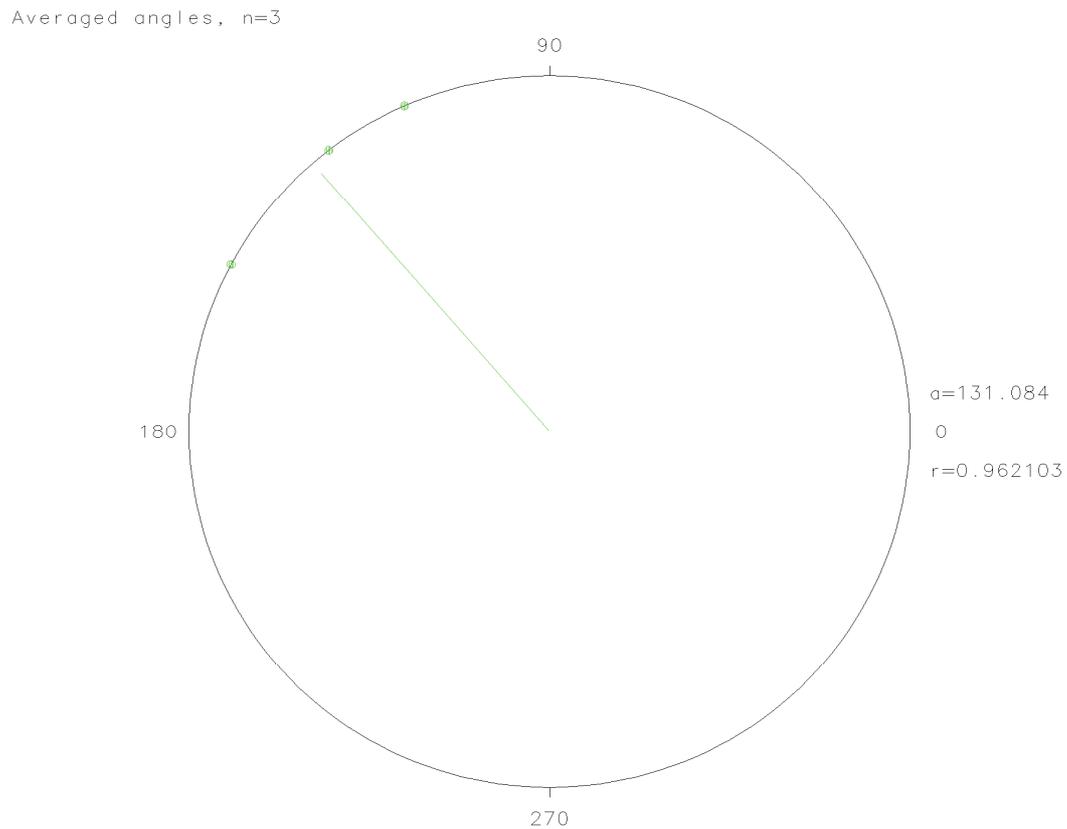


Figure 21.1: Three MLR evoked locomotion control runs, showing 4 rectified and filtered ENGs

Is the coupling between flexors and extensors the same during the control runs? This is the first question, so let's create the average plots for these three run files. Open one file at a time and set the range during which the stimulation was applied and set the cycles for L CP (flexor) and L TIB (extensor). (Note that in order to improve the cycle setting, the waveforms were blanked, rectified and filtered). Then set the waveform number for cycle based analysis for as the L CP (SCWN: 18). This will ensure that L CP will be used as the reference to which the other waveform will be compared (in other words, the zero on the polar plot will represent the onset of L CP). Either of the following two commands will produce the same plot file:

```
gensspp -c -yp3 -a 18 19 20141127_run2[123].prm > pprun21_22_23.plt
```

```
gensspp -c -yp3 -a 18 19 20141127_run21 20141127_run22 20141127_run23 > pprun21_22_23.plt
```



amp 8 L TIB, < 5 Hz activity start time: 3 averages vs X=#18 amp 7 l CP, < 5 Hz Cycle: 51

Figure 21.2: Averaged angle polar plot from three control runs

Now we try running the circular statistics tests. We use the `-an` option rather than `-a`, and we can remove all the options that relate to plotting.

```
[exp@larry exp23_20141127]$ gensspp -an 18 19 20141127_run2[123].prm
Analysing 20141127_run21.prm...
Analysing 20141127_run22.prm...
Analysing 20141127_run23.prm...
Loading required package: boot

Attaching package: 'circular'

The following object(s) are masked from 'package:stats':

sd, var
```

Watson-Williams test for homogeneity of means

data: angles by groups

F = 70.4295, df1 = 2, df2 = 124, p-value < 2.2e-16

sample estimates:

Circular Data:

Type = angles

Units = degrees

Template = none

Modulo = asis

Zero = 0

Rotation = counter

mean of 20141127_run21 mean of 20141127_run22 mean of 20141127_run23

113.6897	127.7061	152.0526
----------	----------	----------

Warning message:

In `watson.williams.test.default(angles, groups)` :

Concentration parameters: 8.444519.147522.5625 not equal between groups. The test might not be applicable

Shapiro-Wilk test of normality: data in 20141127_run21

may not be normally distributed, $p = 2.483312e-09$, $p < .05$

Shapiro-Wilk test of normality: data in 20141127_run22

may not be normally distributed, $p = 0.005581452$, $p < .05$

Shapiro-Wilk test of normality: data in 20141127_run23

may not be normally distributed, $p = 0.004846819$, $p < .05$

Press [Enter] to continue

With the `-an` option, no plot file is generated. Instead, the results appear as text in the terminal window, showing the results of the tests. Also, a window pops up on the X Window display showing Q-Q plots generated for the angle data for each run. In the results of the Watson-Williams test, a warning may show up if it calculates different concentration parameters between groups. This indicates that the test *might* not be applicable. Following this test, a Shapiro-Wilk test of normality is performed on each group (i.e. each run). If the calculated p value is below the threshold of 0.05, a warning appears. Again, this suggests the data *might* not be normally distributed, which means the Watson-Williams test *might* not be applicable. There is no certainty in these tests. The Q-Q plots show that most samples in all three groups are close to a theoretical normal distribution, with a few outliers in the first group. Taking all three tests into account, we can make a judgement call as to whether the Watson-Williams test is appropriate for these runs. The F and p values for this test appear near the start of the output.

After pressing the **Enter** key in the terminal window, the pop-up window with the Q-Q plots goes away. There is currently no other output option for this graph, but a screen capture (before dismissing the window) is always possible.

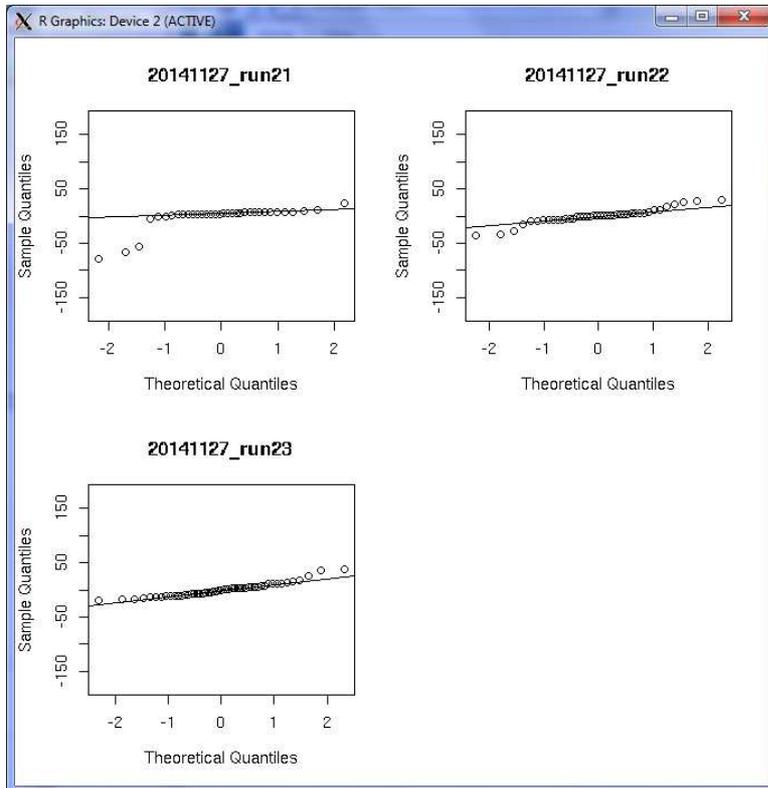


Figure 21.3: Q-Q plots showing how close samples are to theoretical quantiles in a normal distribution