The following text describes how the data for the spreadsheets **in SMTbls_PTPilotStudy** were obtained and analyzed:

Analysis of phrasing of pulse trains in species with discontinuous calling songs.
For each species to be analyzed, five song recordings archived in MLNS were sought. To minimize variation attributable to geographic locality and temperature, the recordings selected for a species were from a single county and at temperatures close to 25°C. To make the recordings selected more representative of the sample for the county's population of the species, no more than one recording per male was allowed. To avoid subjectivity in selecting among multiple qualifying songs, those appearing earliest in a numerical list of TW's *Anaxipha* recordings (TW_TapedSongsMiscSpp.xls) were given priority. For three of the species/county combinations chosen, only four qualifying songs that were of a quality permitting analysis were found.

Selected songs were retrieved into Cool Edit 2000 from the MLNS archive [actually TW used his MLNS copy of his contributions to it]. If needed to facilitate measuring, the song was filtered at its dominant frequency, sometimes with a change in amplitude. Measurements were made by moving Cool Edit's cursor to the beginning and end of each pulse train in the recording and entering in a spreadsheet the three-decimal-place times (from the beginning of the recording) that the software displayed. The duration of each pulse train was calculated by subtracting the time of beginning of the train from the time at its ending. The duration of each pulse train interval was calculated by subtracting the time of ending of the pulse train from the time of beginning of the next pulse train. For recordings sufficiently long, 24 pulse trains and their intervals were measured. If fewer than 14 of each were available for measuring in a recording, the recording was disqualified. In measurements of durations of pulse trains, gaps equivalent to one or two wing-stroke cycles were ignored rather than classed as minimal intervals; in measurements of intervals, a single pulse within the interval was ignored rather than classed as a minimal pulse train.].

The two sets of measured durations (pulse trains and their intervals [PT and PTi) and their sums (pulse-train periods=PTP) were individually sorted by magnitude and the following statistics calculated: median, mean, standard deviation, and coefficient of variation (CV=100xSD/mean). [see SMTbls_PTPilotStudy, msTable n+1]

The data were statistically analyzed by Mihai Giurcanu of the UF-IFAS Statistical Group, but the conclusions were of little help in specifying how to identify species by pulse train phrasing, principally because the songs are highly variable in this respect and TW's samples were too small for the results to be stronger. Here is a copy of an email that reveals the outcome of the project:

On 2/29/2012 11:22 AM, Walker,Thomas J wrote:
> Mihai,
>
> Thank you for your analysis of the data I sent you re the phrasing of
> pulse trains in six N. Amer. Anaxipha species and for attempting to
> explain to me what the analysis meant

Welcome. I think that correct interpretation of statistical results is one of the most difficult steps in data analysis. I think that all analyses have a little (subjective) bias, and in some cases, we do not have any other choice. However, I strongly believe that we have to acknowledge possible limitations of the results and try to diminish the (subjective) bias as much as we can.

>
> Thank you also for pointing out that there must be an error in the
> data for 611-27, one of the ROS individuals. I found the problem and
> made the corrections in the attached version of the workbook. If you
> look at worksheet "phrasing data (3)" you will see the correct numbers
> in red and the former numbers to the left (in black, except for the
> zeros that caught your attention). I got the correct values by
> re-pasting the needed values from worksheet “phrasing data (2)” into
> worksheet “phrasing data (3)” and re-sorting the values one column at
> a time. I entered the new values for 611-27 in the summary table (the
> first worksheet in the workbook) and the changes made little
> difference in what one could conclude from examining the individual
> values or the means for each species.

I thought that there might be an error, but there exists measurement error in almost everything we measure, and I was thinking that the value, although not zero, it might had been so small, that it was approximately zero. This is why I removed the value from the data set (usually I do not remove observations, but since our data set has approximately 750 values, we expect that removing one or two values should not have a dramatic effect on the results).

>
> I hoped that the zeros were reliable indicators of the error I had
> made with the 611-27at data set. To somewhat re-assure myself, I
> re-did the creation of the PT and PTi intervals for 611-10-e and
> 611-36 and found perfect agreement with the original determinations.
> (The re-created values for those individuals are in blue in the
> attached workbook.)

Thank you. I recreated the data set, and now there are no more zeros indeed.

>
> I am not certain I know how to interpret the SAS graphs that displayed
> the results of the GLIMMIX Procedure in your comparisons of the
> pulse-train durations and the pulse-train-interval durations by species.
>
> Is it correct to conclude that, from the data sets you analyzed, all
> 15 comparisons of the six species _except FUL vs. IMI_ are separable
> by pulse-train duration alone, with the probability of a Type I error
> being
> 0.05 or less?

This is a great question. Tukey's multiple testing procedure controls the Family-wise error rate, i.e., the probability of making one or more false discoveries (i.e., the probability to wrongly reject one or more

true null hypotheses). Equivalently, Tukey's multiple testing procedure controls the Type 1 errors among all the hypotheses of the multiple pairwise tests. The assumptions of Tukey's procedure are:

1) the observations are normally distributed (which holds in our case);
2) the design is balanced. However, if the design is not balanced, then the Tukey-Kramer adjustment is used in proc GLIMMIX.
3) the observations are independent (which holds in our case).
4) the errors are homogeneous (this does not hold). However, there is an option SIMULATE in proc GLIMMIX which should take into account all these aspects based on the maximum of a multi-variate t-variable.
The results seem to be the same. However, since the maxim is not a smooth function, I do not know how accurate this method is (this is an area of my research, to study the accuracy of Monte-Carlo methods for non-regular models).

>
> Is it also correct to conclude that, from the data sets you analyzed,
> _FUL vs. IMI_ are separable by pulse-train-interval duration?
>

Yes, your interpretation is correct. Looking more closely at the results, I realized that the means of FULL vs IMI are not separated; however, FULL and IMI are separated in terms of a parameter close to their medians. This follows from the observation that the power -1/4 of the median of PTI is approximately the mean of the PPTI (PPTI is the power-1/4 transform of PTI). However, disregarding the error in this approximation, we can claim that, although not separated in terms of their means, the IMI and PTI groups are separated in terms of a parameter "close" to their median PTI.

> If either of these is incorrect could you refer me to an explanation of
> how that type of graph should be interpreted?
>
> Thanks again and best regards,
>
> Tom
Thank you. Attached you find the analysis of the corrected data set.
I apologize if my statements are vague, but if you need, I could take
a closer look into this problem which is of interest to me as well.

Sincerely,
Mihai Giurcanu

The data set analyzed and the details of the analysis are in three zipped files that TW will send to anyone interested and that he plans to archive on UF's Institutional Repository (IR@UF) under his name [he will do this before submission].