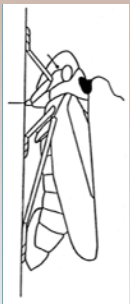


EPG Data Analysis 101

Lecture 2: Variables and Programs

EPG
Work-
shop



Calculating Variables

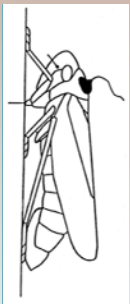
- Psyllids have 6 waveforms arranged in probing and non-probing intervals.

Np, C, G, C, Np, C, D, EI, C, D, EI, E2, C, G, Np, C, G, C, G, C, D, EI, Np

First Probe Second Probe Third Probe

- There are 2 C in the first probe, 3 in the second, and 3 in the third.
 - This insect has 8 C events, or $(2+3+3)/3$ C events per probe

EPG
Work-
shop



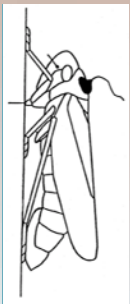
Variables

Np, C, G, C, Np, C, D, EI, C, D, EI, E2, C, G, Np, C, G, C, G, C, D, EI, Np

First Probe Second Probe Third Probe

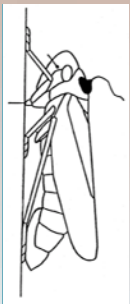
- Each C event has a duration.
 - Total Duration of C: add up the duration of all C events.
 - Mean Duration of C: Take the above sum and divide by 8.
 - Mean duration of C per probe: Sum the first two C, divide by two. Sum the next three C and divide by three. Sum the last three and divide by three. Sum these values and divide by three.
 - Duration of the C before first EI. This is a single value.

EPG
Work-
shop



Sequential and non-sequential

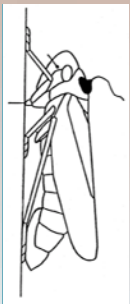
- Non-sequential variables are ones where order doesn't matter.
 - Mean duration of C: I can switch the duration of the 1st C and the 8th C and I will get the same answer.
 - Mean Duration of C in the first probe: I can change the order of C events within the first probe and I will get the same answer.
- Sequential variables are ones where order is important.
 - Mean duration of C per probe: if I switch the 1st and 8th C event, I will not get the same answer unless the two events happen to be the same duration.



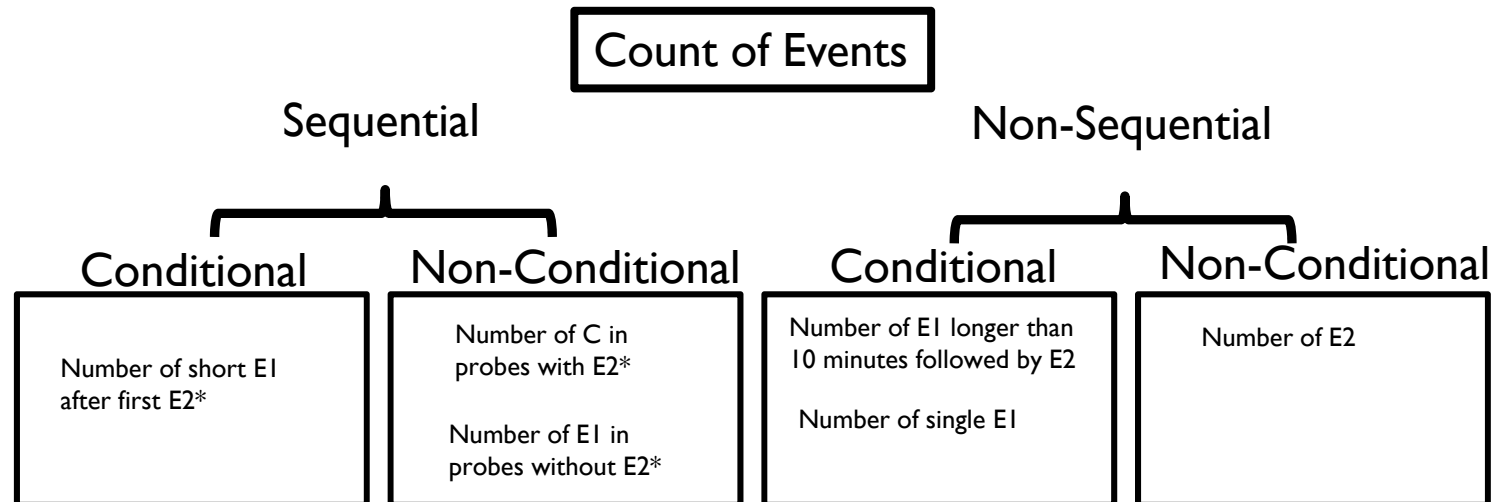
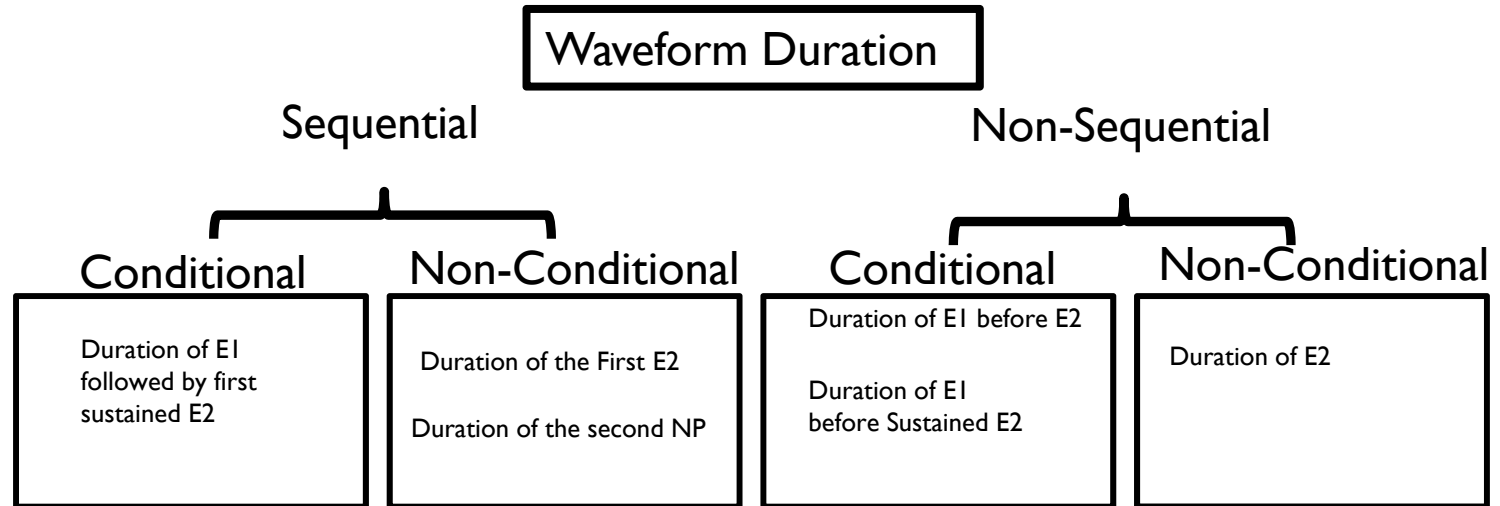
Conditional vs non-conditional

- Non-conditional
 - No restrictions on which values are chosen. If there are eight phloem ingestion events then all of them are used.
 - Mean duration of C: This calculation uses all C events in the recording.
- Conditional
 - Specific events are excluded from the calculation
 - Mean Duration of C in the first probe: This calculation is restricted to only C events in the first probe.

EPG
Work-
shop

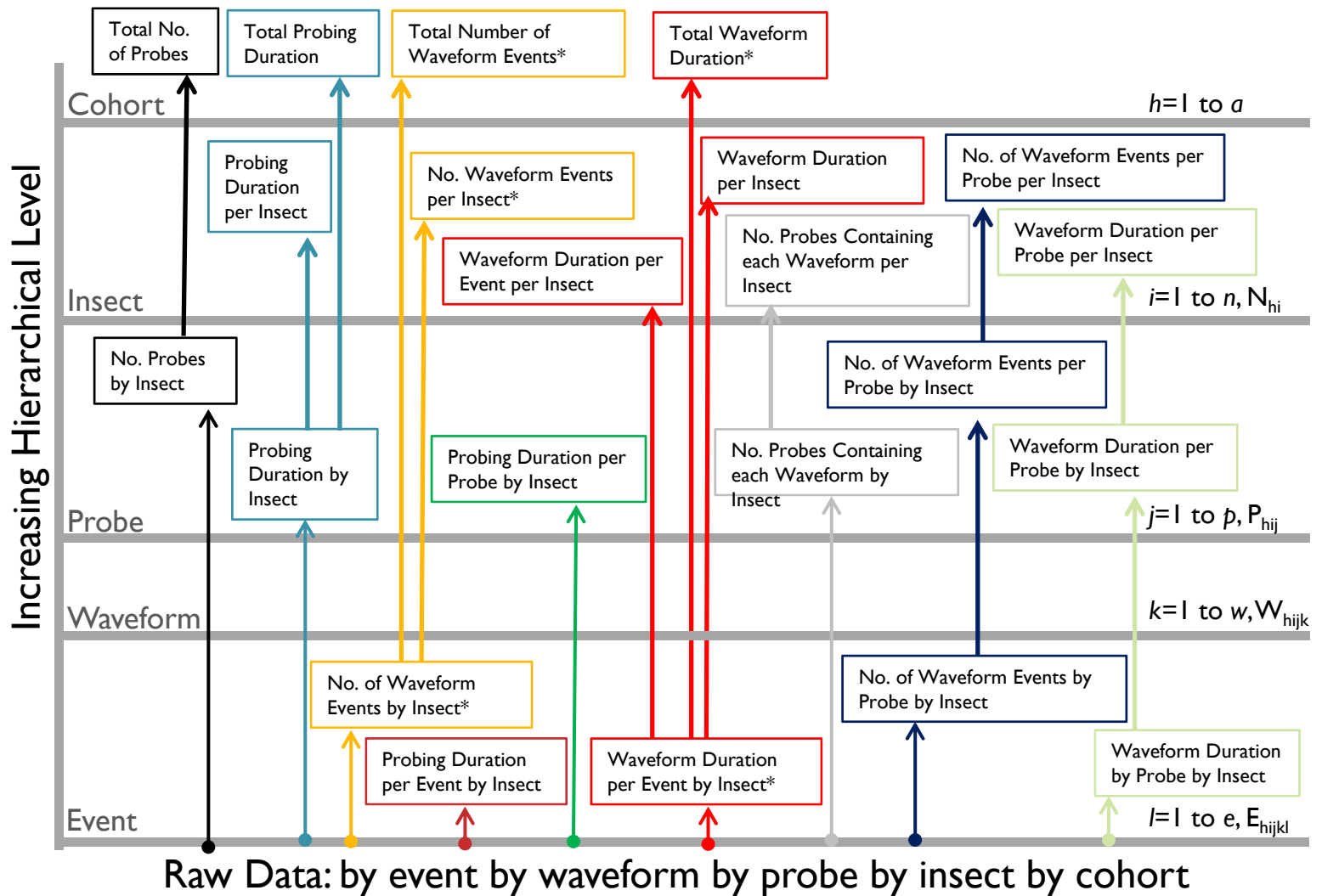


Examples: Conditional::Sequential



* These variables are not currently part of any analysis program.

Hierarchical Structure: Introduction



*See next figure :::: not all intermediate steps are included

** Ebert et al. 2015. A new SAS program for behavioral analysis of electrical penetration graph data. Computers and Electronics in Agriculture 116:80-87

Hierarchical Structure: Experiment

- The Whole Experiment
- Within each Cohort (treatment)
 - Cohort is treatment if there is a treatment.
 - Cohort may be observational, like sex, color or genotype.
- Within each insect
- Within each probe
- A single event

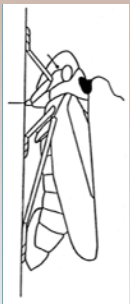
EPG
Work-
shop



Hierarchical Structure: Event

- An event
- Duration of C before first sustained E2
- There is at most a single value for each insect.
- I can average across all insects within a cohort, divide by the number of insects, and get a mean.
- I now have mean duration of C before first sustained E2.

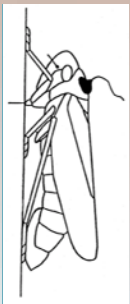
EPG
Work-
shop



Hierarchical Structure: Probe

- Probe
- Duration of C per probe.
- For each insect, calculate the duration of C for each probe, then divide by the number of probes for that insect.
- Calculate means for each cohort.
- Mean duration of C per probe.

EPG
Work-
shop



Hierarchical Structure: Insect

- Insect
- Total duration of G
- Sum the duration of G for each insect.
- Sum this value for all insects within a treatment and divide by the number of insects.
- Mean total duration of G

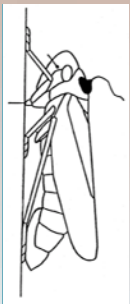
EPG
Work-
shop



Hierarchical Structure

- Experimental unit is the insect.
- One can calculate totals for cohort or experiment.
 - Total Duration of E2 in treatment 2
(where the total is all E2 without respect to which insect each event came from)
 - No statistics

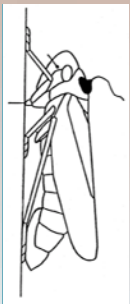
EPG
Work-
shop



A Bounty of Variables

- Mean duration of each waveform
- Number of each waveform
- Average number of each waveform per probe
- Time to first occurrence
- Time to sustained ingestion (E2 and G)
- By hour (Duration of NP in 1st hour)
- Proportions: E2 as a proportion of E1 + E2
- Number of Probes to first
- And many more

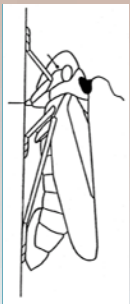
EPG
Work-
shop



EPG Data Analysis 101

Lecture 2: Variables and Programs

EPG
Work-
shop



Primary Available Options

- There are five main programs for calculating EPG variables.
- **Sarria Workbook:**
 - Sarria, E., M. Cid, E. Garzo, A. Fereres. 2009. Workbook for automatic parameter calculation of EPG data. *Computers and Electronics in Agriculture*. 67: 35-42.
- **Backus 1.0:**
 - Distributed through a workshop or by contacting Elaine Backus.
- **EPG-Calc:**
 - <http://link.springer.com/article/10.1007/s11829-014-9298-z>
- **Ebert 1.0:**
 - <http://www.crec.ifas.ufl.edu/extension/epg/sas.shtml>
- **JKL:**
 - available through epgsystems.eu

EPG
Work-
shop



Sarria Workbook

- All are calculated for each insect
- Easy to use
- Used many times
- Output must be moved to a statistical package for analysis
- Best used for aphids where behaviors A and B are not important.

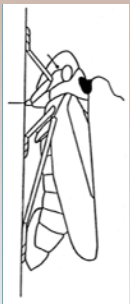
EPG
Work-
shop



Backus 1.0

- Widely used
- Useful on any insect with any behavior
- Calculates variables at different hierarchical levels.
- Is limited to nonsequential and nonconditional variables.
- Is paired with parts of Ebert 1.0 to enable reading text files and error checking.

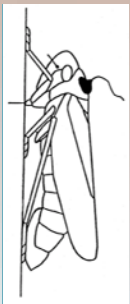
EPG
Work-
shop



Ebert 1.0 (and other versions)

- Completes all data analysis steps using a single software platform (SAS).
- Is open source. You modify the code to fit your project. Current versions are:
 - Ebert 1.0: a mimic of the Sarria Workbook for use with aphids.
 - Ebert 2.0: an adaptation of Ebert 1.0 for use with psyllids.
- It is not easy to use.
- It is expensive.

EPG
Work-
shop



EPG-CALC and JKL

- Similar to others, yet with a few different variables.
- JKL is an Excel workbook.
 - Calculates cumulative “by hour” where Sarria calculates by hour.
- All results must be transferred to a statistical analysis software package.

EPG
Work-
shop





Other Options

- There are a few other programs that were not listed.
- Program the computer to do the analysis yourself.
- The benefit to the last option is that it provides the greatest flexibility to get exactly the output you need.
- The problem is that it is difficult to program. The listed programs have been checked for accuracy in several ways. How will you provide that level of insurance that your new program is providing the correct output?



Disclaimer

- All of the programs discussed (Sarria, JKL, etc...) are free. They have been checked extensively. That does not mean that they are error free. The user is responsible for any problems arising from the use of these programs.
- If you find an error to any program, I am sure the owner would appreciate knowing about it.



Software

- You will be downloading software, and extracting files in the next section.
- You need SAS already installed.
- You need Microsoft Excel[®] and Microsoft Word[®] installed, or OpenOffice[®] installed. I will use Microsoft products, but OpenOffice is equally as good.