

Error Checking: The Issues

- There are many ways problems can arise in the data. Here is a list of those problems for psyllids.
 - The only behavior allowed after Np is A.
 - Unless A is recorded as part of C.
 - The only behavior allowed before E2 is E1.
 - The only behavior allowed before E1 is D.
 - No consecutive waveforms with the same behavior.
 - There can only be one non-probing behavior. There are work-arounds for this in some cases.
 - All recordings start with the non-probing behavior.
 - All durations are positive.

EPG
Work-
shop





Error Checking: outline

- The relevant program is “Error Checker”
- Problems are present when the program fails to produce the correct output, or gives an error message.
- The table that the program should produce helps identify problems with inappropriate transitions: e.g. Np going directly to E2.

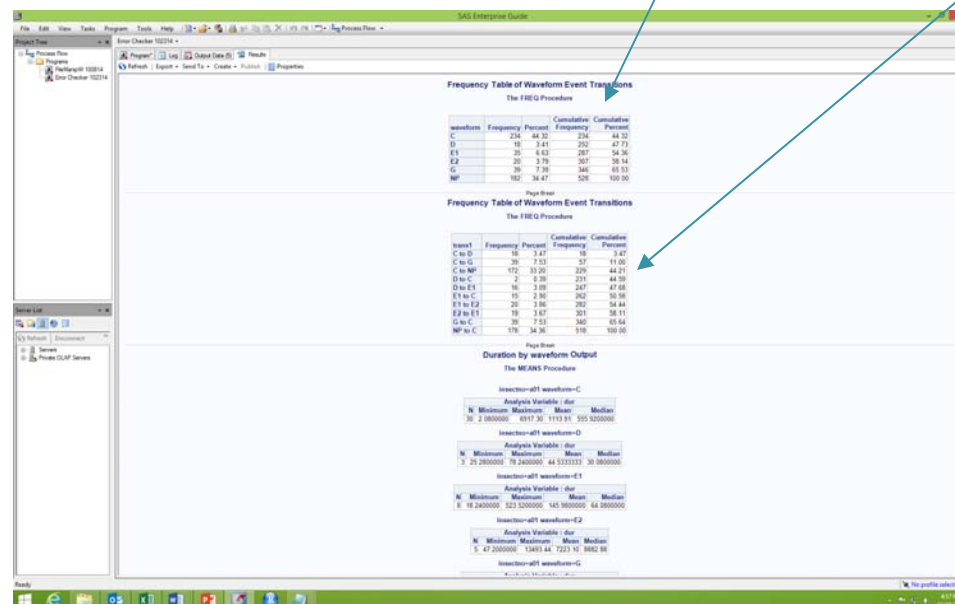


Run the Program

- Open the program “Error Checker.”
- Change the file name in the infile statement.
- Make sure the PsyllidData1.csv file is not open in any other program.
- Run the program.
- These data have no errors that the program can detect.

The Output

- If there are no errors, then the output will have two tables at the top.



Reading the output tables

- The first table

waveform	Frequency	Percent	Cumulative	Cumulative
			Frequency	Percent
C	234	44.32	234	44.32
D	18	3.41	252	47.73
E1	35	6.63	287	54.36
E2	20	3.79	307	58.14
G	39	7.39	346	65.53
NP	182	34.47	528	100

- Waveform is a list of all waveforms in the file. Capitalization matters, so NP is not the same as Np in this table.
- Frequency is the number of times that waveform appears in the data.
- Percent is the relative contribution of each waveform to the total of all observed behaviors.



The first table

- This table is useful for finding typos for Windaq users. Windaq will not consider waveform X as a mistakenly entered version of waveform C.

The second table

- Look carefully at all the transitions (first column).

trans1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
C to D	18	3.47	18	3.47
C to G	39	7.53	57	11
C to NP	172	33.2	229	44.21
D to C	2	0.39	231	44.59
D to E1	16	3.09	247	47.68
E1 to C	15	2.9	262	50.58
E1 to E2	20	3.86	282	54.44
E2 to E1	19	3.67	301	58.11
G to C	39	7.53	340	65.64
NP to C	178	34.36	518	100

- The first row shows that there are 18 cases where the psyllid was in C and went to D.



One last error

- This process does not check to make sure that all recordings start with N_p .
- This is a relatively minor error and will have minimal consequences.
- The fastest fix is to go through Excel.



Start with NP

- Open PsyllidData1.csv in Excel.
- In column D, cell 2 type in the formula
`=IF(A2=A1,"",IF(B2="C",1,""))`
- Move cursor to cell C2 and hit “End” and then ↓.
- Move cursor to cell D259, hit “Shift End ↑” to select all cells (you have to hold the shift key, the end key does not have to be held).



Finishing

- Fill down (Control D)
- In cell D1 type in =Sum(D2:D259)
- The result is zero, so there are no problems.
- Solution: If NP is not an important behavior to your research, I would suggest adding an NP to the first insect and selecting a duration of 0.5 seconds.
- Delete column D, and save the file.