Stations and White Light Sunspot Region Questions

(Questions in black. Response in red.)

1. What stations should we use? All of them that are available, or only a subset (either can be implemented).
   1. All of them that are available
   2. In our approach to extracting station records, we can offer the option of grouping sunspot data by station. That may be helpful for validating station output. Can you anticipate other ways to extract/group this data that are useful for you?
   3. Yes, Judith will provide code to read other sunspot area records (other than USAF). Eventually, we’d like the potential to compare/contrast the sunspot blocking function based on these different records. The differences could be used to define an uncertainty.
2. Duplicates in the sunspot records? Judith’s code checks for these. Can we assume there are no duplicates (after Bill Denig’s automated code steps)?
   1. No, we should check for duplicates. It would have science impact, if we counted a sunspot area twice.
   2. Include a flag indicator to carry through the code as a quality assurance control; essentially alert us to questionable quality data.
3. Observations per day by station: Can we assume each station only makes on observation of a sunspot group per day? (Note, we haven’t seen >1 obs per day).
   1. The USAF dataset does provide a time. Judith isn’t aware of a station that reports more than once per day, but we should monitor for that.
   2. Again, let’s implement a flag indicator if we observe this condition
4. The station, MWIL, does not compute a sunspot area. Should we always exclude this station?
   1. This is expected occurrence for MWIL station. However, as part of separate funding project, someone at Mt. Wilson did (for awhile) did produce a sunspot index.
5. Other stations with missing area. Should we exclude the whole station (i.e all sunspot areas from that station), or just that particular sunspot group? Does excluding just the sunspot grouping introduce a bias in the area averaging?
   1. We should count the area as ‘0’, but not exclude the whole station. We should also use the flag indicator to mark this condition. In past analysis, Judith found more harm than good in estimating an area for the sunspot. But, in our approach to data extraction and QA flagging we could, in the future, investigate this more rigorously and use the findings to better report uncertainty.
6. Currently, the date, lat/lon, sunspot area, and station id are extracted from Bill Denig’s files. Would you like any other data extracted if it’s useful for you (station quality flag, etc.) in validating station output.
   1. ??
7. Potential for excluding data: Sometimes sunspot region data is organized by sunspot grouping and not by date. We have accounted for this in code. But, how should we deal with the fact that a yearly data file (for example, 1992) will have early 1993 data in it? Currently, the code would exclude the 1993 data when processing year 1992, but it would also miss including this early-1993 data (stored in 1992 data file) when it processes year 1993.
   1. The reason for continuing into the next year with the sunspot data record, is the continuity of monitoring a sunspot grouping (i.e. it takes some time to rotate off the disk).
   2. If we don’t correctly merge data from different years, we could introduce an error/uncertainty on a particular day.
   3. I didn’t speak with Judith/Marty/Peter about the various ways to merge the data (i.e. points d and e below).
   4. Our current IDL procedures can’t handle this easily. (But, in LaTiS it’s simple).
   5. But, we could improve the IDL code to check for this and handle it rigorously.
   6. Other options:
   7. Could Bill Denig address this in his creation of the data files?
   8. See point a) above. The inclusion of data from a follow-on year is intentional to provide continuity for people who are monitoring a particular sunspot group.
   9. Doug Lindholm says we can offer to Bill Denig to run the LaTiS server on his server and we could present the dataset as if it were one single file (this would be ~15 minute effort for Doug to implement).
   10. See h)
8. Handling Time: We made a function to convert Judith’s ‘yymmdd’ string format to Julian date. By doing this, we can a) apply the betasun conversion (LASP has a function to compute this wrt Julian date) without the extra logic required to handle leap years using a fixed input file, and b) for output, we can convert from Julian date back to calendar time in ISO format (or any format desired) that the user requires.
   1. I brought up part of this with Judith. The computation of the betasun angle with Julian date is great. Judith, are there any other formats for date/time output you would like (in addition to the ISO calendar time that NOAA would like?)
9. Format of output: Does the header currently have all the information you want when outputting the sunspot blocking function? What would you prefer differently in the header, if anything? Can we automate the header output? We can make it so the output can be created for an arbitrary time range. Would you want the output of the SSB code to include anything else in addition to date, mean ssb, stddev ssb, mean ssbuv, stddev ssbuv.
   1. Judith?