EARTH OBSERVING LABORATORY

Memorandum:

14 October 2015

TO:	HIPPO reprocessing file
FROM:	William Cooper
SUBJECT:	Workflow to find sensitivity coefficients

1 Purpose

This memo describes the workflow for finding new sensitivity coefficients characterizing measurements of angle of attack using the radome gust system. The example used is “AKRDforHIPPO1.Rnw”, a file containing both text and R code. That file generates the memo “AKRDforHIPPO1.pdf” and the archive file “WI-HIPPO1.zip” that contains the generating file, the document in PDF format, and this workflow memo.

The final authority regarding workflow is the ".Rnw" document itself, but this overview and diagram is intended to help explain the workflow at a general level and so should substitute for reading the R and LaTeX code. The intent is to describe the workflow in sufficient detail to support replication of the analysis and figures presented in “AKRDforHIPPO1.pdf” and also to enable changes based on new data or new analysis approaches.

There are references and citations for the tools used (R, RStudio, knitr) in this technical note: Characterization of Uncertainty in Measurements of Wind from the NSF/NCAR GV Research Aircraft. The workflow document for that technical note also contains some additional explanations of the procedures that are used again in generating the present memo and the steps taken to ensure reproducibility of the results.

2 Acquisition of the primary data

The best measurements for the purpose of this analysis are those obtained during “speed runs,” constant-altitude flight maneuvers where the flight speed is varied through the flight envelope of the aircraft,, preferably with modest rates of acceleration and deceleration. Those maneuvers cause the angle of attack to chance with airspeed through the normal range of measurements. Two aspects of the maneuver are crucial, that the vertical wind be near-zero and that the *geometric* altitude be constant.

The measurements used in this case were collected using the NSF/NCAR GV research aircraft during the first set of flights in the HIPPO project. A description of the data-acquisition process and on-board data system, as well as post-processing software, is contained in this technical note: Characterization of Uncertainty in Measurements of Wind from the NSF/NCAR GV Research Aircraft (called the “Wind Uncertainty Technical Note” in the remainder of this document) and in the workflow memo for that technical note. The resulting data files have netCDF format and contain measurements in scientific units along with variable attributes that give brief descriptions of each measurement. These are the data archives used for the processing described by this workflow document.

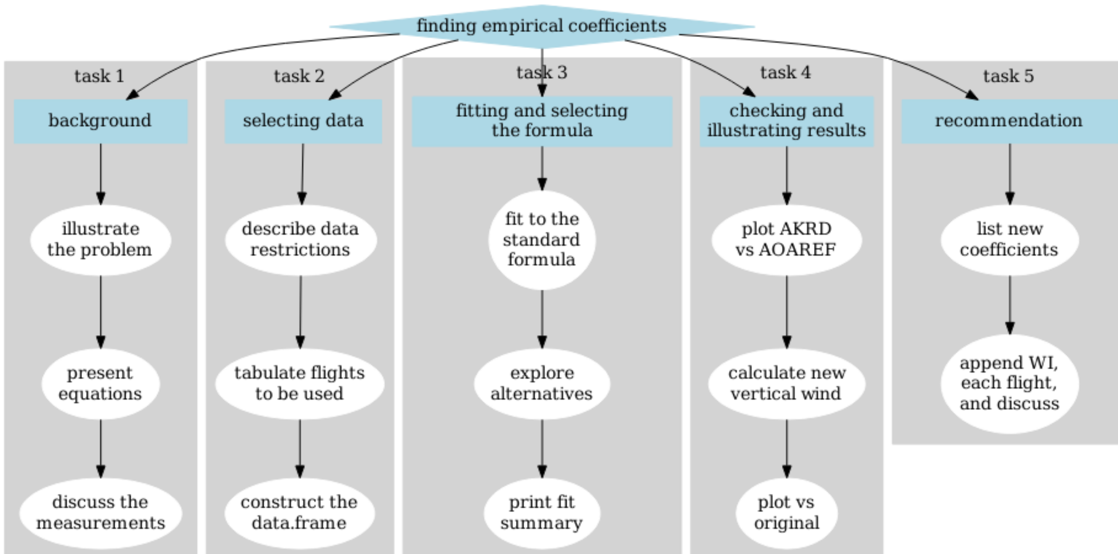

Figure 1: Flow chart describing the workflow.

3 Overview of the workflow

A flow chart describing the workflow is provided by Fig. 1. The workflow is organized into five tasks, each of which is discussed below. Everything in this flow chart, both generating the text document and performing the calculations, is embedded in AKRDforHIPPO1.Rnw. Where particular tasks require calculations using R, the R code is isolated in "chunks" that have header titles, so those titles are referenced below where appropriate. The first chunk is "initialization," which loads some required R packages. An important one is "Ranadu", a set of routines used for accessing, manipulating, and plotting variables in the archived data files. This package is available at this github link and must be installed. Also required is "knitr" so that the text document can be constructed. A list of required variable names is also contained in this chunk; later, when data.frames are constructed for analysis, this list is used to determine the variables to load from the netCDF files.

In the case of HIPPO-1, the rate-of-climb variable is VSPD_A, provided by the GPS measurements from the GV avionics system. After the SPRITES-II project in July 2013, the variable GGVSPD became available, and this is preferable to use when it is present because it does not have filters and associated delays imposed as does the avionics value. For projects after this date, the routine should be modified by replacing VSPD_A by GGVSPD.

3.1 Task 1: Background

1. This task starts by constructing two plots to show the problem that is being addressed. The plots are constructed in the R chunk “vw-rf05”:

   (a) The first is generated via a call to “Ranadu::plotWAC()”, after constructing a new smoothed variable using the function “SmoothInterp()” defined in the “initialization” chunk. The plot is placed automatically in the generated LaTeX file by “knitr” and is referenced by the figure number “\ref{fig:vw-rf051}” where the last digit refers to the first plot generated by the chunk “vw-rf05”. A mean-value line is also added to this plot to show the degree to which the vertical wind appears to be biased. Note that parts of Task 2 must precede the generation of this plot because assembling the data for plotting relies in the data-access steps described in Task 2.

   (b) The second compares the variable AKRD to the reference value provided by (2), to show the unacceptable offset and also to show that there is a dependence on Mach number not addressed correctly by the standard fit. This plot is generated by standard R graphics call “plot()” and referenced in the text via “\ref{fig:vw-rf052”. Note that the figure captions for these two figures are provided by “FCap1” defined in the “get-data” chunk.

2. Next the equations that provide the basis for the fit are presented. These are discussed in more detail in the Wind Uncertainty technical note referenced above and in the Processing Algorithms technical note, section 4.7.

3. Some discussion of the variable used for rate-of-climb is added to the text. This should be modified if GGVSPD is used in future cases, in which case the discussion of differentiating GGALT to obtain an alternate variable should be omitted. The Wind Uncertainty technical note discusses the properties of the different measurements of rate-of-climb.

3.2 Selecting data

1. *Data restrictions:* Some study of the measurements will be needed to select data to which an equation like (1) will be fitted. In the case of HIPPO-1 I selected a subset of flights that excluded flights (1, 2, 8) where plots of the angle-of-attack indicated problems, either because of slow flight (when flaps might be deployed), flight in turns (where generally the quality of wind measurements decreased), or where the measurement ADIFR showed indications of blockage. Many of the latter cases were already flagged as missing in the data files, so those regions did not need to be excluded in any special way. This is described and justified in the text.

2. *Tabulated flight periods.* For reference, the data periods used are shown in a LaTeX table. This was constructed by using “Ranadu::getStartEnd()” to find the start and end of flights after data restrictions were imposed. This step was carried out outside the normal processing

of this file, so if a table like this is desired it must be generated by finding the data periods to use and manually editing the LaTeX table.

3. *Construct the data.frame.* The R chunk "get-data" was then used to read the selected flights and construct two data.frames, one containing all the flights and all measurements and a second containing only periods meeting the restrictions imposed for airspeed, roll, and flight number. To construct the data.frame, the script "Ranadu::getNetCDF()" was used to read the desired variables for each individual flight, and the results were concatenated using "rbind()" to get a single data.frame. Within that data.frame, each flight was tagged with a variable "RF" set to, e.g., 1 for flight rf01, so that subsets of the data.frame could be selected to examine individual flights.

   (a) The first time the data.frame was constructed it was saved in a local file named "WI-HIPPO1.Rdata".

   (b) For subsequent processing, this file was recovered via a 'load()' command to save time. The .Rdata file is also saved as documentation of the status of the data archives at the time of this processing, to protect against changes that may occur when files are reprocessed. Within the saved data.frame (named "Data"), the attributes associated with the netCDF file and with the individual variables are also saved so this file provides a reference for the status of processing that produced the archive.

3.3 Fitting and selecting the formula to use

1. *Fit to the standard formula.* This task begins by finding the best fit coefficients using the standard formula (1). The R chunk "fits" performs this fit, after the subset data.frame "DF" is defined to apply the data restrictions to the data.frame "Data". The linear-model R function "lm()" is used to find the coefficients, and the resulting coefficients are incorporated into the text via "\Sexpr{}" expressions included in the LaTeX text so that, if there is any change, the coefficients quoted will be updated automatically.

2. *Explore alternatives.* If the fit does not appear to be adequate, alternative approaches should be explored here. They might involve using different parameters in a formula like (1) or even applying filtering to the resulting vertical wind in cases where no good fit procedure is available. (This was the case, for example, in HIPPO-5.) In the present case, the standard deviation (0.1 °) was quite good in comparison to most cases and the fit provided coefficients with low uncertainty, so this fit seems fully acceptable. If it had been unacceptable, some other alternatives like those explored in the Wind Uncertainty Technical Note. Some of the possibilities explored in connection with the present study are listed in the text, but none led to sufficient improvement to justify the added complexity of the formula.

3. *Print fit summary.* A special function "SummarizeFit()", defined in the chunk "summarize-fit," is used to write a summary of the results into the text file.

3.4 Checking and illustrating results

1. *Plot AKRD vs AOAREF.* To illustrate the nature of the fit, a scatterplot of measurements of AKRD from (1) with new coefficients vs. AOAREF determined from (2) is shown as Fig. 3. This plot was generated in the R chunk "summarize-fit".

2. *Calculate new vertical wind.* The revised coefficients are used to calculate new values of the vertical wind, here called "WIX", in the R chunk "new-vw". There is code included in this chunk to calculate the vertical wind in two ways, one by correcting WIC as expected for the correction to AKRD and the second, as a check, repeating the full transformation-matrix calculation as in the standard processing. These did not differ, so the variable WIX is used to illustrate the new vertical wind. This value of WIX was calculated for all flights and all data periods, not just those used for the fit.

3. *Plot vs. original.* To show the improvement resulting from this new fit, the new vertical wind was plotted along with the old vertical wind for one flight (rf05) in Fig. 4. That plot is generated by a "Ranadu::plotWAC()" call in R chunk "new-vw".

3.5 Recommendation

The final task is to provide the recommended coefficients to use for processing HIPPO-1 flights. These again reference the results of the fit using "\Sexpr{}" statements to ensure consistency.

Appended information

The end of the memo includes a "Reproducibility" table that documents what is archived, where it is located, and how anyone could retrieve the program, text, and data and repeat the analysis. This would need to be changed if this file is applied to a new project, and care should be taken to avoid overwriting the github archive by failing to change the names of the archived files. The program itself generates the file "WI-HIPPO1.zip" but this and the other archived files must be uploaded to github manually.

To provide additional information for those reviewing these flights and this reprocessing, plots showing the new vertical wind for each flight are appended to this memo. These are generated in the chunk "all-flights", which appears after the closing statement for the memo so that they will be appended after that and after the "reproducibility" information. As for other plots in this memo, the function "Ranadu::plotWAC()" is used to generate these plots.

– End of Memo –