

WSRC-STI-2007-00613

# **Summary of Data Processing for the 2002-2006 SRS Meteorological Database (U)**

Washington Savannah River Company  
Savannah River Site  
Aiken, SC 29808

---

Prepared for the U. S. Department of Energy under contract no. DE-AC09-96SR18500

**Disclaimer**

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement recommendation or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## Summary of Data Processing for the 2002-2006 SRS Meteorological Database (U)

E. D. Kabela and A. H. Weber

Approved by:

 12/13/07

**E. D. Kabela, Author**  
**SRNL Atmospheric Technologies Group**

 12/13/07

**A. H. Weber, Author**  
**Consultant, SRNL Atmospheric Technologies Group**

 12/13/07

**M. J. Parker, Technical Reviewer**  
**SRNL Atmospheric Technologies Group**

 12/13/07

**C. H. Hunter, Manager**  
**SRNL Atmospheric Technologies Group**

Washington Savannah River Company  
Savannah River Site  
Aiken, SC 29808

## Summary of Data Processing for the SRS 2002-2006 Meteorological Database

### Introduction

Every five years since the mid-1970s, the Atmospheric Technologies Group (ATG) of the Savannah River National Laboratory (SRNL) has compiled an updated meteorological database in order to facilitate dosimetric calculations of accident or routine release scenarios for onsite and offsite populations. This meteorological database includes wind speed, wind direction, temperature, dew point, and horizontal and vertical turbulence intensities. The two most recent databases prior to the current one were completed for the time period 1997-2001 (Weber, 2002) and for 1992-1996 (Weber, 1998). The current database covers the period 2002-2006. The advantage of updating the database every five years is that meteorological observations are steadily growing more complete and less subject to errors with the implementation of better electronic data archiving software and hardware, and improved data quality assurance procedures. Also, changes in the region's climate may be manifest.

The five year database for the time period January 1, 2002 through December 31, 2006, like its immediate predecessor 1997-2001, consists of one-hour time averages of temperature and dew point; and wind speed, direction, and turbulence. The dataset includes a complete time-history for the H-Area tower. (Substitutions from other data sources were used whenever missing or erroneous data were flagged on the H-Area tower.) The relatively greater importance of the H-Area tower is due to its proximity to major site facilities and to the center of SRS. The data bases for the remaining towers (A, C, D, F, K, L, and P) are as complete as possible without any data substitutions.

Meteorological data collection processes described by Addis and Parker (1993) result in fifteen-minute averages of temperature, dew point, wind speed and direction, and lateral turbulence intensity from meteorological instruments on 200 foot towers throughout SRS. The data are stored as fifteen-minute averages the database.

Two ATG field engineers and one trained technician examined plots of the meteorological variables and eliminated bad data for the entire five year (2002-2006) for the current database. Instrument icing (which occurred in 2004 and 2005), lightning damage to sensors, calibration periods, etc., cause erroneous data that are eliminated. Quality assured data capture rates usually exceeded 97% for all towers.

### Description of Procedures and Algorithms

The final database was constructed in a series of more than 20 steps that were applied to the raw data stored in SRNL's ORACLE® database for the site's meteorological towers. During these steps the fifteen-minute-averaged data were examined to ensure they fell within possible measurement limits of the instrumentation and were averaged over hourly periods. As noted above, missing data were substituted for the H-Area tower from other

site towers whenever necessary. If no onsite tower data was available, then linear interpolations of each variable were substituted for periods of time less than or equal to 12 hours. Data were substituted from the meteorological tower at the National Weather Service in Columbia, SC for periods of time exceeding 12 hours (Weber et al. 2003). This resulted in a database with an entry for each variable for every hour during the five-year period for the H-Area tower.

The data were processed using a sequence of computer codes on a computer in the Atmospheric Technologies Center running under the Windows operating system. The primary codes for H-Area database processing total 20 members with a SAS® (Statistical Analysis System) extension. All computer codes in the sequence have an alphanumeric character at the end of the first-level name (starting with \_A) indicating the order in which the codes were run. Figure 1 represents the flow diagram for the steps taken to process the current five-year database.

SAS® is a high-level computer language and statistical processing system for data analysis that is easy to comprehend and follow. SAS® is virtually self-documented for a reasonably computer-literate individual and the codes will be briefly discussed here for the purpose of showing the flow of processing the data and additional documentation. The individual modules are discussed in sequence beginning with the “A” module.

### Details of Code Modules (Fig. 1)

EXTRACT\_ALL\_TOWERS\_A.SAS extracts all the quality controlled 15-minute averaged variables including wind speed, wind direction, temperature, dew point, standard deviation of wind direction  $\sigma_A$ , and standard deviation of elevation angle  $\sigma_E$  from the database containing all 61-m Area towers (A, C, D, F, H, K, L, and P). It ensures data entries or missing values for every possible time period during the year (or leap year) are present. The code also sets the values as missing if any temperature, dew point, or wind speed are close to zero (0.01 degrees Celsius or less for temperature and dew point, and 0.1 meters per second or less for wind speed). This portion of the code is a legacy when sensor failures often appeared as zero values from previous five year dataset compilations and was kept for consistency between each five year dataset. Setting temperature, dew point, and wind speed as missing when close to zero was found to have little affect on the dataset. The values of the variables are range-checked against expected ranges. Acceptable data is then stored in a permanent SAS® data set named INT\_NAME.ALL61\_02,03,04,05,06 (61 refers to the 61-m level where the observations were made). This code also identifies the origin of the data with a one-character variable ORI\_TW for the towers A, C, D, F, H, K, L, and P.

The SAS® statement selection criterion for the data is show below.

```
(LOCID in (1,2,3,4,5,6,7,8) ) and  
(DATYPID in (1,3,5,7,9,11) ) and  
(DQ_CODE is NULL or DQ_CODE='G') ;
```

The LOCID variable identifies the tower: A=1, C=2, D=3, F=4, H=5, K=6, L=7, and P=8.

The DATYPID identifies the variable as follows:

1. 61-m wind direction azimuth (from 15-minute vector average)
3. 61-m  $\sigma_A$  (standard deviation from the 15-minute scalar average)
5. 61-m  $\sigma_E$  (standard deviation from the 15-minute scalar average)
7. 61-m wind speed (scalar average)
9. 61-m temperature (scalar average)
11. 61-m dew point (scalar average)

The statement `DQ_CODE IS NULL OR DQ_CODE='G'` selects only data that the QA team judged as “good data” (Kurzeja, 2002).

(The temperature and dew point were selected at 61-m. There is no dew point at 2-m, only a temperature, so both these variables were selected to be from the same level.)

*The codes discussed from this point forward refer only to H-Area database processing.*

The second code in the sequence, `PLUG_INT0_H_B.SAS`, takes the 15-minute averaged values from `INT_NAME.ALL61_02,03,04,05,06` and inserts values from the tower hierarchy list into any missing 15-minute averaged variables. The tower hierarchy list is as follows: F, C, L, P, K, A, and D. Each 15-minute variable substitution is tracked with a variable, e.g., `ORI_SPD='K'`, `ORI_AZ='L'`, etc. The data set is written to an ASCII file named `H_OUT_YR.DAT` (where YR represents 02 through 06) e.g., as follows:

```
PUT @1DATIME      DATETIME13.
   @15 SPEED 7.2  @22 ORI_SPD  $1.
   @24 AZ     7.2  @31 ORI_AZ   $1.
   @33 SIG_AZ 7.2  @40 ORI_SIGA $1.
   @43 SIG_EL 6.2  @49 ORI_SIGE $1.
   @51 TEMP  7.2  @58 ORI_TEMP $1.
   @60 DP    7.2  @67 ORI_DP   $1.
```

The @ character designates the column of the record and the decimal number following the variable name indicates the format (7.2 means 7 columns wide with 2 decimals). The \$ sign designates a character variable.

The next code in the sequence, `LIST_AREA_SUBS_C.SAS`, reads data from the step above and counts the number of substitutions from the area towers that were entered into the H-Area data set. The purpose of this code is to provide an early warning if an unusually large amount of data is missing or has been substituted. The percentage of substitutions for the year is shown as an integer or as a percentage and can be examined in the output display as a result of running the SAS® code.

The code `HOUR_AVE_WINDS_D.SAS` is the most complicated and important member in the sequence. This code reads in the data from `H_OUT_YR.DAT`, calculates one-hour averages on the variables, and writes the results in a file `HOUR_HYR.DAT`, as follows,

PUT

```

@1  FINETIM DATETIME13.
@15 MSPEED      5.2
@20 MINV        6.3
@27 MVVAZ       5.1
@33 MSIG_AZ     5.1  @39 MSIG_EL   5.1
@45 MEANDER     $1.
@51 MTEMP       6.2
@58 MDEW_PT    6.2  @65 MTSIG_AZ  5.1
@73 ORI_SPD    $1.  @74 ORI_AZ    $1.
@75 ORI_SIGA   $1.  @76 ORI_SIGE  $1.
@77 ORI_TEMP   $1.  @ORI_DP     $1.      ;

```

(similar to the earlier format.) The code also converts the wind direction to a flow vector MVVAZ that lies along the direction the wind is blowing. The code must perform a selection for the ORI\_ variables since in a given hour the four 15-minute averages could have come from different towers. The default ORI\_ variable is set to the last of the four, or the next to last, if the last 15-minute average is missing for some reason. The code uses the lagged values of the ORI\_ variable (called GRI and HRI) to store the source of the data. (If there are one or two 15-minute missing values in a one-hour block, the code selects the last one of the available 2 or 3 to identify the substitution source. Hour periods with less than 30 minutes of data are rejected.) It is important to note the variables FINETIM and DATIME are in Universal Time Coordinates (UTC) and not Eastern Standard (EST) or Eastern Daylight Time (EDT).

Scalar time averages are computed for the variables SPEED, TEMP, and DP. A vector average is computed for the wind vector. The standard deviation values, SIG\_AZ and SIG\_EL (for a one-hour period) are combined by taking the square root of the sum of the squares of the 15-minute values of SIG\_AZ or SIG\_EL, and adding to those the squares of the changes in the mean flow direction (Harris, 1992).

The code TRANSPOSE\_AZM\_LOOK\_E.SAS produces an output that lists all cases of remaining missing hourly values. The output listing must be examined for the relatively small number of times when the missing azimuth periods pass through north and for cases that are too long to be linearly interpolated (greater than 12 hours). Data was missing for greater than 12 hours during two separate ice storms early in 2004 and 2005. It is necessary to hand-calculate the azimuth when it sweeps through north. This is done because, for example, the average of a wind speed from 350° for two fifteen minute periods and a wind from 10° for two other fifteen minute periods in an hour is not equal to an average wind from 180° for the hour.

The code TRANSPOSE\_AZM\_LOOK\_E.SAS also stores a permanent SAS® data set called INT\_NAME.HOUR\_YR which contains the original data with the ORI\_ codes, plus the MIS\_SPD, etc., variables that keep track of all missing hourly data at this point. In a later step (PATCH\_CAE\_DATA\_M.SAS) additional missing data are patched in from the Columbia National Weather Service tower data.

The other codes with similar names in the sequence, e.g., TRANSPOSE\_SPD\_F, TMP\_G, DEW\_H, SGE\_I, AND SGA\_J each work on the one-hour averages of the five variables including wind speed, temperature, dew point,  $\sigma_A$ , and  $\sigma_E$ . Each code counts the missing values of the variables and linearly interpolates between missing values then creates a permanent data set, e.g., INT\_NAME.SPDTRYR, etc, for each of the five variables in turn. A sixth code, TRANSPOSE\_AZM\_K.SAS, works exactly the same as the others except that it also enters any hand-calculated values of wind vector azimuth that were identified from the LOG file of step E above.

The code UPDATE\_L.SAS reads in the original file, e.g., INT\_NAME.HOUR\_YR (created in TRANSPOSE\_AZM\_LOOK\_E.SAS above, along with the ORI\_ variables). UPDATE\_L.SAS takes each of the six stored files (INT\_NAME.SPDTRYR, TMPTRYR, DEWTRYR, SGETRYR, SGATRYR, and AZMTRYR) and updates the variables in the stored file INT\_NAME.FINAL\_YR. During this data step, the variables that have been linearly interpolated, such as the wind speed, are designated with a variable to indicate the origin e.g., ORI\_SPD='X' where X means interpolated. The azimuth values passing through north or any other variable that was hand-calculated are given a value such as ORI\_AZ='W' indicating "hand-calculation".

PATCH\_CAE\_DATA\_M.SAS reads in the data from INT\_NAME.FINAL\_YR from the preceding step and substitutes CAE (Columbia National Weather Service) data if needed. The selected time periods for substitution (identified from TRANSPOSE\_AZM\_LOOK\_E.SAS) are entered into PATCH\_CAE\_DATA\_M.SAS and the code takes the corresponding time period from the CAE data and substitutes them into INT\_NAME.FINAL\_YR.

SUMMARY\_STATS\_N.SAS counts the frequency and computes the percentage of occurrence of the origin of each of the variables (ORI\_SPD, ORI\_AZ, ORI\_SIGA, ORI\_SIGE, ORI\_TEMP, and ORI\_DP) for the five-year period.

FREQS\_POST\_ALI\_O.SAS takes the 8760 data values (8784 for leap years) and counts the frequency of occurrence by sector, Pasquill stability class, and speed class. This code also puts the frequencies in the format needed for dosimetric studies.

### **Pasquill Stability Class Determination**

The Pasquill stability classes for each one-hour period are determined from the value of the variable MTSIG\_AZ that represents the standard deviation of the horizontal component of wind direction. As discussed earlier, the code module HOUR\_AVE\_WINDS\_D.SAS increases the values of  $\sigma_A$ , and  $\sigma_E$  to account for changes in the mean wind direction from one 15-minute period to the other.



The lines of code to determine Pasquill stability class are shown below (EPA, 2000):

```

IF (MTSIG_AZ GE 22.5) THEN DO; PAS_CLAS=1;
    CLASS_A+1; END;
IF (MTSIG_AZ GE 17.5 AND MTSIG_AZ LT 22.5) THEN DO; PAS_CLAS=2;
    CLASS_B+1; END;
IF (MTSIG_AZ GE 12.5 AND MTSIG_AZ LT 17.5) THEN DO; PAS_CLAS=3;
    CLASS_C+1; END;
IF (MTSIG_AZ GE 7.5 AND MTSIG_AZ LT 12.5) THEN DO; PAS_CLAS=4;
    CLASS_D+1; END;
IF (MTSIG_AZ GE 3.75 AND MTSIG_AZ LT 7.5) THEN DO; PAS_CLAS=5;
    CLASS_E+1; END;
IF (MTSIG_AZ GE 2.1 AND MTSIG_AZ LT 3.75) THEN DO; PAS_CLAS=6;
    CLASS_F+1; END;
IF (MTSIG_AZ GE 0.0 AND MTSIG_AZ LT 2.1) THEN DO; PAS_CLAS=7;
    CLASS_G+1; END;

```

The measurements of  $\sigma_A$  are taken with bi-directional wind vanes (bivanes). The SRS measurements are taken in compliance with the ANSI/ANS 3.11 (ANSI, 2005) standards for meteorological or nuclear facilities. The bivanes used at SRS are extremely sensitive to fluctuations in wind direction as long as there is wind flow of at least  $0.5 \text{ m sec}^{-1}$  ( $\approx 1 \text{ mph}$ ). The wind speed at 61 m is normally much higher (only 0.42% have one-hour wind speed averages less than  $0.5 \text{ m s}^{-1}$ ) so the bivanes should represent the actual conditions about 99.58% of the time.

The Environmental Protection Agency (EPA, 2000) has issued monitoring guidance for regulatory modeling applications that are recommended in situations where the wind direction fluctuation measurements  $\sigma_A$  are collected at 10 meters above the surface and the roughness length is  $\approx 15 \text{ cm}$  (EPA, 2000). The rationale behind the EPA recommendations is somewhat complex. At least in part their reasoning is based on the fact that wind vanes often do not function correctly in low wind speeds near the threshold of the instrument (that are likely to occur at night for low heights near the ground such as 10 meters). In fact, at these low speeds a bivane may droop and both the bivane and the more common wind vane overshoot their equilibrium position so that incorrect  $\sigma_A$  and  $\sigma_E$  values are measured. Also, at the surface during daytime convective conditions in strong winds, the turbulence should become more nearly neutral, in spite of the fact that the wind vane may be undergoing wide oscillations due to the passage of convective cells.

To compensate for these perceived measurement problems of wind vanes, the EPA applies a second stage correction to the stability categories after the initial Pasquill categories are determined by the method shown above. During this second stage, the daytime stability categories are forced toward the neutral (D category) in above average wind speed conditions during the daytime, and towards stable conditions (E or F category) at night in light winds, regardless of the actual  $\sigma_A$  measurements on the tower. This happens in every case regardless of the actual  $\sigma_A$  measurements of the height of the vane.

## Statistical Summaries and Area Towers

The code `FREQS_ALL_FIVE_NEW_P.SAS` takes the stored data set from all five years and computes the statistics for each. These results are shown in Tables A, B, and C. These tables, particularly Table C, show that the data capture and quality has remained steady for each variable in the past five years. The wind speed missing data has remained steady and decreased slightly from 1.54% in 2002 to 1.32% in 2006. The wind direction missing data has remained steady and decreased slightly from 1.56% in 2002 to 1.31% in 2006. The  $\sigma_A$  missing data has remained steady and decreased slightly from 0.75% in 2002 to 0.50% in 2006. The  $\sigma_E$  missing data has increased from 0.83% in 2002 to 13.50% in 2003 and decreased from 13.50% in 2003 to 0.51% in 2006. The temperature missing data was steady from 2002 to 2004 (3.22%, 2.97%, and 4.46%, respectively), and decreased to 0.48% in 2006. Dew point temperature missing data increased from 2002 to 2004 (0.29% to 2.47%, respectively) and then decreased overall from 2004 to 2006 (2.47% to 0.64%, respectively).

The code `HOUR_AVE_AREAS_R.SAS` takes all the area towers other than H-Area and computes hourly averages as before. It does not attempt to keep track of missing variables nor does it attempt to substitute for missing values. The code `LIST_AREA_MISS_S.SAS` counts the missing values from the area towers and shows the numbers and percentages in Tables D and E. These two tables show that the best Area towers other than H-Area (in terms of percentage of data captured for all variables) are L, C, P, D, A, F, and K.

The code `FREQS_ALL_FIVE_AREAS_T.SAS` takes the stored data set from all five years and for each of the Area Towers and computes the joint frequency distribution for each. The joint frequency distributions for each of the Area Towers are stored in files named e.g., `FREQS_A_YR.DAT`, and ultimately into the five-year composites named `A_0206.DAT`, `C_0206.DAT`, etc.

The code `FREQS_FOR_PAT_CAP88.SAS` takes the five-year composite of the joint frequency distribution for each tower and calculates the weighted mean of the joint frequency. Once calculated, the file is saved as `K_PAT_CAP88.DAT`, `H_PAT_CAP88.DAT`, etc., and can be used in the CAP88 program.

## Summary

Steps taken to process the meteorological database for the years 2002-2006 are documented. The data were analyzed to show, overall, data capture and quality for measured variables during the five year span remained steady. The wind speed missing data has remained steady and decreased slightly from 1.54% in 2002 to 1.32% in 2006. The wind direction missing data has remained steady and decreased slightly from 1.56% in 2002 to 1.31% in 2006. The  $\sigma_A$  missing data has remained steady and decreased slightly from 0.75% in 2002 to 0.50% in 2006. The  $\sigma_E$  missing data has increased from 0.83% in 2002 to 13.50% in 2003 and decreased from 13.50% in 2003 to 0.51% in 2006. The temperature missing data was steady from 2002 to 2004 (3.22%, 2.97%, and 4.46%, respectively), and decreased to 0.48% in 2006. Dew point temperature missing data

increased from 2002 to 2004 (0.29% to 2.47%, respectively) and then decreased overall from 2004 to 2006 (2.47% to 0.64%, respectively).

The data analysis also showed that the best Area towers other than H-Area (in terms of percentage of data captured for all variables) are L, C, P, D, A, F, and K (from highest to lowest). Two separate ice storms in 2004 and 2005 contributed to a data loss for more than 12 hours which was patched in from the Columbia, SC National Weather Service Forecasting Office.

## Reference

- Addis, R. P., and M. J. Parker: Meteorological Monitoring Program (U), WSRC-TR-93106, WSRC, SRTC, Aiken, SC 29808.
- ANSI 2005: Determining Meteorological Information at Nuclear Facilities, ANSI/ANS 3.11 (2005), American Nuclear Society, La Grange Park, IL 60525.
- EPA 2000: Meteorological Monitoring Guidance for Regulatory Modeling Applications. U.S. Environmental Protection Agency, Office of Air Quality Standards, Research Triangle Park, NC 27711. EPA – 454/R-99-005, 168 pp.
- Harris, S. P., Combining Variances (U), SCS-ASG-92022, March 20, 1992.
- Kurzeja, R. J., 2002: Meteorological Monitoring Procedures, Quality Assurance of Meteorological Data, Manual L15.3, SRTC/NTS, 10/09/02, 9 pp.
- Weber, A. H., 1998: Documentation for (92-96) Five-year Database Software, Interoffice Memorandum, March 12, 1998, SRT-ETS-930073, WSRC, Aiken, SC 29808, 8 pp.
- Weber, A. H., 2002: Summary of Data and Steps for Processing the 1997-2001 SRS Meteorological Database (U), WSRC-TR-00445, WSRC, Aiken, SC 29808, 11 pp.
- Weber, A. H., R. L. Buckley, M. J. Parker, R. P. Harvey, and D. M. Hamby, 2003: The Creation of an Historical Meteorological Database for Environmental Dose Assessment, *Environmental Monitoring and Assessment*, 83, 255-281 pp.

## Appendix A

**Data Substitutions** (Tower hierarchy from most desired to least desired: L, C, P, D, A, F, K, V=CAE (formerly Vogtle), W=Weber (hand calculation for azimuth), X=linear interpolation).

**Table A. Number of hour substitutions in the H-Area data set from each of the towers for the variables by year. The column under H shows the total number of hours initially present.**

Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2002 speed	0	0	0	90	0	0	0	0	0	45	8625
2002 azimuth	0	0	0	91	0	0	0	0	0	46	8623
2002 $\sigma_A$	4	0	0	23	0	0	0	0	0	39	8694
2002 $\sigma_E$	0	0	0	26	0	1	0	0	0	46	8687
2002 temp	0	13	0	269	0	0	0	0	0	0	8478
2002 dew point	0	0	0	25	0	0	0	0	0	0	8735
2003 speed	0	0	2	18	0	1	1	0	0	8	8730
2003 azimuth	0	0	1	19	0	1	1	0	0	10	8728
2003 $\sigma_A$	0	1	0	18	1	0	0	0	0	9	8731
2003 $\sigma_E$	0	2	0	1170	1	1	0	0	0	9	7577
2003 temp	0	0	0	260	0	0	0	0	0	0	8500
2003 dew point	0	15	0	6	0	0	0	0	0	0	8739
2004 speed	0	1	1	28	0	2	0	40	0	12	8700
2004 azimuth	0	1	1	28	1	2	0	40	0	14	8697
2004 $\sigma_A$	1	1	0	21	1	0	0	40	0	12	8708
2004 $\sigma_E$	2	1	1	570	0	0	2	40	0	13	8155
2004 temp	0	0	0	392	0	0	0	0	0	0	8392
2004 dew point	0	0	0	217	0	0	0	0	0	0	8567
2005 speed	0	0	0	23	0	1	0	30	0	0	8706
2005 azimuth	0	0	0	62	0	0	1	30	0	1	8666
2005 $\sigma_A$	0	0	0	64	0	0	0	30	0	0	8666
2005 $\sigma_E$	0	1	0	78	0	0	1	30	0	0	8650
2005 temp	0	0	0	20	0	0	0	0	0	0	8740
2005 dew point	0	0	0	22	0	0	0	0	0	0	8738
2006 speed	0	1	0	115	0	0	0	0	0	0	8644
2006 azimuth	0	0	0	115	0	0	0	0	0	0	8645
2006 $\sigma_A$	0	0	0	44	0	0	0	0	0	0	8716
2006 $\sigma_E$	0	7	0	38	0	0	0	0	0	0	8715
2006 temp	0	0	0	42	0	0	0	0	0	0	8718
2006 dew point	0	0	0	56	0	0	0	0	0	0	8704

**Table B. Percentage of hour substitutions in the H-Area data set from each of the towers for the variables by year. The column under H shows the percentage of hours initially present.**

Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2002 speed	0.00	0.00	0.00	1.03	0.00	0.00	0.00	0.00	0.00	0.51	98.46
2002 azimuth	0.00	0.00	0.00	1.04	0.00	0.00	0.00	0.00	0.00	0.53	98.44
2002 $\sigma_A$	0.05	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.45	99.25
2002 $\sigma_E$	0.00	0.00	0.00	0.30	0.00	0.01	0.00	0.00	0.00	0.53	99.17
2002 temp	0.00	0.15	0.00	3.07	0.00	0.00	0.00	0.00	0.00	0.00	96.78
2002 dew point	0.00	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00	0.00	99.71
Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2003 speed	0.00	0.00	0.02	0.21	0.00	0.01	0.01	0.00	0.00	0.09	99.66
2003 azimuth	0.00	0.00	0.01	0.22	0.00	0.01	0.01	0.00	0.00	0.11	99.63
2003 $\sigma_A$	0.00	0.01	0.00	0.21	0.01	0.00	0.00	0.00	0.00	0.10	99.67
2003 $\sigma_E$	0.00	0.02	0.00	13.36	0.01	0.01	0.00	0.00	0.00	0.10	86.50
2003 temp	0.00	0.00	0.00	2.97	0.00	0.00	0.00	0.00	0.00	0.00	97.03
2003 dew point	0.00	0.17	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00	99.76
Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2004 speed	0.00	0.01	0.01	0.32	0.00	0.02	0.00	0.46	0.00	0.14	99.04
2004 azimuth	0.00	0.01	0.01	0.32	0.01	0.02	0.00	0.46	0.00	0.16	99.01
2004 $\sigma_A$	0.01	0.01	0.00	0.24	0.01	0.00	0.00	0.46	0.00	0.14	99.13
2004 $\sigma_E$	0.02	0.01	0.01	6.49	0.00	0.00	0.02	0.46	0.00	0.15	92.84
2004 temp	0.00	0.00	0.00	4.46	0.00	0.00	0.00	0.00	0.00	0.00	95.54
2004 dew point	0.00	0.00	0.00	2.47	0.00	0.00	0.00	0.00	0.00	0.00	97.53
Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2005 speed	0.00	0.00	0.00	0.26	0.00	0.01	0.00	0.34	0.00	0.00	99.38
2005 azimuth	0.00	0.00	0.00	0.71	0.00	0.00	0.01	0.34	0.00	0.01	98.93
2005 $\sigma_A$	0.00	0.00	0.00	0.73	0.00	0.00	0.00	0.34	0.00	0.00	98.93
2005 $\sigma_E$	0.00	0.01	0.00	0.89	0.00	0.00	0.01	0.34	0.00	0.00	98.74
2005 temp	0.00	0.00	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.00	99.77
2005 dew point	0.00	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	99.75
Tower→ Year, variable↓	A	C	D	F	K	L	P	V	W	X	H
2006 speed	0.00	0.01	0.00	1.31	0.00	0.00	0.00	0.00	0.00	0.00	98.68
2006 azimuth	0.00	0.00	0.00	1.31	0.00	0.00	0.00	0.00	0.00	0.00	98.69
2006 $\sigma_A$	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	99.50
2006 $\sigma_E$	0.00	0.08	0.00	0.43	0.00	0.00	0.00	0.00	0.00	0.00	99.49
2006 temp	0.00	0.00	0.00	0.48	0.00	0.00	0.00	0.00	0.00	0.00	99.52
2006 dew point	0.00	0.00	0.00	0.64	0.00	0.00	0.00	0.00	0.00	0.00	99.36

**Table C. Total percentage of substitutions for the H-Area data set from other towers and sources for each variable by year.**

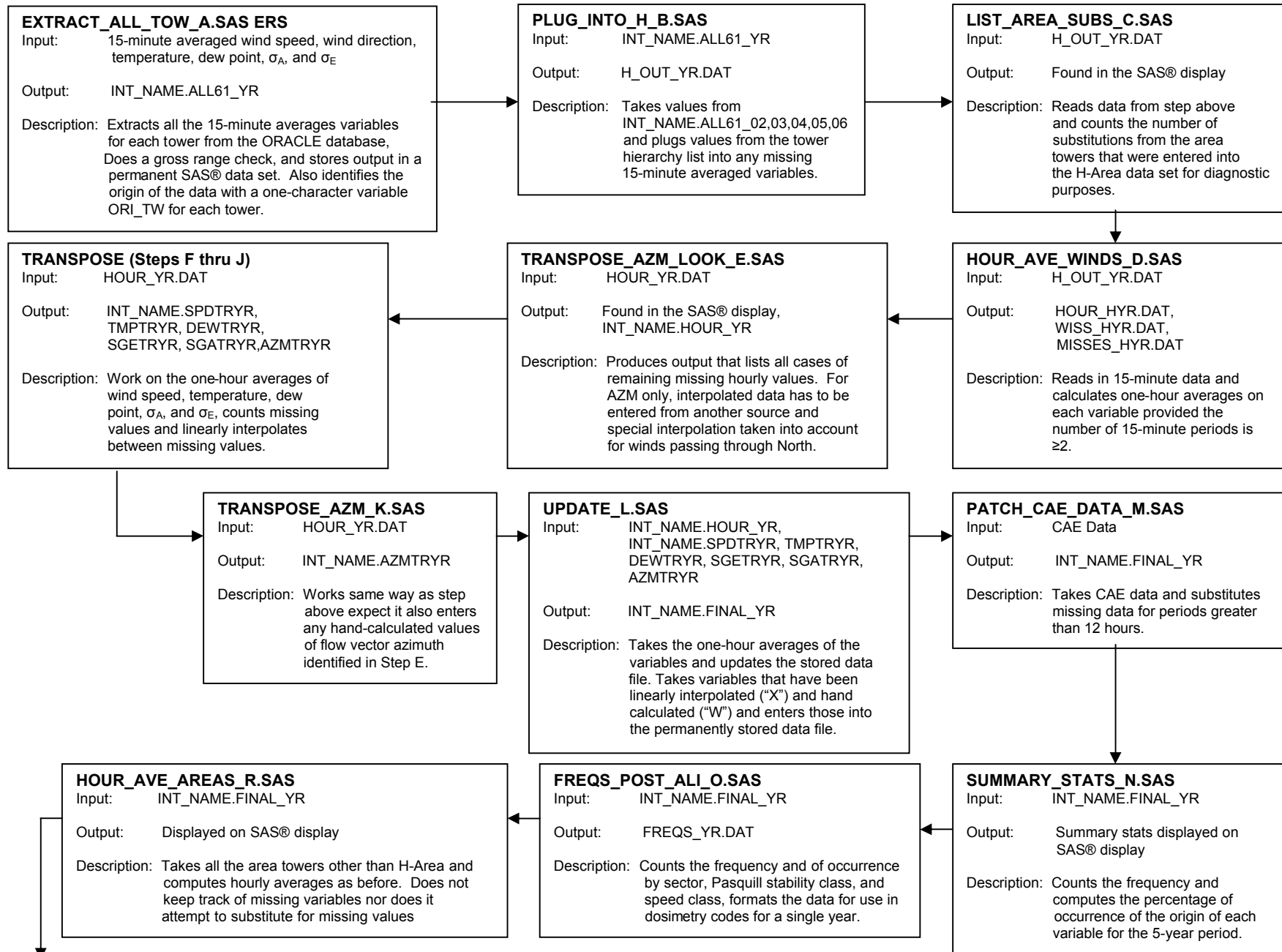
Variable→ Year↓	Speed	Azimuth	$\sigma_A$	$\sigma_E$	Temp	Dew Point
2002	1.54	1.56	0.75	0.83	3.22	0.29
2003	0.34	0.37	0.33	13.50	2.97	0.24
2004	0.96	0.99	0.87	7.16	4.46	2.47
2005	0.62	1.07	1.07	1.26	0.23	0.25
2006	1.32	1.31	0.50	0.51	0.48	0.64

**Table D. Number of missing values for the Area Towers (other than H-Area) for each variable for 2002-2006.**

Variable→ Tower↓	Speed	Azimuth	$\sigma_A$	$\sigma_E$	Temp	Dew Point
A	320	327	295	358	653	5864
C	485	499	355	1598	172	3685
D	953	957	356	391	234	4697
F	623	677	347	2467	180	5010
K	525	530	388	399	2736	8081
L	477	533	350	1500	888	159
P	406	433	411	1382	496	4070

**Table E. Percentage of missing values for Area Towers (other than H-Area) for each variable for 2002-2006.**

Variable→ Tower↓	Speed	Azimuth	$\sigma_A$	$\sigma_E$	Temp	Dew Point	Totals	Rank
A	0.73	0.75	0.67	0.82	1.49	13.38	17.84	5
C	1.11	1.14	0.81	3.65	0.39	8.41	15.51	2
D	2.17	2.18	0.81	0.89	0.53	10.72	17.30	4
F	1.42	1.54	0.79	5.63	0.41	11.43	21.22	6
K	1.20	1.21	0.89	0.91	6.24	18.44	28.89	7
L	1.09	1.22	0.80	3.42	2.03	0.36	8.92	1
P	0.93	0.99	0.94	3.15	1.13	9.29	16.43	3



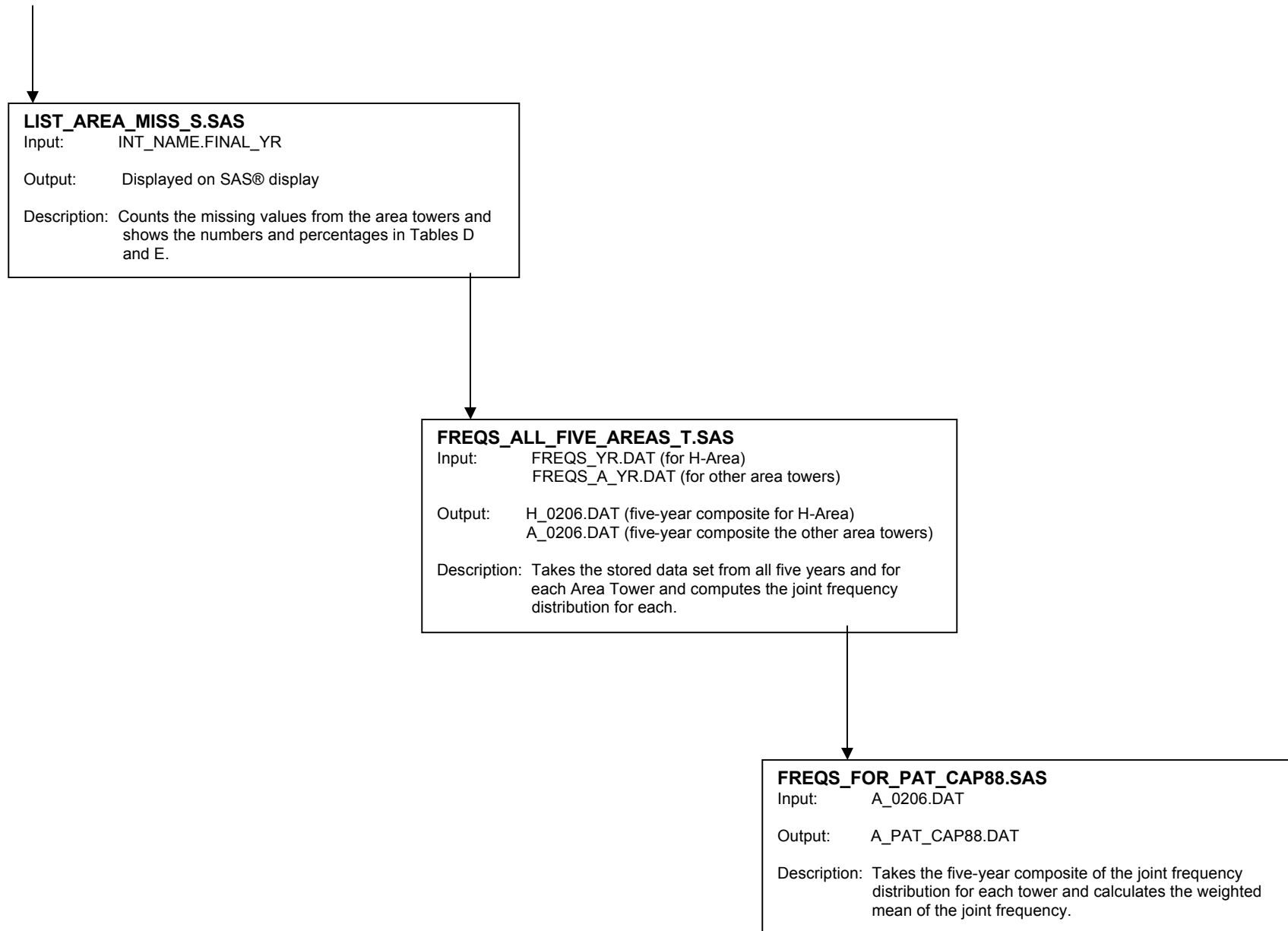


Figure 1. Flow diagram for the steps taken to process the 2002-2006 meteorological database.