WSRC-TR-2002-00445

# Summary of Data and Steps for Processing the SRS 1997-2001 Meteorological Database **(U)**

**Westinghouse Savannah River Company**
**Savannah River Site**
**Aiken, SC 29808**

**Prepared for the U.S. Department of Energy under Contract No. DE-AC09-96SR18500**

This document was prepared in conjunction with work accomplished under Contract No. DE-AC09-96SR18500 with the U. S. Department of Energy.

Summary of Data and Steps for Processing the SRS 1997-2001 Meteorological Database

## Executive Summary

Every five years since the mid-1970s DOE has requested an update on the meteorological conditions at SRS in order to provide dose calculations for accident or routine release scenarios for onsite and offsite populations. The 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 in 1998 for the time period 1992-96 (Weber, 1998) and one for 1987-91 (Parker, et. al., 1992). The current database covers the period 1997-2001. 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, 1997 through December 31, 2001, like its immediate predecessor 1992-96, consists of one-hour time averages of temperature and dew-point; and wind speed, direction, and turbulence. The database has a complete time-history for H-Area Tower with substitutions from other data sources whenever missing or suspicious data were flagged (on the H-Area tower). The data bases for the remaining towers A, C, D, F, K, L, and P are as complete as possible without any data substitution.

Three meteorologists, two field engineers, and two trained technicians examined plots of the meteorological variables and eliminated bad data for the entire five years (1997-2001) for the current database, leaving only missing data to be filled in. Instrument icing, lightning damage to sensors, calibration periods, etc., cause bad data that, for some applications, has to be filled in from alternate onsite towers or offsite measurements. A new Environmental Systems Corporation (ESC) data logger was installed in March 1998 that proved to significantly increase the data capture rate for all towers.

These data are used for running dosimetry codes such as AXAIR, CAP88, and ISC. The data are also used with atmospheric dispersion models in planning, design, and impact assessment of new or existing facilities at SRS. Users of the new data base include the Atmospheric Technologies Group, the Environmental Analysis Section, other sections within the Savannah River Technology Center, as well as outside contractors including Westinghouse Safety Management Solutions.

## 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 SRTC'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 averaged over hourly periods. As noted above, missing data were substituted for the H-Area tower (near the

center of SRS) 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 Plant Vogtle (across the Savannah River from SRS) for periods of time exceeding 12 hours. 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 collection processes described by Addis and Parker (1993) and Parker, et al. (1992) result in fifteen-minute averages of temperature, dew-point, wind speed and direction, and lateral turbulence intensity from meteorological instruments on 200 ft towers throughout SRS. The data are stored as fifteen-minute averages in an ORACLE database on the VAX computer cluster maintained by SRTC. The ESC data logger installed in March 1998 proved to significantly increase the data capture rate for all towers.

The data were processed using a sequence of computer codes on the SRTC1 computer where the ORACLE database resides. The primary codes for H-Area database processing total 20 members with a SAS® (Statistical Analysis System) extension indicating that they are SAS® codes running under the VMS operating system on the SRTC1 computer. 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.

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-documenting 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.

EXTRACT_ALL_TOWERS_A.SAS extracts all the 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). The value of the variables are range-checked to ensure quality data and stored in a permanent SAS® data set named INT_NAME.ALL61_97,98,99,00,01 (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 shown 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 a fifteen-minute vector average)

3. 61-m $\sigma_A$ (standard deviation from the fifteen-minute scalar average)
5. 61-m $\sigma_E$ (standard deviation from the fifteen-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 (including B. Antonicelli, M. Parker, J. Hamilton, R. Holland, S. Patterson, and R. Kurzeja) 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 code EXTRACT_ALL_TOWERS_A.SAS prevents inclusion of time periods when the tower data set contained two records for the same 15-minute time period. This is a holdover from earlier database processing when for some reason two records could be present for the same 15-minute time period. This is not a concern any longer with the new ORACLE database but this code section has been retained just in case the situation should reoccur. This code also prevents exact zero values from entering the data set since these have been found most often to have originated from bogus data. The code sets zero values to missing in case they occur (with tolerance levels as follows).

```
IF (-0.01 <= TMP61M <= 0.01) THEN        TMP61M   = .;
IF (-0.01 <= DPT61M <= 0.01) THEN        DPT61M   = .;
IF (-0.1 <= WD61M   <= 0.1)  THEN        WD61M    = .;
IF (-0.1 <= WS61M   <= 0.1)  THEN        WS61M    = .;
IF (-0.1 <= SIGA61M <= 0.1)  THEN        SIGA61M  = .;
IF (-0.1 <= SIGE61M <= 0.1)  THEN        SIGE61M  = .;
```

The code also stores the 15-minute averaged data in a permanent SAS® data set named INT_NAME.ALL61_97,98,99,00,01 (the last two digits indicating the year).

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

The second code in the sequence, PLUG_INTO_H_B.SAS, takes the 15-minute averaged values from INT_NAME.ALL61_97,98,99,00,01 and plugs 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, D. Each 15-minute variable substitutions 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_97.DAT e.g., as follows:

```
PUT @1 DATIME       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 .LIS file 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_98.DAT, calculates one-hour averages on the variables, and writes the results in a file HOUR_H98.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. @78 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-minutes 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-minutes 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.)

Scalar time averages are computed for the variables SPEED, TEMP, DP. A vector average is computed for the flow vector. The standard deviation values, SIG_AZ and SIG_EL (for a one-hour period) are combined by summing 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). It is necessary to hand-calculate the azimuth when it sweeps through north. For 1997, for example, there were 2 cases, one for 19 March and the other for 16 August.

The code TRANSPOSE_AZM_LOOK_E.SAS also stores a permanent SAS® data set called INT_NAME.HOUR_97 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_VOG_DATA_M.SAS) additional missing data are patched in from Vogtle 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 variable and linearly interpolates between missing values then creates a permanent data set, e.g., INT_NAME.SPDTR98, 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 flow 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_98 (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.SPDTR98, TMPTR98, DEWTR98, SGETR98, SGATR98, and AZMTR98) and updates the variables in the stored file INT_NAME.HOUR_97. A following step creates a new permanent file INT_NAME.FINAL_98. During this data step the variables that have been linearly interpolated such as 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_VOG_DATA_M.SAS reads in the data from INT_NAME.FINAL_97 from the preceding step and substitutes Vogtle data if needed. The selected time periods for substitution (identified from TRANSPOSE_AZM_LOOK_E.SAS) are entered in PATCH_VOG_DATA_M.SAS and the code takes the corresponding time period from the Vogtle data and substitutes in INT_NAME.FINAL_98.

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 by the Environmental Analysis Section for use with AXAIR or other codes running on the IBM mainframe or PC.

**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 $\sigma_A$ 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:

```
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, 2000) 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 m/sec ($\approx 1$ mph). The wind speed at 61 m is normally much higher (only 0.42% have one-hour wind speed averages less than 0.5 m/s) 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 $\approx 15$ 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 droops 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 $\sigma_A$ measurement on the tower. This happens in <u>every case</u> regardless of the actual $\sigma_A$ measurement or the height of the vane.

We prefer the current SRS scheme, since at SRS the bivanes are highly sensitive and are felt to measure quite accurately the $\sigma_A$ needed in the basic atmospheric diffusion equations. Also, the measurements are taken at 61-meters, the level from which most stack releases are expected to occur. Pasquill stability classes derived directly from these measurements should more nearly reflect the true diffusive capability of the atmosphere.

## 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 improved for each variable in the past five years. The wind speed missing data has decreased from 4.1% in 1997 to 0.2% in 2001. The wind direction missing data has decreased from 4.5% in 1997 to 0.2% in 2001. The $\sigma_A$ missing data has decreased from 4.1% in 1997 to 0.2% in 2001. The $\sigma_E$ missing data has decreased from 4.5% in 1997 to 2.2% in 2001. The temperature missing data has decreased from 4.4% in 1997 to 1.1% in 2001.The  Also, the improvement for dew point temperature has been especially dramatic, decreasing from a high of 14.4% in 1998 to 0.4% in 2001.

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 F, P, C, K, A, L, and D.

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_98.DAT, and ultimately into the five-year composites named A_9701.DAT, C_9701.DAT, etc.

## Summary

Steps taken to process meteorological database for the years 1997-2001 are documented. The data were analyzed to show improvements in the data capture and quality for measured variables during this five year span. The wind speed missing data has decreased from 4.1% in 1997 to 0.2% in 2001. The wind direction missing data has decreased from 4.5% in 1997 to 0.2% in 2001. The standard deviation of wind azimuth

($\sigma_A$) missing data has decreased from 4.1% in 1997 to 0.2% in 2001. The standard deviation of elevation angle ($\sigma_E$) missing data has decreased from 4.5% in 1997 to 2.2% in 2001. The temperature missing data has decreased from 4.4% in 1997 to 1.1% in 2001. The improvement for dew point temperature has been especially dramatic, decreasing from a high of 14.4% in 1998 to 0.4% in 2001. The majority of these improvements are more than an order of magnitude, which can be credited of better electronic data archiving software and hardware, improved data quality assurance procedures, and those SRTC meteorologists and technicians who oversee the towers.

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 F, P, C, K, A, L, and D (from highest to lowest).

## References

Addis, R. P., and M.J. Parker: Meteorological Monitoring Program (U), WSRC-TR-93106, WSRC, SRTC, Aiken, SC, 29808.

ANSI 2000: Determining Meteorological Information at Nuclear Facilities", ANSI/ANS 3.11 (2000), 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.

S. P. Harris, Combining Variances (U), SCS-ASG-92022, March 20, 1992.

R. J. Kurzeja, 2002: Meteorological Monitoring Procedures, Quality Assurance of Meteorological Data, Manual L15.3, SRTC/NTS, 10/09/02, 9pp.

Parker, M. J., R. P. Addis, C. H. Hunter, R. J. Kurzeja, C. P. Tatum, and A. H. Weber, 1992: The 1987-1991 Savannah River Site Meteorological Data Base (U). WSRC-RP-92-598. WSRC, 9/15/92, Savannah River Technology Center, Aiken SC 29808, 19pp.

Weber, A. H., 1998: Documentation for (92-96) Five-Year Database Software, Inter-Office Memorandum, March 12, 1998, SRT-ETS-930073, WSRC, Aiken, SC 29808, 8pp.

## Appendix A.

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

**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 | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1997 speed | 0 | 0 | 0 | 316 | 8396 | 0 | 1 | 0 | 0 | 0 | 47 |
| 1997 azimuth | 0 | 0 | 0 | 349 | 8360 | 0 | 0 | 0 | 0 | 12 | 39 |
| 1997 $\sigma_A$ | 0 | 0 | 0 | 318 | 8394 | 0 | 1 | 0 | 0 | 0 | 47 |
| 1997 $\sigma_E$ | 0 | 0 | 0 | 351 | 8362 | 0 | 0 | 0 | 0 | 0 | 47 |
| 1997 temp | 0 | 0 | 0 | 350 | 8367 | 0 | 1 | 0 | 0 | 0 | 42 |
| 1997 dew point | 0 | 8 | 0 | 1074 | 7636 | 0 | 0 | 0 | 0 | 0 | 42 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1998 speed | 3 | 0 | 10 | 140 | 8503 | 1 | 1 | 0 | 80 | 0 | 22 |
| 1998 azimuth | 3 | 0 | 10 | 141 | 8501 | 1 | 1 | 0 | 80 | 2 | 21 |
| 1998 $\sigma_A$ | 3 | 0 | 11 | 95 | 8549 | 1 | 1 | 0 | 80 | 0 | 20 |
| 1998 $\sigma_E$ | 3 | 0 | 10 | 95 | 8548 | 1 | 1 | 0 | 80 | 0 | 22 |
| 1998 temp | 0 | 0 | 0 | 138 | 8529 | 0 | 1 | 0 | 80 | 0 | 12 |
| 1998 dew point | 0 | 23 | 0 | 1142 | 7502 | 0 | 0 | 1 | 80 | 0 | 12 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1999 speed | 0 | 1 | 0 | 118 | 8641 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1999 azimuth | 0 | 1 | 0 | 113 | 8646 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1999 $\sigma_A$ | 0 | 0 | 0 | 91 | 8669 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1999 $\sigma_E$ | 0 | 6 | 0 | 96 | 8658 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1999 temp | 0 | 8 | 0 | 157 | 8595 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1999 dew point | 0 | 17 | 0 | 1528 | 7215 | 0 | 0 | 0 | 0 | 0 | 0 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2000 speed | 0 | 16 | 0 | 49 | 8683 | 1 | 0 | 0 | 24 | 0 | 11 |
| 2000 azimuth | 0 | 16 | 0 | 51 | 8681 | 0 | 0 | 0 | 24 | 0 | 12 |
| 2000 $\sigma_A$ | 0 | 0 | 0 | 71 | 8689 | 0 | 0 | 0 | 24 | 0 | 0 |
| 2000 $\sigma_E$ | 0 | 0 | 0 | 22 | 8735 | 0 | 2 | 0 | 24 | 1 | 0 |
| 2000 temp | 0 | 1 | 0 | 52 | 8707 | 0 | 0 | 0 | 24 | 0 | 0 |
| 2000 dew point | 0 | 0 | 0 | 56 | 8704 | 0 | 0 | 0 | 24 | 0 | 0 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2001 speed | 0 | 0 | 0 | 18 | 8742 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2001 azimuth | 0 | 0 | 0 | 20 | 8739 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2001 $\sigma_A$ | 0 | 0 | 0 | 20 | 8740 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2001 $\sigma_E$ | 0 | 0 | 0 | 189 | 8571 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2001 temp | 0 | 0 | 0 | 94 | 8666 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2001 dew point | 0 | 0 | 0 | 34 | 8726 | 0 | 0 | 0 | 0 | 0 | 0 |

**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 | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1997 speed | 0.00 | 0.00 | 0.00 | 3.61 | 95.90 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.54 |
| 1997 azimuth | 0.00 | 0.00 | 0.00 | 3.99 | 95.49 | 0.00 | 0.00 | 0.00 | 0.00 | 0.14 | 0.45 |
| 1997 $\sigma_A$ | 0.00 | 0.00 | 0.00 | 3.63 | 95.88 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.54 |
| 1997 $\sigma_E$ | 0.00 | 0.00 | 0.00 | 4.01 | 95.50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.54 |
| 1997 temp | 0.00 | 0.00 | 0.00 | 4.00 | 95.56 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.48 |
| 1997 dew point | 0.00 | 0.09 | 0.00 | 12.27 | 87.21 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.48 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1998 speed | 0.03 | 0.00 | 0.11 | 1.60 | 97.07 | 0.01 | 0.01 | 0.00 | 0.91 | 0.00 | 0.25 |
| 1998 azimuth | 0.03 | 0.00 | 0.11 | 1.61 | 97.04 | 0.01 | 0.01 | 0.00 | 0.91 | 0.02 | 0.24 |
| 1998 $\sigma_A$ | 0.03 | 0.00 | 0.13 | 1.08 | 97.59 | 0.01 | 0.01 | 0.00 | 0.91 | 0.00 | 0.23 |
| 1998 $\sigma_E$ | 0.03 | 0.00 | 0.11 | 1.08 | 97.58 | 0.01 | 0.01 | 0.00 | 0.91 | 0.00 | 0.25 |
| 1998 temp | 0.00 | 0.00 | 0.00 | 1.58 | 97.36 | 0.00 | 0.01 | 0.00 | 0.91 | 0.00 | 0.14 |
| 1998 dew point | 0.00 | 0.26 | 0.00 | 13.04 | 85.64 | 0.00 | 0.00 | 0.01 | 0.91 | 0.00 | 0.14 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1999 speed | 0.00 | 0.01 | 0.00 | 1.35 | 98.64 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1999 azimuth | 0.03 | 0.00 | 0.11 | 1.29 | 98.70 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1999 $\sigma_A$ | 0.00 | 0.00 | 0.00 | 1.04 | 98.96 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1999 $\sigma_E$ | 0.00 | 0.07 | 0.00 | 1.10 | 98.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1999 temp | 0.00 | 0.09 | 0.00 | 1.79 | 98.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1999 dew point | 0.00 | 0.19 | 0.00 | 17.44 | 82.36 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2000 speed | 0.00 | 0.18 | 0.00 | 0.56 | 98.85 | 0.01 | 0.00 | 0.00 | 0.27 | 0.00 | 0.13 |
| 2000 azimuth | 0.00 | 0.18 | 0.00 | 0.58 | 98.83 | 0.00 | 0.00 | 0.00 | 0.27 | 0.00 | 0.14 |
| 2000 $\sigma_A$ | 0.00 | 0.00 | 0.00 | 0.81 | 98.92 | 0.00 | 0.00 | 0.00 | 0.27 | 0.00 | 0.00 |
| 2000 $\sigma_E$ | 0.00 | 0.00 | 0.00 | 0.25 | 99.44 | 0.00 | 0.02 | 0.00 | 0.27 | 0.01 | 0.00 |
| 2000 temp | 0.00 | 0.01 | 0.00 | 0.59 | 99.12 | 0.00 | 0.00 | 0.00 | 0.27 | 0.00 | 0.00 |
| 2000 dew point | 0.00 | 0.00 | 0.00 | 0.64 | 99.09 | 0.00 | 0.00 | 0.00 | 0.27 | 0.00 | 0.00 |

| Tower→ Year, variable↓ | A | C | D | F | H | K | L | P | V | W | X |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2001 speed | 0.00 | 0.00 | 0.00 | 0.21 | 99.79 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2001 azimuth | 0.00 | 0.00 | 0.00 | 0.23 | 99.76 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| 2001 $\sigma_A$ | 0.00 | 0.00 | 0.00 | 0.23 | 99.77 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2001 $\sigma_E$ | 0.00 | 0.00 | 0.00 | 2.16 | 97.84 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2001 temp | 0.00 | 0.00 | 0.00 | 1.07 | 98.93 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 2001 dew point | 0.00 | 0.00 | 0.00 | 0.39 | 99.61 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

**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 |
|---|---|---|---|---|---|---|
| 1997 | 4.10 | 4.51 | 4.12 | 4.50 | 4.44 | 12.79 |
| 1998 | 2.93 | 2.96 | 2.41 | 2.42 | 2.64 | 14.36 |
| 1999 | 1.36 | 1.30 | 1.04 | 1.16 | 1.88 | 17.64 |
| 2000 | 1.15 | 1.17 | 1.08 | 0.56 | 0.88 | 0.91 |
| 2001 | 0.21 | 0.24 | 0.23 | 2.16 | 1.07 | 0.39 |

**D. Number of missing values for the Area Towers (other than H-Area) for each
variable for 1997-2001.**

| Variable→ Tower↓ | Speed | Azimuth | $\sigma_A$ | $\sigma_E$ | Temp | Dew Point |
|---|---|---|---|---|---|---|
| A | 637 | 1069 | 838 | 712 | 639 | 3303 |
| C | 683 | 810 | 725 | 750 | 1547 | 1230 |
| D | 2638 | 2785 | 2672 | 2531 | 2675 | 2587 |
| F | 693 | 888 | 606 | 456 | 505 | 474 |
| K | 828 | 1144 | 973 | 719 | 1961 | 1243 |
| L | 460 | 838 | 735 | 3545 | 433 | 3741 |
| P | 1119 | 1160 | 390 | 398 | 416 | 1418 |

**E. Percentage of missing values for Area Towers (other than H-Area) for each
variable for 1997-2001.**

| Variable→ Tower↓ | Speed | Azimuth | $\sigma_A$ | $\sigma_E$ | Temp | Dew Point | Totals | Rank |
|---|---|---|---|---|---|---|---|---|
| A | 1.45 | 2.44 | 1.91 | 1.62 | 1.46 | 7.54 | 16.42 | 5 |
| C | 1.56 | 1.85 | 1.65 | 1.71 | 3.53 | 2.81 | 13.11 | 3 |
| D | 6.02 | 6.35 | 6.10 | 5.78 | 6.10 | 5.90 | 36.25 | 7 |
| F | 1.58 | 2.03 | 1.38 | 1.04 | 1.15 | 1.08 | 8.26 | 1 |
| K | 1.89 | 2.61 | 2.22 | 1.64 | 4.47 | 2.84 | 15.67 | 4 |
| L | 1.05 | 1.91 | 1.68 | 8.09 | 0.99 | 8.54 | 22.26 | 6 |
| P | 2.55 | 2.65 | 0.89 | 0.91 | 0.95 | 3.24 | 11.19 | 2 |