Central California Air Quality Study
Data Transmittal Format Document for CCOS

Release: CCOS021401A

1.0 Managing CCAQS Data

The Central California Air Quality Studies (CCAQS) comprise two studies, the California Regional Particulate Air Quality Study (CRPAQS) and Central California Ozone Study (CCOS). Data collection for these two studies occurred during a 14-month period, between December 1999 and February 2001. The CCAQS Data Management System was developed to provide Study sponsors participants with quality assured data from these studies. The data are for 3-dimensional meteorological and air quality modeling and data analysis. This system provides data that is compiled, evaluated, documented, and distributed by the Modeling Support Section, Planning and Technical Support Division (PTSD) of the California Air Resources Board (CARB). Common data management and validation conventions have been assembled. To the extent possible, CCAQS field data structures, processing, validation and delivery procedures incorporate new technology and methodology but will be consistent with those established for the long-term database and other ARB data sets from recent air quality studies. The CCAQS Data Transmittal Format Document (CCAQS DTFD) includes the full data file format specification for CCAQS data files that are submitted by data providers. This document is used in conjunction with the CCAQS Reference Tables. These tables provide a listing of data elements (Support_ID, Method_Code, etc.) defined below in this specification.

Data Contractors: An ability to explain or replicate the data processing and quality assurance (QA) that were utilized from Level 0 (raw data) to Level 1A is needed. It is very likely that at some point in the future there will be a need to explain or defend the steps that were followed in the data QA process, which may require some level of reproduction. An electronic copy of your process for each "type" of data will be needed as part of the data submittal process, as it may vary from one type of data to the next. Additionally, you are expected to send a copy of all of the raw data collected during the Study in its 'raw' format along with documentation of the format.

2.0 CCAQS Reference Tables

CCAQS data management conventions and methods build on experience from the 1990 San Joaquin Valley Air Quality Study (SJVAQS) (Blumenthal et al., 1993), the 1995 Integrated Monitoring Study (IMS-95) (Solomon and Magliano, 1997), and the 1997 Southern California Oxidant Study (SCOS-97) (Fujita et al., 1997). The following specifications are maintained by the CCAQS Data Manager and will be available to all

project participants via the Internet. The "lookup" reference tables and data described below will be available at http://www.arb.ca.gov/airways/crpaqs/default.htm. The four most important reference tables for the CCAQS data providers are the Supports, Parameters, and Instruments and Methods tables. These tables will be required for reference to properly format data files for submittal.

Supports Table: This table lists all instrument "supports" used during the Study. Monitoring stations for CCAQS are identified with an ID and code which are accompanied by name, address, coordinates, elevation, etc. (To verify the coordinates and elevations for each site, the field managers will use a Global Positioning System (GPS), pressure-based altimeter, and topographical maps). This table includes information identifying aircraft, such as airplanes and blimps. It also includes the identification of monitoring towers.

Parameters Table: Each observed air quality or meteorological parameter is assigned a unique identification and code in this table. Parameter descriptions, units, averaging time, applicable temperature and pressure adjustments accompany these.

Instrument Tracking Table: This table has identification for where an instrument was used during CCAQS to make air measurements. It also includes Support_ID, range, accuracy, and Instrument_ID of instruments used to acquire measurements whether from a station, airplane, blimp or tower.

Methods Table: The table includes all the characteristics of the CCAQS "Methods". This essentially "describes" how an instrument was utilized to obtain an observation data value (Obs_Value). A method is defined by a concatenated string of the following specific codes 1- 8 below. Included is an example set of codes that could comprise a Method.

1. Parameter_Specie_Code: CR (Chromium)
2. Parameter_Property_Code: ELE (Element)
3. Size_Code: PU0000002500 (Fine Particle Size Fraction)
4. Device_Type_Code: XRA (X-Ray Fluorescence Analyzer)
5. Analysis_Method_Code: XRF (X-Ray Fluorescence
6. Media_Code: TEF (Teflon Membrane Filter)
7. Sampling_Frequency_Code: D1 (Every day 0-23)
8. Sampling_Duration_Code: H24 (24 hr)

Study Flags Table: A table of validation flags (Study_Flags) has been developed as part of the CCAQS database. Data providers and contractors are requested to use this set of flags. New flags can be added to the CCAQS database as needed. A document, *Data Flow, Application of Data Quality Flags, and Data Validation Processes*, provides more information and diagrams that will assist in the application of these flags. A draft of this document is available on the CCAQS web site (on or before Feb. 9, 2001).

Data Sources: Each "source" of data is assigned a unique two-character code that is used to identify it. Data transmittals will carry this code as part of the filename when they are submitted to the CCAQS Database System.

Obs Types Table: This table provides the codes for all of the observation types found during the monitoring of air quality. For example, "SFGAS" is the code for 'surface gaseous'.

Sampling Frequencies Table: This table provides the codes for all of the sampling frequencies found during the monitoring of air quality. For example, the "D3" is the code for 'Every third day, 0000-2300 PST'.

Flight Patterns Table: Defines each type of flight pattern and establishes an associated code.

3.0 Data Conventions

Missing and Invalid Data: For missing data, the value is to be reported as a "Null" value accompanied by a validity flag of "MIS". Invalidated records are to retain the invalidated data value. They are not to be replaced by –99 or null values, but instead are to be flagged as invalid using "INV".

Date and Time: Times are expressed in Pacific Standard Time (PST). Years are reported using four-digit codes, e.g., 1999, 2000, 2001.

4.0 Data Validation Levels

Mueller (1980), Mueller et al., (1983), and Watson et al. (1983, 1989, 1995) define a three-level data validation process that should be mandatory in any environmental measurement study. Data records are designated as having passed these levels by the QA_Level column of each observation data record maintain in the database. These levels, and the validation codes that designate them, are defined as follows:

Level 0 (ZERO): These data are obtained directly from the data loggers that acquire data in the field. Averaging times represent the minimum intervals recorded by the data logger, which do not necessarily correspond to the averaging periods specified for the database files. Level 0 data have not been edited for instrument downtime, nor have procedural adjustments for baseline and span changes been applied. Level 0 data are not contained in the CCAQS Database System, although they are consulted on a regular basis to ascertain instrument functionality and to identify potential episodes prior to receipt of Level 1A data. (Levels 0A and 0B have been used by data contractors and data providers to define data prior to the 1A level.)

[**Level 0 for CCAQS:** In addition to data loggers, CCAQS data are also derived from the collection of air samples that are sent to the laboratory for analyses. At a later time

the resultant "raw" data values may be averaged (e.g., 1 minute data averaged to 5 minute data values, 5 minute data averaged to 1 hour data values, etc.). Therefore, the raw data sampling frequency will often differ depending on what parameter is being sampled and as a result will not correspond to the averaging periods specified for the CCAQS database files. This is particularly true for "supplemental" data like those from AIRS, RAWS, etc. Level 0 data are not contained in the CCAQS database. Level 0 data were used to identify potential episodes for sampling.]

Level 1A (1A): These data have passed several validation tests applied by the measurement investigator prior to data submission. The general features of Level 1A are:

1. Removal of data values and replacement with -99 when monitoring instruments did not function within procedural tolerances; [This is not used in CCAQS data validation process.]

2. Flagging measurements when significant deviations from measurement assumptions have occurred;

3. Verifying computer file entries against data sheets;

4. Replacement of data from a backup data acquisition system in the event of failure of the primary system;

5. Adjustment of measurement values for quantifiable baseline and span or interference biases; and

6. Identification, investigation, and flagging of data that are beyond reasonable bounds or that are unrepresentative of the variable being measured (e.g. high light scattering associated with adverse weather).

[Level 1A for CCAQS: In the CCAQS Database System no removal of data values occurs. Instead the use of data flagging is specified when data is missing or invalid, i.e., there is no replacement of missing values or invalid values with –99. For missing data, the value is to be reported as a "Null" value.

Level 1B (1B): Pre-programmed consistency and reasonability tests are applied by the data manager prior to integration into the CCAQS Database System. Consistency tests verify that file naming conventions, data formats, site codes, variable names, reporting units, validation flags, and missing value codes are consistent with project conventions. Discrepancies are reported to the measurement investigator for remediation. When the received files are consistent, reasonability tests are applied that include:

1. Identification of data values outside of a specified minimum or maximum value;

2. Values that change by more than a specified amount from one sample to the next; and

3. Values that do not change over a specified period.

Level 2 (2): Level 2 data validation takes place after data from various measurement methods have been assembled in the master database. Level 2 validation is the first step in data analysis. Level 2A tests involve the testing of measurement assumptions (e.g. internal nephelometer temperatures do not significantly exceed ambient temperatures), comparisons of collocated measurements (e.g. filter and continuous sulfate and absorption), and internal consistency tests (e.g. the sum of measured aerosol species does not exceed measured mass concentrations). Level 2 tests also involve the testing of measurement assumptions, comparisons of collocated measurements, and internal consistency tests.

Level 3 (3): Level 3 is applied during the model reconciliation process, when the results from different modeling and data analysis approaches are compared with each other and with measurements. The first assumption upon finding a measurement, which is inconsistent with physical expectations, is that the unusual value is due to a measurement error. If, upon tracing the path of the measurement, nothing unusual is found, the value can be assumed to be a valid result of an environmental cause. The Level 3 designation is applied only to those variables that have undergone this reexamination after the completion of data analysis and modeling. Level 3 validations continue for as long as the database is maintained.

A higher validation level assigned to a data record indicates that those data have gone through, and passed, a greater level of scrutiny than data at a lower level. All data in the CRPAQS data set will achieve Level 1B status prior to use in data analysis and modeling. The validation tests passed by Level 1B data are stringent by the standards of most air quality and meteorological networks, and few changes are made in elevating the status of a data record from Level 1B to Level 2. Since some analyses are applied to episodes rather than to all samples, some data records in a file will achieve Level 2 designation while the remaining records will remain at Level 1B. Only a few data records will be designated as Level 3 to identify that they have undergone additional investigation. Data designated as Levels 2 or 3 validations are not necessarily "better" than data designated at Level 1B. The level only signifies that they have undergone additional scrutiny as a result of the tests described above.

5.0 Data Transmittal

For the next two years CCAQS field contractors will be collecting, quality assuring and submitting data to the CCAQS Data Manager for inclusion into the database system. The Data Manager will be interacting with numerous data source providers and CCAQS contractors in an attempt to make data available as early as possible. To make the job of

collecting and processing data into the database more efficient and accurate, data transmittal file formats and transmittal file naming conventions have been established. It is important for CCAQS field contractors to adhere to these conventions for transmittal of Study data, as it will significantly improve the data availability turnaround time. The database system is being designed to automate the handling of known CCAQS data formats. This means that efforts can be concentrated on sources of supplemental data that are outside the Study umbrella, where it is impossible to control the data format.

Study data will be submitted in electronic form using file transfer protocol (FTP) after being validated to Level 1A by the field contractor. The data format will be expected to follow the specifications below and the data files must be approved by the CCAQS Data Manager before acceptance. The built-in database routines will assist the process of screening and quality controlling data files.

The CCAQS Program Manager has approved the data transmittal format for the database described in this section, which supercedes all earlier versions of the Data Transmittal/Submittal Document. All data submitted must be free of inconsistencies with the specified format described in this section. Otherwise, data files will be returned to the field contractor(s) for correction.

5.1 Transmittal File Naming Convention

The CCAQS field contractors and investigators are expected to adhere to the CCAQS File Naming Convention as outlined below. Data file naming conventions include the data source code, date information, sequence identifier number, measurement platform and data validation QA level.

TABLE 1. FTP Transmittal file names are of the form CCYMMDDS.PLL

Filename Position	Indicator	Description	Valid Codes
1-2	**CC**	Data Source Code	See Reference Tables for Data Sources (e.g., **NO**=NOAA Boulder)
3	**Y**	Year code	Last digit for the Year the file was created
4-5	**MM**	Month Number	**01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12**
6-7	**DD**	Transmit Date	Date data file was transmitted, reported 01 thru 31
8	**S**	Sequence Identifier	An alphanumeric character used to uniquely identify a file when multiple files are sent on the same Transmit Date. Use 1-9 (and then A–Z if needed). If only one file is submitted on a given date the Sequence Identifier = 1. If two, the second file would be 2, …, the 10th file would be A, the 35th would be Z.
9	**.**	File extension separator	
10	**P**	Meas. platform	**S** = surface, **U** = upper air, **A** = aircraft

11-12	**LL**	Data validation level	**OA, OB, 1A, 1B, 2A, 03** (see Data Validation Levels section above). If file contains other than **1A** data use a file note to describe the process or provide a electronic copy of the process used.

5.2 Transmittal Files

The CCAQS database system will be automated to enable screening and review of data files in this format to ensure files conform to the following format. The transmittal file layout is similar to what is found currently in banking system electronic file transfers (EFTs). A typical EFT file is comprised of several record types, i.e., a header, body records and footer. The following sections define a similar format and record layout for CCAQS related data submittals. Specific examples of the transmittal file records are also included.

As previously indicated, data will be acquired at set intervals via file transfer protocol (FTP) to a pre-defined location specifically set up for each data field contractor and other data providers. The files will be deposited in the "Incoming/CCAQS" Directory of the FTP site ftp://ccaqsftp@ftp.arb.ca.gov and an email sent notifying the data manager (**gobrien@arb.ca.gov**) of the transmittal. The data providers will submit data files as ASCII comma delimited text files, defined as follows:

1) Records begin at file position 0

2) Records end with ASCII characters 13 (carriage return) and 10 (line feed)

3) Fields are delimited with ASCII character 44 (comma)

4) Character data are enclosed with ASCII character 34 (double quote)

5) The End of File (EOF) is indicated by ASCII character 26 (control z)

A convention for the data transmittal file contents has been designed to identify any potential transmission issues in the observation data files sent by data providers. To ensure that each record along with the transmitted file is accurately received in their entirety, "validity checkpoints" have been incorporated into the file format.

For this system, an observation may contain associated observation notes for individual records or for the more likely scenario where a file note pertains to the entire collection of observations. Each record type encountered in a data file has a special identification code shown in Table 3 below.

TABLE 2. CCAQS "EFT" Record Identification Codes

Code	Record Type	Required or Optional
1	File Header	Required
2	(This code is reserved)	(not used)
3	File Note	Optional (with data file transmittal)
4	(This code is reserved)	(not used)
5	Obs Note Header	Required (with Obs Note transmittal)
6	Obs Note	Optional
7	Obs Note Footer	Required (with Obs Note transmittal)
8	Observation	Required
9	File Footer	Required

5.3 "File Header" Record Layout

The File Header Record identifies the record type, data source, interval, transmittal date, a sequence number, measurement platform, validation level, and the number of air observation records contained in the data file. A data source code will be assigned for each CCAQS data provider. The Transmit_Date is the file creation/submittal date in YYYYMMDD format. The Sequence_Identifier identifies the order of the file in the transmitted sequence of files sent on a specific transmit date.

Note: The field position numbers (FP#) in the left-most column of these tables are not part of the transmittal record and are for reference only.

TABLE 3. Record Layout - File Header

File Header						
FP#	**Field Name**	Data Type	Length	Decimals	Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: File Header = **1**
2	DATA_SOURCE_CODE	Char	2	0	No	Code identifying a source for CCAQS data. (e.g., "**NO**" = NOAA Boulder)
3	SUBMITTAL_TYPE	Char	1	0	No	Indicates that the data records are either directly from the field or from the lab **F** = field; **L** = lab. **Defns.:** *Data records* (**F** or **L**) are observation values (Obs_Value) versus an air sample taken with a canister, filter, etc., which may or may not have an associated observation value(s). Air sample submittals without observation values are "**F**".
4	OBS_TYPE_CODE	Char	8	0	No	Use Obs_Type_Code within the Obs_Types Reference

						Table. **Example:** surface gaseous = "**SFGAS"**
5	AVERAGING_INTERVAL	Char	1	0	No	R = Raw data (not averaged)
						A = 3 hour
						B = 6 hour
						C = 12 hour
						D = 24 hour
						H = 1 hour
						J = Jumps (every other hour)
						V = Hourly but varies, possibly less than 24 measurements per day
						I = Instantaneous (< 1 min.)
						F = 5 minute
						T = 10 minute
						M = 15 minute
						N = 30 minute
						P = Partial hour (< 60 min.)
5	TRANSMIT_DATE	Char	8	0	No	Date the data provider sent file, reported as YYYYMMDD. (years are reported using four-digit codes, e.g., 1999, 2000, 2001).
6	SEQUENCE_IDENTIFIER	Char	1	0	No	Alphanumeric character that uniquely identifies a file when multiple files are submitted on the same Transmit_Date (Start with 1-9 then use A-Z).
7	MEASUREMENT_ PLATFORM	Char	1	0	No	S = surface; U = upper air; A = aircraft
8	VALIDATION_LEVEL	Char	2	0	No	Data (file) validation level as defined within this document, e.g., OA, OB, 1A, 1B, 2A, 03 (see Data Validation Levels section above)
9	OBS_RECORDS	Num	7	0	No	A count of the number of Obs Records in the file.

Example 1. File Header

Contractor = NO (for NOAA Boulder). On Sept. 28, 2000, NOAA with 57,347 1-hour surface observations coming directly from the field submitted, the third transmittal file for the day. The following record shows the proper identifying File Header structure.

Note: Space is added between fields for clarity purposes only; a standard ASCII record would not contain spaces.

1, "NO", "F", "SFGAS", "H", "20000928", "3", "S", "1A", 57347

5.4 "File Note" Record Layout

To include a "File Note" within the transmitted file that pertains to the entire data file, then a record in the submitted file must identify its presence. When a File Note is included, it cannot exceed 200 bytes (200 characters) in length and must be broken into multiple records as needed. Typical PC based data processing maximum file record length limits vary, but all systems can handle 255 bytes. This system will use this as its maximum length; therefore, to allow for overhead, notes can be no more than 200 characters long. If multiple File Note records are required then the Subnote_Sequence_Num number is incremented for each portion of the note. The records will be reassembled in the sequence order, with the note fields being concatenated together and stored within the File_Note field of the Submittal_Log Table within the database.

Each File Note record must contain a record type code, the data source, transmittal date, a transmittal sequence identifier for the file, the subnote sequence identifier and the note(s) contents. If multiple file note records (subnotes) make up a single note then the subnote sequence number is incremented for each subnote record.

TABLE 4. Record Layout - File Note

File Notes						
FP#	**Field Name**	Data Type	Length	Decimals	Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: File Note = **3**
2	DATA_SOURCE_CODE	Char	2	0	No	Contractor or data source code for CCAQS. (e.g., "**NO**" = NOAA Boulder)
3	TRANSMIT_DATE	Char	8	0	No	Date data source sent file: YYYYMMDD
4	SEQUENCE_IDENTIFIER	Char	1	0	No	Alphanumeric character that uniquely identifies a file when multiple files are submitted on the same Transmit_Date (Start with 1-9 then use A-Z).
5	SUBNOTE_ SEQUENCE_NUM	Num	3	0	No	Sequence number for the subnote record comprising a larger complete note
6	FILE_NOTE	Char	200	0	No	A subnote (or complete note) 200 characters or less in length

Example 2. File Notes

If a file note, comprised of two subnotes, was submitted (along with the File Header example above) in the third file sent that day, the record would look like that below with two subnotes (highlighted):

1, "NO" , "F", "SFGAS", "H", "20000928" , "3", "S" , "1A" , 57347
3, "NO", "20000928", "3", 1, "This is a note about how this data was collected"

3, "NO", "20000928", "3", 2, "and since its a long one it has two records."
…

5.5 "File Footer" Record Layout

The File Footer record summarizes the content of the transmitted data file. The last record in the transmitted data file is the File Footer. It acts as an end-of-file (EOF) marker and recaps what has been received in the transmission. It contains the same information as the file header record and serves as a demarcation for the EOF . The File Footer will be used to ensure that the transmitted file was received in its entirety.

TABLE 5. Record Layout - File Footer

File Footer						
FP#	**Field Name**	Data Type	Length	Decimals	Allow Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: File Footer = **9**
2	DATA_SOURCE_CODE	Char	2	0	No	Contractor ID for CCAQS (e.g., "**NO**" = NOAA Boulder)
3	TRANSMIT_DATE	Char	8	0	No	Date source sent file as YYYYMMDD
4	SEQUENCE_IDENTIFIER	Char	1	0	No	Alphanumeric character that uniquely identifies a file when multiple files are submitted on the same Transmit_Date (Start with 1-9 then use A-Z).
5	OBS_RECORDS	Num	7	0	No	A count of the number of Obs Records in the transmittal file.

Example 3. File Footer

Building upon the file and record information provided in the above examples, the File Footer record is the last record in a transmitted data file and would look as follows (omit spaces):

…
…
9, "NO","20000928", "3", 57347

5.6 Observation Notes

Each data observation value or set of observations within the submitted file can have multiple 200-character "Obs Note" records comprising a single note. Since there can be many air observations grouped as a data set within a multiple of subnotes, a one to many relationship can exist within the ASCII file, that is, one complete note for many data records. A one-to-one relationship can also exist where a single observation has a note (comprised of one or more subnotes) special for that one observation. This format

specification allows for observations to be submitted in a file with or without accompanying observation notes. The use of observation notes is encouraged.

Observation notes will be located together at the top of the data file. The Observation Record Layout, as described below, allows for the use of up to three Obs Notes for providing information on for each data observation record (data point) within the data file. This is a means of associating observation records together as a data set with one note, and providing individual notes within the data set.

5.7 Observation Note Header Record Layout

The "Observation Note Header" record acts as a 'beginning-of-set' indicator. They also identify the number of note records that follow for a given observation. The Observation Note Header summarizes the number of note records being sent for the current observation.

TABLE 6. Record Layout - Observation Note Header

Obs Note Header						
FP#	**Field Name**	Data Type	Length	Decimals	Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: Obs Note Header = **5**
2	NOTE_NUMBER	Num	2	0	No	The note number assigned By the contractor/data provider
3	SUBNOTE_COUNT	Num	2	0	No	Count of notes or subnotes comprising this record

Example 4. Obs Note Header

If observation notes were submitted with the above example, the records containing the note header would look like this (highlighted):

1, "NO", "F", "SFGAS", "H", "20000928", "3", "S", "1A", 57347
3, "NO", "20000928", "3", 1, "This is a note about how this data was collected"
3, "NO", "20000928", "3", 2, "and since its a long one it has two records."
5, 1, 2
…

5.8 "Observation Note" Record Layout

The "Obs Note" records follow the Observation Header and increments the sequence identifier for each record. There should be as many records as indicated by the Subnote_Count the Observation Note Header record.

TABLE 7. Record Layout - Observation Note

Obs Note						
FP#	**Field Name**	Data Type	Length	Decimals	Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: Obs Notes = **6**
2	NOTE_NUMBER	Num	2	0	No	The note number assigned by the contractor/data provider
3	SUBNOTE_ SEQUENCE_NUM	Num	1	0	No	Sequence number for the subnotes comprising an Obs Note
4	OBS_NOTE	Char	200	0	No	Contractor observation notes

Again, if an air observation note is longer than 200 characters, then a second record (subnote or child note record) continuing the first note record needs to be created. Like File Notes, there can be multiple 200-character note records comprising a single note. These notes will be incorporated into the CCAQS Database System.

Example 5. Obs Note

If observation notes were submitted with the above example the records would look like this (highlighted):

1, “NO”, “F”, “SFGAS”, “H”, ”20000928”, “3”, ”S”, “1A”, 57347
3, “NO”, ”20000928”, “3”, 1, “This is a note about how this data was collected”
3, “NO”, ”20000928”, “3”, 2, “and since its a long one it has two records.”
5, 1, 2
6, 1, 1, ”This is a note for an observation record(s), and since it is more ”
6, 1, 2, ”than 200 bytes it must be broken up into two records.”
…

In the above example, the fifth and sixth records contain a note broken across two records. The fourth record is the Obs Note Header, which indicates that two notes should follow that makes up what is ‘Note 1’.

5.9 “Observation Note Footer”

The Note Footer also summarizes how many subnote records should have been received. The “Observation Note Footer” summarizes the information about each note record being sent. It acts as an end-of-note marker and recaps the number of 200 character subnotes that should have been received and associated with this note. It reiterates the information in the Observation Note Header record.

TABLE 8. Record Layout - Observation Notes Footer

Obs Notes Footer						
FP#	**Field Name**	Data Type	Length	Decimals	Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: Obs Note Footer = **7**
2	NOTE_NUMBER	Num	2	0	No	The note number given to a note by The contractor/data provider
3	SUBNOTE_COUNT	Num	2	0	No	Number of obs subnotes for this Record

Example 7. Obs Notes Footer

If observation notes were submitted with the above example, the records would include a note footer that looks like this (highlighted):

1, “NO”, “F”, “SFGAS”, “H”, ”20000928”, “3”, ”S”, “1A”, 57347
3, “NO”, ”20000928”, “3”, 1, “This is a note about how this data was collected”
3, “NO”, ”20000928”, “3”, 2, “and since its a long one it has two records.”
5, 1, 2
6, 1, 1, “This is a note for an observation record(s), and since it is more ”
6, 1, 2, “than 200 characters it must be broken up into two records.”
7, 1, 2
…

5.10 "Observation” Record Layout

The actual “Observation (Obs)” data record will comprise a single row in the data file, one for each observation data point or Obs_Value (e.g. concentration) or for each air sample (e.g., filter of canister media) where no observations are included. It serves multiple purposes. The following table identifies the content for each column of the observation record. (There are examples provided for each column/field “Description” below. They are not meant to provide an actual or expected observation record and should be consider independent examples).

TABLE 9. Record Layout - Observation (Obs)

Observation (Obs)						
FP #	**Field Name**	Data Type	Length	Decimals	Allow Nulls	Description
1	RECORD_TYPE	Num	1	0	No	Record type: **8** (for “Observation”)
2	AIR_SAMPLE_NUM	Char	15	0	Yes	The identifier used by the contractor/data provider to identify data samples uniquely (e.g. a filter or canister sample). Some types of direct data observations will not use this field. This field is for reporting and tracking all filter and canister type samples, whether they are analyzed or not. **Example: “AB2001”**

3	SUPPORT_ID	Num	4	0	No	Support_ID from the Supports Table. This is used to identify what "supported" the instruments used to take the samples or measurements. This Support_ID table includes specific monitoring stations ('sites'), towers, aircraft, mobile vans, etc. **Note**: For upper air observations (aircraft) this is an identification for the specific aircraft. For data from a tower this is an identification for a specific tower height and combined with location. It also identifies each unique GBS or ground based monitoring station (or site). **Example: 25**
4	SUPPORT_CODE	Char	8	0	No	Support_Code from the SupportsTable corresponding to the Support_ID. **Example: "ANG50"**
5	START_DATE	Date	8	0	No	Date the observation data or sample was started as YYYYMMDD. **Example: "20000928"**
6	END_DATE	Date	8	0	No	Date the observation or sample ended as YYYYMMDD. Useful for data averaged over multiple days. It accommodates averaging periods that overlap multiple days. **Example: "20000929"**
7	NOTE_A_NUMBER	Num	2	0	Yes	First note number associated with this observation obtained from the Obs Notes listed at the top of the submittal file. Three notes (A, B, & C) maximum per observation. **Example: 1**
8	NOTE_B_NUMBER	Num	2	0	Yes	Second note number associated with this Observation obtained from the Obs Notes listed at the top of the submittal file. Three notes (A, B, & C) maximum per observation. **Example: 9**
9	NOTE_C_NUMBER	Num	2	0	Yes	Third note number associated with this Observation obtained from the Obs Notes listed at the top of the submittal file. Three notes (A, B, & C) maximum per observation. **Example: 5**
10	TIME_ZONE_REF	Char	3	0	No	Time Zone reference for the START_TIME and END_TIME. The entry should be Pacific Standard Time (PST). PST is the default for CCAQS and all observations should be referenced to PST. **Example: "PST"**
11	START_TIME	Time	8	0	Yes	Observation start time for monitoring with a start and end time as HH:MM:SS Times are expressed in Pacific Standard Time (PST). **Note:** Both Start _Time and End_Time cannot be 'null'. **Example: "12:05:01"**
12	END_TIME	Time	8	0	Yes	Observation end time for monitoring that has a start and end time as HH:MM:SS (e.g., aircraft). Times are expressed in Pacific Standard Time (PST). **Note:** Both Start _Time and End_Time cannot be 'null' **Example: "12:10:05"**

13	PARAMETER_ID	Num	4	0	Yes	Parameter_ID from the Parameters table. Identification number associated with the Parameter measured. This number might refer to Mass Concentration FINE Size Fraction, for example). **Example: 415**
14	METHOD_ID	Num	4	0	Yes	Method_ID from the Methods table. The identification number associated with the Method_ID (lab or field) used to obtain the Obs_Value reported below. **Example: 1**
15	METHOD_CODE	Char	60	0	Yes	Code associated with the measurement Method_ID (lab or field) used to obtain the Obs_Value reported below. **Example: "CR_ELE_PU0000002500_XFA_XRF_TEF_D1_H24"**
16	INSTRUMENT_ TRACKING_ID	Num	4	0	Yes	Instrument_Tracking_ID from the Instruments_Tracking table. **Example: 35** (Through this number the instrument used to obtain the air sample or observation value can be identified down to the serial number and location and support.)
17	PRIMARY_FLAG	Char	3	0	Yes	Primary_Flag as V0, V1, V2, S, I or M as defined in the Study_Flags_PrimSec reference table. **Example: "S"** = Suspect; Reported value is suspect
18	SECONDARY_FLAG	Char	3	0	Yes	Secondary_Flag as defined in the Study_Flags_PrimSec reference table. **Example: "SCN"** = Suspected Contamination, Lab analysis or Field
19	ACTIVITY_FLAG	Char	3	0	Yes	The Activity_Flag is for exceptional events as defined in the Study_Flags_Activity reference table. **Example: "EXE"** = Forest Fire
20	OBS_VALUE	Num	8	4	**Yes***	The actual observation value. It can be meteorological or air quality data value. **Example: 8.11** *** Nulls allowed only for air sample file submittals that are for sample tracking purposes (e.g. filters and canisters) that are without associated Air_Obs (samples that have not been analyzed). Note: 'Air Samples Only' files use this same observation record format (Table 10).**
21	OBS_UNCERTAINTY	Num	8	4	Yes	Uncertainty value associated with the Obs_Value reported. **Example: 0.05**
22	SAMPLING_FREQ_CODE	Char	3	0	No	Raw Data Collection Frequency (pre-averaged or pre-rollup) by Hour, Min, Sec. This pertains to all data whether the result of a roll-up or not. Refer to the Sample_Frequency Table for the listing of possible entries. **Example: "Min"**
23	START_LATITUDE	Num	8	4	Yes	Start Latitude of Station. **Example: 38.9567**
24	START_LONGITUDE	Num	9	4	Yes	Start Longitude of Station. **Example: -121.0430**
25	START_ELEVATION	Num	4	0	Yes	Elevation at start of upper air observation in meters above mean sea level. **Example: 1300**
26	END_LATITUDE	Num	8	4	Yes	Last reported Station latitude during

						sampling. This applies to upper air/aircraft observation). **Example: 39.1700**
27	END_LONGITUDE	Num	9	4	Yes	Last reported Station longitude during sampling (for upper air observation). **Example: -121.0900**
28	END_ELEVATION	Num	4	0	Yes	Elevation at end of upper air observation in meters above mean sea level. **Example: 1600**
29	FLIGHT_PATTERN	Char	8	0	Yes	Flight pattern taken during sampling (e.g., spiral, dolphin). **Example: "SPIRAL_UP"**
30	FLIGHT_NUMBER	Char	10	0	Yes	Unique flight number associated with the collection of the Air_Sample_Num and/or Obs_Value. Data contractor (aircraft contractor) to provide this number. **Not null if aircraft data. Example: "A1005"**

A Record_Type of 8 identifies actual observation data records as defined in Table 10. The total number of observation records in a file must match the number declared within the File Header record as Obs_Records. Each row in this block will have a record type code of 8.

Note: <u>All unused fields must be accounted for by using blank "comma-comma" entries in the data record.</u>

<u>Example 8.</u> Observation (Record)

Observations submitted with the above example would look like this (highlighted). Note that the START_DATE and END_DATE fields may be different if a data value is a result of averaging over multiple dates.

1, "NO", "F", "SFGAS", "D", "20010628", "3", "S", "1A", 57347
3, "NO", "20010628", "3", 1, "This is a note about how this data was collected"
3, "NO", "20010628", "3", 2, "and since its a long one it has two records."
5, 1, 2
6, 1, 1, "This is a note for an observation record(s), and since it is more "
6, 1, 2, "than 200 characters it must be broken up into two records."
7, 1, 2
…
5, 5, 2
…
5, 9, 3
…
8, "AB2001", 25,"ANG50","20000928","20000929", 1,9,5,"PST","12:05:01", … ,,,,,,

<u>Example 8.</u> Full CCAQS Data Transmittal File Sample

A CCAQS data file transmittal would be submitted having a format like the following:

1, “NO”, “F”, “SFGAS”, “H”, ”20000928”, “3”, ”S”, “1A”, 57347
3, “NO”, ”20000928”, “3”, 1, “This is a note about how this data was collected”
3, “NO”, ”20000928”, “3”, 2, “and since its a long one it has two records.”
5, 1, 2
6, 1, 1, ”This is a note for an observation record, and since it is more ”
6, 1, 2, ”than 200 characters it must be broken up into two records.”
7, 1, 2
5, 2, 3
6, 2, 1, ”This is a note pertains to 1,100 observation records, and since it is more ”
6, 2, 2, ”than 200 characters it must be broken up into multiple records,”
6, 2, 3, ”in this case it took three records.”
7, 2, 3
5, 3, 3
6, 3, 1, ”This is a note pertains to 201 observation records, and since it is more ”
6, 3, 2, ”than 200 characters it must be broken up into multiple records,”
6, 3, 3, ”in this case it took three records.”
7, 3, 3
5, 4, 1
6, 4, 1, “I think this is enough in the way of notes”
7, 4, 1
…
8, “AB2001”, 25, “ANG50”,“20000928”,“20000928”, 1, 9, 5, “PST”,“12:05:01”, … ,,,,,
…
8, “00001”, 737, “BAL1”, “20001010”, “20001011”, 5, 6, 3, “PST”, … ,,,,,,,
8, “00005”, 737, “BAL1”, “20001010”, “20001011”, 5, 6, 3, “PST”, … ,,,,,,,
…
8, “A7347”, 122, “FRES100”, “19991215”, “19991215”, 1, 2, 3, “PST”, … ,,,,,,,,
…
8, “0002”, 454, “MV1”, “20001522”, “20001522”, 5, 2, 3, “PST”, … ,,,,,,,,
8, “0003”, 109, “AS1”, “20000104”, “20010104”, 6, 2, 3, “PST”, … ,,,,,,,,
…
8, “SA2001”, 204, “A1”, “20000909”, “20000909”, 3, 4, 3, “PST”, … ,,,,,,,,
8, “SA2002”, 204, “A1”, “20000909”, “20000909”, 3, 4, 3, “PST”, … ,,,,,,,,
…
8, “0004”, “2001”, “SAT1”, “20000624”, “20000624”, 7, 2, 3, “PST”,……,,,,,,,,
9, “NO”, ”20000928”, “3”, 57347

REFERENCES

1. Blumenthal, D. et al., *Field Program Plan for the San Joaquin Valley Air Quality Study (SJVAQS) and the Atmospheric Utility Signatures, Predictions, and Experiments (AUSPEX) Program.* Sonoma Technology, Inc., October 1993.

2. Fujita, E.M., et al., *1997 Southern California Ozone Study (SCOS97) Field Draft Plan.* DRI, June 24, 1996.

3. Mueller, P.K., "Comments on the Advances in the Analysis of Air Contaminants". JAPCA 30:988, 1980.

4. Mueller, P.K., Hidy, G.M., Baskett, R.L., Fung, K.K., Henry, R.C., Lavery, T.F., Nordi, N.J., Lloyd, A.C., Trasher, J.W., Warren, K.K., and Watson, J.G. *Sulfate Regional Experiment (SURE): Report of Findings*. EPRI Report EA-1901, Electric Power Research Institute, Palo Alto, CA, 1983.

5. Solomon, P., et al., *Integrated Monitoring Study (IMS-95),* 1997

6. Watson, J.G., Roth, P.M., Ziman, S.D., Neff, W.D., Magliano, K.L., Pederson, J.R., Solomon, P.A., and Thuiller, R.H. *Data Analysis Plan for the San Joaquin Valley Air Quality Study/Atmospheric Utilities Signatures, Predictions and Experiments (SJVAQS/AUSPEX) Program.* Report No. DRI Document 8932-001.1F1, Prepared for Pacific Gas & Electric Co., San Ramon, CA; San Joaquin Valleywide Air Pollution Study Agency, Fresno, CA; California Air Resources Board, Sacramento, CA, Desert Research Institute, Reno, NV, 1993.