South Coast Air Quality Management District
Science and Technology Advancement

Monitoring and Analysis Division
Atmospheric Measurements Branch

STANDARD OPERATING PROCEDURE

FOR

Data Management and Validation for Continuous Instruments

SOP 00083
Version 1.0
March 31, 2010

PREPARATION, REVIEWS AND APPROVALS

Data Management and Validation for Continuous Instruments

Prepared By: Mike Hamdan Date: 4/2/10
Mike M. Hamdan
Senior Air Quality Instrument Specialist

Reviewed By: Rene Bermudez Date: 4/2/10
Rene M. Bermudez
Principle Air Quality Instrument Specialist

Approved By: Date: 4-2-10
Philip M. Fine
Atmospheric Measurements Manager

Quality Assurance Approval:

Approved By: Jason C. Low Date: 04/02/10
Jason C. Low, Ph. D
Quality Assurance Manager

REVISION HISTORY

Data Management and Validation for Continuous Instruments

Version	Date
1.0	March 31, 2010

REVISION CHANGES FROM PREVIOUS VERSION

Data Management and Validation for Continuous Instruments

Section	Revisions
N/A	N/A

TABLE OF CONTENTS

Data Management and Validation for Continuous Instruments

1.0 GENERAL INFORMATION

1.1 Purpose:

The purpose of this Standard Operating Procedure (SOP) is to provide a set of written instructions that document the processes and procedures used to collect, transmit, store, validate, and report continuous instrument data from the AQMD ambient air network.

Note: Data management for manual sampling methods such as PM2.5 FRM, PM10 FRM, Pb, PM2.5 speciation and VOCs is documented in laboratory data management procedures.

1.2 General Description

The monitoring network consists of more than 30 air monitoring stations that house criteria pollutant analyzers (Ozone, CO, SO_2, and NO_x), meteorological sensors (wind, solar radiation, relative humidity, and temperature), particulate matter (PM) samplers and continuous PM monitors. These processes and procedures have been developed to maintain the data integrity and validity in accordance with State and Federal requirements as appropriate.

2.0 DATA MANAGEMENT PROCEDURES

2.1 Data Recording

Data recording devices consist of an Environmental System Corporation (ESC) data logger, which is considered the primary data acquisition component, as well as a Eurotherm Chessell electronic chart recorder, used at all sites to simultaneously record the analog output of the instruments. This secondary data-recording device is used to augment the data integrity and to verify suspect data points in the data base. The ESC data logger collects and computes the minute and hourly averages but only the hourly averages are stored by the ESC. Some of the gaseous pollutant analyzers and PM monitors have their own internal data loggers which can be accessed manually to download data into a laptop and transferred to a central storage location at AQMD.

2.2 Data Transmittal

Data transmittal is accomplished using a private internet protocol (IP) data network which links the primary (ESC) data logger through a router and terminal server at the air monitoring sites with the telemetry data acquisition server designated FORTRAN at AQMD headquarters.

The telemetry system polls each air monitoring station once every minute and the hourly averages are generated from the minute data. An hourly average is considered valid only if there are at least 45 valid 1-minute data points for the hour. At the end of each month, FORTRAN also generates a missing hourly average report. A manual polling of the missing hours is performed acquire available hourly averages from the

data logger to back-fill missing hourly averages in FORTRAN that are available on the logger but were not transmitted due to communication issues or other malfunctions. FORTRAN also has an on-demand polling capability where any member of the air monitoring staff can access any site to review real time readings, historical hourly averages, instrument's status and calibration data. Every hour, the hourly averages calculated by FORTRAN are stored locally and transmitted to the central relational data base on a separate database server.

All data edits are performed on the central database server. When editing is completed for each Month, the data for that period is locked form further editing, and an E-mail is sent to the Information Management (IM) division to update the telemetry system (FORTRAN) from the central database and generate the EPA Air Quality System (AQS) file.

In addition, data from the Eurotherm Chessell Video chart recorder at each monitoring station is downloaded monthly via an analog modem. The Chessell data is reviewed by the data validation staff for outliers and constant values, and used as needed to back-fill missing data in the central database. This back fill process is manually initiated by the data validation staff and is performed by taking the top-of-the-minute (snap shot) value from the chart for the missing data period, storing this value in a file that is uploaded into the central database

Data from internal data loggers is also used to back-fill missing data. Most of the continuous PM and gaseous instruments have internal data loggers. Data from the continuous PM-2.5 and PM-10 Met One Bam monitors is done monthly by field staff. The data is saved as CSV format file to network file servers at AQMD headquarters. This data is used for comparison and to back-fill any missing PM-2.5 and PM-10 data in the central database.

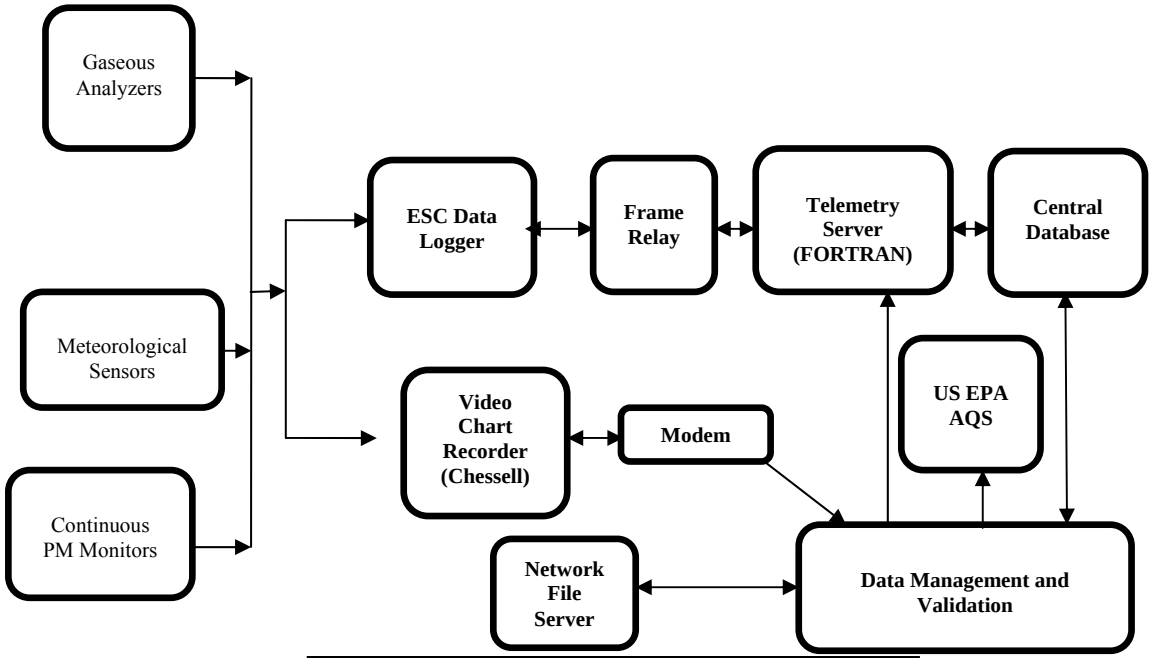


Figure 1.
Data Validation Process

2.3 Data Storage

All data is stored in a central relational database on site and on tapes once it gets too large, a backup copy is also stored at an offsite storage location. Data storage is managed in accordance to the QMP Section 8 (AQMD, 2009).

2.4 Data Validation

All of the analytical instruments measuring ambient concentrations of criteria pollutants undergo periodic maintenance, audits, and calibrations. These procedures are outlined in the "Criteria Gas Pollutant Monitoring Program", Quality Assurance Project Plan, or Divisional Website. Data collected during performance audits and calibrations are used to ascertain the accuracy, precision, and repeatability of each instrument to perform its required function. These data are also used by the data validation staff to determine if the ambient data is valid or not. All calibration or audit data is automatically flagged by the data acquisition system and not reported as ambient data. After a successful calibration/repair, the instrument is put back on line.

2.4.1 Continuous Gaseous Monitoring

Zero and span checks are performed six days/week. A precision check is performed once a week on Tuesday. The average reading from each step of the check is calculated and compared with the true value (expected reading) for each pollutant. If the difference is found to exceed the control limits in Table 1, a report is generated by FORTRAN that includes the site name, parameter and % of deviation. Every working day, a member of the data validation staff reviews this report and logs all the pollutants that exceed validation limits. If deviant data continues to be reported, an E-mail is sent to the Operation and Support lead staff to alert them so that the continuing deviation can be resolved for that instrument.

Table 1. Criteria Pollutant Action When % of Span Deviates from Actual Value

Pollutant	Warning Limit	Invalidating Limit
O3	**5%**	**7%**
CO	**7%**	**10%**
SO2	**7%**	**10%**
NOx	**7%**	**10%**

When reviewing ambient data, any data from an instrument that continues to deviate from daily span and precision values will be invalidated. When it is determined that the deviation is a result of a problem with the calibration

system, the ambient data is reviewed and may be considered valid in some cases. If the exceedance is a result of zero drift, then the ambient is adjusted for the zero shifts and considered valid. If the deviation is as a result of a zero drift the ambient data will be adjusted according to Figure Xx.2

A daily negative value report is generated by the telemetry system and reviewed by the data validation staff to determine if ambient data needs to be adjusted or invalidated based on the criteria in Table 2. Ambient data is adjusted only when the drift continues for several days and when an adjustment value can be applied without affecting the ambient data validity. All zero adjustments are documented in the data action log that indicates Date/Time for data impacted, adjustment value, and a comment. The same zero adjustment procedure is followed based on a review of the maintenance sheet and chart trace.

Table 2. Criteria Pollutants Zero Drift Adjustment/Invalidation Table

	No zero Adjustment	Zero Adjustment	Invalidate data
Pollutant			
O3	0 to +/-5 ppb	+/- 5 to +/- 15 PPB	> +/- 15 ppb
CO	0 to +/-5 ppb	+/- 5 to +/- 15 PPB	> +/- 15 ppb
SO2	0 to +/-5 ppb	+/- 5 to +/- 15 PPB	> +/- 15 ppb
NOx	0 to +/-5 ppb	+/- 5 to +/- 15 PPB	> +/- 15 ppb

Table 3. Typical maximum ambient readings

Variable	Typical Maximum reading
O3	220 PPB
CO	11.0 PPM
NO	605 PPB
NOX	750 ppb
NO2	165 PPB
SO2	165 PPB
PM-2.5	900 ug/m3
PM-10	900/m3

WS	82.5 mph
WD	361 deg.
T-Ambient	125. deg. C
BARP	33.0 in/hg
SR	1320 w/m2
UVR	110 w/m2
Shelter Temp.	30 deg. C
RH	101%

Any observation of an instrument malfunction and data transmission issue observed by the data validation staff is recorded in a log with the date/time impacted, data action and a comment, this log is used when validating the data. Maintenance sheets from the operation staff are delivered monthly and review by the data validation staff.
NO2 data is calculated by the telemetry system by subtracting NO from NOx, whenever either NO or NOx data is invalidated all NO, NO2, and NOx data is invalidated for that period.

2.4.2 Continuous PM Monitoring

The continuous PM-2.5 instrument in the field is the Met One BAM that outputs hourly readings and continues to display the same reading until the next hourly update. The telemetry system takes one of the minute readings and displays it as an hourly average. If the instrument flow is outside the control limits, the instrument channel is disabled and the data is considered invalid. In addition, the data from the instrument internal data logger is downloaded monthly and stored to be compared with the telemetry system data when there is an issue or to back-fill any missing data in the database.

The continuous PM-10 instruments in the field are either PM-10 Met One Bam or Thermo PM-10 TEOM. For the PM-10 TEOM, the minute data is polled by the telemetry system and an hourly average is calculated. Any time the flow is outside the control limits shown in Table 4, the data is invalidated. The PM-10 Met One BAM data is handled similarly to the PM-2.5 TEOM data. Any time the daily average value exceeds the NAAQS as a result of an exceptional event such as wild fire, high winds, or fire works, all of the hourly average data for the affected day are flagged with the appropriate “request for exclusion” flag in AQS, accompanied by a comment documenting the situation. Refer to the qualifier code list in Appendix Xx.x for the proper code.

Table 4. Continuous PM-2.5 and PM10 flow tolerances

Variable	Flow	Acceptance Criteria	Action
PM-2.5 Met One BAM	16.67 L/Min	16.67 L/Min	+/- 4% of standard
PM-10 Met One BAM	16.67 L/Min	16.67 L/Min	+/- 4% of standard

2.4.3 Meteorological Data

All meteorological data are screened by data validation staffs to identify missing data, unchanged (sticking) values, spikes, and values that go above or below the range of each instrument. If determined to be invalid, the data will be tagged invalidate and documented in the log with a note indicating the period of invalidation and the reason for invalidating the data.

Normally the meteorological data is validated and verified at the same time as pollution data, since there is interaction between ambient air pollution and meteorology.

Hourly averages of wind speed, wind direction, resultant wind speed, and resultant wind direction are calculated by the telemetry system using the minute wind speed and wind direction data collected by the data logger and polled by the telemetry system. The hourly averages are then uploaded into the central database. When either wind speed or wind direction data is erroneous, all four parameters are invalidated.

The monthly summaries of wind speed and vector wind speed are spot checked to verify that the vector wind speed is always less than or equal the average wind speed. If a discrepancy exists, the minute-average data and the hourly average calculation are reviewed to resolve the issue.

Wind direction and vector wind direction, which should generally be within 5-10 degrees of each other, is also spot checked to confirm adherence to this rule-of-thumb. Below is a check list that is used when reviewing the wind data.

Meteorological Data Screening Checklist

WSV refers to Vector Wind Speed (Resultant Wind Speed)
WDV refers to Vector Wind Direction (Resultant Wind Direction)

- Was percent data recovery criteria met (i.e., greater than 75%)
- Do WSV and WDV have same data recovery?
- Do WSV and WDV have same hour invalid?
- Review missing data summary
- Vector WS ≤ Average WS?

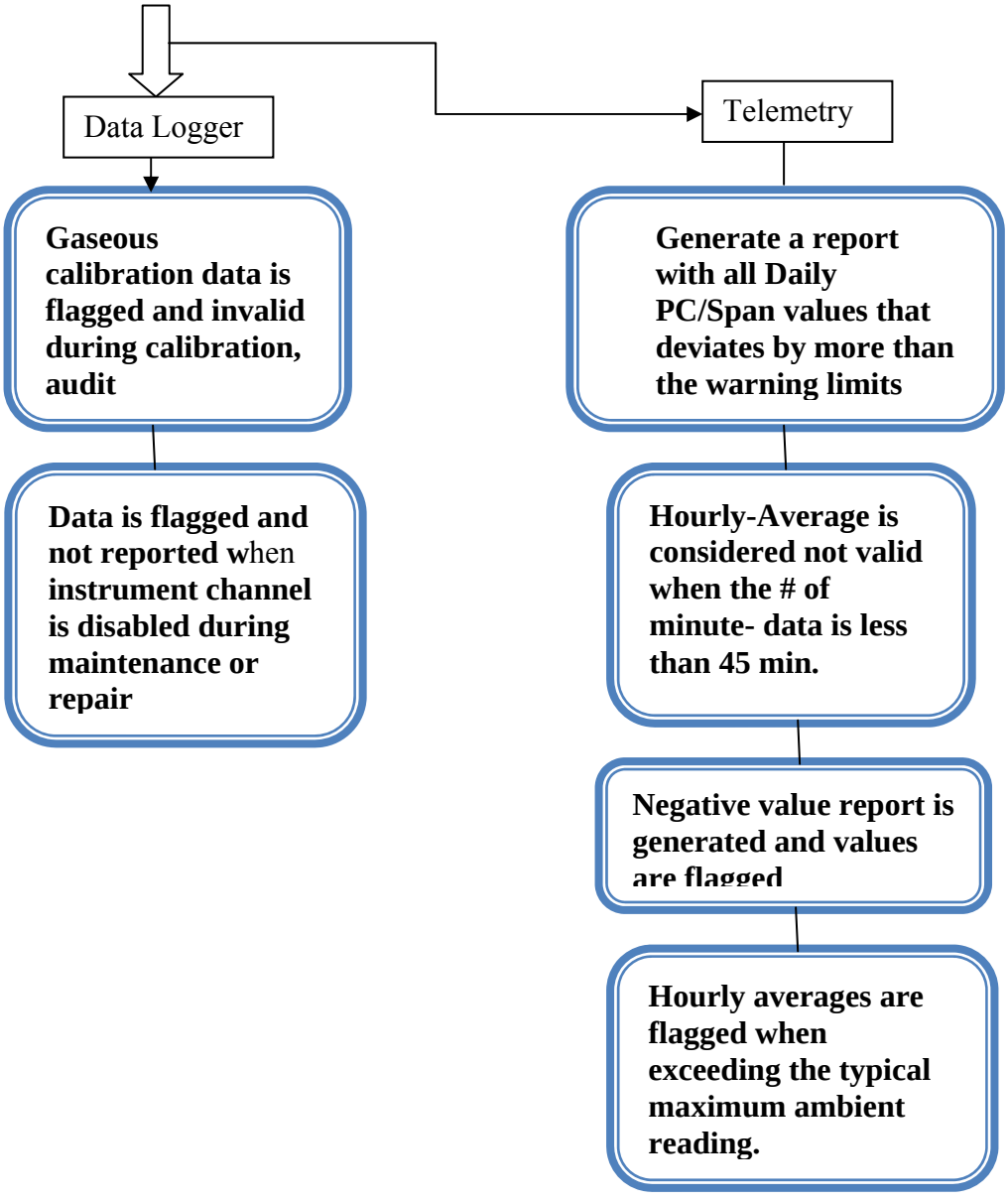
- Compare average WD to vector WD?
- Check dependent parameters:
- WSA affects WSV, WDV
- WDA affects WDV, WSV
- Random check of all meteorological parameters – (chart vs. data)
- Are all calibrations documented?

2.4.4 Data Validation Levels

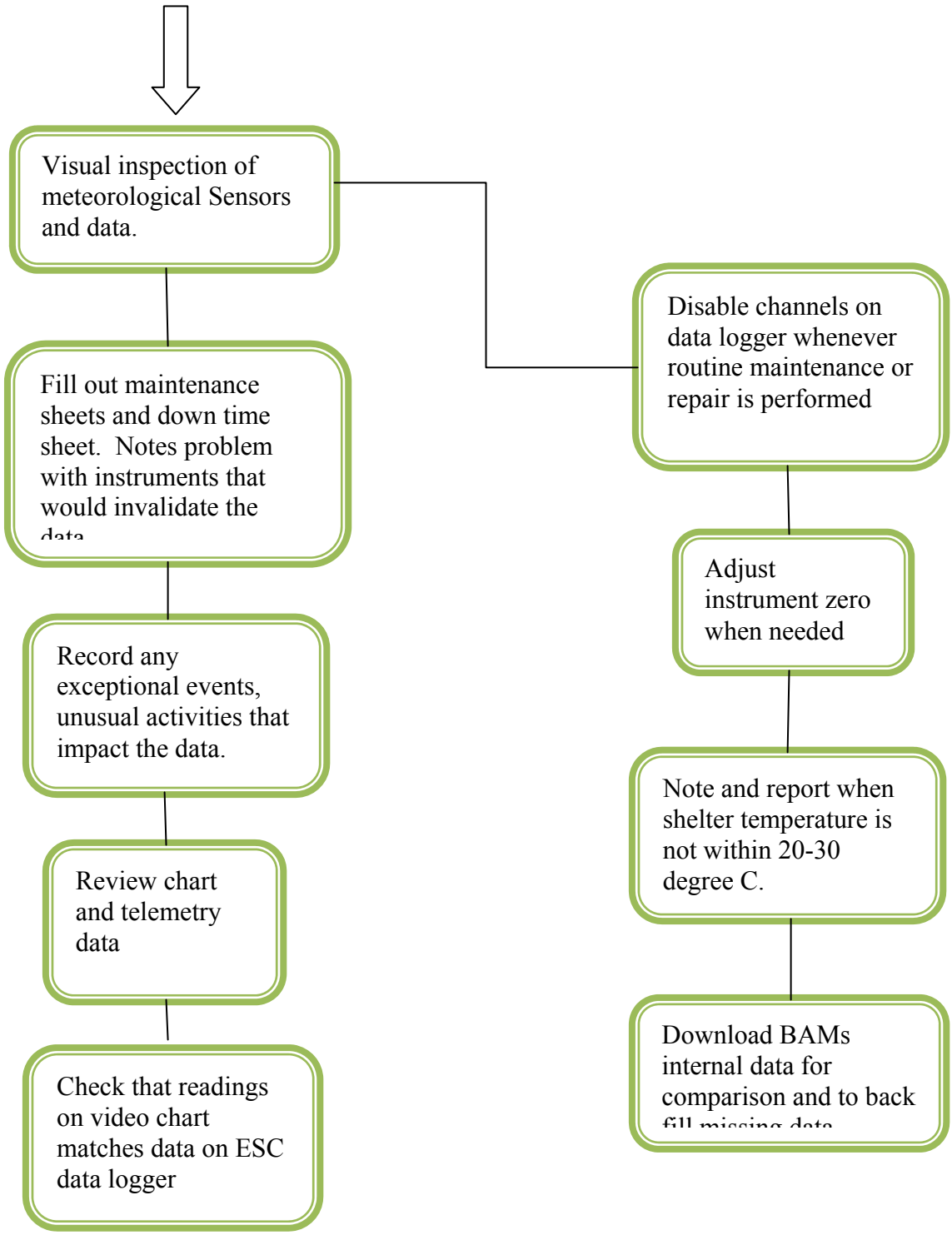
There are four data validation levels within the data screening process.

- **Level 0** is performed by the FORTRAN and the ESC data logger.
- **Level I** is performed by the station operation staff.
- **Level II** and **Level III** are performed by the data validation staff.

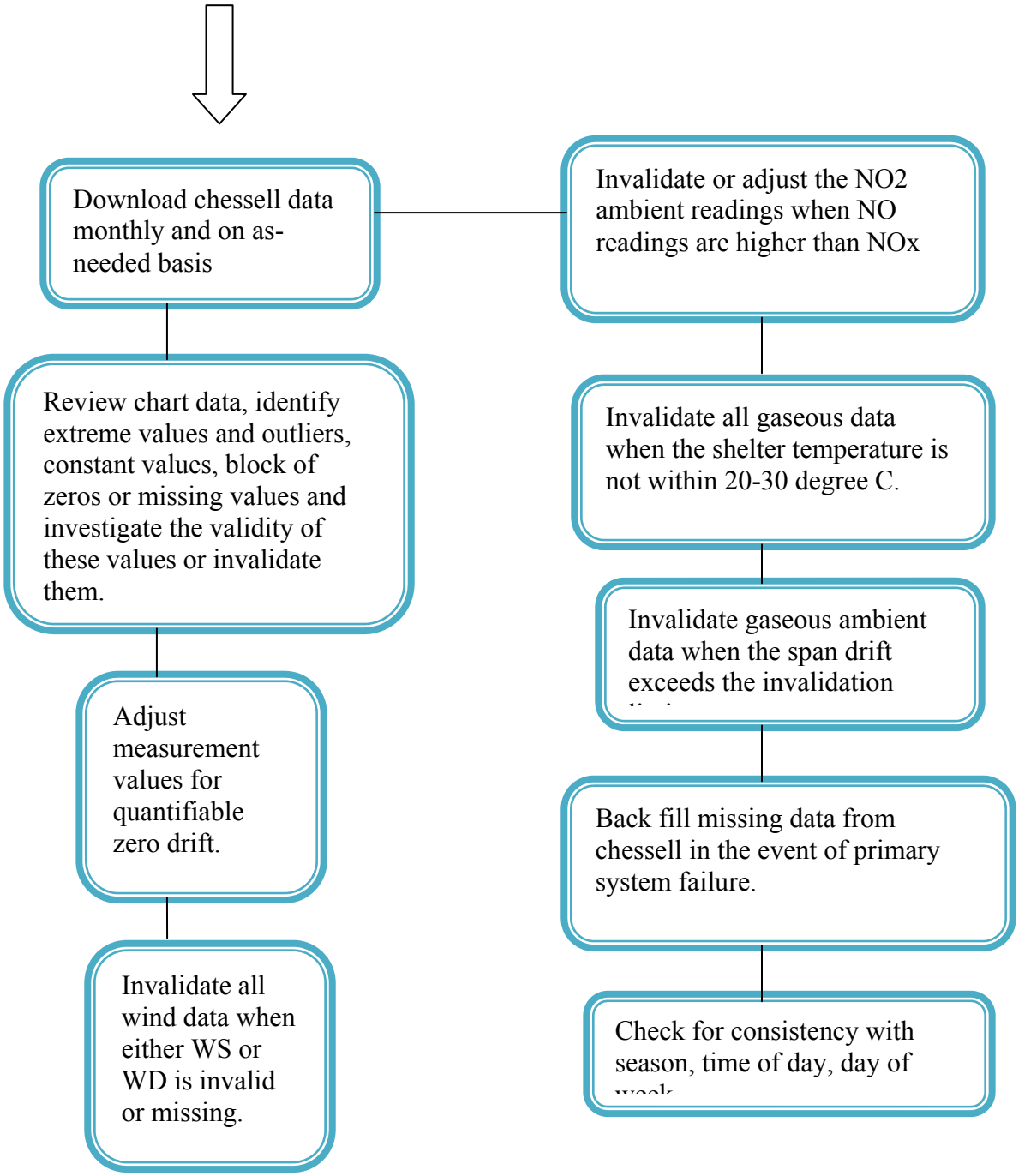
Level 0 Validation



Level I Validation for Continuous Instruments

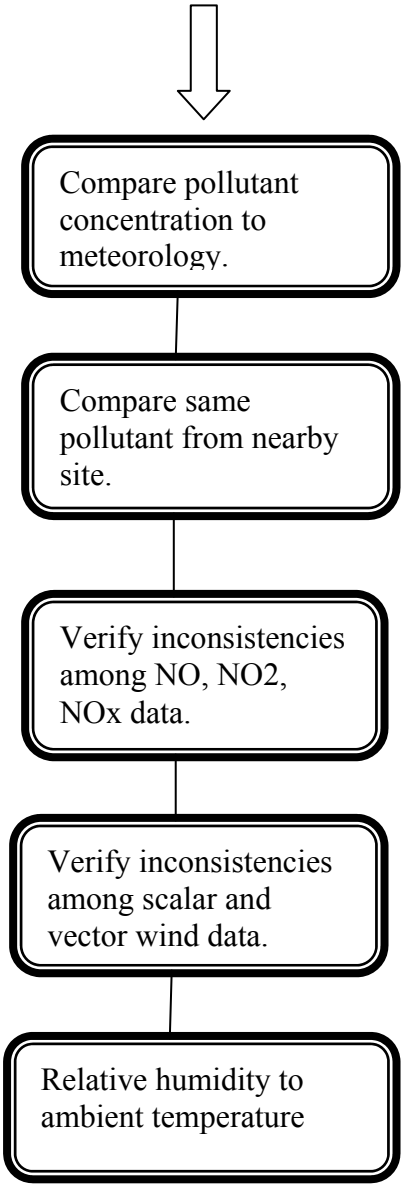


Level II Data Validation



Level III data validation

Data validation staff often review the tabular data by graphing it or by looking at the graphs from the Chessell (chart recorder), especially when the data is suspect since it is easier to see trends and relationship examples of these relationships are.



2.4.4 EPA AQS Data Submittal

Once all the edits on the data in the central database are completed for a whole month, an AQS formatted data file is generated by IM for submittal to EPA-AQS and a copy of the file is archived on the E:\drive for future reference. A Scan Report is run on AQS and reviewed by data validation staff to identify any values that exceed the historical maximum readings in AQS for investigation. If it is found to be invalid or needs to be edited, the changes are done in AQS and the data base then documented in the AQS changes log indicating the time period for the data in question, the action taken and the reason for edit.

Data validation staff review the AQS data summary sheet to input flags/invalidations that have been identified by Planning staff (exceptional events) or QA (results from corrective action or performance evaluation etc.).

At the end of every quarter, data validation staff submits to the EPA Air Quality System (AQS) bi-weekly precision data for all gaseous instruments and quarterly audit results that were generated by the QA field auditor. Every year before the data certification deadline is approached, the EPA-required reports are generated from AQS and reviewed by both the data validation staff and the QA manager, any flagging, missing, or invalidation issues are discussed and the proper action is taken before the certification letter is completed and the reports attached to it.

2.4.5 Enhancements and Future Developments

As part of AQMD's commitment to continually improve data quality and quality assurance systems, a new Data Management System (DMS) is currently being developed and tested. The DMS will include auto-QC capabilities, time-series graphs, chain of custody and many other features that will improve and streamline the data handling process while enhancing quality control and documentation.

APPENDIX A
Exceptional Events Qualifier Codes

Qt Qualifier Type	Qualifier Code	Qualifier Desc	Old Code	Status Ind
EX	K	Agricultural Tilling		I
EX	H	Chem. Spills & Industrial Accidents		I
EX	RC	Chem. Spills & Industrial Accidents		P
EX	R	Cleanup After a Major Disaster		I
EX	RD	Cleanup After a Major Disaster		P
EX	J	Construction/Demolition		I
EX	RE	Demolition		P
EX	RH	Fireworks		P
EX	L	Highway Construction		I
EX	O	Infrequent Large Gatherings		I
EX	RK	Infrequent Large Gatherings		P
EX	RL	Other		P
EX	Z	Other event		I
EX	Q	Prescribed Burning		I
EX	RM	Prescribed Fire		P
EX	M	Rerouting of Traffic		I
EX	P	Roofing Operations		I
EX	D	Sandblasting		I
EX	N	Sanding/Salting of Streets		I
EX	F	Structural Fire		I
EX	RP	Structural Fire		P
EX	RQ	Terrorist Act		P
EX	RR	Unique Traffic Disruption		P
EX	I	Unusual Traffic Congestion		I
INFORM	IA	African Dust		P
INFORM	IB	Asian Dust		P
INFORM	IC	Chem. Spills :INDUST Accidents		P
INFORM	ID	Cleanup After a Major Disaster		P
INFORM	IE	Demolition		P
INFORM	IF	Fire - Canadian		P
INFORM	IG	Fire - Mexico/Central America		P
INFORM	IH	Fireworks		P
INFORM	II	High Pollen Count		P
INFORM	IJ	High Winds		P
INFORM	IK	Infrequent Large Gatherings		P
INFORM	IL	Other		P
INFORM	IM	Prescribed Fire		P
INFORM	IN	Seismic Activity		P
INFORM	IO	Stratospheric Ozone Intrusion		P
INFORM	IP	Structural Fire		P
INFORM	IQ	Terrorist Act		P
INFORM	IR	Unique Traffic Disruption		P

INFORM	IS	Volcanic Eruptions	P
INFORM	IT	Wildfire-U. S.	P
INFORM	IU	Wildland Fire Use Fire-U. S.	P
NAT	RA	African Dust	P
NAT	RB	Asian Dust	P
NAT	RF	Fire - Canadian	P
NAT	RG	Fire - Mexico/Central America	P
NAT	E	Forest Fire	I
NAT	G	High Pollen Count	I
NAT	RI	High Pollen Count	P
NAT	A	High Winds	I
NAT	RJ	High Winds	P
NAT	U	Sahara Dust	I
NAT	S	Seismic Activity	I
NAT	RN	Seismic Activity	P
NAT	RO	Stratospheric Ozone Intrusion	P
NAT	B	Strattospheric Ozone Intrusion	I
NAT	C	Volcanic Eruptions	I
NAT	RS	Volcanic Eruptions	P
NAT	RT	Wildfire-U. S.	P
NAT	RU	Wildland Fire Use Fire-U. S.	P

APPENDIX B
Site Codes

4-Letter Name	Station ID	County	City AQS Code
MLVB	33165	Riverside	8005
AZUS	70060	Los Angeles	0002
MLHS	33163	Riverside	0004
CRES	36181	San Bernardino	9004
ANAH	30176	Orange	0007
BNAP	33164	Riverside	0012
GLEN	70591	Los Angeles	0016
WSLA	70091	Los Angeles	0113
BURK	70069	Los Angeles	1002
CSTA	30195	Orange	1003
RIVM	33146	Riverside	1003
UPLA	36175	San Bernardino	1004
CELA	70087	Los Angeles	1103
RESE	70074	Los Angeles	1201
LYNN	70084	Los Angeles	1301
CMPT	70112	Los Angeles	1302
PICO	70085	Los Angeles	1602
POMA	70075	Los Angeles	1701
INDI	33157	Riverside	2002
FONT	36197	San Bernardino	2002
PASA	70088	Los Angeles	2005
MSVJ	30812	Orange	2022
LGBH	70072	Los Angeles	4002
RDLD	36204	San Bernardino	4003
SLBH	70110	Los Angeles	4004
LAHB	30177	Orange	5001
PLSP	33137	Riverside	5001
LAXH	70820	Los Angeles	5005
PERI	33149	Riverside	6001
SCLR	70090	Los Angeles	6012
RIVR	33144	Riverside	8001
SNBO	36203	San Bernardino	9004
ELSI	33158	Riverside	9001
ANST	70031	Los Angeles	N/A
HDSN	70033	Los Angeles	N/A

APPENDIX C
County Codes

County	County Code
Los Angeles	**037**
Orange	**059**
Riverside	**065**
San Bernardino	**071**

APPENDIX D
Parameter Codes

Name	FORTRAN Code	Units	AQS Code
BARP	72	Inch-HG	64101
CO	15	ppm	42101
NO	26	pphm	42601
NO2	38	pphm	42602
NOx	28	pphm	42603
O3	36	pphm	44201
P10B	64	ug/m3	81102
P10C	97	ug/m3	81102
P10T	66	ug/m3	81102
P25B	65	ug/m3	88101
P25C	95	ug/m3	88502
RH	76	%	62201
RWD	N/A	DEG. COMPASS	61104
RWS	N/A	MPH	61103
SO2	24	pphm	42401
SR	78	W/M2	63301
T	74	DEG-F	62101
UVR	79	W/M2	63302
WD	86	DEG. COMPASS	61102
WS	84	MPH	61101