CENTRAL POLLUTION CONTROL BOARD

(MINISTRY OF ENVIRONMENT AND FORESTS, GOVT. OF INDIA)

PARIVESH BHAVAN, EAST ARJUN NAGAR, DELHI - 110032

Ph: +91 2230 1071, 4310 2300, 4310 2341

TOI, Delhi dt 23/1/2014

A-CPCB/Air Lab/ETS/2013-14

EXPRESSION OF INTEREST

Subject: Development of DAHS Software Programs for Pilot ETS Project

Proposals are invited from eligible firms for Development of DAHS (Data Acquisition and Handling) Software programs, viz. (1) ETS Bridge Software (2) DAHC Server Software which will be broadly responsible for collection of real-time PM emissions data from 1000+ industries, data collection at respective servers located at 3 SPCBs and CPCB including data viewing, data analysis, data validation, report generation, etc.

For further details refer to **Complete EOI** and **Requirement specifications** published at **www.cpcb.nic.in** or govt. tender website **www.tenders.gov.in** The Interested agencies may submit their EoI to the undersigned on or before (21 days from the date of publication) (3 PM), which will be opened at 3.30 PM on the same day.

davp 13113/11/0013/1314

Member Secretary

'T. Times of India', N.Delhi' dated 23/1/1

24.01.2014

Tender Document for the Development of DAHS Software Program for pilot ETS Project
A-CPCB/Air Lab/ETS/2013-14

CENTRAL POLLUTION CONTROL BOARD
(Ministry of Environment and Forests, Govt. of India)
Parivesh Bhawan, East Arjun Nagar, Delhi-110032
+91 22301071, 43102300, 43102341

Notice inviting Expression of Interest (EoI) for "DAHS (Data Acquisition and Handing System) Software" Development to support "Continuous Emission Monitoring and Pilot Emission Trading Scheme (ETS) for Particulate Matter from Stationary Sources (Stack) in India (pilot in three states Gujarat, Maharashtra and Tamil Nadu)"

Background

The pilot emissions trading scheme (ETS) for particulate matter (PM) is collaboration between MoEF, CPCB, GPCB, MPCB, and TNPCB to design, implement, and evaluate an innovative approach to pollution control in India. Continuous Emissions Monitoring Systems (CEMS) form the foundation for a successful ETS program, such that 1000 participating industries (300+ in each state) would be transferring real-time data to the SPCB and CPCB servers.

As per the protocol, industries will install CEMS device and setup DAS (Data Acquisition System – a PC with internet). DAS will have two software programs installed, CEMS Software from vendors to convert the data from CEMS in a prescribed format and ETS Bridge Software to read the file in real-time and send the same to the servers. Servers are installed at each State Pollution Control board, in a room—hereafter called the DAHC(Data Acquisition and Handling Centre). The software installed at DAHC Server, hereafter, DAHC Server Software will accept data from ETS Bridge Software installed in 1000 DAS's, store the data, validate the data and will have features likereal-time data viewing,report generation and admin features for regulator and industry.

To perform the abovementioned applications, the DAHC Server Softwareand ETS Bridge Software are necessary for CPCB and each SPCB to oversee the ETS pilot project and ensure it is being supported by the highest-quality data. Expression of Interest is invited from eligible software vendors for participation in bidding for the development of the DAHS Software—comprising of ETS Bridge Software at industry-side and DAHC Server Software at CPCB and each SPCB.

1. Scope of Work

A DAHC will be set up in Gujarat, Maharashtra, Tamil Nadu, and Delhi, so the geographical scope of work is limited to these regions only. Broadly, the DAHS Software can be divided into two components:-

1) ETS Bridge Software (EBS)
2) DAHC Server Software (DSS)

The eligible software vendor will be responsible for the supply and installation of EBS (responsible for the data transfer to the server) and DSS (capable of data collection in real-time from EBS, storage,

data viewing, data validation, analysis, reporting, login feature, user interfaces etc.). In particular, the scope of work encompasses the following:

1.1. Categories of Work

Table 1 : Scope of Work	
1.	Development of Design Documents for the entire project activities
2.	Supply of the proposed EBS carrying out all the requirements as specified in the SRS document
3.	Supply and installation of the proposed DSS and its various component carrying out all the requirements as specified in the SRS document
4.	Up-gradation and maintenance of the software for a period of five years after commissioning. Support to all SPCBs which will adopt the SW, later on, without additional cost.
5.	Manpower support of one technically competent SW development experts for five years after commissioning for minor change requirements. The developer shall be posted in CPCB and shall report to CPCB IT Division. Additional support if required to sort out the problems, shall be provided by the firm/company.
6.	In-person training of comprehensive software operation for 20 DAHC staff after the go-live phase.
7.	Entire source code with proper comments, data flow diagrams needs to be provided
8.	Module wise systematically written user manuals and training material of the software.
9.	The scope of work for software developer will include incorporating functionalities to enable a linkage to trading platform to use the emission data from DAHS for trading purposes etc..

1.2. Requirements for ETS Bridge Software

1.	**Real-time data collection (reading)**
1.1.	Collection of the data i.e. reading the file in prescribed format (created by CEMS Vendor Software) as per ETS Bridge Software interface in real-time.
1.2.	Tracking the received data from CEMS Vendor Software against time
2.	**Data Storage**
2.1	Storage (for 2 years) of the read data in the encrypted form in the hard-drive of the DAS
3.	**Real-time data transfer to DAHC Server**
3.1.	Real-time data transfer to DAHC Servers using HTTP protocol
3.2.	Tracking the data sent and received by the DAHC Servers at CPCB and SPCB.
4.	**Non Functional Requirements of EBS**
4.1.	ETS Bridge should be a light weight standalone application that is very easy to install and upgrade.
4.2	ETS Bridge should run on Windows XP, Windows 7 (32 bit and 64 bit) and Windows 8 (32 bit and 64 bit) platforms and on other platforms shall come in future till project duration.

1.3. Functional requirements for DAHC Server Software

1	**Data collection and acknowledgement**
1.1	Real-time collection of data from ETS Bridge Software installed at DASs.
1.2	Acknowledging the data received (every minute) to the ETS Bridge Software.
1.3	Collection of calibration data and other administrative data from Industry Web Portal.

	Collection of administrative data from Regulator Web Portal and System Integration Portal.
2.	**Data Storage**
2.1	Storage of raw data received from DAS on the Server Database.
2.2	Storage of the validated data, data post MDRP (Missing Data Replacement Protocol) and WDRP (Wrong Data Replacement Protocol).
2.3	Storage of administrative information like user ID's and passwords.
2.4	Storage of data obtained from Industry Web Portal like Industry details, CEMS details, calibration factors, CEMS operations activities details etc.
2.5	Storage of data obtained from Regulator I (SPCB) Web Portal like administrative actions, acceptance or rejection of industry comments and others.
3.	**Calculations, data analysis and carrying out various activities of CEMS Operations**
3.1	Calculations of measurement data parameters in real-time (such as calculations of calibrated data based on uncalibrated data)
3.2	Data analysis of PM emissions (average PM conc., total load) as per various parameters like sector, cluster, size, and duration (daily, weekly, monthly, etc.)
3.3	Fixed and customizable queries
3.4	Report generation (as per formats provided by CPCB)
3.5	Other requirements as needed for various activities of CEMS Operations.
4.	**Data validation and replacement as per the (pre-defined but customizable) protocol**
4.1	Validation of the data according to the data validation protocol (DVP)
4.2	Replacement of data as per Missing Data Replacement Protocol (MDRP)
4.3	Replacement of data as per Wrong Data Replacement Protocol (WDRP)
5.	**Hosting user-friendly web portals for data collection, viewing, reporting, and processing run-time queries.**
5.1	Web Portals to view the real-time raw data, validated data, alarms, CEMS operations data for regulators and industries.
5.2	Web site for public with ETS/CEMS relevantstatistical reports, articles and news
5.3	System Integration Test Portal for CEMS Vendor's Software.
6.	**Administrative features**
6.1	An Administration Utility to add / update various industries, devices, users etc. and generating and storing User ID's for regulator and industries (with different features)

7.	**Non Functional requirements for DAHC Server Software**
7.1	DAHC Server should have a scalable and robust web-based architecture
7.2	DAHC Server Software should be made on Windows Server 2012 platform.
7.3	The web portals and other UIs should be intuitive to use and easy to maintain. The web portals should be able to handle simultaneous request from multiple clients
7.4	Adequate security measures should be built into the system as per NIC guidelines
7.5	Data sync feature to combine the data validation efforts and display through single portal on real time basis.

2. Reference Document

The Software requirement specification (SRS) document is attached below as Annexure-I.

3. Submission of EOI

Interested software vendors fulfilling the below mentioned eligibility criteria (Section 4.1) may submit an EoI in a sealed envelope by speed post/or by Hand clearly super-scribing**Expression of Interest (EoI) for DAHS Software Development for Pilot Emissions Trading Scheme (ETS)for Particulate Matter(PM)** along with the information specified in Proposal (Section 4.2) to the Member Secretary, Central Pollution Control Board, Parivesh Bhawan, East Arjun Nagar,Delhi- 110032 by **12.02.2014 up to 3:00 PM**.

4. Vendor Eligibility Criteria (VEC)

4.1. TECHNICAL

4.1.1. The Vendor shall be a firm/consortium having IT certifications such as SEI-CMM Level-III or SEI CMM Level-V or ISO for the software development.

4.1.2. The Vendor shall have successfully designed, tested and implemented at least one project of Development of Software from scratch for data communication on real time basis in last **three years** reckoned from 12.02.2014.

4.1.3. The Vendor shall have experience of having successfully carried out and completed similar project during last **three years** reckoned from 12.02.2014, as per either of the following:-

I. **One (1) similar (*) completed project costing not less than the following amount (In Rs.) in lakhs**	**OR**	**II.** **Two (2) similar (*) completed projects each costing not less than the following amount (In Rs.) in lakhs**	**OR**	**III.** **Three (3) similar (*) completed projects each costing not less than the following amount (In Rs.) in lakhs**
32 Lakhs		**20 Lakhs**		**16 Lakhs**

Definition of Similar completed Software development Projects:

Have developed operational software implementing communication protocols like http, ftp etc. and thereby making real time data available from different locations and its web display using maps with mathematical analytics

Or

Having developed software for data communication without human intervention with web display using maps with mathematical analytics

Or

Have developed ERP kind of software through which maintained activities of another institution or firm with web portal display & mathematical analytics

Or

Have integrated GPRS/GSM tracking systems to manage data availability at a specific location and displayed it through web display using maps with mathematical analytics

Or

Have developed software systems using SCADA by integrating measuring parameters desired and data transmitted to a central location for web display using maps with mathematical analytics.

The system should have been developed using standard practices of software engineering: Feasibility study, Software Requirement Study, System analysis, System Design, Prototype product development, Testing, Implementation, Maintenance, and Review.

4.1.4 The Vendor shall have software development group of at-least **ten software professionals**and a network group to support networking issues in the project.

4.1.5. The vendor should not have been black listed by any of the Government agency as on date.

4.2 FINANCIAL

4.2.1. Average Annual financial turnover: Average Annual financial turnover during the last three years, ending on 31st March 2013, should be at least 12(twelve) Lakhs. The vendor should have strong financial base and profit making.

4.3 INFORMATION/ DOCUMENTS REQUIRED

4.3.1. Vendor shall meet the Vendor Eligibility Criteria (VEC) as above and shall produce at the time of submission of EOI, the following DULY SIGNED DOCUMENTS BY AN AUTHORIZED PERSON OF THE FIRM WITH documentary evidence.

a) Technical: Documents pertaining to VEC clause 4.1.1, 4.1.2, 4.1.3, 4.1.4, & 4.1.5:

1. Copies of Work Order by Owner along with value.
2. Copy of corresponding software completion Certificate/Commissioning Certificate/Performance Certificate issued by Owner as point 4.3.1 a).1 above.
3. CVs of software development group.
4. For criteria no. 4.1.5 above self-declaration on vendor's letterhead should be provided.

b) Financial: Audited Annual reports including Balance sheet & Profit & Loss accounts statement for the preceding three financial years as required under clause no. 4.2.1.

c) Other documents required:

1. A company having its development centre in India and operating for at least past three years. Documentary evidence to be attached.
2. Should have executed similar projects involving real time data collection and display. Give information of at least one such project along with client testimonials in your proposal. Documentary evidence to be attached.
3. Additionally a product/ solution in Air Quality domain will be desirable, but is not a must. If yes, provide certificate.
4. Company should have adequate quality processes in place preferably certified by a quality / standards body. If Yes, provide certificate.
5. Company should not have any direct / legal relationship with manufacturer / distributor of any PM CEMS device. This is to ensure an independent, 3rd party software not favouring a particular CEMS hardware. Self-declaration on vendor's letterhead should be provided.

5.0 Timeline

Planned schedule of activities involved in the selection process is as under:

Activity ID	Activity	Date
1	EOI published with link to full specifications	24.01.2014
2	Vendor's submission of proposal, with documents as mentioned in section 4.3.	12.02.2014 up to 3:00 pm
3	Completion of technical evaluation	Will be notified
4	Discussion with short listed Vendors	Will be notified

6.0 Terms and conditions

To ensure that CPCB is provided with complete and accurate information in response to the attached EOI, we request that each software vendor responding to the EOI agrees to read the software specifications in detail, quote the time required, and adhere to the following terms and conditions.

1. **Adherence to timeline** – The vendor agrees to adhere to the timeline for the software development and delivery which is 4 months from the issue of work-order.
2. Subletting of the work is not allowed without prior written permission from CPCB.
3. **Costs and Expenses** – Any and all costs (and expenses) incurred by the vendor in responding to the EOI, engaging in any other activity required in connection to this EOI are to be solely borne by the vendor; these will not be reimbursed by CPCB.
4. **Right to modify the list of functionalities and scenarios** – The list of scenarios / functionalities / requirements detailed in Section 1 is not exhaustive. CPCB, at its sole discretion, may modify or delete any of the existing scenarios, or provide additional scenarios. Any such modification / addition shall be duly communicated to the vendors, and they shall be provided a time of one week for presenting demo on the modified / new scenarios.
5. **Right to interview** – As a part of the evaluation process, CPCB may interview the vendors who participated in the EOI. The interviews may be conducted over telephone, video conference, or face to face. CPCB reserves the right to reject any vendor.
6. **Right to Share EOI Responses with designated personnel** – The response to the EOI submitted by the authorised representative of the vendor, along with all the supporting documents / materials shall become the property of CPCB and shall not be returned to the vendor. CPCB does not undertake to hold the content of the responses to this EOI and any subsequent information or contractual documents related thereto ("Vendor Information") in confidence. Further, CPCB reserves the right to disclose any and all Vendor Information on a need to know basis to its employees, agents and subcontractors.
7. **Right of Refusal -** The vendor understands and agrees that CPCB reserves in its absolute discretion the right to select or reject any vendor any time during or after the EOI process or any subsequent evaluation or contractual process. The vendor further understands and agrees that any such selection or rejection may be based on the vendor's responses to this EOI, on any subsequent information or contractual documents related thereto, or for any other reason whatsoever.
8. **Ownership of documents and copy rights** - Complete DAHS Software developed under this project, documentation and other work products will be fully owned by CPCB. CPCB will get unlimited rights to modify, enhance, install and otherwise use the software as they deem fit.
9. **Transfer of ownership -** The vendor shall grant the purchaser a perpetual license to use the software without any additional payment or obligations to enter into a contact for maintenance or support. There is no limit on the number of users for ETS Bridge Software and for the DAHC

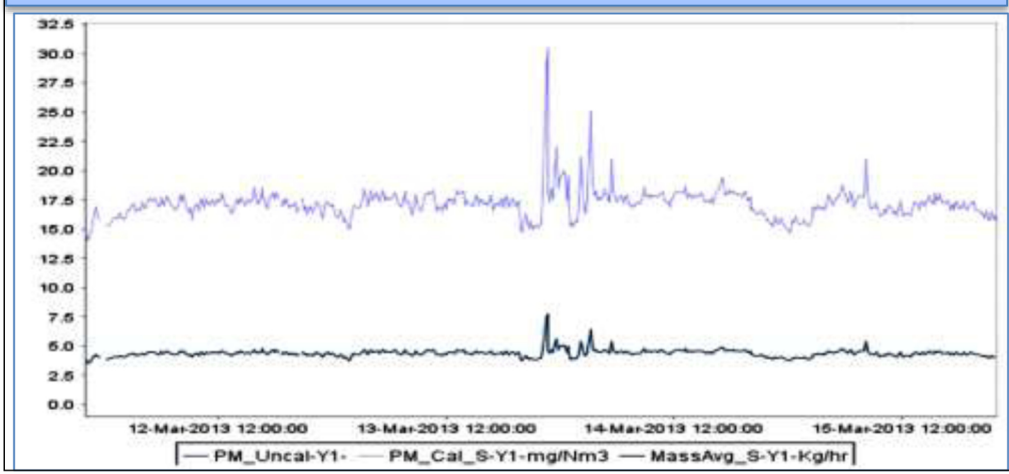
Server Software. Purchaser can freely give the ETS Bridge Software to industries for data transfer. There is no limit on the number of user ids for industry, regulator etc. Concurrent number of servers may be used if required. All study documents, data and specification prepared by the Bidder shall be the property of CPCB.

10. **Compatible with NIC and NICSI –** Vendor should not be using any 3rd party tool which is not allowed by NIC or NICSI while hosting in their domain. It's the responsibility of the vendor to provide documentation and engage in the coordination activities for setting a part of the software or full at the NIC server.
11. In case of disintegration or dissolution of firm due to any reason what so ever it is, the individual members will be accountable for deliverable to CPCB. If firm is purchased by another firm or agency, that new owner or agency shall be responsible for deliverables to CPCB.
12. If there are any queries, clarifications required on above document may please be forwarded to Dr. D. Saha, Scientist 'D', dsaha.cpcb@nic.in, and Sh. Aditya Sharma, Scientist 'C' aditya.cpcb@nic.in on or before 10.02.2014.

Continuous Emissions Monitoring and Pilot Emissions Trading Scheme (ETS) for Particulate Matter from Stationary Sources in India.

Software Requirement Specifications for ETS Bridge Software and DAHC Server Software

Date: January 23, 2014



Ministry of Environment and Forests
GOVERNMENT OF INDIA

Gujarat Pollution Control Board

Maharashtra Pollution Control Board

TNPCB
Tamil Nadu Pollution Control Board

IFMR Research
ABDUL LATIF JAMEEL POVERTY ACTION LAB

Table of Contents

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviations	**Explanation**
CEMS	Continuous Emissions Monitoring System
CPCB	Central Pollution Control Board
CVS	CEMS Vendor Software
DAHC	Data Acquisition and Handling Centre (at CPCB and SPCB)
DAHS	Data Acquisition and Handling System (Entire IT system to implement and support CEMS regime)
DAS	Data Acquisition System
DSS	DAHC Server Software
DVP	Data Validation Protocol
EBS	ETS Bridge Software
ETS	Emissions Trading Scheme
GPCB	Gujarat Pollution Control Board
MPCB	Maharashtra Pollution Control Board
MDRP	Missing Data Replacement Protocol
SW	Software
Time format	DD/MM/YYYY
ToC	Type of Configuration
TNPCB	Tamil Nadu Pollution Control Board
WDRP	Wrong Data Replacement Protocol

GLOSSARY OF TERMS

Calibrate:	To check, adjust, or determine by comparison with a standard (the graduations of a quantitative measuring instrument). In the context of this document, it is the activity of calibrating the PM CEMS device.
CEMS Audit:	Audit performed at least once a year to make sure CEMS device are performing as per the standard. Under the activity of CEMS audit, CEMS-Audit Performance Test is performed.
CEMS Vendor Software:	CEMS Vendor Software is the software installed by the vendor at the DAS that takes the data form the CEMS device and saves in a prescribed format, so that ETS Bridge Software can read it. This software needs to pass the System Integration Test to test its compatibility with the ETS Bridge Software.
Cluster:	A cluster is a collection of industries in the same geographic region designated by the State Pollution Control Board to participate in the pilot Emission Trading Scheme.
DAHC:	Data Acquisition and Handling Centre is the Centre located at CPCB and SPCB's which collects data from the industry.
DAHC Server:	Data Acquisition and Handling Center Server is the server set up at the regulator (CPCB & SPCBs) DAHC.
DAHC Server Software:	Data Acquisition and Handling Center (DAHC) Server Software is the software installed at the DAHC Server for data collection, data storage, data visualization, administration, and other features.
DAS:	Data Acquisition System is a PC at the participating industry which has two software programs: (i) CEMS Vendor Software and (ii) ETS Bridge Software installed on it. It takes the data from the CEMS device and sends it to the DAHC Server
ETS:	Emissions Trading Scheme (ETS) is a regulatory tool used to reduce pollution emissions. In such a scheme, the regulator sets the overall amount of emissions but does not decide what amount a particular source will emit. Industrial plants and other polluters, rather than being told a fixed emissions limit or concentration standard, face a price for their emissions and choose how much to emit, within reasonable limits, taking this price into account. The cost placed on the emissions makes it costly to operate with high pollution and thus gives polluters an incentive to cut back the pollution levels.

ETS Bridge Software:	Free Software or Freeware to be provided by the CPCB and SPCBs for the installation at DAS (industry site). It will act as a bridge and will transfer the data form CEMS Vendor Software to the DAHC Server.
Functional Test:	Tests used to ensure that CEMS device has been installed and is operating correctly.
Industry Web Portal:	Web Portal with login feature for every participating industry via which industries can view their emissions data and other related data, submit their administrative forms, comment on their data, etc.
Moving Average:	A moving average is a type of finite impulse response filter used to analyze a set of data points by creating a series of averages of different subsets of the full data set. A moving average is commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends or cycles.
Parameter Specification:	List of parameters measured by the CEMS device and calculated at the DAS for data transmission to the DAHC Server.
Performance Requirements:	Once CEMS is installed and calibrated, set of requirements (must) to be fulfilled by the participating industry with respect to the performance of the CEMS device. It consists of a Post-Calibration Performance Test and CEMS Audit Performance Test.
Re-calibration:	Calibrating the PM CEMS device again to override the previous calibration equation. This can be triggered because of various conditions.
Regulator Web Portal:	Web Portal with login feature for the regulating boards via which they can view emissions data, flagged, related reports and other administrative data. The web portal will have data analysis and data visualization features, etc.
System Integration Portal:	Web Portal with login feature for the vendors which will have the feature of System Integration test, through which vendors can make their CVS compatible with ETS interface and certify it.
System Integration Test:	A test performed on the CEMS Vendor Software to test if it is compatible with the ETS Bridge Software.

PROJECT TITLE: Development of software programs to establish Data Acquisition and Handling System (DAHS) to support Continuous Emission Monitoring System (CEMS) and Pilot Emissions Trading Scheme (ETS) for particulate matter from stationary sources in India.

PROJECT DURATION: Twelve months after which SW is commissioned.
This includes Development (~ Six months), Deployment & Go Live (~ one month), and User testing (~five months).

1.0. Introduction

The pilot emissions trading scheme (ETS) for particulate matter (PM) is collaboration between MoEF, CPCB, GPCB, MPCB, and TNPCB to design, implement, and evaluate this innovative approach to pollution control in India.

The pilot emissions trading scheme will rely on continuous emissions monitoring systems (CEMS) for measurement. The term CEMS refers to the hardware (measuring instrument, computer, etc.), software and related peripherals required to measure emissions from a stationary source on a continuous basis, such that the total load of particulate matter (PM) emitted from each stationary source can be calculated. Total PM emitted from the industry can then be reconciled against permit holdings in the trading scheme.

At present, the CEMS and ETS program will be piloted in 1000 industries (300+ each in state of Gujarat, Maharashtra, and Tamil Nadu). However, post pilot, this will be scaled up to more industries and states. The hardware, post pilot will be the responsibility of CPCB however the software development should be scalable to support any number of industry users, SPCB users, any no. of states getting involved later in the project.

1.1. Purpose

Continuous Emissions Monitoring Systems (CEMS) forms the foundation for a successful ETS program. For the pilot project, each of the 1000 participating industries would be transferring data on a real-time basis to the SPCB and CPCB servers as per Figure 3: DAHS Architecture.

A monitoring and regulatory infrastructure, further referred to as CEMS regime is established to obtain accurate and reliable measurement of particulate matter through the CEMS device, within the context of this pilot trading scheme.

Data Acquisition and Handling System (DAHS) comprising of software and hardware components is designed to support the CEMS regime. The DAHS is responsible for collecting data, converting it in a predefined fixed format, storing it, validating it, preparing reports for internal and external use, etc.

This document details structure and requirements of two major components of DAHS (which need to be designed by a vendor):
1) ETS Bridge Software (hereinafter referred as **EBS**), &
2) Data Acquisition and Handling Centre (DAHC) Server Software (hereinafter referred as **DSS**)

As per the protocol as set forth in "CEMS Specifications", industries will install CEMS device and setup DAS (Data Acquisition System – a PC with internet) at their site. The DSS will accept data from EBS installed in 1000 DASs at respective industry sites, will store the data, will validate the data and will have various other features as mentioned in Section 4.0 for different users as mentioned in Section 4.1.

1.2. Scope

Data Acquisition and Handling Center (DAHC) will be set up in Gujarat, Maharashtra, Tamil Nadu, and Delhi at the Pollution Control Board offices and is likely to expand in other states. All the required hardware for software operation, data storage, validation, analysis of data and hosting the web portals will be setup in these four DAHCs and shall be the responsibility of CPCB.

The entire software development work is the responsibility of the selected firm and can be divided into two main components:-

1) Development and up-gradation of ETS Bridge Software (EBS)
2) Development, up-gradation and maintenance of DAHC Server Software (DSS)

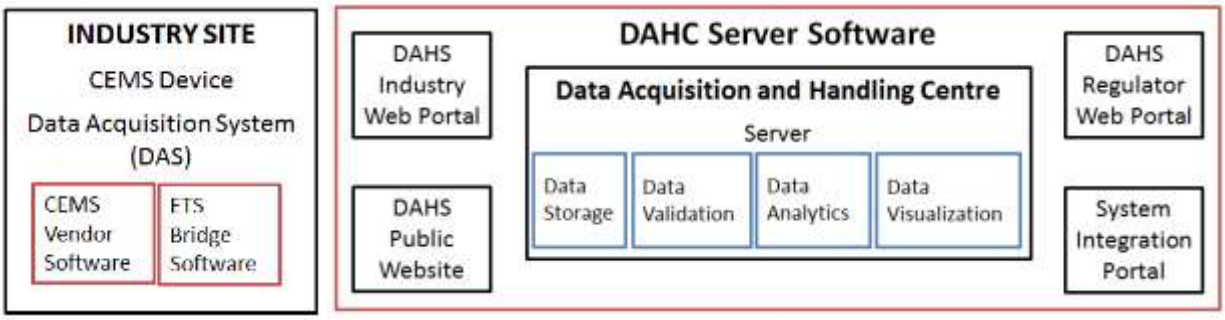


Figure 1: Representation of different software programs, their various components and their locations.

Table 1 : Scope of Work	
1.	Development of Design Documents for the entire project activities
2.	Supply of the proposed EBS (software responsible for the data transfer from industry DASs to the server) needed for carrying out all the requirements as specified in this document, with provision for easy installation such that industries can install it at their specific locations by themselves.
3.	Supply and installation of the proposed DSS (software capable of real-time data collection from EBS, data storage, viewing of real-time data, data validation, data analysis, and hosting the web portals for various users, etc.), and its various component carrying out all the requirements as specified in this document
4.	Up-gradation and maintenance of the software for a period of five years after commissioning. Support to all SPCBs which will adopt the SW, later on, without additional cost.
5.	Manpower support of one technically competent SW development experts for five years after commissioning for minor change requirements.
6.	In-person training of comprehensive software operation for 20 DAHC staff after the go-live phase.
7.	Entire source code with proper comments, data flow diagrams needs to be provided
8.	Module wise systematically written user manuals and training material of the software.
9.	The scope of work for software developer will include incorporating functionalities to enable a linkage to trading platform to use the emission data from DAHS for trading purposes etc..

*In addition, before final commissioning the software firm needs to get the software audited through CERT-In empanelled firms for cyber security, so that the web portals and website are eligible for hosting at NIC domain.

1.3. References

1. Notice inviting Expression of Interest (EoI) for "Development of DAHS Software for Pilot ETS project" issued by CPCB on 23.01.2014. [Thus the responses of this EoI should reach CPCB on or before February 12, 2014 (3.00 PM) & shall be opened at 3.30 PM on the same day]

2. 'SPECIFICATIONS AND GUIDELINES FOR CONTINUOUS EMISSIONS MONITORING SYSTEM (CEMS) FOR PM MEASUREMENT IN EMISSIONS TRADING SCHEME' (hereby referred to as "CEMS Specifications"). The role that CEMS will play in the upcoming pilot regulatory system of ETS, the technical requirements for the CEMS Hardware, Standard Operating Procedures (SOP) of various activities in the CEMS regime, the technical requirements for the CEMS Software, etc. are mentioned in "CEMS Specifications". Various functionalities of EBS and DSS, as laid down in the current document, have to be built as per the procedures, calculations and details laid down in the CEMS Specifications. The minute procedural details laid down in "CEMS specifications" will be provided to the work awardee firm by CPCB at final software development time.

   Some excerpt of the CEMS Specifications: 1) ETS Bridge Interface (Appendix C), 2) CEMS Operating Flowchart (Appendix D) and 3) System Integration Test Flowchart (Appendix E) are added in the document, to understand complete scope of work.

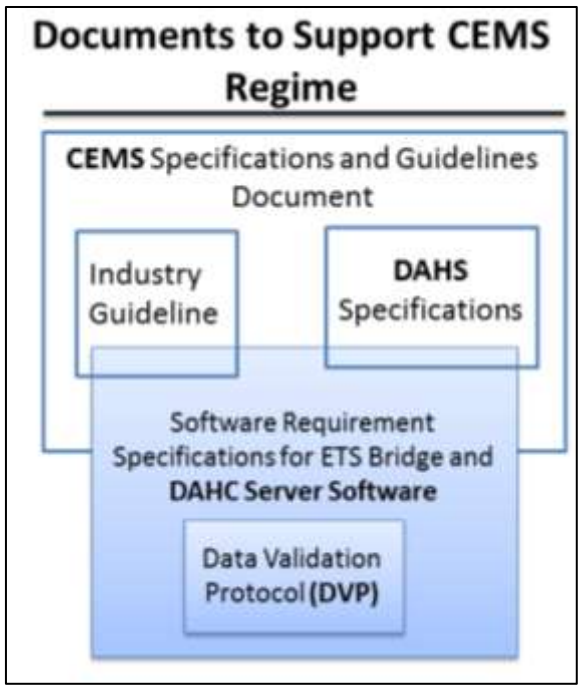


Figure 2: Diagram representing various documents to support CEMS Regime and their relationship

1.4. Document Overview

This document describes the requirements of the software needed to be developed by Software firms for establishment of DAHS. Section 2, gives the overall overview of the DAHS (architecture, schematics, parameters, etc.). Section 3, describes the requirements of the ETS Bridge Software. Section 4, describes the requirements of the DAHC Server Software. Section 3 and 4 contain details of the functional as well as non-functional requirements of the software programs along with the details of the users. Other relevant information is given in the appendices.

2.0. Data Acquisition and Handling System Overview

Data Acquisition and Handling System (DAHS) comprising of software and hardware components is designed to support the CEMS regime. Data Acquisition and Handling system (DAHS) will include the following components:-

Table 2: Components and Responsible Stakeholders of DAHS

S.No.	Component	Responsible stakeholder
1.	PM CEMS Device installed at industry	Industry
2.	Data Acquisition System (DAS) setup at the industry site.	Industry
2.1	CEMS Vendor Software will be installed in the DAS	CEMS Device Vendor
2.2.	ETS Bridge Software will be installed in the DAS	*To be developed by SW Development firm*
3.	Data Acquisition and Handling Centre (DAHC) located at Central Pollution Control Board (CPCB) and State Pollution Control Board (SPCB)	
3.1	DAHC Server Software will be installed on the Server at DAHC. *The functionalities at each site may vary slightly in order to reflect the functions of state and central regulators.	*To be developed by SW Development firm*
4.	Web Portals from which the CEMS data can be viewed and analyzed with various features by different users and stakeholders.	*To be developed by SW Development firm*

2.1 Brief Description of Software Programs

1. **CEMS Vendor Software** (CVS) is the software provided by the vendor to capture the data from the CEMS device and further communicate with the ETS Bridge Software on a file-based protocol [as detailed in CEMS Specifications] to send the data. Each CEMS vendor participating in the pilot ETS project must incorporate this interface into their CEMS software package and pass the System Integration Test to ensure compatibility with ETS Bridge Software. This software can be bought off the shelf or can be outsourced to another IT firm, but it must pass the System Integration Test[1] and thus be compatible with ETS Bridge software interface.

2. The **ETS Bridge Software (EBS)** should be a standalone application that will reside on the DAS at each industry. This software should have a standard interface[2] (file and parameter specification, and data transfer protocol as defined in CEMS Specifications) for receiving real-time data from (or communicating with) the CVS residing on the same machine. It should further send this data to the DAHC Server at the concerned SPCB (and to CPCB server in a parallel) in a secure and reliable manner. The bridge should also store this data locally for a user-specified period.

3. The **DAHC Server Software (DSS)** will comprise of multiple components i.e.

    1) Software which will be installed on 4 servers (subject to increase in future) in respective DAHC at the PCBs (CPCB and 3 SPCB). This component will be responsible for storage of the raw data coming from EBS in real time, its validation, the analysis of data, generation of reports, and other functionalities as defined by CPCB.
    2) User interface for the System Administrator
    3) ETS Website for Public
    4) Web portal for Industries (1000 at present, subject to increase)
    5) Web portal for Regulators
        a. Regulator I: SPCB staff
        b. Regulator II: CPCB Staff

        The differences between portals for CPCB and SPCB are detailed out in Section 4.1. On a broad basis, the functionalities of CPCB staff will be a subset of the functionalities of SPCB staff, and this difference of functionalities will need to be worked on the Web-portals as well.
    6) System Integration Portal for CEMS Device Vendors

The software installed on the DAHC Servers will host these web portals and will be responsible for collecting and sending relevant data from them, and as and when NIC domain server is made available, the web portals and website will be placed in the NIC domain. A provision for this must be made by the software development firm from the beginning.

[1] Refer to Appendix E : SYSTEM INTEGRATION TEST FLOWCHART
[2] Refer to Appendix C : ETS BRIDGE INTERFACE

2.2 DAHS Architecture and Schematics

The proposed DAHS architecture forms a well-defined network of data collection and transfer which together protect data quality and provide a record of accurate data to support the pilot scheme. In the proposed system, the PM CEMS device sends emissions data to the CEMS Software which is installed on the DAS. Among other parameters, the CEMS device will send raw emissions measurements (uncalibrated data), diagnostic parameters such as the status and health of the instrument (called diagnostic flags), flow measurements, and temperature readings. The complete list of parameters and their prescribed format is discussed in Specifications and Guidelines for Continuous Emissions Monitoring System (CEMS) for PM Measurement in Emissions Trading Scheme. Upon accepting emissions data from the CEMS device, the CEMS Software immediately converts it into the prescribed format, after which the ETS Bridge Software reads the data, stores it locally, and sends it to the SPCB and CPCB server on real time basis.

Figure 3: DAHS Architecture

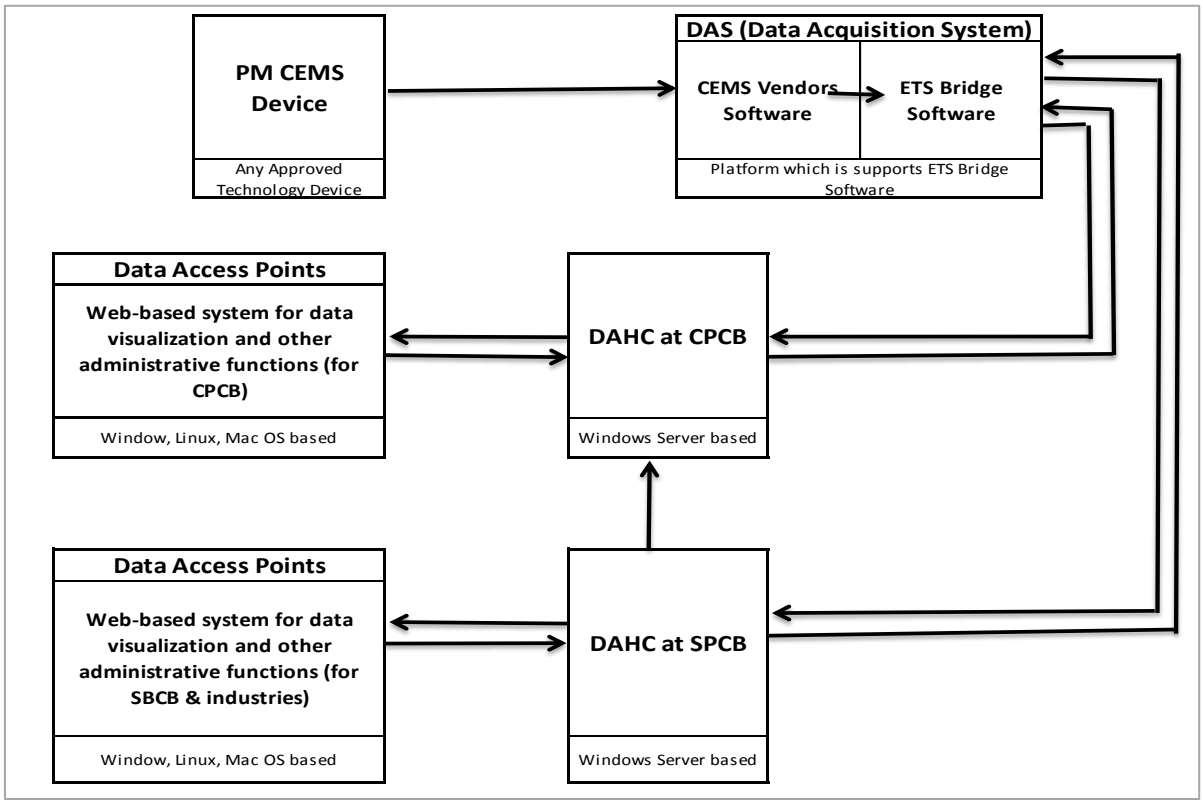


Figure 4: The figure overleaf broadly shows the data flow and various activities of the entire Data Acquisition and Handling System.

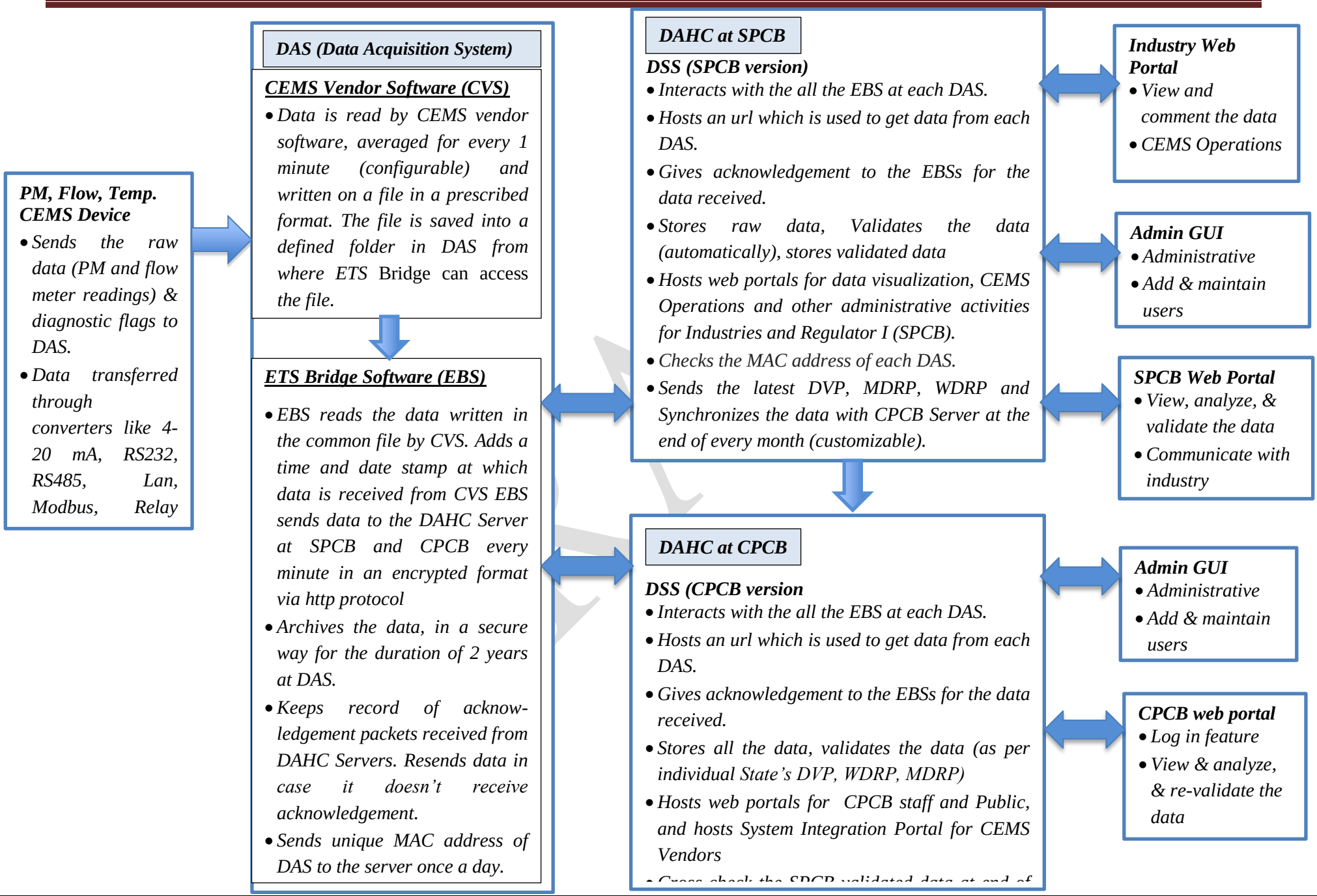
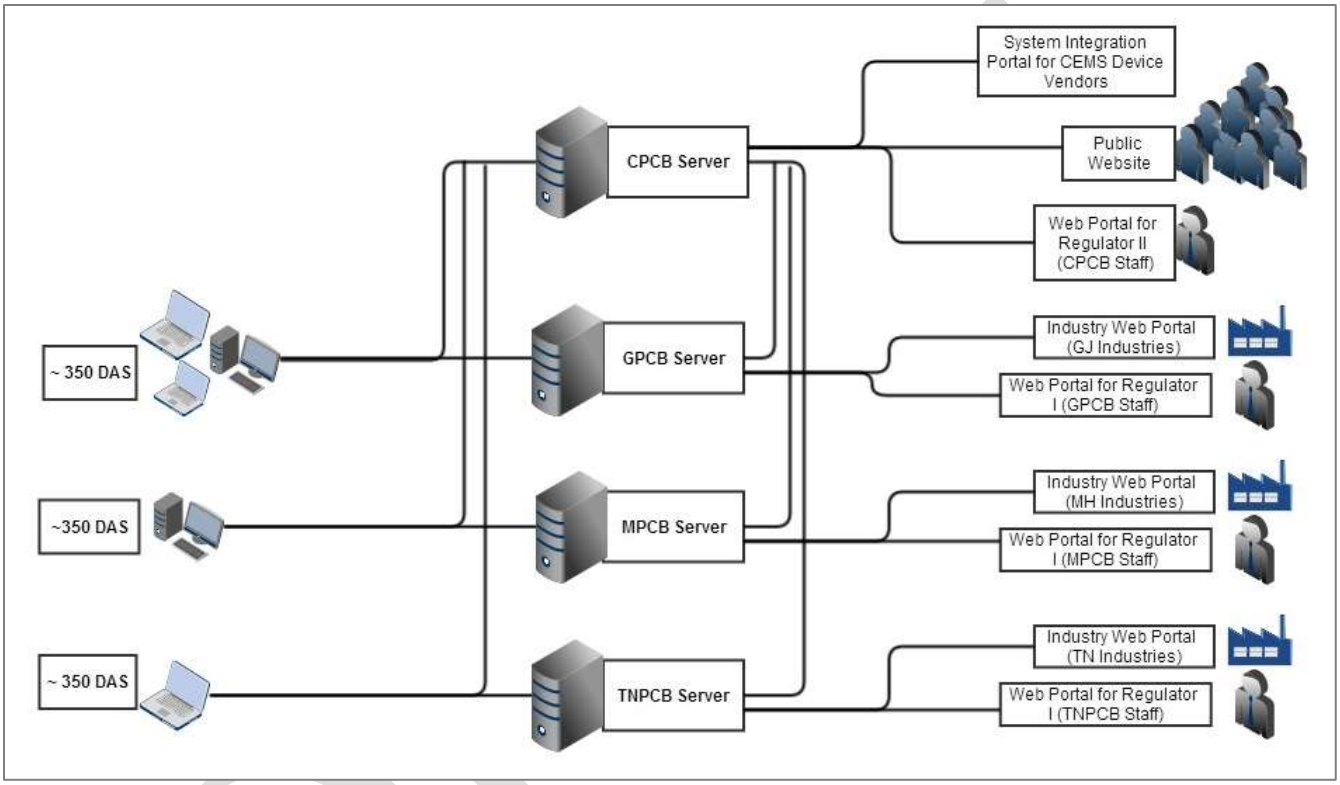

PM, Flow, Temp. CEMS Device
• Sends the raw data (PM and flow meter readings) & diagnostic flags to DAS.
• Data transferred through converters like 4-20 mA, RS232, RS485, Lan, Modbus, Relay
DAS (Data Acquisition System)
CEMS Vendor Software (CVS)
• Data is read by CEMS vendor software, averaged for every 1 minute (configurable) and written on a file in a prescribed format. The file is saved into a defined folder in DAS from where ETS Bridge can access the file.
ETS Bridge Software (EBS)
• EBS reads the data written in the common file by CVS. Adds a time and date stamp at which data is received from CVS EBS sends data to the DAHC Server at SPCB and CPCB every minute in an encrypted format via http protocol
• Archives the data, in a secure way for the duration of 2 years at DAS.
• Keeps record of acknow-ledgement packets received from DAHC Servers. Resends data in case it doesn't receive acknowledgement.
• Sends unique MAC address of DAS to the server once a day.
DAHC at SPCB
DSS (SPCB version)
• Interacts with the all the EBS at each DAS.
• Hosts an url which is used to get data from each DAS.
• Gives acknowledgement to the EBSs for the data received.
• Stores raw data, Validates the data (automatically), stores validated data
• Hosts web portals for data visualization, CEMS Operations and other administrative activities for Industries and Regulator I (SPCB).
• Checks the MAC address of each DAS.
• Sends the latest DVP, MDRP, WDRP and Synchronizes the data with CPCB Server at the end of every month (customizable).
DAHC at CPCB
DSS (CPCB version
• Interacts with the all the EBS at each DAS.
• Hosts an url which is used to get data from each DAS.
• Gives acknowledgement to the EBSs for the data received.
• Stores all the data, validates the data (as per individual State's DVP, WDRP, MDRP)
• Hosts web portals for CPCB staff and Public, and hosts System Integration Portal for CEMS Vendors
Industry Web Portal
• View and comment the data
• CEMS Operations
Admin GUI
• Administrative
• Add & maintain users
SPCB Web Portal
• View, analyze, & validate the data
• Communicate with industry
Admin GUI
• Administrative
• Add & maintain users
CPCB web portal
• Log in feature
• View & analyze, & re-validate the data

Figure 5: The diagram broadly shows different types of servers, web portals and the relationships between them

(Note :

1) Features for each type of user i.e. users in different web portals will be different.
2) The DSS should be developed such that in case of emergency/failure at SPCB servers, CPCB server should be able to host their respective portals. The CPCB server, essentially, will serve as a backup and emergency system.)



2.3. Brief Description of Data

The term 'data' can refer to multiple kinds of variables in the CEMS system. The table given below classifies different types of variables along with their sources.

Table 3 : Description of DAHS Data

S.no.	Variable	Name	Source
1.0	**Administrative Data**		
1.1	Industry Details (Industry registration form)		Manual data collection at SPCB
1.2	CEMS Registration Form		Industry Web Portal
1.3	CEMS Installation Checklist		
1.4	CEMS Functional Checklist		
1.5	Isokinetic Sampling Reporting Form		
1.6	Comments on validated data		
1.7	Administrative action (accept or reject) of Industry comments on validated data		Regulator Web Portal
2	**Configuration Data**		
2.0	Type of Pollutant	ToP	Industry Web Portal
2.1	Type of Configuration	ToC	
2.2	Comp. ID	CompID	Calculated by the DSS
2.3	Stack Dimension	Type of stack Length/Breadth Dia	Industry Web Portal
3	**Calibration Data**		
3.1	Average Pressure during Sampling	Pres_Smpl	1) Industry Web Portal 2) Calculated by the DSS
3.2	Average temp. during sampling	Temp_Smpl	
3.3	Calibration Factor (m)	CF_m	
3.4	Calibration factor (c)	CF_c	
4	**Measurement Data**		
4.1	Date and Time Stamp	DTS	1) PM CEMS Device 2) CVS 3) EBS
4.2	Pressure at Stack	Pres	
4.3	Temp. at stack	Temp	
4.4	PM CEMS raw Uncalibrated average	PM_Uncal	
4.5	PM CEMS Calibrated average	PM_Cal	
4.6	Velocity	Velocity	
4.7	Normalized Flow	N_flow	
4.8	PM mass	PM_mass	
4.9	Device Mode	DM	
4.10	Power Status	PS	
4.11	Alarm	Alrm	
4.12	Maintenance Alarm	M_Alrm	

[Note: In addition, other variables, as project progresses will need to be added. Hence, a flexible or modular system should be developed.]

3.0. Requirement of ETS Bridge Software

I	Functional Requirements of ETS Bridge Software
1.	**Real-time data collection (reading)**
1.1.	Collection of the data[3] i.e. reading the file in prescribed format[4] (created by CEMS Vendor Software) as per ETS Bridge Interface[5] on a real-time basis.
1.2.	Tracking the received data from CEMS Vendor Software against time.
2.0	**Data Storage**
2.1.	Storage (for 2 years) of the CVS read data in the encrypted form in the hard-drive of the DAS
3.0.	**Real-time data transfer to DAHC Server**
3.1.	Real-time data transfer to DAHC Servers using HTTP protocol in an encrypted format[6]
3.2.	Tracking the data sent and received by the DAHC Servers at CPCB and SPCB
II	**Non-functional Requirements of ETS Bridge Software**
1.	ETS Bridge should be a light weight standalone application that is very easy to install and upgrade.
2.	ETS Bridge should be compatible with Windows XP, Windows 7 (32 bit and 64 bit) and Windows 8 (32 bit and 64 bit) platforms. EBS shall me made available on the Industry Web Portal for downloading.
3.	Compatibility to the latest technologies (or latest operating systems and other software programs) should be maintained though out the project duration.

3.1. Details for ETS Bridge Software User(s)

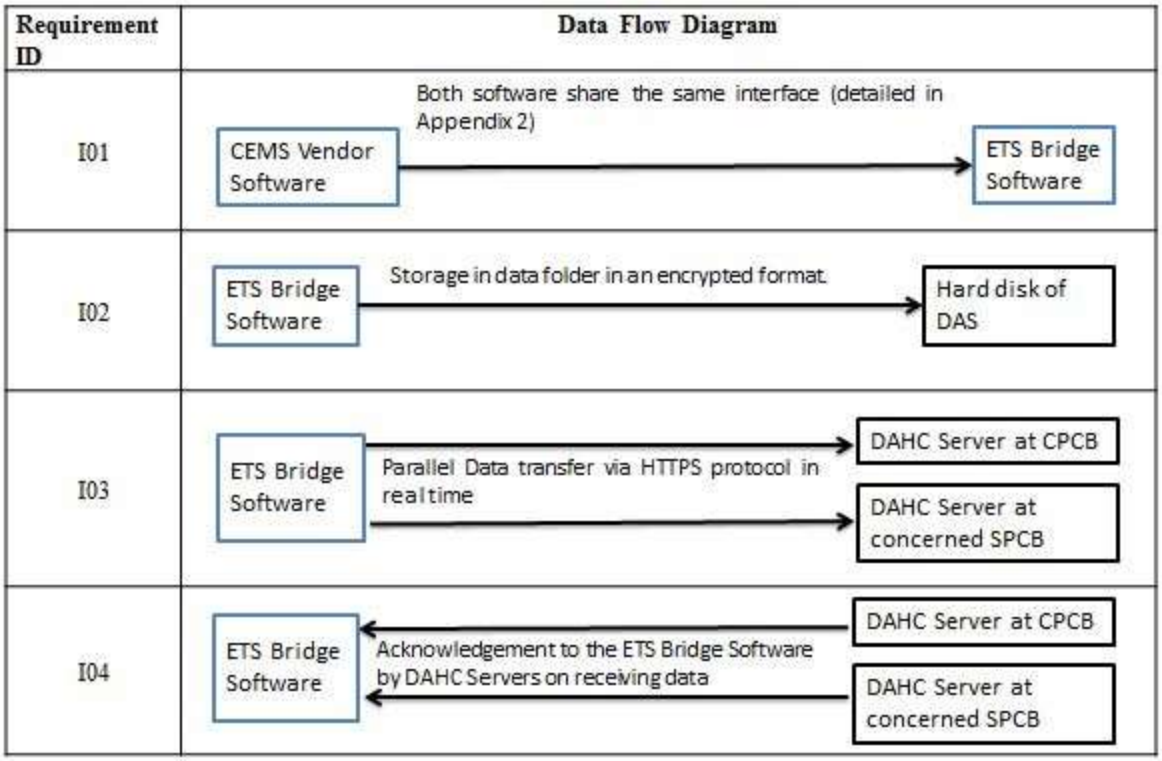
ETS Bridge software has one primary user, i.e. the Industries (also known as operator), the requirements of which are listed below:

Require-ment ID	Description	Other involved entities
I01	Data Collection	CVS
I02	Tracking data received against time	CVS
I03	Data Storage in encrypted format	None
I04	Data Transfer	DSS
I05	Receiving the confirmation of successful data transfer from DSS	DSS

[3] Description of the data has been detailed in the CEMS Specifications.
[4] Refer Appendix C2.3
[5] Refer to Appendix C
[6] Encryption format to be decided by the software firm on consultation with CPCB

3.2. Flow diagrams for ETS Bridge Software

Figure 6: Dataflow for ETS Bridge



4.0 Specifications of DAHC Server Software

I	Functional Requirements of DAHC Server Software
1.	**Data collection and acknowledgement**
1.1.	Real-time collection of data from ETS Bridge Software installed at each DAS.
1.2.	Acknowledging[7] the data received (every minute) to the ETS Bridge Software.
1.3.	Collection of calibration data and other administrative data [8] from Industry Web Portal.
2.	**Data Storage**
2.1.	Storage of raw data received from DAS on the Server Database.
2.2.	Storage of the validated data (after running series of predefined checks on raw data as per Data Validation Protocol (DVP))
2.3.	Storage of the data post MDRP (Missing Data Replacement Protocol) and WDRP (Wrong Data Replacement Protocol), procedures defined to address issues identified during validation.
2.4.	Storage of administrative information like User IDs and Passwords.
2.5.	Storage of data[9] obtained from Industry Web Portal like Industry details, CEMS details, calibration factors, CEMS operations activities details etc.
2.6.	Storage of data obtained from Regulator I (SPCB) Web Portal like administrative actions, acceptance or rejection of industry comments.
2.7.	Tracking change in data post Data Replacement Protocol, post Industry comments, and post regulator acceptance/rejection.
2.8.	Tracking change in the DVP, MDRP and WDRP with time.
3.	**Calculations[10], data analysis[11] and CEMS Operations[12] Activities**
3.1.	Calculations of measurement data parameters in real-time (such as calculations of calibrated data based on uncalibrated data)
3.2.	Calculations for various CEMS Operations Activities such as calibration factor based on Isokinetic Sampling Reporting form at Industry Web Portal.
3.3.	Performing the background calculations of Post-Calibration performance test.
3.4.	Performing the background calculations of CEMS Audit performance test.
3.5.	Data analysis of PM emissions (average PM concentration, total emissions) as per various parameters like sector, sub-area, size of industry, duration (daily, weekly, monthly or any duration) etc.
3.6.	Report generation[13] and fixed queries.
3.7.	Monthly, weekly reports to be generated (in pdf format).

[7] Format of acknowledgement messages is to be decided.
[8] The functionality needs to be carried out in accordance to the format of data as stated in the CEMS Specifications.
[9] Description of the data has been detailed in the CEMS Specifications.
[10] This functionality need to be carried out in a/c to the calculations detailed in CEMS Specifications.
[11] Details of data analysis routines will be partially provided by CBCB during development phase and partially will need to be worked out by SW development firm.
[12] Refer Appendix D: CEMS Operations Flowchart for details
[13] Format of reports will be provided by CPCB after award of work during the development phase itself.

3.8.	Run analysis routines.[14]
3.9.	DSS and Various web portals need to be able to assist in performing CEMS Operations Activities such as:- 1) PM CEMS Installation, Registration and Configuration. 2) Calibration and Post-Calibration Performance Test. 3) Re-calibration. 4) CEMS Audit and CEMS Audit Performance Test. 5) Commenting on the CEMS Data by Industry.
4.	**Data Validation[15] and replacement[16] as per the protocol**
4.1.	Validation of the data according to the specified data validation protocol (DVP).
4.2.	Replacement of data as per Missing Data Replacement Protocol (MDRP) and Wrong Data Replacement Protocol (WDRP).
4.3.	All the above mentioned data validation and replacement functionalities (4.1, 4.2, 4.3) need developed as a module with a set of feature corresponding to individual segments mentioned in DVP, MDRP or WDRP which can be chosen/changed/removed by the Server Admin. Hence, the software implementation module should have a provision to customize the DVP, MDRP or WDRP by the Server Admin, after which the validation and replacement will happen automatically as per the chosen protocol.
4.4.	Reflecting the results of DVP, alarms for the flagged values and the results of MDRP and WDRP on the Industry, Regulator I and Regulator II web portals.
4.5.	Communicating the chosen DVP of every SPCB to the CPCB server at the beginning of financial year.
4.6.	Synchronizing the validated and commented data from SPCB server with CPCB server at the end of every month.
5.	**Hosting web portals for data collection, viewing, and reporting**
5.1.	Web based GUI to view the real-time raw data, validated data, alarms, and CEMS operations data for regulator, industries and public.
5.2.	User friendly platform to write and process run-time queries. Some of the examples of the GUI interfaces are in Appendix A.
5.3.	System Integration Portal to carry out System Integration test[17] of CVS to assure compatibility of EBS and CVS.
5.4.	In case of emergency at SPCB DAHC (which holds Regulator I portal & Industry Portal), provision should be made that portals can be hosted at CPCB DAHC for temporary duration.
6.	**Administrative features**
6.1.	An Administration Utility to add / update various industries, devices, users etc.
6.2.	Generating and storing User IDs for regulator and industries (with different features)

[14] Refer Appendix B3: Analysis Routines for details
[15] Refer Appendix B1: Data Validation Protocol for details
[16] Refer Appendix B1: Data Replacement Protocol for details
[17] Refer Appendix E: System Integration Test Flowchart for details. Further details have been laid down in CEMS Specifications.

II	Non-functional Requirements of DAHC Server Software
1.	DAHC Server should have a scalable and robust web-based architecture
2.	DAHC Server Software should be made on Windows Server 2012 platform.
3.	Appropriate exception / error handling / input validations must be done throughout the system.
4.	The web portals and other UIs should be intuitive to use and easy to maintain.
5.	The web portals and websites should be compatible to mostly used browsers like Windows Explorers, Firefox, Google chrome and Safari etc.
5.	The web portals should be able to handle simultaneous request from multiple clients.
6.	Adequate security measures should be built into the system as per NIC guidelines.

4.1. Details for DAHC Server Software User(s)

Below are the requirements for six different types of user:

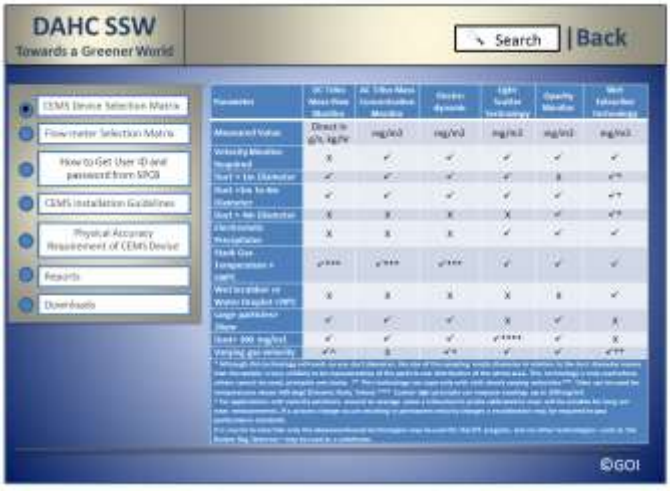
Require-ment ID	Description	Component of DAHS which will be involved in the requirement
1) USER : INDUSTRY (also referred to as Operator)		
I01	Selection of CEMS Device	Public Website
I02	Generating User IDs for the Industry users	DSS (after industries submit the forms manually at the SPCBs)
I03	Registration of CEMS device	Industry Web Portal, Storage at DSS
I04	Filling Installation Checklist	
I05	Filling Functional Checklist	
106	Downloading and Installing ETS Bridge Software	Industry Web Portal.
I07	Configuration of the CEMS Vendor Software	Industry Web Portal, Retrieval of data from DSS
108	Iso-kinetic sampling for calibration of PM CEMS Device, Filling the Isokinetic Sampling reporting form	Industry Web Portal, Storage at DSS
I09	Post calibration performance test	Calculations at DSS, Retrieval of data from DSS
I10	Regular Emissions Data Transfer	ETS Bridge Software
I11	Viewing emissions data (raw and validated), flags and alarms.	Industry Web Portal, Retrieval of data from DSS
I12	Commenting on validated emissions data within a defined time period. (The time period has to be customizable by the DAHC Admin)	Industry Web Portal, Retrieval of data from DSS, Storage of comments on the DSS
I13	Re-calibration alarm (if required or if applicable)	Industry Web Portal, Retrieval of data from DSS
I14	**Re-calibration** : comprises of Step I09 and I10 in addition to storage of new calibration data in the DAHC Server	Industry Web Portal, Calculations at DSS, Retrieval of result from DSS
I15	CEMS Audit and Result display	

Note: Industry users can view their own data only through their web-portals after logging in with their username and password.

Require-ment ID	Description	Component of DAHS which will be involved in the requirement
2) USER: REGULATOR I (SPCB Staff)		
S01	Generating User IDs for the CPCB Staff users	DSS
S02	Raw Data View	Regulator I Web Portal, Retrieval of data from DAHS at SPCB
S03	Validated Data View	
S04	Accepting or Rejecting the Industry comments to the Validated data. Accepting or Rejecting the Industry Admin forms.	Regulator II Web Portal, Retrieval of data from DAHS at SPCB, Storage of data on the SPCB DAHS
S05	Data Analytics	Regulator I Web Portal
S06	Viewing Reports	
S07	CEMS Audit and Result display	Industry Web Portal, Calculations at DSS, Retrieval of data (result) from DSS
3) USER: REGULATOR II (CPCB Staff)		
C01	Generating User IDs for the CPCB Staff users	DSS
C02	Raw Data View	Regulator II Web Portal, Retrieval of data from DAHS at CPCB
C03	Validated Data View	Regulator II Web Portal
C04	Data Analytics	
C05	Viewing Reports	
C06	CEMS Audit Result display	Industry Web Portal, Calculations at DSS, Retrieval of data (result) from DSS
4) USER: DAHC ADMIN (SPCB/CPCB)		
A01	Generating username and passwords and keeping track of it	DSS
A02	Selecting DVP, MDRP, WDRP	Admin UI of DSS
A03	Data View	
A06	Approval of commented data for change in the data base.	Admin UI of DSS, Regulator I/II Web Portal
5) USER: CVS VENDORS		
V01	Register the CVS	System Integration Portal
V02	Carry out System Integration Test to make CVS ETS Compatible	
6) USER: PUBLIC		
P01	Display of relevant information and reports on web	ETS Public Website
P02	Providing different relevant reports and documents on web for downloading (which should be available in different relevant formats like excel, word or PDF)	

Note: All the web portals and websites need to be developed in a consultation with CPCB regarding aesthetics, contents and functionalities.

APPENDIX A: SAMPLE USER INTERFACES OF DAHC SERVER SOFTWARE WEB PORTALS




DAHC SSW
Towards a Greener World
INDUSTRY PORTAL
Industry Name| Home|
EMISSIONS CHART
120
100
80
60
40
20
0
Industry Emission Data
Comparison Parameters
Zoom-in Parameters
Emissions Summary
Your Load:
Your Sector Load:
Cluster Load:
Trade your Emissions
Current Calibration Factors:
Last SATA Data and Result:
Latest Reconfiguration:
Last Calibration date and Results:
Download ETS Bridge Software
©GOI


DAHC SSW
Towards a Greener World
INDUSTRY PORTAL
Industry Name| Home|
Last month load:
Total Load for this year:
Avg. conc.:
% QA Data:
% Available Data:
Emissions Data
Calibration Data
Audit Results
Trade your Emissions
Calibration/Reconfiguration
Date
RMSE
Calibration Range
©GOI


DAHC SSW
Towards a Greener World
Welcome to the DAHC SW Web Portal
INDUSTRY PORTAL
REGULATOR PORTAL
PUBLIC WEBSITE
SYSTEM INTEGRATION PORTAL
Click on the button to go to the relevant portal
© GOI


DAHC SSW
Towards a Greener World
REGULATOR PORTAL
Regulator Name| Home|
EMISSIONS CHART
120
100
80
60
40
20
0
Emissions Summary
Cluster Load:
Total A level Administrative actions
Comparison Parameters
Zoom-in Parameters
- Industries
- Time segment
- Time aggregation
- Hourly
- Daily
- Weekly
- Monthly
- Quarterly
©GOI


DAHC SSW
Towards a Greener World
REGULATOR PORTAL
Regulator Name| Home|
DIMENSION FILTERS
INDUSTRIES
Sector
Size
Cluster
DATA
Data Quality
Total Flagged Data
EMISSIONS CHART
120
100
80
60
40
20
0
Zoom-in Parameters
- Hourly
- Daily
- Weekly
- Monthly
- Quarterly
Account Setting
Logout
Help
©GOI

[Note: The format of data validation protocol is fixed entity and can be programmed as hard code, but most value of parameters and the relevant subset from the pool of parameters are flexible. Hence, the software should be written in such a way, that values of the parameters can be customized as over the experience of data received. The protocol defined might require some adjustments later on hence; a modular approach should be adopted in developing this section so as to be able to accommodate any changes in these protocols at a later stage.]

APPENDIX B1: DATA VALIDATION PROTOCOL (DVP)

The automated nature of the DVP proposed for the pilot ETS distinguishes it from most other ETS programs, where data validation is manual and usually entails third-party auditors visiting industries to review and verify emissions data. Compared to its manual counterpart, automated data validation typically involves less time and labour, and high transparency. To achieve this, many important technical specifications must be addressed in order to formulate robust validation software and optimize its performance. For the pilot ETS DVP, these include:

1) Description of invalid data
2) Method of detecting invalid data

B1.1. Data Checks for Pilot ETS

Incoming emissions data at the SPCB servers may include errors such as invalid formats, negative values, noise, drift, outliers, etc., in addition to the possible measurement deviations related to the sensor performance itself. This 'raw' data needs to be validated in order to ensure its reliability for supporting the ETS, such as for determining permit compliance. The aim of data validation, then, is to identify errors, discrepancies, or inconsistencies in the values monitored and reported for each variable, so that appropriate action can be taken to ensure only data of the highest quality is used for enforcing ETS regulatory requirements.

B1.2. Data Checks

Data checks are structured to run on each incoming emissions data on a point-by-point basis. That is, as each individual data is inputted into the DAHC Server database, it will go through a series of data checks aimed at testing its validity. Each data check takes the form of a piece of computer code which specifies criteria by which a value is deemed valid. If the value for a variable does not pass the test, it is marked as invalid. Thus, after data is processed by the initial data checks, they will be classified as (i) **valid** or (ii) **invalid** (0=valid; 1=invalid). The initial data checks consist of the following broad categories, each containing specific checks:

i. Missing value check
ii. Format Checks
iii. Consistency Checks
iv. Constant value checks

[Note that data checks (i)-(iii) look at each point independently while (iv) is a time-series check and considers points in relation to others around it.]

B1.2.1 Missing Value Check

For the pilot ETS program, industries are accountable for their emissions during each hour of operation, since compliance is determined by comparing total mass of PM to the number of permits held. Thus, a complete per minute basis record of PM emission is required, and data must be reported every minute without exception. In real-life situations PM data may not be available for some time periods due to instrument malfunction, downtime for maintenance, or any other reasons. If data records received by

SPCB contain missing PM mass flow or concentration, the data is rendered invalid. Thus, the purpose of the missing value checks is to detect missing values and invalidate them on a point-by-point basis. If a value is missing, it will be invalidated and a corresponding dummy variables called "missing" will be marked with a "1".

B1.2.2 Format Checks

Format checks are simple tests for ensuring:

1) All fields are in their predefined valid format
2) Not applicable fields are correctly identified

Format checks will run over each data point as it is accepted at the DAHC server. If a variable does not pass the format check, it will be invalidated. For example, if the composite ID of the industry does not match the accepted format, the full data row will be invalidated since the mismatch could indicate that the data is coming from an unauthorized source.

B1.2.3 Consistency Checks

Consistency checks test whether data values are consistent with a given range for the associated variable, and whether it conforms to any methodological requirements. In particular, consistency checks are primarily designed for ensuring:

1) Negative values are invalidated
2) Range checks: data values are within given reasonable range for the associated variable. For example, PM mass flow cannot continuously exceed 120% of the maximum mass flow during certification.
3) Redundancy check: no duplicate row exists in the database

An especially important case is the **out-of-range check**, which is designed to identify when PM data measurements continuously exceed the calibration range. In particular, if PM data exceed 120% of the calibration range for 10% of the week for 5 weeks in a row, the data is rendered invalid. For any one week, if data exceeds 120% of the calibration range 40% (or more) of the time, it is marked invalid. The purpose of detecting out-of-range values is to determine a case when recalibration is required. Thus, when PM data do not pass the out-of-range check, the data is both invalidated and re-calibration is required.

B1.2.4 Constant Value Checks

Constant values can be of three types:

1) Constant at a ceiling
2) Constant at zero
3) Constant anywhere within instrument range

Strings of constant values of any of the above types will be invalidated on a point-by-point basis, considering the values of previous points. SPCB regulators can view the constant values on their user interface and replace them according to a replacement protocol (refer Section B2 for details). One way to assess the constant values is to check the device/power status at the time of constant readings.

B1.2.4.1 Ceiling Check

A ceiling check is actually a special form of the constant value check. A ceiling, or ceiling flatland, describes an interval in a time-series dataset where values level off in a plateau fashion at the maximum instrument range. A technical characteristic of PM CEMS is that when PM concentration (or mass flow) is greater than the maximum range of the instrument, the reading will appear as the maximum detectable value itself. For example, if the instrument range is 0-200 mg/s and the actual PM mass flow is 250 mg/s,

the CEMS will read 200. Because of this phenomenon, it is not known whether a data point equal to the maximum detectable value is the actual reading, or an out of range value replaced by the maximum value. A ceiling flatland, where successive values at the maximum range are observed, is of greater concern since it is highly unlikely that the CEMS gives successive readings which are exactly equal to the maximum value (rather, the true PM values are probably beyond the detectable limit for the interval).

The simplest way to detect and flag ceilings is to first flag points which are equal to the maximum instrument range. Next, if the slope between this point and the point before or/and after is zero, a ceiling flatland has been detected. In mathematical terms, ceilings can be detected using the following techniques:

1) y =max: if detected, first flag as error level 1 (point of concern)
2) At y=max, slope between point before or after=0: if detected, change flags for both points to error level 2 (doubtful).

The above process will be looped to detect all the data points in the ceiling. Wherever a constant string of values is detected at a ceiling, all data in the constant range will be invalidated.

B1.2.4.2 Constant zero check

The zero check detects continuous readings at zero. A continuous string of PM emissions at zero is rare because the CEMS should be continuously switched on. Even if the plant is shut off, the CEMS will detect small amounts of stray PM floating in the stack. Thus, emissions may vary near zero, but should not show exactly zero for successive minutes.

The zero check is simple and identifies zero readings. If the reading in the next minute is also zero, the data is invalidated. This process is looped to identify and invalidate constant zeros of any length. SPCB regulators can act on this data according to their replacement methodology.

B1.2.4.3 Constant Values Anywhere Within Calibration Range

Beyond constant readings at ceiling values and zero, constant values anywhere else will be invalidated. Given the inherent random nature of the PM emissions, and limitations associated with the sensors, successive constant values are suspiciously rare phenomena. Even if in two successive minutes, PM emissions are truly the same (or extremely similar), the CEMS reading for each minute will differ due to random noise associated with the instrument. For this reason, successive data points containing the exact same value will be invalidated and marked accordingly.

B1.3 Frequency of Data Checks

Data checks will be executed in real time on a point-by-point basis. That is, when each data row from the text file is inputted into the SPCB database, it will immediately be checked by the data validation software. A separate timeline has been designed for further processing and analysing the data and flagging potential errors. These analysis routines will be run on the data in 2-week cycles.

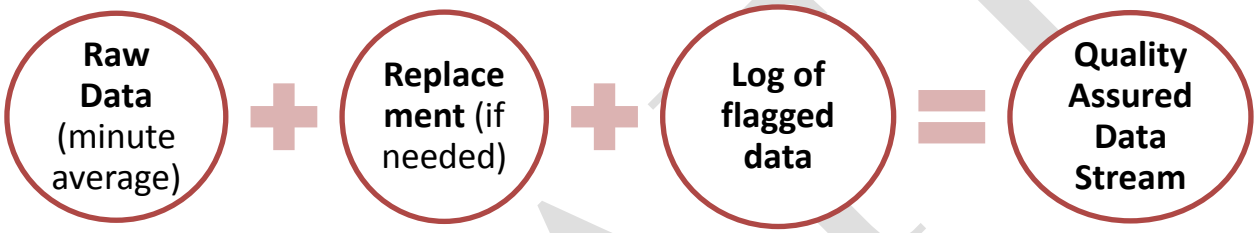
B1.4 Operationalizing Data Checks

The database at the SPCB will contain dummy variables corresponding to the outcome from each data check. As each data row from CEMS is transmitted to the SPCB database, the data checks will immediately be deployed and dummy variables for each check will be updated accordingly depending on whether the data is deemed valid or invalid. For example, if a data row comes in with missing PM emissions and a ToC>3, then the respective dummy variables for each check (missing check and range check) will be updated with "1". In this way regulators can view the indicator variables for each check and accurately trace the source of data invalidation.

APPENDIX B2: DATA REPLACEMENT PROTOCOL

B2.1 Protocol for dealing with flagged data

Since the success of ETS requires full accounting of total load from each industry, a system must be in place for dealing with erroneous CEMS data. Total load is calculated by aggregating the minute-by-minute data for a defined period, so accurate values are needed for each minute even if they must be approximated in cases where the raw data is flagged. Data are flagged when they are either missing or contain an error (type 1, 2, or 3). Because missing and type 2 and 3 errors cannot be used for compliance determination, respective data replacement protocols have been designed for substituting data which best represent the operating process. The missing data protocol is designed to substitute data when no readings are given, while the wrong data protocol is designed to replace measured values containing type 2 and 3 errors. Replaced values will be stored in separate variables while the flagged values will be logged and stored for reference. The replacement procedure thus follows the following process:

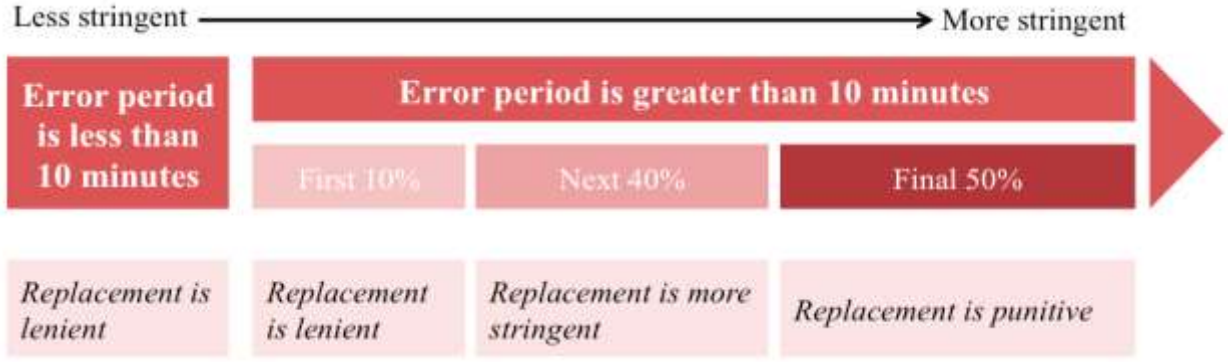


The data points that are valid and pass through all the data validation checks are then assigned to be Quality Assured (QA). The amount of previously QA data pertaining to any industry is taken into account while making replacements for flagged data. This also incentivises the industrial units to send QA data from their end, and thus operate and maintain the CEMS device as per the best practices.

B2.2 Wrong Data Protocol

The wrong data protocol is invoked when the CEMS device gives a reading whose value fails a validation check. The severity of the replacement depends on the length of the error period. If the value is flagged– say, for hitting a ceiling reading—for less than T_w minutes, the replacement would not be punitive. If the error lasts longer than T_w minutes, a step system of replacement is used in which subsequent blocks of data are replaced with increasingly conservative values. This gives industries the incentive to maintain their devices and ensure the period of error is at a minimum. The diagram below illustrates the nature of the step-like replacement procedure.

Figure 7: Step-wise Replacement Procedure, when T_w =10 minutes



With this in mind, the wrong data replacement protocol contains specific replacement procedures for each validation check and for each of the above "steps" as follows:

Table 4 : WRONG DATA REPLACEMENT PROTOCOL

Category	Check	Amount of Data	Error Type	Replacement Procedure/Action Taken
General - Format Checks	Data Format	All	1	Industry and regulator notified immediately
	N/A field giving value		3	Data flagged, logged, and not used. Industry notified
	Field showing "err"		2	Replaced a/c to MDRP
General – Consistency Checks	Negative values	10%	1	Average of value before and after
		Next 40%	2	Greater of either '90th percentile of last 24 hours of data' or 'average of values before/after'
		Next 50%	3	95th percentile of last 24 hours
	Out of Range	10% in a week for 5 weeks or more	3	Industry notified for recalibration
		40% in one week	3	
		Less than above amounts	2	150%*max. PM mass flow reading during calibration
	Duplicate rows	All	3	Duplicate row logged but data not used in an calculations
Time-Series – Raw data	Constant Value	10%	1	Average of value before and after
		Next 40%	2	Greater of either '90th percentile of last 24 hours of data' or 'average of values before/after'
		Next 50%	3	95th percentile of last 24 hours
	Ceiling Check	10%	1	110%*instrument range
		Next 40%	2	140%*instrument range
		Next 50%	3	150mg/nm3*max. designed flow for plant
Time-series – Smoothed data	Outlier Check	10%	1	Same as ceiling check
		Next 40%	2	
		Next 50%	3	
	Step-Change check	All	2	Inquiry into industry operations

[*Note: When step-replacement procedure used, error period length of greater than 10 minutes is assumed. If error length period is less than 10 minutes, replacement procedure is the average of value before and after.]

B2.3 Missing Data Procedure

A protocol has been decided for the purpose of substituting values for the missing data. This protocol refers to a set of formulas for calculating various CEMS operating variables, primarily for the purpose of applying missing data substitution values, in accordance to the quality of previously provided data and quantity of missing data. This protocol constitutes of a set of procedures that gauge the overall quality of the incoming data stream in order to determine how conservative or lenient approximations will be made for substituting values during missing data periods.

The main variables generated for this purpose are:

1) Percent Monitor Availability (PMA) = ratio of the number of quality-assured data hours to the total number of operating hours
2) Missing data period length (MDL)

Two distinct sets of missing data algorithms have been categorized: 1) "initial" and 2) "standard" routines. The initial missing data algorithms are temporary "spin-up" procedures that are used until 720 hours of quality assured CEMS data have been obtained. In particular, if no quality-assured (QA) data exists within 720 hours of certification, the maximum mass flow will be substituted. If QA data do exist within 720 hours of certification, the average of the value before and after will be substituted. After 720 hours have passed since certification, the standard missing data procedures will begin to be applied. Table 4 describes the initial missing data procedure.

Figure 8: Missing Data Replacement Procedure

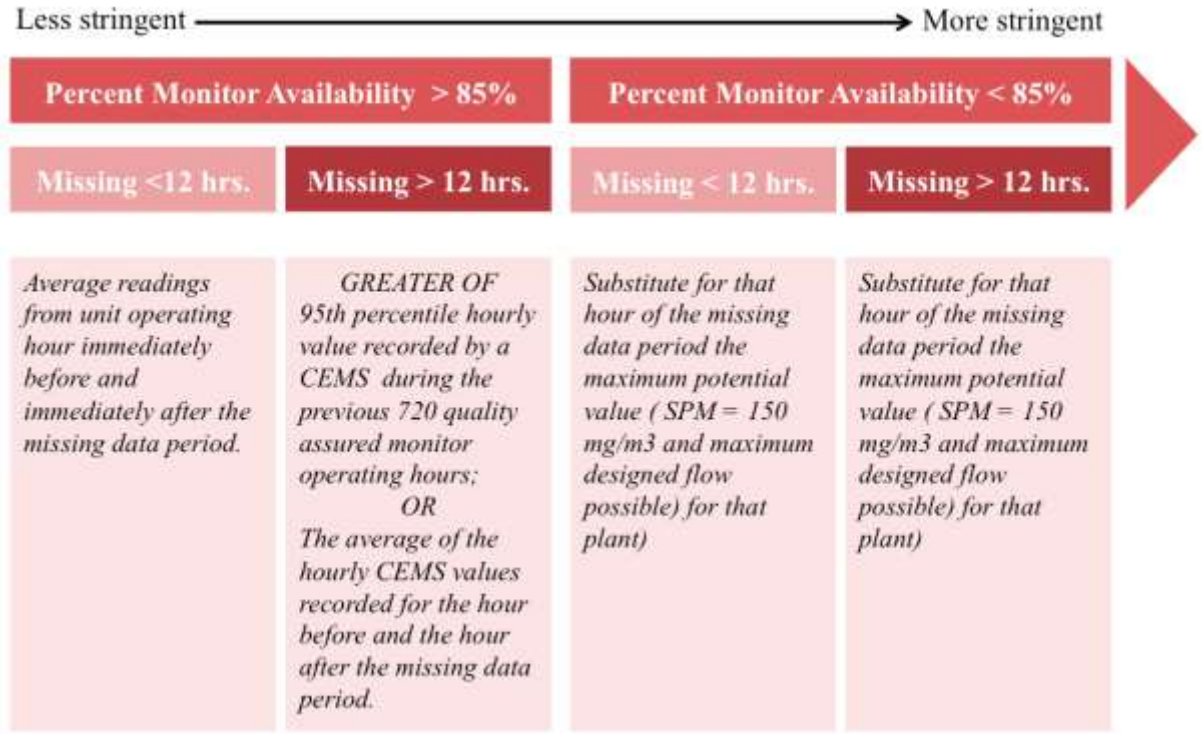


Table 5: INITIAL MISSING DATA PROCEDURE		
Condition on Time	Condition on Quality-Assured Data Availability	Substitution Procedure
< 720 hours of QA data	No previous QA data exist	Replaced with PM= 150 mg/Nm3 (conc. standard maximum) and flow = maximum flow possible.
< 720 hours of QA data	Previous QA data exist	Averages from unit operating hour immediately before immediately after the missing data period.

The standard missing data routines use a tiered approach that takes into account both the PMA and the length of the missing data period. When the number of QA hours is between 720 and 8760, PMA is based on the number of operating hours since certification (calculated as number of QA hours divided by number of operating hours since certification). When the number of QA hours crosses 8760, PMA is based on the number of QA hours in the last 8760 hours. When the PMA is high (≥ 95%) and the missing data period is relatively short (≤ 24 hr), the standard missing data algorithms are nearly identical to the initial missing data routines. However, as the PMA decreases, the substitute data values become increasingly conservative, to ensure that emissions are not under-reported. For example, when the PMA of a CEMS is between 80% and 90%, the substitute data value will be the maximum value observed by looking back through the last 720 hours of historical, quality-assured emissions data. But if the PMA drops below 80%, regardless of the length of the missing data period, the maximum potential concentration must be reported. Table 5 below describes the standard missing data procedure.

Table 6: STANDARD MISSING DATA PROCEDURE	
Condition on QA Data Availability	**PMA Calculation Procedure**
>720 & < 8760 hours of QA data	Calculate PMA based on quality assured data since certification
> 8760 hours of QA data	Calculate PMA based on quality assured data in the last 8760 hours.
Condition on missing data time and PMA	**Substitution Procedure**
If PMA > 85 % & missing data < 12 hours	Average readings from unit operating hour immediately before and immediately after the missing data period.
If PMA > 85 % & missing data > 12 hours	GREATER OF 95th percentile hourly value recorded by a CEMS during the previous 720 quality assured monitor operating hours; OR The average of the hourly CEMS values recorded for the hour before and the hour after the missing data period.
If PMA < 85 %	Substitute for that hour of the missing data period the maximum potential value (SPM = 150 mg/m3 and maximum designed flow possible) for that plant)

APPENDIX B3: ANALYSIS ROUTINES

B3.1 Time-series Checks

A time-series is a series of data points measured at successive points in time at uniform intervals. For this document, the data series of interest is PM mass flow[18], which is measured continuously and reported at 1-minute intervals at the industry site. Time-series checks, then, are procedures for detecting unusual trends and outliers in the data stream over time. The principal difference between time-series and the initial checks in the DVP is that the latter considers each incoming data point independently while the former considers data points in relation to either the full data stream or a neighbourhood of points around it. The main purpose of applying time-series checks is to extract meaningful characteristics of the data stream, thereby providing an additional way for assessing and protecting data quality beyond what can be achieved by the general checks. For the purpose of this document time-series checks will be applied on raw minute-by-minute data and a smoothed dataset in order to detect both individual outliers and also unusual trends in the smoothed data.

B3.1.1 Standard Deviation Check

A common technique for identifying outliers in time-series datasets uses the standard deviation, a statistical measure of the variation in values from the mean. A low standard deviation indicates that data are generally close to the mean, while a high standard deviation indicates data are spread out over a large range of values. Because the great majority of values will fall within two standard deviations of the mean (95% of values if the data are normally distributed), standard deviation can be used to identify emissions values that are very unlikely to have occurred under normal operations alone.

Run at the end of each week, the Standard Deviation Check will flag any value that falls more than two standard deviations away from the daily mean value. The mean (m) and standard deviation (sd) of PM mass flow values will be calculated at the end of the day. The daily mean mass flow is then subtracted from each value to standardize the dataset to a mean of zero. Using these parameters, all standardized values greater than 2*sd will be flagged. In mathematical terms, a PM mass flow data point is flagged as an outlier if:

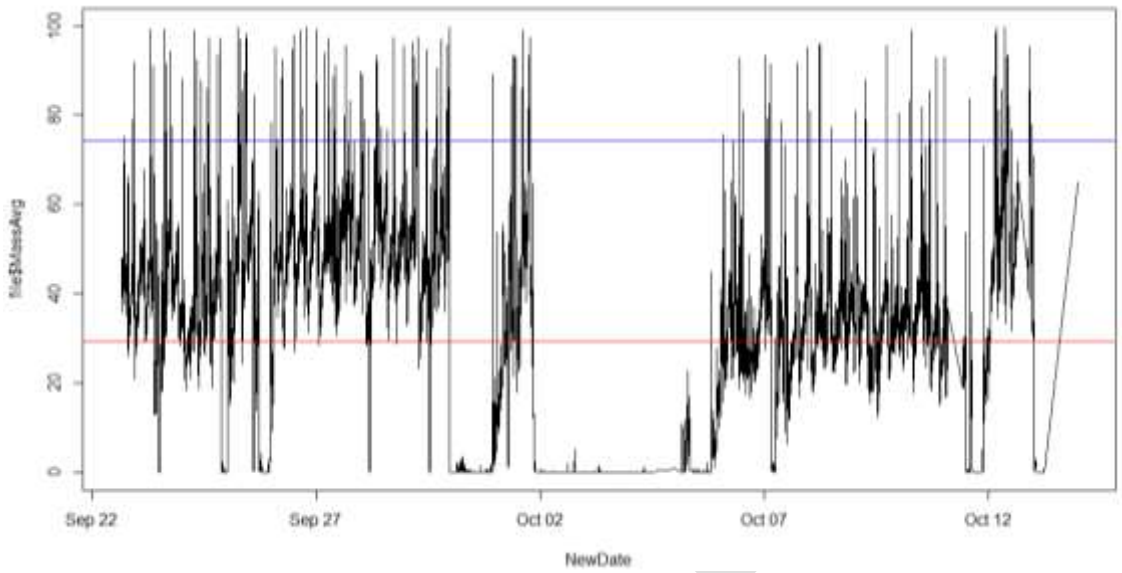
$$|x-\bar{x}|\geq 2\sigma \qquad \bar{x}=\text{daily mean} \qquad \sigma=\text{standard deviation}$$

Because industries' emissions values may not be normally distributed, it is possible that more than 2.5% (i.e. half of 5%) of values will fall above the mean by 2 standard deviations. To ensure that only the most extreme values are invalidated, if more than 2.5% of an industry's values are flagged as outliers in a week, only the largest 2.5% of values will be invalidated.

Figure below shows the standardized PM mass flow values derived from the smoothed dataset. The horizontal lines indicate the standard deviation bands (2σ), and flagged values are those which lie outside.

[18] PM mass flow is the main variable of interest in the time-series because it is the one used for determining total annual PM load, which is the parameter used for determining compliance with permit holdings.

Figure 9: Standard Deviation Check



B3.1.2 Near-Zero Check

The Near-Zero Check is intended to identify strings of exceptional values close to zero that pass initial data validation. A continuous period of emissions measurements fluctuating near zero is an especially important case to flag for two reasons. First, it is an indicator of plant operation status (e.g. if the plant is switched off). Second, it is an indicator of device status (e.g. if the device has been switched off or removed from the stack). Thus, it is important to flag emissions near zero in this case, either to confirm plant/device status, or to mark a suspicious phenomenon where the plant and device are operating normally but emissions are fluctuating around zero (this can indicate tampering).

The near-zero check will flag values that are between 0 and 2% (range is subject to variation) of each device's maximum calibrated value continuously. In mathematical terms, the calculation is performed as follows:

$$.02c_i - x_{raw} > 0$$

c_i = device calibrated maximum value

x_{raw} = raw emissions value

i = industry $1, 2, 3, \ldots, n-1, n$

According to the above procedure, any individual point near zero will be flagged. However, it is a continuous period of near zero values which is exception. It is at the discretion of SPCB regulators to monitor data flagged by this check and determine the course of action based on the length of the near-zero fluctuations. If a regulator sees that three successive values are flagged for being near zero action may not be taken, but if 30 minutes of emissions data is near zero then they would investigate further by checking the device mode and plant status during the corresponding interval.

B3.1.3 Step Change Check

Step change detection refers to the process of finding abrupt changes in the mean level of a time-series. In the case of PM mass flow data, under normal circumstances, a sharp decrease in mean readings can occur

when prime parameters like air pollution control devices are switched on or off (or additional ones are installed) or industries switch to cleaner fuel. In the latter case, industries are required to re-calibrate the CEMS device. If such a step decrement is detected, but no recalibration is documented, the readings following the decrement will be flagged since the data may not appropriately represent actual values given the changes in conditions.

Similarly, a sharp increment could be representative of failures of the air pollution control devices, a bad quality fuel batch, or even a production rise. This situation is of greater concern for the industrial unit itself, but taking into account problems associated with the data consistency, it will be flagged as a concern by the data validation software and left to the discretion of SPCB to decide which action to take.

Because the PM time series data is corrupted by noise, this makes the problem challenging because step changes may become hidden. Although noise is significantly reduced by using a moving average filter, small step changes in the raw data may still not be visibly reflected in the smoothed data trend (see Appendix A6.2 for details on the moving average). For this reason, we employ the CUSUM (cumulative sum of deviations) method—the most common and well-known procedure for step-change detection. The steps for performing relevant calculations are as follows:

1) Calculate the deviation of each raw data point from the corresponding moving average value

$$\text{Deviation}=X_{raw}-X_{MA}$$

2) Generate and plot another data series i.e. cumulative sum of the deviations calculated in step (1).

$$\text{Cumulative sum}=\sum_{i}^{m} x(raw)-x_{MA}$$

If the PM mass flow measurement process remains in control, the cumulative sum will randomly vary around zero. However, if raw values begin to deviate away from the corresponding moving average in the positive direction, then an upward trend will begin to quickly develop in the cumulative sum. A step change is detected when the cumulate sum of deviations begins to continuously increase or decrease. When this occurs, corresponding raw data will be flagged. At SPCB discretion, if the trend increases past a threshold period, data will be invalidated and a follow-up may be required on the quality of the industry CEMS instrument.

B3.2 Data Smoothing

The raw data stream—consisting of 1-minute averages—is characterized by a wildly fluctuating and choppy pattern, primarily due to random noise from the CEMS device. Since it is difficult to detect trends in such noisy data, this protocol employs the use of moving averages to smooth the data before final data presentation and generating weekly validation reports. Also, the step-change check is based on the moving average.

A moving average is a procedure to "smooth out" short-term fluctuations and highlight long-term trends in a time-series dataset. This is done by generating a parallel dataset where each data point is the average of a corresponding subset of a larger set of data points. For the purpose of the pilot ETS, a one-hour single point moving average will be used for data smoothing. This window is chosen because iso-kinetic sampling for CEMS calibration is done over 30-60 minutes, and considering this time window, an individual data point is a reflection of the data pattern over 45 minutes (as done by moving average). The dataset of moving averages can be calculated using equation 1.

$$\big(x(t=45)\big)=\frac{1}{45}*\sum_{k=1}^{45} x(t)$$

APPENDIX C: ETS BRIDGE INTERFACE

C1. Interface Overview

The ETS Bridge Software has to collect and send data from the CEMS device to the SPCB and CPCB server. Hence it is imperative that there is a standard interface for the communication between the CEMS Software and the ETS Bridge Software. All CEMS vendors should ensure that their software provides the required data as per the ETS Bridge interface.

C2. Interface Details

The ETS Bridge Software has a '**File based Interface**'. This means that a file in the prescribed format is generated by the CEMS Software and read by the ETS Bridge Software. This is to ensure that all vendors can easily incorporate any required changes in their software to comply with this interface.

C2.1. Interface Configuration

The following items need to be configured for the interface between the CEMS Vendor Software and ETS Bridge Software, i.e. both of them must refer to the same file. This configuration file will be stored in the same folder as the ETS Bridge, will be named ETSBridge.ini, and will contain following information:

1. Folder where data files will be generated by CEMS vendor software
2. Archive folder where data files will be moved after a file is completely processed by ETS Bridge Software
3. Averaging interval in seconds (which is 60 seconds, as of now)

The format of this configuration file will be like an .ini file with a single section called [Config] and name=value pairs under it:

Format	Example
[Config] DataFolder=<data folder path> ArchiveFolder=<archive folder path> AvgInterval=<averaging interval in seconds>	[Config] DataFolder=C:\\ETS\data ArchiveFolder=C:\\ETS\archive AvgInterval=60

CEMS Software should create data files in the ETS data folder that has to be "C:\\ETS\data".

C2.2. Data File Naming and Creation

A separate data file will be created for each date for each CEMS device connected to the DAS. The data file will be stored in the folder specified in interface configuration file (.ini file).

Format	Example
<Creation date>_<CompID given in configuration data section>.csv	The data file for stack 001 at Sri Ramesh Chemicals in Jalna (having unique code 21908) created on 14th Dec 2012
<yyyymmdd>_<CompID>.csv	**20121214_ MHJLN219080111.csv**

Whenever the CEMS Vendor Software runs, it will look for the data file with the current date for each instrument in the ETS data folder. If found, it will open and append new data to this file, otherwise it will create a new file. It will also create a new file for each instrument when the date changes. The CEMS Vendor Software should always open the file in such a way that ETS Bridge Software is able to open same file simultaneously and read data from it.
[Note: The new day starts at 00:00 hours.]

C2.3. File Format

Each file will have a header row, calibration data row, multiple number of data rows (ideally 1440 rows: 1 per minute) and a footer row.

Section	Format	Example
Header	The **1st** row of the file will always be a header row.	
	<ToP,ToC,CompID>	1,1,**MHJLN219080111**
Calibration data row	The **2nd** row of the file will contain the calibration data.	
	<CF-m,CF-c,Pres_Smpl, Temp_Smpl>	5.1,0.024,100.324,310.1
Data Rows	Subsequent rows in the file will contain measurement data Note: – Comma is used as a delimiter between parameters – No spaces should be used at any point in the file – If a parameter is not applicable for the given instrument, write NA in place of value – If there is an error in the instrument / DAS while reading a parameter, write 'err' in place of value – Each row must end with a CRLF (carriage return/line feed) character	
	<DTS,PM_Uncal,PM_Cal,Velocity,N_Flow,PM_Mass, Pres,Temp,PM_DM,PM_PS,PM_Alrm,PM_M_Alrm,F_DM,F_PS,F_Alrm,F_M_Alrm>	20121214153000,10.1,55.6,NA,NA,55.6, 99.21,150.2,0,0,0,0,0,0,0,0<CRLF>
Footer	Footer row indicates that CEMS vendor software has completed writing this file and the next data set will be written in a new file. This is always the last row in the file.	
	Eof	**Eof**

For the file, first 2 rows won't be repeated. From 3rd row onwards, a row of data will be added every minute.
Some examples of valid and invalid data rows –

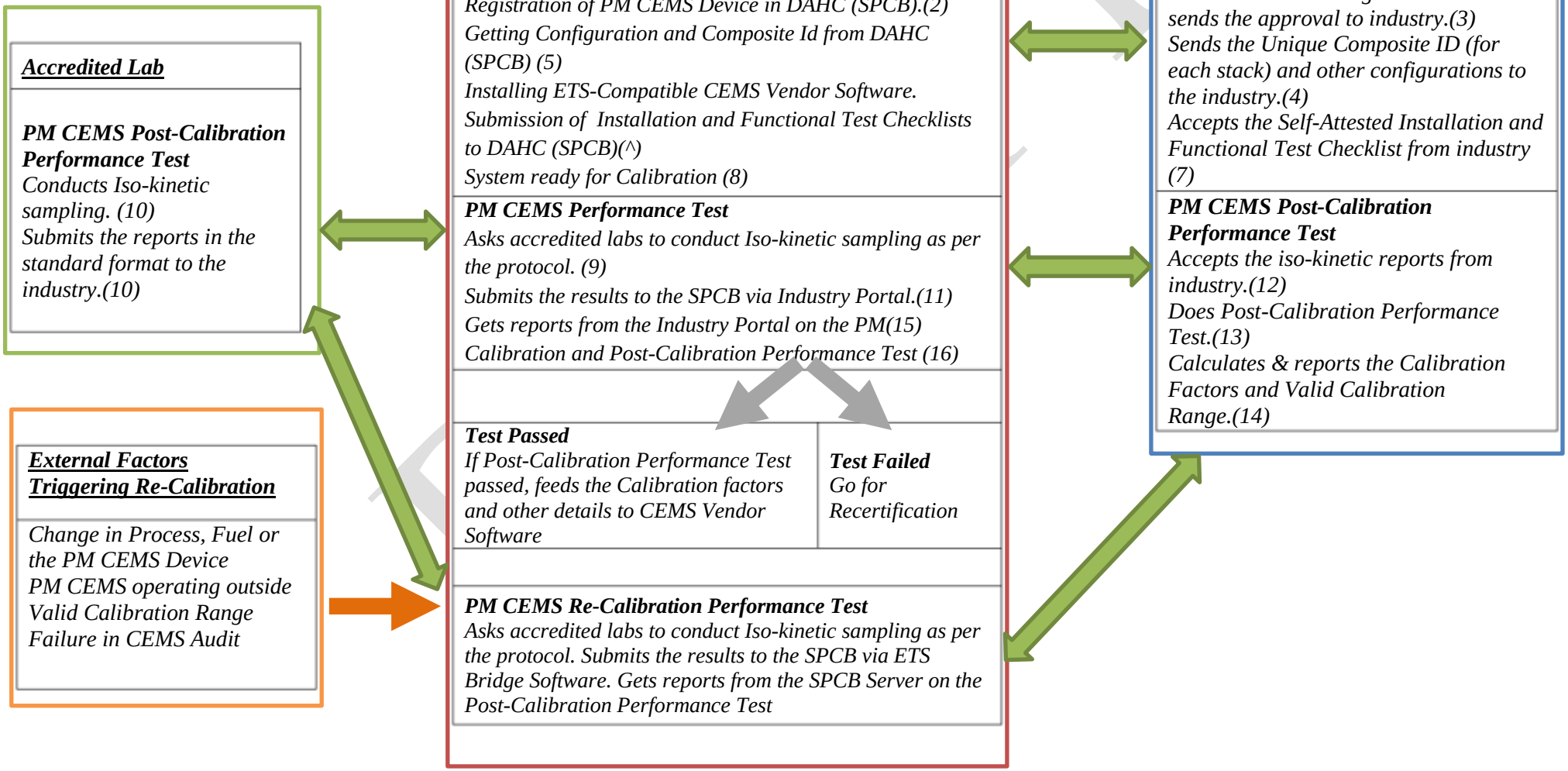
Data	Result
20121214153000,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,0,0,NA,NA <CRLF>	Valid
20121214223100,9.0,50.1,NA,NA,50.1,97.122,err,0,0,0,0,0,0,0,0<CRLF>	**Valid,** Error in reading temperature
20121214153000,10.1,55.6,NA,,55.6, 99.21,150.2,0,0,0,0,0,0,0,0<CRLF>	Invalid - Space after comma - Empty instead of NA
20121214153000,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,0,0,0,0,0,20121214153100,5,5.1,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,0,0,0,0,0,0<CRLF>	Invalid - No CRLF after 1st row and hence 2nd row (red) in continuation with 1st row
20121214153000,5,5.1,10.1,55.6,NA,NA,55.6,99.21,150.2,2,0,0,0,0,0,0,0<CRLF>	**Invalid,** PM CEMS device in Maintenance Mode
20121214153000,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,0,0,0,1,0,0<CRLF>	**Invalid,** Flow-meter power is off
20121214153000,5,5.1,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,1,0,0,0,0,0<CRLF>	**Invalid**, PM Alarm 'On' i.e Fault mode of device or data invalid indicated by device (though device is in normal operation mode)
20121214153000,5,5.1,10.1,55.6,NA,NA,55.6,99.21,150.2,0,0,0,1,0,0,0,0<CRLF>	Valid, Maintenance Alarms from PM CEMS Device On. Device needs maintenance.

C3. Data Exchange Process

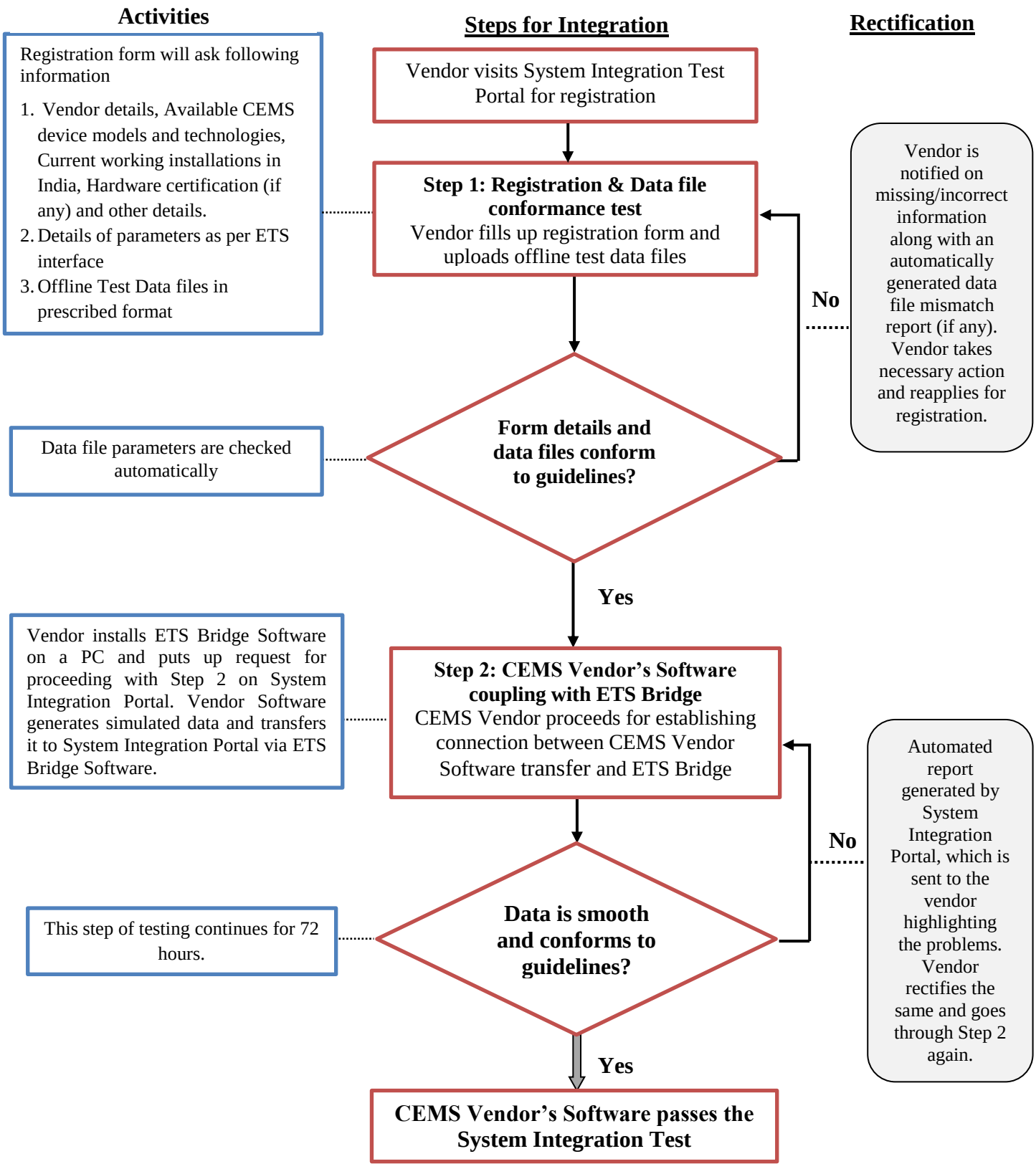
1. Whenever the CEMS Vendor Software starts, it will look for the file with the current date for each instrument connected to it. If found, it will open this file, otherwise it will create a new file. It will also create a new file for each instrument when the date changes.
2. When the ETS Bridge Software starts, it will load information about the number and type of CEMS instruments connected to the DAS via the CEMS vendor software. This information will be obtained from the DAHC server.
3. Then ETS Bridge Software will look for files from each instrument and open these files for reading.
4. It will read all new rows already existing in the file since last reading.
5. Then it will wait for new data rows to arrive in the file. As soon as a new row is appended to the file, the ETS Bridge Software will read that row.
6. ETS Bridge Software will add one more date and time stamp at the beginning of each data row. This time stamp specifies when ETS Bridge Software read the row. If difference between the ETS Bridge Software time stamp and CEMS Software time stamp is more than a threshold, the server will add a 'late' flag for this row. This indicates that the data was not received by the ETS Bridge Software in real time.
7. ETS Bridge Software will send the data to the DAHC servers at SPCB and CPCB, every minute as soon as it is read from file. In case data cannot be sent to one or both servers, ETS Bridge will keep the data locally and send to server(s) as soon as possible. The bridge will keep track of rows sent to each server and ensure that there is neither data loss nor data duplication.
8. ETS Bridge Software will also store all data locally in a different folder in encrypted format.
9. ETS Bridge Software will also generate an error file logging all errors encountered.
10. When the date changes, the CEMS Software will write a footer row in the previous day's file and close that file. After this, the CEMS Software cannot write any new data in this file.
11. After encountering a footer row in a file, the ETS Bridge Software will close that file and move it to an archive folder specified in the configuration file.

If the CEMS Software cannot read data in real time from an instrument due to whatever reason, it MUST read this data later and write it to the appropriate files before a footer row is written. This is to ensure that there is no data loss. In this case, all this data will be treated as 'late' by the server. However, this must be done before the CEMS Software starts writing real time data to the current file. Thus, the CEMS Software cannot keep an older file open while writing data in a newer file.

APPENDIX D: CEMS OPERATIONS FLOWCHART



APPENDIX E: SYSTEM INTEGRATION TEST FLOW CHART



This is a blank page