6.2 AUTOMATED QUALITY CONTROL OF ATMOSPHERIC RADIATION MEASUREMENT PROGRAM (ARM) DATA FROM THE SOUTHERN GREAT PLAINS (SGP), NORTH SLOPE OF ALASKA (NSA), AND TROPICAL WESTERN PACIFIC (TWP) CLOUD AND RADIATION TESTBED (CART) SITES

Karen L. Sonntag*, Andrew R. Dean, Chad M. Shafer, Randy A. Peppler
University of Oklahoma, Norman, Oklahoma

Chad P. Bahrmann
Pennsylvania State University, State College, Pennsylvania

1. INTRODUCTION

One of the primary goals of the ARM Program is to provide data streams of reasonable quality for scientific research. Traditionally, data quality issues have been addressed within ARM by several groups, including instrument mentors, site scientists, site operators, value-added product scientists, and ARM Science Team members. Maintaining data quality for a program with the size and complexity of ARM is a significant challenge.

The ARM Data Quality Office (DQO) was established in July 2000 to help coordinate data quality efforts within the ARM Program. The DQO is responsible for ensuring that quality assurance results are communicated to data users as well as to the ARM Operations and Engineering Groups to facilitate improved instrument performance. To address these responsibilities, the DQO developed a Data Quality Processing Algorithm (DQPA) that automatically performs range checks as well as other user-defined checks on the data. The results of these checks are displayed in tables, which are available on a series of web pages (http://dq.arm.gov/). These tables have options to view static plots and to create plots interactively using a tool called NCVweb (Moore and Bottone, 2003). This comprehensive web package is known as the Data Quality Heath and Status (DQ HandS) system. It provides access to complimentary databases within ARM that allow analysts to access metadata regarding instrument maintenance reports, calibration information, and instrument logs, and allows for problem reporting both to site operators and data users. In continued collaboration, instrument mentors, site scientists, and DQO scientists in particular use this package to inspect and assess data quality on a daily and weekly basis.

* *Corresponding author address:* Karen L. Sonntag, CIMMS, University of Oklahoma, 100 E. Boyd Street, Room 1110, Norman, OK 73019; e-mail: ksonntag@ou.edu

2. DATA QUALITY INSPECTION TOOLS

2.1 *Data Quality Processing Algorithm (DQPA)*

The DQPA includes a set of algorithms of varying sophistication that determine the percentage of data falling within specified quality tolerances. The quality tolerances include simple measures such as minimum, maximum, and delta checks, and in some cases include higher order checks such as comparing measurements to model output, objective analysis, and cross-instrument comparisons. The results of these automated checks represent a first line of defense for our quality control efforts, and are displayed in color-coded tables by day and by hour (see Figure 1 for an example of a color table by hour). Green is indicated for an hour when all values for a measurement passed the automated checks. Yellow appears when 75-nearly 100% of the values pass the automated checks. Red appears when more than 25% of the values fail a check(s). Black is indicated when data values are physically missing from the data file, and gray is shown when a missing value code such as –9999 is indicated for a data value (may indicated missing data or a special circumstance).

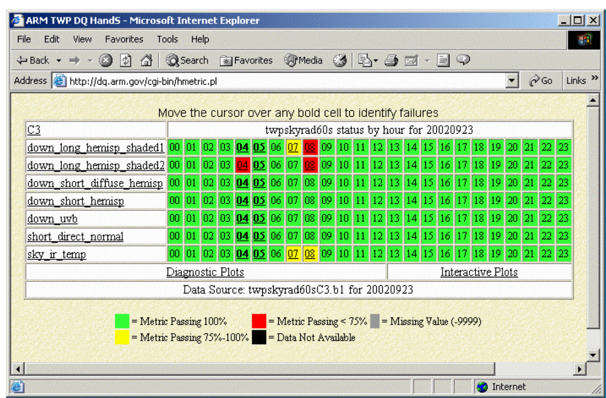


Figure 1. Color-coded table of automated quality control check results for TWP-Darwin downwelling radiation measurements on 23 September 2002.

When performing a "mouse-over" of a red or yellow color box (representing one hour for a particular measurement), a pop-up appears indicating which check was violated and what percentage of data violated the check (Figure 2). A mouse-over of the measurement name will explain that measurement in plain English (Figure 3).

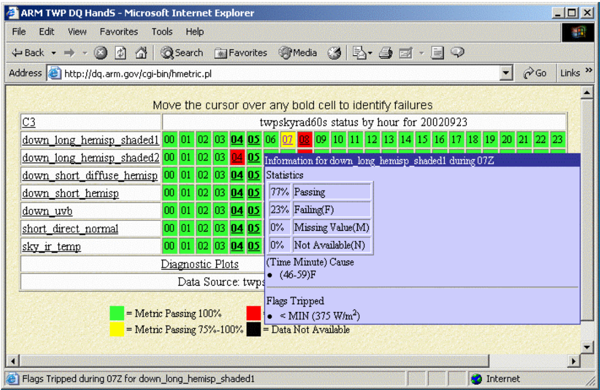


Figure 2. Same as Figure 1, but showing a pop-up of the Hour 7 yellow box for down_long_hemisp_shaded1 (downwelling longwave hemispheric irradiance – shaded pyranometer). The minimum flag (375 Wm^{-2}) was violated by 23% of the values within that hour, specifically during Minutes 46-59.

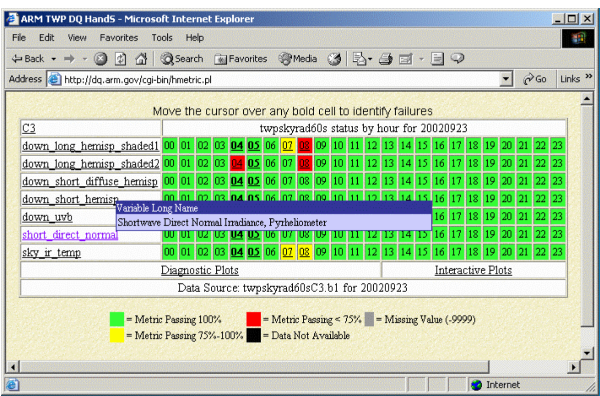


Figure 3. Same as Figures 1 and 2, but showing a pop-up of the description of “short_direct_normal” (shortwave direct normal irradiance, pyrheliometer).

For most instruments, simple min/max checks are not adequate for assessing data quality. Sophisticated checks are currently in development to provide an additional layer of data quality beyond the simple range tests. These include internal consistency checks for a given instrument and external consistency checks between two or more instruments with comparable measurements.

One example of an internal consistency check is the SERI QC flag system for the SIRS (Solar Infrared Radiation Station) platform developed by the National Renewable Energy Laboratory (NREL, 1993). The SERI QC system examines different components of solar and terrestrial radiation to see if they are in good agreement with each other, in the context of the expected radiation budget.

The surface objective analysis (OBAN) system currently under development is an example of an external consistency check for the various ARM platforms that measure surface meteorology variables such as temperature, pressure, and relative humidity. If an observation deviates by a certain amount from the background field, it is flagged as suspicious. The background field is calculated using a Barnes scheme that utilizes both ARM data and surface observations available from the National Weather Service. In addition to data quality flags, it is possible to generate plots based on the OBAN output (Figure 4).

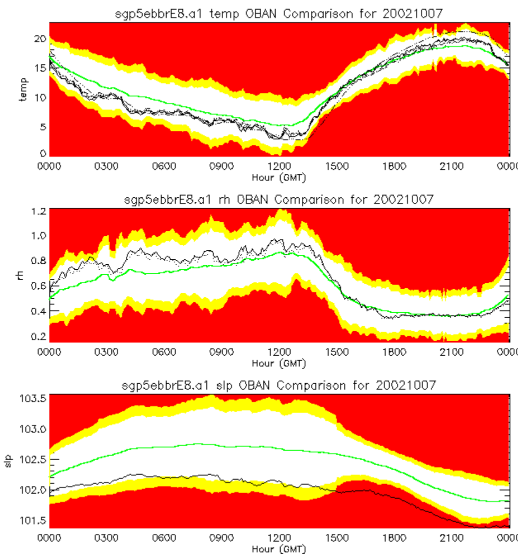


Figure 4: Example of an objective analysis plot for an EBBR (Energy Balance Bowen Ratio) station. The green line represents the expected value from the OBAN while the black line represents the observation. If the black line strays into the red area, it is deemed to have failed the OBAN test. Temperature, RH, and sea-level pressure are shown on this plot.

These more advanced automated checks should prove very useful, but in order to ultimately determine data quality, it is necessary to look at plots of the data itself, which are described below.

2.2 *Data Quality Diagnostic Plots*

Data quality diagnostic plots (Figure 5), produced automatically within DQ HandS when the data reach our file server, provide a tool for further inspecting problems identified initially by the automated checks. These graphical displays plot key geophysical parameters and cross-instrument comparisons. The plots are a necessary supplement to the automated flags, allowing the user to further identify problems indicated by the flags and other problems through visual inspection.

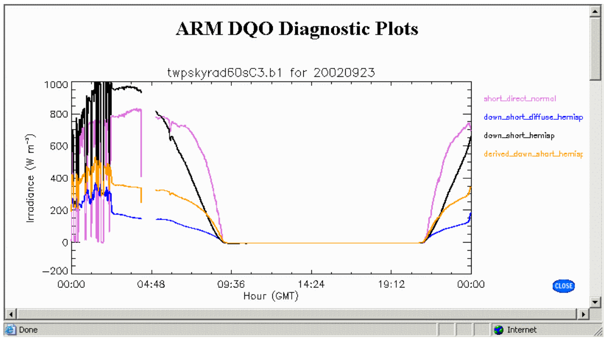


Figure 5. Plots of downwelling radiation components for TWP-Darwin on 23 September 2002 (direct normal - purple, diffuse - blue, global - black, and a derived global - gold).

2.3 *Interactive Plotting*

To manipulate particular data of interest, an interactive plotting tool called NCVweb was developed for use in DQ HandS by Mission Research Corporation. It allows a data quality analyst to visually pinpoint trouble spots to allow more informed problem reporting. It allows for scale and variable changes. An example of a NCVweb plot is shown in Figure 6. Please also see P1.45 in the 19th Conference on IIPS (Moore and Bottone, 2003)

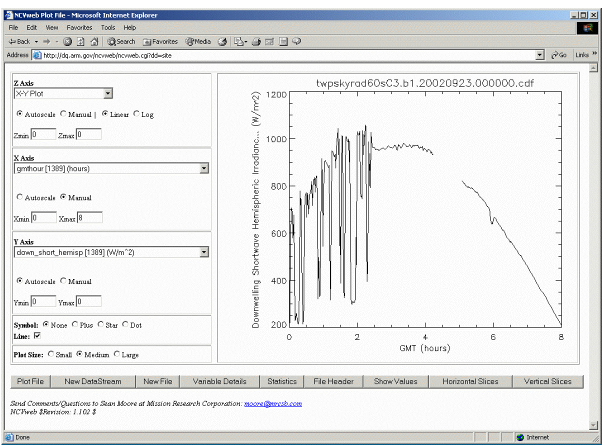
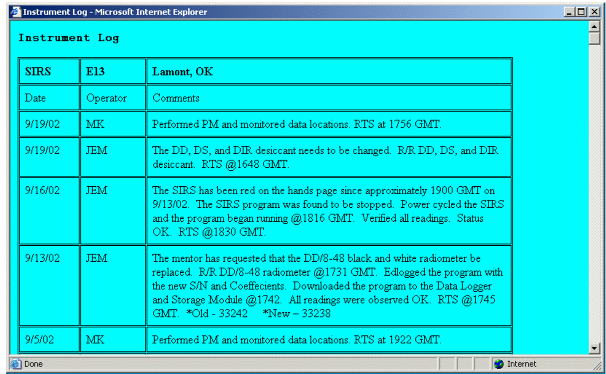


Figure 6. NCVweb plot for downwelling shortwave hemispheric irradiance during Hours 0-8 at TWP-Darwin on 23 September 2002.

2.4 *Supporting Information*

The ARM Program saves supporting metadata that is useful for data quality assessment activities. Such information includes instrument maintenance data, engineering logs, and calibration results. This information is vital for assessing data quality because it alerts the DQ analyst to activities that have been performed on the instruments by site operators or instrument mentors that may affect data quality. An example of such information, as obtained from the SGP Operations Management Information System (OMIS), is shown in Figure 7.



Instrument Log

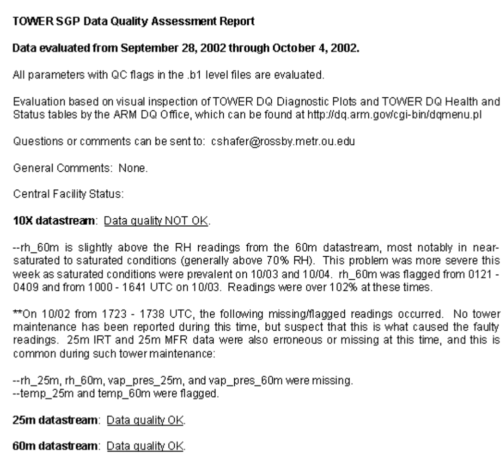
SIRS	E13	Lamont, OK
Date	Operator	Comments
9/19/02	MK	Performed PM and monitored data locations. RTS at 1756 GMT.
9/19/02	JEM	The DD, DS, and DIR desiccant needs to be changed. R/R DD, DS, and DIR desiccant. RTS @1648 GMT.
9/16/02	JEM	The SIRS has been red on the hands page since approximately 1900 GMT on 9/13/02. The SIRS program was found to be stopped. Power cycled the SIRS and the program began running @1816 GMT. Verified all readings. Status OK. RTS @1830 GMT.
9/13/02	JEM	The mentor has requested that the DD/8-48 black and white radiometer be replaced. R/R DD/8-48 radiometer @1731 GMT. Edlogged the program with the new S/N and Coeffecients. Downloaded the program to the Data Logger and Storage Module @1742. All readings were observed OK. RTS @1745 GMT. *Old - 33242 *New – 33238
9/5/02	MK	Performed PM and monitored data locations. RTS at 1922 GMT.

Figure 7. Instrument maintenance log for the radiometer suite at the SGP-Lamont facility during September 2002.

3. DATA QUALITY REPORTING TOOLS

3.1 *Data Quality Assessment Reports*

The first step in documenting data quality is the filing of a Data Quality Assessment Report (DQAR). The DQAR is a weekly, monthly, or periodic summary of instrument performance issued by the DQO, instrument mentors, or site scientists. These reports provide first-level documentation used in subsequent reports to site operators and/or data users. An example DQAR is shown in Figure 8.



TOWER SGP Data Quality Assessment Report

Data evaluated from September 28, 2002 through October 4, 2002.

All parameters with QC flags in the .b1 level files are evaluated.

Evaluation based on visual inspection of TOWER DQ Diagnostic Plots and TOWER DQ Health and Status tables by the ARM DQ Office, which can be found at http://dq.arm.gov/cgi-bin/dqmenu.pl

Questions or comments can be sent to: cshafer@rossby.metr.ou.edu

General Comments: None.

Central Facility Status:

10X datastream: Data quality NOT OK.

--rh_60m is slightly above the RH readings from the 60m datastream, most notably in near-saturated to saturated conditions (generally above 70% RH). This problem was more severe this week as saturated conditions were prevalent on 10/03 and 10/04. rh_60m was flagged from 0121 - 0409 and from 1000 - 1641 UTC on 10/03. Readings were over 102% at these times.

**On 10/02 from 1723 - 1738 UTC, the following missing/flagged readings occurred. No tower maintenance has been reported during this time, but suspect that this is what caused the faulty readings. 25m IRT and 25m MFR data were also erroneous or missing at this time, and this is common during such tower maintenance:

--rh_25m, rh_60m, vap_pres_25m, and vap_pres_60m were missing.
--temp_25m and temp_60m were flagged.

25m datastream: Data quality OK.

60m datastream: Data quality OK.

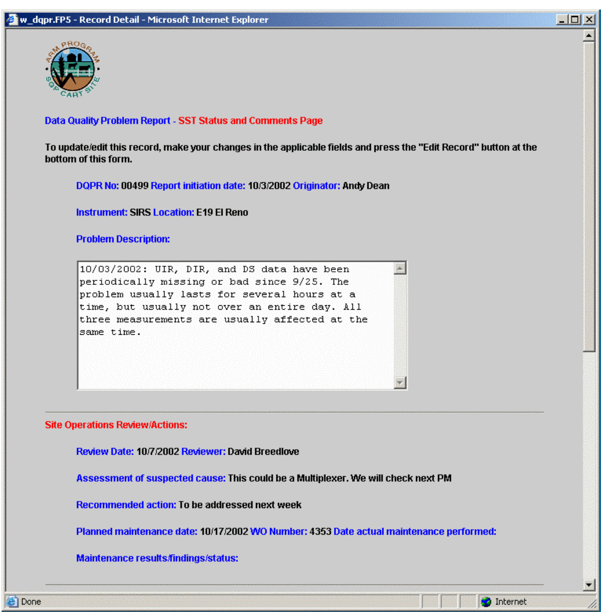
Figure 8. DQAR for 60-m tower at SGP-Lamont facility.

3.2 *Data Quality Problem Reports*

The second step in documenting data quality is to report instrument related problems to ARM site operations through submission of a web-based Data Quality Problem Report (DQPR). The DQPR database in OMIS was developed as a tool to inform site operations and instrument mentors of instrument problems and to allow tracking of the problem status until it is resolved. DQPR tracking involves the originator, the site scientist, the site operator, and the particular instrument mentor. All have the opportunity to comment on the problem and suggest a possible fix. The site scientist manages the DQPR process. The DQPR is reserved for problems that are deemed solvable within the scope of site operator abilities. A sample DQPR is shown in Figure 9.

For catastrophic problems or problems beyond the scope of site operator intervention, an ARM Problem Identification Form (PIF) is used to alert the ARM Problem Review Board, which meets weekly via teleconference. The PRB goes over all open problems and assigns them to someone for solution. Once a problem has been solved, a Corrective Action Report (CAR) is filed.

The end result of both the DQPR and PIF processes is a Data Quality Report (DQR), described in the next section.



w_dqpr.FP5 - Record Detail - Microsoft Internet Explorer

Data Quality Problem Report - SST Status and Comments Page

To update/edit this record, make your changes in the applicable fields and press the "Edit Record" button at the bottom of this form.

DQPR No: 00499 Report initiation date: 10/3/2002 Originator: Andy Dean

Instrument: SIRS Location: E19 El Reno

Problem Description:

10/03/2002: UIR, DIR, and DS data have been periodically missing or bad since 9/25. The problem usually lasts for several hours at a time, but usually not over an entire day. All three measurements are usually affected at the same time.

Site Operations Review/Actions:

Review Date: 10/7/2002 Reviewer: David Breedlove

Assessment of suspected cause: This could be a Multiplexer. We will check next PM

Recommended action: To be addressed next week

Planned maintenance date: 10/17/2002 WO Number: 4353 Date actual maintenance performed:

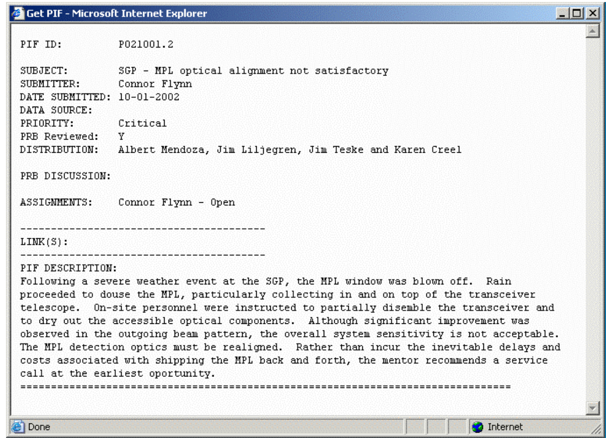
Maintenance results/findings/status:

Done Internet

Figure 9. DQPR for SGP-El Reno radiometer issues. In this case, the DQPR was opened by the DQO analyst who inspected the data. The SGP site maintenance manager has responded with a possible cause. He has also scheduled a maintenance visit with a work order. Once the results of this maintenance visit are known, they will be entered into the DQPR form and if the problem was solved, the DQPR will be closed.

3.3 *ARM Data Quality Reports*

The final step in documenting data quality is an ARM Data Quality Report (DQR). The DQR (Figure 10) is used to inform the data user about the quality of ARM data that she/he has ordered from the ARM Data Archive at Oak Ridge National Laboratory. All information from DQARs and DQPRs is contained within the DQR, which is delivered to data users along with their data file.



Get PIF - Microsoft Internet Explorer

PIF ID: P021001.2

SUBJECT: SGP - MPL optical alignment not satisfactory
SUBMITTER: Connor Flynn
DATE SUBMITTED: 10-01-2002
DATA SOURCE:
PRIORITY: Critical
PRB Reviewed: Y
DISTRIBUTION: Albert Mendoza, Jim Liljegren, Jim Teske and Karen Creel

PRB DISCUSSION:

ASSIGNMENTS: Connor Flynn - Open

LINK(S):

PIF DESCRIPTION:
Following a severe weather event at the SGP, the MPL window was blown off. Rain proceeded to douse the MPL, particularly collecting in and on top of the transceiver telescope. On-site personnel were instructed to partially disemble the transceiver and to dry out the accessible optical components. Although significant improvement was observed in the outgoing beam pattern, the overall system sensitivity is not acceptable. The MPL detection optics must be realigned. Rather than incur the inevitable delays and costs associated with shipping the MPL back and forth, the mentor recommends a service call at the earliest oportunity.

Done Internet

Figure 10. Problem Identification Form regarding a problem with the Micropulse Lidar at SGP-Lamont.

4. FUTURE ACTIVITIES

The DQ HandS system is providing an efficient way to diagnose and document DQ problems. This method has improved instrument performance and has led to an overall enhancement in ARM data quality. DQ HandS is now available for all three CART sites.

A next generation of data quality reporting tools is currently under development. At present, the various reporting mechanisms such as DQAR, DQPR, and PIF/CAR/DQR are not related in a database sense, and it would be nice to be able to pass certain parameters from one report to another to minimize entry errors. The Data Quality Reporter is being developed to coordinate all of these activities. It is anticipated to be available in spring 2003. It will be directly assessable from DQ HandS and as a stand-alone. Figure 11 shows the logic of how we anticipate the DQ Reporter to work.

5. REFERENCES

Moore, S.T., and S. Bottone, 2003: NCVweb: An interactive web-based tool for viewing Atmospheric Radiation Measurement (ARM) data. *Preprints, 19th Conference on IIPS*, Amer. Meteor. Soc., 9-13 February, Long Beach, CA., in press.

NREL, 1993: *User's Manual for SERI_QC Software-Assessing the Quality of Solar Radiation Data.* NREL/TP-463-5608. Golden, CO, National Renewable Energy Laboratory.

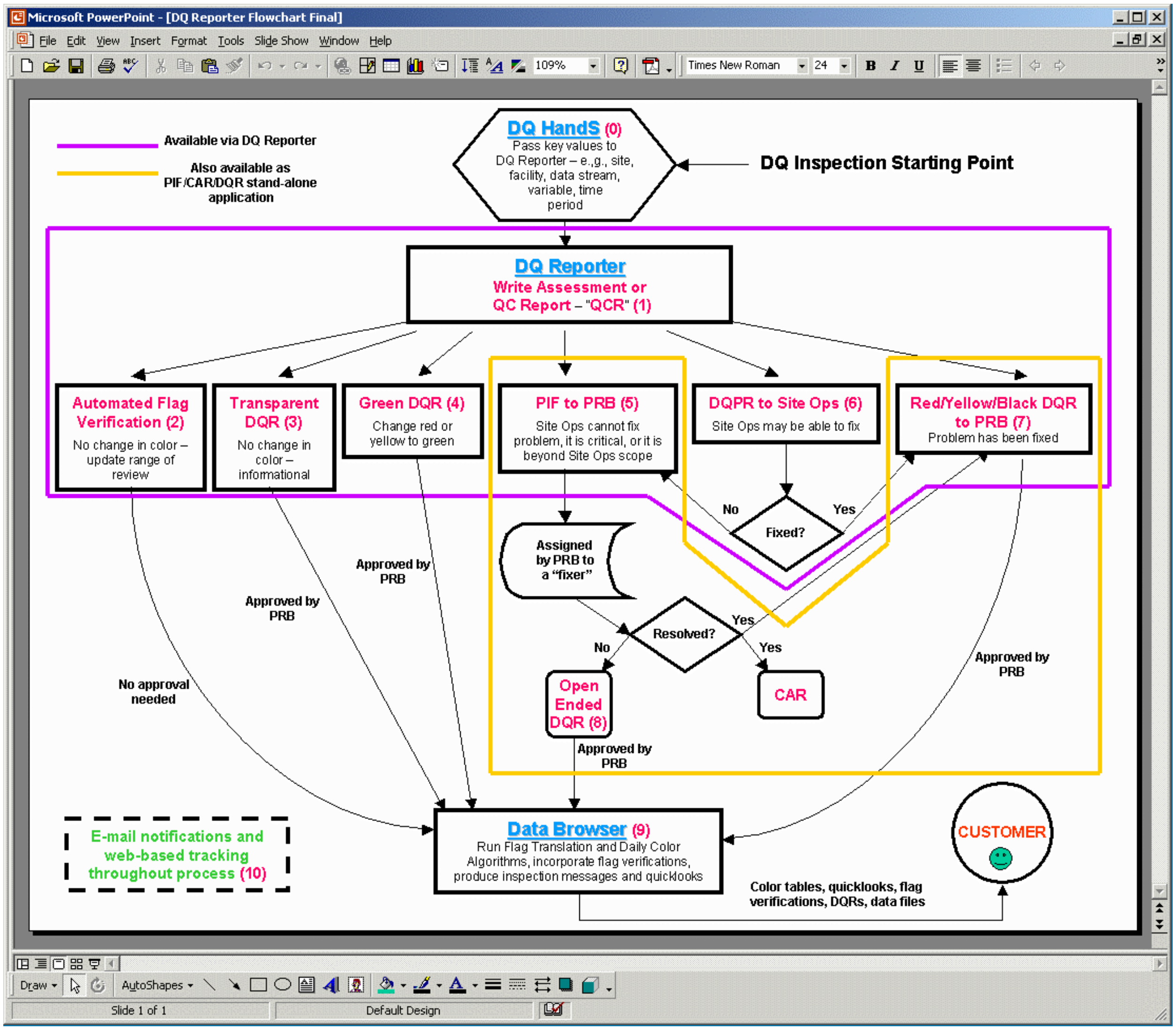


Figure 11. Flow logic for the Data Quality Reporter.