JP1.12 STORING AND ORGANIZING ARM PROGRAM MEASUREMENTS DOCUMENTATION FOR DATA QUALITY PURPOSES

Kenneth E. Kehoe*, Randy A. Peppler, Karen L. Sonntag
CIMMS/Univ. of Oklahoma, Norman, OK

Sean T. Moore
Mission Research and Technical Services,
Santa Barbara, California

1. INTRODUCTION

The ARM Data Quality Office (DQO) inspects on average 5000 measurements per day. To properly accomplish this task, DQO analysts need to have a good understanding of the data collection process and what to expect while analyzing plots and figures. This requires an understanding of many broad meteorological concepts as well as important nuances of individual instruments. Previously, this critical information was stored in numerous locations and not fully accessible by everyone. To correct this gap the DQO has implemented a wiki system allowing those involved in the program access to such information, plus the ability to quickly and easily add new information they discover. This new method of archiving ARM data quality knowledge is organized in an open format viewable from an Internet web browser and easily updatable by any registered user.

2. WIKI THEORY

A wiki is a collaborative platform designed to allow multiple users the ability to edit web pages from any computer. These instant changes are viewable by all users without the need of a central web designer. Wiki pages are directly editable via a web browser, independent of the users operating platform, and permits users to create or edit any page. Some benefits of this method include fixing subtle mistakes, adding new information, updating outdated information and adding new pages to further the documentation.

An example of a wiki platform that has gained large popularity is Wikipedia (www.wikipedia.org). This wiki was created in January 2001, and has grown to over 5 million articles in over 250 languages through voluntary users creating and editing wiki pages (Wikipedia 2006). While the ARM Data Quality Office has no intention of creating documentation that substantial, the same potential for creating and displaying a wealth of data quality information exists.

2.1 ARM Data Quality Office Wiki

There are a number of different wiki collaboration platforms to choose from. The DQO has chosen the TWiki open source platform (TWiki 2006). Figure 1 shows the DQO wiki home page and its initial organization of topics. The TWiki platform stores version histories of each page, including versions of figures and pictures. This is a core principle behind the wiki philosophy, for it allows users to edit any page with a retrievable audit trail. Previous versions of the page or only differences are easily displayed allowing another user to revert to a previous version. In addition to the data quality documentation section of the wiki, each user is given space to create their own personal pages that describe who they are, their interests and their area of expertise.

Upon opening the DQO wiki, each user is prompted to enter their individual user name and password. The login requirement function helps keep the wiki secure, and tracks individual user's changes. This provides wiki administrators the ability to ensure all comments are from appropriately knowledgeable users. There is also the option for a daily e-mail notification of all recently changed pages. This is a good tool to inform users of newly-posted information.

* Corresponding author Address: Kenneth E. Kehoe National Weather Center, Room 2100, 120 David L. Boren Blvd., Norman, OK 73019
e-mail: kkehoe@ou.edu

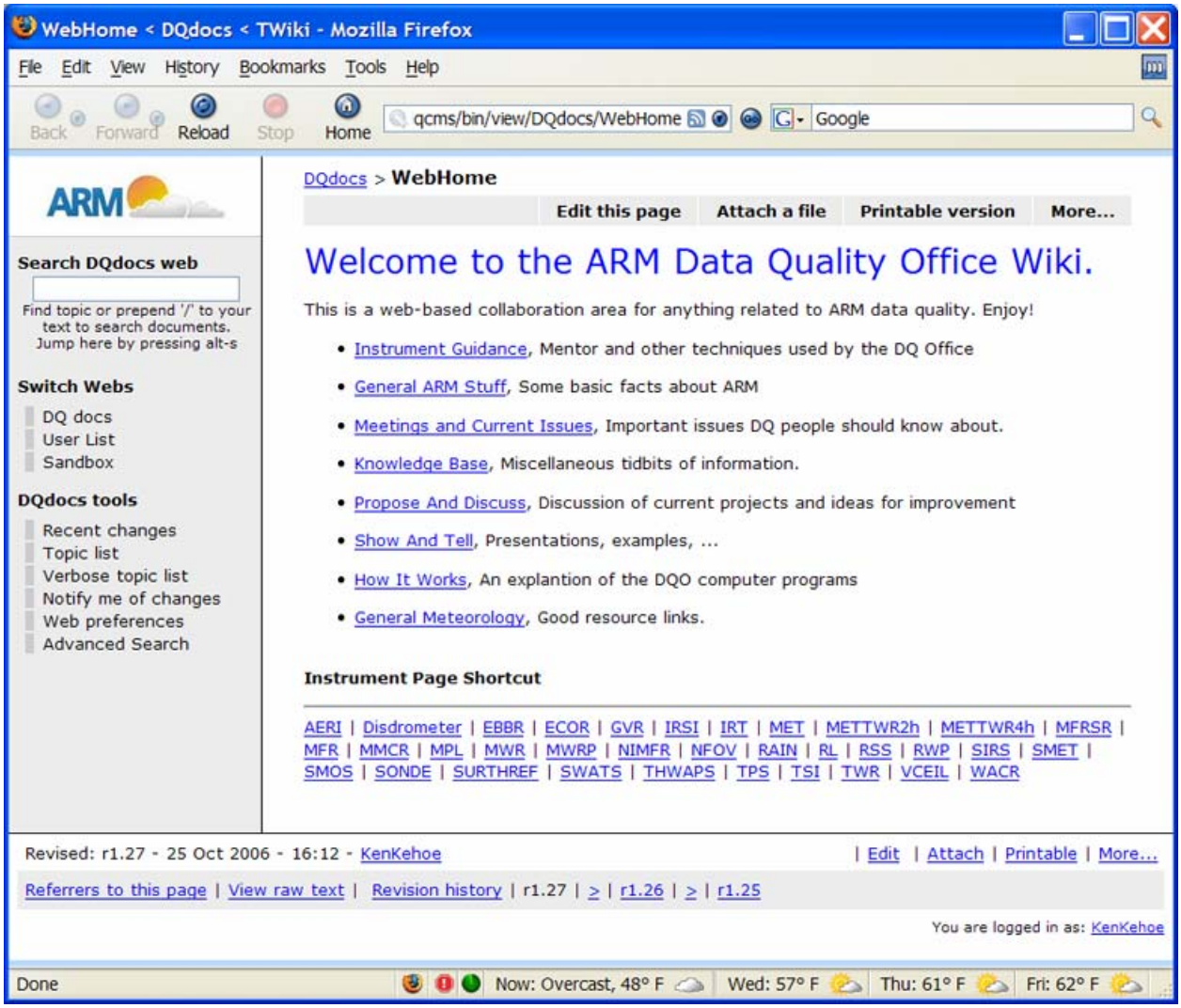

Figure 1 Home page view of ARM Data Quality Office wiki. Notice user's login displayed in lower right corner, and links to view 'Revision History', 'Edit' current page, 'Attach' images, or show the page in a 'Printable' view without the side bar, header or footer section.

Each page of the wiki is organized by a wiki word in CamelCase, where the words are joined without spaces and each word is capitalized (CamelCase 2006). This allows easy creation of new pages, or links to existing wiki pages without additional programming syntax.

Web page development can appear daunting with complicated HTML syntax. To encourage others to use a new documentation system, the major requirement was simplicity. Fortunately the DQO wiki uses simplified syntax allowing a novice to become a proficient editor. One example of the simplicity in creating wiki pages is the use of tables. In HTML, generating complicated tables can become confusing with the requirement of multiple opening and closing tags to set up the table structure. The DQO wiki simplifies the insertion of tables by only requiring one tag (the bar character, |), while still maintaining the possibility of intricate table designs (Figure 2). In addition, all formatting syntax is easily available through a pop up window explaining text formatting rules.

3.DATA QUALITY OFFICE DOCUMENTATION

The DQO wiki is divided into categories pointing users to the general area of interest. Some main topics include instruments, computer systems, DQO specific software, general ARM information, discussions of upcoming meetings and a repository for presentations, papers or

other work. Attachments can encompass any form. A user can also find the appropriate pages by using a search box located in the upper left corner of each page (Figure 1). The user can search for a specific page or a text string within a page. The results of the search are returned in descending order of most-likely matched.

Rain Typical Values

Description	Rain Rate	Precipitation
Drizzle	0.25 mm/hr	0.004 mm (trace)
Light Rain	1.0 mm/hr	0.02 mm
Moderate Rain	4.0 mm/hr	0.05 mm
Heavy Rain	16.0 mm/hr	0.25 mm
Thunderstorm	35.0 mm/hr	0.6 mm
Intense Thunderstorm	100.0 mm/hr	1.6 mm

```
---++%GREEN%Rain Typical Values%ENDCOLOR%
|  *Description*  |  *Rain Rate*  |  *Precipitation*  |
|  Drizzle  |  0.25 mm/hr  |  0.004 mm (trace)  |
|  Light Rain  |  1.0 mm/hr  |  0.02 mm  |
|  Moderate Rain  |  4.0 mm/hr  |  0.05  mm  |
|  Heavy Rain  |  16.0 mm/hr  |  0.25 mm  |
|  Thunderstorm  |  35.0 mm/hr  |  0.6 mm  |
|  Intense Thunderstorm  |  100.0 mm/hr  |  1.6 mm  |
```

Figure 2 An example of the table structure using the DQO wiki to describe typical rain values, with the wiki syntax shown below the table. Users can alphabetically arrange the table by clicking any of the column headings for ascending or descending order.

3.1 Instrument Wiki Pages

An instrument specific page exists for each ARM instrument the DQO monitors. These pages are designed to be used as both a reference for specific pieces of information and as a tool for teaching new analysts quality control inspection techniques.

Each instrument page starts with a dynamically-generated Table of Contents with links to headings on that page. A new link is automatically included when a new section heading is introduced. This allows each page to evolve as users make additions without the need for a structured document. However, each page is designed to have a few required sections including: links to other similar instrument wiki pages, an external link to the official ARM instrument web page, and a general description of the instrument. Because each instrument page is tailored to its specific instrument, the most pertinent information is presented in an orderly manner without additional clutter. This also allows each page to evolve over time as users add/subtract/modify the page whenever they see an opportunity for improvement.

A picture of the instrument is also provided to help the analysts understand the physical arrangements and limitations of the instrument. These images have the ability to clearly describe why an instrument may appear to be malfunctioning, but is actually performing exactly as expected. An example of this is the lack of continuous high latitude Multi Filter Rotating Shadowband Radiometer (MFRSR) data near local solar midnight during continuous daylight seasons. The problem is caused by the physical limitations of the rotating shadowband that would otherwise strike the support arm of the instrument (Figure 3). This example of miscellaneous information shows where knowledge retention is critical; in this case, determining the already-known reason for missing data. A wiki is adept at storing these pieces of information, for as pages evolve through addition and subtraction of content, the best arrangement of this information is sorted out.

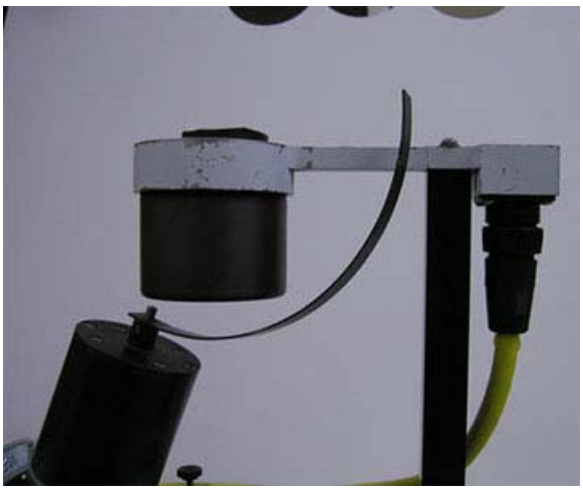

Figure 3 This image clearly shows why the MFRSR reports missing data during days without sunset at high latitudes. The shadowband would strike the support arm of the instrument if allowed to continue rotating.

An effective way to spot problems is to understand the history of an instrument. The ARM Program has been operating many instruments for over 10 years. This provides a knowledge base of how an instrument is expected to function (Figure 4) and any issues with the instrument. When the DQO initially set

up the wiki, basic documentation and guidance was requested from each instrument mentor. This information was transformed into a wiki page including specific examples of past problems. By having a repository of examples and the accompanying explanation, analysts are able to quickly scan the plots to see if the suspect data is exhibiting a known issue or a new problem. This greatly decreases time spent determining if a problem exists and provides suggestions for resolution.

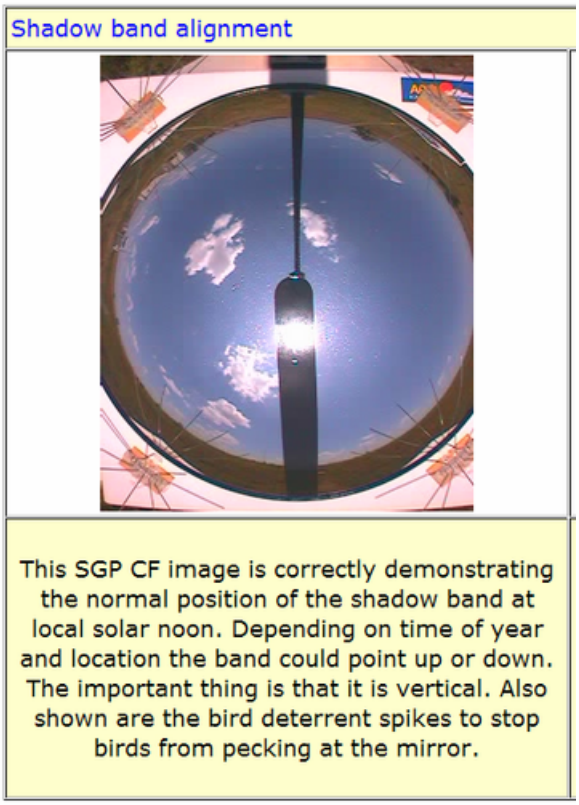

Figure 4 This is an example of using the wiki to demonstrate the correct position of the ARM Total Sky Imager (TSI) shadowband at local solar noon, an optimal time to check alignment of the system.

3.2 Computer Processing Method

With multiple computer programmers involved on multiple projects, sharing information among co-workers is an important priority. The ARM DQO processing system has grown to include many different programs and scripts to manage the large amount of processing and files. Wiki pages help describe the process with flow charts, tables of DQO specific programs, organizational diagrams, and areas of scratch space to share current ideas between developers. The discussion space has become an important area to ensure everyone is properly informed. E-mail is a very fast and efficient method of sharing ideas, but creates the danger of inadvertently excluding contributors. The discussion area allows new and old ideas to be listed for all to read and comment.

4. SUMMARY

Sharing information produced by ARM instrument mentors with data quality analysts has been crucial for the success of the ARM Data Quality Office. Until recently, accessing this information was sometimes difficult and had the possibility of disappearing as staff departed for other positions. The implementation of a wiki system that allows free and open exchanges of ideas and information has greatly increased the DQO's ability to train new analysts and update the training of current analysts.

The current wiki has over 100 pages with 34 instrument specific pages. These pages vary in length and depth of documentation, but have already evolved through innovative ideas to store and display the information. Each DQO member has provided creative edits enabling the wiki to evolve into a very effective document.

The wiki has become such a useful tool, the development of this paper and abstract were tracked with an "AnnualAMS2007Meeting" wiki page that traces and displays edits from multiple users.

5. ACKNOWLEDGEMENTS

This research was supported by the Office of Science (BER), U.S. Department of Energy, Grant Numbers DE-AC06-76RLO and DE-AC05-76RL01830

6. REFERENCES

[CamelCase 2006] Anonymous. Wikipedia. *CamelCase*. October 25, 2006. <http://en.wikipedia.org/wiki/CamelCase>

[TWiki 2006] Peter Thoeny. *TWiki*. October 1, 2006 <http://www.twiki.org/>

[Wikipedia 2006] Anonymous. Wikipedia. *Wikipedia*. October 25, 2006. <http://en.wikipedia.org/wiki/Wikipedia>