4-1-2010

Information Technology Implementation Decisions to Support the Kentucky Mesonet

D. Michael Grogan
Western Kentucky University, michael.grogan@wku.edu

INFORMATION TECHNOLOGY IMPLEMENTATION DECISIONS TO SUPPORT THE KENTUCKY MESONET

A Thesis
Presented to
The Faculty of the Department of Geography and Geology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Geoscience

By
D. Michael Grogan

May 2010

INFORMATION TECHNOLOGY IMPLEMENTATION DECISIONS TO SUPPORT THE KENTUCKY MESONET

Date Recommended ____April 24, 2010____

Dr. Rezaul Mahmood

Director of Thesis

Dr. Stuart A. Foster

Dr. James E. Gary

Dr. Greg Goodrich

____

Dean, Graduate Studies and Research   Date

ACKNOWLEDGMENTS

There are so many people in my life who have contributed to my personal and professional development and I would like to thank them all here, but space does not allow. There are a few key people, though, whom I must make room to acknowledge. First, from a personal standpoint, I thank my wife and childhood sweetheart Beth and our children – both those here and those whom we never got to meet – for their unconditional love and for supporting me in my pursuit of meteorological education and employment, including several cross-country moves, and my mom for always being there for me. I thank my dear friends Chad Ringley and Crystal Bergman for always providing much needed advice and encouragement in both my personal and professional life. I also want to thank the members of my graduating class from the University of Oklahoma (OU) meteorology program, with whom I know I share a life-long bond and an understanding that, if we could make it through that, we can do anything and meet any challenge.

On a professional level, I would like to thank Dr. Rezaul Mahmood, Dr. Stuart A. Foster, Dr. James E. Gary, and Dr. Greg Goodrich from Western Kentucky University (WKU) for serving on my graduate committee and for taking the time to trudge through my lengthy thesis. I thank Kevin Cary, WKU GIS Manager, for his always-available technical advice and camaraderie and Wendy Decroix, Department Office Manager, for always having the administrative answers I need! I acknowledge the staff of the Kentucky Mesonet for their dedication to the network and greatly appreciate the staff of the five local National Weather Service (NWS) offices serving Kentucky for their support. I also appreciate my peers in state government information technology for recognizing the Kentucky Mesonet with a Best of Kentucky Technology Award. I thank

Darryl Martin, former Meteorologist-in-Charge of the Nashville NWS office, for inspiring me to return to school to study meteorology and Bob Baron, founder of Baron Services, for playing a key part in the start of my career in meteorological data systems.

I must also thank some key educators who have played a pivotal role in my academic development. I was privileged to have two courses at OU with Dr. Kelvin Droegemeier, who I think taught me something about every aspect of meteorology in just my first introductory class and helped me tackle difficult dynamics concepts in yet another. Also at OU is Dr. Mark Laufersweiler, who helped inspire the direction my career has taken. I can never repay Mrs. Horner, my 8th grade Algebra teacher, for transforming me from an average student into a young person with a drive to succeed. Finally, I thank Ms. Tilley, my high school Senior English teacher, who, in a time when the Internet was barely known, taught me the basic literature research methods I still use today.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INFORMATION TECHNOLOGY IMPLEMENTATION DECISIONS TO SUPPORT THE KENTUCKY MESONET

D. Michael Grogan May 2010 208 pages

Directed by: Rezaul Mahmood, Stuart A. Foster, James Gary, and Gregory Goodrich

Department of Geography and Geology Western Kentucky University

The Kentucky Mesonet is a high-density, mesoscale network of automated meteorological and climatological sensing platforms being developed across the commonwealth. Data communications, collection, processing, and delivery mechanisms play a critical role in such networks, and the World Meteorological Organization recognizes that "an observing system is not complete unless it is connected to other systems that deliver the data to the users." This document reviews the implementation steps, decisions, and rationale surrounding communications and computing infrastructure development to support the Mesonet. A general overview of the network and technology-related research is provided followed by a review of pertinent literature related to in situ sensing network technology. Initial infrastructure design considerations are then examined followed by an in-depth review of the Mesonet communications and computing architecture. Finally, some general benefits of the Mesonet to the citizens of Kentucky are highlighted.

CHAPTER 1. INTRODUCTION

1.1 Kentucky Mesonet overview

The Kentucky Mesonet is a high-density, mesoscale network of automated meteorological and climatological sensing stations being deployed across the commonwealth. Through a joint resolution by the state legislature, the Kentucky Mesonet was established as "the official source of climatological observations for the state" and is operated under the direction of the state climatologist at Western Kentucky University (Kentucky Legislature 2006). The state climatologist's office is a function of the Kentucky Climate Center (KCC), Department of Geography and Geology, Western Kentucky University (WKU) in Bowling Green. The network is operated in partnership with seven other higher education institutions – Eastern Kentucky University, Kentucky State University, Morehead State University, Murray State University, Northern Kentucky University, the University of Kentucky, and the University of Louisville – that, together with WKU, compose the Kentucky Mesonet Consortium.

Each Mesonet sensing platform includes a set of instruments located on or near a 10 m tower which measure precipitation accumulation, 1.5 m air temperature, relative humidity, solar radiation, 10 m wind speed & direction, and wetness – an indicator of ongoing precipitation. Planning for the initial deployment and testing of soil moisture and temperature sensors at select sites is underway. A photograph of a typical AC-powered Mesonet site is shown in Figure 1-1, while the layout for a typical solar-powered site is given in Figure 1-2.

Figure 1-1. Typical AC-powered Kentucky Mesonet site. Site "LSML", 7 miles south of Frankfort, KY in Franklin County. (Photo source: Stephen Struebig).

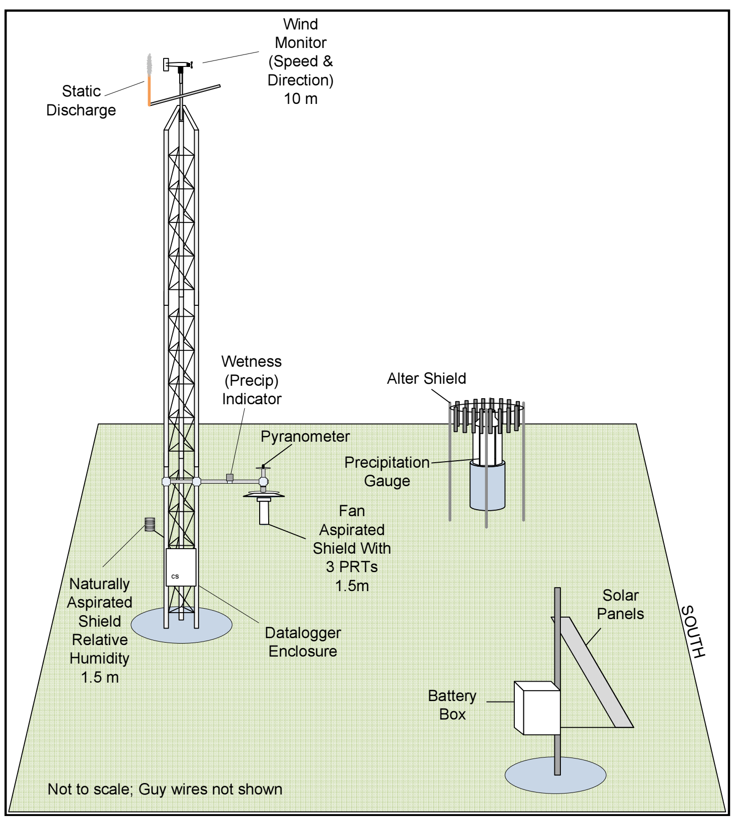


Figure 1-2. Layout for a solar-powered Kentucky Mesonet site (Struebig et al. 2010).

As defined by Orlanski (1975), “mesoscale” refers to phenomena covering between approximately 2 and 2,000 km horizontally which typically last from several minutes to a week. These systems include tornadoes, thunderstorms, squall lines, and fronts. In order to effectively capture these phenomena at the surface, Mesonet stations continue to be placed as uniformly as possible across the commonwealth. The first Mesonet site was established just south of Bowling Green in May, 2007 and the network has grown quickly since then to 46 sites online as of 28 February 2010. Figure 1-3 shows site locations for three successive Januarys, while Figure 1-4 graphs installation progress.

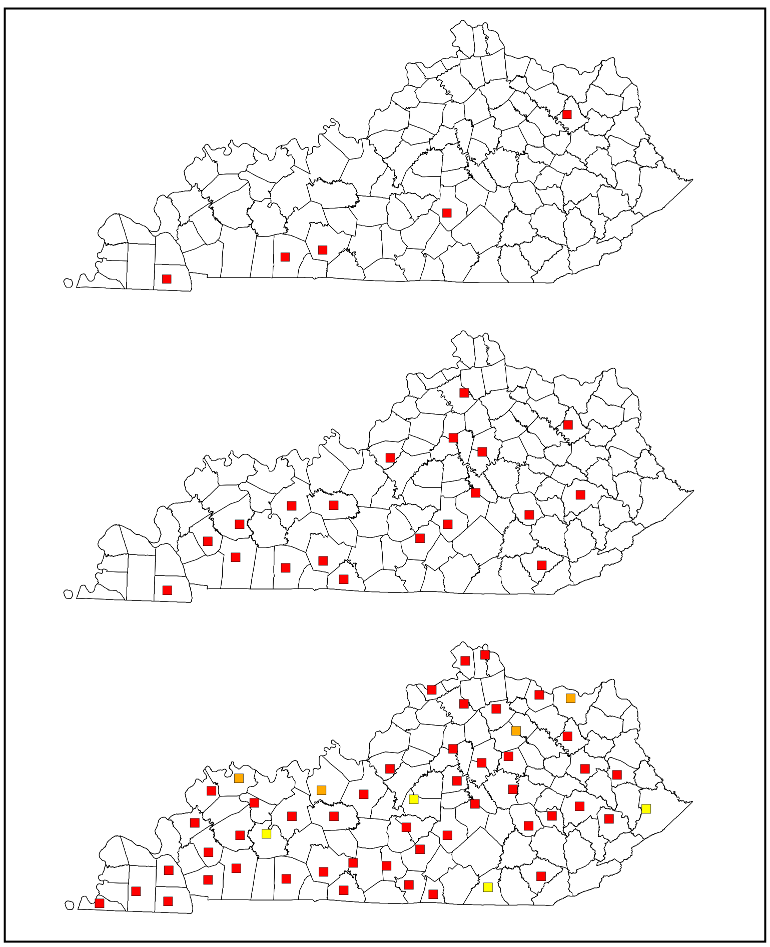


Figure 1-3. Mesonet locations (red) as of Jan., 2008 (top), 2009 (center), and 2010 (bottom). Bottom also includes sites under construction (orange) and sites with a use agreement (yellow).

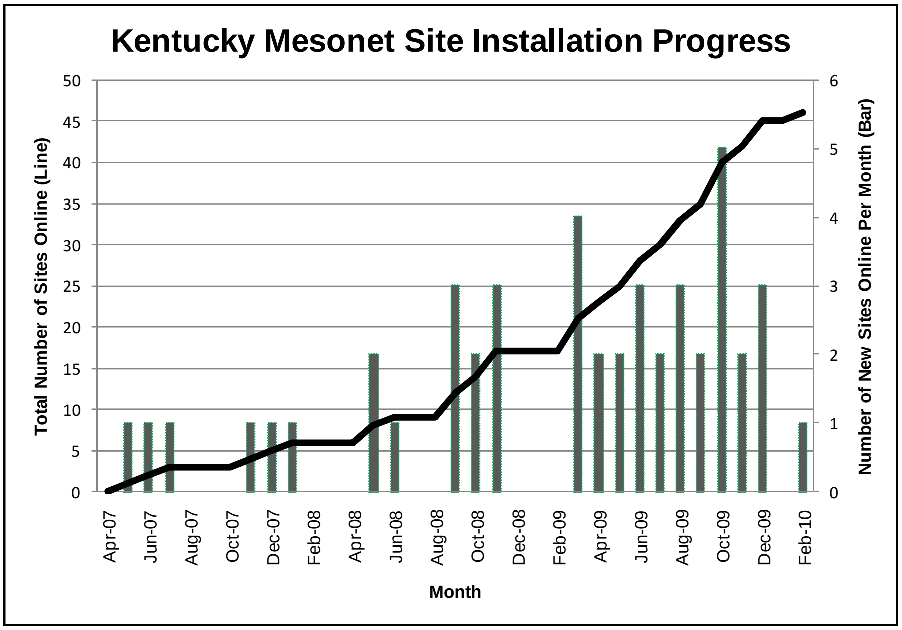


Figure 1-4. Mesonet site installation progress. Line graph indicates total number of sites online. Bar graph indicates number of sites installed per month.

With only 100 total sites planned for the network, it will not be possible to place a sensing station in each of Kentucky's 120 counties. Ideally, though, every location in Kentucky will eventually have at least one Mesonet site within 20 miles. In some cases, placement of sites near county borders helps serve multiple county interests while keeping with the network's placement priorities.

In addition to its field systems, the Kentucky Mesonet is built on and supported by at least 19 core or ancillary information technology (IT) systems, including a robust enterprise-grade communications solution; site survey, metadata, and observational database storage systems; websites; availability assurance mechanisms; and an automated quality control system. These systems support and make possible the use of Mesonet data by both the general public and critical operational partners such as the National Weather Service (NWS), broadcast media, and state government. Development of these IT systems is the core focus of this thesis.

Funding for construction and initial operation of the Mesonet program was provided via a combination of federal earmarks secured by U.S. Senator Mitch McConnell and direct grants from the National Oceanic and Atmospheric Administration (NOAA) managed under NOAA grant # NA06NWS4670010. The support of key constituents including members of the Kentucky Mesonet Consortium, local governments, private land owners, local NWS forecast offices, and other local interests has also been critical.

1.2 Research purpose and motivation

With some coverage from stations in neighboring states, prior to the establishment of the Kentucky Mesonet weather data of substantial research and operational quality were widely distributed and easily available from only 18 surface sites in the state, most of which are Automated Surface Observing Stations, and all of which report their observations in aviation routine weather report (METAR) format. Thus, the Mesonet has to date effectively tripled the number of high quality sites in Kentucky and promises a six-fold increase when full deployment is realized. Figure 1-5 shows the marked spatial improvement of total (METAR + Mesonet) surface network coverage, assuming each station at least roughly representative of a buffer zone with 20 mile radius. Of note is that the Mesonet has substantially improved the timeliness of Kentucky's routinely available meteorological data from once per hour with METARs to at least four times at Mesonet stations; more observations are available during active weather.

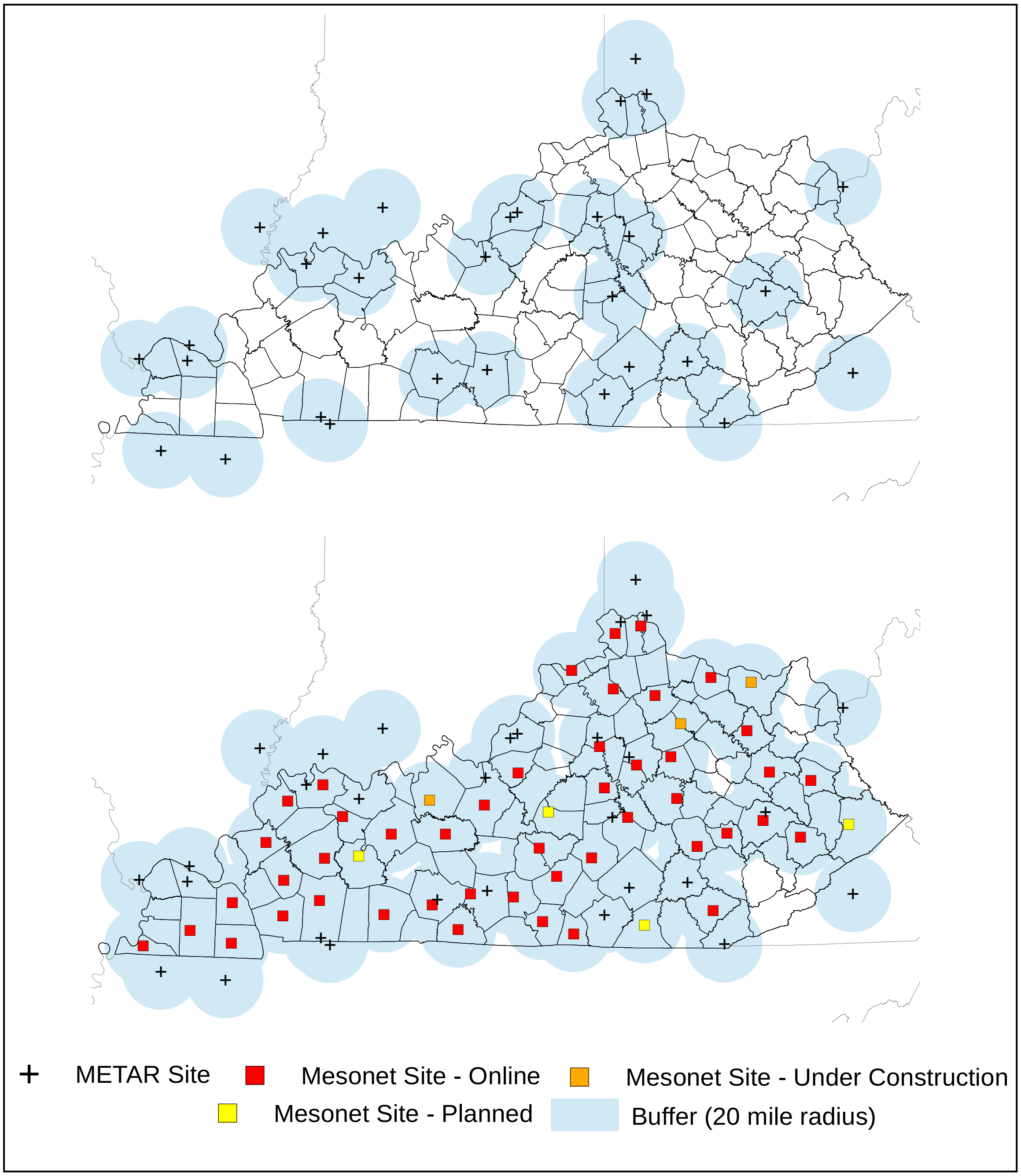


Figure 1-5. Spatial improvement of high quality research and operational surface meteorological sensing sites in and near Kentucky. Top: METAR sites (Thompson 2010). Bottom: METAR sites + Kentucky Mesonet sites (red = online, orange = under construction, yellow = use agreement in place) as of 28 February 2010. Blue polygons show a buffer of 20 mile radius around all sites.

The Kentucky Mesonet has been designed as a dual-purpose network to serve both as a critical operational meteorological sensing network and as a long-term, research quality climatological data collection medium. Each of these core uses of the network presents unique challenges for the network's supporting computing and communications infrastructure. Operational users of the Mesonet include emergency managers, broadcast meteorologists, the National Weather Service, agricultural interests, and the general public. These users need continuous, near-real-time access to network data, which requires a robust technology implementation. Research users of the network include those studying both long-term climatological and shorter-term meteorological phenomena. Research credibility demands not only collection of values measured by the network but also collection of a broad set of metadata describing those data's characteristics.

The United States Senate Appropriations Committee, Subcommittee on Commerce, Justice, Science, and Related Agencies, has stressed the critical importance of mesoscale sensing networks, has recognized the National Oceanic and Atmospheric Administration's interest in the development of a "national mesonet", and has looked to the National Research Council to provide a framework for development and operation of such a national network (S Rep. No. 110-397, 2008; S Rep. No. 111-34, 2009). This framework has been outlined by the NRC (2009) study *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks* (NNoN), commissioned by the Departments of Commerce (DOC), Transportation (DOT), and Homeland Security (DHS), the Environmental Protection Agency (EPA), and the National Aeronautics and Space Administration (NASA). The NRC (2009) study notes that national priorities demand meteorological observations at much finer spatial and temporal resolutions than are widely available today and advocates creation of a “national mesonet" by harnessing the energy and enthusiasm of state and local networks into a well-organized “network of networks”. The NRC (2009) study stresses the need for a high-quality infrastructure to support data collection, data access, quality assurance, and metadata archiving.

A review of pertinent scientific literature reveals a wealth of information applicable to information technology implementation in support of in situ sensing networks. National Research Council (1999) climate monitoring principles, World Meteorological Organization (2006) guidelines, and technology implementations for other networks all provide relevant guidance that has proven useful in the design of the Kentucky Mesonet’s technology infrastructure (Brock et al. 1995; Hubbard et al. 2005; Schroeder et al. 2005; Miller et al. 2007; Splitt et al. 2002). While literature covering networks similar to Kentucky’s certainly overviews the information technology implemented in them, sometimes few specifics are provided concerning the steps, processes, and rationale used in system design and implementation.

The Kentucky Mesonet, in addition to these existing networks, is becoming a leading example for the construction and operation of a large-scale, real-time, surface sensing network, including requisite supporting technology. The author is a member of both the Architecture and Research & Development Testbed working groups for the American Meteorological Society’s advisory efforts on building a NNoN. Kentucky’s emerging network is being watched closely by the groups. Lessons learned from its design promise to help fill in gaps in written scientific literature covering in situ network construction and to provide an updated perspective on existing knowledge.

The purpose and motivation of the author's applied research, therefore, has been to:

(i) significantly increase the spatial coverage, amount, timeliness, availability, and use of original, quality surface meteorological data in Kentucky;

(ii) develop the core information technology systems necessary to support both mission-critical operational and research use of the Kentucky Mesonet;

(iii) show that core information technology-related competencies required by a national network of networks are achievable at the local level, even with a small staff;

(iv) and to provide in the literature an updated perspective on building the IT-related infrastructure to support a statewide in situ surface sensing network, especially in the areas of communications, data ingest, and processing systems.

1.3 Document overview

As this document will show, the research and development (R &D) efforts of the author as the Mesonet's Lead Systems Architect – with assistance from other Mesonet personnel – have substantially solved the challenges or fulfilled the goals surrounding each of these purposes. Though the R & D results examined here include some highly technical discussions, this document serves not as a basis for some type of internal systems or operational guide for the Mesonet but, instead, focuses on technology implementation processes and rationale.

Though it does include each important element, due to the breadth of the work examined this thesis does not follow a traditional "Introduction – Literature Review – Methodology – Results – Conclusion" format. Following the requisite literature review, which covers both general computing and in situ surface sensing network concepts, this document examines the initial design considerations and processes of the Kentucky Mesonet. Network communications choices and implementation are then detailed, followed by a brief examination of overall computing code design approach. The Mesonet's core, geographic information, and ancillary computing systems are then reviewed, followed by a look at some of the early benefits and uses of Mesonet data. The typical discussions and conclusions round out the document.

1.4 Important background information and considerations

The Kentucky Mesonet's organizational structure consists of three principle divisional foci:

(i) field and instrument operations,
(ii) information technology,
(iii) and quality assurance and control.

The Field and Instrumentation division consists of field meteorologists and technicians responsible for the design and construction of Mesonet sites, the calibration and maintenance of the network's instrumentation, and the programming and operation of the network's dataloggers which are used for data collection. The Information Technology division, including the author, a full-time developer, and a part-time student, is responsible for all technical project infrastructure and operations outside of the instruments and logger, including critical site communications. Finally, though sometimes considered part of the Information Technology division in terms of function, the Quality Assurance (QA) and Control division – operated by a QA specialist and student operators – is responsible for the overall quality of Mesonet data. The functions of the two non-IT divisions, except in areas of crossover with or support by the IT division, are generally out of scope for discussion in this document.

Outside of typical desktop support services from WKU's Information Technology helpdesk and some assistance with site communications procurement services from its Communications Technology division, the Mesonet's IT division is on its own in the building, maintenance, and support of the critical infrastructure needed to support the network. Unfortunately, the division's Application Developer position has experienced high turnover, with three different people holding this position in the last three years.

Given that he leads the Information Technology Division, the author of this thesis is a full-time, professional employee of the Kentucky Mesonet. In his official capacity, he has therefore directly supervised or actively guided the work of other Mesonet personnel and contractors in the development of some of the systems described within this document. While overall this thesis represents the cumulative work of the author, it must for completeness cover some work performed by these other parties. A notation is provided in the description of all systems or processes that have significantly benefited from the contributions of such other persons or entities. For systems detailed without such a notation, it should be assumed that the author provided the overwhelming majority of effort in their design or implementation.

In the review of this work, it is critical for the reader to remember that the Kentucky Mesonet is a mission-critical, operational entity complete with its own decisions and priorities which are not predicated on the need of an individual program employee to finish an academic degree. Such a degree, however, obviously has inherent deadlines including those for completion of a written thesis. While some Mesonet systems have not been developed to their full potential due to operational challenges, degree deadlines have required the author to describe a "snapshot" of these systems and programs as they existed at the time of thesis preparation. Improvement to systems design and implementation will continue well after thesis completion. For most references made by this document, the network is described as it stood on 28 February 2010.

In consideration of the public nature of this document, a significant portion of the author's work cannot be included. Specifically, detailed information about the Mesonet's computing network topology, server configurations, and security practices are considered too sensitive for publication.

Finally, two types of networks are referenced throughout this document. One network is the meteorological sensing network that is the Kentucky Mesonet, while the other network is the supporting computing network. Except in cases where the computing network is specifically referenced as such, the referenced network should be assumed to be the sensing network.

CHAPTER 2. LITERATURE REVIEW

2.1 Introduction

Measurements are the foundation of meteorological and climatological studies. Rotch (1904) noted that meteorology became a science upon the invention of principal meteorological instruments in the seventeenth century, which made possible the collection of "exact and comparable" observations at many places on the globe. Just fifty years after the first synchronous observations across a "considerable" territory were telegraphed to a central office to aid in weather forecasting, Rotch (1904) also suggested that meteorology's data collection infrastructure was "tolerably complete", except for a few gaps on the Antarctic continent and the interior of Africa.

More than 100 years later, though, expansion of meteorological and climatological measurement platforms continues, in what Miller and Barth (2003) sees as a response to the need for more frequent, densely spaced, real-time observations to aid in agricultural monitoring, energy and transportation planning, emergency management, fire management, and meteorological research and education. Communications, computing power, and other technological resources have played an integral part in this expansion, especially in recent decades. Such resources will certainly continue their role as principal expansion facilitators for many years to come.

Knowledge of the history and best practices of both meteorological/climatological observation and computing systems is critical for the construction of the Kentucky Mesonet. This document reviews relevant literature and other sources used in search of guidance toward the development of the Mesonet's computing and communications infrastructure. Specifically, it provides a general overview of the history of, need for, and types of in situ surface observing systems in the United States. General computing system requirements are then examined, followed by a review of relevant national and international standards for in situ measurement systems. Finally, several existing measurement / data collection programs similar in scope to that of Kentucky's effort are examined in detail.

2.2 In situ surface observing systems in the United States

Fiebrich (2009) provides a fascinating review of the history of surface observations in the United States, noting that instrumented observations began in colonial Massachusetts at Cambridge in 1715 and Boston in 1725. By the mid-1700s, several American colonists were making regular observations, looking for connections between weather and social issues such as diseases. Some famous U.S. presidents were certainly interested in the weather. Fiebrich (2009) remarks that George Washington's last written words were likely used to detail the weather in Mount Vernon, Virginia. Thomas Jefferson also kept a daily record of weather conditions from 1776 to 1816. The number of observations has obviously increased significantly since then. A review of both federal and non-federal roles in this effort is given below.

2.2.1 History and status of federal efforts

The federal government has long been directly involved in the collection and reporting of in situ observations in the United States. In 1819 the Army organized a system to make weather observations part of the regular routine at its U.S. posts. Established by congressional resolution in 1870, the U.S. Weather Bureau – operating under the Signal Corps – made the first 24 synchronous weather observations at 7:35 a.m. on 1 November of that year. By the time Congress transferred the Weather Bureau to the Department of Agriculture in 1890, it was being realized that collecting sufficient climate records was requiring greater funding than available at the time (Fiebrich 2009). From this realization sprang the Cooperative Observer Network (COOP).

Winkler (2004), which provides an excellent review of federal efforts, notes that the COOP program is the oldest and largest official network in the U.S., with more than 11,000 volunteers recording and reporting daily measurements of maximum and minimum temperature, liquid equivalent of precipitation, snowfall, and other climate-related variables. COOP observations are generally not provided in real time but, instead, are first sent to the National Climatic Data Center (NCDC) where observations are quality controlled before being made available to the public.

The mission of the Automated Surface Observing System (ASOS) – now the primary federal automated observing system in the country – on the other hand is to provide routinely updated data for weather forecasting and aviation needs, including measurements of temperature, relative humidity, pressure, wind speed & direction, rainfall, visibility, cloud ceiling, and precipitation type. Installed in the early 1990s, with the majority of stations commissioned after 1996, ASOS replaced more conventional methods of observation, typically performed by a human observer. As of 2004, ASOS was comprised of 569 sites operated by the Federal Aviation Administration and 313 sites operated by the National Weather Service (NWS), the modern day descendent of the Weather Bureau (Winkler 2004).

ASOS was certainly not the first automated observation system in the U.S. That claim belongs to the U.S. Navy which established the first such station, weighing one ton and powered by a gasoline electric plant, in 1941. In support of aviation interests, the NWS designed the Remote Automatic Meteorological Observing System (RAMOS) in 1969 to collect, process, and transmit information on a number of meteorological variables (Fiebrich 2009).

Winkler (2004) notes that certain inhomogeneities have arisen in both the COOP and ASOS networks due to differences in or biases from instrumentation, station location, and distance between obstacles and measurement platforms. To prevent such inhomogeneities from polluting the official climate record, the U.S. Historical Climatology Network (USHCN) was developed from a subset of approximately 1200 COOP stations, which were selected based on long periods of record, small percentages of missing data, and a minimum number of changes in station location, instrumentation, and observing time. With reasoning similar to that for the USHCN, the U.S. Climate Reference Network (USCRN) effort began in 2001 to provide long-term, high-quality climate observations over the next 50 to 100 years (Hubbard et al. 2005). USCRN, however, focuses on construction of new stations rather than utilization of existing ones. Its motivation and infrastructure are reviewed more extensively later in this document.

Finally, an initiative to construct NOAA's Environmental Real-Time Observation Network (NERON) began in 2004, as part of an effort to create an Integrated Surface Observing System (ISOS) by simultaneously modernizing the COOP program while providing real-time data from the USHCN. By December 2006, one hundred new surface stations had been installed in New England and eastern New York toward the goal of creating a network of surface observing stations at a nominal density of one per every 400 square miles (Crawford and Essenberg 2006; OCS 2006). Though it was making progress, NERON appears to have been mostly abandoned. Servers supporting the project, developed under a grant to the Oklahoma Climatological Survey, were delivered to National Weather Service headquarters and were essentially shelved after testing. The remaining functions of NERON have reportedly been merged with other USHCN expansion efforts. Data from newly installed NERON sites appear to be no longer publicly available.

2.2.2 History and status of nonfederal efforts

While in name NERON may have passed by the wayside, its contributions are fortunately not completely lost – or at least unknown – as they were largely based on the development of the Oklahoma Mesonet, a state-led effort to deploy a mesoscale network of surface and sub-surface sensing stations. Perhaps decreased bureaucratic pressure allows for easier network development at the state, local, and private level, as Meyer and Hubbard (1992) noted a tremendous growth in nonfederal automated weather stations (AWSs) across the United States and Canada in the 1980s. By 1983, some type of AWS had been developed in Colorado, Florida, Idaho, Louisiana, New Mexico, New York, Oregon, and South Dakota. Major installation campaigns were also thought to be underway in Alabama, Georgia, Nevada, and Oklahoma.

The networks examined by Meyer and Hubbard (1992) were “fueled by the need for more specific meteorological data in real or near-real time” than were available from federal sources. “First-order” stations at that time, such as those operated by the National Weather Service, did not provide the spatial density necessary for many research purposes and often did not provide the specialized data sought by the nonfederal networks, such as information directly applicable to agricultural interests. The majority (51%) of the networks examined by Meyer and Hubbard (1992) consisted of five or fewer stations, 35% had between 6 and 20 stations, while only 14% had more than 20 stations. The number of nonfederal stations totaled 608.

By 2007, the Meteorological Assimilation Data Ingest System (MADIS), an aggregator of meteorological data sets, was collecting data from over 20,000 AWS stations operated by local, state, and federal agencies and private firms (Miller et al. 2007). MesoWest, a data aggregator from the western U.S., was collecting data from at least 2,800 stations (Splitt et al. 2002). While both of these aggregators are treated more extensively later in this review, their station tallies are briefly examined here to highlight continued growth in surface observing systems.

Trenberth et al. (2002) notes that such expansion in these types of networks has been justified by their increased role in monitoring and modeling climate change and by their use to reduce climate and weather-related risks in the protection of life and property. Aside from operational meteorology uses, though, true climatological value in such systems comes from providing continuous data over a large area for a long period of time.

Though Meyer and Hubbard (1992) noted what were thought to be considerable sensor network installation efforts in several states, there appears to be only a few examples in relevant literature of nonfederal networks operated by a single entity over a large area with uniform spatial density for any significant length of time. For instance, a review of over 1,600 surface observing stations in the western U.S. by Tucker (1997) showed that many networks operated 20 or fewer stations while stations in networks with larger numbers tended to be more closely clustered together. Many other nonfederal providers appear to be data aggregators, similar to MADIS and MesoWest, or private networks which do not reveal their network statistics in the scientific literature.

The best examples of nonfederal entities attempting to operate large networks with uniform spatial density appear to be the Oklahoma Mesonet, the West Texas Mesonet, aggregator-turned-operator the Delaware Environmental Observing System, and the upstart Kentucky Mesonet (Brock et al. 1995; Schroeder et al. 2005; Legates et al. 2005; Brown et al. 2008; Grogan et al. 2010). Both the Oklahoma and West Texas networks are examined extensively in subsequent sections. As noted, the subject of this review itself was to aid in the development of the Kentucky network.

Finally, Tucker (1997) notes that the distinction between federal and nonfederal networks is often blurred, as many projects receive federal funds to either construct or operate their network. Such is the case for the Kentucky Mesonet, which has received a substantial amount of construction funds from NOAA via both federal earmarks and direct grants.

2.2.3 The emerging Nationwide Network of Networks

As described in the review of federal efforts, NERON's push to develop a nationwide meoscale observing network solely owned and operated by the federal government has been abandoned, but the push to create a "national mesonet" has not. Instead, the National Research Council (2009) supports aggregating individually owned networks – both federal and nonfederal – into a virtual network whose infrastructure supports systematic, nationwide collection and dissemination of observations. In this so called "nationwide network of networks" (NNoN), individual operators would continue to serve their specific missions as they do now, but would be subject to new standards and practices from which would be derived a collective benefit (p. 159). Data from these individual networks would be aggregated and a limited set of national products based on raw observations would be made available (p. 7). Though leadership from the federal government is important for the NNoN, it is desired that nonfederal operators play a large role in its conception and operation. The American Meteorological Society's Ad Hoc Committee on the NNoN is currently providing scientific input regarding many network topics; the author currently serves on the committee's Architecture and R & D Testbed working groups.

2.3 Design considerations – general computing systems concepts

The World Meteorological Organization (WMO) holds that "an observing system is not complete unless it is connected to other systems that deliver the data to the users" (WMO 2006). Since this document is intended to support and describe development of the communications and computing systems which facilitate such a connection, a brief examination of general information technology project practices is useful. Though far from comprehensive, the role of information technology (IT) in an enterprise is reviewed below, along with certain applicable IT definitions, decisions, and design principles. The review moves from higher- to lower-level concepts, beginning with overall considerations of IT within an organization, moving to a treatment of overall IT project and network design principles, and ending with a brief consideration of an individual software application design concept.

2.3.1 Role of IT in the enterprise

Dewett and Jones (2001) notes that the availability and use of information services and technology has grown almost to the point of being commodity-like in nature, becoming nearly as ubiquitous as labor. The article holds that information technology leads to information efficiencies (an increase in amount and quality of information) and information synergies (performance gains via collaboration) in the enterprise by contributing in unique ways:

(i) IT codifies the knowledge base by facilitating organizational memory and making knowledge easy to communicate, assimilate, store, and retrieve.

(ii) IT increases boundary spanning by allowing employees to quickly access useful knowledge and data.

(iii) IT promotes efficiency by providing the ability to store and retrieve lots of information quickly and inexpensively.

(iv) Given that it determines the way information is stored, transmitted, communicated, processed, and acted upon, IT promotes innovation by moderating many aspects of the process of bringing new ‘problem solving ideas’ into use.

2.3.2 Information technology definitions

Before information technology design practices and principles are examined, it is perhaps appropriate to offer some definitions of general information technology concepts, starting with information technology itself and ending with a definition of an enterprise application.

2.3.2.1 Information Technologies, Information Systems, and Information Technology

Dewett and Jones (2001) defines information systems as the enterprise-wide systems designed to manage all major functions of the organization as well as general purpose systems targeted toward specific uses. Information technologies are described as a broad array of communications media and devices which link people with the information systems. Because information systems and information technologies are inextricably linked, Dewett and Jones (2001) suggest they be collective referred to as information technology (IT).

2.3.2.2 Systems Architecture and Engineering

Fowler (2003) notes that architecture is a term that many IT practitioners attempt to define, but with little agreement, while suggesting that architecture consists of two common elements. The first element is a breakdown of a system into its parts, the other being the decisions made about systems design. Martin (2006) states that systems architecture is concerned with an overall integrative, multi-level systems perspective that includes both component level and application level engineering while suggesting that systems engineering is a broader concept that includes information technology hardware development and policy implementation. Finally, Zachman (1987), the most widely referenced authority on the definition of systems architecture, holds that there is not *a* systems architecture but a *set* of additive and complimentary architectures, including architectures for describing data, IT processes, and computing networks.

2.3.2.3 Enterprise Applications

Fowler (2003) defines enterprise applications and systems generally as those that handle lots of persistent data and multiple, concurrent access to that data.

2.3.3 IT project success and failure

Martin (2006) examined the factors that influence the success and failure of IT application projects and identified at least three factors associated with high or improved performance and three associated with poorer performance. Factors found to be associated with high or improved performance are:

(i) a mature, well-planned approach to project architecture;

(ii) externally sourced information technology systems ‘behaving as expected’;

(iii) and, ironically, strong non-functional requirements for high performance, reliability, and security.

Factors found to be associated with poorer performance include:

(i) application of a conservative technology strategy,

(ii) changes in project requirements or staff,

(iii) and requirements for application portability across multiple platforms.

Dewett and Jones (2001) identified the role that time plays in the success of information technology implementations, noting a progression of IT use and success within an organization over time. At first, organizations may be less successful in their IT projects as they learn how to use and implement IT to its fullest potential – so called ‘first-order’ learning. With time, though, IT does become successfully engrained in an enterprise, and related implementation activities transition to ‘second-order’ learning, where technologies are modified to better match the organizational environment.

2.3.4 Project configuration decisions – variables and drivers

Management decisions and configuration constraints are obvious additional drivers leading to the success or failure of IT projects. While noting that resource availability, time scale, and supply of available employees or contractors can greatly constrain project configuration, Martin (2003) suggests that there are three key management decision variables associated with project configuration:

(i) IT architecture – the planned, integrated choice of computing systems (hardware, software);

(ii) the resources and skills of the people implementing the project;

(iii) and application of appropriate methodologies and practices.

The actual drivers of project configuration are:

(i) project requirements – the need for the project to satisfy functional and non-functional requirements within its time scale and budget;

(ii) strategic objectives – an organization's strategic objectives strongly influence configuration;

(iii) risk management – risk management may dictate that a project remain behind the leading edge of technology in well-chartered waters;

(iv) experience – application experience is important in the process of program design;

(v) and pragmatic resolution – the need for delivering to specification within time and budget must be practically reconciled with strategic guidelines for architecture, resources, and methodology.

2.3.5 Project design principles

Richardson et al. (1990) suggests a set of experienced-based principles for guiding the design of information technology which revolve around the enterprise, data, and applications. For the enterprise, they hold that IT professionals need to report either directly or indirectly to the person responsible for the IT function within a business unit and that IT functions be organized to make the most effective use of IT as a strategic tool. They suggest that successful application development, based upon formal planning methodologies, requires proactive user and sponsor involvement to ensure proper functionality and ultimate success. Finally, they believe that data should be viewed as a corporate asset and should be managed as such.

2.3.6 Network design considerations

Computer network management is key to the success of any IT initiative. Murhammer et al. (1999) holds that any network should be designed around eight fundamental principles:

(i) scalability – A well designed computing network is one that is scalable, or able to grow and accommodate new requirements;

(ii) open standards – The computing infrastructure and equipment used in the network should employ open standards to ensure compatibility with other devices. Proprietary features of network infrastructure should be avoided as they can severely limit flexibility, especially in the future;

(iii) availability – Availability generally refers to the amount of time a computing network is accessible and capable of performing its required tasks. Logical and physical redundancy are key to ensuring the availability of a computing infrastructure;

(iv) modularity – Modularity is the division of a complex system into smaller, more manageable parts. In a modular architecture, failure of one computing or network system does not cause the entire infrastructure to fail. Also, the addition of a network segment does not require readdressing of all hosts in the network;

(v) security – Obviously, security is of utmost importance in any computing network. Security risks must be considered during the design phase of a network instead of being an afterthought. Security considerations are critical when a computing system will be accessible from the internet;

(vi) network management – The ability to manage an IP-based network should be considered at the outset of network design. Network management design should include methods to monitor the health of the network, to ascertain operating conditions, to isolate faults, and to configure devices;

(vii) performance – There are two types of performance important for a computing network: throughput and response time. Throughput is the amount of data that can be sent/received by the network in the shortest time possible, while response time is the amount of time a user must wait before a result is returned by the network;

(viii) and finally, economics – One of the most difficult challenges of computer network design is balancing costs while meeting all other requirements of the network. Some fancy features may have to be dropped in order to meet cost requirements, but care should be taken to still meet other basic network requirements.

2.3.7 Application design considerations

Finally, Fowler (2003) provides guidance toward the creation of specific applications within an IT project or organization, suggesting that a layered design approach be utilized. The three principal layers are:

(i) the presentation layer, which is primarily responsible for the display of information to the user and the interpretation of commands from the user into actions which operate on the data source and domain layers;

(ii) the data source layer, which is responsible for communicating with other systems that carry out tasks on behalf of the application and for which an enterprise database is usually the biggest member;

(iii) and the domain layer, also referred to as the business logic layer, where the actual work that the application needs to do is performed according to rules specific to the enterprise.

2.4 Design considerations – in situ surface network requirements

Langdon (2003) suggests that any information systems architecture must be very cognizant of business needs and must include methods where business requirements and information systems capabilities are matched. Failure to consider enterprise-specific requirements throughout the design of all computing and communications systems could lead to, at best, an architecture poorly matched to the needs of the enterprise or, at worst, to complete project failure. In other words, business requirements are the principal concern behind systems design.

As with any computing network being designed for a specific purpose, in situ meteorological sensing networks place specialized design demands on their supporting computing infrastructure. The most stringent requirements arise when the meteorological network will also be used to build a long-term, research quality climatological record. Once again, relevant scientific literature saves the day and provides guidance on the demands required of information technology by in situ networks, or at least on the functional requirements which the IT infrastructure should support. This literature is reviewed below.

2.4.1 General literature

Karl (1993) examines the requirements for databases derived from long-term measurements that can be used to document and help understand historical and ongoing climate variations and change, noting that inhomogeneities in the data must be avoided. While station histories can be easily ignored since they are not a requirement for station functionality, Karl (1993) holds that continuous documentation about station location, types of instruments used, their exposure and elevations above ground, information about local surroundings, observing schedules, and maintenance procedures are critical.

Trenberth et al. (2002) extends this concept, expressing that a climate observing system must focus not only on the climate observations themselves, but also on the processing and support systems which ultimately lead to reliable and useful products. They maintain that a real-time quality control system must be implemented to guard or warn against biases, errors, or missing data, advancing the idea that the absence of a commitment to reliability ultimately leads to an archive incapable of delivering quality data.

Certain obstacles though, such as availability of funding or other resources, certainly constrain the ability of a sensing network to fully implement all desired information technology systems and can lead to an adjustment of implementation plans. Realizing this, Trenberth et al. (2002) suggests that the following priorities, from highest to lowest, be maintained:

(i) data collection and archiving,

(ii) distribution of the raw data in near-real time,

(iii) quality control of the data in delayed mode and archiving of datasets,

(iv) development and maintenance of data access tools (e.g., web sites),

(v) and follow-on processing to produce analyses and reanalyses.

As part of an international workshop, Brown and Hubbard (2000) provided guidance based on important lessons learned in the development of automated weather stations, noting that planning is the most important aspect of developing and operating a network, that it should begin at network conception and continue throughout the life of a network, and that it should certainly extend to data retrieval, processing, and quality control procedures. Noting the same funding obstacles as Trenberth et al. (2002), Brown and Hubbard (2000) also stressed the importance of educating administrators about the cost of running a weather network, holding that funding based on short-term grants can lead to a "feast or famine" funding cycle which can create a loss of network focus and make key technical personnel retention difficult. In terms of a network's technical architecture, Brown and Hubbard (2000) stressed the development of automated quality checks on incoming and processed data, the development of value-added analysis products, and the essential creation of network awareness by potential stakeholders via outreach activities.

2.4.2 National Research Council (1999) climate monitoring principles

Certain national and international standards certainly provide guidance toward the creation of in situ networks and their attendant information technology functions. The first of these considered here is the result of the work of the National Research Council (NRC 1999) Panel of Climate Observing Systems Status. The panel noted that climate researchers often rely upon existing, operational networks for data but that confidence in research results can be severely limited by deficiencies in the accuracy, quality, and continuity of network records.

To help prevent those limitations, the NRC (1999) panel adopted ten climate monitoring principles that should be applied to climate monitoring systems. Five of these principles – metadata, data quality and continuity, continuity of purpose, data/metadata access, and climate monitoring requirements – directly impact computing network design and are therefore examined in greater detail below.

2.4.2.1 Metadata

Metadata is essentially data about the data. McGuirk and May (2003) defines it as everything a researcher would need to know in order to process a network's climate data. NRC (1999) principles require that each observing system and its operating procedures be fully documented. Such documentation must cover all facets of the sensing network, including instruments, instrument sampling time, calibration, validation, processing algorithms, station location, exposure, local environmental conditions, and other platform specifics that could influence the data history. NRC (1999) holds that metadata collection should be a mandatory network function and that metadata should be archived with the original data.

2.4.2.2 Data quality and continuity

A climate monitoring network should assess data quality and homogeneity as part of its routine operating procedures. The assessment should include routine evaluation of long-term, high resolution data capable of revealing and documenting important extreme weather events.

2.4.2.3 Continuity of purpose

NRC (1999) holds that a climate monitoring network must maintain a stable, long-term commitment to its observations. Long-term data storage provisions should be made and the data record should be insulated from bumps associated with uncertain funding situations.

2.4.2.4 Data and metadata access

NRC (1999) encourages climate monitoring networks to develop data management systems that facilitate data access, use, and interpretation of data and data products by users. High importance is placed on freedom of and low cost access to data through directories, catalogs, browsing functions, etc. Access to metadata on station and sensor histories should also be made available. Also, “quality control should be an integral part of data management.”

2.4.2.5 Climate monitoring requirements

Finally, NRC (1999) actually recognizes the need for complete understanding of an in situ network's 'business' requirements in the design of its information systems. Specifically, network designers, operators, and engineers should fully understand climate monitoring requirements at the outset of network design.

2.4.3 Global Climate Observing System (GCOS) design principles

The Global Climate Observing System (GCOS), co-sponsored by the World Meteorological Organization (WMO), the Intergovernmental Oceanographic Commission (IOC), the United Nations Environment Program (UNEP), and the International Council for Science (ICSU), was established in 1992 to ensure that the observations and information needed to address climate-related issues are obtainable by and made available to all potential users (GCOS 2003). In the course of GCOS development, twenty climate monitoring principles have been adopted, five of which directly apply to supporting information systems:

(i) The details and history of local conditions, instruments, operating procedures, data processing algorithms, and other factors pertinent to interpreting data (i.e. metadata) should be documented and treated with the same care as the data themselves.

(ii) Data management systems that facilitate access, use, and interpretation of data and products should be included as essential elements of climate monitoring systems.

(iii) Data systems needed to facilitate user access to climate products, metadata and raw data, including key data for delayed-mode analysis, should be established and maintained.

(iv) Operational production of priority climate products should be sustained and peer-reviewed new products should be introduced as appropriate.

(v) The quality and homogeneity of data should be regularly assessed as a part of routine operations.

2.4.4 World Meteorological Organization (WMO) standards

In its *Guide to Meteorological Instruments and Methods of Observation* (WMO 2006), the World Meteorological Organization provides a substantial wealth of guidance toward the operation of automated weather stations. Guidance and details related to their supporting information systems – specifically, information related to general systems design, data collection, metadata, and quality assurance/control – are reviewed below.

2.4.4.1 General design considerations

The WMO (2006) expresses that the specification of functional and technical requirements of the IT systems which support a network of automated weather stations is a complex and often underestimated task, noting that it requires close cooperation between in situ network designers, specialists in telecommunications, software specialists, and data users. It holds that the centralized computing system should facilitate data acquisition; remote control and housekeeping of sensing stations; network monitoring; data archiving, quality control, and processing; and data transfer to internal or external users.

2.4.4.2 Data communications and transmission

As WMO (2006) asserts, data transmission and communications provide the link from a sensing station to the outside world while holding that the appropriate means of transmission depends on the site(s) in question and the most readily available transmission equipment. As the document details, data transmission between a sensing station and the central computing system can operate in different modes – in response to external commands, at periodic time intervals, or in emergency conditions when certain meteorological thresholds are crossed. Both one- and two-way communications are potential options for data collection, with two-way operations being more powerful as they enable the central computing systems to send command messages to the field to initiate a change in mode of operation or to upload new operating software. Two-way communications also allow for data to be collected at non-routine times.

WMO (2006) examines potential choices of communications technology as well, including:

(i) Integrated Services Digital Network (ISDN), which offers very high security and data speeds adequate for climatic data transfer;

(ii) Wide Area Network (WAN), where the sensing station and central computing system are nodes on the network, with data divided into packets according to specific transmission protocols;

(iii) Virtual Private Networking (VPN), where data flowing between sensing sites and central operations is encrypted on a public telecommunications network;

(iv) and finally, dedicated circuits, where central computing facilities are directly connected to sensing sites.

Admittedly, many of these technologies overlap and there are certainly more choices available, some of which are hybrids of those suggested above.

2.4.4.3 Metadata

WMO (2006) maintains that the central computing systems of quality in situ sensing networks must enable the collection and availability of detailed information concerning the observing system itself and all changes to it that occur during the time of its operation. Specifically, the metadata database should include:

(i) network information, such as the operating authority, and the type and purpose of the network;

(ii) station information, such as administrative information, location, descriptions of surroundings and obstacles, instrument layout, facilities (communications, power supply, cabling), and climatological description;

(iii) and individual instrument information, including manufacturer, model, serial number, operating principles, performance characteristics, calibration data and time, siting and exposure, etc.

2.4.4.4 Quality assurance and control

WMO (2006) notes that quality control aims to achieve assured quality and consistency of output “through a carefully designed set of procedures focused on good maintenance practices, repair, calibration, and data quality checks”. It advocates a robust automated quality control system, facilitated by ‘appropriate’ hardware and software routines, which minimizes the number of inaccurate and missing observations. Quality control algorithms may be applied in either real time, where data are checked during initial acquisition or processing stages as close to the time of observation as possible, or in delayed mode, where more robust statistical and spatial data checks are possible. Recommended quality control checks include intra-sensor comparisons, inter-sensor comparisons, hardware checks, etc.

2.5 Review of similar projects

Using recommendations from the sources and national / international standards described above as a guide, a review of literature and other information has been conducted to reveal meteorological and climatological industry best practices that are applicable to the design and operation of the Kentucky Mesonet information technology infrastructure. Five operations – the Oklahoma Mesonet, the U.S. Climate Reference Network, the West Texas Mesonet, the Meteorological Assimilation Data Ingest System (MADIS), and MesoWest – involved in the collection, correction, and dissemination of in situ surface meteorological data have been chosen for review. These operations were selected based on their similarity with Kentucky Mesonet goals in terms of size, scope, organization, and/or functionality, even though some of them are federally operated. Though the last two projects reviewed are only data "aggregators", their functional requirements are similar enough to those of the Kentucky Mesonet to warrant review. Where applicable and available, each operation is examined in terms of its general description, general computing system design, data collection & storage mechanisms, metadata databases, quality assurance / quality control systems, data access systems, and availability assurance mechanisms.

2.5.1 Oklahoma Mesonet

2.5.1.1 General description

Begun in 1987 and operational on January 1, 1994, the Oklahoma Mesonet (Oklahoma) sets a high standard for dual-use, high-density meteorological/climatological sensing networks and is quite possibly the best known and operated system of its kind. It consists of 115 sensing stations across Oklahoma, with at least one station in every county. The network was built and is maintained by the University of Oklahoma, the Oklahoma State University, and the Oklahoma Climatological Survey (OCS) (McPherson et al. 1999, 2007). Meteorological parameters measured by Oklahoma (Shafer et al. 2000) include:

(i) 10 m wind speed & wind direction,
(ii) 9 m temperature,
(iii) 1.5 m temperature & relative humidity,
(iv) 2 m wind speed,
(v) 1.8 m solar radiation,
(vi) leaf wetness,
(vii) rainfall,
(viii) barometric pressure,
(ix) soil moisture,
(x) and 5, 10, and 30 cm soil temperature.

Daily operations for Oklahoma fall under the jurisdiction of OCS, which was established in 1980 to provide climatological services and research for the citizens of Oklahoma. As McPherson et al. (1999) notes, OCS's primary outreach activities before Mesonet creation involved judging science fair projects, speaking at career days and other school events, and providing summaries of state climate data to interested parties. With the Mesonet, OCS is now able to provide five-minute meteorological observations from across the state and has collected over 3.5 billion weather and soil observations since its inception (McPherson et al. 2007).

OCS staffing levels have grown substantially since creation of the Oklahoma network. In 1990, the Survey employed four scientists, two administrative assistants, and four to six students. By mid-1998, it had 30 full-time and 33 student employees (McPherson et al. 1999). A check of the OCS website[1] shows approximately the same number of employees today.

2.5.1.2 General computing system design

In its earliest days, Oklahoma's central computer system consisted of a field communications PC, a data logger PC, a pair of DEC VAX machines, and a pair of data dissemination PCs (Brock et al. 1995). Today, the Oklahoma network's computing infrastructure includes approximately 30 "x86-style" computers, mostly running a Linux operating system, which perform both in situ data tasks and administrative functions (Wolfinbarger 2006).

[1] http://climate.ok.gov/aboutocs/directory.php

2.5.1.3 Data collection

In 1995, Oklahoma collected five-minute temporal resolution data at fifteen-minute intervals. Today, it collects on five-minute intervals. To retrieve data from remote sites, it utilizes the Oklahoma Law Enforcement Telecommunications System (OLETS), a statewide radio communications network composed of city, county, state, federal, and military law enforcement agencies. A direct radio link exists from each remote site to a nearby OLETS terminal in a sheriff's office or other similar location. Messages from OLETS are then routed to the Oklahoma Mesonet operations center (Brock et al. 1995; McPherson et al. 2007).

The determining factors for Oklahoma's communications choice were "statewide coverage, reasonable cost, high reliability, a full two-way link, and moderate bandwidth" (Brock et al. 1995). Satellite communication via GOES[2] was considered, but did not pass Oklahoma's requirement for two-way communications. Two-way operations allow OCS operators to retrieve missed data from sites and to perform administrative functions such as setting clocks or uploading new datalogger programs (McPherson et al. 2007).

In addition to the OLETS-based system, Oklahoma also uses two small 900-MHz spread-spectrum radio systems. One system operates in an area where transmission difficulties have been experienced; the other is used for direct data transfer – bypassing OLETS – for stations in line of site of Sarkey's Energy Center, a 15-story building on the Oklahoma campus.

Oklahoma's field sites use Campbell Scientific dataloggers, either the CR10X-TD or CR23X-TD model, though an upgrade is reportedly in progress. At its operations

---

[2] Geostationary Operational Environment Satellite

center, Oklahoma utilizes eight computers for data collection via Campbell Scientific's (2009) LoggerNet software. Four machines are used as direct data collection servers, while an additional four machines are utilized for network administration[3]. In the Oklahoma design, no more than approximately 50 sites are assigned to each collection/administration computer pair due to performance concerns. Data observations and automated quality assurance flags are stored in NetCDF (UCAR 2009c) files (Wolfinbarger 2006).

2.5.1.4 Metadata databases

Oklahoma's instrumentation database contains information such as sensor serial numbers, locations, and operational status. For sensor calibrations, results of pre-calibration checks, instrument upgrades, and post-calibration checks are maintained. A "residence-time" reporting system, which records how long a particular sensor has been in a particular location, is also a part of Oklahoma's database system (Brock et al. 1995; Shafer et al. 2000).

Oklahoma's database system also serves as the primary engine for a sensor trouble ticketing system, which is used to report and record sensor problems and resolutions. The trouble ticketing system utilizes a web-based front-end. Trouble tickets include information such as station, parameter, problem description, date/time of problem onset, etc. (Shafer et al. 2000).

Oklahoma currently utilizes a pair of database servers for instrument, calibration, and maintenance tracking purposes. It previously utilized an Oracle database solution but

[3] administration of the sensor network, not the computer network

has migrated to a MySQL solution (Wolfinbarger 2006). For organizational purposes, its database is separated into four interrelated components: a user module, a network site module, an equipment module, and a quality assurance module (McPherson et al. 2007).

2.5.1.5 Quality assurance and control

Brock et al. (1995) holds that data faults must be detected rapidly and corrective action must be initiated in a timely manner in order to maintain high data quality. Oklahoma's quality assurance (QA) system is designed to never alter recorded data but to set 'status bits' indicating suspected data quality issues. Flagged data are available for research purposes but are not generally available for operational use.

Oklahoma utilizes four distinct types of methods in its QA process: laboratory calibration and testing, field intercomparison[4], automated routines, and manual inspection. Automated flags are set using a three-step process that applies filter checks such as those that indicate the presence of a technician on-site, more robust statistical algorithms, and a "decision maker" step which sets the final QA flag. General bounds and integrity checks are applied to data as they are received, while more intricate step, persistence, spatial, and like-instrument comparisons are applied on a delayed basis. For real-time data, up to eight quality control (QC) tests are run per observation and completed within one minute of data receipt. Up to 13 QC tests are run on each variable during delayed tests. Additionally, manual QC checks are recorded throughout the day and night by mesonet operators (Shafer et al. 2000; McPherson et al. 2007).

---

[4] Comparison with a special, reference remote station collocated with a mesonet site. The reference station contains higher quality instruments (Brock et al. 1995).

Oklahoma's data are all tagged with one of four QA flags: "good", "suspect", "warning", or "failure". Only "good" and "suspect" data are delivered in real time to users. "Warning" and "failure" data are withheld from public display (McPherson et al. 2007).

Like the Kentucky Mesonet, Oklahoma has a mission to share its data with multiple federal, state, and local government interests as well as with private users (Brock et al. 1995). Datasets are made available to researchers through multiple methods, including via ftp, CD, and DVD. For public use, Oklahoma's website (http://www.mesonet.org) is the main source for both observations and information about the network. Specialized web products have also been developed for media, public safety, and other interests. As part of its data dissemination efforts, Oklahoma developed WeatherScope, a custom data visualization software package, which can be used to display weather and geographical information from sources within and outside Oklahoma (McPherson et al. 2007). Oklahoma uses multiple servers to support data access and dissemination, including two machines for product generation[5], two load balancers, and four machines for web serving (Wolfinbarger 2006).

2.5.1.6 Availability and reliability

Brock et al. (1995) notes that "reliability is absolutely critical for a system that supplies real-time data for emergency management." Redundancy has always been evident in Oklahoma's computing infrastructure. The original VAX machines were designed in a paired configuration with common disk files and failover capability.

[5] such as GIF images, data graphs, meteorograms, etc.

Today, Oklahoma's infrastructure depends on pairings of primary/backup computers for critical functions. The network's operations center continuously monitors incoming data and systems for irregularities. Additionally, monitoring servers are utilized to automatically check the health of Oklahoma's networks, both computing and sensing, and to issue alerts if necessary (Wolfinbarger 2006; McPherson 2007).

2.5.2 United States Climate Reference Network

2.5.2.1 General description

The U.S. Climate Reference Network (USCRN), operated by NOAA's National Climatic Data Center (NCDC) in Asheville, NC, is a national surface monitoring program aimed at providing long-term, high-quality climate observations – especially for air temperature and precipitation – over the next 50 to 100 years (Hubbard et al. 2005). "The objective of the United States Climate Reference Network is to measure, record, and report with the highest possible quality a thoroughly documented set of surface environmental observations, representative of the climate of the United States" (NOAA 2003).

Plans for USCRN include approximately 300 locations (NOAA 2003). As of 2006, there were 80 operational sites (Phillips 2006), but that number has recently increased[6] to 114. While primary measured parameters are air temperature and precipitation accumulation, secondary measured parameters include wind speed, solar radiation, and ground surface skin temperatures (NOAA 2003).

[6] per check of USCRN website at http://www.ncdc.noaa.gov/crn/

While USCRN’s mandate to operate a climatological, research quality sensing network is similar to that of the Kentucky Mesonet, it is important to remember that USCRN does not operate as a real-time, operational meteorology network.

2.5.2.2 General computing system design

NCDC is charged with incorporating ingest, inventory, quality control, maintenance initiation, and long-term observation storage into its routine base of activities (NOAA 2003). One of the most important things to remember when reviewing USCRN’s computing infrastructure stems from this mandate; USCRN extensively leverages NCDC and NOAA’s existing computing infrastructure. For instance, NCDC’s existing database experts administer and backup USCRN’s databases on existing Oracle servers. Long-term raw observational data storage is incorporated into NCDC’s existing storage infrastructure. USCRN utilizes existing NCDC UPS, power, rack space, and bandwidth. Finally, NCDC leverages the existing telecommunications infrastructure at NOAA’s Silver Spring, MD headquarters for some data ingest functions (Phillips 2006; Hall 2006).

2.5.2.3 Data collection

NCDC is the central data collection facility for USCRN; it ingests and processes reports from all USCRN field sites (NOAA 2003). USCRN collects a data stream from each site once an hour. This stream includes the current hour’s measurements, plus a repeat of measurements from the previous two hours. This redundant data transmission is necessary as incomplete messages are sometimes ingested due to the use of one-way communications via GOES (Phillips 2006).

The one-way GOES communication from the sites causes several drawbacks (Phillips 2006; Hall 2006) including:

(i) Bandwidth availability limits the number of parameters measured.

(ii) Transmissions are limited to once-per-hour.

(iii) Remote data logger reprogramming or addressing is not possible.

(iv) USCRN does not directly control the entirety of its data handling process.

(v) Data not transmitted by a site in the three hour transmission "window" described above must be manually collected in the field.

NCDC uses three redundant ingest methods to help guarantee automated data receipt. After data are transmitted by the sites to GOES, they are received by NCDC via NOAAPORT/GTS, DOMSAT, and FTP from the National Weather Service's Telecommunications Gateway (Hall 2006).

The raw data stream from each sensing site is a sequential element list that includes a time/date stamp, a datalogger program version number, and a series of numeric observations. The program version number is needed to determine the number and order of observations in the data stream; the field dataloggers are not identically programmed and can transmit different measurements in a different order. Original USCRN data streams consisted of clear ASCII text. However, a transition to a binary format more conducive to compression is underway (Phillips 2006).

Datastream ingest and initial processing is carried out on an IBM AIX server (4 processors, 8 GB RAM, 2 TB local storage). Once processed, data are stored in an observational Oracle database running on a SUN server. Redundant ingest servers are located in Boulder, CO (Hall 2006).

Raw data received by NCDC are stored in its archives as the official sensor site's climatological observation record. Processing and normalized database storage are viewed as convenient data access facilitators (Phillips 2006).

2.5.2.4 Metadata databases

USCRN collects and stores metadata for all of its instruments. Equipment serial numbers, calibration history, failure reports, and maintenance records are all maintained in an Oracle database. The full complement of an instrument's metadata is stored once in the database. Metadata changes are then permanently stored on a change-by-change basis, allowing a full instrument metadata history to be constructed through SQL database queries (McGuirk and May 2003; Phillips 2006).

The official repository for USCRN sensing station metadata (latitude/longitude, elevation, site maintenance, etc.) appears to have undergone a series of transitions in the last few years. At the time of McGuirk and May (2003), the repository was an Oracle database named CRNSITES, but was being migrated to a new system called MI3 (Metadata Integration and Improvement Initiative). However, Phillips (2006) indicated a possible transition to an Oracle-based Integrated Station Information System (ISIS) currently under design.

2.5.2.5 Quality assurance and control

Philips (2006) indicated that the automated quality control processes applied to data directly by USCRN involve mainly basic range checks and field intercomparisons against collocated sensors. The range checks are made more robust by incorporating seasonal and regional changes for each parameter. These varying ranges are stored in a database, as are any QC "flags" generated by the checks. In consideration of disk space and database clutter, no flag indicating an observation's passing of QC checks is stored.

USCRN does not directly apply spatial QC checks to its data. However, NCDC applies spatial QC to a wide range of in situ data. It generates several gridded datasets for use in the QC process. The PrecipVal product, which is used for assessing precipitation data quality, integrates ASOS, radar, satellite, and model data into a single gridded product. A similar product exists for temperature analysis and a snowfall product is being developed (DelGreco 2006).

USCRN's quality control manager monitors data from all sites for potential instrumentation problems. Once identified, these problems are entered into an anomaly tracking system that is part of the network's metadata database (McGuirk and May 2003).

2.5.2.6 Data access

NCDC has a mandate to "provide timely access to the [USCRN] data, station history, and all other documentation to a worldwide clientele. All [USCRN] observational data, attached respective 'flags', metadata, and all documentation shall be posted to the web-accessible [USCRN] database for direct on-line access" (NOAA 2003).

NCDC operates a load balancer plus several public web servers[7] which allow access to USCRN data. Public data access is enabled by a read-only connection from the web servers to the USCRN Oracle databases. An interactive graphing application, built with Java and coded by Phillips (2006), provides data visualization. Tabular data are also available.

2.5.2.7 Availability and reliability

USCRN relies on NCDC's existing best practices and robust infrastructure for maintaining availability and reliability of its supporting computing systems.

2.5.3 West Texas Mesonet

2.5.3.1 General description

The West Texas Mesonet (WTM) is an in situ sensing network designed to provide "free, timely, and accurate" meteorological and agricultural data about the South

[7] See http://www.ncdc.noaa.gov/oa/climate/uscrn/

Plains/Rolling Plains region in western Texas. It is modeled after the Oklahoma Mesonet and consists of more than fifty[8] automated surface sensing stations, two atmospheric profilers, and one upper-air sounding system (Schroeder et al. 2005). WTM is operated by the Wind Engineering Research Center in the College of Engineering at Texas Tech University. Parameters measured by WTM's surface sensors (Schroeder et al. 2005) include:

(i) 10 m wind speed and direction,
(ii) 9 m temperature,
(iii) 2 m solar radiation,
(iv) 2 m wind speed,
(v) 1.5 m temperature and relative humidity,
(vi) rainfall,
(vii) leaf wetness,
(viii) and soil temperature and moisture.

2.5.3.2 Data collection

WTM collects meteorological data (air temperature, humidity, etc.) in real time every five minutes. Agricultural data, such as soil moisture content, are collected every 15 minutes. WTM's primary communications system is a project-developed Extended Line of Site Radio System (ELOS). Similar to the Oklahoma Mesonet, WTM attempted

[8] Per website, http://www.mesonet.ttu.edu

to gain access to its state's law enforcement telecommunications system, but was denied. Satellite communication was deemed unacceptable by the project (Schroeder et al. 2005).

WTM's ELOS includes antennae on 73 m towers and two antennae at the 61 m level for radio base stations. At the time of Schroeder et al. (2005), 10 out of 28 WTM stations utilizing ELOS also served as communications repeaters. Two additional communications repeaters – not collocated with a mesonet sensing station – were also a part of the network. The use of ELOS made communications signal strength a key WTM site survey condition.

As the WTM program progressed, the use of cellular telephone technology proved more useful. WTM found that cellular communications provides "acceptable bandwidth, short connection times, and affordable cost" (Schroeder et al. 2005). At least eight WTM stations utilize cellular communications. Other sites use regular phone and internet connections, if available.

2.5.3.3 Quality assurance and control

WTM's QA/QC tests are similar to those employed by Oklahoma. Initial tests are executed to flag suspicious or potentially bad data. A custom developed FORTRAN application is then utilized to apply Barnes analysis, range tests, step tests, persistence tests, etc. to the data. Similar to Oklahoma and USCRN, QA/QC flag information, raw, and corrected data files are separately maintained (Schroeder et al. 2005).

2.5.3.4 Data access

WTM's mission is to make all of its data freely available in real time via the internet. WTM utilizes web server(s) and product generation systems (like GEMPAK) to accomplish this mission. Additionally, Unidata's Local Data Manager (LDM) and standard internet file transfer protocol (FTP) are used to distribute data to other users (Schroeder et al. 2005). WTM's website provides tables of recent observations, as well as summary information. Time-series visualization is provided through the use of meteograms[9].

2.5.4 Meteorological Assimilation Data Ingest System (MADIS)

2.5.4.1 General description

The Meteorological Assimilation Data Ingest System (MADIS), operated by NOAA's Earth Systems Research Laboratory's Global Systems Division (ESRL/GSD), does not focus on direct operation of an in situ surface sensing network. Instead, MADIS acts as a data aggregator, collecting data from more than 150 separate surface networks in addition to other data from radiosonde soundings, aircraft reports, upper-air profilers, and both operational and experimental satellite observations and products. The goals of MADIS are "to promote comprehensive data collection and distribution of operational and experimental systems … and to make the integrated observations easily accessible and usable to the greater meteorological community" (Miller et al. 2007). Though

---

[9] These are provided by Oklahoma Mesonet-developed software.

already in extensive use, MADIS is still officially a research system. It will, however, make the transition to an operational NOAA/National Weather Service system in NOAA's 2010 fiscal year (Miller 2008).

2.5.4.2 General computing system design

In a manner similar to USCRN's utilization of NCDC computing resources, MADIS relies on the existing computing facilities of GSD's Information and Technology Services staff for the operation and monitoring of its ingest, processing, and distribution functions. MADIS utilizes a system of 21 computers, using Intel processors and the Linux operating system, to carry out its mission. Many of its servers are configured in 'high-availability' (HA) pairs (Miller and Barth 2003).

2.5.4.3 Data collection

MADIS data collection activities center around retrieval of data from the various networks integrated into the system. As Miller and Barth (2003) indicates, most data are retrieved from participating networks via the internet through an FTP or web server as simple text, often in a comma-separated-value (CSV) format. MADIS combines these data with observations from other providers, integrates them with NOAA datasets, and merges them into a uniform format consisting of standard observational units and time stamps (Miller et al. 2007).

2.5.4.4 Quality assurance and control

MADIS performs quality control checks on all incoming data and stores a series of flags indicating the results of these checks alongside raw data in its observational database. Static checks, which include single-station, single-time checks consisting of validity, internal intercomparison, and consistency checks, are applied every five minutes to incoming surface observations. Dynamic checks, run on a sub-hourly basis, include position, temporal, and spatial consistency algorithms (Miller et al. 2007).

Spatial consistency tests are performed using Optimal Interpolation (OI) techniques where differences in magnitude or other statistics are calculated for the same parameter from spatially related sites. If the resulting statistic falls outside of acceptable bounds, data are reanalyzed with a one-by-one elimination until the suspect data point is found and flagged (Miller et al. 2005).

Single character data descriptors for each observation, as well as an “overall opinion of the quality of the observation” are provided in MADIS’s integrated data sets (Miller et al. 2007).

2.5.4.5 Data access

GSD, formerly known as the Forecast Systems Laboratory (FSL), received funding in 1997 to build and implement the Local Data Acquisition and Dissemination (LDAD) system for the National Weather Service’s (NWS) Advanced Weather Interactive Processing System (AWIPS), which is the primary computing system used in

NWS weather forecast offices (WFOs). Therefore, MADIS data have always been accessible by the WFOs via LDAD in NetCDF format (Miller and Barth 2003; Miller et al. 2007).

MADIS data are available to non-NWS users via internet FTP, Unidata's Local Data Manager (LDM) software (UCAR 2009b), or through the use of web-based Open Source Project for Network Data Access (OPeNDAP) clients. MADIS also provides an Application Programming Interface (API) which hides the underlying NetCDF data format and allows users to read, interpret, and process the system's observations and quality control flags.

2.5.4.6 Availability and reliability

MADIS is concerned with and monitors both its internal computing processes and data streams from participating networks. High-availability computing pairs provide redundancy in the event of a computing failure, while each MADIS dataset is monitored with a "combination of automated and human operator procedures." When an incoming dataset has been unavailable for a sustained period of six hours, appropriate personnel at the dataset's owning network are notified via e-mail (Miller and Barth 2003).

2.5.5 MesoWest

2.5.5.1 General description

Similar to MADIS, MesoWest operates as a data aggregator for surface data in both the western and broader United States. Begun in 1994 as a collaborative effort between the University of Utah and the Salt Lake City National Weather Service forecast office, the network collects and integrates data from 47 public and 23 commercial sources (over 2800 stations), including ASOS observations (Splitt et al. 2002).

Per Splitt et al. (2002), the objectives of MesoWest are:

(i) to improve timely access to real-time weather observations for NWS operations,

(ii) to improve integration of observations for use in forecasting operations and verification,

(iii) and to provide access to data resources for research and education.

2.5.5.2 Data collection

MesoWest retrieves data from participating networks via the internet using FTP, web retrievals, or Unidata's LDM software. Data collection is scheduled every 15 minutes and is managed by a "master script" which controls data ingest, insertion of data into a database, and graphics generation. MesoWest experiences some significant average delays for some datasets, such as 74 minutes for data from the SNOTEL network. These delays are usually due to configuration choices made by individual networks, not by MesoWest (Splitt et al. 2002).

Once data are received, they are stored in a MySQL (Sun 2008) relational database whose table schema are designed with consideration of measurement type. As an aggregator for different networks, MesoWest must deal with the potential for different types of sensors for each measured parameter, such as unheated tipping buckets or weighing gauges for precipitation measurement. Its database is designed to handle these differences (Splitt et al. 2002).

2.5.5.3 Metadata database

MesoWest stores metadata information alongside operational data in its observation database. Minimum metadata requirements are station name, latitude, longitude, elevation, parameter type, and measurement units. Additional metadata are stored for many stations in northern Utah (Splitt et al. 2002).

2.5.5.4 Quality assurance and control

MesoWest applies real-time quality control to incoming data, assigning a “good”, “caution”, or “suspect” flag depending on algorithm results. Interestingly, MesoWest applies this flag to the entirety of an observation set and all of its data (temperature, relative humidity, etc.), not just to a single measured parameter. As the network admits, this can be problematic and can cause good data to be discarded (Splitt et al. 2002).

2.5.5.5 Data access

MesoWest data are available primarily via the internet through web, FTP, and Unidata LDM technologies. Data from individual stations are available upon receipt. Text summaries, time series, and spatial maps are also available. Data are disseminated to NWS offices in the western region through the region's Wide Area Network and are also made available to MADIS, described in the previous review (Splitt et al. 2002).

2.6 Summary

Meteorological and climatological sciences are based on measurements and observations. The United States has a long, rich history of operating quality in situ surface sensing networks, both at the federal and nonfederal levels. The role of automated and centralized computing systems in the collection, correction, and dissemination of network data continues to increase. Therefore, best practices and standards for both computing and sensing networks must be considered in the design of supporting computing networks. An understanding of these practices, along with the history of U.S. observing systems, has proven crucial for the Kentucky Mesonet in building its own reputable network.

CHAPTER 3. INITIAL PLANNING AND EARLY DESIGN

Dewett and Jones (2001) stress that an important role of information technology in an organization is to make knowledge easy to communicate, assimilate, store, and retrieve. Fulfilling such a role cannot be approached haphazardly. As Brown and Hubbard (2000) caution, planning is an integral part of any successful in situ surface sensing network. That planning process must include a network's critical information technology functions. Martin (2006) notes a mature, well-planned approach to project architecture along with a detailed and early awareness of a project's difficult hurdles – including requirements for high performance, reliability, and security – are critical factors associated with IT project success. Core functionality for the Kentucky Mesonet's computing and communications infrastructure was planned early in the network's development (Grogan 2007). This chapter reviews key requirements – both Kentucky Mesonet-specific and those common to in situ surface sensing networks – that drove the planning process. It also details some early design choices recommended in and ultimately resulting from the plan, including some of their results.

3.1 Kentucky Mesonet-specific requirements

The first and most obvious challenge for the Mesonet's computing architecture is that it must help the network fulfill its mission established by the state legislature as the official source of climatological observations for the state (Kentucky Legislature 2006). The most critical requirements, though, tie back to the network's dual-purpose nature. Operational users of the Mesonet – including the National Weather Service, emergency managers, broadcast meteorologists, and the general public – need continuous, near-real-time access to the data being collected; this requires a robust computing operation. Research users, on the other hand, require that data collected by the Mesonet be well documented and that they be subjected to a quality assurance and control process.

Other network-specific requirements identified at the outset of the planning process included:

(i) Data collection interval – The Kentucky Mesonet's computing and communications infrastructure must be capable of collecting and processing data of five-minute temporal resolution. The infrastructure should be able to collect and process these data in near-real time within fifteen minutes of parameter measurement;

(ii) Availability – Use of the Mesonet by emergency managers, the National Weather Service, and other critical decision makers requires that the Mesonet's computing and communications infrastructure be as continuously functional as possible;

(iii) Outreach – The Mesonet's computing infrastructure must accommodate the network's outreach mission, which includes facilitating data use by data partners & agricultural interests and enhancing educational experiences through student engagement and research opportunities;

(iv) Consortium Access – The Kentucky Mesonet has built a consortium of interested higher education users across the Commonwealth. Its computing infrastructure must support data access by consortium users;

(v) Compatibility – The Kentucky Mesonet desires to be compatible with other regional and national in situ networks. Its computing infrastructure must aid compatibility with those networks;

(vi) Revenue source – The computing infrastructure should support a variety of possible revenue sources, including for-fee data access by unaffiliated research interests; custom, value added environmental network hosting and modeling systems for consortium or external interests; and contract & freelance work for similar networks with a communications network expandable outside of Kentucky;

(vii) and Centralized operations – Mesonet employees do not share common office space. The computing infrastructure must support the ability to store common data so they are readily accessible by all program employees.

3.2 General in situ network requirements

Though Kentucky-specific needs were a big factor in development of the initial information technology plan, the scientific literature and personal conversations detailed in Chapter 2 played a large part in the Mesonet's IT architecture plan. From these references a set of core IT requirements for supporting an in situ surface sensing network were determined, including systems for communications, data ingest, observational data storage, metadata, quality assurance and control, data access and distribution, availability assurance, and ancillary functionality.

3.3 Initial information technology architecture plan

As Zachman (1987) and Martin (2006) note, there is not a systems architecture but a set of additive and complimentary architectures that spring from an integrative perspective covering both component level and application level details. The initial Mesonet information technology architecture plan was developed with this perspective in mind and took into account both Kentucky-specific and general in situ sensing network requirements. That plan (Grogan 2007) made several functional system recommendations, many of which are detailed below.

3.3.1 Site communications architecture

While it recognized that satellite data collection via GOES could extend communications across the country, the IT plan noted that experiences of the USCRN showed the method would conflict with data collection timeliness requirements, primarily due to one-way transmission limitations. While two-way satellite communications options were noted as being available, initial discussions seemed to indicate they would not be economical. Use of Kentucky's law enforcement and/or emergency management telecommunications system was considered but found initially to be contrary to the Mesonet's requirement that communications be expandable outside of the state. Furthermore, the system – the Kentucky Emergency Warning System – was in the process of being upgraded which would have complicated its adoption by the Mesonet. Finally, direct, hardline internet connections or phone connections to individual stations were considered too unwieldy, in terms of managing multiple connections, to be useful.

To facilitate data collection requirements, the initial plan recommended use of a commercial cellular communications platform for secure, two-way communications between Mesonet computing systems and dataloggers at remote sites. Experimentation with the cellular platform had started at the Mesonet before the author's employment and full adoption was recommended by his initial plan, though the door was left open for possible use of alternative communications methods at "cellular-poor" but "climatologically-rich" sites. Chapter 4 of this document is devoted exclusively to communications.

3.3.2 Site survey database

At the time of plan development, Mesonet graduate students were busy canvassing the commonwealth in search of climatologically-suitable sites on which to locate Mesonet stations. The plan recommended development of a site survey database to track and display their findings and called for tracking of geographic site characteristics (latitude, longitude, etc.), site scoring information, site contact information, site photographs, and other digital files (spreadsheets, documents, etc.) related to each site.

3.3.3 Metadata database

The initial plan called for development of methods to track a number of metadata. A site database to track information about operational Mesonet sites such as maintenance and environment changes was included, as was the ability to track instrument information, calibration, and relocation. A trouble-ticketing component was also included.

3.3.4 Data ingest systems

The Kentucky Mesonet had already adopted Campbell Scientific's datalogger platform prior to plan development. Therefore, the plan recommended that Campbell Scientific's LoggerNet software suite (Campbell Scientific 2009) be used for remote data collection. The experiences of the Oklahoma Mesonet (Wolfinbarger 2006) indicated that data ingest by LoggerNet could not be accomplished through the use of a single server. Therefore, multiple servers were recommended to carry out this mission. However, a single ingest server has thus far proven sufficient.

3.3.5 Observational data storage / database system

The original architecture plan recognized the Kentucky Mesonet's mission to develop a long-term, research quality climatological dataset and recommended an observational database system be developed to handle storage of and access to program data. Specifically, the plan recommended a system that would, at minimum, facilitate easy data storage & recall; storage of raw, unaltered data as received from field sites; and storage of quality assurance / quality control flags. It recognized design questions to be answered for the system, including:

(i) Will raw data be stored in flat files or directly in a database structure?

(ii) Will a relational database be used for all data storage purposes, or will some other form of meteorological data storage be used?

(iii) How much data will need to be stored?

(iv) How will data tables and schema be normalized?

3.3.6 QA / QC system

Recognizing that in situ network best practices and guiding principles showed a definite need for a sufficient quality assurance / control system, the plan recommended implementation of a suite of automated quality control analyses and statistical techniques which it indicated could require an extra level of robustness from the supporting computing infrastructure. It called for the ability to handle both automated and manual data quality flags in a database or other data access system and for the flags to be easily relatable to the observations they describe.

3.3.7 Data access systems

Understanding that network data would be made available to a wide range of users including the general public, the initial IT plan recommended a number of data access systems be developed that would tap the obvious ubiquity of commodity internet access for data delivery. It called for web server(s) and server software to provide data via the web and recognized the importance to security of segregating public access systems from critical project systems. Included in the data access systems plans were recommendations for product generation system(s) to handle “heavy lifting” of tasks such as dynamic image creation, meteogram development, and other computationally intense tasks. The plan recognized the need for internal data access systems for Mesonet employees and the need for specialized external data distribution systems for key partners.

3.3.8 Availability assurance systems and methods

Recognizing that basic computing network design principles (Murhammer et al. 1999), continuity of purpose principles (NRC 1999), best practices of other in situ networks, and self-imposed goals all require a high level of data and systems availability, the preliminary IT plan stressed the development of availability assurance systems and other methods to ensure that operational data are continuously available and that the climate record from past observations is protected from loss. It called for the monitoring of critical computing infrastructure using specialized tools able to notify computing systems administrators and other Mesonet personnel in the event of critical failures. Similar systems were recommended for monitoring in situ sensing sites for critical sensor, instrumentation, and communications failure. Regular backup procedures were also stressed.

3.3.9 Ancillary systems

In addition to the key operational systems detailed above, the initial IT plan also recommended a number of supporting systems, including:

(i) a concurrent versioning system to maintain and retain the critical computing and datalogger code base,

(ii) a time server to synchronize and correct the time of all network dataloggers and computer servers to a reference traceable to the National Institute of Standards and Technology,

(iii) development servers for non-operational research & development use among developers and researchers,

(iv) a name server to provide domain name to IP address resolution,

(v) and a map sever to host and serve data for a program-developed geographic information system.

It also recommended that direct support for program-owned desktop computers used by Mesonet employees be provided directly by WKU’s information technology department instead of by Mesonet IT personnel.

3.4 Early design decisions

Key pieces of Kentucky Mesonet architecture articulated in the initial information technology plan were developed over a span of time covering approximately the last three years and are given detailed treatment in subsequent sections of this document. However, some important decisions were made early in the life of the Mesonet, either at the recommendation of the initial plan or in the process of developing it and are not covered elsewhere. Some of these decisions and some of their resulting consequences are detailed below.

3.4.1 Mission-critical / enterprise-grade approach

Given the important nature of some uses of the Mesonet, especially in operational settings, a mission-critical approach was adopted early in the design process. Within budgetary constraints, all Mesonet servers and systems used in the collection, storage, processing, and distribution of operational data are enterprise-grade, complete with redundant power supplies, redundant storage (RAID), and critical support plans with four-hour vendor technician response times. Though financial and space limitations have precluded the purchase and operation of fully redundant systems – meaning a one-to-one spare for each server – a single spare server and communications router were purchased for standby. Finally, for most systems, fully licensed and supported enterprise-grade operating systems were installed.

3.4.2 Network operations environment decisions

Again justified by critical uses of its data, stringent guidelines for the Mesonet’s network operations environment were adopted early on in the program. The initial IT plan called for the Mesonet to locate its computing infrastructure in a network operations center that exhibits as many of the following qualities as possible:

(i) Provides high quality internet access and available bandwidth, with at least two separate paths to the internet backbone

(ii) Provides 24 x 7 x 365 network or facility failure resolution, with any on-call response times no longer than 20 minutes

(iii) Provides emergency generator power capable of powering Mesonet computing infrastructure for a minimum of a 3-week period. Facilities with natural gas or other continuously-fueled generators are preferred

(iv) Provides proper ventilation and cooling to Mesonet computing systems

(v) Provides only secured, verified physical access to Mesonet computing systems

(vi) Allows 24 x 7 x 365 physical access for Mesonet computing administrators

(vii) Provides a dedicated block of static IP addresses to the Mesonet and allows the Mesonet full name resolution control over those addresses, including its own domain names such as kymesonet.org and others

(viii) Is located within the Commonwealth of Kentucky, though an out-of-state backup facility should be considered

With the Mesonet office space obviously meeting few if any of the network operations center requirements, the initial desire was to host the network's computing systems in WKU's primary campus data center. Though WKU's campus information technology leadership were supportive in initial discussions, they were hesitant that the academic nature of the data center would be able to support the Mesonet's mission-critical requirements. Indeed, the fact that the academic data center and/or computing network have been taken completely offline in planned outages – usually during holidays or other academic breaks – five days since Mesonet inception and for a half day of unplanned downtime shows that not locating in the academic data center was a wise choice. Such a statement is not a negative commentary on WKU's IT division; it just illustrates the different needs of academic and 24 x 7 mission-critical systems.

Instead of being located on campus, the Mesonet maintains a contract for server co-location and internet connectivity with Bowling Green Municipal Utilities (BGMU), which operates a fully redundant municipal fiber optic network. BGMU's network provides connection to multiple internet backbone providers and supports both critical municipal public safety interests and commercial operations. Nine rack units, or 9U, of space (Figure 3-1) and 2 Mbps of symmetrical commercial-grade internet service are leased in BGMU's access-controlled, generator-supported, fire-suppression-equipped network data center.

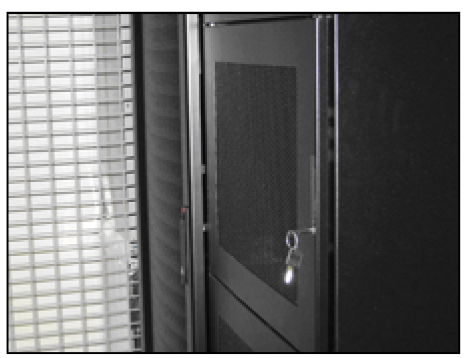


Figure 3-1. Rack space rented from Bowling Green Municipal Utilities. (Source: BGMU 2010).

The reliability of power and connectivity at BGMU has been far better than that which could have been achieved in WKU's campus data center. Based on automated external monitoring services employed by the Mesonet, for the one year period ending on 28 February 2010 the Mesonet's primary external systems were unreachable for a cumulative total of 121 minutes, yielding an uptime percentage of 99.977%. That unreachable amount includes all times when either BGMU or Mesonet systems were unavailable due to both scheduled and unscheduled downtime. Since BGMU-caused outages are not distinguishable in the monitoring service data, it should be noted that BGMU's uptime percentage likely well exceeded 99.977% in the period, as some downtime was certainly due exclusively to Mesonet-related issues. The Mesonet's contract allows program computing systems to be relocated to a new WKU-owned, BGMU-managed commercial data center being developed on WKU's research and development campus.

3.4.3 Operating system choice

As desktop computers other than development machines are supported by university IT personnel, Microsoft Windows (Microsoft 2007) operating systems are used on them. However, Linux is the operating system choice for both operational and developmental servers, systems, and hosts. Several factors played into this decision:

(i) a widespread use of Linux systems in meteorological settings, including for the main NWS computing system, the Advanced Weather Interactive

Processing System, and in the Kentucky Climate Center's Climate Research Laboratory;

(ii) a lack of Windows support for many meteorological applications such as Unidata's GEMPAK and Local Data Manager software (UCAR 2009a, b);

(iii) the experience of the author, who has an extensive Linux server skill set;

(iv) and the desire to use open source software whenever possible, of which a large amount is available for Linux systems.

For the majority of systems, RedHat Enterprise Linux (RedHat 2008) is used. However, CentOS (CentOS 2008), a binary equivalent derivative of RedHat Linux, is used on some systems, especially development hosts.

3.4.4 Time considerations

For a meteorological observation network, time and timestamps are obviously important considerations, especially when sites are split across a time zone boundary. Figure 3-2 shows Kentucky Mesonet sites and the boundary between Central and Eastern time zones. Note that the site in Taylor County, which appears to be directly on the boundary, is about 1,000 meters into the Eastern zone.

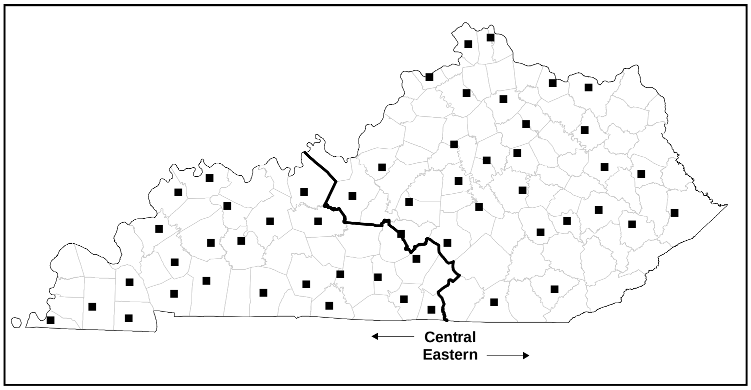


Figure 3-2. Kentucky Mesonet sites and time zone boundary. Sites are those operational, currently planned, or under construction and are shown as points. Boundary between Central and Eastern Time shown as solid black line.

The official time of the Kentucky Mesonet is Coordinated Universal Time (UTC). All field dataloggers and servers are set to UTC. While using UTC helps establish a common time across locations, certain data – especially climate data – must have a reference to local time, both that advanced for Daylight Saving Time (DST) when applicable and that never advanced for DST. As is discussed in Section 6.4.2 below, the Mesonet observation database stores all three types of timestamps, which makes querying by time much easier, and the network's code library understands and handles data requests in all three. Except where specifically noted, UTC should be assumed for all times and dates referenced in this document.

3.5 Discussion

As expected with any infrastructure implementation, a few changes to the original information technology goals have been made over the last few years. However, the original plan and early design decisions remain pretty well in force. Figure 3-3 depicts the general design of the Mesonet architecture in terms of data flow, including site communications, centralized data operations, and external data distribution mechanisms. Though the simplified diagram depicts mainly physical components in both the sensing and computing networks, it should prove a useful reference for the remainder of this document.

With finite staff resources, all Mesonet computing systems, of course, have not been simultaneously implemented. Instead, Trenberth et al.'s (2002) step-by-step priorities for in situ network design have been followed, with the Mesonet having reached at least level 4 of those goals; the follow-on processing priority to produce analyses and reanalyses remains. However, Dewett and Jones' (2001) views regarding first- and second-order IT-related learning by an organization certainly seem to apply to the Mesonet, as there is a need to transition to 'second-order' learning, where technologies are modified to better match the organizational environment.

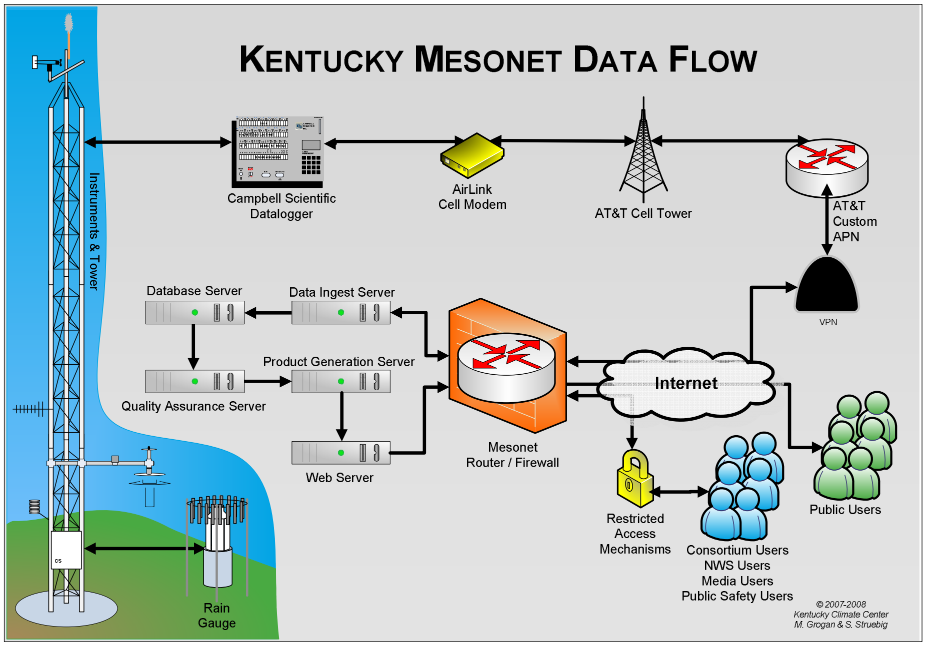


Figure 3-3. Kentucky Mesonet data flow. Simplified data flow diagram typically provided in public documents and presentations.

As the Introduction promised, the remainder of this work focuses on a snapshot of where Mesonet systems and technology stood before an arbitrary academic deadline. That snapshot certainly reveals that a substantial amount of progress has been made toward meeting the goals of the initial IT plan. In some areas, though, it also reveals where the Mesonet has admittedly stumbled and makes recommendations for improvement. The next chapter is completely dedicated to site communications technology, while the subsequent four chapters examine computer coding decisions then mission-critical, geographic information, and ancillary systems, respectively.

CHAPTER 4. COMMUNICATIONS

About 10 years ago as part of a workshop on automated weather stations (AWS) for applications in agriculture and water resources, a working group (Horton et al. 2000) developed a comparison table for communications, shown below in Table 4-1. Of interest is that direct internet connections were not included in the table, though as shown in the Literature Review the method has become a viable option. As discussed in Section 3.3.1, though, direct internet connections, phone line connections, and VHF/UHF (state systems) were rejected as an initial communications choice for the Kentucky Mesonet. From the start, the Mesonet desired a full two-way communications method that would be easily deployable, that would minimize field technician maintenance requirements, and that would keep the number of external communications-related contacts small.

Table 4-1. Horton et al.'s (2000) automated weather station communications comparison.

	Phone Line	Short Haul	Cell Phone	GOES	Meteor Burst	Spread Spectrum	VHF UHF
Skills needed	low	low	med	high	high	high	high
Affected by land topo.	low	low	high	low	low	high	high
Affected by vegetation	low	low	high	low	low	high	low/med
Communication dist.	high	low	high	high	high	low	low
Base station	no	no	no	no	yes/no	yes	yes
Capitol cost	low	low	med	high	high	med	med
Operating cost	low	low	variable	variable	variable	low	low
Power	low	low	high	high	high	high	high
Possible access rate	high	high	med	low	low	high	high
Data throughput	high	high	med	low	low	med	med
2-way communication	yes	yes	yes	no	yes	yes	yes
Stable technology	yes	yes	no	yes	yes	no	no
Affected by population	low	low	high	low	low	low	low
License required	no	no	no	yes	no	no	yes/no

After weighing its options, the Mesonet chose a cellular-based communications method. The communications comparison shown in Table 4-1 suggested that cell-based communications was an unstable technology in 2000. For the same workshop that generated the table, Grant and Toby (2000) analyzed cell-based communications and found the following advantages:

(i) maximum flexibility in locating stations;
(ii) minimal risk from mechanical damage due to farm machinery;
(iii) minimal risk of lightning strike damage;
(iv) minimized costs of installation at locations distant from existing phone lines;
(v) and minimized costs of moving sites due to changing farm/researcher needs.

Disadvantages were found to be:

(i) service being limited to regions with cell towers;
(ii) relatively high power needs (2.15 A during transmission);
(iii) relatively low data transmission rates;
(iv) and rapidly changing technology.

An analysis of cellular-based data retrieval for the Arizona Meteorological Network (Brown et al. 2000) noted similar challenges; power consumption for that network's communications devices was around 1.2 A. Most importantly, both Grant and Toby (2000) and Brown et al. (2000) noted a large hurdle in getting cellular-based communications providers to fully understand and adequately support AWS needs.

The remainder of this chapter is dedicated to a review and analysis of the use of cellular-based communications for the Kentucky Mesonet over the past three years. It shows that, while some of the support-related headaches remain, cellular-based communications is now proving to be a decently stable data retrieval method. Power consumption requirements have improved dramatically with a change from analog to digital transmission and, taken on the whole, reliability percentages are respectably high.

4.1 Choice of cellular provider

Choosing the vendor for the Kentucky Mesonet's cellular connectivity was an admittedly straightforward and obvious process. As shown in Figure 4-1 AT&T, which had recently acquired Cingular wireless, had the largest licensed coverage area[10] in Kentucky in 2007, just as the Mesonet began constructing its communications and computing infrastructure and deploying its initial sites. After its acquisition of Cellular One in 2008, AT&T's licensed cellular coverage area included all of Kentucky. By choosing AT&T as its single provider, the Mesonet avoided the complications and confusion of having to deal with different vendors for different sites.

Fortunately, WKU already had an existing enterprise-level contract with AT&T for cellular services which the Mesonet was able to use for its in situ network needs. This has allowed the Mesonet to procure cellular service through the university's Communication Technologies department and to receive and pay for communications via existing internal billing systems.

[10] Licensed coverage area is not the same as service or signal availability.

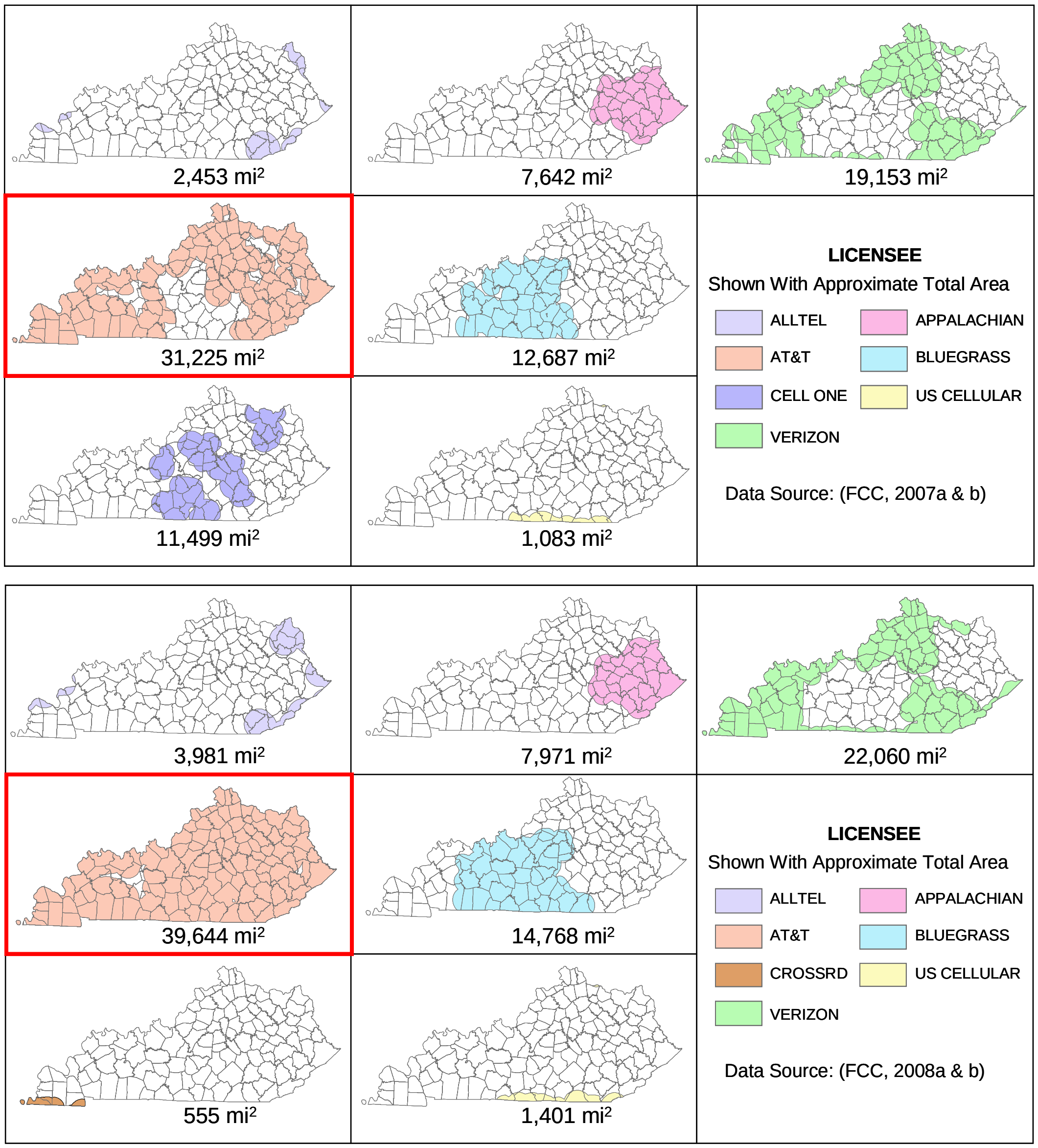


Figure 4-1. Cellular (non-PCS) licensed coverage area analysis. Top: 2007. Bottom: 2008. AT&T highlighted in red box. Data (FCC 2007a,b; 2008a,b) represent licensed coverage area, not signal strength.

4.2 Technical implementation

A number of key technologies and related implementation choices form the complete communications architecture used by the Kentucky Mesonet. These are described below.

4.2.1 Transport methodology

Figure 4-2 depicts in simple form AT&T's Commercial Connectivity Service (CCS), which the Kentucky Mesonet chose as the underlying supporting technology for its site-to-data-center data transport. Unlike consumer-grade connectivity options, CCS provides a method wherein Mesonet site communications devices can remain part of the program's <u>internal</u> computing network (AT&T 2005). The Mesonet's data ingest systems connect via an internet-transported Virtual Private Network (VPN) to an AT&T data center. A virtual routing instance – known as a Custom Access Point Name (APN) – segregates Mesonet data from other cellular data, ensuring privacy and security of the transport method.

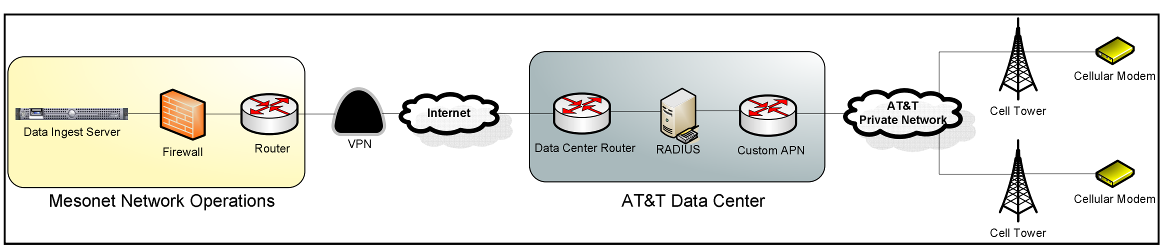


Figure 4-2. Simplified depiction of Kentucky Mesonet's use of AT&T's Commercial Connectivity Service.

As the CCS/APN solution allows Mesonet sites to be an extension of the program's internal communications network, private non-routable IP addresses are assigned to the cellular modems at each site. Usernames and passwords for a total of 2,046 usable addresses were generated by the Mesonet during the initial CCS provisioning process. An AT&T-hosted Remote Access Dial In User Service (RADIUS) is used to dynamically assign persistent IP addresses to the devices upon initialization or reset. The use of private addresses with the CCS and APN allow Mesonet servers to initiate communications with the sites – which is the standard procedure for the Mesonet – and vice-versa. Had public addresses and/or consumer-grade connectivity options been used only site-initiated communications would have been supported, essentially making full two-way communications not possible.

Though the technical process of CCS and APN setup was relatively smooth, some customer service-related aspects were somewhat lacking. The standard imposed waiting period – 84 days – between paperwork completion and service provisioning seemed rather high. Had connectivity testing on the magic 84th day failed, provisioning may have been delayed a month or two more. Also, account representatives assigned to service WKU's contract were somewhat unfamiliar with the CCS and APN technology. Questions about the technology were referred to others in the company.

4.2.2 Device choice

At the recommendation of Campbell Scientific, the Mesonet's datalogger manufacturer, cellular data communications devices (Figure 4-3) from Sierra Wireless[11]

[11] formerly AirLink; purchased by Sierra Wireless

were chosen for use at remote sites. Thirty-five Mesonet sites use the AirLink Raven EDGE E3214 modem, which was discontinued in November, 2008. Remaining sites use its replacement, the AirLink Raven XT G2212-C.

Figure 4-3. Kentucky Mesonet cellular data communications devices from Sierra Wireless. Left: AirLink Raven EDGE E3214. Right: AirLink Raven XT G2212-C. (Photo source: Sierra Wireless 2010a, b).

Though they share underlying technology with other AirLink Raven devices, the E3214 and G2212-C models have radio modules specifically designed to communicate with AT&T's [12]GSM-based EDGE network, which "provides end-to-end packet data services with an enhanced connectivity building on [13]GPRS technology" (Sierra Wireless 2008). EDGE technology – commonly referred to as 2G when referring to AT&T – facilitates transmission speeds up to 384 kbit s^{-1}.

Unlike their analog predecessors referenced in Grant and Toby (2000) and Brown et al. (2000), a major advantage of the digital EDGE devices is their power consumption. Instead of needing 1.2 – 2 .5 A during transmission, the E3214 and G2212-C typically require only 250 and 350 mA, respectively (Sierra Wireless 2008, 2009a). The lower power requirements allow the devices to fit well within the Mesonet's site power budget.

[12] Global System for Mobile Communications
[13] General Packet Radio Service

Communications with site dataloggers is accomplished via the modems' serial server technology, which essentially exposes the loggers' RS-232 connection as a TCP/IP port reachable via the AT&T CCS. Campbell Scientific-provided configuration templates for the modems are customized with Mesonet-specific values – typically CCS and APN related – and are written to the devices via Sierra Wireless's AceManager utility (Sierra Wireless 2009b). One of the most important configurations applied to the modems is the Keepalive feature which is set to automatically reset the devices' radio modules after a 22-minute period of data throughput inactivity, but only if the devices are unable to communicate with Mesonet servers. The feature has proven invaluable at preventing technician truck rolls to reset "stuck" modems.

4.2.3 Domain name resolution

To simplify access to site modems by the Mesonet's data ingest systems, an internal domain name resolution service (DNS) was setup to map the IP addresses used with the CCS to domain names in the form *xxxx.sites.kymesonet.org*, where *xxxx* is a four-letter identification abbreviation assigned to each site.

4.3 Signal strength and site selection

Availability of a usable AT&T cellular signal is an important factor in the placement of Mesonet sites. Some concepts related to this factor, including Mesonet site evaluation processes, are discussed below.

4.3.1 Signal strength concepts

The primary measure of cellular signal strength on the Sierra Wireless AirLink devices is the RSSI, or Received Signal Strength Indicator, value which is measured logarithmically relative to one milliwatt and reported in [14]units of dBm. Reported as a negative number, values closer to 0 indicate a stronger signal. An RSSI of -51 is ten times stronger than an RSSI of -61 dBm. The manufacturer recommends an RSSI between -60 and -80 dBm (Sierra Wireless 2009c), which generally holds well with Mesonet experience. It is important to remember that RSSI indicates <u>received</u> signal strength at the modem, not the cellular tower. The limited power of the digital modems can prevent a usable signal from making it back to the tower from the modem.

4.3.2 Site survey process

Met / climate sensing sites should, ideally, be chosen solely based on their suitability for that purpose. Reality, however, dictates that resource factors including communications availability play a part in the selection process. The Kentucky Mesonet is no exception. Therefore, analysis of RSSI values at potential sites has always been an important part of the network's site survey process. Site surveyors, typically Mesonet student research assistants, record RSSI values in eight passes around an eight-point compass with a directional antenna, as shown in Table 4-2.

[14] The value is really a unitless proportion.

Table 4-2. RSSI values (dBm) from site survey at the Columbia Transpark in Adair County. (Source: Ramsey Quarles).

	N	NE	E	SE	S	SW	W	NW
1	-71	-85	-71	-71	-71	-71	-71	-83
2	-81	-81	-69	-69	-69	-81	-67	-67
3	-67	-79	-79	-63	-63	-73	-73	-73
4	-73	-73	-73	-83	-73	-85	-75	-75
5	-65	-87	-65	-65	-65	-75	-75	-65
6	-65	-81	-81	-81	-69	-87	-75	-75
7	-75	-75	-75	-75	-75	-75	-63	-63
8	-63	-81	-77	-77	-87	-77	-65	-65
Average:	**-70**	**-80**	**-74**	**-73**	**-72**	**-78**	**-71**	**-71**

4.3.3 Marginal signals

Before investing time, money, and effort to place a Mesonet site in a spot whose cellular signal may not sustain operations, a more in-depth signal test is sometimes conducted over time in the proposed location. Such was the case with a proposed site near Harlan County's Pine Mountain in southeastern Kentucky. Early analysis of the site, including difficulty in placing cellular voice calls, indicated it was questionable at best in terms of signal. Figure 4-4 shows the difficult terrain surrounding the proposed site, with the nearest AT&T cellular tower over 13 km down the valley to the southeast along state highway 221.

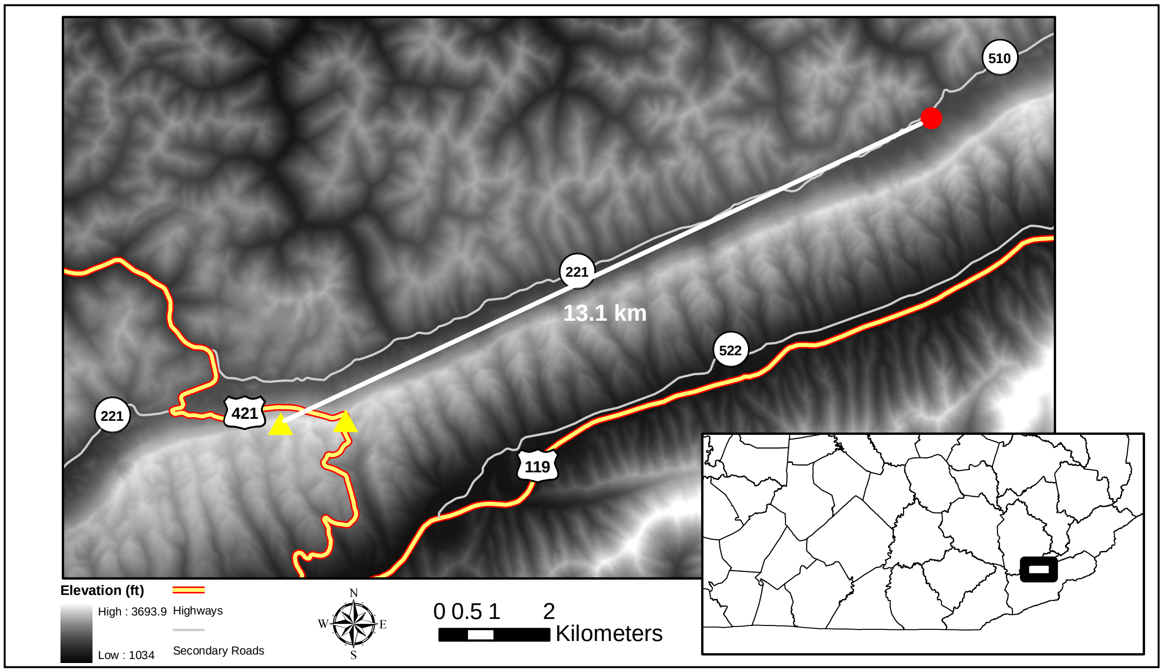


Figure 4-4. Proposed Kentucky Mesonet site (red dot) near Pine Mountain in Harlan County. Cell towers are yellow triangles, with AT&T's the furthest west. Elevation and highway data from KY Division of Geog. Info. Tower data from FCC (2009).

To fully assess the site, a datalogger, AirLink Raven EDGE modem, directional antenna, and solar panel were placed at the site beginning 2 June 2009. A script was written and executed on a data ingest server to connect to the modem every five minutes and collect cellular diagnostic data, including RSSI. For the period lasting until 16 July 2009, RSSI values typically ranged between -83 and -95 dBm. However, there were also extensive periods – many lasting multiple hours – when the signal was too poor to support a connection to the modem.

In an attempt to "save" the climatologically rich site, a cellular amplifier was placed inline between the modem and antenna; RSSI improved on average by 10 dBm. Unfortunately, hours-long periods of modem inaccessibility continued and the 1 A power requirements of the amplifier proved too much for the solar panel and accompanying battery. Alternate communications methods, such as satellite, are being investigated for this site.

4.3.4 Antenna choice

Depending on signal strength, the Kentucky Mesonet uses either an omnidirectional (omni) antenna or a higher gain directional (yagi) antenna. Omni antennae are used at 60% of Mesonet sites; yagis are used at 40%.

4.4 Support hurdles

The Mesonet's relationship with WKU's Communication Technologies (CT) department is a strong one and is vital to the setup and operation of its field data devices. For quite some time, however, the Mesonet found that Grant and Toby (2000) and Brown et al.'s (2000) views in terms of difficulty in acquiring support held true with AT&T, seemingly from a lack of understanding on AT&T's part of the program's needs. In the last three years WKU has been assigned multiple primary account managers, each of whom have needed some "training" by the Mesonet regarding its data usage. For technical support purposes, WKU does qualify for AT&T's enterprise-grade high tier support services through the Mobility Enterprise Customer Maintenance Center (MECMC). Select Mesonet personnel are authorized by the CT department to directly obtain support from MECMC in critical situations; this has proven invaluable during critical outages – especially overnight.

The placement of the Kentucky Mesonet in a proverbial “academic box” due to its affiliation with WKU, though, seemed to limit the level and urgency of support available to the program, which led to a feeling of uneasiness among program principals in terms of having control over one of the most vital parts of the network. This issue became critical at a point when AT&T began a changeover of billing systems ahead of schedule that caused Mesonet devices to drop off the network one-by-one in the order they had been provisioned – from oldest site to newest site. This prompted WKU’s CT department to force AT&T to move Mesonet devices to a special, segregated, “do not touch” account.

Though the author had been querying WKU’s AT&T account representative for over a year about support concerns, no movement on AT&T’s part was seen until a high ranking AT&T executive was pressed by the author to help rectify the situation. This executive was able to arrange conference calls between critical Mesonet personnel and AT&T engineers. Most importantly, though, he was able to break the Mesonet out of the academic box by declaring it to be a public safety agency based on its critical work with the National Weather Service in the severe weather warning and verification process and its availability to emergency managers in other hazardous situations. That designation has been crucial at times in obtaining critical technical support responses.

4.5 Reliability & resources

While there have been some significant communications support hurdles to jump in the last few years, a by-the-numbers analysis shows that the reliability of the AT&T

cellular network for Mesonet data transport has been satisfactory. The results of several ongoing tests of resource use and reliability are provided below.

4.5.1 Data transfer

Instead of paying for more costly [15]unlimited data plans, for most sites the Mesonet opts for less expensive limited plans. To track data transfer usage, scripts are executed once every three hours to poll counters on data ingest system firewalls. Except for an approximate one-month period, each byte transferred between field sites and data ingest servers between July 2007 and the present has been counted. Based on analysis of those data for the one year period[16] ending 28 February 2010, an average of <u>6.46 MB</u> of total data transfer is needed in one month to collect 27 floating point values measured every five minutes and collected at least once every 15 minutes[17] via LoggerNet (Campbell Scientific 2009) server-initiated connections. Interestingly, due to TCP/IP and Campbell Scientific PakBus transmission protocol overhead and handshaking, similar analyses show a nearly threefold increase in required transfer for collecting the same amount of <u>usable data</u> every five minutes versus every fifteen.

---

[15] now typically capped at 5 gigabytes per month
[16] excluding November 2009, when data were accidentally not collected
[17] data used in calculating the average include times when the network was in “5 minute mode”

4.5.2 Signal strength

Utilizing methods similar to the Pine Mountain signal quality study described in Section 4.3.3, cellular diagnostic variables have been collected at least twice daily (8 and 20 UTC) from Mesonet field data modems since 29 June 2009. Variables include:

(i) channel – the cellular channel assignment;

(ii) RSSI – the received signal strength indication in dBm;

(iii) roaming – a Boolean value indicating if the device is “roaming” between cellular carriers;

(iv) cell ID – the identification number of the cell being used;

(v) and LAC – the location area code that, taken with cell ID, uniquely identifies a particular cell.

Figure 4-5 gives an analysis of those diagnostic variables, which were reviewed for each site for the period ending 28 February 2010. For the eight month period – or less for sites which came online more recently – each site’s average RSSI value, the number of cells (via cell ID and LAC) to which it had ever connected, and the cellular bands (via channel number) to which it had ever been assigned were analyzed. Of course, average RSSI is somewhat of a self-determined or -fulfilling value, as sites are purposely placed in locations with higher RSSI. As the Mesonet’s directional antennae are tuned for approximately 850 MHz, the band assignment of each site modem over time is important. The majority of the program’s devices have only ever operated on 850 MHz, though the number receiving 1900 MHz assignments has been increasing over time.

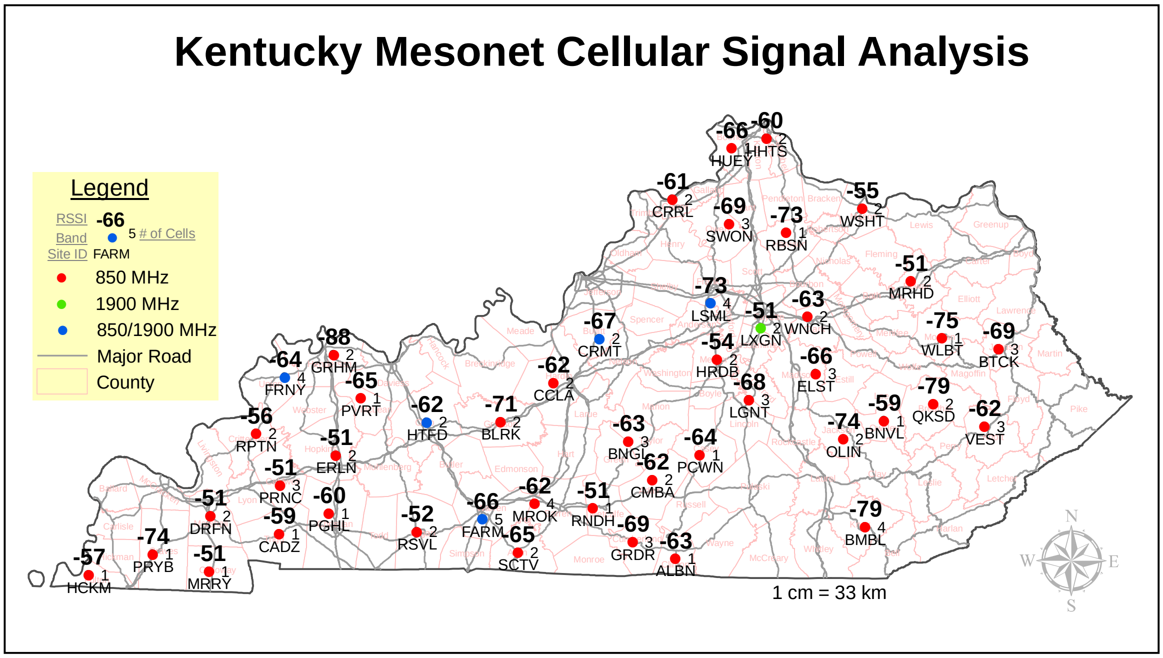


Figure 4-5. Kentucky Mesonet cellular signal data analysis for the period ending 28 February 2010.

4.5.3 Uptime availability

Perhaps the most useful and telling statistics concerning communications performance can be found in an analysis of site uptime or availability. A Mesonet site is considered to be "up" whenever a set of observations are available which are no older than 20 minutes. Mesonet availability assurance mechanisms are used to constantly track site uptime performance via the Nagios IT infrastructure monitoring platform (Nagios 2008). Table 4-3 provides site uptime information for all Mesonet sites for the 1-, 3-, 6-, 9-, and 12-month periods ending 28 February 2010. To be included in a particular period's statistics, a site must have been online for at least 67% of that period. Since site "GRHM" only recently came online, it is not included in the analyses.

Table 4-3. Kentucky Mesonet site uptime availability percentages for the 1-, 3-, 6-, 9-, and 12-month periods ending 28 February 2010.

SITEID	COUNTY	1 MONTH	3 MONTHS	6 MONTHS	9 MONTHS	1 YEAR
ALBN	CLINTON	100.000%	99.995%	N/A	N/A	N/A
BLRK	GRAYSON	99.828%	99.111%	99.428%	99.550%	99.592%
BMBL	KNOX	95.277%	97.724%	98.726%	99.078%	99.292%
BNGL	TAYLOR	99.988%	99.996%	99.681%	99.766%	99.504%
BNVL	OWSLEY	100.000%	99.973%	99.875%	N/A	N/A
BTCK	JOHNSON	100.000%	99.950%	99.886%	99.889%	N/A
CADZ	TRIGG	100.000%	99.996%	N/A	N/A	N/A
CCLA	HARDIN	100.000%	99.992%	99.858%	N/A	N/A
CMBA	ADAIR	100.000%	99.973%	99.664%	99.749%	99.813%
CRMT	BULLITT	100.000%	99.969%	99.514%	99.655%	99.741%
CRRL	CARROLL	94.929%	98.214%	98.683%	99.087%	99.142%
DRFN	MARSHALL	99.294%	99.742%	99.388%	99.520%	N/A
ELST	MADISON	99.938%	99.908%	99.759%	99.771%	N/A
ERLN	HOPKINS	99.715%	99.885%	99.781%	99.816%	99.861%
FARM	WARREN	99.864%	99.611%	99.695%	99.792%	99.771%
FRNY	UNION	99.876%	99.657%	99.655%	99.618%	99.618%
GRDR	CUMBERLAND	99.938%	99.954%	99.660%	99.741%	99.741%
GRHM	HENDERSON	N/A	N/A	N/A	N/A	N/A
HCKM	FULTON	99.926%	99.954%	N/A	N/A	N/A
HHTS	CAMPBELL	99.963%	99.965%	N/A	N/A	N/A
HRDB	MERCER	100.000%	99.950%	99.899%	99.873%	99.877%
HTFD	OHIO	99.344%	99.693%	99.676%	99.764%	99.798%
HUEY	BOONE	100.000%	99.978%	N/A	N/A	N/A
LGNT	LINCOLN	100.000%	99.981%	99.875%	99.846%	99.855%
LSML	FRANKLIN	100.000%	100.000%	99.651%	99.762%	99.822%
LXGN	FAYETTE	99.938%	99.981%	99.885%	99.835%	99.857%
MRHD	ROWAN	100.000%	99.256%	99.485%	99.655%	99.595%
MROK	BARREN	100.000%	99.981%	99.906%	99.925%	99.943%
MRRY	CALLOWAY	99.975%	99.892%	99.811%	99.800%	99.813%
OLIN	JACKSON	100.000%	99.908%	99.884%	99.894%	99.921%
PCWN	CASEY	100.000%	99.981%	99.777%	99.689%	99.766%
PGHL	CHRISTIAN	100.000%	99.942%	99.818%	99.857%	99.829%
PRNC	CALDWELL	99.715%	99.703%	99.738%	99.766%	99.819%
PRYB	GRAVES	99.988%	99.996%	N/A	N/A	N/A
PVRT	MCLEAN	99.888%	99.965%	99.187%	99.430%	99.507%
QKSD	BREATHITT	100.000%	99.823%	99.842%	95.016%	90.081%
RBSN	HARRISON	100.000%	99.969%	N/A	N/A	N/A
RNDH	METCALFE	99.987%	99.992%	N/A	N/A	N/A
RPTN	CRITTENDEN	98.960%	99.634%	99.661%	99.680%	N/A
RSVL	LOGAN	99.864%	99.846%	99.621%	99.603%	99.634%
SCTV	ALLEN	99.938%	99.961%	99.892%	99.915%	99.935%
SWON	OWEN	100.000%	99.981%	99.876%	99.904%	99.924%
VEST	KNOTT	100.000%	100.000%	N/A	N/A	N/A
WLBT	MORGAN	100.000%	99.742%	99.802%	99.808%	N/A
WNCH	CLARK	100.000%	100.000%	99.677%	N/A	N/A
WSHT	MASON	100.000%	99.996%	99.824%	99.829%	N/A
* AVERAGE		99.692%	99.794%	99.668%	99.715%	99.730%

* QKSD Excluded from 9 month & 1 year averages due to site flooding

It is very important to note that site availability percentages can be affected by much more than communications outages. For instance, the lower percentages at the "QKSD" site in the nine-month and one-year analyses were caused by the site being offline for approximately one month due to major flooding which destroyed much of the equipment there. Similarly, site "CRRL" had its uptime percentage lowered due to a day-long datalogger failure. All sites also had small amounts of downtime for routine maintenance. Knowing those caveats, the table can be used to reasonably assess communications performance, as the majority of site downtime[18] is typically caused by communications failures. The typical Mesonet site[19] was available for a respectable 99.794% in the three-month period ending 28 February 2010 and was available for 99.730% of the one-year period ending the same date.

Though outside of the period of analysis for the table, important to note is that approximately five sites, mostly in western Kentucky, were taken completely offline for at least a day due to a communications outage resulting from a devastating ice storm beginning 27 January 2009. The outage, caused by loss of critical fiber optics and power, impacted not only cellular communications but also took local National Weather Service forecast offices offline. No outage of its kind has since been experienced.

4.6 Discussion and summary

From its inception, the Kentucky Mesonet has desired an easily deployable, full two-way communications method that minimizes field technicians' efforts while keeping

[18] During communications failures, data are still measured and are collected once communications are restored. They are just not available to count as "fresh" data against the 20 minute threshold.
[19] QKSD excepted due to flooding.

the number of vendors to a minimum. Based on licensed coverage area, the Mesonet chose AT&T at its cellular provider, benefiting from an existing WKU contract with the company. To take maximum advantage of its investment, the Mesonet chose to utilize AT&T's Commercial Connectivity Service, which provides functionality well beyond what is possible with consumer-grade data service options. Though there have been some support hurdles similar to those experienced in the past by others utilizing cellular-based data collection, experience in Kentucky shows that the barrier to cellular-based data collection for automated weather stations has certainly been lowered over the last decade. In particular, data throughput has increased and power requirements have decreased thanks to the replacement of analog technologies by digital counterparts. Furthermore, analysis of uptime statistics shows the method to be decently reliable and stable, though some prolonged outages did occur during a major ice storm. Though cell-based data collection has been useful for the Mesonet and will likely remain the predominant method for the foreseeable future, alternative communications are being examined for areas which are "climatologically-rich" but "cellular-poor".

Before jumping into a thorough examination of the Kentucky Mesonet's computing systems architecture in the next three chapters, an overview of its general code approach may prove helpful. Figure 5-1 significantly aids that overview.

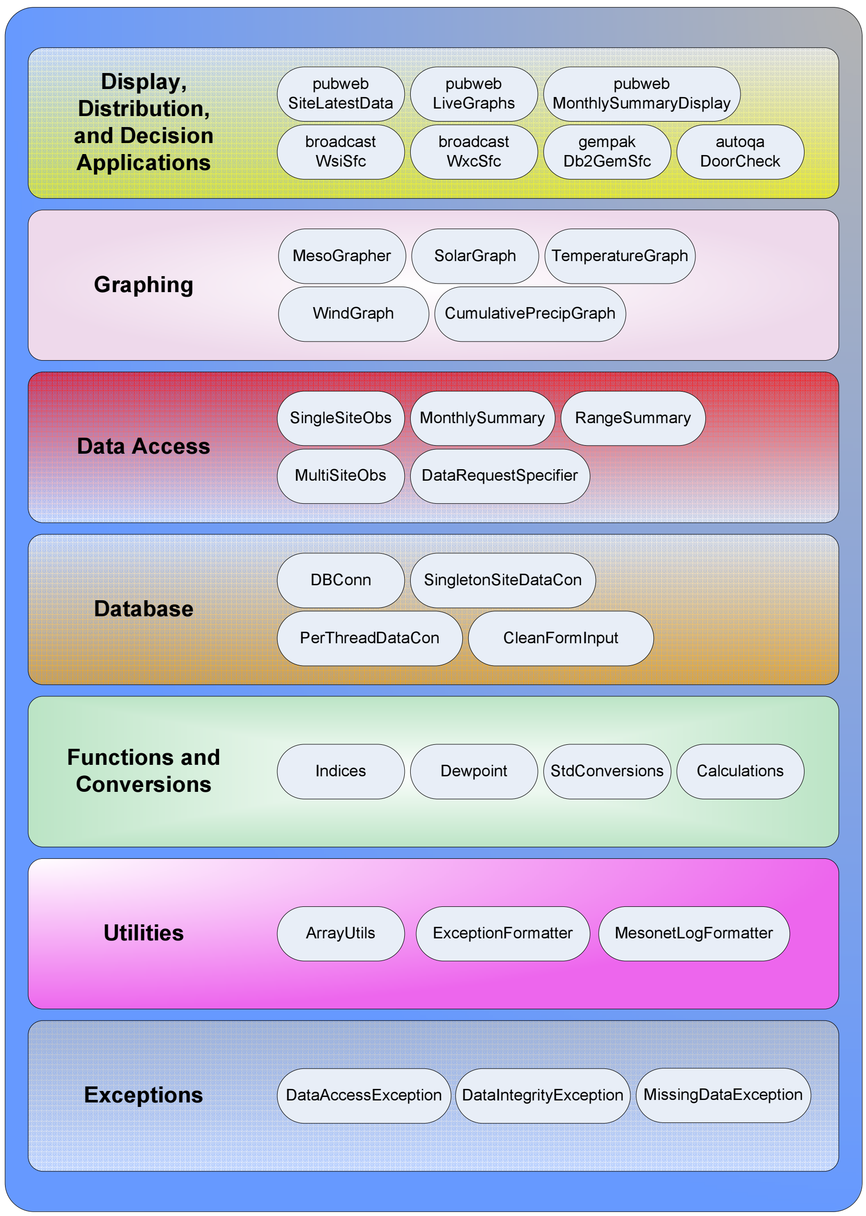


Figure 5-1. Kentucky Mesonet custom code libraries.

While many externally-developed applications and utilities have certainly been used in the creation of the Mesonet IT architecture, a substantial project-developed code base has also been implemented, with [20]approximately 60% of that code being produced by the author and 40% by other program personnel. Fowler (2003), in keeping with widely accepted best practices, notes that a layered approach should be utilized in the creation of specific applications within an IT project or organization. While not necessarily adopting the exact basic layers given by Fowler (2003), an object-oriented, modular approach has been used to develop the principal applications for the Mesonet. A good bit of procedural-based code is also used, mostly in the form of small- to moderately-sized scripts designed to carry out specific tasks, usually on a repetitive basis.

Figure 5-1 gives a non-exhaustive graphical overview of many of the modular code libraries – along with some example code classes – which have been developed to support Mesonet applications and which allow for increased coding efficiency through code reuse. While there is not an aversion to other languages, the majority of project-developed display, distribution, and decision application code is written either in PHP or Java (PHP Group 2009; Sun 2010), with Java being used for more complicated applications. As many of the libraries were developed in some form for both languages, distinctions between the languages are not stressed in either the figure or the discussion within this chapter. A delineation of specific languages used to develop certain applications is provided in the following three chapters, where individual systems and applications are covered.

[20] based on a count of code modifications in the Mesonet version control system

The remainder of this chapter reviews the main modular code libraries. An overview of script code is also provided, as are the network's views on the importance of code comments.

5.1 Modular libraries

A number of modular code libraries and packages have been developed in both PHP and Java to support display, distribution, and decision applications – including the main Mesonet website, critical data distribution methods for external partners, and the network's automated quality control system. Several of these modular libraries, depicted graphically in Figure 5-1, are overviewed below. The overview, however, is not exhaustive and is meant only to provide a general sense of the characteristics of the code base.

5.1.1 Display, distribution, and decision applications

Display, distribution, and decision application code generally consists of individual applications designed for a particular purpose, such as displaying near-real-time data on the Mesonet website (SiteLatestData), generating specialized data products for distribution to broadcast weather partners (WsiSfc, WxcSfc), and generating map-based data plots (Db2GemSfc). Following a modular approach, these applications invariably utilize code from one or many of the other libraries described below.

5.1.2 Graphing

Extending the JpGraph (Aditus 2008) object-oriented library for PHP, the graphing library is used to generate graphs of Mesonet data, primarily for the main website. The SolarGraph, TemperatureGraph, and WindGraph classes produce, not surprisingly, graphs of Solar Radiation, Air Temperature, and Wind, respectively. Other classes produce other graph types.

5.1.3 Data access

With a couple of exceptions, instead of being hard coded with direct access to the Mesonet observational database, applications and the graphing library use data access classes to retrieve and summarize Mesonet observational data in a uniform, predetermined manner. SingleSiteObs and MultiSiteObs classes allow access to individual observations, while MonthlySummary and RangeSummary provide statistical summaries for specific time periods.

5.1.4 Database

Classes within the database libraries create and provide the actual connections to Mesonet observational database(s) and also provide some security-related validation of query parameters. Instead of creating database connections themselves, classes within the data access layer rely on classes in the database layer.

5.1.5 Functions and conversions

Classes in the functions and conversions libraries provide standardized and typically static methods for particular meteorological calculations, such as calculation of dewpoint, and for standardized conversions (StdConversions) between observational units, such as meters per second and miles per hour. The Indices class provides methods for calculating Heat Index and Wind Chill values.

5.1.6 Utilities

Utility libraries generally provide some basic formatting and manipulation functions such as working with arrays (ArrayUtils) and formatting error messages (ExceptionFormatter) & application log entries (MesonetLogFormatter).

5.1.7 Exceptions

The exceptions library, which is only currently available for Java applications, is used to create a standardized set of exception types for errors in Mesonet-developed applications.

5.2 Scripts

In addition to the display, distribution, and decision code base described above, a number of critical scripting applications and utilities have been developed. These scripts, mostly Linux bash shell and Perl (2009) based, include:

(i) backup scripts – scripts used to automate data and system backups;

(ii) broadcast text generation – scripts used to automate broadcast data partner product generation;

(iii) ingest scripts – scripts used to automate the transport and graphing of raw data observations;

(iv) ldmp2db – scripts used to populate the Mesonet observational database with data collected via Campbell Scientific's LoggerNet software;

(v) Nagios scripts – an extensive set of scripts used to monitor individual systems and processes as part of Mesonet availability assurance methods;

(vi) and bandwidth accounting & signal test scripts – scripts used to conduct communications reliability and resource analyses, as used in Section 4.5.

5.3 Importance of comments

Before moving to the lengthy discussion of individual IT systems and applications in the next three chapters, this chapter will close with a note on the importance of well-commented code. Comments can provide a detailed description of complicated code and can make sharing and maintenance of code between different developers much easier.

Therefore, strong code commenting practices are encouraged and enforced for Kentucky Mesonet applications and systems. Figure 5-2 shows an example of opening class-level comments for an automated quality control algorithm.

```
/**
 * This algorithm is used to perform an intercomparison
 * between the 3 air temperature sensors at a Kentucky Mesonet
 * site.  If the difference between any two sensors is greater
 * than the threshold, both of those sensors are
 * marked SUSPECT.  Then, for an individual sensor, if its
 * differences between both other sensors are BOTH greater than
 * the threshold, it is marked as
 * WARNING.  Note that the derivation of a final, derived air
 * temperature value and a final, derived air temperature qa
 * flag is not accomplished here, but should rather be done
 * in other data access code.
 * <br><br>
 * Difference Tests:
 * <br> D12     D12 > THRESHOLD    TA01 & TA02 MARKED SUSPECT
 * <br> D23     D23 > THRESHOLD    TA02 & TA03 MARKED SUSPECT
 * <br> D31     D31 > THRESHOLD    TA03 & TA01 MARKED SUSPECT
 * <br><br>
 * Individual Sensor Tests:
 * <br>D12 & D31 > THRESHOLD TA01 MARKED WARNING
 * <br>D12 & D23 > THRESHOLD TA02 MARKED WARNING
 * <br>D23 & D31 > THRESHOLD TA03 MARKED WARNING
 *
 * <br><br>Rules for use:
 * <li> QaTarget must have 1 and only 1 network.
 * <li> Network must have 1 and only 1 network site.
 * <li> Network site must have ONLY the TBL_5min ObGroup
 * <li> The ob group must have same start and stop time
 * <li> MUST have TA01, TA02, and TA03 are variables
 * <li> The time step in ob group is irrelevant
 * <br><br>
 * @author Mike Grogan, Kentucky Mesonet
 * @author Andrew Quilligan, Kentucky Climate Center
 *
 */
```

Figure 5-2. Example code comments.

CHAPTER 6. CORE IT SYSTEMS

According to Dewett and Jones (2001), information technology moderates many aspects of bringing new problem solving ideas into use by determining the way information is stored, transmitted, communicated, processed, and perceived. If this is the case, which is certainly supported by the remainder of this document, it means that the IT systems developed over the course of the last three years have not only played a pivotal part in the way the network operates but have also greatly shaped the character of the Kentucky Mesonet. The three key IT management decisions noted by Martin (2003) – the architecture plan, resources & skills of practitioners, and application of appropriate methodologies and practices – have culminated in the development of nine core Mesonet IT systems, each of which are reviewed in this chapter, plus three geographic information and six ancillary systems reviewed in the next two chapters. Again, as Zachman (1987) notes, there is not a single systems architecture but a set of them. The “systems” reviewed in this chapter, therefore, include specific applications, servers, services spanning multiple servers, databases, scripts, and more complex code. The intent is not to describe every last technical detail but rather to provide for each system a general overview, a review of technical implementation steps and rationale, and a discussion of possible areas for improvement.

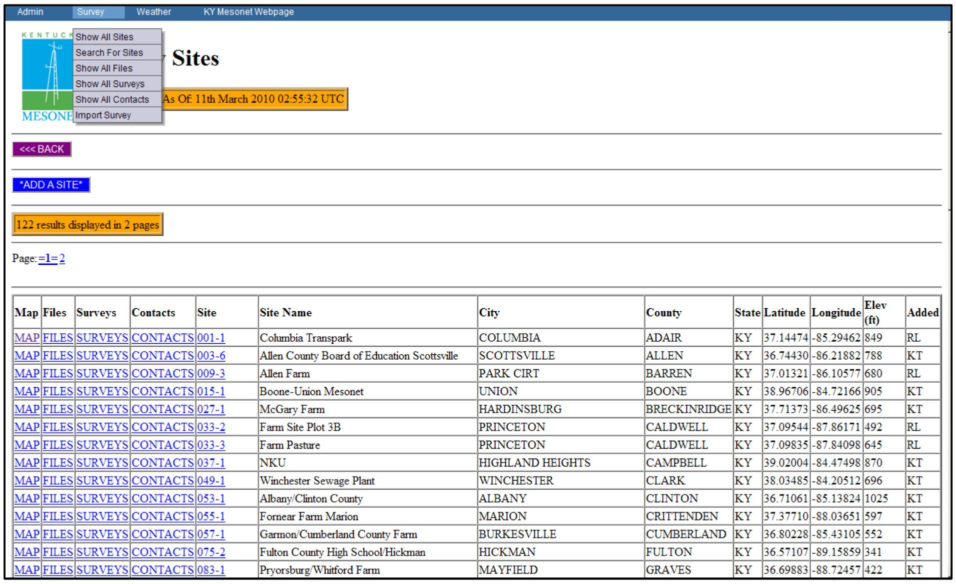
6.1 Site survey database

One of the main NRC (1999) observing principles is that network documentation should include information on station location, exposure, and local environmental conditions. This is echoed in WMO (2006), which notes the importance of including site surroundings, obstacles, and instrument layout in such documentation. As it desires to locate sites in quality locations suitable for long-term measurement of met/climate variables, the Kentucky Mesonet has always utilized a candidate site survey process to collect many data about potential sites, including:

(i) site suitability scores for meteorological variables and obstructions;
(ii) site contact and travel directions;
(iii) site geographical information;
(iv) resource availability, such as AC power and whether or not the landowner will allow guy wires for the network's 10 m towers;
(v) site communications statistics, as described in Section 4.3.2;
(vi) and site photographs taken of the site.

6.1.1 General overview and need

Through professional courtesy, the network had been using survey forms from NERON and storing survey data in a NERON database (OCS 2006). However, the discontinuation of that network shortly after the author's hiring meant a replacement survey database had to be quickly developed. Built as a no-frills, web-based, database-backed storage and retrieval system, the front page (Figure 6-1) of the network-developed site survey system provides a quick listing of all surveyed sites and a dropdown menu for system navigation. It also provides links to and the ability to edit all of the individual surveys, contacts, and site information detailed above.



Map	Files	Surveys	Contacts	Site	Site Name	City	County	State	Latitude	Longitude	Elev (ft)	Added
MAP	FILES	SURVEYS	CONTACTS	001-1	Columbia Transpark	COLUMBIA	ADAIR	KY	37.14474	-85.29462	849	RL
MAP	FILES	SURVEYS	CONTACTS	003-6	Allen County Board of Education Scottsville	SCOTTSVILLE	ALLEN	KY	36.74430	-86.21882	788	KT
MAP	FILES	SURVEYS	CONTACTS	009-3	Allen Farm	PARK CIRT	BARREN	KY	37.01321	-86.10577	680	RL
MAP	FILES	SURVEYS	CONTACTS	015-1	Boone-Union Mesonet	UNION	BOONE	KY	38.96706	-84.72166	905	KT
MAP	FILES	SURVEYS	CONTACTS	027-1	McGary Farm	HARDINSBURG	BRECKINRIDGE	KY	37.71373	-86.49625	695	KT
MAP	FILES	SURVEYS	CONTACTS	033-2	Farm Site Plot 3B	PRINCETON	CALDWELL	KY	37.09544	-87.86171	492	RL
MAP	FILES	SURVEYS	CONTACTS	033-3	Farm Pasture	PRINCETON	CALDWELL	KY	37.09835	-87.84098	645	RL
MAP	FILES	SURVEYS	CONTACTS	037-1	NKU	HIGHLAND HEIGHTS	CAMPBELL	KY	39.02004	-84.47498	870	KT
MAP	FILES	SURVEYS	CONTACTS	049-1	Winchester Sewage Plant	WINCHESTER	CLARK	KY	38.03485	-84.20512	696	KT
MAP	FILES	SURVEYS	CONTACTS	053-1	Albany/Clinton County	ALBANY	CLINTON	KY	36.71061	-85.13824	1025	KT
MAP	FILES	SURVEYS	CONTACTS	055-1	Fornear Farm Marion	MARION	CRITTENDEN	KY	37.37710	-88.03651	597	KT
MAP	FILES	SURVEYS	CONTACTS	057-1	Garmon/Cumberland County Farm	BURKESVILLE	CUMBERLAND	KY	36.80228	-85.43105	552	KT
MAP	FILES	SURVEYS	CONTACTS	075-2	Fulton County High School/Hickman	HICKMAN	FULTON	KY	36.57107	-89.15859	341	KT
MAP	FILES	SURVEYS	CONTACTS	083-1	Pryorsburg/Whitford Farm	MAYFIELD	GRAVES	KY	36.69883	-88.72457	422	KT

Figure 6-1. Opening screen of the site survey system.

Two of the most important features of the system involve the upload and import of site- and survey-related files. Upon completion of a survey, a network surveyor can use the system's automated Excel (Microsoft 2003)-to-database translator (Figure 6-2) to quickly populate most data fields from a standard scoring spreadsheet. They can also quickly import other binary documents, such as the obstruction drawing shown in Figure 6-3, and associate those with the site through the system's file upload utility.

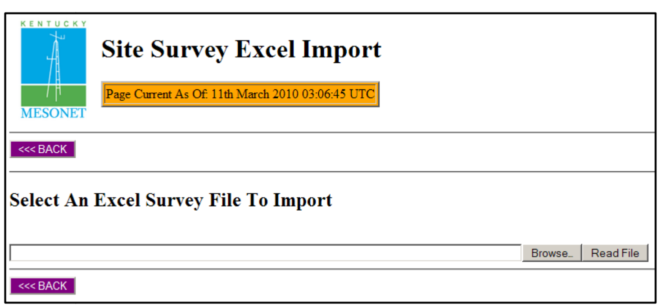


Figure 6-2. Site survey system Excel based importer.

Form Date: 08/14/2006

The site is located about 1.5 miles from town; Greenville. There are rolling hills off to the west and northward direction with relatively flat areas off to the south and southeast. A past the residential homes to the sites immediate west this is a main highway about 140 meters out. North of the site this is about 13 meters there is an active crop field.

B.4.3 SITE OBSTRUCTIONS DRAWING

(Use only **BLACK INK** to facilitate scanning)
Draw each obstruction within 100 meters (330 feet) of the center of the plot, label its bearing from the center of the plot in degrees relative to true north, its angular height, and its distance from the center of the plot in meters below. The center of the circle below indicates the center of the plot and the edge of the circle represents the extent of the 100-meter range. Each range ring indicates 25 meters (82.5 ft.). In addition, label the locations of other significant terrain features that could affect instrument measurements, such as roads, parking lots, concrete slabs, and bodies of water.

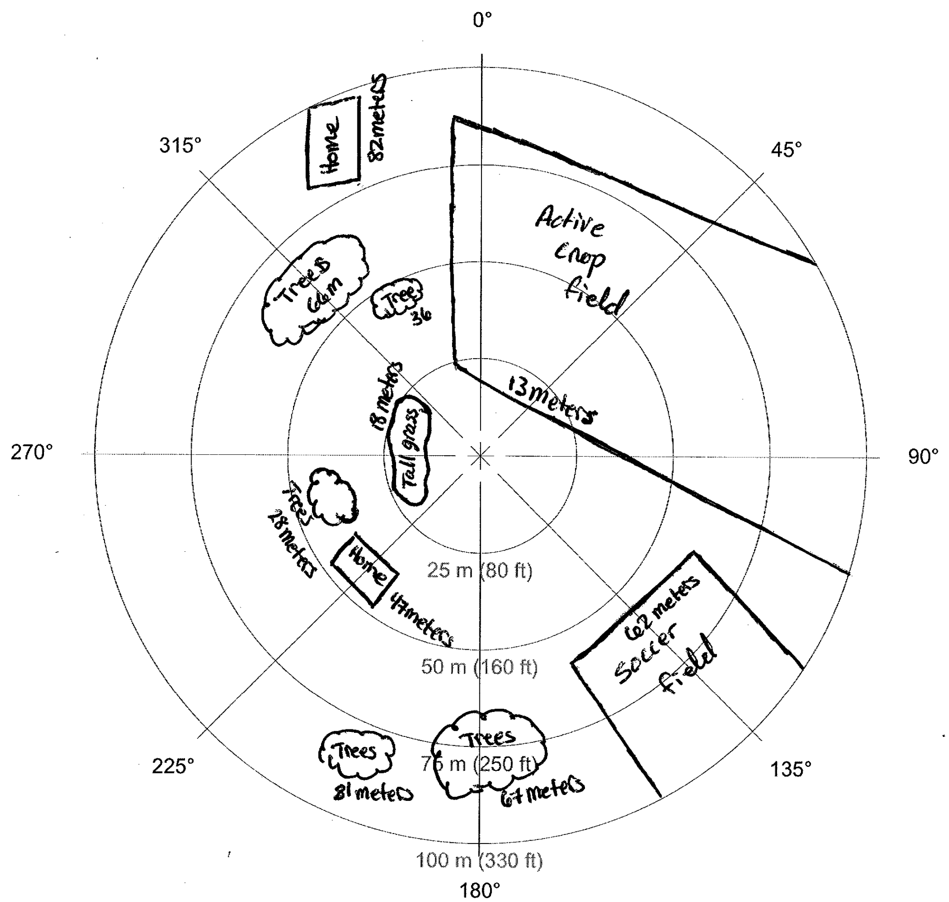


Figure 6-3. Example site survey obstruction drawing. (Source: Ronnie Leeper).

6.1.2 Technical implementation

Because it was needed before any of the Mesonet's computing infrastructure was really in place, the site survey system was built on a desktop PC running a version of the Linux Fedora Core operating system (RedHat 2007). The Apache web server (Apache 2010a) is used to host the system's web pages, and access to the system is restricted to Mesonet offices. A user privileging system, shown in Figure 6-4, controls what an individual user may do on the system.

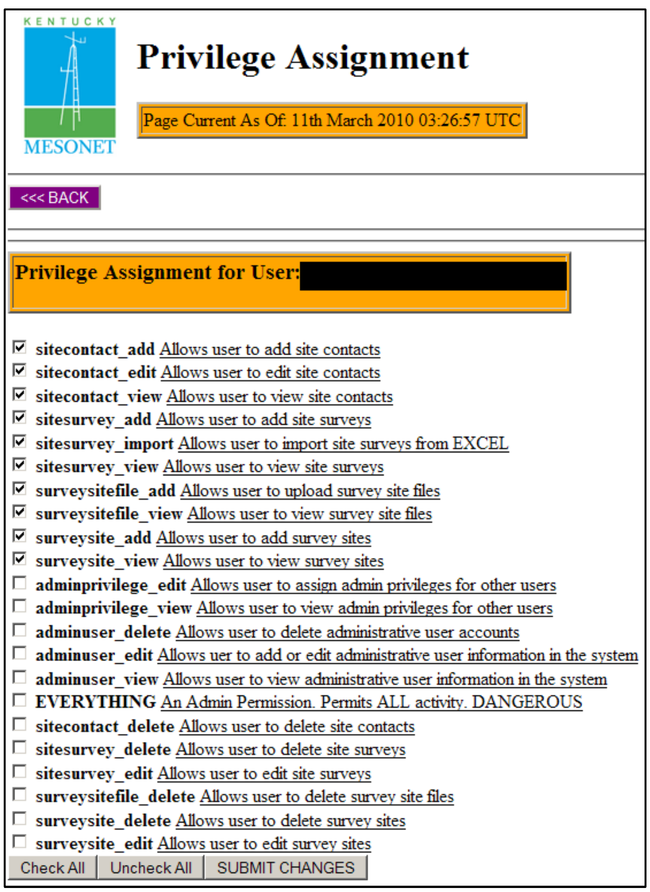


Figure 6-4. Site survey system privilege assignment.

The PHP (PHP Group 2009) code that powers the system uses more of a procedural-based coding method rather than an object-oriented approach. In addition to tables related to user privileges and security, the MySQL (Sun 2008) database backing the system contains the following:

(i) logentry – a table with a running log of actions taken by users;

(ii) counties – a table with relationships between counties and their Federal Information Processing Standards (FIPS) id;

(iii) sitesurvey – a table containing data from individual site surveys;

(iv) surveysite – a table containing general information about surveyed sites whose auto incrementing ID serves as primary foreign keys for other tables;

(v) surveysitecontact – a table containing contact information for a site;

(vi) and surveysitefile – a table containing tracking and relationship information for uploaded files.

The Excel-based importer uses the PHP-ExcelReader designed by Tkachenko and Harris (2007). Other uploaded binary files such as photographs are not stored in the database. Instead, they are stored on local disk in standard directories while references to the files are tracked in the database.

6.1.3 Needs for improvement

The site survey database needs to be moved from the Mesonet offices and incorporated into the network's centralized co-location facility. Additionally, it and the metadata database should be merged into a single system rather than being separate parts.

6.2 Metadata database

The meteorological sources in the Literature Review all stress the importance of metadata, which can often be just as important as actual observational data. Given this importance, there exists a need for the Kentucky Mesonet to track information about site locations, instrumentation, maintenance, calibration, and other related information in order to know "what was where when" and how it performed. The program's metadata system was designed by a former application developer who was responsible for about 75% of the design effort, with a student developer responsible for 15% and the author for the remainder. Since metadata are so important, a review of the system is included here for completeness. However, as the author was not the principal developer, the review may not be as detailed or thorough as for other systems. More details are available from internal program documentation (Brown 2008a). As was the case during initial development, the metadata system has historically been and will likely continue to be the responsibility of the person holding the Application Developer position. Turnover in this position and the need to keep other systems running has, unfortunately, led somewhat to a loss of focus on the system. However, a renewed maintenance and development effort is underway.

6.2.1 General overview

There are three main categories to the metadata system: sites, equipment and administration. As shown in Figure 6-5, site-related components are used to maintain an overview of the equipment currently assigned to a site and the measurement that each is assigned to take. The main “sites” screen can also be used to assign equipment to a site.

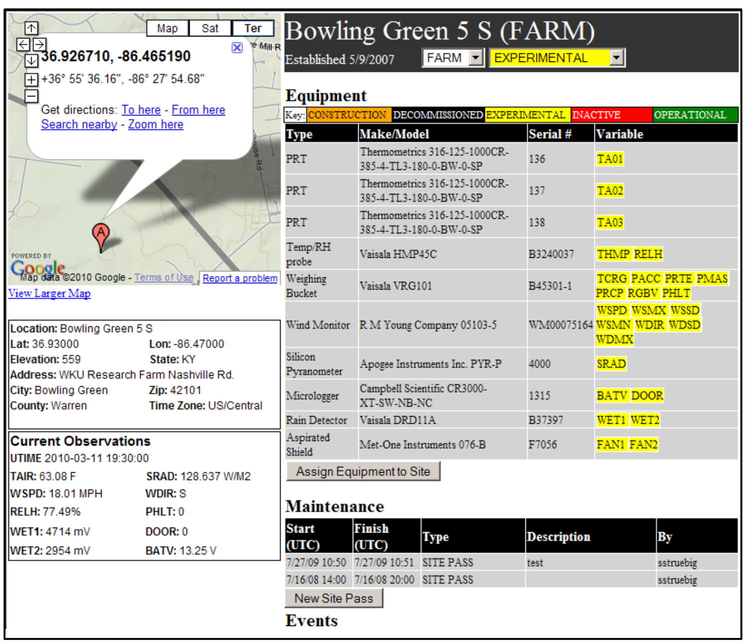
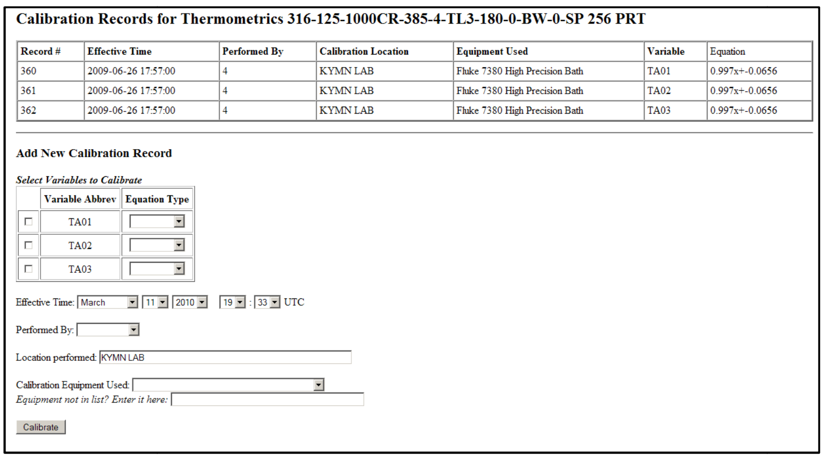


Figure 6-5. Metadata system "Sites" screen.

Equipment-related system functionality is intended to track the inventory, calibration, and maintenance of both measurement and ancillary equipment assigned to Mesonet sites. It is used to change the location of a piece of equipment, assign its measurements to a particular variable, and to record calibration information. For instance, Figure 6-6 shows the calibration records for a Thermometrics Platinum Resistance Thermometer calibrated in a Fluke 7830 high precision temperature bath in the Mesonet instrumentation and calibration lab. A variable selection is provided as some instruments, such as the program's wind monitors and relative humidity sensors, can measure multiple variables.



Calibration Records for Thermometrics 316-125-1000CR-385-4-TL3-180-0-BW-0-SP 256 PRT

Record #	Effective Time	Performed By	Calibration Location	Equipment Used	Variable	Equation
360	2009-06-26 17:57:00	4	KYMN LAB	Fluke 7380 High Precision Bath	TA01	0.997x+-0.0656
361	2009-06-26 17:57:00	4	KYMN LAB	Fluke 7380 High Precision Bath	TA02	0.997x+-0.0656
362	2009-06-26 17:57:00	4	KYMN LAB	Fluke 7380 High Precision Bath	TA03	0.997x+-0.0656

Add New Calibration Record

Select Variables to Calibrate

	Variable Abbrev	Equation Type
☐	TA01	
☐	TA02	
☐	TA03	

Effective Time: March 11 2010 19 : 33 UTC

Performed By:

Location performed: KYMN LAB

Calibration Equipment Used:
Equipment not in list? Enter it here:

Calibrate

Figure 6-6. Sample metadata system calibration record.

Other than leading all of the Mesonet's architecture development efforts, the author's primary contribution to the metadata system was the user privileging system, largely borrowed from the site survey database. Functionality of that component of the metadata system is very similar to that shown in Figure 6-4 for the site survey system.

6.2.2 Technical implementation

The metadata system is a completely on-line, web-based system available on the internal Mesonet computing network. It is built using a MySQL (Sun 2008) database with PHP (PHP Group 2009) code running on an Apache (2010a) webserver on a Linux host. Brown (2008a) notes that the code was primarily written in a procedural manner instead of using an object-oriented approach, as most of the PHP perform a unique, single task and “do not lend themselves to an object-oriented framework”. The display of current observations as shown in Figure 6-5, though, uses the object-oriented data access code library previously described.

The MySQL (Sun 2008) database structure includes about 64 different tables and also incorporates database triggers which are used to update some “snapshot” tables for convenience purposes. This is done so some simple web pages showing only current status and location information need not involve complicated or long-running queries (Brown 2008b).

6.2.3 Needs for improvement

There are several areas of the metadata system that need improvement and, as noted, the system has suffered due to turnover in the Application Developer position. Going forward, there must be a renewed focus on maintenance and updating of the database, functionality, and bug fixes. The current Application Developer is beginning that process now.

There also needs to be an effort to better tie this system into the site survey database and the observational database. At the very least, metadata must be made more available in the network’s data display and distribution mechanisms. This will certainly be required by the emerging Nationwide Network of Networks.

6.3 Data ingest system

At the core of any met/climate observation network are, obviously, the observations. Therefore, the data ingest system is one of the most critical systems deployed by the Kentucky Mesonet. An overview, technical details, and needs for improvement are given below.

6.3.1 General overview

As Campbell Scientific dataloggers are used for field data collection, the company's LoggerNet for Linux server software and its attendant LoggerNet Admin and LoggerNet Remote (Campbell Scientific 2009) utilities are used for site data retrieval. A screenshot of LoggerNetAdmin is shown below in Figure 6-7.

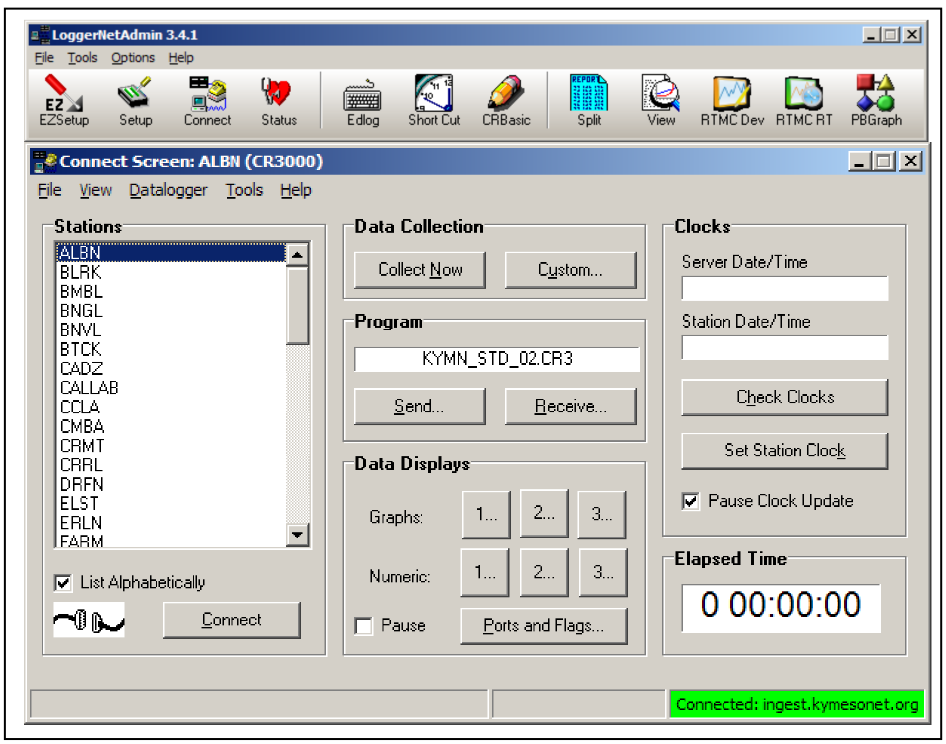


Figure 6-7. Campbell Scientific LoggerNetAdmin screen capture.

While Campbell Scientific's dataloggers and software can be configured to enable the logger to send observations to the server through callback methods, the Mesonet instead utilizes scheduled, server-initiated data collection. This method ensures that there are no collectible data holes and that all data are retrieved in temporal order. The Mesonet's ingest server is configured to contact each site once every 15 minutes to retrieve all data measured by each logger since last collection. This retrieval resolution is easily changed to five-minute collection during times of severe weather; see Section 4.5.1. A retry interval is configured so that the server will reattempt collection should it initially fail.

The ingest server is setup to store data in comma-delimited flat text files. These files are not used for routine storage, access, and display but are intended for raw data archival purposes and serve as a lowest-level backup for the operational observation database. Some basic data output and graphs, however, are generated from these via custom developed code. These graphs (Figure 6-8) are mainly used in evaluating a new site before it comes online or for viewing diagnostic variables.

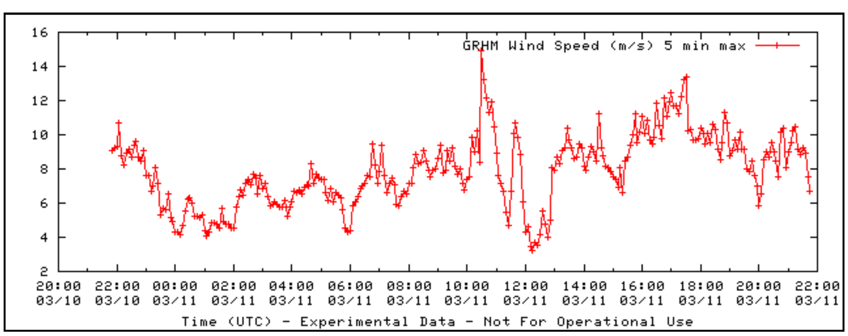


Figure 6-8. Rudimentary, initial graph created by Mesonet data ingest systems.

Population of the operational observation database is achieved via custom author-developed scripts which form a bridge between LoggerNet and the database. Data are written to the database immediately upon collection, without using the raw text files. While newer versions of LoggerNet can be configured to write to a database directly, the custom method developed allows for greater flexibility in database schema design.

6.3.2 Technical implementation

Campbell Scientific's LoggerNet (Campbell Scientific 2009), version 3.4, is installed on a server dedicated to data ingest. A Linux version of the software, for which the Kentucky Mesonet was among the first users, was chosen for consistency with the operating system of other computing network servers and hosts. Configuration of the server is accomplished through LoggerNetAdmin or LoggerNetRemote running on Windows (Microsoft 2007) desktops in the Mesonet offices.

Though not used operationally, the raw text files are important. Upon data collection, the server is configured to append data to a raw text file – one file for each site. For backup purposes, these files are transferred via the Linux rsync command to a PC in the Mesonet office – the same PC used for the site survey system. This machine uses custom bash shell scripts and GnuPlot (Williams and Kelley 2004) to produce graphs like that shown in Figure 6-8.

To populate the observation database, a custom written Perl (2009) script named ldbm2db is executed as a continuously running daemon which connects to a port monitored by LoggerNet's Logger Data Monitor Protocol version 2, or LDMP2. Upon connection, the script sends configuration information to the LDMP concerning the sites for which it wants data and the desired time period. As data are collected by the server, they are sent over the socket connection in a comma-delimited form, which is parsed by the custom Perl code. As it is the bridge between ingest and database, the custom code also maintains a connection to observation database systems. As data are streamed from LoggerNet, SQL queries are formed and sent to the database for operational data storage. Ldmp2db also maintains a connection to the automated QC system (Section 6.9) to signal to it the availability of new data. Fairly robust error handing mechanisms are incorporated into ldmp2db to allow for reconnection to the ingest server, database system, and automated QC and to allow for rollbacks of failed queries through the use of SQL transactions. Logs from ldmp2db are closely monitored by Mesonet availability assurance mechanisms.

6.3.3 Needs for improvement

Possible areas of improvement include breaking the ldmp2db daemon up into multiple instances rather than using a single instance for all sites. Newer versions of LoggerNet may also be investigated and tested.

6.4 Observation data storage system

Dewett and Jones (2001) note that IT promotes efficiency by providing the ability to store and retrieve lots of information quickly and easily; it codifies the knowledge base by facilitating organizational memory and making knowledge easy to communicate, assimilate, store and retrieve. Richardson et al. (1990) holds that data should be viewed and managed as a corporate asset. Data are at the core of Trenberth et al.'s (2002) priorities for in situ sensing networks, and data continuity & access are at the core of NRC (1999), GCOS (1993), and other guiding works. Therefore, to say that the observation data storage system lies at the heart of the Kentucky Mesonet's IT infrastructure would be an understatement. That system is reviewed below.

6.4.1 General overview

As of 28 February 2010, the Mesonet's observation database contained over five million groups of five-minute observations. As Table 6-1 shows, 20 individual meteorological measurements are included in each observation group, yielding a database with over 100 million total meteorological data points. Over 130 million measurements have been made when diagnostic variables are included, and about 250,000 additional measurements from non-standard equipment at select sites are included.

Table 6-1. Observation database record counts.

Year	Total 5 Minute Observation Groups	Met. Measure. per Ob. Group	Diag. Measure. per Ob. Group	Total Meteorological Measurements	Total Diagnostic Measurements	Total Measurements
*2010**	769,928	20	6	15,398,560	4,619,568	20,018,128
2009	3,047,596	20	6	60,951,920	18,285,576	79,237,496
2008	1,033,741	20	6	20,674,820	6,202,446	26,877,266
2007	194,271	20	6	3,885,420	1,165,626	5,051,046
Total	5,045,536	20	6	100,910,720	30,273,216	131,183,936

* ending February 28, 2010

When developing its observation storage system, the Mesonet had a choice between flat file storage mechanisms and the use of a relational database. One format, Unidata's network Common Data Form , or NetCDF, has seen healthy acceptance for scientific data storage (UCAR 2009c). However, based on Murhammer et al.'s (1999) design characteristics, most importantly economics, a MySQL (Sun 2008) database-backed system was chosen instead. While NetCDF is available at no cost, the availability of potential student and full-time employees with knowledge of MySQL was felt to be higher than availability of NetCDF-versed candidates.

The observation database contains only measurements and some statistical calculations, all in their original units. Derived measurements, summary statistics for defined periods, and unit-converted values are calculated on-the-fly by various classes in the Mesonet code libraries, and are often generated at display time.

6.4.2 Technical implementation

The observation storage system resides on a Linux server purchased and configured for this exclusive purpose. As mentioned, MySQL (Sun 2008) is used as the supporting database. For security and backup purposes, the database is mirrored on the program's public web server. The database's primary tables are shown in Table 6-2. The database is designed to potentially handle observations from multiple networks.

Table 6-2. Observation database primary tables.

Table	Description
data_KYMN_Aux_Dev_YYYY	Auxillary observations from the year YYYY for the KYMN network.
data_KYMN_TBL_5min_YYYY	Primary observations from the year YYYY for the KYMN network.
network	Table of networks, such as KYMN for the Kentucky Mesonet.
network_site_names	Names and abbreviations for network measurement sites.
QA_KYMN_Aux_Dev_YYYY	Quality controlled auxillary observations and flags from the year YYYY for the KYMN network.
QA_KYMN_Derived_TBL_5min_YYYY	Derived, quality controlled observations and flags from the year YYYY. For future use for the KYMN network.
QA_KYMN_flag_log_YYYY	Quality control and assurance flags from the year YYYY for the KYMN network.
QA_KYMN_Preset_Flags_YYYY	Manual quality assurance and control flags from the year YYYY for the KYMN network.
QA_KYMN_TBL_5min_YYYY	Quality controlled primary observations and flags from the year YYYY.
site_geog	Basic site geographic metadata, such as lat/lon/elevation.
site_time_zone_history	History of time zone changes for a site.
units	Measurement units.
variables	Measurement variables.

To speed query performance carefully constructed database indices are applied to the data. To aid queries where the site is the most important variable and time is of secondary importance – such as in querying all observations for a site or all observations for a site between a certain start and stop point – *site + observation time* indices are used. To aid queries where time is the most important query parameter – such as querying all observations for all sites between a certain start and stop point – *observation time + site* indices are used. Coordinated Universal Time (UTC) is the official timestamp for all observations but, to speed queries based on local time, both standard and daylight-saving-time-advanced timestamps are also stored.

To fully support transactions and enforce foreign key rules, the InnoDB database engine for MySQL is used (Sun 2008). Data are stored in a less normalized fashion than was originally wanted, as initial calculations in the design phase indicated that it would be quicker to keep individual measurements from an observation grouped together versus splitting them into separate rows with a variable ID field. These calculations were based on queries which regrouped fully normalized data at the command line. As Mesonet development grew, though, data access code libraries were increasingly used to pull observations from the database, likely rendering the timing issue moot. Deep down, the less-normalized database form felt like a mistake at the time and probably was. A more fully normalized schema appears now to have been more appropriate.

6.4.3 Needs for improvement

Some initial steps at a database redesign are underway now, including looking at databases more supportive of GIS applications and adjusting schema to a more fully-normalized form. Hopefully the object-oriented code approach used in Mesonet applications will help with this, as only underlying data access libraries should need modification for a schema update, not entire applications.

6.5 Product generation system

To help meet Trenberth et al.'s (2002) call for distribution of data in near-real time and to aid development of tools to satisfy NRC (1999) requirements for data access, a product generation system was developed. It is described below.

6.5.1 General overview

The primary intent of the product generation system is to perform automated, scripted product generation for both the public website and specialized data distribution systems used with key Mesonet partners. These data products need to be regularly updated but do not have to be dynamically generated on-demand, making them ideal candidates for automated creation.

One main function of the system is to create plots of data for the public website, such as wind speed & direction and radar reflectivity, as shown in Figure 6-9. As other websites sometimes co-opt these images, the Mesonet's primary web address is also output on plots of network data.

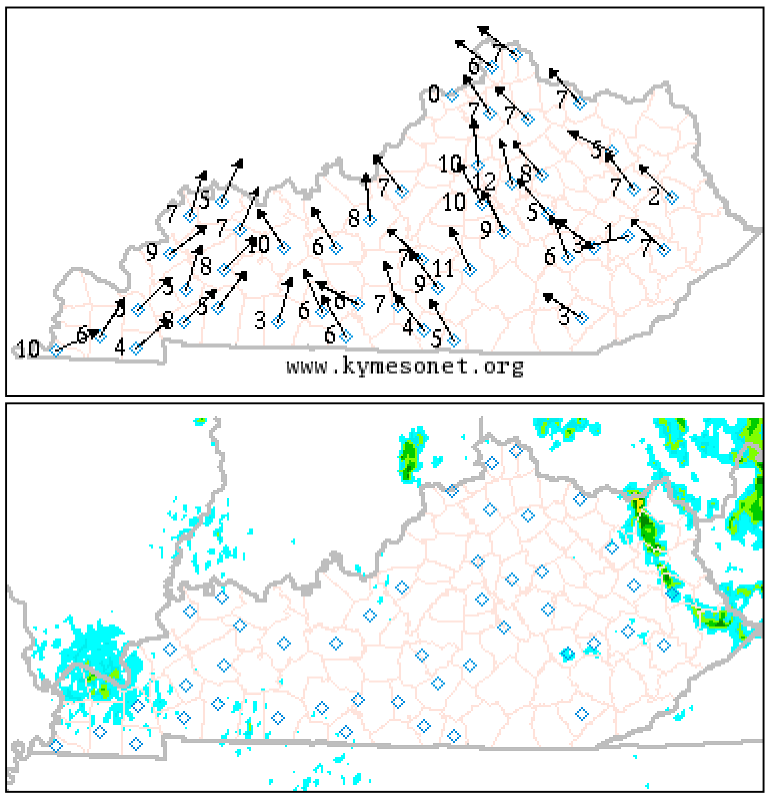


Figure 6-9. Graphic of wind speed & direction (top) and composite radar reflectivity data (bottom) generated for public website.

Partners in the broadcast industry in multiple markets receive feeds of network data which are directly importable into their weather display systems. The product generation system is used to generate data output in proprietary formats for both Weather Central and WSI systems. These feeds were developed in close cooperation with the broadcasters and their vendors.

6.5.2 Technical implementation

Unidata's GEMPAK software (UCAR 2009a) is used to create website plots of network and radar data. To take advantage of existing data access libraries, GEMPAK commands are executed by shell calls in Mesonet-developed PHP (PHP Group 2009) code. Data are retrieved from the observation database, are formatted into a binary GEMPAK surface file via its sfcfil and sfedit routines, and are then plotted with the sfmap_gf command. A virtual frame buffer is utilized by sfmap_gf, and in doing so some memory leaks have been experienced. After creation of data plots, images are then cropped for desired website size, labeled with "www.kymesonet.org", and copied over the internal network to the public website. A file containing UNIX epoch timestamps is created for each image and also copied to the website. It should be noted that the FORTRAN source code of the sfmap routine was modified for precipitation plots in order to show decimal places.

For radar data, a Perl (2009) script is used to download base reflectivity data for 12 radars in or near Kentucky via FTP from the National Weather Service and is executed every five minutes. Before download, online file timestamps are checked against files previously downloaded and the download is skipped if the file has been previously retrieved. GEMPAK's gdradr utility is used to create a gridded composite of all the radar files, followed by execution of its gdplot2 utility to create an image of the composite for the web. That image, along with an accompanying timestamp file, is then copied over the internal network to the public website, just like plots of Mesonet data.

PHP code is used to create every five minutes the necessary broadcast text output in proprietary formats. Classes in the data access library are used to retrieve the most current data for each Mesonet site and to retrieve summary data such as high and low temperature and precipitation. Code from the functions and conversions library is used to calculate parameters such as wind chill, heat index, and 16-point cardinal wind direction. Generated data files are then copied to the partner data distribution website to be made available to broadcast partners with data access agreements.

6.5.3 Needs for improvement

There is great desire to convert products generated by the product generation system into more interactive, GIS-based tools such as those shown in Figure 7-4. As that happens, the product generation system may morph into more of a product support system, moving away from GEMPAK map creation toward becoming a GIS map server. When and if that happens, though, care must be taken to protect external sites linking to existing graphics, and generation of broadcast data may need to be moved elsewhere.

6.6 Public website

NRC (1999) and GCOS (1993) principles call for data systems to facilitate user access to climate products and raw data. The Kentucky Mesonet's website (http://www.kymesonet.org) serves as the public face of the program and provides the main public access to the network's met/climate data. The site, of course, overviews the network and provides details on instrumentation, network quality, etc., but its focus is to provide both near-real-time data and summary statistics from all Mesonet locations.

6.6.1 General overview

The front page (Figure 6-10) includes a clickable map – generated by the product generation system – with which users can choose a site for which to display data in the "blue box" of current observations. The variable displayed on the map can be changed with the menu underneath it.

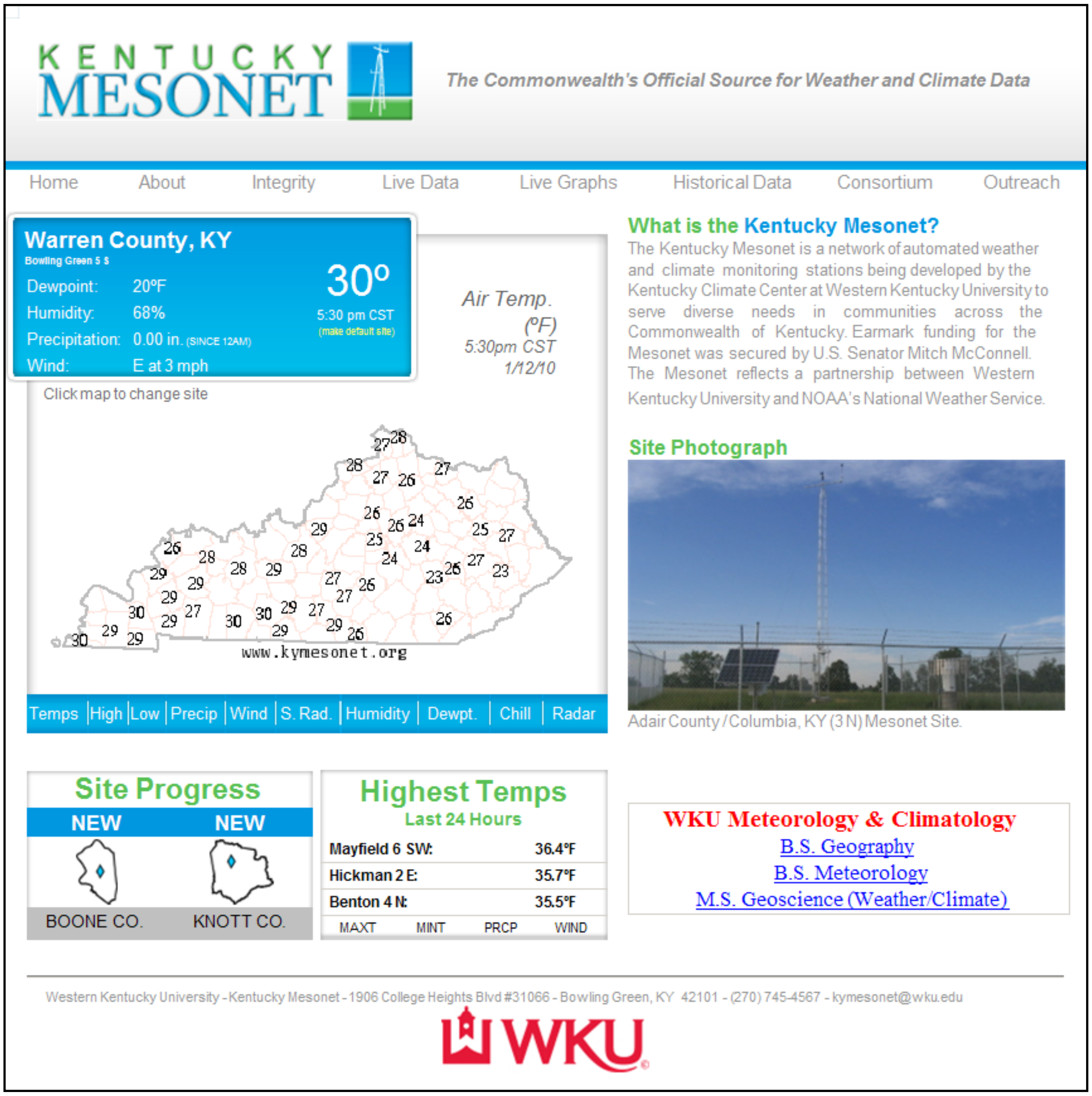


Figure 6-10. Front page of Kentucky Mesonet public website.

Meteograms for each network site are available from the “Live Graphs” page, shown in Figure 6-11. The graphs depict temperature, dewpoint, relative humidity, solar radiation, wind speed & direction, and precipitation. The latest 24 hours[21] of individual observations from each site are also available in tabular form from the “Live Data” page, not shown.

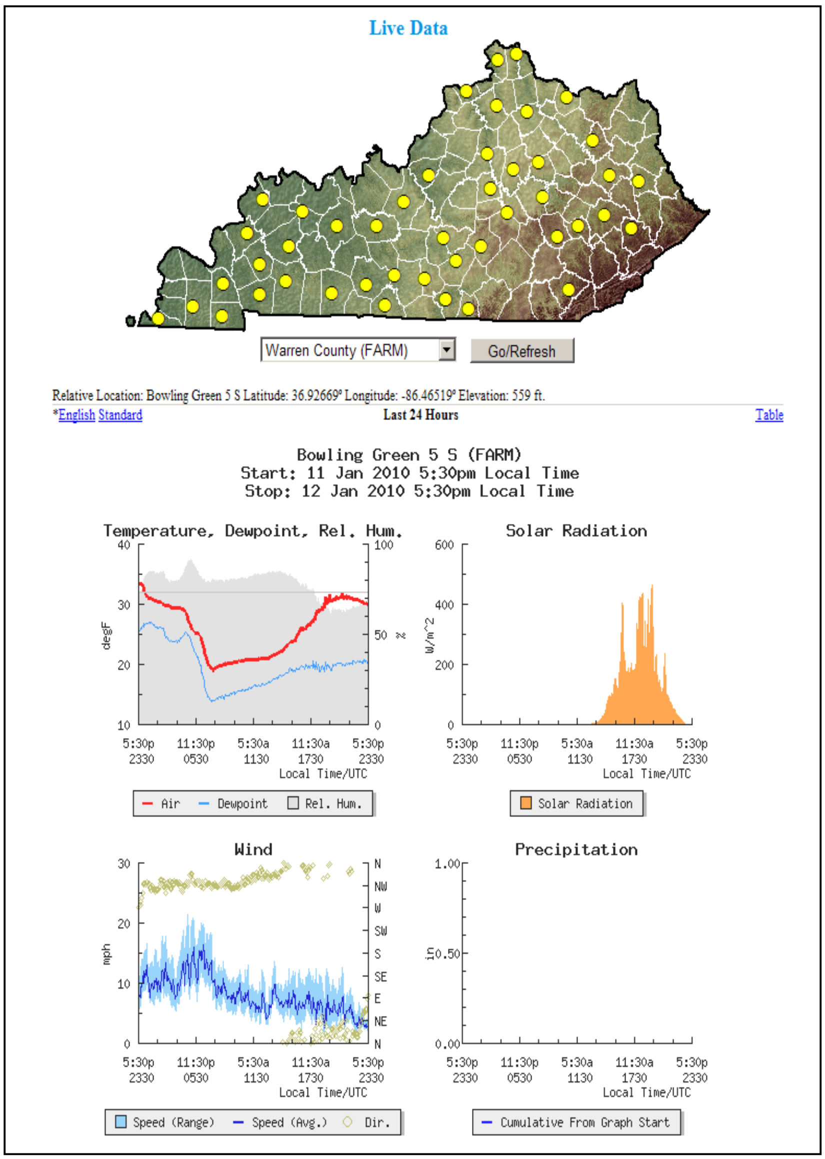


Figure 6-11. Website data graphs.

[21] From the internal network, this list shows up to 30 days of data.

The Monthly Climatological Summary website function, shown in Figure 6-12, provides a monthly summary of met/climate statistics for each site. Not shown, a dropdown menu allows the user to choose the month and site for which they want statistics.

Monthly Climatological Summary
Experimental Version
Full Quality Control Not Applied
(12/2009)
KENTUCKY MESONET

Station ID:	RPTN
Relative Location:	Marion 4 NE
County:	Crittenden County
***Location:**	Lat: 37.38°; Lon: -88.04°
Elevation:	594 ft.
Observation Day:	Central Standard Time

*To Nearest Hundredth

Day	Date	Temperature (°F)				Degree Days		Humidity (%)		Precip (inch)	Wind Speed (mph) and Direction				Solar (MJ/m^2)
		Max	Min	Avg	Avg Dwpt	HDD	CDD	Max	Min		Res Dir	Res Spd.	Avg Spd.	Max 3-sec	
TUE	1	50.0	32.6	41.3	31.1	24	0	91	42	0.00	S	4.1	5.9	14.9	10.1
WED	2	43.2	36.2	39.7	37.8	25	0	98	75	0.60	N	6.1	10.5	34.1	1.3
THU	3	40.3	25.8	33.0	28.7	32	0	97	60	0.00	NW	8.5	8.7	23.2	5.4
FRI	4	34.9	21.3	28.1	19.4	37	0	93	47	0.00	NNW	3.3	3.9	11.3	10.9
SAT	5	36.1	17.4	26.7	17.1	38	0	96	37	0.00	NNE	0.1	3.1	11.2	11.8
SUN	6	37.4	22.2	29.8	20.2	35	0	84	47	0.00	SSE	4.6	4.9	13.2	6.7
MON	7	42.5	33.4	37.9	34.0	27	0	97	71	0.11	SW	2.6	5.8	19.1	4.9
TUE	8	56.0	36.1	46.1	41.5	19	0	99	85	1.21	SE	7.6	11.5	31.7	0.9
WED	9	52.2	20.2	36.2	26.3	29	0	95	56	0.00	W	18.6	19.3	45.1	9.0
THU	10	28.8	16.0	22.4	9.2	43	0	79	32	0.00	W	9.3	9.9	23.2	11.8
FRI	11	38.1	19.1	28.6	13.1	36	0	89	26	0.00	SSW	4.8	5.6	17.4	10.3
SAT	12	43.3	25.0	34.1	19.3	31	0	90	34	0.02	SSE	7.2	7.4	18.6	4.0
SUN	13	48.0	41.4	44.7	43.1	20	0	98	78	0.16	SW	5.3	7.4	19.8	3.7
MON	14	58.3	41.6	49.9	46.6	15	0	94	72	0.00	SW	8.6	12.5	30.8	4.7
TUE	15	41.6	22.1	31.8	24.4	33	0	91	66	0.00	NNW	11.6	11.9	26.5	7.8
WED	16	35.4	18.1	26.8	16.1	38	0	90	40	0.00	E	1.2	3.7	10.3	11.2
THU	17	46.7	21.1	33.9	21.0	31	0	87	33	0.00	SE	4.5	4.9	11.5	8.3
FRI	18	50.6	32.7	41.7	29.2	23	0	95	34	0.04	SE	1.3	3.1	12.5	5.9
SAT	19	38.6	31.3	35.0	32.2	30	0	99	87	0.04	NW	11.7	11.8	28.3	1.5
SUN	20	31.5	29.0	30.3	26.9	35	0	91	83	0.00	WSW	6.6	8.2	20.4	2.4
MON	21	33.1	25.9	29.5	29.0	35	0	100	90	0.00	S	1.5	3.6	11.1	2.6
TUE	22	53.3	30.1	41.7	35.8	23	0	101	57	0.00	SE	5.8	6.3	13.7	7.4
WED	23	56.7	44.8	50.7	40.9	14	0	90	61	0.04	SE	8.4	8.7	21.9	5.6
THU	24	59.3	45.2	52.2	44.9	13	0	97	53	0.89	SE	14.7	15.4	43.3	3.8
FRI	25	53.5	28.2	40.9	30.4	24	0	96	57	0.15	SSW	15.0	16.3	38.5	1.6
SAT	26	39.2	27.1	33.2	20.9	32	0	70	45	0.00	SW	12.2	12.9	35.4	10.4
SUN	27	34.0	21.5	27.7	21.9	37	0	87	63	0.00	W	8.9	10.5	28.7	2.6
MON	28	33.2	20.1	26.7	21.1	38	0	89	69	0.00	W	8.1	9.7	24.0	5.5
TUE	29	35.0	23.4	29.2	19.3	36	0	83	50	0.00	NE	4.1	5.8	21.3	10.6
WED	30	42.7	28.5	35.6	28.8	29	0	99	53	0.14	S	6.8	7.6	25.5	3.2
THU	31	41.1	23.4	32.2	35.4	33	0	99	77	0.03	NW	7.6	8.9	23.8	2.1
Monthly Average		43.0	27.8	35.4	27.9			92	58		SW	2.8	8.6	22.9	
Monthly Total						917	0			3.41					188.2

Figure 6-12. Monthly climatological summary.

Due to a lack of available time during Mesonet startup, and to take advantage of existing university partnerships, the graphical design, look, and feel of the website were handled by a local web design company, HitCents, located in WKU's Center for Research and Development. The content and technical functionality, however, were completely designed, coded, and implemented by Mesonet personnel. In the form presented here, the author contributed about 50% of the coding effort, with a student developer contributing the other 50%.

6.6.2 Technical implementation

The website is hosted on a Linux-based Apache (2010a) web server and, at the time of document creation, was a pretty straightforward, basic site. Non-data functionality is mostly implemented in basic Hypertext Markup Language (HTML) with cascading style sheets (CSS). Some JavaScript (Sun 2009) is used to load a common menu on the top of each page and is sometimes use for displaying a common informational banner across each page.

PHP (PHP Group 2009) applications provide the majority of dynamic data capability. Each of these heavily relies on the data access code library for retrieving individual observations and summary statistics. The functions and conversions library is used extensively to convert data at display time from Mesonet standard units, typically metric, to English units. Graph creation utilizes the Mesonet graphing library, which is based on JpGraph (Aditus 2008); see Section 5.1.2. To cut down on processing time, a type of dynamic caching capability designed by the author is used to create graphs as an image upon first request. That cached image is then re-served upon subsequent requests for the same station and time period.

Several versions of the website are typically in use at any given time. One is the released public version, while others are versions under maintenance and development.

6.6.3 Needs for improvement

Modifications to the website are currently underway and it will likely have changed before this thesis is approved in final form. Changes to the site should bring increased interactivity and more dynamic data displays via Asynchronous JavaScript and XML, or AJAX, based methods. It will include larger mapping to accommodate more stations as well as additional topical sections.

6.7 Partner distribution systems

The Kentucky Mesonet works with a number of critical partners, including the National Weather Service, broadcast media, state government, universities, and agriculture interests to provide both specialized feeds of Mesonet data and on-demand data retrieval from the observation database. These are detailed below.

6.7.1 General overview

Extensible Markup Language (XML) based data feeds, which are really dynamic data pulls from a special website for data partners, were first developed with the National

Weather Service office in Jackson, KY and the University of Kentucky Agricultural Weather Center. The available feeds include:

(i) the network sites feed, which contains basic site metadata like latitude, longitude, and elevation;

(ii) the latest data feed, which contains an XML-based listing of recent data collected from Kentucky Mesonet sites;

(iii) the latest GeoRSS feed, which is a GeoRSS (2010) version of (ii) developed for the Kentucky Division of Geographic Information;

(iv) and the range summary feed, which allows partners to request data summaries (such as max/min temperature, precipitation, etc.) for a user-defined period.

The Bulk Data Retrieval Interface, shown in Figure 6-13, allows internal Mesonet personnel and data partners to retrieve archived network data in a more human-readable and usable format. Data can be output in an HTML table or as comma separated values.

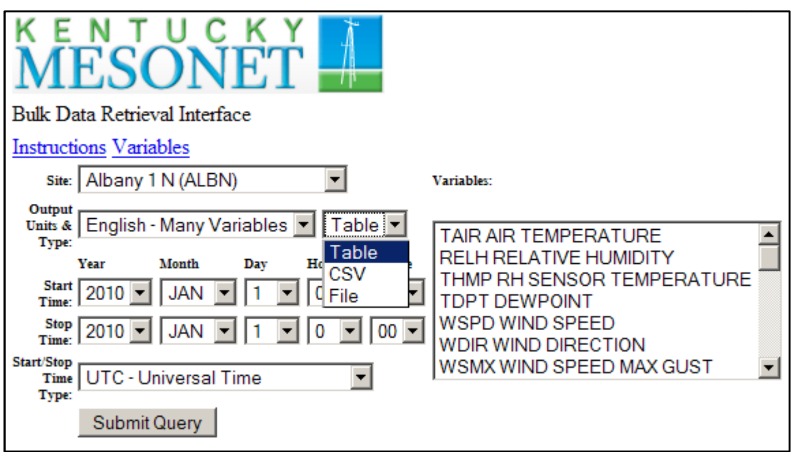


Figure 6-13. Bulk Data Retrieval Interface

The “File” option prompts the interface to return data with a specialized header that causes or allows them to be opened directly into a user’s default spreadsheet program, if installed. Covered already in Section 6.5, partner data distribution systems also host Mesonet data in proprietary, text-based broadcast weather system formats.

6.7.2 Technical implementation

Partner data feeds are hosted on an Apache (2010a) web server, with PHP (PHP Group 2009) code powering each. All of them take advantage of the data access code library, specifically those classes that deal with station data retrieval and summary products. For the XML-based feeds, the PHP generates XML-based output in both custom XML and GeoRSS forms, whereas the PHP for the Bulk Data Retrieval interface generates tables or comma separated values wrapped in HTML format. If “File” is the output type chosen by the user, then PHP is used to generate a special header that causes a web browser to open the data in the user’s default spreadsheet application. The PHP applications are dynamic in that they allow data partners to specify sites and temporal coverage of the output. For the bulk retrieval interface, they can also specify the variables. Due to processing time, especially for summary statistics, the data retrieval period is limited. Broadcast data generation is as described in Section 6.5.

6.7.3 Needs for improvement

Like the broadcast text generation systems, all partner data distribution methods may need to be slightly redesigned with a consideration toward caching the most commonly requested datasets, which typically are requests for the latest observations from all Mesonet sites. By caching the results of these requests, load on distribution and database systems may be significantly reduced.

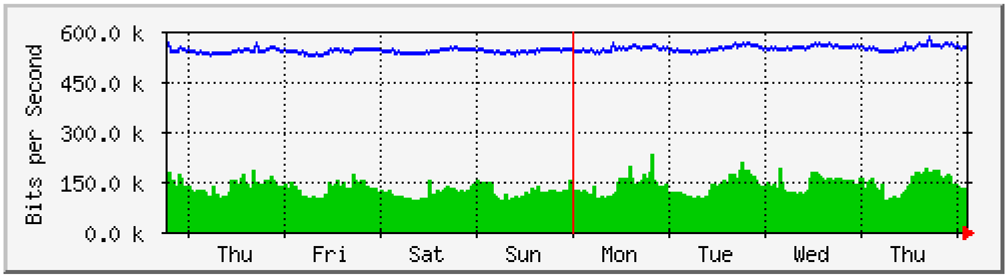
6.8 Availability assurance systems

One of Murhammer et al.'s (1999) principles is that computer network management should include methods to monitor the health of systems to ascertain operating conditions and to isolate faults. This principle, along with the Mesonet's own requirements for high availability, dictates the operation of availability monitoring and assurance systems. Several systems have been developed to support the Mesonet and are described below.

6.8.1 General overview

The simplest availability assurance mechanism is an audible system which monitors a custom web-based status page for error messages or connection failures. A synthesized voice in the main Mesonet office alerts computing staff to errors. It also reads aloud on-the-hour the temperature at the Bowling Green site.

Network traffic graphs continuously monitor and record usage statistics from computing network devices. Figure 6-14 shows traffic both terminating at and originating from the program's web server for a week and month.



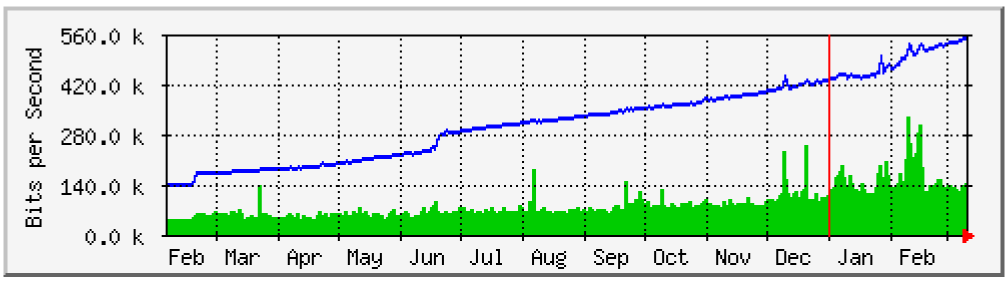


Figure 6-14. Weekly (top) and yearly (bottom) website traffic statistics ending 11 March 2010. Outbound traffic is in blue, with inbound traffic in green.

The Nagios (2008) IT Infrastructure Monitoring system has been implemented to maintain a constant vigil over sensor network data availability, computer network health, physical server health, and data product availability. Most importantly, the system alerts computing personnel by both e-mail and text message about critical outages. The front page of the Nagios system is shown in Figure 6-15, while a sample notification alert is given in Figure 6-16. In addition to alerting computing staff quickly about problems, the Nagios system can be used to generate a number of availability statistics. This capability was used to generate the site uptime statistics given in the Communications chapter.

Because the Nagios system sits within the internal computing network, a method to monitor the network itself for availability is required, else there would be no way to receive notifications if the entire computing network were down. Therefore, the Mesonet uses an external, fee-based monitoring service from SiteUptime, LLC. to perform rudimentary checks of Mesonet website and computing network availability. A graph from that service (Figure 6-17) shows uptime for the month ending 28 February 2010.

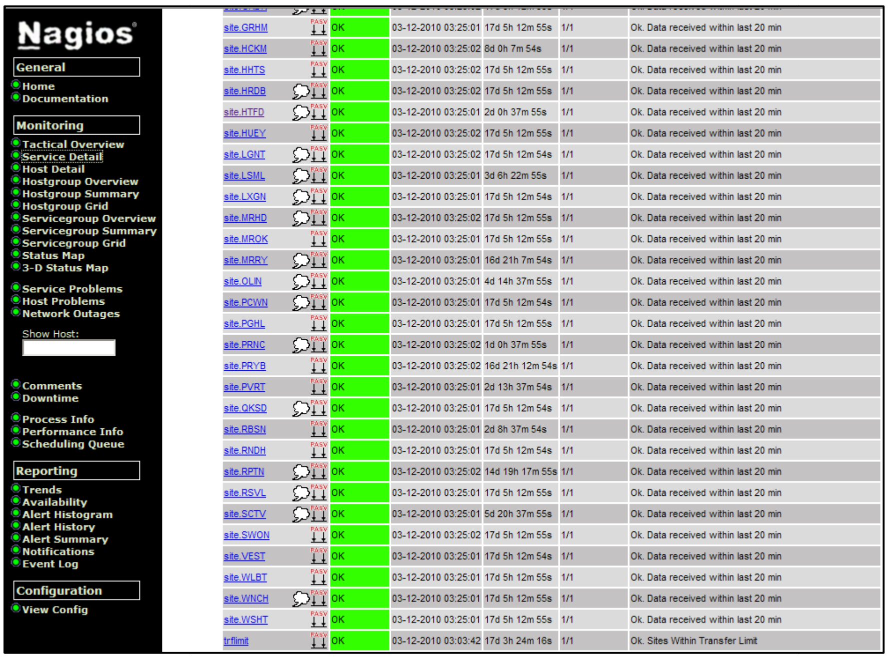


Figure 6-15. Front page of Nagios (2008) IT infrastructure monitoring system.

```
***** Nagios 2.9 *****
Notification Type: PROBLEM
Service: site.PRNC
Host: ingest
Address: [redacted]
State: CRITICAL
Date/Time: Thu Mar 11 02:40:16 GMT 2010
Additional Info:
PRNC datafile 24.4 mins old
```

Figure 6-16. Sample Nagios (2008) alert notification.

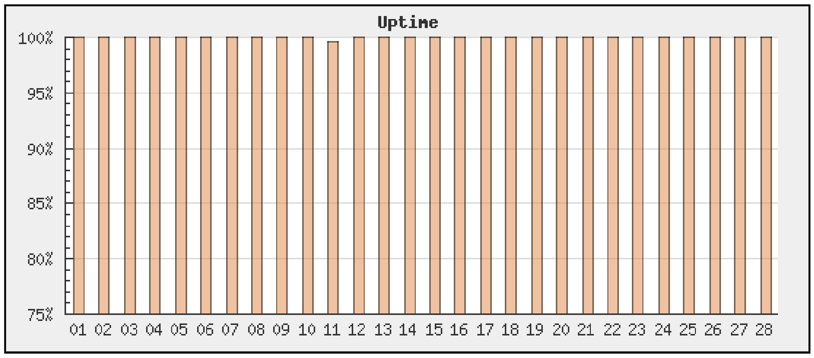


Figure 6-17. Uptime graph from external monitoring service, SiteUptime, LLC. Period is 1 month, ending 28 February 2010.

6.8.2 Technical implementation

Synthesized voice monitoring is implemented using the Festival Speech Synthesis System from the University of Edinburgh (2004). Perl (2009) scripts are executed to download status and observation information from the Mesonet website. If those data are not available, or if they indicate errors, the scripts are designed to have Festival audibly alert and prompt personnel to "Check Mesonet Systems."

To create network traffic graphs, Oetiker's (2006) Multi Router Traffic Grapher (MRTG) is used. Mesonet router(s) and switch(es) are monitored by MRTG via the Simple Network Management Protocol (SNMP). MRTG analyzes and monitors SNMP data and automatically produces temporally-relevant statistics.

Two monitoring approaches, active and passive checks, are used with the Nagios system. With active checks, Nagios itself performs checks of the availability of servers and services. For passive checks, Perl scripts on individual servers are written to monitor services, products, etc., in a highly customized manner and to report back to the Nagios software running on a centralized server. These scripts are designed to send results of their checks via NSCA, or the Nagios Service Check Adaptor. While the scripts can themselves signal a problem to Nagios, if the software does not hear from a script within a configurable amount of time an alert is also generated.

The health of Mesonet servers, all of which are Dells, is monitored by the manufacturer's OpenManage Server Administrator, or OMSA (Dell 2008). OMSA monitors data about the servers, including information on their fans, intrusion attempts, memory, power supplies, processors, temperatures, voltages, hardware logs, and batteries. It also monitors the health of RAID storage devices. While OMSA is available on each server through the internal network via a web browser, custom scripts have been written to query OMSA and send its results back to Nagios, making for cleaner and completely centralized alerting operations.

The external monitoring service from SiteUptime, LLC., is configured to monitor both the availability of the Kentucky Mesonet website and to look for a customized health message on a special status page. If the public website is either unavailable or the customized health message indicates a problem, Mesonet personnel are sent e-mails and text messages.

6.8.3 Needs for improvement

Availability assurance mechanisms are already fairly robust as currently developed. Improvements may include creating some map-based reporting functions for observation sites. Though a delicate matter, the possibility of reducing the number of alert messages generated may also be investigated.

6.9 Automated quality control system

While the last section focused on maintaining the overall quality of the network, including its servers, product availability, and communications systems, this final section of the chapter examines the Mesonet's automated quality control system. The entirety of the reviewed literature on in situ surface networks stresses the importance of an overall quality assurance program and quality control of network data. From its beginning the Mesonet has utilized manual quality assurance (QA) and control (QC) techniques and inspection implemented by the program's QA Specialist and the student operators overseen by the specialist (Ferris et al. 2010). At the time of writing, an automated QC system had been developed in Java (Sun 2010) and was being carefully woven into operational use. The system could possibly form the basis of a thesis of its own and is described generally but at some length here. The author designed and coded approximately 90% of the system, with 10% of the effort coming from a student developer. Former Application Developers contributed some ideas and advice on the system.

6.9.1 General overview

The automated QC system has been designed to handle multiple types of QC, defined by the Mesonet to include real-time checks on measurements in a single observation, hourly QC on an hour's worth of data, daily QC on a day's worth of data, and so forth. The base system has also been designed to handle spatial QC with some modifications. At this time, however, upper level program management and the QA specialist have decided to apply only real-time QC to observations from individual sites. Some details behind this decision are in Ferris et al. (2010). Though the general overview of the system only details its real-time aspects, the technical implementation section examines the entirety of current system capabilities.

As Figure 6-18 shows, real-time QC is designed to run a set of automated algorithms on multiple variables that are part of a single observation. The following algorithms are executed on each five-minute observation:

(i) Uncertainty Alter – Values of relative humidity (RELH) and solar radiation (SRAD) are checked to see if they are outside of physical reality. Specifically, RELH values greater than 100% are truncated to 100% and values of SRAD less than 0 are changed to 0. A final QC value, not the original value, is modified but only if it is still within instrument range and accuracy. Otherwise, the next algorithm is used to catch the bad value.

(ii) Range Check – 1.5 m air temperature values (TA01-3); relative humidity (RELH) and relative humidity sensor temperature (THMP); 10 m wind speed (WSPD), direction (WDIR), peak speed (WSMX), direction at peak speed (WDMX), minimum speed (WSMN), direction standard deviation (WDSD), and speed standard deviation (WSSD); solar radiation (SRAD); and five-minute precipitation accumulation (PRCP) are all checked against a known and expected range of values unique to each variable.

(iii) Fan Comparison – Air temperature values are validated against proper operation of aspirated shield fans (FAN1, 2).

(iv) Vaisala Health – Five-minute precipitation (PRCP), total mass (PMAS), total accumulation (PACC), and intensity (PRTE) are validated against a health diagnostic value (PHTL) provided by the weighing bucket gauge's central processing unit.

(v) PRT Intercomparison – Three separate air temperature values are inter-compared for consistency.

(vi) Precip Vs Wetness – Precipitation values are checked against wetness sensor values (WET1) from the previous 20 minutes. WET1 essentially indicates if precipitation was or was not actually occurring.

(vii) Door Check – Multiple variables, the same ones examined in the Range Check algorithm, are validated against the presence of a site technician.

Figure 6-18 shows the general progression of algorithm execution from top to bottom. Care must be taken to not run multiple algorithms on the same variable at the same time for the same site. However, the system has been designed to use multi-threading capabilities to take advantage of situations where multiple algorithms can be run on different variables simultaneously[22]. Algorithms grouped together in the figure are executed in parallel, as the variables they each examine are not common between them.

[22] at least pseudo-simultaneously, depending on operating system and processor assignment

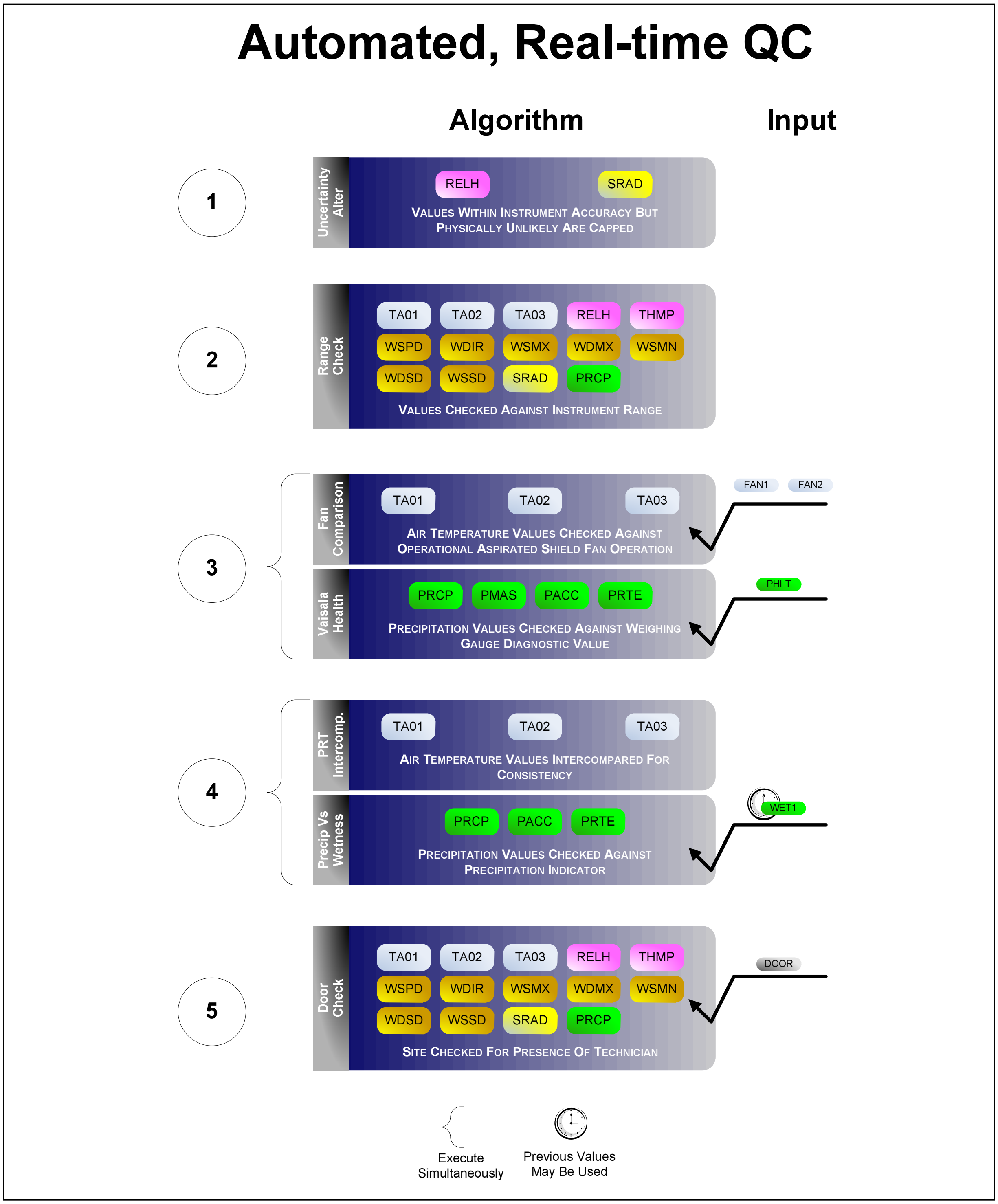


Figure 6-18. Automated, real-time QC algorithms. Variables are identified by a four-letter abbreviation.

The result of each algorithm is a set of QC flags for each variable. The Mesonet's QA Specialist defined the following four tiers of flags:

(i) GOOD – variable passed check by algorithm;

(ii) SUSPECT – variable considered suspect by algorithm, though it may still be used in data displays and calculations;

(iii) WARNING – variable integrity considered to be very questionable and its value may not be used in data displays and calculations;

(iv) FAILURE – variable integrity considered negligible, possibly due to complete instrument failure.

The flags returned by each algorithm along with a description of algorithm results are logged in the observation database. The worst flag for a variable from all algorithms becomes the final flag assigned to it in the database. As noted in the algorithm list, select algorithms can modify a variable's final QC'd value – which will be used in display and calculations – but not its original value. Table 6-3 shows example algorithm results, where WARNING flags were set for all data at the "LSML" site for 1930 UTC 6 January 2010. At that point, the site's datalogger enclosure had been opened to indicate the presence of Mesonet personnel who were giving a site tour.

Table 6-3. Sample results from the Kentucky Mesonet automated quality control system.

STID	UTME	Variable	Algorithm	Flag	Reason
LSML	1/6/2010 19:30	RELH	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WDSD	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WDIR	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	TA02	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WSMN	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	THMP	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	SRAD	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WSMX	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	TA03	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WSSD	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WSPD	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	PRCP	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	TA01	DoorCheck	WARNING	DOOR open
LSML	1/6/2010 19:30	WDMX	DoorCheck	WARNING	DOOR open

The automated QC system creates some vulnerability in the network data flow, as it is essentially intended to be the gatekeeper between measurements made at network sites and the publication of those data via other systems. If a serious, unchecked system-level error were to occur, the QC system could bring critical Mesonet operations to a halt. Therefore, it has been in testing before initial release for the last five to six months.

Automated QC is now considered fully developed in its initial form and may be judged as such, though all operational systems have not yet been modified to take advantage of it. This process is likely to have been completed or be well underway at the time of thesis defense. The remainder of this section examines the technical implementation of the system.

6.9.2 Technical implementation

To make the most out of Java'a multithreading capability, the automated QC system – a "console" application – is implemented on its own dual core four processor server with 8 GB of RAM. It maintains remote connection to and works with the observation database via classes in the Mesonet data access code layer, previously described. While only four processors and eight total cores do not allow all QC algorithms to truly be run simultaneously, the multithreaded approach taken in the design of the system does allow many simultaneous operations with an aim of increasing performance.

At startup, the application's main class, [23]QaStartup, checks to make sure it is the only instance of the system running, resets a database-backed algorithm processing queue, initializes an XML-based configuration management object, and then uses an ExecutorService in Java's concurrency packages to launch the application's main threads in a cached thread pool. Figure 6-19 depicts those initial threads and begins to show how they, in turn, create and manage other threads.

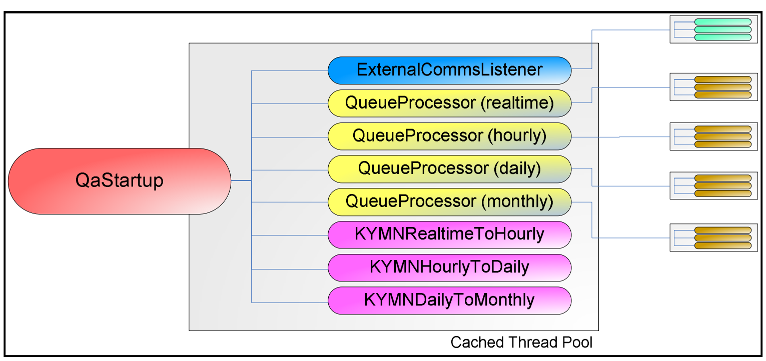


Figure 6-19. Automated QC system's startup threads.

[23] Instead of QaStartup, this should really be named QcStartup, but is described as implemented.

The threads started via the QaStartup class are critical application management threads and are intended to be continuously available. The QaStartup class, therefore, monitors each of these threads for availability and restarts any should they fail. The application's error handling mechanisms should prevent this, so this functionality is implemented as a safeguard. Each of these startup threads are eventually discussed in some fashion in the remainder of this section.

The ExternalCommsListener thread / object is responsible for monitoring and handling external TCP connections to the system, which may be made either from the data ingest system's ldmp2db process or via basic telnet-type connections from other allowed systems. As shown in Figure 6-20, the ExternalCommsListener creates a separate ExternalCommsConn thread – up to 10 total – to handle each incoming connection.

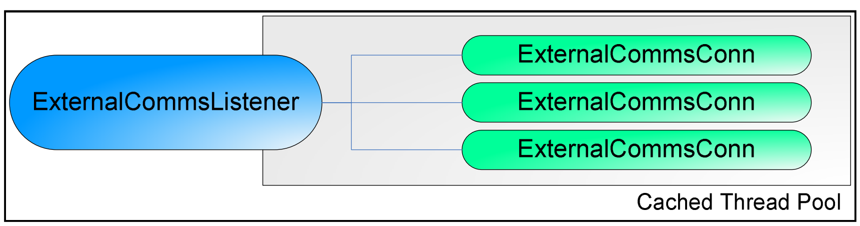


Figure 6-20. ExternalCommsListener and related ExternalCommsConn threads.

The ExternalCommsConn objects handle two types of messages passed to the QC system, both data availability or "sched" messages and QC system control or "command" messages. Sched messages are specially formatted JavaScript Object Notation (JSON) messages sent by the ldmp2db ingest process to indicate the availability of new data in the observation database for real-time QC processing. The ExternalCommsConn thread uses two additional classes, the JSONToProcessQueueMember and QueueInserter, to schedule this processing with the process queue.

Command messages, also JSON-based, are used to remotely control the operation of the QC system. The JSONToQueueCommand class is used to set what are essentially Boolean switches monitored by other objects and threads. These switches control the starting and stopping of QC processing and also the re-reading and re-initialization of configuration files.

Running in the same thread pool as the ExternalCommsListener are multiple QueueProcessor threads, one each for real-time, hourly, daily, and monthly QC. All of the threads are objects of the same class but their behavior is determined by a parameter set by QaStartup. With some helper classes, each QueueProcessor monitors the system's processing queue for data scheduled for its specific QC type. As Figure 6-21 shows, the QueueProcessor sets off a Sequencer thread for each member of the queue available for processing.

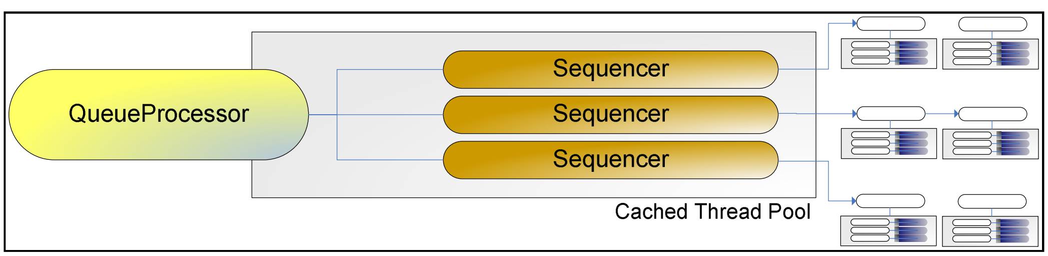


Figure 6-21. QueueProcessor thread.

Because the number of members in the processing queue can quickly build up, the queue is implemented as a local MySQL (Sun 2008) database, which consists of just two tables. One contains the QC type (real-time, etc.) to be run and the time for which it is to be executed. The other holds the network sites for which the QC should be processed. Implementing the queue on a disk-based database versus in memory allows the queue to survive system shutdowns.

The Sequencer threads, shown in Figure 6-22, are used to actually kickoff individual algorithms like those in Figure 6-18. The Sequencer works with the system's XMLManager and associated classes, which manage XML-based configuration files, to determine the appropriate order of algorithm execution. The Sequencer steps through what are known as SequenceGroups and SequenceElements in the XML. SequenceGroups essentially correspond to the numbered items in Figure 6-18.

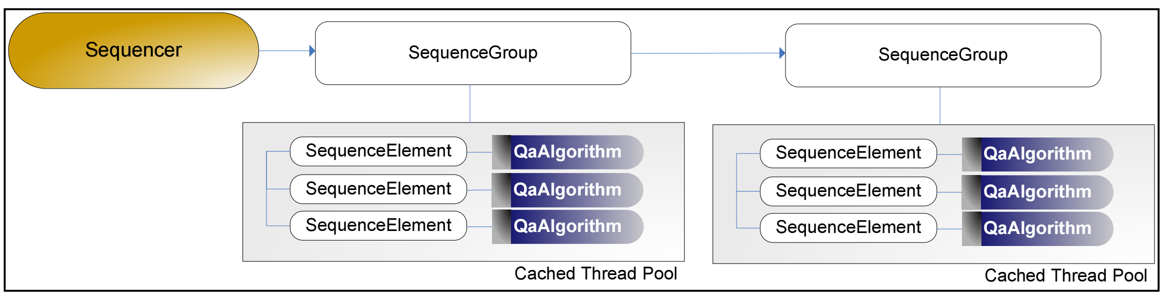


Figure 6-22. Sequencer threads.

For each SequenceGroup, the Sequencer adds to a cached thread pool the appropriate QaAlgorithms[24] defined by SequenceElements in the XML. Each algorithm in the group is allowed to run simultaneously, and the Sequencer waits for each to complete before stepping to the next SequenceGroup. Based on the results of each completed algorithm, a QaFlagAndStatusSetter object from the data access library is used

[24] These, too, should be more appropriately named QcAlgorithms.

to set flags in the observation database. Another class handles any modified QC values. When all SequenceElements in each SequenceGroup have been exhausted, the Sequencer then sets a status indicator in the observation database to mark a particular observation as having completed a particular type of QC.

Individual QC algorithms are all designed to implement an author-developed Java interface, QaAlgorithm, so that they can all be expected to behave in a uniform manner in terms of how they are initialized & executed and in terms of how they return flags and modified data. Upon initialization, each algorithm is provided details about the site(s) and variables for which it is executing. Each has full access to the observation database via the data access library, which allows the algorithm to retrieve all data needed for execution. Figure 6-23 gives a flowchart for a representative algorithm.

While all of the processing threads are running, a number of QC staging threads known collectively as "chainers" are also executing. These threads are shown as KYMN-RealtimeToHourly, -HourlyToDaily, and -DailyToMonthly in Figure 6-19. Because of the way the Mesonet retrieves its data (Section 6.3) in temporal order, the measurements for a particular observation can be seen as setting off a chain of events shown in Figure 6-24 that continues until all QC processing steps have been run.

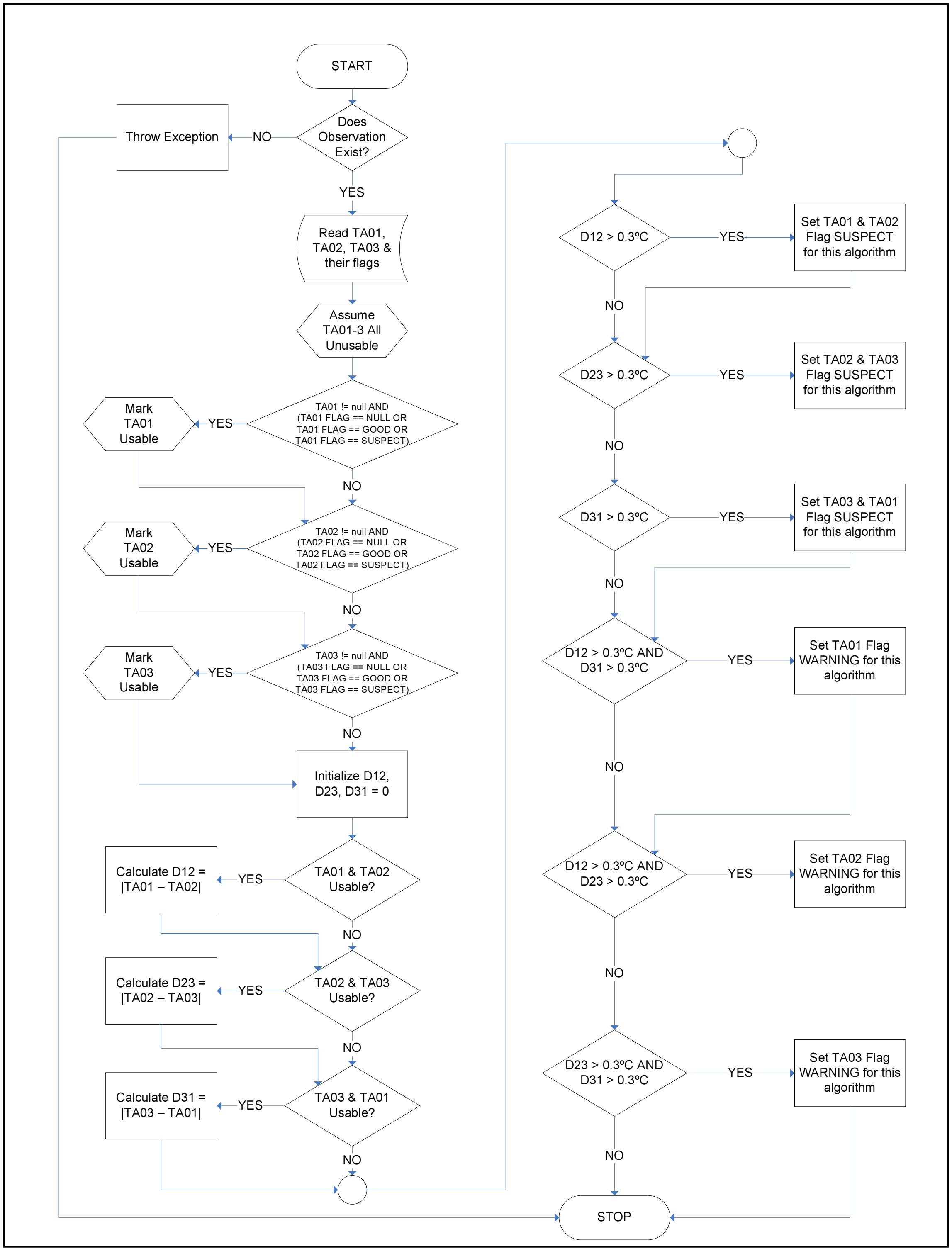


Figure 6-23. Flowchart of PRTIntercomparison QC algorithm.

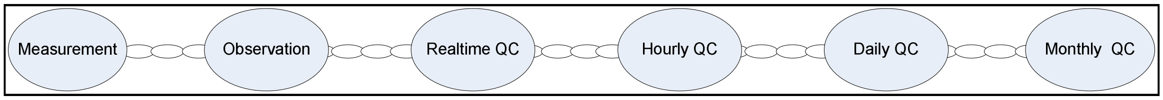


Figure 6-24. Conceptual drawing of the Mesonet data chain.

The "chainer" threads are used to populate the QC processing queue for each QC type other than real-time. For example, the RealtimeToHourly chainer constantly[25] monitors the status flag of all observations in the observation database to determine which ones have completed real-time QC. Because data are known to only be collected and subsequently processed by the QC system in temporal order, as soon as the first observation for, say, the second hour of the day has completed real-time QC, the chainer knows it can schedule hourly QC for the day's first hour.

Several references to the data access library and the observation database have been made in describing the automated QC system, as they are critical to its operation. Some major components of the Java-based Mesonet data access library were developed specifically to support the QC system. Since constant creation and destruction of connections to the observation database by individual algorithms would be too computationally expensive, classes in the data access library were created to manage a pool of connections which are shared among threads. The PerThreadDataCon class is used to establish and manage database connectivity to the observation database from a connection resource pool, coded by the author, using a connection-per-thread implementation model. For the description of the class, it is important to remember that a

[25] The period between executions is based on the type of chainer.

thread may consist of multiple objects that require database access, but since they are running in the same thread they cannot simultaneously access it.

Given the modular nature of application programming for the Mesonet, the data access library must allow any Java object to request access to the observation database. Furthermore, having the initial Runnable object in a thread establish a database connection and pass it to subsequently created objects would break modularity. This creates an impasse where database transactions are concerned, as a transaction can only be wrapped around database queries made on the same connection. PerThreadDataConn alleviates this impasse by using ThreadLocal constructs to ensure that the same database connection is used by all objects running in a thread, even if those objects have no knowledge of each other. This allows the initial Runnable object in a thread to wrap in a single transaction all queries made by the objects it creates – even if it has no direct knowledge of those queries – and to rollback them back should an Exception occur.

Speaking of Java Exceptions, because of the critical nature of the automated QC system and its position in the network dataflow, robust error handling and logging features were included in the design. Error handling mechanisms are coded to keep the system running in the event of disconnects or other failures. The logs produced by the system are monitored by the availability assurance mechanisms discussed in Section 6.8 and Mesonet personnel are notified of any unexpected errors.

CHAPTER 7. GIS DEVELOPMENT

While the initial Kentucky Mesonet website and data access methods reviewed thus far are interactive, they lack graphical spatial interaction in terms of user-defined visualization domains, base maps, and data layers. To rectify this, the author has developed or assisted in the development of several geographic information system (GIS)-based interactive visualization tools, two of which are in use operationally and a third whose development was being finalized at the time of this writing. As GIS courses have been an important part of the author's degree work, in addition to his core courses and research, and as the developed systems are an important contribution to the Mesonet, each of them is reviewed in this chapter.

7.1 KEMAP & Kentucky Weather Mapping application

The Kentucky Event Mapping and Analysis Portal (KEMAP) and the related Kentucky Weather Mapping application (KY DGI 2009a, b) were the low hanging fruit, so to speak, in the development of GIS tools for Mesonet data display, as the Kentucky Division of Geographic Information (DGI) developed base services and hosts the applications.

7.1.1 General overview

Available online and now based on Adobe's Flash and Flex technology (Adobe 2010), KEMAP allows critical state government agencies to display Mesonet data alongside layers representing critical infrastructure, such as oil & gas pipelines and electric utility transmission information. This marriage of critical information promises to be invaluable to KEMAP users, which include the Kentucky National Guard, Emergency Management, State Police, Homeland Security, and other entities.

A similar public tool, the Kentucky Weather Mapping application, has been stripped of KEMAP's restricted datasets but still includes many useful layers such as political boundaries, transportation, topography, and aerial photography. In addition to Mesonet weather data, the application includes National Weather Service radar data & severe weather information, traffic webcams from a variety of applications, and Kentucky 511 road condition and construction alerts.

The public application is available on-line at http://kygeonet.ky.gov/kyweather/ and allows for selection of displayable layers and a few geoprocessing tools from a graphical menu. Mesonet sites are symbolized by the Mesonet logo and their data are displayable in list format upon a "mouseover" of the logo. Figure 7-1 is a screenshot from this application showing Mesonet data over aerial photography, which is part of DGI's Commonwealth Map dataset (KY DGI 2009c).

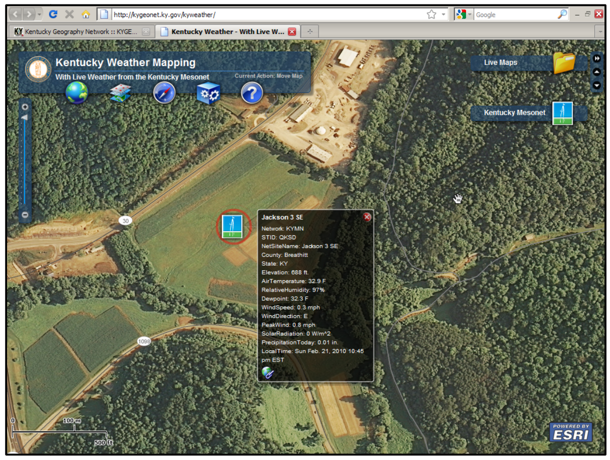


Figure 7-1. Mesonet data in the KY Division of Geographic Information's Kentucky Weather Mapping application. Site (QKSD) shown is 3 miles south of Jackson, KY in Breathitt County (KY DGI 2009b).

7.1.2 Technical overview

As previously mentioned, DGI did the majority of the heavy lifting for these applications and hosts them on state government operated web and GIS server platforms. Though KEMAP has long existed as an ArcIMS (ESRI 2008) application, users are being migrated to a new version based on ArcGIS Server and Adobe Flash/Flex technologies. Weather data are only available in this new version of KEMAP and the Kentucky Weather Mapping application exists solely as a Flex application. Initial application design by DGI was based largely on ESRI-provided templates supplied by the unfortunately named company Weather Underground, Inc. Mesonet data are provided to DGI via an author-developed PHP application (PHP Group 2009) in an XML-based GeoRSS (GeoRSS 2010) format, which is described in Section 6.7. DGI connects to this utility via an HTTP request once every five minutes then re-hosts the XML data on its own server. Its Flex applications are designed to access, decode, and display these GeoRSS-based data.

7.1.3 Possible improvements

As far as Mesonet data are concerned, the most significant possible improvement to KEMAP and its related public application would be the direct plotting of observation values on the map rather than solely in the summary box displayed upon mouseover. As the Mesonet is only responsible for the GeoRSS feed, though, modifying these two applications to include this capability will have to be the responsibility of DGI.

7.2 ArcGIS Engine & Objects application

To support spatially-based inspection of both current and archived data by internal Mesonet personnel, a custom desktop application based on ArcGIS Engine & Objects technologies (ESRI 2009a, b) called the “Simple Data Viewer” (SDV) was developed. That application is described below.

7.2.1 General overview

The SDV application allows for the display of Mesonet observational data over a set of both locally-stored and on-line base maps. Local map layers consist of:

(i) county & state outlines and county name annotations;

(ii) hydrography polygons;

(iii) boundary between Eastern & Central time zones;

(iv) populated places data & corporate boundaries;

(v) transportation data, including local & state roads, U.S. highways, Kentucky parkways, interstates, and active railroads;

(vi) National Weather Service county warning areas and River Forecast Center basin IDs;

(vii) and elevation and hillshade data.

Remote mapping data, provided by the Kentucky Division of Geographic Information's Commonwealth Map (KY DGI 2009c), provides a wealth of additional layers including hospital, school, and other structure points; landcover; and orthophotography.

Figure 7-2 depicts the user interface for the SDV. Map navigation and simple measurement utilities are provided in the toolbar (1) while map layers are selectable from the table of contents (2). Selection of observation times and variables are via the list boxes in (3) and (4). Control of text and station marker size and the amount of rounding applied to the data are available from controls (5) through (9). The "Arrows" checkbox (10) allows for the display of wind direction arrows, while the "Latest" checkbox (11) overrides the time from (3) and displays the latest data for each Mesonet site. Finally, the buttons in (12) allow for the user to step through time in predefined increments. The latitude and longitude of the mouse pointer are given in a text dialog (13), while data timestamp(s) are provided in (14). Mesonet data are also exportable by right clicking their table of contents entry in (2).

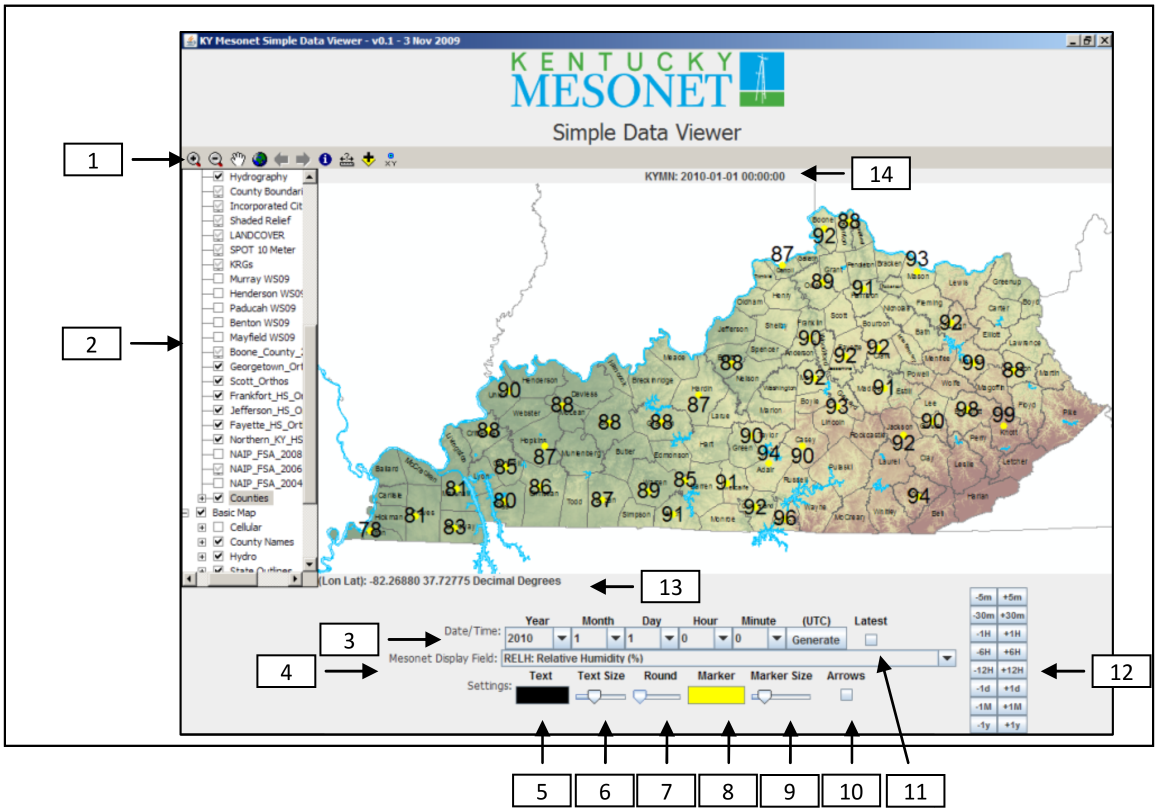


Figure 7-2. ArcEngine & Objects-based "Simple Data Viewer" application.

7.2.2 Technical overview

The SDV application was built using ESRI's ArcGIS Engine 9.3.1 (ESRI 2009a) for Java (Sun 2010) which is based on its ArcObjects software (ESRI 2009b) component libraries – the same components that are the foundation of all ArcGIS products. The Eclipse (Eclipse 2009a) integrated development environment (IDE), which is shown in Figure 7-3, was used to design and code the application. Graphical design was aided by the Visual Editor (Eclipse 2009b) plugin for Eclipse, which allowed for straightforward layout and "wiring" of native Java Swing components & controls and ArcObjects map-related controls exposed as Java Beans. Java Beans are reusable software components designed for easy manipulation in tools like the IDE and Visual Editor.

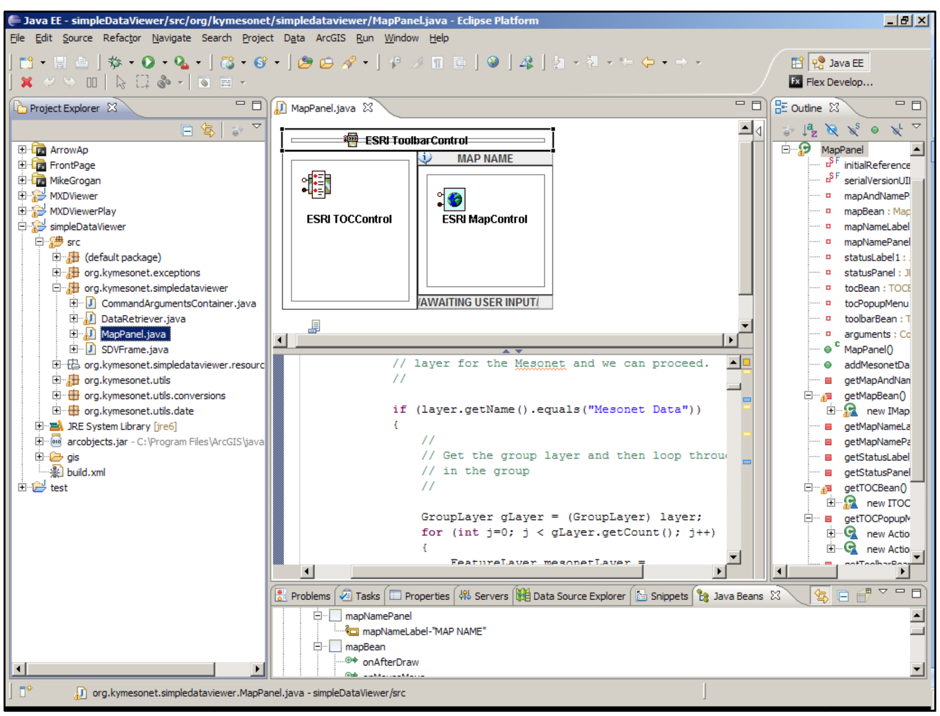


Figure 7-3. Eclipse Integrated Development Environment (Eclipse 2009a).

Following object-oriented (OO) coding practices, the application was designed modularly using multiple classes, each with its own semi-independent functionality. The application's main class, SDVFrame, creates the main layout for the application and also handles all user input from the interfaces in Figure 7-2, except for the map toolbar (1) and table of contents (2), which are handled by ArcObjects (ESRI 2009b) controls. The application's MapPanel class is used to encapsulate all map functionality and is built from a number of ArcObjects controls exposed as Java Beans. The DataRetriever class, unsurprisingly, is used to actually retrieve Mesonet observational data. Finally, a convenience class called the CommandArgumentsContainer serves as an application-wide container of command-line arguments supplied when the application is started which, admittedly, probably breaks a bit from true OO programming philosophy. A number of other Mesonet-developed Java code packages from various libraries are used in the SDV, including packages related to error handling, general formatting utilities, unit conversions, meteorological calculations, and time.

Data for the SDV are retrieved from the Mesonet web server via the DataRetriever class. Using separate but similar code to that used to populate KEMAP in Section 7.1, the server returns an XML-based data array which is parsed by the class. Unlike the data for KEMAP, the XML data returned for the SDV include a number of diagnostic variables such as battery voltage and a door flag, which is used to indicate the presence of Mesonet technicians at a particular site.

Though the application may be used cross-platform, an installation program was created for the Windows (Microsoft 2007) platform via the Launch4J utility (Kowal 2008). While this helps streamline the installation process for the SDV, it is critical to note that the ArcGIS Engine 9.3.1 (ESRI 2009a) Runtime for Java (Sun 2010) and Windows (Microsoft 2007) must be installed before the SDV can be used.

7.2.3 Technical details

To begin the development process, the ArcGIS Engine 9.3 Software Development Kit (SDK) for the Java Platform on Windows was installed. A license for the kit was obtained through the ESRI Enterprise Developer Network, to which the author subscribes via authorization available to Western Kentucky University through a state higher education licensing agreement. Along with ArcGIS Desktop applications on the author's computer, the SDK was then updated to version 9.3.1. In order to use the SDK in the Eclipse IDE, the Visual Editor plugin was then installed.

There are a few interesting SDV design decisions to highlight. First, upon startup, the SDV is designed to consume a license at the ArcView level using whatever license manager is setup on the local PC. This allows for Mesonet personnel to use the existing WKU GIS license server instead of licensing individual instances of the ArcGIS Engine Runtime. After retrieving data in XML format from the Mesonet server, the DataRetriever class creates an <u>in memory</u> point Feature Class via an ArcObjects InMemoryWorkspaceFactory; observational data are set as attributes to the points, which are created from latitude/longitude data of Mesonet sites.

The MapPanel's updateMesonetLabels method is used to symbolize the numerical observation values and has to include a kludge to handle missing data, as null values are not allowed in an ESRI point Feature Class. The method also includes a workaround for zero-sized labels, which are apparently not allowed in ArcObjects; they are converted to fully transparent labels instead. Code is also included to convert from the Mesonet's standard units to English units.

To symbolize wind directional arrows the updateMesonetMarkers method utilizes a UniqueValueRenderer combined with rotation effects to draw markers. When wind arrows are not being displayed, a SimpleRenderer is used to draw a basic circle denoting site location. The method includes a similar workaround as above but for zero-sized symbols.

Finally, as the displayed map may be in a variety of projections, map coordinates determined using the onMouseMove event dispatched from the map control are projected dynamically during mouse motion into NAD1983 geographic coordinates before they are displayed in the application.

7.2.4 Possible improvements

Improvements to the SDV may be realized by implementation of a better asynchronous event dispatching mechanism, coupled with a data or map loading progress bar, which would solve some problems with perceived freezing of the application in its current form. Additionally, the interface could be modified to allow for the selection of data display units rather than forcing an English-only display. Finally, full implementation of image export functionality should be added.

7.3 Interactive web data display

Intended to add map-based interactivity to the Mesonet website and replace the GEMPAK product generation described in Section 6.5, an ArcGIS Server application was developed based on Adobe's Flash and Flex (Adobe 2010) technology and ESRI's application programming interface (API) for Flex (ESRI 2009c, d). At the time of this writing, the tool was being tested and incorporated into the Mesonet's public website but was not yet fully implemented there. While some changes are likely before operational release, the utility's underpinnings should remain the same and are, therefore, described below.

7.3.1 General overview

This Flex-based GIS application allows public data users to quickly visualize and explore the majority of current network data over a variety of base maps, including custom Mesonet layers, street maps, and aerial imagery. As the user moves the mouse over plots of data on the map, shown in Figure 7-4, the application's "blue box" (1) is populated with data from individual sites positioned under the mouse cursor. In addition to using a mouse scroll wheel or a rubber-band zoom box (shift key + left mouse button), users can zoom the map in or out using the slider control shown in (2). The variable plotted on the map is user-selectable via the array of buttons shown in (3). Finally, the user can switch between base map types – such as basic, street, and imagery – with the buttons in (4).

For the majority of data types a simple numerical value is displayed. However, for wind data, a numerical plot of the speed is given along with a wind vector, which is drawn parallel and in proportion to the wind at each particular site. Data are not plotted on the map if they are older than 30 minutes. Instead, a simple diamond marker symbol is drawn in their place, though they are still shown in the blue box, which changes its background to red. No data older than one hour are shown there. Data are automatically refreshed once per minute.

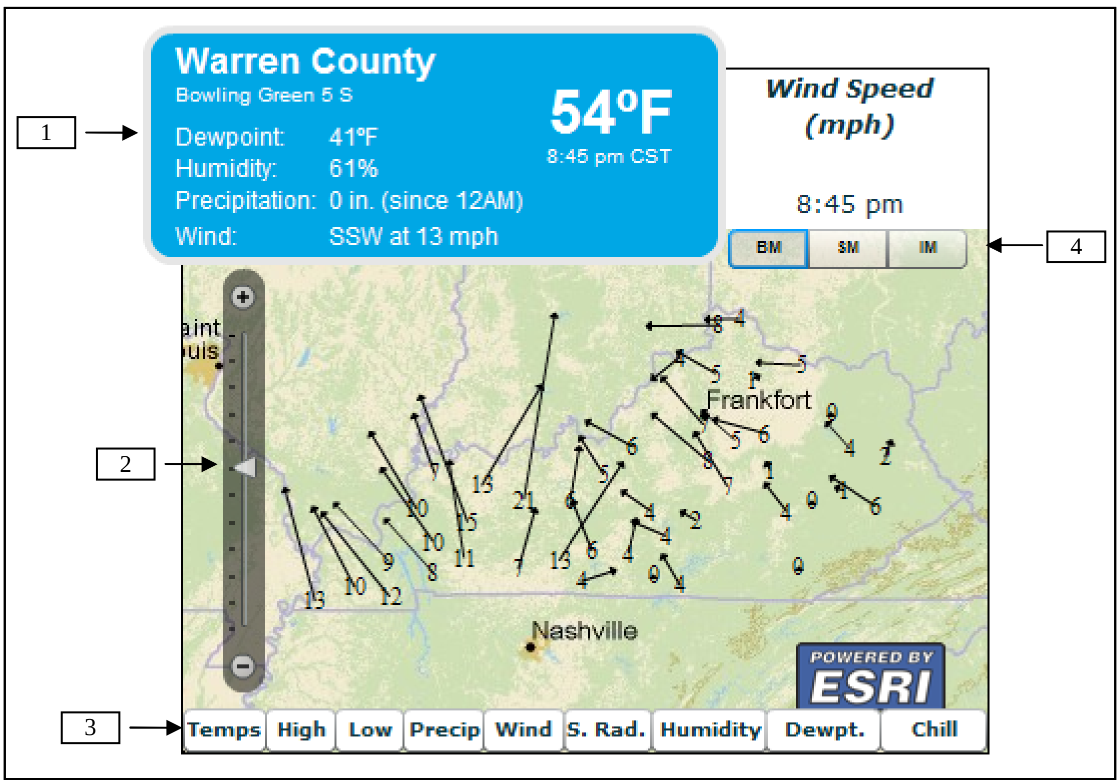


Figure 7-4. Flash and Flex-based interactive data mapping application.

7.3.2 Technical overview

The interactive web mapping program was developed on the Adobe Flex (Adobe 2010) platform, which allows for consistent application deployment across disparate platforms, be they different browsers, operating systems, etc., by creating applications "playable" in Adobe's Flash player. ESRI provides a freely available API (ESRI 2009c) for the Flex platform which allows for the access and navigation of a variety of remote map services such as those hosted on an ArcGIS Server (ESRI 2009d). However, the Flex development environment is not free, though Adobe provides it at no charge to faculty, staff, and students of academic institutions; a Flex license was obtained through this program.

The Flex development environment is based on the Eclipse IDE (Eclipse 2009a), the same development tool used to create the Java-based Simple Data Viewer application in Section 7.2. Instead of Java, however, Flex is based on the ActionScript language and the [26]Macromedia (MXML) markup language, a convenience language whose commands are eventually converted to ActionScript. Though not the same, the ActionScript language is very similar to Java with some syntactical differences. Coding in ActionScript "feels" very much like coding in Java.

An object-oriented approach was used to design and code the Flex application. The main application component, FrontPage.mxml, establishes the layout for the application and ultimately controls, via dispatched events, interaction between the map, the blue data box, and the buttons used to select mapped data type. It also requests and handles observational data from Mesonet servers, which are returned in XML form. The

[26] Flash technology was purchased from Macromedia by Adobe.

BlueBox.mxml component creates and handles, not surprisingly, the blue-colored individual station data display, while MapBox.mxml implements the actual interactive mapping controls. A utility class, WindVectorGenerator.as (.as = ActionScript) is used to create vectors for the wind display. Some ESRI-supplied code is used to draw the arrowheads on the vectors. CoordinateTranslator.as is a utility class used to convert between a screen coordinate system with origin based in the upper-left-hand corner and an origin centered on the middle of the application screen.

To satisfy certain requirements for an independent study course, a basic tiled mapping service was created on an author-administered ArcGIS 9.3.1 Server (ESRI 2009d) for Linux (CentOS 2008) installation. A screenshot showing output from this service within the application is given in Figure 7-5.

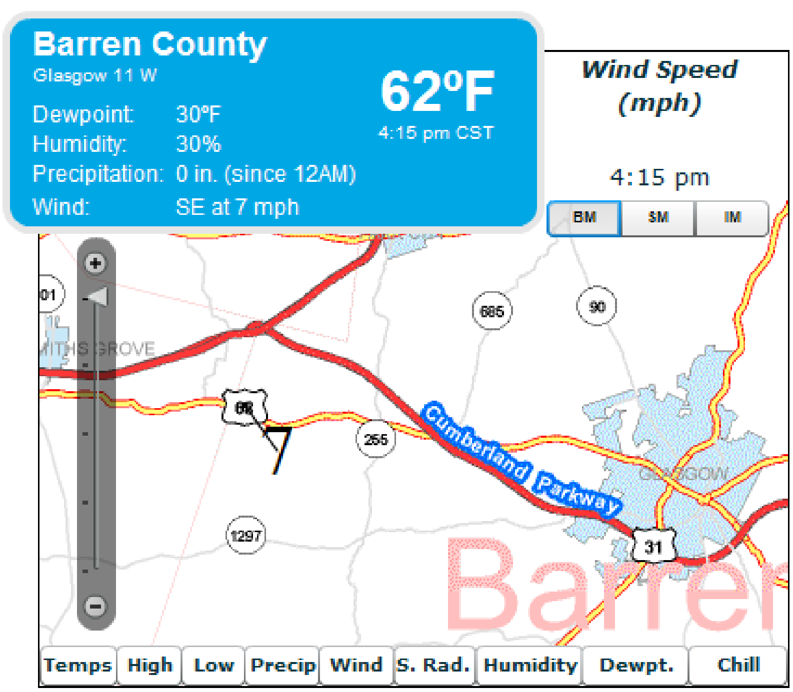


Figure 7-5. Author-created tiled mapping service consumed by Flex application.

The tiled mapping service was setup on an author-controlled development box in the main Mesonet office. To utilize Mesonet-created base mapping layers on the operational website, an ArcGIS Server instance will need to be installed on operational Mesonet servers. At the time of this writing, base maps freely available to non-commercial users from ESRI's ArcGIS Online services were being utilized.

7.3.3 Technical details

While the majority of Flex code authored for the application was straightforward, two classes proved to be an interesting challenge. The WindVectorGenerator.as and CoordinateTranslator.as classes used in the drawing of wind vectors step through a potential mine field of coordinate transformations, from map coordinates to screen coordinates to a regular Cartesian coordinate system with origin at center screen to geophysical coordinates and back again!

As the Mesonet utilizes only Linux operating systems for its operational servers, the author chose to install ArcGIS Server for Linux, which is coded in Java. This was a painstaking process that included not only installation of base Server 9.3 applications but also patching and updating of those applications to work with RedHat Linux 5.0[27] (RedHat 2008) at the ArcGIS 9.3.1 release level.

To create the tiled mapping service, mapping data and layers were first added to and symbolized in an ArcGIS Desktop MXD file on a Windows machine (Figure 7-6). A Kentucky state outline, outlines of other states, county name annotations, roads, city

[27] Note: CentOS Linux is a binary-identical derivative of RedHat Linux. ArcGIS Server installation was performed on a CentOS machine.

borders, and county polygons were all added to a single data frame in the MXD, whose full extent was zoomed slightly out over Kentucky to keep the server from having to create tiles for the entire U.S. In keeping with recent changes to ESRI's free online basemaps, the custom maps were projected in WGS_1984_WEB_MERCATOR_ AUXILIARY_SPHERE. Because the Windows computer and the Linux-based ArcGIS Server did not share a common network file system, the MXD files and related data were manually copied to the Linux machine.

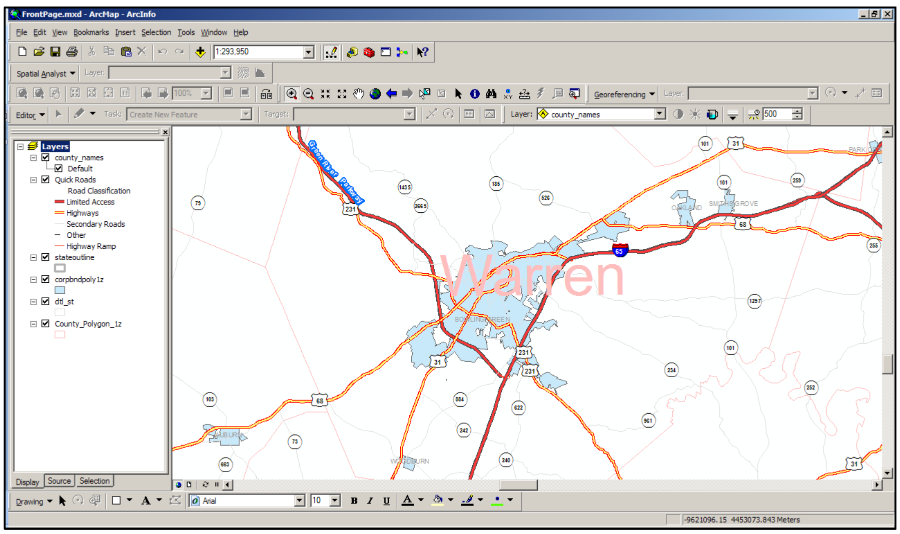


Figure 7-6. Map creation for tiled map service. (Source: ESRI (2009e) ArcMap software).

In the web-based ArcGIS Server Manager application, a new mapping service (frontpgserver) was then created from those files and its capabilities were set to "Mapping" and "KML". A pooled service was created using the default values of a minimum of one instance and a maximum of two. As shown in Figure 7-7, the map service was then edited to create a cache of map tiles at a few representative scales.

Though a remote Windows-based ArcCatalog instance can be used to kickoff tile generation[28], it can also be started via Linux shell scripts in the java/tools/caching application subdirectory of the ArcGIS Server installation. In this example, the ManageMapServer CacheTiles.sh script was executed to create the tiles which, for the small amount of mapping data, areas, and scales, took only two minutes to execute.

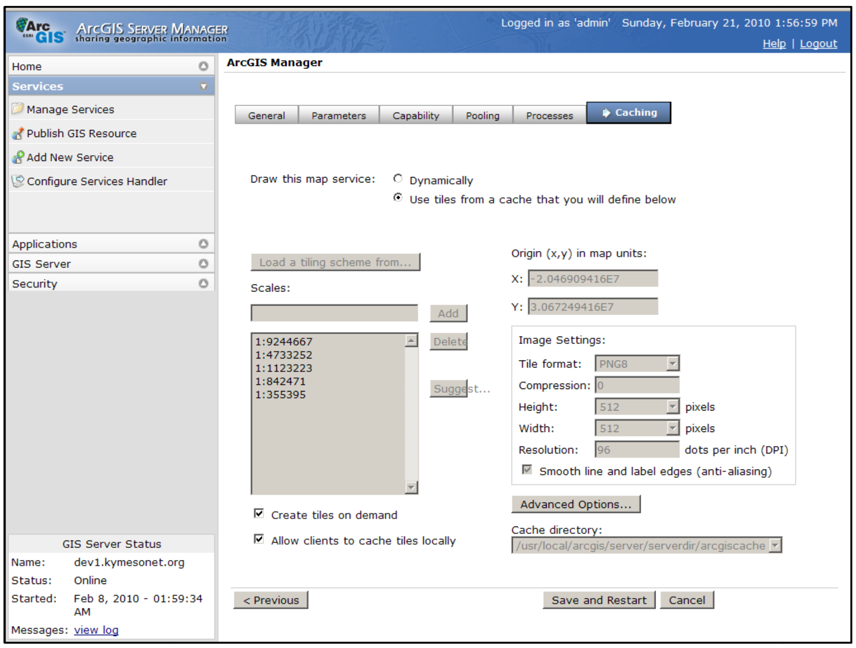


Figure 7-7. ArcGIS Server (ESRI 2009d) map tile creation.

7.3.4 Possible improvements

There are a number of possible improvements for the Flex mapping application. First, before being deployed operationally, Mesonet administrators will define the exact

[28] The Windows user's username and password must be added on the Linux machine in ArcGIS Server Manager as an administrator.

look, feel, and functionality they desire from the application. In all likelihood the size of the application will increase and additional data types will be added. Some additional technical improvements will likely include improved error handling and more appropriate scaling of wind vectors.

The biggest area of improvement in the tool, though, must be the eventual inclusion of color raster fields based on Mesonet observations. However, Kentucky's terrain complicates the accurate creation of such interpolated fields and a Mesonet graduate research assistant is currently analyzing best approaches for solving the problem. To generate and utilize the rasters operationally for current data, an ArcGIS Server-based image service may be used in conjunction with server-side interpolation scripts based on ArcObjects. For historical data, an ArcGIS geoprocessing service may be used to recreate rasters on demand.

CHAPTER 8. ANCILLARY SYSTEMS

In addition to the core IT and geographic information systems detailed in the previous two chapters, there are several ancillary systems that also play important roles in the operation of the Kentucky Mesonet. These are reviewed in this brief chapter.

8.1 Version control system

A version control system is used to track and archive every modification to the Kentucky Mesonet codebase, including code using in centralized computing operations and on data loggers in the field. Subversion (Apache 2010b) is the choice of version control software for the program. Like some others, it allows for all versions of code to be forever retrievable. Most importantly, its logging facility (Figure 8-1) can be used by developers to describe the rationale behind code changes. On Linux servers, Subversion command line tools are used, while Windows desktops use Tortoise SVN (Collabnet 2010). Each of these tools connects to a centralized, server-based code repository.

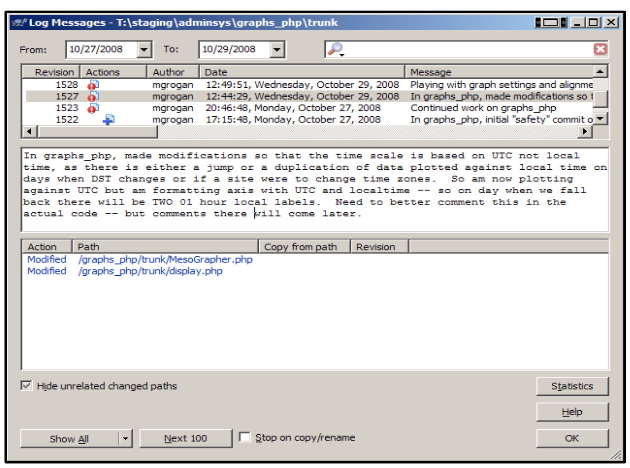


Figure 8-1. Subversion version control system. Shown is TortoiseSVN (Collabnet 2010).

8.2 Time synchronization

As mentioned in Section 3.4.4, time is very important for an observation network. To keep time across centralized and field systems synchronized and highly accurate, the Mesonet operates a Network Time Protocol (NTP) server. The NTP daemon program "sets and maintains the system time of day in synchronism with internet standard time servers" (NTPD 2010). When used properly, the NTP daemon can help maintain a monotonic clock, meaning the time on a server moves naturally forward and is not set backward during the synchronization process. Only the slew, a measure of clock drift, is adjusted.

NTP operates by exchanging polling messages with upstream time servers at specified intervals. The Mesonet time server polls multiple upstream servers operated by both the U.S. Naval Observatory and the National Institute of Standards and Technology (NIST) for time synchronization traceable to NIST. Instead of connecting directly to these upstream servers, other Mesonet servers use the Mesonet time server as their upstream source, as is best practice for NTP.

Field data loggers do not use NTP. However, their clocks are synchronized at least once per day by LoggerNet, whose time is constantly synchronized to the time server.

8.3 Name servers

To provide domain name to IP address resolution (such as www.kymesonet.org to 12.180.242.91), the Mesonet operates multiple Domain Name System (DNS) lookup services running on Berkeley Internet Name Domain (BIND) software (ISC 2010). An internal BIND instance provides name resolution services to internally-networked systems for both external public and internal private domains, such as the sites.kymesonet.org domain. An externally accessible BIND instance provides name resolution for publicly accessible services such as the public website and partner distribution services.

8.4 Local Data Manger

For some external data exchange functions, such as data transport from a NOAAPORT weather data satellite system at Mesonet offices, a Unidata Local Data Manager service is run. The LDM "is a collection of cooperating programs that select, capture, manage, and distribute arbitrary data products" (UCAR 2009b).

8.5 Development hosts

Several development hosts – both in the main Mesonet office and within centralized computing operations – are maintained for non-operational research and development use. These hosts allow for R & D experimentation and code development without impacting operational systems.

8.6 Virtual servers

Some ancillary systems and a couple of operational ones are implemented using VMWare Server, “a hosted virtualization platform … that partitions a physical server into multiple virtual machines” (VMware 2010). Virtualization allows for the operation of multiple hosts, even with different operating systems, on one physical server.

CHAPTER 9. NETWORK USE, BENEFITS, AND PARTNERSHIPS

By 1920, meteorological observations were being regularly taken at more than 200 stations of the U.S. Weather Bureau and by cooperative observers around the country. Even then, as Marvin (1920) notes, the value of such efforts was described as "incalculable" due to their affects on and benefits for "the entire people." To be the most worthwhile, then, the efforts of a met/climate sensing network must have benefits for the people of the area or region it serves. Some of the many comments received by the Mesonet from its users and stakeholders indicate that Kentucky's new in situ surface sensing network is already bringing this kind of benefit:

> *"I wanted to make you aware of how invaluable the Kentucky Mesonet data [have] been during our major winter storm over the past 36 hours. The wind data from the Mesonet stations allowed us to provide better forecasts and services to the taxpayers."*
>
> – Meteorologist-in-Charge, National Weather Service

> *"The first time I saw the Kentucky Mesonet web site I knew we had to find a way to incorporate this data feed into our applications. Having this type of data at the fingertips of first responders is essential, and making it available to the citizens of the state is just icing on the cake."*
>
> – Technical manager, Commonwealth of Kentucky government

> *"The Mesonet … was a TREMENDOUS help during the [high wind] event. I can't wait for the day you have all 100+ up and running!"*
>
> – Warning Coordination Meteorologist, National Weather Service

While the information technology infrastructure supporting the network has been thoroughly discussed in this document, the work done to build it would be pointless without the end use it facilitates and without the partnerships it has helped foster. As this writing nears a close, therefore, this chapter examines the early benefits of the network to the public and to operational meteorology, looks at a couple of representative research uses of the network, and closes by touching on the local partnerships that have made the Mesonet possible.

9.1 Public and operational use

As discussed in Section 6.6, the Kentucky Mesonet website is the main, direct channel by which the general public accesses and uses network data. Figure 9-1 gives the number of website visits and viewed pages per month.

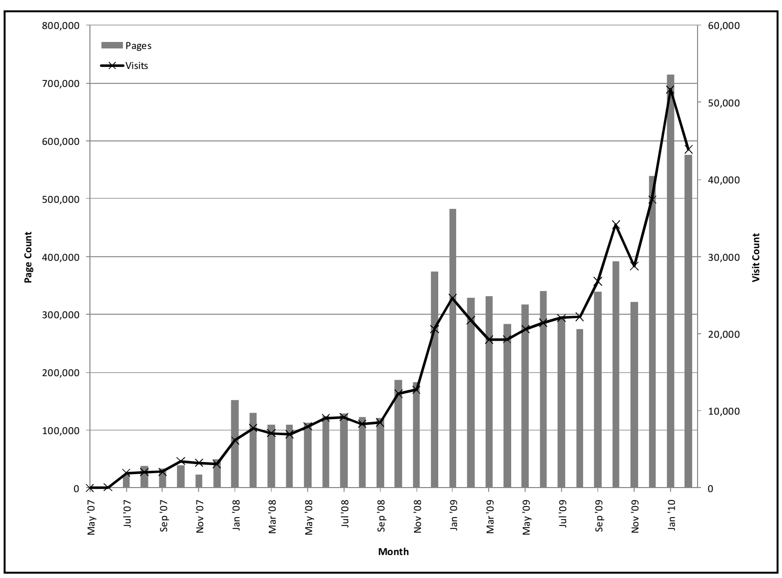


Figure 9-1. Kentucky Mesonet website visits (line) and viewed pages (bar) by month.

Website usage is certainly on the rise, as 2.2 million of the total 7.6 million viewed pages occurred in the four month period ending February 2010. Over thirty percent of the site's half million plus visits occurred in that same period. A visit is defined as a single, user-initiated web session which may consist of multiple viewed pages and which has been active in the last 30 minutes. It is important to note that counted visits do not include content and images that are often displayed on third party websites – even when they are directly served and hosted on the Mesonet website – nor do they include visits to the important weather applications hosted by the Kentucky Division of Geographic Information (Section 7.1). The Webalizer (Barrett 2009) log file analysis program was used to generate the web usage statistics.

Recognizing the unique ability of broadcast media to reach and benefit the citizens of Kentucky, the Mesonet maintains data usage agreements with television stations in multiple broadcast markets, as discussed in Sections 6.5 and 6.7. Figure 9-2 shows a typical way in which network data are shown to users by the broadcast media.

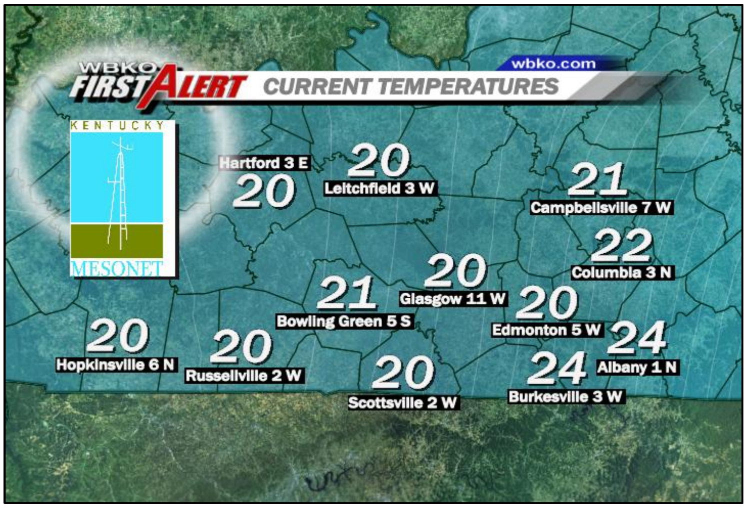


Figure 9-2. Kentucky Mesonet data as shown by a broadcast weather display system. (Source: WBKO Television, Bowling Green).

As evidenced by the quotes which opened this chapter, a critical way in which network data are utilized to benefit the public is through their use by the National Weather Service (NWS). As discussed in Section 6.7**,** local NWS offices serving Kentucky have access to a specialized Mesonet data feed and can display network data directly in their principal workstation, the Advanced Weather Interactive Processing System (AWIPS), as shown in Figure 9-3.

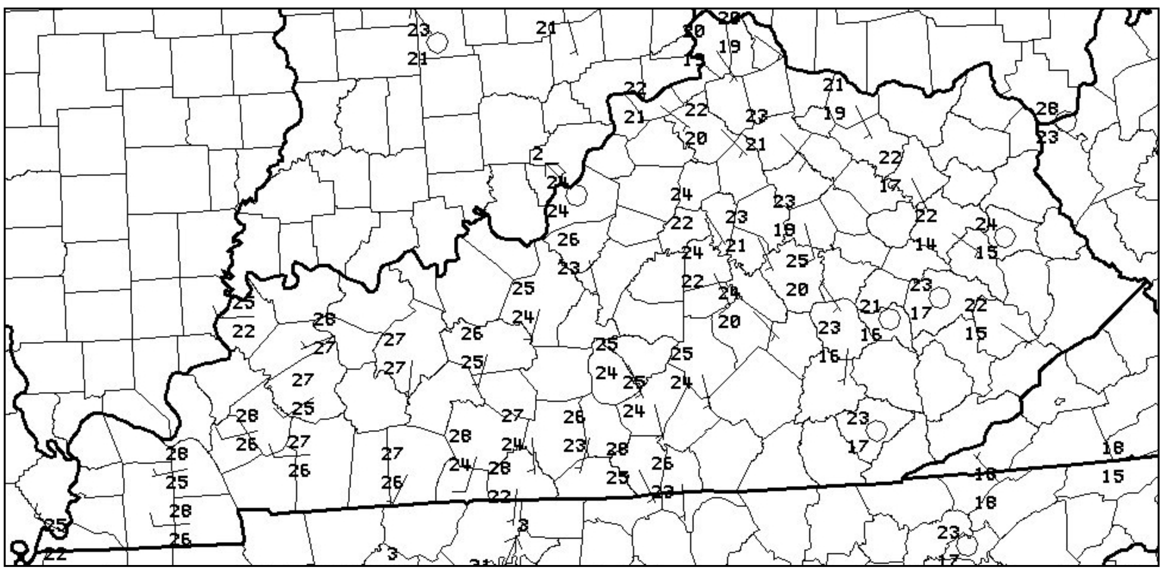


Figure 9-3. Kentucky Mesonet data in the NWS AWIPS. Some data from other networks are shown in surrounding states. All but one data plot in Kentucky is from the Mesonet. (Source: NWS Jackson, KY).

Widespread usage of network data by NWS personnel can be seen in the direct references to the network in official NWS products and bulletins, including those related to routine forecasting and alerting operations, as shown in Figure 9-4 through Figure 9-7.

```
000
FXUS63 KPAH 091732 AAA
AFDPAH

AREA FORECAST DISCUSSION
NATIONAL WEATHER SERVICE PADUCAH KY
1130 AM CST WED DEC 9 2009

.UPDATED...

FOR 18Z AVIATION

&&

.DISCUSSION...
STRONG SURFACE LOW PRESSURE WAS CENTERED OVER WEST CENTRAL ILLINOIS
AT 06Z. THIS LOW WILL CONTINUE TO DEEPEN TO AROUND 975 MB BY THE
TIME IT REACHES NORTHERN LOWER MICHIGAN THIS AFTERNOON. AS ONE MIGHT
EXPECT WITH SUCH A POWERFUL STORM...A TIGHT PRESSURE GRADIENT EXISTS
OVER A WIDE AREA OF THE EASTERN U.S. AND SE CANADA.

MAIN FORECAST CONCERN REMAINS STRONG WINDS TODAY. LOW LEVEL LAPSE
RATES QUICKLY STEEPENED AROUND 06Z AS COLD ADVECTION COMMENCED. THIS
ALLOWED STRONGER WINDS TO REACH THE SURFACE...AS HIGH AS 45 MPH AT A
NEW KENTUCKY MESONET SITE IN FULTON COUNTY. VERY IMPRESSIVE MIXING

...
```

```
000
FXUS63 KLMK 111732
AFDLMK

AREA FORECAST DISCUSSION
NATIONAL WEATHER SERVICE LOUISVILLE KY
1232 PM EST THU FEB 11 2010

...UPDATED AVIATION DISCUSSION...

...FORECAST UPDATE...

HIGH PRESSURE NOW IS CENTERED OVER THE MID MS RIVER VALLEY AND WILL
CONTINUE TO DRIFT EAST

...

SKIES WILL BE PARTLY CLOUDY TONIGHT WITH NEARLY CALM WINDS. WENT
BELOW GUIDANCE FOR LOWS OVER THE SNOWPACK. AT 0730Z THIS MORNING
UNDER THE CLOUDS THE KENTUCKY MESONET SITE IN WARREN COUNTY WAS
REPORTING 25 DEGREES...WHILE RIGHT NEXT DOOR IN LOGAN COUNTY WHERE
SKIES WERE CLEAR THE KENTUCKY MESONET SITE WAS REPORTING 16.

...
```

Figure 9-4. Kentucky Mesonet references in NWS Area Forecast Discussions, truncated to emphasize Mesonet references.

```
000
WGUS83 KPAH 050019
FLSPAH

FLOOD ADVISORY
NATIONAL WEATHER SERVICE PADUCAH KY
719 PM CDT SAT JUL 4 2009

KYC035-047-221-050315-
/O.NEW.KPAH.FA.Y.0046.090705T0019Z-090705T0315Z/
/00000.N.ER.000000T0000Z.000000T0000Z.000000T0000Z.OO/
CHRISTIAN KY-TRIGG KY-CALLOWAY KY-
719 PM CDT SAT JUL 4 2009

THE NATIONAL WEATHER SERVICE IN PADUCAH HAS ISSUED AN

* URBAN AND SMALL STREAM FLOOD ADVISORY FOR...
  CHRISTIAN COUNTY IN SOUTH CENTRAL KENTUCKY...
  CALLOWAY COUNTY IN WESTERN KENTUCKY...
  TRIGG COUNTY IN WESTERN KENTUCKY...

* UNTIL 1015 PM CDT.

* AT 712 PM CDT THUNDERSTORMS WITH VERY HEAVY RAIN WERE MOVING EAST
ACROSS EASTERN TRIGG AND CHRISTIAN COUNTY...AS WELL AS INTO WESTERN
CALLOWAY COUNTY. RADAR ESTIMATED RAINFALL IN EXCESS OF 2 INCHES HAD
FALLEN IN NORTHWEST TRIGG COUNTY IN THE LAST HOUR. THE KENTUCKY
MESONET SITE AT MURRAY MEASURED 1.33 INCHES OF RAINFALL IN THE LAST
HOUR WHILE 1.04 WAS MEASURED AT MESONET SITE JUST NORTH OF
HOPKINSVILLE.

HEAVY RAINFALL WILL CONTINUE ACROSS MUCH OF CALLOWAY...TRIGG AND
CHRISTIAN COUNTIES THROUGH ABOUT 9 PM WITH FLOODING OF SMALL STREAMS
AND LOW LYING AREAS.

PRECAUTIONARY/PREPAREDNESS ACTIONS...

DO NOT DRIVE YOUR VEHICLE INTO AREAS WHERE THE WATER COVERS THE
ROADWAY. THE WATER DEPTH MAY BE TOO GREAT TO ALLOW YOUR CAR TO CROSS
SAFELY. MOVE TO HIGHER GROUND.

&&

LAT...LON 3676 8814 3676 8812 3687 8816 3698 8771
      3711 8765 3710 8751 3712 8735 3703 8728
      3694 8733 3667 8736 3668 8809 3658 8806
      3651 8807 3650 8849 3675 8848

$$
SHANKLIN
```

Figure 9-5. Kentucky Mesonet data referenced in an NWS Flood Advisory.

000
WUUS53 KPAH 242106
SVRPAH
KYC047-059-107-149-177-242145-
/O.NEW.KPAH.SV.W.0027.100424T2106Z-100424T2145Z/

BULLETIN - EAS ACTIVATION REQUESTED
SEVERE THUNDERSTORM WARNING
NATIONAL WEATHER SERVICE PADUCAH KY
406 PM CDT SAT APR 24 2010

THE NATIONAL WEATHER SERVICE IN PADUCAH HAS ISSUED A

* SEVERE THUNDERSTORM WARNING FOR...
  NORTH CENTRAL CHRISTIAN COUNTY IN WESTERN KENTUCKY...
  DAVIESS COUNTY IN WESTERN KENTUCKY...
  EASTERN HOPKINS COUNTY IN WESTERN KENTUCKY...
  MCLEAN COUNTY IN WESTERN KENTUCKY...
  NORTHERN MUHLENBERG COUNTY IN WESTERN KENTUCKY...

* UNTIL 445 PM CDT.

* AT 401 PM CDT...TRAINED WEATHER SPOTTERS REPORTED A LINE OF SEVERE
  THUNDERSTORMS CAPABLE OF PRODUCING DAMAGING WINDS IN EXCESS OF 60
  MPH. THESE STORMS WERE LOCATED ALONG A LINE EXTENDING FROM CALHOUN
  TO MORTONS GAP...OR ALONG A LINE EXTENDING FROM CALHOUN TO 8 MILES
  SOUTH OF MADISONVILLE...AND MOVING NORTHEAST AT 70 MPH. **THE
  KENTUCKY MESONET REPORTING STATION REPORTED 62 MPH AT 4 PM CDT.**

* LOCATIONS IN THE WARNING INCLUDE...
  MORTONS GAP...
  NORTONVILLE...
  GRAHAM...
  CALHOUN...
  CENTRAL CITY...
  LIVERMORE...
  MASONVILLE...
  KNOTTSVILLE...

  **DAMAGING WINDS UP TO 62 MPH WERE REPORTED 4 MILES SOUTHWEST OF
  MADISONVILLE WITH THIS STORM.**

PRECAUTIONARY/PREPAREDNESS ACTIONS...

A TORNADO WATCH REMAINS IN EFFECT UNTIL 900 PM CDT SATURDAY EVENING
FOR SOUTHERN ILLINOIS AND SOUTHWEST INDIANA AND WESTERN KENTUCKY.

&&

LAT...LON 3763 8690 3763 8696 3756 8704 3748 8711
      3737 8710 3737 8704 3731 8699 3711 8755
      3764 8731 3786 8691 3783 8691 3784 8686
      3783 8685
TIME...MOT...LOC 2105Z 222DEG 61KT 3763 8726 3727 8743
WIND...HAIL 60MPH <1.00IN

Figure 9-6. Kentucky Mesonet wind data referenced in an NWS Severe Thunderstorm Warning.

```
000

NWUS53 KLMK 091623
LSRLMK

PRELIMINARY LOCAL STORM REPORT
NATIONAL WEATHER SERVICE LOUISVILLE KY
1122 AM EST WED DEC 09 2009

..TIME..    ...EVENT...     ...CITY LOCATION...      ...LAT.LON...
..DATE..    ....MAG....     ..COUNTY LOCATION.. ST.. ...SOURCE....
            ..REMARKS..

1120 AM     NON-TSTM WND GST 3 N HARRODSBURG         37.81N  84.85W
12/09/2009  M56 MPH          MERCER             KY   MESONET


&&

EVENT NUMBER LMK0900207

$$

CMC
```

Figure 9-7. Very strong gradient winds at the Mesonet Site in Mercer County referenced in an NWS Local Storm Report.

9.2 Research use

As a heavy user of Mesonet data operationally, the NWS also has a strong interest in using network data for research purposes to examine impacts on and processes of mesoscale meteorology. A currently ongoing study being conducted jointly by the NWS office in Jackson, KY and the Kentucky Climate Center – including the author – is using both Mesonet and Automated Surface Observing System (ASOS) data to better understand the setup and destruction of ridge/valley temperature splits during which valley locations quickly decouple near nightfall, with their temperatures plummeting compared to their ridge top counterparts. Figure 9-8 and Figure 9-9, both from Grogan et al. (2010), illustrate an occurrence of a significant split on 26-27 December 2008, when the ridge top was as much as 13ºC (23ºF) warmer than the valley.

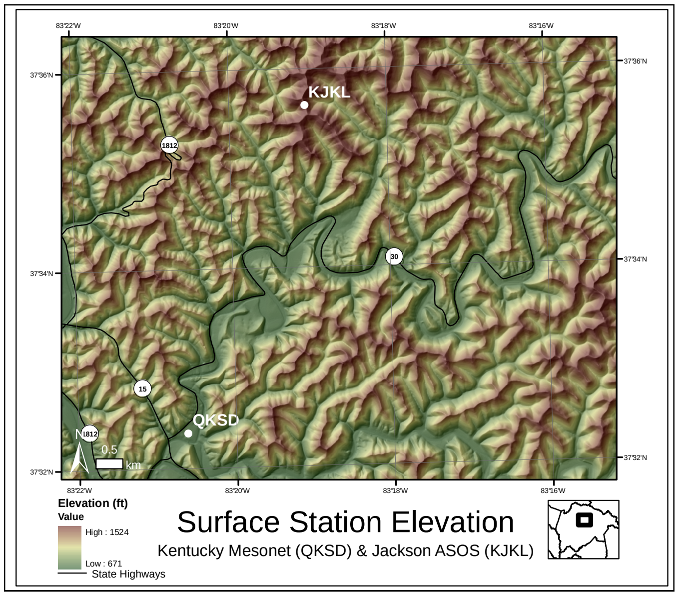


Figure 9-8. Elevation difference between Jackson, KY ASOS (KJKL) and Kentucky Mesonet station (QKSD) in Breathitt County. (Grogan et al. 2010).

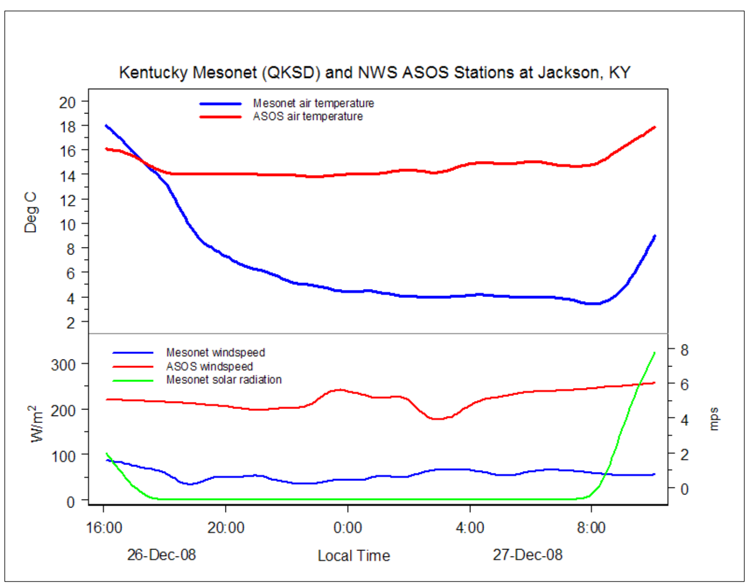


Figure 9-9. Ridge (ASOS) / valley (Mesonet) temperature split of 26-27 December 2008. Times are EST. (Grogan et al. 2010).

Agriculture is an important part of Kentucky's economy and meteorological data can be critical to farmers and growers. Mesonet data, therefore, are provided to both the University of Kentucky Agricultural Weather Center and the Fusarium Head Blight Prediction Center (FHBPC) – a cooperative effort between Penn State, Ohio State, Kansas State, Purdue, North Dakota State, and South Dakota State universities – for use in researching and operating predictive models of crop disease. Figure 9-10 shows use of Kentucky Mesonet data in the FHBPC's Risk Assessment Tool.

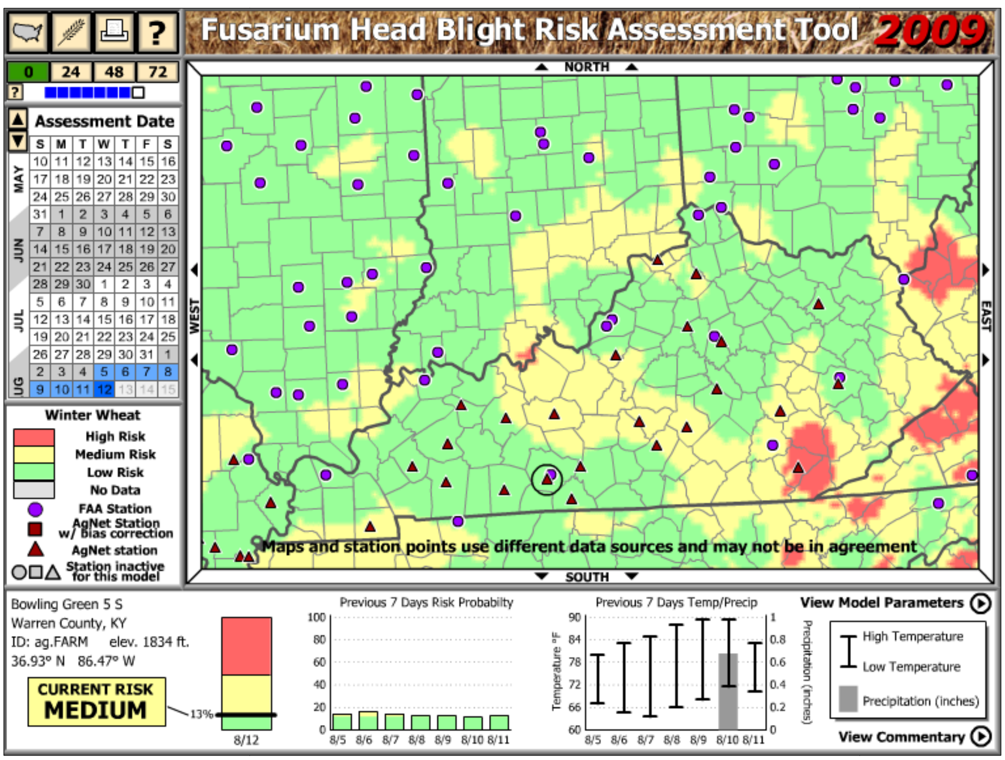


Figure 9-10. Fusarium Head Blight Prediction Center Risk Assessment Tool. Kentucky Mesonet sites shown as triangles. (Source: http://www.wheatscab.psu.edu/riskTool_2009.html).

9.3 Key local partnerships

Though admittedly out of scope for a technology-based thesis, it feels appropriate that this last section of core content focuses on some literally-foundational partnerships that have made the Kentucky Mesonet possible. After all, local stakeholders have put trust and faith in the network to fulfill the promises made to obtain their participation. The network's information technology infrastructure must support the activities that make fulfilling these promises possible.

Local relationships have been credited for the success of the world-renowned Oklahoma Mesonet, with McPherson et al. (1999) noting "in the mind of a student or emergency manager, a feeling of ownership in this Mesonet weather information has incalculable results." Without exception the local interests and land owners who have been willing to work with the Kentucky Mesonet to host stations – often in prime locations – have done so out of a belief that their contributions will have a significant, positive impact on their local communities.

As described in the network overview, the Kentucky Climate Center formed the Kentucky Mesonet Consortium with all public universities in the state to leverage the value of the network for the benefit of the citizens of Kentucky. While the consortium has certainly made contributions in terms of land for placing stations, a wide array of other local interests have provided significant assistance in the locating of Mesonet sensing stations in areas mostly well suited for long-term climate monitoring. Figure 9-11 shows the number of sites for which each local entity type has aided the location search and survey process.

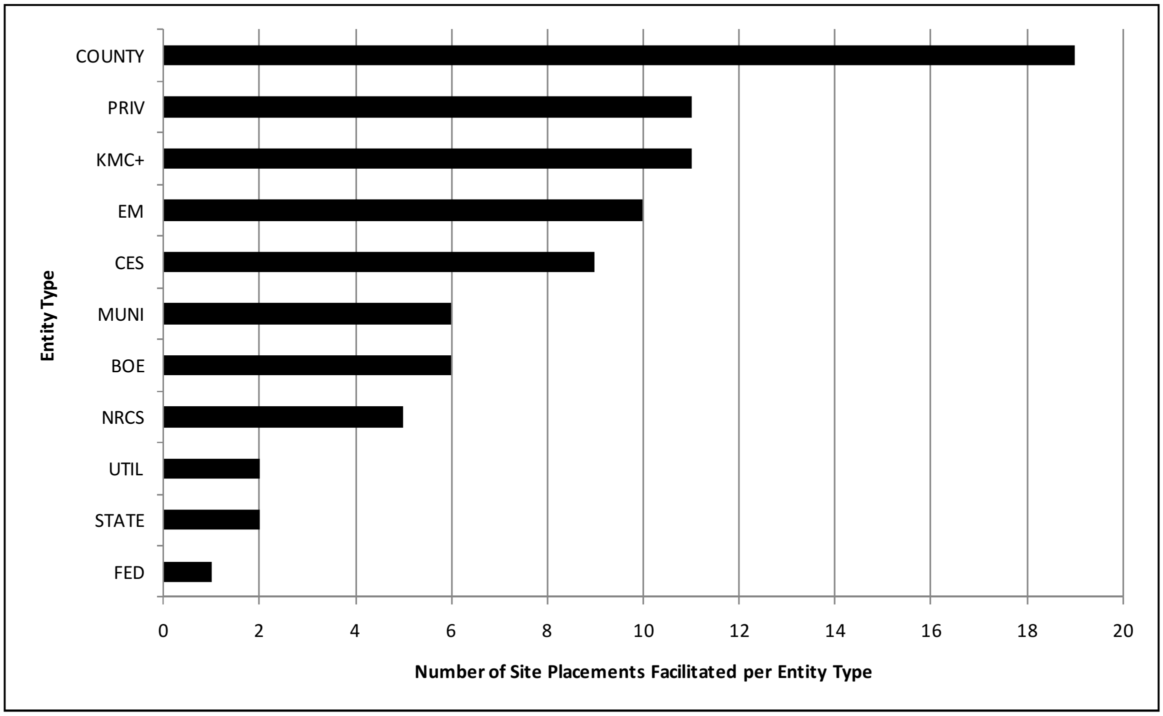


Figure 9-11. Site placement aided by local interests. COUNTY = county gov.; PRIV = private; KMC+ = KY Mesonet Consortium or other higher ed.; EM = emerg. mgr.; CES = U. of KY Coop. Ext. Svc; MUNI = city gov.; BOE = board of ed.; NRCS = Nat'l Resources Conservation Svc.; UTIL = utility; STATE = state gov.; FED = federal gov. (above & beyond ongoing NWS assistance). Many sites were facilitated by multiple entities, which are each given "credit" for that site in the graph.

Local interests have, in addition to their assistance with locating prime locations, been directly involved throughout the planning and installation phases of Mesonet sites. In many cases, such as sites on publicly accessible lands, local governments or other entities have participated in a cost share to help fund erection of a security fence and/or digging of trenches for power conduits. In some cases, even when a Mesonet site could not be located on public land, local government officials facilitated the siting of a station on private property. While it executes a site license agreement for each station, to date the Mesonet has had to pay no rental or usage fees to any property owner. Figure 9-12 shows the percentage breakdown of Mesonet site locations by land owner type.

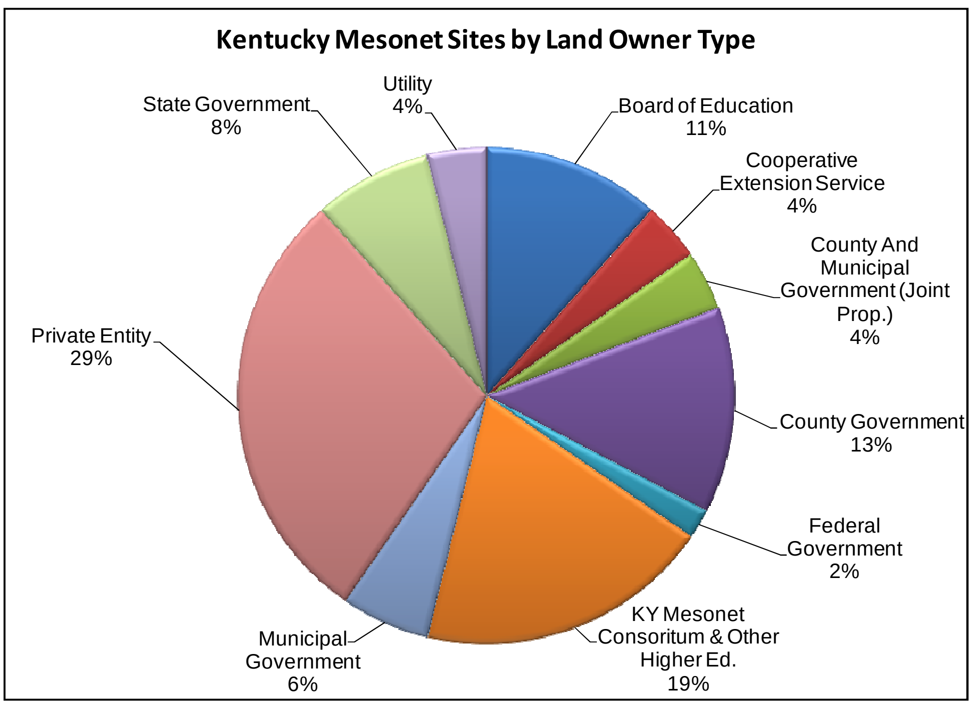


Figure 9-12. Kentucky Mesonet sites by land owner type. Total sites represented = 53, including those online, planned, or already under construction.

CHAPTER 10. DISCUSSION AND CONCLUSIONS

With the core elements of the author's applied research now addressed, this document turns to a discussion of that work's objectives in the context of contributions to the knowledge of in situ surface network design and benefits to the people of Kentucky. A look back at some hurdles to achieving those objectives is then provided, followed by a look at future work that must occur to improve the network.

10.1 Reflection on research goals

By this point, the reader hopefully has a good sense as to how well the author's research goals have been achieved. This section restates those goals and provides a contextual discussion of the author's work toward an evaluation of how well the research purposes and motivations presented in Chapter 1 have been met.

10.1.1 Increase in quality, original data

The first goal of the last three years' effort has been to "significantly increase the spatial coverage, amount, timeliness, availability, and use of original, quality surface meteorological data in Kentucky." Figure 1-5 shows the marked spatial improvement of research-grade, operational surface network coverage made in the state. Prior to the deployment of Mesonet stations, there were vast holes where no research-grade station was available nearby. Today, there are only a couple of small holes with no research-grade station within 20 miles. As the network grows toward its 100 site goal, coverage will become even better.

The Mesonet's computing network and supporting infrastructure have made a substantial contribution to operational and research meteorology through handling and distributing original meteorological and climatological data for Kentucky. As of 28 February 2010, over five million five-minute observations had been collected by Mesonet IT systems and the observation database contained over 100 million individual meteorological measurements.

Not only are those measurements being taken, they are seeing substantial use by both the public and key operational users. While the Mesonet website distributes data directly to a significant number of public users, the citizens of Kentucky also benefit from data use by broadcast media during regular newscasts or times of inclement weather. Local National Weather Service offices, core partners in the development of the network, are prime users of Mesonet data as evidenced in their official forecast and advisory products.

Finally, the IT systems and the network in general have and will continue to make contributions to meteorological and meteorology-dependent research. Quantitative studies of ridge/valley temperature splits in eastern Kentucky are being used to improve forecast skill for these phenomena, while agricultural use of the data is contributing to better predictions of crop disease.

Given the above, it is concluded that the author's first research objective has been met.

10.1.2 Development of IT infrastructure

The second research goal presented in Chapter 1 was to "develop the core information technology systems necessary to support both mission-critical operational and research use of the Kentucky Mesonet." Nineteen core, geographic information (GIS), and ancillary systems have been developed in the last three years, nearly half of which could likely form the basis for their own thesis. Systems developed range from core field communications to GIS-based access mechanisms to code management systems.

Most importantly, the systems and methods developed are quality systems that have been designed with mission criticality in mind from the start. Some simple statistics tell a good bit of the story of just how well those systems have been constructed. For the 12 month period ending 28 February 2010, the uptime availability percentage – largely influenced by the site communications method – averaged 99.730%, which is certainly respectable considering some targets for the "best" networks in the emerging National Network of Networks (NNoN) are around 98% (AASC 2010). For the same period, the Mesonet computing network was accessible and available 99.977% of the time and was unreachable due to outage or maintenance by Mesonet or co-location personnel for only 121 minutes.

Given the number of systems developed, the performance of the overall technology infrastructure, and the fact that an extensive supporting, object-oriented, modular code base has been developed, it is concluded that the second research objective has also been achieved.

10.1.3 Core competencies at the local level

The author's third research objective was to "show that core information technology-related competencies required by a national network are achievable at the local level, even with a small staff." Pushed for by congress, the emerging NNoN will harness the energy and enthusiasm of state and local networks. While in its final form the NNoN may provide assistance with some IT-related needs, the NRC (2009) study and follow-on meetings stress the importance of local competencies in network operation and design.

As the core systems for the Kentucky Mesonet have been designed, built, and maintained by only a full-time architect, a student developer, and an on-again/off-again Application Developer with only minimal university-level technical support, the experiences of the Kentucky Mesonet show that, indeed, these core competencies can be and are available at the local level, even with a small staff. Critical to achieving such a goal, though, is that the local staff possess a high level of dedication and professionalism, that they share and help develop the vision of the network, and that they take ownership of their role in it.

10.1.4 An updated perspective

The fourth and final research purpose has been to "provide in the literature an updated perspective on building the IT-related infrastructure to support a statewide in situ surface sensing network, especially in the areas of communications, data ingest, and processing systems." As this document's literature review shows, Kentucky is certainly not the first state to endeavor to build such a network. However, the experiences in Kentucky do offer an updated perspective and are being looked to as an example for other efforts, including by the various American Meteorological Society working groups charged with providing guidance on the emerging NNoN.

Of course, it is hoped that this thesis itself will be a positive contribution to the literature concerning design and operation of an in situ network. However, multiple peer-reviewed papers are being planned from this work and the Kentucky Mesonet's IT experiences have already been widely shared in the form of multiple conference presentations and papers.

10.2 Past, current, and future directions

While the intended research goals can be judged to have been met, in no way should the IT infrastructure of the Mesonet be considered complete. Realistically, for the network to grow and change, its infrastructure must continuously grow and change with it. However, efforts to date place the Mesonet somewhere probably between level 4 and 5 of Trenberth et al.'s (2002) surface network priorities given in Section 2.4.1. Getting to this point has not been easy and the bumps experienced along the way have certainly kept IT development from reaching a higher level possible without them.

Martin (2006) noted that a significant factor associated with poorer performance of IT projects is a change in project requirements or staff. While not intended as commentary on the skills of those who currently do or who have previously held the position, the network has certainly experienced a significant hurdle in attracting and keeping qualified practitioners for the Application Developer position. The typically lower academic pay and WKU's not being in a major city are thought to be contributing factors.

Brown and Hubbard (2000) stressed that funding in situ networks on short-term grants can lead to a "feast or famine" funding cycle which can create a loss of network focus and make key technical personnel retention difficult. The experience of this situation for the Kentucky Mesonet has certainly led to the shifting priorities Martin (2006) warned about. The requirement to build a robust and scientifically respectable IT infrastructure as planned in Chapter 3 has been constantly juxtaposed against the desire to develop magic bullet applications to attract long-term funding, with the irony being that those applications cannot exist without the supporting core infrastructure. Given the realities of the current economy, and the original goal to develop them, the importance of those applications is by no means discounted. However, the funding-model-caused loss of network focus described in Brown and Hubbard (2000) is at least somewhat visible in the IT infrastructure. Even with – and especially in the face of – changing personnel and shifting focus, the development of the IT infrastructure to date can rightfully be judged a success.

While a lot of effort has been expended to build existing Kentucky Mesonet systems, future work is just as critical to the continued growth and operation of the network. This future work must be carried forth in three core areas:

(i) in making the needed improvements described in Chapters 6 through 9 to the individual core, geographic information, and ancillary systems;

(ii) in the redoubling of efforts to indeed develop more value-added applications to help support particular funding models;

(iii) and finally in taking a renewed look at the entirety of Mesonet information technology for broad improvements, particularly toward a more unified implementation in terms of databases, systems, and overall architecture.

CHAPTER 11. SUMMARY

The Kentucky Mesonet is a high-density, mesoscale network of automated meteorological and climatological sensing stations deployed across the commonwealth which measure a suite of atmospheric surface parameters, including 1.5 m air temperature, relative humidity, solar radiation, 10 m wind speed & direction, and wetness – an indicator of ongoing precipitation. The network has grown fairly quickly, with the first site established just south of Bowling Green in May 2007 and the 46th site established near Henderson in February 2010. Funding for construction and initial operation of the network was provided by a combination of federal earmarks and direct grants from the National Oceanic and Atmospheric Administration. As Lead Systems Architect for the network, the author has worked to meet multiple information technology (IT) research and development goals:

(i) to significantly increase the spatial coverage, amount, timeliness, availability, and use of original, quality surface meteorological data in Kentucky;

(ii) to develop the core information technology systems necessary to support both mission-critical operational and research use of the Kentucky Mesonet;

(iii) to show that core information technology-related competencies required by a national network of networks are achievable at the local level, even with a small staff;

(iv) and to provide in the literature an updated perspective on building the IT-related infrastructure to support a statewide in situ surface sensing network, especially in the areas of communications, data ingest, and processing systems.

In the last three years, nineteen core or ancillary IT systems have been developed, including a robust enterprise-grade communications solution; site survey, metadata, and observational database storage systems; websites; availability assurance mechanisms; an automated quality control system; and various geographic information system (GIS)-based data visualization tools. These systems support and make possible the use of Mesonet data by both the general public and critical operational partners such as the National Weather Service (NWS), broadcast media, and state government.

Development of the network's IT systems has been rooted in well-established standards and best practices for meteorological surface sensing networks and has generally followed Trenberth et al.'s (2002) implementation priorities. The Mesonet has achieved a level somewhere between the fourth and fifth of these five priorities:

(i) data collection and archiving
(ii) distribution of the raw data in near-real time;
(iii) quality control of the data in delayed mode and archiving of datasets;
(iv) development and maintenance of data access tools (e.g., web sites);
(v) and follow-on processing to produce analyses and reanalyses.

The computing network and supporting infrastructure developed thus far have made a substantial contribution to operational and research meteorology & climatology by collecting, storing, handling, and distributing over five million five-minute observations containing over 100 million individual meteorological measurements. The citizens of Kentucky have benefited by network operations through direct access to data on the official program website and through data use by broadcast media during regular newscasts or times of inclement weather. Local National Weather Service offices have been prime users of Mesonet data, routinely referencing them in their official forecast and advisory products. Research use of network data is supplementing understanding of mountain/valley interactions and aiding with predictions of crop disease.

Mesonet communications and computing systems have been designed to be as mission-critical as possible within budgetary constraints. An enterprise-class, cellular-based communications method implemented with AT&T has provided a respectable average site uptime of at least 99.730% over the last year. Choice of co-location internet provider, server technology, and implementation approach has yielded a network availability percentage of 99.977% for that same period.

Pushed for by the U.S. congress, an emerging Nationwide Network of Networks is planned to harness the energy and enthusiasm of state and local networks. Through development of its own critical systems, the Kentucky Mesonet has shown that the core competencies needed for participation in the NNoN can be and are available at the local level, even with a relatively small staff. The experiences of the Kentucky Mesonet are being or will be shared with the broader scientific community through the author's participation in multiple NNoN working groups established by the American Meteorological Society, through multiple conference papers and presentations, and through planned peer-reviewed publications.

The road for systems development has been somewhat bumpy, with unfortunate turnover in an important personnel position and a somewhat shifting focus hindering it from reaching its full potential. Supported by a well-designed implementation plan, and even in the face of those difficulties, the development of the network's IT infrastructure to date can still be rightfully judged a success, having positively met the four research and development goals. Continued development of the network's IT infrastructure is critical to its continued growth, development, and success and must include a redoubling of efforts to develop more value-added applications and work toward a more unified system of databases and overall architecture.

BIBLIOGRAPHY

AASC [American Association of State Climatologists], 2010: Standards for a surface-based national Mesonet: Achieving climate/weather-quality data from a tiered network-of-networks. AASC Tech Rep. Draft Version 1.1, 46 pp.

Aditus Consulting, 2008: JpGraph graphing library for PHP Version 2.3.3.

Adobe Systems, Inc., 2010: Flex Builder software development kit Version 3.0.2.

Apache Foundation, 2010a: HTTP server project.

_____, 2010b: Subversion version control system.

AT&T, 2005: Cingular Wireless Commercial Connectivity Service Connectivity Reference Document. AT&T Tech Rep. Version 2.1, 48 pp.

Barrett, B., 2009: Webalizer log file analysis program.

BGMU [Bowling Green Municipal Utilities], 2010: BGMU Fiber Optics Division – Overview [Available online at: http://www.bgmu.com/optics1_overview.htm.] (accessed 28 February 2010).

Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma mesonet: A technical overview. *J. Atmos. Oceanic Technol*., **12**, 5-19.

Brown, P. W. and K. G. Hubbard, 2000: Lessons from the North American experience with automated weather stations. *Proc. of Int. Workshop on Automated Weather Stations for Applications in Agriculture and Water Resources Management*, Lincoln, NE, High Plains Climate Center and World Meteorological Organization, 21-28.

_____, T. Machibya, and B. T. Russell, 2000: Use of cell phones for data retrieval in Arizona. *Proc. of Int. Workshop on Automated Weather Stations for Applications in Agriculture and Water Resources Management*, Lincoln, NE, High Plains Climate Center and World Meteorological Organization, 49-54.

Brown, Z. G., 2008a: Metadata PHP programming documentation. Kentucky Mesonet, Western Kentucky University.

_____, 2008b: ADS2 table descriptions documentation. Kentucky Mesonet, Western Kentucky University.

_____, S. A. Foster, R. Mahmood, S. Struebig, D. Grabowski, D. M. Grogan, and M. M. Ferris, 2008: An overview of the Kentucky Mesonet. *Extended Abstracts, 17th Conf. on Applied Climatology*, Whistler, BC, Amer. Meteor. Soc., 1.3.

Campbell Scientific, 2009: LoggerNet and LoggerNet for Linux server software.

CentOS, 2008: CentOS Linux Operating System Version 5.1.

Collabnet, 2010: TortoiseSVN interface for Subversion Version 1.6.7.

Crawford, K. C., and G.R. Essenberg, 2006: COOP modernization: NOAA's Environmental Real-time Observation Network in New England, the Southeast, and Addressing NIDIS in the West. *Extended Abstracts, 22nd International Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* Atlanta, GA, Amer. Meteor. Soc., J5.9.

DelGreco, Stephen, National Climatic Data Center, 2006: personal meeting. 21 Nov. 2006.

Dell, 2008: OpenManage Server Administrator for Linux software.

Dewett, T., and G. R. Jones, 2001: The role of information technology in the organization: a review, model, and assessment. *Journal of Management*, **27**, 313-346.

Eclipse Foundation, 2009a: Eclipse Integrated Development Environment Version 3.4.2.

_____, 2009b: Visual Editor.

ESRI, 2008: ArcIMS Version 9.3.1.

_____, 2009a: ArcGIS Engine Version 9.3.1.

_____, 2009b: ArcObjects Version 9.3.1.

_____, 2009c: ArcGIS API for Flex Version 1.3.

_____, 2009d: ArcGIS Server Version 9.3.1.

_____, 2009e: ArcMap Version 9.3.1.

FCC [Federal Communications Commission], 2007a: Federal Communications Commission Licensing Database Extracts: Cellular Service Area Boundaries (aggregate by callsign). [Available online at: http://wireless.fcc.gov/geographic/index.htm?job=licensing_database_extracts & http://wireless.fcc.gov/geographic/data/db/cellular_aggregate_callsign_sab.zip.] 1 July 2007.

______, 2007b: Federal Communications Commission Universal Licensing System: Cellular – 47 CFR Part 22 – Licenses. [Available online at: http://wireless.fcc.gov/uls/data/complete/l_cell.zip and http://wireless.fcc.gov/uls/index.htm?job=transaction&page=weekly.] 14 Sep 2007.

______, 2008a: Federal Communications Commission Licensing Database Extracts: Cellular Service Area Boundaries (aggregate by callsign) [Available online at: http://wireless.fcc.gov/geographic/index.htm?job=licensing_database_extracts & http://wireless.fcc.gov/geographic/data/db/cellular_aggregate_callsign_sab.zip.] 23 July 2008.

______, 2008b: Federal Communications Commission Universal Licensing System: Cellular – 47 CFR Part 22 – Licenses. [Available online at: http://wireless.fcc.gov/uls/data/complete/l_cell.zip and http://wireless.fcc.gov/uls/index.htm?job=transaction&page=weekly.] 23 July 2008.

_____, 2009: Federal Communications Commission Universal Licensing System: Extract of Cellular Radiotelephone Service Sites. [Available online at: http://wireless.fcc.gov/geographic/data/db/cellular.zip.] 26 July 2009.

Ferris, M. M., S. Struebig, M. Grogan, S. Foster, and R. Mahmood, 2010: Quality assurance methods for the Kentucky Mesonet. *15th Symposium on Meteorological Observation and Instrumentation,* Atlanta, Georgia, Amer. Meteor. Soc., **3.3**.

Fiebrich, C. A., 2009: History of surface weather observations in the United States. *Earth-Sci Rev.*, **93**, 77-84.

Fowler, M., 2003: *Patterns of enterprise application architecture*. Addison-Wesley, 533 pp.

GCOS [Global Climate Observing System], 2003: GCOS climate monitoring principles. [Available online at http://www.wmo.int/pages/prog/gcos/documents/ GCOS_Climate_Monitoring_Principles.pdf.] (accessed 25 September 2008).

GeoRSS, cited 2010: GeoRSS Main Page. [Available online at: http://www.georss .org/Main_Page].

Grant, R.H., and D. Toby, 2000: Cellular communications for automated weather stations. *Proc. of Int. Workshop on Automated Weather Stations for Applications in*

Agriculture and Water Resources Management, Lincoln, NE, High Plains Climate Center and World Meteorological Organization, 41-48.

Grogan, D. M., 2007: Computing infrastructure design considerations and recommendations. Kentucky Mesonet, Western Kentucky University.

_____, S. A. Foster, and R. Mahmood, 2010: The Kentucky Mesonet: Perspectives on data access, distribution, and use for a mesoscale surface network. *Extended Abstracts, 26th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Atlanta, GA, Amer. Meteor. Soc., **14A.4**.

Hall, Alan D., National Climatic Data Center, 2006: personal meeting. 21 Nov. 2006.

Horton, G., K. Blauvelt, R. Carcoana, J. Gibbons, I. Nichols, and K. Rhodes, 2000: Communications with Weather Stations (and among Networks). *Proc. of Int. Workshop on Automated Weather Stations for Applications in Agriculture and Water Resources Management*, Lincoln, NE, High Plains Climate Center and World Meteorological Organization, 244.

Hubbard, K. G., X. Lin, and C. B. Baker, 2005: On the USCRN temperature system. *J. Atmos. Oceanic Technol.*, **22,** 1095-1101.

ISC [Internet Systems Consortium], 2010: BIND Domain Name System server software.

Karl, T. R., R. G. Quayle, and P. Y. Groisman, 1993: Detecting climate variations and change: New challenges for observing and data management systems. *J. Climate*, **6**, 1481-1494.

KY DGI [Kentucky Division of Geographic Information], cited 2009a: Kentucky Event Mapping and Analysis Portal. [Available online at: http://kygeonet.ky.gov/kemap_flex/].

_____, cited 2009b: Kentucky Weather Mapping. [Available online at: http://kygeonet.ky.gov/kyweather].

_____, cited 2009c: The Commonwealth Map [Available online at: http://kygeonet.ky.gov/tcm/tcm.htm].

Kentucky Legislature, 2006: *A joint resolution recognizing the Kentucky Mesonet as the official source of climatological observations for the state.* SJR 228.

Kowal, 2008: Launch4j version 3.0.1.

Langdon, C. S., 2003: Information systems architecture styles and business interaction patterns: Toward theoretic correspondence. *Information Systems and E-Business Management.* **1**, 283-304.

Legates, D. R., D. J. Leathers, T. L. DeLiberty, G. E. Quelch, K. Brinson, J. Butke, R. Mahmood, and S. A. Foster, 2005: DEOS: The Delaware Environmental Observing Systems. *Extended Abstracts, 21st International Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* San Diego, CA, Amer. Meteor. Soc., 18.10.

Martin, A., 2003: What drives the configuration of information technology projects? Exploratory research in 10 organizations. *Journal of Information Technology,* **18**, 1-15.

______, 2006: Successful IT application architecture design: an empirical study. *Information Systems and E-Business Management,* **4(2)**, 107-135.

Marvin, C. F., 1920: The status and problems of meteorology. *Proc. Nat. Acad. Sci. U.S.A.,* **6(10),** 561-572.

McGuirk, M. and E. May, 2003: Draft USCRN metadata management – survey to operations. NDCD report, 14 pp. [Available online at: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/metadata/ MetaDataMangt.doc].

McPherson, R. A., K. C. Crawford, J. M. Wolfinbarger, A. D. Melvin, and D. A. Morris, 1999: From rags to riches – seven years in the life of a climate office's educational programs. Preprints, *8th Symposium on Education*, Dallas, TX, Amer. Meteor. Soc., J1.5.

______, and Coauthors, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.,* **24**, 301-321.

Meyer, S. J., and K. G. Hubbard, 1992: Nonfederal automated weather stations and networks in the United States and Canada: A preliminary survey. *Bull. Amer. Meteor. Soc.,* **73(4)**, 449 – 457.

Microsoft Corporation, 2003: Microsoft Excel spreadsheet software.

_____, 2007: Microsoft Windows XP Professional operating system.

Miller, P. A., and M. F. Barth, 2003: Ingest, quality control, and distribution of observations from state transportation departments using MADIS. Preprints, *19tjh International Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* Long Beach, CA, Amer. Meteor. Soc., 10.11.

______, ______, and L. A. Benjamin, 2005: An update on MADIS observation ingest integration, quality control, and distribution capabilities. *Extended Abstracts, 21st International Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* San Diego, CA, Amer. Meteor. Soc., J7.12.

______, ______, ______, R. S. Artz, and W. R. Pendergrass, 2007: Preprints, *14th Symposium on Meteorological Observations and Instrumentation*, San Antonio, TX, Amer. Meteor. Soc., JP2.5.

______, 2008: personal meeting. 5 Nov. 2008.

Murhammer, M. W., K. Lee, P. Motallebi, P. Borghi, and K. Wozabal, 1999: IP network design guide. IBM Rep. SG24-2580-01, 324 pp. [Available online at: http://www.redbooks.ibm.com/redbooks/pdfs/sg242580.pdf].

Nagios Enterprises, 2008: Nagios IT Infrastructure Monitoring software.

NRC [National Research Council], 1999: *Adequacy of Climate Observing Systems*. National Academy Press, 66 pp.

_____, 2009: *Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks*. National Academy Press, Prepublication Form, 234 pp.

NTPD [Network Time Protocol Daemon], 2010: NTPD Linux on-line manual pages. [Available in RedHat Linux help system].

NOAA [National Oceanic and Atmospheric Administration], 2003: United States Climate Reference Network (USCRN): Functional requirements document. NOAA Rep. NOAA-CRN/OSD-2003-0009R0UD0, 13 pp. [Available online at: http://www1.ncdc.noaa.gov/pub/data/uscrn/documentation/program/X040_d0.pdf].

Oetiker, T., 2006: Multi Router Traffic Grapher software.

OCS [Oklahoma Climatological Survey], 2006: *Operations user manual.* Vol. 2, *The data ingest, quality assurance, and monitoring system for NOAA's Environmental Real-Time Observing Network (NERON)*. Oklahoma Climatological Survey, Norman, OK, 266 pp.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. *Bull. Amer. Meteor. Soc.,* **56(5)**, 527-530.

Perl, 2009: Perl programming language. Version 5.8.5.

Phillips, Mark, National Climatic Data Center, 2006: personal meeting. 21 Nov. 2006.

PHP Group, 2009: PHP Hypertext Processor.

RedHat, 2007: Fedora Core Linux operating system.

_____, 2008: RedHat Enterprise Linux 5.1 operating system.

Richardson, G. L., B. M. Jackson, and G. W. Dickson, 1990: A principles-based enterprise architecture: Lessons from Texaco and Star Enterprise. *MIS Quarterly*. **14(4)**, 385-402.

Rotch, A. L., 1904: Present problems of meteorology. *Science*. **20**(521), 872-878.

S. Rep. No. 110-397, 2008: A report to accompany S. 3182, Departments of Commerce and Justice, Science, and Related Agencies Appropriations Bill, 2009, U.S. Senate, 110th Congress, 2nd sess.

S. Rep. No. 111-34, 2009: A report to accompany H. R. 2847, Departments of Commerce and Justice, and Science, and Related Agencies Appropriations Bill, 2010, U.S. Senate, 111th Congress, 1st sess.

Schroeder, J. L., W. S. Burgett, K. B. Haynie, I. Sonmez, G. D. Skwira, A. L. Doggett, and J. W. Lipe, 2005: The West Texas Mesonet: A technical overview. *J. Atmos. Oceanic Technol.* **22,** 211-222.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredrickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma mesonetwork. *J. Atmos. Oceanic Technol.* **17,** 474-494.

Sierra Wireless, 2008: *Raven EDGE for AT&T User Guide*. Sierra Wireless, 98 pp.

_____, 2009a: *Raven XT for AT&T User Guide.* Sierra Wireless, 35 pp.

_____, 2009b: AceManager software. Version 3.3.0.187.

_____, cited 2009c: Using a Sierra Wireless AirLink Raven X or Raven-E with a Cisco Router. [Available online at:http://www.sierrawireless.com/en/Support/AirLink/~/media/Support_Downloads/AirLink/docs/AppNotes/CiscoRouter.ashx].

_____, 2010a: AirLink Raven EDGE for AT&T [Available online at: http://www.sierrawireless.com/en/sitecore/content/Sierra%20Wireless/Support/Downloads/AirLink/M2M_Gateways/AirLink_Raven_EDGE_for_ATT.aspx.] (accessed 28 February 2010).

_____, 2010b: AirLink Raven XT GPRS for AT&T [Available online at: http://www.sierrawireless.com/en/sitecore/content/Sierra%20Wireless/Support/Downloads/

AirLink/M2M_Gateways/AirLink_Raven_XT_GPRS_for_ATT.aspx] (accessed 28 February 2010).

Splitt, J. H., L. Dunn, J. Pechmann, B. White, C. Ciliberti, S. Lazarus, J. Slemmer, D. Zaff, and J. Burks, 2002: MesoWest: cooperative mesonets in the western United States. *Bull. Amer. Meteor. Soc.*, **83(2)**, 211-225.

Struebig, S., D. Grabowski, S. A. Foster, and R. Mahmood, 2010: Equipment and instrumentation used at a Kentucky Mesonet environmental monitoring station. *15th Symposium on Meteorological Observation and Instrumentation.* Atlanta, GA. Amer. Meteor Soc. P257.

Sun Microsystems[29], 2008: MySQL database.

_____, 2009: JavaScript programming language.

_____, 2010: Java programming language.

Thompson, G., cited 2010: National Center for Atmospheric Research surface weather station list. [Available online at: http://www.rap.ucar.edu/weather/ surface/ stations.txt].

Tkachenko, V. and T. Harris, 2007: PHP-ExcelReader.

Trenberth, K. E., T. R. Karl, and T. W. Spence, 2002: The need for a systems approach to climate observations. *Bull. Amer. Meteor. Soc.,* **83(11)**, 1593-1602.

Tucker, D. F., 1997: Surface mesonets of the western United States. *Bull. Amer. Meteor. Soc.*, **78(7)**, 1485-1495.

UCAR [University Corporation for Atmospheric Research], cited 2009a: Network Common Data Form software library. [Available online at: http://www.unidata.ucar.edu/software/netcdf/.]

_____, cited 2009b: Local Data Manager software & website [Available online at: http://www.unidata.ucar.edu/software/ldm/].

_____, 2009c: GEMPAK/N-AWIPS software version 5.11.4.

University of Edinburgh, 2004: Festival Speech Synthesis System.

VMWare, cited 2010: VMware Server website. [Available online at: http://www.vmware .com/products/server/].

Williams, T. and C. Kelley, 2004: Gnuplot.

[29] now owned by Oracle

Winkler, J. A., 2004: The impact of technology upon in situ atmospheric observations and climate science. *Geography and Technology*, S. D. Brunn, S. L. Cutter, and J. W. Harrington, Jr., Eds. Kluwer Academic Publishers, 461-490.

WMO [World Meteorological Organization], 2006: *Guide to Meteorological Instruments and Methods of Observation.* Prelim 7th ed. WMO 8, Geneva, Switzerland.

Wolfinbarger, Michael, Oklahoma Mesonet, 2006: phone conversation. 9 Oct. 2006

Zachman, J. A., 1987: A framework for information systems architecture. *IBM Systems Journal*, **26(3)**, 276-292.