

## RESEARCH PAPER

# The Oklahoma Mesonet: A Pilot Study of Environmental Sensor Data Citations

Betsy Van der Veer Martens<sup>1</sup>, Bradley G. Illston<sup>2</sup> and Christopher A. Fiebrich<sup>2</sup><sup>1</sup> School of Library and Information Studies, University of Oklahoma, 208 Schusterman Library, Tulsa, OK 74135, US<sup>2</sup> Oklahoma Mesonet, Oklahoma Climatological Survey, University of Oklahoma, 120 David L. Boren Blvd., Suite 2900, Norman, OK 73072, USCorresponding author: Betsy Van der Veer Martens, Ph.D. ([bvmartens@ou.edu](mailto:bvmartens@ou.edu))

This pilot study of 110 scientific papers utilizing environmental sensor data from the Oklahoma Mesonet during its first two decades of operations demonstrates the diversity of potential purposes in scientific research for a robust, rigorously maintained, accessible source of environmental sensor data, as well as the challenges involved in identifying uses of that data within scientific papers. The study authors selected three publication years (1995, 2005, 2015) from an extensive corpus of peer-reviewed journal publications, identified each paper's specific citation of and uses of the Mesonet's environmental sensor data, and derived a typology of those usages (assimilation, experimentation, observation, simulation, utilization, validation) found to be most common. The rapid increase in data assimilation research projects today is discussed in terms of the difficulty and importance of correct attribution to individual data sources in these complex research projects. The study examines the possible role played by highly-cited papers that describe the quality assurance procedures in sensor data sources, which may serve as surrogates to signal the quality of the data provided by such sources, and which may also provide a useful contribution towards understanding data citation as a special form of scholarly citation.

**Keywords:** data citation; environmental sensors; mesoscale networks; Oklahoma; weather

## 1. Introduction

Data, both experimental and observational, have always played a central role in science, as Strasser (2012) points out, but recently the creation and curation of data, datasets, and datastreams have become of intense interest to information scientists as well as to a new generation of data scientists (Halevi & Moed 2012). The question of how data from any specific source are found, used, and cited remains understudied, as compared to the numerous studies of citations to individual papers, scientists, and theories, all of which have been greatly facilitated by the increasing standardization of those citation practices over time (Williams 2011), which has yet to become the case for data citation practices (Altman *et al.* 2015; Mayernik 2012; Mooney & Newton 2012).

Acknowledging the use of externally-obtained data was once considered merely a 'scholar's courtesy' (Cronin, 1995); today, such acknowledgment is often elicited by either disciplinary or publication pressures, though it remains far from universal, despite the increase in funding agency and institutional requirements for data attribution and sharing (Kim & Stanton 2016). Even such sophisticated approaches to identifying datasets as the development of personalized, contextualized dataset search engines (Singhal, Kasturi, & Srivastava 2014; Singhal & Srivastava 2013; Singhal & Srivastava 2017) are hampered by this lack of uniformity.

According to Parsons & Fox (2013), actual data sources are much more diverse than is often recognized, ranging from an individual scientist's own development of small datasets, through scientific support systems that manage the data production for entire laboratories, through large-scale installations that produce massive datastreams for different disciplines or in different locations. There is also wide variation among data, ranging from qualitative to quantitative data types, from potentially replicable experimental

data to unique, unrepeatable observational data, from small datasets to enormous datastreams, making the interrelated questions of data attribution (Hou & Mayernik 2016; Hou & Mayernik 2017), data citation (Borgman 2015), and data provenance (Tilmes, Yesha & Halem 2010) even more complex.

Although such recent data-centric initiatives as the Joint Declaration of Data Citation Principles (Data Citation Synthesis Working Group 2014), DataCite (Rueda, Fenner & Cruse 2016), and the Data Citation Index (Robinson-García, Jiménez-Contreras & Torres-Salinas 2016) have helped in the creation and communication of these identification standards, their adoption and use continue to be limited within many scientific communities (Goldstein, Mayernik & Ramapriyan 2017). Borgman *et al.* (2014) observe, however, that even very large 'knowledge infrastructures' that produce and utilize data in such well-established areas as astronomy, for instance, vary in their robustness, as telescopic arrays and other astronomical instrumentation, both ground and space-based, must rely on economic, policy, and scientific support to maintain their viability over time. Tracking data uses is becoming an important part of these efforts to gain this support, as well as communicating best practices in data curation and data citation.

Accordingly, institutional studies from major knowledge infrastructures, such as that of Apai *et al.* (2010) on the use of Hubble Telescope data in scientific papers over a ten year period, Parsons *et al.* (2011) on the management of diverse, interdisciplinary data within the International Polar Year project, Belter (2014) on the value of oceanographic datasets from the National Oceanic & Atmospheric Administration, Peng *et al.* (2015) on the development of a unified data stewardship maturity model at NOAA's Cooperative Institute for Climate and Satellites, and Mayernik (2016) on the institutional data curation and dissemination practices at organizations such as the University Corporation for Atmospheric Research are now gaining visibility, with varying suggested solutions to these data-related issues, and should help to promote standardization across disciplines.

Adding to the challenges faced by knowledge infrastructures, the rapid increase in the number of environmentally embedded wireless sensors providing real-time, heterogeneous data for a wide variety of applications has contributed significantly to the so-called 'data deluge' (Porter, Hanson, & Lin 2012). Connected sensors of all types have been predicted to number a trillion by 2025 (Johnson 2015). Studies focusing on the various factors involved in the creation, curation, and citation of sensor data, particularly sensor datastreams in the sciences, include work by Bates, Lin & Goodale (2016), Borgman (2015), Cragin, Palmer & Chao (2010), Ganguly *et al.* (2007), and Mayernik, Wallis & Borgman (2013).

The present article is a contribution to this literature on the citation of specific sensor datastreams, as it examines the use of environmental sensor data over the past 20 years from a well-established knowledge infrastructure, the extensive mesoscale network (or 'Mesonet') for weather-related information throughout the state of Oklahoma. The history of the standardization of weather-related data in the United States has been described by Edwards (2010), Fiebrich (2009), Fleming (2016), and Klemm & McPherson (2017), while the evolution of global weather prediction techniques and tools has been detailed by Bauer, Thorpe & Brunet (2015) and Lynch (2008). A study of scientific papers focusing on the various uses of sensor data from a statewide mesoscale network such as that of Oklahoma, however, has not previously been performed, and should be of value in understanding the opportunities and challenges presented by such studies.

## 2. The Oklahoma Mesonet

This section describes the history and current operations of the Oklahoma Mesonet. Discussions concerning the mesonet installation studied here began in the 1980s, as the result of an ongoing collaboration between agricultural researchers at Oklahoma State University and meteorological researchers at the University of Oklahoma to develop a near real-time, extremely reliable source of surface observational data about local weather conditions across the state (Brock 2013; Crawford 2013). This need was felt to be particularly critical for Oklahoma, given the state's long history of drought and drainage issues. Losses due to crop failure (Ding, Hayes & Widhalm 2011) and urban flooding (Waite 2011) ran into the billions of dollars. The initiative garnered support throughout the state, largely due to the efforts of those involved in this unprecedented partnership among scientists at Oklahoma's two major research universities. Improvements in both research into and mitigation of weather impacts were the intended outcomes of this extensive project.

As a result, the Oklahoma Mesonet, a statewide network of 121 automated environmental monitoring stations, was officially launched in March of 1994. These 10-meter-tall towers, one or more of which is located within each of Oklahoma's 77 counties, provide regular measurements of air and soil temperature, barometric pressure, rainfall, relative humidity, solar radiation, soil moisture, and wind speed and direction, both direct and calculated, including instrumentation data, all of which are then packaged into 'observations'

that are regularly transmitted every 5 minutes to the Oklahoma Climatological Survey, where the data quality is immediately verified and then made available to Mesonet users.

The Oklahoma Mesonet has been termed the 'gold standard' among statewide climate and weather networks because of its well-known attention to quality assurance and quality control (National Research Council 2008: ix). As explained by Campbell *et al.* (2013: 575), quality assurance involves specific processes to ensure that the sensor networks and associated protocols are developed and adhered to in a way that minimizes inaccuracies in the data produced. This proactive approach attempts to avoid system problems that may produce suspect data, while the subsequent quality control measures are employed to identify and flag any suspect data after they have been generated. It should be noted that the explosive growth in sophisticated mathematical models for numerical weather prediction and environmental problem solving increases rather than lessens the importance of the quality of the data being utilized (Apte 2015).

Mesonet data and products are stored and delivered in a number of ways. The official data archive is stored in self-describing Network Common Data Form (Unidata 2017). From the NetCDF archive, approximately 1300 products are routinely produced, many of which are updated every five minutes. Data files are delivered in several text formats (*e.g.*, comma delimited or fixed column length) for users to download via Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP), or Local Data Manager (LDM).

Although some commercial and media entities are licensed users that receive Mesonet data on a paid basis, the data are also made available online free of charge at [www.mesonet.org](http://www.mesonet.org) to other users who wish to access them (Klockow, McPherson & Sutter 2010). Examples of Mesonet data users include emergency and public safety personnel, other state and local government agencies, aviation and other transportation firms, farmers, ranchers, construction firms, water experts, weather forecasters, and a variety of others vitally interested in Oklahoma's severe storms, dry line movements, soil moisture, drought conditions, heat impact, and other weather-related phenomena (Ziolkowska *et al.* 2017).

This direct distribution of data is in addition to the Oklahoma Mesonet's participation in various data distribution channels, such as the National Oceanic & Atmospheric Administration's National Mesonet Program, the U.S. Department of Agriculture's National Agricultural Statistics Service, and the U. S. Geological Survey's National Soil Moisture Network, and various other provisions of data made upon request to the Oklahoma Mesonet.

The Oklahoma Mesonet's requested acknowledgment statement in its agreement for researchers using its data (as shown under the heading 'Oklahoma Mesonet Data Access Policy' on its website) is that 'Oklahoma Mesonet data are provided courtesy of the Oklahoma Mesonet, which is jointly operated by Oklahoma State University and the University of Oklahoma. Continued funding for maintenance of the network is provided by the taxpayers of Oklahoma.'

Beginning in June 2014, the 'For Researchers' section of the website has listed two technically-oriented articles about the Mesonet (Brock *et al.* 1995 and McPherson *et al.* 2007) as useful potential citations by any researchers working with Oklahoma Mesonet data. These articles are also often informally recommended by the Oklahoma Mesonet staff to anyone wishing to learn more about the technical details of the Mesonet.

Beginning in March 2016, the 'About the Mesonet' section of the website has listed the following DOIs for Oklahoma Mesonet data and datasets: 10.15763/dbs.mesonet (for all Mesonet products) and 10.15763/dbs.mesonet.standard (for standard Mesonet observations). These are the first permanent identifiers of Oklahoma Mesonet data made available on its website.

### 3. Methodology: Oklahoma Mesonet Sensor Data Publication Analysis

The specifics of the study and its methods are described in this section. The overarching goal of the research effort is to identify, collect, and analyze all peer-reviewed scientific articles published in the English language that may have utilized sensor data from the Oklahoma Mesonet during the first decades of its existence, to determine who used the data, how it was used, and which specific data and data site(s) were used. Although the Oklahoma Mesonet has been routinely compiling a bibliography of publications pertinent to its efforts since its inception, the study described here is an initial attempt toward a formal analysis of the peer-reviewed scientific publications, their uses of the Oklahoma Mesonet data, and toward further efforts to understand and facilitate such Mesonet data usage.

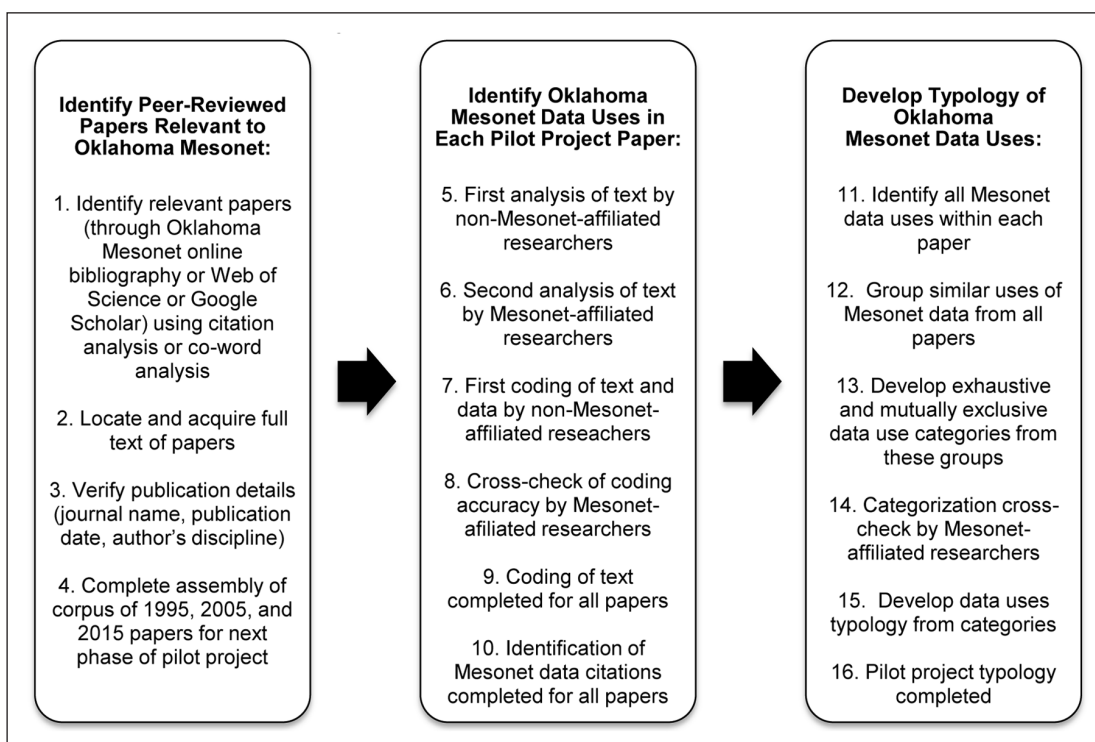
Mayernik *et al.* (2017: 1344) assert that 'Manual processes for collecting publications that cite data sets . . . are fraught with problems given that they are time-consuming, require multiple tools (*e.g.*, WoS, Google Scholar, and literature databases), and do not scale.' This project was undertaken with the understanding that efforts to identify and analyze all published journal articles using the Mesonet data will be time-consuming and complex, so a pilot project would be a useful initial step towards a full analysis, which would in turn further

inform and potentially impact Mesonet data operations and delivery. The initial pilot project processes are shown in **Figure 1**.

Prior to the start of this project in May 2015, the Oklahoma Mesonet online bibliography listed 509 peer-reviewed journal articles published between 1994 and 2014. The bibliography, although extensive, was known to be incomplete, as its collecting emphasis had been on research in the agricultural and atmospheric sciences and closely related areas. No 2015 journal publications had been identified and listed in the bibliography at that time. An important focus of the project, therefore, has been to identify, obtain, and add all relevant peer-reviewed scientific papers published between 1994 and the present day to the online bibliography.

The second step was to plan a pilot project that would inform future analysis of the corpus being constructed from articles obtained for the bibliography. In order to obtain a historical perspective on the development and diffusion of these articles, as well as an understanding of whether and how the Mesonet data and its uses have been described in various disciplines, three different publication years were selected for this initial analysis: the first full year of Mesonet operations (1994–5) hereafter referred to as 1995 for brevity, the end of the first decade of Mesonet operations (2005), and the end of the second decade of Mesonet operations (2015).

This study's methodology combines the standard bibliometric techniques of citation analysis and co-word analysis, with a close reading of the identified articles' text, graphic content, and references. Beginning with the online bibliography of articles routinely maintained by the Oklahoma Mesonet itself, an initial set of 11 journal articles from the Mesonet's first operational year was identified, providing the preliminary pilot project corpus. (Four of these articles were found outside the bibliography and added during the pilot project.) The second set of 21 articles from 2005 was also identified. (Three of these articles were found outside the bibliography and added during the pilot project.) Next, additional research using author, full text, keyword, and topic searches for the use of the phrases 'Oklahoma Mesonet,' 'Oklahoma Mesonet network,' and 'Oklahoma Climatological Survey' were performed in the Web of Science and Google Scholar for the publication year of 2015 to identify any peer-reviewed journal articles in English potentially using Oklahoma Mesonet data. (As the Mesonet DOIs mentioned above were not publicly available prior to March 2016, these were not part of the pilot project search process.) Additionally, any citations to the 1995 articles in 2005 and 2015 publications were examined, as possible sources of additional Oklahoma Mesonet data uses. All articles found in these searches were examined to determine whether they included explicit mention of the



**Figure 1:** Pilot project processes for Oklahoma Mesonet data.

use of Oklahoma Mesonet data. Those articles not making explicit use of Mesonet data were removed from the pilot project corpus, though still retained in the bibliography for future study.

All initial identification, categorization and coding of data usage were first done by project researchers not affiliated with the Oklahoma Mesonet, then checked for accuracy by researchers affiliated with the Mesonet, then re-checked by the researchers not affiliated with the Mesonet. This was to ensure that both scientific and non-scientific aspects of the data usage were captured correctly.

As shown in **Table 1**, the number of articles analyzed for each year is as follows: 1995 (11 articles), 2005 (21 articles), 2015 (78 articles) for a total of 110 articles using Oklahoma Mesonet data. Of these 110 articles, 85 were new additions to the online bibliography as well.

The 110 articles in this analysis were published in 62 different journals, in such diverse fields as agriculture (representing 18% of the 62 journals), biological sciences (13%), environmental sciences (27%), hydrology (16%), meteorology (18%), and other interdisciplinary areas (8%). The use of the data is clearly not confined to a single research specialty or discipline.

## 4. Findings

The findings from the pilot study of these 110 scientific publications, their citations of, and uses of the Oklahoma Mesonet sensor data are described in detail in this section.

### 4.1 Analysis of Oklahoma Mesonet Pilot Project Publications

#### 1995 Publications

As shown in **Table 1**, there were 11 peer-reviewed journal articles published during the initial year of Oklahoma Mesonet operations. These earliest publications focused on the economic considerations, technological specifications, and potential uses of the mesoscale network and its data for agriculture and meteorology.

The most highly-cited article from this period (Brock *et al.* 1995), co-authored by eight authors from the University of Oklahoma and Oklahoma State University, all closely involved in the development of the Oklahoma Mesonet, is a technical overview of the network itself, including its physical composition and procedures for quality assurance. The second most-highly-cited article (Rasmussen *et al.* 1994), co-written by seven authors from the National Oceanic & Atmospheric Administration's National Severe Storms Laboratory, also located in Norman, Oklahoma, describes the planned use of Mesonet data in the study of tornadogenesis and associated severe weather phenomena. The next three most-cited articles (Morrissey *et al.* 1995; Ryzhkov & Zrnic 1995a; Ryzhkov & Zrnic 1995b) deal with techniques for validating precipitation estimates through comparisons with empirical data from Mesonet rain gauges.

Thus, unsurprisingly, most early attention to the possibilities provided by data from the Oklahoma Mesonet came from Oklahoma-affiliated scientists interested in evaluating and extending existing measurement instruments and techniques, particularly in hydrology and meteorology.

#### 2005 Publications

The next decade of publications utilizing data from the Oklahoma Mesonet showed broadening in scientific scope by the 21 papers published in 2005. The Brock *et al.* (1995) paper continued to be well-cited for its technical information, for instance in Schroeder *et al.*'s (2005) description of the development of the West Texas Mesonet, which was modeled after the Oklahoma Mesonet. It was also cited continuously in atmospheric research such as Demoz *et al.*'s (2005) determination of the causes of undular bores in the boundary layer, in agricultural research such as MacKown & Carver's (2005) study of dual-purpose winter wheat cultivars, and in interdisciplinary investigations like those of Haugland & Crawford (2005) and McPherson & Stensrud (2005) on the impact of the replacement of tallgrass prairie by agricultural cropland on daily atmospheric conditions at the mesoscale.

**Table 1:** Analyzed articles from Oklahoma Mesonet corpus.

Year	Main
1995	11
2005	21
2015	78
<b>Total:</b>	<b>110</b>

A number of publications focused on uses of Oklahoma Mesonet data in other disciplines, such as biology (*e.g.*, Major *et al.*'s (2005) study of algal mass microbial communities in the Oklahoma salt flats), ecology (*e.g.*, Verberg *et al.*'s (2005) experimental study of soil CO<sub>2</sub> efflux within a tallgrass prairie ecosystem), and geology (*e.g.*, Harris, Tapp & Sublett's (2005) research on remediation of oil-field brine impacted soil). Mesonet data was also employed in several data assimilation projects, notably Dunne & Entekhabi (2005)'s extensive use of precipitation data to compare improvements in estimations by various filter ensembles for satellite remote sensing of soil moisture.

The geographic distribution of scientists using Mesonet data continued to broaden as well. While a third of the papers published in 2005 were authored by scientists at Oklahoma universities, the other two-thirds showed lead authorship from six other states, as well as from the National Aeronautics & Space Administration's Goddard Space Center in Maryland and the European Centre for Middle-Range Weather Forecasting in Great Britain.

## 2015 Publications

The further expansion in scope for use of Oklahoma Mesonet data is readily apparent in the 78 papers published in 2015. Interestingly, both the Brock *et al.* (1995) and the McPherson *et al.* (2007) articles were cited by 25% of these 2015 papers, though it was not possible to ascertain whether this was due more to their prevalence in the literature prior to June 2014 or to their recommendation as potential citations on the Oklahoma Mesonet website beginning in June 2014.

Atmospheric research, especially in regard to severe weather phenomena, exemplified by such work as that of Brotzge & Lutrell (2015) on interactions between boundary layers and mesocyclonic winds in tornadogenesis, Xu *et al.* (2015) on vortex wind analysis in tornadic mesocyclones, and French *et al.* (2015) on bulk "hook echo" precipitation drop size distributions in tornadic and nontornadic supercells, all continued extensive use of the atmospheric pressure, wind, and precipitation data. A growing number of papers used the soil moisture data to explore terrestrial impacts, particularly warning signs for flash droughts (Ford, *et al.* 2015) and for wildfires (Krueger *et al.* 2015), as well as for the effects of long-term drought (Bajgain *et al.* 2015).

Relatedly, environmental research using Oklahoma Mesonet data was performed for various species, ranging from the increasingly endangered greater prairie chicken (Hovick *et al.* 2015), to the more common Western chicken-turtle (MacKnight *et al.* 2015), to the wide assortment of crustaceans found in seemingly desiccated Oklahoma prairie playa (Bright & Bergey 2015).

Agricultural uses of the data continued to be strong, especially for research on improving management of resources such as cattle (Scasta *et al.* 2015), crops (Lollato & Edwards 2015), and timber (Bendixsen, Hallgren & Frazier 2015). Increased attention to potential energy resources, such as biofuels (Yimam, Ochsner & Kakani 2015), solar (McGovern *et al.* 2015), and wind (Stadler, Dryden & Green 2015) also drew upon Mesonet sensor data for economic and technical model building.

The most striking increase, however, was in the number of data assimilation projects, in which Mesonet surface data was combined with a variety of remote sensing data in order to develop models for various complex systems such as those involving boundary-layer (Endo *et al.* 2015), convection (Jones *et al.* 2015), erosion (Garbrecht & Zhang 2015), groundwater (Guzman *et al.* 2015), precipitation (Liu, Sorooshian & Gao 2015), river basin (Islam & Gan 2015), soil moisture (Kim, Mohanty & Shin 2015), watershed (Goldstein & Tarhule 2015) and wind speed (Genton, Padoan & Sang 2015) research. Validation efforts, such as the comparison by Kuster, Lunday & McPherson (2015) between North American Regional Climate Change Assessment Program output and Oklahoma Mesonet observations, emphasized the necessity of attending to the data themselves in all these endeavors.

Geographic diversity of lead authorship for these 78 papers using Mesonet data increased substantially in 2015 as well, indicating diffusion of data usage among different institutions during the decade since 2005. While Oklahoma-based researchers were lead authors of approximately 55% of the papers, the remaining 45% showed primary authorship by scientists from thirteen other states and four foreign countries.

## 4.2 Analysis of Oklahoma Mesonet Data Citation and Data Uses Categorization

This section describes how the Oklahoma Mesonet data were referenced, and how the data were used in the 110 papers from the pilot study. In 'How to Cite Earth Science Datasets,' Parsons (2012) observes that current data citation practices in earth sciences publications include the following: citing a traditional publication that contains the data; not mentioning the source of data, just using the data in tables or figures; referencing the name or source of data in text; including a Uniform Resource Locator (URL) in text (with variable degrees of specificity); citing a related paper; citing an actual data set, typically using a recommended citation format given by data center; or citing a data set including a persistent identifier, typically a DOI.

Since the Oklahoma Mesonet made its dataset DOIs available on its website in 2016, after the 2015 papers were published, no use of these DOIs was observed in this pilot study. Of the other six practices, the study found only three: referencing the name or source of data in text, including a URL in text (with variable degrees of specificity), or citing a related paper. The most commonly-used phrases used to refer to the Mesonet data sources in text were 'Oklahoma Mesonet' and 'Oklahoma Climatological Survey,' although these were only sometimes cited as sources in the reference lists of the papers. A number of papers (2 in 2005; 33 in 2015) provided the URL of either the Oklahoma Climatological Survey or the Oklahoma Mesonet in the text or the reference list. Finally, the Brock *et al.* 1995 paper often appeared as a reference accompanying an initial mention of the 'Oklahoma Mesonet' and could be considered to be a 'related paper.'

As noted above, citations to Oklahoma Mesonet data have varied to such an extent over the past 20 years, even within disciplines, that it remains a time-consuming task to identify these, and to determine their respective formats. However, as a general observation, it can be said that most of the researchers utilized only highly select portions of the entire datastream available to them, and most did not comment on why these specific selections were made, presumably because it was thought to be obvious to the reader within the particular scientific context.

Because the Mesonet datastreams include multiple observations (*e.g.*, temperature, windspeed, etc.) from specific station locations at specific times, and because the papers in the pilot project seldom included all details of the data actually used, further investigation would be necessary in most cases in order to identify and match the actual Mesonet data from specific site locations and times to recreate or replicate the research performed. This problem of so-called 'data granularity' (that is, the level of depth to which data can be referred) is endemic to large datasets and datastreams, however, and, in the case of the Mesonet, may eventually be addressed through new data curation and identification techniques. At present, however, it remains a very labor-intensive process to 'match' the stored Mesonet data to particular studies.

Uses of the Mesonet data shown in the pilot study were diverse, though the most prevalent uses tended to be event-oriented, focused on phenomena such as severe storms; geographically-oriented, focused on phenomena such as prairie environments; or time-oriented, focused on phenomena such as long-term droughts. Importantly, explicit use of the data also was seldom repeated or replicated by different studies, with the exception of a few specific severe storm events. This may be a result of the 'data granularity' issue discussed in the preceding paragraph.

Based on the descriptions given of various uses of the specific data within each of the 110 papers citing Oklahoma Mesonet data, we empirically derived a broad typology of data uses across disciplines, noting of course that different scientists may use the same data in different ways. The five major categories are *assimilation*, *experimentation*, *observation*, *simulation*, and *validation*. A sixth category, *utilization*, includes all other uses not otherwise categorized.

**Assimilation** describes the use of Oklahoma Mesonet sensor data in an ensemble with a variety of other data for purposes of system-level prediction or similar large-scale pattern analyses.

**Experimentation** involves the use of Oklahoma Mesonet sensor data as a controlled factor within an experimental or quasi-experimental environment.

**Observation** involves the use of Oklahoma Mesonet sensor data as a non-controlled factor within an experimental, quasi-experimental, or other environment.

**Simulation** involves the use of the Oklahoma Mesonet sensor data with or without other data for purposes of modeling.

**Utilization** indicates a direct utilization of Oklahoma Mesonet data not otherwise categorized here.

**Validation** involves the use of the Oklahoma Mesonet sensor data to validate other data or to ground-truth other instruments.

**Table 2** provides textual samples showing the variety in data citation formats and the six derived categories (assimilation, experimentation, observation, simulation, utilization, validation) from the data usages.

As shown in **Table 3**, an increasing variety of applications were made of the data as its availability and utility became more widely known among different research communities.

## 5. Conclusion

This pilot study of 110 papers from various scientific disciplines using Oklahoma Mesonet data over the past 20 years demonstrates the value of understanding the diversity of uses possible for a robust, rigorously maintained, easily available source of current and historical environmental data that can be used for experimentation, explanation, and prediction in a wide variety of disciplines. The pilot study shows that

**Table 2:** Textual samples of Oklahoma Mesonet data citation categorization.

Category	Source	Text of Data Citation
Assimilation	(Tanamachi, Heinselman & Wicker 2015: 511–512)	Twenty-eight Oklahoma Mesonet (Brock et al. 1995) stations were operating in the analysis domain on 24 May 2011 . . . recording observations of relative humidity RH, aspirated air temperature, wind speed, wind direction, and atmospheric pressure every 1 min. Of particular interest are the data from the El Reno Mesonet station [ <a href="http://ticker.mesonet.org/select.php?mo505&amp;da527&amp;yr52011">http://ticker.mesonet.org/select.php?mo505&amp;da527&amp;yr52011</a> ]. As the tornado passed, the atmospheric pressure decreased 17 hPa at 2120 UTC . . . while the 1-min-average ensemble of pressure traces from the lowest model scalar level (125 m AGL) to the surface using the hydrostatic equation and the surface pressure in the initial sounding (947 hPa). wind speed at 10 m AGL increased from 10 to 51 ms <sup>21</sup> . . . We derived simulated near-surface observations for comparison by interpolating the model variables at 125 m AGL from the grid points closest to the El Reno Mesonet station . . . We adjusted the simulated ensemble of pressure traces from the lowest model scalar level (125 m AGL) to the surface using the hydrostatic equation and the surface pressure in the initial sounding (947 hPa).
Experimentation	(Verburg et al. 2005: 1721–1722)	Starting on 10 February, 2002, temperatures inside the EcoCELLs were maintained based on 8-year (1993–2000), 5-min averages from a MESONET station (Brock et al., 1995) less than 1.6 km from the excavation site [in Purcell, Oklahoma] . . . Using a heating and cooling underneath the monoliths, soil temperatures in the deepest horizons (1.45 m depth) were kept as close as possible to the 1993–2000 mean annual ambient air temperature (16 °C) measured at the excavation site to simulate realistic soil temperature profiles . . . Precipitation was applied using an overhead rain simulator with a fixed intensity. We based frequency and amounts on measurements from the MESONET site (same 8-year period used for temperature) with precipitation at each watering applied at the mean monthly total divided by the mean frequency for that month.
Observation	(Tanner et al. 2015: 477–478)	Our objectives were to determine if sympatric populations of bobwhite and scaled quail respond behaviorally to artificial surface-water sources in a semiarid region at the species distribution extremes. . . Over the course of our study (2012–2014), average temperatures in summer ranged from 19.56–22.28, 25.72–27.22, and 26.78–30.06 °C during May, June, and July, respectively . . . Average temperatures in the winter ranged from –0.83 to 2.17, 1.28–1.33, and –0.33 to 2.39 °C during December, January, and February, respectively . . . Annual precipitation was 34.44, 50.29, and 39.42 cm in 2012, 2013, and 2014, respectively. . . Climate data were obtained from the Beaver Mesonet station (Brock et al., 1995; McPherson et al., 2007).
Simulation	(Stadler, Dryden & Greene 2015: 210–211)	Wind measurements at individual wind farm turbines would have been ideal for use in this research, but wind farm owners are notoriously protective of such data, which could be used by competitors to plot the efficiency of performance. Therefore, we used proxy data from the Oklahoma Mesonet network. The Oklahoma Mesonet network (Mesonet) is a network of 120 automated meteorological observation stations across all 77 counties in Oklahoma [36]. A full suite of data is reported and quality-controlled for five-minute time increments. Wind was measured at the standard 10-m height by the RM Young wind monitor with an accuracy of ±0.3 m s <sup>-1</sup> [37]. Each wind farm in Oklahoma is within 20 kilometers of a Mesonet station. We chose two wind farms in the western Oklahoma within areas of the most dramatic projected wind velocity changes in the NARCCAP output. Each wind farm was paired with the closest Mesonet station to represent that wind farm with a year of five-minute data (Figure 3). The Centennial wind farm in the northern portion of the study domain was paired with the Buffalo Mesonet station, and the Weatherford Wind Energy Center and the Weatherford Mesonet station were paired further south.

(Contd.)



Utilization	(Kenkel & Norris 1995: 365)	The two objectives of the study were to estimate average willingness to pay and to determine the characteristics of producers who would pay to access and use mesoscale weather information. To this end, two maximum likelihood models were estimated: one for the raw [Oklahoma] Mesonet data and one for the raw data/value-added information combination . . . Variables representing payments for agricultural publications, full- versus part-time farming, gross sales, use of irrigation, and weather-related crop income losses were found to significantly impact the willingness to pay for raw mesoscale weather data.
Validation	(Bajgain et al. 2015: 153)	Daily precipitation and soil moisture data from 2000–2013 at the Oklahoma Mesonet Marena and El Reno stations were downloaded from the Oklahoma Mesonet website ( <a href="http://www.mesonet.org/index.php/weather/daily_data_retrieval">www.mesonet.org/index.php/weather/daily_data_retrieval</a> ). The daily data were aggregated into 8-day periods to match with the temporal resolution of the Moderate-Resolution Imaging Spectroradiometer (MODIS) derived VIs. Three different soil moisture data products (soil water potential, fractional water index and volumetric water content) are available at the Mesonet website. These soil moisture data products were derived based on the calibrated change in soil temperature over time after a heat pulse is introduced (Illston et al., 2008). In our analysis, we used volumetric soil water content (SWC) collected at three different soil profiles (5, 25, and 60 cm depth.)

**Table 3:** Typology of Oklahoma Mesonet data uses.

Use Typology	1995	2005	2015	TOTAL
Assimilation	0	7	25	32
Experimentation	3	4	20	27
Observation	0	3	8	11
Simulation	0	1	15	16
Utilization	2	0	0	2
Validation	6	6	10	22
TOTAL	11	21	78	110

the Mesonet data are utilized in multiple scientific communities around the world, not only by the region's agricultural and meteorological researchers, as was originally anticipated by the Mesonet's founders and funders, and can potentially contribute to a much broader range of endeavors, making it even more essential. The development of an initial typology of Mesonet data uses is an important step in that direction, as it shows that these uses seem to be changing over time.

Besides the obvious value of Oklahoma Mesonet data for forecasting purposes, as well as in current meteorological and hydrological uses, the increasing historical value of the Mesonet data has become apparent, as shown by the numerous climate-related papers in the corpus. The recent implementation of Oklahoma Mesonet DOIs was done with the intent of making it easier to facilitate research uses in future and to increase stakeholder awareness of their importance. However, the current growth in data simulation and assimilation studies also makes evident that any single infrastructure such as the Mesonet that routinely supplies data for these rapidly evolving, data-intensive computerized models should be cognizant of the other stakeholders involved in such efforts, and attempt to work together on mutually beneficial outcomes, just as the researchers themselves do.

Efforts to correctly attribute assimilated data, as discussed by Hou & Mayernik (2013, 2017) in reference to climate data models, and by Tilmes, Yesha & Halem (2010) in reference to data transfer and transformations throughout the earth sciences in general, will likely become even more critical to data providers and other stakeholders. This realization has new implications for the Mesonet in terms of both its administrative and advocacy goals. The results of the pilot study, therefore, have already been of practical value, and an incentive toward further investigation of the past, present and potential uses of the Oklahoma Mesonet sensor data.

In an extensive review of research efforts similar to the Oklahoma Mesonet pilot project described here, Mayernik *et al.* (2017: 1354) have suggested that practical, in-depth studies of specific 'research infrastructures' that routinely provide significant research assets, such as datasets and datastreams, also offer unique

opportunities to contribute to theoretical knowledge of citation purposes in general, though this is seldom an explicit purpose of such studies. They note that data citations and acknowledgments “do not fall cleanly into either the traditional scholarly ‘citation as reward’ or ‘citation as persuasion’ categories but may represent a mix of both. Citing a data source may thus serve to recognize the contribution of the data provider and also to illustrate ‘how data of sufficiently high visibility or quality underlie a particular scientific result.’ As an example, they offer several analyses of successive ‘release articles’ related to a specific oceanographic and atmospheric data set which showed that roughly 80% of the citations to the release articles were data-usage related, which may indicate that certain published articles can be highly cited as proxies for data sets.

Although the Oklahoma Mesonet pilot study was not designed to focus on the Brock *et al.* 1995 article in this way, the findings do suggest that it may serve not only as a proxy for the data source, but also as an indication of the high quality of the data source. The key role of quality assurance in sensor data research, as shown by the continued references to this article and to other Mesonet articles that provide the technical and instrumental sensor specifications for data collection and communication, may be viewed as implicit signals of the quality of the data source and, accordingly, to that of the datastreams and to the data used in each particular study. A future study focused on this ‘signal’ aspect of sensor data citation is now being planned, with special attention to its possible significance and the various channels by which such ‘signal papers’ may be diffused.

Another planned study involving the preliminary typology of data uses derived from the articles in the pilot study is its further use and potential modification during the analysis of all journal articles in the full corpus. A comparison of these uses to those described in peer-reviewed journal publications using sensor data from the growing number of similar mesonetworks in other states, offering an opportunity for more collaboration with other researchers interested in this type of study, is also being planned.

In conclusion, as the Oklahoma Mesonet can provide a uniquely valuable source of environmental sensor data to those who study a wide range of weather-related phenomena, it can also provide an invaluable source of information to those researchers interested in how that data, its uses, and its users may be studied.

## Author Note

The online bibliography of Oklahoma Mesonet-related journal articles is available at <https://www.mesonet.org/index.php/site/bibliography>.

## Acknowledgements

The authors wish to thank the reviewers of the initial version of this article for their helpful comments and criticisms. The research assistance of Ms. Katherine Morris of the University of Oklahoma School of Library & Information Studies is also gratefully acknowledged.

## Competing Interests

BVM received one month of summer salary support from the Oklahoma Climatological Survey in 2015 to help fund her work on this project. CAF and BGI are both senior scientists employed by the Oklahoma Climatological Survey, which operates the Oklahoma Mesonet.

## Author Information

Betsy Van der Veer Martens is an associate professor at the School of Library & Information Studies at the University of Oklahoma, and has written on bibliometrics in the *Annual Review of Information Science & Technology* and the *Journal of the Association for Information Science & Technology*. Bradley G. Illston is a senior research scientist at the Oklahoma Mesonet. His research focuses on urban meteorology and soil moisture, and has appeared in such journals as the *Bulletin of the American Meteorological Society*, the *Journal of Atmospheric & Oceanic Technology*, and the *Soil Science Society of America Journal*. Christopher A. Fiebrich is the Associate Director of the Oklahoma Climatological Survey and the Executive Director of the Oklahoma Mesonet. His research interests include the history and operation of climate observational networks, and he has published extensively in the *Bulletin of the American Meteorological Society*, *Earth-Science Reviews*, the *Journal of Atmospheric & Oceanic Technology*, the *Journal of Environmental Quality*, and elsewhere.

## References

- Altman, M, Borgman, C, Crosas, M and Matone, M** 2015 An introduction to the joint principles for data citation. *Bulletin of the American Society for Information Science and Technology*, 41(3): 43–45. DOI: <https://doi.org/10.1002/bult.2015.1720410313>

- Apai, D, Lagerstrom, J, Reid, I N, Levay, K L, Fraser, E, Nota, A and Henneken, E** 2010 Lessons from a high-impact observatory: The Hubble Space Telescope's science productivity between 1998 and 2008. *Publications of the Astronomical Society of the Pacific*, 122(893): 808–826. DOI: <https://doi.org/10.1086/654851>
- Apte, A** 2015 An introduction to data assimilation. In: Sarkar, S. et al. (eds.), *Applied mathematics. Springer proceedings in mathematics & statistics*, 146: 31–42. New Delhi: Springer. DOI: [https://doi.org/10.1007/978-81-322-2547-8\\_4](https://doi.org/10.1007/978-81-322-2547-8_4)
- Bajgain, R, Xiao, X, Wagle, P, Basara, J and Zhou, Y** 2015 Sensitivity analysis of vegetation indices to drought over two tallgrass prairie sites. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108: 151–160. DOI: <https://doi.org/10.1016/j.isprsjprs.2015.07.004>
- Bates, J, Lin, Y W and Goodale, P** 2016 Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society*, 3(2): 2053951716654502. DOI: <https://doi.org/10.1177/2053951716654502>
- Bauer, P, Thorpe, A and Brunet, G** 2015 The quiet revolution of numerical weather prediction. *Nature*, 525(7567): 47–55. DOI: <https://doi.org/10.1038/nature14956>
- Belter, C W** 2014 Measuring the value of research data: A citation analysis of oceanographic data sets. *PLoS ONE*, 9(3): e92590. DOI: <https://doi.org/10.1371/journal.pone.0092590>
- Bendixsen, D P, Hallgren, S W and Frazier, A E** 2015 Stress factors associated with forest decline in xeric oak forests of south-central United States. *Forest Ecology and Management*, 347: 40–48. DOI: <https://doi.org/10.1016/j.foreco.2015.03.015>
- Borgman, C L** 2015 *Big data, little data, no data: Scholarship in the networked world*. Cambridge MA: MIT Press.
- Borgman, C L, Darch, P T, Sands, A E, Wallis, J C and Traweek, S** 2014 The ups and downs of knowledge infrastructures in science: Implications for data management. In *Proceedings of the IEEE/ACM joint conference on digital libraries. September 8–12, 2014*, 257–266. London, U.K.
- Bright, E G and Bergey, E A** 2015 Spatial distribution of resting stages of crustaceans and other aquatic invertebrates in playa wetlands. *Journal of Crustacean Biology*, 35(4): 515–521. DOI: <https://doi.org/10.1163/1937240X-00002347>
- Brock, F** 2013 Oral history interview with Fred V. Brock. Available at: <http://dc.library.okstate.edu/cdm/ref/collection/mesonet/id/266> [Last accessed 27 July 2017].
- Brock, F V, Crawford, K C, Elliott, R L, Cuperus, G W, Stadler, S J, Johnson, H L and Eilts, M D** 1995 The Oklahoma Mesonet: A technical overview. *Journal of Atmospheric and Oceanic Technology*, 12(1): 5–19. DOI: [https://doi.org/10.1175/1520-0426\(1995\)012<0005:TOMATO>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2)
- Brotzge, J A and Luttrell, C M** 2015 Genesis of the Chickasha, Oklahoma, tornado on 24 May 2011 as observed by CASA radar and Oklahoma Mesonet. *Journal of Operational Meteorology*, 3(6): 59–69. DOI: <https://doi.org/10.15191/nwajom.2015.0306>
- Campbell, J L, et al.** 2013 Quantity is nothing without quality: Automated QA/QC for streaming environmental sensor data. *BioScience*, 63(7): 574–585. DOI: <https://doi.org/10.1525/bio.2013.63.7.10>
- Cragin, M H, Palmer, C L and Chao, T C** 2010 Relating data practices, types, and curation functions: An empirically derived framework. *Proceedings of the American Society for Information Science and Technology*, 47(1): 1–2. DOI: <https://doi.org/10.1002/meet.14504701426>
- Crawford, K** 2013 Oral history interview with Ken Crawford. Available at: <http://dc.library.okstate.edu/cdm/ref/collection/mesonet/id/293> [Last accessed 27 July 2017].
- Cronin, B** 1995 *The scholar's courtesy: The role of acknowledgment in the primary communication process*. London: Taylor Graham.
- Data Citation Synthesis Group** 2014 Joint Declaration of Data Citation Principles. Martone, M (ed.), San Diego CA: FORCE11. Available at: <https://www.force11.org/group/joint-declaration-data-citation-principles-final> [Last accessed 27 July 2017].
- Demoz, B B, Starr, D O C, Evans, K D, Lare, A R, Whiteman, D N, Schwemmer, G, Bisson, S E, et al.** 2005 The cold front of 15 April 1994 over the central United States. Part I: Observations. *Monthly Weather Review*, 133(6): 1525–1543. DOI: <https://doi.org/10.1175/MWR2932.1>
- Ding, Y, Hayes, M J and Widhalm, M** 2011 Measuring economic impacts of drought: A review and discussion. *Disaster Prevention and Management*, 20(4): 434–446. DOI: <https://doi.org/10.1108/096535611111161752>

- Dunne, S** and **Entekhabi, D** 2005 An ensemble-based reanalysis approach to land data assimilation. *Water Resources Research*, 41(2): W02013. DOI: <https://doi.org/10.1029/2004WR003449>
- Edwards, P N** 2010 *A vast machine: Computer models, climate data, and the politics of global warming*. Cambridge MA: MIT Press.
- Endo, S, Frielind, A M, Lind, W, Vogelmann, A M, Toto, T, Ackerman, A S, Liu, Y**, et al. 2015 RACORO continental boundary layer cloud investigations: 2. Large-eddy simulations of cumulus clouds and evaluation within situ and ground-based observations. *Journal of Geophysical Research: Atmospheres*, 120(12): 5993–6014. DOI: <https://doi.org/10.1002/2014JD022525>
- Fiebrich, C A** 2009 History of surface weather observations in the United States. *Earth-Science Reviews*, 93(3): 77–84. DOI: <https://doi.org/10.1016/j.earscirev.2009.01.001>
- Fleming, J R** 2016 *Inventing atmospheric science: Bjerknes, Rossby, Wexler, and the foundations of modern meteorology*. Cambridge MA: MIT Press.
- Ford, T W, McRoberts, D B, Quiring, S M** and **Hall, R E** 2015 On the utility of in situ soil moisture observations for flash drought early warning in Oklahoma, USA. *Geophysical Research Letters*, 42(22): 9790–9798. DOI: <https://doi.org/10.1002/2015GL066600>
- French, M M, Burgess, D W, Mansell, E R** and **Wicker, L J** 2015 Bulk hook echo raindrop sizes retrieved using mobile, polarimetric Doppler radar observations. *Journal of Applied Meteorology and Climatology*, 54(2): 423–450. DOI: <https://doi.org/10.1175/JAMC-D-14-0171.1>
- Ganguly, A R, Omitaomu, O A, Fang, Y, Khan, S** and **Bhaduri, B L** 2007 Knowledge discovery from sensor data for scientific applications. In: Gama, J, et al. (eds.), *Learning from data streams: Processing techniques in sensor networks*, 205–228. Berlin: Springer. DOI: [https://doi.org/10.1007/3-540-73679-4\\_13](https://doi.org/10.1007/3-540-73679-4_13)
- Garbrecht, J D** and **Zhang, X C** 2015 Soil erosion from winter wheat cropland under climate change in central Oklahoma. *Applied Engineering in Agriculture*, 31(3): 439–454. DOI: <https://doi.org/10.13031/aea.31.10998>
- Genton, M G, Padoan, S A** and **Sang, H** 2015 Multivariate max-stable spatial processes. *Biometrika*, 102(1): 215–230. DOI: <https://doi.org/10.1093/biomet/asu066>
- Goldstein, J C, Mayernik, M S** and **Ramapriyan, H K** 2017 Identifiers for earth science data sets: Where we have been and where we need to go. *Data Science Journal*, 16: 23. DOI: <https://doi.org/10.5334/dsj-2017-023>
- Goldstein, J C** and **Tarhule, A** 2015 Evaluating the impacts of climate change and switchgrass production on a semiarid basin. *Hydrological Processes*, 29(5): 724–738. DOI: <https://doi.org/10.1002/hyp.10159>
- Guzman, J A, Moriasi, D N, Gowda, P H, Steiner, J L, Starks, P J, Arnold, J G** and **Srinivasan, R** 2015 A model integration framework for linking SWAT and MODFLOW. *Environmental Modelling & Software*, 73: 103–116. DOI: <https://doi.org/10.1016/j.envsoft.2015.08.011>
- Halevi, G** and **Moed, F** 2012 The evolution of big data as a research and scientific topic: Overview of the literature. *Research Trends*, 30(1): 3–6. Available at: <https://www.researchtrends.com/issue-30-september-2012/the-evolution-of-big-data-as-a-research-and-scientific-topic-overview-of-the-literature/> [Last accessed 27 July 2017].
- Harris, T A, Tapp, J B** and **Sublette, K L** 2005 Remediation of oil-field brine-impacted soil using a subsurface drainage system and hay. *Environmental Geosciences*, 12: 101–113. DOI: <https://doi.org/10.1306/eg.11160404041>
- Haugland, M J** and **Crawford, K C** 2005 The diurnal cycle of land-atmosphere interactions across Oklahoma's winter wheat belt. *Monthly Weather Review*, 133(1): 120–130. DOI: <https://doi.org/10.1175/MWR-2842.1>
- Hou, C Y** and **Mayernik, M** 2016 Recognizing the diversity of contributions: A case study for framing attribution and acknowledgement for scientific data. *International Journal of Digital Curation*, 11(1): 33–52. DOI: <https://doi.org/10.2218/ijdc.v11i1.357>
- Hou, C Y** and **Mayernik, M** 2017 Formalizing an attribution framework for scientific data/software products and collections. *International Journal of Digital Curation*, 11(2): 87–103. DOI: <https://doi.org/10.2218/ijdc.v11i2.404>
- Hovick, T J, Elmore, R D, Fuhlendorf, S D** and **Dahlgren, D K** 2015 Weather constrains the influence of fire and grazing on nesting greater prairie-chickens. *Rangeland Ecology & Management*, 68(2): 186–193. DOI: <https://doi.org/10.1016/j.rama.2015.01.009>
- Islam, Z** and **Gan, T Y** 2015 Hydrologic modeling of the Blue River basin using NEXRAD precipitation data with a semidistributed and a fully distributed model. *Journal of Hydrologic Engineering*, 20(10): 04015015-1-04015015-9. DOI: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001179](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001179)

- Johnson, R C** 2015 Roadmap to Trillion Sensor Forks: System-Side Here, Sensor-Side Not for 6–12 Months. *EE Times*, December 10, 2015. Available at: [http://www.eetimes.com/document.asp?doc\\_id=1328466](http://www.eetimes.com/document.asp?doc_id=1328466) [Last accessed 27 July 2017].
- Jones, T A, Stensrud, D, Wicker, L, Minnis, P and Palikonda, R** 2015 Simultaneous radar and satellite data storm-scale assimilation using an ensemble Kalman filter approach for 24 May 2011. *Monthly Weather Review*, 143(1): 165–194. DOI: <https://doi.org/10.1175/MWR-D-14-00180.1>
- Kenkel, P L and Norris, P E** 1995 Agricultural producers' willingness to pay for real-time mesoscale weather information. *Journal of Agricultural and Resource Economics*, 20(2): 356–372. <https://www.jstor.org/stable/40987744>.
- Kim, J, Mohanty, B P and Shin, Y** 2015 Effective soil moisture estimate and its uncertainty using multimodel simulation based on Bayesian model averaging. *Journal of Geophysical Research: Atmospheres*, 120(16): 8023–8042. DOI: <https://doi.org/10.1002/2014JD022905>
- Kim, Y and Stanton, J M** 2016 Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, 67: 776–799. DOI: <https://doi.org/10.1002/asi.23424>
- Klemm, T and McPherson, R A** 2017 The development of seasonal climate forecasting for agricultural producers. *Agricultural and Forest Meteorology*, 232, 384–399. DOI: <https://doi.org/10.1016/j.agrformet.2016.09.005>
- Klockow, K E, McPherson, R A and Sutter, D S** 2010 On the economic nature of crop production decisions using the Oklahoma Mesonet. *Weather, Climate, and Society*, 2(3): 224–236. DOI: <https://doi.org/10.1175/2010WCAS1034.1>
- Krueger, E S, Ochsner, T E, Engle, D M, Carlson, J D, Twidwell, D and Fuhlendorf, S D** 2015 Soil moisture affects growing-season wildfire size in the Southern Great Plains. *Soil Science Society of America*, 79(6): 1567–1576. DOI: <https://doi.org/10.2136/sssaj2015.01.0041>
- Kuster, E L, Lunday, C E and McPherson, R A** 2015 A comparison of North American regional climate change assessment program output and Oklahoma Mesonet observations: Precipitation and temperature. *The Geographical Bulletin*, 56(2): 102–121.
- Liu, H, Sorooshian, S and Gao, X** 2015 Assessment of the spatial and seasonal variation of the error-intensity relationship in satellite-based precipitation measurements using an adaptive parametric model. *Journal of Hydrometeorology*, 16(4): 1700–1716. DOI: <https://doi.org/10.1175/JHM-D-14-0219.1>
- Lollato, R P and Edwards, J T** 2015 Maximum attainable wheat yield and resource-use efficiency in the southern Great Plains. *Crop Science*, 55(6): 2863–2876. DOI: <https://doi.org/10.2135/cropsci2015.04.0215>
- Lynch, P** 2008 The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, 227(7): 3431–3444. DOI: <https://doi.org/10.1016/j.jcp.2007.02.034>
- MacKown, C T and Carver, B F** 2005 Fall forage biomass and nitrogen composition of winter wheat populations selected from grain-only and dual-purpose environments. *Crop Science*, 45(1): 322–328. DOI: <https://doi.org/10.2135/cropsci2005.0322>
- Major, K M, Kirkwood, A E, Major, C S, McCreddie, J W and Henley, W J** 2005 In situ studies of algal biomass in relation to physicochemical characteristics of the Salt Plains National Wildlife Refuge, Oklahoma, USA. *Saline Systems*, 1(11). DOI: <https://doi.org/10.1186/1746-1448-1-11>
- Mayernik, M S** 2012 Data citation initiatives and issues. *Bulletin of the American Society for Information Science and Technology*, 38(5): 23–28. DOI: <https://doi.org/10.1002/bult.2012.1720380508>
- Mayernik, M S** 2016 Research data and metadata curation as institutional issues. *Journal of the Association for Information Science and Technology*, 67(4): 973–993. DOI: <https://doi.org/10.1002/asi.23425>
- Mayernik, M S, Hart, D L, Maull, K E and Weber, N M** 2017 Assessing and tracing the outcomes and impact of research infrastructures. *Journal of the Association for Information Science and Technology*, 68(6): 1341–1359. DOI: <https://doi.org/10.1002/asi.23721>
- Mayernik, M S, Wallis, J C and Borgman, C L** 2013 Unearthing the infrastructure: Humans and sensors in field-based scientific research. *Computer Supported Cooperative Work*, 22(1): 65–101. DOI: <https://doi.org/10.1007/s10606-012-9178-y>
- McGovern, A, Gagne, D J, II, Basara, J, Hamill, T M and Margolin, D** 2015 Solar energy prediction: An international contest to initiate interdisciplinary research on compelling meteorological problems. *Bulletin of the American Meteorological Society*, 96(8): 1388–1393. DOI: <https://doi.org/10.1175/BAMS-D-14-00006.1>


- McPherson, R A, Fiebrich, C A, Crawford, K C, Elliott, R L, Kilby, J R, Grimsley, D L, Martinez, J E, Demko, D B**, et al. 2007 Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology*, 24(3): 301–321. DOI: <https://doi.org/10.1175/JTECH1976.1>
- McPherson, R A** and **Stensrud, D J** 2005 Influences of a winter wheat belt on the evolution of the boundary layer. *Monthly Weather Review*, 133(8): 2178–2199. DOI: <https://doi.org/10.1175/MWR2968.1>
- Mooney, H** and **Newton, M P** 2012 The anatomy of a data citation: Discovery, reuse, and credit. *Journal of Librarianship and Scholarly Communication*, 1(1): eP1035. DOI: <https://doi.org/10.7710/2162-3309.1035>
- Morrissey, M L, Maliekal, J A, Greene, J S** and **Wang, J M** 1995 The uncertainty of simple spatial averages using rain-gauge networks. *Water Resources Research*, 31(8): 2011–2017. DOI: <https://doi.org/10.1029/95WR01232>
- National Research Council** 2008 *Observing weather and climate from the ground up: A nationwide network of networks*. Washington DC: National Academies Press.
- Parsons, M A** 2012 How to cite earth science datasets. In: Uhler, P (ed.), *For attribution—Developing data attribution and citation practices and standards*, 117–124. Washington DC: National Academies Press.
- Parsons, M A** and **Fox, P A** 2013 Is data publication the right metaphor?. *Data Science Journal*, 12: WDS32–WDS46. DOI: <https://doi.org/10.2481/dsj.WDS-042>
- Parsons, M A, Godøy, Ø, LeDrew, E, De Bruin, T F, Danis, B, Tomlinson, S** and **Carlson, D** 2011 A conceptual framework for managing very diverse data for complex, interdisciplinary science. *Journal of Information Science*, 37(6): 555–569. DOI: <https://doi.org/10.1177/0165551511412705>
- Peng, G, Privette, J L, Kearns, E J, Ritchey, N A** and **Ansari, S** 2015 A unified framework for measuring stewardship practices applied to digital environmental datasets. *Data Science Journal*, 13: 231–253. DOI: <https://doi.org/10.2481/dsj.14-049>
- Porter, J H, Hanson, P C** and **Lin, C C** 2012 Staying afloat in the sensor data deluge. *Trends in Ecology & Evolution*, 27(2): 121–129. DOI: <https://doi.org/10.1016/j.tree.2011.11.009>
- Rasmussen, E N, Straka, J M, Davies-Jones, R, Doswell, C A, III, Carr, F H, Eilts, M D** and **MacGorman, D R** 1994 Verification of the origins of rotation in tornadoes experiment: VORTEX. *Bulletin of the American Meteorological Society*, 75(6): 995–1006. DOI: [https://doi.org/10.1175/1520-0477\(1994\)075<0995:VOT OOR>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<0995:VOT OOR>2.0.CO;2)
- Robinson-García, N, Jiménez-Contreras, E** and **Torres-Salinas, D** 2016 Analyzing data citation practices using the Data Citation Index. *Journal of the Association for Information Science and Technology*, 67(12): 2964–2975. DOI: <https://doi.org/10.1002/asi.23529>
- Rueda, L, Fenner, M** and **Cruse, P** 2016 DataCite: Lessons learned on persistent identifiers for research data. *International Journal of Digital Curation*, 11(2): 39–47. DOI: <https://doi.org/10.2218/ijdc.v11i2.421>
- Ryzhkov, A V** and **Zrnic, D S** 1995a Comparison of dual-polarization radar estimators of rain. *Journal of Atmospheric and Oceanic Technology*, 12(2): 249–256. DOI: [https://doi.org/10.1175/1520-0426\(1995\)012<0249:CODPRE>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0249:CODPRE>2.0.CO;2)
- Ryzhkov, A V** and **Zrnic, D S** 1995b Precipitation and attenuation measurements at a 10-cm wavelength. *Journal of Applied Meteorology*, 34(10): 2121–2134. DOI: [https://doi.org/10.1175/1520-0450\(1995\)034<2120:PAAMAA>2.0.CO;2](https://doi.org/10.1175/1520-0450(1995)034<2120:PAAMAA>2.0.CO;2)
- Scasta, J D, Engle, D M, Talley, J L, Weir, J R, Fuhlendorf, S D** and **Debinski, D M** 2015 Drought influences control of parasitic flies of cattle on pastures managed with patch-burn grazing. *Rangeland Ecology & Management*, 68(3): 290–297. DOI: <https://doi.org/10.1016/j.rama.2015.03.001>
- Schroeder, J L, Burgett, W S, Haynie, K B, Sonmez, I, Skwira, G D, Doggett, A L** and **Lipe, J W** 2005 The West Texas Mesonet: A technical overview. *Journal of Atmospheric and Oceanic Technology*, 22(2): 211–222. DOI: <https://doi.org/10.1175/JTECH-1690.1>
- Singhal, A, Kasturi, R** and **Srivastava, J** 2014 DataGopher: Context-based search for research datasets. In *Proceedings of the 2014 IEEE 15th international conference on information reuse and integration*, 1: 749–756. DOI: <https://doi.org/10.1109/IRI.2014.7051964>
- Singhal, A** and **Srivastava, J** 2013 Data extract: Mining context from the web for dataset extraction. *International Journal of Machine Learning and Computing*, 3(2): 219–223. DOI: <https://doi.org/10.7763/IJMLC.2013.V3.306>
- Singhal, A** and **Srivastava, J** 2017 Research dataset discovery from research publications using web context. *Web Intelligence*, 15(2): 81–99. DOI: <https://doi.org/10.3233/WEB-170354>
- Stadler, S, Dryden, J M** and **Greene, J S** 2015 Climate change impacts on Oklahoma wind resources: Potential energy output changes. *Resources*, 4(2): 203–226. DOI: <https://doi.org/10.3390/resources4020203>

- Strasser, B J** 2012 Data-driven sciences: From wonder cabinets to electronic databases. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43(1): 85–87. DOI: <https://doi.org/10.1016/j.shpsc.2011.10.009>
- Tanner, E P, Elmore, R D, Fuhlendorf, S D, Davis, C A, Thacker, E T and Dahlgren, D K** 2015 Behavioral responses at distribution extremes: How artificial surface water can affect quail movement patterns. *Rangeland Ecology & Management*, 68(6): 476–484. DOI: <https://doi.org/10.1016/j.rama.2015.07.008>
- Tilmes, C, Yesha, Y and Halem, M** 2010 Tracking provenance of earth science data. *Earth Science Informatics*, 3(1–2): 59–65. DOI: <https://doi.org/10.1007/s12145-010-0046-3>
- Unidata** 2017 Network Common Data Format. DOI: <https://doi.org/10.5065/D6H70CW6>
- Verburg, P S J, Larsen, J, Johnson, D W, Schorran, D E and Arnone, J A** 2005 Impacts of an anomalously warm year on soil CO<sub>2</sub> efflux in experimentally manipulated tallgrass prairie ecosystems. *Global Change Biology*, 11(10): 1720–1732. DOI: <https://doi.org/10.1111/j.1365-2486.2005.001032.x>
- Waite, J L** 2011 Using the urban regime framework to analyze hazard mitigation sustainability at the local scale. Ph.D. Dissertation, University of Maryland, College Park. Available at: <http://drum.lib.umd.edu/handle/1903/12290> [Last accessed 27 July 2017].
- Williams, R B** 2011 Citation systems in the biosciences: A history, classification and descriptive terminology. *Journal of Documentation*, 67(6): 995–1014. DOI: <https://doi.org/10.1108/00220411111183564>
- Xu, Q, Wei, L, Nai, K, Liu, S, Rabin, R M and Zhao, Q** 2015 A radar wind analysis system for nowcast applications. *Advances in Meteorology*, Article ID 264515. DOI: <https://doi.org/10.1155/2015/264515>
- Yimam, Y T, Ochsner, T E and Kakani, V G** 2015 Evapotranspiration partitioning and water use efficiency of switchgrass and biomass sorghum managed for biofuel. *Agricultural Water Management*, 155, 40–47. DOI: <https://doi.org/10.1016/j.agwat.2015.03.018>
- Ziolkowska, J R, Fiebrich, C A, Carlson, J D, Melvin, A D, Sutherland, A J, Kloesel, K A, Reyes, R, et al.** 2017 Benefits and beneficiaries of the Oklahoma Mesonet: A multi-sectoral ripple effect analysis. *Weather, Climate & Society*, 9(3): 499–519. DOI: <https://doi.org/10.1175/WCAS-D-16-0139.1>

**How to cite this article:** Van der Veer Martens, B, Illston, B G and Fiebrich, C A 2017 The Oklahoma Mesonet: A Pilot Study of Environmental Sensor Data Citations. *Data Science Journal*, 16: 47, pp.1–15, DOI: <https://doi.org/10.5334/dsj-2017-047>

**Submitted:** 01 March 2017      **Accepted:** 20 September 2017      **Published:** 04 October 2017

**Copyright:** © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

 *Data Science Journal* is a peer-reviewed open access journal published by Ubiquity Press.

**OPEN ACCESS** 