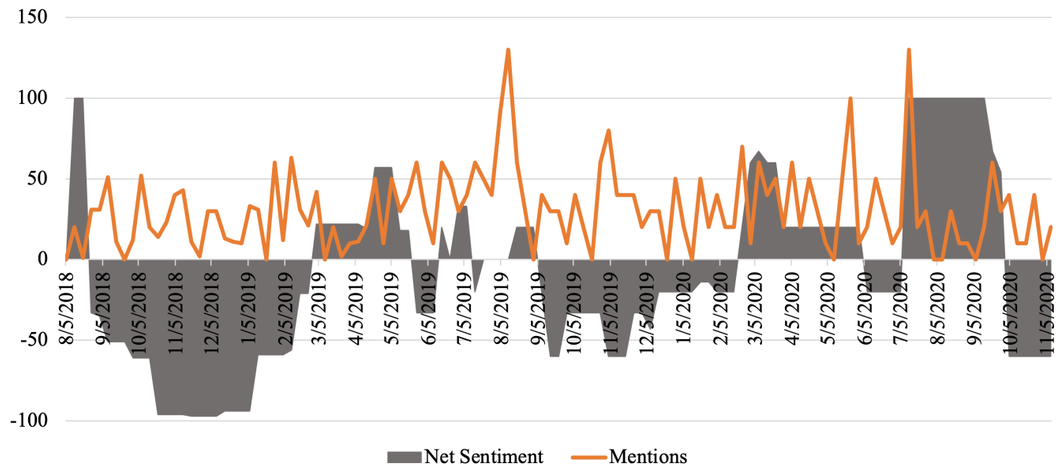

Figure 3. Net sentiment and Mentions over time.

earthquakes. For now at least, actual testing of water samples remains the best way to detect issues with tap water quality.

Conclusions

Although social media data provides interesting information, it is unlikely to serve as a measure or indicator of water quality any time soon. Many of the discussions that were happening on social media were of a national, or general interest nature, with very little information about the poster's personal water quality. The EPA water quality violations indicates that most counties in Oklahoma have fairly good quality water, with most counties having less than 15 violations during the time period studied. It is interesting to note that rural counties had fewer violations than urban counties, a point for further research and discussion.

Acknowledgement: The authors acknowledge Dr. Nicole Widmar in the Department of Agricultural Economics at Purdue University for providing access to the Netbase platform through collaboration within her ongoing related research.

References

Chatfield, Akemi, and Brajawidagda, Uuf. 2012. Twitter tsunami early warning network: a social network analysis of Twitter information flows. Faculty of Engineering and Information Sciences - Papers: Part A. 136.

Crowley, James. 2019. Tap water at trump national golf course contaminated with toxic forever chemicals. Available from: https://www.newsweek.com/trump-national-golf-course-new-jersey-toxic-chemicals-tap-water-1470189

Fouriezos, Nick. 2019. Love thy neighbor: The bible belt is becoming a dumping ground. OZY. Available from: https://www.ozy.com/the-new-and-the-next/love-thy-neighbor-the-bible-belt-is-becoming-a-dumping-ground/93854/ Accessed 12-25-20

Netbase (2020). Overview. Available from: https://netbasequid.com/wp-content/uploads/NetBaseQuid_overview.pdf Accessed 12-14-20.

Sakaki Takeshi, Okazaki Makoto, Matsuo Yutaka. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. Proceedings of the 19th international conference on World wide web. ACM.

USDA. 2020. Rural Urban Continuum code documentation. Available from: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation/ Accessed 2-4-21

USEPA (The United States Environmental Protection Agency), 2020. SDWIS Federal Reports Advanced Search. Available from: https://ofmpub.epa.gov/apex/sfdw/f?p=108:9 Accessed 11-9-2020.

USEPA (The United States Environmental Protection Agency), 2018. Drinking Water Contaminant Human Health Effects Information, 2018. Online at https://www.epa.gov/sdwa/drinking-water-contaminant-human-health-effects-information (accessed online November 12, 2020)

USEPA (The United States Environmental Protection Agency), 1999. 25 years of drinking water act history and trends. EPA 816-R-99-007, December.

Zhang, Hailin. (2017). Drinking Water Testing. PSS-2912 https://extension.okstate.edu/fact-sheets/drinking-water-testing.html

Stainley, Tim. (2021). How much do home and whole house water filter systems cost? Available online at https://www.best-osmosis-systems.com/water-filtration-system-cost/ (Accessed online January 27,2021)

Winter, Thomas C., Harvey, Judson W., Franke, O. Lehn, and Alley, William M. (1998). Ground water and surface water a single resource U.S. Geological Survey Circular 1139.

Harter, Thomas. (2003). Groundwater quality and groundwater pollution. UC Agriculture & Natural Resources. http:/doi.org/10.3733/ucanr.8084

Oklahoma State University, as an equal opportunity employer, complies with all applicable federal and state laws regarding non-discrimination and affirmative action. Oklahoma State University is committed to a policy of equal opportunity for all individuals and does not discriminate based on race, religion, age, sex, color, national origin, marital status, sexual orientation, gender identity/expression, disability, or veteran status with regard to employment, educational programs and activities, and/or admissions. For more information, visit https:///eeo.okstate.edu.

Issued in furtherance of Cooperative Extension work, acts of May 8 and June 30, 1914, in cooperation with the U.S. Department of Agriculture, Director of Oklahoma Cooperative Extension Service, Oklahoma State University, Stillwater, Oklahoma. This publication is printed and issued by Oklahoma State University as authorized by the Vice President for Agricultural Programs and has been prepared and distributed at a cost of 20 cents per copy. 03/2021 GH.

EXTENSION

Considering Water Quality in Oklahoma

March 2021

Courtney Bir
Farm Management Specialist

Lixia H. Lambert
Assistant Professor

Oklahoma Cooperative Extension Fact Sheets are also available on our website at: **extension.okstate.edu**

Introduction

Often referred to as "tap water" drinking water from public water systems is an important resource. This current report seeks to evaluate the perceptions of water quality. Two measures are used including the number of water violations per county in Oklahoma, and social media data related to tap water. In combination, we assess whether social media can be useful for looking at local water quality perceptions and as a potential warning, warranting additional testing.

The safe drinking water act (SDWA), which was passed in 1974, is the primary legislation that sets legal and health standards for drinking water in the U.S. Before legislation, only 60% of public water systems could meet federal standards. Great improvements have been made, with over 90% of tap water now meeting water quality standards. However, there are still issues with providing the same high quality of water to all Americans, especially those in rural areas. In the U.S., there are over 148,000 public water systems. Groundwater makes up over 90% of these systems, with the rest reliant on surface water. Groundwater in general is of higher quality than surface water with less toxic contaminants caused by human activities (Winter et al., 1998). However, human activities, such as the disposal of chemicals in soils or the direct injection of waste materials into groundwater can deteriorate groundwater quality. Once contaminated, it may take years to clean up the water system and the cleanup requires expensive treatment methods to meet drinking water quality standards (Harter, 2003).

Public water systems that are under state and U.S. Environmental Protection Agency (EPA) enforcement are required to report any health threatening water violations within 24 hours. It is important to note that testing and quality monitoring of water from a private well is the responsibility of the owner. Water quality standards are a great equalizer. If there is poor quality tap water, those who can afford it will simply buy bottled water, or invest in expensive filtration systems. A home water purification system could cost from $20 (filter for water pitcher or faucet mounted, requires frequent replacements) to more than $1000 (whole house water filtration system) (Stainley, 2021). Not all consumers can afford to make those purchases, making water quality an important issue. There also may be water quality disparities between rural and urban communities, among other factors.

EPA data

Many people are unaware that water quality violations are publically available. In Oklahoma, the Department of Environmental Quality (DEQ) is responsible for assuring the safety of Oklahoma water. If a violation with potential serious harmful health implication occurs, the information must be made public in 24 hours. Other violations should be made public in a local newspaper and on monthly water bills to customers until the violation is corrected. Information about additional tests that can be performed and how to complete such tests can be found through OSU extension PSS-2912 entitled Drinking Water Testing. In this current report, we have focused on violations for total coliform, treatment rule and nitrate, arsenic, lead and copper, and "other" violations as reported on the EPA website. "Other violations" included stage 1- disinfectants and disinfection byproducts by rule, and stage 2- disinfectants and disinfection byproducts rule, inorganic chemicals, volatile organic chemicals, synthetic organic chemicals and radionuclides. The time period we included was August 18, 2018 through November 11, 2020.

During the time period studied, most Oklahoma counties had a very low number of violations (Figure 1). Higher violation counts appear in the south central and north-eastern counties where the population density is higher as well as the number of public water systems. Pottawatomie and Cherokee counties had the highest violation count during this period. The USDA uses the rural-urban continuum code system to evaluate rurality at the county level. This scale ranges from 1 (counties in metro areas of 1 million population or more) to 9 (completely rural or less than 2,500 urban population, not adjacent to metro area) (USDA, 2020). Using correlation analysis we found a statistically significant negative relationship between the Oklahoma rurality codes and the number of EPA water quality violations present. This indicates that rural counties tend to have lower violation counts during the time period studied (Figure 2). One of the reasons rural counties have lower violation numbers may be that there are fewer public water systems in rural areas due to population needs. Additionally, many of the smaller public water systems purchase water from other larger public water systems. However, it is unclear if simply having more water systems within a county results in more violations. For example, one bad public water system could have more violations than 10 good ones.

Social media

Online content including social media posts, blogs, reviews etc. is a new frontier for researchers. In many cases, social and other online media can be used as a "canary in the coal mine." Information provided in social media posts have been found to be a good indicator and source of early warning for earthquakes, tsunamis, and other natural disasters (Sakaki, 2010; Chatfield, 2012). We hypothesized that by focusing on tap water related

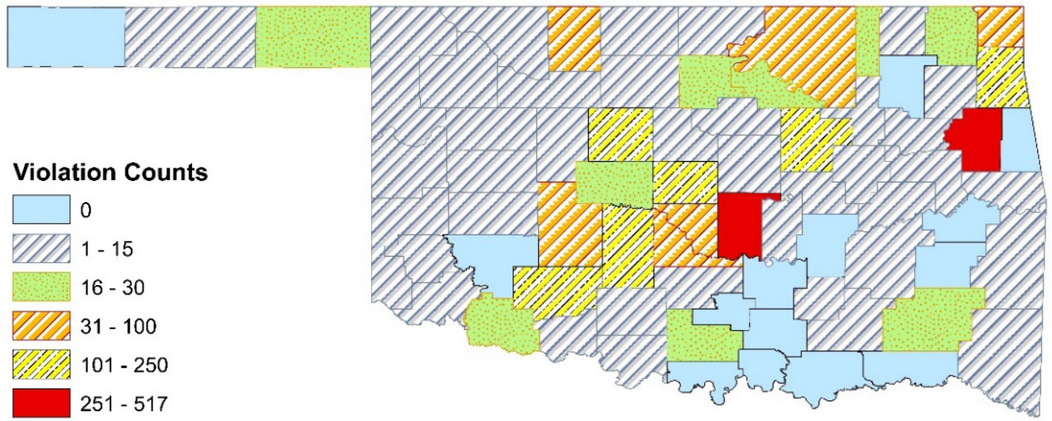

Figure 1. Map of Reported EPA violations by county.

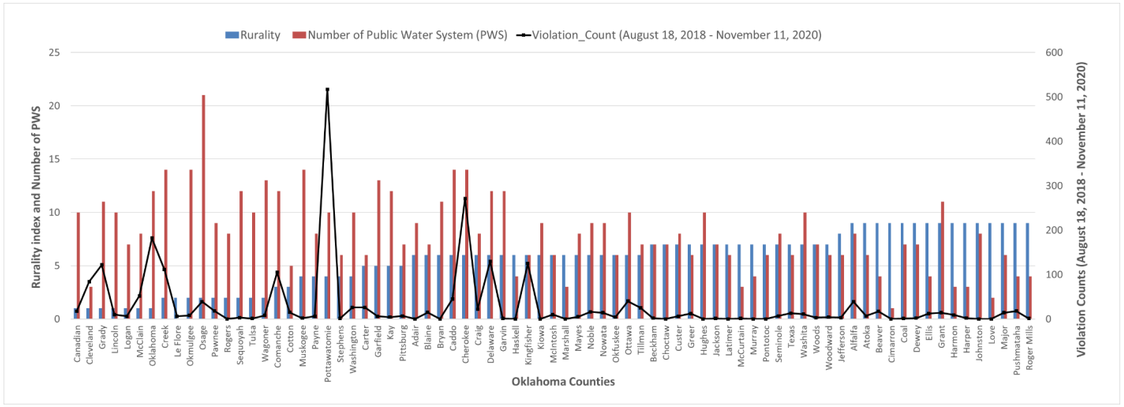

Figure 2. Rurality index, number of PWS, and violation counts by Oklahoma counties.

posts online, we could gain insight into people's perceptions of their water quality. We collected social media data from those in the state of Oklahoma, using the platform Netbase. Netbase is a social media analytics software that allows for the identification and linguistic analysis of online and social media content (Netbase, 2020). To gather social media posts related to tap water quality, a query including 13 terms was developed. Terms included: *tap water*, *piped water*, *tap-water*, *#tapwater*, *city water*, *#citywater*, *public water*, *#publicwater*, *water from the tap*, *#tap-water*, *faucet water*, *mains water* and *water from the faucet*. The same date range as the EPA data, August 18, 2018 through November 11, 2020 was studied.

Although social media is a powerful tool, there are many limitations. Despite choosing a rather large time period to study, the search resulted in only 3,628 posts that were related to tap water in Oklahoma. Demographic data is limited to only those who were using Twitter. The percentage of male and female twitter posters was roughly equal (Table 1). Looking at the age breakdown, given the lower number of older people who are active on social media, many older people were talking about tap water online. Top interests of posters included family, politics, and religion. Unsurprisingly, given the topic of tap water, food and drink also made the top five. A source gives a broad idea of where a post appeared, for example on a news cite. A domain is a more detailed example of where the post appeared, for example cnn.com. The majority of posts related to water were on Twitter. Interestingly, other domains included recreational vehicle (RV) websites, soil related websites, and a website that sells parts for vehicles and other machinery. Tap water and water quality would be an important factor for those looking for information on these additional domains.

The top emotions, terms, and hashtags that appeared in posts are available in Table 2. Most of the positive emotions included phrases like *good* or *best*. Negative emotions pointed towards issues with tap water including *warn*, *bad*, and *not like*. The top positive and negative term was the word *drink*. It is important to note that the words used around the term are analyzed to determine if a term is being referenced in a negative or positive context. Many of the emotions, terms, and hashtags were not particularly useful in monitoring water quality in a particular location. For positive terms, the next most frequently used term was *Mexican tap water*. Many people refer to their drinking water or other things as being "better than Mexican tap water." The term *controlavirus numbers* is nonsensical out of context. A controlavirus is a fictional term that references the mind control of a population. Many people were stating "tap water was better than the controlavirus." The use of tap water as a common colloquial item of comparison is an interesting trend. One of the most obvious indicators that the quality of tap water was being discussed to some extent online came from the top hashtags which included #arsenic.

Net sentiment is the number of positive posts minus the number of negative posts divided by the total number of posts multiplied by 100. Net sentiment must be between 100 and -100, with zero being neutral. For the entire state, the number of posts with sentiment indicators was 681, and the net sentiment was -1. This indicates that in general, the conversations surrounding tap water online were slightly negative. Looking at individual cities in Oklahoma (Table 3) Edmond had the highest net sentiment (30), followed by Oklahoma City (27) and Tulsa (25). All other cities (Norman, Broken Arrow, and Lawton) had net sentiment scores of 0 (neutral). The 6 cities in Table 3 were the only cities in Oklahoma that had enough social media data related to tap water to conduct analysis. The amount of data, and the net sentiment varied greatly over time (Figure 3). Much of the negative net sentiment around January 27, 2019 was related to people talking about the water issues in Flint Michigan. Around that time, the report outlining the crucial errors made by the Department of Environmental Quality's drinking water office that resulted in tragic water quality violations was released (Fouriezos, 2019). Another spike in mentions, and decrease in net sentiment occurred around November 3, 2019, when it was reported that tap water at the Trump golf course was contaminated with high levels of perflourooctanoic acid and other perfluorinated chemicals (Crowley, 2019).

The conversations online regarding tap water varied between national level news stories, and colloquial phrasing. Very little conversation regarding the quality of the posters' personal tap water appeared in the search. Social media can be used to evaluate trends related to the use of the word "tap water", and national issues, but there is little evidence that it can be used to point towards drinking water quality violations. The mundane nature of tap water may make it a less interesting topic for social media posting when compared to other issues such as

Table 1. Social media data demographics, domains and sources n=3,628.

Category	*Percentage of posts*
Gender	
Male	58
Female	42
Implied Age	
<18	9
18-24	12
25-34	16
35-44	16
45-54	16
55-64	20
65+	13
Interests	
Family	34
Politics	29
Religion	28
Food and Drink	13
Pets	11
Profession	
Creative Arts	61
Education	12
Technology	10
Entrepreneurship	7
Science and Technology	7
Domains	
Twitter.com	99
Instagram.com	<1
Irv2.com	<1
Forums.pelicanparts.com	<1
bobistheoilguy.com	<1
Sources	
Twitter	99
Forums	1
Instagram	<1
Consumer Reviews	<1

Table 2. Attributes, emotions, terms and hashtags from social media data.

	Positive	*Percent*	*Negative*	*Percent*
Top emotions n=160				
	Good	31	Warn	6
	Best	13	Probably	6
	Adequate	6	Bad6	
	Enjoy	6	Not like	6
	Dope	6	No best	6
Top terms n=1498				
	Drink	9	Drink	9
	Mexican tap water	6	Using	3
	Trust	5	Chemicals	3
	Controlavirus numbers	5	Contaminated	2
	Good	5	Free	1
Top hashtags n=33				
	NA		#drinkingwater	30
	NA		#arsenic	30
	NA		#powerwashing	3
	NA		#watersnob	3
	NA		#oddlysatisfying	3

NA indicates insufficient data for reporting.

Table 3.

City	*Net Sentiment*
Oklahoma City n=709	27
Tulsa n=542	25
Norman n=193	0
Broken Arrow n=51	0
Edmond n=130	30
Lawton n=10	0

Net sentiment is the number of positive posts minus the number of negative posts divided by 100. This number ranges from positive 100 to negative 100, with 0 being neutral.