Microbial Quality of Agricultural Water Used in Pre-harvest Production on the Eastern Shore of Virginia

Laura Nicole Truitt

Major Project/ Report submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Online Master of Agricultural and Life Sciences
In
Food Safety and Biosecurity

Laura K. Strawn - Food Science and Technology
Robert Williams
Steven Rideout

(Date of Submission – 04/30/2018)

Keywords: *Salmonella*, *Escherichia coli*, prevalence, concentration, diversity

Copyright (Optional)

Microbial Quality of Agricultural Water Used in Pre-harvest Production on the Eastern Shore of Virginia

Laura Nicole Truitt

Extended Abstract

Several produce-borne outbreaks have been associated with the use of contaminated water during pre-harvest applications. *Salmonella* has been implicated in a number of these outbreaks. The purpose of this study was to evaluate the microbial quality of agricultural surface water used in pre-harvest production on the Eastern Shore of Virginia in accordance to the Food Safety Modernization Act's Produce Safety Rule water standards. The study also examined the prevalence, concentration, and diversity of *Salmonella* in those water sources. Water samples (1 L) from 20 agricultural ponds were collected during the 2015 and 2016 growing seasons (n=400). Total aerobic bacteria, total coliforms, and generic *E. coli* were enumerated for each sample. Population levels of each microorganism were calculated per 100 mL sample and log transformed, when necessary. Samples (250 mL) were also enriched for *Salmonella*. Presumptive *Salmonella* isolates were confirmed by PCR of the *invA* gene and were serotyped. In 2016, the concentration of *Salmonella* in each sample was also estimated by MPN. Indicator bacteria, and environmental and meteorological factors were analyzed for their association with the detection of a *Salmonella*-positive water sample using logistic regression analysis. Seventeen of the 20 ponds met the FSMA PSR standards for production agricultural water. Three ponds did not meet the standards because the statistical threshold value exceeded the limit. *Salmonella* was detected in 19% of water samples in each year (38/200; 2015 and 38/200; 2016). Of the 118 *Salmonella* isolates serotyped, 14 serotypes were identified with the most prevalent being *S.* Newport. Generic *Escherichia coli* concentration, total aerobic bacteria concentration, and farm were significantly associated with the likelihood of detecting a *Salmonella*-positive sample. The

average concentration of *Salmonella* in all samples was 4.44 MPN/100 mL, with the limit of detection being 3 MPN/100 mL. The highest concentration of *Salmonella* was 93 MPN/100 mL. These data will assist in understanding the risks of production water poses to produce contamination events.

Table of Contents

Chapter 1: Introduction 5
1.1 Background and Setting 5
1.2 Purpose of the Project 7
1.3 Project Objectives 8
1.4 Definition of Terms 8
1.5 Significance of the Problem 9
Chapter 2: Review of Literature 9
2.1 Introduction 9
2.2 Indicator Organisms 10
2.3 The Food Safety Modernization Act Produce Safety Rule 12
2.4 Increasing Concern of Foodborne Illness Related to Agricultural Water 14
2.5 Taking Action to Reduce the Risk of Contamination 16
2.6 Summary 17
Chapter 3: Project Methodology 18
Chapter 4: Results 22
Chapter 5: Discussion 28
Acknowledgements 36
References 38
Appendices 45
A. Tables 45
B. Figures 52
C. Supplemental Material 55

Chapter 1: Introduction

1.1 Background and Setting

Non-typhoidal *Salmonella enterica* is a major pathogen of concern in agricultural environments (*4, 6, 7, 35, 43, 46*). *Salmonella* is the leading cause of bacterial foodborne illness in the United States (*44*) and the Centers for Disease Control and Prevention estimate it causes 1.2 million cases of foodborne illness and 450 deaths each year (*14*). Contaminated agricultural surface water used in pre-harvest production has been identified as the source of several produce-associated outbreaks in recent years (*17, 29*). Furthermore, *Salmonella* has been implicated in several multi-state outbreaks associated with produce, including the 2008 outbreak of *Salmonella* Saintpaul from serrano and jalapeno peppers (*15*), and the 2002, 2005, and 2006 *Salmonella* Newport outbreaks from tomatoes (*16*), where outbreak strains have been isolated from agricultural water (*29*). The multi-state outbreaks of *Salmonella* Newport associated with tomatoes have been traced to tomatoes grown on the Eastern Shore of Virginia (ESVA) (*5, 16, 23*).

Several studies (*6, 26, 28, 49, 50*) have investigated factors that influence the likelihood of contamination events in the pre-harvest environment; for example, wildlife intrusion, manure application, application of surface water to crops, and extreme weather events all increase *Salmonella* detection in fields. In fact, untreated surface water used in produce production environments is considered one of the most likely pathways for produce contamination events (*55*). Prior studies (*5, 22, 39, 52*) have examined the prevalence of *Salmonella* in surface water sources in different produce producing regions (e.g., California, Georgia, Florida, Virginia, Mid-Atlantic) and some (*4, 28, 36, 38, 58*) have determined the concentration of *Salmonella* in water samples; however, minimal data exists on prevalence, concentration, and diversity of *Salmonella* in surface water on the ESVA. Bell and colleagues (*5*) studied the prevalence and diversity of

Salmonella in tomato production areas of the ESVA. An overall prevalence of 8.4% was detected in creek and creek sediment samples; however, the concentration of *Salmonella* was not quantified in this study. Of the 234 *Salmonella* isolates that were serotyped, 21 different serotypes were identified with *Salmonella* serotype Newport being found most frequently (*5*).

The Food Safety Modernization Act Produce Safety Rule (FSMA PSR) was published in the Federal Register on November 27, 2015 (*55*). The PSR describes the science-based minimum standards for growing, harvesting, packing, and holding produce safely (*55*). The PSR is focused on reducing produce contamination events in the pre- and post-harvest environment by setting criteria for worker health and hygiene and agricultural water, among others (*55*). The standards for agricultural water are based on levels of generic *Escherichia coli*, as generic *E. coli* is traditionally used as an indicator of fecal contamination. Indicator organisms are typically non-pathogenic and are frequently found in higher concentrations than pathogens (*51*). The PSR criteria used to evaluate generic *E. coli* levels are geometric mean (GM) and statistical threshold value (STV). For untreated surface water that meets the definition of agricultural water (i.e., water that comes in direct contact with the harvestable portion of the produce during growing activities (*47*) used in production, the GM and STV must be less than or equal to 126 and 410 colony forming units (CFU) of *E. coli* per 100 mL of water, respectively. Initially, each untreated surface water source must have a microbial water quality profile (MWQP) established, which consists of 20 samples over the course of two-four years. These criteria are used to determine if an agricultural water source may be directly applied to the harvestable portion of produce. If the agricultural water source meets the standards, the water may be used during production. Once a water source achieves a MWQP within PSR water standards, each year a minimum of five water samples must be collected and results are combined with the most recent 15 samples (used to

calculate the baseline MQWP). Corrective measures may be implemented for water sources that do not meet the standards, such as a pre-harvest application interval (time between direct water application and harvest) or treatment (*55*). It is important to note the PSR does not specify testing for foodborne pathogens. Consequently, *Salmonella* and or other human pathogens may be present in agricultural water sources which meet the PSR agricultural production water standards. Data on the relationship between indicator organisms, and pathogen prevalence and concentration are needed to evaluate the risk of agricultural production water contamination events, and to develop appropriate mitigation strategies to reduce such events.

Statement of the Problem

Agricultural water can carry harmful pathogens that can contaminate produce. Untreated, surface agricultural water is considered the most susceptible to contamination (*55*). Little is known about the microbial quality of agricultural surface water on the Eastern Shore of Virginia. Several produce-borne outbreaks of salmonellosis have been linked to contaminated agricultural surface water (*15*, *16*, *29*). FSMA's PSR includes water testing requirements for agricultural surface water, which are based on the levels of generic *E. coli*, but they do not call for the testing of pathogens, like *Salmonella*. Measuring the generic *E. coli* levels gives a general understanding of the water quality in terms of possible fecal contamination, but it does not include pathogen prevalence or concentration. More information on water metrics for the Eastern Shore of Virginia is needed to properly assess public health risks. Growers are also lacking in science-based evaluations of the agricultural water used to irrigate crops.

1.2 Purpose of the Project

This study aims to provide data on the microbial quality of agricultural surface water used in pre-harvest production on the ESVA based on the FSMA PSR. It also includes data on

the prevalence, concentration, and distribution of *Salmonella* in these surface water sources that will provide information on the relationship between microbial indicators and pathogens so that mitigation strategies may be developed. Microbial, environmental, and meteorological factors were also analyzed to determine their effect on detecting *Salmonella*. This will also aid growers in determining when and how to use their water sources.

1.3 Project Objectives

The first objective of this study was to evaluate the microbial quality of agricultural water, specifically surface water used in production on the ESVA, based on currently published FSMA PSR water standards (*20*). The second objective of the study was to evaluate the prevalence, concentration, and diversity of *Salmonella* in the tested surface water samples. The third objective was to identify microbial, environmental, and meteorological factors that were associated with detection of a *Salmonella*-positive surface water sample.

1.4 Definition of Terms

Agricultural Water - water that is intended to, or likely to, contact the harvestable portion of covered produce or food-contact surfaces (*20*)

Geometric Mean – an average that represents the central tendency of generic *E. coli* levels in a water source (*55*)

Produce Safety Rule – science-based minimum standards for the safe growing, harvesting, packing, and holding of fruits and vegetables grown for human consumption (55)

Statistical Threshold Value – accounts for the variability of generic *E. coli* levels in a water source (55)

Salmonella – a gram negative, pathogenic bacteria comprised of 2 species, 6 subspecies, and over 2,500 serotypes that can cause gastrointestinal illness or typhoid fever (*56*)

1.5 Significance of the Problem

Findings from this study will provide crucial baseline data on total aerobic bacteria, coliforms, generic *E. coli* and *Salmonella* in agricultural pond water on the ESVA, as well as insight into the relationship between indicator bacteria and *Salmonella*, a human pathogen. Additionally, data will assist growers in development of mitigation strategies to reduce produce contamination events, such as minimizing direct application of pond water to the harvestable portion of crops.

Chapter 2: Review of Literature

2.1 Introduction

Agricultural water has been identified as a possible source of contamination for produce during growing, harvesting, and post-harvest activities (*20*). Because of this, agricultural water quality standards have been put into place to reduce the risk of foodborne illness from contaminated water sources. The FSMA PSR established key requirements for agricultural water that is applied to produce based on generic *E. coli* levels present in the water source (*20*). Generic *E. coli*, which is non-pathogenic, was chosen because it is considered an indicator organism for fecal contamination (*55*). Fecal contamination can lead to the presence of pathogens, so indicator organisms should give insight on the microbial quality of the water source (*40*).

An increasing number of foodborne illness outbreaks have occurred due to produce that was irrigated with contaminated water. A major pathogen of concern when dealing with agricultural surface water is *Salmonella. Salmonella* has been implicated in several foodborne illness outbreaks associated with produce that has been contaminated by agricultural water (*17, 29*). Produce including jalapeno peppers, serrano peppers, and tomatoes have been involved in

these very large outbreaks (*29*). Several of these outbreaks have been traced back to the Eastern Shore of Virginia (*5, 16, 23*). It is thought that wildlife, water usage, and weather events may affect the probability of detecting *Salmonella* in agricultural water sources (*29*). Because of these outbreaks, and the lack of data on water metrics for the Eastern Shore of Virginia, it is important to study the microbial populations in the agricultural water that is used to irrigate produce.

2.2 Indicator Organisms

Indicator organisms are organisms that are used to show the potential for the presence of pathogens. For this study, indicator organisms in water sources will be discussed. Indicator organisms for specific pathogens are generally chosen based on their abundance, the cost effectiveness of methods for enumeration, and the correlation between the presence of the indicator organism and the pathogen (*40*). Testing for indicator organisms is primarily done because it is too expensive and time consuming to test water samples for individual pathogens (*40*). Indicator organisms are typically non-pathogenic, usually present in pathogen-contaminated water, and are frequently found in higher concentrations than pathogens (*51*). An example of an indicator organism in agricultural surface water is fecal coliform bacteria. Fecal coliforms are a group of bacteria that belong to a larger group of bacteria called total coliform bacteria. Total coliforms are gram negative, rod-shaped, non-spore-forming, facultative anaerobic bacteria that are readily found in the environment. The difference between total coliforms and fecal coliforms is that fecal coliforms are commonly found in the feces of warm blooded animals. Generic *Escherichia coli*, which is a non-pathogenic subgroup of the fecal coliforms, is also considered an indicator organism. It too can be found in animal feces along with harmful enteric pathogens. When fecal coliforms are present in a water supply, it shows that there is fecal contamination and that pathogens may also be present (*40*).

Currently, generic *E. coli* is used as an indicator organism for the FSMA PSR water quality standards. Research is being conducted to determine if testing the levels of indicator organisms is sufficient to determine the quality of water. The results on the effectiveness of indicator organisms on predicting pathogen populations have been conflicting. They continue to be used as predictors because of the cost effectiveness of their analysis. Several studies have shown that the presence of certain pathogens, like *Salmonella*, have been significantly associated with indicator organism populations (*34, 38*). A study conducted in North Carolina to compare the presence of indicator organisms and *Salmonella* showed that the presence of *Salmonella* was significantly correlated with the presence of 5 different indicator organisms. The indicator organisms included fecal coliforms, generic *E. coli*, enterococci, somatic coliphage, and *C. perfringens* (*34*). Another study conducted in Florida to predict *Salmonella* populations based on indicator organisms showed that there was a significant positive relationship between *E. coli* and *Salmonella*, but the correlation was not strong enough to be able to predict *Salmonella* prevalence based on *E. coli* levels. Because of the positive relationship, it was determined that until a more promising indicator organism is discovered, *E. coli* may be used to estimate the likelihood of detecting *Salmonella* (*38*).

Other studies have shown that indicator organisms may not be reliable for predicting pathogen prevalence. An example of this is when water which meets standards that are based on indicator organism levels still contains pathogens. The same study conducted in North Carolina that was previously mentioned showed that of 25 water samples which met the Environmental Protection Agency's water quality standards of 235 CFU *E. coli* per 100 mL of water sample, 14 of them were *Salmonella*-positive. This shows that even when water quality standards based on indicator organisms are met, there may still be a risk to public health (*34*). A study conducted in

Australia also showed that indicator organisms may be inadequate for predicting the prevalence of pathogens in waterways. Water samples were collected on a bi-weekly basis from 3 estuarine locations for a year to determine the effect of environmental factors on pathogen and fecal indicator populations. *E. coli* was detected in 100% of samples, however *Salmonella* was not detected for the first three months of sampling, so it was removed from the study. This data provides evidence that indicator organisms may not always be consistent for predicting pathogen populations (*32*).

2.3 The Food Safety Modernization Act Produce Safety Rule

The Food Safety Modernization Act (FSMA) was signed into law on January 4, 2011 by President Obama. The purpose of this act is to ensure the safety of the food supply in the United States by implementing prevention-based methods (*20*). The Produce Safety Rule (PSR), which is a requirement under FSMA, was first proposed in January 2013. Its purpose is to provide science-based minimum standards for the safe growing, harvesting, packing, and holding of fruits and vegetables that are grown for human consumption. Revisions were made to make the PSR more practical in 2014, and the PSR was made final November 27, 2015. The effective date for the PSR was January 26, 2016. The PSR does not apply to produce that is rarely eaten raw, that is grown for personal consumption, or is not considered a raw agricultural commodity. Compliance dates for the PSR vary depending on the size of the operation (*20*).

The PSR contains 6 key requirements which include agricultural water, biological soil amendments, sprouts, domesticated and wild animals, worker training and health and hygiene, and equipment, tools, and buildings (*55*). The focus for this study is the requirement for untreated agricultural surface water, specifically the surface water that is applied directly to produce. The PSR agricultural water quality standards are based on the levels of generic *E. coli*

which is used as an indicator organism. The levels of generic *E. coli* must meet two criteria. The first is the geometric mean (GM), or central tendency, of the generic *E. coli* must be less than or equal to 126 CFU/100mL of water. The second is the statistical threshold value (STV), or the variability, must be less than or equal to 410 CFU/100mL. The FDA is currently working on an online tool for growers to use to calculate these values for their water sources (*55*).

To determine if an untreated agricultural surface water source meets the PSR standards, a minimum of 20 samples must be taken over the course of 2 to 4 years. After the 20 samples are taken, an initial microbial water quality profile may be made by calculating the GM and STV for generic *E. coli* to determine if the water source meets the PSR standards. After the initial microbial water quality profile is established, 5 water samples must be taken each year, and be combined with the most recent 15 samples, to calculate a new GM and STV and create a rolling microbial water quality profile. After the microbial water quality profile has been established, recommendations may be made to growers regarding the use of their agricultural water sources. If the water quality standards are met, the water source can continue to be used. If the agricultural water source does not meet the standards, corrective measures such as a pre-harvest interval, a postharvest storage interval, or treating the water may be put into place so the water source may continue to be used (*55*).

There is an ongoing debate on whether these standards are sufficient. A study on wash water used for lettuce showed that water which contained higher than 2 log *E. coli*/100mL of water should not be used because 42% of samples exceeding this level contained pathogens. Only 10 percent of samples below the 2 log *E. coli*/100mL level contained pathogens (*33*). This study shows that generic *E. coli* is a good indicator of contamination, but other research has shown contradictory data. Another study done on rainwater samples in Australia showed that

12% of rainwater samples contained one or more pathogens but had less than 1 CFU *E. coli*/100mL of water (*1*). Many other types of indicator organisms such as enterococci, *Bacteroides* spp., *Bifidobacterium* spp., *Clostridium perfringens*, and bacteriophages have been suggested for determining water quality, but no sufficient data exists on these possibilities (*57*).

2.4 Increasing Concern of Foodborne Illness Related to Agricultural Water

An increasing demand for fresh produce has developed based on current health recommendations. From 1970 to 2005, consumption of fresh fruits and vegetables increased by 19 percent (*59*). The United States Department of Agriculture dietary guidelines recommend eating at least 2.5 cups of fruits and vegetables per day for a healthy diet (*53*). With more fruits and vegetables being consumed, this has led to an increase in foodborne illness outbreaks associated with fresh produce (*12*). Because of the increased demand for fresh produce, there is also an increasing demand for irrigation water. Irrigation water, especially when it comes from an above ground, untreated source, can carry harmful pathogens which contaminate produce and cause foodborne illness (*20*).

Salmonella has been implicated in a number of foodborne illness outbreaks caused by contaminated water. Two major multistate outbreaks of *Salmonella enterica* serotype Newport occurred in 2002 and 2005. The 2005 outbreak affected people in 16 states. Seventy-two isolates of *S*. Newport were found to be indistinguishable by Pulsed-Field Gel Electrophoresis (PFGE). Traceback investigation determined the outbreak was caused by contaminated tomatoes that were grown on the Eastern Shore of Virginia. The outbreak strain of *S*. Newport was later identified in a pond on the Eastern Shore of Virginia that was used to irrigate tomato fields. The pond water was applied to the fields using drip irrigation, so it was only in contact with the soil beneath the tomato plants, but it was later discovered that a grower used the pond water for

pesticide applications. Further traceback and epidemiological investigation showed commonalities between this outbreak and a previous *S*. Newport outbreak which occurred in 2002. The 2002 outbreak affected 510 people in 26 different states and was caused by *Salmonella* with the same serotype, PFGE pattern, geographical distribution of patients, association with tomatoes, and isolation from pond water as the 2005 outbreak. These recurrent outbreaks of the same strain of *S*. Newport suggest the Eastern Shore of Virginia may be a source of environmental contamination (*23*). Further research needs to be conducted on the concentration of *Salmonella* on the Eastern Shore of Virginia because previous studies focused mainly on prevalence (*5*), so the actual risk to public health has not been fully assessed.

Other major outbreaks of *Salmonella* caused by contaminated irrigation water have been associated with produce like sprouts, serrano peppers, and cantaloupes. Three outbreaks of *Salmonella* Enteritidis associated with sprouts occurred between 2000 and 2001 because the water used to irrigate sprouts was contaminated with *Salmonella*. It was later found that FDA guidelines for disinfecting irrigation water were not being followed (*29*). In 2008, one of the largest outbreaks of salmonellosis due to fresh produce in the United States occurred. It was caused by *Salmonella* Saintpaul and was found to be associated with serrano and jalapeno peppers grown on a farm in Mexico. The source of contamination was found to be irrigation water used to water the serrano and jalapeno pepper crops. Over 1000 people in 42 different states were sickened during this outbreak. Preliminary investigations showed the source of the outbreak may have been tomatoes, especially because of their association with previous outbreaks. When the final outbreak investigation results showed that serrano and jalapeno peppers were the source of the outbreak, it demonstrated that many different types of produce may be affected by *Salmonella* (*29*). Multistate outbreaks of *Salmonella* Poona associated with

cantaloupes occurred during 2000-2002. The outbreak strains were linked to cantaloupes that were imported from Mexico. FDA investigation of the outbreak identified irrigation water contaminated with sewage as a possible source of contamination (*17*).

2.5 Taking Action to Reduce the Risk of Contamination

Several corrective measures, or management options, may be put in place to reduce the risk of foodborne illness from contaminated irrigation water. If an agricultural water source does not meet the FSMA PSR criteria for the GM and STV of generic *E. coli* levels, corrective measures may be taken so the water source can continue to be used. One of these corrective measures involves implementing a time interval to allow for die-off. A time interval in the field between the last irrigation and harvest may be used to achieve appropriate die-off. A storage interval after harvest may also be used. A second corrective measure that may be used is commercial washing of produce if it achieves log reduction. Water sources may also be inspected for sources of contamination and corrective measures may be implemented to address those issues. The final corrective measure that may be used to continue using a water source that does not meet the PSR standards is to treat the water as long as any disinfectants or sanitizers are used in accordance to their labels (*47*). Treating water sources may sometimes be difficult because treatment options may not be practical or could have high costs. Treatment options other than sanitizers include filtration, ozonation, exposure to ultraviolet light, electric beam processing, and heat treatment. Because these treatment options can often not be used, prevention of contamination of irrigation water is key (*46*).

Several options are available to help prevent contamination of produce by irrigation water. These include following good agricultural practices (GAPs) and using a method of irrigation called drip irrigation. Following GAPs can help to control contamination from point

sources. It is very difficult to control contamination from nonpoint sources like wildlife and birds, so it is important to try to prevent contamination where possible. Point sources of contamination that can be controlled through GAPs include manure used as fertilizer and runoff from feedlots. Methods to control these are to keep livestock away from sources of irrigation, to look for upstream uses of irrigation water, and to ensure that manure does not run off into an irrigation water source (*46*). It is important to work with local watershed committees to learn more about what may be affecting water sources used for irrigation (*42*). The type of irrigation that is used can also help to prevent contamination. Drip irrigation can minimize contamination of produce by irrigation water because the edible portions of plants are not in direct contact with the water. The water is applied at the surface of the soil, so the portion of the plant above the ground does not get wet. Overhead irrigation wets the entire plant, therefore increasing the probability of contamination of the produce (*46*).

2.6 Summary

The increasing concern of agricultural surface water as a source of contamination of produce has caused a need for more research on water sources that are used for irrigation. It is important to understand the microbial quality of these water sources to prevent contamination of produce, which can lead to foodborne illness outbreaks. *Salmonella*, a harmful human pathogen, has been implicated in many of these outbreaks. By following the FSMA PSR key requirements for agricultural water and using GAPs, the incidence of foodborne illness outbreaks associated with agricultural water should be greatly reduced. Prevention-based methods are now replacing reaction-based methods to try to reduce the incidence of foodborne illness from contaminated irrigation water.

Chapter 3: Project Methodology

Study Design. The study design was to sample agricultural water used for production purposes listed in the PSR (e.g. irrigation, crop protective sprays) from produce farms on the ESVA during the growing season. Farms were enrolled based on willingness to participate, and agreement that farms would be kept confidential. Six farms enrolled in the study that were located in one of the two counties on the ESVA: Accomack County (3) and Northampton County (3). To capture diversity of agricultural water sources within farm, up to five water sources were selected as sampling sites. In total, 20 untreated, surface water ponds were sampled on a bi-weekly basis during the 2015 and 2016 growing seasons (10 sampling visits each year from May through September). Characteristics of the ponds (e.g. creek or well fed, size, vegetation, and sunlight exposure, etc.) are listed in Table 1.

Water Sampling. Over the course of the two growing seasons, 400 water samples were collected (n=200; 2015 and n=200; 2016). Water samples were collected in 1 L polypropylene wide mouth sterile bottles (Thomas Scientific, Swedesboro, NJ) using a Swing Sampler sampling pole (Nasco, Fort Atkinson, WI). The sampling pole was used to lower the jars approximately 0.5 m below the surface of the water. Water samples were collected approximately 2-3 m from the shore near the intake pump. Sample visits were completed before 11 am EST (to ensure samples were collected at approximately the same time each visit and reduce the impact of sunlight). Samples were transported to the laboratory on ice and processed within 3 h of collection.

Total Aerobic Plate Count. Surface water samples were serially diluted in 0.1% peptone water (Fisher Scientific, Fair Lawn, NJ), and 100 uL aliquots were plated in duplicate on tryptic soy

agar (TSA; Hardy Diagnostics, Santa Maria, CA). TSA plates were incubated for 24 ± 2 h at 35 ± 2°C. All bacterial colonies were enumerated by hand and CFU/100 mL calculated.

Coliform bacteria and generic *E. coli* Most Probable Number (MPN). The Colilert-18/Quanti-Tray/2000 (IDEXX Laboratories, Inc., Westbrook, ME) method, which has been approved by the FDA for FSMA PSR water testing, was used to estimate the most probable number (MPN) of total coliform bacteria and *Escherichia coli* per 100 mL of each water sample. Quanti-Trays were incubated for 24 ± 2 h at 35 ± 2°C. Yellow color wells indicated the detection of coliform bacteria, and fluorescent wells under ultraviolet light indicated the detection of *E. coli*. Coliform bacteria and generic *E. coli* MPN/100 mL were estimated for each water sample using IDEXX estimation tables.

Microbial Water Quality Profile (MWQP). *E. coli* MPN data was used to create a MWQP for each of the 20 ponds. An Excel tool (version 4.0 http://ucfoodsafety.ucdavis.edu/files/268306.xlsx) developed by the Western Center for Produce Safety, University of California, Davis (*60*) was used to calculate the GM and STV using standard approaches. Results below the limit of detection (LoD) were replaced by half the lower limit (0.5 MPN/100 mL), and results above the LoD by the upper limit (1011.2 MPN/100 mL).

***Salmonella* Detection, Isolation, and Serotyping**. *Salmonella* detection and isolation were performed using a modified version of the Food and Drug Administration's Bacteriological Analytical Manual (FDA BAM) (*38, 54*). Briefly, surface water samples (250 ml) were pre-enriched in double strength lactose broth (Hardy Diagnostics, Santa Maria, CA) and incubated for 24 ± 2 h at 35 ± 2°C. A 1 and 0.1 mL aliquot of each sample was then transferred to Tetrathionate (TT) (HiMedia, Kennett Square, PA) and Rappaport-Vassiliadis (RV) (Becton,

Dickinson and Company, Sparks, MD) selective enrichment broths, respectively (*49*). TT and RV broths were incubated for 24 ± 2 h and 48 ± 2 h at 35 ± 2 h and 42 ± 2°C, respectively. Fifty microliters of each TT and RV broths were streaked onto Xylose Lysine Deoxycholate (XLD) (Hardy Diagnostics, Santa Maria, CA) and CHROMagar *Salmonella* Plus (Paris, France) agars. Both XLD and CHROMagar *Salmonella* agar plates were incubated for 24 ± 2 h at 35 ± 2°C (*49*). Presumptive *Salmonella* colonies were confirmed by PCR amplification of the *invA* gene (a 284 bp amplicon), as previously described by McEgan et al. and Strawn et al. (*38, 49*). The *invA* primers were GAATCCTCAGTTTTTCAACGTTTC and TAGCCGTAACAACCAATACAAATG (Integrated DNA Technologies, Coralville, IA). PCR confirmed *Salmonella* isolates (up to 4 per sample) were stored in a -80°C freezer. One *Salmonella* isolate per sample was sent to the National Veterinary Services Laboratory in Ames, Iowa for serotyping (traditional serology using the Kaufman-White-LeMinor scheme; *24*).

***Salmonella* MPN**. Methods for estimating the most probable number (MPN) of *Salmonella* in each water sample were based on those described by McEgan and colleagues (*38*). Briefly, a three by three tube MPN was used with the following dilutions: 10 mL in 10 mL double strength lactose broth, 1 mL in 9 mL of single strength lactose broth, and 0.1 mL in 9 mL single strength lactose broth. Tubes were incubated for 24 ± 2 h at 35 ± 2°C. A 1 and 0.1 mL aliquot of each sample was transferred to TT and RV selective enrichment broths. TT and RV were incubated for 24 ± 2 h and 48 ± 2 h at 35 ± 2 h at 42 ± 2°C, respectively. Ten microliters of each TT and RV broths were streaked onto XLD and CHROMagar *Salmonella* Plus agars which were incubated for 24 h at 35°C. Presumptive *Salmonella* colonies were confirmed by PCR amplification of the *invA* gene as previously described above. MPN values were estimated using a modified table from the FDA BAM to match the volumes of water used in this study (*54*).

Environmental and Meteorological Data Collection. Proximity data was assessed using Google Maps (https://www.google.com/maps) to determine how far tested water sources were from roads and packing sheds, and to determine what, if any, other water sources fed into the tested source. Environmental and meteorological factors were observed for each site including precipitation, temperature, and reptile presence (S1). Rainfall and temperature data were collected for the day of sampling and for 7 days prior to sampling from the National Oceanic and Atmospheric Administration's National Weather Service database. The average and total precipitation for the week prior to sampling, and average temperature were also recorded. Data was recorded from the weather recording station nearest to the water sampling sites (e.g., farm weather station). Evidence of reptile presence was recorded when observed and was compared to sampling sites that tested positive for *Salmonella*.

Statistical Analysis. All statistical analyses were performed in **R** version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria). Data in Excel sheets were imported in **R** using the *XLConnect* package. *E. coli* concentration values (MPN/100 mL) were analyzed by calculating the 50th and 90th percentile of a lognormal distribution fit to the *E. coli* count data (GM and STV, respectively). *E. coli* concentrations were both right and left censored, with a lower limit of detection of 1 MPN/100 mL and an upper limit of detection of 1011.2 MPN/100 mL. The PSR specifies that the GM and STV must be calculated using the method of moments (MM), by calculating the mean and standard deviation of log-transformed counts. These values were calculated for each site, with all transformations based on natural (base *e*) logarithms. Truncated lognormal and several alternative distributions were also fit to the data using maximum likelihood estimation (MLE) with the **R** package *fitdistrplus*. Alternative distributions considered were the Weibull, generalized Pareto, and Fréchet distributions (the latter two are available in the

actuar **R** package). Akaike Information Criterion (AIC) and visual inspection of plots of the fitted and empirical cumulative distribution function were used to compare the fit of each distribution to empirical data from each site. To evaluate the effects of censoring, the 50th and 90th percentile of a lognormal distribution fit to the *E. coli* count data by the method of moments (MM) and that fit by maximum likelihood estimation considering censoring, were compared.

Univariate logistic regressions were also performed for each potential predictor variable for *Salmonella*, including sampling site location, rainfall (0d – 7d prior, average, and total for 3d and 7d prior), temperature (0d – 7d prior and average for 7 days), reptile presence, intake pump activity, total aerobic bacteria, coliforms, generic *E. coli*, farm, and county. Significant variables ($P < 0.2$) were combined into a multivariate model and analyzed. A P value of ≤ 0.05 was used to determine significance. Confidence intervals (95%) were calculated assuming a binomial distribution. The diversity of *Salmonella* serotypes within sampling season, county, farm, and pond were quantified using Simpson's index of diversity (*D*) (*45*).

Chapter 4: Results

A total of 400 water samples were collected and analyzed from May 2015 to September 2016 from twenty agricultural water sources (ponds). The concentration of *Escherichia coli* (MPN/100 mL), total coliforms (MPN/100 mL), and total aerobic bacteria (log CFU/100 mL), along with prevalence of *Salmonella*, were determined for all 400 samples (Table 2). MPN for *Salmonella* (MPN/100 mL) was only performed in the second year of sampling (n=200) (Table 2).

E. coli. *E. coli* concentrations were estimated using a scientifically verified fluorogenic MPN method (Quanti-Tray), so results are in MPN/100 mL. The PSR regulation requires GM and STV values to be calculated for generic *E. coli* levels in CFU/100 mL by enumerating colonies on an

agar plate. Multiple sources show the difference between CFU and MPN methods are not significant (*10, 11, 19, 21, 25, 30*). Other studies also show that fluorogenic *E. coli* detection methods are better for turbid agricultural water sources because they avoid having clogged filters (*9*). The median *E. coli* concentration was 15.8 MPN/100 ml with a lower quartile of 4.10 MPN/100 mL and an upper quartile of 53.7 MPN/100 mL. The mean *E. coli* concentration was 85.4 MPN/100 mL, but the most frequently measured *E. coli* concentration, or mode, was only 1.00 MPN/100 mL. Only a small number of MPN values for *E. coli* (7/400) reached the upper limit of detection (> 1011.2 MPN/100 mL). On the other hand, there were several MPN values for *E. coli* (29/400) which fell below the limit of detection (< 1.00 MPN/100 mL). While, total aerobic bacteria and total coliform concentrations were similar for the two sampling years, the median *E. coli* concentration in 2016 (22.15 MPN/100 mL) was much higher than the median for 2015 (11.5 MPN/100 mL). Median *E. coli* concentrations also varied greatly between farms (9.10 – 32.1 MPN/100 mL). Farm A had the lowest median concentration of *E. coli* at 9.10 MPN/100 mL and Farm E had the highest median concentration of *E. coli* at 32.1 MPN/100 mL. Pearson correlation coefficients were calculated for each environmental and meteorological factor versus *E. coli* concentration (MPN/100 mL), and all correlation coefficients fell below 0.28, indicating a lack of correlation. Additionally, *E. coli*, total coliforms, and total aerobic bacteria were not significantly correlated ($r < 0.25$).

Microbial Water Quality Profile. Seventeen of the 20 ponds would be in compliance based on the FSMA PSR agricultural water standards (using the MWQP). All 20 ponds met the PSR criteria for the GM (S5), as GM values ranged from 5 to 62 MPN/100 mL for all ponds calculated by standard approaches (i.e. including replacement of censored data by the upper LoD (1011.2 MPN/100 mL) or half the lower LoD (0.5 MPN/100 mL). Three ponds did not meet the

PSR criteria for the STV ≤ 410 CFU *E. coli* /100 mL (S6). STV values ranged from 23 to 384 MPN/100 mL, except ponds 173, 284, and 228. The STV values for ponds 173, 284, and 228 which were 1120, 670, and 776 MPN/100 mL, respectively. All 3 of the ponds that exceeded the STV upper limit had an *E. coli* MPN value that exceeded the limit of detection (1011.2 MPN/100 mL) for at least one sampling. In September 2015 and June 2016, pond 173 had *E. coli* and coliform MPN values that exceeded the limit of detection. The aerobic plate counts for those months were also 7.4 log CFU/mL or above for this pond. Interestingly, on both sampling dates in September 2015 and June 2016, over 76 mm of rain was recorded on the 7 d prior to sampling. Pond 284 had an *E. coli* and coliform MPN value at the limit of detection with an aerobic plate count value of 8.1 log CFU/100 mL in August 2016, but only 15.24 mm of rain was recorded for the week prior to sampling. Pond 228 had 2 samplings in August 2015 and 2016 that were near or at the limit of detection for *E. coli* and coliform bacteria, and a total aerobic plate count over 6.5 log CFU/100 mL. For the week prior to the August 2015 sampling, 36.83 mm of rain was recorded, while in the week prior to the August 2016 sampling, only 10.16 mm of rain was recorded. All three of the ponds that exceeded the PSR STV requirement tested positive for *Salmonella* on the sampling dates with high *E. coli* concentrations. All three of these ponds were located next to the packing shed. Water from ponds 173 and 284 was often turbid (high sediment load), and pond 228 had heavy vegetation near intake pump/bank.

Furthermore, most GM values were greater when calculated using the method of moments with the remaining being equal to estimates using maximum likelihood estimation (S4). This suggests a trend of overestimating the GM when using the method prescribed by the PSR; therefore, ponds may be falsely out of compliance under the currently calculated FSMA PSR GM metric. On the other hand, the majority of STV estimates were found to be greater

when calculated with maximum likelihood estimation (S4), suggesting underestimation by the PSR methods. An underestimation of the STV may allow ponds to be in compliance with the FSMA PSR when they may actually be out of compliance.

Skewness and kurtosis of log-transformed *E. coli* concentration values for each site (measures of distribution symmetry and peakedness, respectively) were calculated (S2) to further inform FSMA PSR agricultural water standards. Sites in this study had skewness values between -0.646 and 0.886 and kurtosis values between 1.725 and 3.714. Skewness values centered around 0 suggest the log-transformed distributions are symmetrical. Kurtosis values centered below 3 suggest the log-transformed distributions are, on average, less peaked (more dispersed) than a normal distribution. AIC values can be found in S3 for each distribution fit to the data using maximum likelihood estimation. From this comparison, the lognormal distribution appears to be the best fit across all sites, with the lowest AIC for 9/20 sites and the second lowest AIC for 7/20 sites.

Total Aerobic Bacteria. The median total aerobic bacteria concentration was 5.2 log CFU/100 mL with a lower quartile of 4.9 log CFU/100 mL and an upper quartile of 6.0 log CFU/100 mL. The mean concentration of total aerobic bacteria was 6.8 log CFU/100 mL. The highest total aerobic bacterial counts were 8.5 and 8.4 log CFU/100 mL, and these both occurred in July of 2015 and 2016. The median concentration of total aerobic bacteria was similar for each year of sampling: 5.0 log CFU/100 mL in 2015 and 5.5 log CFU/100 mL in 2016. The median total aerobic bacterial count for each farm ranged from 5.0 to 5.4 log CFU/100 mL. Pond 199 had the lowest total aerobic bacteria concentration at 5.0 log CFU/100 mL. Pond 279 had the highest total aerobic bacteria concentration at 7.6 log CFU/100 mL. Pearson correlation coefficients were calculated for each environmental and meteorological factor versus total aerobic bacteria

concentration (CFU/100 mL), and all correlation coefficient values fell below 0.12, indicating a lack of correlation.

Total Coliforms. Most of the MPN values for total coliforms (346/400) reached the upper limit of detection (>1011.2 MPN/100 mL). Only one MPN value for coliforms fell below the limit of detection (<1 MPN/100 mL). The median coliform MPN value was 1011.2 MPN/100 mL with both the upper and lower quartiles being 1011.2 MPN/100 mL. The average coliform concentration was 985 MPN/100 mL. The median total coliform concentration was the same for both years of sampling (2015: 1011.2 MPN/100 mL; 2016: 1011.2 MPN/100 mL). Median coliform counts for each farm were also the same for each farm (1011.2 MPN/100 mL). Pearson correlation coefficients were calculated for each environmental and meteorological factor versus total coliform concentration (MPN/100 mL), and all correlation coefficient values fell below 0.20, indicating a lack of correlation.

***Salmonella* Prevalence.** The prevalence of *Salmonella* in each sampling year (2015 and 2016) was 19% (76/400 combined; 38/200 in 2015; 38/200 in 2016). Sampling year (2015 and 2016) and pond (20) had no significant association with the detection of a *Salmonella*-positive water sample in this study. After performing univariate logistic regression for each predictor variable (i.e., factors), *E. coli* concentration, total aerobic bacteria count, precipitation on day of sampling, precipitation on the day prior to sampling, total precipitation for the 3 days prior to sampling, precipitation on the 7th day prior to sampling, temperature on the 4th day prior to sampling, and total coliforms were found to have a significant effect on the probability of detecting a *Salmonella*-positive water sample ($P < 0.2$). Additionally, only two spatial variables, county and farm, were significantly associated with the detection of a *Salmonella*-positive water sample (Table 3). For instance, the prevalence of *Salmonella* was significantly higher in County

2 (26%), compared to County 1 (13%); approximately two-fold difference (Table 3). Farms C and E showed a significantly higher prevalence of *Salmonella* than Farm A (Figure 2). Interestingly, Farms C and E also had large quantities of produce culls located near some of the ponds (< 5 m away), which attracted wildlife, such as birds or rodents. Produce culls were not observed at the other farms. When the predictor variables were combined into a multivariate model, *E. coli* count, total aerobic bacteria count, and farm provided the best model fit for predicting *Salmonella* in a water sample (Figure 2). Interestingly, the significance of all other variables were decreased, and the model was not improved from the univariate models.

***Salmonella* Concentration**. The limit of detection for the MPN method used was < 3.00 MPN/100 mL (< 0.50 log). Of the 200 samples tested in 2016, 158 samples (approximately 79%) were below the limit of detection. The concentration of *Salmonella* ranged from <3 .00 MPN/100 mL (< 0.50 log) to 93.0 MPN/100 mL (1.97 log) with an average of 4.40 MPN/100 mL (0.65 log). The median concentration was 3.00 MPN/100 mL. Pearson correlations for *Salmonella* concentration (MPN/100 mL) and all precipitation and temperature variables (e.g., on day of sampling, seven days prior to sampling) were run, and there were no significant correlations (r < 0.20).

***Salmonella* Diversity**. Serotyping was performed on one representative isolate for each *Salmonella*-positive water sample. Of the 118 *Salmonella* isolates, 13 serotypes were identified in *Salmonella enterica* subspecies *enterica* including Javiana (n=17), Thompson (n=12), Newport (n=42), Infantis (n=4), Hartford (n=4), Typhimurium (n=7), Norwich (n=16), Saintpaul (n=5), Bareilly (n=2), Anatum (n=1), 9,12:I,z28:- (n=1), Berta (n=4), and Senftenberg (n=1). Interestingly, only 1 serotype was identified in *Salmonella enterica* subsp. *arizonae*: III_38:(k):z35 (n=2). *Salmonella* serotype Newport was isolated most frequently, consisting of

approximately 36% of the *Salmonella* isolates from water samples for 2015 and 2016. In 2015, 50% of the *Salmonella* isolates from water samples represented serotype Newport (19/38). In 2016, the proportion of *Salmonella* serotype Newport isolates dropped to 29% (Figure 1). Overall, there was a high diversity of *Salmonella* serotypes detected over the two-year study ($D = 0.82 \pm 0.05$). Both Counties, 1 and 2, also had a high diversity of *Salmonella* serotypes ($D = 0.82 \pm 0.07$ and 0.81 ± 0.06, respectively). Similarly, all six farms sampled had a relatively high diversity of *Salmonella* serotypes (high range: Farm F with $D = 0.89 \pm 0.13$; low range: Farm C with $D = 0.68 \pm 0.17$). Serotype diversity within ponds (each pond was sampled twenty times over two years) was also relatively high (average: $D = 0.74 \pm 0.17$, high range: $D = 1.00 \pm 0.00$, low range: $D = 0.62 \pm 0.17$). Lastly, 10 serotypes were isolated repeatedly over the course of the study from the same water source (Table 4); in fact, 15 water sources had instances of *Salmonella* repeat isolation.

Chapter 5: Discussion

The purpose of this study was to evaluate the microbial quality of untreated agricultural surface water used in pre-harvest production according to the FSMA PSR standards. Additionally, this study aimed to investigate *Salmonella* prevalence, concentration, and diversity in the water sources, and how the detection of *Salmonella* was associated with indicator organisms, and spatio-temporal factors.

Most ponds met the FSMA PSR agricultural water standards. In the study reported here, a high percentage of the agricultural ponds met the FSMA PSR water standards (85%) for untreated surface water sources, which are based on generic *E. coli* levels in 100 mL of water for each source. Other studies conducted in the Mid-Atlantic region did not find the same high rate of compliance in agricultural water. For example, Allard and others found that only 2 of 12 water

sources (approximately 17%), including surface water sources (ponds and rivers) and nontraditional sources (reclaimed wastewater), were compliant using the FSMA PSR water standards (*2*). In the Allard study (*2*), all 10 water sources that were out of compliance had a STV value which exceeded the PSR limit of less than or equal to 410 CFU *E. coli*/100 mL. Additionally, in the Allard study, seven of the 12 water sources exceeded the limit for GM (≤126 CFU *E. coli*/100 mL). In the study described here, all 20 ponds met the PSR limits for GM, and three of the 20 ponds exceeded the PSR limits for STV. Other studies (*36, 52*) have evaluated untreated agricultural surface water in produce-growing regions beyond the Mid-Atlantic, such as Florida and Georgia. Topalcengiz and colleagues sampled six agricultural ponds in Central Florida and evaluated them according to the FSMA PSR standards over a two-year period (*52*). All ponds tested met the FSMA PSR standards. Similarly, another study in Southern Georgia surveyed 10 agricultural ponds that all met the FSMA PSR standards (*36*). These findings suggest that water quality (using concentrations of generic *E. coli*) is difficult to predict at a regional level (e.g., Mid-Atlantic) even if similar practices are used. Instead, microbial water quality is dependent on farm localized specific specific spatio-temporal factors (e.g., wildlife pressures, or rainfall) and water source management (by the farm or local EPA).

Untreated surface water sources that meet PSR standards may contain pathogens upon enrichment; however, generic *E. coli* was significantly associated with the detection of *Salmonella*. It is important to note that achieving compliance with FSMA PSR water standards for microbial quality does not guarantee the absence of pathogens. The 17 agricultural ponds that met PSR MWQP standards, all tested positive for *Salmonella* at least once during the study (each pond was sampled 20 times over the course of two years). Topalcengiz and colleagues observed similar results in Central Florida surface waters, where the six irrigation ponds tested met the

PSR water standards, but each of the ponds also contained virulence genes for multiple pathogens, including *Salmonella* (*52*). Luo and colleagues also observed similar results where the 10 ponds sampled, met the FSMA PSR standards, but all tested positive for *Salmonella* at least once during the 22-month study (*36*). The Topalcengiz, Luo, and Havelaar studies (*31, 36*, *52*), as well as the current study, observed generic *E. coli* to be a predictor for *Salmonella* detection, which provides additional justification for the choice of generic *E. coli* as indicator of agricultural water quality in FSMA PSR. The relationship between *Salmonella* and generic *E. coli* is not universal as different studies find different co-variables such as turbidity or farm to modify this relationship. Several studies have shown that the detection of certain pathogens, like *Salmonella*, have been significantly associated with indicator organisms. For example, a North Carolina study (*34*) observed the detection of *Salmonella* was significantly correlated with the presence of 5 different indicator organisms (fecal coliforms, generic *E. coli*, enterococci, somatic coliphage, and *C. perfringens*) in water samples. Additionally, a Florida study observed there was a significant positive relationship between generic *E. coli* and *Salmonella* prevalence; however, that correlation was not strong enough to predict *Salmonella* prevalence based on generic *E. coli* levels (*38*). More research is clearly needed to address the relationship between indicator microorganisms and pathogens, especially quantitative data on indicators and pathogens. However, our study was limited in scope to twenty ponds over two years, and *Salmonella* concentration was only determined in the second year. Each site in this dataset displayed less variability in *E. coli* concentration values than sites from the similar Topalcengiz and colleagues study performed in central Florida (*31, 52*). Our study's limited number of samples per site (20), may not have captured all the variability present. Censoring (samples above or below the limit of detection) does play a role in reducing apparent variability for water

sources. Censored data points as both a number and a percentage of the water source sites data are listed in S2, with 3 sites having a fifth or more of their data points above or below detection limits. Accounting for censoring increases the spread of the fitted distributions, as could be expected theoretically. It is therefore urgently recommended to apply advanced statistical estimation methods, such as the method described here, then calculate unbiased estimates of GM and STV. Statistical software such as R is well suited for this purpose while providing the opportunity for interactive web interfaces. The current study supports the findings of Havelaar and colleagues (31) that more than 20 water samples are needed to accurately portray the variability of *E. coli* concentrations in agricultural water sources, and further research needs to be conducted to determine the causes of such variability to provide growers with information on optimal water quality management. Risks are associated with peak contamination events and understanding what drives such events will provide growers with effective tools to manage agricultural water quality and provide best use of limited resources.

***Salmonella* is detected in untreated surface waters, but in low concentrations.** The overall prevalence of *Salmonella* from sampled surface water sources was 19%. In another study (*5*) surveying surface waters on the ESVA, Bell and colleagues observed a *Salmonella* prevalence of approximately 40% using the same volume of water tested (250 mL) (*5*). Other studies (*36, 38, 39, 49, 52*) performed to investigate the *Salmonella* prevalence in surface water range from 5 to 100%. For example, the *Salmonella* prevalence in water samples (using 250 mL) collected from New York produce production environments was 9% (*49*). Another study by Micallef and colleagues found the *Salmonella* prevalence in water samples (250 mL) from the Mid-Atlantic region (specific locations not identified) was approximately 8% (*39*). A study performed to evaluate Florida surface waters observed a *Salmonella* prevalence of 100% from 10 L water

samples concentrated by filtration to 250 mL samples (*38*). The significantly higher *Salmonella* prevalence observed in this Florida survey *(38)* is likely a direct result of a larger volume of water sampled. Another study performed in Georgia (*28*) surveying surface water sources, also sampled a larger volume of water (555 mL) and observed a higher *Salmonella* prevalence of approximately 80%. Similarly, Rajabi and colleagues (41) tested a larger volume of water using a 5 dilution MPN method (500 mL, 100 mL, 50 mL, 10 mL, 1 mL – 661 mL total) and found a 96% prevalence of *Salmonella* in Florida surface water samples. However, *Salmonella* prevalence can be misleading when linking to risk of produce contamination. *Salmonella* prevalence may be highly associated with the volume of water tested (e.g., larger volumes of water, higher prevalence); as well as risk of contamination may be dependent on produce farm management practices (e.g., overhead, drip irrigation) and or water quality management. Therefore, it is important when comparing *Salmonella* prevalence to other regions (or studies) to consider the volume of water tested. For example, this was evident in our study, using 2016 data, when 38 *Salmonella* positive water samples were detected (19%) upon primary enrichment (250 mL), compared to 80 *Salmonella* positive water samples were detected (40%) upon primary enrichment and MPN (283.3 mL). When comparing studies (*5, 39, 48, 49, 50*) that sampled the same volume of water (250 mL), the prevalence of *Salmonella* in water samples ranged from 8 to 40% (the study reported here observed a 19% *Salmonella* prevalence in ESVA water samples, which was the median prevalence value in the range).

One of the unique aspects of this study was in year two (2016); the concentration of *Salmonella* was quantified for 200 samples. It is crucial to understand the concentration of *Salmonella* in surface water sources to develop mitigation strategies aimed at reducing produce contamination events. For instance, if *Salmonella* concentrations are low, then avoidance of

direct application of water to the harvestable portion of the produce may be effective, as most research emphasizes likely contamination when *Salmonella* is in contact with the harvestable portion (e.g., blossoms, stem scars) (*27, 61*). In the study reported here, the mean concentration of *Salmonella* was 4.4 MPN/100 mL (0.65 log) with a 19% prevalence. A study *(38)* reported similar findings in a 2010 Florida surface water survey with a mean *Salmonella* concentration of 6.2 MPN/100 mL (0.79 log) with a 100% prevalence. The highest *Salmonella* concentration recorded in this study discussed here was 93 MPN/100 mL (1.97 log) and was found in one pond during one sampling. This was much lower than the highest *Salmonella* concentration reported in other comparable southeastern water surveys, also with high *Salmonella* prevalence (1000 MPN/100 mL (3.0 log), (*38*); 1000 MPN/100 mL (3.0 log), (*18*); 5400 MPN/100 mL (3.7 log), (*41*)). Interestingly, in our study the pond (369) with the highest prevalence of *Salmonella* (40%) had the lowest *Salmonella* concentration (all samples were below the limit of detection < 3 MPN/100 mL (< 0.5)). These studies all use similar MPN methodology and provide evidence that while *Salmonella* prevalence may be high, concentration of *Salmonella* was low.

The source of *Salmonella* remains unknown on the ESVA as *Salmonella* diversity in water samples was high, even though *S*. Newport remains top serotype detected. *Salmonella* diversity in the surface water samples was relatively high during the 2015 and 2016 sampling seasons ($D = 0.82 \pm 0.05$). Of the 118 isolates sent for serotyping, 14 different serotypes were identified. In a study conducted in New York and Florida on *Salmonella* subtype distribution, Strawn and colleagues also found a high diversity of *Salmonella* including 14 and 11 different serotypes identified out of 80 and 32 isolates, respectively (48). Additionally, McEgan and colleagues reported 13 different serotypes from 67 isolates in a Central Florida water study (*37*). Another study performed in the Suwannee River identified eight different *Salmonella* serotypes

from 30 isolates *(41)*. These data demonstrate high levels of *Salmonella* diversity amongst serotypes may not be uncommon in the environment, especially agricultural environments.

While *Salmonella* serotype diversity was high, *Salmonella* Newport was the most prevalent of the serotypes identified (approximately 36%). Various studies (*31, 34, 36*) have shown *Salmonella* Newport to be one of the most commonly isolated serotypes. However, there have been five confirmed *Salmonella* Newport outbreaks linked to produce on the ESVA (n=4) and Eastern Shore of Maryland (n=1). While a number of past *Salmonella* outbreaks have been linked to the Delmarva by some evidence (a region including Delaware, Maryland, and Virginia), those outbreaks were never confirmed and remain classified as outbreak source unknown (*14*). The *Salmonella* Newport outbreaks occurred in 2002, 2005, 2006, 2007 and 2014 and resulted in 333, 72, 115, 65, and 275 reported illnesses in the US, respectively *(3, 16, 17, 23, 27)*. All investigations had epidemiological data which showed tomatoes (2002, 2005, 2006, 2007) or cucumbers (2014) were the likely source. Additionally, in 2002 and 2005, the outbreak strain was also isolated from irrigation ponds *(23)*. Since pathogens, like *Salmonella*, can be found in agricultural environments, using food safety best practices is key to reducing the number of produce contamination events that can lead to foodborne illness outbreaks. For example, tomato growers on the ESVA routinely use drip irrigation and sometimes treat water used from surface water sources (e.g., ponds) with 2-7 ppm free chlorine.

Salmonella diversity may also be influenced by seasonal trends, wildlife pressures, and or other factors that have not been investigated. A study (*28*) on the levels on *Salmonella* in Georgia surface waters showed that there were higher concentrations of *Salmonella* found in the summer months, with the highest concentration in August. Water temperature was also positively correlated with *Salmonella* concentration, which could account for the higher levels in the

summer months (*28*). A different study (*5*) performed on the ESVA detected *Salmonella* from goose feces found near surface water sources. Wildlife pressures may influence *Salmonella* prevalence (*26*). Another study *(26)* also performed on the ESVA evaluated gull feces for *Salmonella* collected at four sites for three months. The study authors (26) strongly suggested that patterns of *Salmonella* Newport were endemic to sites on the ESVA where gulls were sampled. In the study reported here, we did not sample wildlife feces (including birds), so we are unable to determine if gull or goose presence (or feces) may have influenced our *Salmonella* findings. However, all ponds sampled where located on commercial produce farms where wildlife is deterred by decoys, squawkers, air cannons, inflatable air noodles/balloons, reflective tape, among others. Poultry is also a known carrier of *Salmonella* (*8*), and the ESVA does have a large poultry industry. Yet, there was a significantly lower prevalence of *Salmonella* in County 1 (nearly two-fold lower), where the majority of the poultry industry is located on the ESVA. In the study reported here, the highest *Salmonella* prevalence was observed in ponds from Farms C and E. Reptiles, such as turtles and snakes, were also observed frequently (at 10 or more samplings) during sampling at Farms C and E. According to the CDC, there were 15 multistate *Salmonella* outbreaks linked to turtles between 2006 and 2014 (*13*), thus this may be a hypothesized source of contamination for the ponds located on these farms.

Conclusions

Data on the prevalence and concentration of microorganisms, including *Salmonella,* in agricultural surface water is imperative to determine what efforts are needed to reduce the risk of possible produce contamination events. Finding the source of *Salmonella* and other pathogens is also crucial to mitigation strategies; for example, growers may treat agricultural water, apply pre-harvest intervals to maximum time between water application and harvest, or use water delivery

methods that avoid water contact with harvestable portions of crops. This study found a positive relationship between generic *E. coli* levels and *Salmonella* detection and also total aerobic bacteria counts and *Salmonella*; thus, monitoring generic *E. coli* or total microbial presence may help to alert growers when the risk of *Salmonella* presence might be higher. Water sources that meet PSR agricultural water standards may still contain pathogens, so further research is still needed to fully understand the relationship between microbial indicators and pathogens, possibly with other indicators beyond generic *E. coli*. By learning more about microbial indicators like generic *E. coli* or total aerobic bacteria, the risks of contamination from *Salmonella* may be lowered. The volume of water sampled directly affects the detection of *Salmonella*, so research on standardizing sample size for pathogen testing is needed. Even though *Salmonella* was detected in untreated surface water samples (19%), the concentration of *Salmonella* was quite low (average concentration of 4.4 MPN/100 mL). Additionally, this study emphasized the importance of collecting quantitative data on pathogens in environmental sources to assist in evaluation of contamination risk and development of mitigation strategies. More research is needed to determine how pathogen concentration affects risk of produce contamination.

Acknowledgements

This research and extension collaboration was supported by Virginia Tech's Eastern Shore Agricultural Research and Extension Center, the Virginia Agricultural Experiment Station, and the Hatch Program of the National Institute of Food and Agriculture, U.S. Department of Agriculture. We are appreciative of the assistance of our fantastic farm partners. We thank Theresa Pittman, Ursula Deitch, Jacob McClaskey, Amy Harrison, and Cecelia Erwin for all their efforts on the project.

References

1. Ahmed, W., A. Vieritz, A. Goonetilleke, and T. Gardner. 2010. Health risk from the use of roof-harvested rainwater in southeast Queensland, Australia, as potable or nonpotable water, determined using quantitative microbial risk assessment. *Appl. Environ. Microbiol.* 76:7382-7391. doi:10.1128/AEM.00944-10.

2. Allard, S., S. Solaiman, M. T. Callahan, E. Handy, C. East, H. C. Kelbick, R. Murray, A. Bui, J. Haymaker, D. Foust, S. Gartley, A. Vanore, S. Parveen, F. Hashem, M. Taabodi, K. Kniel, M. Sharma, S. Micallef, and A. Sapkota. 2017. Generic *E. coli* levels in surface and non-traditional irrigation water in the Mid-Atlantic in relation to FSMA water quality standards: a CONSERVE study. Abstract. International Association for Food Protection Annual Meeting, Tampa, FL. Available at: https://iafp.confex.com/iafp/2017/webprogram/Paper15628.html. Accessed 26 June 2017.

3. Angelo, K. M., A. Chu, M. Anand, T. Nguyen, L. Bottichio, M. Wise, I. Williams, S. Seelman, R. Bell, M. Fatica, S. Lance, D. Baldwin, K. Shannon, H. Lee, E. Trees, E. Strain, and L. Gieraltowski. 2015. Outbreak of *Salmonella* Newport infections linked to cucumbers – United States, 2014. *Morb. Mort. Wkly. Rep*. 64:144-147.

4. Antaki, E. M., G. Vellidis, C. Harris, P. Aminabadi, K. Levy, M. T. Jay-Russell. 2016. Low concentration of *Salmonella enterica* and generic *Escherichia coli* in farm ponds and irrigation distribution systems used for mixed produce production in Southern Georgia. *Foodborne Pathog. Dis*. 13:551-558. doi:10.1089/fpd.2016.2117.

5. Bell, R. L., J. Zheng, E. Burrows, S. Allard, C. Y. Wang, C. E. Keys, D. C. Melka, E. Strain, Y. Luo, M.W. Allard, S. Rideout, E. W. Brown. 2015. Ecological prevalence, genetic diversity, and epidemiological aspects of *Salmonella* isolated from the tomato agricultural regions of the Virginia Eastern Shore. *Front. Microbiol.* 6:415. doi:10.3389/fmicb.2015.00415.

6. Beuchat, L.R. 2006. Vectors and conditions for preharvest contamination of fruits and vegetables with pathogens capable of causing enteric diseases. *Brit. Food. J*. 106:38-53. doi:10.1108/00070700610637625.

7. Beuchat, L.R. 1995. Pathogenic microorganisms associated with fresh produce. *J. Food Prot.* 59:204-216. doi:10.4315/0362-028X-59.2.204.

8. Braden, C.R. 2006. *Salmonella enterica* serotype Enteritidis and eggs: a national epidemic in the United States. *Clin. Infect. Dis*. 43:512-517. doi:10.1086/505973.

9. Brassill, N. 2013. The assessment of *Escherichia coli* as an indicator of microbial quality of irrigation waters used for produce. Master's thesis, University of Arizona. Available at: http://arizona.openrepository.com/arizona/handle/10150/293596. Accessed 30 November 2017.

10. Buckalew, D., L. Hartman, G. Grimsley, A. Martin, and K. Register. 2006. A long-term study comparing membrane filtration with Colilert® defined substrates in detecting fecal coliforms and *Escherichia coli* in natural waters. *J. Environ. Manage*. 80:191-197. doi: 10.1016/j.jenvman.2005.08.024.

11. Budnick, G. E., R. T. Howard, and D. Mayo. 2001. Comparison of Colilert-18 to the mTEC agar method for the enumeration of *Escherichia coli* in recreational waters. Proceedings from the American Society of Microbiologists Annual Convention May 2001. Available at: http://www.idexx.com/resource-library/water/water-reg-article61.pdf. Accessed 30 November 2017.

12. Callejon, R. M., M. I. Rodriquez-Naranjo, C. Ubeda, R. Hornedo-Ortega, M. C. Garcia-Parrilla, and A. M. Troncoso. 2015. Reported foodborne outbreaks due to fresh produce in the United States and European Union: trends and causes. *Foodborne Pathog. Dis.* 12:32-38. doi:10.1089/fpd.2014.1821.

13. Centers for Disease Control and Prevention. 2017. Take care with pet reptiles and amphibians. Available at: https://www.cdc.gov/features/salmonellafrogturtle/index.html. Accessed 28 June 2017.

14. Centers for Disease Control and Prevention. 2016. *Salmonella*. Available at: https://www.cdc.gov/salmonella/index.html. Accessed 6 April 2017.

15. Centers for Disease Control and Prevention. 2008. Outbreak of *Salmonella* serotype Saintpaul infections associated with multiple raw produce items---United States, 2008. *Morb. Mort. Wkly. Rep.* 57:929-934.

16. Centers for Disease Control and Prevention. 2007. Multistate outbreaks of *Salmonella* infections associated with raw tomatoes eaten in restaurants---United States, 2005--2006. *Morb. Mort. Wkly. Rep*. 56:909-911.

17. Centers for Disease Control and Prevention. 2002. Multistate outbreaks of *Salmonella* serotype Poona infections associated with eating cantaloupe from Mexico---United States and Canada, 2000-2002. *Morb. Mort. Wkly. Rep*. 51:1044-1047.

18. Chapin, T., and M. D. Danyluk. 2017. Isolation of *Salmonella* and detection of generic *Escherichia coli* populations from South Florida surface waters. Abstract. International Association for Food Protection Annual Meeting, Tampa, FL. Available at: https://iafp.confex.com/iafp/2017/webprogram/Paper15524.html. Accessed 8 August 2017.

19. Cowburn, J. K., T. Goodall, E. J. Fricker, K. S. Watler, C. R. Fricker. 1994. A preliminary study of the use of Colilert® for water quality monitoring. *Appl. Microbiol*. 19:50-52. doi:10.1111/j.1472-765X.1994.tb00902.x.

20. Department of Health and Human Services. 2015. Standards for the growing, harvesting, packing, and holding of produce for human consumption. Federal Register 80:74353-74672.

Available at: https://www.gpo.gov/fdsys/pkg/FR-2015-11-27/pdf/2015-28159.pdf. Accessed 7 April 2017.

21. Fricker, C. R., P. S. Warden, M. DeSarno, B. J. Eldred. 2010. Significance of methods and sample volumes for *E. coli* and total coliform measurements. Water Research Foundation. Available at: http://www.waterrf.org/PublicReportLibrary/4024.pdf. Accessed 30 November 2017.

22. Gorski, L., C. T. Parker, A. Liang, M. B. Cooley, M. T. Jay-Russell, A. G. Gordus, E. R. Atwill, R. E. Mandrell. 2011. Prevalence, distribution, and diversity of *Salmonella* enterica in a major produce region of California. *Appl. Environ. Microbiol*. 77:2734-2748. doi:10.1128/AEM.02321-10.

23. Greene, S. K., E. R. Daly, E. A Talbot, L. J. Demma, S. Holzbauer, N. J. Patel, T. A. Hill, M. O. Walderhaug, R. M. Hoekstra, M. F. Lynch, J. A. Painter. 2008. Recurrent multistate outbreak of *Salmonella* Newport associated with tomatoes from contaminated fields, 2005. *Epidemiol. Infect*. 136:157-165. doi:10.1017/S095026880700859X.

24. Grimont, P. A. D. and F. X. Weill. 2007. Antigenic formulae of the *Salmonella* serovars, 9th ed. World Health Organization Collaborating Centre for Reference and Research on *Salmonella*, Pasteur Institute, Paris.

25. Gronewald, A. D., and R. L. Wolpert. 2008. Modeling the relationship between most probable number (MPN) and colony-forming unit (CFU) estimates of fecal coliform concentration. *Water Res.* 42:3327-3334. doi:10.1016/j.watres.2008.04.011.

26. Gruszynski, K., S. Pao, C. Kim, D. Toney, K. Wright, P. G. Ross, K. Colon, and S. Levine. 2014. Evaluating wildlife as a potential source of *Salmonella* serotype Newport (JJPX01.0061) contamination for tomatoes on the Eastern Shore of Virginia. *Zoonoses Public Health* 61:202-207. doi:10.1111/zph.12061.

27. Guo, X., J. Chen, R. E. Brackett, and L. R. Beuchat. 2001. Survival of Salmonellae on and in tomato plants from the time of inoculation at flowering and early stages of fruit development through fruit ripening. *Appl. Environ. Microbiol*. 67:4760-4764. doi:10.1128/AEM.67.10.4760-4764.2001.

28. Haley, B. J., D. J. Cole, and E. K. Lipp. 2009. Distribution, diversity, and seasonality of waterborne salmonellae in a rural watershed. *Appl. Environ. Microbiol*. 75:1248-1255. doi:10.1128/AEM.01648-08.

29. Hanning, I. B., J. D. Nutt, S. C. Ricke. 2009. Salmonellosis outbreaks in the United States due to fresh produce: sources and potential intervention methods. *Foodborne Pathog. Dis*. 6:635-648. doi:10.1089/fpd.2008.0232.

30. Hargett, E., and L. Goyn. 2004. Modified mTEC agar, Colilert®, and M-FC agar – field trial comparison of bacteria enumeration methods in surface waters of Eastern Wyoming. Available at: http://deq.wyoming.gov/media/attachments/Water%20Quality/Surface%20Water%20Monitoring/Publications/wqd-wpp-monitoring_comp_study_e.coli_final2.pdf. Accessed 11 November 2017.

31. Havelaar, A. H., K. M. Vazquez, Z. Topalcengiz, R. Muñoz-Carpena, and M. D. Danyluk. Evaluating the U.S. Food Safety Modernization Act Produce Safety Rule standard for microbial quality of agricultural water for growing produce. *J. Food Prot*. 80:1832-1841. doi:10.4315/0362-028X.JFP-17-122.

32. Henry, R., C. Schang, P. Kolotelo, R. Coleman, G. Rooney, J. Schmidt, A. Deletic, and D. T. McCarthy. 2016. Effect of environmental parameters on pathogen and faecal indicator organism concentrations within an urban estuary. *Estuar. Coast. Shelf Sci.* 174:18-26.

33. Holvoet, K., A. De Keuckelaere, I. Sampers, S. Van Haute, A. Stals, and M. Uyttendaele. 2014. Quantitative study of cross-contamination with *Escherichia coli*, *E. coli* 0157, MS2 phage and murine norovirus in a simulated fresh-cut lettuce wash process. *Food Control* 37:218-227.

34. Krometis, L. H., G. W. Characklis, P. N. Drummey, and M. D. Sobsey. 2010. Comparison of the presence and partitioning behavior of indicator organisms and *Salmonella* spp. in an urban watershed. *J Water Health* 8:44-59. doi:10.2166/wh.2009.232.

35. Li, B., G. Vellidis, H. Liu, M. Jay-Russell, S. Zhao, Z. Hu, A. Wright, and C. A. Elkins. 2014. Diversity and antimicrobial resistance of *Salmonella enterica* isolates from surface water in Southeastern United States. *Appl. Environ. Microbiol*. 80:6355-6365. doi:10.1128/AEM.02063-14

36. Luo, Z., G. Gu, A. Ginn, M. C. Giurcanu, P. Adams, G. Vellidis, A.H. C. van Bruggen, M. D. Danyluk, and A. C. Wright. 2015. Distribution and characterization of *Salmonella* enterica isolates from irrigation ponds in the Southeastern United States. *Appl. Environ. Microbiol*. 81:4376-4387. doi:10.1128/AEM.04086-14.

37. McEgan, R., J. C. Chandler, L. D. Goodridge, and M. D. Danyluk. 2014. Diversity of *Salmonella* isolates from Central Florida surface waters. *Appl. Environ. Microbiol*. 80:6819:6827. doi:10.1128/AEM.02191-14.

38. McEgan, R., G. Mootlan, L. D. Goodridge, D. W. Schaffner, and M. D. Danyluk. 2013. Predicting *Salmonella* populations from biological, chemical, and physical indicators in Florida surface waters. *Appl. Environ. Microbiol*. 79:4094-4105. doi:10.1128/AEM.00777-13.

39. Micallef, S. A., R. E. Rosenberg Goldstein, A. George, L. Kleinfelter, M. S. Boyer, C. R. McLaughlin, A. Estrin, L. Ewing, J. J. Beaubrun, D. E. Hanes, M. H. Kothary, B. D. Tall, J.

H. Razeq, S. W. Joseph, and A. R. Sapkota. 2012. Occurrence and antibiotic resistance of multiple *Salmonella* serotypes recovered from water, sediment, and soil on mid-Atlantic tomato farms. *Environ. Res*. 114:31-39. doi:10.1016/j.envres.2012.02.005.

40. New Hampshire Department of Environmental Sciences. 2003. Environmental Fact Sheet: Fecal Coliform as an Indicator Organism. Available at: http://www.des.nh.gov/organization/commissioner/pip/factsheets/wwt/documents/web-18.pdf. Accessed 17 May 2017.

41. Rajabi, M., M. Jones, H. Michael, G. Rodrick, and A. C. Wright. 2011. Distribution and genetic diversity of *Salmonella* enterica in the upper Suwannee River. *Int. J. Microbiol*. 2011:1-9. doi:10.1155/2011/461321.

42. Rangarajan, A., E. A. Bihn, R. B. Gravani, D. L. Scott, and M. P. Pritts. 2000. Food Safety Begins on the Farm: A Grower's Guide. Cornell Good Agricultural Practices Program. Cornell University, Ithaca, NY. Available at: https://gaps.cornell.edu/sites/gaps.cornell.edu/files/shared/documents/FSBFEngLOW.pdf. Accessed 6 June 2017.

43. Rudolfs, W., L. L. Falk, and R. A. Ragotzkie. 1950. Literature review on the occurrence and survival of enteric, pathogenic, and relative organisms in soil, water, sewage, and sludges, and on vegetation: I. Bacterial and Virus Diseases. *Sewage Ind. Waste*. 22:1261-1281.

44. Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin. 2011. Foodborne illness acquired in the United States – major pathogens. *Emerg. Infect. Dis*. 17:7-15. doi:10.3201/eid1701.P11101.

45. Simpson, E. H. 1949. Measurement of diversity. *Nature* 163:688. doi:10.1038/163688a0.

46. Steele, M., and J. Odumeru. 2004. Irrigation water as a source of foodborne pathogens on fruits and vegetables. *J. Food. Prot*. 67:2839-2849. doi: 10.4315/0362-028X-67.12.2839.

47. Stoeckel, D., G. Wall, and B. Bihn. 2016. FSMA Produce Safety Rule water requirements: insights to get you organized! Produce Safety Alliance. Cornell University College of Agriculture and Life Sciences, Ithaca, NY. Available at: https://producesafetyalliance.cornell.edu/resources/educational-materials/fsma-produce-safety-rule-water-requirements-insights-get-you-organized. Accessed 31 May 2017.

48. Strawn, L. K., M. D. Danyluk, R. W. Worobo, and M. Wiedmann. 2014. Distributions of *Salmonella* subtypes differ between two U.S. produce-growing regions. *Appl. Environ. Microbiol*. 80:3982-3991. doi:10.1128 /AEM.00348-14.

49. Strawn, L. K., E. D. Fortes, E. A. Bihn, K. K. Nightingale, Y. T. Gröhn, R. W. Worobo, M. Wiedmann, and P. W. Bergholz. 2013. Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. *Appl. Environ. Microbiol*. 79:588-600. doi:10.1128/AEM.02491-12.

50. Strawn, L. K., Y. T. Grohn, S. Warchocki, R. W. Worboro, E. A Bihn, and M. Wiedmann. 2013. Risk factors associated with *Salmonella* and *Listeria monocytogenes* contamination of produce fields. *Appl. Environ. Microbiol*. 79:7618-7627. doi:10.1128/AEM02831-13.

51. Thomas, J. C., M. A. Lutz, J. L. Bruce, D. J. Graczyk, K. D. Richards, D. P. Krabbenhoft, S. M. Westenbroek, B. C. Scudder, D. J. Sullivan, and A. H. Bell. 2007. Water-quality characteristics for selected sites within the Milwaukee Metropolitan Sewerage District planning area, Wisconsin, February 2004-September 2005: U.S. Geological Survey Scientific Investigations Report 2007-5084. Available at: https://pubs.usgs.gov/sir/2007/5084/pdf/SIR_2007-5084.pdf. Accessed 17 May 2017.

52. Topalcengiz, Z., L. K. Strawn, and M. D. Danyluk. 2017. Microbial quality of agricultural water in Central Florida. PLoS One 12:e0174889. doi:10.1371/journal.pone.0174889.

53. U.S. Department of Agriculture and U.S. Department of Health and Human Services. 2010. Dietary Guidelines for Americans, 2010. Available at: https://health.gov/dietaryguidelines/dga2010/DietaryGuidelines2010.pdf. Accessed 24 May 2017.

54. U.S. Food and Drug Administration. 2017. Bacteriological Analytical Manual (BAM). Available at: https://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm2006949.htm. Accessed 7 June 2017.

55. U.S. Food and Drug Administration. 2017. FSMA Final rule on produce safety. Available at: https://www.fda.gov/food/guidanceregulation/fsma/ucm334114.htm#key. Accessed 7 April 2017.

56. U.S. Food and Drug Administration. 2012. Bad Bug Book: Foodborne Pathogenic Microorganisms and Natural Toxins Handbook, *Salmonella* species. Available at: https://www.fda.gov/downloads/Food/FoodborneIllnessContaminants/UCM297627.pdf. Accessed 7 April 2017.

57. Uyttendaele, M., L. Jaykus, P. Amoah, A. Chiodini, D. Cunliffe, L. Jacxsens, K. Holvoet, L. Korsten, M. Lau, P. McClure, G. Medema, I. Sampers, and P. R. Jasti. 2015. Microbial hazards in irrigation water: standards, norms, and testing to manage use of water in fresh produce primary production. *Compr. Rev. Food Sci. Food Saf*. 14:336-356.

58. Walters, S. P., N. Gonzalez-Escalona, I. Son, D. C. Melka, L. M. Sassoubre, A. B. Boehm. 2013. *Salmonella enterica* diversity in Central Californian coastal waterways. *Appl. Environ. Microbiol*. 79:4199-4209. doi:10.1128/AEM.00930-13.

59. Wells, H. F., and J. C. Buzby. 2008. Dietary assessment of major food trends in U.S. food consumption, 1970-2005. United States Department of Agriculture: Economic Information Bulletin No. 33. Available at:

https://www.ers.usda.gov/webdocs/publications/44217/12199_eib33_1_.pdf?v=41055. Accessed 24 May 2017.

60. Western Center for Food Safety. 2016. Agricultural Water and the Produce Safety Rule. University of California, Davis, Davis, CA. Available at: http://ucfoodsafety.ucdavis.edu/files/268306.xlsx. Accessed 14 June 2017.

61. Zheng, J., S. Allard, S. Reynolds, P. Millner, G. Arce, R. J. Blodgett, and E. W. Brown. 2013. Colonization and internalization of *Salmonella enterica* in tomato plants. *Appl. Environ. Microbiol*. 79:2494-2502. doi:10.1128/AEM.03704-12.

Appendices

A. Tables

TABLE 1. Description of water sampling sites on the Eastern Shore of Virginia.

Pond	Size of Pond[a]	Heavy Vegetation[b]	Creek Fed[c]	Sunlight Exposure[d]	Distance to Road[e]	Distance to Packing Shed[f]
852	153	no	no	high	482	483
159	475	no	yes	high	483	483
486	654	no	yes	medium	1930	1930
279	53.0	no	no	medium	97.0	483
681	212	no	no	low	1130	1290
173	35.0	no	no	medium	16.0	5150
543	157	no	yes	medium	16.0	644
340	157	no	yes	medium	1610	3710
199	211	no	no	high	161	19200
652	683	somewhat	yes	low	145	145
803	49.0	somewhat	yes	medium	97.0	161
369	80.0	no	no	high	805	32.0
771	95.0	no	no	medium	1130	322
915	121	no	no	high	1610	2890
841	293	no	yes	low	16.0	6280
326	157	no	no	high	322	3710
284	146	no	no	high	113	6.00

653	776	no	no	low	966	3710
672	87.0	no	no	low	322	5470
228	142	heavy	yes	high	1290	81.0

[a, e, f] variables measured in meters using Google Maps
[b, d] variables measured by visual observation at time of sampling
[c] variables determined by combination of visual observation at the time of sampling and Google Maps Satellite View

TABLE 2. Median concentration of generic *E. coli*, total coliforms, total aerobic bacteria, and *Salmonella*, and prevalence of *Salmonella* for each pond.

	Median of				
Pond	*Generic E. coli*[a]	Total Coliforms[b]	Total Aerobic Bacteria[c]	*Salmonella*[d]	*Salmonella*[e] (%)
159	4.60	1010	5.20	< 3.00	5.00
173	93.3	1010	6.98	3.00	30.0
199	5.75	1010	4.82	3.00	5.00
228	73.3	1010	5.45	3.00	35.0
279	13.0	1010	6.60	3.00	10.0
284	53.3	1010	6.70	3.30	30.0
326	2.55	1010	4.96	< 3.00	10.0
340	50.8	1010	4.96	< 3.00	20.0
369	47.8	1010	4.67	3.00	40.0
486	32.2	1010	5.25	< 3.00	10.0
543	16.4	1010	4.99	3.00	5.00
652	14.0	1010	5.22	3.00	20.0
653	12.1	1010	5.39	3.00	2.00
672	26.3	986	4.85	3.00	30.0
681	18.3	1010	5.02	3.00	25.0
771	9.75	1010	4.9	3.00	40.0
803	15.8	1010	5.42	3.00	5.00

841	7.95	1010	5.19	3.00	25.0
852	4.05	1010	5.93	3.00	10.0
915	8.60	1010	5.11	3.00	20.0

[a, b, d] Median concentration (across all 20 sampling visits) in MPN/100 mL

[c] Median concentration (across all 20 sampling visits) in log CFU/100 mL

[e] Prevalence (across all 20 sampling visits) in percent

TABLE 3. Effect of factor (year, county, and farm) on frequency of positive *Salmonella* surface water samples.

Factor (No. of samples)	Frequency (%)[a]
Year	
2015 (200)	38 (19)
2016 (200)	38 (19)
County	
1 (220)	29 (13)^{A}
2 (180)	47 (26)^{B}
Farm	
A (80)	7 (9)^{A}
B (100)	17 (17)^{AB}
C (60)	20 (33)^{B}
D (80)	14 (18)^{AB}
E (40)	13 (33)^{B}
F (40)	5 (13)^{AB}

[a] Different superscript letters represent statistically significant differences with *P*-value <0.05. No letters represent values that are not significantly different.

TABLE 4. Distribution of *Salmonella* serotypes detected at each sampling site during the 2015 and 2016 growing seasons combined.

Serotype	Pond																				
	681	173	543	340	199	672	228	841	326	284	653	369	771	915	852	159	486	279	652	803	Total
Newport	2	4	2	1	2	4	2	2	2	1		6	5	2	1	1	1	1	1	2	42
Javiana	3	1	1				2			4		2		1				2	1		17
Norwich						2	2		1	6		1	2	1					1		16
Thomson	1	5		1		2	2												1		12
Typhimurium	3		1									1		1				1			7
Saintpaul							2				1		1							1	5
Infantis															1		1				4
Hartford								4													4
Berta								3			1										4
Bareilly							1						1								2
III_38:(k):z35																			1	1	2
Anatum															1						1
9,12:I,z28:-																		1			1

Senftenberg	1	1

B. Figures

Figure Legends

Figure 1. The distribution of the *Salmonella* serotypes isolated from water samples during the 2015 and 2016 growing seasons. The x axis represents the number of times each serotype was isolated during the 2015 and 2016 growing seasons combined. The y axis represents each serotype that was identified in the study. *Salmonella* serotype Newport was isolated most frequently (42 of the 118 isolates).

Figure 2. Logistic regression model of the relationship between farm and probability of *Salmonella* presence. The x axis represents the concentration of generic *E. coli* in MPN/100 mL and the y axis represents the probability of detecting *Salmonella* based on generic *E. coli* levels, total aerobic bacteria levels, and farm. Each letter represents a different farm. The lines represent median total aerobic bacteria levels and the upper and lower bounds of the ribbon are the 95% confidence limits of TAPC value. Farms C and E had a significantly higher probability of detecting *Salmonella* than Farm A ($P < 0.05$).

Figure 1

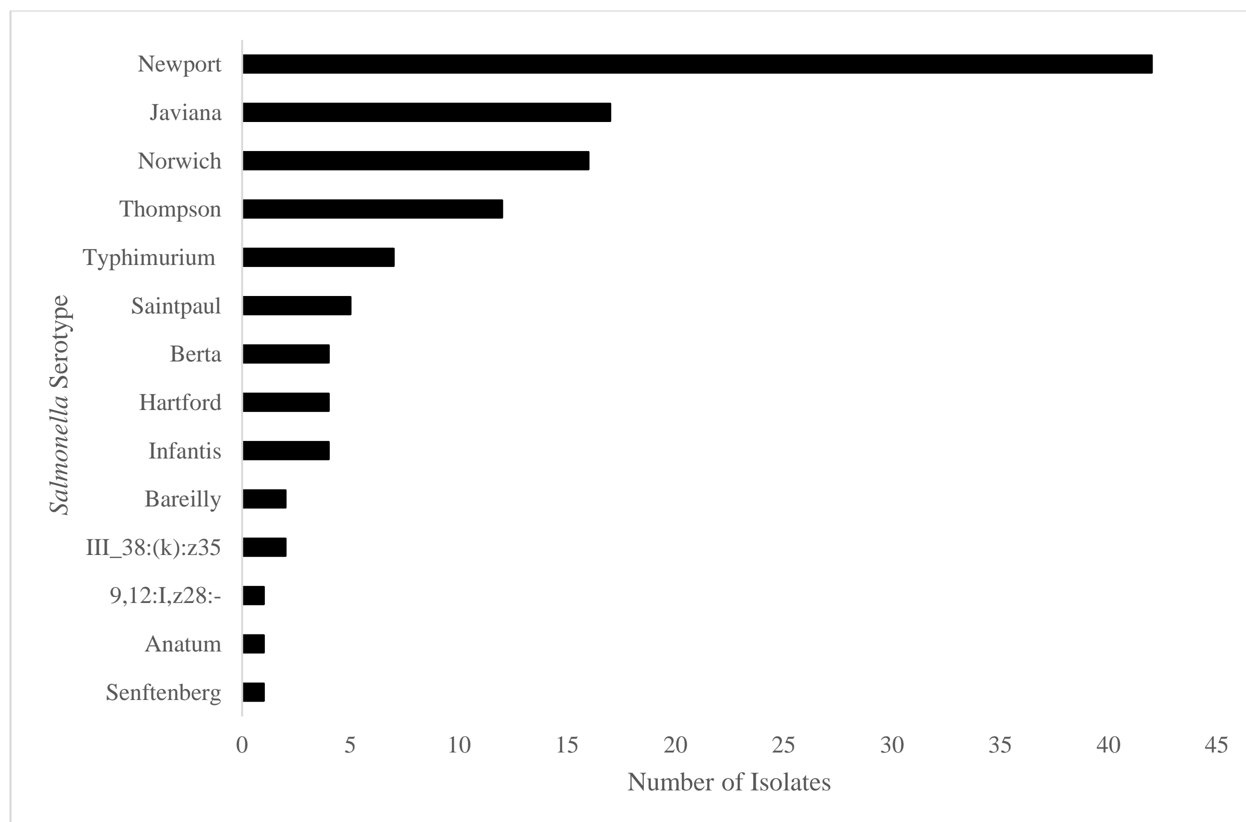
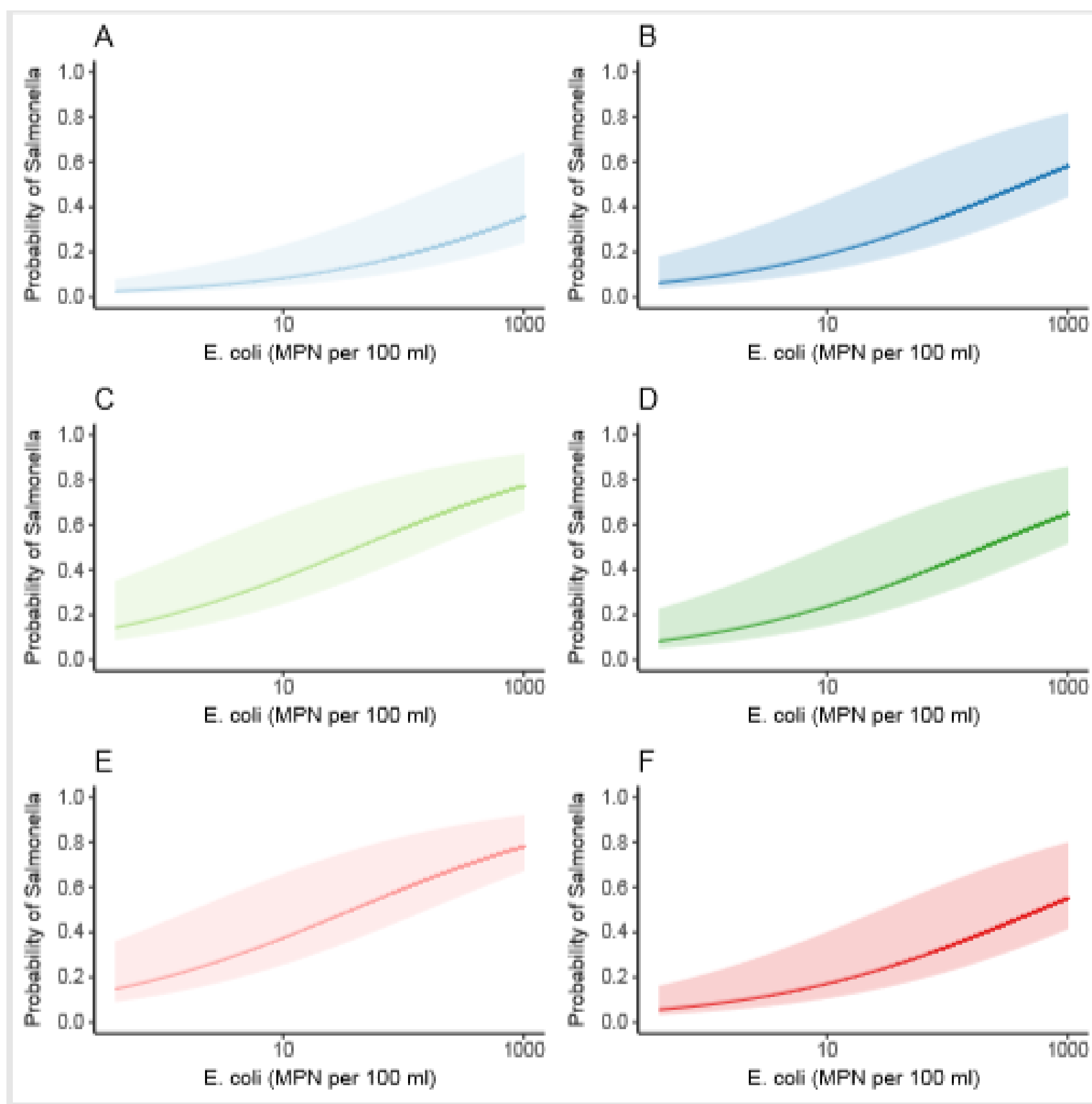


Figure 2



C. Supplemental Material

S1. TABLE 1. Names, Descriptions, and Units of Measurement (where applicable) of the Environmental and Meteorological Factors Observed at Each Site.

Variable name	Description	Unit[a]
Precipitation (11 sub-variables)	Amount of precipitation on each day for 7 days prior to sampling, total precipitation for the 7 days prior to sampling, total precipitation for 3 days prior to sampling, and average precipitation for 7 days prior to sampling	mm
Temperature (9 sub-variables)	The temperature for each day for 7 days prior to sampling and the average temperature for the 7 days prior to sampling	°C
Reptile Presence	Reptile presence or absence was observed at each location for each sampling	NA
Proximity to Packing Shed	Distance from the sampling site to the packing facility	m
Proximity to Impervious Surfaces	Distance from the sampling site to a road	m

[a] NA is not applicable.

S2. TABLE 2. Statistics of the lognormal (base e) distribution of *E. coli* concentrations (MPN per 100 ml) for all sites. All results are based on 20 observations per site. Censored data (i.e. values above and below the limits of detection) were replaced with the limit of detection.

Site	Censored Data Points[a]	Censored Data Points[b]	Mean	Standard Deviation	Skewness	Kurtosis
852	6	30	1.924	2.096	0.886	2.8
159	2	10	1.741	1.399	0.355	2.185
486	1	5	3.032	2.029	0.165	2.373
279	1	5	3.004	1.973	0.182	2.058
681	1	5	2.692	1.806	0.394	2.577

173	3	15	4.123	2.261	-0.485	2.101
543	1	5	2.619	1.666	0.457	3.714
340	1	5	4.076	1.371	0.522	2.245
199	3	15	1.639	1.162	0.003	2.002
369	0	0	3.794	1.679	-0.475	2.751
771	1	5	2.412	1.475	0.116	2.094
915	2	10	2.059	1.819	0.826	2.928
841	4	20	2.035	1.713	0.209	2.715
326	5	25	1.509	1.583	0.7	2.342
284	2	10	4.013	1.949	-0.152	2.237
653	1	5	2.349	1.612	0.173	2.295
672	0	0	2.865	1.512	-0.058	2.34
228	1	5	3.815	2.215	-0.22	1.725
652	1	5	2.538	1.196	-0.646	3.028
803	0	0	2.781	1.278	-0.381	2.427

[a] Number of censored data points for each sampling site

[b] Percentage of censored data points per sampling site

S3. TABLE 3. Comparison of distribution fitting by maximum likelihood estimation through AIC comparison. The lowest AIC for each site is highlighted in blue and the second lowest is highlighted in orange. *

Site	Lognormal	Weibull	Pareto	Fréchet
852	161.05	162.46	161.61	161.27
159	143.49	144.95	144.15	144.75
486	209.32	212.65	210.64	210.12
279	207.06	208.45	208.26	208.71
681	191.20	194.32	191.44	192.10
173	257.24	254.00	260.62	262.55
543	185.26	189.93	184.26	187.15
340	235.39	240.13	237.56	233.37
199	133.64	132.02	132.26	137.56
369	232.23	231.51	232.05	237.48
771	172.44	173.47	173.15	174.71
915	165.45	169.66	165.32	164.07
841	161.48	161.08	162.35	164.91
326	136.20	136.78	137.10	137.09
284	247.37	246.52	248.36	251.85
653	173.13	174.94	173.73	174.96
672	190.85	192.82	191.57	193.20
228	244.14	243.67	247.66	246.58
652	169.87	166.75	166.75	176.98

803	180.77	180.30	180.27	185.18

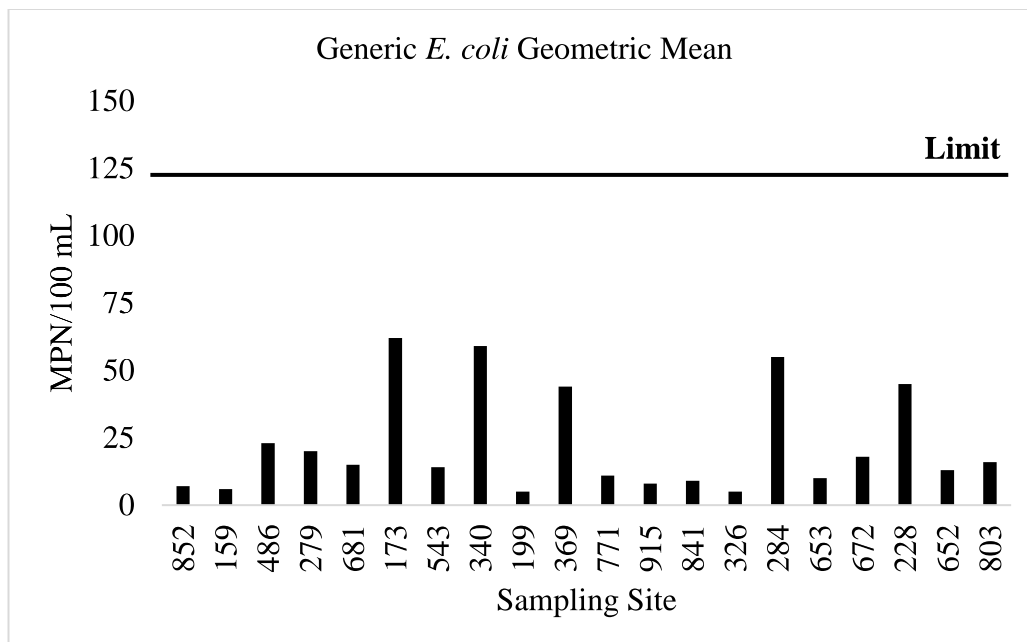
S4. TABLE 4. Comparison of GM and STV (50th and 90th percentile of a lognormal distribution) found through the method of moments (MM) and maximum likelihood estimation (MLE) with 95% confidence intervals included for MLE. For each comparison, the higher value is highlighted in blue. Where values are equal, both are highlighted in orange.*

Site	GM		STV	
	MM	MLE	MM	MLE
852	6.85	4.61 (1 - 15)	100.18	123.39 (22 - 478)
159	5.71	5.29 (3 - 10)	34.18	35.57 (14 - 75)
486	20.74	20.74 (9 - 50)	278.35	261.44 (84 - 786)
279	20.16	19.29 (8 - 45)	252.05	253.05 (80 - 675)
681	14.76	14.16 (6 - 34)	148.97	149.27 (41 - 424)
173	61.72	58.92 (19 - 160)	1115.17	1130.75 (366 - 2248)
543	13.72	13.22 (7 - 28)	115.73	116.27 (35 - 358)
340	58.89	58.89 (33 - 101)	340.65	326.52 (121 - 679)
199	5.15	4.68 (2 - 8)	22.79	24.8 (13 - 43)
369	44.45	44.45 (21 - 87)	381.3	362.02 (137 - 679)
771	11.16	10.81 (6 - 21)	73.74	74.15 (29 - 150)
915	7.84	7.09 (3 - 17)	80.46	84.05 (22 - 299)
841	7.65	6.15 (2 - 15)	68.5	80.74 (28 - 218)
326	4.52	3.36 (1 - 8)	34.29	40.46 (12 - 111)
284	55.3	53.28 (21 - 125)	669.77	680.97 (227 - 1593)
653	10.47	10.09 (5 - 20)	82.49	82.61 (28 - 184)
672	17.54	17.54 (9 - 32)	121.47	115.93 (51 - 227)

228	45.37	45.37 (16 - 123)	772.8	721.68 (212 - 1634)
652	12.65	12.37 (7 - 21)	58.5	59.17 (34 - 95)
803	16.13	16.13 (9 - 26)	82.82	79.61 (39 - 130)

S5. Figure 1. The geometric means (GM) for generic *E. coli* for each pond in MPN/100 mL. For the bar graph, the x axis represents each sampling site and the y axis represents the GM for each site in MPN/100 mL. All 20 sites were below 126 MPN/100 mL generic *E. coli*, and therefore met the FSMA PSR standards for GM.

S5. Fig. 1



S6. Figure 2. The statistical threshold values (STV) for generic *E. coli* for each pond in MPN/100 mL. For the bar graph, the x axis represents each sampling site and the y axis represents the STV for each site in MPN/100 mL. Seventeen of the 20 sites met the FSMA PSR standards by having a STV below 410 MPN/100 mL. Three sites did not meet the standards because the STV exceeded the limit.

S6. Fig. 2