ABSTRACT

Title of Dissertation: EVALUATION AND MODELING OF RISK FACTORS ASSOCIATED WITH MICROBIAL CONTAMINATION IN PRODUCE PRE-HARVEST ENVIRONMENT

Hao Pang, Doctor of Philosophy, 2017

Dissertation directed by: Assistant professor, Abani K. Pradhan, Department of Nutrition and Food Science

Produce (fruits and vegetables) has been frequently linked to foodborne disease outbreaks in the United States and worldwide. Produce-related outbreaks have been traced back to contamination occurring at pre-harvest production level. The overall goal of this dissertation was to identify risk factors for produce pre-harvest contamination and develop models to predict the introduction, survival, and persistence of enteric foodborne bacteria in produce at the pre-harvest level.

Produce from mixed farms, where vegetable crops and animals are grown at the same premise, is potentially at higher risk of microbial contamination due to its proximity to environmental reservoirs such animal enclosures and composting facilities. Such contamination can be affected by meteorological factors such as temperature, precipitation, and wind speed. By integrating microbial sampling and meteorological data, the effects of meteorological factors on prevalence and concentration of *Listeria* species and generic *Escherichia coli* in samples collected

from a mixed produce and dairy farm were analyzed using logistic regression and tree-based methods. The developed models have robust predictive ability and can be used to estimate the risk of microbial contamination in mixed farms under different weather conditions.

Survival and persistence of pathogens in field soil is a food safety concern as soil can serve as a source and route for microbial contamination in produce. Regression models were developed to evaluate the effects of meteorological factors, cover cropping, and farming system on the survival and persistence of generic *E. coli* and *L. innocua* in produce field soil. The models revealed that survival of *E. coli* and persistence of *L. innocua* were predominately influenced by temperature, precipitation, and relative humidity.

Further, data from a large microbial sampling study were used to determine the effects of a variety of meteorological, environmental, and farm management factors on the presence and concentration of food safety and quality bacteria indicator in tomatoes and tomato environmental samples. The results suggest that microbial contamination in tomatoes and in tomato production environments can be significantly affected by certain meteorological conditions, environmental factors, and farm management practices.

In conclusion, this study identified potential risk factors associated with the presence, concentration, survival, and persistence of enteric foodborne bacteria in produce and in produce production environments. The developed models can be used to predict the risk of microbial contamination in produce farms under different meteorological conditions, geographical regions, and farm management practices.

Such information and tools will help growers to improve farm management practices to reduce potential contamination of produce.

EVALUATION AND MODELING OF RISK FACTORS ASSOCIATED WITH MICROBIAL CONTAMINATION IN PRODUCE PRE-HARVEST ENVIRONMENT

by

Hao Pang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Assistant Professor Abani K. Pradhan, Chair
Professor Robert L. Buchanan
Professor Kathryne L. Everts
Assistant Professor Shirley A. Micallef
Dr. Vijay K. Juneja

Acknowledgements

First and foremost, my deepest gratitude goes to my advisor Dr. Abani Pradhan. He offered me the opportunity to pursue my dream of getting a graduate degree in the U.S. Without the guidance and support from him, I would not be the researcher and person I am today. Dr. Pradhan has always pushed me to strive for a higher standard through my graduate study which greatly helped in establishing work habits as a researcher and overcoming encountered difficulties as a Ph.D. student. I am always grateful for his constant support, caring, and patience. I am very fortunate to have Dr. Pradhan as my advisor.

I would like to express my sincere thanks to Dr. Robert Buchanan, for countless invaluable suggestions on my research and sharing his wisdom of life. I would also like to thank Dr. Shirley Micallef for the guidance, perspective, as well as time and support she provided during many discussions of my research. I am also grateful to Dr. Vijay Juneja from the U.S. Department of Agriculture for his support and guidance on my research, and his thoughtful assist and advice on my future career. Lastly, I would like to thank Dr. Kathryne Everts for her insightful suggestions and her help in gathering data and information needed for my research.

I am extremely grateful to all the help from the members of Dr. Pradhan's lab, both past and present: Dr. Elisabetta Lambertini, Dr. Miao Guo, Dr. Abhinav Mishra, Miao Wang, Yuqing Ying, Huilin Cao, Yinzhi Qu, Surabhi Rani, Jinyao Chen, and Shraddha Karanth. Many thanks also go to all the faculty and staffs in the Department of Nutrition and Food Sciences at the University of Maryland. I could not have

completed my graduate studies without their support and everyone has become my Maryland family.

I would like to thank all my friends for their support throughout my graduate study, I deeply valued their friendship and appreciated all the time, support, and care they provided to me.

My sincere thanks to my parents and my parents-in-law for their constant support, encouragement, understanding, and unconditional love throughout the years. Special thanks to my sister-in-law and her husband for their help and support. Many thanks to my wonderful wife Yangyang Wang for her love, support, and always believing in me. My thanks also go to my two lovely cats Candy and Wukong for being cute and silly, and cheering me up.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIC	Alkaike Information Criterion
APC	Aerobic Plate Count
AUC	Area under the Curve
BHI	Brain Heart Infusion
BLEB	Buffered *Listeria* Enrichment Broth
CART	Classification and Regression Trees
CDC	Centers for Disease Control and Prevention
CMREC	Central Maryland Research and Education Center Facility
CPH	Cox Proportional Hazards
CT	Classification Tree
EPA	U.S. Environmental Protection Agency
FDA	U.S Food and Drug Administration
GAP	Good Agricultural Practice
HUS	Hemolytic Uremic Syndrome
KM	Kaplan-Meier
LESREC	Lower Eastern Shore Research and Education Center
LGMA	Leafy Green Products Handler Marketing Agreement
LOD	Limit of Detection
LR	Logistic regression
NBR	Negative Binomial Regression
NCDC	National Climatic Data Center

NOP	National Organic Program
OR	Odds Ratio
OXA	Oxford Agar
PC	Principal Components
PCA	Principal Components Analysis
PW	Peptone Water
RF	Random Forest
ROC	Receiver Operating Characteristic
RR	Relative Risk
RT	Regression Tree
TC	Total Coliform
UMD	University of Maryland
VIF	Variable Inflation Factor

Chapter 1 Introduction and Literature Review

1.1 Fresh produce: growing supply and consumption

Produce (fruits and vegetables) is an important part of a healthy and nutritious diet, which provides important vitamins, minerals, phyto-nutrients, and serves as an essential source of antioxidants and dietary fiber that are very beneficial for weight loss (1). The Dietary Guidelines for Americans have been continuously highlighting vegetables and fruits as important components of a healthy eating pattern in key recommendations (2). As a result, there are growing interests of pursing a healthy diet and lifestyle among consumers, which in turn stimulate the growth of fresh produce supply and consumption over the past few decades in the U.S. For example, total supply of fresh vegetables have reached 66,628 million lbs. (30,222 kg) in 2014 – a 17% increase since 1994 (3), and in-home consumption of fresh vegetables and fruits continues to grow from 2004 to 2014 (4) (**Figure 1.1**). In the U.S., fresh produce production is a diverse industry with a variety of products, each with a specific system of production and handling, and the final value of fresh produce sold through all marketing channels is estimated to be over $122.1 billion in 2010 (5).

1.2 Foodborne illness burden in the U.S

Foodborne diseases have been a major food safety concern in the United States as foodborne disease outbreaks continue to occur and have been a significant threat to public health. In the U.S., the Centers for Disease Control and Prevention (CDC) estimated that each year 48 million people get sick, 128,000 are hospitalized, and 3,000

die due to foodborne diseases (6, 7). These foodborne diseases also result in a huge economic burden to the society with an estimated economic loss of $77.7 billion each year (8)

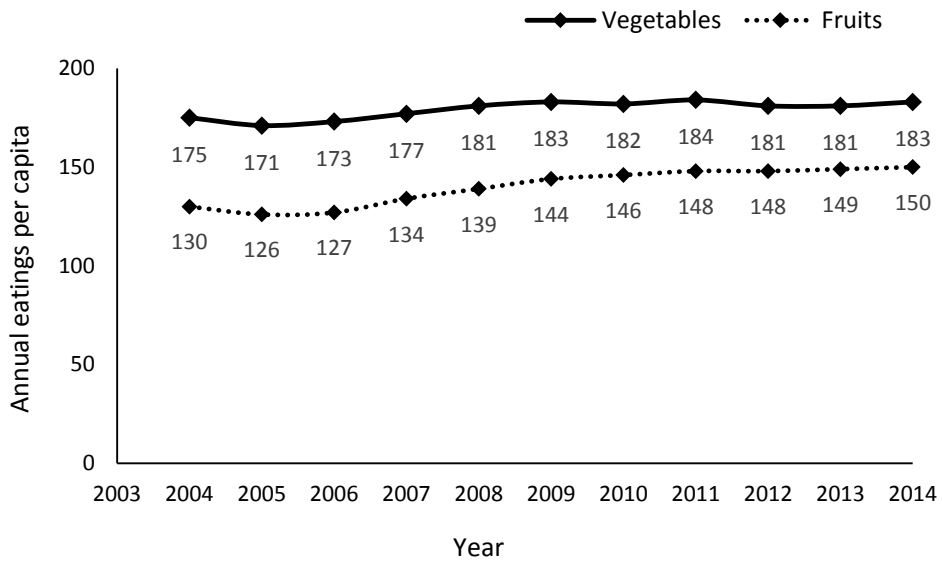

Figure 1.1 In-home consumption of fresh vegetables and fruits, 2004-2014 in the U.S.

1.3 Risk associated with produce

Despite its growing demand and popularity among consumers, fresh produce can serve as a vehicle for foodborne pathogens and has been associated with a number of outbreaks and recalls in recent years in the U.S. (Table 1.1). Produce-related outbreaks have resulted in significant numbers of illnesses, hospitalizations, and deaths in the U.S. - produce commodities accounted for the most foodborne illnesses (46%) during 1998-2008 (9). Despite having a non-pathogenic microbiota, produce can become contaminated during different stages of the farm-to-fork continuum (production field, harvesting, processing, packaging, transportation, handling, and retail/home storage) from a variety of sources. As fresh produce is generally consumed

raw without any heating process as a pathogen kill step, the presence and persistence of pathogens in produce during the farm-to-fork production and supply chain represent a significant public health risk of causing diseases.

Table 1.1 Multistate foodborne outbreaks related to produce in recent years (2011-2016) in the U.S.

Year	Pathogen	No. of illnesses	Produce type	Reference
2011	*Salmonella*	20	Cantaloupe	(10)
2011	*Escherichia coli* O157:H7	58	Romaine lettuce	(11)
2012	*Escherichia coli* O157:H7	33	Spinach	(12)
2012	*Salmonella*	127	Mango	(13)
2012	*Salmonella*	261	Cantaloupe	(14)
2013	*Escherichia coli* O157:H7	33	Ready-to-eat salad	(15)
2013	*Salmonella*	84	Cucumber	(16)
2014	*Salmonella*	115	Bean sprouts	(17)
2014	*Listeria monocytogenes*	5	Bean sprouts	(18)
2015	*Listeria monocytogenes*	35	Caramel apple	(19)
2016	*Listeria monocytogenes*	19	Packaged salad	(20)
2016	*Listeria monocytogenes*	9	Frozen vegetable	(21)

1.4 Emerging foodborne pathogens in produce

As shown in **Table 1.1**, *E. coli* O157:H7, *Salmonella* and *L. monocytogenes* are the main pathogens associated with produce related foodborne disease outbreaks in the U.S. *E. coli* O157:H7 is capable of causing an acute gastrointestinal disease - hemorrhagic colitis characterized by abdominal pain and bloody diarrhea (22). *E coli* O157:H7 can also cause hemolytic uremic syndrome (HUS), a potentially life threatening sequelae characterized by renal failure. It is estimated that *E. coli* O157:H7 is annually responsible for 63,153 cases of foodborne illnesses, 2,138 hospitalizations, and 20 deaths in the U.S. (6). Traditionally, *E. coli* O157:H7 emerged as a human pathogen that has been linked to foods from animal origin; the consumption of

undercooked meat product has been implicated in many foodborne outbreaks (23). More recently, leafy green vegetables have been implicated in *E. coli* O157:H7 outbreaks. For example, in 2006, a major multistate outbreak of *E. coli* O157:H7 was linked to spinach; the outbreak resulted in 199 illnesses, 102 hospitalizations and 3 deaths (24). The *E. coli* O157:H7 and leafy greens pathogen-commodity pair ranks first in the risk-ranking of fresh produce in the U.S. in one recent study in 2011 (25).

The genus *Salmonella* is divided into two species *S. enterica* and *S. bongori*, of which, *S. enterica* is the greatest health concern. *Salmonella* can cause gastrointestinal (non-typhoidal) illness (such as nausea, vomiting, diarrhea, cramps, and fever) and typhoidal illness (high fever) which is a more serious condition (22). In the U.S., *Salmonella* was estimated to cause 1,027,561 cases of non-typhoidal illnesses and 1,821 cases of typhoid fever each year. *Salmonella* was also a leading cause of produce related foodborne disease outbreaks with 32 outbreaks resulting in 1,447 illnesses and 84 hospitalizations in the U.S during 1973-2012 (26).

Listeria monocytogenes is a major food safety concern as it is one of the leading cause of death from foodborne diseases, and is estimated to cause 1,591 illnesses and 1,455 hospitalizations each year in the U.S. *Listeria monocytogenes* is ubiquitous in natural environments (e.g., soil and decaying vegetation) and has been linked to a number of produce-related outbreaks and recalls (27, 28). For example, in 2011, a *L. monocytogenes* outbreak associated with cantaloupe resulted in 147 illness cases, 143 hospitalizations, and 33 deaths across 28 states in the U.S. (27).

1.5 Pre-harvest contamination sources and risk factors

Contamination of produce with pathogenic microorganisms can occur at any stage during the production and supply chain including in-field production, harvesting, post-harvest handling, processing, and distribution. Recent investigations of produce-related outbreaks have suggested contamination at pre-harvest level as a possible source of contamination (29, 30). At pre-harvest, enteric pathogens such as *E. coli* O157:H7 and *Salmonella* can be shed into the environment where they could survive and persist. In addition, *L. monocytogenes* are frequently found in soil. The pathogens persisted in the environment can potentially contaminate produce grown on the fields. The introduction, survival, and persistence of pathogens in produce pre-harvest environment can be affected by a variety of factors including but not limited to meteorological factors (e.g., temperature, precipitation, and wind speed), environmental factors (e.g., landscape, adjacent land use, animal activity), and farm management factors (e.g., manure application, irrigation, cover cropping).

1.5.1 Wildlife and livestock activities

A variety of different pathogens commonly associated with produce have been identified from the waste of domestic/wild animal. These pathogens can be shed into the environment through feces of animals. In fields when produce is grown, deposited fecal waste can be a direct source of contamination. *E. coli* O157:H7 has been found in many domestic and wild animals, with cattle being the major reservoir and up to 50% of the cattle may shed *E. coli* O157:H7 with population levels at approximately 3.3 log CFU/g in feces (31, 32, 33). In addition to cattle, *E. coli* O157:H7 was also frequently

isolated from feces of many other livestock and wild animals, including poultry, sheep, goats, deer, and feral pigs (34, 35, 36, 37).

In addition to direct shedding from feces, pathogens from wild/domestic animals could also be introduced and contaminate produce in the fields indirectly through various routes and vehicles. One possible route is through rodents, insects, or birds, who serve as carriers of pathogens. *E. coli* O157:H7 have been isolated from rats, flies, and wild birds on or near animal farms (38, 39, 40). These carriers may acquire pathogens from feces of infected hosts from the animal farms nearby and subsequently produce grown on the fields may become contaminated if introduced by these carriers.

1.5.2 Irrigation

Irrigation is an essential part of produce production. Water used for irrigation can be drawn from various sources such as rivers, lakes, rainwater, groundwater captured in wells, reclaimed wastewater or potable sources (41). Pathogenic bacteria including *Salmonella* and *E. coli* O157:H7 have been identified in irrigation water from various water sources (42). The microbiological quality of irrigation water is a major influential factor for potential contamination in produce as bacterial pathogens in irrigation water can be transmitted to produce through direct contact of contaminated water during irrigation.

In the U.S., over 50% of the farms use ground water from wells as their primary source of irrigation (43). Although enteric pathogens are generally less prevalent in groundwater due to the filtering mechanism of soil, groundwater can potentially become contaminated through sources including latrines, septic tanks leach fields, land application of wastewater for irrigation, oxidation ponds, leaking sewer lines, and

unlined landfills (44). In general, the microbiological quality of ground water is influenced by depth (from pathogen sources) to the groundwater and improves with distance below surface (41, 44). Water from deep well normally has good microbiological quality due to the longer distance from surface to ground water table that increases the travel time for pathogens to die off and/or be filtered before reaching the ground water system (45). However, it has been suggested that pathogens might be present in shallow aquifers and wells (46).

Irrigation water from surface sources (such as streams, rivers, lakes, and ponds) is a potential source and route of microbial contamination in produce, particularly those located in proximity to livestock, wildlife habitat, humans and their wastes (47). As animals can be attracted by open water bodies, animal feces are the main source of pathogens in irrigation water drawn from surface water. Grazing cattle and livestock production can affect the quality of surface water and has been associated with contamination of pathogens in a variety of produce (48, 49, 50). Surface water can also become contaminated due to runoff from animal farms, manure lagoons, and pasture lands, discharge of sewage water, and leakage from defective septic systems (41, 51). The microbiological quality of surface water can be affected by weather and/or climatic conditions. According to one analysis, about 50% of waterborne outbreaks occurred as a consequence of heavy rainfall (52). Severe climatic events such as flooding also have major impact on the quality of surface water (41, 45).

The type of irrigation system (overhead sprays, drip irrigation systems or flooding of fields through furrows) is another influential factor on the transmission of pathogens from irrigation water to produce. In the U.S., types of irrigation systems

include gravity system (e.g., furrow irrigation, flood irrigation), sprinkler system (overhead irrigation), and low-flow irrigation (drip, trickle, or micro sprinklers) (43). Transfer of *E. coli* from contaminated water to lettuce was reported to occur at a greater rate on plants irrigated by flooding of furrows compared to irrigation through drip system (53). Higher levels of *Salmonella* population were found in cantaloupes and iceberg lettuce irrigated through furrow irrigation compare to surface drip irrigation methods (54). Additionally, higher number of internalized *E. coli* O157:H7 cells were observed when small droplets were applied to spinach leaves other than with mist spraying (55). Subsurface irrigation generally lowers the risk of pathogen transfer to produce crops while overhead irrigation can result in extensive contamination of pathogen in produce due to direct contact between irrigation water and crop surfaces (56).

1.5.3 Manure application

Manure is used widely in produce production to provide organic matter and nutrients. A recent study conducted in Colorado and Texas shows that 795 out of 955 surveyed farmers use manure, and 60% of them use cattle manure (57). Manure is a known reservoir for pathogens and manure application is a possible route of microbial contamination in produce growing in the fields. Bacterial pathogens such as *E. coli* O157:H7 and *Salmonella* have been isolated from in a variety of animal manures. In cattle manure, reported prevalence of pathogenic *E. coli* ranged from 0.7% to 27.8% (58, 59, 60, 61, 62). For poultry manure, prevalence of *Salmonella* range from 8 to 88% (63). Despite the high variability among studies which may be due to regional or

seasonal variation, it is clear that manure can serve as carrier for pathogens and is a source for subsequent contamination in produce.

The survival and persistence of pathogens in manure and manure-amended soils pose a risk of subsequent pathogen contamination of produce. Enteric pathogens have the ability to survive for extended periods in manure (64), and their survival is influenced by a number of factors, including pH, fiber content, temperature, microflora, and aeration (41, 65, 66, 67, 68). Similarly, enteric pathogens can survive for extended periods in manure-amended soils and their survival is affected by a combination of abiotic and biotic factors including temperature, soil microflora, nutrient availability, microbial diversity, and clay content (66, 68, 69, 70, 71).

Pathogens in manure and manure amended soils may colonize plants. Pathogens were found to be present and persist on produce leaves after growing on manure amended soil inoculated with high levels of pathogens (72, 73). Potential internalization of enteric pathogens into the tissue of leaves has also been reported (66, 74, 75). In addition, one study indicates that pathogens present in soil may be transferred to produce leaves thorough harvesting tools (76).

1.5.4 Meteorological factors

Climate and weather conditions are being recognized as influential factors that might be correlated with the incidence and distribution of foodborne diseases (77, 78). Local weather factors such as temperature and precipitation affect the growth and survival of foodborne pathogens in the environment, and also their transmission between produce and environmental reservoirs such as irrigation water and field soil at the pre-harvest level. Understanding how climate/weather factors affect the

introduction and persistence of pathogens in produce production is critical to control potential contamination in produce at pre-harvest level.

Temperature is one of the most important factors affecting the growth and survival of microorganisms. Warmer temperatures favor of the growth of many human pathogens including *E. coli* O157:H7 in environmental reservoirs. Higher prevalence and/or concentration of pathogens have been observed in surface water during warmer months (79, 80, 81). Warmer temperatures are also closely correlated with the survival and growth of microorganisms in manure and manure amended soil. *Salmonella* and *E. coli* levels were higher in soil when manure was applied during warm temperature months(>20°C), compared to application in cold months (<10°C) (72). Higher prevalence of *E. coli* and other bacterial indicators in tomatoes and leafy vegetables were also observed when samples were collected during warmer seasons (82, 83, 84). Conversely, low temperature and temperature fluctuation may have an inhibitory effect on microbial growth and survival. Freeze and freeze-thaw cycles encountered during winter are a major abiotic stress for soil microorganisms and studies have shown significant decrease in soil microbial density result from freeze-thaw cycling (85).

Temperature may also affect the survival and dissemination of pathogens in natural environment indirectly. Temperature is also closely related to insects and pests activity. Increasing activity of insects and pests in warmer seasons in and around produce farms may lead to transfer of human pathogens to farms where produce is grown (41, 86). Temperature change may also lead to increased susceptibility of livestock to animal diseases, which might increase the colonization of enteric pathogens in animal gut (77, 78).

Higher rainfall result in high soil moisture content and is supportive of the survival of microorganisms including foodborne pathogens (87). Elevated precipitation also is correlated with increased prevalence and concentration of *Salmonella* (79) and *E. coli* O157:H7 (81) in water. Excessive rainfall may cause runoff from animal farms or composting facilities, which might serve as a vehicle that carries pathogens to distant area. Heavy rainfall may also cause overflow of urban wastewater that contains human pathogens to open water bodies such as wells, pond, and streams or rivers, which may be sources of water used for irrigation. Heavy rainfall may also contribute to dissemination of microorganisms in the environment (80) and transfer of microorganisms from soil to fresh produce due to splashing (88, 89).

Wind is another risk factor for microbial contamination of produce growing on farm. Wind may cause dust storms that bring dust particles onto produce leaves (41). Human pathogens have been reported to be able to survive in dust for up to 26 months and 10 months for *Salmonella* and *E. coli*, respectively (90, 91). Dust and aerosols carrying pathogens have the ability to travel long distances with the help of wind (92, 93).

1.5.5 Landscape and geographical factor

Each produce farm location is a unique combination of landscape and geographical characteristics. Certain landscape and geographical factors may favor the introduction and persistence of pathogens to produce production environment.

The location of produce farms are important due to the potential cross-contamination of fields with nearby environmental contaminants and sources. Farms located downstream from urban and highly industrialized areas are more susceptible to

potential contaminations as urban and industrial waste may run off into streams which influence the water quality and increases the risk of pathogen spread onto produce farms (41, 94). Distance to environmental reservoirs (e.g., open water, animal operations, septic system, and composting facilities) is an important factor affecting the possibility of contamination in produce farms. In general, produce farms located close to animal production are more likely to become contaminated from animals such as cattle or poultry. Direct contamination of crops, or indirect contamination through soil and water could result from runoff from animal production or fecal deposits from wildlife or domestic animals that intrude into produce fields. Animals may be attracted to produce farms for various reasons, including seeking for food or water, seeking shelter, or incidental passage through produce farms (41). This is especially the case for animal farms that lack of fences or buffer zones (e.g., free- range livestock farms).

1.6 Mitigation strategies to reduce microbial contamination risk in produce at pre-harvest level

Composting is an efficient method that could be used in practice to reduce pathogens in manure. The U.S Food and Drug Administration (FDA) has proposed changes to produce rule to encourage use of compost as a safer and more environmentally friendly alternative to raw manure (95). High temperatures are essential for eliminating pathogens. U.S. Environmental Protection Agency (96) suggested that compost should be maintained at above 40°C for five days, and at above 55°C for at least four hours of the five days to achieve significant reductions of pathogens during composting. The National Organic Program (NOP) recommended

that composting temperatures should reach a critical temperature of 55° C for at least 3 days (97).

Pathogen load is influenced by time between manure application and harvesting, manure types, and manure handling strategies (41). Sufficient time intervals between application of raw manure and harvesting can reduce pathogen levels in manure and reduce the risk of potential transmission of pathogens to vegetable crops. A 120-day interval between application of raw manure and harvest of crops in contact with the soil and a 90 day interval for crops not in contact with the soil were required by NOP standards (97). Survival of pathogens in manure has been reported to be shorter where manure had higher fiber content, high pH, high temperature, where temperature fluctuations were large, where high levels of natural microflora existed, where manure was applied subsurface, and where high levels of aeration occurred (41, 65, 66, 67, 68).

Improvement in the microbial quality of irrigation water is an important way to reduce the risk of microbial contamination in produce. Protection of irrigation water sources (e.g., surface and ground well water) from contamination (wildlife, waste from animal production, agricultural run-off, human activity, sewage, industrial effluent) is essential to maintaining good quality. Monitoring (i.e., sampling and testing) of irrigation water microbial quality is needed on a routine basis to ensure the quality of water. Microbiological standards for testing, testing frequencies, and testing strategies should be adjusted with regard to irrigation water sources, season, and locations. California Leafy Green Products Handler Marketing Agreement (LGMA) recommended that concentration of generic *E. coli* in water used for foliar application

(e.g., overhead irrigation) should not exceed 126 CFU/100 ml (based on a rolling geometric mean n=5) with no individual sample exceeding 235 CFU/100 ml (98).

1.7 Project overview

Preventing pre-harvest contamination of produce is crucial as most produce are consumed raw without heating process as a pathogen killing step. Identification of potential risk factors and understanding their effects on the introduction, survival, persistence, growth, and transmission of pathogens in the produce pre-harvest environment is critical to develop farm-level risk mitigation strategies to effectively control pre-harvest contamination of produce. Contamination can be reduced by implementing good agricultural practices (GAPs) and modifying the conditions that are favorable for pathogen survival and transmission at the pre-harvest level.

The overall goal of this study was to develop predictive models to identify and evaluate risk factors for microbial contamination during produce pre-harvest production, and predict the prevalence and concentration of enteric bacteria in produce pre-harvest environment under different weather conditions, geographic regions, and farm management practices. Specific objectives are:

(1) To identify and evaluate meteorological risk factors associated with pre-harvest contamination of *Listeria* species in a mixed produce and dairy farm. Produce from mixed farms are at high risk of microbial contamination due to its proximity to animal operations. This objective will focus on the investigation of the effect of meteorological factors on the prevalence of *Listeria* spp. in a mixed farm setting.

(2) To evaluate meteorological factors associated with prevalence and concentration of generic *E. coli* in a mixed produce and dairy farm. Foodborne outbreaks have been tracked back to fecal contamination of produce at pre-harvest level. Understanding the impact of meteorological factors on the prevalence and concentration of generic *E. coli* as an indicator of fecal contamination is critical in guiding the development of good agricultural practices (GAPs) to reduce the risk of pre-harvest contamination in produce.

(3) To investigate the effects of cover cropping, farming system, and meteorological factors on the survival and persistence of generic *Escherichia coli* and *Listeria innocua* in produce fields. Soil can serve as a potential reservoir and route of pathogen contamination in produce during production in the field. Understanding the factors influencing the survival and persistence of food safety related bacteria in field soil is essential in the development of intervention strategies to prevent possible contamination in produce.

(4) To evaluate the effects of meteorological, farm management, and environmental factors on microbial contamination of tomatoes and the tomato production environment. Microbial contamination in tomatoes at pre-harvest level can be affected by a variety of factors. There is a need for systematic assessments of meteorological, environmental, and farm management factors on their joint effects on the presence and concentration of food safety and quality indicator bacteria in tomato pre-harvest environment.

These four objectives holistically evaluated the influence of a variety of risk factors on the presence and concentration of enteric and indicator bacteria in produce

pre-harvest production environments. Objective 1 investigated the impact of meteorological factors on presence of microorganisms and developed a modeling framework that was used in objectives 2, 3, and 4. Objective 2 extended the analysis in objective 1 by using both prevalence and count data. Objective 3 evaluated the effect of particular farm management practices along with meteorological factors on the survival and persistence of indicator bacteria in different production systems. Objective 4 provided a systematic assessment of meteorological, environmental, and farm management factors at the pre-harvest level that are potentially responsible for contamination of produce.

Chapter 2 Identifying and modeling meteorological risk factors associated with pre-harvest contamination of *Listeria* Species in a mixed produce and dairy farm

2.1 Abstract

This study sought to investigate the prevalence of *Listeria* species (including *Listeria monocytogenes*) in a mixed produce and dairy farm and to identify specific meteorological factors affecting *Listeria* spp. presence. Environmental samples were collected monthly from locations within the mixed farm over 14 months and were analyzed for *Listeria* spp. Meteorological factors were evaluated for their association with the presence of *Listeria* spp. by using logistic regression (LR) and random forest (RF) analysis. The developed LR model identified wind speed and precipitation as significant risk factors (P – value < 0.05), indicating that higher average wind speed 2 days prior to sampling and higher average precipitation over the previous 25 days before sampling increased the probability of isolation of *Listeria* spp. from the mixed farm. Results from RF revealed that average wind speed at day 2 prior to sampling and average precipitation in the previous 25 days before sampling were the most important factors influencing the presence of *Listeria* spp., which supported the findings from LR. These findings indicate that occurrence of *Listeria* spp. was influenced by wind speed and precipitation, suggesting runoff and wind-driven dust might be possible routes of pathogen transmission on mixed farms. The developed LR and RF models, with robust predictive performances as measured by area under the receiver operating characteristic curves, can be used to predict *Listeria* spp. contamination risk in a mixed farm under different weather conditions and can help with evaluation of farm

management practices and development of control strategies aimed at reducing pre-harvest microbial contamination in a mixed farming system.

2.2 Introduction

The natural microflora of fresh produce is generally considered as enteric pathogen-free, as its microflora is composed of plant-associated microorganisms incapable of causing human illness. During different stages of the farm-to-fork continuum (cultivation, harvest, processing, packaging, transportation, handling, and retail/home storage), microbial contamination can occur from a variety of sources and human pathogenic microorganisms could be introduced to produce. As fresh produce is generally consumed raw without any heating as a pathogen-killing step, the presence and persistence of pathogens in produce during the farm-to-fork supply chain represents significant public health risks.

Listeria monocytogenes is a serious food safety concern among consumers and the produce industry. Further, as a psychrotroph, this bacterium can grow under refrigeration conditions. A number of produce related foodborne disease outbreaks and recalls have been linked to contamination with *L. monocytogene*s. As mentioned previously, a *L. monocytogenes* outbreak associated with cantaloupe in 2011 resulted in 147 illness cases, 143 hospitalizations, and 33 deaths across 28 states in the U.S. (27). In 2014, caramel apples contaminated with *L. monocytogenes* sickened 35 people from 12 states; of these, 34 were hospitalized and 3 died due to listeriosis (19). In addition, *L. monocytogenes* was responsible for a recent large recall of frozen produce involving 456 products under 42 separate brands sold nationwide in the U.S. and in

four provinces in Canada (28). Investigations of produce-borne *L. monocytogenes* outbreaks have suggested contamination at produce pre-harvest stages as a potential cause. For example, during FDA's investigation of a 2014 sprouts outbreak, the associated *L. monocytogenes* strain was isolated from environmental swabs collected from a sprouts production environment (99). The investigation of the 2011 cantaloupe outbreak with *L. monocytogenes* presented the possibility that contamination of low level sporadic *L. monocytogenes* in the growing agricultural environment might be one of the sources of this outbreak (100).

Mixed farms, where produce crop production and livestock operations are integrated, are present worldwide (101, 102). In mixed farming systems, produce farmers benefit from animal-derived manures that serve as source of nutrients applied to soil directly or after composting. In turn, crop residues provide an economical source of feed for livestock (103). However, there are potential risks for fresh produce crops grown in a mixed farm setting. Animal feces, manure, and compost are known reservoirs for foodborne pathogens (87). Due to the proximity to these reservoirs, vegetables grown in a mixed farm may be susceptible to pathogen contamination from these sources. Thus, it is crucial to investigate risk factors affecting the presence and dynamics of pathogens within a mixed farming system to prevent such contamination risk.

It is being recognized that climate and weather conditions are possibly related to the incidence and distribution of foodborne diseases(77, 78). Meteorological factors such as temperature and precipitation can affect the growth and persistence of foodborne pathogens, as well as their transport to and within the farm environment.

Previous studies have been conducted to identify meteorological factors affecting microbial contamination in produce farms (57, 104, 105, 106). Precipitation and rain events that may facilitate spread of microorganisms or increase the survivability of microorganisms have been identified as risk factors that increase the presence of pathogens or indicator bacteria in produce farms (57, 105, 106, 107). In addition, temperature was found to be associated with pathogen isolation from produce farms (104, 105). Considering the unique setting of a mixed farm, it is important to investigate the impact of meteorological factors on the presence of pathogens in a mixed farming system. The objective of this study was to identify specific meteorological factors affecting *listeria* spp. Occurrence in a mixed farm facility integrating dairy cow management, cow manure composting, and a vegetable production area. This study sought to provide scientific evidence that can guide the development of science-based good agricultural practices (gaps) to reduce food safety risks in mixed farming systems.

2.3 Materials and methods

2.3.1 Description of study data

A longitudinal study was conducted at a mixed produce and dairy farm at the Central Maryland Research and Education Center Facility (Clarksville, Maryland) over a 14-month period from February 2014 to April 2015. The farm houses a dairy with pasture, a composting facility for dairy manure, and a vegetable production area (**Figure 2.1**). The composting facility receives manure from the barn floor, when bedding, feces and manure are flushed periodically into a pit where the sludge is separated into solids and liquid manure. The liquid manure is directed to a lagoon where

it is allowed to stand before application to nearby fields planted yearly with corn and soybean. The separated solids are transferred to a composting site where the manure is laid out in windrows and allowed to thermophilically compost, with periodic turning, over a period of weeks. Finished compost is piled in a separate heap before transported off the farm. Samples were collected monthly (excluding April and September 2014) from 12 sites within the mixed farm including around the field and pasture, the dairy barn, and the composting facility. The following sample types were collected from around the barn: fresh cow feces from the dairy barn, cow feed, cow drinking water, and bird feces from the ground along the perimeter of the barn, where birds gather to feed on uncovered cow feed. The following samples were collected from the composting facility: raw liquid manure, water from the lagoon (receiving raw liquid manure), raw separated solids, partially composted material (from windrows) and fully composted material from the finished compost heap. Environmental samples collected from around the farm included surface water from a creek, soil from the cow pasture and soil from the vegetable production area. In total, 159 samples were collected.

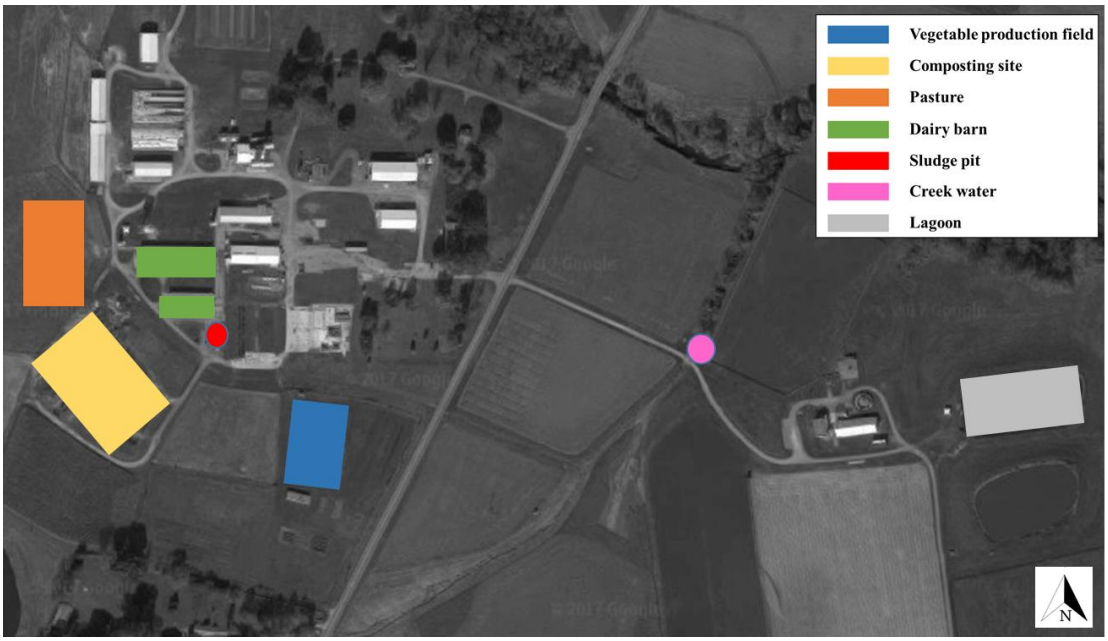

Figure 2.1 Location of the sampling sites and area included in this study.

2.3.2 Sample collection and preparation

All samples were collected in sterile containers and latex gloves were worn for each sample collection. Gloves were disinfected with 70% ethanol prior to sample collection, and changed if soiled, wetted or torn. Approximately 1 liter of water was collected into sterile Nalgene bottles (Thermo Scientific, Rochester, NY). Liquid manure was collected using a grab sample available on site into a sterile Nalgene bottle. Solid samples (soil, compost, feed, and feces) consisted of about 300 g samples collected in sterile WhirlPak bags (Nasco, Fort Atkinson, WI) using sterile scoops (Fisher Scientific, Hampton, NH). All samples were transported in coolers with ice and were transferred to 4°C until processing. All samples were processed within 24 h of collection.

2.3.3 Sample Processing and *Listeria* spp. isolation

For solid samples (soil, feces, dry manure, compost, and cow feed), 10 g of sample were diluted 1:10 (wt/vol) with Buffered *Listeria* Enrichment Broth (BLEB) (EMD Chemicals Inc., Darmstadt, Germany), vortexed for 2 min at high speed and incubated at 30 ± 2°C. After 4 h incubation, nalidixic acid, cycloheximide and acriflavine were added as recommended by the manufacturer and re-incubated at 30°C for 20 h. Suspensions were streaked onto Oxford Agar (OXA) plates (Becton Dickenson and Company (BD), Franklin Lakes, NJ) and incubated at 35°C for 48 h. Up to 10 presumptive black colonies of *Listeria* spp. were picked and streaked for isolation and archived in Brucella Broth (BD) with 15% glycerol, by storing at -80°C until further identification confirmation. Up to 500 ml of liquid samples were filtered (depending on turbidity) through 0.45 µm mixed cellulose ester filters (Millipore, Billerica, MA) using a PALL filtration system (PALL Life Sciences, Ann Arbor, MI). Filters were vortexed and enriched in BLEB, incubated overnight then streaked onto OXA plates and processed as described above.

2.3.4 Confirmation of *Listeria* spp. and *L. monocytogenes* identification

Presumptive *Listeria* spp. isolates were streaked from frozen stock onto Brain Heart Infusion (BHI) agar (BD) plates and grown at 30 °C. DNA was extracted from one colony from each culture using a quick lysis method by suspending the cells in a 7.5% Chelex® 100 resin (Bio-Rad, Hercules, CA) solution and heated for 10 minutes at 105°C, then centrifuged for 1 min at 8,000 *g*, as previously described (108). Genomic DNA obtained from each isolate was subjected to PCR amplification with primers specific to a 350 bp portion of the Bacterial 16S rRNA gene as described in Micallef et

al., (108) to ensure successful DNA extraction. Confirmation of the *Listeria* genus was done using *Listeria*-specific primers UnilisA (5´-GCTACAGCTGGGATTGCGGT-3´) and Lis1B (5´-TTATACGCGACCGAAGCCAA-3´) (109). For identification of *L. monocytogenes*, two genes were targeted; the hemolysin A gene (*hlyA*) using primers A1 5´-GCAGTTGCAAGCgCTTGGAGTGAA-3´ and A2 5´-GCAACGTATCCTCCAGAGTGATCG-3´ (110) and the invasion-associated protein gene (*iap*) with *iap*F 5´-AATCTGTTAGCGCAACTTGGTTAA-3´ and *iap*R 5´-CACCTTTGATGGACGTAATAATACTGTT-3´ (111). PCR was carried out in a total volume of 20 µl containing 1 µl of 10× Standard Taq Reaction Buffer (BioLabs Inc., New England), 0.8 U Taq DNA polymerase (BioLabs), 2 mm $MgCl_2$, 0.4 µm of each dNTP (BioLabs), 0.4 µm of each reverse and forward primer and 50-100 ng pure DNA. The remaining volume was adjusted by adding sterile ultrapure water as needed. DNA was amplified in a C1000Touch™ Thermal Cycler (BIO RAD, Singapore). Initially, DNA denaturation was carried out at 95°C for 2 min followed by 30 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 30 s and extension at 72°C for 45 s, with a final extension step of 10 min. Amplified DNA fragments were analyzed on 1% (w/v) agarose gel (Lonza, Rockland, ME) in Tris–borate–EDTA buffer (BIO-RAD). The amplified DNA fragments were visualized using a Molecular Imager Gel Doc™ XR+ with Image Lab™ Software (BIO-RAD). The size of DNA fragments was established from molecular weight markers included in each gel.

2.3.5 Meteorological data

For each sample collection date, meteorological variables were obtained from the weather station located at the Central Maryland Research and Education Center

Clarksville Facility. In total, 102 different meteorological factors were obtained, including ambient temperature (maximum, minimum, and daily average), precipitation, and wind speed (daily average and maximum). Specifically, maximum, minimum, and daily average temperature, precipitation, and maximum and average daily wind speed on the day of sampling and on day 1 and day 2 prior to sampling day were obtained, representing the instant effect of weather variables. In addition, average levels of the weather variables between sampling and from 1 day to up to 30 days prior to sampling day were obtained to capture possible long term effect of those meteorological variables (104).

2.3.6 Statistical analysis and modeling

All statistical analyses were performed using R (version 3.2; R Core Team, Vienna, Austria). Prevalence of *Listeria* spp. was calculated for each sample type and season (spring, summer, fall, and winter). Fisher's exact tests were used to compare *Listeria* spp. prevalence across seasons. *P* values less than 0.05 were considered statistically significant. The Holm-Bonferroni multiple comparison correction was used to assess significance.

Data from all sample types were used to determine the association between meteorological factors and presence of *Listeria* spp. within the mixed produce and dairy farms. Each individual meteorological variable was first evaluated by univariate analyses using an arbitrary significance level at 0.2 to include all possible influential variables in multivariable analyses (112, 113). In univariate and multivariable analyses, logistic regression (LR) models were developed using "glm" function to determine associations between meteorological factors and the presence of *Listeria* spp. in the

mixed farm. Correlations among significant variables by univariate analyses were assessed by Spearman's rank coefficients. When two or more significant variables considered for multivariate analyses were highly correlated (correlation coefficient > 0.70), variables were considered one at a time in multivariate analysis and the one gave the best model fit was chosen in the final model.

The final multivariable model was build using a backward selection method based on the Alkaike information criterion (AIC) until only significant variables were retained ($P < 0.05$). In the final model, the assumption of a linear relationship between continuous explanatory variables and logit transformation of outcome (log odds) was assessed using Box-Tidwell transformation by adding an interaction term between the explanatory variable and its natural log to the model. The goodness of fit of the final model was assessed by Pearson and deviance Chi-square test and the Le Cessie-van Houwelingen-Copas-Hosmer test. Possible collinearities among explanatory variables in the final model was investigated by calculating the variance inflation factors.

Random forest (RF) was used as an alternative approach to regression method to determine meteorological variables that were associated with presence of *Listeria* spp. and to predict the probability of presence of *Listeria* spp. under different weather conditions. All 102 collected meteorological variables were analyzed in the RF model as predictor variables. The RF model was built in R using the "randomForest" package with 10,000 bootstrap samples and 2 randomly selected variables at each node. The relevance of meteorological variables for presence of *Listeria* spp. in a mixed farm was illustrated by variable importance score (measured by decrease in node impurity) where

large variable importance scores indicates variables that were highly associated with presence of *Listeria* spp.

Meteorological variables tend to be correlated especially among those in the same category (i.e., temperature, precipitation, or wind speed). Thus, principal component analysis (PCA) was performed when significant variables from univariate analysis were correlated. The number of components to retain was determined by considering both the percentage of variance explained and interpretability (114). Retained components should explain at least 80% of the total variance and retained components should have at least three variables with major loadings (correlation coefficients between variables and principle components) on each retained component, the same conceptual meaning among the variables loading on the same component, and simple structure of the rotated pattern with relatively high factor loading of a variable on only one component and relatively small loading on other components. Component scores for each retained components were calculated, and were used as new explanatory variables in multivariable analysis.

To evaluate the robustness of predictive performances, 5-fold cross-validations were conducted for the developed LR and RF models: the whole data set was randomly divided into five subsets with equal sizes and then nine subsets were used as training sets in LR and RF models to generate model coefficients while the remain subset was used as a test set to assess model performances. This process was repeated 5 times, each time with a different subset as test set. In addition, the whole data set was also used for testing models' predictive performance as a way of internal validation. The

area under the curve (AUC) from the receiver operating characteristic (ROC) curves was used as a measurement of predictive performances.

2.4 Results

Overall, *Listeria* spp. was detected from 10 out of 156 of the total samples collected from the farm. Positive samples occurred in March, May and June 2014, and March and April 2015 (**Table 2.1)**. Only 2 out of 10 *Listeria* spp.-positive samples were confirmed for *L. monocytogenes.* Positive sample types were cow feed, fresh cow pies, raw separated solid manure, windrow and finished compost, pasture soil and bird feces. Sample type was not a significant factor associated with prevalence of *Listeria* spp. ($P = 0.13$) while season was significantly associated with prevalence of *Listeria* spp. ($P < 0.001$). The prevalence of *Listeria* spp. in all farm samples was significantly higher in spring than in other seasons (**Table 2.2**).

In the univariate analyses, 19 out of 102 meteorological variables were significantly associated with the presence of *Listeria* spp. within the farm, including 15 precipitation variables and 4 wind speed variables, but none of the temperature related variables were significant (**Table 2.3**). From the univariate analyses, the odds of *Listeria* spp. presence within the mixed farm significantly increased with increasing average and maximum wind speed two days prior to sampling day. Odds of *Listeria* spp. isolation also increased significantly when the farm was exposed to higher average amount of rainfall within up to 30 days prior to sampling day. These odds indicate that higher precipitation and higher wind speed were potential risk factors associated with increasing prevalence of *Listeria* spp. within the mixed farm. In addition, while the

effect of wind speed on the presence of *Listeria* spp. seemed to be instant, precipitation may have a cumulative, long-term effect on the presence of *Listeria* spp. within the farm.

Table 2.1 Frequency of samples positive for *Listeria* spp. isolated from pre-harvest environments of a mixed produce and dairy farm.

Sample type	No. of Samples	No. (%) of positive samples[a]	
		Listeria spp.[b]	*L. monocytogenes*
Cow feed	14	2 (Mar 14, Apr 15)	1 (Apr 15)
Cow drinking water	14	0	0
Cow Pie	12	1 (May 14)	0
Raw Separated Solid	12	1 (Jun 14)	0
Raw Liquid Manure	14	0	0
Lagoon Water	14	0	0
Windrow Compost	14	1 (Mar 15)	0
Finished Compost	14	3 (May 14, Mar 15, Apr 15)	0
Bird Feces	9	1 (Mar 14)	0
Organic Field Soil	14	0	0
Pasture Soil	14	1 (Apr 15)	1 (Apr 15)
Creek Water	14	0	0
Total	159	10	2

[a]Different letters represent values that are significantly different ($P < 0.05$) by Fisher's exact test.

[b]Excluding *L. monocytogenes*

The final multivariable model was comprised of two significant variables: average precipitation over the previous 25 days before sampling, and average wind speed at day 2 before sampling (**Table 2.4**). Wind speed was identified as a risk factor, as odds of isolation of *Listeria* spp. increased with each 1 m/s increase of average wind speed at day 2 before sampling (odds ratio [OR] = 13.5). Increasing precipitation also

increased the odds of isolation of *Listeria* spp.: for each 1 mm increase of average precipitation over the previous 25 day before sampling, the odds of isolation of *Listeria* spp. increased (OR= 7.4). Probability of isolation of *Listeria* spp. can be estimated using the LR model as the function of the two predictors, and it increases with increasing precipitation in the previous 25 days and increasing wind speed in the previous 2 days (**Figure 2.2**). The final LR model showed solid predictive performance as indicated by internal validation (AUC = 81%) and cross validation (mean AUC = 85%, range: 74% to 97%) (**Figure 2.3A**).

Table 2.2 Effect of season on frequency of samples positive for *Listeria* spp. isolated from pre-harvest environments of a mixed produce and dairy farm.

Sample category	No. of samples	No. (%) of positive samples[a]	
		Listeria spp.	*L. monocytogenes*
Season			
Spring	44	9 (20) A	2 (5%)
Summer	34	1(3) B	0
Fall	36	0 B	0
Winter	45	0 B	0
Total	159	10	2

[a]Different letters represent values that are significantly different ($P < 0.05$) by Fisher's exact test.

Table 2.3 Associations between meteorological variables and prevalence of *Listeria* spp. based on the univariate logistic regression models.

Description	OR[a]	95%CI[b]	*P*
Precipitation at day 1 prior to sampling day	1.06	1.00, 1.12	0.05
Precipitation at day 2 prior to sampling day	1.21	1.02, 1.40	0.02
Average precipitation over the previous 1 days before sampling	1.11	1.00, 1.25	0.08
Average precipitation over the previous 2 days before sampling	1.19	1.01, 1.38	0.03
Average precipitation over the previous 3 days before sampling	1.31	1.06, 1.63	0.01
Average precipitation over the previous 4 days before sampling	1.40	1.10, 1.80	0.01
Average precipitation over the previous 5 days before sampling	1.47	1.04, 2.00	0.01
Average precipitation over the previous 7 days before sampling	1.40	0.95, 2.05	0.08
Average precipitation over the previous 8 days before sampling	1.47	0.91, 2.32	0.10
Average precipitation over the previous 9 days before sampling	1.50	0.84, 2.63	0.15
Average precipitation over the previous 10 days before sampling	1.98	1.01, 4.08	0.05
Average precipitation over the previous 15 days before sampling	1.71	1.06, 2.80	0.03
Average precipitation over the previous 20 days before sampling	1.81	0.91, 3.56	0.08
Average precipitation over the previous 25 days before sampling	2.35	1.02, 5.71	0.04
Average precipitation over the previous 30 days before sampling	1.41	0.94, 1.99	0.06
Average wind speed at day two prior to sampling day	2.99	1.51, 6.18	0.00
Average wind speed between sampling day and 2 days before sampling	4.14	1.24, 15.87	0.03
Maximum wind speed at day 2 prior to sampling day	1.30	1.05, 1.62	0.02
Average of the maximum wind speed over the previous 2 days before sampling	1.45	1.01, 2.18	0.05

[a]OR, odds ratio.
[b]CI, confidence interval

Table 2.4 Final multivariable model for the likelihood of isolation of *Listeria* spp. from the mixed farm.

Factor	OR[a]	95% CI[b]	*P* value
Average precipitation over the previous 25 days before sampling	1.49	1.18, 1.92	0.001
Average wind speed at day 2 before sampling	4.02	2.28, 7.49	<0.001

[a]OR, odds ratio.
[b]CI, confidence interval.

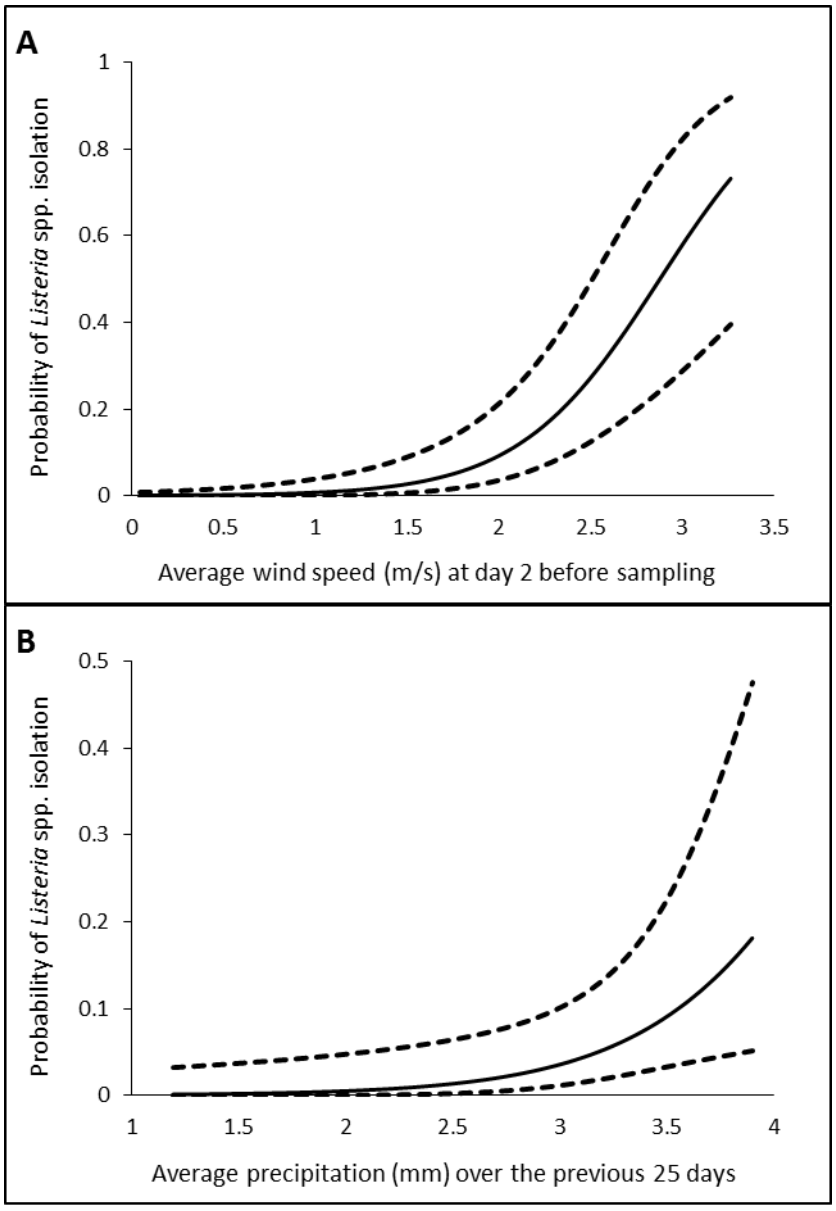

Figure 2.2 Predicted probabilities of *Listeria* spp. isolation (solid lines) and confidence interval (dashed lines) from sampling locations within the mixed farm for different values of average wind speed at day 2 before sampling (A) and average precipitation over the previous 25 days before sampling (B).

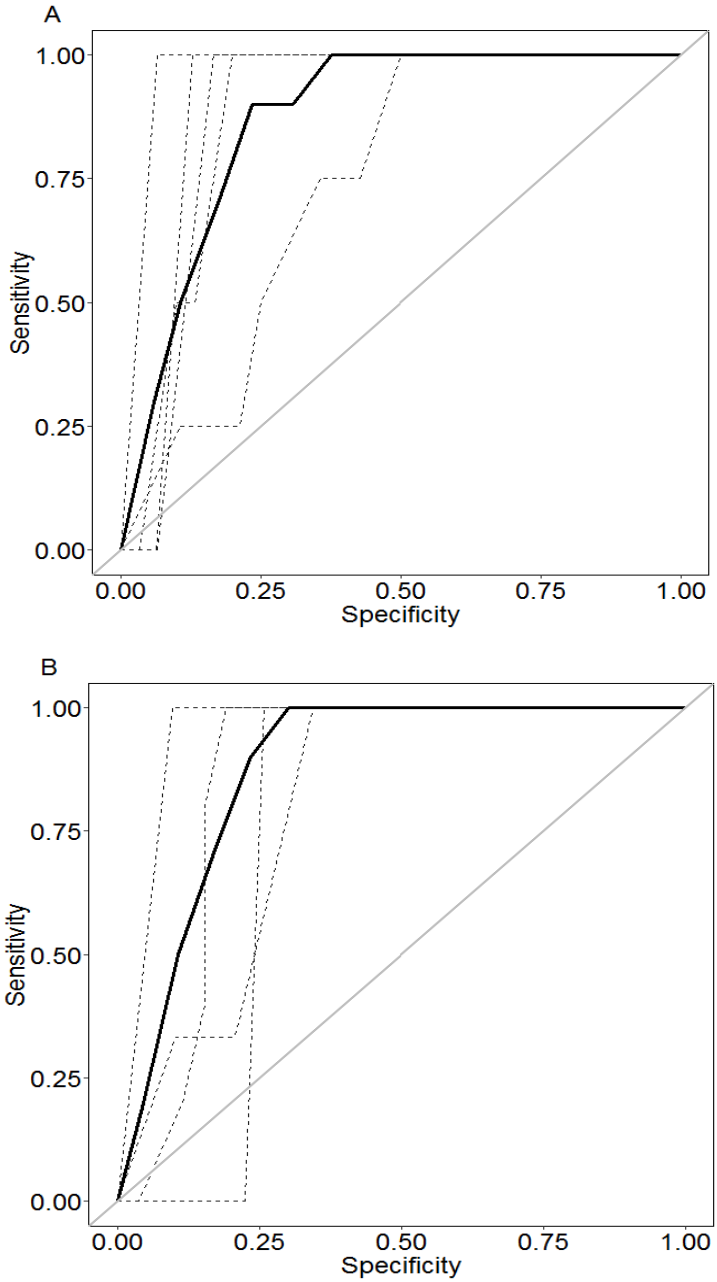

Figure 2.3 Receiver operating characteristic (ROC) curves for 5-fold cross validation (dashed line) and internal validation (solid line) of the developed logistic regression model (A) and random forest (B). The diagonal line is the line of no discrimination (represents completely random guess).

Fifteen out of 17 precipitation variables were significant from univariate analysis and were highly correlated. Thus, all 17 precipitation variables were subjected to PCA. Based on our aforementioned selection criteria, the first three principal components were selected, explaining 89% of the total variability. Variables describing average precipitation over up to 10 days prior to sampling day loaded on the first component, denoted as short-term cumulative effect component. The 3 variables describing the daily precipitation on sampling day or up to two days prior to sampling day loaded on the second component, denoted as instant effect component. The third principal component had major loadings of 4 variables describing average precipitation over 15 to 30 days prior to sampling, and was denoted as long-term cumulative effect component. When principal component scores of the three retained precipitation components were used in multivariable modeling, the final model consisted of three variables: average wind speed at day 2 before sampling, the first component (short-term cumulative effect component), and the third component (long-term cumulative effect component).

The variable importance plot representing the rank of variables with largest variable importance scores is shown in **Figure 2.4**. Precipitation and wind speed factors were highly associated with *Listeria* spp. isolation from the mixed farm, each with 7 factors ranked at the top 15 variable importance scores. Maximum temperature at day 2 prior to sampling ranked 13^{th} among all variables and was the only temperature factor ranked top 15 based on the developed RF model. Average precipitation in the previous 25 days prior to sampling and wind speed at day 2 prior to sampling were the top 2 factors with the highest variable importance scores, indicating their influences in

isolation of *Listeria* spp. To investigate the effect of these two factors, partial dependence plots were generated, which showed the dependence between the outcome and one or more predictor variables, marginalizing over the values of all other variables (**Figure 2.5**). The partial dependence of probability of isolation of *Listeria* spp. on wind speed at day 2 before sampling was monotonic increasing. Probability of *Listeria* spp. generally increased with increasing average precipitation over the previous 25 days before sampling, except when average precipitation was less than approximately 1.5 mm. **Figure 2.6** shows the two-variable partial dependence of probability of isolation of *Listeria* spp. on joint values of average precipitation over the previous 25 days before sampling and average wind speed at day 2 before sampling. Generally, both variables had positive relationships with the outcome, such that increased values in each variable corresponded to an increase in the predicted probability of isolation of *Listeria* spp. As both variables have consistent relations with predicted probabilities across the values of the other variable, there is no substantial evidence for an interaction between these two variables. The developed RF model showed an AUC of 87% from internal validation and a mean AUC of 85% (range: 76% - 93%) from cross validation (**Figure 2.3B**).

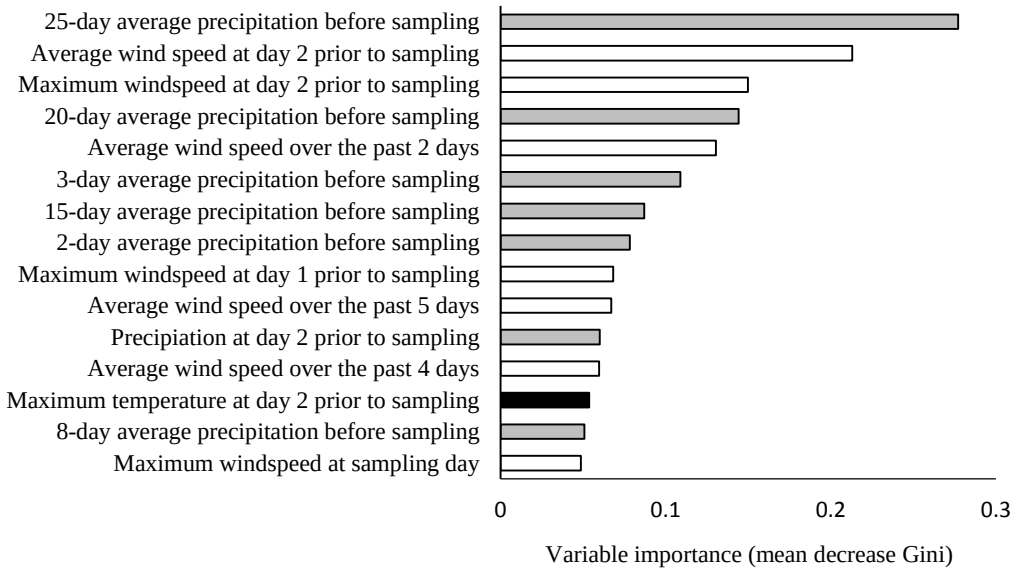

Figure 2.4 Variable importance plot for each meteorological factors based on RF analysis. Grey bars represent precipitation factors; white bars represent wind speed factors; black bars represent temperature factors.

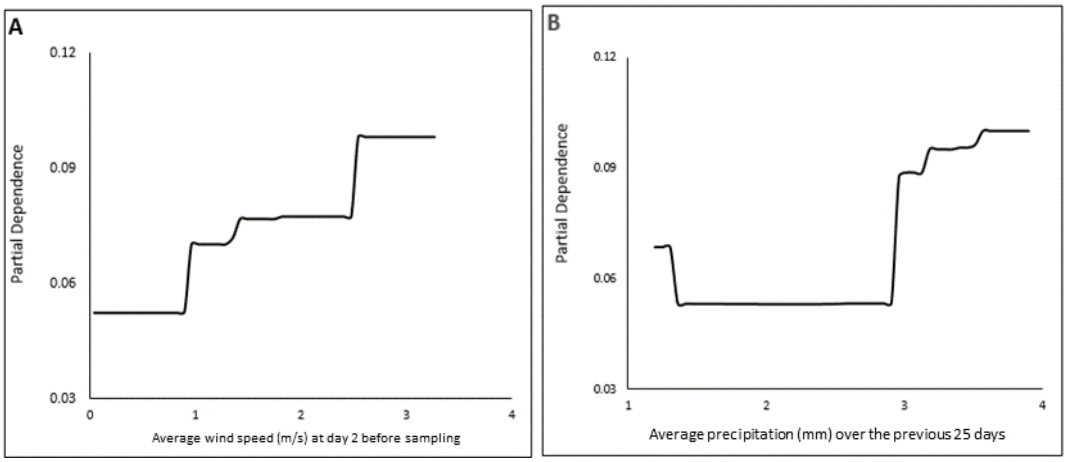

Figure 2.5 Partial dependence of probability of isolation of *Listeria* spp. on wind speed at day 2 prior to sampling (A) and on average precipitation over the previous 25 days (B).

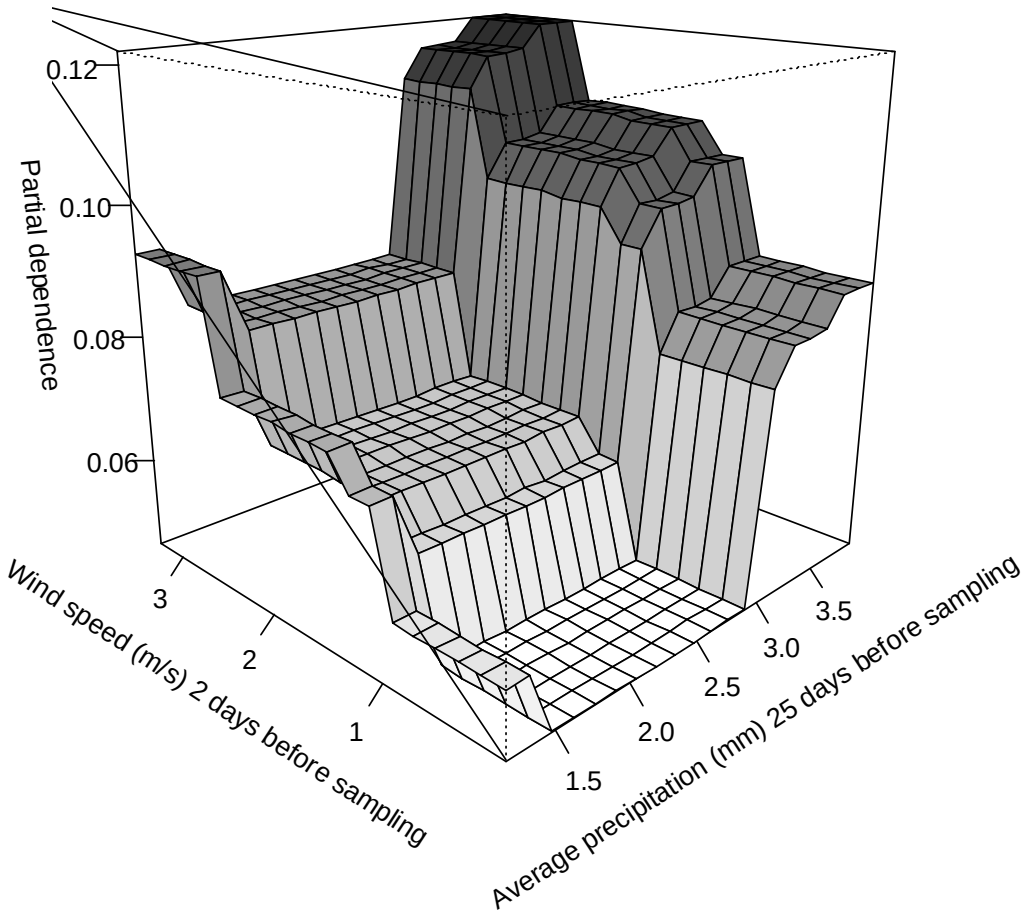

Figure 2.6 Partial dependence of probability of isolation of *Listeria* spp. on wind speed at day 2 prior to sampling and average precipitation over the previous 25 days.

2.5 Discussion

This study evaluated the effect of different meteorological factors on the occurrence of *Listeria* spp. isolated from locations within a mixed produce and dairy farm. Wind speed and precipitation were identified by both LR and RF as risk factors associated with increasing probability of *Listeria* spp. contamination in the mixed farm. The developed LR model and RF showed solid predictive ability and can be applied to

predict the occurrence of *Listeria* spp. within a mixed farm under different weather conditions as a function of wind speed and precipitation. Findings from our study illustrate that contamination of *Listeria* spp. in a mixed farm is influenced by meteorological factors, and meteorological factors should be considered when evaluating management practices and developing GAPs aimed at reducing the risk of such contamination in mixed farming systems.

In this study, sampling season had significant effect on presence of *Listeria* spp. from samples in the mixed farm with samples collected from spring having the highest prevalence of *Listeria* spp. Interestingly, except for one positive sample from June, all *Listeria* spp.-positive samples were collected from March (year 2014 and 2015), April (year 2014 and 2015), and May (year 2014), which were the top 5 months with the highest monthly precipitation during the 14-month study period. Indeed, the developed LR and RF model identified average precipitation over the previous 25 days before sampling as a risk factor associated with increasing occurrence of *Listeria* spp. in the mixed farm. This finding is in accordance with previous studies (57, 104, 105, 106). Increased rainfall may lead to higher soil moisture content which has been reported to increase the survival of *Listeria* spp. (115, 116, 117). Elevated precipitation was also found to be correlated with increased prevalence and concentration of foodborne pathogens in water (79, 81) and excessive rainfall may increase runoff from surface and subsurface water, which might act as a vehicle for introducing pathogens into pre-harvest environments (41). Heavy rainfall may also contribute to the dissemination of microorganisms in the environment (80) and transfer of microorganisms from soil to fresh produce due to splashing (88, 89). Although 15 out of the 17 precipitation

variables were significant in LR univariate analysis, only one describing the long-term cumulative effect of precipitation was retained in the final LR model. In addition, in the variable importance plot, precipitation variables describing long-term cumulative effects (average precipitation over previous 15 or more days before sampling) ranked higher than those describing short-term (average precipitation over previous 10 days or less) or instant effect (precipitation on sampling day or up to 2 days before sampling day) of precipitation. Similarly, in PCA, only the long-term and short-term cumulative effect components were retained in the final LR model using principal component scores. These results suggest average precipitation over a period of time is a better indicator of increased risk and that precipitation has a cumulative effect on isolation of *Listeria* spp. from a mixed farm as increased prevalence of *Listeria* spp. is likely to be associated with higher precipitation occurring over a period of time.

The developed LR and RF model both identified increasing wind speed just before sampling as a risk factor associated with increasing probability of *Listeria* spp. isolation from the mixed farm. During the study period, the dominant direction of wind around the farm was from west (W). Cow pasture and composting fields (liquid manure, separated solids, partially and fully composted material) were located on the west of the farm, while the vegetable production area was located downwind, east of cow pasture and composting fields. Although the average wind speed at day 2 before sampling was only 1.0 m/s observed during our study period, the maximum wind speed at day 2 before sampling averaged at 7.2 m/s, a wind speed of which can cause movement of dust according to the Beaufort wind scale (118). Dust and/or aerosols can travel and spread long distances with the presence of strong wind. Survival of

microorganisms including human pathogens in dust and aerosols have been reported and contaminated dust or aerosols can be transmitted to vegetable growing fields with the help of wind (119, 120). Wind-driven manure dust has been suspected as a possible route for microbial contamination of vegetable crops grown in proximity of animal operations (93, 119, 121, 122, 123). In our study, wind appeared to have caused the increased prevalence observed in samples collected under higher average wind speed before sampling day. Although *Listeria* spp. were never identified in organic field soil, since pathogen reservoirs such as feces, manure, and compost were located upwind, it is possible that stronger winds could blow and spread contaminated dust from these upwind reservoirs to downwind locations including the vegetable production area. Thus, our findings illustrate the potential contamination risk of vegetables from animal operations or composting fields within a mixed farm via dust or aerosols and suggest that possible control could include strategies to minimize dissemination of contaminated dust in areas with strong wind. Buffer zones or a set-back distance have been suggested as a way of reducing the risk of airborne transmission of microorganisms to produce fields, as microbial contamination of produce from animal operation via dust decreases with increasing distance between vegetable crop field and animal feedlots (122, 124).

In LR analysis, none of the 51 temperatures were significantly associated with increasing prevalence of *Listeria* spp. in the mixed farm by the univariate analysis. In RF analysis, only maximum temperature at day 2 before sampling was among the top 15 factors with the largest variable importance scores. Increased temperatures may favor the growth of bacterial pathogens that grow optimally at mesophilic temperature

ranges. For example, higher prevalence and/or concentration of *Salmonella* and *E. coli* O157:H7 have been observed during warmer months in surface water, manure and manure amended soil, tomatoes, and leafy greens (72, 79, 80, 81, 82, 83, 84). The weak associations between temperature variables and isolation of *Listeria* spp. are attributed to the fact that *Listeria* spp. are able to survive and multiply in a wide range of temperatures from 1-2°C to 45°C (125). Moreover, higher prevalence of *Listeria* spp. was observed during lower temperatures or colder months from previous studies (105, 126, 127). This study provides evidence that temperature is a less influential meteorological factor for *Listeria* spp. isolation compared to precipitation and wind speed factors.

RF is an ensemble method consisting of various sub-models (classification or regression trees) that are combined to obtain a prediction of the outcome of interest (128). Due to its lack of interpretability, RF is often considered as a "black-box" method that is good for prediction but not well-suited for inference (129). However, as our study shows here, RF can be interpreted through variable importance measures and partial dependence plots and can be an ideal backup method for traditional regression methods. Variable importance measures based on RF can be of great help in variable selection in the regression process especially in the presence of highly correlated variables. In our study, 17 precipitation variables were highly correlated, instead of arbitrarily selecting among correlated variables or using PCA which lacks interpretability, RF provided a great reference by showing the importance ranking of variables. The variable importance measures and partial dependence plots based on the RF model in our study confirmed the findings from the LR analysis that increasing

precipitation over a period of time and increasing wind speed are risk factors for the presence of *Listeria* spp. in a mixed farm. Both of the developed LR and RF models showed solid predictive performance, indicating their ability to predict the risk of *Listeria* spp. isolation from a mixed farm under different weather conditions, which can be used to provide guidance on farm management practices and development of intervention strategies. For example, growers in mixed farms may adjust the schedule for harvest if and higher wind speed occurred within recent days to reduce the risk of *Listeria* spp. contamination.

In conclusion, presence of *Listeria* spp. in a mixed farm setting is affected by meteorological factors. Specifically, the developed LR and RF model identified increasing wind speed and increasing precipitation as two risk factors for the presence of *Listeria* spp., suggesting rain-facilitated processes such as run-off and wind-driven processes such as blown dust as two possible routes of contamination in a mixed farm setting. Our study demonstrated how meteorological factors affect pre-harvest contamination of *Listeria* spp. in mixed farms and can be used to predict the risk of contamination within a mixed farm under different weather conditions. These findings will assist farmers of mixed farms in evaluating the farm management practices and developing intervention strategies to reduce the risk of pre-harvest contamination.

Chapter 3 Evaluation of meteorological factors associated with pre-harvest fecal contamination risk in a mixed produce and dairy farm

3.1 Abstract

Enteric foodborne pathogens can be shed, survive and multiply in the environment that can subsequently serve as reservoirs or sources of contamination for produce during cultivation. Produce products from mixed farms may be at risk due to its unique setting and practice. It is necessary to investigate risk factors for pre-harvest contamination in mixed farms. This study sought to identify specific meteorological factors affecting the presence and population levels of generic *Escherichia coli* (as an indicator for fecal contamination) in a mixed produce and dairy farm. Over 14 months, environmental samples were collected from locations within a mixed produce and dairy farm, and enumerated for generic *E. coli*. Local weather factors were evaluated for their association with the presence of generic *E.coli* by using logistic regression and classification trees. In addition, negative binomial regression and regression tree method were applied to identify factors affecting population levels of generic *E. coli* from a sample location. The logistic regression and classification tree identified monthly precipitation (OR=4.4, P = 0.0001) and monthly temperature (OR=1.1, P = 0.0003) as risk factors, indicating that the probability of isolation of generic *E. coli* increases with higher average amount of rain (> 1.42 mm) and higher average temperature (> 20.2°C) in the previous 30 days. However, probability of isolation was negatively correlated with rain amount within 2 days of sampling (P < 0.0001). In addition, according to the negative binomial model and regression tree, generic *E. coli*

populations decreased with increasing rainfall and wind speed in the previous 2 days, suggesting that recent rainfall (> 0.51 mm) and high wind speed (> 2.53 m/s) may lower generic *E. coli* population levels within farm environments. Results suggest that presence and population level of *E. coli* on integrated dairy/vegetable farms is influenced by temperature, precipitation and wind speed. Meteorological factors should be considered when evaluating farm management practices to reduce pre-harvest pathogen contamination.

3.2 Introduction

Produce can serve as a vehicle for foodborne pathogens and has been associated with significant number of illnesses, hospitalizations, and death in the U.S. over the years (9). Produce is vulnerable to microbial contamination during pre-harvest production as it is generally grown in open fields and may be exposed to environmental reservoirs of foodborne pathogens such as soil, irrigation water, animal manure, and wildlife or livestock (130). Presence of foodborne pathogens in produce pre-harvest production environment is usually associated with fecal contamination, as enteric foodborne pathogens, such as *E. coli* O157:H7 and *Salmonella*, can be shed into the environment through feces of a variety of animal host. Pathogens that are shed into the environment have shown the ability to persist in environmental reservoirs and may spread and contaminate distant locations within the environment via a variety of vehicles such as runoff from animal operations or wind-driven manure dust (119, 120). A number of recent outbreaks associated with produce were traced back to fecal contamination during pre-harvest production. Investigations of a multistate outbreak of

Salmonella Typhimurium and *Salmonella* Newport linked to cantaloupe isolated the same outbreak strain of *Salmonella* from environmental samples including animal feces (131). A multistate outbreak of *E. coli* O157:H7 linked to romaine lettuce used in ready-to-eat salads caused 33 illnesses and 7 hospitalizations in 2013 (15). The source of this outbreak was suspected to be fecal contamination from animal operations located near the lettuce harvesting field (30). Thus, control and prevention of fecal contamination is of great importance to reduce the contamination risk associated with produce. Generic *E. coli* has been used by food industry and regulatory agencies as an indicator of fecal contamination for microbial quality testing and the evaluation of good agricultural practices (GAPs) and good manufacturing practices (GMPs) (http://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm334114.htm#key). As generic *E. coli* is abundant in human and animal feces but not generally found in other niches (132), its presence can be considered as an indicator of contamination with fecal materials where enteric pathogens may be present. Therefore, generic *E. coli* could be used to identify risk factors associated with pre-harvest produce contamination of fecal materials and related pathogens.

As mentioned in Chapter 2 section 2.2, Microbial contamination of produce can be affected by meteorological conditions. Studies have demonstrated that presence and concentration of indicator bacteria and foodborne pathogens in produce pre-harvest environments are influenced by meteorological factors (105, 133). Understanding the effects of meteorological factors is critical while developing intervention strategies aimed at reducing the risk of pre-harvest produce contamination.

Mixed farms are one of the major food production systems for organic foods (134). By integrating vegetable crop production and livestock operations, mixed farms have the advantages of sustainable agriculture: efficient utilization of resources, maintaining environmental balance, and improvement of soil structure (101, 135). However, produce products out of mixed farms may be at high risk of microbial contamination due to the unique agricultural setting and practices of mixed farming. In a mixed farming system, livestock are reared on grass or crop residues and animal waste will be further as sources of nutrients to fertilize fields for crop production. (103, 134). Soil amendments such as raw animal manure or incompletely composted manure are known reservoirs for enteric pathogens (87) Due to its proximity to these environmental reservoirs within the same facility, produce products grown in a mixed farm are susceptible of cross-contamination from these sources. Weather conditions such as warm temperature and high humidity may also favor the growth of pathogens among these environmental reservoirs. In addition, weather events like heavy rainfall facilitate the movement of pathogen carriers (e.g., fecal material, manure, soil, and water) and the transmission of pathogens between produce growing fields, animal operations, and composting facilities within the same mixed farm facility. Thus, it is of the interest to investigate the role of meteorological factors in the introduction and transmission of fecal materials and associated pathogens within a mixed farm setting. The objectives of this study were to: (1) identify specific meteorological factors associated with the presence and population levels of generic *E. coli* in a mixed produce and dairy farm, and (2) to predict the prevalence and the level of generic *E. coli* contamination in the mixed produce and dairy farm under different weather conditions.

3.3 Materials and methods

3.3.1 Sample collection and preparation

A longitudinal study was conducted at a mixed produce and dairy farm at the Central Maryland Research and Education Center Facility (CMREC) (**Figure 2.1**). Samples were collected monthly over 14-month from February 2014 to April 2015 (excluding April 2014) from different sites within the mixed farm including around the dairy barn, the composting facility, and the environment around the farm. Sample types collected from around the barn includes cow feces, cow feed and drinking water, bird feces. Samples collected from the composting facility include raw liquid manure, lagoon water, raw separated solids, partially composted material, and fully composted material. Samples collected from the environment around the farm includes surface water from a creek, cow pasture soil, and vegetable field soil. Samples were collected following the methods described in Chapter 2 session 2.3.3. In total, 147 samples were analyzed for presence and concentration of generic *E. coli*.

3.3.2 Generic *E. coli* detection and enumeration

For solid samples (soil, feces, dry manure, compost, and cow feed), 1 g of sample were diluted 1:10 (wt/vol) with sterile phosphate buffered saline (PBS). After vortex, each sample was then serially diluted with 0.1% Peptone Water (PW) to make dilutions of 10^{-1} to 10^{-4}. A 1-ml aliquot of each dilution was pipetted onto *E. coli*/coliform Petrifilms (3M Global Headquarters, St. Paul, MN) for quantification of *E. coli*, as recommended by manufacturer. Petrifilms were counted for *E. coli* after incubation at 35 °C for 48 h. Up to 100 ml of liquid samples were filtered (depending

on turbidity) through 0.2 μm mixed cellulose ester filters (Millipore, Billerica, MA) using a PALL filtration system (PALL Life Sciences, Ann Arbor, MI). Membrane filters were aseptically removed and transferred to MI agar (Becton Dickenson and Company, Franklin Lakes, NJ), which was incubated at 35 °C for 24 h. Following incubation, blue colonies were counted under ambient light conditions to obtain the *E. coli* count.

3.3.3 Meteorological data

A total of 102 meteorological variables including temperature, precipitation, and wind speed were obtained from the weather station located at the CMREC. Maximum, minimum, average daily air temperature, daily precipitation, and average and maximum wind speed were acquired for the day of sampling, and day 1 and day 2 before sampling day. In addition, the average temperature (maximum, minimum, and average daily temperature), precipitation, and wind speed (average, maximum) were calculated for various period (from 1 to 30 days) before sampling to capture any potential long term cumulative effect of meteorological factors.

3.3.4 Statistical analyses and modeling

All statistical analyses were performed using R (version 3.2; R Core Team, Vienna, Austria). Prevalence and count of generic *E. coli* was calculated for each sample type (soil, water, fertilizers, dairy barn) and season (spring, summer, fall, and winter). The counts of generic *E. coli* in positive samples were $\log_{10}$ transformed and rounded up to the nearest integer, while *E. coli* negative samples were assigned a value of 0 log CFU/g for regression modeling. Chi-square test or Fisher's exact tests (when

expected frequency in any cell is less than 5) and Tukey's tests were used to compare generic *E. coli* prevalence and count between different sample types and seasons. Individual *P* values were considered significant at ≤ 0.05. The Holm-Bonferroni multiple comparison correction was used to assess significance.

Microbial sampling data (prevalence and count of generic *E. coli*) from all sample types were pooled to represent the overall microbial quality within the mixed produce and dairy farms. Prevalence and count data were analyzed separately for their association with meteorological factors by applying two different methods: regression models, and classification and regression trees (CART). For prevalence analyses, each individual meteorological variable was first evaluated for its association with presence of generic *E. coli* from all samples by using univariate logistic regression (LR) analyses. Significant variables (*P* values < 0.2) in univariate analyses were further evaluated in multivariate LR analyses. When two or more significant variables considered for multivariate analyses were highly correlated (correlation coefficient > 0.70 determined by Spearman's correlation coefficient), variables will be considered one at a time in the multivariate modeling and the one gives the best model fitting will be retained in the final model. For analyses of count data, a standard statistical approach would be Poisson regression. However, as Poisson regression assumes equal mean and variance, its use is often limited in handling complex and over-dispersed ecology data (136). Negative binomial regression (NBR) as a modification of Poisson regression can be used as an alternative approach to assess over-dispersed data (137). For analysis of generic *E. coli* count data, both Poisson regression and NBR models were considered. By comparison of the results from chi-square tests of goodness-of-fit of the two models

based on Pearson residuals, NBR model is more appropriate than Poisson regression model (P value = 0.59 vs. P value < 0.01). Thus, NBR model was selected as the approach to analyze generic *E. coli* count data. The NBR model for predicting generic *E. coli* count in the mixed farm was developed in two steps: univariate and multivariable analyses, following the similar approach as in prevalence analyses described above.

Significant meteorological variables in univariate regression analyses were evaluated in the multivariable LR model ("glm" function in STAS package) and the multivariable NBR model ("glm.nb" function in MASS package). The final multivariable models were build using a backward selection method based on the Alkaike information criterion (AIC) until only significant variables were retained (P < 0.05). In the final model, Box-Tidwell tests were used to assess the assumptions of linear relationships between explanatory variables and the transformation of outcome based on the link function: logit link for LR model and log link for NBR model. The goodness of fit of the final models were assessed by Pearson and deviance Chi-square test and the Le Cessie-van Houwelingen-Copas-Hosmer test. Possible collinearities among explanatory variables in the final logistic regression model were investigated by calculating the variance inflation factor.

Classification tree (CT) modeling was used to determine meteorological factors that classified sampling sites in the mixed farm by generic *E. coli* presence or absence and regression tree (RT) modeling was used to determine meteorological factors that classified sampling sites in the mixed farm by the count of generic *E. coli*. CT and RT were built using the "rpart" package. Significant variables from univariate analyses

were used to develop CT and RT. Growing and splitting of the trees were conducted by minimizing nodes impurity measured by Gini index and sum of squares for CT and RT respectively. A 10-fold cross-validation was used for CT and RT to prune the trees to the optimal sizes that minimize the cross-validated error.

The predictive performances of each of the developed models were evaluated: LR model and CT model developed in prevalence analyses were assessed by receiver operating characteristic curves (ROC) by calculating the area under the curve (AUC), while NBR model and RT model developed in count analyses were assessed by calculating the normalized root mean squared error (NRMSE). A 10-fold cross validation was conducted for each evaluation of predictive performance where the whole data set was randomly divided into ten subsets with equal size and then nine subsets were used as training sets in LR, CT, NBR, and RT models to generate model coefficients while the remain subset was used to as a test set to assess model performance by calculating AUC or NRMSE. This process was repeated ten times, each time with a different subset as test set.

3.4 Results

Overall, 83 out of 147 (56%) samples collected from locations within the mixed farm were positive for generic *E. coli*. The generic *E. coli* counts on positive samples ranged from 1 to 7 log CFU/g. Sample type was significantly associated with prevalence and count of generic *E. coli* ($P < 0.001$) with the highest prevalence and count observed in cow pie, bird feces, and separated solids (**Table 3.1**). Season was significantly associated with the prevalence of generic *E. coli* ($P < 0.001$). The

prevalence of generic *E. coli* was significantly higher in summer and fall than in spring (P = 0.02 and P = 0.005 respectively) (**Table 3.2**). In addition, season was also significantly associated with generic *E. coli* count ($P < 0.001$) and generic *E. coli* count was significantly higher in fall than in spring ($P < 0.001$).

Table 3.1 Prevalence and count of generic *E. coli* isolated from pre-harvest environments of a mixed produce and dairy farm among different sample types.

Sample category	No. of samples	No. of *positve samples (%)*	Average count of generic *E. coli* (log CFU/g)[a]
Sample type			
Cow pie	11	10 (90.9)	4.18
Bird feces	8	7 (87.5)	4.13
Separated solids	11	10 (90.9)	3.55
Lagoon 1	13	10 (76.9)	2.62
Lagoon 2	13	11 (84.6)	2.08
Pasture soil	13	9 (69.2)	1.92
Windrow compost	13	9 (69.2)	1.46
Cow feed	13	7 (53.9)	1.38
Creek water	13	4 (30.8)	0.46
Late compost	13	3 (23.1)	0.31
Organic field soil	13	2 (15.4)	0.31
Cow drinking water	13	1 (7.7)	0.23
Total	147	83 (56.5)	1.80

[a]Including negative samples that were assigned with a value of 0 log CFU/g

Table 3.2 Effect of season on prevalence and count of generic *E. coli* isolated from pre-harvest environments of a mixed produce and dairy farm.

Sample category	No. of samples	No. of positive samples (%)[a]	Average count of generic *E. coli* (log CFU/g)
Season			
Winter	36	22 (48.9) A	1.6 AB
Spring	45	10 (31.3) B	1.0 B
Summer	34	23 (67.6) AB	1.8 AB
Fall	32	28 (77.8) A	2.6 A
Total	147	10	2

[a]Different letters represent values that are significantly different ($P < 0.05$) by Fisher's exact test.

In the LR univariate analyses, 11 out of 102 meteorological variables were significantly associated with the presence of generic *E. coli* within the mixed farm, including two temperature variables, five precipitation variables, and four wind speed variables (**Table 3.3**). The odds of generic *E. coli* isolation from the mixed farm increased with higher monthly average ambient temperature (OR = 1.03) and higher monthly average precipitation (OR = 1.76) over the past 30 days before sampling. On the other hand, odds of generic *E. coli* isolation decreased when the mixed farm was exposed to a larger amount of rain or higher wind speed 2 or 3 days prior to sampling. Similar trend was observed from the results of NBR univariate analyses, where count of generic *E. coli* increased with higher average temperature and higher average precipitation 30 days prior to sampling but decreased when exposed to larger rainfall and higher wind speed 2 or 3 days prior to sampling (**Table 3.4**). Results from univariate analyses indicate that increasing rainfall and increasing wind speed just a few days before sampling may reduce the prevalence and population levels of generic *E. coli* within the mixed farm, however, higher average temperature and larger average rainfall have a long term effect and may increase the prevalence and population levels of generic *E. coli* in the mixed farm.

Table 3.3 Associations between meteorological variables and prevalence of generic *E. coli* in samples from the mixed farm based on the univariate logistic regression models.

Description	OR[a]	95%CI[b]	*P*
Average of the maximum temperature over the previous 30 days before sampling	1.02	0.99, 1.06	0.185
Average temperature over the previous 30 days before sampling	1.03	0.99, 1.06	0.166
Precipitation at day 2 before sampling	0.73	0.53, 0.87	0.006
Precipitation at day 3 before sampling	0.83	0.65, 1.04	0.112
Average precipitation over the previous 3 days before sampling	0.86	0.71, 1.03	0.107
Average precipitation over the previous 9 days before sampling	0.75	0.49, 1.12	0.158
Average precipitation over the previous 30 days before sampling	1.76	1.57, 1.97	0.034
Average wind speed at day 2 before sampling	0.53	0.34, 0.80	0.004
Average wind speed at day 3 before sampling	1.59	1.08, 2.40	0.022
Average wind speed over the previous 1 day before sampling	0.65	0.35, 1.20	0.170
Average wind speed over the previous 2 days before sampling	0.41	0.20, 0.79	0.009

[a]OR, odds ratio.
[b]CI, confidence interval

Table 3.4 Associations between meteorological variables and count of generic *E. coli* in samples from the mixed farm based on the univariate negative binomial regression models.

Description	OR[a]	95%CI[b]	*P*
Average of the maximum temperature over the previous 30 days before sampling	1.02	1.0, 1.06	0.165
Average temperature over the previous 30 days before sampling	1.03	1.0, 1.06	0.166
Precipitation at day 2 before sampling	0.81	0.70, 0.90	0.001
Average precipitation over the previous 30 days before sampling	1.83	1.67, 2.00	0.029
Average wind speed at sampling day	0.84	0.66, 1.07	0.022
Average wind speed at day 2 before sampling	0.53	0.34, 0.80	0.004
Average wind speed over the previous 1 day before sampling	0.65	0.35, 1.20	0.170
Average wind speed over the previous 2 day before sampling	0.57	0.37, 0.86	0.009
Average wind speed over the previous 3 day before sampling	0.70	0.45, 1.09	0.130

[a]OR, odds ratio.
[b]CI, confidence interval

The final LR multivariable model had four meteorological factors: 30 day average precipitation prior to sampling day (OR = 4.4), 30 day average temperature prior to sampling day (OR = 1.1), precipitation at day 2 prior to sampling day (OR = 0.4), and average wind speed at day 2 prior to sampling day (OR = 0.5) (**Table 3.5**). The probability of isolation of generic *E. coli* can be predicted using the final LR model with the equation: p =1-1/{1+ exp (-2.60 + 0.09 × Tam30 + 1.47 × Pm30 – 1.02 × P2 – 0.68 × WSa2}, where p is the probability of isolation of generic *E. coli*, Tam30 is the 30 day average temperature prior to sampling day, Pm30 is the 30 day average precipitation prior to sampling day, P2 is the precipitation at day 2 prior to sampling day, and WSa2 is the wind speed at day 2 prior to sampling day. As illustrated in **Figure 3.1,** 30 day average temperature and precipitation were positively associated with the probability of isolation of generic *E. coli* in the mixed farm while precipitation and wind speed at day 2 prior to sampling day were negatively associated with the presence of generic *E. coli* in the mixed farm. The final NBR model is comprised of the same four factors as in the final LR model: Pm30 (relative risk [RR] = 2.0), Tam30 (RR = 1.1), P2 (RR = 0.6), and WSa2 (RR = 0.6) (**Table 3.5**). The predicted number of generic *E. coli* count increases with increasing 30-day average precipitation and temperature, but decreases with increasing wind speed and precipitation at day 2 prior to sampling day (**Figure 3.2**).

Table 3.5 Final multivariable LR model for prevalence of generic *E. coli* in samples collected from the mixed farm and final multivariable NBR model for count of generic *E. coli* in samples collected from the mixed farm.

Description	OR or RR[a]	95%CI[b]	*P*
Factors for LR model			
30 day average temperature prior to sampling day	1.1	1.0, 1.2	< 0.001
30 day average precipitation prior to sampling day	4.4	2.0, 10.6	< 0.001
Precipitation at day two prior to sampling day	0.4	0.2, 0.5	< 0.001
Average daily wind speed at day two prior to sampling day	0.5	0.3, 0.9	0.027
Factors for NBR model			
30 day average temperature prior to sampling day	1.1	1.0, 1.1	< 0.001
30 day average precipitation prior to sampling day	2.0	1.4, 2.9	< 0.001
Precipitation at day two prior to sampling day	0.6	0.5, 0.7	< 0.001
Average daily wind speed at day two prior to sampling day	0.6	0.5, 0.8	0.001

[a]OR, odds ratio; RR, relative risk.
[b]CI, confidence interval

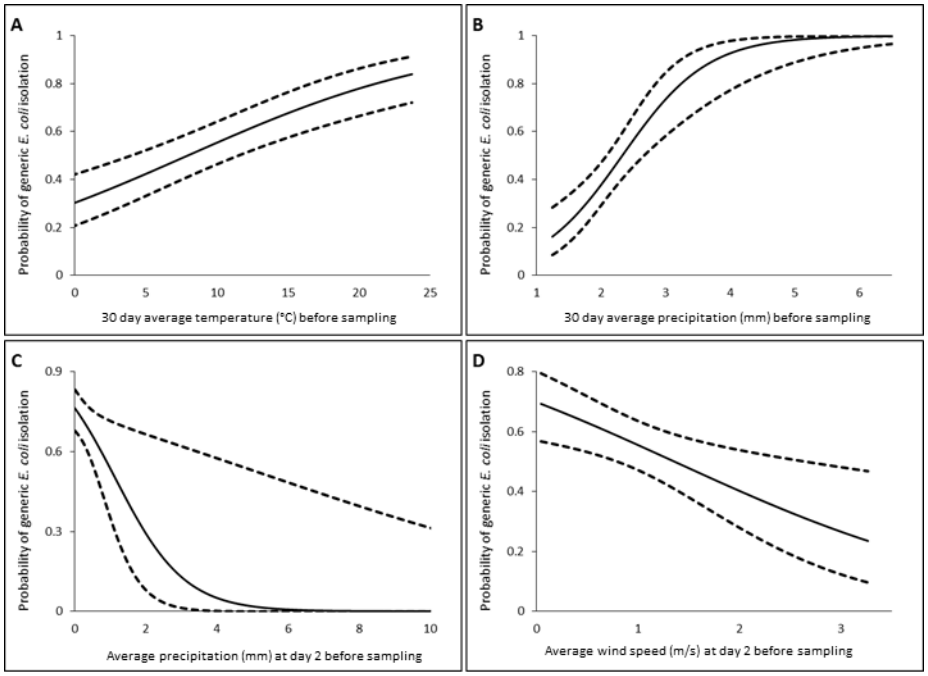

Figure 3.1 Predicted probability of generic *E. coli* isolation from sampling locations within the mixed farm for different values of 30 day average temperature prior to sampling (A); 30 day average precipitation prior to sampling (B); Average precipitation at day 2 prior to sampling (C); and Average wind speed at day 2 prior to sampling (D).

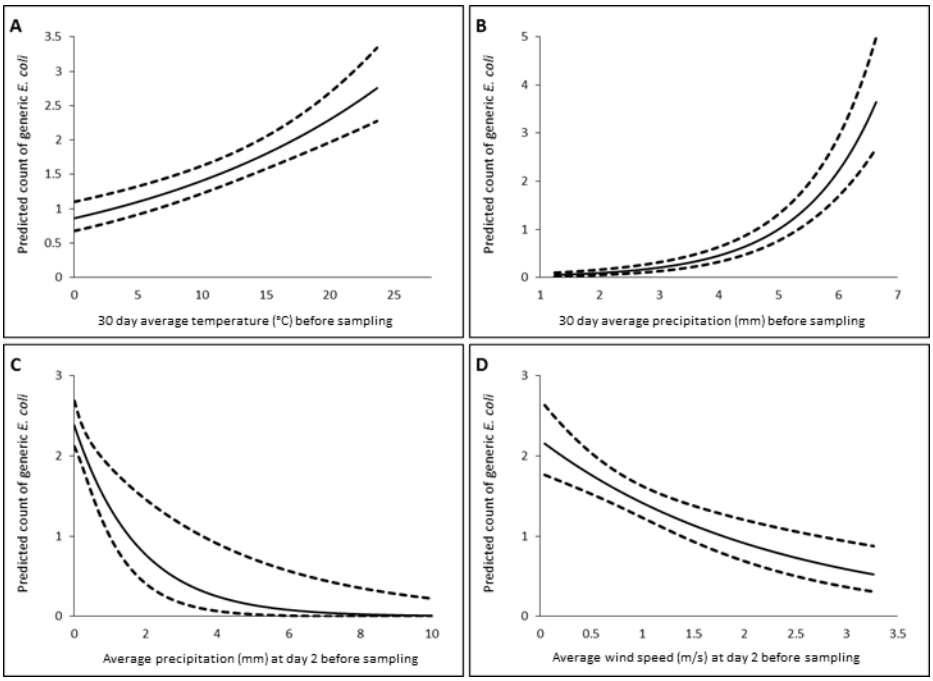

Figure 3.2 Predicted count of generic *E. coli* isolation from sampling locations within the mixed farm for different values of 30 day average temperature prior to sampling (A); 30 day average precipitation prior to sampling (B); Average precipitation at day 2 prior to sampling (C); and Average wind speed at day 2 prior to sampling (D).

The developed CT for presence of generic *E. coli* from sampling locations within the mixed farm is shown in **Figure 3.3**. CT identified two important meteorological variables: 30 day average precipitation before sampling day and precipitation at day 2 before sampling day. According to the CT model, when sampling locations within the mixed farm were exposed to a precipitation over 0.51 mm at day 2 before sampling, the probability of isolation of generic *E. coli* from the mixed farm is 20%. If precipitation at day 2 prior to sampling day is less than 0.51 mm but 30 day average precipitation prior to sampling day is greater than 1.4 mm, the predicted

prevalence of generic *E. coli* contamination increases to 75%, whereas samples exposed to less monthly rainfall were all predicted to be negative for generic *E. coli*.

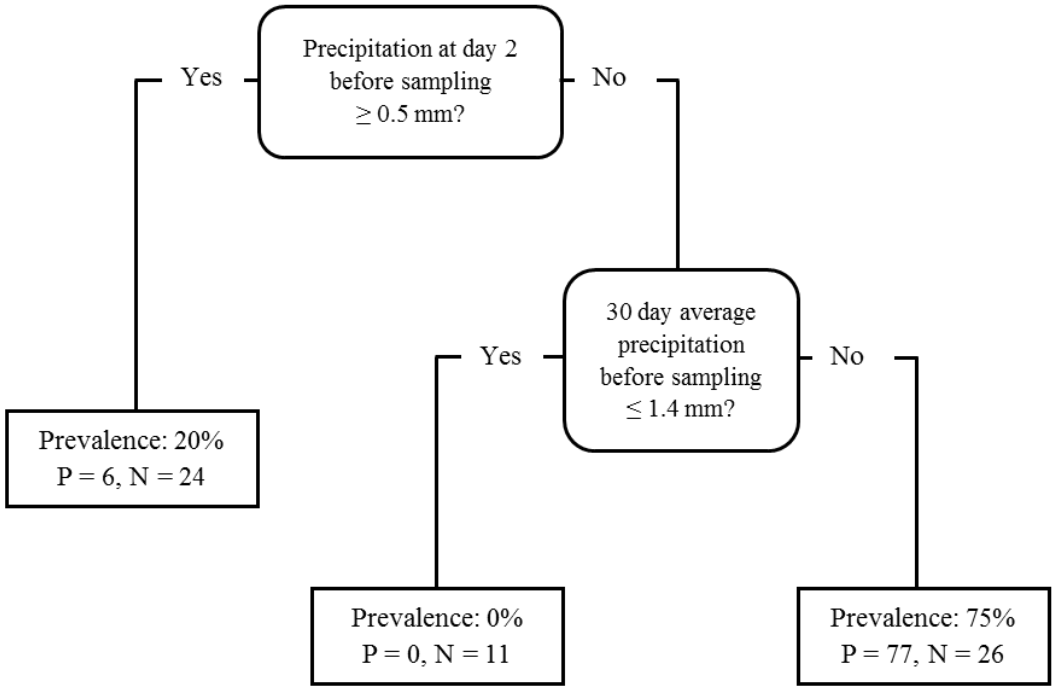

Figure 3.3 Classification tree for isolation of generic *E. coli* from sampling locations within the mixed farm. P, number of generic *E. coli* positive samples; N, number of *E. coli* negative samples.

The RT model for count of generic *E. coli* in samples collected from the mixed farm include three meteorological variables: 30 day average precipitation before sampling day, 30 day average temperature before sampling day, and precipitation at day 2 before sampling day. (**Figure 3.4**). The RT determined that average generic *E. coli* count in samples exposed to less amount of rainfall (< 0.5 mm) at day 2 prior to sampling was 2.1 log CFU/g, while samples exposed to larger amount of rainfall had an average count of 0.7 log CFU/g. When samples were exposed to more than 0.5 mm

of rainfall at day 2 prior to sampling and the 30-day average temperature is lower than 20 °C, the average generic *E. coli* count was 0.29 log CFU/g, but the average count increased to 1.7 log CFU/g among samples collected from warmer temperature. When precipitation at day 2 prior to sampling day was less than 0.5 mm but samples were exposed to larger average amount of rainfall over 30 days (> 1.4 mm) prior to sampling, the count of generic *E. coli* increased to 2.3 log CFU/g, while all samples exposed to less amount of rainfall were negative (0 log CFU/g).

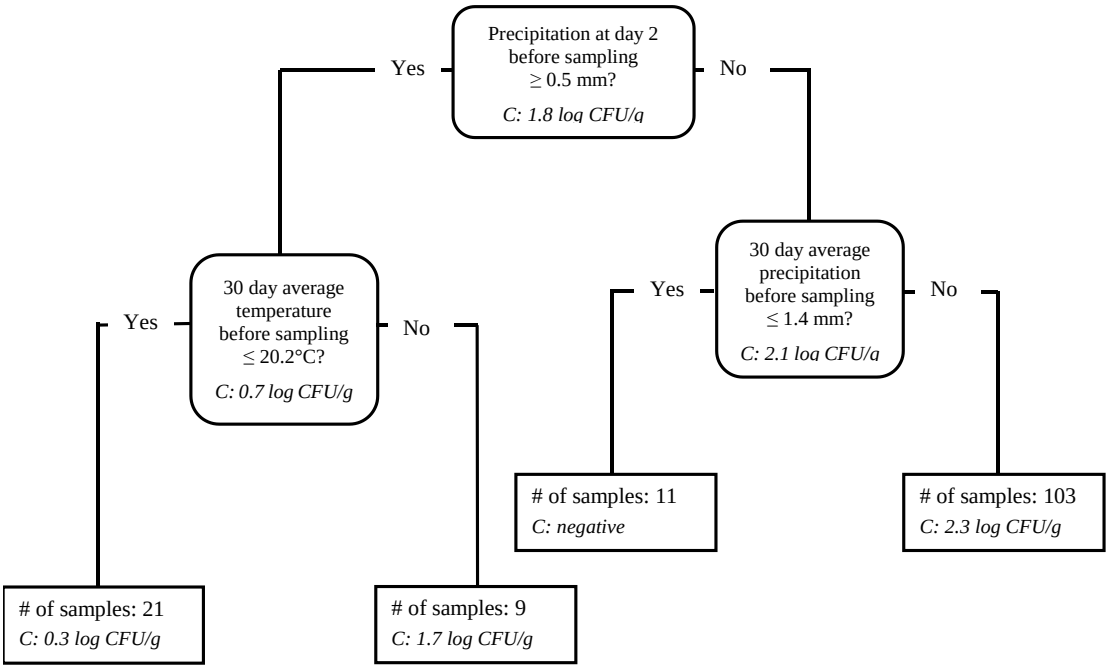

Figure 3.4 Regression tree for count of generic *E. coli* from sampling locations within the mixed farm. C, average count of generic *E. coli* with in each node.

The developed LR model was internal validated using the whole dataset and the AUC was 77.4%, while in cross-validation, the mean AUC of LR model was 81.9% (**Figure 3.5A**). For CT, AUC from internal validation was 75.7% and mean AUC from

cross validation was 77.6% (**Figure 3.5B**). For count models, average NRMSE was 0.6% from cross validation of NBR model and 20.4% for RT model.

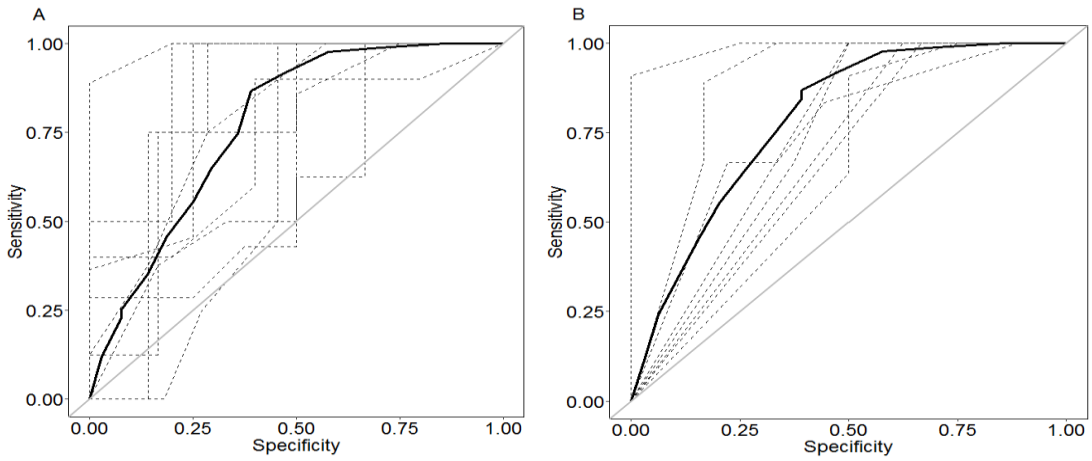

Figure 3.5 Receiver operating characteristic (ROC) curves for 10-fold cross validation (dashed line) and internal validation (solid line) of the developed logistic regression model (A) and classification tree (B). The diagonal line is the line of no discrimination (represents completely random guess).

3.5 Discussion

This study evaluated different meteorological factors associated with contamination risk of generic *E. coli* within a mixed produce and dairy farm. The LR and CT models show that the prevalence of generic *E. coli* from samples collected within the mixed farm is higher when samples were exposed to higher monthly temperature and precipitation prior to sampling. Additionally, wind speed and precipitation at day 2 prior to sampling were associated with an increase of prevalence. The NBR and RT models show that temperature, precipitation, and wind speed were also associated with the extent of generic *E. coli* contamination in the mixed farm.

These findings support the conclusions from precious research that meteorological factors were associated with microbial contamination of produce pre-harvest production environments and these factors should be considered when evaluating farm management practices and developing intervention strategies aimed at reducing such contamination (57, 104, 105, 133).

Among the 54 temperature variables evaluated in our study, two temperature related variables were found significantly associated by univariate analyses with the prevalence and count of generic *E. coli* in the mixed farm. These two temperature variables described average temperature over a long period of time (30 days) and were both retained in the final LR, CT, NBR, and RT models, indicating that higher temperature over the past month increases the odds of generic *E. coli* contamination and the extent of such contamination. Generic *E. coli* can grow under a wide range of temperatures from 4- 45°C and the optimum growth temperature is 37°C. It is well documented that warm temperature facilitate the growth of generic *E. coli* (138), and higher prevalence and concentration of generic *E. coli* have been observed during warmer months in surface water, manure and manure amended soil, tomatoes, and leafy greens (72, 79, 80, 81, 82, 83, 84). In addition, the increase in wildlife activities (and defecation) during warmer temperatures may also contribute to higher occurrence of fecal bacteria (81). Interestingly, none of the other temperature variables describing average temperature over shorter time periods or temperature at just a few days prior to sampling were significantly associated with the prevalence and count of generic *E. coli* by univariate analyses in our study. These results indicate that effect of temperature

on prevalence of and extent of generic *E. coli* contamination in a mixed farm accumulates over a long period of time.

The analyses showed that higher rainfall over longer periods was associated with an increase in both the prevalence and concentration of generic *E. coli* in samples collected from the mixed farm. This find is consistent with previous study by Park et al. (133), where higher 29 day average amount of rain was found significantly increased prevalence and count of generic *E. coli* in spinach among 18 precipitation variables. In another study, increased total rainfall over the past month and increased rainfall during the week before last of sampling day were significantly associated with increased prevalence and concentration of *E. coli* O157:H7 in water while no significant association was found for total amount of rainfall during the week before sampling (81). Runoff from surface water and livestock operations as a result of excessive rainfall may serve as vehicles for the transmission of microorganisms within produce pre-harvest environments (41, 87). Additionally, higher humidity in soil has been reported to support the survival and growth of *E. coli* (139, 140), which may explain the higher observed count of generic *E. coli* in samples exposed to higher precipitation in the past month. Studies have also suggested that seasons with higher precipitation increases wildlife and grazing activities (141), which may increase the probability of fecal contamination. Overall, the findings illustrate that the prevalence and extent of generic *E. coli* contamination in a mixed farm increases when exposed to larger amount of rainfall over a long period of time. Considering that the final LR, CT, NBR, and RT models did not identify any other risk factors, it can be concluded that the increase of prevalence and count of generic *E. coli* in a mixed farm is determine by higher average

temperature (>20°C) and larger average amount of rainfall (> 1.4 mm) in the past 30 days which support the survival and growth of generic *E. coli.*

Interestingly, according to the models in this study, lower prevalence and lower count of generic *E. coli* were expected in samples from the mixed farm that were exposed to a higher amount of rain at day 2 prior to sampling and higher wind speed prior to sampling. Previous studies have demonstrated that high wind may facilitate transportation and spread of contaminated manure dust (119, 120) and wind-driven manure dust may be a possible route of microbial contamination of vegetable crops grown in proximity of animal operations (93, 119, 121, 122, 123). However, in this study, wind speed was not identified as a risk factor as the prevalence and count of generic *E. coli* in the mixed farm decreased with increasing average wind speed at day 2 prior to sampling. Many farm management practices such as irrigation and manure application that may influence the contamination in produce are closely related to weather conditions. For example, irrigation water is a potential source and route of microbial contamination in produce, and irrigation water may be applied more frequently when fields were exposed to less rainfall recently. It is possible that farm management factors such as irrigation and manure application may be a more informative predictor for prevalence and concentration of generic *E. coli* in mixed farm, while meteorological factors such as precipitation and wind speed may act as confounding factors. Thus, results from this suggested that meteorological factors might not be the only and most influential factors for contamination of generic *E. coli* of in a mixed farm. Future studies should take into account the effect of farm

management practices, especially those that may interact with weather conditions, when predicting the contamination in produce pre-harvest environment.

Similar meteorological factors were identified from CT and RT, and the results from CT and RT supported the findings from LR and NBR. In cross validation, LR and NBR model showed good predictive abilities (mean AUC = 81.9% and mean NRMSE = 0.6%). As a comparison, the mean AUC from CT was 77.6% and mean NRMSE was 20.4% for NBR model in cross validation. The relative low predictive performance of RT model is due to its inability to accurately predict samples with higher generic *E. coli* count, the maximum predicted count from RT is 2.4 log CFU/g while the count (3 – 7 log CFU/g) were observed in the samples. Nevertheless, CART as an alternative approach to regression offers great interpretability and has the advantage of its ability to handle and analyze complex, unbalanced data involved in ecological studies (136). Previous studies also recommended the use of CART together with regression when analyzing microbial presence in the environment (104). The CT and RT developed in this study showed robust predictive performance comparable to that of regression methods, indicating their abilities to predict the risk of *Listeria* spp. isolation from a mixed farm under different weather conditions. As the developed CT and RT separate "high-risk date" and "low-risk date" based on specific meteorological rules (temperature, wind speed, and precipitation), it can be used to provide guidance on farm management practices and development of intervention strategies.

In summary, this study identified influential meteorological factors associated with the contamination risk of generic *E. coli* in a mixed produce and dairy farm. The developed models can used to predict the occurrence and the population level of generic

E. coli in a mixed farm under different weather conditions. The findings from this study suggest that while meteorological factors temperature, precipitation, and wind speed affect the prevalence and concentration of generic *E. coli* in the mixed farm, other factors such as farm management practices (e.g., irrigation and manure application) need to be considered when predicting the contamination in a mixed farm and developing intervention strategies to reduce the risk of such contamination.

Chapter 4 Effects of cover cropping and meteorological factors on the survival and population dynamics of generic *Escherichia coli* and *Listeria innocua* in produce fields

4.1 Abstract

As soil can serve as a reservoir and contamination route for foodborne pathogens during produce production, it is necessary to evaluate the effects of various factors on the survival of foodborne pathogens in produce fields. This study sought to investigate the effect of a particular farm practice, cover cropping, along with meteorological factors on the survival of generic *E. coli* and *L. innocua* in organic and transitional organic produce fields. Five cover crops and bare ground (no cover crop) control plots were inoculated with indicator bacteria generic *E. coli* and *L. innocua* in fall 2013 and 2014. Soil samples were collected periodically and were enumerated for *E. coli* and *L. innocua* by a modified MPN method. Survival analysis and Poisson regression were applied to determine the effects of cover crop and meteorological factors on the survival of *E. coli* and *L. innocua* in soil. Survival analysis indicated that cover crop treatment was not a significant factor affecting the survival of *E. coli* in soil and survival of *E. coli* was not significantly different between organic and transitional organic fields. Interestingly, Cox regression revealed that survival of *E. coli* in soil was significantly associated with precipitation (P = 0.001) and temperature (P = 0.006); increasing precipitation increased the survival of *E. coli* in soil and survival of *E. coli* was significantly greater at colder temperature. For *L. innocua*, population levels were significantly higher in transitional organic plots as compared to organic plots.

Significantly higher population levels were also observed under higher monthly precipitation, relative humidity, and temperature. While the effect of cover cropping is minimal, survival of food safety indicator bacteria is greatly influenced by meteorological factors, indicating increasing precipitation and humidity may prolong the survival of *E. coli* and persistence of *L. innocua* in soil in regions with cold weather.

4.2 Introduction

Produce is an essential part of a healthy diet: the newest Dietary Guidelines for Americans highlighted vegetables and fruits as important components of a healthy eating pattern in key recommendations (2). Fresh produce is often consumed raw which makes the safety of produce crucial for the health of consumers and for maintaining the produce industry. Programs such as Good Agricultural Practices (GAPs) are in place to reduce the risk of microbial contamination in production, harvest and handling environments. However, produce-related outbreaks and recalls continues to occur (15, 16, 19, 28, 142), making the produce a growing food safety concern among consumers. During produce production, soil can serve as a reservoir and source of transmission of foodborne pathogens (143). Identification and evaluation of various factors that affect the survival and persistence of foodborne pathogens in produce fields is crucial to prevent such contamination during pre-harvest stage, since remediation or elimination of contamination that occurs before harvest is difficult to achieve during the post-harvest stage (83).

Contamination of produce fields can be affected by farm management practices. For example, specific farm management practices may influence pathogen

contamination in the pre-harvest environment of produce production (144, 145, 146, 147, 148). Cover cropping, the establishment of a crop, typically a small grain or legume, in between cultivations of a cash crop, have been applied for improving crop productivity and soil health and maintaining the sustainability of agroecosystems (149, 150). Cover crops have a positive effect in increasing the soil microbial population and activity, through cover crop plant roots that modify the soil habitat, improve aeration, and serve as a source of nutrients (151, 152). It is of the interest to investigate whether cover crops affect the population dynamics of soil bacteria relevant to food safety.

Foodborne pathogens (such as *E. coli* O157:H7 and *Salmonella*) can be shed into produce field soil through defecation of animal hosts. *Listeria* spp. including *L. monocytogenes* are associated with soil and decaying vegetation (153). Once introduced into soil, survival and persistence of foodborne pathogens can be affected by a variety of abiotic and biotic factors such as soil properties and soil microbial diversity (69, 154). Meteorological factors such as temperature and precipitation can also affect the survival and persistence of foodborne pathogens in soil directly, and by indirectly influencing soil properties and soil microbial communities. Identification of meteorological factors that affect the survival and persistence of foodborne pathogens in field soil is of great importance to reduce the risk of subsequent produce contamination.

In an attempt to investigate the effects of cover crops on microbial population dynamics in produce fields, Reed-Jones et al., (150) found no consistent effect on survival of indicator bacteria relevant to food safety (*Listeria innocua* and generic *Escherichia coli*), due to cover crop species and weather conditions. However, the

effect of meteorological factors on population dynamics of indicator bacteria in soil were not quantitatively assessed. Quantitative analysis and data are essential to allow for identification of risk factors that significantly influence the population dynamics of food safety related bacteria in soil which may help facilitate implementation of science-based preventive controls. Thus, this study aimed to evaluate the joint effect of cover cropping and meteorological factors on the survival and persistence of generic *E. coli* and *L. innocua* in produce field soil.

4.3 Materials and methods

4.3.1 Study data

Sample collection and microbial analysis were described in details previously (150). Briefly, field experiments were conducted on two different fields: a certified organic field (noted as field A) and a transitional (previously conventional) organic field (noted as field B) at the University of Maryland (UMD) Lower Eastern Shore Research and Education Center (LESREC) in Salisbury, MD during 2013 to 2015. General farm characteristics were shown in **Table 4.1.** Cover crops including hairy vetch (HV), crimson clover (C), cereal rye (R), a 2:3 (wt/wt) mixture of hairy vetch and rye (HVR), and a 2:1 (wt/wt) mixture of crimson clover and rye (CR) were sown using a grain drill while bare-ground (BG) plots served as the control and remained fallow throughout the sampling period. A cocktail of the two nonpathogenic strains (generic *Escherichia coli* and *Listeria innocua*) with concentrations at approximately 10^6 CFU/ml were applied to the fields after cover crop seeding. Soil samples were collected immediately prior to field inoculation and within 2 h following inoculation.

Sampling continued periodically post-inoculation for up to 30 weeks. Four soil samples were collected from each of the cover crop and bare ground plots per field. In cover crop plots, samples were collected to a 7-cm depth from the root zone that were nondestructive to the cover crops. In the bare ground plots, soil samples were collected in areas devoid of weeds to a 7-cm depth. All samples were collected in sterile Whirl-Pak bags (Nasco, Ft. Atkinson, WI) using sterile scoops (Fisher Scientific, Hampton, NH). Samples were sealed and transported in coolers with ice packs to the laboratory for microbiological analysis within 24 h. Thirty grams of soil was mixed with buffered peptone water (BPW) (1:5 [wt/vol]), stomached at 200 rpm for 1 min, and allowed to recover for 1h to revive injured cells. Two 96-well plates for a 3-tube most-probable-number (MPN) analysis were prepared, one for *E. coli* with brilliant green bile broth and one for *Listeria* using buffered *Listeria* enrichment broth. Serial dilutions of the samples were prepared in the respective medium. After incubation at 42°C for 24 h and 30°C for 48 h for *E. coli* plates and *Listeria* plates respectively, a 10 µl of culture was either plated on Tryptone bile glucuronic agar (TBX) and incubated at 42°C for 24 h for *E. coli* or plated on Oxford *Listeria* agar (OXA) and incubated at 30°C for 24 to 48 h for *L. innocua*.

Table 4.1 General characteristics of the two experimental fields.

Field	pH	Organic matter	Irrigation	Slope	Prior land use	Organic
A	6.8	<1%	Drip	0 to 5%	Mixed vegetable	Certified
B	5.9	<1%	Overhead	0 to 2%	Field corn and soy bean	Transitional

4.3.2 Meteorological data

Meteorological data were collected from the weather station located on LESREC. Meteorological data including precipitation, temperature, relative humidity, and solar radiance were obtained for each of the sampling date. Additional information and data were also collected or calculated for further analyses (e.g., average values for each meteorological factor between 0 and 30 days before sampling).

4.3.3 Statistical analysis and modeling

Preliminary examination of the dataset showed very different population dynamics between generic *E. coli* and *L. innocua*. Based on data from both years, average concentration of generic *E. coli* declined rapidly to below the limit of detection after inoculation and remained mostly undetectable for the rest of sampling period, while the population levels of *L. innocua,* despite declining with time overall, persisted throughout the study period (**Figure 4.1**). Thus, different statistical and modeling methods were used for analysis of generic *E. coli* and *L. innocua*. Survival of generic *E. coli* were analyzed using survival models to determine the association between predictor variables (cover crop, farming system, and meteorological factors) and survival time of generic *E. coli* in soil. For *L. innocua*, as the data violated the equal mean and variance assumption of Poisson regression, negative binomial regression (NBR) was used to determine the association between predictor variables (cover crop, farming system, time post-inoculation, and meteorological factors) and population level of *L. innocua* in soil. In addition, random forest (RF) method was also applied for analysis of *L. innouca*, as an alternative approach. All statistical analyses were performed using R (version 3.2; R Core Team, Vienna, Austria).

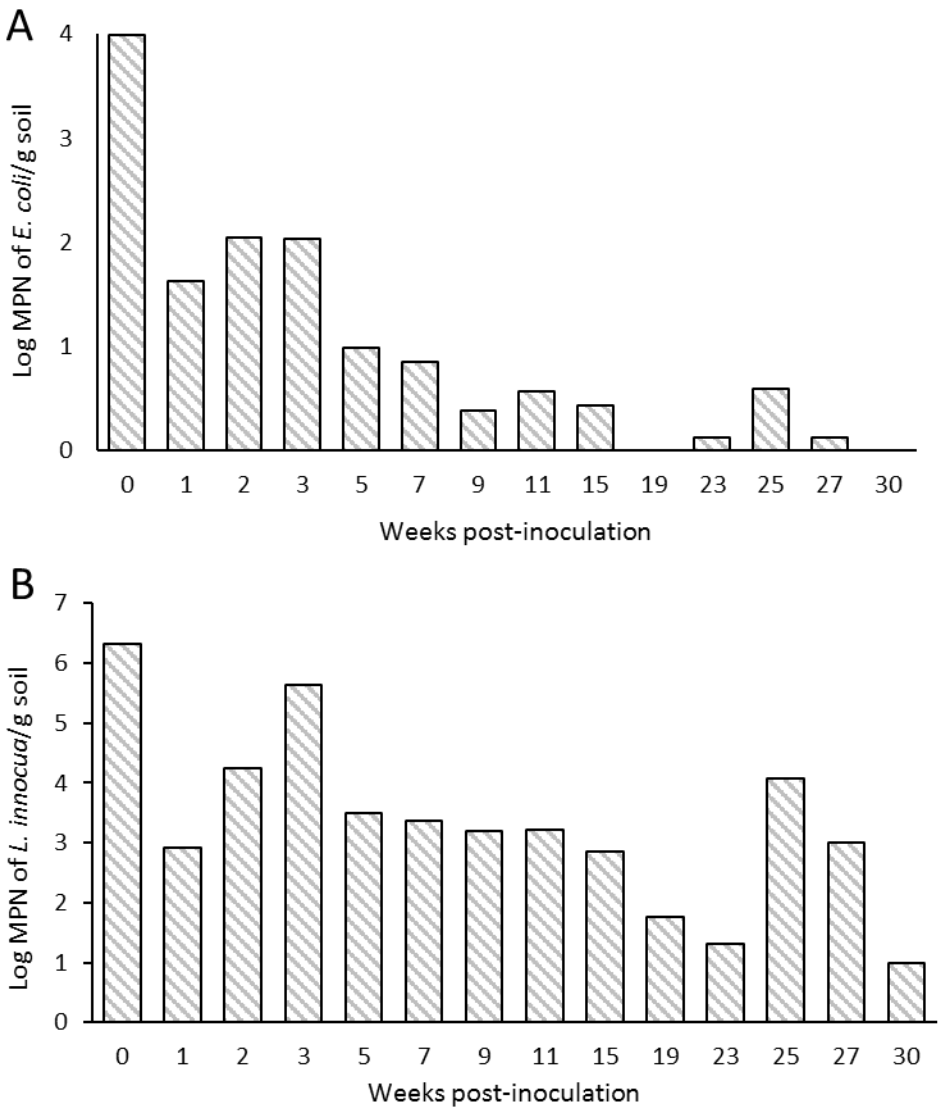

Figure 4.1 Population dynamics of generic *E. coli* (A) and *L. innocua* (B) in field soil during the study period.

4.3.3.1 Survival analysis for generic *E. coli*

Survival models are widely used in medical and veterinary studies to analyze life table data (137, 155). In survival analysis, an event is defined (such as death of a person or an animal) and time to the event is modeled as the response variable. Here, the survival of *E. coli* in produce fields was modeled with time post-inoculation until

average concentration of generic *E. coli* falls below the limit of detection (i.e., survival time of *E. coli*) as the response variable. Data from all cover crop species (including bare ground control), both fields (field A and field B), and both years (2013-2014 and 2014-2015) were used in the survival analysis. Each plot was treated as a subject in the survival analysis, and the survival time of *E. coli* in each plot was the response variable in the survival analysis. Predictor variables included were cover crop species, field, and meteorological related factors to determine their association with the survival time of generic *E. coli* in soil. Meteorological factors considered in the survival model were the average value of each factor (temperature, precipitation, relative humidity, and solar radiance) during the survival period of generic *E. coli* (i.e., from inoculation to below the limit of detection). The overall effects of categorical predictor variables, i.e., cover crop (cover crop treatment or bare ground) and fields (field A or field B) were first evaluated by Kaplan-Meier (KM) estimates of survival function. Log-rank test was performed to test whether difference in the overall survival functions between groups was significant. The associations between generic *E. coli* survival times and each individual predictor variables (both categorical and continuous) were evaluated using a multivariable Cox proportional hazards (CPH) model. The final CPH model were build following Collette's model selection approach (112): (i) fit a univariate CPH model for each individual predictor variables, and identify significant variables at $P < 0.2$ level; (ii) fit a multivariable model for all significant variables from step (i), and use a backward selection method until only significant variables (at $P < 0.05$ level) were retained; (iii) start with the multivariable model in step (ii) and evaluate each of the non-significant variables from step (i) using forward selection method, then

determine the final model by omitting non-significant variables at 0.05 level. The proportional hazards assumption of CPH was tested by Schoenfeld's residuals.

4.3.3.2 Regression analysis for generic *L. innocua*

Cover crop species, fields, and meteorological factors were considered as possible predictors for population level of *L. innocua* in soil samples Meteorological factors considered in the regression analysis were recorded value for each factor (temperature, precipitation, relative humidity, and solar radiance) on sampling dates and the average values for each factor between 5 and up to 30 days before sampling. High levels of correlations were determined among the meteorological factors from the same category (i.e., temperature, precipitation, relative humidity, and solar radiance). Thus, principal component (PC) analyses were performed to identify variables in multivariable regression analysis. For each category of the meteorological factor, the number of components retained was determined by percentage of variance explained and interpretability: retained components should explain at least 80% of the total variance, variables loading on the same component should have the same conceptual meaning, and components should have simple structure of with relatively high factor loading of a variable on only one component and relatively small loading on other components. The PC scores from these retained meteorological PCs were then used as explanatory variables in RF and in NBR multivariable modeling. The final NBR model was selected following Collett's approach as described above. The RF model was developed in "cforest" function in "party" package with 15,000 bootstrap samples and 2 randomly selected variables at each node. Variable importance plot was generated based on the developed RF to illustrate variables that were highly associated with

population levels of *L. innocua* in field soil samples. Additional, representative meteorological variables that had good interpretability from each PC were selected and used for NBR multivariable modeling to quantify the effects of specific meteorological variables. Box-Tidwell tests were used to assess the assumptions of linear relationships between explanatory variables and the log transformation of outcome. The goodness of fit of the final models were assessed by Pearson and deviance Chi-square test and the Le Cessie-van Houwelingen-Copas-Hosmer test. Possible collinearities among explanatory variables in the final NBR model were investigated by calculating the variance inflation factors (VIF).

4.4 Results

4.4.1 Factors for survival of generic *E. coli* in produce fields

The log-rank test results indicated no significant differences between the survival of *E. coli* in soil with different cover crop species (P = 0.502). The KM estimate of survival functions for *E. coli* survival in bare ground plots and in cover crop plots was shown in **Figure 4.2A**, and no significant differences were detected (P = 0.81)**.** The KM survival functions for organic fields and transitional fields (**Figure 4.2B**) showed that the survival probability was higher in field A (organic field) than that in field B (transitional organic field) throughout the study period and the difference was borderline significant (P = 0.057).

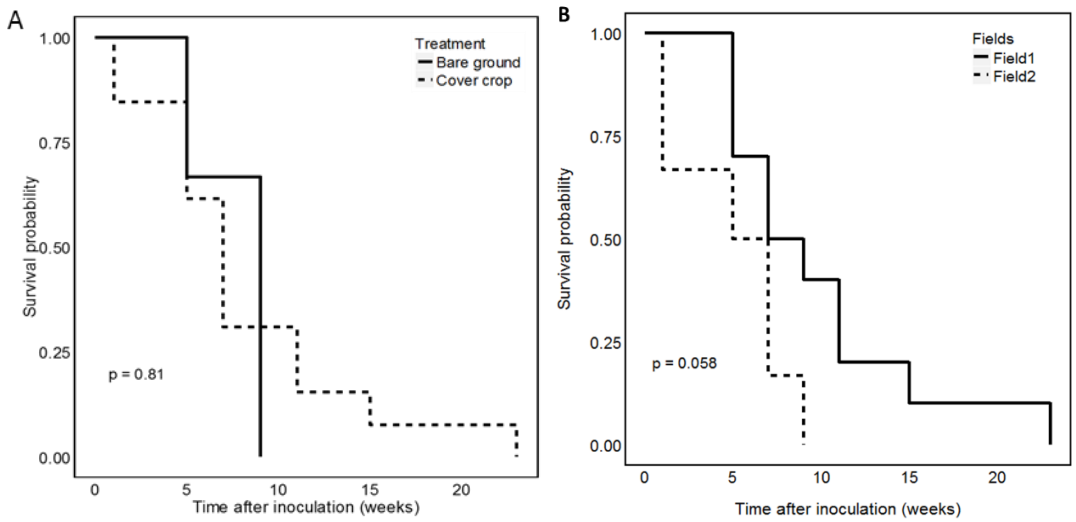

Figure 4.2 Kaplan-Meier estimate of survival functions for generic *E. coli* in bare ground plots versus cover crop plots (A) and in field A versus field B (B); *P*-value determined by log-rank test.

The CPH analyses were conducted to simultaneously determine and evaluate the effect of all predictor variables on survival of generic *E. coli*. Similarly to KM estimates, no significant impact was detected for cover crop application and survival probability of *E. coli* in field A was borderline significantly higher than that in field B (P = 0.073) in CPH univariate analyses (**Table 4.2**). For meteorological variables, average temperature, average precipitation, and average solar radiance during survival were significant factors for *E. coli* survival in univariate analyses (**Table 4.2**). The final CPH model contained two predictor variables: average precipitation (P = 0.001) and average temperature (P = 0.006) during survival period (**Table 4.3**). The model showed that with a 1 mm increase in average precipitation the risk of "death" of *E. coli* dropped by 96% (hazard ratio = 0.04) and with a 1°C increase in average temperature the risk

of *E. coli* "death" increased by 260% (hazard ratio = 2.60). As shown in **Figure 4.3**, survival of *E. coli* in produce fields was predicted to be better when produce fields were exposed to lower temperature or higher precipitation during the survival period.

Table 4.2 Association between variables and survival of generic *Escherichia coli* in produce field soil based on the univariate Cox proportional hazard model.

Factors (reference level)[a]	**Coef**[b]	**HR**[c]	***P*-value**
Fields (Field A)			
Field B	1.06	2.88	0.07
Treatment (Bare ground)			
C	-1.39	0.25	0.28
CR	-0.75	0.47	0.48
HV	-0.21	0.81	0.80
HVR	-0.08	0.92	0.92
R	0.97	2.63	0.26
Average daily air temperature during survival	0.39	1.48	0.01
Average daily precipitation during survival	-1.42	0.24	<0.01
Average daily solar radiance during survival	-3.52	0.03	0.06
Average daily relative humidity during survival	0.32	1.39	0.24

[a]C, crimson clover; CR, mixture of crimson clover and rye; HV, hairy vetch; R, cereal rye (R), HVR, mixture of hairy vetch and rye.

[b]Coef, value of the regression coefficient.

[c]HR, hazard ratio.

Table 4.3 Final Cox proportional hazard model for survival of generic *Escherichia coli* in produce field soil.

Factor	Coef[a]	HR[b]	*P*-value
Average daily air temperature during survival	0.95	2.60	0.005
Average daily precipitation during survival	-3.3	0.04	0.001

[a]Coef, value of the regression coefficient.

[b]HR, hazard ratio.

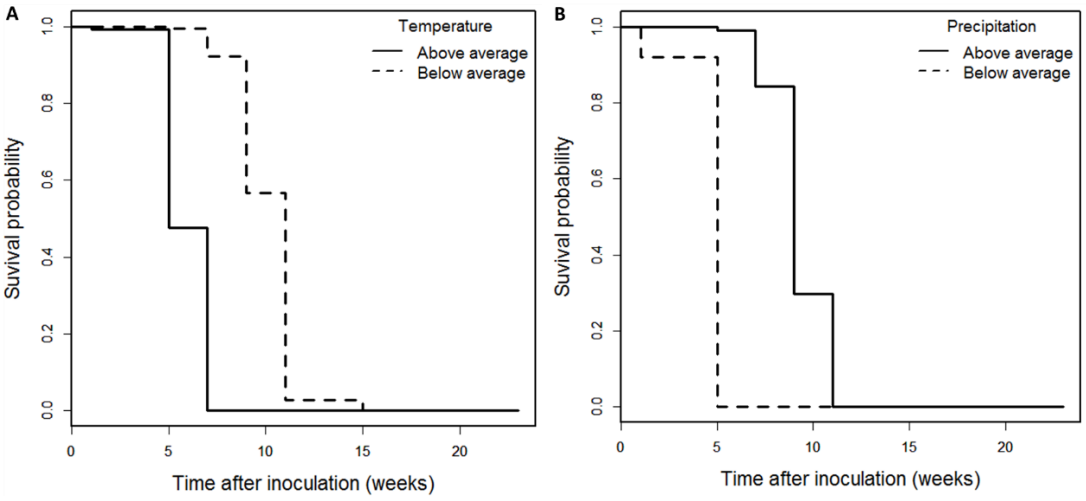

Figure 4.3 Predicted survival curves of generic *E. coli* in produce fields under different temperature (A) and precipitation (B).

4.4.2 Factors for population dynamic of *L. innocua* in produce fields

In PC analysis, the first PC of temperature variables explained 86% of the total variance and retained and subsequently named as "PC.temp". Similarly, the first PC of solar radiance variables that explained 80% was retained and named PC. s. The first

two PCs of relative humidity variables explained 88% of the total variance. Variables describing average relative humidity 5 to 25 days before sampling day had major loadings on the first relative humidity PC and was named "PC.rh1", while variables describing relative humidity on sampling day and the average relative humidity over the previous one day before sampling day was found to be loaded on the second relative humidity PC, which was named "PC.rh2" (**Table 4.4**). Similarly, the first two component from precipitation PC were retained and were subsequently named as "PC.p1" and "PC.p2" (**Table 4.4**). Results from univariate analysis (**Table 4.5**) shows that cover crop treatment was not a significant factor influencing population levels of *L. innocua* in produce fields (P = 0.269). However, population levels of *L. innocua* were significantly higher in field B than in field A. (P = 0.028). The multivariable NBR model using PCs showed that PC.temp, PC.rh1, PC.rh2, and PC.p1 were significantly associated with population levels of *L. innocua* in soil. These four PCs also ranked in the top four variable importance scores in random forest analysis, indicating relative humidity, temperature, and precipitation were the most influential meteorological factors for population level of *L. innocua* in produce field soil (**Figure 4.4**). Four representative variables were selected from each of the four PCs for the final NBR multivariable modeling: average temperature over the past 30 days before sampling, average precipitation over the past 25 days before sampling, relative humidity at the sampling day, and average humidity over the past 30 days before sampling. These four variables were evaluated together with time (weeks after inoculation) and field in the multivariable NBR analysis. The final NBR multivariable model included four variables: weeks after inoculation, average temperature over the past 30 days before

sampling, average precipitation over the past 25 days before sampling, and relative humidity at the sampling day. According to this final model, the population dynamics of *L. innocua* in soil after inoculation were influenced by these four variables: population levels of *L. innocua* declined with time, however, higher levels of *L. innocua* were observed when fields were exposed to higher temperature, precipitation, and relative humidity (**Figure 4.5**).

Table 4.4 Component loadings from principal component analysis of precipitation and relative humidity variables.

Name of variable	**PC1**[a]	**PC2**[a]
Precipitation		
Precipitation at sampling day	-0.01	**-0.59**[b]
Average precipitation over the previous 1 day before sampling day	-0.02	**-0.59**
Average precipitation over the previous 5 day before sampling day	-0.29	**-0.42**
Average precipitation over the previous 10 day before sampling day	**-0.44**	-0.13
Average precipitation over the previous 15 day before sampling day	**-0.46**	-0.02
Average precipitation over the previous 20 day before sampling day	**-0.46**	0.04
Average precipitation over the previous 25 day before sampling day	**-0.42**	0.21
Average precipitation over the previous 30 day before sampling day	**-0.36**	0.28
Relative humidity		
Relative humidity at sampling day	-0.22	**0.69**
Average relative humidity in the previous 1 day before sampling day	-0.26	**0.63**
Average relative humidity over the previous 5 day before sampling day	**-0.34**	-0.16
Average relative humidity over the previous 10 day before sampling day	**-0.39**	-0.10
Average relative humidity over the previous 15 day before sampling day	**-0.40**	-0.16
Average relative humidity over the previous 20 day before sampling day	**-0.40**	-0.09
Average relative humidity over the previous 25 day before sampling day	**-0.39**	-0.16
Average relative humidity over the previous 30 day before sampling day	**-0.38**	-0.16

[a]PC1 and PC2 denotes the first and second principal component of precipitation or relative humidity variables
[b]Bold number indicates the highest component loading.

Table 4.5 Associations between population level of *L. innouca* in field soil and assessed factors in univariate negative binomial analysis.

Variables[a]	Coef[b]	RR[c]	*P*
Time (weeks) after inoculation	-0.03	0.97	<0.001
Treatment			
Cover crop	-0.08	0.92	0.269
Bare ground	0	1	
Field			
Field B	0.134	1.14	0.028
Field A	0	1	Reference
PC.t.score	-0.10	0.90	<0.001
PC.p1.score	0.02	1.02	0.111
PC.p2.score	0.01	1.01	0.573
PC.rh1.score	-0.11	0.90	<0.001
PC.rh2.score	0.03	1.03	0.181
PC.s.score	-0.01	0.99	0.541

[a]PC.t.score, PC.p1.score, PC.p2.score, PCrh1.score, PC. rh2.score, and PC.s.score denotes PC scores from PC.t, PC.p1, PC.p2, PCrh1, PC. rh2, PC.s, respectively.
[b]Coef, value of the regression coefficient.
[c]RR, relative risk.

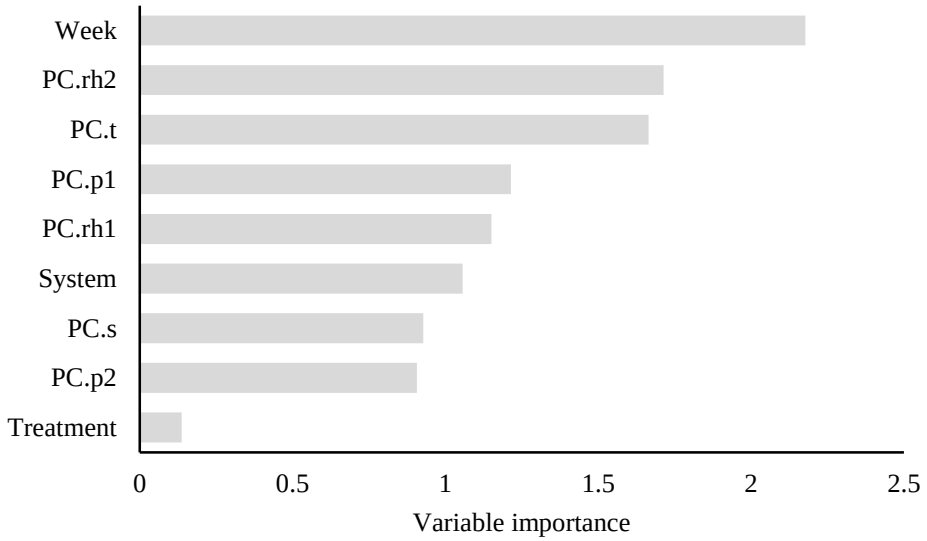

Figure 4.4 Variable importance plot of assessed factors based on random forest analysis for population level of *L. innouca* in field soil.

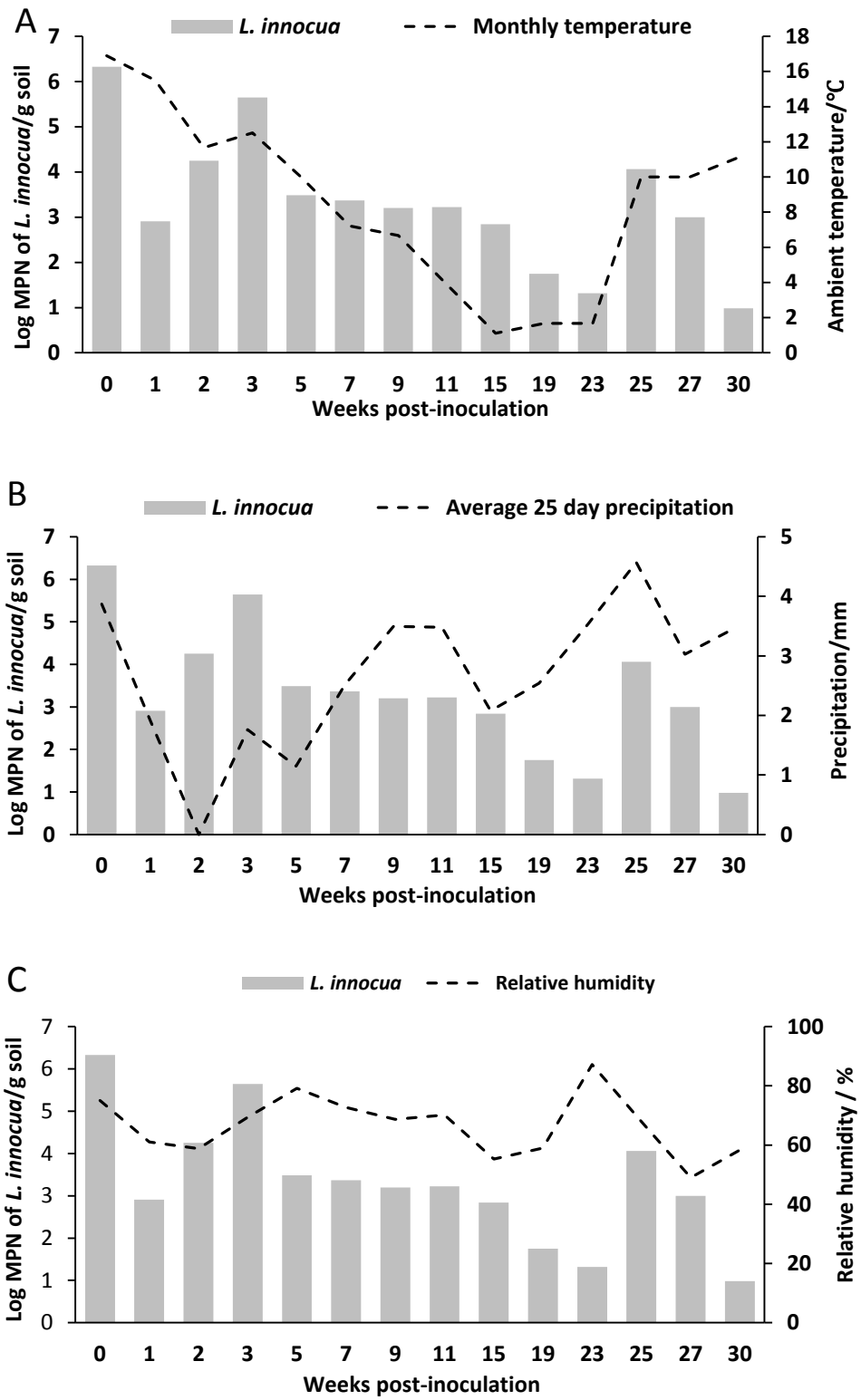

Figure 4.5 Monthly temperature (A), average 25 day precipitation (B), relative humidity on sampling day (C), and population dynamics of *L. innocua* in field soil.

4.5 Discussion

As a bacterial indicator for fecal contamination, generic *E. coli* can be shed into agricultural soil via several ways, and *Listeria* spp. is considered to be ubiquitous in the environmental and known to be associated with soil (104, 115). The survival and persistence of *E. coli* and *Listeria* spp. in agricultural soil are well documented (117, 156, 157, 158, 159) and are likely to play a major role in potential produce pre-harvest contamination. While the effect of soil properties influencing the survival and persistence of *E. coli* and *Listeria* spp. in soil has been investigated in several studies (117, 157, 158, 160, 161), the potential influence of meteorological factors has received less attention. In our study, investigation of the population dynamics of generic *E. coli* and *L. innocua* in produce field soil showed that the survival and persistence are largely dependent on meteorological conditions. We demonstrated that survival of generic *E. coli* in soil is longer when fields were exposed to higher precipitation and lower temperature. In addition, persistence of *L. innocua* in produce fields were significantly increased under higher relative humidity, precipitation, and temperature. Although not retained as a significant factor in both survival and NBR final model, bacterial survival in produce field soil seemed to be different between the two experiment fields with different field characteristics. On the other hand, cover crop application appeared to be less influential to bacterial survival and persistence in produce field soil.

In both fields, while population level of *E. coli* declined rapidly after inoculation, increasing precipitation significantly prolonged the survival of *E. coli* in soil. Precipitation is closely related with soil moisture content which has been suggested as an abiotic factor influencing *E. coli* survival in soil (69). Increasing

precipitation leads to higher moisture content in soil, which has been reported to support the survival of *E. coli* (140, 156, 162). In addition, low soil moisture content and soil dryness/desiccation may cause increasing water stress around the bacteria cells (69), and have been identified as key factors limiting the survival of *E. coli* in soil (159, 163). Similarly, moisture related factors were key determinants of *L. innocua* persistence in soil, with higher population observed under increasing precipitation and relative humidity. This finding supported the conclusions from previous studies where high soil moisture content encouraged the survival of *Listeria* spp. in soil (115, 164). In addition, *Listeria* spp. was isolated more frequently from soil with higher moisture level (104, 105) and *L. monocytogenes* survived better after one week in sealed soil samples than in unsealed soil samples, indicating water loss reduces the survival of *L. monocytogenes* in soil (117).

While increasing precipitation and higher soil moisture prolonged the survival of *E. coli* in soil, greater survival of *E. coli* was observed under lower temperature. This result is consistent with data and observations from previous studies. Sjogren (162) reported a higher die-off rate of *E. coli* in soil at 20°C and 37°C than at lower temperature (5°C and 10°C). Survival of *E. coli* was significant better at 5°C than at 25°C in three different types of soil (loam, loam sand, and sand) (140). Similarly, *E. coli* population in soil declined more rapidly at 15°C than at 6°C. In addition, colder temperature reduced the decline rate of *E. coli* in soil under both constant and alternating temperature conditions (70). As a mesophilic bacterium, *E. coli* is known to grow best at moderate temperature with optimal growth occurs at 37°C, our results indicate different temperature preference for growth and survival of *E. coli* (70).

Compared to intestinal tract of animal hosts, soil is a more hostile environment for survival of *E. coli* due to antagonism, competition of nutrients, and predation from indigenous microflora in soil (69). Survival of bacteria under such stressful environment requires metabolic change (165) and bacteria may survive better when metabolic rate is low as a result of lower temperature (117). In addition, the negative effects of indigenous soil microorganisms towards survival of *E. coli* also appears to be temperature specific. Jiang et al. (64) reported a higher decline rates of *E. coli* O157:H7 in unautoclaved soil than in autoclaved soil and a more rapid decline in unautoclaved soil at 15°C or 21°C than at 5°C. Similarly, effects of temperature on decline rates of *E. coli* were only observed in soils with the presence of indigenous microorganisms (unautoclaved soil) (163, 166). At lower temperature, growth of indigenous soil microorganism antagonistic towards *E. coli* may be suppressed and become less competitive (64). Thus, the observed better survival of *E. coli* in soil under lower temperature was likely caused by reduced activity of indigenous soil microorganisms and consequent antagonism to *E. coli*.

The NBR model identified temperature as a significant factors influencing population of *L. innocua* in produce field soil. In contrast to *E. coli*, population of *L. innocua* remained above limit of detection and showed signs of resurgence in spring when temperature increased. Temperature is one of the most dominant factors for bacteria growth and has been reported to affect the survival of *Listeria* spp. in soil. During the study period, temperature at the sampling fields decreased quickly and dropped below 5°C during winter months. The overall low temperature may have supported the persistence of *L. innocua* in soil during the study period as previous

studies have demonstrated that *L. innocua* and pathogenic *L. monocytogenes* survive better in soil at lower temperature (117, 167). The ability of *Listeria spp.* to survive in produce fields and resurge when temperature rises is of serious food safety concern. Soil is generally considered as a natural reservoir of *Listeria* spp. and presence of *Listeria* spp. in soil has been associated with contamination from manure, sewage, animal feces, and decaying vegetation (153). In addition, the pathogenic strain *L. monocytogenes* has the ability to adapt to a wide array of environments including soil and was able to maintain its pathogenicity during long-term survival, enabling its transition from soil saprophyte to a human pathogen (168, 169). In turn, field soil can act as the source for *Listeria spp.* including pathogenic strain *L. monocytogenes* and result in contamination of fresh fruit and vegetables growing in the field (170).

In univariate analyses, the survival of *E. coli* and the persistence of *L. innocua* in soil appeared to be different between the experiment fields. As shown in **Table 4.1**, the two studied fields are similar in organic matter (<1%) and have similar type of soil (loamy sand soil) but with different soil pH (organic field: 6.8; transitional organic field: 5.9). *E. coli* in soil seems to survive the best under a neutral soil pH, as higher decline rates of *E. coli* in soil were observed in more acidic soils (159, 171, 172) and in more alkaline soils (157, 161, 173). Thus, in our study, the better survival of *E. coli* observed in field A may have been due to the near neutral pH, compared the pH in field B. The two fields also have a different land use history: while field A, a certified organic field, had a history of mixed vegetable production, field B, a transitional organic field, was previously used for field corn and soy bean production. Prior use of land may pose a risk of contamination for the subsequently grown produce crops (41). As the

transitional organic field was used for cultivation of a different crop, it is possible that the field was managed differently before being used as produce cultivation field. For example, soil amendments and irrigation water not appropriate for produce may be applied or applied in an inappropriate manner for produce (41). Moreover, fields under different management history are likely to have different soil property (e.g., clay content) and different soil microbial community which has been reported to greatly influence the survival of *E. coli* and *Listeria* spp. in soil (69, 174). Activity and diversity of soil indigenous microflora has shown a negative effect on survival of *E. coli* and *Listeria*, possibly as a result of predation, competition, and antagonism (64, 68, 69, 117, 154, 174, 175). Considering its positive effect on soil microbial population and activities, cover cropping might be a way of controlling introduced enteric bacteria in soil. However, our study was not able to detect any significant effect of cover cropping on survival and persistence of *E. coli* and *L. innocua* in produce field soil. Soil properties other than pH and organic matter content were not characterized nor analyzed in our study. Factors such as clay content, availability of essential nutrients (e.g., carbon and nitrogen), and chemical properties (e.g., base cation saturation ratio) have been reported to influence the survival and persistence of *E. coli* and *L. innocua* in soil (117, 157, 158, 161). It is likely that the survival and persistence of generic *E. coli* and *L. innouca* are influenced by a combination of soil biotic and abiotic factors, farm management, and meteorological factors.

In conclusion, the results from this study indicate that lower temperature and higher precipitation increased the survival of generic *E. coli* in produce field soil. In addition, *L. innocua* was able to persist through low temperature, and higher population

levels were predicted under higher average precipitation over the previous 25 days and higher relative humidity on the sampling day. Thus, the findings suggested that pathogens may survive in soil for long period in regions with cold weather and high humidity, and point out the need for strategies to reduce possible introduction of pathogens to field soil in those regions to prevent potential contamination in produce.

Chapter 5 Risk of pre-harvest microbiological contamination in tomatoes: effects of meteorological, farm management, and environmental factors

5.1 Abstract

Tomatoes have been linked to several foodborne disease outbreaks in recent years and the source of contamination has been traced back to tomato farms. This study sought to identify and evaluate meteorological, environmental, and farm management factors affecting microbial contamination risk in tomato fruits and in tomato production environments at the pre-harvest level. Over six weeks, tomato (n = 259), irrigation water (n = 72), and field soil (n = 45) samples were collected from 24 farms located in Maryland, Delaware, and New Jersey. Local meteorological information (temperature and precipitation) during the study period were collected from National Climatic Data Center. Farm level environmental factors and management practices were acquired through questionnaires answered by farmers. These factors were evaluated for their association with aerobic plate count (APC), count of total coliform (TC), and presence of generic *E. coli* in tomato, irrigation water, and field soil samples. For tomato samples, prevalence of *E. coli* was significantly reduced by the use of potable water for cleaning, chemical application, and hand washing; however, prevalence of *E. coli* increased with increasing of precipitation on sampling day. In addition, higher TC counts occurred in tomatoes from farms exposed to higher average temperature and higher average precipitation over the previous 10 days before sampling. For field soil samples, none of the evaluated factors were found to have a significant effect on

presence of *E. coli* or APC/TC count. On the other hand, presence of *E. coli* in irrigation water samples increased with increasing average temperature over the previous 25 days before sampling day. In addition, increasing precipitation over the previous 30 days before sampling increased the count of APC and TC in irrigation water samples. Our findings suggest that microbial contamination in in tomatoes and in tomato production environments can be significantly affected by certain meteorological conditions, environmental factors, and farm management practices.

5.2 Introduction

Due to the increasing variety of tomatoes (such as grape tomatoes, and specialty/heirloom varieties), the enduring popularity of salads and bacon-lettuce-tomato (BLT) sandwiches among consumers, and American Dietary Guidelines on the health benefit of fruits and vegetables, the production and consumption of tomato have grown substantially since the 1980s and tomatoes are the second most consumed vegetables and the third most consumed fresh vegetable in the United States (2, 3, 176). However, fruits and vegetables, including tomatoes, can be a means of foodborne pathogen transmission and have been linked to 46% of foodborne illnesses occurred during year 1998-2008 in the U.S. (9). Tomatoes have been associated with 10 multistate foodborne disease outbreaks in the U.S within the last decade, and have caused a total of 806 illnesses and 78 hospitalizations (177).

Several tomato-related foodborne disease outbreaks have been traced back to contamination that occurred in the field during pre-harvest production. For example, in a multistate *Salmonella* outbreak linked to raw tomatoes, the outbreak strain was

isolated from the irrigation pond near the field (178). In a recent review, farm was identified or suspected as the most likely location of contamination for 12 out of the 14 tomato related multistate foodborne disease outbreaks during 1990-2010 (179). Pre-harvest contamination with foodborne pathogens in produce, including tomatoes, can occur via direct contact with (feces of) domestic animals or wildlife (180). In addition, enteric pathogens such *E. coli* O157:H7, *Salmonella*, and *L. monocytogenes* can be shed through domestic animals or wildlife feces into the environment (153, 181). Once introduced, these pathogens may survive, persist, and multiply in environment reservoirs such as irrigation water and field soil, which are potential sources of produce contamination (108, 182). The introduction, population dynamics (survival, persist, and growth), and transmission of pathogens in the environment can be influenced by a variety of factors. Farm management practices can significantly affect the contamination risk in produce (57, 144, 145, 147, 148). For example, application of manure has been associated with higher prevalence of generic *E. coli* in produce (57, 144), and higher prevalence of *Salmonella* and *L. monocytogenes* from environmental samples (106, 148). Landscape factors such as adjacent land use have also influence contamination of produce during pre-harvest stages. For example, farms located near potential environmental pathogen reservoirs such as animal farms and pastures have shown association with increased risk of microbial contamination in produce or field soil (57, 105). Contamination in produce fields were also linked to wildlife fecal deposits (36) and intrusion of wildlife or domestic animals (57, 180). In addition, meteorological conditions have been recognized as a factor that influences the presence and transmission of pathogens to produce production environment. Increasing rainfall

and warmer temperature were correlated with increasing prevalence and elevated level of various pathogens including *E. coli* O157:H7, *Salmonella*, and *L. monocytogenes* in environmental samples such as water (81, 123, 183) and soil (104, 105, 106). Additionally, Park et al. (133) found that temperature and precipitation directly influenced the presence and count of generic *E. coli* in spinach samples.

Preventing pre-harvest produce contamination remains a challenge because each produce farm has distinct geographic location, landscape, climate/weather condition, and farm management and practices that may influence the introduction, survival, and persistence of pathogens (148). Development of effective risk mitigation strategies at pre-harvest level requires holistic approaches that can consider a variety of potential meteorological, environmental, and farm management risk factors and evaluate their influence on potential contamination of produce crops and environmental reservoirs (e.g., irrigation water and soil). Several studies have identified and evaluated risk factors from a combination of meteorological, environmental, and farm management factors for the presence of pathogens including *L. monocytogenes* and *Salmonella* or pathogen surrogates including generic *E. coli* and *Listeria* spp. (57, 104, 105, 106, 133, 184). Most of these studies used environmental samples such as soil, water, and drag swabs of field (104, 105, 106, 184), while spinach samples were used in two studies (57, 133). Yet, no such studies are currently available for pre-harvest contamination in tomato farms. In addition, no study has been conducted to investigate if the same or different risk factors affect the microbial contamination in produce crops and in environment reservoirs (irrigation water and soil).

The objective of this study was to (i) identify specific meteorological, environmental, and farm management risk factors associated with microbial contamination in tomatoes and in tomato farm environments (irrigation water and field soil) using *E. coli* as an indicator for fecal contamination and aerobic plate count (APC) and total coliform (TC) as indicators for microbiological quality, and (ii) predict the risk of contamination in tomatoes and in tomato farm environments under different environment, meteorological conditions, and farm management practices.

5.3 Materials and methods

5.3.1 Sample collection and microbial analyses

Study design and collection and microbial analysis of tomato and environmental samples were described in details in a previous study (84). Briefly, 24 organic and conventional farms were sampled between July and September 2012 in Maryland, Delaware, and New Jersey and analyzed for indicator bacteria: *Escherichia coli*, aerobic plate count (APC), and total coliform (TC). Farms were visited every two weeks during the tomato harvest season from collection of tomato fruit, irrigation water (from deep wells), and field soil samples. All samples were aseptically collected using sterile gloves or sterile scoops (Fisher Scientific, Hampton, NH) and placed into sterile Whirl-pak bags (Nasco, Fort Atkinson,WI) or Nalgene bottles (ThermoScien- tific, Rochester, NY) from each farm at each sampling trip. All samples were sealed, transported in coolers with ice packs to the laboratories and were processed within 24 h of collection. Tomato fruit samples were mixed in a 1:1 weight by volume (w/v) ratio

with buffered peptone water (BPW) (Becton Dickenson and Company (BD), Franklin Lakes, NJ). The sample bag was hand rubbed for 2 min and incubated at room temperature for 1 h for recovery of injured cells. Rinsate from each sample was then serially diluted with 0.1% peptone water (PW) (BD) and then a 1ml aliquot of each dilution was plated onto aerobic plate count (APC) Petrifilms (3M Global Headquarters, St.Paul, MN) for enumeration of aerobic mesophiles, and *E. coli*/coliform Petrifilms (3M) for enumeration of *E. coli* and total coliforms. Petrifilms were counted after incubation at 35 °C for 24h for coliforms and 48 h for *E. coli* and APC. 10 g of field soil samples were diluted in a 1:10 w/v ratio with BPW and then serially diluted with 0.1% PW. A 1 ml aliquot of each dilution was then pipetted onto APC or *E. coli*/coliform 3M Petrifilms for enumeration. Standard membrane filtration was used for processing of irrigation water samples. Ten-fold serial dilutions were prepared and then a 100 ml and 250 ml aliquot of each original water sample and a 10 ml aliquot of each dilution was passed through a 0.45 μm mixed cellulose ester filter (Millipore, Billerica, MA) using a PALL filtration system (PALL Life Sciences, Ann Arbor, MI). The filters from the 100ml and the 10ml aliquots were then placed onto MI agar (EPA Method Number 1604) (BD followed by incubation at 37 °C for 24 h for enumeration, with blue colonies counted as *E. coli* count under ambient light and this number added to fluorescent colonies counted under long wavelength fluorescent light (365 nm) to as TC count.

5.3.2 Farm practice and environmental data

As described in (84), For each participating farm, information about on-farm practices and environmental factors was obtained by questionnaires answered by

individual farmers through one or more conversations (via email, phone or in-person). The questionnaire was designed to obtain data about on-farm practices and environmental factor that are possible risk factors influencing pre-harvest contamination. Questions included were (i) four on cultivation site characteristics (e.g., is the cultivation site fenced?); (ii) six on adjacent land use (e.g., Are the any composting operations within sight of the field?); (iii) five on water/irrigation (e.g., what is the source of water that is used for irrigation?); (iv) five on crops (e.g., what was the last crop in this location prior to current season?); (v) two on soil amendments (e.g., what type of soil amendments are applied to this field?); (vi) three on intrusion (e.g., have there been instances of human trespassers in this field in the past year?). Additional details were asked for certain questions. A complete list of questionnaire questions and options can be found in Appendices. Data and information collected from the questionnaires were entered into Excel (Microsoft, Redmond, WA), and then coded to create variables to be further used in statistical analysis and modeling.

5.3.3 Meteorological data

For each sampled farm, data on meteorological factors (air temperature and precipitation) were collected from the National Climatic Data Center (NCDC) (http://www.ncdc.noaa.gov/). Specifically, the nearest weather stations to each of the farms (based on zip codes and county information) that had air temperature and precipitation information for the study period were identified. Then, maximum, minimum, average daily air temperature, and daily precipitation were acquired for the day of sampling. In addition, the average temperature (maximum, minimum, and average daily temperature) and precipitation were calculated for various period (from

1 to 30 days) before sampling to capture any potential long term cumulative effect of meteorological factors. In total, 32 meteorological variables were collected for each farm.

5.3.4 Statistical analysis and modeling

Separate statistical analysis and modeling for presence of *E. coli* and count of APC and TC in tomato samples were performed in R software (The R Project for Statistical Computing, version 3.2.5, https://www.r-project.org/). Collected meteorological, farm management, and environmental variables were analyzed as possible predictors of *E. coli* presence in tomato using a logistic regression model. The final multivariable model was selected following the approach suggested by Collette (112): (i) each individual predictor variable was evaluated one at a time using univariate logistic regression models, and significant variables were identified at $P < 0.2$ level; (ii) significant variables from step (i) were analyzed simultaneously in a multivariable model, and a backward selection method was applied until only significant variables (at $P < 0.05$ level) were retained; (iii) each of the non-significant variables from step (ii) were evaluated using a forward selection method, to then determine the final model by omitting non-significant variables at 0.05 level. Correlation between variables were investigated by phi coefficients between two individual explanatory categorical variables and Spearman's rank coefficients when one or both of the explanatory variables were continuous. When two or more significant variables considered for multivariate analyses were correlated (correlation coefficient > 0.70 determined by Spearman's correlation coefficient), then the variables were considered one at a time in the multivariate modeling and the one that gives the best model fit was retained in

the final model. Interactions between predictor variables were investigated. In the final model, Box-Tidwell tests were used to assess the assumptions of linear relationships between explanatory variables and the transformation of outcome (log odds). The goodness of fit was assessed by Pearson and deviance Chi-square test and the Le Cessie-van Houwelingen-Copas-Hosmer test. In the final model, Variance inflation factor (VIF) was used to diagnose possible collinearities among explanatory variables.

For analysis of count data, as the distributions of positive counts of APC and TC were highly skewed, the counts were transformed to logarithm scale (base 10). Poisson regression is a standard statistical approach for count analysis, however, its assumption of equal mean and variance may limit its use when dealing with complex and over dispersed ecology data (136, 137). Negative binomial regression has been suggested and applied as an alternative approach to Poisson regression for analyzing microbial count data showing dispersion (123, 137). Thus, both Poisson regression and negative binomial models were considered to determine the association between explanatory variables and count of APC and TC in tomato samples. Suitability of models were investigated by dispersion test of the Poisson regression model and by comparison of Akaike information criterion (AIC) results. Final models for predicting APC and TC count in tomato samples were developed separately, following the same model selection approach described above.

Similarly, logistic regression and Poisson (or negative binomial models) were developed for analyzing the presence of *E. coli* and APC/TC count in environmental samples: irrigation water and field soil. Additional analyses were performed to

investigate if presence of *E. coli* and concentration of APC/TC in tomato samples were associated with those in irrigation water and field soil samples.

5.4 Results

5.4.1 Factors associated with presence of *E. coli* in tomato samples

E. coli was detected from 11 out of 259 (4%) tomato samples. Variables that are associated with presence of *E. coli* in tomato samples based on univariate analysis were listed in **Table 5.1**. Among 16 meteorological factors, two precipitation-related variables were significantly associated with presence of *E. coli* in tomatoes. Tomatoes were more likely to become contaminated with *E. coli* when exposed to higher precipitation on sampling day and higher average precipitation between sampling day and a day before. For environmental and farm management factors, *E. coli* was more likely to be isolated from samples collected from conventional farms and farms located in rural area. However, when potable water was used for equipment cleaning, chemical application, and hand washing (referred to as "use of potable water"), the odds of *E. coli* contamination in tomatoes was significantly reduced. The final multivariable model for presence of *E. coli* in tomato samples was consisted of two factors (**Table 5.2**): the use of potable water and precipitation on sampling day. According to this model, the use of potable water reduced the odds of contamination of *E. coli* in tomatoes by approximately 80% (OR=0.20), however, the odds increased by approximately 2% with every 1 mm increase of precipitation on sampling day (OR=1.02). Based on the model, the probability of tomatoes contaminated with *E. coli* (p) can be predicted as a function of these to variables: $p = 1-1/\{1+\exp(-2.33+0.02*X_1-$

1.61*X_2)}, where X_1 is the precipitation on the sampling day and X_2 is whether or not the farm uses potable water (yes noted as 1 versus no noted as 0). For example, as illustrated in **Figure 5.1**, tomatoes collected from a farm where potable water was used with no rainfall on the day of sampling are predicted to have a probability of 1.9% of getting contaminated by *E. coli*, whereas the probability increases to 11.2% when tomatoes were collected from a farm where potable water was not used and the farm was exposed to 15 mm of precipitation on the sampling day.

Table 5.1 Variables significantly associated with prevalence of generic *E. coli* in tomato samples based on univariate logistic regression.

Factors	**Coef**[a]	**SE**[b]	**OR**[c]	***P*-value**
Precipitation on the sampling day	0.02	0.01	1.02	0.026
Average precipitation over the previous 1 day before sampling	0.01	0.01	1.01	0.095
Farming system				
Conventional	0		1	
Organic	-1.55	0.79	0.21	0.050
Use of potable water				
No	0		1	
Yes	-1.58	0.63	0.21	0.011
Farm location				
Rural	0		1	
Suburban	-1.12	0.79	0.32	0.154

[a]Coef, value of the regression coefficient.
[b]SE, standard error.
[c]OR, odds ratio

Table 5.2 Final multivariable model for presence of *E. coli* in tomato samples.

Factors	**Coef**[a]	**SE**[b]	**OR**[c]	***P*-value**
Precipitation on the sampling day	0.02	0.01	1.02	0.024
Use of potable water				
No	0		1	
Yes	-1.61	0.64	0.20	0.012

[a]Coef, value of the regression coefficient.
[b]SE, standard error.
[c]OR, odds ratio

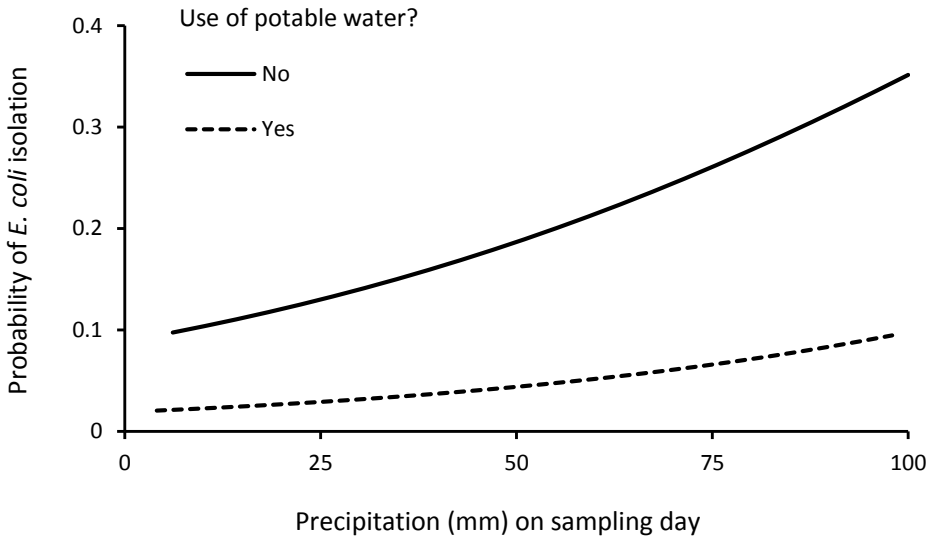

Figure 5.1 Predicted probability of isolation of generic *E. coli* in tomato samples influenced by precipitation on sampling day and the use of potable water for cleaning, chemical application, and hand washing.

5.4.2 Factors associated with APC and TC counts on tomato samples

The dispersion test indicated no significant sign of over dispersion for both Poisson regression models (intercept only) for APC and TC (*P*-value >0.05). The Poisson models also had lower AIC values compared to negative binomial models.

Thus, Poisson regression was selected as the approach for univariate and multivariable modeling of APC and TC counts in tomato samples. For APC, none of the 16 meteorological variables nor the 25 environmental and farm management variables were found significant in univariate analysis. For count of TC in tomatoes, significant variables from univariate analysis were shown in **Table 5.3**. Twelve of the 16 meteorological variables, including 8 temperature variables and 4 precipitation variables were significantly associated with TC count in tomatoes. In general, TC counts in tomatoes increased when tomatoes were exposed to higher precipitation and higher temperature. In addition to meteorological variables, TC counts in tomatoes were significantly reduced when the fields were completely fenced. The final multivariable model for TC in tomatoes included two meteorological factors: the average daily precipitation over the previous 10 days before sampling (relative risk [RR] = 1.02); and the mean average temperature over the previous 10 days before sampling (RR = 1.08) (**Table 5.4**). As shown in **Figure 5.2**, according to the model, the predicted counts of TC (c) can be calculated using the equation: $c = \exp(-1.08+0.07*X_1-0.02*X_2)$, where X_1 is the mean average temperature over the previous 10 days (°C) before sampling and X_2 is the average daily precipitation over the previous 10 days before sampling (mm). For example, if tomatoes were collected from a farm with no rainfall and an average temperature of 20°C over the previous 10 days, the counts of TC on tomatoes are expected to be 1.50 log CFU/g. If over the previous 10 days the average daily precipitation was 10 mm and average temperature was 30°C, the predicted TC count increases to 3.76 log CFU/g.

Table 5.3 Significant variables for TC count in tomato samples based on univariate analysis.

Factors	**Coef**[a]	**SE**[b]	**RR**[c]	***P-value***
Average temperature on the sampling day	0.032	0.014	1.032	0.027
Average daily temperature over the previous 1 day before sampling	0.029	0.016	1.029	0.066
Average daily temperature over the previous 5 day before sampling	0.038	0.024	1.041	0.105
Average daily temperature over the previous 10 day before sampling	0.056	0.026	1.058	0.029
Average daily temperature over the previous 15 day before sampling	0.05	0.028	1.052	0.065
Average daily temperature over the previous 20 day before sampling	0.052	0.029	1.053	0.074
Average daily temperature over the previous 25 day before sampling	0.054	0.03	1.055	0.077
Average daily temperature over the previous 30 day before sampling	0.059	0.03	1.06	0.052
Average precipitation over the previous 1 day before sampling	0.002	0.001	1.003	0.068
Average precipitation over the previous 5 day before sampling	0.006	0.004	1.006	0.109
Average precipitation over the previous 10 day before sampling	0.012	0.007	1.012	0.059
Average precipitation over the previous 15 day before sampling	0.013	0.009	1.013	0.134
Is the field fenced?				
No	0		1	
Partially	-0.034	0.113	0.966	0.759
Completely	-0.355	0.156	0.701	0.003

[a]Coef, value of the regression coefficient.
[b]SE, standard error.
[c]RR, relative risk

Table 5.4 Final multivariable model for TC count in tomato samples.

Factors	**Coef**[a]	**SE**[b]	**RR**[c]	***P*-value**
Average precipitation over the previous 10 day before sampling	0.02	0.01	1.02	0.009
Average temperature over the previous 10 day before sampling	0.07	0.03	1.08	0.005

[a]Coef, value of the regression coefficient.
[b]SE, standard error.
[c]RR, relative risk

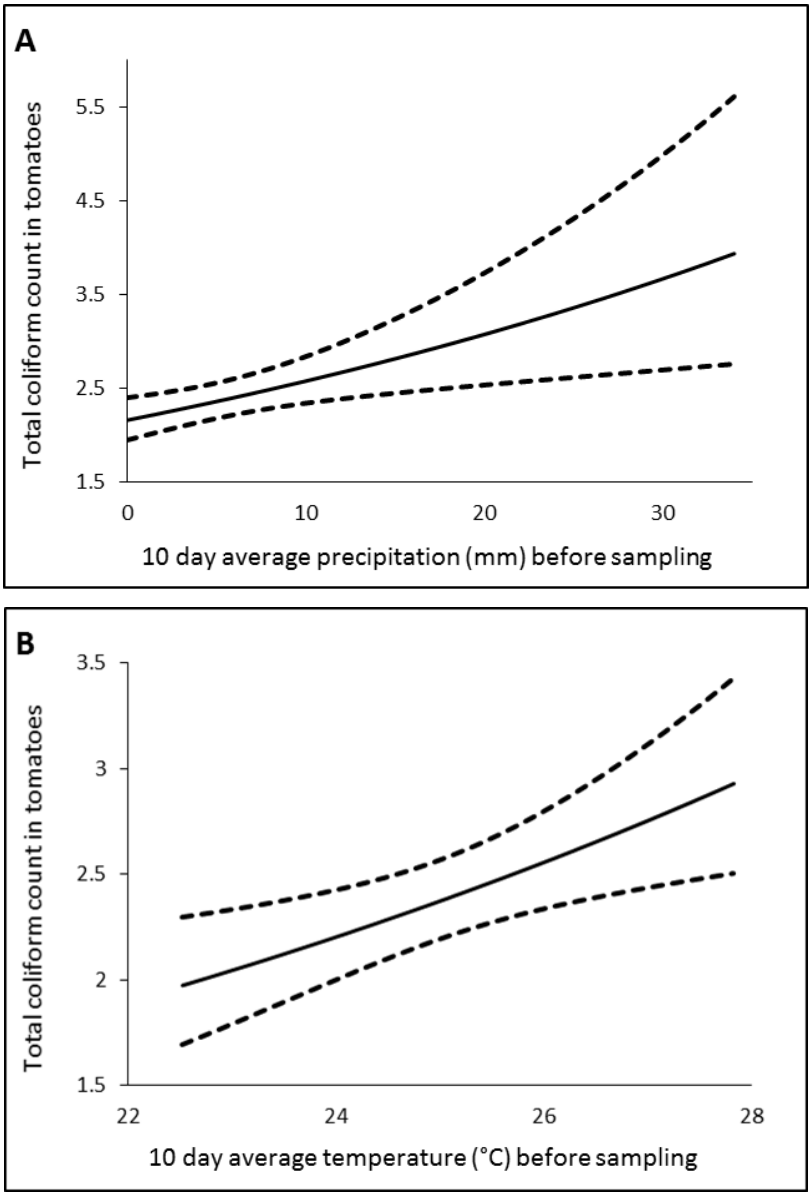

Figure 5.2 Predicted total coliform count in tomato samples as affected by average precipitation (mm) over the previous 10 days before sampling (A) and average temperature (°C) over the previous 10 days before sampling (B).

5.4.3 Factors associated with presence of *E. coli* and count of APC/TC in environmental samples.

Meteorological, environmental, and farm management variables were investigated separately for their potential association with presence of *E. coli*, concentration of APC, and TC count in field soil samples. However, none of the evaluated variables were found significant. For irrigation water samples collected from deep wells, seven temperature variables were significantly associated with presence of *E. coli*, two precipitation variables were significantly associated with APC count, and one temperature and five precipitation variables were significantly associated with TC count in univariate analyses (**Table 5.5**). The final model presence of *E. coli* in irrigation water samples included one single temperature variable and showed that the odds of *E. coli* contamination in irrigation water increased with increasing average temperature over the previous 25 days before sampling (OR=8.63). Similarly, the final models for count of APC and TC indicated that APC count as well as TC count in irrigation water samples increased with increasing precipitation over the previous 30 days before sampling (RR=1.09 and 1.12 respectively).

5.4.4 Association between presence of *E. coli* and count of APC/TC in tomato and environmental samples.

In general, tomato and environmental samples showing higher levels of APC and TC counts were more likely to contain *E.* coli. Specifically, in tomato samples, APC count was borderline significantly associated with presence of *E. coli* ($P = 0.07$), while in field soil samples TC count was borderline significantly associated with presence of *E. coli* ($P = 0.09$). In irrigation water samples, both APC and TC count were borderline significantly associated with presence of *E. coli* ($P = 0.07$ and $P =$

0.06). Across different sample types, APC count in tomato samples was borderline significantly associated with APC count in field soil samples (P = 0.07). No other significant associations between presence of *E. coli* or count of APC/TC in tomato samples and environmental samples were found.

Table 5.5 Factors associated with presence of *E. coli*, aerobic plate count, and total coliform count in irrigation water samples based on univariate analyses.

Factors[a]	Presence of *E. coli*			Aerobic plate count			Total coliform count		
	Coef[b]	OR[c]	*P*	Coef	RR[d]	*P*	Coef	RR	*P*
Tam1	0.81	2.25	0.089	–	–	–	–	–	–
Tam5	0.75	2.11	0.156	–	–	–	–	–	–
Tam10	1.43	4.19	0.126	–	–	–	–	–	–
Tam15	1.97	7.14	0.067	–	–	–	–	–	–
Tam20	1.85	6.36	0.073	–	–	–	0.14	1.15	0.177
Tam25	2.16	8.63	0.053	–	–	–	–	–	–
Tam30	1.95	7.06	0.063	–	–	–	–	–	–
Pm10	–	–	–	–	–	–	0.03	1.012	1.03
Pm15	–	–	–	–	–	–	0.04	1.013	0.116
Pm20	–	–	–	–	–	–	0.06	1.06	0.07
Pm25	–	–	–	0.07	1.07	0.047	0.09	1.09	0.026
Pm30	–	–	–	0.09	1.09	0.024	0.11	1.12	0.008

[a]Tam1-Tam15 denotes average temperature in the previous 1 to 15 days before sampling respectively; Pm10-Pm30 denotes average precipitation in the previous 1 to 15 days before sampling respectively.
[b]Coef, value of regression coefficient; –, indicates non-significance.
[c]OR, odds ratio.
[d]RR, relative risk.

5.5 Discussion

This study provided a systematic assessment on the influence of a variety of factors on the prevalence and concentration of food safety indicator bacteria from tomato and farm environmental samples. The results indicate that microbial

contamination in tomatoes can be jointly affected by meteorological, environmental, and farm management factors. In environmental samples from tomato farms, risk of microbial contamination was predominantly influenced by meteorological factors, suggesting that different risk factors were associated with contamination in tomatoes and contamination in tomato farm environments such as irrigation water and field soil.

In univariate analysis of tomato samples, farm location, farming system, and use of potable water were identifies as three environmental and farm management risk factors affecting the presence of *E. coli*. Farms located in rural area were predicted to have higher prevalence of *E. coli* in tomatoes than those located in suburban area. Farms located in rural area are more likely to be surrounded by wildlife habitat such as forest, animal operations such dairy barns, pasture lands for grazing of domestic animals, and open water bodies such as pond, creek, or stream that may attract domestic and wildlife animals. Thus, farms located in rural area are likely to have higher frequencies of domestic and wildlife presence and intrusion. Indeed, based on the response from farmers, while all the farms (22/23, with one missing observation) located in rural or suburban area reported wildlife intrusion, domestic animal intrusion was more frequently observed in farms located in rural area (36%, 5/14) than farms located suburban area (30%, 3/10). The presence of domestic animals and wildlife in fields may pose a risk of produce contamination because their feces are a known reservoir for foodborne pathogens such *E. coli* O157:H7 (36, 185, 186). These pathogens can be shed into the environments through defecation and can be a direct source of contamination in produce crops in the fields. Domestic animals and wildlife may also be involved in the contamination of produce indirectly by fecal contamination

of the environment around produce fields such as surface water or field soil. In addition, our study shows that complete fencing of cultivating field significantly reduced population level of TC in tomato samples, indicating the usefulness of fencing as a possible preventive measure for contamination of microorganisms including pathogens. Fence, as a physic barrier, can prevent domestic animals or wildlife from entering produce fields, reduce the frequency of animal intrusions, and therefore reduce the risk of contamination of produce. It is likely that tomatoes in our study were contaminated by domestic animal or wildlife intrusions to cultivation fields. Control measure could be focused on preventing animal intrusions, especially in rural areas.

The use of potable water for equipment cleaning, chemical application, and hand washing was shown to reduce the odds of isolation of *E. coli* from tomato samples. Microbial quality of water for agricultural uses has great impact on microbial food safety of produce. Irrigation water has been identified by numerous studies as a major source and a route for transmission of pathogens to produce fields and produce crops (41, 51, 53, 187, 188). Microbial quality of water used for other purposes such as application of foliar treatments (e.g., pesticides and fertilizer sprays) other than irrigation during produce production received less attention and their potential influence on produce contamination is not well understood. Although of low volume, microbial quality of water used for these purposes is of food safety concern because the water is in direct contact with edible part of the produce crops (e.g., tomato fruits or lettuce leaves) (51). The investigation of a *Cyclospora* outbreak linked to raspberry has suspected water used in pesticide solutions as the most likely source for this outbreak and potable water was recommended for use in pesticide solutions and

handwashing (189). Source of water used as solutions for foliar application is not well documented. In this study, majority (75%, 18/24) of the participating farms use potable water for chemical applications. Among the remaining six farms, five farms use water from irrigation well and one farm use water from stream. Surface water and ground well water can become contaminated via a variety of ways such as fecal contamination from domestic animals and wildlife, run-off from animal production or composting facilities, discharge of waste water, and leakage from defective septic systems (41, 51). Microbial analysis shows that 70% (21/30) of the surface water and 2.8% (2/72) of the ground water samples collected during our study were positive for fecal contamination indicator *E. coli*. Our analysis suggest that water used as solutions in produce foliar chemical application may be a source of contamination in tomatoes during pre-harvest production.

Meteorological factors temperature and precipitation were found to be associated with contamination of indicator bacteria in tomatoes samples and in irrigation water samples. Higher prevalence of *E. coli* was expected when fields were exposed to higher precipitation on the day of sampling. Previous studies have shown that a rain event can cause contamination in produce through splash (89) and aerosols formed after rain (190). Rain also increases the humidity which has been reported to be supportive for growth and survival of microorganisms on produce surfaces (146, 191). In irrigation water samples, higher prevalence of *E. coli* was found to be associated with higher monthly temperature before sampling. This finding supported previous studies which observed higher prevalence and concentration of generic *E. coli* during warmer months in surface water (79, 80). Increasing precipitation has also been

associated with increasing prevalence and concentration of foodborne pathogens in irrigation water source (79, 81). However, in our study no significant associations were found between 8 evaluated precipitation variables and presence of *E. coli* in irrigation water samples. In general, precipitation and temperature were positively associated with APC and TC count in tomato and irrigation water samples, indicating that warmer temperature and higher precipitation may favor the persistence of microorganisms in tomatoes or irrigation water.

None of the evaluated meteorological, environmental, and farm management factors were significantly associated with presence of *E. coli* and APC and TC counts. This is possibly due to the relatively small sample size (n = 45) and low number of *E. coli* positive samples (n = 4). Nevertheless, the presence and persistence of foodborne pathogens in soil is a food safety concern as soil can serve as a source of foodborne pathogens and a vehicle for transmission of pathogen to produce. Additional research is needed to identify potential risk factors that affect the presence and persistence of food safety related microorganisms in produce field soil.

In conclusion, this study identified increasing precipitation and increasing temperature as risk factors for increasing the probability and extent of microbial contamination in tomato and irrigation water. Additionally, domestic/wildlife animal intrusions and water used for application of foliar treatment were identified as two possible routes of contamination in produce. The findings suggest that microbial contamination in tomatoes and in tomato production environments can be significantly affected by certain meteorological conditions, environmental factors, and farm management practices. Moreover, this study provided information that will help

growers in evaluating their farm management and preventive measures such as fencing of tomato cultivation fields and using potable water for chemical application to reduce potential contamination in produce at the pre-harvest level.

Chapter 6 Summary and future studies

6.1 Summary

Produce has been frequently associated with foodborne disease outbreaks and a significant number of illnesses, hospitalizations, and deaths in the U.S and worldwide. Many of these produce-related outbreaks were suspected to be caused by contamination occurred at the pre-harvest level. This dissertation systematically investigated a combination of meteorological, environmental, and farm management factors that may influence the presence, survival, and persistence of food safety related microorganisms in produce and in produce production environments, and provided valuable information to aid the development of risk mitigation strategies during pre-harvest production.

Chapter 2 was focused on the effects of meteorological factors on the presence of *Listeria* spp. in a mixed farm, and precipitation and wind speed were identified as the two important risk factors. Potential run-off after rainfall from animal farms or composting facilities and wind-blown contaminated dust were suggested as two possible routes of contamination in the mixed farm. Furthermore, the developed logistic regression and random forest models showed solid predictive ability and can be used to predict the risk of contamination in a mixed farm under different weather conditions.

Following the modeling framework that was developed in chapter 2, chapter 3 was extended to further data analyses by using bacterial prevalence data as well as bacteria count data. Using regression and tree-based methods, meteorological factors affecting the presence and concentration of generic *E. coli* in samples collected from a mixed produce and dairy farm were identified.

Soil can serve as a reservoir for enteric pathogens and their survival and persistence in soil pose a risk of subsequent microbial contamination in produce. Understanding the factors that affect the survival and persistence of pathogens in soil is critical in developing mitigation strategies to reduce such contamination risk. Thus, chapter 4 evaluated the effect of a particular farm management practice, cover cropping, along with meteorological factors on the survival and persistence of food safety indicator bacteria, generic *E. coli* and *Listeria innocua*, in field soil. According to the developed model, predicted survival of *E. coli* in soil is better when fields were exposed to higher precipitation and lower temperature. *L. innocua* was persistent through the study period, and higher population levels were observed when fields were exposed to higher temperature, precipitation, and relative humidity.

Chapter 5 provided a systematic assessment of the combination of meteorological, environmental, and farm management factors. Models were developed separately for tomatoes samples and environmental samples (irrigation water and field soil) to identify potential risk factors and predict the contamination risk. In general, meteorological factors were found to influence the prevalence of *E. coli* and APC/TC count in both tomato samples and environmental samples, with higher prevalence and higher count expected under higher temperature and precipitation. In addition, farm management practices, specifically fencing and use of potable water for equipment cleaning, chemical application, and hand washing were found to significantly reduce contamination in tomato samples. Based on the findings, microbial contamination in tomatoes at pre-harvest level can be significantly affected by meteorological, environmental, and farm management practices, indicating these factors should be

considered together when evaluating good agricultural practices and developing risk mitigation strategies.

Collectively, the dissertation provided a systematic assessment on the effects of a variety of factors that influence the risk of microbial contamination in produce at the pre-harvest level. Potential risk factors were identified and models were developed to predict the microbial contamination risk under different meteorological conditions, farm environment, and farm management practices. Information from this project provided data and tools that may allow growers to make better informed food safety decisions.

6.2 Future studies

This dissertation represents the development and use of various statistical analyses and modeling approaches to understand and predict the presence, survival, and persistence of enteric bacteria of food safety concern in produce pre-harvest environment under different conditions. Some possible areas of future research are proposed as follows.

(1) More extensive sampling and surveys on the microbiological quality (i.e., presence and concentration of indicator bacteria or pathogens) of produce and produce environmental samples are needed. Data currently available have low numbers of positive samples which may affect the predictability of the statistical models.

(2) More studies are needed to investigate the use of bacteria indicators such as *E. coli, L. innocua,* aerobic plate count, or total coliform and to determine if they can be used to make inferences on foodborne pathogens such as *E. coli* O157:H7, *Salmonella,* and *L. monocytogenes* that are frequently associated with produce contamination. For this purpose, more extensive sampling is needed to possibly acquire sufficient pathogen positive samples as contamination of pathogens in produce seems rare.

(3) Current studies on pre-harvest risk factors associated with produce contamination were conducted in limited regions. Thus, the results and findings may not be applicable to other produce growing regions. More studies are needed to determine if any region specific factors may influence the contamination risk in produce pre-harvest environment.

(4) More comprehensive information on farm-level environmental factors and farm management practices will be helpful in identify other potential influential factors on produce contamination. Alternative information collection methods (more "objective" measures) other than questionnaires are needed to better collect information on farm-level environmental factors such as wildlife intrusions and farm management factors such as irrigation frequency to improve the accuracy of information and to avoid individual bias.

(5) Validation studies of the developed models are needed. After validation, user-friendly tools can be developed based on the developed models, which

would allow produce grower to predict the risk of microbial contamination under different weather conditions and farm-level environmental factors or farm practices, and then subsequently make informed food safety decisions.

(6) Microbial sampling studies, and collection of farm environmental and farm management information and data targeted for other produce products are needed. When become available, these information and data could be used to analyze potential risk factors for microbial contamination in pre-harvest environment of other produce products following the modeling frameworks developed in this project.

Appendices

Questionnaire in Chapter 4

I. Cultivation Site Characteristics

Question	Answer	Comments
How would you best describe the composition of the soil?	_____ Sandy _____ Loamy _____ "Mucklands" _____ Other	If "Other", please describe.
How would you best describe the area in which the cultivation site is located? (Pick one only.)	_____ Rural _____ Suburban _____ Urban _____ Mixed use	
Is the cultivation sited "fenced"	____Yes ____ No	If yes, approximately how high is the fence? _____ feet
Is there a history of flooding at this cultivation site?	____Yes ____ No	If yes, please estimate the percentage of the field that is flooded. _____%

II. Adjacent Land Use

Question	Answer	Comments
Are there any compost operations?	____Yes ____ No	If yes, approximately what is the distance? _____ feet
Are there any open water bodies (e.g., ponds, streams, canals)?	____Yes ____ No	If yes, approximately what is the distance? _____ feet
Are there any homes, dwellings, parks, or other foci of human activity?	____Yes ____ No	If yes, approximately what is the distance? _____ feet
Are there any confined animal feeding operations?	____Yes ____ No	If yes, approximately what is the distance? _____ feet

		If yes, is the location of the feed operation uphill or downhill from the cultivation site? _____
Are there adjacent lands devoted to grazing of domestic animals?	____Yes ____ No	If yes, approximately what is the distance? _____ feet If yes, is the location of the feed operation uphill or downhill from the cultivation site? _____
Are there woodlands, fallow fields, or other uncultivated areas?	____Yes ____ No	If yes, approximately what is the distance? _____ feet

III. *Water/Irrigation*

Question	Answer	Comments
What type of irrigation is employed at this cultivation site? (Check all that apply)	_____ Overhead _____ Flood/furrow _____ Drip _____ Other _____ Do not irrigate	
What is the source of water that is used for irrigation? (Check all that apply.)	_____ Pond water or stream _____ Shallow wells _____ Deep wells _____ "Reclaimed" water _____ Municipal water _____ Do not know _____ Other	If other, please specify source:
Is the irrigation water treated in any fashion before it is applied?	____Yes ____ No	If yes, please specify type of treatment:
Is there a readily available source of potable water for equipment cleaning, pesticide application,	____Yes ____ No	

nutrient applications, etc?		
How often is the source of irrigation tested?	____ Once per year _____ At the beginning of each growing season _____ Multiple times during the growing season _____ After major weather incidents _____ Other	If other, please specify the frequency of testing?

IV. Crops

Question	Answer	Comments
What is the tomato crop currently being cultivated in this location?	_____ Round Red _____ Plum _____ Cherry _____ Grape _____ Other variety	If "Other variety", please specify:
What crop was cultivated in this location prior to the current season?	_____Tomatoes _____ Vegetable _____ Fruit _____ Cereal grain _____ Corn _____ Other _____ Not planted _____ Used for foraging	If "Other," please specify:
How many growing seasons in the last five (including the current year) have tomatoes been cultivated in this location?	_____ 1 _____ 2 _____ 3 _____ 4 _____ 5	

V. Soil Amendments

Question	Answer	Comments
What types of soil amendments are applied at this location? (Check all that apply.)	_____ Inorganic fertilizers _____ Raw or partially composted animal manure or	If other, please specifiy:

	_____ Raw or partially composted "green waste" _____ Fully composted animal manure _____ Fully composted green waste _____ Compost teas _____ Fish or blood meals _____ Other	
Are non-inorganic soil amendments tested for microbial contaminants?	____Yes ____ No	If yes, please check of the following that apply. _____ Coliforms _____ Fecal coliforms _____ Generic *Escherichia coli* _____ *Escherichia coli* O157:H7 _____ *Salmonella* _____ Other (please specify)

VI. Intrusions

Question	Answer	Comments
Is there history of domestic animal intrusions at this cultivation site?	____Yes ____ No	If yes, please check all that apply. _____ Cattle, Dairy cows _____ Pigs _____ Poultry _____ Sheep, goats _____ Other (Please specify)
Is there history of wild animal intrusions at this cultivation site?	____Yes ____ No	If yes, please check all that apply. _____ Deer _____ Birds _____ Rabbits _____ Pigs _____ Foxes, Coyotes, Wild dogs _____ Reptiles _____ Insects _____ Other (Please specify)
Is there a history of human intrusions at this cultivation site?	____Yes ____ No	If yes, please describe nature of the intrusions.

Reference

1. Mercanoglu Taban, B., and A. K. Halkman. 2011. Do leafy green vegetables and their ready-to-eat [RTE] salads carry a risk of foodborne pathogens? *Anaerobe* 17:286–287.
2. U.S. Department of Health and Human Services and U.S. Department of Agriculture. 2015. 2015 – 2020 Dietary Guidelines for Americans. Available at: http://health.gov/dietaryguidelines/2015/guidelines/ Accessed 10 March 2017.
3. U.S. Department of Agriculture. 2015. Food availability (per capita) data system. Available at: http://www.ers.usda.gov/data-products/food-availability-(per-capita)-data-system/.aspx. Accessed 10 March 2017.
4. Produce for Better Health Foundation. 2015. Slate of the plate: 2015 Study on america's consumption of fruit and vegetables. Available at: http://www.pbhfoundation.org/pdfs/about/res/pbh_res/State_of_the_Plate_2015_WEB_Bookmarked.pdf. Accessed 10 March 2017.
5. Cook, R. L. 2011. Fundamental forces affecting U.S. fresh produce growers and marketers. *Choices*. Available at: http://ucce.ucdavis.edu/files/datastore/234-2158.pdf. Accessed 10 March 2017.
6. 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.
7. Scallan, E., P. M. Griffin, F. J. Angulo, R. V. Tauxe, and R. M. Hoekstra. 2011. Foodborne illness acquired in the United States-unspecified agents. *Emerg. Infect. Dis.* 17:16-22.

8. Scharff, R. L. 2012. Economic burden from health losses due to foodborne illness in the United States. *J. Food Prot.* 75:123-131.

9. Painter, J., R. M. Hoekstra, T. Ayers, R. V. Tauxe, C. R. Braden, F. J. Angulo, and P. M. Griffin. 2013. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerg. Infect. Dis.* 19:407-415.

10. Centers for Disease Control and Prevention. 2011. Multistate outbreak of *Salmonella* Panama infections linked to cantaloupe (final update). Available at: https://www.cdc.gov/salmonella/2011/cantaloupes-6-23-2011.html. Accessed 10 March 2017.

11. Centers for Disease Control and Prevention. 2012. Multistate outbreak of *E. coli* O157:H7 infections linked to romaine lettuce (final update). Available at: https://www.cdc.gov/ecoli/2011/romaine-lettace-3-23-12.html. Accessed 10 March 2017.

12. Centers for Disease Control and Prevention. 2012. Multistate outbreak of shiga toxin-producing *Escherichia coli* O157:H7 Infections linked to organic spinach and spring mix blend (final update). Available at: https://www.cdc.gov/ecoli/2012/O157H7-11-12/index.html. Accessed 10 March 2017.

13. Centers for Disease Control and Prevention. 2012. Multistate outbreak of *Salmonella* Braenderup infections associated with mangoes (final update). Available at: https://www.cdc.gov/salmonella/braenderup-08-12/index.html. Accessed 10 March 2017.

14. Centers for Disease Control and Prevention. 2012. Multistate outbreak of *Salmonella* Typhimurium and *Salmonella* Newport infections linked to cantaloupe (final update). Available at: https://www.cdc.gov/salmonella/typhimurium-cantaloupe-08-12/index.html. Accessed 10 March 2017.

15. Centers for Disease Control and Prevention.. 2013. Multistate outbreak of shiga toxin-producing *Escherichia coli* O157:H7 infections linked to ready-to-eat salads (final update). Available at: https://www.cdc.gov/ecoli/2013/O157H7-11-13/index.html. Accessed 10 March 2017.

16. Centers for Disease Control and Prevention. 2013. Multistate outbreak of *Salmonella* Saintpaul infections linked to imported cucumbers (final update). Available at: https://www.cdc.gov/salmonella/saintpaul-04-13/index.html. Accessed 10 March 2017.

17. Centers for Disease Control and Prevention. 2015. Multistate outbreak of *Salmonella* Enteritidis infections linked to bean sprouts (final update). Available at: https://www.cdc.gov/salmonella/enteritidis-11-14/index.html. Accessed 10 March 2017.

18. Centers for Disease Control and Prevention. 2015. Wholesome soy products, inc. sprouts and investigation of human listeriosis cases (final update). Available at: https://www.cdc.gov/listeria/outbreaks/bean-sprouts-11-14/index.html. Accessed 10 March 2017.

19. Centers for Disease Control and Prevention. 2015. Multistate outbreak of listeriosis linked to commercially produced, prepackaged caramel apples made

from bidart bros. apples (final update). Available at: https://www.cdc.gov/listeria/outbreaks/caramel-apples-12-14/. Accessed 10 March 2017.

20. Centers for Disease Control and Prevention. 2016. Multistate outbreak of listeriosis linked to packaged salads produced at springfield, Ohio Dole processing facility (final update). Available at: https://www.cdc.gov/listeria/outbreaks/bagged-salads-01-16/index.html. Accessed 10 March 2017.

21. Centers for Disease Control and Prevention. 2016. Multistate outbreak of listeriosis linked to frozen vegetables (final update). Available at: https://www.cdc.gov/listeria/outbreaks/frozen-vegetables-05-16/index.html. Accessed 10 March 2017.

22. U.S. Food and Drug Administration. 2012. Bad bug book: Foodborne pathogenic microorganisms and natural toxins handbook. Available at: http://www.fda.gov/downloads/Food/FoodSafety/FoodborneIllness/FoodborneIllnessFoodbornePathogensNaturalToxins/BadBugBook/UCM297627.pdf. Accessed 10 March 2017.

23. Riley, L. W., R. S. Remis, S. D. Helgerson, H. B. McGee, J. G. Wells, B. R. Davis, R. J. Hebert, E. S. Olcott, L. M. Johnson, N. T. Hargrett, P. A. Blake, and M. L. Cohen. 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* 308:681-685

24. Centers for Disease Control and Prevention. 2006. Multistate outbreak of *E. coli* O157:H7 infections linked to fresh spinach. Available at:

https://www.cdc.gov/ecoli/2006/spinach-10-2006.html. Accessed 10 March 2017.

25. Anderson, M., L. A. Jaykus, S. Beaulieu, and S. Dennis. 2011. Pathogen-produce pair attribution risk ranking tool to prioritize fresh produce commodity and pathogen combinations for further evaluation (P3ARRT). *Food Control* 22:1865-1872.

26. Herman, K. M., A. J. Hall, and L. H. Gould. 2015. Outbreaks attributed to fresh leafy vegetables, United States, 1973-2012. *Epidemiol. Infect.* 143:3011-3021.

27. Centers for Disease Control and Prevention. 2012. Multistate outbreak of listeriosis linked to whole cantaloupes from jensen farms, Colorado. Available at: http://www.cdc.gov/listeria/outbreaks/cantaloupes-jensen-farms/index.html. Accessed 10 March 2017.

28. U.S. Food and Drug Administration. 2016. fda investigated *Listeria* outbreak linked to frozen vegetables. Available at: http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm499157.htm. Accessed 10 March 2017.

29. U.S. Food and Drug Adminstration. 2013. FDA investigation summary: multistate outbreak of *Salmonella* Typhimurium and *Salmonella* Newport infections linked to cantaloupe grown at chamberlain farms in southwest Indiana. Available at: https://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm315879.htm. Accessed 10 March 2017.

30. U.S. Food and Drug Administration. 2013. FDA investigation summary -

multistate outbreak of *E. coli* O157:H7 illnesses linked to ready-to-eat salads. Available at: https://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm374327.htm. Accessed 10 March 2017.

31. Berg, J., T. McAllister, S. Bach, R. Stilborn, D. Hancock, and J. LeJeune. 2004. *Escherichia coli* O157:H7 excretion by commercial feedlot cattle fed either barley- or corn-based finishing diets. *J. Food Prot.* 67:666-671.

32. Stanford, K., D. Croy, S. J. Bach, G. L. Wallins, H. Zahiroddini, and T. A. McAllister. 2005. Ecology of *Escherichia coli* O157:H7 in commercial dairies in southern Alberta. *J. Dairy Sci.* 88:4441-4451.

33. Callaway, T. R., M. A. Carr, T. S. Edrington, R. C. Anderson, and D. J. Nisbet. 2009. Diet, *Escherichia coli* O157:H7, and cattle: a review after 10 years. *Curr. Issues Mol. Biol.* 11:67-79.

34. Branham, L. A., M. A. Carr, C. B. Scott, and T. R. Callaway. 2005. *E. coli* O157 and *Salmonella* spp. in white-tailed deer and livestock. *Curr. Issues Intest. Microbiol.* 6:25-29.

35. Doane, C. A., P. Pangloli, H. A. Richards, J. R. Mount, D. A. Golden, and F. A. Draughon. 2007. Occurrence of *Escherichia coli* O157:H7 in diverse farm environments. *J. Food Prot.* 70:6–10.

36. Jay, M. T., M. Cooley, D. Carychao, G. W. Wiscomb, R. A. Sweitzer, L. Crawford-Miksza, J. A. Farrar, D. K. Lau, J. O'Connell, A. Millington, R. V Asmundson, E. R. Atwill, and R. E. Mandrell. 2007. *Escherichia coli* O157:H7 in feral swine near spinach fields and cattle, central California coast. *Emerg.*

Infect. Dis. 13:1908-1911.

37. La Ragione, R. M., A. Best, D. Clifford, U. Weyer, L. Johnson, R. N. Marshall, J. Marshall, W. A. Cooley, S. Farrelly, G. R. Pearson, and M. J. Woodward. 2006. Influence of colostrum deprivation and concurrent Cryptosporidium parvum infection on the colonization and persistence of *Escherichia coli* O157 : H7 in young lambs. *J. Med. Microbiol.* 55:819-828.

38. Alam, M. J., and L. Zurek. 2004. Association of *Escherichia coli* O157:H7 with houseflies on a cattle farm. *Appl. Environ. Microbiol.* 70:7578-7580.

39. Nielsen, E. M., M. N. Skov, J. J. Madsen, J. Lodal, J. B. Jespersen, and D. L. Baggesen. 2004. Verocytotoxin-producing *Escherichia coli* in wild birds and rodents in close proximity to farms. *Appl. Environ. Microbiol.* 70:6944-6947.

40. De Jesús, A. J., A. R. Olsen, J. R. Bryce, and R. C. Whiting. 2004. Quantitative contamination and transfer of *Escherichia coli* from foods by houseflies, Musca domestica L. (Diptera: Muscidae). *Int. J. Food Microbiol.* 93:259-262.41. FAO/WHO. 2008. Microbiological Hazards in Fresh Leafy.

42. Food and Agriculture Organization / World Health Organization. 2008. Microbiological hazards in fresh leafy vegetables and herbs. Available at: http://www.fao.org/3/a-i0452e.pdf. Accessed 10 March 2017.

43. U.S. Department of Agriculture. 2013. Farm and ranch irrigation survey. Available at: https://www.agcensus.usda.gov/Publications/2012/Online_Resources/Farm_and_Ranch_Irrigation_Survey/. Accessed 10 March 2017.

44. Gerba, C. P. The role of water and water testing in produce safety, p. 129-142.

In Microbial safety of fresh produce (eds X. Fan, B. A. Niemira, C. J. Doona, F. E. Feeherry and R. B. Gravani). Wiley-Blackwell, Oxford, UK.

45. Pachepsky, Y., D. R. Shelton, J. E. T. McLain, J. Patel, and R. E. Mandrell. 2011. Irrigation waters as a source of pathogenic microorganisms in produce a review. *Adv. Agron.* 113:75-141.

46. Borchardt, M. A., P. D. Bertz, S. K. Spencer, and D. A. Battigelli. 2003. Incidence of Enteric Viruses in Groundwater from Household Wells in Wisconsin. *Appl. Environ. Microbiol.* 69:1172-1180.

47. Leifert, C., K. Ball, N. Volakakis, and J. M. Cooper. 2008. Control of enteric pathogens in ready-to-eat vegetable crops in organic and "low input" production systems: a HACCP-based approach. *J. Appl. Microbiol.* 105:931-950.

48. Wachtel, M. R., L. C. Whitehand, and R. E. Mandrell. 2002. Prevalence of *Escherichia coli* associated with a cabbage crop inadvertently irrigated with partially treated sewage wastewater. *J. Food Prot.* 65:471-475.

49. Johnson, J. Y. M., J. E. Thomas, T. A. Graham, I. Townshend, J. Byrne, L. B. Selinger, and V. P. J. Gannon. 2003. Prevalence of *Escherichia coli* O157:H7 and *Salmonella* spp. in surface waters of southern Alberta and its relation to manure sources. *Can. J. Microbiol.* 49:326-335.

50. Rzezutka, A., R. A. B. Nichols, L. Connelly, A. Kaupke, I. Kozyra, N. Cook, S. Birrell, and H. V Smith. 2010. Cryptosporidium oocysts on fresh produce from areas of high livestock production in Poland. *Int. J. Food Microbiol.* 139:96-101.

51. Suslow, T. V., M. P. Oria, L. R. Beuchat, E. H. Garrett, M. E. Parish, L. J. Harris,

J. N. Farber, and F. F. Busta. 2003. Production practices as risk factors in microbial food safety of fresh and fresh-cut produce. *Compr. Rev. Food Sci. Food Saf.* 2:38–77.

52. Curriero, F. C., J. A. Patz, J. B. Rose, and S. Lele. 2001. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. *Am. J. Public Health.* 91:1194-1199.

53. Song, I., S. W. Stine, C. Y. Choi, and C. P. Gerba. 2006. Comparison of crop contamination by microorganisms during subsurface drip and furrow irrigation. *J. Environ. Eng.* 132:1243–1248.

54. Stine, S. W., I. Song, C. Y. Choi, and C. P. Gerba. 2005. Application of microbial risk assessment to the development of standards for enteric pathogens in water used to irrigate fresh produce. *J. Food Prot.* 68:913-918.

55. Erickson, M. C., J. Liao, A. S. Payton, D. G. Riley, C. C. Webb, L. E. Davey, S. Kimbrel, L. Ma, G. Zhang, I. Flitcroft, M. P. Doyle, and L. R. Beuchat. 2010. Preharvest internalization of *Escherichia coli* O157:H7 into lettuce leaves, as affected by insect and physical damage. *J. Food Prot.* 73:1809–1816.

56. Hamilton, A. J., F. Stagnitti, R. Premier, A.-M. Boland, and G. Hale. 2006. Quantitative microbial risk assessment models for consumption of raw vegetables irrigated with reclaimed water. *Appl. Environ. Microbiol.* 72:3284-3290.

57. Park, S., S. Navratil, A. Gregory, A. Bauer, I. Srinath, B. Szonyi, K. Nightingale, J. Anciso, M. Jun, D. Han, S. Lawhon, and R. Ivanek. 2014. Farm management, environment, and weather factors jointly affect the probability of spinach

contamination by generic *Escherichia coli* at the preharvest stage. *Appl. Environ. Microbiol.* 80:2504-2515.

58. Sargeant, J., M. Sanderson, R. . Smith, and D. Griffin. 2003. *Escherichia coli* O157 in feedlot cattle feces and water in four major feeder-cattle states in the USA. *Prev. Vet. Med.* 61:127-135.

59. Dunn, J. R., J. E. Keen, and R. A. Thompson. 2004. Prevalence of Shiga-toxigenic *Escherichia coli* O157:H7 in adult dairy cattle. *J. Am. Vet. Med. Assoc.* 224:1151-1158.

60. Alam, M. J., and L. Zurek. 2006. Seasonal prevalence of *Escherichia coli* O157:H7 in beef cattle feces. *J. Food Prot.* 69:3018-3020.

61. LeJeune, J. T., D. Hancock, Y. Wasteson, E. Skjerve, and A. M. Urdahl. 2006. Comparison of *E. coli* O157 and Shiga toxin-encoding genes (stx) prevalence between Ohio, USA and Norwegian dairy cattle. *Int. J. Food Microbiol.* 109:19-24.

62. Jacob, M. E., K. M. Almes, X. Shi, J. M. Sargeant, and T. G. Nagaraja. 2011. *Escherichia coli* O157:H7 genetic diversity in bovine fecal samples. *J. Food Prot.* 74:1186-1188.

63. van Kessel, J. S., J. S. Karns, D. R. Wolfgang, E. Hovingh, B. M. Jayarao, C. P. V. A. N. Tassell, and Y. H. Schukken. 2008. Environmental sampling to predict fecal prevalence of *Salmonella* in an intensively monitored dairy herd. *J. Food Prot.* 71:1967-1973.

64. Jiang, X., J. Morgan, and M. P. Doyle. 2002. Fate of *Escherichia coli* O157:H7 in manure-amended soil. *Appl. Environ. Microbiol.* 68:2605-2609.

65. Franz, E., A. D. Van Diepeningen, O. J. De Vos, and A. H. C. Van Bruggen. 2005. Effects of cattle feeding regimen and soil management type on the fate of *Escherichia coli* O157 : H7 and *Salmonella enterica* serovar Typhimurium in manure, manure-amended soil, and lettuce. *Appl. Environ. Microbiol.* 71:6165-6174.

66. Franz, E., A. Visser, A. D. Van Diepeningen, M. M. Klerks, A. J. Termorshuizen, and A. H. C. van Bruggen. 2007. Quantification of contamination of lettuce by GFP-expressing *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium. *Food Microbiol.* 24:106-112.

67. Shepherd, M. W., P. Liang, X. Jiang, M. P. Doyle, and M. C. Erickson. 2007. Fate of *Escherichia coi* O157:H7 during on-farm dairy manure-based composting. *J. Food Prot.* 70:2708-2716.

68. Semenov, A. V., A. H. C. van Bruggen, L. van Overbeek, A. J. Termorshuizen, and A. M. Semenov. 2007. Influence of temperature fluctuations on *Escherichia coli* O157:H7 and *Salmonella enterica* serovar Typhimurium in cow manure. *FEMS Microbiol. Ecol.* 60:419-428.

69. van Elsas, D., A. V. Semenov, R. Costa, and J. T. Trevors. 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J.* 580:173-183.

70. Habteselassie, M., M. Bischoff, E. Blume, B. Applegate, B. Reuhs, S. Brouder, and R. F. Turco. 2008. Environmental controls on the fate of *Escherichia coli* in soil. *Water. Air. Soil Pollut.* 190:143-155.

71. Fremaux, B., C. Prigent-Combaret, M. L. Delignette-Muller, B. Mallen, M.

Dothal, A. Gleizal, and C. Vernozy-Rozand. 2008. Persistence of Shiga toxin-producing *Escherichia coli* O26 in various manure-amended soil types. *J. Appl. Microbiol.* 104:296-304.

72. Natvig, E. E., S. C. Ingham, B. H. Ingham, R. Cooperband, T. R. Roper, and L. R. Cooperband. 2002. *Salmonella enterica* serovar Typhimurium and *Escherichia coli* contamination of root and leaf vegetables grown in soils with incorporated bovine manure. *Appl. Environ. Microbiol.* 68: 2737-2744.

73. Islam, M., M. P. Doyle, S. C. Phatak, P. Millner, and X. Jiang. 2004. Persistence of enterohemorrhagic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *J. Food Prot.* 67:1365-1370.

74. Solomon, E. B., C. J. Potenski, and K. R. Matthews. 2002. Effect of irrigation method on transmission to and persistence of *Escherichia coli* O157:H7 on lettuce. *J. Food Prot.* 65:673-676.

75. Klerks, M. M., M. van Gent-Pelzer, E. Franz, C. Zijlstra, and A. H. C. van Bruggen. 2007. Physiological and molecular responses of Lactuca sativa to colonization by *Salmonella enterica* serovar Dublin. *Appl. Environ. Microbiol.* 73:4905-4914.

76. Yang, Y., Y. Luo, P. Millner, E. Turner, and H. Feng. 2012. Assessment of *Escherichia coli* O157:H7 transference from soil to iceberg lettuce via a contaminated field coring harvesting knife. *Int. J. Food Microbiol.* 153:345-350.

77. Tirado, M. C., R. Clarke, L. a. Jaykus, a. McQuatters-Gollop, and J. M. Frank. 2010. Climate change and food safety: A review. *Food Res. Int.* 43:1745-1765.

78. Liu, C., N. Hofstra, and E. Franz. 2013. Impacts of climate change on the microbial safety of pre-harvest leafy green vegetables as indicated by *Escherichia coli* O157 and *Salmonella* spp. *Int. J. Food Microbiol.* 163:119-128.

79. 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.

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

81. Gu, G., Z. Luo, J. M. Cevallos-Cevallos, P. Adams, G. Vellidis, A. Wright, and A. H. C. van Bruggen. 2013. Factors affecting the occurrence of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed. *Can. J. Microbiol.* 59:175-182.

82. Ailes, E. C., J. S. Leon, L.-A. Jaykus, L. M. Johnston, H. a Clayton, S. Blanding, D. G. Kleinbaum, L. C. Backer, and C. L. Moe. 2008. Microbial concentrations on fresh produce are affected by postharvest processing, importation, and season. *J. Food Prot.* 71:2389-2397.

83. Marine, S. C., S. Pagadala, F. Wang, D. M. Pahl, M. V. Melendez, W. L. Kline, R. A. Oni, C. S. Walsh, K. L. Everts, R. L. Buchanan, and S. A. Micallef. 2015. The growing season, but not the farming system, is a food safety risk determinant for leafy greens in the Mid-Atlantic Region of the United States. *Appl. Environ. Microbiol.* 81:2395-2407.

84. Pagadala, S., S. C. Marine, S. a. Micallef, F. Wang, D. M. Pahl, M. V. Melendez, W. L. Kline, R. a. Oni, C. S. Walsh, K. L. Everts, and R. L. Buchanan. 2015. Assessment of region, farming system, irrigation source and sampling time as food safety risk factors for tomatoes. *Int. J. Food Microbiol.* 196:98-108.

85. Larsen, K. S., S. Jonasson, and A. Michelsen. 2002. Repeated freeze–thaw cycles and their effects on biological processes in two arctic ecosystem types. *Appl. Soil Ecol.* 21:187-195.

86. Epstein, P. R. 1995. Emerging diseases and ecosystem instability: new threats to public health. *Am. J. Public Health* 85:168-172.

87. Beuchat, L. R. 2006. Vectors and conditions for preharvest contamination of fruits and vegetables with pathogens capable of causing enteric diseases. *Br. Food J.* 108:38-53.

88. Girardin, H., C. E. Morris, C. Albagnac, N. Dreux, C. Glaux, and C. Nguyen-The. 2005. Behaviour of the pathogen surrogates *Listeria innocua* and Clostridium sporogenes during production of parsley in fields fertilized with contaminated amendments. *FEMS Microbiol. Ecol.* 54:287-295.

89. Cevallos-Cevallos, J. M., M. D. Danyluk, G. Gu, G. E. Vallad, and A. H. C. van Bruggen. 2012. Dispersal of *Salmonella* Typhimurium by rain splash onto tomato plants. *J. Food Prot.* 75:472-479.

90. Davies, R. H., and C. Wray. 1996. Seasonal variations in the isolation of *Salmonella* typhimurium, *Salmonella* enteritidis, *Bacillus cereus* and *Clostridium perfringens* from environmental samples. *J. Vet. Med. Ser. B* 43:119-127.

91. Varma, J. K., K. D. Greene, M. E. Reller, S. M. DeLong, J. Trottier, S. F. Nowicki, M. DiOrio, E. M. Koch, T. L. Bannerman, S. T. York, M.-A. Lambert-Fair, J. G. Wells, and P. S. Mead. 2003. An outbreak of *Escherichia coli* O157 infection following exposure to a contaminated building. *JAMA* 290:2709-2712.

92. Griffin, D. W., V. H. Garrison, J. R. Herman, and E. A. Shinn. 2001. African desert dust in the Caribbean atmosphere: Microbiology and public health. *Aerobiologia (Bologna).* 17:203-213.

93. Baertsch, C., T. Paez-Rubio, E. Viau, and J. Peccia. 2007. Source tracking aerosols released from land-applied class B biosolids during high-wind events. *Appl. Environ. Microbiol.* 73:4522-4531.

94. Keraita, B. 2003. Influence of urban wastewater on stream water quality and agriculture in and around Kumasi, Ghana. *Environ. Urban.* 15:171-178.

95. U.S. Food and Drug Administration. 2015. Standards for the growing, harvesting, packing, and holding of produce for human consumption. Available at: https://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm334114.htm. Accessed 10 March 2017.

96. U.S. Environmental Protection Agency. 2012. Standards for the use or disposal of sewage sludge. Available at: https://www.gpo.gov/fdsys/pkg/CFR-2012-title40-vol31/pdf/CFR-2012-title40-vol31-part503-appB.pdf. Accessed 10 March 2017.

97. U.S. Department of Agriculture. 2011. National organic program: organic production and handling requirements. Available at: https://www.nofany.org/files/USDA_Organic_Production___Handling_Standa

rds.pdf. Accessed 10 March 2017.

98. California Leafy Green Products Handler Marketing Agreement. 2010. Comodity specific food safety guidelines for the production and harvest of produce and leafy green. Available at: http://www.lgma.ca.gov/wp-content/uploads/2014/09/California-LGMA-metrics-08-26-13-Final.pdf. Accessed 10 March 2017.

99. U.S. Food and Drug Adminstration. 2015. FDA investigated *Listeria monocytogenes* in sprouts from wholesome soy products, inc. Available at: http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm422562.htm. Accessed 10 March 2017.

100. U.S. Food and Drug Adminstration. 2011. Environmental assessment: factors potentially contributing to the contamination of fresh whole cantaloupe implicated in a multi-state outbreak of listeriosis. Available at: http://www.fda.gov/Food/RecallsOutbreaksEmergencies/Outbreaks/ucm276247.htm. Accessed 10 March 2017.

101. Schiere, H., and L. Kater. 2001. Mixed crop-livestock farming : a review of traditional technologies based on literature and field experiences. Available at: http://www.fao.org/docrep/004/Y0501E/y0501e00.htm. Accessed 10 March 2017.

102. Biswas, D. 2013. Ecology of *Campylobacter* and *Salmonella* in Pasture Poultry/Mixed Farm and Their Control with Natural Organic Antimicrobials (abstract). In Internation Association for Food Protection 2013 annual meeting, Charlotte, North Carolina.

103. Kersbergen, R. 2007. Integrating livestock production into a vegetable cropping system: pros and cons, dos and don'ts. Available at: http://www.newenglandvfc.org/pdf_proceedings/integrating_Livestock.pdf. Accessed 10 March 2017.

104. Ivanek, R., Y. T. Gröhn, M. T. Wells, J. Lembo, B. D. Sauders, and M. Wiedmann. 2009. Modeling of spatially referenced environmental and meteorological factors influencing the probability of *Listeria* species isolation from natural environments. *Appl. Environ. Microbiol.* 75:5893-7909.

105. Strawn, L. K., E. D. Fortes, E. 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.

106. Weller, D., M. Wiedmann, and L. K. Strawn. 2015. Spatial and temporal factors associated with an increased prevalence of *Listeria monocytogenes* in spinach fields in New York State. *Appl. Environ. Microbiol.* 81:6059-6069.

107. Xu, A., R. L. Buchanan, and S. A. Micallef. 2016. Impact of mulches and growing season on indicator bacteria survival during lettuce cultivation. *Int. J. Food Microbiol.* 224:28-39.

108. Micallef, S. A., R. E. Rosenberg Goldstein, A. George, L. Kleinfelter, M. S. Boyer, C. R. McLaughlin, A. Estrin, L. Ewing, J. Jean-Gilles 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.

109. Bubert, A., J. Riebe, N. Schnitzler, A. Schönberg, W. Goebel, and P. Schubert. 1997. Isolation of catalase-negative *Listeria monocytogenes* strains from listeriosis patients and their rapid identification by anti-p60 antibodies and/or PCR. *J. Clin. Microbiol.* 35:179-183.

110. Paziak-Domańska, B., Bogusławska, E., W.-S., R. M., Kotłowski, B. Różalska, M. Chmiela, J. Kur, W. Dąbrowski, and W. Rudnicka. 1999. Evaluation of the API test, phosphatidylinositol-specific phospholipase C activity and PCR method in identification of *Listeria monocytogenes* in meat foods. *FEMS Microbiol. Lett.* 171:209-214.

111. Rodríguez-Lázaro, D., M. Hernández, M. Scortti, T. Esteve, J. A. Vázquez-Boland, and M. Pla. 2004. Quantitative detection of *Listeria monocytogenes* and *Listeria innocua* by real-time PCR: assessment of *hly*, *iap*, and *lin02483* targets and AmpliFluor Technology. *Appl. Environ. Microbiol.* 70:1366-1377.

112. Collett, D. 2003. Modelling survival data in medical research, 2nd ed. Chapman & Hall/CRC, Boca Raton, Florida, USA.

113. Hosmer, D. W., and S. Lemeshow. 2000. Applied logistic regression. John Wiley & Sons, Inc., Hoboken, NJ, USA. John Wiley & Sons, Inc. Hoboken, New Jersey, USA.

114. O 'rourke, N., and L. Hatcher. A step-by-step approach to using SAS ® for factor analysis and structural equation modeling second edition. Available at: https://www.researchgate.net/file.PostFileLoader.html?id=57437b9a615e271c7c620522&assetKey=AS%3A365025819676672%401464040345894.

Accessed 10 March 2017.

115. Welshimer, H. J. 1960. Survival of *Listeria monocytogenes* in soil. *J. Bacteriol.* 80:316-320.

116. Dowe, M., E. Jackson, J. Mori, and B. Colin. 1997. *Listeria monocytogenes* Survival in soil and incidence in agricultural soils. *J. Food Prot.* 10:1158-1286.

117. McLaughlin, H. P., P. G. Casey, J. Cotter, C. G. M. Gahan, and C. Hill. 2011. Factors affecting survival of *Listeria monocytogenes* and *Listeria innocua* in soil samples. *Arch. Microbiol.* 193:775-785.

118. Francis Beaufort. 1805. Beaufort wind scale. Available at: http://www.spc.noaa.gov/faq/tornado/beaufort.html. Accessed 10 March 2017.

119. Ravva, S. V, C. Z. Sarreal, and R. E. Mandrell. 2011. Bacterial communities in aerosols and manure samples from two different dairies in central and Sonoma valleys of California. *PLoS One*. 6: e17281.

120. Oni, R. A., M. Sharma, and R. L. Buchanan. 2015. Survival of *Salmonella enterica* in dried turkey manure and persistence on spinach leaves. *J. Food Prot.* 78:1791-1799.

121. Oni, R. The role of organic matrices (dried turkey manure) in the contamination and survival of *Salmonella* spp. on baby spinach leaves (master's thesis). Retrieved from: http://drum.lib.umd.edu/handle/1903/13565.

122. Yanamala, S., M. F. Miller, G. H. Loneragan, S. E. Gragg, and M. M. Brashears. 2011. Potential for microbial contamination of spinach through feedyard air/dust growing in close proximity to cattle feedyard operations. *J. Food Saf.* 31:525-529.

123. Benjamin, L., E. R. Atwill, M. Jay-Russell, M. Cooley, D. Carychao, L. Gorski, and R. E. Mandrell. 2013. Occurrence of generic *Escherichia coli*, *E. coli* O157 and *Salmonella* spp. in water and sediment from leafy green produce farms and streams on the Central California coast. *Int. J. Food Microbiol.*165:65-76.

124. Berry, E. D., J. E. Wells, J. L. Bono, B. L. Woodbury, N. Kalchayanand, K. N. Norman, T. V. Suslow, G. Lopez-Velasco, and P. D. Millner. 2015. Effect of proximity to a cattle feedlot on *Escherichia coli* O157:H7 contamination of leafy greens and evaluation of the potential for airborne transmission. *Appl. Environ. Microbiol.* 81:1101-1110.

125. Junttila, J. R., S. I. Niemelä, and J. Hirn. 1988. Minimum growth temperatures of *Listeria monocytogenes* and non-haemolytic *Listeria*. *J. Appl. Bacteriol.* 65:321–327.

126. Guerini, M. N., D. M. Brichta-Harhay, S. D. Shackelford, T. M. Arthur, J. M. Bosilevac, N. Kalchayanand, T. L. Wheeler, and M. Koohmaraie. *Listeria* prevalence and *Listeria monocytogenes* serovar diversity at cull cow and bull processing plants in the United States. *J. Food Prot.* 70:2578-2582.

127. Wilkes, G., T. A. Edge, V. P. J. Gannon, C. Jokinen, E. Lyautey, N. F. Neumann, N. Ruecker, A. Scott, M. Sunohara, E. Topp, and D. R. Lapen. 2011. Associations among pathogenic bacteria, parasites, and environmental and land use factors in multiple mixed-use watersheds. *Water Res.* 45:5807–5825.

128. Breiman, L. 2001. Random forests. *Mach. Learn.* 45:5-32.

129. Jones, Z., and F. Linder. 2015. Exploratory Data Analysis using Random Forests. *73rd Annu. MPSA Conf.* 1-31.

130. Atwill, E. R., J. A. Chase, D. Oryang, R. F. Bond, S. T. Koike, M. D. Cahn, M. Anderson, A. Mokhtari, and S. Dennis. 2015. Transfer of *Escherichia coli* O157:H7 from simulated wildlife scat onto romaine lettuce during foliar irrigation. *J. Food Prot.* 78:240-247.

131. U.S. Food and Drug Administration. 2013. FDA investigation summary: multistate outbreak of *Salmonella* Typhimurium and *Salmonella* Newport infections linked to cantaloupe grown at Chamberlain farms in southwest Indiana. Available at: https://www.fda.gov/food/recallsoutbreaksemergencies/outbreaks/ucm315879.htm. Accessed 10 March 2017.

132. Feng, P., S. D. Weagant, M. A. Grant, and W. Burkhardt. 2002. Bacteriological analytical manual: enumeration of *Escherichia coli* and the coliform bacteria. Available at: https://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm064948.htm. Accessed 10 March 2017.

133. Park, S., S. Navratil, A. Gregory, A. Bauer, I. Srinath, B. Szonyi, K. Nightingale, J. Anciso, M. Jun, D. Han, S. Lawhon, and R. Ivanek. 2015. Multifactorial effects of ambient temperature, precipitation, farm management, and environmental factors determine the level of generic *Escherichia coli* contamination on preharvested spinach. *Appl. Environ. Microbiol.* 81:2635-2650.

134. Salaheen, S., N. Chowdhury, I. Hanning, and D. Biswas. 2015. Zoonotic bacterial pathogens and mixed crop-livestock farming. *Poult. Sci.* 94:1398-1410.

135. Franzluebbers, A. J. 2007. Integrated crop – livestock systems in the southeastern USA. *Agron. J.* 99:361-372.

136. De'ath, G., and K. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178-3192.

137. Dohoo, I., W. Martin, and H. Stryhn. 2003. Veterinary Epdimiological Research. VER, Inc, Prince Edward Island, Canada.

138. Cooper, V. S., A. F. Bennett, and R. E. Lenski. 2001. Evolution of thermal dependence of growth rate of *Escherichia coli* populations during 20,000 generations in a constant environment. *Evolution (N. Y).* 55:889-896.

139. Ishii, S., T. Yan, H. Vu, D. L. Hansen, R. E. Hicks, M. J. Sadowsky, T. Zhang, H. Wang, L. Wu, J. Lou, J. Wu, P. C. Brookes, J. Xu, S. Ishii, W. B. Ksoll, R. E. Hicks, M. J. Sadowsky, M. L. Williams, J. T. LeJeune, B. McSpadden Gardener, S. Vidovic, H. C. Block, and D. R. Korber. 2010. Factors controlling long-term survival and growth of naturalized *Escherichia coli* populations in temperate field soils. *Microbes Environ.* 25:8-14.

140. Cools, D., R. Merckx, K. Vlassak, and J. Verhaegen. 2001. Survival of *E. coli* and *Enterococcus* spp. derived from pig slurry in soils of different texture. *Appl. Soil Ecol.* 17:53-62.

141. Cooley, M. B., B. Quiñones, D. Oryang, R. E. Mandrell, and L. Gorski. 2014. Prevalence of shiga toxin producing *Escherichia coli*, *Salmonella enterica*, and *Listeria monocytogenes* at public access watershed sites in a California central coast agricultural region. *Front. Cell. Infect. Microbiol.* 4:1-13.

142. Centers for Disease Control and Prevention. 2016. Multistate outbreak of *Salmonella* Poona infections linked to imported cucumbers. Available at: https://www.cdc.gov/salmonella/poona-09-15/index.html. Accessed 10 March 2017.

143. Santamaría, J., and G. A. Toranzos. 2003. Enteric pathogens and soil: a short review. *Int. Microbiol.* 6:5–9.

144. Mukherjee, A., D. Speh, and F. Diez-Gonzalez. 2007. Association of farm management practices with risk of *Escherichia coli* contamination in pre-harvest produce grown in Minnesota and Wisconsin. *Int. J. Food Microbiol.* 120:296-302.

145. Mukherjee, A., D. Speh, and E. Dyck. 2004. Preharvest evaluation of coliforms, *Escherichia coli*, *Salmonella*, and *Escherichia coli* O157 :H7 in organic and conventional produce grown by Minnesota farmers. *J. Food Prot.* 67:894-900.

146. Doyle, M. P., and M. C. Erickson. 2008. Summer meeting 2007 - the problems with fresh produce: an overview. *J. Appl. Microbiol.* 105:317-330.

147. Park, S., B. Szonyi, R. Gautam, K. Nightingale, J. Anciso, and R. Ivanek. 2012. Risk factors for microbial contamination in fruits and vegetables at the preharvest level: a systematic review. *J. Food Prot.* 75:2055-2081.

148. Strawn, L. K., Y. T. Gröhn, S. Warchocki, R. W. Worobo, E. Bihn, and M. Wiedmann. 2013. Risk factors associated with *Salmonella* and *Listeria monocytogenes* contamination of produce fields. *Appl. Environ. Microbiol.* 79:7618-7627.

149. Fageria, N. K., V. C. Baligar, and B. A. Bailey. 2007. Role of Cover Crops in

Improving Soil and Row Crop Productivity. *Commun. Soil Sci. Plant Anal.* Taylor & Francis Group 36:2733– 2757.

150. Reed-Jones, N. L., S. C. Marine, K. L. Everts, and S. a Micallef. 2016. Effect of cover crop species and season on population dynamics of *Escherichia coli* and *Listeria innocua* in soil. *Appl. Environ. Microbiol.* 82: 1767-1777.

151. Doran, J. W., and M. R. Zeiss. 2000. Soil health and sustainability: managing the biotic component of soil quality. *Appl. Soil Ecol.* 15:3-11.

152. Ramos, C., L. Molbak, and S. Molin. 2000. Bacterial activity in the rhizosphere analyzed at the single-cell level by monitoring ribosome contents and synthesis rates. *Appl. Environ. Microbiol.* 66:801-809.

153. Ivanek, R., Y. T. Gröhn, and M. Wiedmann. 2006. *Listeria monocytogenes* in multiple habitats and host populations: review of available data for mathematical modeling. *Foodborne Pathog. Dis.* 3:319-336.

154. Vivant, A.-L., D. Garmyn, and P. Piveteau. 2013. *Listeria monocytogenes*, a down-to-earth pathogen. *Front. Cell. Infect. Microbiol.* 3:87.

155. Hosmer, D. W., S. Lemeshow, and S. May. 2008. Applied survival analysis : regression modeling of time-to-event data. Wiley. Hoboken, New Jersey, USA.

156. Mubiru, D. N., M. S. Coyne, and J. H. Grove. 2000. Mortality of *Escherichia coli* O157:H7 in two soils with different physical and chemical properties. *J. Environ. Qual.* 29:1821-1825.

157. Ma, J., A. M. Ibekwe, D. E. Crowley, and C. H. Yang. 2012. Persistence of *Escherichia coli* O157:H7 in major leafy green producing soils. *Environ. Sci. Technol.* 46:12154-12161.

158. Locatelli, A., A. Spor, C. Jolivet, P. Piveteau, and A. Hartmann. 2013. Biotic and abiotic soil properties influence survival of *Listeria monocytogenes* in soil. *PLoS One* 8:1-8.

159. Williams, M. L., J. T. LeJeune, and B. McSpadden Gardener. 2015. Soil conditions that can alter natural suppression of *Escherichia coli* O157:H7 in Ohio specialty crop soils. *Appl Env. Microbiol* 81:4634-4641.

160. Sidorenko, M. L., L. S. Buzoleva, and N. M. Kostenkov. 2006. The effect of soil properties on the preservation and reproduction of *Listeria* and *Yersinia*. *Eurasian Soil Sci.* 39:211-217.

161. Ma, J., A. Mark Ibekwe, D. E. Crowley, and C. H. Yang. 2014. Persistence of *Escherichia coli* O157 and non-O157 strains in agricultural soils. *Sci. Total Environ.* 490:822-829.

162. Sjogren, R. E. 1994. Prolonged survival of an environmental *Escherchia Coli* in laboratory soil microcosms. *Water. Air. Soil Pollut.* 75:389-403.

163. Vidovic, S., H. C. Block, and D. R. Korber. 2007. Effect of soil composition, temperature, indigenous microflora, and environmental conditions on the survival of *Escherichia coli* O157:H7. *Can. J. Microbiol.* 53:822-829.

164. Weis, J., and H. P. Seeliger. 1975. Incidence of *Listeria monocytogenes* in nature. *Appl. Microbiol.* 30:29-32.

165. Klein, D. A., and L. E. Casida Jr. 1967. *Escherichia coli* die-out from normal soil as related to nutrient availability and the indigenous microflora. *Can. J. Microbiol.* 13:1461-1470.

166. Moynihan, E. L., K. G. Richards, K. Ritz, S. F. Tyrrel, and F. P. Brennan. 2013.

Impact of soil type, biology and temperature on the survival of non-toxigenic *Escherichia coli* O157. *Biol. Environ. Proc. R. Irish Acad.* 113:1-6.

167. Jiang, X., M. Islam, J. Morgan, and M. P. Doyle. 2004. Fate of *Listeria monocytogenes* in bovine manure – amended soil. *J. Food Prot.* 67:1676-1681.

168. Freitag, N. E., G. C. Port, and M. D. Miner. 2009. *Listeria monocytogenes* - from saprophyte to intracellular pathogen. *Nat. Rev. Microbiol.* 7:623-628.

169. Bruno, J. C., and N. E. Freitag. 2011. *Listeria monocytogenes* adapts to long-term stationary phase survival without compromising bacterial virulence. *FEMS Microbiol. Lett.* 323:171-179.

170. Nightingale, K. K., Y. H. Schukken, C. R. Nightingale, E. D. Fortes, A. J. Ho, Z. Her, Y. T. Grohn, P. L. McDonough, and M. Wiedmann. 2004. Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. *Appl. Environ. Microbiol.* 70:4458-4467.

171. Zhang, T., H. Wang, L. Wu, J. Lou, J. Wu, P. C. Brookes, and J. Xu. 2013. Survival of *Escherichia coli* O157:H7 in soils from Jiangsu province, China. *PLoS One* 8: e81178.

172. Wang, H., T. Zhang, G. Wei, L. Wu, J. Wu, and J. Xu. 2014. Survival of *Escherichia coli* O157:H7 in soils under different land use types. *Environ. Sci. Pollut. Res.* 21:518-524.

173. Wang, H., T. Zhang, G. Wei, L. Wu, J. Wu, and J. Xu. 2014. Survival of *Escherichia coli* O157:H7 in soils under different land use types. *Environ. Sci. Pollut. Res.* 21:518-524.

174. Vivant, A. L., D. Garmyn, P. A. Maron, V. Nowak, and P. Piveteau. 2013.

Microbial diversity and structure are drivers of the biological barrier effect against *Listeria monocytogenes* in Soil. *PLoS One* 8:1-11.

175. Unc, A., J. Gardner, S. Springthorpe, A. Unc, J. Gardner, and S. Springthorpe. 2006. Recovery of *Escherichia coli* from soil after addition of sterile organic wastes. *Appl. Environ. Microbiol.* 72: 2287-2289.

176. U.S. Department of Agriculture. 2016. Tomatoes-fresh tomato industry. Available at: https://www.ers.usda.gov/topics/crops/vegetables-pulses/tomatoes.aspx. Accessed 10 March 2017.

177. Centers for Disease Control and Prevention. 2016. Foodborne outbreak online database (food tool). Available at: http://wwwn.cdc.gov/foodborneoutbreaks/. Accessed 10 March 2017.

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

179. Bennett, S. D., K. W. Littrell, T. A. Hill, M. Mahovic, and C. B. Behravesh. 2015. Multistate foodborne disease outbreaks associated with raw tomatoes, United States, 1990-2010: a recurring public health problem. *Epidemiol. Infect.* 143:1352-1359.

180. Orozco R, L., M. H. Iturriaga, M. L. Tamplin, P. M. Fratamico, J. E. Call, J. B. Luchansky, and E. F. Escartin. 2008. Animal and environmental impact on the presence and distribution of *Salmonella* and *Escherichia coli* in hydroponic tomato greenhouses. *J. Food Prot.* 71:676-683.

181. Franz, E., a V Semenov, and a H. C. van Bruggen. 2008. Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. *J. Appl. Microbiol.* 105:1569-1584.

182. Pang, H., E. Lambertini, R. L. Buchanan, D. W. Schaffner, and A. K. Pradhan. 2017. Quantitative microbial risk assessment for *Escherichia coli* O157:H7 in fresh-cut lettuce. *J. Food Prot.* 80:302-311.

183. Cooley, M., D. Carychao, L. Crawford-Miksza, M. T. Jay, C. Myers, C. Rose, C. Keys, J. Farrar, and R. E. Mandrell. 2007. Incidence and tracking of *Escherichia coli* O157:H7 in a major produce production region in California. *PLoS One* 2:e1159.

184. Chapin, T. K., K. K. Nightingale, R. W. Worobo, M. Wiedmann, and L. K. Strawn. 2014. Geographical and meteorological factors associated with isolation of *Listeria* species in new york state produce production and natural environments. *J. Food Prot.* 77:1919-1928.

185. Langholz, J. A., and T. J. R. Michele. 2013. Potential role of wildlife in pathogenic contamination of fresh produce. *Human-Wildlife Interact.* 7:140-157.

186. Lyautey, E., A. Hartmann, F. Pagotto, K. Tyler, D. R. Lapen, G. Wilkes, P. Piveteau, A. Rieu, W. J. Robertson, D. T. Medeiros, T. a Edge, V. Gannon, and E. Topp. 2007. Characteristics and frequency of detection of fecal *Listeria monocytogenes* shed by livestock, wildlife, and humans. *Can. J. Microbiol.* 53:1158-1167.

187. Fonseca, J. M., S. D. Fallon, C. A. Sanchez, and K. D. Nolte. 2011. *Escherichia*

coli survival in lettuce fields following its introduction through different irrigation systems. *J. Appl. Microbiol.* 110:893-902.

188. Steele, M., and J. Odumeru. 2004. Irrigation water as source of foodborne pathogens on fruit and vegetables. *J Food Prot.* 67:2839-2849.

189. Centers for Disease Control and Prevention. 1997. Water in pest spray likely source of *Cyclospora* in outbreak. Available at: http://extoxnet.orst.edu/newsletters/n172_97.htm. Accessed 10 March 2017.

190. Cevallos-Cevallos, J. M., G. Gu, M. D. Danyluk, N. S. Dufault, and A. H. C. van Bruggen. 2012. *Salmonella* can reach tomato fruits on plants exposed to aerosols formed by rain. *Int. J. Food Microbiol.* 158:140-146.

191. Dreux, N., C. Albagnac, F. Carlin, C. E. Morris, and C. Nguyen-The. 2007. Fate of *Listeria* spp. on parsley leaves grown in laboratory and field cultures. *J. Appl. Microbiol.* 103:1821-1827.