QUANTITATIVE MICROBIAL RISK ASSESSMENT FOR FRESH ORANGES SPRAYED WITH LOW MICROBIAL QUALITY WATER IN CENTRAL FLORIDA

By

GABRIEL MOOTIAN

A dissertation is submitted to

Graduate School New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Food Science

Written under the direction of

Dr. Donald W. Schaffner

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey
May 2016

ABSTRACT OF DISSERTATION

Quantitative Microbial Risk Assessment For Fresh Oranges Sprayed With Surface Water From In Central Florida

By

Gabriel Mootian

Advisor: Dr. Donald Schaffner

Surface waters are an important source of agricultural water. About 58% of total surface water withdrawals are used for agricultural purposes in the US, thus the microbial quality of surface water is important for the safe production of fresh fruits and vegetables. Studies have demonstrated that surface waters are prone to contamination and may contain viral, bacterial and protozoal water and food borne pathogens. The efficacy of biological and physicochemical indicators for *Salmonella* in surface waters in central Florida was evaluated in chapter 2 of this dissertation. *E. coli* and solar radiation (SR) levels were found to significantly correlate well with *Salmonella* levels but the R^2 were low (< 0.1). The influence of key weather variables on the survival of *E. coli* (surrogate for *Salmonella*) on the surface of oranges in grove following application of contaminated spray water was evaluated in the 3rd chapter. Weather variables relative humidity (RH), solar radiation (SR) and temperature and time were found to be significantly correlated to the log change in *E. coli* levels/h on the surface on an orange; in addition time was also found to correlate well with log change in *E. coli* levels/h. However the R^2 values were

found to be low (< 0.4). Logistic regression analysis showed that there was a high probability of a decline in *E. coli* levels at high SR and low RH levels, and a low probability of an increase in *E. coli* levels even at the lowest SR and the highest RH levels. Chapter 4 studies the predicted levels of *E. coli* on the surface of an orange under the influence of the four significant variables RH, SR, temperature and time. An empirical model based on upper and lower boundary limits for RH, SR, temperature and time variable was developed. IF logic statements, probability distributions and Monte Carlo simulations were used to describe the variability and uncertainty in log *E. coli* change rate for the most constricting variable. The mean predicted *E. coli* concentrations at day 1 after spraying were very close to the actual field trial data but the model under predicted the risk for 3, 7 and 14 days post spraying. The FDA produce rule on microbial die off rate (0.5 log/day of generic *E. coli* for a maximum of 4 days) was also evaluated. The rule was found to over predict actual data at day 1 and 3, but under predicted at day 7 and 14. The FDA rule for a pre harvest interval of less than 4 days is fail safe and has the potential to reduce the risk of exposure to pathogens on surfaces of pre harvest crops by restricting the use of contaminated sources of agricultural water.

Dedication

I dedicate this dissertation to my wife and son for their love, support and understanding throughout my studies especially when I have had to be away on long academic conferences.

I would also like to dedicate this work to my parents, siblings and relatives in Kenya for their support and encouragement.

Acknowledgment

I would like to thank my advisor Dr. Donald Schaffner for this generous opportunity to work in his lab for the last 7 years. My sincere and heartfelt thanks for your financial support and academic guidance.

I would also like to thank my committee members Dr. Karl Matthews, Dr. Mukund Karwe and Dr. Michelle Danyluk for serving in my dissertation committee. Thank you all for your financial support and allowing me to work in your labs on very fascinating research projects. I am also grateful to Dr. Lawrence Goodridge and Rutgers University dining services for providing a graduate assistantship.

Finally I would like to acknowledge my labs mates and undergraduate students whom I have had the pleasure to work with.

Table of Contents

List of Figures

List of Tables

Chapter 1: Literature review

I.1 Surface waters

Surface waters are defined as waters that are found on the surface of the earth. Examples of surface waters include lakes, rivers, streams, ponds and marshes. Ground water is defined as water below the earth's surface. Such water is found in an aquifer and soil pore spaces. Surfaces waters are different than ground waters in terms of both chemical and biological contaminant risks as well as the speed of flow. Since surfaces are open to pollution, they may contain higher concentrations of contaminants than ground water. This is the reason why ground water is commonly used as a source of drinking water by people without access to treated municipal water. Ground water is filtered through layers of rock and soil and as a result has lower concentrations of biological contaminants. Ground and surface water are interconnected by the hydrologic cycle and may exchange contaminants under the right conditions. In 2005, about 80% of all the water (public supply, hydroelectric power generation, irrigation, mining and industrial purposes) used in the USA came from surface water sources with the remaining 20% coming from ground water sources (*85*). Popular uses of surface waters include generation of hydroelectric power, industrial and mining purposes, irrigation, recreation and also public drinking water supply. Since all of the water used for hydroelectric power (53%) is most often returned back to the source, ignoring hydroelectric power withdrawals, irrigation accounts for most of surface water withdrawals (58%) followed by public supply and industrial uses (*85*).

I.2 Sources and types of biological contamination

Natural and anthropogenic contaminants of surfaces waters sources are categorized into two primary groups 1) Point sources and 2) Non point sources. Point sources are defined as discrete sources whose inputs into aquatic systems can be described in spatially explicit manner through measurement of chemical residues and or epidemiological data such as mortality and morbidity data (*70*). Common point sources of contamination include municipal waste water treatment plants (WWTPs), industrial effluent (food processing plants), sewage storm water overflows (discharge from overloaded WWTPs) and land disposal sites (leachates or discharge from land fills, septic tanks and industrial impoundments). In contrast non point sources cannot be explicitly distinguished spatially because of their diffuse nature and also the broad and geographical extent in which they occur. Examples of non point sources include agricultural runoffs (from crop production, animal feed lots, pastures) storm water and urban runoff (from impervious surfaces including streets and other paved areas) (*70*).

Besides nutrients, the most significant form of pollution from sewage and agricultural runoffs to surface waters are pathogens. Numerous studies on the microbial quality of surface water conducted in different geographical regions have demonstrated that surface waters are vulnerable to pollution and can contain varying levels and prevalence rates of enteric pathogens (viruses, bacteria and protozoa). Levantesi et al. (2012) conducted a review of literature (over 10 years) on the prevalence, diversity and survival of *Salmonella* in surface and drinking water. *Salmonella* detection frequency ranged between 3% to 100% while the some of the highest *Salmonella* levels recorded in India were between 10^4 to 10^6 CFU/ml and were enumerated from river water.

Campylobacter species have been isolated from surface waters in various countries (*1, 12, 17, 42, 77*). Protozoan pathogens such as *Cyclospora, Giardia* and *Cryptosporidium* cysts and oocysts have been reported in numerous studies of surface waters from Europe, South America and North America (*17, 42, 73, 77, 79*). Human enteric viruses such norovirus and rotavirus have been detected from surface water studies conducted in Finland, Amsterdam, Netherlands and USA (*30, 42, 50, 70*). Other enteric pathogenic bacteria isolated from surface waters include pathogenic *E. coli, Vibrio* spp and *Shigella* (*1, 17, 70, 77*).

I.3 Persistence of pathogens in surface water

The occurrence and elevated survival of enteric pathogens in non-host environments presents a serious public health challenge with respect to safe drinking and recreational water and the safety of minimally processed fresh fruits and vegetables. In the majority of surface water studies mentioned above, sampling and microbial analysis conducted over months or years indicate that pathogens may remain viable in aquatic environments for long periods. For example, high levels of *Salmonella* (1 to 4 log CFU/100 ml) have been reported in river water impacted by raw sewage in Italy over a sampling period of 2 years (*49*). Thus the influence and proximity to point and non point sources may be a cause of the persistence of pathogens in surface waters through continuous pollution. Survival and persistence of enteric pathogens in surface waters has also been associated with seasons and weather conditions. Lower levels, lower frequencies of detection and serotype diversity in *Campylobacter, Salmonella, E. coli* O157:H7, *Giardia* and *Cryptosporidium* have been correlated to lower temperatures

during autumn and winter seasons (*32*, *36*, *42*, *64*, *81*). Enteric viruses appear to be more prevalent during colder temperatures in winter compared to bacterial and protozoal pathogens (*42*, *50*). Some studies have reported a lack of seasonal influence in the persistence of some pathogens in surface waters. Some authors have indicated that there may be other factors besides seasonality driving the prevalence and persistence of pathogens in surface waters. In a study conducted in Finland, birds were found to be the cause of the rather high numbers of *Campylobacter* and not the high temperatures (>18 °C) recorded during the study (*42*, *88*). An increased persistence during a particular season may also coincide with an increased shedding of the pathogen by humans or animals (*42*).

I.4 Food and waterborne gastroenteritis associated with surface waters

Epidemiological investigations, experimental studies and surface water microbial surveys have provided evidence that surface waters can be sources of water and foodborne illness. Laboratory studies have shown that foodborne pathogens can be transmitted to edible portions of a plant during application of contaminated agricultural water. Solomon et al. (2002) and Mootian et al. (2009) have demonstrated that *E. coli* O157:H7 can be transmitted to lettuce plants through contaminated irrigation water without the plant necessarily coming into contact with contaminated soil. Islam et al. (2004), and Lapidot et al. (2009) have shown that *Salmonella* can be transmitted from contaminated irrigation water to carrots, lettuce, radish and parsley. These studies are important in demonstrating the role of contaminated irrigation water in the transmission of pathogens to fruits and vegetables at pre-harvest stages.

Outbreaks associated with consumption of fruits and vegetables have been traced back to either contaminated irrigation water or process water. In a review of *Salmonella* prevalence in surface waters, Levantesi et al. (2012) described multistate outbreaks associated with Serrano and jalapeno peppers, tomatoes, basil, lettuce, cantaloupe and mangoes where epidemiological investigations traced back outbreak strains of *Salmonella* to on farm irrigation and processing waters. The largest outbreak reported by CDC occurred in 2008 where multiple raw produce items contaminated with *Salmonella* Saint Paul infected over 1400 individuals in the US and Canada. *E. coli* O157:H7 isolated from irrigation water has also been reported to cause outbreaks from consumption of lettuce (*40, 77*). Protozoan pathogens may also pose a risk. *Cyclospora* was implicated in an outbreak involving raspberries imported from Guatemala to the US and Canada, and was traced back to the use of contaminated water for the preparation of fungicides and insecticides (*79*). Reports of outbreaks traced back to enteric viruses isolated from surface waters are more rare, but still do occur. Heaton et al. (2008) and Beuchat (1996) reported outbreaks associated with Hepatitis A virus in lettuce and spring onions and norovirus in celery respectively. These viral outbreaks were linked to the use of irrigation water contaminated with untreated sewage.

Although there are numerous review articles on outbreaks associated with minimally processed fresh fruits and vegetables, not many implicate the use of contaminated surface waters. Other sources of enteric pathogens such as contaminated biological amendments, poor hygiene by food handlers, contaminated soil, feces from birds and animals appear to have contributed more to a higher incidence of food borne illnesses compared to agricultural water. This may be due to less use of sprinkler or

overhead irrigation and more use of subsurface type of irrigation, which has been shown to reduce the risk of contaminating pre harvest crops (*22*). There are also studies that have quantified the reduction of risk of illness associated with holding pre harvest crops for a certain period of time after the last episode of irrigation to allow for the natural death of pathogens transferred during irrigation (*37*). However there is also evidence of the persistence and survival of pathogens in traditional non host environments including surface waters which means use of such sources of agricultural water may pose a risk to the health of consumers of fresh fruits and vegetables (*49*).

I.5 *Salmonella*

Salmonella is a facultative anaerobic gram-negative rod shaped bacteria belonging to the family *Enterobacteriaceae.* The bacterium grows optimally at 37 °C and can break down glucose to produce gas and acid. *Salmonella* are also oxidase negative and catalase positive, grow on citrate as the sole carbon source, produce hydrogen sulphide, decarbolyze lysine and ornithine and do not hydrolyze urea (*18*).

A number of taxonomic schemes have been proposed for the nomenclature of genus *Salmonella*. The most widely divides the genus into 2 species *Salmonella enterica* and *Salmonella bongori*. *Salmonella enterica* is divided into 6 six sub species I. *S. enterica* subspecies *enterica,* II, *S. enterica* subspecies *salamae*; IIIa, *S. enterica* subspecies *arizonae*; IIIb,*S.enterica* subspecies *diarizonae*; IV, *S.enterica* subspecies *houtenae*; and VI, *S. enterica* subspecies *indica.* Biochemical identification is generally coupled with serological confirmation involving agglutination of bacterial surface antigens with salmonella specific antibodies. These include somatic (O)

lipopolysaccharides (LPS) on the external surface of the bacterial outer membrane, flagellin (H) antigens for peritrichous flagella and the capsular antigens (Vi) antigen, which occurs only in serovars Typhi, Paratyphi C, and Dublin. Currently, there are over 2400 serotypes of *Salmonella*, but only 50 are associated with infections in humans and warm-blooded animals, all within the subspecies *enterica* (*49*). Based on clinical syndrome, *Salmonella* serotypes can be classified into 2 distinct groups typhoidal and non-typhoidal. Typhoidal serotypes namely *S.* Typhi and Paratyphi are associated with enteric fevers typhoid and paratyphoid, which are the severe form of *Salmonella* infections. Typhi and Paratyphi serotypes only infect humans and are common in developing countries where the lack of clean drinking water is a major problem. In contrast non-typhoidal *Salmonella* serotypes are commonly associated with food borne gastroenteritis than water borne illnesses and are found in both humans and animals (*49*). These zoonotic serotypes of *Salmonella* cause acute but usually self-limiting gastroenteritis.

I.6 Salmonellosis

Salmonella is reported to cause 1 million illnesses, 19,000 hospitalizations and 380 deaths every year in the United States (*13*). The CDC in its 2011 estimates reported *Salmonella* as the leading pathogen in the number of deaths and hospitalizations. Historically *Salmonella* has been associated with meat, poultry, dairy and egg products (*18*). In recent years outbreaks of human salmonellosis have been linked to consumption of fresh tomatoes, raw almonds, melons, mangoes, peppers, orange juice, cucumbers,

lettuce, alfalfa sprouts, mixed salads and low moisture foods such a spices and peanut butter (*13*).

I.7 Quantitative microbial risk assessment (QMRA)

Quantitative microbial risk assessment (QMRA) is a scientifically based process that is used to estimate the likelihood (probability) of illness to humans as a result of exposure to food or water contaminated with pathogenic microorganisms or toxins (*24*, *83*). QMRA involves the following steps: i) Hazard identification; ii) Hazard characterization; iii) Exposure assessment and iv) Risk characterization.

I.7.a Hazard identification

The hazard identification component of a microbial risk assessment identifies the pathogenic microorganism that may be present in a food, water or a group of foods, that is capable of causing adverse public health effects (*24*). The hazard identification component may also identify specific population or subpopulation of individuals impacted by consumption of the target microorganism.

I.7.b Hazard characterization

Hazard characterization describes the relationship between the level of exposure of a pathogen (dose) and the likelihood of an adverse health effect (response). This relation is in a form of a mathematical model referred to as a dose response model and is derived from epidemiological surveillance data (outbreak data), human clinical feeding trials or animal feeding studies (*25*). Dose response model development can occur as a "stand alone" process or as part of a quantitative risk assessment. Dose responses for a

particular pathogen may be used between risk assessments (of different foods) for the same pathogen (*25*).

I.7.c Exposure assessment

The exposure assessment component of a microbial risk assessment is an evaluation of the likelihood of ingesting pathogenic microorganism through food or water and the likely level of exposure. Levels of a microbial hazard may be obtained through microbiological analysis of raw material or a finished food product. Since the level of a hazard at the time of consumption may be different from that when the food is being produced an exposure assessment considers all the production processes (from farm to fork) and their effect on levels of the hazard, to estimate the likely levels at the time of consumption. Where data is unavailable, quantitative exposure assessment models may be built taking into consideration various factors (for e.g. temperature, relative humidity, pH, salt conc. water activity) and their interactions to estimate the distribution and levels of the hazard in a food at consumption. Predictive growth, inactivation and cross contamination models are have also been used in the development of exposure assessment models (*15, 16, 63, 72*). Data from predictive modeling and microbiological analysis is combined with food consumption patterns to assess exposure to the hazard over a period of time (*23*).

I.7.d Risk characterization

Risk characterization is the final step in a QMRA. This step integrates information from the other three steps hazard identification, hazard characterization and exposure assessment. The result of risk characterization is a risk estimate (*26*). Risk estimates can be reported in form of risk per serving of a particular food, risk per individual or population risk. The risk per serving requires a defined quantity that

constitutes a serving (e.g. number of oysters, grams of lettuce, ml of orange juice). Risk per individual can refer to an individual risk in specified population or an individual risk among a population that consumes a specific product i.e. this risk applies to individuals in specific category of a population or individuals who consume a particular product. Population risk considers the risk distributed over a population (*26*). Population risk is often reported as the total number of food borne illness expected in a population within a year.

I.7.e Deterministic versus stochastic

Microbial risk assessments can take deterministic or a stochastic (probabilistic) approach. Deterministic or single point approach involves using a single "best guess" estimate of all the variables within a model (*87*). The outcome of a deterministic model approach is a single risk estimate. Point estimates do not take into account the variability or uncertainty of variable. In contrast, in a stochastic approach a probability distribution is used to account for every possible value that each variable can take and it's probability of occurrence. The output of a stochastic model is a probability distribution of all the possible outcomes. Most QMRA adopt the stochastic approach because probability distributions allow for characterization of variability and uncertainty. QMRAs can use Monte Carlo simulation technique where random sampling of each probability distribution within a model is done hundreds or thousands of time to produce hundreds or thousands of scenarios (iterations or trials) (*87*).

Chapter II: Predicting *Salmonella* Populations from Biological, Chemical, and Physical Indicators in Florida Surface Waters

Rachel McEgan[1], Gabriel Mootian[2], Lawrence D. Goodridge[3], Donald W. Schaffner[2], and Michelle D. Danyluk[1]*

[1] Department of Food Science and Human Nutrition, Citrus Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Lake Alfred, Florida, USA

[2] Department of Food Science, School of Environmental and Biological Sciences, Rutgers, New Brunswick, New Jersey, USA

[3] Center for Meat Safety and Quality, Food Safety Cluster, Department of Animal Sciences, Colorado State University, Fort Collins, Colorado, USA

Received 11 March 2013 Accepted 23 April 2013

Published ahead of print 26 April 2013.Address correspondence to Michelle D. Danyluk, mddanyluk@ufl.edu.

II.1 Abstract

Coliforms, *Escherichia coli*, and various physicochemical water characteristics have been suggested as indicators of microbial water quality or index organisms for pathogen populations. The relationship between the presence and/or concentration of *Salmonella* and biological, physical, or chemical indicators in Central Florida surface water samples over 12 consecutive months was explored. Samples were taken monthly for 12 months from 18 locations throughout Central Florida (n=202). Air and water temperature, pH, oxidation-reduction potential (ORP), turbidity, and conductivity were measured. Weather data were obtained from nearby weather stations. Aerobic plate counts and most probable numbers (MPN) for *Salmonella, E. coli,* and coliforms were performed. Weak linear relationships existed between biological indicators (*E. coli*/coliforms) and *Salmonella* levels (R^2 <0.1) and between physicochemical indicators and *Salmonella* levels (R^2 <0.1). The average rainfall (previous day, week, and month) before sampling did not correlate well with bacterial levels. Logistic regression analysis showed that *E. coli* concentration can predict the probability of enumerating selected *Salmonella* levels. The lack of good correlations between biological indicators and *Salmonella* levels and between physicochemical indicators and *Salmonella* levels shows that the relationship between pathogens and indicators is complex. However, *Escherichia coli* provides a reasonable way to predict *Salmonella* levels in Central Florida surface water through logistic regression.

II.2 Introduction

Water quality can be characterized by three types of analysis, index microorganisms, indicator microorganisms, and various physicochemical water characteristics. Indicator microorganisms are used to suggest, or "indicate," the possible presence of fecal contamination (*75*). Index organisms represent the presence and behavior of a pathogen in a given environment (*57*), and one organism can be both an indicator and an index organism. Physicochemical water characteristics include turbidity, temperature, pH, and oxidation-reduction potential (ORP) (*2*).

Physicochemical measurements have the advantage of being considerably more rapid than microbial-based measurements, but such measurements may not correlate with microbiological quality. The United States produce industry currently relies on testing 100 ml of water for indicator organisms, specifically, generic *Escherichia coli,* as a means to monitor microbial water quality used for edible horticultural crop production (*11, 27, 82*) There are a wide variety of bacterial genera, groups, and species, viruses, and bacteriophages that have been used or proposed for use as indicator microorganisms (*2, 3*, *21*, *33*). Coliforms, either total or fecal, are a common choice of indicator organism. Total coliforms are aerobic and facultatively anaerobic, gram-negative, non-spore-forming, rod-shaped bacteria that pro- duce gas and acid upon lactose fermentation within 48 h at 35 °C; fecal coliforms are a subset of coliforms that also ferment lactose at 44 °C. Generic *E. coli* has been proposed as both an indicator and an index microorganism for enterohemorrhagic *E. coli*, *Salmonella* spp., and *Shigella* spp. (*48*). When contaminated agricultural water comes into contact with the edible portion of a plant, for example, during irrigation, pesticide and fertilizer application, and/or frost protection, the safety of fresh produce items has been affected

(*20*, *43*, *55*). In 2005, a large-scale outbreak of *Salmonella enterica* subsp. *enterica* serovar Newport from tomato consumption occurred, and the trace-back investigation revealed that the same serovar was in the irrigation pond of the farm from which the suspect tomatoes were harvested (*34*). Another large outbreak in 2009 was epidemiologically linked to Serrano peppers, which came from a farm where a matching *Salmonella enterica* subsp. *enterica* serovar Saintpaul strain was isolated from the irrigation pond (*46*). Both outbreaks demonstrate that the quality of water applied to fresh produce is of at most importance in ensuring food safety. Contradictory results have been reported as to the efficacy of index or indicator organisms in predicting the presence and/or prevalence of human pathogens, such as *Salmonella,* in surface waters (*10*, *14*, *69*). While any specific pathogen may not be present in the sample being tested, the presence of index microorganisms is meant to suggest that pathogens have a reasonable likelihood of being present (*75*). The tests for index microorganisms are less cumbersome and time- consuming, as well as typically less expensive, than screening for the presence of individual pathogens. This makes the detection of index microorganisms a much more economical and practical choice than screening for individual pathogens.

Physicochemical water characteristics, such as turbidity, temperature, pH, and ORP, have all been used to monitor water quality in rivers and lakes (*2*), but the presence or absence of correlations between physicochemical measurements and microbial measurements of water is not discussed by the American Public Health Association (APHA) in their methods manual. Physical measurements like temperature and rainfall ($R^2 = 0.317$) have been correlated with *Salmonella* isolation frequency in seawater (*84*). The main advantage of monitoring a physicochemical characteristic is that nearly

instantaneous results can be obtained and used to predict water quality so that a risk management decision can be made in a timely manner. The relationship between *Salmonella* concentration and biological index organisms or physicochemical indicators in Central Florida surface waters over 12 consecutive months was explored.

II.2 Materials And Methods

II.2.a Water sampling

Eighteen surface water sites across Central Florida were sampled monthly for 12 consecutive months beginning in August 2010. Sites in rural agricultural areas, away from animal agriculture, including ponds, creeks, rivers, and canals, were selected. Sampling locations included two lakes, one pond, six creeks, two streams, one river, and six canals. The land use and catchment watershed for each site are detailed in **Table II.a.** Water was collected in 10-liter sterile carboys (Nalgene, Rochester, NY) fitted with 4 kg of lead weights attached to a rope. The rope was used to lower the mouth of the carboy to a depth of 20 cm into the water source. Samples from all but one location were collected away from the shore, using a bridge or other outcropping. Samples were collected on three separate days each month (six samples per day). All samples were taken before solar noon (the moment when the sun reaches its highest elevation on a given day at a given place) from a shaded area of the water source. Chemical and physical water characteristics were measured immediately after sampling as described below. Samples were transported to the laboratory and stored at 4 °C for up to 24 h before microbial testing as described below. Samples were transported at ambient temperature, protected from the sun. No more than 4 h elapsed between sample collection and 4 °C storage.

Table II.a: Sampling site water source type and catchment area usage

Site no.	Water source type	Immediate area land use(s)	% of watershed used for: Agriculture	Rangeland	Urban/suburban
1	Lake	Agriculture, wetlands, residential	33	<1	29
2	Lake	Agriculture, wetlands	33	<1	29
3	Pond	Agriculture	33	<1	29
4	Creek	Wetlands	33	<1	29
5	Creek	Wetlands	33	<1	29
6	Stream	Agriculture	32	3	20
7	Creek	Agriculture	26	<1	49
8	Creek	Agriculture	32	12	25
9	Creek	Agriculture	32	12	25
10	Stream	Agriculture	32	12	25
11	River	Wetlands, agriculture	32	12	25
12	Creek	Agriculture	45	13	10
13	Canal	Agriculture	55	NA[a]	21
14	Canal	Agriculture	55	NA	21
15	Canal	Agriculture	55	NA	21
16	Canal	Agriculture	55	NA	21
17	Canal	Agriculture	55	NA	21
18	Canal	Agriculture	55	NA	21

[a] NA, not available.

II.2.b Chemical and physical water characteristics

Turbidity, temperature (air and water), pH, and ORP were measured. Each measurement was repeated in triplicate. Turbidity was measured in formazin attenuation units (FAU) using a portable colorimeter (DR/850;Hach Company, Loveland, CO) according to the manufacturer's instructions. Water and air temperatures were measured with a portable temperature probe (SH66A; Cooper Instrument Corporation, Middlefield, CT). The pH value and ORP were measured with a portable pH/ORP meter (pH 6 Acorn series; Oakton, Vernon Hills, IL).

II.2.c Total aerobic plate count

Surface water samples were spread plated in duplicate onto tryptic soy agar (TSA; Difco, Becton, Dickinson, Sparks, MD) following serial dilutions in 0.1% peptone water (Difco, Becton, Dickinson, Sparks, MD) to determine total aerobic counts. Plates were incubated for 24 ± 2 h at 35 ± 2 °C, and all bacterial colonies were enumerated by hand.

II.2.d Coliform and *E. coli* MPN

Colisure Presence Absence Snap Packs (IDEXX Laboratories, Inc., Westbrook, Maine, USA) were used to determine coliform and *Escherichia coli* Most Probable Numbers (MPN) in a five by three MPN configuration (10, 1, 0.1 ml dilutions). Tubes were incubated 24 h at 35 ± 0.5 °C. A yellow color indicated coliforms and *E. coli* was determined by observing fluorescence using a 6-watt UV lamp. The MPN/100 ml was determined from the table in Standard Methods for the Examination of Water and Wastewater, 18th ed. (*2*).

II.2.e *Salmonella* detection

Salmonella was determined in each water sample as previously described by McEgan et al. (2013). Briefly, each water sample was concentrated using tangential flow filtration (TFF) to a final retentate volume of ca. 250 ml. A KrosFlo Research II Pump (vendor, location) was used in combination with Masterflex easy-load pump head (Spectrum Labs, Rancho Dominguez, CA, USA). The inlet flow rate was 1000 ml/min. A Mini Kros Plus Tangential Flow Filter Module (Spectrum Labs) made of polyethersulfone with a pore rating of 0.2 µm and a surface area of 1050 cm^2 was used as a tangential flow filter. The TFF was run at a transmembrane pressure of 67 Pa. Double strength lactose broth (250 ml; Difco, Becton Dickinson) was added to the retentate, which was then stored for 1 h at room temperature and incubated 24 ± 2 h at 35 ± 2 °C, to pre-enrich the sample.

Pre-enrichment was followed by an immuno-capture of *Salmonella* using the Pathatrix system (Matrix MicroScience, Golden, CO, USA). The method was modified such that the entire pre-enrichment was placed in a sterile stomacher bag (17.8 x 30.5 cm; Fisher Scientific, New Jersey, USA; instead of the typical 100 ml conical tube). Pathatrix

Salmonella capture beads (50 µl) were added to the system and used according to the manufactures' instructions in a pre-set 30 min cycle.

The beads (40 µl) were used in a DNA extraction using the MoBio UltraClean DNA kit (MoBio, Carlsbad, CA, USA). The extracted DNA was used in real time PCR, using Applied Biosystems' MicroSEQ *Salmonella* spp. Detection Kit (Applied Biosystems, Carlsbad, CA, USA). The real time PCR kit was used according to manufacturers' instructions in a BioRad CFX96 RealTime thermocycler.

II.2.f Modified MPN for *Salmonella* positive water samples

A 50 ml retain sample of each water sample was removed prior to processing. A modified *Salmonella* MPN method was used to determine *Salmonella* concentration (*84*) for each retained sample. A three-by-three tube MPN was set up and dilutions were as follows: 10 ml in 10 ml double strength lactose broth, 1 ml in 9 ml single strength lactose broth, and 0.1 ml in 9 ml single strength lactose broth. These were incubated for 24 ± 2 h at 35 ± 2 °C. Selective enrichment was done by transferring a one ml aliquot to tetrathionate broth (TT broth; Difco, Becton Dickinson) and 0.1 ml to Rappaport Vassiliadis broth (RV broth; Difco, Becton Dickinson); selective enrichment broths were incubated 24 ± 2 h at 35 ± 2 °C and 48 ± 2 h at 41 ± 2 °C, respectively. Ten microliters were streaked onto XLT4 and Chromagar *Salmonella* Plus and incubated for 24 ± 2 h at 35 ± 2 °C following enrichment. Colonies displaying typical *Salmonella* phenotypes were confirmed biochemically on lysine iron agar slants (LIA; Difco, Becton Dickinson) and triple sugar iron agar slants (TSI; Difco, Becton Dickinson).

One representative biochemically confirmed *Salmonella* colony from each plate was transferred to TSA and incubated for 24 ± 2 h at 35 ± 2 °C. One colony was transferred to tryptic soy broth and incubated 24 ± 2 h at 35 ± 2 °C and DNA extraction

using the MoBio UltraClean DNA kit. *Salmonella* was genetically confirmed by PCR of the *invA* and *oriC* genes. The *invA* primers were GTGAAATTATCGCCACGTTCGGGCAA and TCATCGCACCGTCAAAGGAACC, giving a PCR product of 284 bp; the *oriC* primers were TTATTAGGATCGCGCCAGGA and AAAGAATAACCGTTGTTCAC, giving a PCR product of 163 bp. Both primer sets were as described by Malorny et al. (2003). All reagents were obtained from the Fisher exACTGene Complete PCR kit (Fisher Scientific, New Jersey, USA) and primers were used at a concentration of 20 µM. The PCR reaction mix was as follows: 34.75 µl water; 5 µl 10X PCR buffer; 1 µl of each primer; 0.25 µl Taq DNA polymerase; 5 µl template DNA. The optimized PCR conditions were: 3 min at 94 °C melting, followed by 30 cycles of 30 s at 94 °C, 30 s at 58 °C, 1 min at 72 °C; a final elongation of 5 min at 72 °C. Gel electrophoresis was done using a 1.8% agarose gel with 0.5X TBE buffer. Gels were run at 70 V for 90 min.

II.2.g Weather data collection

Rainfall and solar radiation data were collected from the Florida Automated Weather Network (http://fawn.ifas.ufl.edu/). The closest weather recording station to each individual sampling site was determined and the data applied accordingly.

II.2.h Statistical analysis

Correlations were determined using JMP Pro 9. Multiple linear regression analysis of *Salmonella* vs. all parameters was done using SAS 9.3 (SAS Institute Inc., North Carolina, USA) software. A p value of < 0.05 was set as the condition for entry of any parameter into a model.

To determine the probability of enumerating *Salmonella*, binomial logistic regression analysis using SAS 9.3 was done with all the other parameters as independent

variables. Binomial logistic regression assumes a response variable with two possible outcomes (e.g. *Salmonella* is present or *Salmonella* is not present). Since all water samples tested had *Salmonella* counts above detection limit, the simple absence or presence of *Salmonella* could not be used as a criterion. To overcome this limitation, we modeled the probability of enumerating a concentration of *Salmonella* exceeding a particular value. Six levels of *Salmonella* were selected for modeling: 3, 5, 10, 15, 20, and 60 MPN/100 ml. Using the lowest concentration as an example, *Salmonella* observations were converted to a value of 1 if an observation was ≥ 3 MPN/100 ml and a value 0 if an observation was < 3 MPN/100 ml. The same procedure was repeated for 5, 10, 15, 20 and 60 MPN/100 ml. This binomial logistic regression analysis was done with the coded *Salmonella* data as the dependent variable and all the other variables (turbidity, conductivity, air temperature, water temperature, pH, ORP, coliforms and *E. coli*) as independent variables.

II.3 Results

Two hundred and two samples were collected and analyzed over the twelve-month sampling period. Fourteen samples were missed due to laboratory error or inaccessibility of the water source due to unrelated factors (e.g. nearby herbicide application precluded access due to investigator safety). The correlation coefficients for each characteristic vs. *Salmonella* log MPN/100 ml are displayed in **Table II.b**. When each R^2 value is calculated for individual sampling sites 152 of the 162 of the R^2 values fall below 0.4 indicating a lack of correlation. Each characteristic will be discussed further in the appropriate section below.

II.3.a *Salmonella* MPN

All 202 water samples were positive for *Salmonella* using the described TFF-immuno-capture Real Time PCR method. The detection level for the described methodology has a detection limit of ca. 1 CFU/l in 10 l. All 10 l surface water samples collected had greater than 1 CFU/l *Salmonella* spp. in each 10 l sample volume collected and processed.

Salmonella concentrations for each site by sampling date can be seen in **Figure 2.1**. The median *Salmonella* concentration was 0.79 log MPN/100 ml for all samples taken; the upper quartile was 1.2, the lower quartile was 0.48 log MPN/100 ml; and 139 of 202 samples had *Salmonella* concentrations below 1 log MPN/ 100 ml. The highest concentration of *Salmonella,* 3.0 log MPN/100 ml, was identified in July (**Fig 2.1)**; however, the month of sampling was not strongly correlated with the concentration of *Salmonella* ($R^2 = 0.20$).

II.3.b Coliform and *E. coli* MPN

Values for coliform and *E. coli* log MPN/100 ml can also be seen in **Figure 2.1.** Many of the coliform (128/202), and some of the *E. coli* (17/202), MPNs were at or over the upper limit of detection (≥ 3.2 log MPN/100 ml) for the MPN method used. The median coliform concentration was ≥ 3.2 log MPN/100 ml, with the lower quartile at ≥ 2.7 log MPN/100 ml. *E. coli* had a median concentration of 1.7 log MPN/100 ml, an upper quartile of 2.4 and lower quartile of 0.60 log MPN/100 ml. Neither coliform log MPN/100 ml ($R^2 = 0.004$; **Table II.b**), nor *E. coli* log MPN/100 ml ($R^2 = 0.078$; **Table II.b**) values were strongly correlated with the *Salmonella* MPN values. However, at sites 11 and 18, *E. coli* log MPN/100 ml resulted in higher correlations, $R^2 = 0.606$ and $R^2 =$

0.678, respectively. Interestingly, coliform and *E. coli* log MPN/100 ml were not strongly correlated ($R^2 = 0.364$; **Table II.c**) with each other.

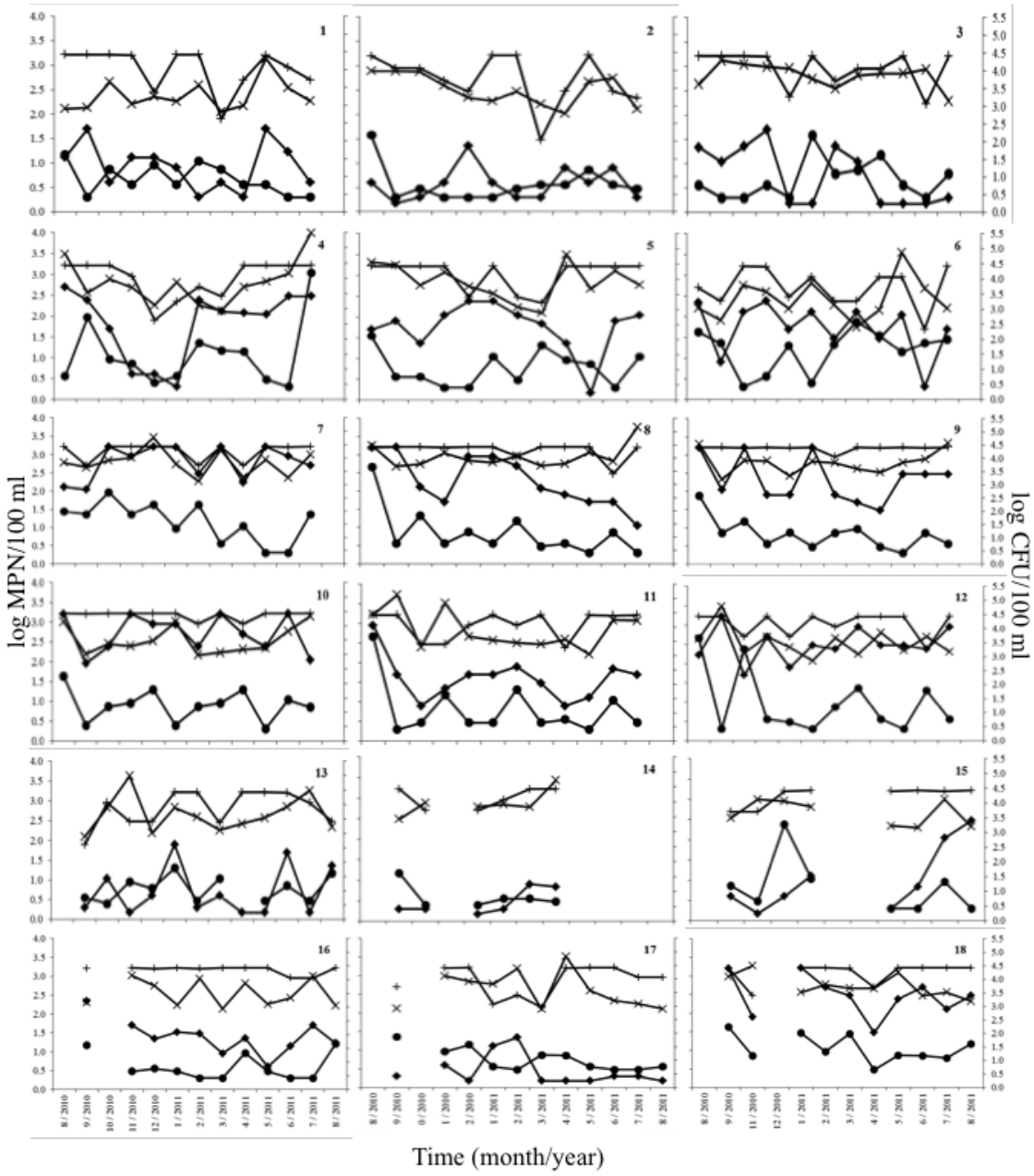

Figure 2.1: Populations of *Salmonella*, *E. coli* and coliforms (left axis), and aerobic plate for 18 Central Florida sites sampled monthly for a continuous 12-month period (right axis). *Salmonella* (circle), *E. coli* (diamond) and coliforms (+), reported in log MPN/100

ml, and aerobic plate counts (x), reported in log CFU/100 ml, for 18 Central Florida sites sampled monthly for a continuous 12 month period.

Table II.b: Coefficients of determination for log MPN *Salmonella*/100 ml versus each of the physical, chemical, and biological water characteristics recorded from each sampling by sampling site

	R^2 for log MPN *Salmonella*/100 ml versus:								
Site no.	*E. coli* (log MPN/100 ml)	Coliforms (log MPN/100 ml)	APC (log CFU/100ml)	Water temp (°C)	Air temp (°C)	pH	ORP (mV)	Turbidity (FAU)	Conductivity (µS/cm)
1	0.108	0.009	0.001	0.000	0.158	0.003	0.007	0.220	0.021
2	0.000	0.059	0.119	0.205	0.235	0.006	0.008	0.011	0.210
3	0.082	0.079	0.178	0.101	0.043	0.123	0.146	0.141	0.033
4	0.149	0.087	0.144	0.099	0.123	0.006	0.027	**0.691**	**0.405**
5	0.034	0.003	0.003	0.069	0.103	0.068	0.080	0.152	0.036
6	0.097	0.254	0.344	0.014	0.127	0.177	0.165	0.195	0.039
7	0.066	0.039	0.023	0.026	0.076	0.009	0.271	0.042	0.265
8	0.263	0.003	0.007	0.031	0.019	0.052	0.240	0.078	0.194
9	0.150	0.001	0.092	0.065	0.016	0.021	**0.424**	0.055	0.315
10	0.278	0.037	0.038	0.037	0.035	0.071	**0.555**	0.359	0.269
11	**0.606**	0.003	0.099	0.005	0.039	0.272	0.070	**0.584**	0.355
12	0.230	0.079	0.020	0.119	0.136	0.204	0.346	0.019	0.024
13	0.372	0.019	0.000	0.153	0.303	0.010	0.005	0.109	0.018
14	0.017	0.322	0.354	**0.408**	0.120	0.191	0.154	0.023	0.030
15	0.016	0.016	0.377	0.138	**0.430**	0.075	0.022	0.010	0.065
16	0.131	0.168	0.069	0.148	0.041	0.013	0.037	0.024	0.054
17	0.107	0.004	0.000	0.022	0.087	0.185	0.102	0.152	**0.450**
18	**0.678**	0.247	0.002	0.001	0.020	0.000	0.101	0.237	0.053
Overall	0.078	0.000	0.004	0.000	0.006	0.076	0.070	0.010	0.015

[a] Values greater than 0.4 are in bold.

II.3.c Aerobic plate count

The values for each aerobic plate count for each site can be seen in **Figure 2.1**. The median aerobic plate count was 3.7 log CFU/100 ml; the upper quartile was 4.0 log CFU/100 ml, and the lower quartile was 3.2 log CFU/100 ml. The two highest aerobic plate counts, 5.5 and 5.2 log CFU/100 ml (**Figures 2.1-4 and 2.1-8**), both occurred during the month of July. The aerobic plate count, in log CFU/100 ml, did not correlate with the *Salmonella* log MPN/100 ml ($R^2 = 0.004$; **Table II.b**).

II.3.d Air and water temperature

Air and water temperatures are displayed in **Figure 2.2**. Not surprisingly, recorded air and water temperatures showed the strongest correlation of any two variables ($R^2 = 0.680$; **Table II.c**). Neither the air nor water temperature correlated with the

Salmonella log MPN/100 ml (R^2 = 0.000 and R^2 = 0.006, respectively; Table **II.b**). However, when calculated on an individual site bases, site 14 (R^2 = 0.405; **Table II.b**) and site 15 (R^2 = 0.403; **Table II.b**) *Salmonella* concentrations showed relatively higher correlation with water and air temperatures, respectively.

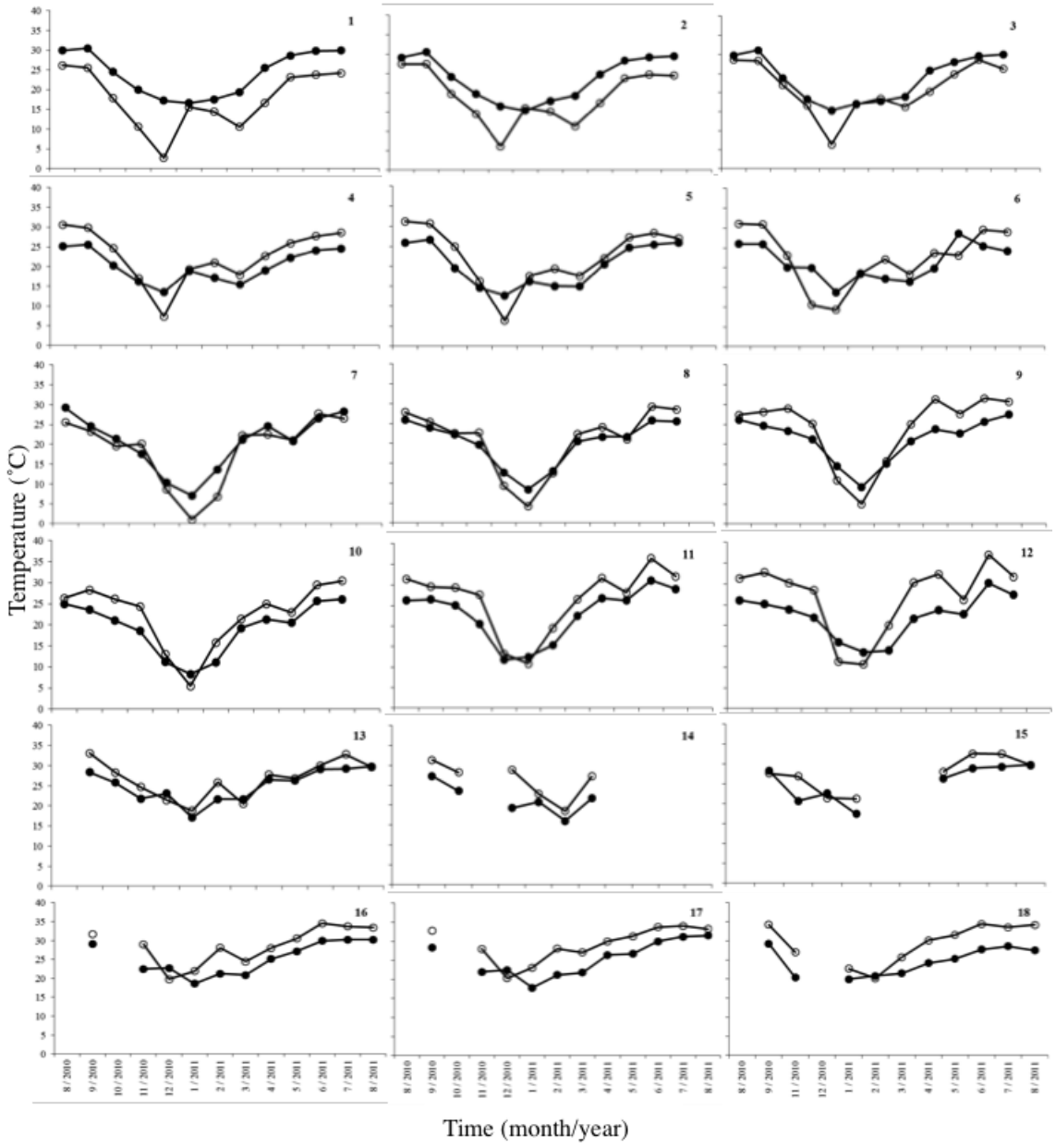

Figure 2.2: Air (open circle) and water (dark circle) temperatures, reported in °C, measured using a temperature probe at the time of sampling for 18 central Florida sites sampled monthly for a continuous 12-month period.

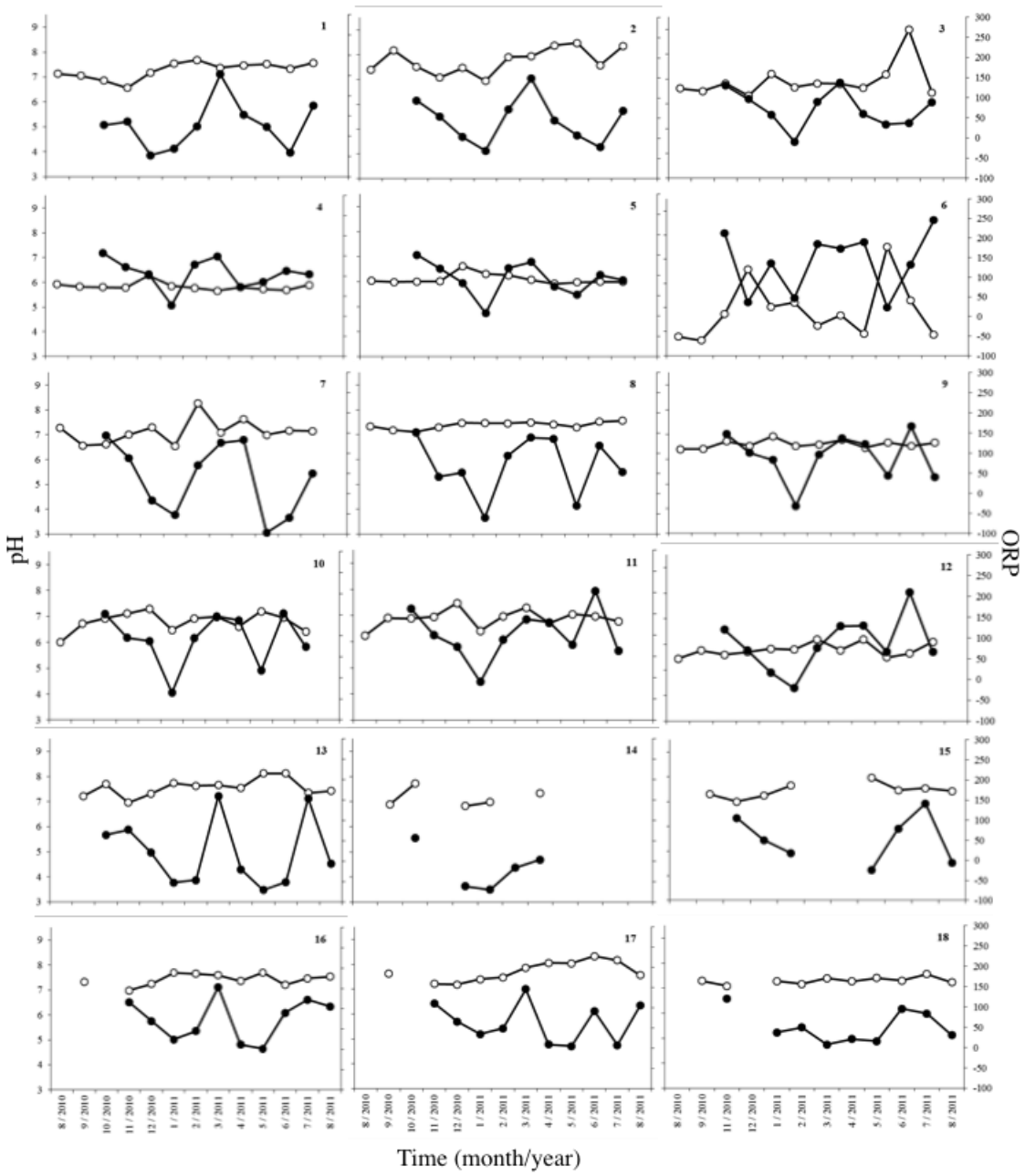

Figure 2.3: pH (open circle; left axis) and ORP (dark circle; right axis), measured at the time of sampling for 18 central Florida sites sampled monthly for a continuous 12-month period.

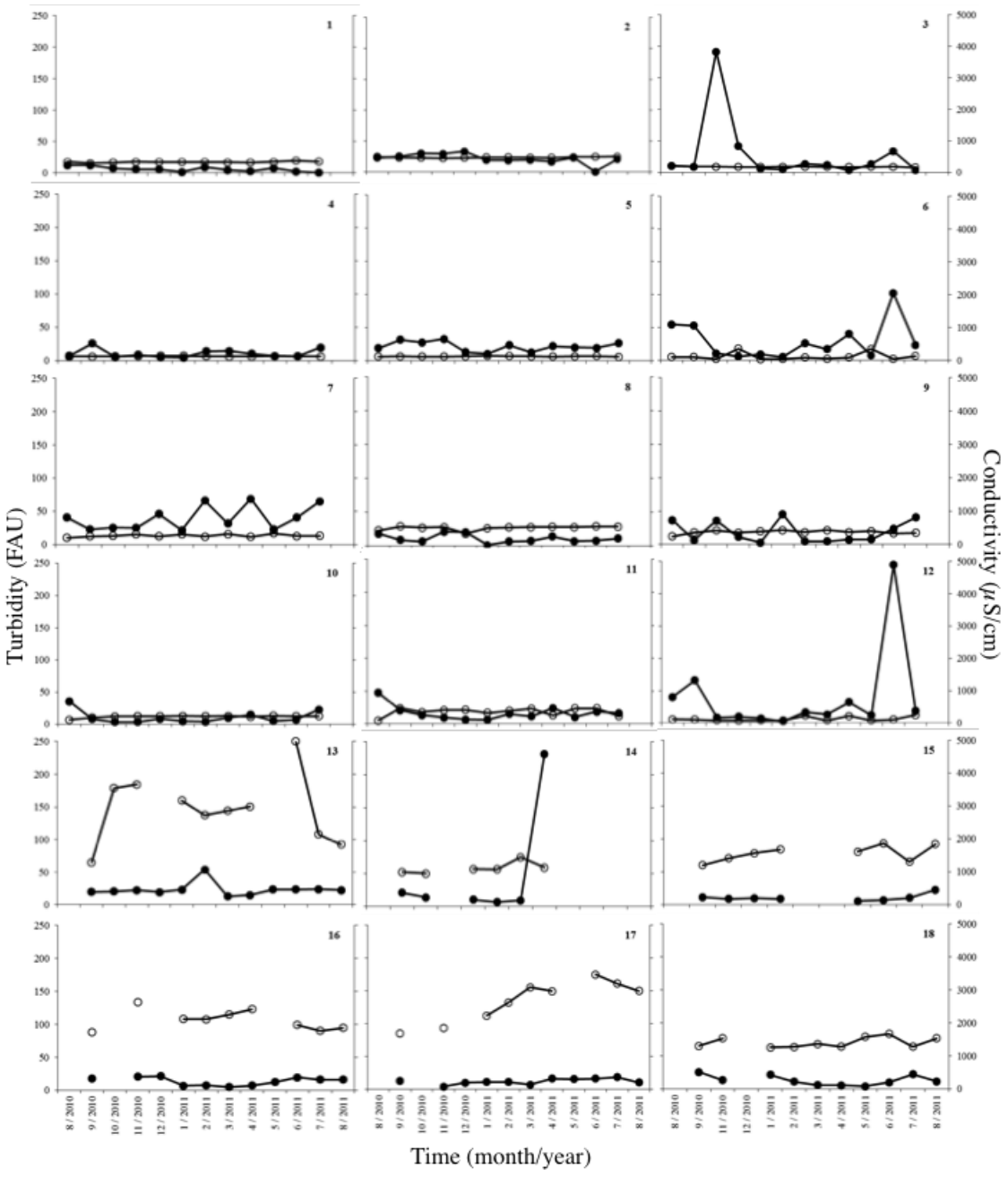

Figure 2.4: Turbidity (dark circle; left axis) and conductivity (open circle; right axis), measured at the time of sampling for 18 central Florida sites sampled monthly for a continuous 12-month period.

II.3.e pH, ORP, turbidity, and conductivity

pH and ORP values by sampling site over time are displayed in **Figure 2.3**; and similarly turbidity and conductivity by sampling site over time are displayed in **Figure 2.4**. None of these physical water characteristics correlated with *Salmonella* log MPN/100 ml; as show by the R^2 values in **Table II.b**. However, when calculated by site, ORP had a relatively high correlation with *Salmonella* concentration at site 9 ($R^2 = 0.424$) and site 10 ($R^2 = 0.555$); while turbidity had a relatively high correlation with *Salmonella* concentration at site 4 ($R^2 = 0.691$) and site 11 ($R^2 = 0.584$); and conductivity had a relatively high correlation with *Salmonella* concentration at site 4 ($R^2 = 0.405$) and site 17 ($R^2 = 0.450$). pH correlations were below 0.4 at all sites.

II.3.f Rainfall and solar radiation

The average rainfall (in the previous day, week or month) before sampling did not correlate with *Salmonella*, coliform, *E. coli*, or aerobic plate count levels (data not shown). In addition solar radiation was also poorly correlated with *Salmonella,* coliforms, *E. coli* or aerobic plate counts levels (data not shown).

II.3.g Multiple linear and logistic regression analysis

A multiple linear regression analysis with pH, ORP, conductivity, turbidity, water temperature, air temperature, coliforms, and aerobic plate count as independent variables and *Salmonella* as the dependent variable, showed that *E. coli* and solar radiation were the only variables that were significantly correlated with *Salmonella* concentration ($p < 0.05$), however the resulting R^2 was very low (< 0.1). **Table II.c** shows Pearson product moment correlation coefficients and p values for all the parameters. Correlation coefficients closer to 1 or -1 indicated a strong linear relationship between variables, similar to the relationship indicated by R^2. Positive correlation coefficients indicated that

the two variables are positively correlated, while a negative correlation coefficient indicates the opposite. **Table II.c** shows that *E. coli* and solar radiation levels were positively correlated ($p < 0.05$) with *Salmonella* concentrations but with low correlation coefficients.

Table II.c: Pearson product moment correlation coefficients (r) and p values, determined between each of the physical, chemical, and biological water characteristics for all sampling sites and months combined

Parameter	Water temp (°C)	Air temp (°C)	pH	ORP (mV)	Turbidity (FAU)	Conductivity (µS/m)	Coliforms	*E. coli*	Aerobicplate count	Solar Rad. (W/m2)	*Salmonella*
Water temp (°C)	**r**	**0.825[a]**	**0.208[a]**	**-0.0117**	**0.144[a]**	**0.1300**	**-0.0165**	**−0.144[a]**	**-0.0125**	**-0.0330**	**0.1210**
	p	0.0000	0.0031	0.8800	0.0418	0.0707	0.8170	0.0432	0.8620	0.6410	0.0946
Air temp (°C)	**r**		**0.0127**	**0.0254**	**0.1410**	**0.225[a]**	**0.0307**	**-0.0744**	**0.0297**	**-0.0696**	**0.0538**
	p		0.8580	0.7450	0.0473	0.0017	0.6680	0.2990	0.6790	0.3250	0.4590
pH	**r**			**-0.1330**	**-0.0686**	**0.423[a]**	**-0.0479**	**-0.0141**	**0.0398**	**-0.0101**	**0.0506**
	p			0.0867	0.3370	0.0000	0.5040	0.8450	0.5790	0.8860	0.4880
ORP(mv)	**r**				**-0.0209**	**-0.0566**	**0.0799**	**-0.0499**	**-0.0151**	**0.1010**	**-0.0578**
	p				0.7880	0.4720	0.3020	0.5200	0.8450	0.1870	0.4640
Turbidity(FAU)	**r**					**-0.0878**	**−0.172[a]**	**-0.0297**	**-0.0336**	**-0.0819**	**0.0824**
	p					0.2210	0.0151	0.6780	0.6390	0.2460	0.2570
Conductivity(µS/m)	**r**						**-0.0592**	**−0.317[a]**	**-0.0568**	**−0.273[a]**	**-0.1210**
	p						0.4130	0.0000	0.4320	0.0001	0.0994
Coliforms	**r**							**0.364[a]**	**0.212[a]**	**-0.0670**	**0.0032**
	p							0.0000	0.0029	0.3450	0.9650
E. coli	**r**								**0.0386**	**0.151[a]**	**0.248[a]**
	p								0.5920	0.0330	0.0005
Aerobic plate count	**r**									**-0.0132**	**0.0314**
	p									0.8520	0.6680
Solar Rad.(W/m^2)	**r**										**0.165[a]**
	p										0.0216

Aerobic plate count was in log CFU/100 ml, *Salmonella*, *E. coli* and coliforms were (log MPN/100 ml).

Given the absence of a strong linear relationship between *Salmonella* concentrations and any variable investigated, multiple logistic regression analysis was used to predict the probability of *Salmonella* concentration exceeding a given concentration as a function of any of the variables investigated. *E. coli* concentration was the only variable that was significantly correlated with the probability of *Salmonella* concentration exceeding a certain value ($p < 0.05$). A model relating the probability of

Salmonella occurring at levels ≥ 3, 5, 10, 15, 20, or 60 MPN/100ml was developed. The model has the form:

Logit P = C + K * *E. coli* (Log MPN/100 ml)…….. (1)

Where logit P is defined as the natural log (ln) of (P/1-P), C is a constant and K is the model coefficient. Individual plots for each criteria used to predict the probability of occurrence of *Salmonella* given any level of *E. coli* in surface water are shown in **Figure 2.5**. With higher *E. coli* concentrations in surface water the probability of enumerating *Salmonella* at any of the six concentrations (3, 5, 10, 15, 20 and 60 MPN/100 ml) rises in a roughly proportionate manner from the lowest level of *E. coli* observed (-1 log MPN/100 ml) to the highest (3.2 log MPN/100 ml). The probability of *Salmonella* exceeding a particular value was also arrayed in an essentially logical manner. For example when the *E. coli* concentration is 1 log MPN/100 ml, the model predicts an 80% chance of enumerating ≥ 3MPN/100 ml, a 43% chance of enumerating ≥ 5MPN/100 ml *Salmonella*, 24% chance for enumerating *Salmonella* ≥ 10MPN/100ml, 20% chance of enumerating *Salmonella* ≥ 15 MPN/100 ml, 17% chance of enumerating ≥ 20 MPN and a 1% chance of enumerating *Salmonella* ≥ 60 MPN. All of the models are statistically significant at $p<0.05$, with the exception of the *Salmonella* ≥ 3MPN/100 ml model ($p = 0.815$).

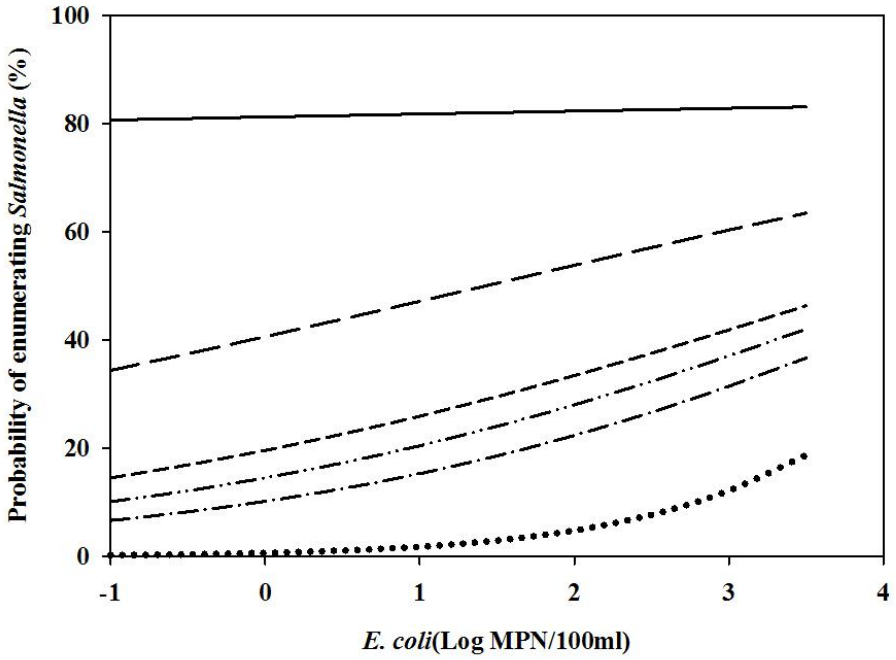

Figure 2.5: Logistic regression model for predicting the probability of enumerating *Salmonella* ≥ 3 MPN/100 ml (solid line, p=0.815), 5 MPN/100 ml (long dashes, p=0.035), 10 MPN/100 ml (short dashes, p=0.014), 15 MPN/100 ml (dash two dots, p=0.010), 20 MPN/100 ml (dash one dot, p=0.009) and 60 MPN/100 ml (dotted, p=0.020) MPN/100 ml.

II.4 Discussion

Salmonella is present in central Florida surface waters throughout the year. Surveys of surface waters for *Salmonella* in the Southeastern United States have reported similar findings; in North Carolina, surface water had a *Salmonella* prevalence of 54.7% (94/172) in 25 ml water samples (*62*), and in a southern Georgia watershed, 79.2% of water samples

were *Salmonella* positive using a five replicate, three dilution (100, 10, and 1 ml) MPN method (*36*). A survey of the Suwannee River in northern Florida reported that 96% (106/110) of water samples were *Salmonella* positive using a three replicate, five dilution (500, 100, 50, 10, and 1 ml) MPN method (*66*) These findings are similar to those found by other researchers performing surface water samplings over extended periods of time at the same set of sites in more geographically diverse locations. Two rivers and one creek, all in one Southern Ontario, Canada, water shed, had 78.4% (91/116) *Salmonella* positive samples where locations were sampled twice monthly for 2.5 years. Other studies have reported lower *Salmonella* prevalence in surface waters. Surface waters in the Salinas Valley of California, surveyed using either Moore swabs or a 100-ml sample, report 7.1% (18/252) *Salmonella*-positive samples (*32*). The prevalence of *Salmonella* in the coastal waters of southern Morocco during 4-year monthly sampling of six sites was 4.1% (10/243) (*74*). Similarly, *Salmonella* was present in 11% (16/145) of 100 ml samples of surface waters collected from fruit and vegetable farms in New York state (*80*). The presence of *Salmonella* throughout the year in Central Florida surface waters is similar to that in surface water surveys done throughout the Southeastern United States, especially when considering the increased sample volume and low limit of detection used in the current study.

No regulatory limit currently exists for *Salmonella* in agricultural waters, and as larger volumes of waters are screened for *Salmonella*, the limit of detection decreases accordingly (*53*). Protocols for *Salmonella* detection are expensive and time-consuming compared to standard water testing of 100 ml for generic *E. coli*. Mitigation strategies suited for treating typical volumes of water used in agricultural production when

Salmonella might be detected are very limited. Therefore, the ability of index organisms, such as *E. coli,* to predict the probability of enumerating *Salmonella* levels in surface waters intended for agricultural use is desirable.

None of the index microorganisms measured strongly linearly correlated with the prevalence of *Salmonella* in the same Central Florida surface water sample. Many of the coliform log MPN/100 ml counts were at or above the upper limit of detection; this may be responsible for the lack of correlation between *Salmonella* and coliform populations. In the cases in which coliform populations were at or above the upper limit of detection, the population could have been higher and would not be appropriately reflected by the data point. All water samples in this study with coliform counts above 3.2 log MPN/100 ml, the upper limit of detection, were recorded as 3.2 log MPN/100 ml, adding error into the calculations of the correlation coefficients.

The presence of weak correlations between index microorganisms and pathogens in surface water (*12*, *17*, *19*, *33*, *42*, *64*) Moderate correlations were obtained when *E. coli* or coliform data from the current study were analyzed for correlations with *Salmonella* populations for individual sites (**Table II.b**) however, when the data were aggregated for all site correlations between *E. coli* or coliforms and *Salmonella*, both had R^2 values of 0.1, indicating that other factors beyond those studied here influence the relationships between pathogens and index microorganisms. Studies that report similar results conclude that high correlations between pathogens and index microorganisms may be temporal, random, site specific, or time specific (*64*). High correlations between pathogens and index microorganisms often occur at point sources, such as surface waters impacted by improperly treated sewage or runoffs from livestock farms, where the levels

of index microorganisms and pathogens are consistently high for a longer period of time (*64*). Our study was conducted in rural agricultural Florida, where the influence of raw sewage effluent is not expected and was not observed. The occurrence of *Salmonella* in the environment may be at a much higher frequency than previously assumed; under some conditions, such as those in Central Florida, *Salmonella* may not be transient in nature within surface waters as was previously reported (*9*).

In tropical and subtropical watersheds, *E. coli* may be autochthonous rather than an indication of fecal contamination (*21*, *44*). Results reported here, along with those of other studies carried out in geographically similar locations (*36*, *62*, *66*), suggest that the presence of *Salmonella* in the environment may not result solely from a recent fecal contamination. Correlations between *E. coli* and *Salmonella* populations in Central Florida surface waters may not be explicit evidence of fecal contamination but rather of conditions favorable to the survival of both organisms.

Not only did our analysis find no strong correlations between *E. coli* or coliform populations and *Salmonella*, possibly due to no influence of improperly treated sewage or runoff from livestock, *Salmonella* levels were not correlated with rainfall or seasons as has previously been reported (*33*). In our analysis, average rainfall, whether aggregated for 24 h, 1 week, or 1 month (prior to sampling), did not correlate with *Salmonella*, *E. coli*, or coliform levels. Conflicting results have been reported in several other studies that also attempted to determine a correlation between rainfall and pathogen prevalence. A southern Ontario, Canada, surface water study found that the lowest occurrence of *Salmonella* did coincide with the month (February) with the lowest recorded precipitation, air temperature, and water temperature (*81*). A positive correlation between *Salmonella*

prevalence and rainfall (mm/day) was reported, without actual R^2 values being present, during a study in southern Morocco (*74*). The Morocco study analyzed 100 ml of seawater collected monthly at six sites for a 4 year duration. In New York state, *Salmonella* had a higher prevalence in surface waters when measurable precipitation within 3 days prior to sampling; however, this was observed only in areas of poorly drained soils (*80*). The differing results may suggest that rainfall does have some effect on *Salmonella* prevalence in surface waters, not as a direct correlation, but perhaps in a multifactorial way, including more characteristics than those noted in the current study.

A study in Puerto Rico attempted to correlate rainfall from 24 h, 48 h, and 1 week prior to water sampling for fecal coliforms and found no correlation between fecal coliforms and precipitation in any of the 10 sampling sites (*71*). Santiago-Rodriguez et al. (2012) suggested that rainfall may cause correlations to increase due to resuspension of fecal coliforms from sediments but that rainfall may have a possible dilution effect on fecal coliform concentrations, which may lower correlations. Some studies have indicated that turbulence caused by rainfall may resuspend pathogens or indicator microorganisms from sediments, raising their levels in the waters by 100 or 1,000-fold and that the depth and size of the water body dampen any rainfall dilution effect on the indicator/pathogen levels (*33*). Our study did not find a correlation between rainfall and *Salmonella* or *E. coli* and coliform populations. At no point during the water sampling procedure was the sediment disturbed; all sampling sites had great enough depth, and care was taken not to disturb the sediment during sampling. The depth of water varied at and between each of the sampling locations. The consistently low turbidity (**Figure 2.4**) evidences the lack of sediment disturbance. Higher sample turbidity did not correlate with larger *Salmonella*

populations. This may indicate that overt sediment disturbance is not a major contributing factor to the presence of *Salmonella* in Central Florida surface waters. Correlations were not influenced by season, possibly due to the relatively small temperature changes in Central Florida's subtropical climate. However, in southern Ontario and New York State, areas with much larger seasonal temperature variations, *Salmonella* positive samples were not significantly different between seasons (*80*, *81*). The rise and fall in the levels of biological characteristics was random and did not coincide with increasing temperatures in the summer and reduced temperatures in the winter. This is contrary to the finding of Gorski et al. (2011), who noted a noticeable seasonal trend in the prevalence of *Salmonella* in surface water samples from the Salinas Valley of California. Further work is required to acquire a better understanding of whether, and over what range, water temperature would have an effect on *Salmonella* prevalence and whether geographic location or other climacteric factors also have an effect.

Physicochemical water characteristics have been used to monitor the chemical and microbiological qualities of drinking and recreational water. Data from some sites (**Table II.b**) showed that turbidity, ORP, and conductivity can predict the levels of *Salmonella* in surface water, but when aggregated across all sites, none of these parameters produced acceptable correlations. Although R^2 values were low, indicating the possible impact of additional un measured factors, such as dissolved oxygen, total assimilable carbon, or biological oxygen demand. **Table II.c** describes interesting and significant ($P < 0.05$) relationships between the measured parameters and the microbial concentration. For instance, when *E. coli* or solar radiation levels increase, these increases are correlated

with increases in *Salmonella* levels. The positive correlation of solar radiation with *Salmonella* may be partially explained by *Salmonella* being more resistant to solar radiation than are other bacterial pathogens, such as *E. coli*, *Shigella flexneri,* and *Vibrio cholerae* (*4*). As solar radiation increases, less resistant bacteria are inactivated, leaving fewer bacteria to compete for nutrients; *Salmonella* would have access to a greater amount of nutrients as solar radiation increase. When conductivity levels increase, there is a highly statistically significant ($P < 0.0001$) inverse correlation with the *E. coli* concentration. Conductivity of surface water is an indirect measure of salinity and total dissolved solids and may be affected by storm water and runoff (*31*). An inverse correlation between indicator bacteria (*E. coli*, enterococci, and *Bacteroides*) and conductivity or salinity was reported for estuary waters in eastern North Carolina (*31*). The inverse correlation noted in the current study may be due to storm water and runoff effects in which the addition of the storm water or runoff in carrying salinity or other total dissolved solids from the surrounding area also dilutes the *E. coli* concentration. The sampling sites used in the current study were not influenced by seawater. Though these relationships are statistically significant, low coefficient of correlation (1 or -1) values indicate that other as yet unknown factors are also involved.

The complex nature of the index or indicator and pathogen relationship makes predicting the levels of pathogens through index/indicator microorganisms challenging; simple, linear relationships cannot be relied upon for predicting pathogen levels from indicator populations. Logistic regression analysis can be used where a linear relationship between variables is weak. Logistic regressions provide a measure of predictability and have been used to predict the efficiency of a water treatment process in

which the pathogen of interest is poorly correlated to its index/indicator microorganisms (*64*). *E. coli* concentration has a fair ability to predict the *Salmonella* concentration (**Figure 2.5**) in surface water in Central Florida. The model is limited to predictions with the range of *E. coli* concentrations used to develop the model, ranging from the limit of detection (1 log MPN/100 ml) to the maximum possible enumeration level for *E. coli* (3.2 log MPN/100 ml). Predictive models (linear) for *E. coli* have been reported for individual sites for inland recreational lakes in Ohio; explanatory variables included rainfall, turbidity, wind, and water temperature (*28*). Some site models were validated by data collected from an independent year, while others were not able to be validated due to climatic changes (*28*). In the same study, no predictive models were generated for *Salmonella*, but the authors did note that water samples that were PCR positive for *Salmonella* had higher median concentrations of *E. coli* (*28*). The performance of our *Salmonella* predictive model was able to be evaluated by data collected during an independent year or from independent locations.

As part of the implementation of the Food Safety Modernization Act (FSMA), the U.S. FDA has proposed produce safety rules for safe growing, harvesting, packing, and holding of fruits and vegetables grown for human consumption. The proposed produce safety rule recommends collecting 5 samples. Those samples should have a mean *E. coli* concentration of 126 MPN/100 ml, and no single sample should exceed 235 MPN generic *E. coli*/100 ml *(82).* Based on the model for surface water in Central Florida, 126 MPN/100 ml and 235 CFU/100 ml criteria translate to probabilities of containing 5 MPN *Salmonella* /100 ml of 54.4% and 56.2%, respectively. In addition, the levels of generic *E. coli* in our surface water samples

exceeded the recommended criteria ca. 30% of the time, suggesting that, according to proposed standards, some sites tested may be unsuitable for agricultural use.

The absence of strong correlations (low R^2) between index/indicator microorganisms or chemical or physical indicators and *Salmonella* levels limits their ability to predict the prevalence of *Salmonella*. However, significant positive relationships occur between *E. coli* and *Salmonella* levels and between solar radiation and *Salmonella* levels. These may offer a fast but qualitative indication of the degree of *Salmonella* risk. Logistic regression analysis appears to provide the best alternative to a weak quantitative linear model. Thus, pending the discovery of new, more reliable index/indicator microorganisms, *E. coli* levels can be used to predict the probability of enumerating a standard *Salmonella* level in Central Florida surface water and provide a preliminary measure of risk.

Given the dynamic, heterogonous, and complex nature of surface water ecosystems, the challenge remains to find an indicator parameter that is easy, rapid, and inexpensive to test for and that has a more intimate relationship, beyond causal association, with the pathogen and is therefore less sensitive to extraneous factors.

Chapter III: Quantifying the Influence of Weather on the Survival of *E. coli* on the Surface of Oranges Following the Application of Low Microbial Quality Water

Gabriel Mootian[2], Loretta M. Friedrich[1], Travis Chapin[1], Timothy M. Spann[1] Michelle D. Danyluk[1] and Donald W. Schaffner[2*]

[1] Department of Food Science and Human Nutrition, Citrus Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Lake Alfred, Florida, USA

[2] Department of Food Science, School of Environmental and Biological Sciences, Rutgers, New Brunswick, New Jersey, USA

Running title: Influence of weather on survival of microbes on pre harvest crops

*Correspondent Footnote: Corresponding author electronic mail address: schaffner@Aesop.rutgers.edu

To be submitted to: Journal of Food Protection

III.1 Abstract

The influence of weather on microbes deposited onto pre-harvest crops is poorly characterized. This study quantifies the effect of key weather attributes on the fate of *E. coli* introduced onto the surface of oranges in a grove through foliar spraying. Three orange trees were sprayed with low microbial quality water (ca.10^6 CFU *E. coli*/ ml) and three trees (control) were sprayed with well water in 27 monthly field trials. Three replicates of 10 oranges each were harvested from each of the 3 trees at 0, 2, 6 and 24 h intervals and then 1,2 or 3 day intervals until *E. coli* could not be detected by enrichment. This procedure was repeated with 0.1% copper hydroxide (pesticide) added to contaminated spray water in 8 monthly trials. *E. coli* populations were enumerated by plate count and most probable number techniques. Solar radiation (SR), temperature, relative humidity (RH) and rainfall data were also obtained from the Florida Automated Weather Network (FAWN) and plotted against log change in *E. coli* concentration between sampling points for each month. Change in RH, SR, rainfall, temperature and time interval between sampling points did not show strong linear correlations with log change in *E. coli* between sampling points ($R^2 < 0.4$). However solar radiation, relative humidity, temperature and time were significantly ($p < 0.001$) correlated with log change in *E. coli*/h. Logistic regression analysis showed that log reductions in *E. coli* concentration were higher and more likely to occur due to dry weather (high SR and low RH) compared to log increase due to wet weather (low SR and high RH). *E. coli* populations on the surface of oranges were reduced by 2.6 logs when sprayed with water containing 0.1% copper hydroxide at time 0.

Key words: Modeling, *E. coli*, correlation coefficient, likelihood, solar radiation,

III.2 Introduction

Agricultural water is considered one of the common sources of bacterial, viral and protozoal pathogen contamination of fruits and vegetables (*20, 32, 39, 40, 49, 77*). There have been reports of foodborne illness associated with fruits and vegetables traced back to the application of contaminated agricultural water (*6, 39, 77*). Direct application of low microbial quality water to pre harvest crops has implications on the microbial safety of fruits and vegetables. Examples of direct application of agricultural water include dilution of agrochemicals for foliar spray applications, frost protection, hydrocooling and post harvest wash. In 1996 and 1997, an outbreak of Cyclosporiasis associated with Guatemalan raspberries was reported in North America. Investigations revealed that the raspberries had been contaminated through insecticides and fungicides diluted with contaminated water (*41*). *Salmonella* outbreaks associated with tomatoes and Serrano peppers in 2005 and 2009 respectively, were traced to water from irrigation ponds (*34, 46*).

The persistence of food borne pathogens on pre-harvest crops surfaces is dependent on intrinsic characteristics of the pathogen, surface characteristics of the plant, extrinsic ecological conditions and production practices for the crop (*8, 38*). Attachment structures such as fimbriae, curli and flagella, type III secretion system, biofilm formation, bacterial hydrophobicity, virulence genes, cross protection against environmental stresses, horizontal gene transfer, are some intrinsic factors that have been used to explain the establishment and fitness of bacterial pathogens on pre harvest crops (*8, 68*). Plant surface structures such as lenticals, trichomes and stomata, tissue injury, wax, presence of microsites with nutrients and moisture, competition from resident

epiphytic bacteria may affect colonization of plant surfaces by enteric pathogens (*8, 78*). Enteric pathogens may also encounter extrinsic conditions on the surfaces of plants that may affect their survival. Rapid and wide fluctuations in relative humidity (RH), temperature, solar radiation (SR), and rainfall have been reported to cause rapid death of cell populations on leaf and fruit surfaces (*5*, *8*, *78*).

Laboratory based model systems (growth chambers, controlled green houses) have been used to study the influence of weather factors on the survival of enteric pathogens on pre harvest crops. Stine et al. (2005) evaluated the effects of two relative humidity levels on the survival of 8 viral and bacterial pathogens on live lettuce, green peppers and cantaloupes in a controlled environment chamber. The study found that inactivation rates of microorganisms under different relative humidity conditions were variable. Whereas with lab based model systems it is possible to select and control the variables of study, it is difficult to simulate weather factors such as wind, rainfall, fluctuating relative humidity, solar radiation and temperature, soil chemistry variation, predation of enteric pathogens by other microorganisms and production practices that characterize the natural agricultural environment (*38*). Although field based studies are conducted under most natural conditions and using the appropriate agricultural practices, it is difficult to study the effects of any variable on the survival of pathogens. The environment is highly heterogeneous with many potentially correlated and confounding variables at play. One additional source of complexity is that use of pathogens associated with outbreaks presents significant biocontainment and decontamination challenges (*38*). As a result, field studies are rare and most of the available literature on the survival of pathogens on pre harvest crops comes from experiments with lab based model systems.

Our study evaluates and quantifies the effects of key weather parameters (RH, SR, rainfall and temperature) on the survival of surrogate bacteria (generic *E. coli*) on the surface of oranges under natural conditions in a grove. This study also evaluates the effectiveness of copper hydroxide based pesticides (used for prevention of citrus canker) as a risk mitigating step against low microbial quality water sources used in the preparation of foliar sprays.

III.3 Materials And Methods

III.3.a Inoculation of spray water with cow manure

Well water wasg mixed with fresh cow manure (from a free range cattle ranch) to simulate contamination of oranges through application of low microbial quality spray water. A 10 g of representative sample was mixed with 90 ml of 0.1% peptone (Difco, Becton Dickinson & Co., Sparks, MD) water to determine how much manure to add to well water to produce a low (2 log CFU *E. coli*/ml) and a high inoculum level (4 log CFU *E. coli*/ml). Serial dilutions were made as needed with 0.1% peptone water and 100 µl were plated on tryptic soy agar (TSA; Difco, Becton Dickinson & Co., Sparks, MD) and Chromagar™ ECC (Microbiology, Paris, France). Plates were incubated at 37 °C for 24 h. Pink (coliforms) and green (*E. coli*) colonies were counted. The results were used to calculate the amount of manure needed, which was mixed with 38 liters of well water in a 13 gallon bin (Newell Rubbermaid Inc., Atlanta, GA). The manure/water mixture was filtered using a fiberglass screen (New York wire, Grand Island, NY) to remove grass and other debris prior to use. The slurry was further filtered through a cheese cloth (Thermo Fisher Scientific Inc., Waltham, MA) to prevent clogging into 15 gallon sprayer

equipment (Model:15SW102HLB1G0N 36 X 16. 25 inch 1.0 GPM, SMV industries, Council Bluffs, IA).

III.3.b Enumeration of *E. coli* and coliforms in spray water

Samples of low and high concentration inoculated spray water were collected and tested for *E. coli*, coliforms and total plate count by making appropriate dilutions in 0.1% peptone water and plating 100 µl on Chromagar™ ECC and tryptose soy agar. Plates were incubated at 37 °C for 24 h. Generic *E. coli* and coliform counts were determined by counting blue and pink colonies on Chromagar™ ECC respectively.

III.3.c Spraying orange trees in a grove

Orange trees used for this research were located in a grove at the Citrus Research and Education Center (CREC) in Lake Alfred, Florida. Three trees were selected for spraying and appropriately marked, while 3 other trees were marked as controls. A buffer zone of one row of orange trees was used to separate the test trees from the control trees to prevent contamination of the control group during spraying. The test trees were sprayed (ca. 200-250 l/acre) with manure inoculated well water until "drip off" (i.e. leaves and oranges were saturated with water). Control trees were sprayed with inoculated well water until drip off. Spraying was done during the harvest seasons for the Hamlin (September-February) or Valencia (March-June) orange varieties. This procedure was repeated for the spray water that was inoculated with low levels of *E. coli*. Twenty seven monthly field trials were conducted between 2012 and 2015. Personnel changes that occurred during the 2015 season necessitated the use of only a single inoculum level, as well as raising of the detection limit for some of the monthly trials.

III.3.d Application of copper hydroxide in spray water

Copper hydroxide (Kocide 3000, Wilmington, DE) was added to high and low inoculum water prior to spray application. Eight monthly field trials were conducted with copper hydroxide treated water between October 2014 to May 2015. The rate of 0.1% copper hydroxide used was equivalent to a typical field spray concentration for citrus prepared by adding 2.26 lb. of Kocide in 125 gallons of well water.

III.3.e Harvesting and enumeration of *E. coli* from oranges

Three replicates of 10 fruits each (a total of 30 fruit) were harvested from each tree, using sterile gloves between each replicate and placed in separate sterile plastic bags (Fisher scientific, Pittsburgh, PA 15275). The fruit was taken to the laboratory for processing. Each orange (from each replicate) was placed in a sterile whirl pack bag (Nasco, Fort Atkinson, WI, USA) with 100 ml of sterile 0.1% peptone water and massaged by hand for 1 min until all 10 oranges were processed. The stock solutions were diluted with 0.1% peptone water, plated onto TSA and ECC Chromagar. Plates were incubated at 37 °C for 24 h. Sampling was done at time 0, 2, 6, 24 h and at subsequent 24 h intervals. Coliforms and *E. coli* levels were determined by plate count method, but when counts below detection limit were attained, an MPN count was done using Colilert and QuantiTray®/2000 (IDEXX Laboratories Inc., Westbrook, Maine). Sampling was continued until when the detection limit (-1 log MPN/orange and 1 log CFU/orange in year 2015) was attained in 2 sequential samplings.

III.3.f Weather data

Solar radiation, rainfall, temperature and relative humidity were obtained from the Florida Automated Weather Network (FAWN) website (http://fawn.ifas.ufl.edu/).

III.3.g Statistical analysis

Non-linear curve fitting for *E. coli* survival curves was also done using the Microsoft® Excel (Microsoft, Redmond, CA) add-in GInaFit (*29*). The biphasic model consisting of two log linear phases (equation 2) and Weibull model (equation 3) were fitted to *E. coli* survival curves.

$$Log_{10}(N) = LogN_o - (t/\Delta)^p \quad \text{.....(2)}$$

Where delta (**Δ)** is a scale parameter and can be denoted as the time for the first decimal reduction, **p** is a shape parameter, final cell density (Log_{10} (N)) at time ***t***, and initial cell density (Log_{10} (N_0)).

$$Log_{10}(N) = LogN_o - Log_{10}(f.e^{-K_{max1}t} + (1-f).e^{-K_{max2}t}) \quad \text{.....(3)}$$

Where *f* is the fraction of the initial population (major subpopulation), (*1-f*) is the fraction of the initial population (minor subpopulation which is more heat resistant than the major population), and $kmax_1$ and $kmax_2$ (1/time unit) are the specific inactivation rates of the two populations, respectively. Linear regression analysis between the average monthly trial weather data and model parameters was done.

Linear regression, correlation coefficient and logistic regression analysis were done using Sigma® plot statistical software (Systat software, San Jose, CA) and Microsoft Excel (Redmond, WA). Classical log linear inactivation model was fitted on log change in *E. coli* populations/h and average change in rainfall, temperature, relative humidity, time and solar radiation data between sampling points. Log change in *E. coli* populations/h was determined by dividing the log change in *E. coli* populations by the time interval between sampling points. Logistic regression models are probabilistic and semi quantitative compared to linear regression models but serve as suitable alternatives in the absence of strong linear regression models. They show the most likely outcome through

probability of occurrence and significance (p values). Log *E. coli* change/h data were converted into binary data (0 or 1). For log reduction, a criterion of between 0.01 and 1.0 log *E. coli* change/h was set where observations above the criteria were coded 1 and observations below the criteria were coded 0. For log increase the criterion was set at 0.001 to 0.01 log *E. coli*/h. Logistic regression analysis was conducted with the binary data as the dependent variable and weather factor as the independent variable. The equation for the model was: -

Logit P = C + K * (Weather variable data)…..……... (4)

Where logit P is defined as the natural log (ln) of (P/1-P), C is a constant, K is the model coefficient and weather variables were SR and RH.

III.4 Results

Data on all the field trials are available as supplemental material, and presented in appendices VI.1 (field trial 2012), VI.2 (field trial 2013), VI.3 (field trial 2014) and VI.4 (field trial 2015) of this dissertation.

Field trial 2012 (see Appendix VI.1: Figures 5.1 to 5.4): In April and May, *E. coli* populations declined to detection limit at day 22 and day 15 respectively. There were no significant (> log 0.5) increases between sampling points. In June, there were significant increases in *E. coli* populations on days 2 (1 log) and 3 (0.8 log) respectively, but thereafter, populations declined to detection limit at day 23. In October, *E. coli* populations increased significantly on day 2 (0.7 log) and remained detectable up to day 22 for high inoculation levels; for low inoculation levels *E. coli* levels declined to detection limit on day 6; however there was an increase (log 0.8) on day 3. In November

and December, populations of *E. coli* became undetectable on days 12 and 3 respectively for high inoculation levels; at low inoculation levels, *E. coli* was undetectable at day 2 and 6 hours in November and December respectively. Overall (see Appendix VI.1: Figures 5.1 to 5.4) there were no visual correlations between hourly RH, rainfall, SR and temperature patterns with *E. coli* survival patterns except for the month of June and October (Appendix VI.1: Figure 5.4C and 5.4D). Increase in *E. coli* levels correlated (visually) well with increase in rainfall levels (< 10 days). Similarly increase in *E coli* levels was also correlated visually with low fluctuations in RH whereas high fluctuations in RH correlated well with decline in *E. coli* levels (Appendix VI.1:Figure 5.2C and 5.2D).

Field trial 2013 (see Appendix VI.2: Figure 6.1 to 6.4): In January, *E. coli* populations declined and were undetectable at day 7 and day 15 for both low and high inoculation levels respectively. In February, *E. coli* populations declined to detection limit at day 7 and day 10 for low and high *E. coli* populations respectively. In March, at low inoculum level, *E. coli* populations declined rapidly and were undetectable after 2 hours, whereas at high inoculation, *E. coli* was undetectable at day 1. In April, *E. coli* populations declined to detection limit at day 16 and day 10 for high and low inoculation levels respectively; there was a significant increase at day 13 for high inoculation level. In June, *E. coli* populations declined to detection limit at day 9 for high inoculation level whereas at low inoculation level sampling was discontinued at day 3; there were significant increases at day 3 (log 0.6) and day 1(1.4 log) for high and low inoculation levels respectively. In October, *E. coli* populations declined to detection limit at days 8 and 1 for high and low inoculation levels respectively. In November, *E. coli* populations

declined to detection limit on day 2; there was however an increase at 6 h by 1.7 log. For low inoculation levels, *E. coli* populations were at undetectable levels even at 2, 6, and 24 h. In December, *E. coli* populations declined to detection limit at day 14 for high inoculation level; there was an increase at day 3 by 1 log. For low inoculation levels, *E. coli* populations declined gradually to detection limit at day 7. Overall there were no visual correlations between *E. coli* survival patterns with hourly weather data except for April, June and November (see Appendix VI.2: Figures 6.4 D, 6.4E and 6.4G), where rainfall correlated well with increase in *E. coli* levels. Low fluctuations in RH also correlated visually well with increase in *E. coli* levels (Appendix VI.2: Figure 6.2E and 6.2G) whereas high fluctuations correlated (visually) well with rapid decline in *E. coli* levels (Appendix VI.2: Figure 6.2A, 6.2C).

Field trial 2014 (See Appendix VI.3: Figure 7.1 to 7.4): In January *E. coli* populations declined to detection limit at day 1 and day 8 for low and high inoculation levels respectively; at day 3 there was a significant (0.9 log) increase in *E. coli* populations for high inoculation level. In February, *E. coli* populations declined rapidly to detection limit at 6 h for both low and high inoculation. There was a slight but insignificant increase in *E. coli* populations at day 1 and 2 for high inoculation but the levels declined back to detection limit at day 3;populations remained undetectable even at day 3 for low inoculation. In March, *E. coli* populations declined post spraying but there were significant increases at day 2 by 3.4 logs and at day 14 for high inoculation; there was significant increase (0.9 logs) for low inoculation at day 2. *E. coli* populations declined to detection limit at day 3 and day 17 for low and high inoculation levels respectively. In April, *E. coli* populations declined rapidly to detection limit after 2 h for

low inoculation and day 3 for high inoculation. In May, *E. coli* populations declined rapidly to detection limit after 2 h and at day 3 for low and high inoculations respectively. Overall, survival patterns for *E. coli* didn't correlate visually well with hourly RH, rainfall, SR and temperature patterns except for the months of January (and then only slightly) and March, October, and December where increase in rainfall correlated visually well with rise in *E. coli* levels (Appendix VI.3; Figures 7.4A, 7.4C, 7.4F and 7.4H). As a result of an increase in rainfall, low relative humidity fluctuations correlated visually well with increase in *E. coli* populations in January, March, October and December (Appendix VI.3: Figures 7.2A, 7.2C, 7.2F and 7.2H). High fluctuations in RH correlated visually well with decline in *E. coli* populations particularly in the later stages of sampling for January, October and December (Appendix VI.3: Figures 7.2 A, 7.2F and 7.2H).

Field trial 2015 (see Appendix VI.4: Figure 8.1 to 8.4): In January, *E. coli* populations declined to detection limit (-1 log MPN/orange) at low inoculation level but after 20 h, there were increases at both high and low inoculation levels. Subsequently *E. coli* at both high and low levels then declined to the detection limit at 72 and 48 h respectively. In February, March and April, *E. coli* at high inoculation level declined to a revised detection limit (1 log CFU/orange) after 48 h; low inoculation levels for *E. coli* was not used. There was no significant increase in *E. coli* populations at below 48 h. In May, *E. coli* populations at high inoculation levels declined to 1 log CFU/orange after 24 h. In summary, rainfall (Appendix VI.4: Figure 8.4A) and RH (Appendix VI.4: Figure 8.2A) patterns for January correlated well with increase in *E. coli* populations; SR and temperature patterns didn't correlate visually well with *E. coli* survival patterns.

Figure 3.1 shows representative survival curves for indicator bacteria and hourly weather data patterns for 4 field trials.

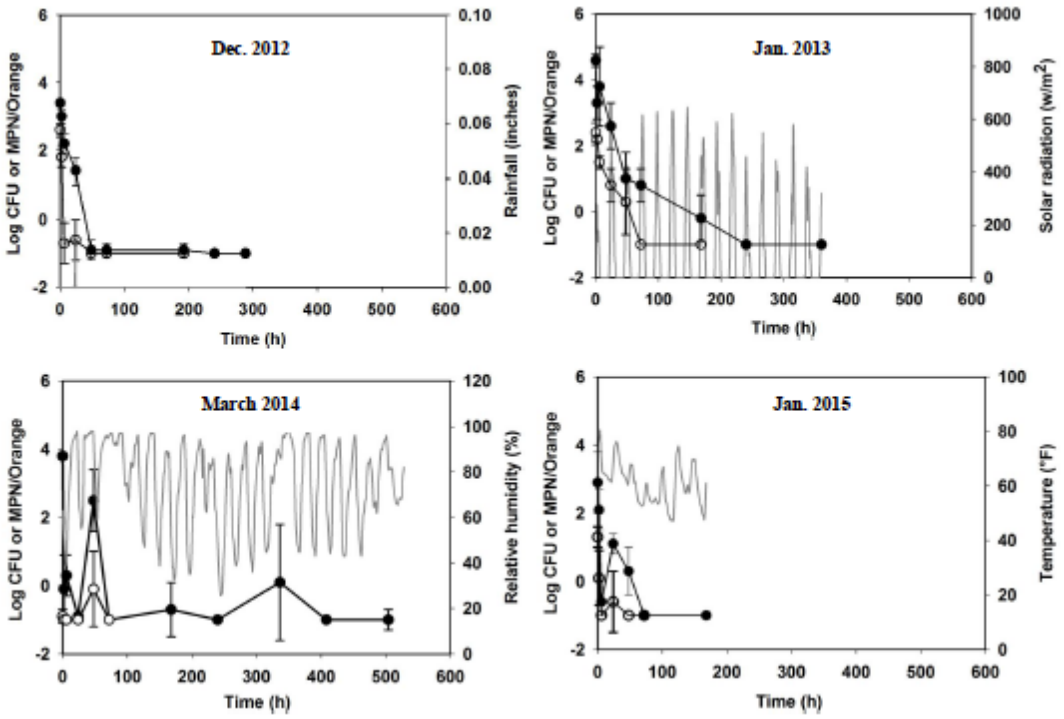

Figure 3.1: Representative plots for the influence of weather factors on *E. coli* population changes. High *E. coli* inoculation level (dark circle) and low *E. coli* inoculation (open circle), against rainfall (upper left), relative humidity (lower left), solar radiation (upper right) and temperature (lower right), where all weather factors are represented by grey lines

III.4.a Modeling the effect of weather on the survival of *E. coli* on oranges

The R^2 values for the Weibull (equation 2) and biphasic (equation 3) models fitted to the monthly trial data were between 0.5 and 0.99 indicating a high goodness of fit for the models (**Table III.a**). **Figure 3.2** shows an example of Weibull model fitting of April 2012 monthly trial data. When the **Λ** (time for first decimal reduction) and **p** (shape parameter) for the Weibull model and $\mathbf{kmax_1}$ (inactivation rate in initial phase), $\mathbf{kmax_2}$

(inactivation rate in final phase**)** and **f** (fraction of microbial population) parameters for the biphasic model were plotted against the monthly weather averages (SR, RH, rainfall and temperature) the R^2 values were low (< 0.3, data not shown). This demonstrated that the weather variables couldn't be used to predict parameters for both models and therefore both models could not be used quantify the effects of the weather

Figure 3.3 shows scatter plots of the SR, temperature, rainfall, RH and time (between sampling points) against log *E. coli* change/h at low and high *E. coli* inoculation levels. Linear regression analysis showed that weather and time (between sampling point) variables had weak linear relationships with log change in *E. coli* populations/h (R^2 < 0.3). A multiple linear regression analysis with the weather and time variables as the independent variables and log change in *E. coli* populations/h as the dependent variable also showed that a linear combination of the 5 variables couldn't be used to predict log change in *E. coli* populations/h (R^2 < 0.2); of the 5 variables, only solar radiation (SR) was significant ($p < 0.05$). **Table III. b** shows correlation coefficients and p values between SR, RH, rainfall, temperature, time (between sampling points) and log change in *E. coli* populations/h for combined low and high *E. coli* inoculation data. SR and temperature were significantly negatively correlated with log *E. coli.* However based on their p and r values SR (p = 7.39E-22, r = -0.55) was more significant (had a higher influence on log reduction) compared to temperature (p = 5.44E-07, r = -0.31). RH and time were significantly positively correlated with log change in *E. coli*/h. However based on p and r values, RH (p = 8.92E-09, r = 0.35) was more significant (had a higher influence on log increase) compared to time (p = 2.58E-02, r = 0.26).

The first time the level of *E. coli* fell below the limit of detection (-1log

MPN/orange or 1log MPN/orange) the log change between that time and the prior time was included in the analysis. The log change between the first and second occurrence of an *E. coli* level below the detection limit was not included in the analysis, since an accurate determination of the log change was not possible.

Logistic regression analysis was done using the two most influential variables on log change in *E. coli*/h. SR and RH were significant ($p < 0.05$) in predicting the likelihood of a log reduction in *E. coli* per hour (**Figure 3.4 and 3.5).** SR and RH were significant ($p > 0.05$) in predicting the likelihood of a log increase of ≥ 0.001 to 0.1 log *E. coli* per hour. Overall the likelihood of a log reduction in *E. coli* at low RH and high SR was higher compared to the likelihood of a log increase at low SR and high RH.

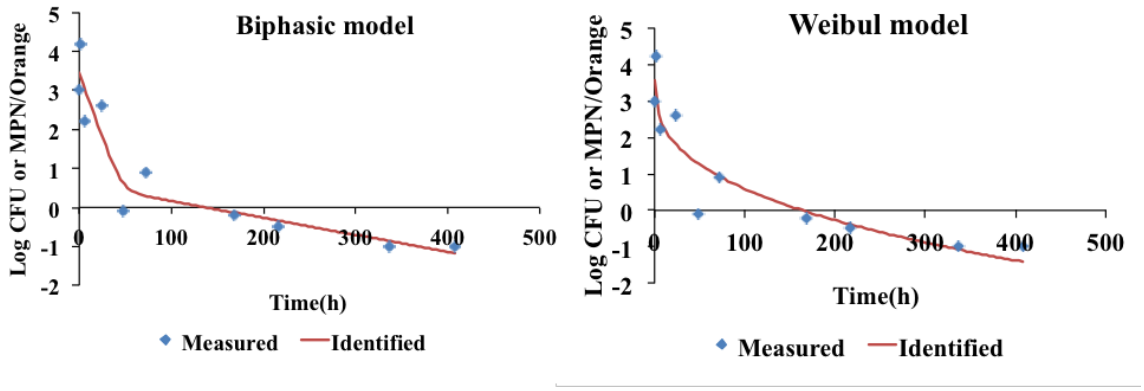

Figure 3.2: Biphasic and Weibull model fitting of December 2013 *E. coli* survival data

Table III.a: R^2values for Weibull and Biphasic model fitting of *E. coli* survival monthly trial data for high inoculation level

Month	**R^2 (Weibull Model)**	**R^2 (Biphasic model)**
Apr-12	0.93	0.33
May-12	0.98	0.66
Jun-12	0.75	0.74
Oct-12	0.94	0.94
Nov-12	0.86	0.97
Dec-12	0.84	0.99
Jan-13	0.96	0.92
Feb-13	0.92	0.78
Mar-13	0.79	0.97
Apr-13	0.74	0.97
Jun-13	0.46	0.96
Oct-13	0.90	0.93
Nov-13	0.63	0.78
Dec-13	0.84	0.90
Jan-14	0.58	No fit
Feb-14	0.51	0.99
Mar-14	0.58	No fit
Apr-14	0.45	No fit
May-14	0.45	0.99
Oct-14	0.92	No fit
Nov-14	0.86	0.99
Dec-14	0.77	0.87
Jan-15	0.71	0.82
Feb-15	0.86	0.98
Mar-15	0.86	0.98
Apr-15	0.60	No fit
May-15	0.76	No fit

No fit-Data points less than the minimum required for curve fitting

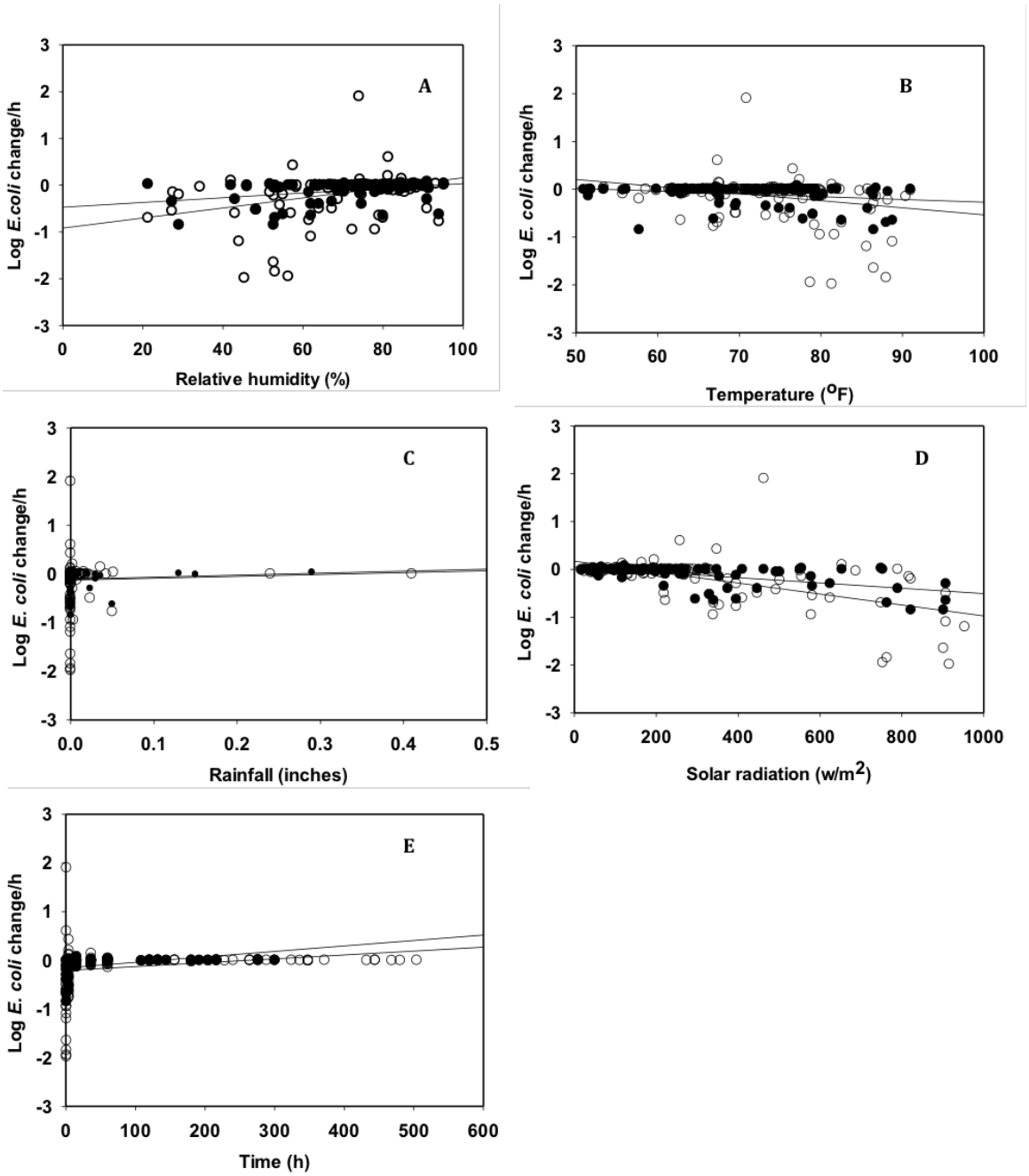

Figure 3.3: Linear regression analyses between log *E. coli* change/h at both high and low *E. coli* inoculation levels against solar radiation (D), temperature (B), relative humidity (A), rainfall (C) and sampling time (E). The slopes were close to zero**.**

Table III.b: Correlation coefficients (r) and p values (p) for log *E. coli* change/h against solar radiation (SR), temperature, relative humidity (RH), rainfall and sampling time for all *E. coli* inoculation levels**.**

		Temp (°F)	RH (%)	Rain (inches)	SR (W/m^2)	*E.coli* change h^{-1}
Time (h)	r	-0.15	0.137	-0.0585	-0.254	0.255
	p	**8.88E-03**	0.017	0.308	**6.78E-06**	**6.42E-06**
Temp (°F)	r		-0.236	0.0115	0.52	-0.31
	p		**3.14E-05**	0.842	1.64E-22	**3.16E-08**
RH (%)	r			0.154	-0.737	0.355
	p			**7.02E-03**	**1.75E-53**	**1.63E-10**
Rain (inches)	r				-0.0882	-0.0668
	p				0.124	0.245
SR (W/m^2)	r					-0.562
	p					**8.84E-27**

N = 256 observations, p values in bold were significant ($p < 0.05$)

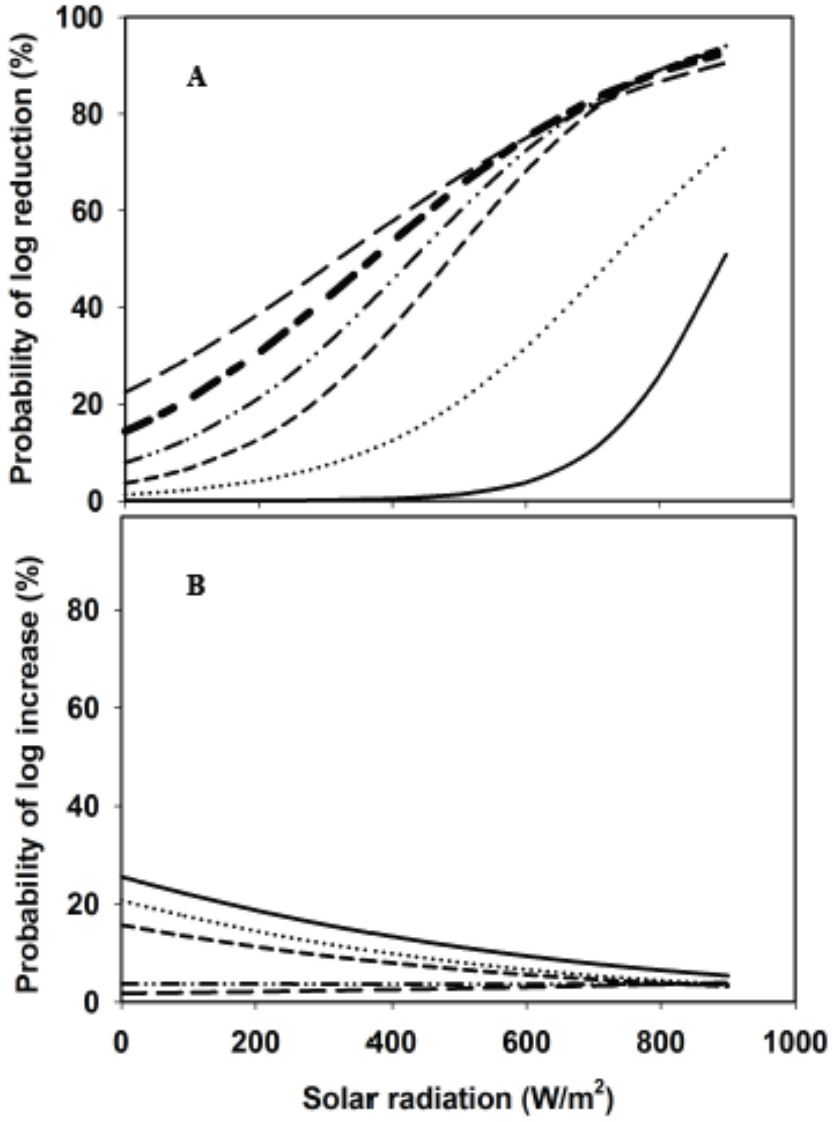

Figure 3.4: Logistic regression analysis between solar radiation (SR) and log change in *E. coli* populations/h. A shows the probabilities of log reduction/h ≥ 1.0 (black line), 0.5 (dotted line), 0.1 (short dashes), 0.05 (dash dot dot), 0.02 (bold long dash short dash), 0.01 (long dashes) at high SR levels. B shows the probability of log increase/h ≥ 0.001 (black line), 0.005 (dotted line), 0.01 (short dashes), 0.05 (dash dot dot) and 0.1 (long dashes) in *E. coli* at low SR levels. Log reductions were significant more likely ($p < 0.001$) to occur compared to log increase

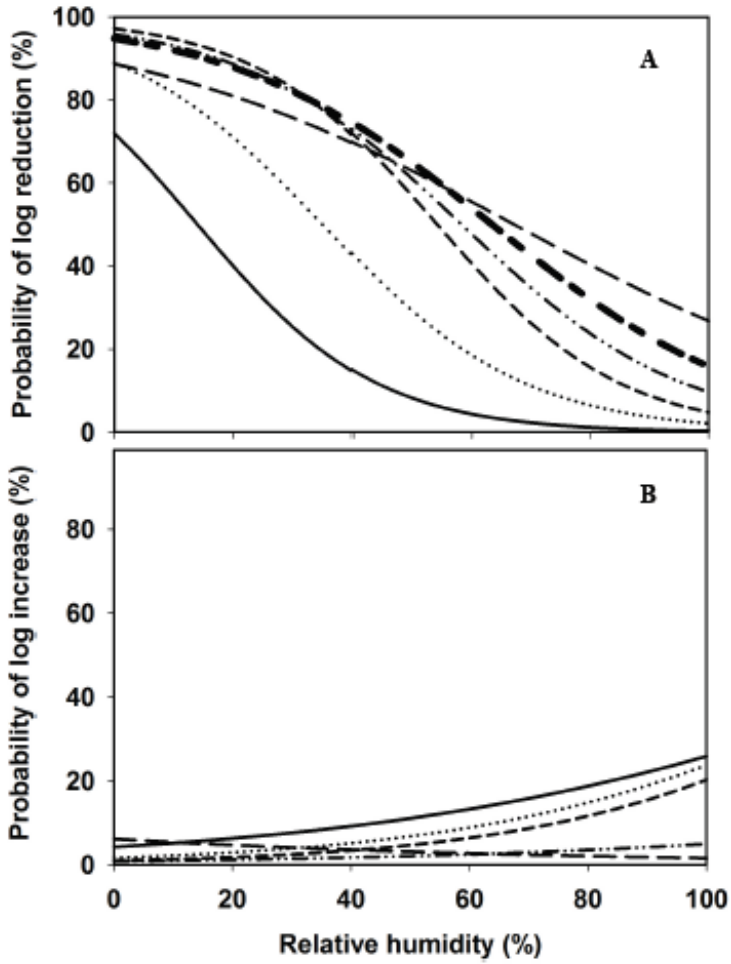

Figure 3.5: Logistic regression analysis between relative humidity (RH) and change in log *E. coli* populations/h. A shows the probability of log reduction/h ≥ 1.0 (dark line), 0.5 (dotted line), 0.1 (short dashes), 0.05 (dash dot dot), 0.02 (bold long dash short dash) and 0.01(log dashes) at low RH levels. B shows the probability of log increase/h ≥ 0.001 (dark line), 0.005 (dotted line), 0.01 (small dashes), 0.05 (dash dot dot) and 0.1(long dashes) in *E. coli* at high RH levels. Log reductions were significant ($p < 0.001$) and more likely to occur compared to log increase

III.4.b Quantifying the effect of copper hydroxide on *E. coli* concentration in spray water

To quantify the effect of the addition of 0.1% copper hydroxide in inoculated spray water, the average of the difference between concentrations of *E. coli* from inoculum sprays (with and without $CuOH_2$) at time zero for both high and low *E. coli* levels was calculated. The mean difference was found to be 2.6 ± 0.9 SD. This means that adding 0.1% $CuHO_2$ to spray water reduces the concentrations of indicator microorganisms and that of associated bacterial pathogens by ca. 2.6 log reductions.

III.5 Discussion

A visual inspection of the monthly trial data showed a few instances where rapid decline in *E. coli* populations appeared to occur co-incident with fluctuations in RH and absence of rainfall, which were characteristic of dry weather. We speculate that the high fluctuations in RH during dry weather may have stressed the bacteria resulting in rapid decline in *E. coli* levels, as has been reported in the literature (*5*, *8*, *78*). Conversely, increases in *E. coli* populations were sometimes seen to occur coincident with minimal fluctuations in RH and in times of rainfall characteristic of wet weather, which is also consistent with the published literature (*78*). Unfortunately these phenomena were not consistently observed in across all dry or wet months. In some occasions, *E. coli* populations declined even in the presence of high rainfall and high RH. Fluctuations in temperature and solar radiation patterns didn't also correlate with survival curves of *E. coli*.

Although the Weibull and biphasic models provided the best fit (compared to classical log linear model) for *E. coli* survival curves, the average monthly weather data

couldn't be used to predict the models parameters (delta, p, $kmax_1$ and $kmax_2$) owing to the low R^2 values. Consequently the models couldn't be used to predict the effects of weather on *E. coli* populations. Mckellar et al. (2014) evaluated the survival of *E. coli* O157:H7 on contaminated field lettuce using several data sets collected from field based experiments. The authors found that the pathogen decay pattern in most cases fitted the biphasic model although the Weibull model also showed a good fit. We found that the Weibull model fitted our data better than the biphasic model. Classical linear regression analysis between SR, temperature, rainfall, RH and sampling time and log *E. coli* change/h produced low R^2 (< 0.4) indicating weather variables couldn't adequately predict log *E. coli* change/h. Mckellar et al. (2014) similarly observed a lack of a good fit in the linear regression analysis of *E. coli* O157:H7 survival data on field lettuce.

Correlation coefficients (**Table III.b**) were useful in showing the direction and the significance of the relationship between independent and the dependent variables. The pairs of variables with positive correlation coefficient (r) and p values below 0.05 tend to increase together. For the pairs with negative correlation coefficients and p values below 0.05, one variable tends to decrease while the other increases. For pairs with p values greater than 0.05, there is no significant relationship between the two variables. Correlation coefficients close to 1 or -1 were indicative of strong linear relationships between variables. Our study has demonstrated that relative humidity (RH), solar radiation (SR), temperature and time had an influence on the survival of *E. coli* on the surface of oranges **(Table III.b)** consistent with the literature (*5, 8, 78*). SR and temperature were significantly negatively correlated to reduction in *E. coli* populations but since SR had a comparatively lower p value and a higher correlation coefficient (p =

7.39E-22, r = -0.6), we considered SR to be a more influential variable on log reduction. RH, rainfall and time were correlated with increase in *E. coli* populations, however only RH and time were significant. Based on p value and correlation coefficient (p = 8.92E-09, r = 0.35) we considered RH to have had the strongest effect on log increase in *E. coli* populations. Contrary to expectation, time was positively correlated to log increase in *E. coli* populations (**Table III.b**). This may have been occasioned by the rapid decline in *E. coli* populations after spraying as bacteria were acclimatizing to the new environment. Other interesting significant correlations were observed between temperature and relative humidity, temperature and solar radiation, temperature and rainfall, solar radiation and relative humidity and solar radiation and rainfall (**Table III.b**). The significance of these correlations was suggestive of a confounding effect between the variables. This is particularly important for lab based studies as it demonstrates the challenge of maintaining a variable constant e.g. increasing light intensity levels will raise temperature levels as well (*38*). Thus a considerable investment in equipment may be necessary to achieve constancy in the confounded variable.

Similar to our recent study (*53*) where linear models between biological and physicochemical parameters against microbial levels were found to have low R^2 values, logistic regression analysis was used in this study as an alternative to linear predictive modeling. **Figures 3.4 and 3.5** show probabilities of the specified log reductions and log increase in *E. coli*/h at high and low SR and at high and low RH levels. For instance during dry and sunny weather, characterized by high SR levels (ca.800 W/m^2) and low RH (ca. 30%) the models predicted an 80% and 85% chance of a ≥ 0.01 log reduction in *E. coli*/h respectively. During wet cloudy weather characterized by low SR levels (ca. 200

w/m^2) and high RH level (100%) the models predict a 20% and 25% chance for ≥ 0.001 log increase in *E. coli*/h respectively. Overall the likelihood and magnitude of log reductions were higher compared to log increase.

Stine et al. (2005) evaluated the influence of high (mean 85.7 to 90.3%) and low humidity (mean, 45.1 to 48.4%) on the survival of 8 viral and bacterial microorganisms on the surface of lettuce, cantaloupe and bell peppers in a controlled environment chamber. The authors found the survival of microorganisms to be dependent on differences in the surface of the crop and the microorganism's susceptibility to environmental stress. In cantaloupe, *Salmonella* and *E. coli* O157:H7 had lower inactivation rates under humid conditions while PRDI coliphage and FVC (feline calicivirus) survived longer under dry conditions. Similarly, our results showed that high RH had the potential to support microbial survival including microbial growth on the surface of oranges in a grove (**Table III.b and Figure 3.5**). Stine (2005) also evaluated the effect of light intensity (ca. 300 w/m^2) on microbial survival at 3 light conditions, full light exposure, shaded exposure and no direct exposure, on the survival of *E. coli* inoculated on petri plates. On full exposure 99.9% reduction occurred in less than 1 day while for shaded exposure and no exposure, 99.9% reduction occurred in 3 days. These results confirm our observations of a highly significant relationship between SR and log reduction in *E. coli* populations/h on the surface of oranges in the grove **(Table III.b and Figure 3.4**). Other laboratory studies have evaluated the effects of UV radiation (responsible for the germicidal effect of SR), temperature and RH on microbial survival on fruits and vegetables (*60*, *67*). Their results are in agreement with our correlation models.

For agricultural water that doesn't meet proposed microbial quality requirements (Statistical threshold value (STV) of 410 or geometric mean of 126 CFU/100 ml of generic *E. coli*), the FDA proposes a microbial die off rate of 0.5 log CFU/day on pre harvest crops as an alternative for achieving similar public health safety as the microbial water standards (*82*). Our logistic regression models have demonstrated that log reduction or log increase rates are not constant. The probability of either log reduction or a log increase is dependent on increasing or decreasing SR or RH levels. Our data does not agree with the FDA's published rule (0.5/day = 0.02/h) which assumes the probability of 0.02 log reduction/h rate occurs 100% independent of changes in SR or RH levels.

Narciso et al. (2012) demonstrated the effect of copper hydroxide spray used for citrus control on the survival of *E. coli* in broth and on grape fruit leaves. *E. coli* exposed to 0.1% copper hydroxide in vitro were eradicated from the broth within 6 to 8 h. No *E. coli* survived (detection limit of 1 CFU/ml) on leaf surfaces for both controls and leaves sprayed with copper beyond 48 h. In our study a 2.6 log reduction on surface of oranges in a grove was shown to occur at time 0 after application of 0.1% copper hydroxide spray (well water manure mixture). The 2.6 log reduction in *E. coli* that occurs as result of copper used in foliar sprays, coupled with predicted die off rates due to dry weather conditions, may allow large scale citrus growers to use surface water sources that do not meet the proposed produce water standards.

III.6 Conclusion

Many studies have documented the variety of pathways through which pre harvest fruits and vegetables can be contaminated with pathogens. They include contact with

contaminated manure, irrigation water, application of agrochemicals (prepared from contaminated water), feces from wild animals, reptiles and birds, dust, insect vectors, raw sewage, precipitation, surface runoffs, and climate (*6*, *39*, *40*). Our main objective for this study was to evaluate the change in microbiological risk posed by weather conditions on oranges in a grove contaminated through application of water of low microbial quality. Initial concentration of microorganisms on the surface of oranges in a grove after foliar spray application will be dependent on microorganism concentration in surface water. Thus high levels of enteric pathogens in surface water would result in high pathogen levels transferred to the surface of citrus fruits during foliar spray applications. This could pose a serious public health risk to consumers of fresh oranges and unpasteurized orange juice. Our study has shown that weather plays a significant role in the microbiological safety of fresh oranges. The role of dry weather (high SR and low RH) may be the reason behind the rare incidence of illness associated with the consumption of fresh oranges from central Florida.

Chapter IV: A Quantitative Microbial Risk Assessment On The Effect of Key Weather Factors and Time on the Survival of *E. coli* on Oranges Following Foliar Spray Applications Prepared From Low Microbial Quality Water

Gabriel Mootian[2], Loretta M. Friedrich[1], Timothy M. Spann[1] Michelle D. Danyluk[1] and Donald W. Schaffner[2*]

[1] Department of Food Science and Human Nutrition, Citrus Research and Education Center, Institute of Food and Agricultural Sciences, University of Florida, Lake Alfred, Florida, USA

[2] Department of Food Science, School of Environmental and Biological Sciences, Rutgers, New Brunswick, New Jersey, USA

Running title: Risk of Salmonellosis from Oranges in Central Florida

*Correspondent Footnote: Corresponding author electronic mail address: Schaffner@Aesop.rutgers.edu

To be submitted to: Journal of Food Protection

IV.1 Abstract

Enteric pathogens may be transferred to surfaces of fresh fruits and vegetables before harvest through the application of contaminated foliar sprays or irrigation water. This research estimates the likely levels of generic *E. coli* (as a *Salmonella* surrogate) on oranges by modeling the effects of weather and time on the survival of *E. coli*. Data from 27 monthly field trials (2012 to 2015) was used to create linear regression models for upper and lower boundary limits of log *E. coli* rate/h for solar radiation, relative humidity, temperature and time. An empirical model based on IF logic statements, Monte Carlo simulations and probability distributions were used to estimate log change in *E. coli*/h between the minimum and maximum boundary limits of the most restricting variable. The sum of predicted log reduction rates were applied to the initial concentration of *E. coli* to determine the concentration at subsequent sampling points. The model was compared to the experimental data, and FDA fresh produce rule assumption of a microbial decline rate of 0.5 log/day (for a maximum of 4 days) in *E. coli* on pre harvest crops. The predicted mean, 5^{th} and 95^{th} percentile of *E. coli* concentrations at days 1, 3, 7 and 14 sampling points showed the variable predictive quality of the model. The mean predicted values were close to the actual data for day 1 but under predicted for day 3, 7 and 14 for most of the monthly trials. The FDA rule over predicted actual data for day 1 and day 3 but under predicted actual data for day 7 and 14 for most of the monthly field trials. For a pre harvest hold time of not more than 4 days, the FDA model is a fail safe model compared to our model.

IV.2 Introduction

Risk factors associated with contamination of fresh fruits and vegetables have been well studied. Research has shown that pathways of contamination of fresh fruits and vegetables include contaminated agricultural water, soil, biological soil amendments and poor post harvest handling (*6, 61*). The use of low microbiological quality water for agricultural purposes may transfer pathogens onto the surface of pre harvest crops. Pesticides reconstituted with contaminated water and sprayed on to pre harvest fruits and vegetables have been linked to food borne outbreaks. (*35, 51, 59, 79, 86*). Animal and bird feces have also been reported as potential sources of contamination of pre harvest crops with enteric pathogens (*6, 45*). The survival and persistence of pathogens on the surface of pre harvest crops is dependent on the intrinsic characteristics of the pathogen, the type of crop itself and extrinsic ecological factors surrounding the crop (*8*).

Ottoson et al. (2011) conducted a quantitative microbial risk (QMRA) with *E. coli* O157:H7 on lettuce in a controlled climate chamber. The authors found *E. coli* O157:H7 sensitive to light intensity and temperature. The effect of irrigation water standards, hold times after last irrigation event and a 15 s rinse in cold water on the risk of infection were also evaluated. Holding harvest after last irrigation for 1, 2, 4 and 7 days reduced risk by 3, 8, 8 and 18 times respectively while a 15 s rinse reduced the risk six fold. A QMRA study by Hamilton et al. (2006) found that to achieve a risk of one illness or less in 10,000 exposures per year, vegetables irrigated with non disinfected secondary effluent containing enteric viruses had to be held for at least 14 days after the last irrigation event. Stine et al. (2011) conducted a microbial risk assessment for Hepatitis A virus (HAV) and *Salmonella* transferred to the surface of cantaloupe, iceberg lettuce and bell peppers

through spraying with diluted pesticide solutions. These authors found that where harvesting is done on the same day as pesticide application, concentrations of 1.59×10^{-3} CFU/ 100 ml *Salmonella* or 2.79×10^{-7} MPN/100 ml HAV in the water used for spray application, would result in 1:10,000 annual risk of infection to consumers of fresh produce.

Quantitative microbial risk assessment (QMRA) is a scientific process that involves i) hazard identification ii) hazard characterization iii) exposure assessment and iv) risk characterization (*26*). Quantitative risk assessment studies take a stochastic or probabilistic rather than a deterministic approach where variables are assigned distributions instead of single values. There are a very limited number of QMRA studies assessing the risk posed by weather factors on the safety of pre-harvest fruits and vegetables. Although the effects of ambient temperature, relative humidity and solar radiation on the survival of viral and bacterial pathogens on the surface of pre-harvest crops have been quantified, no relationship between risk of illness and these factors has been determined.

The aim of this manuscript is to model the survival of *E. coli* on fresh oranges contaminated through foliar sprays using a Monte Carlo simulation based approach. The probabilistic survival model uses weather variables (relative humidity (RH), solar radiation (SR) and temperature and time to predict *E. coli* survival. The eventual overall objective is to be able to predict the risk of salmonellosis from exposure to contaminated fresh oranges at given weather conditions and also predict the number of days before harvest that would be optimal to minimize the risk of salmonellosis.

IV.3 Methods

IV.3.a Overview of the model

The model outlined here incorporated the effect of the four significant ($p < 0.001$) variables (solar radiation (SR), relative humidity (RH), temperature and time) on *E. coli* survival. The model brings together the effects of the 4 variables using logic statements rather than the more conventional multivariate linear regression analysis. Note that a linear combination of the four variables (SR, RH, temperature and time) was not found to be statically significant ($p < 0.05$).

Our approach was an empirical model that defined the maximum and minimum boundary limits for log increase and log reduction rates for each variable as described below. We then selected the most constricting variable for log increase (minimum of the maximum from the 4 variables) and most constricting variable for log reduction (maximum of the minimum from the 4 variables). A uniform distribution was used to simulate between the selected maximum and minimum boundary limits (for the most restricting variable) for log change in *E. coli* rate/h using the @risk add-in (Palisade Software, Ithaca, NY) for Microsoft Excel (Redmond, WA).

IV.3.b Maximum and minimum boundaries for log *E. coli* change/h

Maximum log increase and minimum log reduction boundary limits for SR, RH, temperature and time were determined by sorting the data for each variable together with their corresponding log change in *E. coli*/h rates from smallest to the highest values. Each sorted data set was divided into portions of equal number of observations (Either n = 8, 16, 32, 64). Minimum and maximum log increase, minimum and maximum log reduction rates and averages weather variable values were determined for each portion of the data

set. The upper boundary limit (for each variable) was constructed by plotting a regression line using maximum log reduction rate values against the average weather variable values from each portion. The lower boundary limit was constructed by plotting regression lines using the minimum log reduction rate values and average weather variable value from each portion. The maximum and minimum boundaries for the time variable were determined using slightly different criteria from the other 3 variables. Most monthly field trial data showed that *E. coli* declined rapidly after spraying during the first 24 hours and thereafter declined at lower rates at later sampling intervals (Appendices VI.1 to V1.4). We split the log change/h in *E. coli* data into four different time intervals (0 to 15 h, 15 to 60 h, 60 to 108 h and 108 to 504 h) to accommodate this pattern. Maximum and minimum log increase and log reduction boundary limits for the time intervals were then determined in a similar manner as the other 3 variables.

IV.3.c Predicting log *E. coli* change/h from RH, SR, temp. and time variables

Regression lines that defined the highest maximum and minimum boundary limits for the 4 variables were used to predict maximum and minimum values for log change in *E. coli*/h. Six regression equations were plotted for the environmental variables (i.e. SR, RH and temperature for both greatest increase and greatest decline). Similarly, an additional eight regression equations were plotted to model the effect of time (0-15 h, 15-60 h, 60-108 h and 108-504 h for both greatest increase and greatest decline in log *E. coli*/h rate) (**Figures 4.1 to 4.4**). The actual weather data for each monthly trial was input into a spreadsheet template, and model predictions were generated for each hourly interval. The limits for both greatest increase and greatest decline for use in simulation modeling were selected using MIN, MAX functions and IF logic statements in Microsoft Excel. The @risk RiskUniform function was used to select the simulation result for that

hour from 10,000 Monte Carlo iterations. Predicted hourly log change in *E. coli* rates were summed over those time intervals where experimental observations were collected to determine the final concentration of *E. coli*. Final *E. coli* concentrations could also be determined for time intervals where no experimental observations were collected (**Table IV.a).**

Table IV.a: Overview of model, parameters and simulations

Cell	Variable	Value	Units	Source
D3	Solar radiation (SR)	a−	W/m2	User Input
D4	Relative humidity (RH)	a−	%	User Input
D5	Temperature	a−	°F	User Input
D6	Time	a−	h	User Input
D7	Concentration of *E. coli* on orange at spray time (t=0)	b−	Log CFU or MPN/orange	This study
D8	**Modeling for upper boundary limit (log increase)**			
D9	SR maximum	=6E-05*D3 + 0.0988	Log change/h	This study
D10	RH maximum	=0.0003*D4 + 0.0983	Log change/h	This study
D11	Temperature maximum	=-0.0028*D5 + 0.3232	Log change/h	This study
D12	Time maximum			
D13	0 to 15 h	=-0.0345*D6 + 0.603	Log change/h	This study
D14	15 to 60 h	=-0.0004*D6 + 0.1196	Log change/h	This study

D15	60 to 108 h	=-0.0019*D6+ 0.1955	Log change/h	This study
D16	108 to 504 h	=-5E-06*D6+ 0.0017	Log change/h	This study
D17	> 504 h	0	Log change/h	User input
D18	**Modeling the lower boundary limits (Log reduction)**			
D19	SR minimum	=-0.0025*D3- 0.3097	Log change/h	This study
D20	RH minimum	=0.0333*D4 - 3.6005	Log change/h	This study
D21	Temperature minimum	-0.0494*D5+ 2.5852	Log change/h	This study
D22	Time minimum			
D23	0 to 15 h	=0.1119*D6 - 1.7122	Log change/h	This study
D24	15 to 60 h	-0.0002*D6 - 0.1256	Log change/h	This study
D25	60 to 108 h	=0.0027*D6 - 0.3108	Log change/h	This study
D26	108 to 504h	=3E-05*D6 - 0.0199	Log change/h	This study
D27	> 504 h	0	Log change/h	User input
D28	Time (max)	=IF (D6<15,D13, IF (D6<60,D14, IF (D6<108,D15,IF(D6<504,D16,D17))))	Log change/h	Calculated
D29	Time (min)	=IF (D6<15,D23, IF (D6<60,D24, IF (D6<108,D25, IF (D6<504,D26, D27))))	Log change/h	Calculated
D30	Selected minimum of the maximums	=Min (D9, D10, D11, D28)	Log change/h	Calculated
D31	Selected maximum of the minimums	=Max (D19, D20, D21, D29)	Log change/h	Calculated
D32	Predicted log change in *E. coli*/h	=RiskUnifom (D30, D31)	Log change/h	Calculated
D33	Sampling interval (t)	c−	h	User input

D34	Predicted final concentration log change at time interval t	Sum of log change rates/h for t hours	log	Calculated
D35	Predicted final concentration on orange surface	=RiskOutput (D7-D34)	Log CFU or MPN/orange	Calculated

a- Data obtained from weather stations or forecasts
b- Concentrations obtained from irrigation water (assuming 100% transfer) or microbial analysis of fruit surface after spray application.
c- Sampling interval or pre harvest time interval.

IV.3.d Validation of model using field trial data

The predicted *E. coli* population at each sampling point was compared to actual data for each monthly trial to validate the model. The Monte Carlo simulated predicted 5^{th} percentile, mean and 95^{th} percentile of log *E. coli* populations levels from 10,000 iterations at each sampling point were plotted together with actual field trial data. The assumption used in FDA's fresh produce rule (0.5 log CFU reduction/day in *E. coli* levels for a maximum of 4 days) (*82*) for crops irrigated with low microbial quality water was also compared with field trial data and model simulation results. Since actual field trial data for day 4 after spraying were not available (no sampling was done on that day) day 3 results were used instead. Thus actual field data for day 1, 3, 7 and 14 were compared to predicted *E. coli* concentration data.

IV.4 Results

IV.4.a Upper and lower boundary limits for SR, RH, temperature and time models

Figures 4.1, 4.2, 4.3, 4.4 and 4.5 show the regression lines for the upper boundary limit (describing the most likely maximum log increase rate in *E. coli* populations) and the lower boundary limit (most likely maximum log reduction rate in *E. coli* populations) for the four variables**.** The results for the three environmental variables, SR, RH, and temperature (**Figure 4.1, 4.2 and 4.3**) show similar patterns. As solar radiation increased

(**Figure 4.1**), the potential for a greater log reductions increased, and for the greatest solar radiation value observed (ca. 900 W/m^2) the greatest log reduction was observed (ca. 2 log CFU/h), although at that same solar radiation value lower reductions were also observed. Solar radiation appeared to have had little effect on *E. coli* increases, although one very high increase occurred at an intermediate SR value (450 W/m^2).

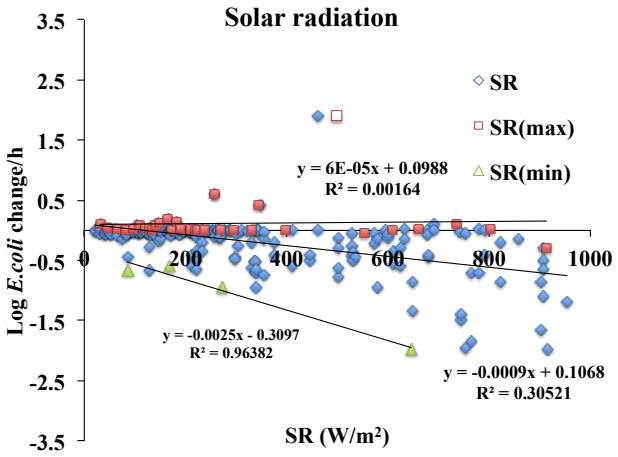

Figure 4.1: Upper (red squares) and lower (green triangles) boundaries for log change in *E. coli* rates for solar radiation data (blue diamonds)

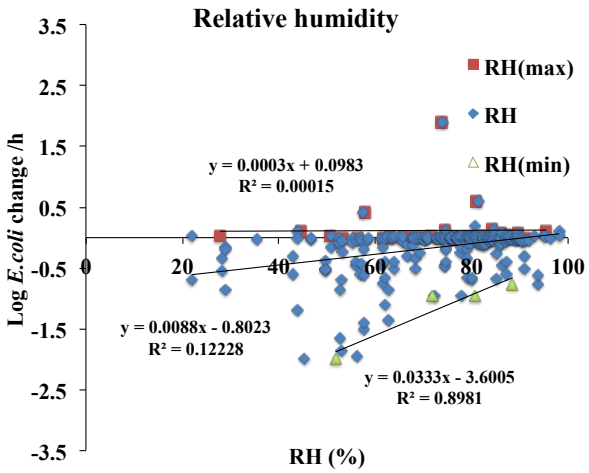

Figure 4.2: Upper (red squares) and lower (green triangles) boundaries for log change in *E. coli* rates for relative humidity data (blue diamonds)

The plot of relative humidity (RH) versus log *E. coli* change/h (**Figure 4.2**) shows a similar pattern, except that greater *E. coli* reductions were seen at lower RH values. The lowest RH values observed (ca. 20% RH) were not associated with the greatest log reductions, perhaps because either these low RH levels rarely occurred (thus there were limited observations). It could also be that the expected log reductions would be so great that the *E. coli* populations would be below the detection limit such that a reduction would not be observed. Under those most common RH conditions (40-100% RH) the relationship between a decreased RH and great possibility of a decline in *E. coli*/h was evident and vice versa.

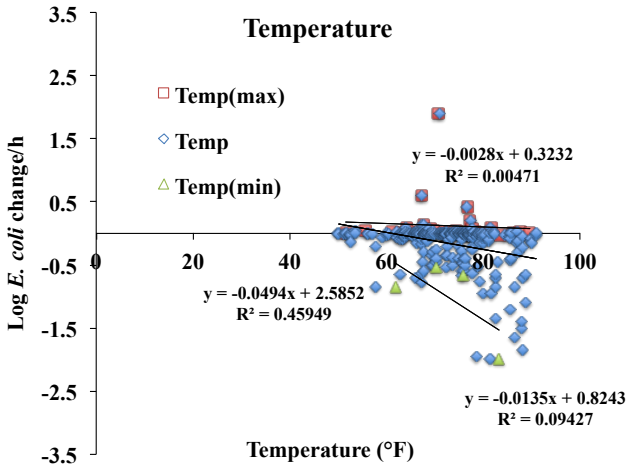

Figure 4.3: Upper (red squares) and lower (green triangles) boundaries for log change in *E. coli* rates for temperature data (blue diamonds)

The association between temperature and log *E. coli* change/h is shown in **Figure 4.3**. This figure shows increasing log reductions in *E. coli* as temperature increased from 60 °F to 90 °F. The highest log reductions (> log 1.5) occurred at temperature range 90 °F and above. There were lower reductions recorded at higher temperatures and this may

have resulted from reduced initial *E. coli* concentrations (at later stages of sampling) being exposed to rising temperatures. The model also showed log reduction at 50 °F. The lower temperatures (i.e. winter months) appear to have induced a bacteriostatic state (minimal growth or death). The moderate to optimal temperatures (60 °F to 80 °F) did not result in a consistent increase in *E. coli* concentrations, with most increases limited to ≤ 0.5 log; although one increase of ca. 2 log was observed at ca.70 °F. Since there were no data observed at temperatures ≤ 50 °F range, model predictions ≤ 50 °F were assumed to be no change (i.e. 0 log CFU/h *E. coli* change).

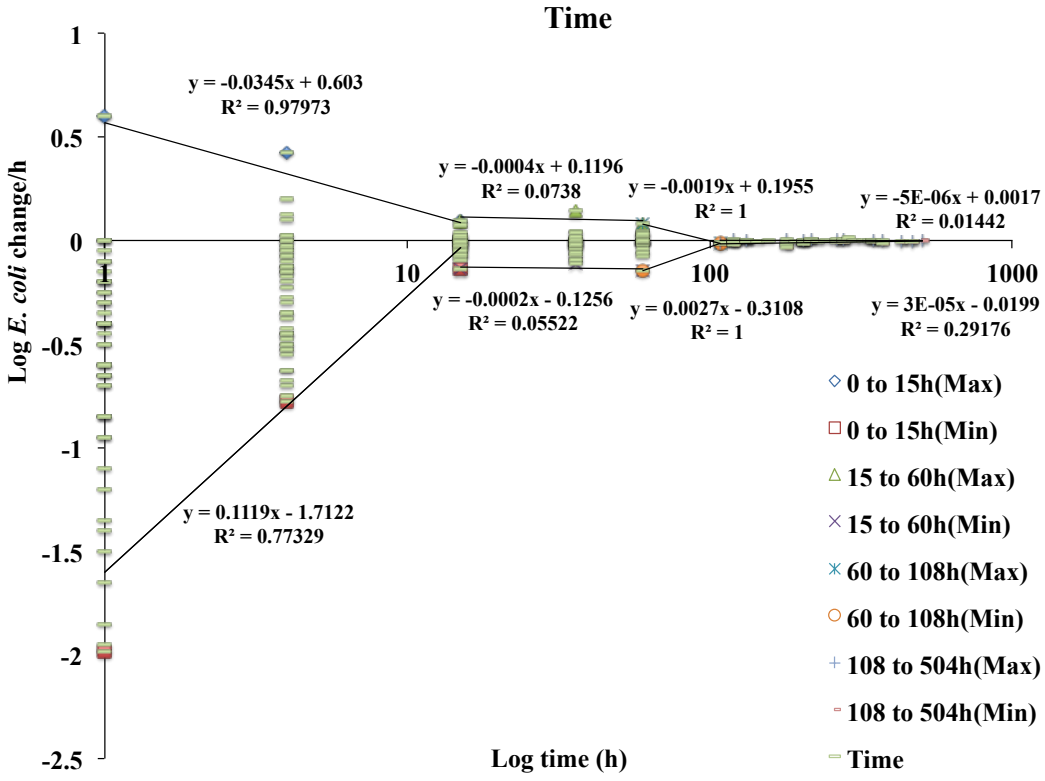

Figure 4.4: Log change in *E. coli* concentration vs. Log time (h) represented by green medium dash, divided into 4 time ranges (0 to 15 h, 15 to 60 h, 60 to 108 h and 108 to 504 h). Maximum boundaries for 0 to 15 h (blue diamond), 15 to 60 h (green triangle), 60

to 108h (blue star), 108 to 504 h (+) and minimum boundaries 0 to 15 h (red square), 15 to 60h (x), 60 to 108 h (orange circle), 108 to 504 h (pink small dash)

Figure 4.4 shows the relationship between log change in *E. coli*/ h with log time. It is quite clear from this figure that time since the application of the foliar spray had a profound effect on the potential increase or decrease in *E. coli* concentration. The greatest changes occurred in the first 24 h, much smaller changes occur in the next 2 days, and almost no changes occurred after 4 days.

IV.4.b Predicted 5th percentile, mean and 95th percentile vs. actual field trial data and FDA proposed microbial die off rate

The predicted 5th percentile, mean and 95th percentile, the actual monthly trial data and the predicted *E. coli* concentrations based on the proposed FDA rule of 0.5 log microbial die off rate/day (for a maximum of 4 days) were plotted for day 1 (**Figure 4.5**) and day 3 sampling times (**Figure 4.6**). Additional day 7 (**See Appendix VI.5, Figure 9.1)** and day 14 (See **Appendix VI.5, Figure 9.2**) were plotted to evaluate the performance of model and the FDA's rule past the recommended maximum 4 days of pre harvest holding. Since sampling times were different from one monthly trial to another the day 7 figure includes results from days 6-9 and day 14 figure includes results from days 12-16.

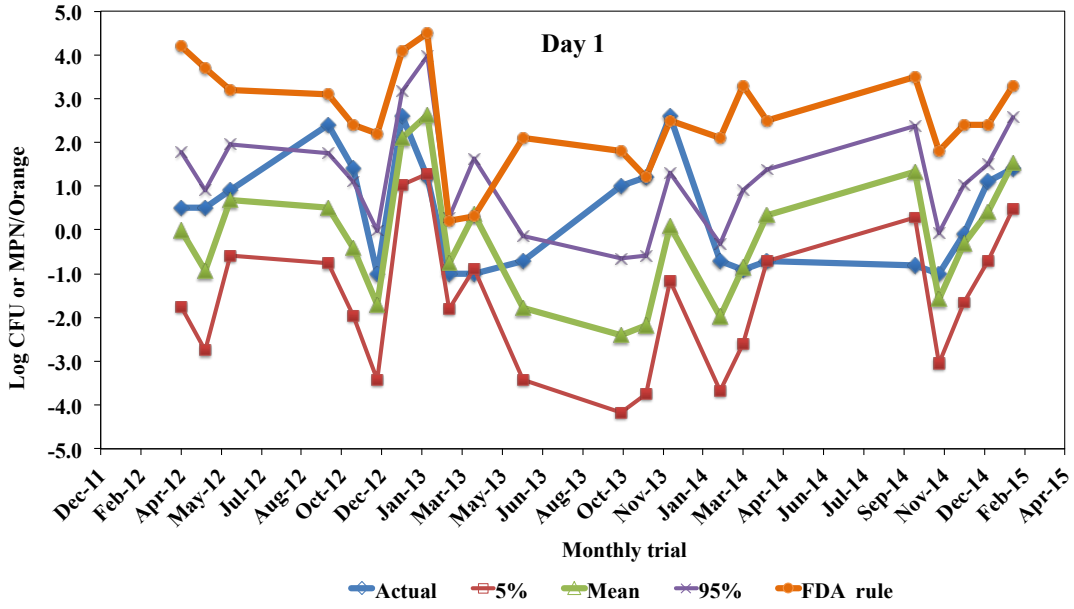

Figure 4.5: Predicted 5% (red line), mean (green line) and 95% (purple line) vs. actual (blue line) and FDA 0.5 log/day (orange line) for day 1 (24 h) post foliar spray

Day 1 results are shown in **Figure 4.5**, and it is immediately apparent our model under predicted actual data in almost every case whereas the FDA assumption rule over predicted the actual results. The model generally under predicted for the 2012 field trial data except in Dec. 2012. In 2013, the mean predicted *E. coli* concentrations for January and March 2013 were very close to the actual data but the model under predicted for the rest of that year. In 2014, the model over and under predicted the actual data except for November and December 2014 where the mean predicted data was close to the actual data. In 2015, the mean predicted value for February was close to the actual value. The predicted 95th percentile values were close to actual data for FT 2012, late 2014 and 2015. The FDA rule over predicted actual data except April, November and December

2013. April to May 2015 field trials were not evaluated as the detection limit was attained at either the 2 h or 6 h sampling times.

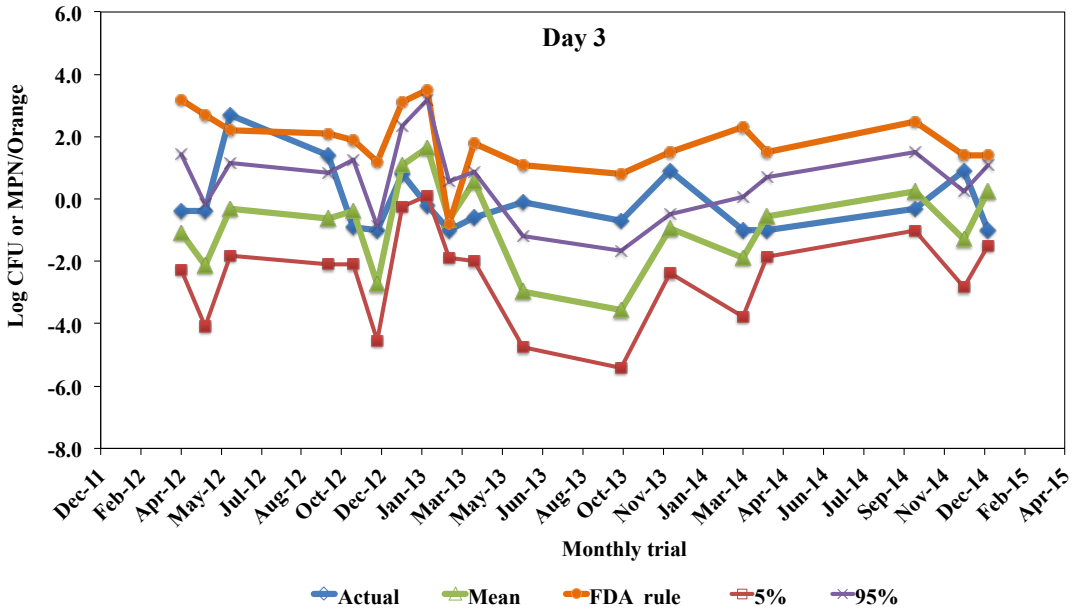

Figure 4.6: Predicted 5% (red line), mean (green line) and 95% (purple line) vs. actual (blue line) and FDA 0.5 log/day (orange line) for day 3 post foliar spray

Figure 4.6 shows FDA assumption, model predictions, and actual field trial data for *E. coli* populations for day 3 post foliar spray application. The model tended to under predict throughout 2012 except for November, where the mean predicted value were close to the actual monthly trial data. The FDA proposed rule predictions were much higher than actual values compared to our model. For field trial in 2013, 2014 and 2015, the model tended to over and under predict in most of the monthly trials except for Jan 2013, March 2013 and April 2014 where the mean predicted values were close to the actual data. The predicted 95th percentile was close to actual values for most of field trials between 2012 and 2015 i.e. less than 5% of the time the model predicted close to actual values. The FDA rule over predicted actual data from 2012 to 2015 but remained close to

the actual values (< log 0.5 difference) for June 2012, March 2013 and December 2014.

Figure 9.1 (See **Appendix VI.5, Figure 9.1**) shows FDA assumption, model predictions, and actual data for *E. coli* populations for day 7 ± 2 post foliar spray application. The model tended to under predict throughout 2012 except for November, where the mean predicted value was close to the actual monthly trial data. For the monthly trials before November 2012, only the predicted 95th percentile came close to the actual values i.e.ca. less than 5% of the time the model predicted close to actual values. The FDA proposed rule predictions were much closer to the actual values compared to our model and were very close in October and November 2012. For trials in 2013, 2014 and 2015, the model tended to under predict in most of the monthly trials except for Jan 2013 and October 2014, where the mean predicted values were close to the actual data. The FDA rule over predicted in April 2013 but remained close to the actual values for the rest of the field trials in 2014 and 2015.

Figure 9.2 (See **Appendix VI.5, Figure 9.2**) shows FDA assumption, model predictions, and actual data for *E. coli* populations for day 14 ± 2 post foliar spray. The model under predicts for all field trial data except January 2013; the predicted 95th percentile values were close to actual values for all field trials data. The FDA assumption under predicted actual data for all the field trials except for January and February 2013 where an over prediction occurred.

IV.5 Discussion

Our model combines 4 different variables (three for weather: relative humidity, solar radiation, and temperature) and one for time. Each of the four variables was found

to have a statistically significant effect on the survival of *E. coli* sprayed onto the surfaces of grove oranges. The challenges of modeling the combined effect of these variables can be partly attributed to the strong and significant correlations between them. For instance SR was significantly negatively correlated to RH with a high coefficient of correlation (r) and an extremely low p value (Table III.a). Studies that have looked at the survival of pathogens on surfaces of pre harvest crops have reported inactivation rates based on first order kinetics. Stine et al. (2011) estimated the inactivation rates of hepatitis A virus (HAV) and *Salmonella* on the surfaces of cantaloupe, lettuce, and bell peppers to be 0.01, 0.12, and 0.11 day^{-1}. Another publication from the same lab reported that the effects of high and low RH on the survival of 8 microorganisms on cantaloupe, lettuce, and bell pepper were variable (*78*). Petterson et al. (2001) reported estimated viral decay rates (*k*) to be 0.45 and 0.69 day^{-1}. These bacterial and viral decay estimates have been used in QMRAs for estimation of microbial standards for irrigation water and pre harvest hold times to achieve 1 illness in 10,000 exposures (*37*, *79*). In all of these studies the influence of environmental factors (temperature, light intensity and relative humidity) though reported to be significant were not quantified or modeled.

As demonstrated in chapter III of this dissertation, the Biphasic and Weibull non-linear regression modeling of *E. coli* survival data produced the best fits compared to linear regression modeling. Our survival curve fitting results were consistent with those of Mckellar et al. (2014) though we observed a better fit by Weibull than Biphasic models. Unfortunately as we noted, model parameters (kmax, delta and p) did not correlate well with monthly averages of rainfall, RH, SR and temperature data.

To address these challenges, we designed an empirical model, which incorporated

the log change in *E. coli*/h data associated with the 4 variables (RH, time, SR and temperature). The major assumption for this model was that at every hour, one of the 4 variables would be the most constricting variable in influencing log change in *E. coli* populations. The model predicts the data well 24 h after spraying but under predicts at day 3. The model also under predicts actual data at longer time intervals post spray (7 or 14 days). The FDA assumption from the fresh produce rule (0.5 log reduction/day for a maximum of 4 days) on pre harvest crops, over predicts the actual data on day 1 and day 3, predicts close to the actual data at day 7 and under predicts at day 14 (See Appendix V1.5). In summary, neither FDA produce rule assumption nor our model (which incorporates weather factors) may accurately predict the final concentrations of pathogens on the surface of pre harvest at different pre harvest times. However since it was observed that our model under predicted risk compared to the FDA assumption rule (which over predicted risk for most of the field trial data), the rule may result in stricter microbial standards for agricultural water. Our analysis also supports the FDA's recommendation of a pre harvest hold of not more than 4 days because the rule begins to under predict risk from day 7 and beyond, while our empirical model is effective for pre harvest hold times of 1 day and below.

IV.6 Conclusions

The development of an exposure assessment model for predicting the likely pathogen exposure levels at given weather conditions is highly complex. Microorganisms on the surface of pre harvest crops are subject to many interacting factors including fluctuating weather conditions, nutrient availability and type, predation or competition

from resident epiphytic microorganisms, biofilm formation, DNA degradation, osmotic, oxidative and desiccation stress (*8*). To model the individual effect or the effects of a combination these factors on the survival of pathogens on pre harvest crop surfaces is complex, as interactions may be antagonistic or synergistic. Our study supports the existence of strong interactions between weather factors. This compounds the design and cost of experiments required to model the effect of each variable while holding the rest constant. The change in concentration of microorganisms on the surface of crops may also be linear or non-linear depending on the environmental factors that are at play. The inadequacy and inconsistencies in our model and also that proposed by the FDA on microbial die off rates attest to the reality of insufficient knowledge on the science of survival of pathogens under field conditions. However from our study, we can conclude that a fail safe model based on a 0.5 log reduction in generic *E. coli*/day for a maximum of 4 days will reduce the risk of exposure to pathogens by limiting the usage of contaminated sources of agricultural water.

V. Bibliography

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

2. American Public Health Association. 1992. Standard methods of examining water and wastewater, 18th ed. American Public Health Association Washington DC.

3. Bae, S., and S. Wuertz. 2012. Survival of host-associated bacteroidales cells and their relationship with Enterococcus spp., *Campylobacter jejuni, Salmonella enterica serovar* Typhimurium, and adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Appl. Environ. Microbiol.* 78:922-932.

4. Berney, M., H. U. Weilenmann, A. Simonetti, and T. Egli. 2006. Efficacy of solar disinfection of *Escherichia coli, Shigella flexneri, Salmonella* Typhimurium and *Vibrio cholerae*. *Journal. Of. Applied. Microbiology.* 101:828-836.

5. Beuchat, L. R. 2002. Ecological factors influencing survival and growth of human pathogens on raw fruits and vegetables. *Microbes. And. Infection.* 4:413-423.

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

7. Beuchat, L. R., and C. A. Hwang. 1996. Comparison of dilution schemes for enumerating fungi in ground pepper: a collaborative study. *International. Journal. Of. Food. Microbiology.* 29:157-160.

8. Brandl, M. T. 2006. Fitness of human enteric pathogens on plants and implications for food safety. *Annu. Rev. Phytopathol.* 44:367-392.

9. Buchanan, R. L. 2000. The development of science-based food safety regulations in the United States. *Irish. Journal. Of. Agricultural. And. Food. Research.* 39:331-342.

10. Burton, G. A., D. Gunnison, and G. R. Lanza. 1987. Survival of pathogenic bacteria in various freshwater sediments. *Appl. Environ. Microbiol.* 53:633-638.

11. California Leafy Greens Marketing Agreement. 2012. http://www.lgma.ca.gov/wp-content/uploads/2014/09/California-LGMA-metrics-08-26-13-Final.pdf.

12. Carter, A. M., R. E. Pacha, G. W. Clark, and E. A. Williams. 1987. Seasonal occurrence of Campylobacter spp. in surface waters and their correlation with standard indicator bacteria. *Appl. Environ. Microbiol.* 53:523-526.
13. CDC. Foodborne Outbreak Online Database (FOOD Tool). http://wwwn.cdc.gov/foodborneoutbreaks/.

14. Chandran, A., A. A. Mohamed Hatha, and . 2005. Relative survival of Escherichia coli and *Salmonella* typhimurium in a tropical estuary. *In. Water. Research.* 39:1397-1403.

15. Combase. Combase in a nutshell. http://www.combase.cc/index.php/en/about-combase.

16. Danyluk, M. D., and D. W. Schaffner. 2011. Quantitative assessment of the microbial risk of leafy greens from farm to consumption: preliminary framework, data, and risk estimates. *J. Food. Prot.* 74:700-708.

17. Dechesne, M., and E. Soyeux. 2007. Assessment of source water pathogen contamination. *J. Water. Health.* 5 Suppl 1:39-50.

18. Doyle, M. P., and M. C. Erickson. 2008. Summer meeting 2007, the problems with fresh produce: an overview. *Journal. Of. Applied. Microbiology.* 105:317-330.

19. Economou, V., P. Gousia, A. Kansouzidou, H. Sakkas, P. Karanis, and C. Papadopoulou. 2012. Prevalence, antimicrobial resistance and relation to indicator and pathogenic microorganisms of *Salmonella enterica* isolated from surface waters within an agricultural landscape. *Int. J. Hyg. Environ. Health.* 216:435-444

20. Erickson, M. C. 2010. Microbial risks associated with cabbage, carrots, celery, onions, and deli salads made with these produce items. *Comprehensive. Reviews. In. Food. Science. And. Food. Safety.* 9:602-619.

21. Feng, F., D. Goto, and T. Yan. 2010. Effects of autochthonous microbial community on the die-off of fecal indicators in tropical beach sand. *FEMS. Microbiol. Ecol.* 74:214-225.

22. 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.

23. Food and Agriculture Organization and World Health Organization. 2008. Exposure assessment of microbiological hazards in food - Guidelines. http://www.fao.org/docrep/010/a0251e/a0251e00.HTM.

24. Food and Agriculture Organization and World Health Organization. 2007. Food safety risk analysis - A guide for national food safety authorities. http://www.fao.org/docrep/012/a0822e/a0822e00.htm.

25. Food and Agriculture Organization and World Health Organization. 2003. Hazard characterization for pathogens in food and water guidelines. http://www.fao.org/docrep/006/y4666e/y4666e00.htm.

26. Food and Agriculture Organization and World Health Organization. 2009. Risk characterization of microbiological hazards in food. http://www.fao.org/docrep/012/i1134e/i1134e00.htm.

27. Florida Department of Agriculture and Consumer services. 2007. Tomato best practices manual. http://www.freshfromflorida.com/Divisions-Offices/Food-Safety.

28. Francy, D. S., E. A. Stelzer, J. W. Duris, A. M. Brady, J. H. Harrison, H. E. Johnson, and M. W. Ware. 2013. Predictive Models for *Escherichia coli* Concentrations at Inland Lake Beaches and Relationship of Model Variables to Pathogen Detection. *Appl. Environ. Microbiol.* 79:1676-1688.

29. Geeraerd, A. H., V. P. Valdramidis, and J. F. Van Impe. 2005. GInaFiT, a freeware tool to assess non-log-linear microbial survivor curves. *Int. J. Food. Microbiol.* 102:95-105.

30. Gibson, K. E., and K. J. Schwab. 2011. Detection of bacterial indicators and human and bovine enteric viruses in surface water and groundwater sources potentially impacted by animal and human wastes in lower Yakima Valley, Washington. *Appl. Environ. Microbiol.* 77:355-362.

31. Gonzalez, R. A., K. E. Conn, J. R. Crosswell, and R. T. Noble. 2012. Application of empirical predictive modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters. *Water. Res.* 46:5871-5882.

32. 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

33. Goyal, S. M., C. P. Gerba, and J. L. Melnick. 1977. Occurrence and distribution of bacterial indicators and pathogens in canal communities along the Texas coast. *Appl. Environ. Microbiol.* 34:139-149.

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

35. Guan, T. Y., G. Blank, A. Ismond, and R. Van Acker. 2001. Fate of foodborne bacterial pathogens in pesticide products. *Journal. Of. The. Science. Of. Food. And. Agriculture.* 81:503-512.

36. Haley, B. J., D. J. Cole, and E. K. Lipp. 2009. Distribution, diversity, and seasonality of waterborne *Salmonella* in a rural watershed. *Appl. Environ. Microbiol.* 75:1248-1255.

37. 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.

38. Harris, L. J., J. Bender, E. A. Bihn, T. Blessington, M. D. Danyluk, P. Delaquis, L. Goodridge, A. M. Ibekwe, S. Ilic, K. Kniel, J. T. LeJeune, D. W. Schaffner, D. Stoeckel, and T. V. Suslow. 2012. A framework for developing research protocols for evaluation of microbial hazards and controls during production that pertain to the quality of agricultural water contacting fresh produce that may be consumed raw. *Journal. Of. Food. Protection.* 75:2251-2273.

39. Harris, L. J., J. N. Farber, L. R. Beuchat, M. E. Parish, T. V. Suslow, E. H. Garrett, and F. F. Busta. 2003. Outbreaks associated with fresh produce: incidence, growth, and

survival of pathogens in fresh and fresh-cut produce. *Comprehensive. Reviews. In. Food. Science. And. Food. Safety.* 2:78-141.

40. Heaton, J. C., and K. Jones. 2008. Microbial contamination of fruit and vegetables and the behaviour of enteropathogens in the phyllosphere: a review. *J. Appl. Microbiol.* 104:613-626.

41. Herwaldt, B. L., and M. J. Beach. 1999. The return of *Cyclospora* in 1997: another outbreak of cyclosporiasis in North America associated with imported raspberries. *Cyclospora* Working Group. *Ann. Intern. Med.* 130:210-220.

42. Horman, A., R. Rimhanen-Finne, L. Maunula, C-H. Von Bonsdorff, N. Torvela, A. Heikinheimo, and M-L. Hanninen. 2004. *Campylobacter spp., Giardia spp., Cryptosporidium* spp., Noroviruses, and indicator organisms in surface water in Southwestern Finland, 2000-2001. *Appl. Environ. Microbiol.* 70:87-95.

43. Ijabadeniyi, O. A., L. K. Debusho, M. Vanderlinde, and E. M. Buys. 2011. Irrigation water as a potential preharvest source of bacterial contamination of vegetables. *Journal. Of. Food. Safety.* 31:452-461.

44. Ishii, S., and M. J. Sadowsky. 2008. *Escherichia coli* in the environment: implications for water quality and human health. *Microbes. Environ.* 23:101-108.

45. Islam, M., J. Morgan, M. P. Doyle, S. C. Phatak, P. Millner, and X. Jiang. 2004. Fate of *Salmonella enterica* serovar Typhimurium on carrots and radishes grown in fields treated with contaminated manure composts or irrigation water. *Appl. Environ. Microbiol.* 70:2497-2502.

46. Klontz, K. C., J. C. Klontz, R. K. Mody, and R. M. Hoekstra. 2010. Analysis of tomato and jalapeno and Serrano pepper imports into the United States from Mexico before and during a national outbreak of *Salmonella* serotype Saintpaul Infections in 2008. *J. Food. Prot.* 73:1967-1974.

47. Lapidot, A., and S. Yaron. 2009. Transfer of *Salmonella enterica* serovar Typhimurium from contaminated irrigation water to parsley is dependent on curli and cellulose, the biofilm matrix components. *J. Food. Prot.* 72:618-623.

48. Leclerc, H., D. A. Mossel, S. C. Edberg, and C. B. Struijk. 2001. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annu. Rev. Microbiol.* 55:201-234.

49. Levantesi, C., L. Bonadonna, R. Briancesco, E. Grohmann, S. Toze, and V. Tandoi. 2012. *Salmonella* in surface and drinking water: Occurrence and water-mediated transmission. *Food. Research. International.* 45:587-602.

50. Lodder, W. J., and A. M. D. Husman. 2005. Presence of noroviruses and other enteric viruses in sewage and surface waters in The Netherlands. *Appl. Environ. Microbiol.* 71:1453-1461.

51. Lopez-Velasco, G., A. Tomas-Callejas, D. Diribsa, P. Wei, and T. V. Suslow. 2013. Growth of *Salmonella enterica* in foliar pesticide solutions and its survival during field production and postharvest handling of fresh market tomato. *J. Appl. Microbiol.* 114:1547-1558.

52. Malorny, B., J. Hoorfar, M. Hugas, A. Heuvelink, P. Fach, L. Ellerbroek, C. Bunge, C. Dorn, and R. Helmuth. 2003. Interlaboratory diagnostic accuracy of a *Salmonella* specific PCR-based method. *International. Journal. Of. Food. Microbiology.* 89:241-249.

53. McEgan, R., G. Mootian, L. D. Goodridge, D. W. Schaffner, and M. D. Danyluk. 2013. Predicting *Salmonella* populations from biological, chemical, and physical indicators in Florida surface waters. *Appl. Environ. Microbiol.* 79: 4094–4105

54. McKellar, R. C., F. Peréz-Rodríguez, L. J. Harris, A-L. Moyne, B. Blais, E. Topp, G. Bezanson, S. Bach, and P. Delaquis. 2014. Evaluation of different approaches for modelling *Escherichia coli* O157: H7 survival on field lettuce. *International. Journal. Of. Food. Microbiology.* 184: 74-85.

55. Miles, J. M., S. S. Sumner, R. R. Boyer, R. C. Williams, J. G. Latimer, and J. M. McKinney. 2009. Internalization of *Salmonella enterica* serovar Montevideo into greenhouse tomato plants through contaminated irrigation water or seed stock. *J. Food. Prot.* 72:849-852.

56. Mootian, G., W. H. Wu, and K. R. Matthews. 2009. Transfer of *Escherichia coli* O157:H7 from soil, water, and manure contaminated with low numbers of the pathogen to lettuce plants. *J. Food. Prot.* 72:2308-2312.

57. Mossel, D. A. A., J. E. L. Corry, C. B. Struijk, and R. M. Baird. 1995. Essentials of the microbiology of foods: a textbook for advanced studies. John Wiley & Sons, New York, NY.

58. Narciso, J. A., C. M. Ference, M. A. Ritenour, and W. W. Widmer. 2012. Effect of copper hydroxide sprays for citrus canker control on wild-type *Escherichia coli*. *Lett. Appl. Microbiol.* 54:108-111.

59. Ng, P. J., G. H. Fleet, and G. M. Heard. 2005. Pesticides as a source of microbial contamination of salad vegetables. *Int. J. Food. Microbiol.* 101:237-250.

60. Ottoson, J. R., K. Nyberg, R. Lindqvist, and A. Albihn. 2011. Quantitative microbial risk assessment for Escherichia coli O157 on lettuce, based on survival data from controlled studies in a climate chamber. *J. Food. Prot.* 74:2000-2007.

61. 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.

62. Patchanee, P., B. Molla, N. White, D. E. Line, and W. A. Gebreyes. 2010. Tracking *Salmonella* contamination in various watersheds and phenotypic and genotypic diversity. *Foodborne. Pathog. Dis.* 7:1113-1120.

63. Pathogen Modeling Program. http://ars.usda.gov/Services/docs.htm?docid=6786.
64. Payment, P., and A. Locas. 2011. Pathogens in water: value and limits of correlation with microbial indicators. *Ground. Water.* 49:4-11.

65. Petterson, S. R., and N. J. Ashbolt. 2001. Viral risks associated with wastewater reuse: modeling virus persistence on wastewater irrigated salad crops. *Water. Science. And. Technology.* 43:23-26.

66. Rajabi, M., M. Jones, M. Hubbard, G. Rodrick, and A. C. Wright. 2011. Distribution and genetic diversity of *Salmonella enterica* in the Upper Suwannee River. *International. Journal. Of. Microbiology.* 2011:
67. Rathinasabapathi, B. 2004. Survival of *Salmonella* Montevideo on tomato leaves and mature green tomatoes. *J. Food. Prot.* 67:2277-2279.

68. Reina, L. D., H. P. Fleming, and F. Breidt. 2002. Bacterial contamination of cucumber fruit through adhesion. *J. Food. Prot.* 65:1881-1887.

69. Rhodes, M. W., and H. Kator. 1988. Survival of *Escherichia coli* and *Salmonella* spp. in estuarine environments. *Appl. Environ. Microbiol.* 54:2902-2907.

70. Ritter, L., K. Solomon, P. Sibley, K. Hall, P. Keen, G. Mattu, and B. Linton. 2002. Sources, pathways, and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton inquiry. *J. Toxicol. Environ. Health. A.* 65:1-142.

71. Santiago-Rodriguez, T. M., R. L. Tremblay, C. Toledo-Hernandez, J. E. Gonzalez-Nieves, H. Ryu, J. W. Santo Domingo, and G. A. Toranzos. 2012. Microbial quality of tropical inland waters and effects of rainfall events. *Appl. Environ. Microbiol.* 78:5160-5169.

72. Schaffner, D. W., and S. Smith. 2004. Indicator organisms. *Encyclopedia. Of. Meat. Sciences. Elsevier. Amsterdam.*773-779.

73. Schets, F. M., J. H. van Wijnen, J. F. Schijven, H. Schoon, and A. M. de Roda Husman. 2008. Monitoring of waterborne pathogens in surface waters in amsterdam, The Netherlands, and the potential health risk associated with exposure to cryptosporidium and giardia in these waters. *Appl. Environ. Microbiol.* 74:2069-2078.

74. Setti, I., A. Rodriguez-Castro, M. P. Pata, C. Cadarso-Suarez, B. Yacoubi, L. Bensmael, A. Moukrim, and J. Martinez-Urtaza. 2009. Characteristics and dynamics of *Salmonella* contamination along the coast of Agadir, Morocco. *Appl. Environ. Microbiol.* 75:7700-7709.

75. Smith, R. M. M., F. Drobniewski, A. Gibson, J. D. E. Montague, M. N. Logan, D. Hunt, G. Hewinson, R. L. Salmon, and B. O'Neill. 2004. *Mycobacterium bovis* Infection, United Kingdom. *Emerging. Infectious. Diseases.* 10:539-541.

76. Solomon, E. B., S. Yaron, and K. R. Matthews. 2002. Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. *Appl. Environ. Microbiol.* 68:397-400.

77. Steele, M., A. Mahdi, and J. Odumeru. 2005. Microbial assessment of irrigation water used for production of fruit and vegetables in Ontario, Canada. *J. Food. Prot.* 68:1388-1392.

78. Stine, S. W., I. Song, C. Y. Choi, and C. P. Gerba. 2005. Effect of relative humidity on preharvest survival of bacterial and viral pathogens on the surface of cantaloupe, lettuce, and bell peppers. *J. Food. Prot.* 68:1352-1358.

79. Stine, S. W., I. Song, C. Y. Choi, and C. P. Gerba. 2011. Application of pesticide sprays to fresh produce: A risk assessment for Hepatitis A and *Salmonella*. *Food. Environ. Virol.* 3:86-91.

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

81. Thomas, J. L., R. M. Slawson, and W. D. Taylor. 2012. *Salmonella* serotype diversity and seasonality in urban and rural streams. *J. Appl. Microbiol.* 114: 907-922

82. United Stated Food and Drug Administration (FDA). 2015. Food Safety Modernization Act (FSMA). http://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm334552.htm#E.

83. United States Department of Agriculture (USDA). 2012. Microbial risk assessment guideline. http://www.fsis.usda.gov/wps/portal/fsis/topics/science/risk-assessments/microbial-risk-assessment-guideline/microbial-risk-assessment-guideline.

84. United States Food and Drug Administration. 2014. Bacteriological Analytical Manual. Chapter 5 *Salmonella*. http://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm070149.htm.

85. United States Geological Survey. 2005. Estimated Use of Water in the United States 2005 http://pubs.usgs.gov/circ/1344/.

86. Verhaelen, K., M. Bouwknegt, S. A. Rutjes, and A. M. de Roda Husman. 2013. Persistence of human norovirus in reconstituted pesticides--pesticide application as a possible source of viruses in fresh produce chains. *Int. J. Food. Microbiol.* 160:323-328.

87. Vose, D. 2000. Risk Analysis: A Quantitative Guide. 2. John Wiley & Sons, Chichester, NY.

88. Waldenström, J., T. Broman, I. Carlsson, D. Hasselquist, R. P. Achterberg, J. A. Wagenaar, and B. Olsen. 2002. Prevalence of *Campylobacter jejuni*, *Campylobacter lari*, and *Campylobacter coli* in different ecological guilds and taxa of migrating birds. *Appl. Environ. Microbiol.* 68:5911-5917.

VI. Appendices

Appendix VI.1: Figures of solar radiation (SR), relative humidity (RH), temperature and rainfall for field trial 2012

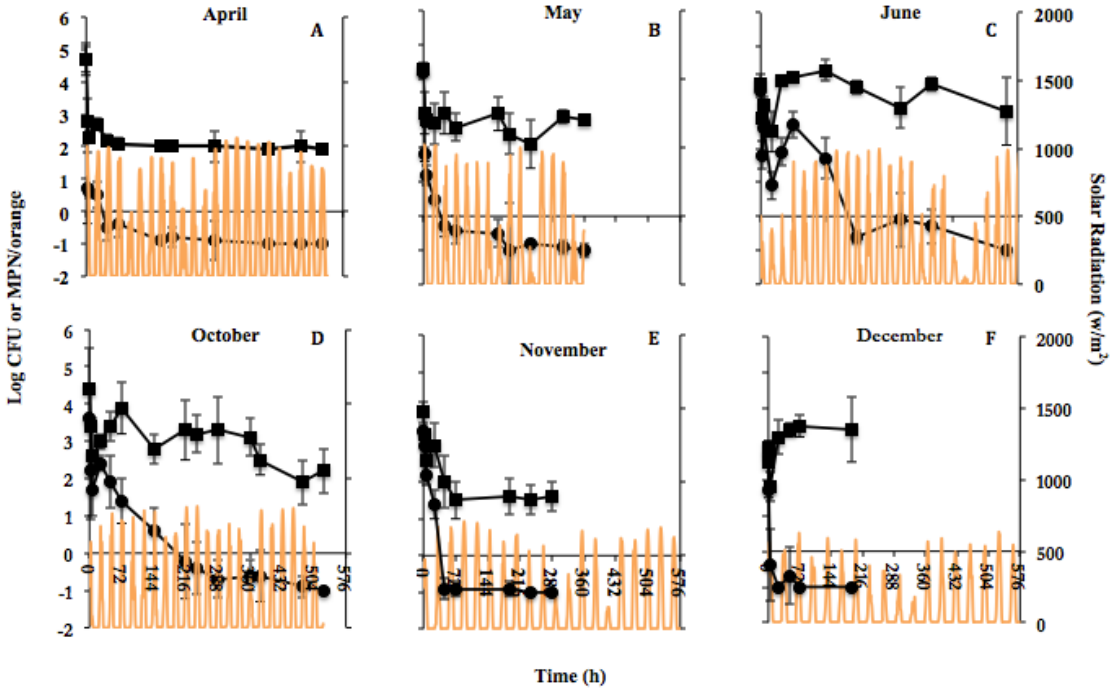

Figure 5.1: High coliforms (box) and high *E. coli* (triangle) innoculation levels plotted against hourly SR data for FT 2012

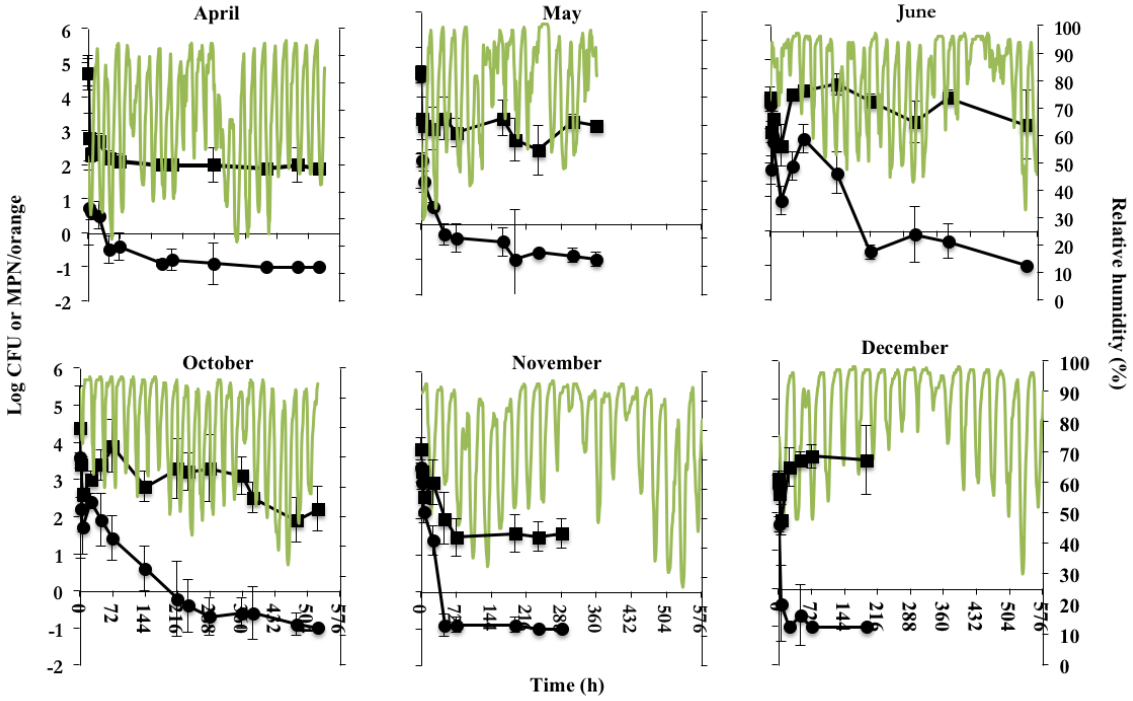

Figure 5.2: High coliforms (box) and high *E. coli* (triangle), innoculation levels plotted against hourly RH data for FT 2012

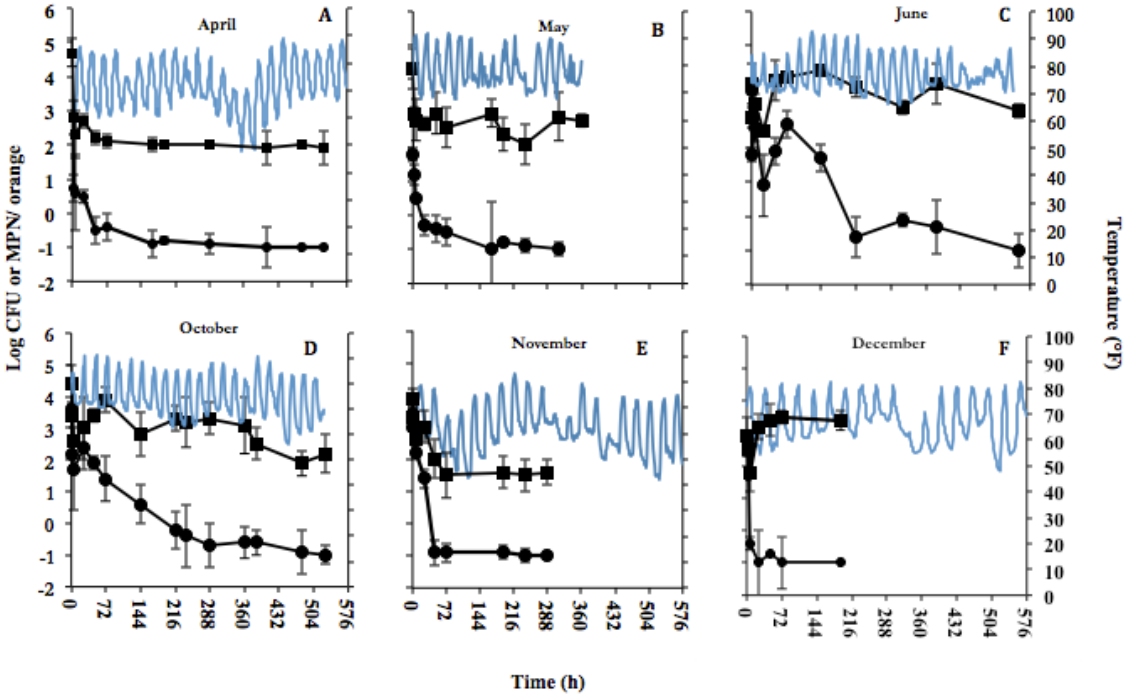

Figure 5.3: High coliforms (box) and high *E. coli* (triangle) innoculation levels plotted against hourly temperature data for FT 2012

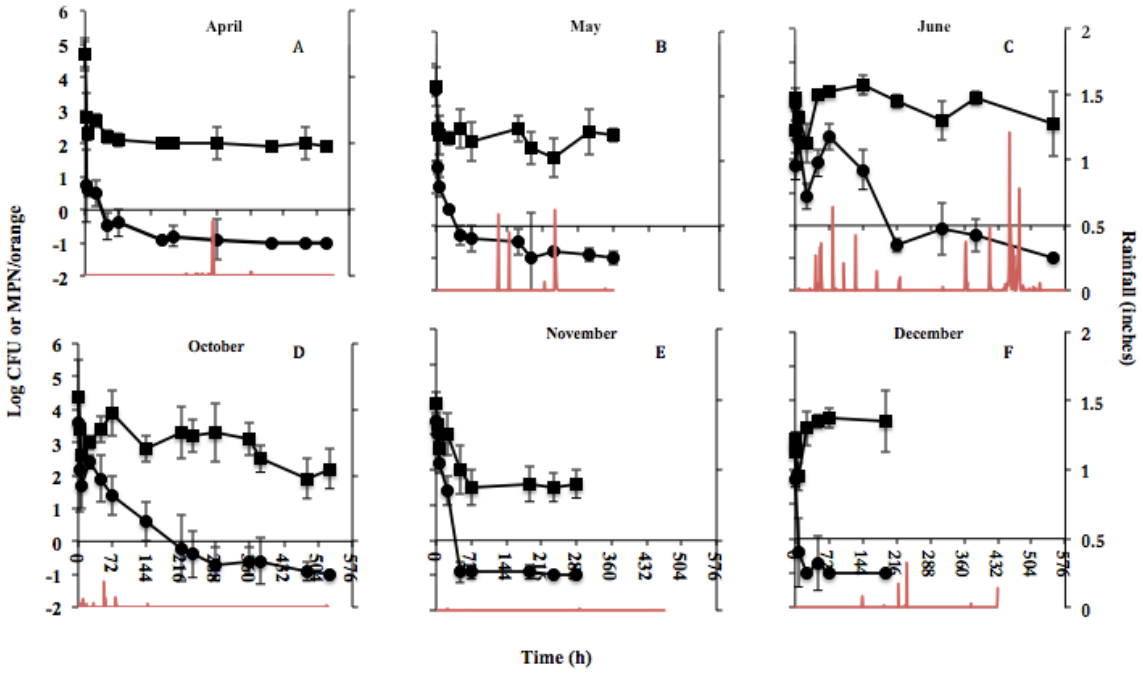

Figure 5.4: High coliforms (box) and high *E. coli* (triangle), innoculation levels plotted against hourly rainfall data for FT 2012

Appendix VI.2: Figures of solar radiation (SR), relative humidity (RH), temperature and rainfall for field trial 2013

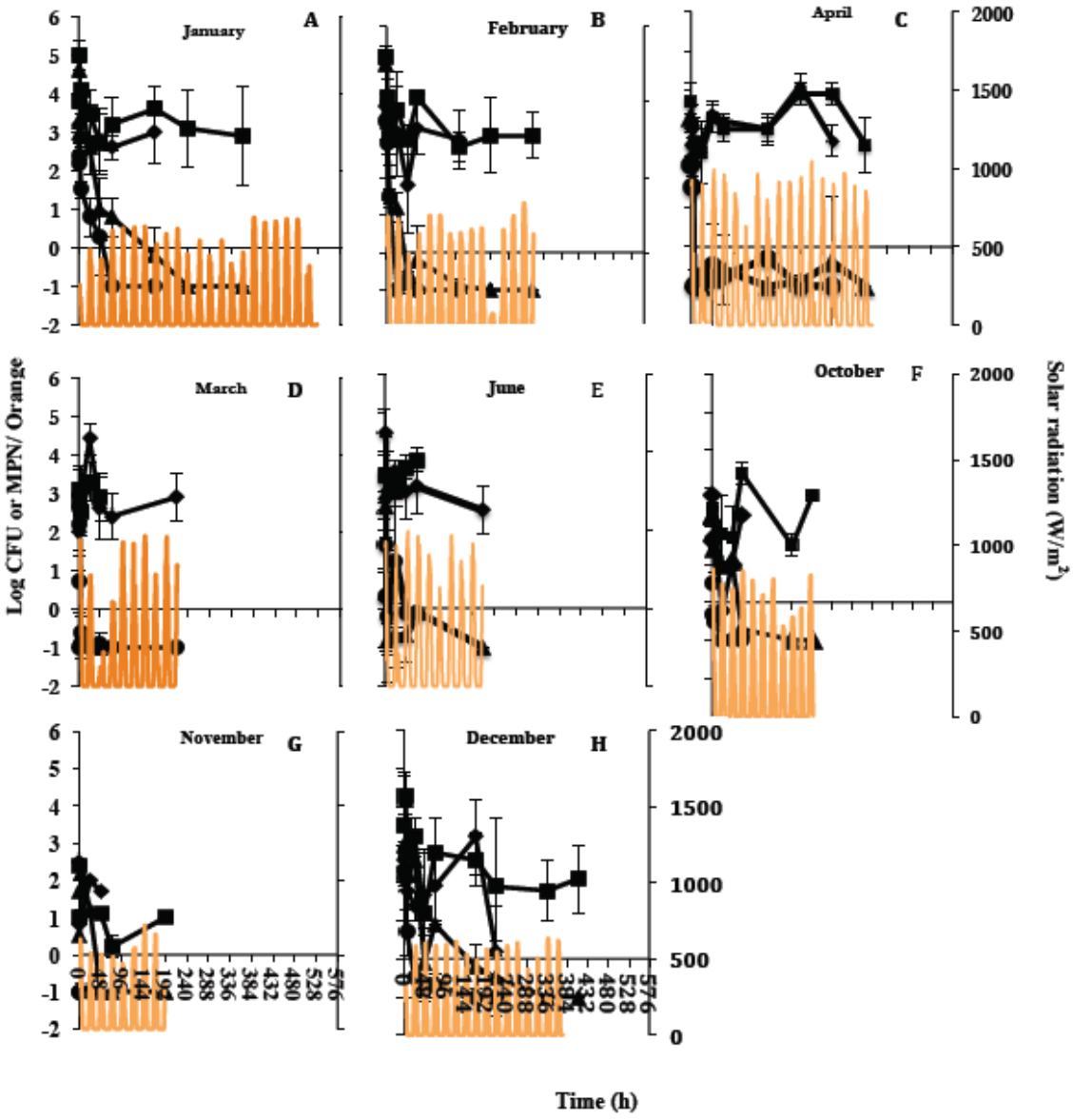

Figure 6.1: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly SR data for FT 2013

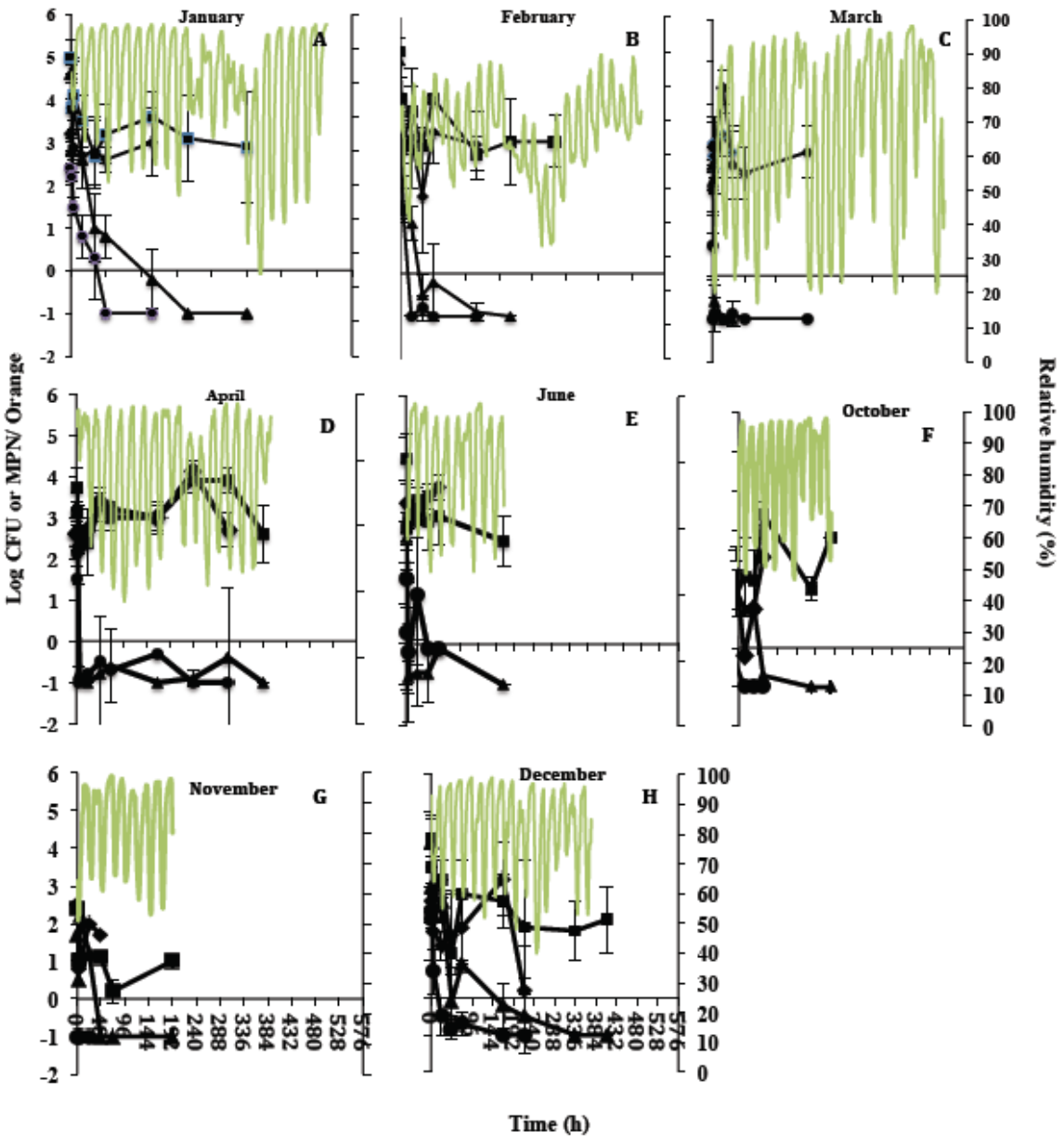

Figure 6.2: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly RH data for FT 2013

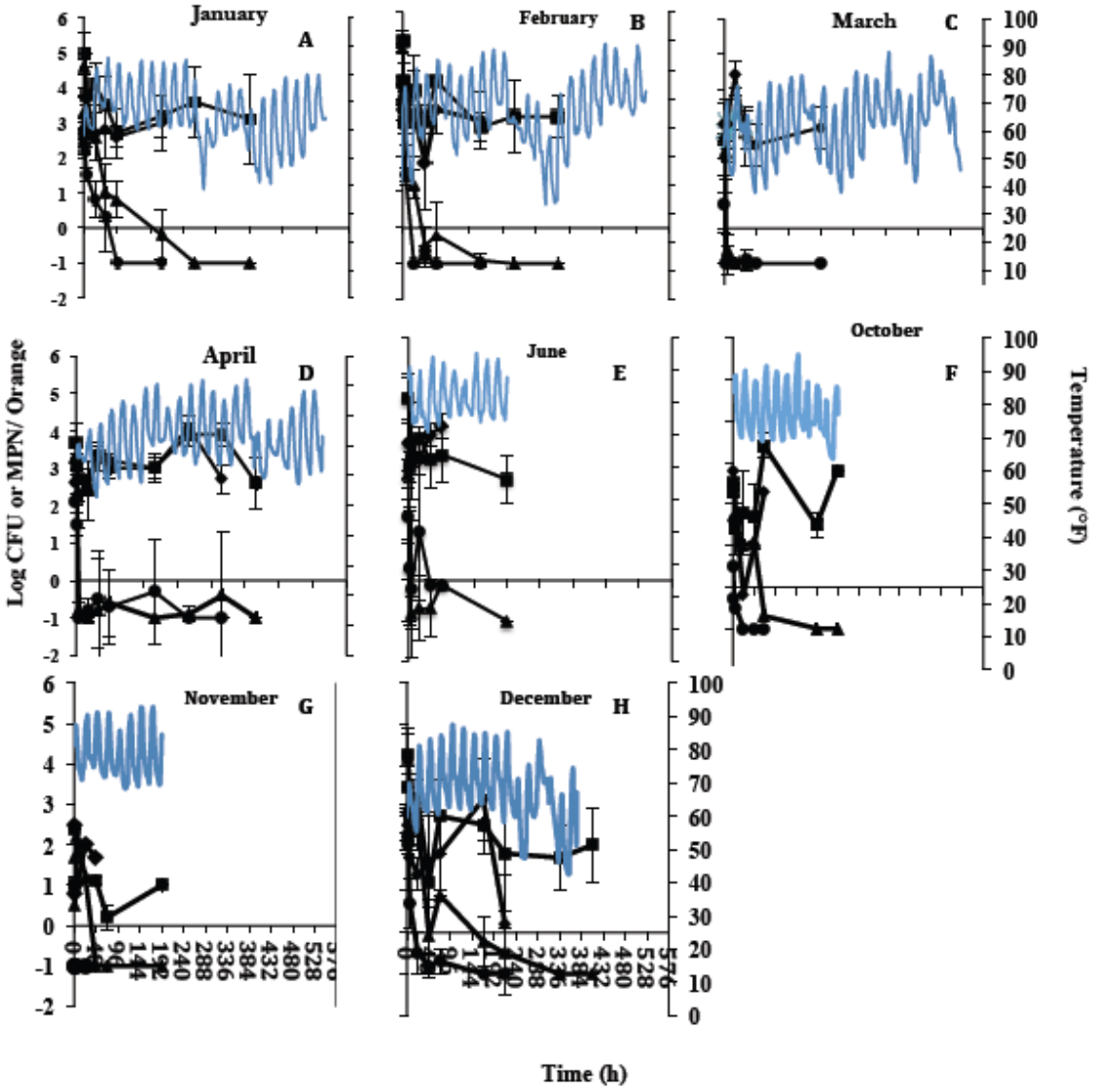

Figure 6.3: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly temperature data for FT 2013

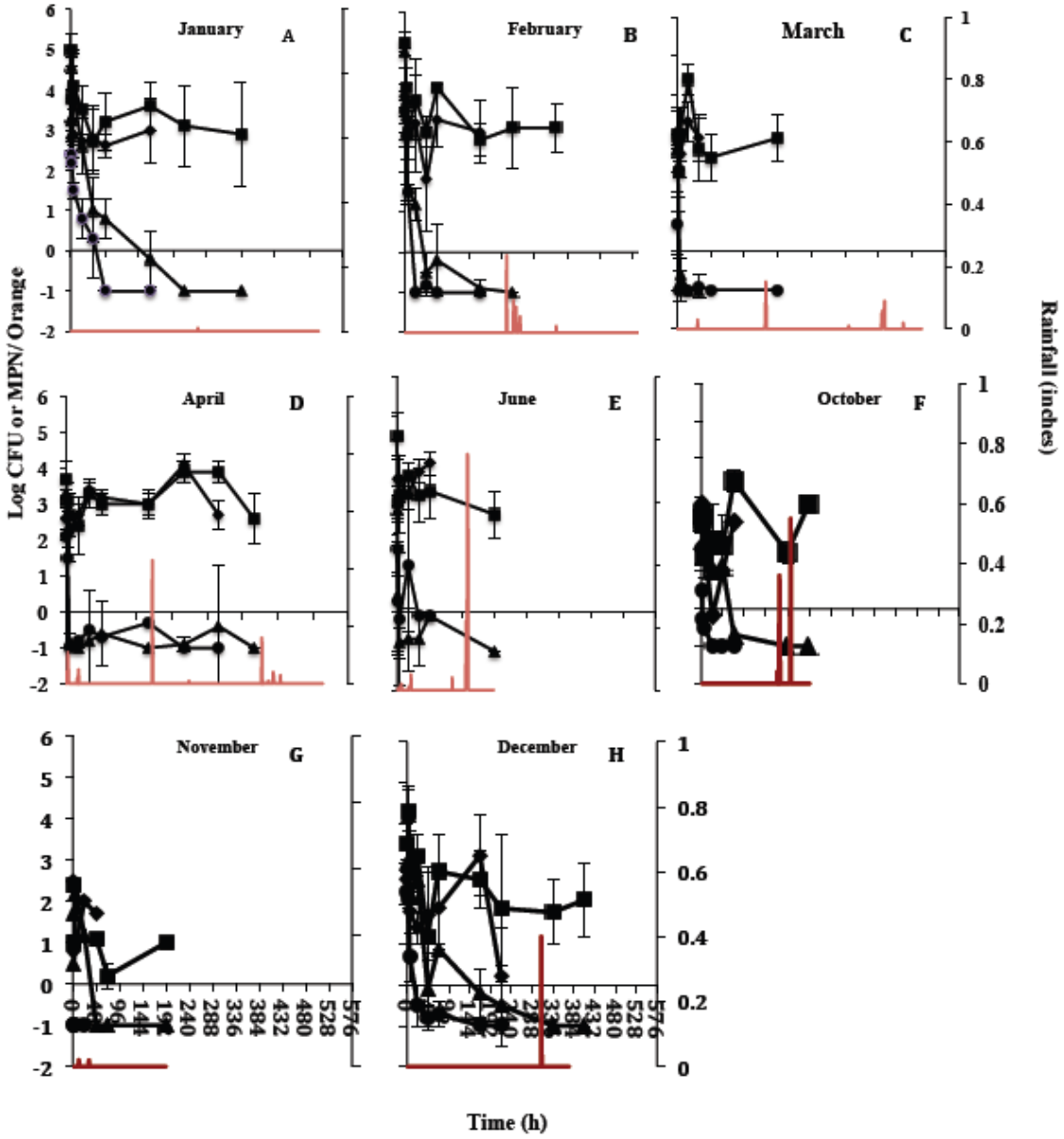

Figure 6.4: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly rainfall data FT 2013

Appendix VI.3: Figures of solar radiation (SR), relative humidity (RH), temperature and rainfall for field trial 2014

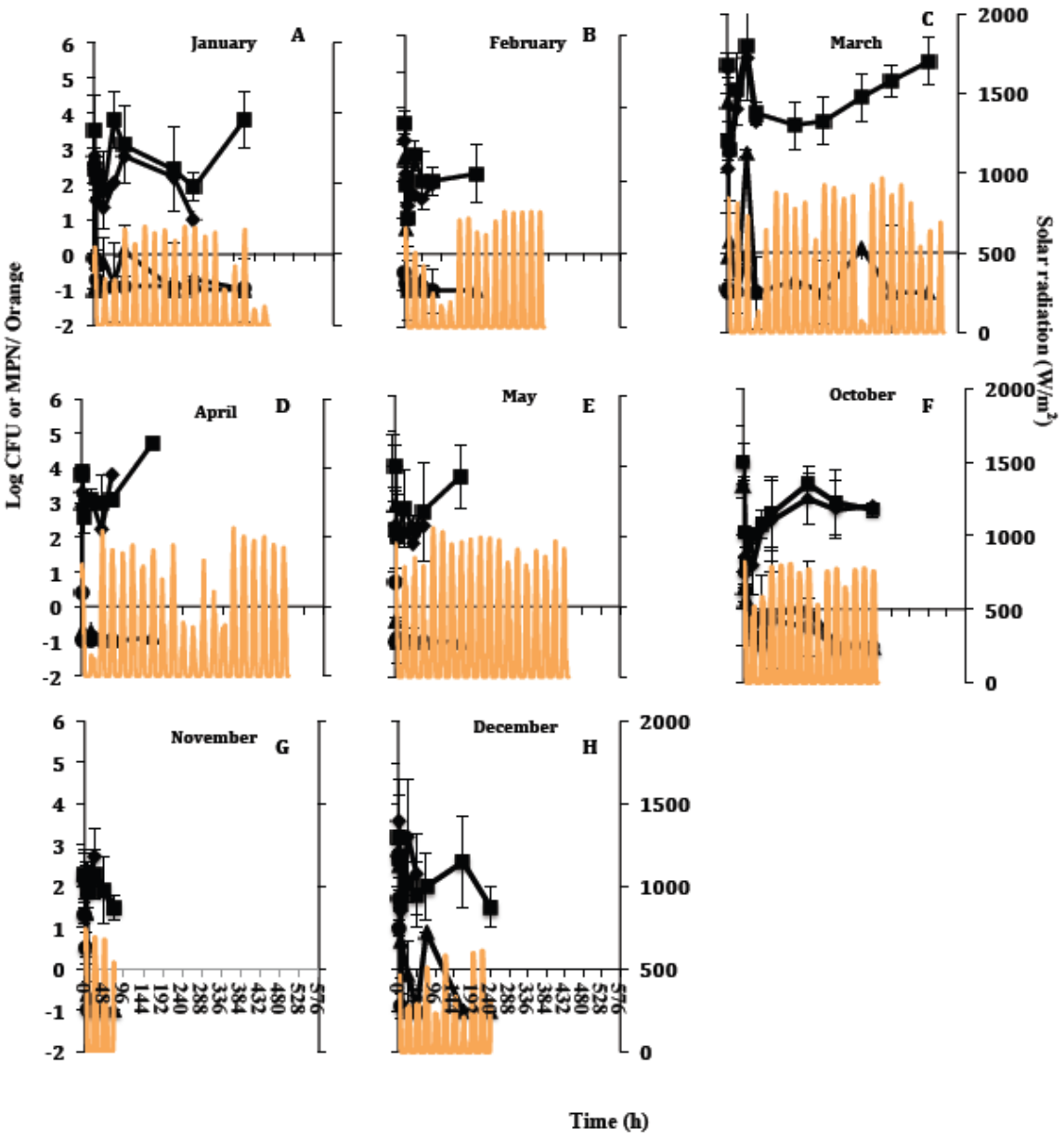

Figure 7.1: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly SR data FT 2014

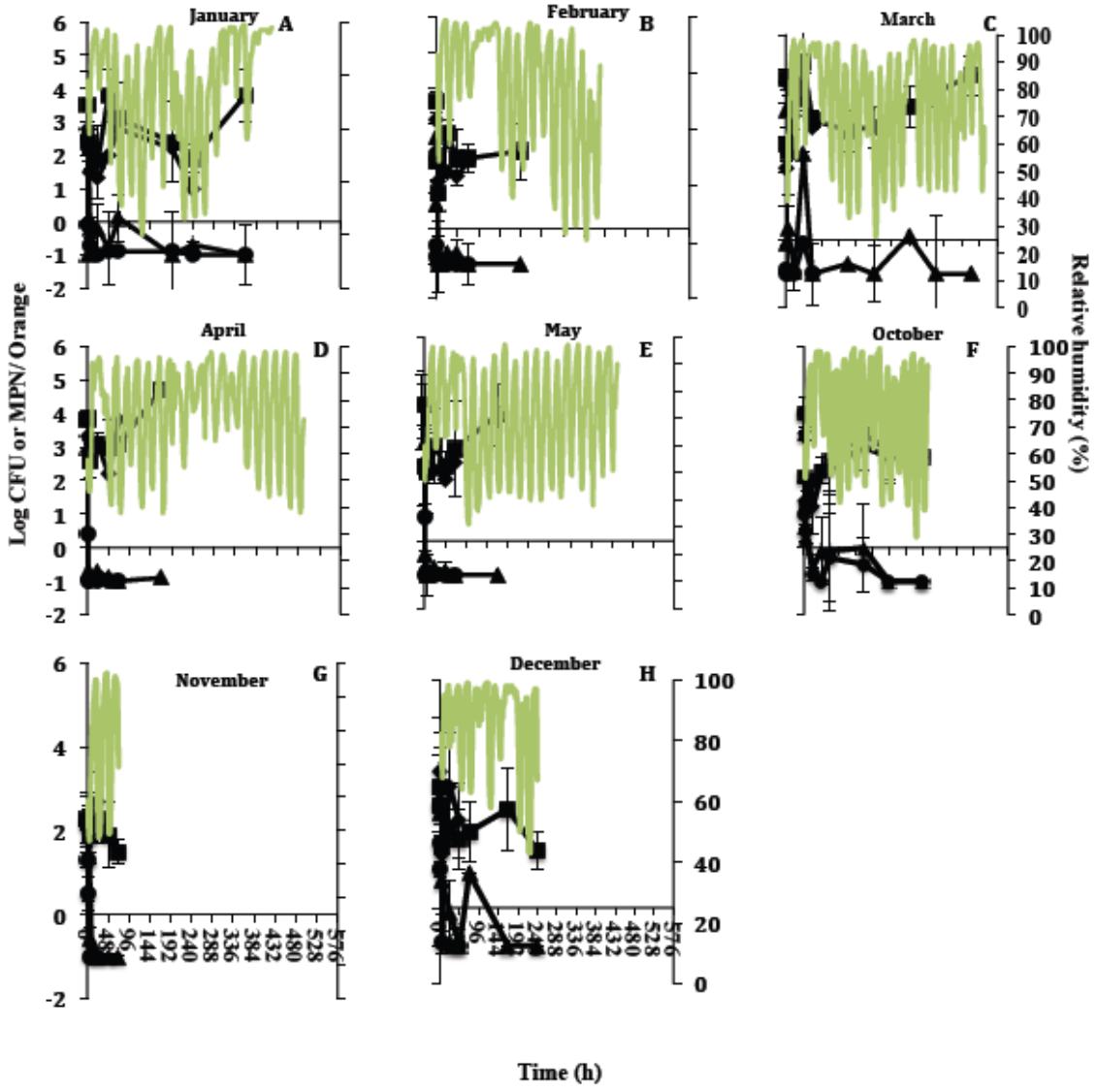

Figure 7.2: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly RH data FT 2014

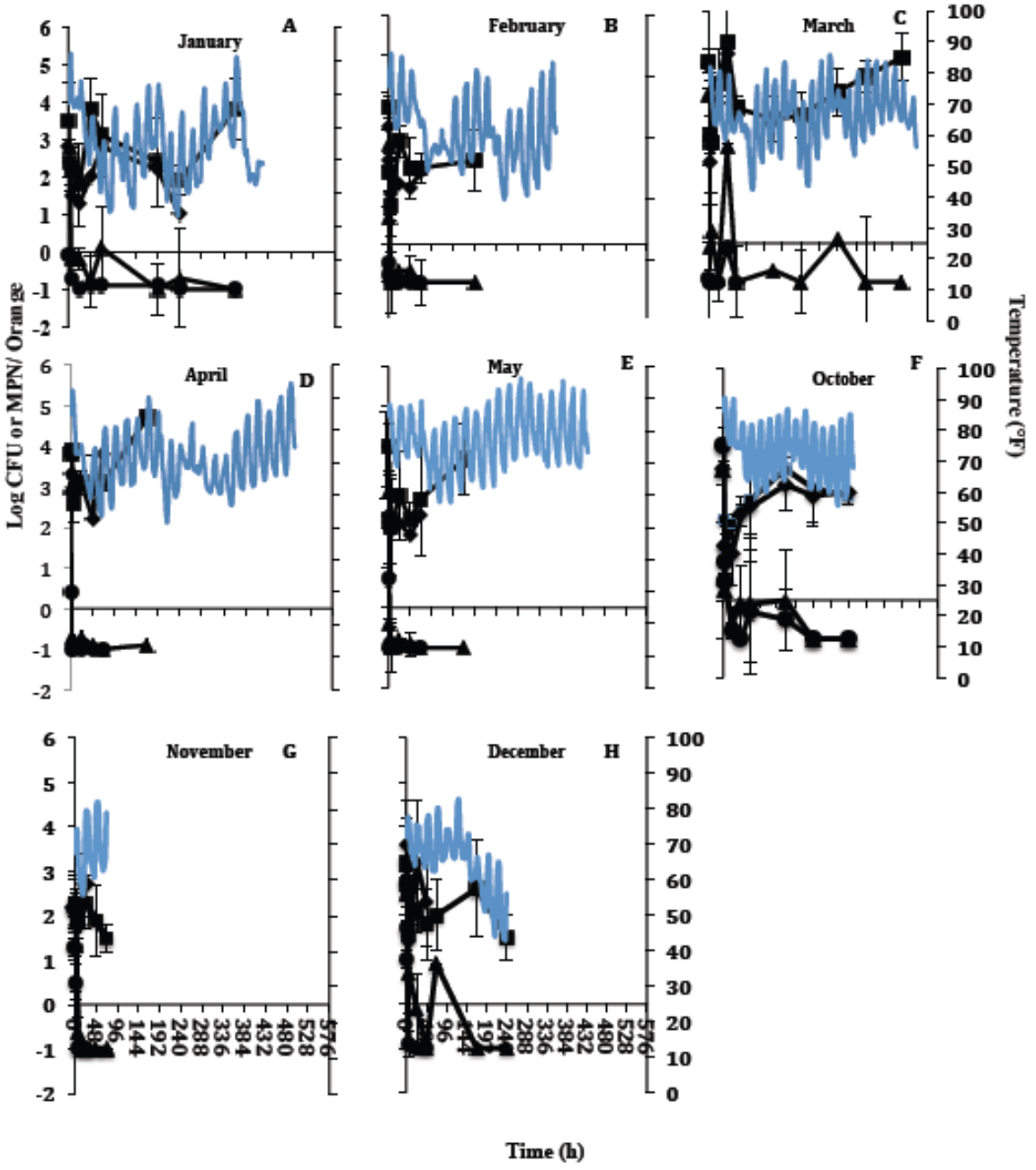

Figure 7.3: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly temperature data FT 2014

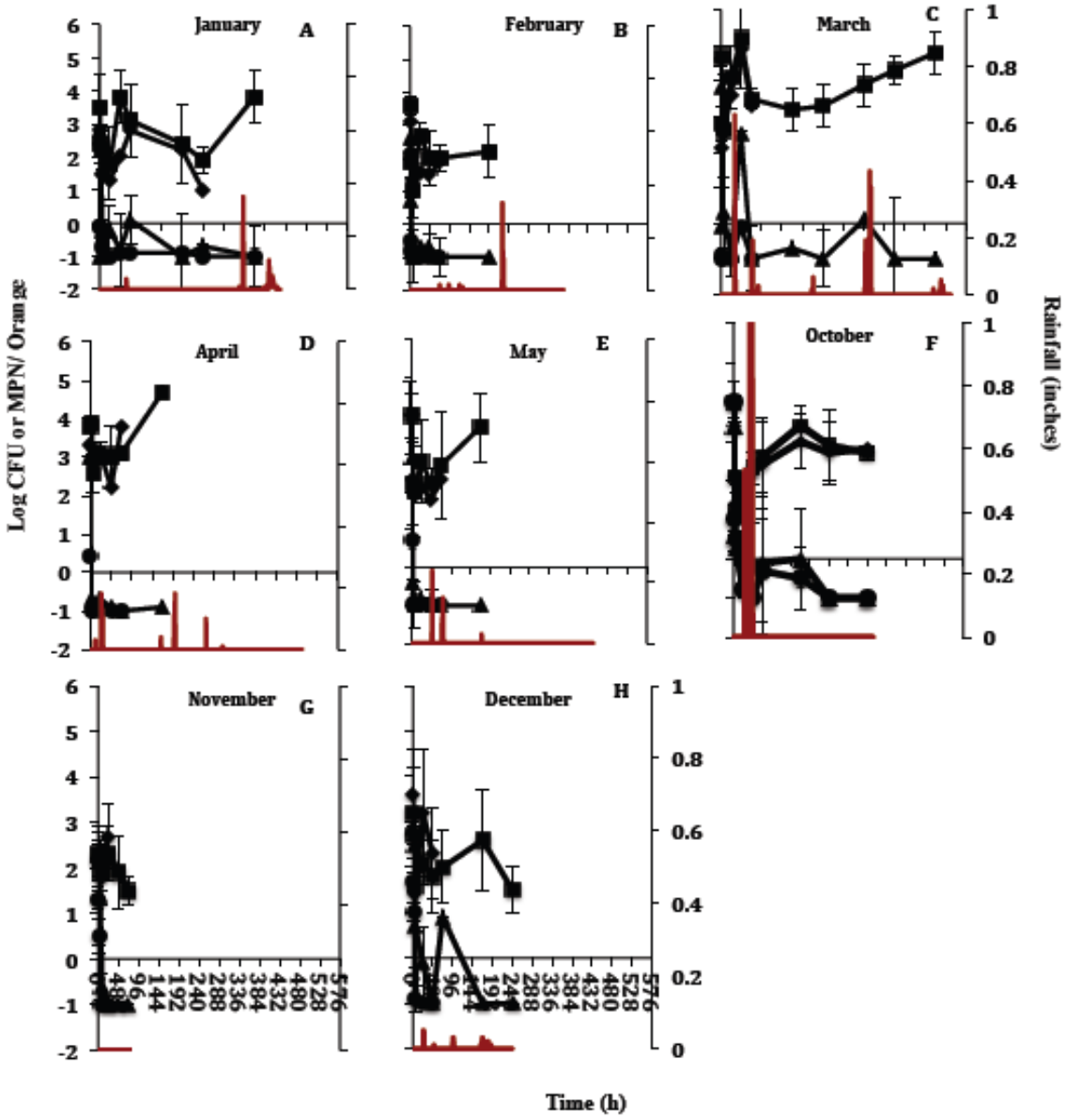

Figure 7.4: High coliforms (box), low coliform (diamond), high *E. coli* (triangle), low *E. coli* (circle) innoculation levels plotted against hourly rainfall data for FT 2014

Appendix VI.4: Figures of solar radiation (SR), relative humidity (RH), temperature and rainfall for field trial 2015

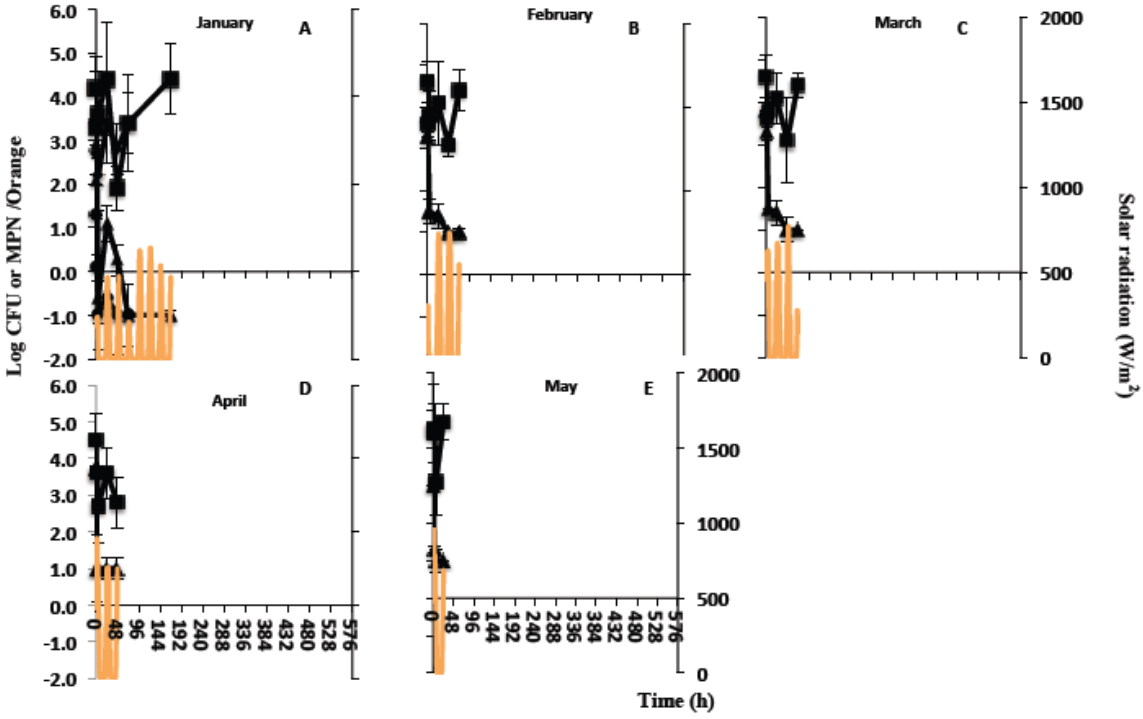

Figure 8.1: High coliforms (box) and high *E. coli* (triangle) innoculation levels plotted against hourly solar radiation data for FT 2015

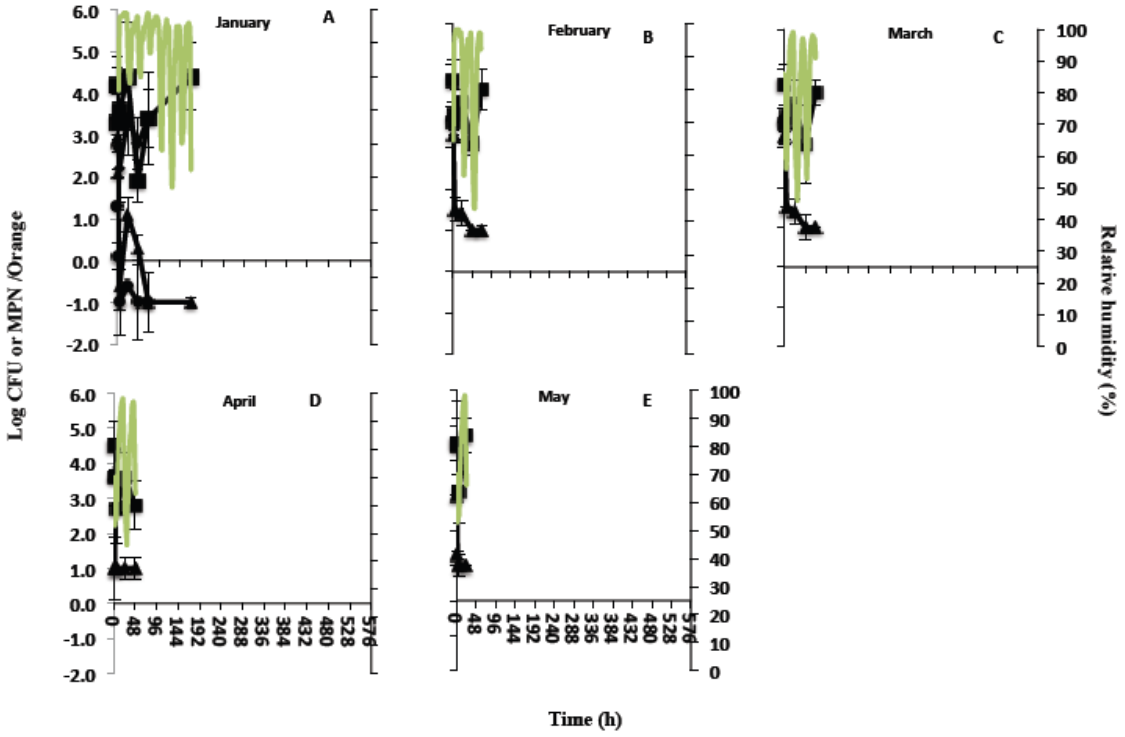

Figure 8.2: High coliforms (box) and high *E. coli* (triangle) innoculation levels plotted against hourly RH data for FT 2015

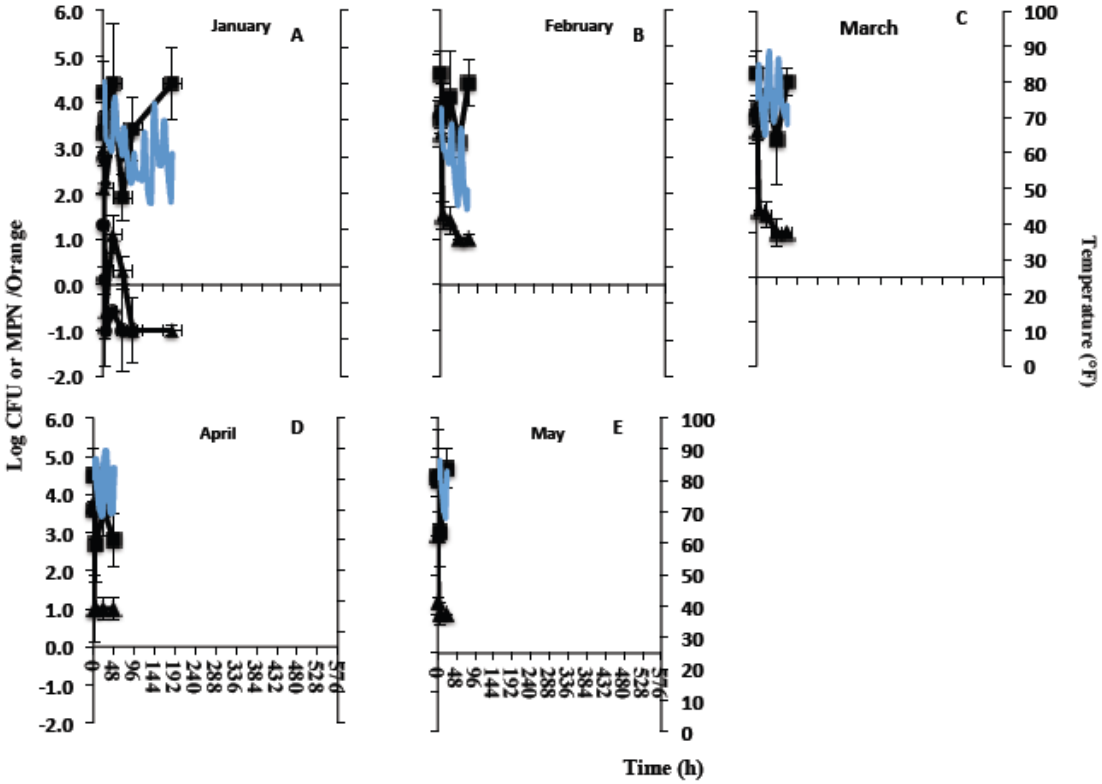

Figure 8.3: High coliforms (box) and high *E. coli* (triangle) innoculation levels plotted against hourly temperature data for FT 2015.

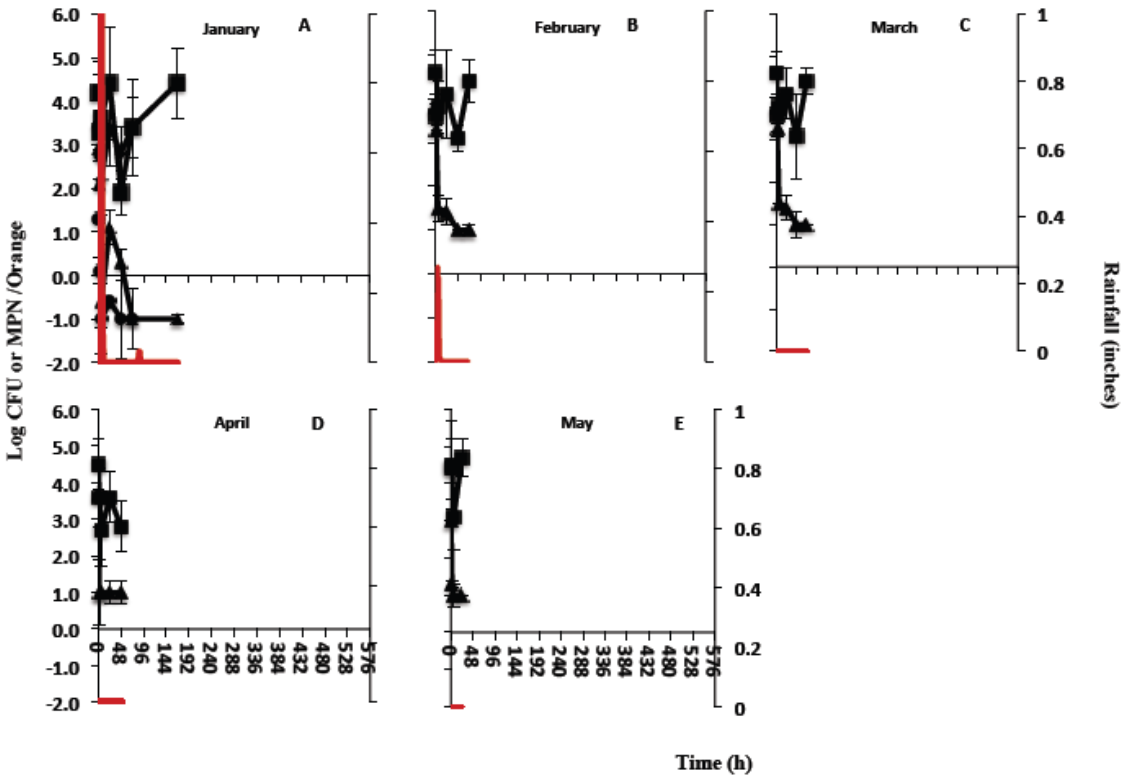

Figure 8.4: High coliforms (box) and high *E. coli* (triangle) innoculation levels, plotted against hourly rainfall data for FT 2015**.**

Appendix VI.5: Figures of predicted 5th percentile, mean and 95th percentile vs. actual field trial data and FDA rule (0.5 log/day for max. of 4 days)

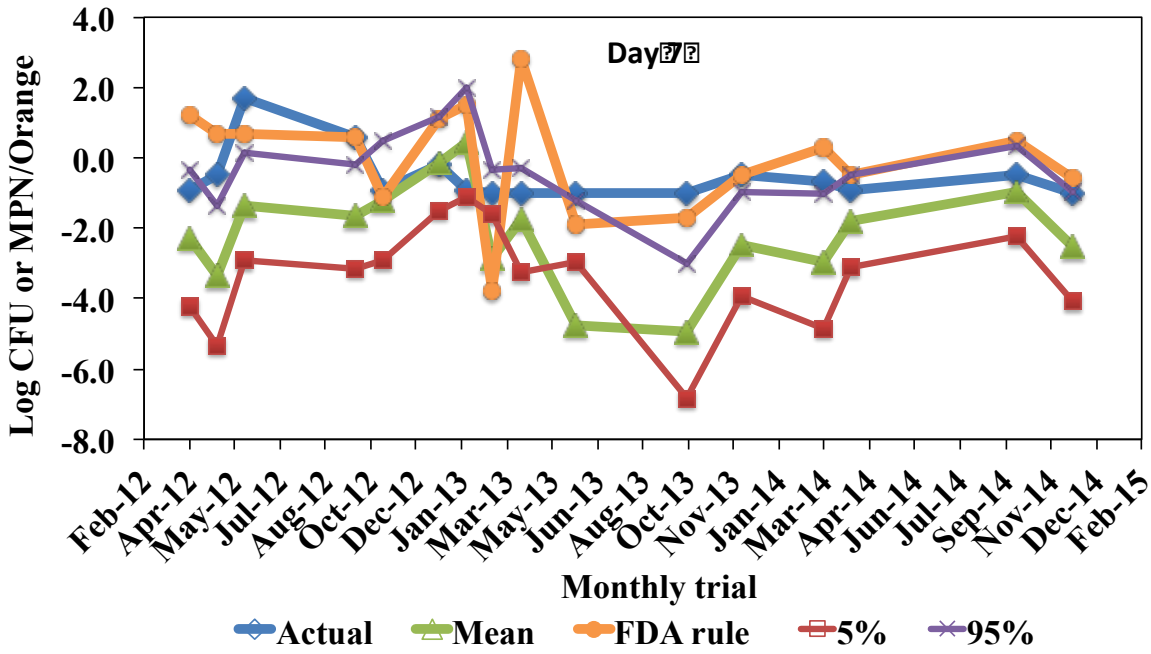

Figure 9.1: Predicted 5% (red line), mean (green line) and 95% (purple line) vs. actual (blue line) and FDA 0.5 log/day (orange line) for day 7 ± 2 day post foliar spray.

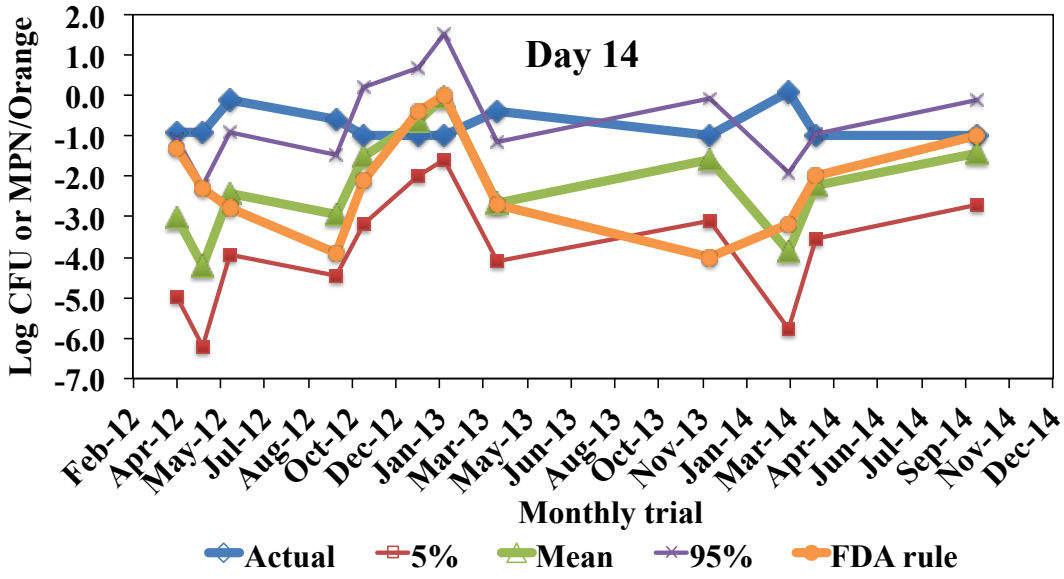

Figure 9.2: Predicted 5% (red line), mean (green line) and 95% (purple line) vs. actual (blue line) and FDA 0.5 log/day (orange line) for day 14 ± 2 day post foliar spray