

Impacts of Climate and Management Variables on the Contamination of Preharvest Leafy Greens with *Escherichia coli*

CHENG LIU,^{1*} NYNKE HOFSTRA,¹ AND EELCO FRANZ²

¹Environmental Systems Analysis Group, Wageningen University, P.O. Box 47, 6700AA Wageningen, The Netherlands; and ²National Institute for Public Health and the Environment (RIVM), Centre Infectious Disease Control, 3720 BA Bilthoven, The Netherlands

MS 15-255: Received 10 June 2015/Accepted 24 August 2015

ABSTRACT

The observed seasonality of foodborne disease suggests that climatic conditions play a role and that changes in the climate may affect the presence of pathogens. However, it is hard to determine whether this effect is direct or whether it works indirectly through other factors, such as farm management. This study aimed to identify the climate and management variables that are associated with the contamination (presence and concentration) of leafy green vegetables with *E. coli*. This study used data about *E. coli* contamination from 562 leafy green vegetables (lettuce and spinach) samples taken between 2011 and 2013 from 23 open-field farms in Belgium, Brazil, Egypt, Norway, and Spain. Mixed-effect logistic and linear regression models were used to study the statistical relationship between the dependent and independent variables. Climate variables and agricultural management practices together had a systematic influence on *E. coli* presence and concentration. The variables important for *E. coli* presence included the minimum temperature of the sampling day (odds ratio = 1.47), region, and application of inorganic fertilizer. The variables important for concentration ($R^2 = 0.75$) were the maximum temperature during the 3 days before sampling and the region. Temperature had a stronger influence (had a significant parameter estimate and the highest R^2) than did management practices on *E. coli* presence and concentration. Region was a variable that masked many management variables, including rainwater, surface water, manure, inorganic fertilizer, and spray irrigation. Climate variables had a positive relationship with *E. coli* presence and concentration. Temperature, irrigation water type, fertilizer type, and irrigation method should be systematically considered in future studies of fresh produce safety.

Leafy green vegetables (LGVs) have been identified as the fresh produce commodity group of highest concern from a microbiological safety perspective (8) because they are often grown in the open field and are vulnerable to contamination from manure used as fertilizer, soil, water used for irrigation, and wildlife feces (8). Moreover, LGVs are consumed raw and in large volumes. Pathogenic *Escherichia coli* strains are one of the main concerns with respect to foodborne disease associated with LGVs (11, 12, 35, 39).

The incidence of foodborne disease is generally correlated with climate conditions (19, 27, 40). Roughly one-third (population attributable fraction) of salmonellosis cases in England, Wales, Poland, The Netherlands, the Czech Republic, and Switzerland can be linked to higher temperatures (34). In Australia, the rate of salmonellosis also increases with decreasing latitude and, consequently, with increasing average annual temperatures (13). The correlation between foodborne disease and climatic conditions is (partly) reflected in the strong seasonality of many foodborne diseases. The seasonal salmonellosis patterns have been statistically correlated with the mean monthly temperature of the previous month (7). Similarly, in the

Australian subtropical and tropical regions, temperature and precipitation have been positively associated with the number of salmonellosis cases (45).

The mechanisms underlying the observed seasonality in foodborne disease are not fully understood, but they are likely a complex interplay of different factors. Besides climatic conditions, these factors include human behavior and consumption patterns (43, 46), farm management practices, pathogen prevalence in the animal reservoir, and pathogen environmental survival patterns. The risk of foodborne disease associated with LGVs is directly related to the likelihood of occurrence and the subsequent level of contamination. The observed seasonality suggests that climatic conditions influence the presence and/or level of pathogens (22). Improved understanding of this is important for better control and surveillance of LGV contamination, particularly in the face of ongoing climate change. Determination of whether the effect of climate on LGV contamination is direct or indirect (through other factors, e.g., farm management) remains elusive. These farm management practices are affected by climate and could be influenced even more or in different ways by climate change. Though uncertain, the effects of seasonality and climate on produce food safety should not be ignored because they may result in higher risks (22).

* Author for correspondence. Tel: +31 317 575270; Fax: +31 317 419000; E-mail: cheng.liu@wur.nl.

Changes in temperature, distribution of precipitation (including more extreme events, such as floods and droughts), UV radiation, and moisture content are already observed worldwide (25). Temperature has increased since the start of observations in 1654 (4). Consistent with precipitation changes, runoff has been notably reduced in southern Europe and has increased in Southeast Asia and at high latitudes. The larger simulated runoff changes reached a 20% increase compared with 1980 to 1999 mean values (37). Droughts have already become more common, especially in the tropical and subtropical regions since the 1970s (37). These changes will likely become more apparent in the future (37). Climate changes will mainly impact the contamination sources and pathways of bacteria onto LGVs during the preharvest phase. Other phases of the food chain will be less affected, because processing and transport generally occur in controlled environments.

Good agricultural management practices are essential for food safety control, and they are often applied in response to particular climatic conditions (23). Response strategies have been developed to adapt to the pressures on fresh produce safety due to climate change (20). The use of animal manure for fertilization of production fields increases the risk of *E. coli* contamination on both organic and conventional farms significantly (29, 31, 32). Farmers should follow good agricultural practices for handling animal manure to minimize the risk of introducing pathogens. Such practices include manure composting and minimizing direct or indirect contact between manure and produce (41). The use of contaminated surface water may also lead to contamination (17, 18, 36). Groundwater is generally less likely to be contaminated with pathogens than is surface water. Shallow wells and wells that are old or improperly constructed are more likely to be susceptible to contamination (6, 41). Infected persons who work with fresh produce may also transmit foodborne illnesses (41). Farmers should understand and follow basic hygiene principles to reduce the likelihood of contamination of fresh produce and water supplies and the spread of contamination among workers.

A review of the impacts of climate change on microorganisms (22) has led to improved qualitative understanding, which will now be used to study these impacts quantitatively. Farmers are likely to change management practices to adapt to climate change (20). Although a better understanding of the quantitative changes due to management schemes and climate change is important, such quantitative analyses are sparse due to a lack of available data. Therefore, aggregation of the available information in a meta-analysis may be useful to achieve higher statistical power and generalizability.

This study aimed to explore the relative contribution of climate and agriculture management factors to *E. coli* contamination on LGVs across different regions. To identify a combination of statistically significant variables that best explain the observed variation in *E. coli* presence and concentration throughout regions, we applied statistical modeling to data from production fields in different regions. The meta-analysis in this study combines findings from independent studies from different regions within the Veg-i-

Trade project; the results of logistic and linear regressions with climate and management variables are presented and summarized; and, finally, data complexity and limitations are discussed, concluding with lessons learned in this meta-analysis.

MATERIALS AND METHODS

Data. The data used in this study were collected in the Veg-i-Trade project, which studied the impact of climate change and globalization on the safety of fresh produce. A Horticultural Assessment Scheme (HAS) was developed in the Veg-i-Trade project to assess the microbiological quality of LGVs. HAS was a systematic approach to sample, analyze, and standardize the sampling scheme in various regions within Veg-i-Trade. HAS defined the identification of critical sampling locations, the selection of microbiological parameters, the assessment of sampling frequency, the selection of sampling method and method of analysis, and, finally, data processing and interpretation (14). All Veg-i-Trade sampling data used in this study were collected and analyzed under HAS.

This meta-analysis included raw sampling data from Holvoet et al. (15), Ceuppens et al. (6), Uyttendaele et al. (42), and Castro-Ibáñez et al. (5). Our study used *E. coli* contamination data from 562 LGV samples taken from 23 open-field farms from six regions (Fig. 1): Belgium ($n = 160$), Brazil ($n = 69$), Egypt ($n = 18$), Norway ($n = 99$), and Spain ($n = 216$). All farms grew lettuce, except for the farms in Spain, which grew spinach. The data were collected from 2011 to 2013 by different laboratories of the local universities or research institutes within the Veg-i-Trade project. Each laboratory used its own detection limits (Table 1). All samples were taken at the moment of harvest or 1, 2, or 3 weeks before harvest. Climate variables included temperature (T) and precipitation (Pr): daily average (T_{avg} , $T_{\text{avg}3}$, $T_{\text{avg}7}$), minimum (T_{min} , $T_{\text{min}3}$, $T_{\text{min}7}$), and maximum (T_{max} , $T_{\text{max}3}$, $T_{\text{max}7}$) temperatures of the sampling day and of 3 and 7 days before the sampling day and daily precipitation (Pr) and total precipitation for 3 (Pr3) and 7 days (Pr7). Management variables included the categorical variables region and toilet distance (ToiletD) and binary variables: drinking water (DrinkingW); rainwater (RainW); groundwater (GroundW); surface water (SurfaceW); drip irrigation (Drip); spray irrigation (Spray); flood irrigation (Flood); composted manure (derived) (Manure), which included composted manure and a mixture of fertilizer with composted manure; inorganic fertilizer (Inorganic); nonanimal organic fertilizer (NonAOrganic); and farm animal presence (FarmA). Missing values were omitted because regression models can only run with a complete set of observations. Temperature and precipitation data were collated from the weather station nearest each farm. The data sources are summarized in Table 1. Management information was collected by means of a questionnaire given to farmers. All variables are defined in Table 2.

Statistical model development in general. The data were analyzed using the statistical software package R version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria). All statistical tests were assessed for significance at the 95% confidence level ($P < 0.05$), except for the univariate analysis ($P < 0.25$).

The data were checked for collinearity by the variance inflation factor between categorical and binary variables and the phi coefficient between binary variables. In case of collinearity (variance inflation factor > 2 or phi coefficient > 0.6), the variable

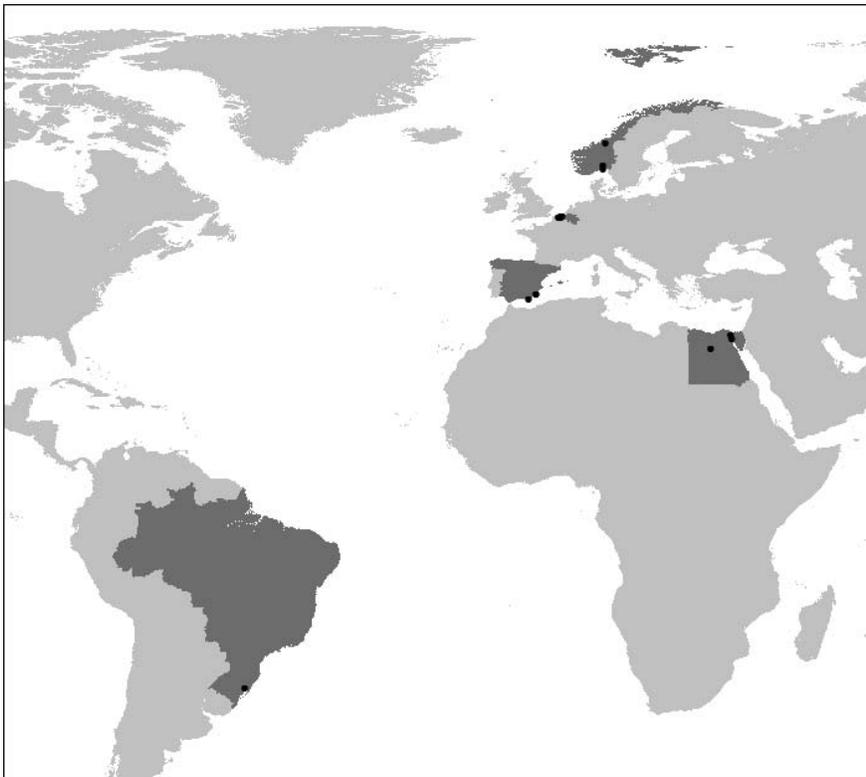


FIGURE 1. Sampling farms distribution. Black dots represent sampling farms that cooperated in the Veg-i-Trade project.

with the least biological relevance was omitted from further analysis.

We developed models for *E. coli* presence (logistic regression) first and for *E. coli* concentration (linear regression) afterwards. We focused on assessing the relationship of *E. coli* with climate variables first and then assessed management variables. This way, impacts of climate or management variables can be analyzed separately to further understand how climate or management individually influences the safety of LGVs. After that, all variables were combined in the final model to study the overall effects of climate and management influence on LGV safety.

To select variables for the logistic regression models, the stepwise selection method developed by Hosmer and Lemeshow (16) was used. Spearman's rank correlation was used to assess correlations between numeric and ordinal variables. In the first step, univariate analysis was applied by fitting a univariable regression model to obtain the estimated coefficient, the estimated standard error, the likelihood ratio test for the significance of the coefficient, and the univariable Wald statistic. Any variable with a likelihood ratio test that has a P value <0.25 is a candidate for the multivariable model. Using a more tolerant significance level ($P <0.25$ instead of $P <0.05$) allowed for inclusion of variables that are of potential importance at the model building stage (3, 26). In the second step, the candidate variables from step one were entered into the backward selection method to select the variables for the multivariable model. The overall importance of each categorical variable included in the multivariable model was verified by an examination of the Wald statistic. In the third step, variables that were not selected for the multivariable model were added back into the model. By doing this, we could identify the variables that, by themselves, are not significantly related to the *E. coli* present but that make an important contribution in the presence of other variables (16). Interactions and quadratic terms were checked for the variables in the model.

After the variables were selected, a mixed-effect model was applied, because of the hierarchical structure in the data (with

repeated sampling at the same farm). We used the “lme4” package (2) and “lmeTest” package (21) in R software (R Core Team, R Foundation for Statistical Computing) for the mixed-effect model. In this study, all samples were combined and treated as one data set. Region was treated as a fixed variable to cover the differences in sampling and detection limits among regions.

Logistic regression model. To investigate the presence or absence of *E. coli* on LGVs in the open-field farms, data ($n = 562$) were fitted to a logistic regression model, combining the different variables and locations together. This model aimed to separately assess the relative contributions of the climate and management variables to the observed variation in *E. coli* presence.

In the logistic regression model, the Akaike Information Criterion (AIC) was used to compare and select the best model. AIC measures the relative quality of a model for model selection (1). After the final model was chosen, the odds ratio was calculated from the parameter coefficients. The parameter coefficients give the change in the log odds of *E. coli* presence for a one-unit increase in the predictor variable.

To assess the robustness of a model's predictive ability, a 10-fold cross-validation was conducted. The data were randomly divided into 10 subsets of equal size; nine subsets were used to train the model, and the 10th was used to test the model's predictive ability. This process was repeated 10 times, each time with a different test subset. So, all observations were used for both training and testing, and each observation was used for validation exactly once. The mean area under the curves was calculated. An area of 100% represents a perfect test, and an area of 50% represents a worthless test.

Linear regression model. To investigate the relationship between climate and management variables and the *E. coli* concentration on preharvest LGVs, the *E. coli*-positive data ($n = 59$) were fitted to a linear mixed-effect model combining the

TABLE 2. Definitions of explanatory variables

Variable	Unit	Description
T_{avg}	°C	Avg air temp of the sampling day
T_{avg3}	°C	Avg air temp of 3 days before the sampling day
T_{avg7}	°C	Avg air temp of 7 days before the sampling day
T_{min}	°C	Minimum air temp of the sampling day
T_{min3}	°C	Minimum air temp of 3 days before the sampling day
T_{min7}	°C	Minimum air temp of 7 days before the sampling day
T_{max}	°C	Maximum air temp of the sampling day
T_{max3}	°C	Maximum air temp of 3 days before the sampling day
T_{max7}	°C	Maximum air temp of 7 days before the sampling day
Pr	mm	Amt of precipitation on the sampling day
Pr3	mm	Total precipitation of 3 days before the sampling day
Pr7	mm	Total precipitation of 7 days before the sampling day
Rainwater	Yes/no	Does the farmer use rainwater for irrigation?
Surface water	Yes/no	Does the farmer use surface water for irrigation?
Drinking water	Yes/no	Does the farmer use surface water for irrigation?
Groundwater	Yes/no	Does the farmer use groundwater for irrigation?
Spray irrigation	Yes/no	Does the farmer use spray irrigation?
Drip irrigation	Yes/no	Does the farmer use drip irrigation?
Flood irrigation	Yes/no	Does the farmer use flood irrigation?
Composted manure (derived)	Yes/no	Does the farmer use composted animal manure or mixed fertilizer with composted manure?
Inorganic fertilizer	Yes/no	Does the farmer use inorganic fertilizer?
Nonanimal organic	Yes/no	Does the farmer use compost or mixture of mineral and organic fertilizer?
Farm animal presence	Yes/no	Does the farmer also own farm animals?
Toilet distance	m	The distance between the fields and the nearest toilet (groups of distances were created)

different variables as fixed-effect variables and Farm as a random effect. Linear regression models assess the relative contributions of the climate and management variables to the observed variation in *E. coli* concentration. We used log-transformed *E. coli* concentration levels to approximate data normality. Visual inspection of residual plots (standardized residuals with fitted value) for the homoscedasticity test and a quantile-quantile plot for the normality test were performed to check the assumptions for linear regression. The *F* test was used to select the best model fit. Root mean square error and mean absolute error were calculated to evaluate models. Lower values for both performance indicators are better.

RESULTS

Data. In total, 59 LGV samples were positive for *E. coli*. Very few samples (in total 18) were tested for *E. coli* on farms in Egypt, but more than half of them were positive (Table 1). Among 99 samples in Norway, only three were *E. coli* positive (Table 1). Presence ranged from 3% in Norway to 55.6% in Egypt. Note that differences in presence can be due to differences in the lower detection limit among studies, ranging from 0.7 log CFU/g in Belgium to 2 log CFU/g in Spain (Table 1).

The mean (median) concentration of the positive samples was 1.91 (2.00) log CFU/g, with a standard deviation of 0.81. In general, the highest *E. coli* concentrations were found in samples from Brazil (Fig. 2). The observed *E. coli* concentrations on LGVs ranged from the detection limits to 3.9 log CFU/g in Brazil (Fig. 2). *E. coli* concentrations below the detection limits are indicated as 0 log CFU/g in Figure 2.

The T_{avg} in Brazil was very high (up to 30°C, Table 1). Brazil also had the highest Pr among all regions (Table 1). In

contrast, Egypt had no rain on the sampling days. In general, the days on which *E. coli*-positive samples were found were dry in all regions, except for Brazil (Table 1). The variation in climate was due not only to the different geographic locations, but also to different growing seasons. For example, farmers in Spain grow spinach in their wintertime, from September to March, to avoid their high summer temperatures, whereas farmers in the rest of the regions grow lettuce during their summertime. Consequently, the temperature range during the sampling period in Spain was similar to that during the sampling period in other regions, such as Norway (Table 1). In total, about 60% of the sampling days had precipitation amounts lower than 0.1 mm.

In this data set, management variables were region specific. Some of the regions happened to have only one type of irrigation water or irrigation method. For instance, flood irrigation was applied only in Egypt. All samples from Belgium used rainwater, and all samples from Egypt and Spain used composted manure or a mixture with composted manure. The details for each management variable are summarized in Table 1.

***E. coli* presence on LGVs.** Logistic regression was applied separately for both climate variables and management variables to assess relationships with *E. coli* presence on LGVs. Results are presented accordingly in this section.

For climate variables, univariate analysis results showed that all variables in this analysis had a *P* value of <0.25, except for Pr (Table 3). Because temperature and precipitation variables were not independent, only one variable should be selected for temperature and one for precipitation. The univariate Wald test suggested that Pr3 was the only

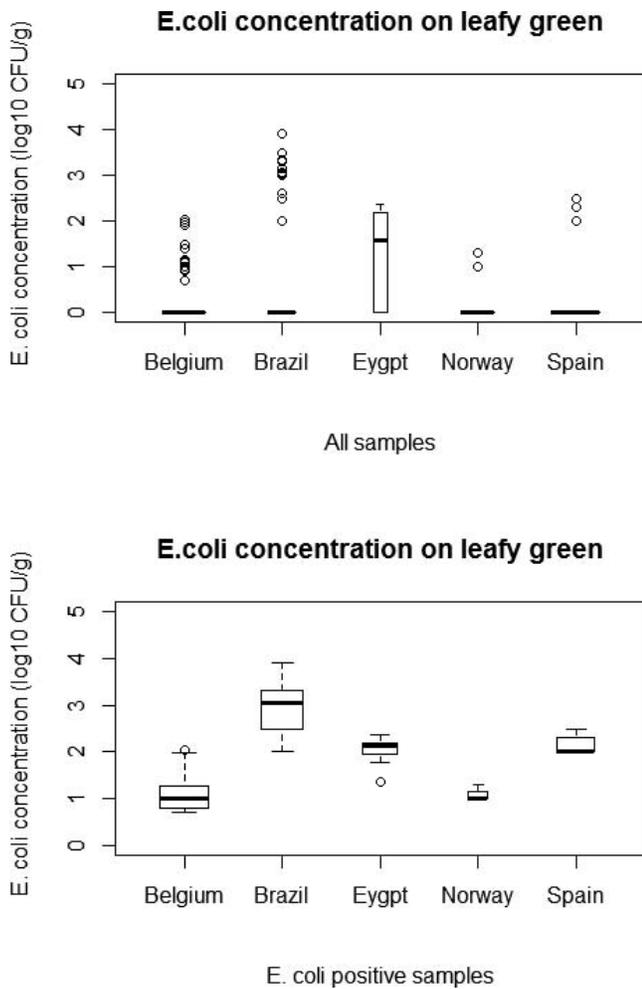


FIGURE 2. *E. coli* concentrations (top: all samples, bottom: positive samples only) on lettuce sampled from farms in Belgium, Brazil, Egypt, and Norway and on spinach from farms in Spain. The box width is proportional to the square root of the number of observations in each region.

significant precipitation variable ($P < 0.05$). T_{\min} had the lowest AIC value. So, T_{\min} , Pr3, and Region were chosen after the univariate analysis.

Upon completion of the univariate analyses, first the variables T_{\min} , Pr3, and Region were selected for the multivariable analysis. With backward selection, the Wald test for importance of the categorical variable, and interaction check, the best model was the following:

$$\ln \frac{P}{(1-P)} = \beta_0 + \beta_1 T_{\min} + \beta_2 \text{Pr3} + \beta_3 \text{Region} + \varepsilon_{\text{Farm}} \quad (1)$$

where P is the probability of having an *E. coli*-positive sample, β_i are constants, T_{\min} is the minimum temperature of the sampling day in °C, Pr3 is the total precipitation amount of 3 days before the sampling day, $\beta_3 \text{Region}$ is the dummy variable for region, and $\varepsilon_{\text{Farm}}$ is the random effect for farm. The β estimates are available in Table 4.

The estimated coefficients showed that the odds of having *E. coli*-positive samples on LGVs increased when temperature and cumulative precipitation amount increased

by one measurement unit (°C or mm). For a 1°C increase in daily minimum temperature, the odds of having *E. coli*-positive samples on LGVs increased by a factor of 1.48 (95% CI, 1.27 to 1.73) (Table 4). For a 1-mm increase in 3 days' cumulative precipitation, the odds of having *E. coli*-positive samples on LGVs increased by a factor of 1.02 (95% CI, 1.01 to 1.03) (Table 4). T_{\min} and Pr3 had a statistically significant relation with *E. coli* presence.

For management variables, both univariate analysis results and Wald test results showed that all variables in this analysis had a P value of < 0.25 , except for NonAOrganic (Table 3). DrinkingW and Flood irrigation were not used for any samples after the missing values were removed. DrinkingW and Flood were, therefore, dropped from multivariate analysis. All other management variables were selected as candidates for the multivariate model ($n = 520$). The variance inflation factor results showed that collinearity exists between Region and RainW, SurfaceW, and Spray. This meant that the Region effect in the *E. coli* presence model may be a proxy for these management practices. Although we would have preferred to keep Region in the model to explain the differences in the sampling effort and detection limits among regions, it would be a huge sacrifice to avoid so many relevant variables. Therefore, Region was not included in the selection.

Upon completion of the univariate analyses, all variables were selected for the multivariable analysis. With backward selection and the Wald test for importance, the following model was the best, with Farm as a random effect term:

$$\ln \frac{P}{(1-P)} = \beta_0 + \beta_1 \text{RainW} + \beta_2 \text{Spray} + \varepsilon_{\text{Farm}} \quad (2)$$

where P is the probability of having an *E. coli*-positive sample, β_i are constants, $\beta_1 \text{RainW}$ is the dummy variable for rainwater, $\beta_2 \text{Spray}$ is the dummy variable for spray irrigation, and $\varepsilon_{\text{Farm}}$ is the random effect for farm. Spray irrigation showed a protective effect (Table 4). Using rainwater for irrigation versus other irrigation water types increased the odds of having *E. coli* contamination on LGVs by a factor of 12.14 (Table 4).

The backward selection was performed once more, combining significant ($P < 0.25$) climate and management variables to predict *E. coli* presence on LGVs. Region was again not included in the backward selection. The combined model has the following form:

$$\ln \frac{P}{(1-P)} = \beta_0 + \beta_1 T_{\min} + \beta_2 \text{SurfaceW} + \beta_3 \text{Inorganic} + \varepsilon_{\text{Farm}} \quad (3)$$

where P is the probability of having an *E. coli*-positive sample, T_{\min} is the minimum temperature in °C, β_i are constants, $\beta_2 \text{SurfaceW}$ is the dummy variable for surface water, $\beta_3 \text{Inorganic}$ is the dummy variable for inorganic fertilizer, and $\varepsilon_{\text{Farm}}$ is the random effect for farm. Because Region had collinearity with SurfaceW, the joint model was run again, using the variable Region instead of SurfaceW. Compared with equation 3, a significantly lower AIC was found in that model with the following form:

TABLE 3. Variables associated with the presence and concentration of *E. coli*^a

Variable	Baseline or unit	<i>E. coli</i> presence				<i>E. coli</i> concn		
		Coeff	SE	OR (95% CI)	<i>P</i> value	Coeff	SE	<i>P</i> value
<i>T</i> _{avg}	°C	0.14	0.03	1.15 (1.09, 1.21)	0.000	0.10	0.02	0.000
<i>T</i> _{avg3}	°C	0.10	0.03	1.11 (1.05, 1.17)	0.000	0.09	0.02	0.000
<i>T</i> _{avg7}	°C	0.10	0.03	1.11 (1.05, 1.17)	0.000	0.05	0.02	0.032
<i>T</i> _{max}	°C	0.07	0.02	1.07 (1.03, 1.12)	0.002	0.07	0.01	0.000
<i>T</i> _{max3}	°C	0.08	0.02	1.09 (1.04, 1.14)	0.000	0.10	0.02	0.000
<i>T</i> _{max7}	°C	0.07	0.02	1.08 (1.03, 1.13)	0.001	0.04	0.02	0.020
<i>T</i> _{min}	°C	0.18	0.03	1.19 (1.13, 1.27)	0.000	0.09	0.03	0.002
<i>T</i> _{min3}	°C	0.09	0.03	1.09 (1.04, 1.15)	0.001	0.03	0.02	0.090
<i>T</i> _{min7}	°C	0.09	0.03	1.09 (1.03, 1.16)	0.004	0.04	0.03	0.174
Pr3	mm	0.01	0.01	1.01 (1.00, 1.03)	0.073	-0.01	0.01	0.291
Pr7	mm	0.01	0.00	1.01 (1.00, 1.01)	0.068	-0.01	0.00	0.026
Rainwater	Yes/no	0.68	0.28	1.97 (1.13, 3.47)	0.017	-0.57	0.233	0.019
Surface water	Yes/no	-0.63	0.29	0.53 (0.30, 0.94)	0.031	0.25	0.26	0.337
Groundwater	Yes/no	1.28	0.69	3.60 (0.78, 12.61)	0.061	1.34	0.45	0.004
Spray irrigation	Yes/no	-1.28	0.69	0.28 (0.08, 1.28)	0.061	-0.46	0.25	0.072
Drip irrigation	Yes/no	1.28	0.69	3.60 (0.78, 12.61)	0.061	0.53	0.32	0.104
Composted manure (derived)	Yes/no	-0.55	0.31	0.57 (0.31, 1.06)	0.074	1.01	0.17	0.000
Inorganic fertilizer	Yes/no	0.66	0.32	1.94 (1.02, 3.61)	0.039	-1.29	0.15	0.000
Nonanimal organic	Yes/no	-0.07	0.49	0.93 (0.31, 2.25)	0.882	0.46	0.38	0.229
Farm animal presence	Yes/no	1.162	0.31	3.19 (1.72, 5.93)	0.000	-0.63	0.24	0.011
Toilet distance (100–200 m)	0–100 m	0.63	0.35	1.89 (0.93, 3.76)	0.072	1.27	0.17	0.000
Toilet distance (200–500 m)	0–100 m	-0.54	0.48	0.58 (0.21, 1.42)	0.263	1.84	0.24	0.000
Region (Brazil)	Belgium	0.58	0.38	1.78 (0.83, 3.76)	0.132	1.77	0.14	0.000
Region (Egypt)	Belgium	2.17	0.53	8.75 (3.11, 25.56)	0.00	0.90	0.16	0.000
Region (Norway)	Belgium	-1.52	0.63	0.22 (0.05, 0.66)	0.016	-0.03	0.26	0.897
Region (Spain)	Belgium	-0.89	0.38	0.41 (0.19, 0.86)	0.020	1.02	0.15	0.000

^a Coeff, regression coefficient; OR (95% CI), odds ratio (95% confidence interval). Only variables with *P* < 0.25 are shown.

$$\ln \frac{P}{(1 - P)} = \beta_0 + \beta_1 T_{\min} + \beta_{2, \text{Region}} + \beta_{3, \text{Inorganic}} + \varepsilon_{\text{Farm}} \tag{4}$$

where *P* is the probability of having an *E. coli*-positive samples, *T*_{min} is the minimum temperature in °C, β_{*i*} are constants, β_{2,Region} is the dummy variable for regions, β_{3,Inorganic} is the dummy variable for inorganic fertilizer, and ε_{Farm} is the random effect for farm. The odds ratio (Table 4) shows that a 1°C increase in minimum temperature increases the odds of *E. coli* contamination on LGVs by a factor of 1.47, assuming that management remains the same. In both equations 3 and 4, *T*_{min} had a significant influence on estimating *E. coli* contamination. Although other variables were not significant (*P* < 0.05), they improved the model fit significantly according to the AIC test. Therefore, management variables were also included in the final mixed model. Although equation 4 was the best model fit for this meta-analysis, we also presented equation 3 because it provided more understanding of the variable Region and would be more useful than equation 4 for future studies on a prediction model in a specific region. Cross-validation analysis of equation 4 showed that the mean area under the curves was 88% (range 79 to 95%) (Fig. 3). So, the model had a good predictive value.

Because *T*_{min} was the only significant variable in the joint model, we concluded that, although climate and management variables together influence *E. coli* presence

on LGVs, *T*_{min} had stronger influence on *E. coli* presence than management variables.

***E. coli* concentration on LGVs.** Visual inspection of normality plots of standardized residuals and the homoscedasticity test (residual plot of standardized residuals with fitted values) did not reveal any obvious deviation from normality. The random effect Farm was dropped because its effect was negligible in the smaller data set (59 samples from 16 farms). To assess the effect of the climate variables and management variables on *E. coli* concentration, *E. coli*-positive samples were fitted to the linear regression model. Results of climate and management variables are presented in this section.

Based on univariate analysis, variables *T*_{avg}, *T*_{avg3}, *T*_{avg7}, *T*_{max}, *T*_{max3} (highest *R*² 0.38), *T*_{max7}, *T*_{min}, and Pr7 were significant (*P* < 0.25) (Table 3). Because the temperature variables were not independent, only one variable should be selected. So *T*_{max3}, Pr7, and Region were chosen after the univariate analysis. With backward selection, the Wald test for importance of the categorical variable, the *F* test to select the best model fit, and the interaction check, the final model had the following form:

$$Y = \beta_0 + \beta_1 T_{\max 3} + \beta_{2, \text{Region}} \tag{5}$$

where *Y* is the *E. coli* concentration in log CFU/g, β_{*i*} are constants, *T*_{max3} is the maximum temperature on 3 days before the sampling day in °C, and β_{2,Region} is the dummy

TABLE 4. Final mixed-effect logistic regression models to estimate *Escherichia coli* absence or presence on lettuce and spinach

	β_i estimate	Odds ratio (95% CI)	<i>P</i> value
Climate model (<i>n</i> = 559):			
Variance (SD) 3.06 (1.75) for random effect Farm			
β 0	-8.339	0.00 (0.00–0.01)	<0.000
β 1 (T_{\min})	0.391	1.48 (1.27–1.73)	<0.000
β 2 (Pr7)	0.019	1.02 (1.01–1.03)	0.003
β 3,Belgium	Reference	Reference	Reference
β 3,Brazil	-1.203	0.30 (0.02–3.70)	0.348
β 3,Egypt	3.092	2.20 (1.64–295.47)	0.020
β 3,Norway	-2.088	0.12 (0.00–3.61)	0.225
β 3,Spain	0.327	1.39 (0.06–29.69)	0.834
Management model (<i>n</i> = 520):			
Variance (SD) 1.33 (1.15) for random effect Farm			
β 0	-1.184	0.31 (0.02–4.28)	0.379
β 1,RainW	2.497	12.14 (1.80–81.99)	0.010
β 2,Spray	-3.315	0.04 (0.00–0.86)	0.040
Joint model (<i>n</i> = 520):			
Variance (SD) 1.81 (1.35) for random effect Farm			
β 0	-9.016	0.00 (0.00–0.00)	<0.000
β 1 (T_{\min})	0.401	1.51 (1.29–1.76)	<0.000
β 2,SurfaceW	0.844	2.33 (0.23–23.78)	0.476
β 3,Inorganic	1.334	3.80 (0.50–29.01)	0.198
β 0	-9.433	0.00 (0.00–0.00)	<0.000
β 1 (T_{\min})	0.388	1.47 (1.26–1.72)	<0.000
β 2,Belgium	Reference	Reference	Reference
β 3,Brazil	0.664	1.94 (0.14–27.11)	0.621
β 4,Norway	-17.589	0.00 (0.00–0.00)	0.997
β 5,Spain	2.273	9.71 (0.00–168.09)	0.118
β 6,Inorganic	2.714	1.51 (0.92–247.23)	0.057

variable for region. Maximum temperature was significantly and positively correlated with *E. coli* concentrations (Table 5). This model gave a root mean square error of 0.38 and a mean absolute error of 0.30, indicating high accuracy. The

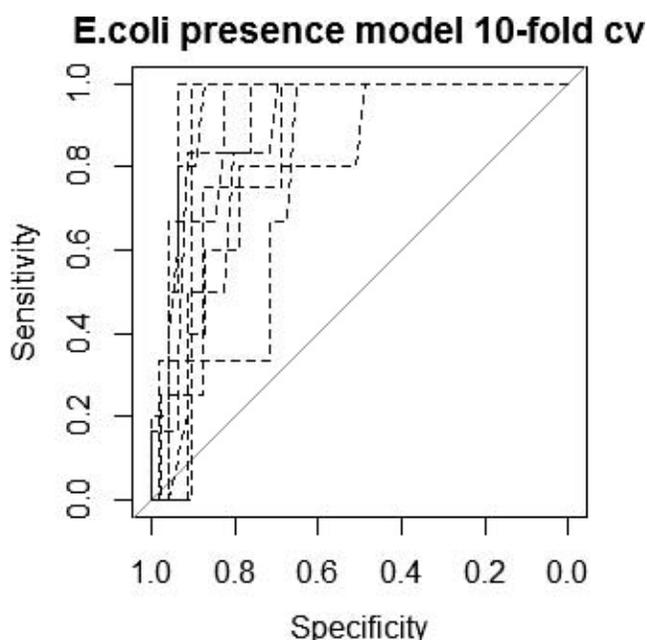


FIGURE 3. Receiver operating characteristic curves for each of the 10-fold cross-validations.

adjusted R^2 was 0.75 with an associated P value of <0.00. Figure 4 graphically shows the regression of $T_{\max}3$ and *E. coli* concentration for each region. This regression gave an adjusted R^2 of 0.38 and an associated P value of 0.00 (Fig. 4), indicating that *E. coli* concentration had a significant positive correlation with $T_{\max}3$.

In the management model, variables FarmA and ToiletD were excluded in the univariate analysis, because the missing values in these two variables were 17 and 25% of the total samples. Too many samples would have to be excluded for the regression analysis if these two variables were included in the data set. Three samples from Norway were excluded from the data set due to missing values in irrigation water type. Upon completion of the univariate analysis, all variables had a P value <0.25 except for SurfaceW and Flood (Table 3). The variance inflation factor results showed that Region had multicollinearity with RainW, Manure, Inorganic, and Spray. This meant that the region effect may be a proxy for these management practices. Therefore, Region was not included in the selection. With backward selection, the F -test, and the interaction check, the final model (*n* = 56) had the following form:

$$Y = \beta_0 + \beta_{1,\text{RainW}} + \beta_{2,\text{Spray}} + \beta_{3,\text{GroundW}} + \beta_{4,\text{Inorganic}} \quad (6)$$

where Y is the *E. coli* concentration in log CFU/g, β_i are constants, $\beta_{1,\text{RainW}}$ is the dummy variable for rainwater, $\beta_{2,\text{Spray}}$ is the dummy variable for spray irrigation, $\beta_{3,\text{GroundW}}$

TABLE 5. Parameter estimates, SE, and probability in the final linear regression models to estimate *Escherichia coli* (log CFU/g) on lettuce and spinach

	β_i estimate	SE	P value
Climate model ($n = 59$, adjusted $R^2 = 0.75$)			
β_0 (intercept)	0.258	0.432	0.553
β_1 ($T_{\max 3}$)	0.041	0.020	0.043
β_2 ,Belgium	Reference	Reference	Reference
β_2 ,Brazil	1.432	0.215	<0.000
β_2 ,Egypt	0.765	0.169	<0.000
β_2 ,Norway	0.287	0.294	0.333
β_2 ,Spain	1.090	0.151	<0.000
Management model ($n = 56$, adjusted $R^2 = 0.67$)			
β_0 (Intercept)	1.561	0.236	<0.000
β_1 ,RainW	0.471	0.185	0.014
β_2 ,Spray	0.594	0.195	0.004
β_3 ,GroundW	1.663	0.357	<0.000
β_4 ,Inorganic	-1.544	0.169	<0.000
Joint model ($n = 56$, adjusted $R^2 = 0.69$)			
β_0	-0.502	0.589	0.399
β_1 ($T_{\max 3}$)	0.083	0.016	<0.000
β_2 ,Manure	0.613	0.249	0.017
β_3 ,Inorganic	-0.557	0.266	0.041
β_4 ,Spray	0.391	0.168	0.024

is the dummy variable for groundwater, and $\beta_{4,\text{Inorganic}}$ is the dummy variable for inorganic fertilizer. The influences of rainwater, spray irrigation, groundwater, and inorganic fertilizer were significantly different from all other irrigation water types and fertilizer types (equation 6). This model gave a root mean square error of 0.44 and a mean absolute error of 0.32, indicating high accuracy. Adjusted R^2 was 0.67, with $P < 0.00$. The model parameter coefficients are given in Table 5. Inorganic fertilizer gave a protective effect compared with other fertilizer types (Table 5). *E. coli* concentrations were positively related with RainW, Spray, and GroundW.

Climate variables and management variables were then combined to predict *E. coli* concentrations. All variables selected for the multivariable analysis in the previous two models were combined for the stepwise selection. According

to this joint model, the *E. coli* concentrations on LGVs were estimated with the following equation:

$$Y = \beta_0 + \beta_1 T_{\max 3} + \beta_{2,\text{Manure}} + \beta_{3,\text{Inorganic}} + \epsilon_{4,\text{Spray}} \quad (7)$$

where Y is the *E. coli* concentration in log CFU/g, β_i are constants, $\beta_{2,\text{Manure}}$ is the dummy variable for composted manure (derived), $\beta_{3,\text{Inorganic}}$ is the dummy variable for inorganic fertilizer, and $\beta_{4,\text{Spray}}$ is the dummy variable for spray irrigation. In the joint model, $T_{\max 3}$, Manure, Inorganic, and Spray were selected to estimate *E. coli* concentrations. This model gave a root mean square error of 0.42 and a mean absolute error of 0.31, indicating high accuracy. Adjusted R^2 was 0.69, with a P value of <0.00 . Inorganic fertilizer again had a protective effect compared with other fertilizer types (Table 5). *E. coli* concentrations were positively related to higher maximum temperature and use of composted manure (derived) and spray irrigation.

This model had a lower adjusted R^2 than equation 5. Multicollinearity was present among Region, Manure, Inorganic, and Spray. Equation 5 was the best model fit to estimate *E. coli* concentrations on LGVs based on this meta-regression analysis; the variable Region was masking three other management variables. However, equation 7 was useful for future prediction modeling in a specific region.

We concluded that both climate and management variables influence *E. coli* concentrations significantly. T_{\max} had the strongest influence (adjusted R^2 of 0.38) among all variables.

DISCUSSION

Our study identified a combination of statistically significant variables that best explained observed variation in *E. coli* presence and concentration. A two-step approach

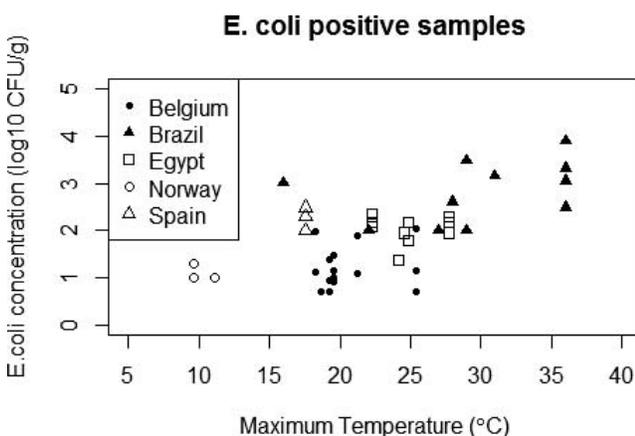


FIGURE 4. Correlation of *E. coli* concentrations with daily average temperature for each region.

was taken; first, to study the relationship between climate and management variables and *E. coli* presence and, second, to study the relationship between climate and management variables and *E. coli* concentrations. The climate variables T_{\min} and Pr3 were important for presence; T_{\max} 3 was important for concentration. The management variables RainW and Spray were important for both presence and concentration. In addition to these two variables, GroundW and Inorganic were also important in estimating *E. coli* concentration. When climate and management variables were combined, both temperature and management practices influenced *E. coli* presence and concentration together. Temperature had a stronger influence (shown by significant parameter estimate and highest R^2) than did management practices on *E. coli* presence and concentration on LGVs. Inorganic fertilizer had a positive parameter estimate in the *E. coli* presence model. This differed from our expectation because inorganic fertilizer should be sterile. However, the parameter estimate was not significant, meaning that, although inorganic fertilizer significantly explained data variation, the parameter estimate was not representative for indicating directional changes in *E. coli* contamination. One hypothesis is that the counterintuitive risk factor of inorganic fertilizer is due to the nature of the data set. The farms applying inorganic fertilizer happened to have a higher percentage of positive samples than the ones using composted manure. Almost all of these positive samples were from the same farm in Belgium. The contamination could be caused by poor hygiene conditions on the farm or by other factors that cannot be explained by our data set. Another hypothesis is that *E. coli* takes advantage of an easily metabolized nutrient source (10) and, thereby, increases its growth rate in the soil. This process is known to be favored by high soil temperatures (10). Composted manure contains comparatively more nutrients that are not readily available (e.g., the nutrients are locked in complex structures such as cellulose and lignin) (10), and, thus, inorganic fertilizer may promote *E. coli* survival compared with composted manure. Nonsignificant variables should not be used to predict future *E. coli* presence, whereas we think that the *E. coli* concentration model with an adjusted R^2 of 0.75 is applicable for scenario analysis.

In this study, log-transformed *E. coli* concentration data were used to transform a multiplier effect (nonlinear relationship) to an additive effect (linear relationship). The log-transformed model still indicates the same directional trend, but it uses a linear relationship instead of a nonlinear relationship. The log-transformed regression coefficients should be interpreted as the ratio of the expected median of the untransformed data instead of the expected mean. We want to emphasize this difference for interpretation of the untransformed data.

The current study has some minor limitations. First, the LGV samples used in the modeling may not represent the situation on large farms due to the high uncertainty of the measurements. Second, the length of the measurement period was short for observation of changing management behavior due to climate change. However, a broad temperature range in the data was able to cover that. Third, the precipitation data used may not fully represent the

precipitation over the farms. Although the weather station nearest to each LGV farm was chosen, the distance still ranged up to 50 km away from the farm. Fourth, little precipitation was observed on the sampling days, and we have, thus far, not looked at rain amounts over more than 7 days. Medina-Martínez et al. (24) found that lettuce was contaminated after a flood in Spain and that bacteria disappeared after 7 days due to UV radiation.

Our results for *E. coli* presence and concentration are slightly different than those of other studies. Park et al. (32, 33) found that *E. coli* presence was determined by farm management (manure application) and by the region (state). Once a contamination event had occurred, the count of *E. coli* on spinach was determined by weather only (mean precipitation of the past 29 days and mean maximum temperature over the past 9 days) (33). Strawn et al. (38) also found that manure application, irrigation water, temperature, and precipitation increase the risk of pathogen contamination. Pagadala et al. (30) stated with their univariate analysis that the irrigation water source was a significant variable for all indicator bacteria on tomatoes. And Region was a significant variable for total coliforms levels. In line with our results, Castro-Ibáñez et al. (5) also found that coliform counts were positively related to temperature. We found that the influence on *E. coli* presence and concentration was dominated by temperature variables. Differences with other studies can be explained by different model objectives and experimental setup. For instance, Park et al. (33) included cumulative weather effects and considered the survival of *E. coli* between contamination and sampling. We did not take the dynamics in the period between contamination and sampling into consideration due to lack of information on UV, which is the most important factor for bacteria survival.

Increasing temperature due to climate change may increase the presence and concentration of *E. coli* in the future, although the actual increase in *E. coli* concentration with a 1°C increase is low. A direct positive effect of temperature on *E. coli* presence and contamination is, however, not expected, given the general observed negative relation between temperature and environmental persistence (9). From a microbiology perspective, *E. coli* is expected to have reduced survival with increasing temperature in soil, manure, or water (9, 22, 28, 44). Temperature may affect environmental factors such as wildlife intrusion, insect activity, and irrigation frequency, which, in turn, directly affect *E. coli* presence and concentration. These environmental factors should be included in future sampling and analyses to cover more potential contamination pathways. Figure 5 illustrates the relationships found in this study for *E. coli* concentration, climate, and management variables. Both climate and management variables had a positive relationship with *E. coli* presence and concentration. In addition, climate may indirectly relate to management variables and then influence the *E. coli* concentration. However, management practices vary in different temperature zones, often due to different socioeconomic conditions in these zones. Fertilizer type and irrigation water type may also be influenced by the gross domestic product, which happens to be lower in those countries in this study that have higher temperatures. We

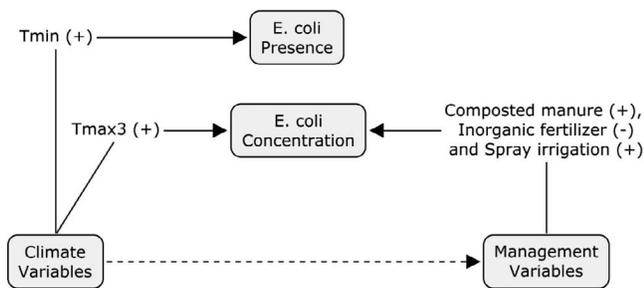


FIGURE 5. Diagram shows statistically significant effects of climate and management variables on *E. coli* presence and concentration. Solid arrows indicate quantified significant effects from this study. Dashed arrows indicate unquantified potential associations. Plus (minus) signs indicate positive (negative) relationships between variables and *E. coli* (presence and concentration).

do not have sufficient data to reach a conclusion about this. Further research is needed to prove that climate variables influence *E. coli* concentrations strongly via management practices.

Meta-analysis provides opportunities to perform statistical analysis with limited positive samples from different regions. Including the regions with different climate conditions and *E. coli* concentrations enlarges the temperature range for the analysis. Although Region appears to be an important variable for *E. coli* contamination in this study, the meta-analysis allows a generic model to identify the statistically significant variables for *E. coli* contamination throughout the regions. From this study we have learned several lessons for future meta-analysis. (i) Experimental design in meta-analysis has to be standardized as much as possible. In our study all samples were taken according to the HAS developed in the Veg-i-Trade project. Each partner used the same sampling method and questionnaire. However, the sampling and analysis in different regions were not designed for a meta-analysis from the start. Even though the sampling was performed according to the same scheme, we had to work with two produce types (lettuce and spinach) and, more importantly, different detection limits. (ii) Meta-analysis has more strict sampling requirements than individual studies. The combined data sets need to have a representation that is as balanced as possible for all levels of management variables. If a study is set up for a meta-analysis, then every management variable would ideally have the same amount of samples in each region. The use of specific management variables in studies outside the Veg-i-Trade project made it impossible to include these other studies in this meta-analysis. This highlights the need for a coordinated future international sampling collection effort and for development of a study design and reporting standards to assure that the data collected and the results reported in different regions are comparable and could be used in subsequent meta-analyses. (iii) Some of the regional differences are not defined in the meta-analysis, and they should not be ignored. The differences in joint models with and without Region show that Region explained not only the regional variations in many

management practices but also additional regional differences. We recommend that the model boundary be enlarged in future studies by including these additional differences (e.g., variation in detection limits, experimental material and equipment, local hygiene, socioeconomic development levels, presence of wildlife intrusion, insect activity, irrigation frequency, soil type, and slope and topography) to complete the system analysis of LGVs safety.

This is the first large-scale meta-analysis of *E. coli* presence and concentration on LGVs. This study combined climate and management variables from 23 farms and included 562 samples. The current study sets the baseline for future monitoring of climate and contamination relationships. The significant climate and management variables (temperature, types of fertilizer and irrigation water, and irrigation methods) determined in this study should be considered systematically in fresh produce safety studies in the future.

ACKNOWLEDGMENTS

The authors thank Dr. Evert-Jan Bakker for his statistical advice and help in this study, Dr. Siele Ceuppens for collecting data and many valuable discussions during the data analysis, and an anonymous contributor for the constructive discussion and valuable statistical advice. This research is funded by the European Union FP7 Veg-i-Trade project (Grant agreement no. 244994).

REFERENCES

1. Akaike, H. 1974. A new look at the statistical model identification. *IEEE (Inst. Electr. Electron. Eng.) Trans. Automat. Control* 19:716–723.
2. Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. Fitting linear mixed-effects models using lme4. Available at: <http://arxiv.org/abs/1406.5823>. Accessed 3 December 2014.
3. Bendel, R. B., and A. A. Afifi. 1977. Comparison of stopping rules in forward “stepwise” regression. *J. Am. Stat. Assoc.* 72:46–53.
4. Camuffo, D., and C. Bertolin. 2012. The earliest temperature observations in the world: the Medici Network (1654–1670). *Clim. Change* 111:335–363.
5. Castro-Ibañez, I., M. Gil, J. Tudela, and A. Allende. 2014. Microbial safety considerations of flooding in primary production of leafy greens: a case study. *Food Res. Int.* 68:62–69.
6. Ceuppens, S., C. T. Hessel, R. de Quadros Rodrigues, S. Bartz, E. C. Tondo, and M. Uyttendaele. 2014. Microbiological quality and safety assessment of lettuce production in Brazil. *Int. J. Food Microbiol.* 181:67–76.
7. D’Souza, R. M., N. G. Becker, G. Hall, and K. B. A. Moodie. 2004. Does ambient temperature affect foodborne disease? *Epidemiology* 15:86–92.
8. Food and Agriculture Organization of the United Nations, World Health Organization. 2008. Microbiological hazards in fresh fruits and vegetables. Meeting report. Microbiological risk assessment series. Food and Agriculture Organization of the United Nations, World Health Organization, Geneva.
9. Franz, E., J. Schijven, A. M. de Roda Husman, and H. Blaak. 2014. Meta-regression analysis of commensal and pathogenic *Escherichia coli* survival in soil and water. *Environ. Sci. Technol.* 48:6763–6771.
10. Franz, E., A. V. Semenov, A. J. Termorshuizen, O. J. de Vos, J. G. Bokhorst, and A. H. C. van Bruggen. 2008. Manure-amended soil characteristics affecting the survival of *E. coli* O157:H7 in 36 Dutch soils. *Environ. Microbiol.* 10:313–327.
11. Friesema, I., G. Sigmundsdottir, K. van der Zwaluw, A. Heuvelink, B. Schimmer, C. de Jager, B. Rump, H. Briem, H. Hardardottir, A. Atladottir, E. Gudmundsdottir, and W. van Pelt. 2008. An international outbreak of Shiga toxin-producing *Escherichia coli*

- O157 infection due to lettuce, September–October 2007. *Euro Surveill.* 13:19065.
12. Gajraj, R., S. Pooransingh, J. I. Hawker, and B. Olowokure. 2012. Multiple outbreaks of *Salmonella* Braenderup associated with consumption of iceberg lettuce. *Int. J. Environ. Health Res.* 22:150–155.
 13. Hall, G. V., R. M. D'Souza, and M. D. Kirk. 2002. Foodborne disease in the new millennium: out of the frying pan and into the fire? *Med. J. Aust.* 177:614–618.
 14. Holvoet, K., L. Jacxsens, I. Sumpers, and M. Uyttendaele. 2011. Horticultural assessment scheme: insight in prevalence and distribution of microbial contamination to evaluate water management in fresh produce processing industry. Abstr. Annu. Meet. IAFP 2011. International Association for Food Protection, Des Moines, IA.
 15. Holvoet, K., I. Sumpers, B. Callens, J. Dewulf, and M. Uyttendaele. 2013. Moderate prevalence of antimicrobial resistance in *Escherichia coli* isolates from lettuce, irrigation water, and soil. *Appl. Environ. Microbiol.* 79:6677–6683.
 16. Hosmer, D. W., Jr., and S. Lemeshow. 2004. Applied logistic regression. John Wiley & Sons, New York.
 17. Islam, M., M. P. Doyle, S. C. Phatak, P. Millner, and X. P. Jiang. 2004. Persistence of enterohemorrhagic *Escherichia coli* O157:H7 in soil and on leaf lettuce and parsley grown in fields treated with contaminated manure composts or irrigation water. *J. Food Prot.* 67:1365–1370.
 18. Islam, M., J. Morgan, M. P. Doyle, S. C. Phatak, P. Millner, and X. Jiang. 2004. Persistence of *Salmonella enterica* serovar Typhimurium on lettuce and parsley and in soils on which they were grown in fields treated with contaminated manure composts or irrigation water. *Foodborne Pathog. Dis.* 1:27–35.
 19. Jacxsens, L., P. A. Luning, J. G. A. J. van der Vorst, F. Devlieghere, R. Leemans, and M. Uyttendaele. 2010. Simulation modelling and risk assessment as tools to identify the impact of climate change on microbiological food safety—the case study of fresh produce supply chain. *Food Res. Int.* 43:1925–1935.
 20. Kirezieva, K., L. Jacxsens, M. A. van Boekel, and P. A. Luning. 2015. Towards strategies to adapt to pressures on safety of fresh produce due to climate change. *Food Res. Int.* 68:94–107.
 21. Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2014. lmerTest: tests in linear mixed effects models. R package version 2.0-20. Available at: <http://CRAN.R-project.org/package=lmerTest>. Accessed 3 December 2014.
 22. Liu, C., N. Hofstra, and E. Franz. 2013. Impacts of climate change on the microbial safety of pre-harvest leafy green vegetables as indicated by *Escherichia coli* O157 and *Salmonella* spp. *Int. J. Food Microbiol.* 163:119–128.
 23. Manning, L., and R. N. Baines. 2004. Effective management of food safety and quality. *Br. Food J.* 106:598–606.
 24. Medina-Martínez, M. S., A. Allende, G. G. Barberá, and M. I. Gil. 2015. Climatic variations influence the dynamic of epiphyte bacteria of baby lettuce. *Food Res. Int.* 68:54–61.
 25. Meehl, G. A., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, S. C. B. Raper, I. G. Watterson, A. J. Weaver, and Z.-C. Zhao. 2007. Global climate projections, chap. 10. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (ed.), *Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and New York.
 26. Mickey, R. M., and S. Greenland. 1989. The impact of confounder selection criteria on effect estimation. *Am. J. Epidemiol.* 129:125–137.
 27. Miraglia, M., H. J. P. Marvin, G. A. Kleter, P. Battilani, C. Brera, E. Coni, F. Cubadda, L. Croci, B. De Santis, S. Dekkers, L. Filippi, R. W. A. Hutjes, M. Y. Noordam, M. Pisante, G. Piva, A. Prandini, L. Toti, G. J. van den Born, and A. Vespermann. 2009. Climate change and food safety: an emerging issue with special focus on Europe. *Food Chem. Toxicol.* 47:1009–1021.
 28. Mukherjee, A., S. Cho, J. Scheffel, S. Jawahir, K. Smith, and F. Diez-Gonzalez. 2006. Soil survival of *Escherichia coli* O157:H7 acquired by a child from garden soil recently fertilized with cattle manure. *J. Appl. Microbiol.* 101:429–436.
 29. Mukherjee, A., D. Speh, and F. Diez-Gonzalez. 2007. Association of farm management practices with risk of *Escherichia coli* contamination in pre-harvest produce grown in Minnesota and Wisconsin. *Int. J. Food Microbiol.* 120:296–302.
 30. Pagadala, S., S. C. Marine, S. A. Micallef, F. Wang, D. M. Pahl, M. V. Melendez, W. L. Kline, R. A. Oni, C. S. Walsh, K. L. Everts, and R. L. Buchanan. 2015. Assessment of region, farming system, irrigation source and sampling time as food safety risk factors for tomatoes. *Int. J. Food Microbiol.* 196:98–108.
 31. Park, S., S. Navratil, A. Gregory, A. Bauer, I. Srinath, M. Jun, B. Szonyi, K. Nightingale, J. Anciso, and R. Ivanek. 2013. Generic *Escherichia coli* contamination of spinach at the preharvest stage: effects of farm management and environmental factors. *Appl. Environ. Microbiol.* 79:4347–4358.
 32. Park, S., S. Navratil, A. Gregory, A. Bauer, I. Srinath, B. Szonyi, K. Nightingale, J. Anciso, M. Jun, D. Han, S. Lawhon, and R. Ivanek. 2014. Farm management, environment, and weather factors jointly affect the probability of spinach contamination by generic *Escherichia coli* at the preharvest stage. *Appl. Environ. Microbiol.* 80:2504–2515.
 33. Park, S., S. Navratil, A. Gregory, A. Bauer, I. Srinath, B. Szonyi, K. Nightingale, J. Anciso, M. Jun, D. Han, S. Lawhon, and R. Ivanek. 2015. Multifactorial effects of ambient temperature, precipitation, farm management, and environmental factors determine the level of generic *Escherichia coli* contamination on preharvested spinach. *Appl. Environ. Microbiol.* 81:2635–2650.
 34. Semenza, J. C., and B. Menne. 2009. Climate change and infectious diseases in Europe. *Lancet Infect. Dis.* 9:365–375.
 35. Söderström, A., P. Österberg, A. Lindqvist, B. Jönsson, A. Lindberg, S. B. Ulander, C. Welinder-Olsson, S. Löfdahl, B. Kaijser, B. De Jong, S. Kühlmann-Berenzon, S. Boqvist, E. Eriksson, E. Szanto, S. Andersson, G. Allestam, I. Hedenström, L. L. Muller, and Y. Andersson. 2008. A large *Escherichia coli* O157 outbreak in Sweden associated with locally produced lettuce. *Foodborne Pathog. Dis.* 5:339–349.
 36. Steele, M., and J. Odumeru. 2004. Irrigation water as source of foodborne pathogens on fruit and vegetables. *J. Food Prot.* 67:2839–2849.
 37. Stocker, T. F., D. Qin, G.-K. Plattner, L. V. Alexander, S. K. Allen, N. L. Bindoff, F.-M. Bréon, J. A. Church, U. Cubasch, S. Emori, P. Forster, P. Friedlingstein, N. Gillett, J. M. Gregory, D. L. Hartmann, E. Jansen, B. Kirtman, R. Knutti, K. K. Kumar, P. Lemke, J. Marotzke, V. Masson-Delmotte, G. A. Meehl, I. I. Mokhov, S. Piao, V. Ramaswamy, D. Randall, M. Rhein, M. Rojas, C. Sabine, D. Shindell, L. D. Talley, D. G. Vaughan, and S.-P. Xie. 2013. Technical summary, p. 33–115. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (ed.), *Climate change 2013: the physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge and New York.
 38. Strawn, L. K., E. D. Fortes, E. A. Bihn, K. K. Nightingale, Y. T. Gröhn, R. W. Worobo, M. Wiedmann, and P. W. Bergholz. 2013. Landscape and meteorological factors affecting prevalence of three food-borne pathogens in fruit and vegetable farms. *Appl. Environ. Microbiol.* 79:588–600.
 39. Takkinen, J., U. Nakari, T. Johansson, T. Niskanen, A. Siitonen, and M. Kuusi. 2005. A nationwide outbreak of multiresistant *Salmonella* Typhimurium in Finland due to contaminated lettuce from Spain, May 2005. *Euro Surveill.* 10:2734.
 40. Tirado, M. C., R. Clarke, L. A. Jaykus, A. McQuatters-Gollop, and J. M. Frank. 2010. Climate change and food safety: a review. *Food Res. Int.* 43:1745–1765.
 41. U.S. Food and Drug Administration, U.S. Department of Agriculture, Centers for Disease Control and Prevention. 1998. Guidance for industry: guide to minimize microbial food safety hazards for fresh fruits and vegetables. U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD.

42. Uyttendaele, M., A. A. Moneim, S. Ceuppens, and F. El-Tahan. 2014. Microbiological safety of strawberries and lettuce for domestic consumption in Egypt. *J. Food Process. Technol.* 5:308.
43. Van Staveren, W. A., P. Deurenberg, J. Burema, L. C. De Groot, and J. Hautvast. 1986. Seasonal variation in food intake, pattern of physical activity and change in body weight in a group of young adult Dutch women consuming self-selected diets. *Int. J. Obes.* 10:133.
44. Wang, G., and M. P. Doyle. 1998. Survival of enterohemorrhagic *Escherichia coli* O157:H7 in water. *J. Food Prot.* 61:662–667.
45. Zhang, Y., P. Bi, and J. E. Hiller. 2010. Climate variations and *Salmonella* infection in Australian subtropical and tropical regions. *Sci. Total Environ.* 408:524–530.
46. Ziegler, R., H. Wilcox, T. Mason, J. Bill, and P. Virgo. 1987. Seasonal variation in intake of carotenoids and vegetables and fruits among white men in New Jersey. *Am. J. Clin. Nutr.* 45:107–114.