

# AN INVESTIGATION OF THE SEASONAL RELATIONSHIPS BETWEEN METEOROLOGICAL FACTORS, WATER QUALITY, AND SPORADIC CASES OF LEGIONNAIRES' DISEASE IN WASHINGTON, DC

ALEXANDER KIRPICH<sup>1</sup>, ALEKSANDR SHISHKIN<sup>1</sup>, PEMA LHEWA<sup>1</sup>, CHEN YANG<sup>2</sup>, MICHAEL E. VON FRICKEN<sup>3</sup>, MICHAEL H. NORRIS<sup>4</sup>, AND THOMAS A. WEPPELMANN<sup>5</sup>



1. DEPARTMENT OF POPULATION HEALTH SCIENCES, GEORGIA STATE UNIVERSITY; 2. DEPARTMENT OF BIostatISTICS, HARVARD T.H. CHAN SCHOOL OF PUBLIC HEALTH, HARVARD UNIVERSITY; 3. DEPARTMENT OF GLOBAL AND COMMUNITY HEALTH, COLLEGE OF HEALTH AND HUMAN SERVICES, GEORGE MASON UNIVERSITY; 4. SCHOOL OF LIFE SCIENCES, UNIVERSITY OF HAWAII AT MANOA; 5. DEPARTMENT OF INTERNAL MEDICINE, UNIVERSITY OF SOUTH FLORIDA;



## INTRODUCTION

### Motivation

- Since the discovery of LD only limited progress has been made concerning the epidemiology of sporadic cases.
- For selected outbreaks air conditioning and potable water systems were confirmed as point sources of community-acquired LD.

### Challenges

- Measuring the associations between water quality and LD incidence has remained challenging due to the heterogeneity of water systems over the geographically large study areas.
- While seasonal trends in incidence have already been linked to increased rainfall and warmer temperatures, the geographic units remained large.
- Large areas caused large distances between weather stations and the actual locations of reported cases.
- The goal is to explore the relationships of incidence with rainfall and temperatures on a more accurate scale.

### Study

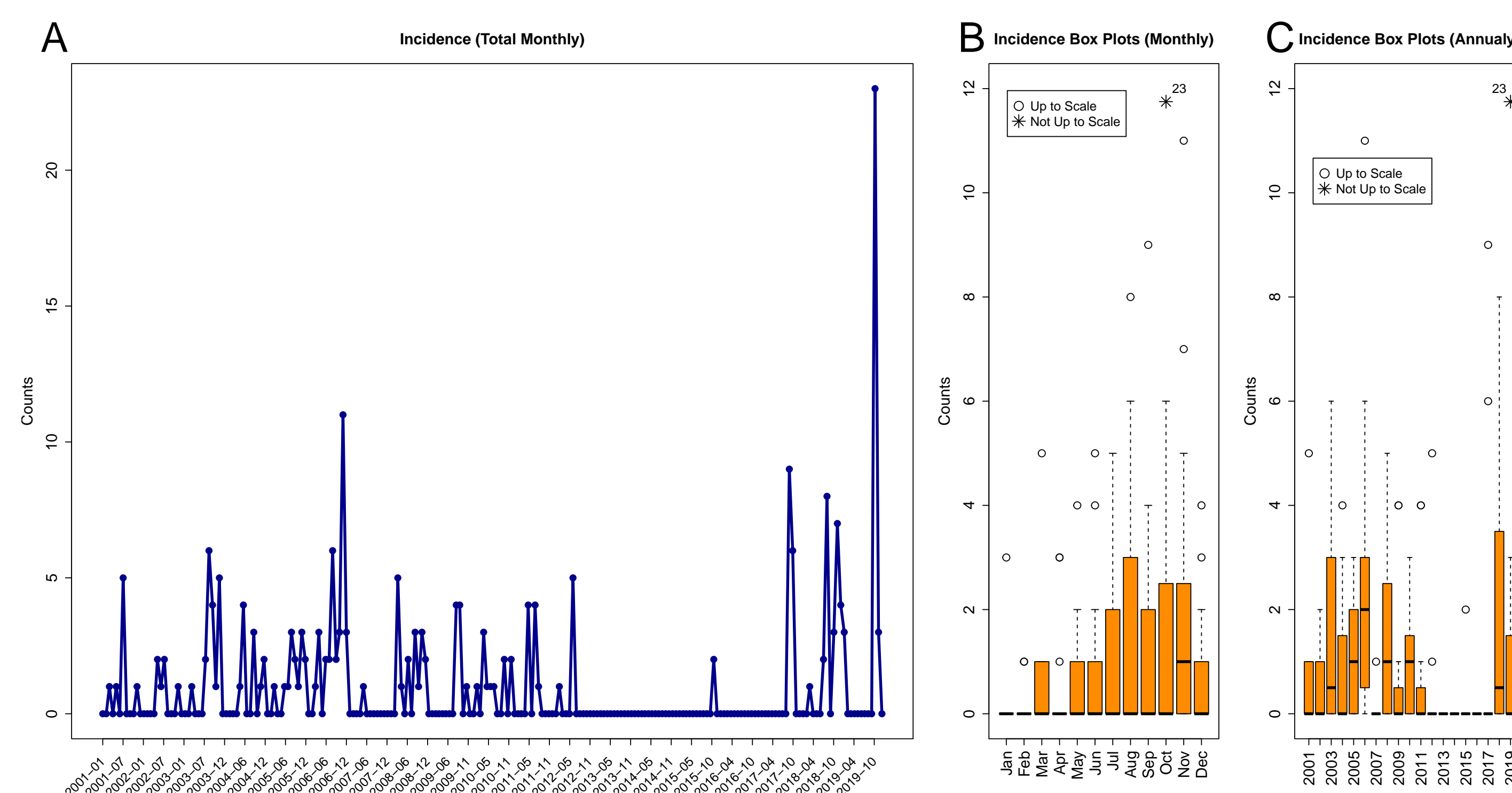
- Retrospective ecological study in Washington, DC using monthly reported cases of LD from 2001 to 2019 together with meteorological and water quality data.
- This study aimed to ascertain if changes in municipal or surface water components concentrations as well as meteorological factors and their seasonality were associated with the increase in sporadic cases of LD.

## MOTIVATION

- The municipality of Washington, DC was selected to overcome the challenge of large geographic units.
- Study area is located in the Northeastern Mid-Atlantic region which had a history of outbreaks earlier.
- Study area is geographically homogenous and is located within the same climatic region with precise meteorological data.
- Study has only a single source of municipal water drawn from the Potomac River during 2001-2019.
- Washington, DC maintained detailed water quality records of both surface water and water processed by treatment plants.
- Since LD is a reportable disease, cases for Washington, DC were aggregated and reported weekly by the CDC.

## DATA SUMMARIES

The raw summaries of case counts (Panel A) and box and whisker plots both monthly across all studied years (Panel B) and yearly across all months within each year (Panel C).



The raw summaries of case counts by year (Table 1) and monthly across years (Table 2).

Table 1. Monthly reports of Legionnaires' disease cases for Washington, DC from January 2001 until December 2019.

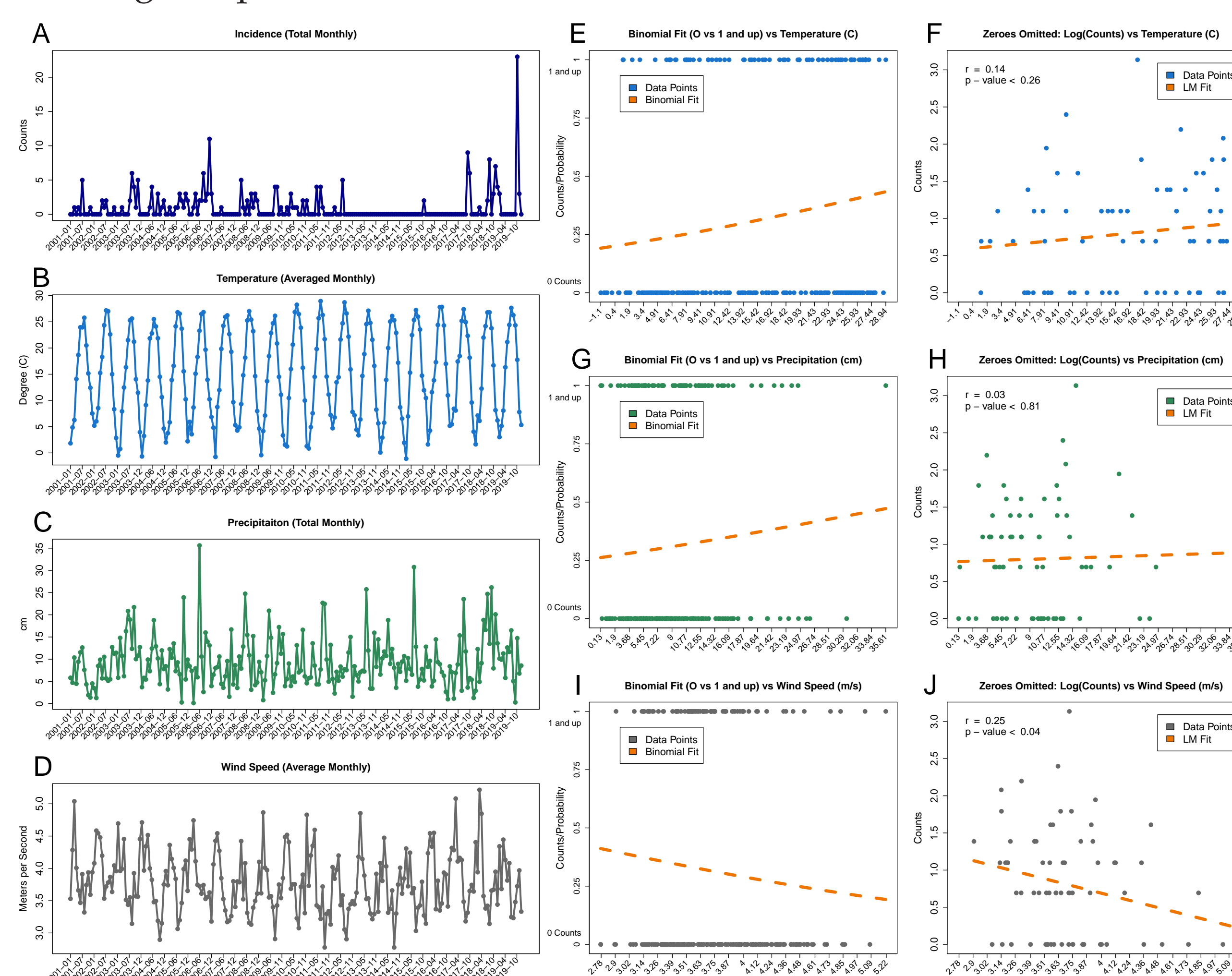
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2001	0	0	1	0	1	0	5	0	0	0	1	0
2002	0	0	0	0	2	1	2	0	0	0	1	0
2003	0	0	1	0	0	2	6	4	1	5	0	0
2004	0	0	0	0	1	4	0	0	3	0	1	2
2005	0	0	1	0	0	1	1	3	2	1	3	2
2006	0	0	1	3	0	2	2	6	2	3	11	3
2007	0	0	0	0	1	0	0	0	0	0	0	0
2008	0	0	5	1	0	2	0	3	1	3	2	0
2009	0	0	0	0	0	0	0	4	4	0	1	0
2010	0	1	0	3	1	1	1	0	0	2	0	2
2011	0	0	0	0	4	0	4	1	0	0	0	0
2012	0	1	0	0	0	5	0	0	0	0	0	0
2013	0	0	0	0	0	0	0	0	0	0	0	0
2014	0	0	0	0	0	0	0	0	0	0	0	0
2015	0	0	0	0	0	0	0	0	0	0	2	0
2016	0	0	0	0	0	0	0	0	0	0	0	0
2017	0	0	0	0	0	0	0	0	9	6	0	0
2018	0	0	1	0	0	0	2	8	0	3	7	4
2019	3	0	0	0	0	0	0	0	23	3	0	0

Table 2. Monthly summaries of Legionnaires' disease cases with the corresponding quantiles across studied years for Washington, DC from January 2001 until December 2019.

Month	Mean	St. Dev.	Min (0%)	Lw. Qrt. (25%)	Median (50%)	Up. Qrt. (75%)	Max (100%)	Total
Jan	0.16	0.69	0	0	0	0	3	3
Feb	0.11	0.32	0	0	0	0	1	2
Mar	0.53	1.17	0	0	0	1	5	10
Apr	0.37	0.96	0	0	0	0	3	7
May	0.53	1.02	0	0	0	1	4	10
Jun	0.84	1.46	0	0	0	1	5	16
Jul	1	1.49	0	0	0	2	5	19
Aug	1.63	2.59	0	0	0	3	8	31
Sep	1.32	2.33	0	0	0	2	9	25
Oct	2.21	5.3	0	0	0	2.5	23	42
Nov	1.95	2.91	0	0	1	2.5	11	37
Dec	0.68	1.25	0	0	0	1	4	13

## RESULTS

The pairwise relationships between the considered predictors of interest and for the logit and log components of the model.



- 1) **Seasonality.** LD incidence with the majority of cases from June to December. The largest number of cases were August, October and November.
- 2) **Temperature.** Increase was associated with the presence of LD incidence.
- 3) **Surface water.** The increase in concentrations of manganese, iron and strontium were positively associated and for aluminum and orthophosphate were negatively associated with LD incidence.
- 4) **Treatment plant water.** The increase in concentrations of total organic carbon, aluminum, barium and chlorine were positively associated and strontium, zinc and orthophosphate were negatively associated with the LD incidence.
- 5) The results for orthophosphates and turbidity were ambiguous and additional studies were recommended for those compounds.

## DATA SOURCES

- **Reported Cases** in Washington, DC from January 2001 to December 2019 were obtained from the CDC Notifiable Infectious Disease Data Tables. The cases reported by the CDC are defined as clinically compatible which have laboratory confirmed infections.
- **Water quality** measurements were obtained from the US Army Corps monthly reports. The records from January 2001 to December 2019 included measurements from surface water of the Potomac River and the treated water from Dalecarlia and McMillan treatment plants.
- **Surface water** measurements included pH, concentrations of nitrate, orthophosphate, manganese, strontium, barium, nickel, total organic carbon, turbidity, aluminum, zinc, and iron.
- **Treated water** measurements included pH, concentrations of nitrate, orthophosphate, manganese, strontium, barium, nickel, total organic carbon, turbidity, aluminum, total coliform, heterotrophic plate count (HPC), zink, iron, and chlorine.
- **Meteorological data** were obtained from the National Oceanographic and Atmospheric Association Global Summary of the Day generated reports. Data from Reagan Airport monitoring station adjacent to Washington, DC were used for the analysis.

## ANALYSIS

- Hurdle versions of Poisson and negative binomial models were used for the analysis.
- In the hurdle model, there are two components:
  - i) The binary hurdle component, which accounts for excessive zeroes in the distribution (i.e. it models zeroes vs. non-zeroes) with the logit link.
  - ii) The truncated Poisson (or negative binomial) distribution which is used for the remaining counts starting from one (i.e. omitting zero) with the log link.

## CONTACT

Alexander Kirpich Ph.D.  
 Department of Population Health Sciences  
 School of Public Health  
 Georgia State University  
 140 Decatur Street SE  
 Suite 407  
 Atlanta, GA 30303  
 akirpich@gsu.edu  
<https://alexanderkirpich.com>