

University of Rhode Island

DigitalCommons@URI

---

Civil & Environmental Engineering Faculty  
Publications

Civil & Environmental Engineering

---

2022

## A High-Resolution Earth Observations and Machine Learning-Based Approach to Forecast Waterborne Disease Risk in Post-Disaster Settings

Farah Nusrat  
*University of Rhode Island*

Musad Haque

Derek Rollend

Gordon Christie

Ali S. Akanda  
*University of Rhode Island, akanda@uri.edu*

Follow this and additional works at: [https://digitalcommons.uri.edu/cve\\_facpubs](https://digitalcommons.uri.edu/cve_facpubs)

Creative Commons License



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

---

### Citation/Publisher Attribution

Nusrat, F.; Haque, M.; Rollend, D.; Christie, G.; Akanda, A.S. A High-Resolution Earth Observations and Machine Learning-Based Approach to Forecast Waterborne Disease Risk in Post-Disaster Settings. *Climate* 2022, 10, 48. <https://doi.org/10.3390/cli10040048>  
Available at: <https://doi.org/10.3390/cli10040048>

This Article is brought to you for free and open access by the Civil & Environmental Engineering at DigitalCommons@URI. It has been accepted for inclusion in Civil & Environmental Engineering Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

## Article

# A High-Resolution Earth Observations and Machine Learning-Based Approach to Forecast Waterborne Disease Risk in Post-Disaster Settings

Farah Nusrat <sup>1</sup>, Musad Haque <sup>2</sup> , Derek Rollend <sup>2</sup>, Gordon Christie <sup>2</sup> and Ali S. Akanda <sup>1,\*</sup> 

<sup>1</sup> Department of Civil and Environmental Engineering, University of Rhode Island, Kingston, RI 02881, USA; farah\_nusrat@uri.edu

<sup>2</sup> Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723, USA; musad.haque@jhuapl.edu (M.H.); derek.rollend@jhuapl.edu (D.R.); gordon.christie@jhuapl.edu (G.C.)

\* Correspondence: akanda@uri.edu; Tel.: +1-401-874-7050

**Abstract:** Responding to infrastructural damage in the aftermath of natural disasters at a national, regional, and local level poses a significant challenge. Damage to road networks, clean water supply, and sanitation infrastructures, as well as social amenities like schools and hospitals, exacerbates the circumstances. As safe water sources are destroyed or mixed with contaminated water during a disaster, the risk of a waterborne disease outbreak is elevated in those disaster-affected locations. A country such as Haiti, where a large quantity of the population is deprived of safe water and basic sanitation facilities, would suffer more in post-disaster scenarios. Early warning of waterborne diseases like cholera would be of great help for humanitarian aid, and the management of disease outbreak perspectives. The challenging task in disease forecasting is to identify the suitable variables that would better predict a potential outbreak. In this study, we developed five (5) models including a machine learning approach, to identify and determine the impact of the environmental and social variables that play a significant role in post-disaster cholera outbreaks. We implemented the model setup with cholera outbreak data in Haiti after the landfall of Hurricane Matthew in October 2016. Our results demonstrate that adding high-resolution data in combination with appropriate social and environmental variables is helpful for better cholera forecasting in a post-disaster scenario. In addition, using a machine learning approach in combination with existing statistical or mechanistic models provides important insights into the selection of variables and identification of cholera risk hotspots, which can address the shortcomings of existing approaches.

**Keywords:** earth observations; machine learning; waterborne disease forecast; post-disaster



**Citation:** Nusrat, F.; Haque, M.; Rollend, D.; Christie, G.; Akanda, A.S. A High-Resolution Earth Observations and Machine Learning-Based Approach to Forecast Waterborne Disease Risk in Post-Disaster Settings. *Climate* **2022**, *10*, 48. <https://doi.org/10.3390/cli10040048>

Academic Editors: Mário Gonzalez Pereira and Rajib Shaw

Received: 26 January 2022

Accepted: 18 March 2022

Published: 22 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Impact of Disasters on Humans (Health, Economic, and Social)

The frequency and intensity of natural disasters are increasing with time due to the vital influence of climate change on environmental and natural disasters along with its influence in shifting extreme weather events [1,2]. The International Federation of Red Cross (IFRC) reports that a sharp increase of 35% has been observed in the number of climate- and weather-related extreme events across the world [3]. Although the severity and frequency of natural disasters impact many countries [4], people in underdeveloped and developing countries are the most vulnerable to natural disasters [5] due to a lack of essential resources to overcome the adverse impact of these events. The IFRC report states that around 1.7 billion people were impacted in the last 10 years due to climate- and weather-related disasters, and that approximately 410,000 deaths have been attributed to these events in low- or middle-income countries [3].

One of the most common consequences of natural disasters is the effect on health, which is closely related to water and sanitation crises that propagate following such

disasters [6]. As existing water supply and sanitation systems are damaged during natural disasters, it causes difficulty in accessing safe water supply and sanitation systems for the populations living in that area. Their physical health is also at risk during and post-disaster scenarios by the displacement of people who subsequently have scarce opportunities to practice personal hygiene along with having sufficient access to safe water. As a result, the likelihood of an increase in waterborne disease is more prevalent in a post-disaster scenario due to poor water, sanitation, and hygiene (WASH) conditions [7]. Besides, traumatic experiences during and after disasters also hamper mental health and have long-term health consequences [8]. These include the disruption in their full-functioning lives, the demise of family members or friends, loss of resources, and social support [9]. Apart from health issues, natural disasters significantly interrupt the economic and social development of a region and take considerable time periods to recover and reinstall necessary services [10]. Food insecurity and malnutrition often occur after natural disasters due to loss of agricultural lands and crops amid disasters, which elevates the risk of cholera and other diarrheal disease outbreaks [11]. Mainly, the transmission of cholera happens through the consumption of water and foods contaminated with *Vibrio cholerae*, the causative agent of the cholera disease [12]. The disease is mostly seen in populations that are deprived of proper water supply and sanitation systems [13–19]. Globally, around 100,000–150,000 deaths occur per year due to cholera [20,21].

### 1.2. Waterborne Disease after Natural Disasters

According to the Intergovernmental Panel for Climate Change (IPCC), more extreme temperature and precipitation events are expected in the future with higher frequency and intensity due to climate change [22]. Extreme precipitation events can cause flooding and geological hazards like landslides [23]. Floods can increase the transmission risk of waterborne and vector-borne diseases, apart from injury and mortality [24], as heavy rainfall-induced floods can increase human contact with causative pathogens of waterborne diseases [25]. Not only floods, but also droughts are found to influence the increase of diarrheal diseases [26]. Approximately 3.6 billion of the global population is living in water-scarce areas, who are suffering from water scarcity at least one month per year [27]. As water availability decreases, people need to collect water from distant and often compromised shared sources [28], and sometimes are forced to share sources of water with animals [29] and thus sacrifice hygiene behavior [30,31]. Low rainfall intensifies effluent concentration in surface water, which amplifies pathogen concentration in water sources [26,30,32]. Additionally, the pathogen concentration in water increases as the river discharge decreases, and scarce rainfall cannot provide enough dilution of pathogens [32].

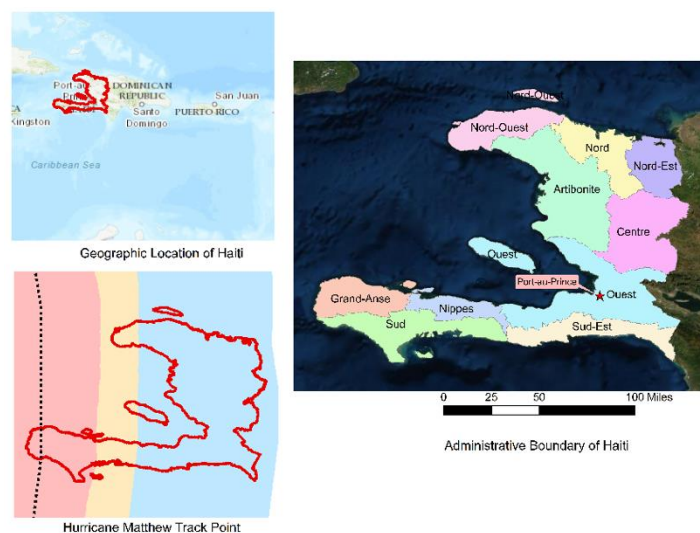
Cholera outbreaks are also observed after large coastal storms, such as hurricanes or cyclones, due to water contamination and disruption in the water distribution systems and lack of proper hygiene facilities, which all create favorable conditions for an outbreak [33–35]. The floods and storm surges resulting from hurricanes damage road networks and river embankments, causing the submergence of ponds and water wells [36] and scarcity of safe drinking water. Unsanitary practices, such as open defecation due to infrastructural damage following a hurricane and living in congested and overpopulated temporary shelters, can additionally elevate the cholera outbreak risk [37–39]. Additionally, salinity is a key environmental factor that affects cholera bacteria growth; thus, seawater intrusion and contamination of coastal water resources following cyclones make coastal areas vulnerable to cholera transmission [40,41]. In 2019, the country of Mozambique was hit by two tropical cyclones, Idai and Kenneth, 6 weeks apart [42]. Cyclone Idai killed 1000 people and destroyed 100,000 homes [43]. Additionally, 600,000 people were displaced due to cyclone-triggered flooding [44] and about 1500 suspected cholera cases were reported a week after the cyclone [45]. Information about cholera cases and mortalities after hurricanes are listed in Table 1.

**Table 1.** Cyclone/hurricane-related cholera outbreaks.

Natural Disasters	Country	People Impacted by Natural Disasters		Year	Cholera Related to Natural Disasters		House Damaged	Health Facility Damaged
		Affected	Death		Cases	Fatalities		
Cyclone Aila	West Bengal, India	6.8 million [46]	138 [46]	26 May 2009	1076 [36] (diarrhea: 85,000) [46]	14 [36] (diarrhea: 28) [46]	945,000 [46]	30% of health sub-centers [46]
Hurricane Matthew	Haiti	2.1 million [47]	546 [47]	4 October 2016 (after Matthew) 2017 2018 2019	8457 [48] 13,747 [47] 3786 [51] 308 [52]	100 [48] 159 [47] 41 [51] 3 [52]	25,160 [49]	36 [50]
Cyclone Kenneth	Mozambique	374,000 [53]	45 [53]	25 April 2019	267 [54]	–	45,000 [53]	19 [53]
Cyclone Idai	Mozambique	1.85 million [55]	Over 602 [55]	14 March 2019	6682 [56]	8	240,000 [53]	94 [53]

### 1.3. Cholera Outbreaks in Haiti

Ten months after the devastating earthquake that struck Haiti in 2010, the country has been facing at least 27,000 cholera cases per year, with more than half of total cases requiring hospitalization due to severity [57,58]. After analyzing the genetic studies of the particular cholera strain in Haiti, it was found to have originated in South Asia and was discharged to Haiti from a single point source and killed around 5000 people in the initial phase of this epidemic [59]. Between 2011 and 2014, the reported cholera cases in Haiti reduced from 350,000 to 21,916 [60], which was thought to be the end of the epidemic. However, the failure to provide financial assistance and essential WASH interventions [58] contributed to the reoccurrence of cholera outbreaks in 2015 and again in 2016, after Hurricane Matthew. The administrative area of Haiti and Hurricane Matthew Track Points are shown in Figure 1.

**Figure 1.** Administrative area of Haiti and Hurricane Matthew Track Points.

### 1.4. Research Questions and Objectives of the Paper

It remains a challenge to forecast the risk of waterborne diseases (such as cholera) in post-disaster scenarios and the corresponding impacts on WASH conditions, which are conducive to an outbreak following a natural disaster. Additionally, infrastructural damages, such as those to road networks, bridges, and hospitals during natural disasters, often restrict access to nearby health facilities, which contributes to the lack of a comprehensive database of disease cases for use by humanitarian agencies, government offices, and academic researchers. Satellite data can be beneficial in disease forecasting, as it provides

high spatio-temporal resolution and better global coverage than station data [61,62]. It provides a substantial opportunity to understand the influence of environmental variables that can make suitable conditions for disease outbreaks and to monitor the impact of spatio-temporal variability on identifying disease hotspots. Access to high-resolution data during and after natural calamities can be challenging as the environmental status is distinct from the usual, i.e., high cloud cover during a cyclone [63]. Thus, creating a predictive cholera risk model using satellite-derived high spatio-temporal resolution data as environmental predictors in post-disaster scenarios can be beneficial for rapid WASH interventions.

While existing work on cholera prediction from measured environmental variables has been guided by human intuition, machine learning (ML) techniques have the potential to automatically determine which variables are the strongest indicators of cholera risk. By learning from previously recorded cholera case counts in post-disaster scenarios, we can build more accurate prediction models to determine whether ML inputs align with human intuition and identify new environmental variables that should be considered.

Using cholera as a signature disease, this study offers insight on cholera outbreaks in Haiti in 2016 following the landfall of Hurricane Matthew. Using high-resolution spatio-temporal datasets of environmental variables that are associated with the hurricane, we developed five unique and comprehensive post-disaster cholera prediction models. The study aims to (a) improve the understanding of the influence of hydroclimatological variables on waterborne disease outbreaks in a post-disaster scenario and (b) select appropriate predictors for improved cholera forecasting models by using better spatio-temporal satellite Earth Observation (EO) datasets. Apart from selecting suitable hydroclimatic drivers for better cholera risk prediction, this analysis will significantly increase our understanding of the threats to WASH access, damage to road networks and infrastructure, population displacement, and the disruptions to essential services following natural disasters that may elevate to waterborne disease threats. In addition, the outcomes of this study will help to improve policy interventions and community engagement approaches to better manage environmental health impacts.

## 2. Materials and Methods

### 2.1. Data

In the geospatial models used in our study, cholera risk is the dependent variable, whereas socio-environmental factors are the independent variables. Model A is created with four basic socio-environmental variables, which use higher-resolution data than the previous study. Model B includes extreme event variables during the hurricane. Each Plus model includes cloud height, cloud top temperature, wind speed, and building damage data, as they are found to be influential in machine learning models. The data used in different geospatial models are population density, precipitation anomaly, temperature anomaly, Hurricane Matthew wind swath and windspeed, elevation, extreme rainfall during Hurricane Matthew, cloud top temperature, cloud height, and building damage. Combinations of input variables of different geospatial models are listed in Table 2.

**Table 2.** Input variables of different geospatial models.

Model A	Model A Plus	Model B	Model B Plus
Precipitation Anomalies (mm/day)	Precipitation Anomalies (mm/day)	Precipitation Anomalies (mm/day)	Precipitation Anomalies (mm/day)
Temperature Anomalies (°C)	Temperature Anomalies (°C)	Temperature Anomalies (°C)	Temperature Anomalies (°C)
Population Densities (count per km <sup>2</sup> )	Population Densities (count per km <sup>2</sup> )	Population Densities (count per km <sup>2</sup> )	Population Densities (count per km <sup>2</sup> )
Wind Swath of Hurricane (miles)	Wind Speed Above Ground (m/s)	Wind Speed (miles/hour)	Wind Speed Above Ground (m/s)
	Cloud Height (m)	Extreme Rainfall (mm/day)	Extreme Rainfall (mm/day)
	Cloud Temperature (k)	Elevation (m)	Elevation (m)
	Building Damage		Cloud Height (m)
			Cloud Temperature (k)
			Building Damage

Gridded population density data were collected from NASA's Socioeconomic Data and Applications Center (SEDAC) website (<https://sedac.ciesin.columbia.edu/data/collection/gpw-v4> (accessed on 3 March 2022)) and the product name is "Gridded Population of the World (GPW), v4" with a spatial resolution of 30 s (approximately 1 km × 1 km). Extreme rainfall that occurred during the hurricane was collected from the Humanitarian Data Exchange (<https://data.humdata.org/dataset/accumulated-gpm-imerg-data-for-haiti-hurricane-matthew-october-3-6th-2016> (accessed on 3 March 2022)), which is the total accumulated rainfall in mm from 3 to 6 October 2016, and the data source is Integrated Multi-satellite Retrievals for GPM (IMERG) data. For model B, the Hurricane Matthew gust footprint and windspeed (with highest wind gusts of 160 mph or 260 km/h) were collected from the Humanitarian Data Exchange but were originally an output from University College London model (<https://data.humdata.org/dataset/hurricane-matthew-gust-footprint-tropical-storm-risk-university-college-london> (accessed on 3 March 2022)). For model A, wind swath data were collected from the National Hurricane Center [62]. Elevation data were taken from HydroSHEDS (<https://www.hydrosheds.org> (accessed on 3 March 2022)), which is derived from the Shuttle Radar Topography Mission (SRTM) with 1 km × 1 km spatial resolution. All datasets were then resampled to 1 km × 1 km spatial resolution.

The current case study focused on the outbreak of cholera in Haiti in the aftermath of Hurricane Matthew in October 2016 [64]. In addition to variables used in previous works, auxiliary remote sensing products were retrieved from the NASA Global Maps collection [65] to provide more environmental measurements from which the ML algorithms could learn. Monthly averaged global map features were used for three months before and three months after Hurricane Matthew (July–December 2016), and example features include aerosol optical depth, cloud fraction, chlorophyll concentrations, and water vapor amounts.

### 2.1.1. Precipitation Anomaly

The difference between the monthly precipitation in a month and long-term mean precipitation is known as the precipitation anomaly. Thus, a positive anomaly means current precipitation is more than the average precipitation of that location, and a negative anomaly implies the mean precipitation is higher than the current precipitation.

The Global Precipitation Measurement (GPM) data with a spatial resolution of  $0.10^\circ \times 0.10^\circ$  (~10 km × 10 km) were used for precipitation of 2016 in all models, whereas two different data sources were used for long-term precipitation mean. For model A and A Plus, Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA/3B43\_v7) data from 1998 to 2019 were used for the long-term mean with a spatial resolution of  $0.25^\circ \times 0.25^\circ$  (~25 km × 25 km), and for model B and model B Plus, IMERG data were used from 2000 to 2019 with a spatial resolution of  $0.10^\circ \times 0.10^\circ$  (~10 km × 10 km) for the long-term mean. IMERG uses an algorithm to combine the previous precipitation data of the TRMM satellite (2000–2015) with recent precipitation data of the GPM satellite (2014–present). All data were collected from the Giovanni-NASA website (<https://giovanni.gsfc.nasa.gov/giovanni/> (accessed on 3 March 2022)).

### 2.1.2. Temperature Anomaly

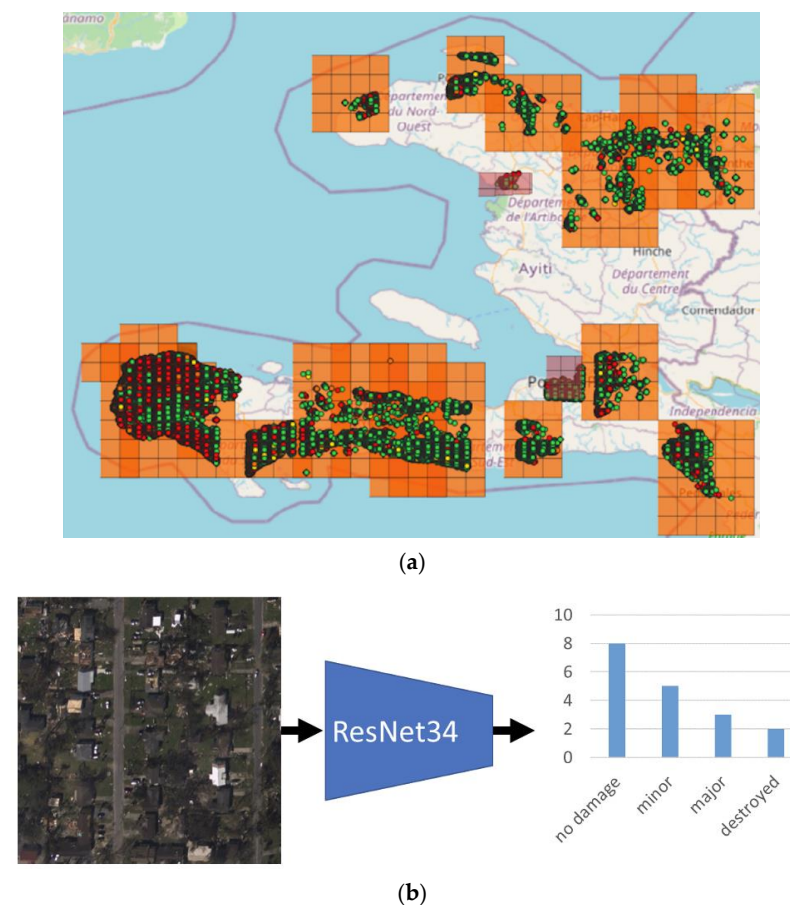
The difference between the average temperature in a month and long-term mean temperature is known as the temperature anomaly. Thus, a positive anomaly shows that the current month's average temperature is more than the long-term average temperature of that month, and a negative anomaly implies the long-term mean temperature is higher than the current average temperature.

In all models, to get a single Land Surface Temperature Analysis product, a combination of the Global Historical Climatology Network (GHCN) and Climate Anomaly Monitoring System (CAMS) Land Temperature Analysis was used from the National Oceanic and Atmospheric Administration (NOAA) with a high spatial resolution of  $0.5^\circ \times 0.5^\circ$  from 1948 to the near present. For the long-term temperature mean, data from 1990 to 2019 were used.

Both the current and long-term average temperature data were collected from NOAA's Physical Sciences Laboratory (<https://psl.noaa.gov/data/gridded/data.ghcncams.html> (accessed on 3 March 2022)) website, and the product name is "GHCN\_CAMS Gridded 2 m Temperature (Land)".

### 2.1.3. Humanitarian Assistance and Disaster Relief (HADR) Data

Deep learning was used to assess building damage and characterized as either "no damage", "minor", "major", or "destroyed" from visible spectrum satellite imagery. This damage was characterized and used as an input to the Plus models. In this study, a deep learning model was trained to estimate the amount of building damage in a given area from visible spectrum satellite imagery (Figure 2a). Previous work has focused on parcel-level damage assessment, but this level of granularity is unnecessary in the current context and including more regional context was found to help reduce false positives. A regional damage regression model based on the ResNet-34 architecture [66] was trained using the xView2 dataset [67]. A  $1024 \times 1024$  pixel image tile ( $\sim 0.5 \text{ km}^2$  with  $\sim 50 \text{ cm}$  Maxar WorldView imagery) was used as the input, and the output was the estimated number of damaged buildings in each xView2 damage category: no damage, minor, major, and destroyed (see Figure 2b). For training, the instance-level xView2 damage labels were converted to image-level summaries, which alleviated the need for building segmentation or pre-existing building vector data. The trained model was then run over all available Digital Globe Open Imagery from Hurricane Matthew over Haiti [68], and the predicted total number of damaged buildings for each  $\sim 0.5 \text{ km}^2$  image tile were used as features for the higher-level ML algorithms.



**Figure 2.** (a) Imagery extents (colored squares) and predicted building damage locations (colored points) for Haiti after Hurricane Matthew. This feature was fed into the model and was generated by using deep learning on visible spectrum satellite imagery. (b) Region-level building damage estimation model overview.

Using all a priori, post-hurricane, and monthly features up to and including the current month resulted in a total of 57 features for prediction of cholera risk in October, 72 for November, and 87 for December.

#### 2.1.4. Cholera Data

There are 10 departments (provinces) in Haiti, made up of 140 communes (villages or towns). Cholera case counts for each department and commune in Haiti were retrieved from the Pan American Health Organization, spanning epidemiological weeks 40–49 (4 October–7 December 2016) for communes and weeks 35–49 for departments (1 September–7 December 2016) [69]. For easier comparison with previous works that estimated cholera risk from (0–1), the recorded cholera case counts were first normalized by commune and department populations and then scaled from 0–1 over the entire country of Haiti. To evaluate commune-level ML model predictions every month, monthly commune risks were estimated by combining monthly department-level recorded cholera cases and commune proportions of total cholera case counts. Total cholera counts from October to December 2016 were available for each commune and department in Haiti. Ratios were computed for each commune by dividing by their encompassing department’s total counts.

#### 2.2. Models

##### Weighted Sum Models

The model structure with input variables is provided in Figure 3a. In this weighted sum model (Figure 3a), the risk for each variable is classified from risk 0 to risk 4 (a value of 4 represents the highest risk) and given a weightage according to the variable’s importance for cholera risk. Then, each variable is multiplied by their respective weightage and summing them together to get an integrated raster layer of cholera risk. For better comparison between time periods, the risk layer is then rescaled between 0 and 1. Different model compositions are available from Tables S1–S4 in the Supplementary Materials.

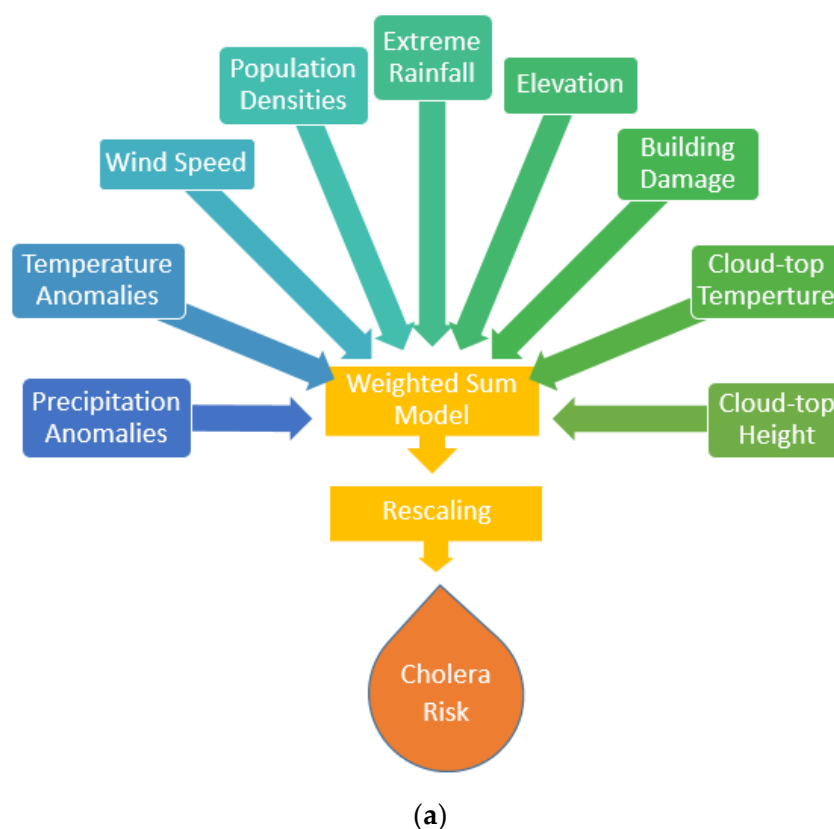
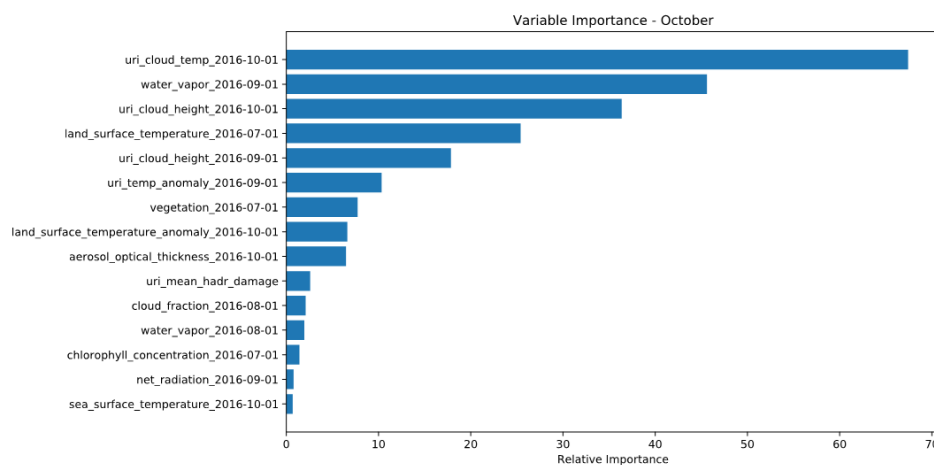


Figure 3. Cont.





(b)

**Figure 3.** (a) Weighted sum model for cholera risk. (b) Relative feature importance for the XGBoost regression model for October.

The machine learning approach was formulated as a regression problem, where the goal was to predict cholera risk on a per-commune basis for each month in the aftermath of Hurricane Matthew (October, November, and December 2016). Gradient-boosted trees [70] were used for the ML models, as they can produce accurate results and provide interpretability as to which features are most indicative of cholera risk. The trees were trained by minimizing the mean squared error between predicted and ground truth cholera risk values. During training, additive tree learning was used to iteratively build decision trees, the goal of which was to optimize the information gained by splitting training exemplars using different input features at every node. This process was repeated for the maximum number of trees in the ensemble, and the ensemble consensus was used as the final model output. In our experiments, the number of trees in the final ensembles varied between 40 and 100 and the maximum depth of the decision trees varied between 2 and 4, across the four separate monthly ML models. Three models were trained (one for each month), and a hyperparameter grid search was performed over key model parameters for each month (since there were a variable number of features for each month).

As the size of the input dataset was limited (cholera cases were recorded for 103 of the 140 total communes in Haiti), 10-fold randomized cross-validation was performed over all available data a total of 10 times, with risk predictions for each commune averaged over all 10 appearances in the validation set. These commune-level risk predictions were used to compute the average risk for each department in Haiti, to compare with the ground truth monthly data. Analysis was performed on the department level to provide a fairer comparison with existing model outputs, which are more informative and accurate when averaged overall points in a given department.

- Feature Importance

Gradient-boosted trees [70] were used for ML model development, as they are able to produce accurate results and provide interpretability as to which features are most indicative of cholera risk. An example feature importance plot is shown in Figure 3b for the October ML risk prediction model. While the conclusiveness of these feature importance values may be limited by the small size of our dataset, they do provide preliminary insights into which variables are most predictive for an ML model. In this case, measurements of cloud properties appear to be indicative of post-hurricane cholera risk. This observation is intuitive as stronger, larger clouds likely produced larger amounts of precipitation and flooding, which are known to influence cholera outbreaks.

### 3. Results

#### 3.1. Correlation with Observed Cholera Cases

There are four post-disaster cholera prediction models: Model A (base model), Model B, Model A Plus, and Model B Plus. All the Plus models have variables of their respective primary models (Model A, Model B) plus four other variables: cloud height, cloud top temperature, wind speed above ground, and building damage data. To validate the models, the actual number of cholera cases in all departments and communes of Haiti (presented in Table 3) was compared with the models' predicted risks for each month using the Pearson Correlation coefficient, as presented in Table 4. The  $p$ -values of the ML model shows that the correlation between the ML model and reported cholera cases is high and statistically significant.

**Table 3.** Total reported cholera cases per month [69].

Month	Total Cholera Cases
September	2461
October	4998
November	3913
December	946

**Table 4.** Monthly Pearson correlation between the actual number of cholera cases and predicted cholera risk from all models.

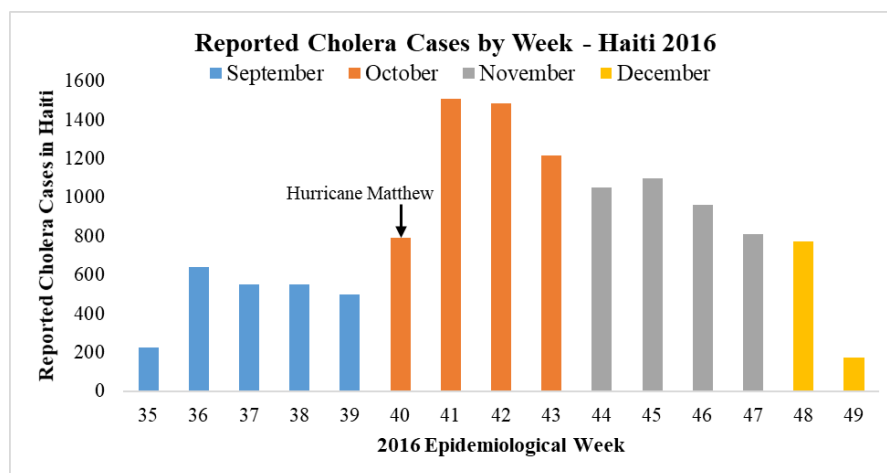
Month	Model A	Model A Plus	Model B	Model B Plus	ML Model
September	−0.044	−0.398	−0.099	−0.692	0.818 (0.0039)
October	0.590	0.574	0.645	0.512	0.811 (0.0044)
November	0.649	0.687	0.748	0.570	0.733 (0.0158)
December	0.690	0.365	0.648	0.364	0.744 (0.0136)
Average (October–December)	0.643	0.542	0.680	0.482	0.760

Cholera case counts for 10 departments and 140 communes in Haiti were retrieved from the Pan American Health Organization, spanning epidemiological weeks 40–49 (4 October–7 December 2016) for communes and weeks 35–49 for departments (1 September–7 December 2016) (Figure 4) [69]. For easier comparison with the base model that estimated cholera risk from (0–1), the recorded cholera case counts were first normalized by commune and department populations and then scaled from 0 to 1 over the entire country of Haiti. To evaluate commune-level ML model predictions every month, monthly commune risks were estimated by combining monthly department-level recorded cholera cases and commune proportions of total cholera case counts. Ratios were computed for each commune by dividing by their encompassing department's total counts, as the total cholera counts from October to December 2016 were available for each commune and department in Haiti.

#### 3.2. Understanding from the Variation of Model Outputs in Different Months

##### 3.2.1. Model A

Model A uses four variables for the October to December models—population density, rainfall anomaly, temperature anomaly, and wind swath during the hurricane—and for September, the model uses all variables except wind swath during the hurricane. Model A is the base model for this analysis, which used the same variables as this paper [62], and the objective is to use minimum variables to predict cholera risk.



**Figure 4.** Reported cholera cases from 10 departments by week in Haiti before and after Hurricane Matthew in 2016 [69].

- September

The model overestimated the risk in the western peninsula and northern part of the country. It estimated the high risk well in the middle of the country near the capital, where population density is high. The correlation of the model output and actual cholera cases showed a negative low correlation ( $-0.044$ ), which implies the model failed to predict risk spatially before the hurricane in September with only three variables (population density, rainfall, and temperature anomaly). No extreme climatic variables were included in this model for September. Population played a major role in the risk map, as its pattern matches with the model output, whereas the pattern of temperature and rainfall anomaly did not match with the model output in most of the locations.

- October

Starting from October to December, all risk layers showed some vertical straight lines, which are due to the wind swath data that measures the distance from the wind path. The correlation between the actual cholera cases and cholera risk was 0.590, which implies this model predicts better than the previous month after including the extreme variable—wind swath. Though it better estimated the high cholera risk in the western peninsula, it overestimated the cholera risk in Nord-Ouest, Nord, Nord-Est, Nippes, Sud-Est, and Centre departments. Except for temperature anomaly, the spatial pattern of all other variables matched with the model output's pattern.

- November

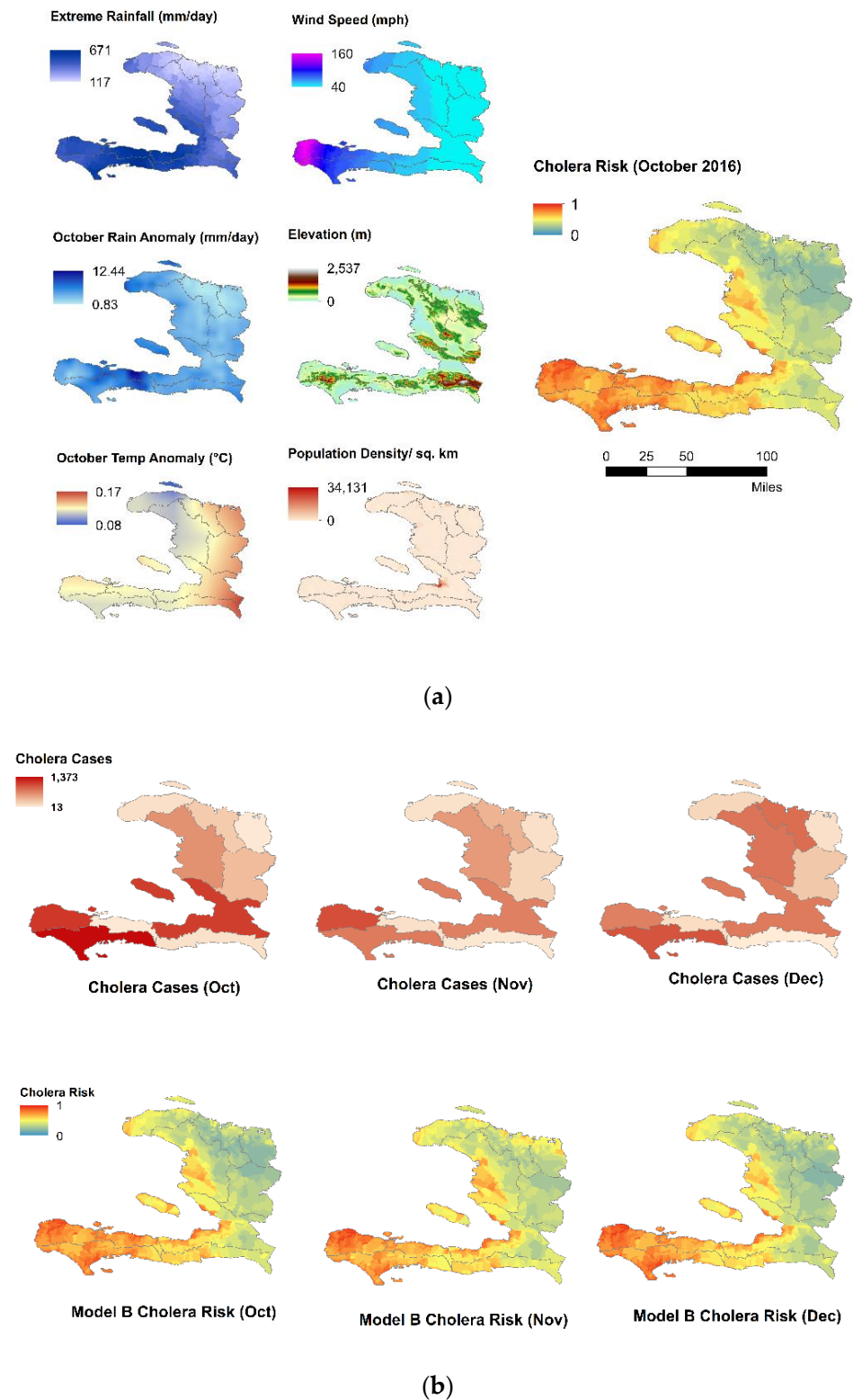
From October to November, the actual cholera cases increased in the north direction, which the model captured very well; thus, the correlation increased to 0.649. Though the high-risk area remained in the western peninsula, the number of cholera cases spread in the northern direction as well. As wind swath and population density are fixed layers in all models, and the rainfall anomaly pattern played a role in November's cholera risk pattern, as higher rainfall was observed in the northern part and western peninsula of Haiti.

- December

Though cholera cases were lower in December than in November, they followed the same direction to the north, with higher cholera cases with respect to December. The results matched with the temperature anomaly trend and can capture it correctly, and thus the correlation increased to 0.690. Temperature plays a role to create favorable conditions for pathogen growth, and low rainfall causes safe water accessibility issue, thus increasing cholera.

### 3.2.2. Model B

Model B uses six variables for October to December models—population density, elevation, rainfall anomaly, temperature anomaly, extreme rainfall, and wind speed during the hurricane (Figure 5). It incorporates two variables that are associated with the hurricane—extreme rainfall and wind speed.



**Figure 5.** (a) Input variables and cholera risk prediction of Model B for October 2016. (b) Spatio-temporal change of reported cholera cases and cholera risk index of Model B.

- September

In September 2016, before the hurricane, some places in both the western and northern parts of the country had high cholera risks including the capital, Port-au-Prince, where the population density per square kilometer is moderately high to very high. So, cholera risk in September was mainly influenced by higher population density. Model B cholera risk showed a very low ( $-0.099$ ) negative correlation with the departments' actual cholera cases, which implies that the model output failed to capture the spatial location of the actual cholera case occurrences. Except Port-au-Prince and nearby urban areas, the model overestimated the cholera risk. This model did not include variables related to the hurricane, i.e., extreme rainfall and wind speed.

- October

The high cholera risk area totally shifted to the western peninsula of Haiti, mostly in three departments—Grand'Anse, Sud, and Nippes—after Hurricane Matthew in October 2016. Windspeed during the hurricane mostly hit the western peninsula. So, the more to the west of the country, the higher the Hurricane Matthew's windspeed risk. The temperature-induced risk was moderately low for October. Rain-induced risk was very high mostly in Nippes and Sud department, and otherwise, it was moderate to low all over Haiti. After the hurricane, environmental variables played a stronger role than social variables like population density. Hurricane-induced extreme rainfall had a higher impact on all five departments in the southwestern part of the country—Grand'Anse, Sud, Nippes, Sud-Est, and Ouest—and the impact was reduced in the northern direction of the country.

Elevation did not play that much of a role in the October cholera risk except in the western part of the Artibonite department and Port-au-Prince. The spatial correlation of actual cholera cases in each department and cholera risk in October was 0.645. It underestimated the cholera risk in Grand'Anse and the western part of the Sud departments. The model overestimated the risk in the eastern part of Nippes, Sud, the western part of the Sud-Est and Ouest departments, and the Nord, Nord-Ouest, and Nord-Est departments. The model captured the risk well in the middle portion of the country.

- November

The spatial correlation of actual cholera cases in each department and cholera risk in November was 0.748. The model overestimated the risk in the eastern part of Nippes, the Sud, the western part of the Sud-Est and Ouest, and the Nord-Ouest and northwestern part of the Artibonite departments. The model predicted well in locations with a higher population density—Port-au-Prince. It underestimated the middle northern part of Haiti. The cholera risk of November showed a similar spatial pattern as October. The influence of wind, extreme rainfall during the hurricane, temperature, elevation, and the population at risk of cholera was like October. Rainfall influenced the higher cholera risk in Grand'Anse and Artibonite departments only. For other locations, lower rainfall in November and higher cholera risk occurred in similar spots. Scarcity of safe water after the hurricane along with low rainfall can be the cause of the inverse spatial pattern between rainfall and cholera risk.

- December

The spatial correlation of actual cholera cases in each department and cholera risk in December dropped from November (0.748) to 0.648. The prediction was close to the actual values for the Grand'Anse and Sud departments, though it underestimated the cholera risk in parts of the Sud and Grand'Anse. The model overestimated the risk in the eastern part of Nippes, the Sud, the western parts of the Sud-Est and Ouest, and the Nord-Ouest departments. The model failed to predict the higher cholera risk that occurred in the middle part of the Artibonite and Nord departments. It appears that rain did not play a role in December's cholera risk. However, after two months of Hurricane Matthew, the temperature seemed to elevate the risk in the southwestern parts of the country, particularly the western peninsula. There was extreme rainfall and high wind during the hurricane,

which caused flooding and destruction to the sanitation infrastructure, respectively. Such a condition along with the increasing temperature in the western part of Haiti creates a more favorable condition for pathogen growth, and the subsequent increase in cholera.

### 3.2.3. Model A Plus

All the plus models have variables of their primary name models (Model A and Model B) plus three other variables—cloud height, cloud top temperature, and building damage data.

- September

Model A Plus failed to capture the low cholera risk in most of the locations of the country, whereas it overestimated the risk in the western peninsula and resulted in a negative correlation of  $-0.398$ . Apart from wind speed during the hurricane—a fixed layer with respect to month—the cloud top temperature and cloud height played a role in the spatial pattern of the model output.

- October

A positive correlation ( $0.574$ ) was observed between actual cholera cases and the predicted cholera risk, which implies that the model successfully captured the spatial trends of cholera cases. Except for the western peninsula and Port-au-Prince and its surrounding areas, the model overpredicted the cholera risk. Apart from wind speed during the hurricane—a fixed layer with respect to month—the cloud top temperature, cloud height, and rainfall anomaly played a role in the spatial pattern of the model output. In the month of the hurricane, both cloud variables affect the model, whereas in the later months, the effect reduced gradually.

- November

The model predicted the spatial extend of cholera cases in the north direction well. So, the correlation value improved to  $0.687$ . Apart from wind speed during the hurricane—a fixed layer with respect to month—the cloud top temperature and rainfall anomaly played a role in the spatial pattern of the model output.

- December

The model fails to capture the northern higher cholera risk and overestimates the risk in the western peninsula. Thus, the correlation value drops to  $0.365$ . Temperature anomaly and windspeed played a role in the spatial pattern of the model's output. Cloud top temperature, cloud height, and building damage data did not affect the output that much.

### 3.2.4. Model B Plus

All the Plus models have variables of their primary name models (Model A and Model B) plus three other variables—cloud height, cloud top temperature, and building damage data.

- September

The correlation was highly negative ( $-0.692$ ), which means that the spatial pattern of the risk was opposite to the actual cholera case pattern.

- October

The model predicted the risk better than the previous month, and thus, the correlation changed from negative to positive and increased to  $0.512$ .

- November

The model predicted the risk better than the previous month, and thus, the correlations increased to  $0.570$ .

- December

The model failed to predict the cholera risk, and as a result, the correlation decreased to 0.364 from 0.570.

### 3.3. Improvement of Different Models from the Base Model A

It is implied from the computed Pearson correlation coefficients between the predicted and ground truth cholera risk that Model B is the best model among all four existing models, as it has the power of capturing the high monthly cholera risk after Hurricane Matthew. The model also showed the highest average correlation value. The accomplishment of all models for better cholera risk prediction with respect to the base model is represented in Table 5.

**Table 5.** Improvement offered by other models from the base model.

Month	Improvement from Base Model A (%)		
	Model A Plus	Model B	Model B Plus
October	−2.717	9.304	−13.108
November	5.807	15.262	−12.220
December	−47.109	−6.012	−47.303
Average (October–December)	−14.673	6.185	−24.210

Model B surpassed the base model prediction by 9.30% and 15.26% in October 2016 and November 2016, respectively. For December 2016, the base model predicted better cholera risk than model B. On average of the post-disaster months in 2016, model B was capable of predicting cholera risk 6.19% better than the base model. The Model B Plus results did not offer any better forecasting for all three months compared to the base model. Model A Plus only showed improvement (5.81%) from the base model in November 2016.

### 3.4. Cholera Risk Prediction Improvement Offered by the Machine Learning Model

Comparisons were made between Model B and the machine learning (ML) model. The monthly improvement (%) of the machine learning model from the best of all four other models is represented in Table 6.

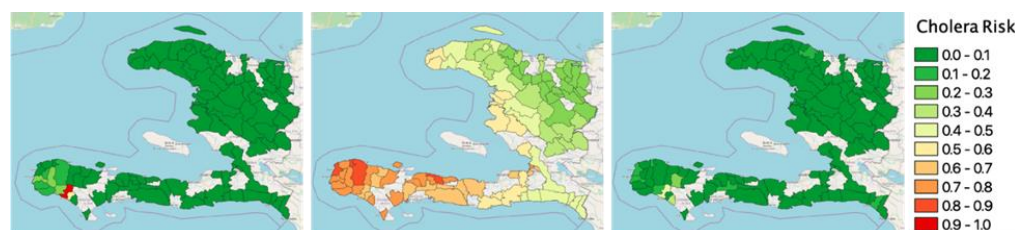
**Table 6.** Improvement (%) offered by the machine learning model from the best geospatial models.

Month	Correlation with Monthly Cholera Cases with Best Existing Model (Model B) Outputs	Correlation with Monthly Cholera Cases with ML Outputs	Improvement from Model B (%)
October	0.645	0.811	25.81
November	0.748	0.733	−2.00
December	0.648	0.744	14.74

Table 6 describes to what extent the ML model offers a better correlation with the actual cholera cases. The largest improvement of the ML model occurred when predicting cholera risk for October, a significant result considering that the largest number of cases was recorded in October (Table 3). The Model B output was better at estimating risk in November (one month after the hurricane), which is likely due to the environmental variables used in the model having sufficient lag time to reflect the altered landscape after Hurricane Matthew. It is not that the ML model provides less accuracy in November, but model B is more capable of capturing the risk with a 1-month lag better than the ML model. The ML model again produced more accurate results in December, demonstrating the promise of using ML to predict disease outbreak further into the future. Using higher-resolution earth observation datasets for multiple environmental variables makes the ML model more robust in forecasting cholera with better accuracy with ground data. Instead of

going with a single model approach, a combined model approach including both machine learning model and the geospatial model can be a smart and effective way for better cholera risk forecasting.

A spatial representation of the ground truth and estimated cholera risk for October can be seen in Figure 6. The existing model generally overestimated the cholera risk, especially in the eastern and northern parts of Haiti, but correctly showed higher levels of risk in the southwestern peninsula where the path of Hurricane Matthew passed through Haiti. The ML model predictions showed significant improvements, as they can identify input features that are most indicative of cholera risk. The ML model was better at not over-predicting risk compared to the existing geospatial model, Model B. In a real-world scenario, this improved accuracy would ensure that the required resources (e.g., cholera vaccines and rehydration solutions) can be optimally distributed to the most affected areas before an outbreak occurs.



**Figure 6.** Commune-level cholera risk ground truth (**left**), existing model output (**middle**), and ML model prediction (**right**) for October.

#### 4. Discussion

Machine learning has previously been used to forecast the outbreak of cholera in Yemen [71] and India [72]. The model used for Yemen (XGBoost, similar to the model reported in this paper) was trained over a dataset reporting cholera over an 8-week period in 2017 and 12 weeks of 2018. Rainfall and conflict data were used as predictors, and the results demonstrated an ability to predict cholera within an error of 5 deaths per 10,000 cases. The best performance was reported in the cases predicting an outbreak in the immediate aftermath (0–2 weeks), whereas the model’s performance degraded as the prediction window extended further out (6–8 weeks). A Random Forest classifier was trained on nine years of cholera data (2010 through 2018) using various climate variables, including sea surface temperature, sea surface salinity, and soil moisture. The model performed well, yielding a high accuracy (0.99), but due to an imbalance in their dataset (77 outbreak and 8504 non-outbreak data points), it cannot be fairly compared to the results obtained for Haiti. All these efforts show promising results and highlight the need for research into a model that trains over a diverse set of regions and for broad applicability. However, there may be region-specific factors that contribute to the overall prediction process, and that is a connection that can only be understood with a broader investigation that spans distinct geographical regions.

In this study, we used social and environmental variables for each month for the model of that month only. We did not use the lag effect of the environmental variables for creating a suitable environment for pathogen growth and accelerating cholera outbreaks. However, natural processes may take weeks to months to develop a cause–effect relationship and have a lagged effect on disease outbreak and transmission. Including different lag times in this analysis can be a future step. Additionally, in addition to the monthly risk maps, another model can be introduced using the epidemiological weeks. To convert data of different timestamps (daily, weekly, monthly), we took an average of the data to convert them to a monthly scale, i.e., cloud height and cloud top temperature. A weekly model can be more fruitful in a post-disaster scenario, which may be able to better capture the changes in extreme conditions and determine cholera risk more precisely.

Although there are 140 communes in Haiti, cholera data were only available for 103 communes, and those data were not catalogued in weekly or monthly formats but



aggregated to a total amount of patients from October to December. Having data in a weekly or monthly format for all communes can give us a better insight into the post-disaster outbreaks and allow for better identification of the variables and training for the ML model. The spatial resolution of the cloud top temperature and cloud height temperature is  $0.1^\circ \times 0.1^\circ$ , and for wind speed, the resolution is even more coarse at  $0.25^\circ \times 0.25^\circ$ . Better resolution data from earth observation sources can be more useful to monitor and capture the slightest change on a smaller scale.

Different combinations of variables and weightages were used in this study, as all variables that do not have the same contribution towards high cholera risk. In Model B, we used elevation, wind speed, and extreme rainfall during the hurricane, which are new additions from Model A. The influence of elevation and extreme rainfall together can be considered as flood impact, as lower elevation and higher rainfall together can pose a flood threat to the population, and these have an adverse impact on the water supply and sanitation system. Additionally, high wind speed can destroy houses and outdoor water supply and sanitation structures, which may lead to higher cholera cases. Though temperature and precipitation anomalies were used in both Model A and B, better spatiotemporal resolution data were used in Model B, which leads to better cholera risk prediction. In Model A Plus and Model B Plus, building damage due to hurricane, cloud height, and cloud top temperature was used. The same building damage layer was used for all months and the same respective cloud height and cloud top temperature were used for the corresponding months. Both cloud top temperature and cloud height/thickness were used, as these may have an impact on the precipitation process that is not explicit. The inclusion of lag in these cloud variables can potentially be useful for better model outputs.

Model A, the base model, was prepared based on the methodology described in [62]. However, the spatial resolution of the temperature data in that study was  $2.5^\circ \times 2.5^\circ$ , and thus, we used higher-resolution Land Temperature Analysis data ( $0.5^\circ \times 0.5^\circ$ ), a combination of the Global Historical Climatology Network (GHCN) and Climate Anomaly Monitoring System (CAMS), from the National Oceanic and Atmospheric Administration (NOAA). Global Precipitation Measurement (GPM) was used to obtain 2016 rainfall data, while Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analyses (TMPA) were used to calculate the long-term mean in the base model. Other geospatial models used IMERG data with better spatial resolution than used in a previous study [62]. Adding high-resolution data along with including extreme variables improved Model B's performance over the base model.

The models were compared based on their performances over the study months, and after taking the average of all months' correlations of model outputs and reported cholera cases, Model B had the highest average correlation among the four geospatial models. In the months immediately following the hurricane, the ML model captured the risk very well, but in later months, the geospatial model (Model B) performed better. A combined approach using both the geospatial and ML models would potentially predict the risk more accurately.

The primary beneficiaries of these models will be people related to health care facilities and government officials, who are responsible for providing relief, medical supplies, and vaccinations. If decision making is done by government officials according to the risk index, where 0 is the lowest cholera risk and 1 is the highest risk, then supplies can be provided to health facilities based on this index. As health facilities' data are not available for each pixel ( $1 \text{ km} \times 1 \text{ km}$ ), nearby pixel values can be averaged to determine the risk for each hospital for each health care facility.

## 5. Conclusions

Our models were able to successfully utilize earth observations of environmental variables to determine the spatio-temporal distribution of cholera risk in Haiti following the 2016 Hurricane Matthew landfall. The models captured cholera risk well in the middle portion of the country and especially the hurricane-affected western peninsula. As the

models were designed for extreme conditions, like a hurricane, the models performed better in locations near the path of Hurricane Matthew, where environmental variables showed extreme patterns. The models were less effective in capturing the cholera risk in locations farther from the hurricane's path. The models overestimated the risk in various areas of the Nord, Nord-Ouest, and Nord-Est departments. Though the eastern part of Nippes and the Sud department are situated near the hurricane path, the models did not reflect their risk index properly. As elevation, rainfall, and extreme rainfall together raise the possibility of floods, high risk in any of these variables may have caused overestimation. The results strengthen our hypothesis behind selecting variables like elevation, rainfall, and extreme rainfall, which lower elevations, and higher and extreme rainfall may rapidly cause floods in post-disaster situations, which will worsen the situation with more cholera cases, as water will be contaminated due to flooding. Our results will help develop waterborne disease forecasting frameworks to identify vulnerable regions in post-disaster settings and help to guide efforts to plan health interventions and recovery efforts.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cli10040048/s1>. Table S1. Risk level classification for Model A (base model). Table S2. Risk level classification for Model B. Table S3. Risk level classification for Model A Plus. Table S4. Risk level classification for Model B Plus.

**Author Contributions:** Conceptualization, F.N., M.H. and A.S.A.; methodology, F.N., M.H. and D.R.; software, F.N. and D.R.; validation, F.N., D.R. and A.S.A.; formal analysis, F.N., M.H., D.R., G.C. and A.S.A.; investigation, F.N., D.R. and A.S.A.; resources, M.H. and A.S.A.; data curation, F.N., M.H., D.R. and G.C.; writing—original draft preparation, F.N. and D.R.; writing—review and editing, M.H. and A.S.A.; visualization, F.N., D.R. and G.C.; supervision, M.H. and A.S.A.; project administration, M.H. and A.S.A.; funding acquisition, M.H. and A.S.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** F.N. and A.S.A. gratefully acknowledge the support of the Rhode Island Water Resources Center and the Dean of the College of Engineering, University of Rhode Island.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Mal, S.; Singh, R.B.; Huggel, C.; Grover, A. Introducing Linkages between Climate Change, Extreme Events, and Disaster Risk Reduction. In *Climate Change, Extreme Events and Disaster Risk Reduction*; Springer: New York, NY, USA, 2018; pp. 1–14.
2. Trenberth, K.E.; Fasullo, J.T.; Shepherd, T.G. Attribution of climate extreme events. *Nat. Clim. Chang.* **2015**, *5*, 725–730. [[CrossRef](#)]
3. Natural Disasters Are Increasing. The World's Poorest Are Left to Fend for Themselves. Available online: <https://www.vox.com/21571842/coronavirus-pandemic-climate-change-covid-19-natural-disaster-vaccine> (accessed on 3 March 2022).
4. Sawada, Y.; Takasaki, Y. Natural disaster, poverty, and development: An introduction. *World Dev.* **2017**, *94*, 2–15. [[CrossRef](#)]
5. World Bank; United Nations. *Natural Hazards, Unnatural Disasters: The Economics of Effective Prevention*; The World Bank: Washington, DC, USA, 2011; Volume 48; ISBN 0821380508.
6. Cash, R.A.; Halder, S.R.; Husain, M.; Islam, M.S.; Mallick, F.H.; May, M.A.; Rahman, M.; Rahman, M.A. Reducing the health effect of natural hazards in Bangladesh. *Lancet* **2013**, *382*, 2094–2103. [[CrossRef](#)]
7. Charnley, G.E.C.; Kelman, I.; Gaythorpe, K.A.M.; Murray, K.A. Traits and risk factors of post-disaster infectious disease outbreaks: A systematic review. *Sci. Rep.* **2021**, *11*, 5616. [[CrossRef](#)] [[PubMed](#)]
8. Hackbarth, M.; Pavkov, T.; Wetchler, J.; Flannery, M. Natural Disasters: An Assessment of Family Resiliency Following Hurricane Katrina. *J. Marital Fam. Ther.* **2012**, *38*, 340–351. [[CrossRef](#)]
9. Freedy, J.R.; Shaw, D.L.; Jarrell, M.P.; Masters, C.R. Towards an understanding of the psychological impact of natural disasters: An application of the conservation resources stress model. *J. Trauma. Stress* **1992**, *5*, 441–454. [[CrossRef](#)]
10. Kreimer, A. Social and economic impacts of natural disasters. *Int. Geol. Rev.* **2001**, *43*, 401–405. [[CrossRef](#)]
11. Chigudu, S. The politics of cholera, crisis and citizenship in urban Zimbabwe: "People were dying like flies". *Afr. Aff.* **2019**, *118*, 413–434. [[CrossRef](#)]

12. Colwell, R. Cholera and climate: A demonstrated relationship. *Trans. Am. Clin. Climatol. Assoc.* **2009**, *120*, 119–128.
13. Ali, M.; Emch, M.; Donnay, J.P.; Yunus, M.; Sack, R.B. Identifying environmental risk factors for endemic cholera: A raster GIS approach. *Health Place* **2002**, *8*, 201–210. [[CrossRef](#)]
14. Ali, M.; Emch, M.; Donnay, J.P.; Yunus, M.; Sack, R.B. The spatial epidemiology of cholera in an endemic area of Bangladesh. *Soc. Sci. Med.* **2002**, *55*, 1015–1024. [[CrossRef](#)]
15. Hashizume, M.; Armstrong, B.; Hajat, S.; Wagatsuma, Y.; Faruque, A.S.G.; Hayashi, T.; Sack, D.A. Association between climate variability and hospital visits for non-cholera diarrhoea in Bangladesh: Effects and vulnerable groups. *Int. J. Epidemiol.* **2007**, *36*, 1030–1037. [[CrossRef](#)] [[PubMed](#)]
16. Penrose, K.; De Castro, M.C.; Werema, J.; Ryan, E.T. Informal urban settlements and cholera risk in Dar es Salaam, Tanzania. *PLoS Negl. Trop. Dis.* **2010**, *4*, e631. [[CrossRef](#)]
17. Rajendran, K.; Sumi, A.; Bhattachariya, M.K.; Manna, B.; Sur, D.; Kobayashi, N.; Ramamurthy, T. Influence of relative humidity in *Vibrio cholerae* infection: A time series model. *Indian J. Med. Res.* **2011**, *133*, 138–145. [[PubMed](#)]
18. Reiner, R.C.; King, A.A.; Emch, M.; Yunus, M.; Faruque, A.S.G.; Pascual, M. Highly localized sensitivity to climate forcing drives endemic cholera in a megacity. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 2033–2036. [[CrossRef](#)] [[PubMed](#)]
19. Talavera, A.; Pérez, E.M. Is cholera disease associated with poverty? *J. Infect. Dev. Ctries.* **2009**, *3*, 408–411. [[CrossRef](#)] [[PubMed](#)]
20. Akanda, A.S.; Aziz, A.; Jutla, A.; Huq, A.; Alam, M.; Asham, G.U.; Colwell, R.R. Satellites and cell phones form a cholera early-warning system. *Eos Trans. Am. Geophys. Union* **2018**, *99*. [[CrossRef](#)]
21. Ali, M.; Nelson, A.R.; Lopez, A.L.; Sack, D.A. Updated global burden of cholera in endemic countries. *PLoS Negl. Trop. Dis.* **2015**, *9*, e0003832. [[CrossRef](#)] [[PubMed](#)]
22. Parry, M.L.; Canziani, O.; Palutikof, J.; Van der Linden, P.; Hanson, C. *Climate Change 2007-Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Fourth Assessment Report of the IPCC*; Cambridge University Press: Cambridge, UK, 2007; Volume 4; ISBN 0521880106.
23. Liu, Z.D.; Li, J.; Zhang, Y.; Ding, G.Y.; Xu, X.; Gao, L.; Liu, X.N.; Liu, Q.Y.; Jiang, B.F. Distributed lag effects and vulnerable groups of floods on bacillary dysentery in Huaihua, China. *Sci. Rep.* **2016**, *6*, 29456. [[CrossRef](#)] [[PubMed](#)]
24. Flooding and Communicable Diseases Fact Sheet. Available online: <https://www.who.int/hac/techguidance/ems/FloodingandCommunicableDiseasesfactsheet.pdf> (accessed on 3 March 2022).
25. Fewtrell, L.; Kay, D.; Watkins, J.; Davies, C.; Francis, C. The microbiology of urban UK floodwaters and a quantitative microbial risk assessment of flooding and gastrointestinal illness. *J. Flood Risk Manag.* **2011**, *4*, 77–87. [[CrossRef](#)]
26. Levy, K.; Woster, A.P.; Goldstein, R.S.; Carlton, E.J. Untangling the Impacts of Climate Change on Waterborne Diseases: A Systematic Review of Relationships between Diarrheal Diseases and Temperature, Rainfall, Flooding, and Drought. *Environ. Sci. Technol.* **2016**, *50*, 4905–4922. [[CrossRef](#)]
27. World Water Development Report 2018. Available online: <https://unesdoc.unesco.org/ark:/48223/pf0000261424> (accessed on 3 March 2022).
28. Nusrat, F.; Akanda, A.S.; Johnson, K.; Khan, A.I.; Islam, M.T.; Rashid, M.M.; Qadri, F. Understanding the Water Access and Waterborne diseases vulnerability of women in the context of gender biased social norms. In Proceedings of the AGU Fall Meeting 2021, New Orleans, LA, USA, 13–17 December 2021.
29. Effler, P.; Isaäcson, M.; Arntzen, L.; Heenan, R.; Canter, P.; Barrett, T.; Lee, L.; Mambo, C.; Levine, W.; Zaidi, A.; et al. Factors contributing to the emergence of *Escherichia coli* O157 in Africa. *Emerg. Infect. Dis.* **2001**, *7*, 812–819. [[CrossRef](#)] [[PubMed](#)]
30. Moors, E.; Singh, T.; Siderius, C.; Balakrishnan, S.; Mishra, A. Climate change and waterborne diarrhoea in northern India: Impacts and adaptation strategies. *Sci. Total Environ.* **2013**, *468–469*, S139–S151. [[CrossRef](#)]
31. Rieckmann, A.; Tamason, C.C.; Gurley, E.S.; Rod, N.H.; Jensen, P.K.M. Exploring droughts and floods and their association with cholera outbreaks in sub-saharan africa: A register-based ecological study from 1990 to 2010. *Am. J. Trop. Med. Hyg.* **2018**, *98*, 1269–1274. [[CrossRef](#)] [[PubMed](#)]
32. Senhorst, H.A.J.; Zwolsman, J.J.G. Climate change and effects on water quality: A first impression. *Water Sci. Technol.* **2005**, *51*, 53–59. [[CrossRef](#)] [[PubMed](#)]
33. Ahern, M.; Kovats, R.S.; Wilkinson, P.; Few, R.; Matthies, F. Global health impacts of floods: Epidemiologic evidence. *Epidemiol. Rev.* **2005**, *27*, 36–46. [[CrossRef](#)] [[PubMed](#)]
34. Shultz, J.M.; Russell, J.; Espinel, Z. Epidemiology of tropical cyclones: The dynamics of disaster, disease, and development. *Epidemiol. Rev.* **2005**, *27*, 21–35. [[CrossRef](#)]
35. Watson, J.T.; Gayer, M.; Connolly, M.A. Epidemics after natural disasters. *Emerg. Infect. Dis.* **2007**, *13*, 1–5. [[CrossRef](#)]
36. Bhunia, R.; Ghosh, S. Waterborne cholera outbreak following Cyclone Aila in Sundarban area of West Bengal, India, 2009. *Trans. R. Soc. Trop. Med. Hyg.* **2011**, *105*, 214–219. [[CrossRef](#)]
37. Bhunia, R.; Ramakrishnan, R.; Hutin, Y.; Gupte, M.D. Cholera outbreak secondary to contaminated pipe water in an urban area, West Bengal, India, 2006. *Indian J. Gastroenterol.* **2009**, *28*, 62–64. [[CrossRef](#)]
38. Sur, D.; Dutta, P.; Nair, G.B.; Bhattacharya, S.K. Severe cholera outbreak following floods in a northern district of West Bengal. *Indian J. Med. Res.* **2000**, *112*, 178–182. [[PubMed](#)]
39. Taneja, N.; Kaur, J.; Sharma, K.; Singh, M.; Kalra, J.K.; Sharma, N.M.; Sharma, M. A recent outbreak of cholera due to *Vibrio cholerae* O1 Ogawa in & around Chandigarh, North India. *Indian J. Med. Res.* **2003**, *117*, 243–246.

40. Akanda, A.S.; Jutla, A.S.; Alam, M.; De Magny, G.C.; Siddique, A.K.; Sack, R.B.; Huq, A.; Colwell, R.R.; Islam, S. Hydroclimatic influences on seasonal and spatial cholera transmission cycles: Implications for public health intervention in the Bengal Delta. *Water Resour. Res.* **2011**, *47*, W00H07. [CrossRef]
41. Huq, A.; West, P.A.; Small, E.B.; Huq, M.I.; Colwell, R.R. Influence of water temperature, salinity, and pH on survival and growth of toxigenic *Vibrio cholerae* serovar O1 associated with live copepods in laboratory microcosms. *Appl. Environ. Microbiol.* **1984**, *48*, 420–424. [CrossRef]
42. Cambaza, E.; Mongo, E.; Anapakala, E.; Nhambire, R.; Singo, J.; Machava, E. Outbreak of cholera due to cyclone kenneth in northern Mozambique, 2019. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2925. [CrossRef] [PubMed]
43. 2 Years Since Cyclone Idai and Mozambique Has Already Faced an Additional 3 Cyclones. Available online: <https://reliefweb.int/report/mozambique/2-years-cyclone-idai-and-mozambique-has-already-faced-additional-3-cyclones> (accessed on 3 March 2022).
44. Cholera Is Spreading in Mozambique in the Wake of Cyclone Idai. Available online: <https://www.vox.com/science-and-health/2019/3/29/18287342/mozambique-cyclone-idai-cholera-how-to-help> (accessed on 3 March 2022).
45. Chen, W.H.; Azman, A.S. Mitigating cholera in the aftermath of cyclone Idai. *Am. J. Trop. Med. Hyg.* **2019**, *101*, 960. [CrossRef]
46. India: Cyclone Aila Situation Update, 30 June 2009. Available online: <https://reliefweb.int/report/india/india-cyclone-aila-situation-update-30-jun-2009> (accessed on 3 March 2022).
47. ECHO Factsheet–Haiti–April 2018. Available online: <https://reliefweb.int/report/haiti/echo-factsheet-haiti-april-2018> (accessed on 3 January 2022).
48. HAITI Humanitarian Situation Report #16. Available online: <https://reliefweb.int/report/haiti/unicef-haiti-humanitarian-situation-report-hurricane-matthew-16-15-december-2016> (accessed on 2 January 2022).
49. Organization United Nations Foundation. Agriculture Haiti: Hurricane Matthew SITUATION REPORT–22 November 2016. Available online: <https://www.fao.org/emergencies/recursos/documentos/recursos-detalle/es/c/454208/> (accessed on 3 January 2022).
50. Haiti: Hurricane Matthew–Situation Report No. 14 (21 October 2016). Available online: <https://reliefweb.int/report/haiti/haiti-hurricane-matthew-situation-report-no-14-21-october-2016> (accessed on 3 March 2022).
51. Haiti: Humanitarian Situation Report End of Yea–2018. Available online: <https://reliefweb.int/report/haiti/haiti-humanitarian-situation-report-end-year-2018> (accessed on 4 January 2022).
52. ECHO Factsheet–Haiti–(Last Updated 21/05/2019). Available online: <https://reliefweb.int/report/haiti/echo-factsheet-haiti-last-updated-21052019> (accessed on 5 January 2022).
53. Mozambique: Two Cyclones: Idai and Kenneth (As of 22 May 2019). Available online: <https://reliefweb.int/report/mozambique/mozambique-two-cyclones-idai-and-kenneth-22-may-2019> (accessed on 3 March 2022).
54. UNICEF Mozambique Cyclones Idai and Kenneth Situation Report #11 (1–10 June 2019). Available online: <https://reliefweb.int/report/mozambique/unicef-mozambique-cyclones-idai-and-kenneth-situation-report-11-01-10-june-2019> (accessed on 16 December 2021).
55. UNICEF Cyclone Idai Post-Impact Situation April 2019. Available online: <https://reliefweb.int/report/mozambique/unicef-cyclone-idai-post-impact-situation-april-2019> (accessed on 1 March 2022).
56. Mozambique Cyclone Idai Response: Situation Report #6 21–30 April 2019. Available online: <https://reliefweb.int/report/mozambique/mozambique-cyclone-idai-response-situation-report-6-21-30-april-2019> (accessed on 19 December 2021).
57. Barzilay, E.J.; Schaad, N.; Magloire, R.; Mung, K.S.; Boncy, J.; Dahourou, G.A.; Mintz, E.D.; Steenland, M.W.; Vertefeuille, J.F.; Tappero, J.W. Cholera Surveillance during the Haiti Epidemic—The First 2 Years. *N. Engl. J. Med.* **2013**, *368*, 599–609. [CrossRef] [PubMed]
58. Guillaume, Y.; Ternier, R.; Vissieres, K.; Casseus, A.; Chery, M.J.; Ivers, L.C. Responding to cholera in Haiti: Implications for the national plan to eliminate cholera by 2022. *J. Infect. Dis.* **2018**, *218*, S167–S170. [CrossRef]
59. Enserink, M. Cholera linked to U.N. forces, but questions remain. *Science* **2011**, *332*, 776–777. [CrossRef]
60. Haiti Cholera Response December 2014. Available online: <https://reliefweb.int/report/haiti/haiti-cholera-response-december-2014> (accessed on 29 November 2021).
61. Haque, M.; Rollend, D.; Withee, S.; Christie, G.; Papusha, I.; Nikoukar, R.; Chen, M.; Nusrat, F.; Akanda, A.S. Continuum: A New Observing Strategies-Based Framework for Hyper-Local Situational Awareness. In Proceedings of the AGU Fall Meeting, Virtual Online, 1–17 December 2020.
62. Khan, R.; Anwar, R.; Akanda, S.; McDonald, M.D.; Huq, A.; Jutla, A.; Colwell, R. Assessment of risk of cholera in Haiti following Hurricane Matthew. *Am. J. Trop. Med. Hyg.* **2017**, *97*, 896–903. [CrossRef]
63. Nusrat, F.; Akanda, A.S.; Haque, M.; Rollend, D.; Christie, G. Waterborne Disease Outbreaks in the Post-Disaster Scenario: Can High-Resolution Earth Observations and Smallsat Imagery Enhance Cholera Forecasting? In Proceedings of the AGU Fall Meeting, Virtual Online, 1–17 December 2020.
64. Haiti Faces Fresh Cholera Outbreak after Hurricane Matthew, Aid Agencies Fear. Available online: <https://www.theguardian.com/world/2016/oct/14/haiti-cholera-hurricane-matthew-aid-agencies> (accessed on 20 March 2022).
65. NASA Global Maps. Available online: <https://earthobservatory.nasa.gov/global-maps> (accessed on 3 March 2022).
66. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; Volume 2016, pp. 770–778.

67. Gupta, R.; Goodman, B.; Patel, N.; Hosfelt, R.; Sajeev, S.; Heim, E.; Doshi, J.; Lucas, K.; Choset, H.; Gaston, M. Creating XBD: A dataset for assessing building damage from satellite imagery. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 15–20 June 2019; Volume 2019, pp. 10–17.
68. Digital Globe Open Data Program: Hurricane Matthew. Available online: <https://www.digitalglobe.com/ecosystem/open-data/hurricane-matthew> (accessed on 20 March 2022).
69. Suspected Cholera Cases by Commune in Haiti Associated to Matthew Hurricane. Available online: [http://ais.paho.org/hip/viz/ed\\_haiticoleracases.asp](http://ais.paho.org/hip/viz/ed_haiticoleracases.asp) (accessed on 3 March 2022).
70. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
71. Campbell, A.M.; Racault, M.-F.; Goult, S.; Laurenson, A. Cholera risk: A machine learning approach applied to essential climate variables. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9378. [[CrossRef](#)] [[PubMed](#)]
72. Badkundri, R.; Valbuena, V.; Pinnamareddy, S.; Cantrell, B.; Standeven, J. Forecasting the 2017–2018 Yemen cholera outbreak with machine learning. *arXiv* **2019**, arXiv:1902.06739 2019.