Early Action for Cholera Project

Yemen Case Study

Met Office, University of Florida & University of Maryland

Authors:

Met Office: Rosa Barciela, Tarkan Bilge, Kate Brown, Adrian Champion, Christophe Sarran, Maxine Shields, Helen Ticehurst; **University of Florida**: Antar Jutla, Moiz Usmani; **University of Maryland**; Rita Colwell

Reviewers: Helen Bye, Kathrin Hall, Tim Donovan (Met Office)

02 July 2021

Contents

Glossary

Acronym	Meaning
BHC	Bradford Hill Criteria
CERF	Central Emergency Response Funds
CAM	Crisis Area Model; one of the Met Office's limited area numerical weather prediction models, using boundary conditions from the Global Model.
Case data	Information on cholera cases
CERF	Central Emergency Response Fund (run by UN OCHA)
COF	Climate Outlook Forum
COVID	Coronavirus disease (COVID-19)
CRM	Cholera Risk Model
DFID	Department for International Development
DRR	Disaster Risk Reduction
DREF	Disaster Relief Emergency Fund
EACH	Early Action for Cholera
ECMWF	European Centre for Medium-Range Weather Forecasts
EOC	Emergency Operations Centre
EOR	Emergency Operations Room
EMRO	Eastern Mediterranean Regional Office
Epi data	Epidemiological data sets
EWARS	Early Warning, Alert and Response System
FbA	Forecast Based Action
FBS	Fractions Brier Score; a score function to quantify the accuracy of a forecast, corresponds to the mean square difference between forecast and satellite fields, used in calculation of the FSS.
FCDO	Foreign, Commonwealth & Development Office
FSS	Fractions Skill Score; a score for assessing the skill of precipitation forecasts by evaluating the forecast skill across different spatial scales (Roberts and Lean 2008).
FSS_{ufc}	Fractions Skill Score Useful Forecast Criterion; defined as being the FSS halfway between a random forecast and a perfect forecast, this is half the base rate plus 0.5, which can be approximated to 0.5 when dealing with low base rates.
FTP	File transfer protocol
GFS	Global Forecast System, produced by NCEP
GGU	Global Guidance Unit (Met Office)
GM	Global Model; refers to the global configuration of the Met Office's Unified Model and the global land and global atmospheric science configurations.
GPM	Global Precipitation Measurement; a joint mission between NASA and JAXA aimed at studying precipitation around the globe using an international network of satellites, active 2014-present.
GTFCC	Global Task Force on Cholera Control
IMERG	Integrated Multi-satellitE Retrievals for GPM; algorithm that produces the multi-satellite precipitation product.
ITCZ	Intertropical Convergence Zone; a band of low pressure caused by the convergence of the Northern hemisphere and Southern hemisphere trade winds, often accompanied by a band of precipitation.
IV	Intravenous
JMA	Japanese Meteorological Agency
JOF	Joint Operational Framework
MEL	Monitoring, Evaluation and Learning
MO	Met Office
MOH	Ministry of Health

MoWE	Ministry of Water and Environment
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NGO	Non-Governmental Organisation
NMHS	National Meteorological and Hydrological Services
NWP	Numerical Weather Prediction
OCHA	(United Nations) Office for the Coordination of Humanitarian Affairs
OCV	Oral Cholera Vaccine
Precipitation	A scientific definition of all moisture that falls from the atmosphere, including rainfall, hail and snow. This is the field used by the forecast models analysed in this study.
Rainfall	Precipitation that falls as liquid to the ground and is the form of precipitation that is detected by satellites.
REACH	REACH is a leading humanitarian initiative providing granular data, timely information and in-depth analysis from contexts of crisis, disaster and displacement.
REAP	Risk-informed Early Action Partnership launched at the Climate Action Summit in 2019
RMSE	Root-mean square error
RRT	Rapid Response Team - UNICEF’s teams of volunteers who provide hygiene and sanitation training.
RSCZ	Red Sea Convergence Zone; a convergence of air masses in the Red Sea where north-westerlies from the Mediterranean meet south-easterlies from the Gulf of Aden, often causing precipitation.
SOP	Standard Operating Procedure
UF	University of Florida
UM	Unified Model; the Met Office’s numerical model used for both weather and climate applications, the GM and CAM used in this analysis are both suites of the Unified Model.
UMD	University of Maryland
UN	United Nations
UNICEF	United Nations Children's Fund
UN IOM	United Nations International Organisation for Migration
WASH	Water Sanitation and Hygiene
WHO	World Health Organisation
WMO	World Meteorological Organization
WRA	Weekly Rainfall Assessment or Yemen Rainfall Assessment

Executive Summary

Cholera is a highly contagious diarrhetic infection which is transmitted through the consumption of contaminated water or food. If untreated, it can kill within a matter of hours. Regarded as a disease of inequality, cholera presents a major threat to lower income countries with poor quality drinking water and sanitation systems. Natural disasters and conflict also increase susceptibility to outbreaks of the disease.

Before the war in Yemen, which started in 2016, the country was predisposed to cholera due to its high levels of poverty, frequent droughts, and poor sanitation infrastructure. On top of this, the widespread displacement of people, food shortages and other conflict-related impacts led to one of the worst cholera epidemics in modern times; between October 2016 and January 2020, over 2.3 million cases and nearly 4,000 deaths were reported.

UNICEF is one of the key actors in cholera prevention and response in Yemen and coordinate water, sanitation, and hygiene (WASH) interventions. These include the provision of safe drinking water and sending teams of volunteers into communities to run hygiene awareness campaigns.

The association between environmental factors and cholera has been established (Camacho et al 2018, Eisenburg et al, 2013, Hashizume et al, 2008) so the use of a tool was proposed to help prioritise cholera interventions ahead of the rainy season (which is associated with cholera outbreaks). Additionally, at the beginning of 2018 the Houthi run Ministry of Health proposed a change in the case definition of cholera which risked delaying a response to any outbreak. Additional sources of risk were therefore seen as essential to ensuring the response remained dynamic. Supported by the UK's Foreign, Commonwealth and Development Office (FCDO), UNICEF began to receive weekly reports from the University of Florida's (UF) Cholera Risk Model (CRM) in 2018. They also started receiving weekly rainfall forecasts from the Met Office in 2018.

The CRM provides an indication of cholera risk which is valid for 4 weeks (from issue date). Based upon rainfall and temperature data, information on population density and movement and, (where available) WASH data, the model's algorithm then calculates a risk score for cholera. This is presented in a series of maps along with a brief description on how to interpret the risk values.

The Met Office provides rainfall information to users in Yemen on a weekly basis. This includes a 7-day hindcast, a 7-day forecast, a 4-week forward outlook, and a summary highlighting high-impact weather. It also includes maps showing the spatial distribution of rainfall and tables giving forecast rainfall, by category, for specific locations around the country.

Cholera monitoring and response is coordinated in Yemen by an Emergency Operation Centre which prepares a table of the administrative districts most affected by cholera. The CRM risk scores and rainfall forecasts are considered alongside this data. The districts are then ranked into low-high risk categories according to where there are cases already and where predictions suggest these will increase. The most appropriate action to take in high-risk districts is then identified, based on local contexts.

A significant drop in cholera cases was observed in Yemen during 2018. For example, during one week in 2018, there were 2,500 cases, compared to 50,000 during the same week in 2017. The drop in cases was attributed anecdotally by UNICEF to the forecast based early intervention actions they had been taking, using the information from UF and the Met Office.

The approach taken by UNICEF represents a novel way of tackling infectious diseases by bringing interventions forward, using predictive tools. Whilst well established in humanitarian contexts, the concept of 'anticipatory / early action' is nascent in the field of cholera control.

To understand whether the continued use of these tools in Yemen is appropriate, and to explore the scalability of the approach to other countries, the validity of the CRM and rainfall forecasts in Yemen was assessed.

To validate the CRM, its predictions in 2017, 2018 and 2019 were compared to recorded cases of cholera in Yemen. In the most populous governorates (comprising about 80% Yemeni population), the CRM's predictions were accurate 60% of the time. Assessments of the CRM's performance in other countries also supports these findings. Analysis of sensitivity, specificity, accuracy and precision, and negative predictive value, indicate changes in model risk scores predict change in number of cholera cases locally in Yemen. The CRM had the highest accuracy in 2017, followed by 2019 and lastly 2018. Cholera has occurred consistently in Yemen each week from 2017 to 2019 and this suggests cholera is becoming endemic.

The precipitation from the Met Office forecast models that are used in the weekly rainfall assessments was validated against observation data. As there was an absence of in-situ rainfall observations in Yemen, satellite derived rainfall observations were used to investigate model accuracy. The forecast rainfall was given one of five categories, from light rain to storm, each with a specific threshold. The analysis showed that light rain was typically forecast to be within 11km of the observation, whereas for heavy rain, the location accuracy was at least 160km. The accuracy of the Met Office Global Model is higher, or similar, to models from other National Weather Centres.

It was found that the number of new cholera cases is weakly correlated to forecast rainfall. The statistical modelling suggests that targeted interventions based on the weekly rainfall assessments may have reduced the number of cholera cases, however more data would be needed to validate this.

Based upon the validation of the CRM and rainfall forecasts in the contextual understanding of how these tools are used by UNICEF, the following recommendations are made:

Key Recommendations on how the CRM and rainfall forecasts could continue to be used in Yemen and elsewhere

How the CRM and rainfall forecasts can be used to inform early action:

- In areas where epidemic cholera is expected, preventative measures will most likely already be underway in anticipation of outbreaks. In these contexts, the CRM and rainfall forecasts should be used to **inform planning and preparation activities** and to **intensify early control measures** such as surveillance and reporting, strengthening healthcare systems and community engagement.
- Specific actions that can be informed by the CRM and rainfall forecasts **will vary by context/use case** and should be identified with cholera response stakeholders and providers of the CRM and rainfall forecasts through a process of co-design.
- As Yemen represents the only pilot in which the CRM and rainfall forecasts are used in an operational context (and a relatively extreme one), **further pilots are needed** to test and inform the development of this approach. Sharing insight and learnings from these pilots, with the wider anticipatory action sphere and cholera response community, will be key in enabling the concept of early action in the cholera domain, (which is in its infancy), to develop. Through demonstrating the value of using risk information, the pilots may also have potential to inform implementation of the Early Detection and Surveillance pillar of the Global Cholera Task Force's Roadmap to Ending Cholera by 2030.

Strengthening the CRM and rainfall forecasts:

- Inclusion of the CRM in a new the newly formed WHO Infections Disease (ID) modelling inter-comparison pilot study would enable potential users to understand how and where it can add most value to their decision-making.
- Development of **communications materials** that describe the evidence base and evolution of the CRM to potential users and provide guidance in how to understand and

interpret its outputs would make it easier to use. **Access to the tool** could be improved through the development of a web-based platform.

- The ability to **test water for the *Vibrio Cholerae* bacteria** in areas where the CRM has predicted cholera would provide a 'ground truthing' of the model's predictions and provide further justification for anticipatory action.
- Enhancing understanding of the relationship between rainfall and cholera at local levels will help to determine **where rainfall forecasts are most relevant.**
- Understanding the impacts of inter-annual events such as El Niño and other seasonal patterns on cholera could provide even earlier indications of cholera risk.

Figure ES1 illustrates how using the CRM and rainfall forecasts could influence the epidemic curve of a hypothetical cholera outbreak.

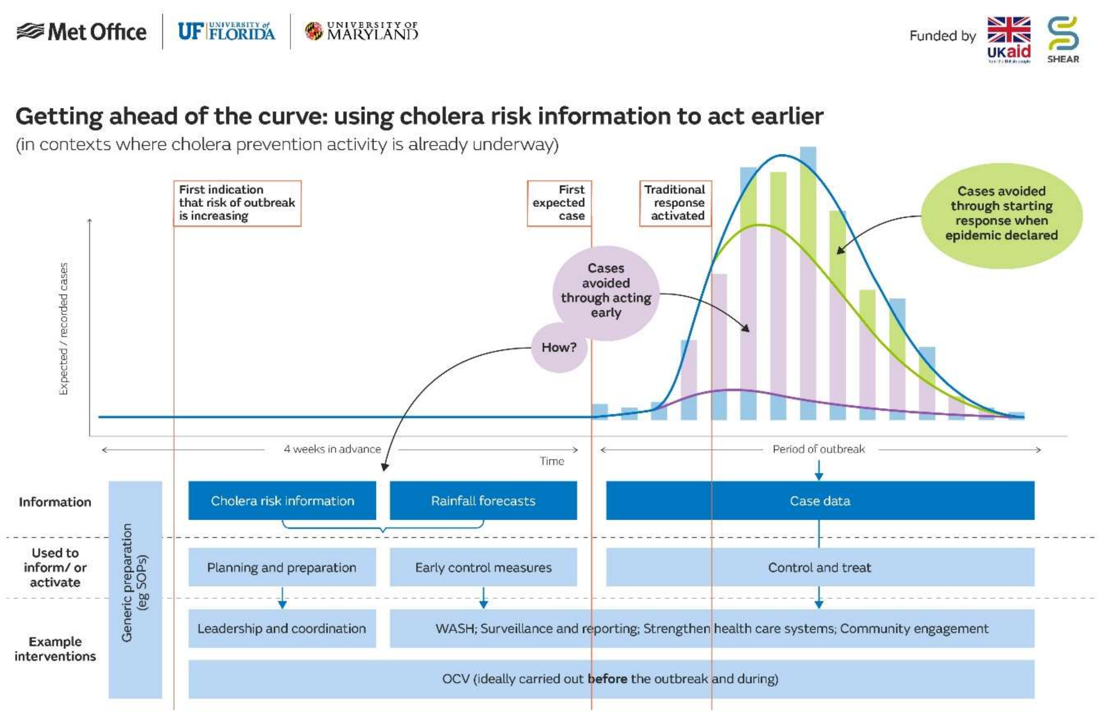


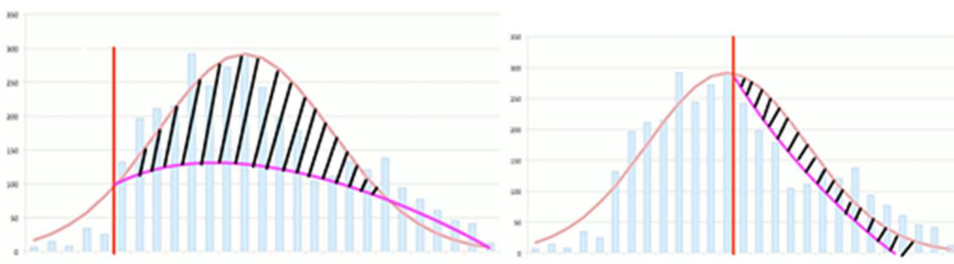
Figure ES1: Proposed use of the CRM and rainfall forecasts to inform cholera decision making and the impact this can have on the epidemic curve.

UNICEF's approach in Yemen represents a primary and promising example of how risk information can inform cholera response. This paper describes how the CRM and rainfall forecasts were used in Yemen and presents validation work on their performance. It makes recommendations on how such tools should be used with the aim of furthering conversation on how a combination of data and models could be used to impact the course of an epidemic and reduce the suffering caused by cholera.

1. Background: Predicting Cholera in Advance

1.1 Prologue

During an infectious disease outbreak, the timing of the response is almost everything. This is true for prevention case detection/treatment, and for control measures. Early prevention, detection and treatment substantially reduces case fatality rates. With regards to control measures, to make a substantial difference to the course of an epidemic, the response needs to be ahead of the epidemic curve.



Scenario A Scenario B

Figure 1a: Schematic representation of the same cholera control measures implemented at the beginning (Scenario A) and after the peak (Scenario B) of an outbreak, and potential cases averted. [Y–axis = incidence of new cases, X-axis = time]. (Source UNICEF Evaluation of level 3 response to cholera epidemic in Yemen: A crises within a crises, 2018)

After the peak of the outbreak, the effect of the response (even if well designed and implemented) will likely be marginal, or at least will have far less effect than if implemented earlier in the outbreak. The difference may be measured in days. It is important to note that during an outbreak, there is not only one curve, but rather many small curves in multiple locations and at different levels (at governorate, district, and even village level). To be most effective, the control response needs to get ahead of multiple epidemic curves, as shown in figure 1 c

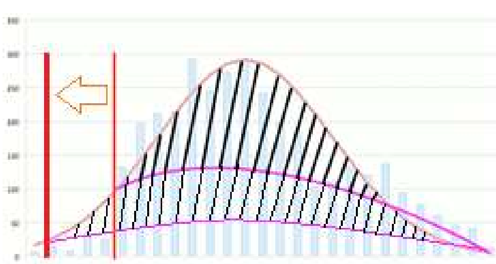


Figure 1b: Getting ahead of multiple epidemic curves (Source: FDCO internal presentation, 2018)

Responding quickly enough to every new outbreak in every new location is very difficult – and can only be achieved through agile and mobile rapid response teams with a very high level of organisation, using recently gathered data as a basis for targeted interventions. Preparedness is crucial to effective response – without having such systems, roles and capacities pre-established, the response is always likely to lag behind the epidemic curve.

This report presents a case in which a prediction model for a highly infectious disease, cholera, is validated using near real time data from Yemen.

1.2 Cholera, its treatment and prevention

Cholera is an infection of the small intestine by strains of the bacterium called *Vibrio cholerae*. It is spread mostly by unsafe water and food that has been contaminated with human faeces containing the bacteria. Symptoms may range from mild to severe with acute cases causing severe watery diarrhoea and dehydration which, if untreated, can be fatal within hours even in a previously healthy person.

Modern sewage and water treatment systems mean that the disease has been eliminated in industrialised countries. However, this "ancient illness that today sickens and kills only the poorest and most vulnerable people" (Global Task Force on Cholera Control ((GTFCC)) remains widespread in parts of Africa, South East Asia, and Yemen. It has a worldwide case load of between 1.3 to 4 million which results in 21,000-143,000 deaths.

Cholera has two forms, epidemic and endemic. Areas described as having endemic cholera are those in which cholera cases have persisted during the last 3 to 5 years with evidence of local transmission, meaning human-to-human interactions dominate as a mechanism of infection. Cholera epidemics, meanwhile, occur in those regions where cholera is not reported and occurs sporadically. This is often associated with natural or anthropogenic shock events such as war, floods, or civil unrest. These shocks often change the way humans live (e.g. they may be forced to live in crowded conditions without adequate access to safe drinking water and sanitation).

Whilst universal access to safe drinking water and adequate sanitation is the long-term solution to cholera, this tends to be linked to economic development and can still be vulnerable to environmental and humanitarian crises. Interventions to control and prevent cholera, once cases are observed, include surveillance and monitoring of cases, water chlorination, provision of oral vaccinations and strengthening education programs such as WASH through social mobilisation. The responsibility for treating, controlling, and preventing cholera at a national level lies between Government Health Ministries and non-governmental organisations (NGOs). These include (but are not limited to) UNICEF, UN IOM, Save the Children, Medair, Médecins Sans Frontiers and Water Aid.

Mitigation efforts are generally synchronised by the GTFCC. Hosted by the World Health Organization (WHO), the Task Force brings together over 50 institutions (including governments, NGOs, academic institutions, and UN agencies) to implement a strategy for ending cholera by 2030. The Ending Cholera—A Global Roadmap to 2030 aims to reduce cholera deaths by 90% over the next decade. It has three pillars: Early detection of and response to outbreaks, integrated prevention tactics, and coordination between countries and partners.

The roadmap provides a concrete path for ending cholera as a public health threat and countries are encouraged to develop National Cholera Plans, whose implementation is overseen and supported by the Task Force.

1.3 Predictive tools for cholera

The GTFCC's Roadmap recognises the need for early detection of cholera outbreaks and suggests this can be achieved through:

> *"The strengthening of integrated early warning surveillance systems, including the confirmation of suspected cholera cases (requiring laboratory culture capacity and rapid diagnostic tests) at the peripheral level."*

This is elaborated in national level roadmaps as *"the need for surveillance systems to be strengthened at community and health facility levels through a reliable alert system and fully fledged laboratory capacity and proper sample management to detect and confirm cholera cases".*

Whilst such activities are essential in identifying when and where clusters of cases are occurring, most activities focus on treatment and containment with some focus on prevention (in response to an outbreak).

In the context of weather forecasting and climate change adaptation, the concept of 'early warning' generally refers to the forecasting of *anticipated* hazards, not observations or proof that the event or hazard is already happening. Whilst early warnings for severe weather have been around for many years, considerable progress has been made more recently in linking specific humanitarian interventions to these forecasts. For example, providing government subsidies for cattle fodder ahead of a harsh winter or setting up evacuation centres before a tropical storm hits. Known as 'anticipatory action,' this approach is now widely used by the International Federation of the Red Cross in their Forecast Based Financing initiatives, and by UN agencies such as the World Food Programme and Food and Agriculture Organisation. Anticipatory windows have also been included in funds such as the Red Cross' Disaster Relief Emergency Fund (DREF) and the UN's Central Emergency Response Fund (CERF). These enable the rapid release of financing to support actions to prepare for a forecasted event.

Studies looking into the merits of acting in advance suggest the benefits of early action include enabling an earlier response and reducing time taken to respond (e.g. delivering food supplies before access routes are destroyed due to storms) and reduced suffering and indignity. Analysis also demonstrates the significant economic gains of investing in more proactive responses to crises (e.g. prepositioning, early procurement, evacuations before a predicted event) through comparing the relative cost of a late versus early humanitarian response (Cabot Venton, 2013).

In the epidemiological domain, there is less familiarity with the concept of using forecasted risk of disease outbreak to take early action, even though there are several diseases which are associated with environmental factors which may offer a means for predicting risk. This is primarily because of a lack of disciplinary focus on the integration of weather and climate information with epidemiological data. This, in part, can be attributed to the absence of predictive tools for diarrheal disease, such as cholera. Models in this domain have tended to track transmission (defined as the human-to-human infection route only) and the impact of interventions, as opposed to highlighting when and where there will be an issue in the future.

A Cholera Risk Model developed by the University of Florida (UF) seeks to address this gap through evaluating the environmental and social factors that are known to be associated with the occurrence of cholera. Data on rainfall, temperature, sanitation infrastructure and population movement are used to produce risk maps of where cholera will occur, up to 4 weeks in advance.

A major cholera outbreak in Yemen in 2017 (which continues today) saw over half a million people infected. Alongside other more direct support to help address the situation, the UK's Department for International Development (now 'FCDO', and FCDO used in the rest of the report) saw an opportunity to explore the application of anticipatory action for cholera with UNICEF, who are responsible for cholera prevention and response in the country.

UNICEF teams started to receive weekly cholera risk data from the UF's CRM and rainfall forecasts from the Met Office in April 2018 and used these to prioritise where their interventions were focused. By August 2018, epidemiological data suggested a significant drop in the cases in 2018 compared to 2017, with services from UF and the Met Office being anecdotally linked to this.

This paper presents an understanding of how UNICEF used these tools to inform their cholera response actions and explores the barriers and opportunities of taking anticipatory action in this new way. It provides an analysis of how reliable these tools were at predicting cholera and rainfall in Yemen from 2017-2019 and explores if the reduction in cases observed in 2018 can be attributed to the forecast-based interventions taken by UNICEF. Finally, it shares recommendations on how these tools should be used in Yemen and elsewhere and presents suggestions for how they can be enhanced.

2. Cholera Risk Model

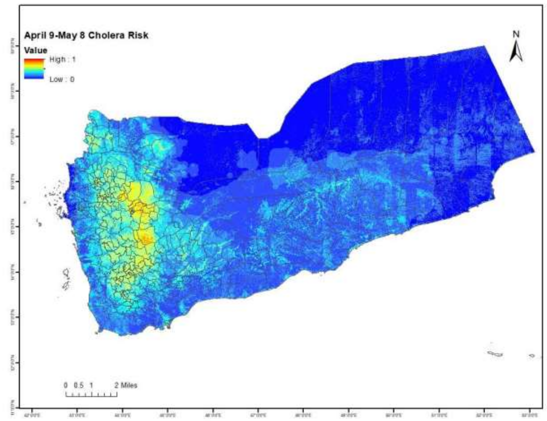
2.1 About the CRM

The Cholera Risk Model (CRM) is an integrated platform that calculates the risk of the trigger (initiation risk of disease in population) and transmission (spread of disease in human population) of cholera. It has a validity period of 4 weeks from the date of issue. The CRM is based on the integration of data on rainfall, temperature and social determinants such as human mobility, water security and access to water and sanitation infrastructure. It is divided into two components: The Trigger Module (TM) and The Transmission Module (TrM). The CRM data sent to UNICEF in Yemen only included the Trigger Module (TM) of the model since the Transmission Module was (and is still) under development.

A widespread outbreak of cholera requires the trigger and transmission mechanisms to act at appropriate times (sequentially for epidemic regions and simultaneously for endemic regions) to result in a public health emergency. The Trigger Module represents the mechanisms which simulate growth, multiplication, and persistence of *Vibrio cholerae* in the aquatic environment, followed by the interaction of the bacteria with the human population. The interaction of pathogenic bacteria usually occurs because of destruction of water and sanitation infrastructure which leads to contamination of drinking water. The model's transmission mechanism represents the pathway by which a widespread outbreak of cholera occurs and involves a complex pathway of interaction of humans with contaminated water, in addition to a prevailing trigger mechanism.

The CRM (trigger and transmission modules) was developed by the UF and the University of Maryland (UMD) and is based on previously published research work (Colwell, 1996; Huq et al., 2005a, 2017a; Jutla et al., 2013; Jutla et al., 2015; Khan et al., 2018a; Singleton et al., 1982). Output from the trigger module of CRM is a risk score at a resolution of 1km x 1km, with predicted risk ranging from high (numerical value of 1) to low (numerical value of 0). Once the risk scores are generated at 1km resolution, it can be averaged over a user specified area of interest.

The CRM is inherently different from traditional epidemiological models (Grad et al., 2012) where the output is usually presented as prevalence or incidence of cholera (Huq et al., 2017a). The motivation to use the score, rather than prevalence or incidence, is to be able to circumvent the lack of epidemiological data during public health emergencies, as is often a challenge due to lack of disease surveillance networks in regions with weak water and sanitation infrastructures. (Khan et al., 2018a). Traditional epidemiological models are variants of Susceptible-Infected-Recovered (SIR) architecture (Grad et al., 2012). Parameterisation of such models presents a challenging uncertainty on when and where cholera may occur. The CRM utilises a state-of –the-art framework of a geographically weighted raster overlay technique (Andersson & Mitchell, 2006), and hence is not bound by the uncertainty of model parameters. Details on methodology are provided in Annex 4. A typical output of the CRM's Trigger Module is shown in Figure 2a.



As requested, this is the cholera risk map at the finest resolution. The map is pixel based instead of county based.

Figure 2a: Example of output from the CRM over Yemen for 9 April – 8 May 2019.

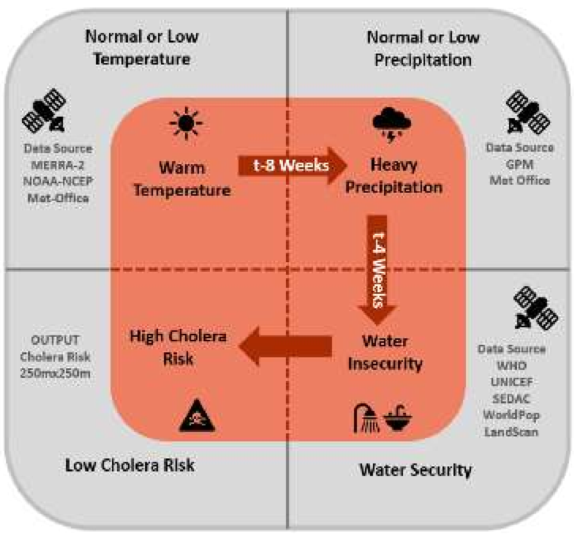


Figure 2b: Cholera Risk Model Trigger Mechanism

2.2 Evolution of the CRM - what causes Cholera?

There are two main schools of thought on how cholera spreads in human populations.

One school suggests that cholera is triggered by the introduction of infection individual(s) to an unaffected populate (Eppinger et al., 2014; Freichs et al., 2012). It assumes that the cholera bacteria (which is carried by humans) finds its way into water systems, thereby contaminating them and resulting in an explosive prevalence of the disease. This hypothesis assumes humans are the carrier of the pathogenic bacteria.

The second school of thought is based on the integration of climate and microbiology of regions (such as Bangladesh, India, South America and Africa), these studies suggest that cholera is initially triggered though the interaction of humans with ponds, rivers etc. which contain autochthonous pathogenic vibrios (Alam et al., 2006; Huq et al., 1983: Jutla et al., 2013; Singleton et al., 1982). Secondary transmission routes (Jutla et al., 2015) then occur through the human-environment-human route, where infected individuals reintroduce the bacteria to the same or other water resources systems (Blokesch et al., 2012; Condeco et al., 2008)

Large cholera outbreaks have been recorded to be associated with natural and anthropogenic disasters, notably when environmental conditions favour growth of the bacterium. The motivation for the CRM is derived from John Snow's pioneering work in identification of a water well that was responsible for cholera outbreaks in England in the early 19th century. A weak water and sanitation infrastructure will often lead to an outbreak of cholera. In simple terms, the CRM is the mathematical representation of Snow's empirical observations.

The CRM is a culmination of several decades of work on understanding microbiology of cholera bacteria, hydro-climatology impacting bacteria and sociological processes.

2.3 The environmental hypothesis of cholera

Based on an analysis of cell surface antigen structures, the etiological agent, which is historically recognised as the cause of ongoing cholera pandemics, is known as *Vibrio cholerae serotype O1 or 0139.* Horizontal transfer of the serotype coding genes (Bik et al., 1995) has led to the recent emergence of serotype O139 outbreaks of cholera. Reports of *V. cholerae* non-O1 strains are now known to carry multiple virulence factors, including the primary virulence factors of *V. cholerae* O1, namely cholera toxin. The toxin co-regulated pilus (Rivera et al., 1995) need to be considered in the context of public health (Hasan et al., 2013). Various *Vibrio cholerae* strains and serotypes led to an investigation to determine a common detecting factor, which was directed towards environmental variables.

Early environmental studies of cholera were unsuccessful in detecting habitat of *V. cholerae*, such as domestic animals or human carriers (Pollitzer, 1954). However, in the late 1960's the bacterium was detected in environmental water samples collected in cholera-free regions (Mukerjee et al., 1965; Pollitzer et al., 1959).

Those *V. cholerae* were subsequently shown to be associated with zooplankton (Kaneko & Colwell, 1975; Kaper et al., 1979; Tamplin, n.d.). The environmental variables that define the habitats of *Vibrio cholerae* were demonstrated in extensive studies carried out globally from 1970-2000 (Bougoudogo, 1998; Colwell, 1996; Colwell & Huq, 1994; Colwell & Huq, 1998; Lobitz et al., 2000; Mata, 1994).

Between outbreaks, and during unfavourable environmental conditions, the bacterium persists in environmental reservoirs in a viable, but non culturable state (Roszak & Colwell, 1987). Based on extensive environmental microbiology studies, it has been established that *V. cholerae* is autochthonous to pond, river, estuary, coastal, and marine ecosystems, with copepods as its host/vector. Copepods are zooplankton, comprising a significant component of the aquatic fauna of rivers, bays, estuaries, and the open ocean, are the major host/vector of cholera (Conner et al., 2016; A. Huq et al., 1983).

Environmental factors that drive the **prevalence** of *V. cholerae* in the environment (and are associated with increased number of cases of cholera in an outbreak), include warmer sea surface and coastal water temperatures (Akanda et al., 2011; Lobitz et al., 2000; Vezzulli et al., 2016).

Several environmental and climate variables are linked to **proliferation** of *V. cholerae* and incidence of cholera, including:

- precipitation (Hashizume et al., 2008);
- flooding (Koelle et al., 2005);
- sea surface temperature and height (Lobitz et al., 2000);
- river level and freshwater discharge (Akanda et al., 2011; Schwartz et al., 2006);
- coastal salinity (Miller et al., 1982);
- dissolved organic material (Neogi et al., 2018);
- chlorophyll (Constantin de Magny et al., 2008), and;
- components of phytoplankton and zooplankton (Constantin de Magny et al., 2008; de Magny et al., 2011). Epidemiological surveillance suggests a link with estuarine ecosystems, namely river and coastal regions (Lipp et al., 2002).

2.4 Endemic and Epidemic Cholera

Based on analysis of cholera records maintained during British India from 1823-1875, cholera has been defined as occurring in two dominant forms:

1) **Epidemic:** characterised by sudden and sporadic occurrence of a large number of cases, and
2) **Endemic:** where cholera cases occur at a baseline level throughout the year, with distinct seasonal peaks(Anwar Huq et al., 2017b; Rakibul Khan et al., 2019, p.).

Epidemic cholera is hypothesised to be related to elevated air temperatures followed by above-average precipitation, in concatenation with insufficient and/or damaged water, sanitation, and hygiene (WASH) infrastructure. This places the human population at a high risk of interaction with cholera bacteria and a subsequent disease outbreak (Huq et al., 2013).

Endemic cholera is associated with a constant occurrence of cholera cases, primarily in regions where coastal or terrestrial water systems create favourable conditions for growth and proliferation of *Vibrio cholerae (Jutla et al., 2013)*. Under certain environmental conditions, a sustained epidemic mode of cholera can evolve into the endemic form in regions where there is enhanced and continuing exposure to, and transmission of, *V. cholerae*.

As a pandemic disease, cholera affects millions in vulnerable human populations (Clemens et al., 2017) and is a dominant water-borne disease in Latin America, sub-Saharan Africa, and Southern Asia (Ali et al., 2015; Jutla et al., 2010a).

2.5 Predicting Cholera

In Haiti, during the months following Hurricane Matthew, WASH infrastructure was extensively damaged which exposed the population to unsafe drinking water (Huq et al., 2017b). Analysis of the epidemiological data showed cholera risk could be predicted successfully by employing environmental and epidemiological factors.

Since March 2015, Yemen, a coastal Middle Eastern country, has suffered violent surges of civil unrest (Sharp & Salaam-Blyther, 2017), and in October 2016, the country reported an outbreak of a few cholera cases. By the end of 2017, Yemen accounted for *ca*. 80% of cholera cases worldwide since 2015 (WHO, 2018).

During the first six months of the outbreak, cholera in Yemen surpassed the number of reported cases in Haiti over a span of seven years (ca. 815,000 cases between 2010-2017), when the Haitian cholera had, until then, been considered historically to be the largest cholera epidemic (Lyons, 2017). The Yemen cholera epidemic was considered one of the worst public health disasters in recorded history (Federspiel & Ali, 2018), until COVID-19.

Cholera is unlikely to be eradicated globally since the disease-causing agent is autochthonous to aquatic environments and plays a role in their carbon and nitrogen cycles (de Magny et al., 2008). Furthermore, there is mounting evidence that warming sea surface temperatures are associated with spread of *Vibrio* spp (a common group of Gram negative bacteria) and emergence of human disease globally (Vezzulli et al., 2016).

Disease predictions can be achieved by recognising that disease progression comprises two components: trigger and transmission. These, together, result in an outbreak and, subsequently, a public health emergency (Khan et al., 2019).

2.6 Operational use of the CRM

The CRM comprises a trigger (Huq et al., 2017c; Jutla et al., 2017b) module, which uses data on precipitation, temperature, population and (WASH) infrastructure to compute a risk score, with values that vary between 1 (high) and 0 (low) in a given region.

The trigger algorithm identifies conditions of anomalous temperature and rainfall, providing an assessment of cholera risk for the following four weeks (Anwar Huq et al., 2013; Antarpreet Jutla et al., 2015) for a given region. Details of model development and algorithmic architecture have been published and referenced in Annex 4 (Khan et al., 2018a).

The CRM includes data from a range of sources:

- **Rainfall data:** Daily and monthly rainfall data at two different resolutions were obtained from the National Aeronautics and Space Administration (NASA).

  Monthly rainfall data, at a resolution of 0.25^{o} X 0.25^{o} from the Tropical Rainfall Measuring Mission (TRMM) were employed to compute the long-term average (from 1998 to 2018).

Daily rainfall data at a spatial resolution of 0.1° X 0.1° were obtained from the Global Precipitation Mission (GPM) and used to determine precipitation variation from long-term average at resampled data points.

- The average correlation over land between GPM and TRMM data is very high (>0.90) with small bias (unidirectional-negative bias) (Liu, 2016). A recent Yemen focussed study (AL-Falahi et al., 2020), which considered only one governorate, found that TRMM precipitation data were statistically significantly correlated with limited available observed gauge data. Therefore, our confidence in using these two datasets remains high, especially given that we are not using absolute precipitation. Instead, we are computing anomalies and thereafter binning the data based on standard deviation of TRMM datasets (details on methodology in Annex 4). **Air temperature:** Daily and monthly data for air temperature on the surface, at a spatial resolution of 0.5° X 0.625° were obtained from the NASA Modern-Era Retrospective analysis Research and Application, Version 2 (NASA-MERRA 2), and used to compute long-term averages and calculate anomalies.

- **Population data:** LandScan population data at a spatial resolution of 1 km x 1 km were obtained from Oak Ridge National Laboratory and used in the model. Model output was resampled at 1km. The population data is a static data, implying that the data does not change over time steps.

- **Epidemiological data (used for evaluation):** Weekly cholera reported cases at the governate level between January 2017 and December 2018 were provided by the Early Warning, Alert and Response System EWARS and between January 2019 and July 2019 by the Assessment Capacities Project ACAPS [WHO, 2020].

3. Rainfall Forecasts

3.1 Yemen's Climatology Relevant to the Study

The Climate of Yemen can be described as a subtropical dry, hot desert climate with low annual rainfall, very high temperatures in summer and a big difference between maximum and minimum temperatures, especially in the inland area.

The regional distribution of rainfall in Yemen is affected by the country's topography, particularly by the Sarawat mountain range which runs down the western coast of the Arabian Peninsula.

Moist winds arriving in Yemen from the Red Sea, or as part of the south-westerly monsoon, are lifted by the mountain range causing significant precipitation. This orographic enhancement of precipitation means that much of the rainfall in Yemen occurs in the west, where much of the population is also located. The seasonal precipitation patterns in this region are largely governed by the locations of the Intertropical Convergence Zone (ITCZ) and the Red Sea Convergence Zone (RSCZ).

The ITCZ is a band of low pressure formed where the Northern Hemisphere trade winds meet the Southern Hemisphere trade winds. This band moves seasonally with the thermal equator between the Tropic of Cancer and the Tropic of Capricorn and is accompanied by a band of precipitation closely aligned with it. The RSCZ refers to the convergence zone in the Red Sea where north-westerlies from the Mediterranean meet with south-easterlies from the Gulf of Aden which also produces precipitation.

The RSCZ is active between March and May, producing precipitation in the west of Yemen; the ITCZ is active over Yemen between July and September (Farquharson, Plinston, and Sutcliffe 1996). Together the RSCZ and the ITCZ produce a bimodal seasonal distribution of precipitation with one peak between March and May and the other between July and September. In this report, a single wet season between April and November is considered. Outside of the two wet periods associated with the RSCZ and ITCZ, Yemen is largely dry. The region does experience Tropical Cyclones, e.g. Chapala, Megh, Luba. However, the specific risk from tropical cyclones has not been assessed in this report as it is out of scope of the aims of the analysis undertaken here.

3.2 Met Office Rainfall Forecasts for Yemen

Alongside UF's CRM products that were sent to UNICEF in Yemen to support their cholera prevention activity, the Met Office started to provide weekly rainfall assessments. The Met Office service comprises of the following two components, Yemen Rainfall Assessment (PDF of MS PowerPoint slide-deck) and Yemen Rainfall Guidance (PDF and MS Word, including a narrative forecast and district level rainfall forecast tables). These are sent to a distribution list provided by UNICEF every Monday morning.

The service is provided by the Global Guidance Unit (GGU) at the Met Office who provide 24-hour guidance for the interests of the UK and other partners globally. The GGU critically assess a wide range of models (including the Met Office Global Model, ECMWF, GFS) and observational data and provide guidance on the 'best possible forecast' on high-impact weather events from the available models. To assess whether the hazard will be impactful, the GGU team are required to understand the local vulnerabilities and exposure to the hazards.

3.2.1 Yemen Rainfall Guidance

The Yemen Rainfall Guidance provides a summary, a 7 day hindcast, a 7-day forecast, a 4-week outlook and ten-day district level rainfall tables. It is a narrative description of the current weather in Yemen. If high impact weather is expected, this will be highlighted.

This is a recent example,

'A period of heavier than usual showers is expected for western Yemen from Thursday to Sunday

Forecast next 7 days

The dry period will continue through to Wednesday. However, from Thursday there will be an increased threat of showers across western Yemen, especially across the Highlands and coastal plain. There could be isolated accumulations of up to 5-10 mm in a day and perhaps as much as 10-20 mm over a 4-day period in a few places. However, much of the country will remain dry.'

This narrative is followed by tables which list the Priority 1, 2, 3 and 4 administrative districts (for cholera) and the associated level of forecast rain. Rainfall is categorised by colour in terms of daily accumulation (as defined by the Global Guidance Unit and described in Table 6).

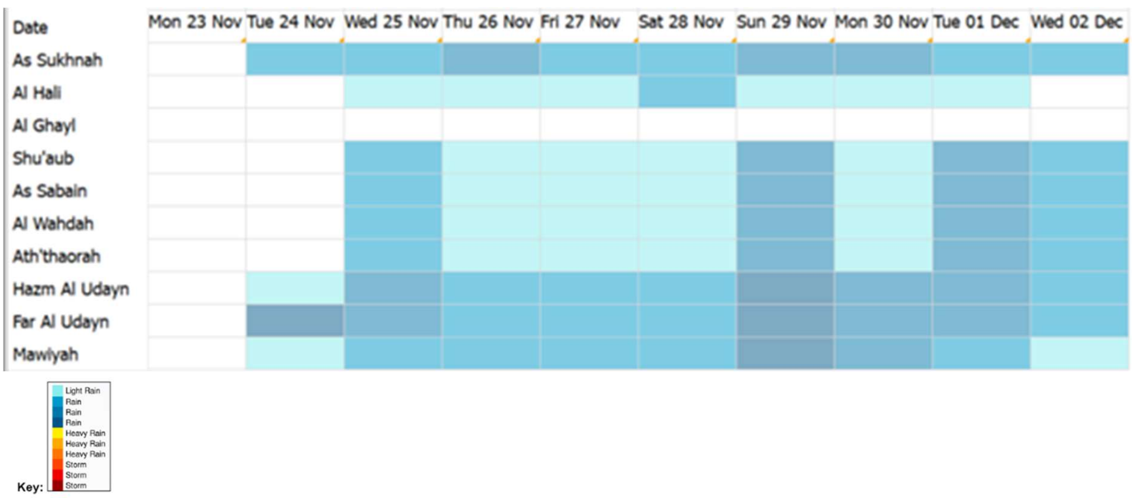


Figure 3a: Example of rainfall predictions for Priority 1 Districts (priority districts defined by cholera risk assessment carried out for the oral vaccination campaign)

3.2.2 Yemen Rainfall Assessment

The Yemen Rainfall Assessment is a set of slides (MS PowerPoint in pdf format) and includes a rainfall accumulation hindcast (images of satellite derived rainfall accumulation), rainfall forecasts, and site- specific (main cities) ensemble rainfall forecast accumulations. The three hindcast maps below show satellite derived rainfall over Yemen during the previous 1, 7 and 30 days.

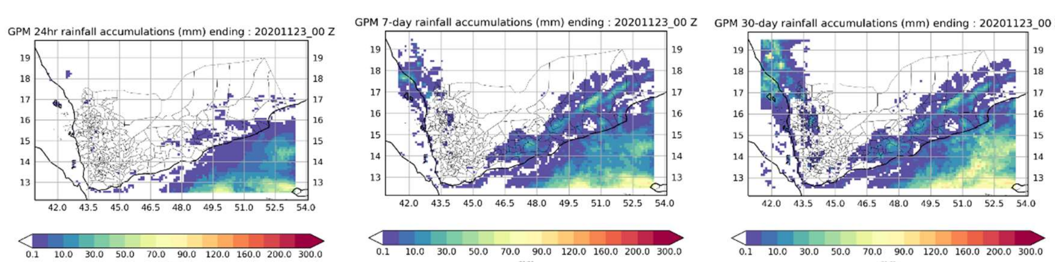


Figure 3b: Satellite derived maps of rainfall accumulations over Yemen

The following maps show the spatial distribution of rainfall over Yemen for the next six days. The first example takes rainfall accumulation from the Global Model that has a spatial resolution of 10 km and shows accumulation up to 6 days. The second example uses a high-resolution local area model (the Crisis Area Model, 4 km) and shows accumulation on day 1. The high-resolution model takes its boundary conditions from the global model and is available out to day 2. The final example uses the Global Model and shows the accumulation on day 3.

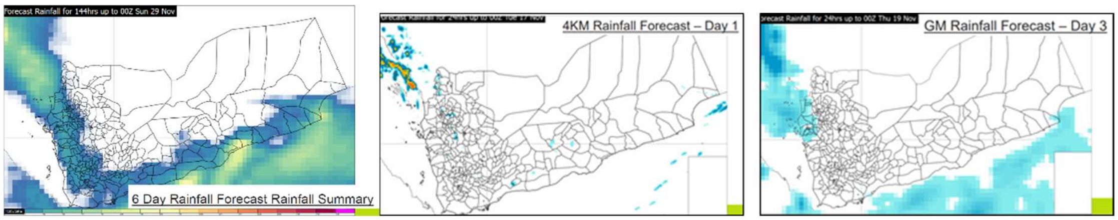


Figure 3c: Examples of maps of forecasted rainfall accumulations over Yemen

The final section has site-specific rainfall accumulation forecast meteograms for the main populated locations, which use the European Centre for Medium-Range Weather Forecasting (ECMWF) ensemble prediction system (EPS). The EPS produces a range of probable forecasts (instead of a single forecast) by running several weather forecasts using different conditions to capture the variability in the atmosphere.

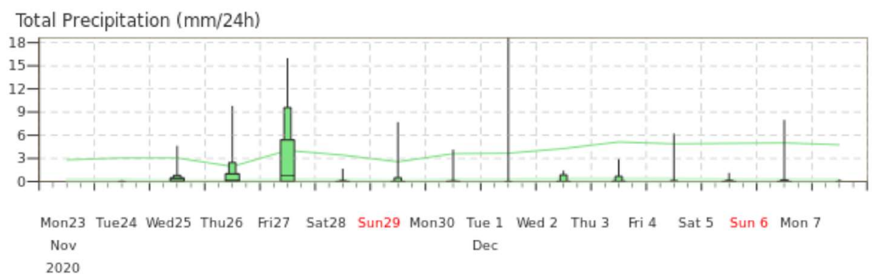


Figure 3d: Example of city level forecast

3.3 UNICEF requirements

These products were set up in April 2018 following discussions with UNICEF and testing of various formats. After prototype forecasts were created and refined with UNICEF over a period of a few weeks, the Word summary forecast and PowerPoint maps were finalised.

After receiving the forecast for a few months, discussions with UNICEF revealed that the most useful element of the forecast were the tables described in section 3.2.1 as these could be considered against cholera case data presented at district levels. These were therefore provided in an Excel format to make them easier to integrate with other data types.

The distribution list for the rainfall forecasts includes UNICEF and WHO staff in Yemen as well as staff at UNICEF headquarters and the FCDO.

3.4 Model performance

The Met Office's Global Model is one of the top five global forecasting models, compared to NOAA, ECMWF, NCEP, GFS, and consistently rated as one of the most accurate through independent and verified methods[1].

The Yemen daily rainfall accumulation forecasts from 3 days to 6 days ahead are produced using the Met Office's Global Model, which covers the world at a resolution of 10 km (2). In this model the convective processes are represented through statistical functions (parameterised convection). The Crisis Area Model, which downscales the global model boundary conditions to provide a 4km resolution is used for days 1-2 ahead (this model is only run for a 48-hour period), is a convection-permitting model. This means that processes that lead to convective rain can be captured within the model, however not at the resolution to resolve the convective rain, which occurs at smaller scales. To be able to capture convective rainfall, the model needs to be convection resolving, which would only be possible on a very high resolution limited area model ~1.5km.

As there were no rain gauges in Yemen providing observations of rainfall during the period of this study, a satellite derived observation data set (NASA Global Precipitation Measurement dataset) was used. The use of satellite data for forecast model verification is discussed in Section 6 and Annex 5.

[1] https://apps.ecmwf.int/wmolcdnv/scores/surface.time_series/tp
[2] The km of the model refers to the size of the gridsquare numerical models use to represent the atmosphere. Higher resolution models are those with fewer km and generally provide a greater level of accuracy

4. Use of the CRM and rainfall forecasts in Yemen (Contextual Analysis)

4.1 Cholera in Yemen

More than five years of conflict between government forces and Houthi rebels (among others) have pushed Yemen to the brink of collapse and caused its people to suffer from one of the world's most complex and destructive humanitarian crises. Access to food is limited, fighting leads to injury and death, millions of children are unable to go to school and deadly diseases spread rapidly.

Even before the war, Yemen was described as a country that was "beset by circumstances that made it ripe for cholera" due to high poverty rates, frequent droughts and water access problems and sanitation for only half the population (Qadri, Islam and Clemens, 2017).

Armed conflict and the resulting displacement of people who do not have adequate food, water, housing, or sanitation exacerbated pre-existing conditions and led to a cholera epidemic which has been described as the worst in recorded history.

The UNICEF and World Health Organization (WHO) executive directors stated in 2017:

> *"This deadly cholera outbreak is the direct consequence of two years of heavy conflict. Collapsing health, water and sanitation systems have cut off 14.5 million people from regular access to clean water and sanitation, increasing the ability of the disease to spread. Rising rates of malnutrition have weakened children's health and made them more vulnerable to disease. An estimated 30,000 dedicated local health workers who play the largest role in ending this outbreak have not been paid their salaries for nearly ten months"*

The earliest cases of the disease in Yemen were in the capital, Sana'a, with some occurring in Aden. By the end of October 2016, cases had been reported in the governorates of Al-Bayda, Aden, Al-Hudaydah, Hajjah, Ibb, Lahij and Taiz and by late November, also in Al-Dhale'a and Amran. Whilst the 2016 wave was relatively limited in scale, the second (from late April 2017) was country-wide and of a different order of magnitude. By June 2017, a total of 268 districts from 20 governorates had reported cases. Over half were from the governorates of Amanat Al Asimah (the capital Sana'a), Al-Hudaydah, Amran and Hajjah, which are all located in the west of the country (WHO EMRO).

77.7% of cholera cases and 80.7% of deaths from cholera occurred in Houthi-controlled governorates, compared to 15.4% of cases and 10.4% of deaths in government-controlled governorates.

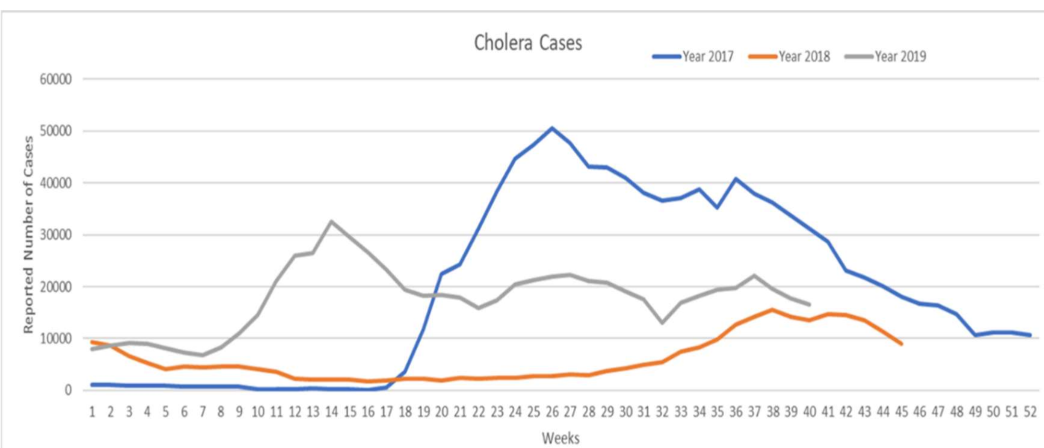


Figure 4a. Reported Cases of cholera throughout the year in 2017, 2018 and 2019.

4.2 Monitoring Cholera in Yemen

The Yemen Health Authorities with support from WHO set up an electronic integrated disease early warning system in 2013, which gradually increased from 98 reporting sites to 1982 by the end of 2016.

When the first cholera case is confirmed, public and private health facilities that were part of this countrywide cholera surveillance system collated suspected cases using a common line-list database (Excel 2010, Microsoft). Using guidelines from the GTFCC, a suspected case was described as any patient presenting with three or more liquid stools with or without vomiting in the past 24 hours. A confirmed case was a suspected case with V cholerae O1 or O139 confirmed by culture. District surveillance officers (for each of the 333 districts) compiled the list from all health facilities in their district. These were sent electronically to the governorate level (23 total governorates) each day. Data were aggregated by the Emergency Operation Centre run by the Yemen Health Authorities and digitally cleaned by WHO surveillance officers.

Subsequent analysis of the monitoring of cholera in Yemen has found that levels of biological testing carried out was low (2,706 biological tests done out of more than 1 million suspected cases). An over-estimation of cholera is therefore suspected due to all acute diarrhoeal cases being registered as cholera when there are several diseases that lead to diarrhoea in Yemen ((Evaluation of the UNICEF Level 3 response to the cholera epidemic in Yemen).

There is also a widespread view that health workers inflated the figures for suspected cholera due to fears that too low a figure, or a declining trend, might result in the closure of the relevant cholera treatment centre. However, UNICEF's evaluation team were not able to verify of quantify this.

4.3 UNICEF's role in Cholera Response in Yemen

UNICEF, the International Committee of the Red Cross, the WHO, Oxfam and Medécins San Frontiers are all involved in cholera prevention in Yemen and work with the government's Ministry of Water and Environment (MoWE).

In Yemen, the key activities in the fight against cholera are categorised by the WHO as:

- vaccination (Oral Cholera Vaccine (OCV) campaign began in 2018);
- the delivery of new health services, such as new treatment centres and oral rehydration 'corners' and nutrition interventions;
- support of the existing health system;
- health infrastructure, such as delivery of IV fluids, fuel for hospital generators and provision of cholera kits;
- WASH interventions such as provision of safe water, water chlorination, distribution of hygiene kits and hygiene promotion/awareness raising;
- staff training;
- epidemiological activities such as disease surveillance;
- management; and
- advocacy.

Of these, UNICEF's role has involved acting as the WASH 'cluster'[3] lead. This has included providing safe water to over 1 million people and deploying 20,000 hygiene promoters to run hygiene awareness campaigns. UNICEF have also supported the health infrastructure through provision of equipment such as medicine, oral rehydration solution, intravenous fluids, and diarrhoea kits.

[3] A 'Cluster' is a partnership aimed at improving coordination

In response to the high rates of cholera in 2017, the Health cluster (run by WHO) and WASH clusters became better integrated so that they could coordinate the management of the outbreak and combine resources. The MoWE set up a network of Emergency Operation Centres (EOC) which bring together governmental, UN and INGO (International Non-Government Organisations) and NGO (Non-Government Organisations) partners from across the WASH and Health sectors to improve the efficacy of the cholera response. Activities of the EOC involve gathering, analysing, and sharing of data to inform strategic and operational decisions.

4.4 Lessons learned from the 2017 wave of the cholera epidemic

An independent evaluation of UNICEF's response to the cholera epidemic in Yemen was published in June 2018. This suggested that because the scale of the epidemic in 2017 was not anticipated, the response system in place was unable to keep pace with its rapid escalation. Whilst the control and prevention measures taken were appropriate, the report describes how "this took time to emerge, and full operating capacity was not reached until the epidemic was already well advanced".

The evaluation recognised that timing of response is almost everything and describes how control measures can make a substantial difference to the course of an epidemic if interventions are taken ahead of the epidemic curve. After the peak of the outbreak, it acknowledges that the effectiveness of the response (even if well designed and implemented) will likely be marginal, or at least far more limited than if implemented earlier in the outbreak.

UNICEF's evaluation report did not consider the use of the CRM and rainfall forecasts to target interventions as these started being used in April 2018 and the report was released in June 2018 and so it was probably being finalised before the reduction in cholera cases was observed.

4.5 The need for predictive tools

The UNICEF evaluation report revealed that it would be better to target resources where they were needed most rather than always 'chasing the epidemic' (UNICEF evaluation report) there was a need to be able to target resources at areas most in need.

The role of environmental factors in relation to cholera risk was considered after a study by Epicentre was published in the Lancet (Camacho et al.2018). This estimated that a weekly rainfall of 25mm was associated with a 42% increase in the chance of a person developing suspected cholera in the following 10 days, compared to a week with no rain. The study suggested that the small first cholera epidemic wave seeded cholera across Yemen during the dry season in 2016/17. When the rains returned in April 2017, they triggered widespread cholera transmission.
Whilst rainfall (and temperature) were known to play a role in the start and spread of cholera, it was also understood that this was not a straightforward association since localised differences also determined rates of disease transmission (based on findings from Haiti's outbreak). A tool which captured rainfall and localised differences was therefore required (Personal communication from UNICEF WASH Cluster Coordinator).

The UF's CRM was (and is, to our knowledge) the only tool of its type which does this and so in April 2018, FCDO supported the provision of weekly CRM data to the Emergency Operations Centre (EOC). As the Yemeni Government was unable to provide weather forecasts, FCDO also arranged for the Met Office to send weekly weather forecasts.

Before the CRM and rainfall forecasts were used operationally, UNICEF shared the data with the EOC. This was done to corroborate the findings of the Lancet study. The EOC also found that in districts where there was heavy rain, high numbers of acute watery diarrhoea and cholera cases appeared, following the rainfall.

The data then started to be used operationally to inform cholera response. Decisions regarding where to prioritise interventions were made according to administrative districts in Yemen. The rainfall data was therefore provided in a tabulated format (for the top 100 high cholera-risk districts), showing the district and forecast rainfall daily accumulation for the next 10 days.

4.6 How forecasts were used

Sub-EOC coordinators prepared a table of the districts most affected in their area (column A in the table below). This was prepared ahead of the EOC weekly meeting. The Met Office's rainfall prediction was then added to column (B). Only darkest blue rain in the 'rain' category, heavy rain and storm prediction were relevant to UNICEF (as shown in Figure 4b below). As UNICEF do not meet to discuss the rainfall predictions until the Wednesday of each week (the forecasts are issued on a Monday), the first 2 days of the forecast were not used

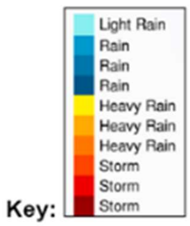


Figure 4b: From figure 3a in section 3, key to rainfall accumulation categorised by colour

The sub-EOC coordinators then indicated the sub-districts with increasing incidence OR lab confirmed case OR acute watery diarrhea (AWD) related death during the previous week. This enabled them to identify the sub-districts most at risk of being negatively affected by the predicted rains. At the weekly EOC meeting, level of risk was then assigned to sub districts, either 1 (low risk), 2 (moderate) or 3 (high) will be given in column D to every sub-district matching the criteria. Each entity or branch then planned what they would do in specific sub-districts (column E). Preventive actions were and are tailored to the local context which actors understood well.

Priority district (A)	Number of heavy days of rain or severe weather in next 10 days (B)	Sub-district with increasing incidence OR Lab confirmed case OR AWD death during previous week (C)	Level of risk (1-3) 3 being the highest (D)	Specific actions to be taken at sub-district level (E)
Eg. As Sabain	1	SD xxx	3	• Increase the chlorination of water systems serving this neighborhood; • Undertake a sewer clearing in one specific zone known to have overflow following rains; • Check 2 known private water networks in the area; • Diffuse message on the radio focused on the risk of diarrhea diseases after rains

Table 1: Example of prioritisation of districts according to cholera risk and associated actions

Two basic interventions were intensified:

- The risk communication and hygiene awareness in areas most at risk. These focused on the risks related to heavy rainfall consequences, such as cesspit /sewage overflow as these would expose people to fecal matter in the street; lead to food in markets being more easily exposed to contamination and increase the potential for unprotected wells to be contaminated by surface water. Households were given information on the appropriate actions they should take to minimise risk.

- Reinforcing the chlorination of water piped systems, both public and private, serving the population living in the neighborhoods most at risk.

Packages of action related to areas with risk 1,2 and 3 scores are described in the table below. These are adapted to the local context.

Packages of actions to be intensified in districts with most cases, increasing number of cases or presence of lab confirmed cases or AWD death	
Risk level 1 (heavy rain predicted in one of the known hotspot)	• Reinforce hygiene awareness messages in sub-districts most affected (Awareness Water Center, Health Education Center - MoH, WASH NGOs): food market and other gathering places, community awareness session via Mosques and schools • Ensure chlorination in water piped systems of the sub-districts most affected: increase frequency of FRC control along water systems, adjust dosage accordingly (LWSC)
Risk level 2 (heavy rain predicted in one of the known hotspots coupled with an increase of suspected cases the previous week)	• Reinforce hygiene awareness in sub-districts most affected (Awareness Water Center, Health Education Centar - MoH, WASH NGOs): food market and other gathering places, community awareness session via Mosques and schools • Ensure chlorination in water piped systems of the sub-districts most affected: increase frequency of FRC control along water systems, adjust dosage accordingly (LWSC) • Reinforce chlorination at point of collection: increase control on private wells supplying trucks, activate on-site chlorination on most frequented water points located in cases clusters (GARWASP, NWRA, NGOs)
Risk level 3 (heavy+++/storm predicted in one of the known hotspots coupled with an increase of cases OR lab confirmed case OR AWD related death)	• Reinforce risk communication through mass media and hygiene awareness in sub-districts most affected (Awareness Water Center, Health Education Centar - MoH, WASH NGOs): food market places, community awareness session (mosques, schools), engage religious leaders and Moshidats on alerts messaging (households should store and treat water for few days, wash/disinfect systematically fruits and vegetables, households to alert LWSC about cesspits and sewage blockage before and during rains) • Ensure chlorination in water piped systems of the sub-districts most affected: increase frequency of FRC control along water systems, adjust dosage accordingly (LWSC) • Reinforce chlorination at point of collection: increase control on private wells supplying trucks, activate on-site chlorination on most frequented water points located in cases clusters (GARWASP, NWRA, NGOs) • Be ready to activate water trucking in affected areas should local water system be disrupted (GARWSP-EU, NGOs) • LWSC to clear sewage systems in flood-prone neighborhoods or where overflowing are frequent, LWSCP to be ready for sewage systems clearing and cesspits emptying during rains

Table 2: Cholera prevention actions according to level of risk of a district.

The map below shows the number of Rapid Response Teams (RRT) operating, in accordance with the rainfall forecast, at district level. *UNICEF Case Study on use of Rainfall Forecasts to Reduce the Spread of Cholera in Yemen – Annex 3*:

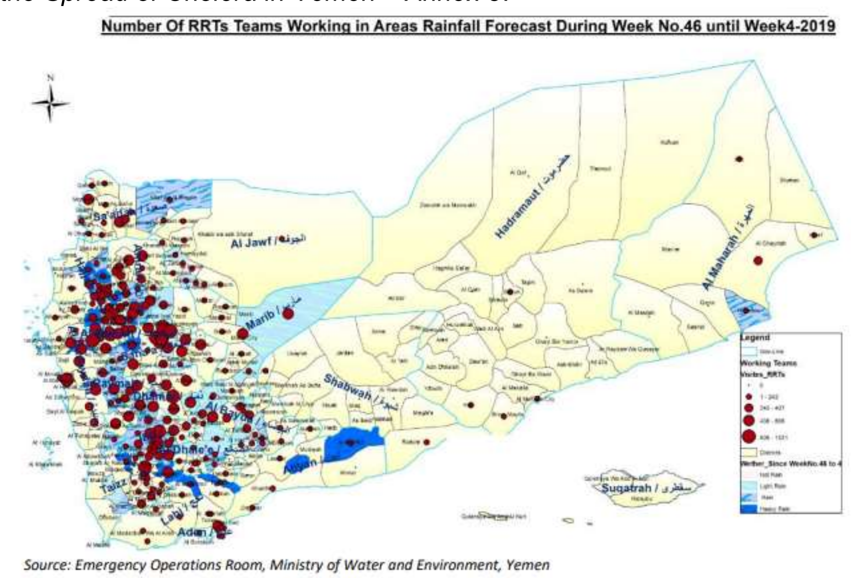


Figure 4c: Map showing Rapid Response Teams in November 2019

In 2018 (January to November), more than 311,000 suspected cholera cases were reported; this was considerably less than the 987,000 suspected cases during the same period in 2017 - a decrease of 63%.

The number of districts affected by cholera also decreased from 305 in 2017 to 287 in 2018 and the attack rate decreased from 364 to 132, respectively (data from the UN Office for the Coordination of Humanitarian Affairs, Humanitarian Needs Overview, Yemen, 2019 and data provided by Emergency Operations Room, MoWE, Yemen).

According to the EOR, the RRT's response in 2018 totalled 37,846 visits, which were targeted to districts affected by rainfall. Whilst several factors were thought to contribute to this reduction in cholera cases, the use of rainfall data to deploy teams in areas where rainfall was predicted was seen to have helped to prevent further outbreaks and the spread of the epidemic.

UNICEF continued to work with the EOC to use this data to inform WASH preparedness activities and other impacts associated with heavy rainfall. For example, the rainfall assessment indicated that Cyclone Luban was making its way quickly towards Yemen. Having this data enabled UNICEF to immediately deploy teams to the targeted areas and ensure that supplies were ready on the ground and awareness and cleaning campaigns were conducted. Additionally, any damaged water pipes and supply systems in the at-risk districts were repaired, to improve sanitation and sewage management and mitigate the health risks from the predicted flooding.

> *"In mid-February 2019, UNICEF received a report of a collapsed and severely damaged sewer transmission line, at risk of bursting, just outside Sana'a old city. Thanks to the data received from the Met Office and University of Florida, UNICEF was aware that the rainy season was expected to start in March. The Old City is in a densely populated area where children often play on the streets. The damaged sanitation system was therefore posting a threat to people's lives, given that cholera transmission increases in periods of heavy rain.*
>
> *UNICEF, together with MoWE, worked immediately on the rehabilitation of the sewer lines and set a timeline on completion of the work by 23 February, followed by reinstatement of the stonework, ensuring the work could be completed before the rain began. Whilst this was a quick fix solution to a potential cholera outbreak, it also provided a long-term solution of strengthening the sanitation system in Sana'a city."*
> *UNICEF Case Study, Annex 3*

4.7 'No Regrets' Interventions

UNICEF describe the actions they take which are informed by the rainfall forecasts (and until 2019 the CRM) as 'no regrets'. In the parlance of anticipatory action, this describes:

> *".....actions by households, communities, and local/national/international institutions that can be justified from economic, and social, and environmental perspectives whether natural hazard events or climate change (or other hazards [in this case cholera] take place or not. "No-regrets" actions increase resilience, which is the ability of a "system" to deal with different types of hazards in a timely, efficient, and equitable manner. (Siegel, P Jorgensen, 2009; UNDP, 2010)."*

When actions are taken but are not necessary (because the hazard does not occur), they are described as having been taking 'in vain'. For time bound actions, there is often a loss associated with this (such as distributing perishable items which are then not needed).

The benefit of ‘no regrets’ actions is that they still have value in these circumstances. For instance, households in areas where a RRT conducted WASH awareness raising will still be better prepared should cholera occur.

For this reason, UNICEF are more concerned about ‘false negative’ predictions of rainfall and cholera than ‘false positive’. For example, saying that it is not going to rain, but it does (false negative) means actions will not have been taken when they were needed, whereas saying that is it going to rain, and it does not (false positive) means actions will have been taken that may end up being ‘in vain’. Because the actions (such as WASH training) are ‘no regrets’, the consequences of acting in vain are limited.

The consequences of ‘acting in vain’ play a role in determining how and when early action should be triggered and in turn, what level of confidence is required in predictions. In Yemen, the scale of the cholera epidemic meant that a tool which could assist with decision making on where to intensify interventions which had more than a 50/50 level of accuracy (“better than tossing a coin”) was helpful (the actual accuracy of the tools used is discussed in sections 5 and 6). In a different context, a higher level of certainty maybe required.

Furthermore, whilst determining if actions have been taken ‘in vain’ is easier to assess in a DRR context where the risks are climatic (a cyclone was predicted but didn’t occur), this is harder to determine for disease control, as a predicted epidemic may not occur *because* of the risk-based interventions taken. This should be considered when assessing the impact of risk-based cholera response.

4.8 Preventative action and fatigue

Whilst the concept of taking no regrets actions seems positive and rates of cholera in Yemen are currently manageable, UNICEF are starting to observe fatigue, both in terms of those who actions are aimed at, and those who deliver them.

For families who are starving, remembering to chlorinate water and wash hands can seem unimportant when cholera does not seem to present an immediate threat in their community. Equally, for teams of responders (who are often volunteers), the need to invest time and resources in preventing a disease which appears to be under control may be questioned.

Demonstrating the value of such actions, to seniors, can also be difficult when their impact is so hard to prove, since a range of social, environmental, and political factors may determine the spread of cholera in an area. Attributing a plateauing or reduction of cases to early action is therefore difficult, especially in the context of increasing rates of vaccination.

Whilst these factors are not affecting cholera response in Yemen, at the current time, they are factors to consider if the use of the CRM and rainfall forecasts becomes ‘business as usual’ in Yemen and, if a similar approach is taken in a different context.

4.9 How UNICEF use the CRM and rainfall forecasts in 2020

In 2020, following provision of the CRM and rainfall products to UNICEF for two years, the Met Office, UF and UNICEF reflected on their use, with the objective of identifying ways to improve these in Yemen and to understand how they might be used by cholera response practitioners in other countries.

It should be noted that due to COVID and the pressures linked to the crisis context of Yemen, the engagement between the EACH team and UNICEF was limited and did not provide the anticipated opportunities to enrich understanding of how the CRM and Rainfall Assessments were being used. The insight reflected here is not as in depth as was as planned so some of the considerations and recommendations (described in section 7) have been inferred from the limited dialogue that did take place.

As the epidemic is four years old, epidemiological trends have been able to provide a good understanding in the country of how the disease behaves during the wet and dry seasons and which populations will be most affected.

UNICEF are therefore finding the CRM to be less useful as it "doesn't provide information on anything that is not already known" (personal communication with member of UNICEF Yemen team). This may be due to the following:

- Cholera incidence data from 2017 to 2019 shows that the region is experiencing cases every month in most of the governorates. This could imply that cholera may be inching towards endemicity in the region. Once cholera becomes endemic, then it has a predictable seasonal pattern of occurrence, such as during monsoon season in Bangladesh or India. The massive outbreaks are limited to severe shocks that further damage water infrastructure in a particular region. If cholera is becoming endemic, the transmission mode of the CRM would be more relevant than the trigger mode as this predicts the human-to-human transmission of the disease as oppose to the outbreak.

- UNICEF expressed concerns about the level of accuracy of the CRM so sharing the validation work described in sections 5 and 6 will be important in order to explore whether using the rainfall assessments in isolation is appropriate when the CRM may have value in informing decisions.

A training session was planned in Jordan to support UNICEF to use the CRM but this was not able to go ahead and subsequent meetings with team members suggested they found the product hard to understand and described the advice offered on what action to take as not being relevant (since this was agreed at by the team and national and local levels). The need for training materials and training videos (which can be delivered in person or virtually) could therefore help to improve uptake of the tool.

Use of the rainfall forecasts for cholera prevention has continued, with these primarily being needed during Yemen's rainy season which runs from April to August. The rest of the year is characterised by light or no rain. Warnings of severe or unexpected weather (as referenced above) are also used all year round to inform action.

UNICEF describe scanning the main body of the forecast product (the narrative description, maps etc) for any severe weather indications, with a particular focus on days 3-6 as this is the timescale which accords with the lead times required for action to be feasibly taken in time. The Excel spreadsheet which lists districts, and their associated levels of rainfall is then used to inform interventions. In the understanding that rainfall does not imply cholera, this data is also compared with other data (such as reported cases, sanitation, and movement of people) which is provided on a portal by REACH (4).

4 REACH is a humanitarian initiative providing granular data, timely information and in-depth analysis from contexts of crisis, disaster and displacement. The work of REACH directly feeds into aid response and decision-making by providing accessible and precise information on the humanitarian situation of crisis-affected populations. Its online platform allows access to reports, factsheets, maps and other information products developed by REACH teams worldwide.

Because of the conflict in Yemen, the Yemeni Meteorology Service (the Civil Aviation and Meteorology Authority) does not currently issue regular forecasts. The rainfall forecast is therefore being used beyond cholera actors (e.g. for disaster risk reduction and for refugee camp management).

UNICEF and REACH are in the process of developing a flood early warning service which will consider rainfall forecasts, water catchment data and the vulnerability of populations exposed to flooding to inform flood preparation activity. The current format of data provided by the Met Office does not support integration into other data sets, so the provision of different formats is being explored. When these are operational, the relevance of the rest of the service currently provided can be determined.

4.10 Other stakeholders interested in early action for cholera

There is significant interest among the cholera community in the pioneering, risk-informed approach that UNICEF have been using in Yemen to prioritise their cholera prevention activity.

The Red Cross, Oxfam and academic organisations are among stakeholders who may have appetite to use similar strategies. Before exploring further, they are interested to see if the validation work, done as part of this study, can show what level of accuracy the CRM and rainfall forecasts have and what this means about how they should be used more generally.

The Red Cross have a sophisticated Forecast Based Financing methodology which is used primarily to take disaster risk reduction (DRR) early action in the event of severe weather. However, there is interest in adapting this for response to infectious diseases. In Uganda, cholera early action plans have been drawn up and triggers for these can be piloted if information on cholera risk is available.

The United Nations Office for the Coordination of Humanitarian Affairs (OCHA) supports UN partners to take anticipatory action through the Central Emergency Response Fund (CERF). The CERF allows UN agencies and other humanitarian partners to have direct access to flexible, timely and predictable funding to cover critical gaps, address unforeseen needs and complement response efforts. It also supports country level humanitarian partners to initiate early action interventions to mitigate the risks of deepening crises.

Like the Red Cross, the CERF is used to support early action in disaster risk reduction contexts but there is appetite to trial use of the fund to support early action for infectious diseases, if appropriate tools for indicating risk are available. Engagement with the cholera community to explore how this could be done suggests that there is appetite to bring interventions forward. However, there may be barriers to overcome in terms of stakeholders' willingness to use predictive tools based on the school of thought that cholera is caused by environmental factors (ref section 2.1).

There is also interest from the joint World Meteorological Organization (WMO) and WHO's working group on the use of predictions for infectious diseases who have set up a pilot study to conduct an inter-comparison of different models available.

A key consideration when exploring use of the CRM with other users will be assessing which interventions in their package of cholera response measures could be brought forward for use in an anticipatory manner. Once these have been identified, the lead time for taking these actions and their associated cost can be explored along with the consequences of taking such actions 'in vain'. The necessary confidence needed in the CRM's predictions can then be assessed and the best way of using this alongside other information sources (such as rainfall forecasts and epi data) and cholera models determined. The way cholera risk information is shared (e.g. web platform, emails) and the interplay between national, regional, and global actors in cholera response also needs to be considered.

5. CRM validation in Yemen

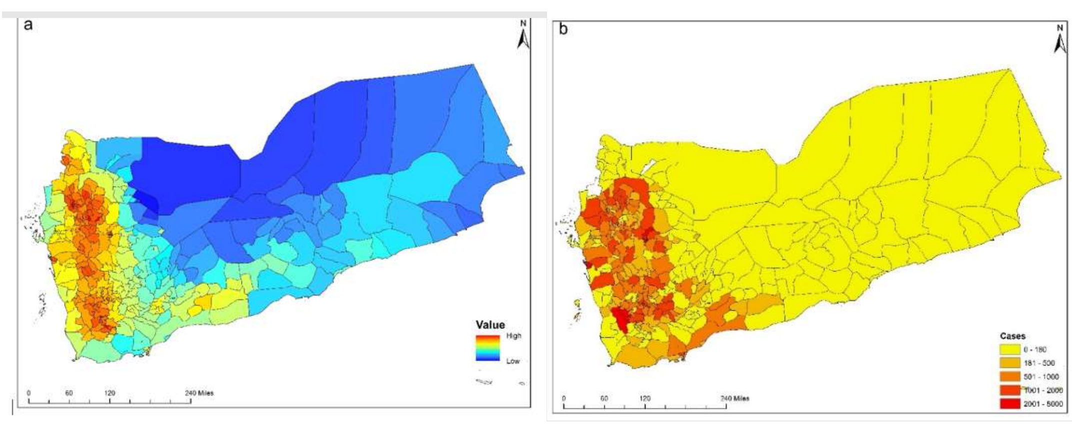
The analysis of the CRM was performed to understand how well predictions of cholera risk generated by the CRM (in trigger mode only) reflected actual cases of the disease in the country.

5.1 Methodology

- Standard correlation tests were used to understand the strength of association between CRM risk scores and cholera cases in all Yemen governorates.
- For each time point in the weekly CRM output, trigger risk scores were computed and compared with total number of cholera cases occurring the following four weeks. That is to say that the computed cholera risk (for the following four weeks) was evaluated in near real-time.
- Epidemiological data for 2017 and 2018 was obtained at governorate levels from the Early Warning and Response Database. Data for 2019 was obtained from ACAPS (an independent information provider operating in Yemen). More detailed data was obtained but not in time for it to be included in this validation work.
- The model's performance was also evaluated against the Bradford Hill Criteria (BHC) (Fedak et al., 2015) which is a framework, although not absolute, comprising a set of ten parameters that provides epidemiological coherency for determining causal relationships between a public health outcome and factors influencing the outcome.

5.2 Validation process and results in full

The first indication of cholera occurring in Yemen was detected in October 2016. Epidemiological prevalence data became available in June 2017, and the trigger module captured the risk of a significant cholera outbreak in Yemen at that time (Figure 5a), with *ca*. 92% spatial match between locations where cholera cases were reported, and high risk (values greater than 0.75) was computed.



Figures 5a&b: (a) Cholera risk map of Yemen for June 2017 produced on May 30, 2017. (b) Actual cholera cases observed in June 2017.

Although the trigger mode of the CRM has been validated using historical data for Sudan, Bangladesh, Mozambique, Algeria, Cameroon, and Haiti, Yemen is unique since epidemiological conditions can be assessed in near real-time.

A well-established public health assessment tool, the Bradford Hill Criteria (BHC) (Fedak et al., 2015; Lucas & McMichael, 2005) was used to validate the CRM. BHC comprises of a set of ten parameters, that provides epidemiological evidence for causal relationships between a public health outcome and factors influencing the outcome, and was used as defining criteria for the CRM's performance. The following sections detail each parameter of the BHC, with respect to model performance.

Strength: This is one of the first parameters of BHC that provides evidence of associational relationship between disease prevalence and factors influencing disease outbreak.

To evaluate strength of the model, correlation was calculated at the governorate level (since the data were available at this scale), namely between cholera prevalence from 2017 to 2019 and risk values computed (for the same time period) using parametric (Pearson) and non-parametric rank correlation coefficient (Kendall Tau scores). Additional details on correlation methods are provided in Annex 4. The Pearson method exhibited significant (correlation and statistical significance values in Table 3) positive correlation for all governorates except Aden, as shown in Figure 5c and Table 3.

Kendall Tau values were statistically significant ($p<0.05$) for all governorates (Table 3). This is indicative of model being able to capture variability in the actual cholera time series, and hence suggest satisfactory model strength.

The three-year correlation analysis (statistically signification correlation values) provided evidence of overall model performance.

	Kendall	Kendall p-values	Pearson	Pearson p-values
Abyan أبين	0.43	0.00000	0.46	0.00000
Aden عدن	0.21	0.00111	0.06	0.49779
Al Bayda البيضاء	0.38	0.00000	0.58	0.00000
Al Dhale'e الضالع	0.43	0.00000	0.49	0.00000
Al Hudaydah الحديدة	0.42	0.00000	0.75	0.00000
Al Jawf الجوف	0.31	0.00000	0.57	0.00000
Al Maharah المهرة	0.21	0.00126	0.32	0.00035
Al Mahwit المحويت	0.43	0.00000	0.58	0.00000
Amanat Al Asimah أمانة العاصمة	0.27	0.00001	0.43	0.00000
Amran عمران	0.27	0.00001	0.46	0.00000
Dhamar ذمار	0.34	0.00000	0.52	0.00000
Hajjah حجة	0.44	0.00000	0.66	0.00000
Haydramaut	0.31	0.00000	0.70	0.00000
Ibb إب	0.39	0.00000	0.48	0.00000
Lahj لحج	0.44	0.00000	0.59	0.00000
Marib مأرب	0.28	0.00001	0.37	0.00002
Raymah ريمة	0.42	0.00000	0.50	0.00000
Sa'ada صعدة	0.20	0.00129	0.32	0.00026
Sana'a صنعاء	0.38	0.00000	0.56	0.00000
Shabwah شبوة	0.42	0.00000	0.72	0.00000
Taizz تعز	0.34	0.00000	0.55	0.00000
P value less than 0.05 implies that the correlation value is 95% statistically significant.				

Table 3: Correlation and p-values of weekly cholera prevalence and CRM outputs (all years taken together)

However, it can be argued that if there were effective intervention strategies, such as robust access to WASH, a decline in model performance over the years would have been observed.

The CRM trigger module is designed to capture disease initiation in a region; therefore, unless there are new outbreak(s), the model performance should decline over the years since transmission dynamics should dictate spread of cholera in a human population.

Accordingly, individual year correlation analysis was performed on a weekly scale. Pearson and Kendall-Tau correlation values for each year is shown in Figures 5d and 5e, respectively. Using Pearson correlation 19 (of 21), 11 (of 21), and 15 (of 20) governorates in 2017, 2018, and 2019, respectively, shows statistically significant ($p<0.05$) association between computed risk and disease prevalence (Figure 5d).

Using Kendall Tau, 19, 8, and 13 governorates showed statistically significant association for the same time period (Figure 5e). In 2017, the model detected increased risk for more than 90% of the governorates, with Aden being the only governorate not determined to be at increased risk.

Decrease in model performance for year 2018 is perhaps an indication of changes in the definition of cholera or impacts of intervention strategies to mitigate cholera in the region. However, increase of model performance for year 2019 was observed. The improvement in model performance is encouraging, however, since the cholera outbreak was an ongoing event in Yemen, it is indicative of two speculative conclusions: (a) there were new outbreaks in the region compared to the prior year which appears to be true since number of cholera cases increased in 2019 from 2018 (Figure 4a) or (b) the lapse of intervention activities that may have resulted in increase in cholera cases in 2019.

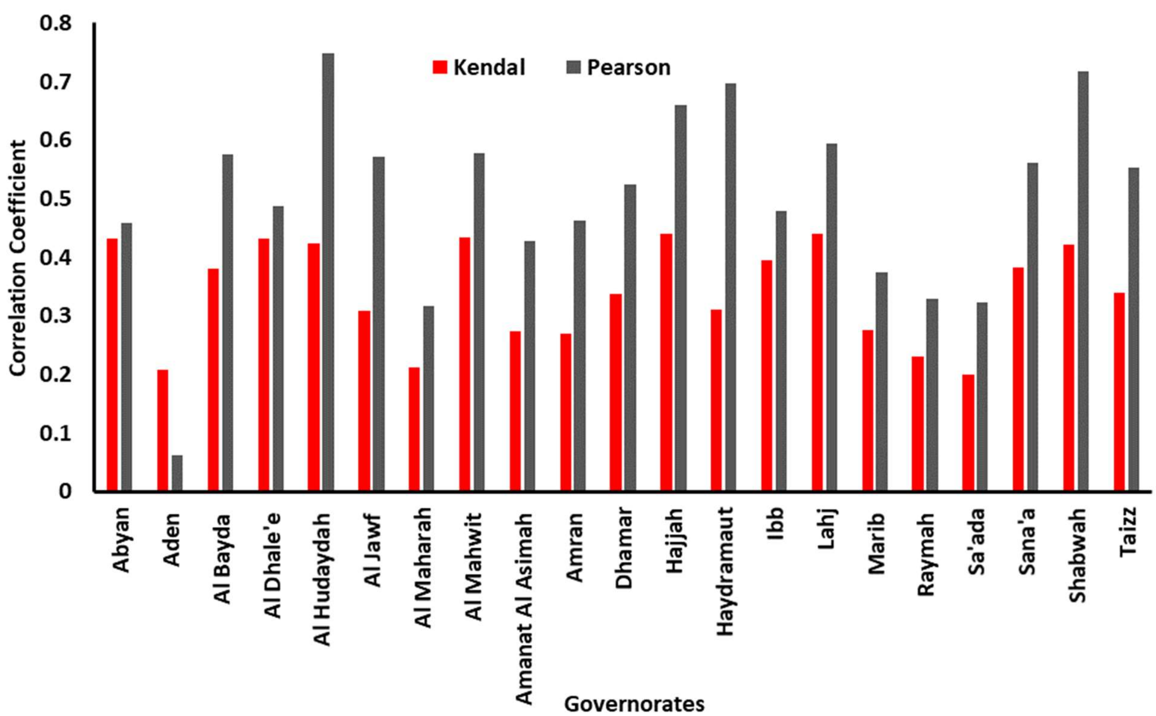


Figure 5c: Correlation coefficients between cholera cases and risk values for all governorates

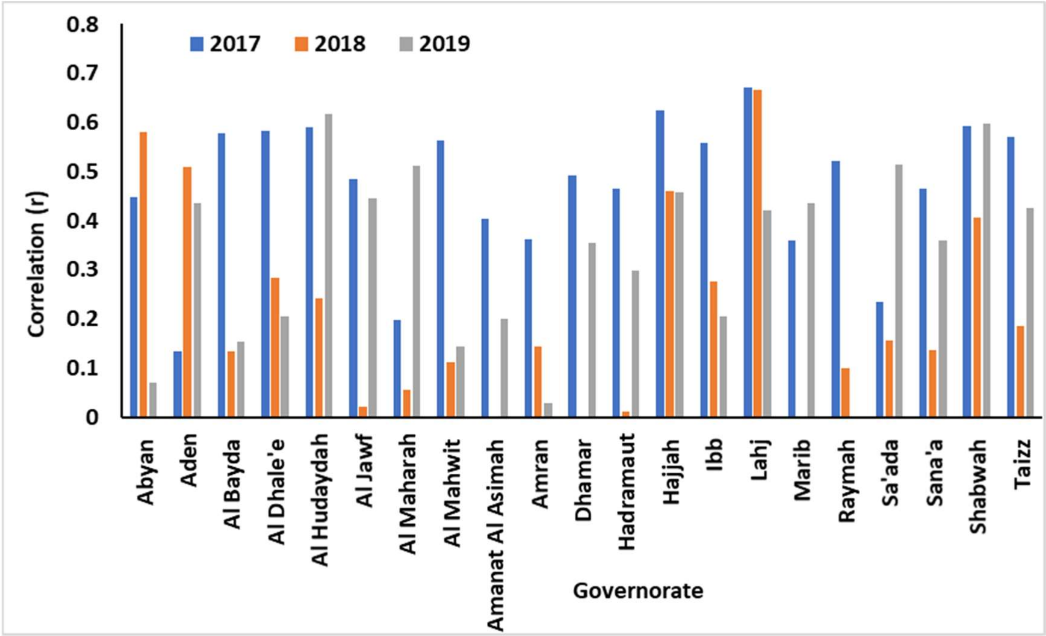


Figure 5d: Pearson correlation coefficient between cholera cases and risk values for all governorates for individual years (2017, 2018, 2019)

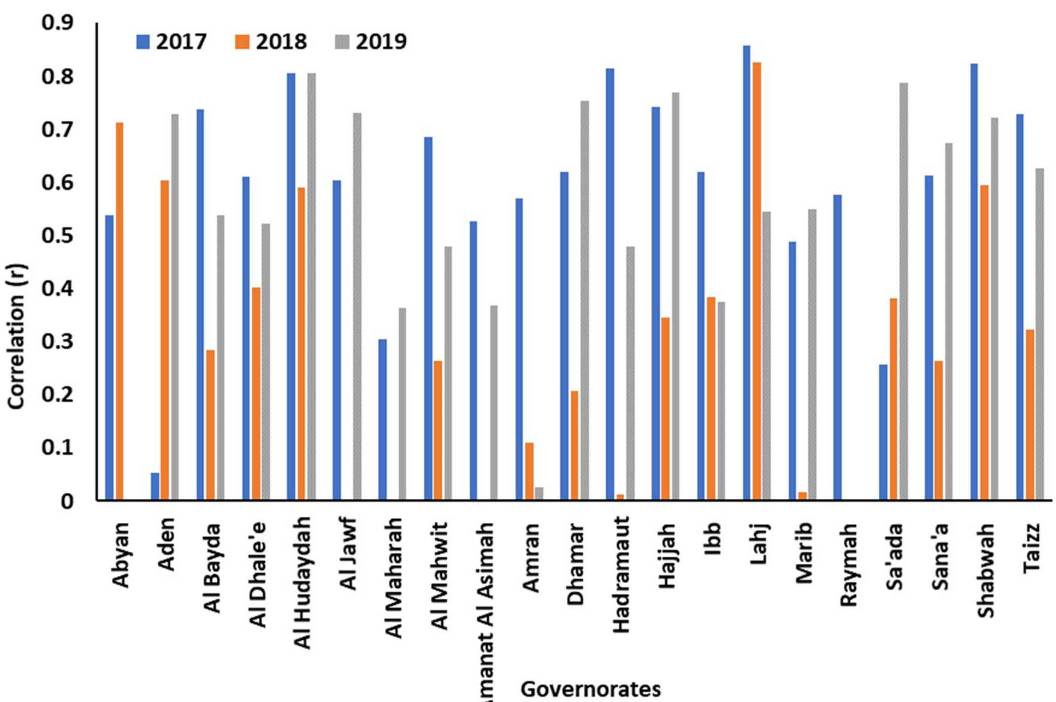


Figure 5e: Kendall Tau correlation coefficient between cholera cases and risk values for all governorates

Specificity: This is the second criterion for evaluating predictive capacity of the CRM and is achieved by quantifying causality of model output with cholera prevalence. Causality is quantified using three statistical metrics: accuracy, sensitivity, and specificity, as defined below:

Accuracy = $(t_p+t_n)/(t_p+t_n+f_p+f_n)$

Sensitivity = $t_p/(t_p+f_n)$

Specificity = $t_n/(t_n+f_p)$

In this study, an increase in risk is considered as a true positive (t_p) if it captures increase in reported cases. A decrease in risk is considered as a true negative (t_n) if it captures the decrease in cases. If increase in computed risk fails to capture the increase in risk, it is considered a false positive (f_p); and if a decreased risk fails to capture the decrease in cases, it is considered a false negative (f_n). The performance matrix with these variables for various governorates is provided in Table 4. As shown in Figure 5f, the cholera risk model met all three criteria for causality more than 60% of time for nine governorates where more than 100,000 cholera cases had been reported and all together these governorates represent about 80% of Yemeni population. Sensitivity and specificity varied from 55% to 67%, with averages of 60% and 61%, respectively, indicating ability of the model to capture increase and decrease in cholera cases in the region. Accuracy varied between 57% to 67%, with an average of 60 % for nine governorates (which represent 80% of Yemeni population) where more than 100,000 cholera cases had been reported. In summary, causality results of the CRM were 60% (averaged) for the key three statistical metrics of accuracy, sensitivity, and specificity for nine governorates. Results for the other 12 governorates are not shown as those regions only represented 20% of population with sparse cholera cases over three years of study.

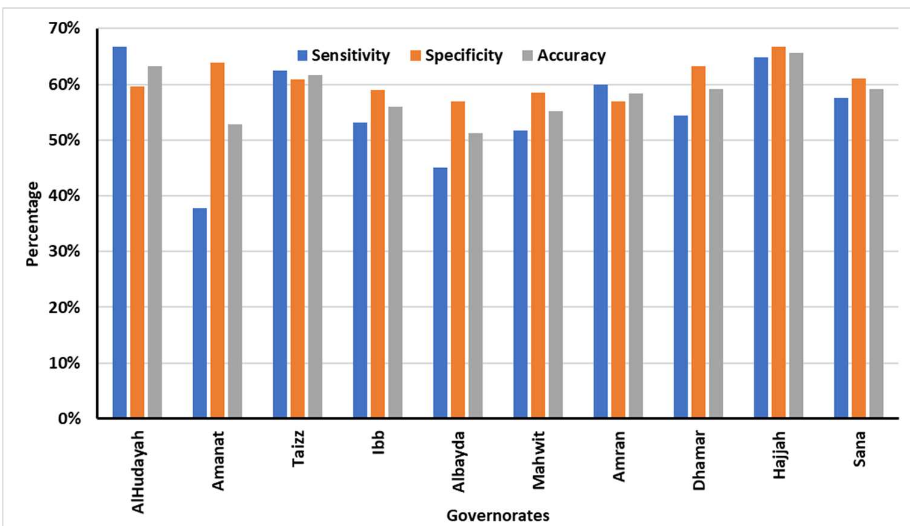


Figure 5f: Sensitivity [$t_p/(t_p+f_n$], Specificity[$t_n/(t_n+f_p$], and Accuracy[$(t_p+t_n)/(t_p+f_p+ t_n+f_n$] of CRM.

Conditions	Al Hudaydah	Amanat	Taizz	Ibb	Mahwit	Amran	Dhammar	Hajjah	Sana
Fn	21	33	21	30	29	24	26	25	28
Fp	25	26	27	25	27	28	25	18	23
Tn	37	46	42	36	38	37	43	36	36
Tp	42	46	35	34	31	36	41	46	38

Table 4: Performance matrix for various governorates in Yemen (2017 – 2019)

Biological Gradient: This parameter is traditionally interpreted as a monotonic gradient, indicating direct proportionality of cause of the increase in a disease burden with exposure risk. It has been argued that BHC should include non-monotonic and complex relationships between cause of the trigger and transmission of the disease with variation in the reported disease cases (Fedak et al., 2015). Thus, in this study, to evaluate the biological gradient, the relationship between change in the computed risk rate using the model with change in reported disease prevalence was explored. To evaluate the incidental (biological) gradient of the model, the positive predictive value (PPV; frequently referred to as precision) and negative predictive value (NPV) of the model outcomes were computed. PPV is a fraction of positive computed risk, which can capture positive change in prevalence. In this study, an increase in relative risk and reported disease was used rather than absolute values. NPV is a fraction of negative (decreased) computed risk, which can capture negative change in prevalence. These two indicators were computed for the same governorates (the nine governorates where more than 100,000 cholera cases had been reported) for which we evaluated specificity (causality) of the model and were calculated as follows:

Positive predictive value = $t_p/(t_p+f_p)$ [4]

Negative predictive value = $t_n/(t_n+f_n)$ [5]

Using the three-year data, we determined PPV and NPV for the model (Figure 5g). PPV and NPV values varied between 57 % to 67 % and 55 % to 67 %, respectively, with an average of 60 %, suggesting *ca*. 60 % of the time, the model correctly responded to increase or decrease in number of reported cholera cases. However, for practical use of the cause-effect relationship, temporality needs to be included, with cause preceding effect with a lead time.

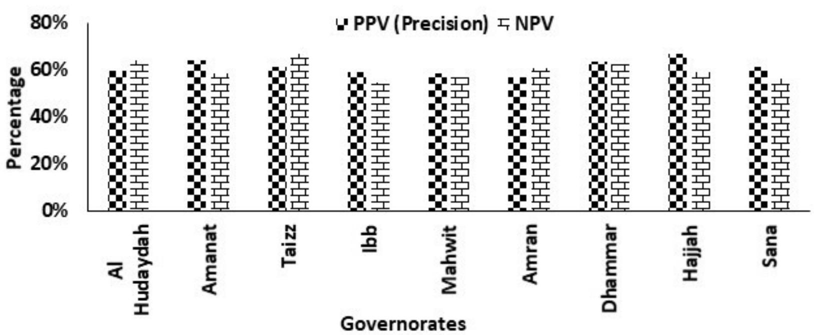


Figure 5g: Positive Predictive Value (PPV) or precision [$t_p/(t_p+f_p)$], and Negative Predictive Value (NPV) [$t_n/(t_n+f_n)$] of the trigger module of cholera prediction system.

Temporality: Epidemiological understanding of temporality is exposure duration and extent of its impact in terms of severity or number of incidences. To assess disease risk prediction, lead time is an essential criterion because it provides time to intervene and limit impact of a disease outbreak.

In risk modelling, these lead times can be evaluated in terms of temporality since risk precedes incidence of disease. Risk computed using the CRM provides an assessment on likelihood of trigger of cholera for the next four weeks (Antarpreet Jutla et al., 2015), providing ample time for intervention and mobilisation of resources. Our previous studies quantified the role of environmental, climatic, and sociological processes that influenced an outbreak of cholera in a population (Anwar Huq et al., 2013; Luque Fernández et al., 2009). The hypothesis presented in Figure 2b shows that cholera cases are generally observed four weeks after anomalous warm temperatures followed by anomalous high precipitation in those locations where there is significant deviation in the behaviour of the population with respect to water use habits caused by damaged WASH infrastructure. The hypothesis for temporality of BHC, has been tested in several regions in Africa (Antarpreet Jutla et al., 2015), Asia (Rakibul Khan et al., 2018a) and Latin America(Anwar Huq et al., 2017c).

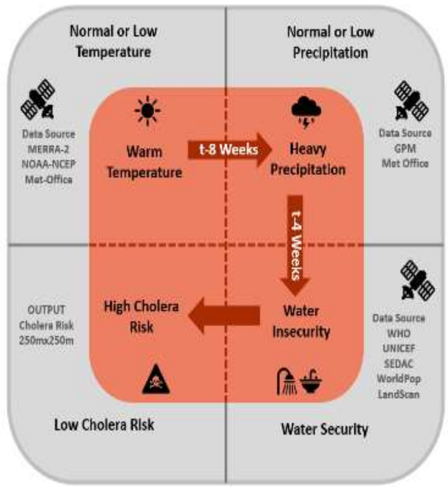


Figure 2b (Repeat): Cholera trigger mechanism

Consistency: UF's key hypothesis in Figure 2b argues that damaged WASH infrastructure and a combination of hydroclimatic processes favour conditions for an outbreak of cholera. This cause and effect relationship has been observed in many studies (Huq et al., 2013; Khan et al., 2018b; Lipp et al., 2002).

Attributing the cholera outbreak in Peru to El-Niño events in Central Pacific was one of the earliest precursors to this hypothesis (Colwell, 1996). Studies conducted using data from Bangladesh (Hashizume et al., 2008) and Haiti (Khan et al., 2017) report a strong relationship between rainfall and incidence of cholera. In Bangladesh, cholera occurs annually in a bimodal cycle. The first peak occurs in the spring, and a larger peak occurs following the fall monsoon season. Cholera seasonality also coincides with warmest temperatures of the year and is reduced to sporadic incidence as the temperature decreases in winter (Lipp et al., 2002). Haiti has been a main focus of cholera research since the 2010-11 outbreaks, which identified rainfall as a critical driver of the disease in that country (Blokesch et al., 2012). Rainfall can have a significant impact on water resource, e.g., nutrient concentration, salinity, pH, river level, and freshwater discharge, which in turn affect growth and persistence of V. cholerae and its zooplankton host in the environment. Various studies have determined dependence of air temperature and precipitation as dominant hydroclimatic variables impacting occurrence and transmission of cholera in various parts of the world (Table 5). The hypothesis has been validated for countries in Africa, Asia, and the Americas, which reinforces its repeatability.

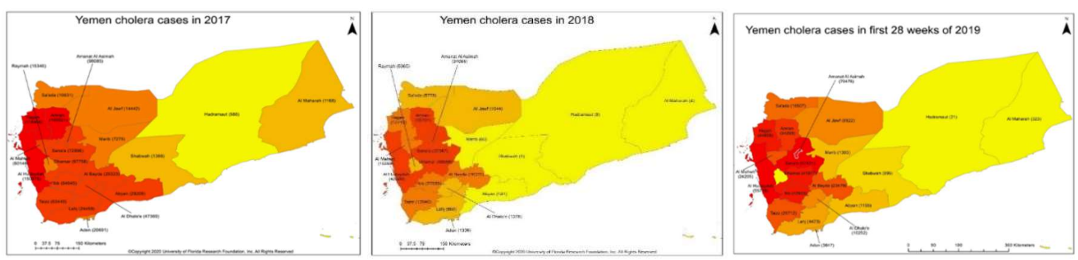


Figure 5h: Total cholera cases in Yemen in 2017, 2018 and first 28 weeks of 2019

Authors	Region	Citation score	Climatic variable
Pascual et al 2000	Bangladesh	587	Temperature
Lobitz et al 1999	Bangladesh	516	Temperature
Rose et al 2001	Bangladesh and Peru	492	Temperature
Huq et al 2005	Bangladesh	342	Temperature, Rainfall
Griffith et al 2006	Africa	252	Rainfall
Louis et al 2003	USA	214	Rainfall, Temperature
Rinaldo et al 2012	Haiti	150	Rainfall

Table 5: Evidence of consistency using BHC

Plausibility: Cholera is a disease caused by *Vibrio cholerae*. A warm climate facilitates growth and proliferation of the bacterium and enhances metabolic activity. When environmental conditions are unfavourable, vibrios have been shown to enter a protective, viable but non culturable (VBNC) state whereby the bacterial cells become metabolically dormant (Colwell, 2000; Roszak & Colwell, 1987). Subsequently, when environmental conditions return to favourable, typically triggered by temperature, salinity, and nutrient modification, VBNC cells regain cultivability and virulence potential (Oliver, 2010; Roszak & Colwell, 1987).

Elevated water temperature causes a density differential amongst layers of water in an aquatic reservoir, contributing to stratification of bacterial populations. In addition to stratification related to temperature, dissolved oxygen, pH, and other physical/chemical parameters determine non-uniform microbial community profiles in the water column (Huq et al., 2005b; Louis et al., 2003).

This non-uniformity contributes to those environmental conditions enhancing bacterial growth and multiplication. These conditions are favourable for multiplication of zooplankton, namely copepods and other chitinous zooplankton, shown by Kaneko and Colwell (1975) to host vibrios, including *V. cholerae*. The three-year study funded by the US National Institutes of Health and carried out by Colwell and Huq (Huq et al., 1996) in Bangladesh that included villages in Bangladesh whereby simple filtration using sari cloth material as filters resulted in a nearly 50% reduction in cholera. The hypothesis was that having proven that copepods were host/vector of *Vibrio cholerae*, by removing copepods and particulate material from household water, cholera case numbers would be reduced.

That cholera is a dose dependent disease, requiring ingestion of *ca*. 10^6 V. cholerae cells/ml water (Cash et al., 1974), was the fundamental principle of the hypothesis that removal of 99% of V. cholerae as had been demonstrated in laboratory experiments (Huq et al., 1996) would reduce the number of cases of cholera. This was the basis of studies carried out in the remote villages of Bangladesh. Thus, employing simple sari cloth filtration by village women effectively removed zooplankton and particulate matter from drinking water and reduced exposure to V. cholera and the number of cases of cholera by nearly 50% (Colwell et al., 2003; Huq et al., 1996).

The ecological parameters enhancing growth and proliferation of cholera bacteria frame the model developed for risk prediction. Heavy rainfall that follows a period of high air temperature aids explosive growth of bacteria in water bodies serving communities as drinking water source (Huq et al., 2013). Thus, an inadequate water supply infrastructure exposes a population to untreated water. Yemen, a Middle Eastern country grappling with war and frequently experiencing floods, is a region with population exposed to poor WASH conditions, a factor by far a dominant sociological cause of the continuing cholera epidemic rampant in most of its governates. Thus, identifying and describing mechanics of the trigger, a rational clarification of the 'black box' between the biology and ecology of infectious agent and disease epidemiology are now possible.

Coherence: In the environment, an increase in *V. cholerae* populations was observed in water and plankton samples collected in a longitudinal, multi-year study carried out in the Chesapeake Bay, Maryland. When water temperature was above 19°C, *V. cholerae* populations in the water column proliferated with elevated temperatures (Louis et al., 2003). Similarly, water samples collected in estuarine zones of the Bengal Delta yielded similar results confirming enhanced growth of *V. cholerae* in warm pond water (Neogi et al., 2018). Furthermore, 5°C increase in water temperature resulted in a 30-fold increased risk of a cholera outbreak with a lag of six weeks (Huq et al., 2005b). Additionally, Pruzzo, Vezzulli, and colleagues reported a global warming trend in sea surface temperature was found to be strongly associated with proliferation of populations of *Vibrio* spp. and emergence of *Vibrio* related disease (Vezzulli et al., 2016). These observations in combination with findings from laboratory experiments conducted in different regions of the world (Table 5), comprise a crucial validatory underpinning experimental evidence supporting the hypothesis of this study (Figure 2a).

Experiment: Various studies have associated the variability in hydroclimatic variables with cholera trigger and transmission risk (de Magny et al., 2008; Huq et al., 2017b; Jutla et al., 2010a; Jutla et al., 2015). As discussed above, laboratory-based investigations showed *V. cholerae* thrives in water with temperatures between 20 and 45°C (Martinez et al., 2010). Experimental studies have shown increased risk of cholera when ambient air temperatures rise to between 19 and 28°C (Hood & Winter, 2006; Louis et al., 2003) and there is also increasing water entrapment (Huq et al., 2013; Khan et al., 2019). Temperature and precipitation are essential factors, but separately do not trigger or control cholera spread (Paz, 2009). The combination of warm temperature convergence with heavy rainfall and inadequate WASH infrastructure (Sasaki et al., 2009; Taylor et al., 2015) leads to outbreak of the disease (Huq et al., 2013; Jutla et al., 2015). Since *V. cholerae* is a naturally occurring inhabitant of the aquatic environment (Grim et al., 2009; Mishra et al., 2011), conditions favouring its growth and multiplication show incorporating a single parameter provides, at best, an incomplete description of disease trigger and transmission of cholera. The CRM incorporates both temperature and precipitation as hydroclimatic variables with experimentally demonstrated significant association with *V. cholerae* proliferation and cholera. CRM provides a tool for policymakers and an intervention committee to implement preventative measures in regions that it is based on trigger and transmission dynamics.

Analogy: It is unfortunate, but cholera outbreaks are a regular phenomenon in regions subjected to positive anomalous precipitation associated with positive anomalous ambient air temperatures (hydroclimatic conditions) (de Magny et al., 2008; Huq et al., 2013; Jutla et al., 2010b; Khan et al., 2017) in regions with damaged WASH infrastructure(Huq et al., 2013; Khan et al., 2018b). Previous results suggest that the odds of occurrence of cholera is 1.5 times higher when precipitation is higher than climatological average (Jutla et al., 2013). Similarly, chances of cholera increase 6-fold if the air temperature was above climatological average for 2 months preceding the disease outbreak (Jutla et al., 2013).

Spatial analyses done using data from India, Bangladesh, Nepal, Mozambique, Cameroon, Central African Republic, Congo, and Zimbabwe exhibit a similar hydroclimatic pattern related to cholera outbreaks (Ali et al., 2015; Jutla et al., 2010b; Khan et al., 2017; Nasr-Azadani et al., 2016; Sharp & Salaam-Blyther, 2017).

Damaged WASH infrastructure accelerates interaction between *Vibrio cholerae,* thereby enhancing characteristics of the available water resource, notably lack of safe water, sanitation, and hygiene, increasing the likelihood of waterborne disease in the population. In 2015, Nepal demonstrated how sufficient WASH infrastructure can reduce an outbreak, even under hydroclimatic conditions favouring an outbreak of cholera (Khan et al., 2018a). To control spread of cholera in a population, policymakers could be recommended to emphasize primarily WASH infrastructure to achieve reduction of cholera, as was observed in 2018

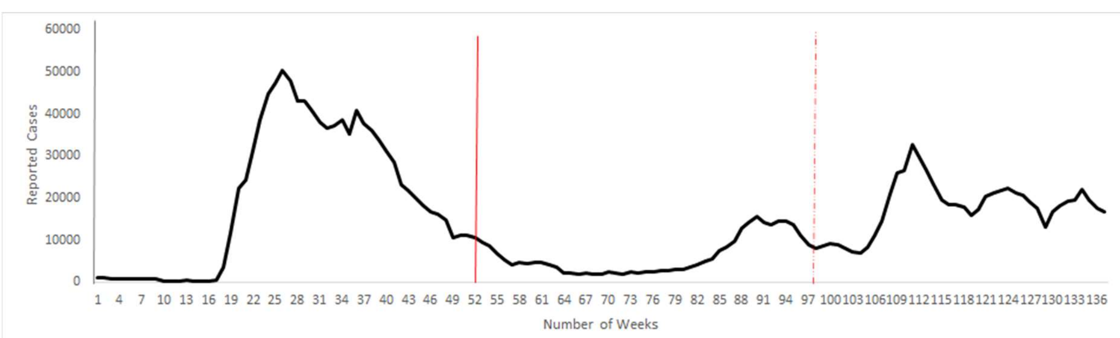


Figure 5i: Total number of cholera cases from 2017 to 2019 in Yemen.

Reversibility: WASH, one of the parameters included in CRM trigger analysis provides a means for testing reversibility of the model. After the 2015 earthquake, Nepal presented environmental conditions that could be considered favourable for a massive cholera epidemic (Khan et al., 2018b). However, the observed cholera cases were lower than anticipated disease burden in that region. The lower number of cholera cases than envisioned was attributed to effective WASH infrastructure and WASH facilities appear to have played a crucial role in controlling the Nepal cholera outbreak. The CRM, by incorporating WASH, precipitation, and ambient air temperature, can test reversibility.

Any variable failing to demonstrate positive anomalous variability from long-term average can be interpreted as indicating minimal cholera risk.

5.3 Summary of methodology and results

- The CRM was able to statistically capture variability in the occurrence of cholera in Yemen in 20 of the 21 governorates (for the time series 2017-2019).
- For governorates comprising 80% of Yemeni population and where total cholera cases are high (more than 100,000), the model was able to predict risk of cholera with 60% of accuracy.
- The objective of this study was to validate the near-real time prediction of cholera risk provided by the CRM from 2017 to 2019. Bradford Hill Criteria (BHC), a widely used epidemiological fact-finding criterion, was employed to assess the performance of the trigger mode of CRM. BHC is composed of ten parameters that provide epidemiological evidence for causal relationship between public health outcome and factors influencing an outcome. The ten criteria factors of BHC shows that outputs from CRM pass an epidemiological test of causation of cholera with environmental and sociological parameters embedded within the architecture of CRM.
- Using BHC, analysis of sensitivity, specificity, accuracy, and precision indicate that changes in model risk scores predict a change in the number of cholera cases.
- The performance of the model, based on correlative (values range between 0.3 to 0.7) assessment, was best in year 2017 followed by year 2019 and then year 2018.
- Only the trigger mode of the CRM was used, and therefore, the ability of the model to associate CRM outputs with 2017 cholera cases is significant since it implies that the model has relevance in forecasting cholera risks for up to the following four weeks. The trigger model captures cases when disease outbreak starts. A limitation is that case definitions change.
- Al Hudayah (one of most densely populated governorates) had a statistically significant association of risk with number of cholera cases throughout the year. At the time of writing this report, we do not have enough information on interventions in this specific region.

5.4 Considerations

- The lack of performance of the model in 2018 is an indication of either effectiveness of intervention activities or the impact of changes in the definition of cholera cases in the region. This highlights the importance of collection and preservation of long-term surveillance data in regions with poor water and sanitation infrastructure. Also, global rule-sets should be adopted as and when cholera cases are reported at a new location so that a consistent recording of data on intervention activities can be ascertained.
- The performance of the CRM in 2019 was lower than 2017 when a cholera outbreak was reported in Yemen. The increase in performance (a greater number of governorates showing statistically significant correlation with risk) is perhaps an indication of either lapse of intervention activities or new cholera outbreaks occurring in the country.
- Assessment of impacts of interventions were not included in the analysis. This was due to non-availability of data on provisions on water, sanitation and hygiene.
- The analysis shows cholera cases being reported consistently in the country each week, from 2017 to 2019, indicating cholera in the region may be transitioning toward endemicity (seasonal outbreaks of cholera over several years). Cholera cases are still being reported (as of March 20, 2020) in Yemen. The CRM outputs provided are currently in the trigger mode of the model which is developed for epidemic cholera. The transmission mode (for endemic cholera) may therefore be more relevant to UNICEF, if understanding on how this is used is considered in collaboration with UNICEF.
- Access to safe water and sanitation is based on assumptions to create a blanket baseline across the country and remains static in the CRM's algorithm. Different communities will have better or worse access to safe water and sanitation than others. If information on

WASH infrastructure was available, it is expected that values for sensitivity, specificity, accuracy, and precision metrics would improve significantly.

5.5 Limitation(s)

- The results presented in this study are based on governorate level. The large spatial averages of epidemiological data on governorate scale limits ability for accurate determination of hot spots of this disease.
- Data on intervention and mitigation activities, as well as WASH was not available. This poses limitations to near real time updates on accessibility of safe drinking water and adequate sanitation facilities.
- The CRM uses static information on population density. Therefore, information on population movement should be encouraged routinely rather than sporadically.
- There is an inherent limitation on the spatial resolution of the model. Currently, the results are averaged on the governorate scales (1km sq.), which may not be appropriate for the decision making at more local levels. Better understanding needs to be gained on how UNICEF Yemen make decisions to take preventative action and what size areas these decisions apply to.

5.6 Recommendations

- In many countries where epidemic cholera is a frequent occurrence, preventative measures will already be underway in anticipation of an outbreak (often on a seasonal basis). In these contexts, the information from the CRM can be used to support planning and preparations and to intensify cholera control measures in areas which the CRM predicts to have a higher risk of the disease.
- Forecast rainfall data should be ingested into the CRM and performance should be evaluated (based on methodology provided in our previous studies: (Huq et al., 2017c; Khan et al., 2018a). This would mean the model would still have a 4 week validity from the date of issue, but forecast data would give a 1 week lead time into this which may enable more meaningful and proactive interventions.
- Intervention and mitigation activities for any water-borne infectious disease are often planned after an outbreak has been reported in a particular region, which is a reactionary approach. Based on its use in the Yemen context, it is proposed that the CRM has the potential to be used to support planning and preparation and earlier control measures in contexts with epidemic cholera and when used as part of a suite of tools used by cholera responders.
- If cholera prevention and intervention activities were taken based on the CRM's cholera risk information, a reduction in the model's performance would have been observed (i.e. the CRM may show cholera risk but it may not be related to cholera cases due to impacts of interventions having taken place). Because the model's predictions of cholera and actual incidence of cholera are statistically significant, one interpretation is that this reflects the understanding that various aid agencies are not currently using the CRM in operational decision making. The need for provision of support in the form of training and guidance in how to use and interpret the CRM (currently in progress) is therefore underlined.
- A prototype of cyberinfrastructure (which acts a repository for epidemiological and CRM data, enables visualisation and provides decision making information) is essential to provide anticipatory and logistical support in case of high risk of cholera. Resources such as CATAPULT or NASA Disaster platforms may be engaged to develop such cyberinfrastructure support. This is critical since there need to be a credible source of information to various health ministries and NGOs.
- Decisions made using cholera risk predictions should be tailored to local contexts and actors.

6. Validation of the Rainfall forecasts

6.1 Introduction

In addition to evaluating the CRM, which was presented in the previous section, the case study evaluates the historical use of meteorological information in Yemen. This is particularly important in terms of understanding the added value, and complimentary role these forecasts played. To understand the value of the meteorological information provided to UNICEF, we undertook a verification of the rainfall forecasts. The precipitation from the forecast models that are used to create the Weekly Rainfall Assessments (WRA), specifically the Met Office's Global Model (GM, Walters et al. 2019) and the Crisis Area Model (CAM) are analysed. They are compared to observational satellite data to assess how accurate the forecasts are and provide an assessment as to what spatial scale they are accurate at, with the view for them to be integrated into an enhanced CRM in the near future. The statistical relationship between the WRA and the weekly cholera cases is examined to determine whether the reduction in cholera can, at least in part, be attributed to the provision of the WRA.

6.2 Methodology

To carry out the verification of rainfall forecasts over the Yemen, daily precipitation accumulation data from the CAM and the GM are verified against satellite-derived precipitation data to assess the spatial skill of the models. The satellite data used for the verification is Global Precipitation Measurement (GPM) data, as used in the CRM. However, the version of the GPM used for the forecast verification is the Level 3 Integrated Multi-satellitE Retrievals for GPM (IMERG) (Huffman et al. 2014). This final run global product goes through additional calibration and integration with available precipitation gauge analyses from around the world. Due to the lack of in situ data in Yemen, there is no calibration of the satellite product over this region. Figure 6a shows the region over which the precipitation forecast verification was performed. A discussion on the limitations of using satellite data, specifically over the Yemen, due to the lack of in situ data, can be found in Annex 5.

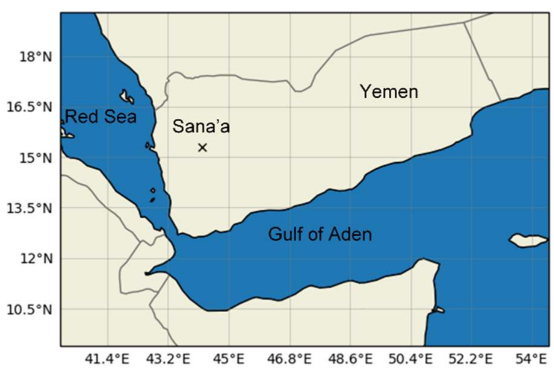


Figure 6a: Area of verification of rainfall forecasts

There are two characteristics to precipitation that are important in conducting a verification analysis, intensity (in this case daily accumulations), and spatial accuracy. In this study the two characteristics are investigated separately, to provide: i) a clearer analysis of the forecasts, ii) better guidance, and iii) increase confidence for the users and to give stronger evidence for recommendations for future model developments. The approach used in this study categorises the daily precipitation accumulation into binary outcomes, depending on whether the accumulation is above or below a specific threshold for each grid cell, namely 20 mm/day for the CAM and 10 mm/day for the GM, the "Take Action" thresholds discussed later. The precipitation events forecast can be compared to the observed precipitation at each grid cell and entered into a contingency (hit-miss) table. The results from this analysis are shown in the next section and discussed in full in Annex 6.

The contingency table approach allows analysis of the same variable in different data products, such as forecast precipitation and observed precipitation. The output is very useful but can be misleading when the forecast and observed precipitation events are far apart with spatial inaccuracies, which would not be identified by the contingency table analysis. Generating many contingency tables for different areas and with a range of regions is not realistic. To overcome this, a more suitable approach to analyse the spatial accuracy of the precipitation events is to calculate the Fractions Skill Score (FSS). This evaluates the accuracy of precipitation forecasts against satellite observations for smaller areas, then average these areas across the region of interest. The results of the FSS analysis are shown in Section 6.2.

The GGU use four precipitation categories in the WRA product, light rain, rain, heavy rain, and storm. An additional category was added, “Take Action”, that corresponds to a threshold used by UNICEF to target WASH interventions (Personal Communication, Bulit, August 2019), which is the threshold investigated in this study. The thresholds (in mm per day) for these categories are shown in Table 6. Due to the models having different spatial resolutions, the thresholds for each rain category are different between the two models.

Precipitation (mm/day)	“Light rain”	“Rain”	‘’Take action’’	“Heavy rain”	“Storm”
CAM	0.1	2	20	50	150
GM	0.1	1	10	25	75

Table 6: Precipitation thresholds used to create a range of binary fields in mm per day.

6.3 Evaluation of the rainfall accumulations from the forecasts - Contingency Tables

The contingency tables analysis considers the Yemen domain (100 x 147 grid cells), as shown in Figure 6a, for each month. The cells of the contingency table contain Hits, False Alarms, Misses, and Correct Rejections and are defined as follows:

- A **Hit** is defined as one or more grid cells, in both the forecast and satellite Yemen domain, where the precipitation exceeds the absolute threshold.
- A **False Alarm** is when none of the grid cells in the satellite data record precipitation above the threshold and at least one grid cell in the forecast has a value greater than the threshold.
- A **Miss** (or **False Negative**) is where none of the grid cells in the forecast have a precipitation event greater than the threshold and at least one of the grid cells in the satellite data recorded precipitation above the threshold.
- A **Correct Rejection** is defined as when there are no grid cells, in either the satellite or forecast data, where the precipitation exceeds the threshold.

A near perfect forecast would have a high count of Hits or Correct Rejections compared to the sample size and a corresponding small number of False Alarms or Misses. Figure 6b shows the results of the contingency table analysis for each month in the period January 2017 to August 2019, with the cholera count for each month shown as a black line, for the whole Yemen domain. Yellow and teal/green bars show when the forecast model and observations agree, with disagreements being shown as purple and dark blue.

Figure 6b shows that the model and observations typically agree over 70% of the time for the whole period (2017 to August 2019), with the day 2 forecasts (which are provided from the CAM) showing the best performance. For the wet season, the model and observations agree over 80 % of the time, with the GM for both the day 3 and day 6 forecasts performing even better than the CAM for days 1 and 2. The performance of the GM is consistent for the length of the forecast, i.e. the forecasts for day 6 performing similarly to the forecasts for day 3. There are also a number of months which show a perfect forecast (hits and correct rejections only), e.g. August 2017, July 2018, and August 2018 (for the GM), with several months showing a near-perfect forecast (very small number of misses and false alarms). This methodology does not consider any differences in the location of the precipitation; this is investigated in the next section.

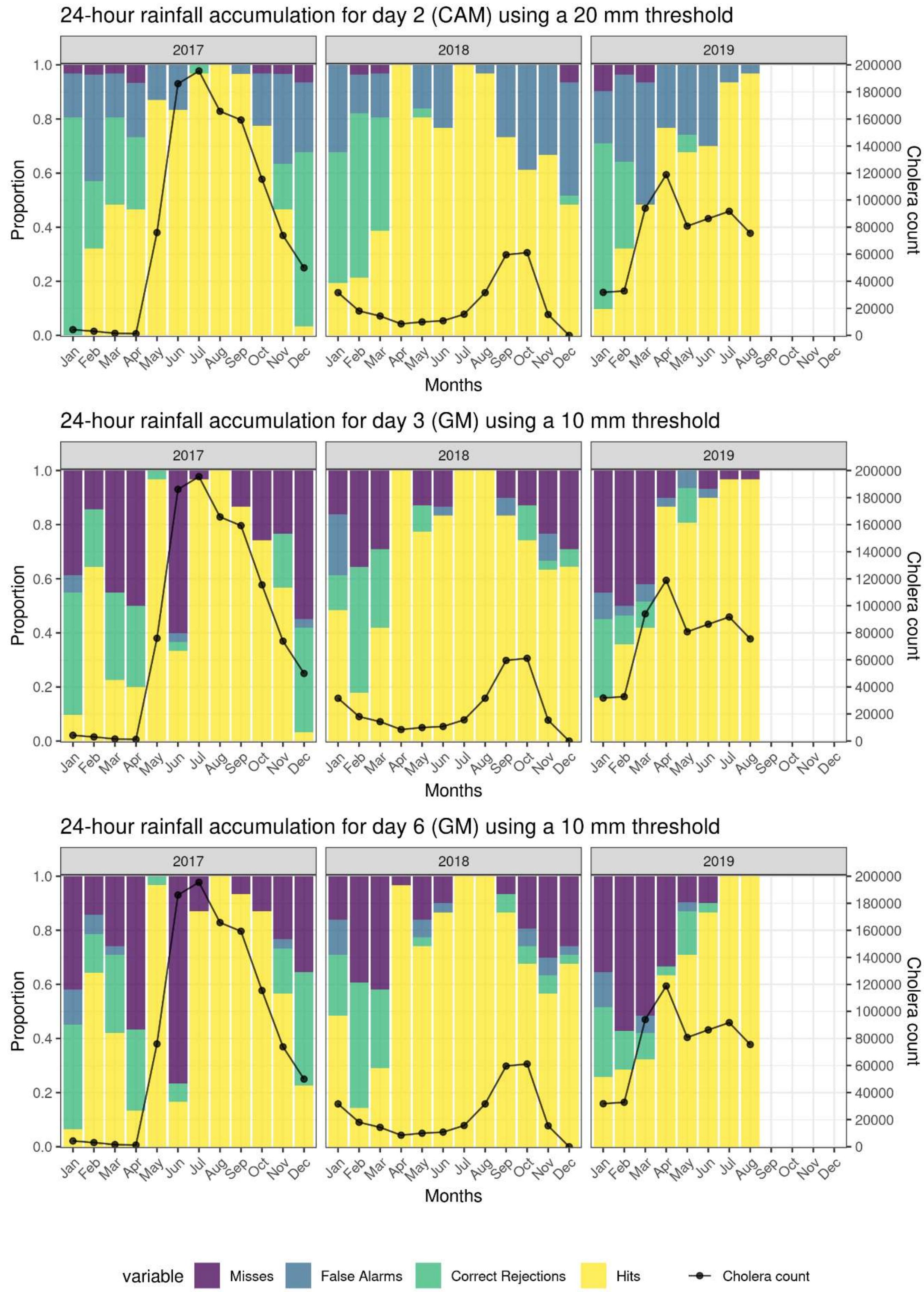


Figure 6b: Stacked histogram showing the proportion of Hits, Correct Rejections, Misses and False Alarms for each month between 2017 and August 2019 for the 24-hour precipitation accumulation for day 2 (top row) using a 20 mm threshold and day 3 (middle row), day 6 (bottom row) using a 10 mm threshold, the "Take Action" thresholds. The dotted line shows the cholera count for that month.

6.4 Evaluation of the spatial accuracy of forecasts - Fractions Skill Score (FSS)

Due to the complex and chaotic nature of precipitation processes, precipitation should not be interpreted at the grid cell resolution alone but as an average over several grid cells. As a result, any scheme that looks to assess the accuracy of precipitation forecasts should consider not just the accumulated amount of precipitation but also the spatial variation of forecast skill around the location of interest. For this purpose, we used the FSS (Roberts and Lean 2008) method, whose full details are presented in Annex 7.

This method has been applied, using a fixed accumulation threshold to define a rain event, to reflect how the data is used by UNICEF. Therefore, a seemingly poor performance of the model than expected may be suggested by this method, especially when compared to other FSS studies. The main reasons are due to double-negative scoring when an area of precipitation in one of either the model or observations do not exceed the fixed threshold, and the effect of observation bias discussed in Annex 5. However, it is still considered an appropriate way to apply this method for this study.

In this study dry days were excluded. This is due to the FSS method not being suitable to compare domains which are predominantly dry, resulting in low scores which do not reflect the true performance of the model. If less than 0.5 % of the domain exceeded the threshold this was considered a dry day and excluded from the FSS analysis. A plot of FSS against the length of the neighbourhood square can be used to identify the spatial scale at which the forecast becomes useful, FSSufc. The displacement error is half the neighbourhood length identified from the FSSufc line. The displacement error indicates how large a difference in location there could be between where the precipitation is forecast and where it is observed.

The size of the domain used for the FSS analysis needs to be sufficiently large to capture the large-scale meteorological processes that are the main drivers of the precipitation (Mittermaier and Roberts, 2010), e.g. the ITCZ and RSCZ. However, this will result in disassociated rain events being analysed and compared, e.g., rain in western Yemen and rain over the Horn of Africa. If either of these are not captured in one of the satellite data or forecast model, due to the bias and using a fixed threshold, this will result in no FSS performance, or an FSS the size of the domain. There is no easy solution to this issue and is the cause for some of the spread in the results (grey lines) shown in Figure 6c. Further analysis of the FSS results is presented in Annex 8.

The daily FSS plots for Day 3 in August 2017, 2018 and 2019, excluding dry days, are shown in Figure 6c. Each grey curve represents the FSS for each day in the month, when more than 0.5 % of the domain exceeds the threshold. The average FSS for these days is represented by the solid red curve, along with the 95 % confidence interval (dotted red curves). Figure 6c shows the FSS increasing as the size of the neighbourhood size is increased, as expected. The blue horizontal line represents FSSufc, the line that demonstrates the spatial scale at which a forecast can be considered useful (Roberts and Lean 2008). To add clarity in Figure 6c, the solid purple vertical lines guide the reader to the approximate useful mean spatial scale (neighbourhood size), along with the bounds (dashed lines) of the upper and lower 95 % confidence intervals (CI). Plots were produced for all months between April and November (the wet season) and are presented in Annex 8.

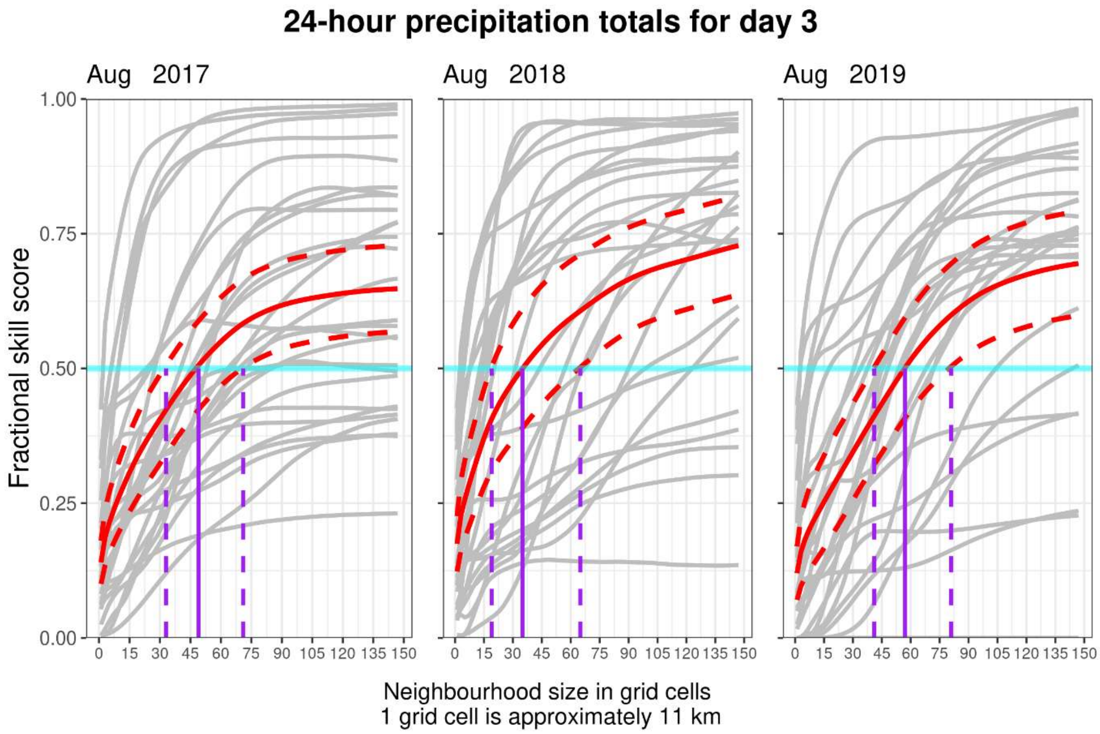


Figure 6c: Daily Fractions Skill Scores for August 2017, 2018 and 2019, using the "Take Action" 10 mm threshold for 24-hour precipitation for day 3 accumulation (GM), excluding days where less than 0.5 % domain exceeds the threshold.

When the average FSS curve (grey line) does not cross the FSSufc line (blue line) is when the forecast is considered useful. Figure 6c shows that, for August, the model has skill in capturing the precipitation between 70 % (2017) and 80 % (2018 and 2019) of the time, shown by the number of grey lines crossing the blue horizontal line. No skill is seen for 20 % (2018 and 2019) to 30 % (2017) of the time, shown by the grey lines not crossing the blue line.

Table 7 shows these statistics for all the months which were analysed. Table 7 shows that a lot of days are excluded during the dry season, as to be expected, but also there are a number of dry days within the wet season. For months with a large number of non-dry days (more than half the month) the model typically shows skill greater than 50 % and up to 96 % of the time. There is greater variation in performance for relatively dry months, again to be expected. The model shows skill during the wet season for 2018 and 2019 greater than 60 % of the time.

% days showing skill (# days included in study)	2017	2018	2019
Annual	48 (129)	65 (158)	71 (92)
Wet season (April to November)	46 (114)	69 (142)	62 (87)
January	- (0)	100 (1)	- (0)
February	100 (6)	0 (3)	- (0)
March	33 (9)	29 (7)	20 (5)
April	60 (5)	75 (20)	59 (17)
May	53 (19)	58 (19)	63 (16)
June	29 (7)	55 (11)	82 (11)

% days showing skill (# days included in study)	2017	2018	2019
July	38 (21)	96 (28)	82 (17)
August	71 (28)	81 (26)	81 (26)
September	29 (21)	0 (15)	N/A
October	22 (9)	93 (14)	N/A
November	50 (4)	56 (9)	N/A
December	- (0)	20 (5)	N/A

Table 7: The percentage within each month of how often the model shows skill in capturing the precipitation, seen by the grey lines cross the blue horizontal line, for day 3 forecasts, with dry days (less than 0.5 % of the domain exceeding the threshold) excluded. The number of days included in the analysis are shown in brackets.

The results show that, for August, displacement errors, an indication of how large a difference in location there could be between where the precipitation is forecast and where it is observed, of between 150 km and 550 km are obtained using the FSS method for day 3 accumulations. This range in values is typical for the results for the wet-season months in this analysis. During the dry season, the displacement errors increase, which is to be expected when using this method. When the FSS method was applied to the ECMWF model over the European domain the displacement error for day 5 precipitation was 375 km (Skok and Roberts, 2016). Whilst the results from this study and the one by Skok and Roberts (2016) cannot be compared directly, due to looking at different domains which have different meteorological drivers to the precipitation and due to different methods of defining a precipitation event, it allows for a comparison of expected FSS results.

For January 2017 to August 2019, both models perform well for days with light rain, i.e. 0.1 mm per day, with relatively poorer performance for heavy rain days. However, the performance for heavy rain days is similar to that of other national forecast centres' models. The performance of the Met Office GM is routinely compared against the ECMWF, NCEP, and JMA global models using four different measures of skill[5]. For the period 2017 to 2019, the Met Office's skill in precipitation for the Tropics was similar to the ECMWF and both centres' models frequently performed similarly or better than both the NCEP and JMA models, for all four measures of skill.

6.5 Epidemiological relationship between forecast rainfall and cholera

A statistical model was developed to examine the relationship between forecast rainfall and cholera incidence at a country level, the same cholera incidence data source was used as in the CRM evaluation but not at a governorate level. The model was fitted using forecast precipitation and cholera incidence data from 2017, shown in Figure 6d; this was prior to the WRA being issued to UNICEF by the Met Office. This model is then evaluated for 2018 and 2019 to determine whether the reduction in cholera incidence can be attributed to the provision of the WRA. Details about the statistical model developed are in Annex 9.

[5] https://apps.ecmwf.int/wmolcdnv/scores/surface.time_series/tp

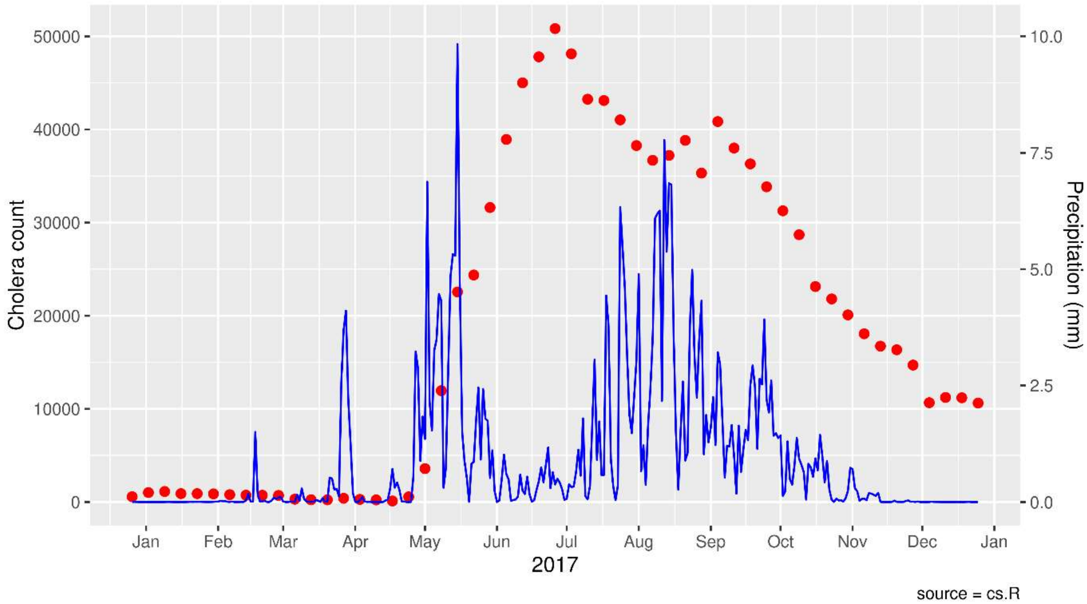


Figure 6d: Weekly counts of cholera cases (red points) and forecast daily precipitation accumulations (blue line) for Yemen in 2017, the period over which the statistical model is fitted.

Following the cholera epidemic in 2017 the number of weekly new cholera cases rose in the second half of 2018 to 15 000 cases, peaked again in March/April 2019 at over 30 000 cases and remained high with two secondary peaks at two-month intervals (Figure 6d). There was a change in the definition of cholera between 2017 and 2018, which is thought to underestimate the number of cases in 2018; this definition changed back in 2019. The precipitation-based cholera risk index was computed for each week by taking the maximum risk index value and using the forecast model dependent 20 mm and 10 mm thresholds. The threshold is represented by the risk index line C = 0.5 (Figure 6e) and the cholera risk index only exceeds this on two occasions in 2018, consistent with the previous observation that precipitation averaged over the whole of Yemen rarely reaches the threshold values used.

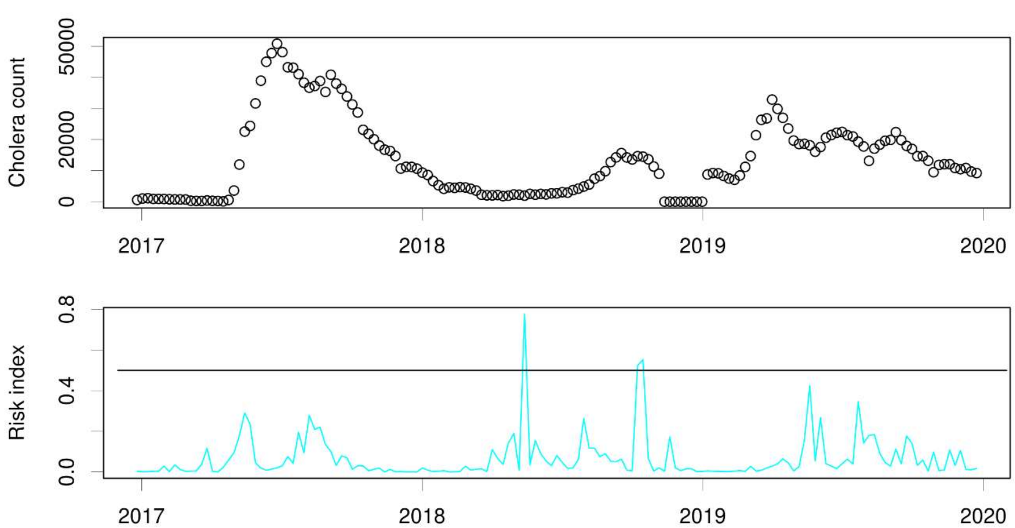


Figure 6e: Weekly count of cholera cases (top) and computed cholera risk index (bottom) using fixed precipitation thresholds of 20 mm for the CAM (days 1 and 2) and 10 mm for the GM (from day 3) precipitation forecasts, for Yemen for 2017 to 2019.

For 2017, the year before the forecasts were being used to help manage the epidemic in Yemen, the RMSE is relatively high at 20,000 (results not shown) and the regression has a Pearson correlation coefficient of 0.39 previously calculated from the nonlinear regression. For 2017, the negative bias (not shown) means that new cholera cases are underestimated by up to 10,000 (on average) when using fixed 20 mm and 10 mm precipitation thresholds.

The results show a higher RMSE and a positive bias in 2018, compared to both 2017 and 2019. For 2019, smaller biases and RMSEs are seen compared to 2017. These results could be explained by a number of different factors that could either act independently or in combination:

1. the precipitation forecasts were used effectively to prevent increases in number of cholera cases;
2. an increase in the number of cholera cases was prevented due to factors other than the precipitation forecasts, e.g. through monitoring water supply microbiology;
3. there is significant scatter in the risk index to cholera count relationship and the results are due to a random effect, natural to this type of data;
4. the annual seasonality which has not been accounted for (only one year, 2017, before the forecasts was available).

An alternative statistical approach to modelling cholera counts, using a Generalised Additive Model and where the predictive model is fitted more closely to the 2017 data, is presented in Annex 10. However, the conclusions drawn are the same.

6.6 Summary of Rainfall Forecast validation methodology and results

6.6.1 Methodology

- The Met Office rainfall forecasts were compared to satellite data, comparing daily accumulations of rainfall.
- Contingency tables (‘hit or miss’) were used to indicate if the forecast daily accumulation and observed daily accumulation was above or below certain thresholds (light rain, rain, heavy rain, storm and ‘take action’).
- Fractions Skill Score (FSS) was used to evaluate the spatial accuracy of forecasts (i.e. where in Yemen did the forecast say it would rain, and where did it rain), using the ‘take action’ threshold to define a rain event.
- A statistical model was developed to examine the relationship between rainfall and cholera incidence at a country level (the same cholera incidence source data as used in the CRM evaluation, although not at the governorate level due to time constraints). The model was fitted using precipitation and cholera incidence data from 2017, shown in Figure 6d prior to the Weekly Rainfall Assessment (WRA) being issued to UNICEF by the Met Office. This model is then evaluated for 2018 and 2019 to determine whether the reduction in cholera incidence can be attributed to the provision of the WRA.

6.6.2 Results

- The rainfall forecasts over Yemen have better spatial accuracy for days with light rain, compared to days with heavy rain. When the forecasts of light rain are compared with satellite data, the expected difference in location between where the rain is forecast and where it is observed is close to 11 km, the resolution of the GM. This is seen for all months for the years 2017 to 2019.
- For days with heavy rain, the difference in location between where it is forecast and where it is observed is at least 160 km, with the greatest skill in the 24-hour forecasts seen in July and August. These differences are smaller than those found for the ECMWF day 5 precipitation forecast for the European domain (Skok and Roberts, 2016).
- The very heavy rain observed over Yemen is typically due to small-scale, convective, processes embedded within the large-scale drivers. To be able to model these processes, forecast models require spatial resolutions of the order of these small-scale processes

(~1.5 km). Light rain is driven by the large-scale processes which can be modelled by coarser resolution models, e.g. the CAM and GM, hence is forecast more accurately.

- The precipitation is weakly correlated (a coefficient of between 0.4 and 0.5) with weekly number of new cholera cases. This is a similar result as found by Camacho et al. (2018).
- The statistical model suggests that the targeted interventions made, acting on the weekly rainfall assessments, have reduced the number of cholera cases, although more epidemiological data prior to 2017 (i.e. more before targeted interventions took place) is needed to validate this. This is based on the results from the statistical model, which showed an increase in the Root-mean square error (RMSE) in 2018 compared to 2017 and 2019, and a positive bias in 2018 compared to negative biases in 2017 and 2019.
- The performance of the Met Office Global Model (GM) in the Tropics for 2017 to 2019 is similar to the ECMWF and both centres' models frequently performed better than, or similarly to, both the NCEP (National Centers for Environmental Prediction) and JMA (Japanese Meteorological Agency) models.

6.6.3 Considerations

Inter-annual climate events (e.g. El Niño) and the anomalous environmental conditions (e.g. rainfall, temperature, etc.) they generate, have been linked to outbreaks or amplification of cholera. The influence of these events on cholera rates in Yemen could be assessed as these events are predictable (often quite far in advance) so could provide an earlier warning to UNICEF and others of what to expect for the purposes of longer-term strategic planning or OCV (Oral Cholera Vaccine) campaigns.

6.6.5 Limitations

The verification of the forecast models used in the WRA used satellite data as the observations data. The limitations of using satellite data are discussed in Annex 5, however the main limitation in this study was not being able to investigate the bias between the forecast models and the satellite data, which would affect the FSS results. The satellite data was also at the resolution of the GM; thus the CAM was not analysed at the high resolution it is run at, which would also affect the FSS and contingency table analysis results.

The statistical model that was used to examine whether the targeted interventions, acting on the weekly rainfall assessments, reduced the number of cholera cases, was developed using a very small sample of epidemiological data. It also made several assumptions as to how WASH interventions were initiated without the WRA, i.e. based on epidemiological case data alone. To improve the understanding of the relationship between the WRA and epidemiological data, more epidemiological data prior to 2017 is required.

6.6.4 Recommendations

The current "Take Action" threshold used by UNICEF is a very high precipitation threshold which all forecast models have comparatively lower skill at forecasting accurately, as shown by the analysis in this report. A recommendation would be to investigate the threshold used by UNICEF, following the performance analysis of the models used in the Weekly Rainfall Assessments. The aim would be to use a threshold where there is more confidence in the performance of the forecast models. It would also be useful to ensure UNICEF understand the limitations of the rainfall information, as identified in this report, in terms of its accuracy in forecasting where there will be rain, as this may influence the way they use the data.

Given the NWP model spatial resolution and the need by UNICEF to action information at fine, governorate scale, it may be useful to explore mature statistical model downscaling options, where the aim is to create local projections of weather for governorates, districts, and cities. Model downscaling, where the spatial resolution of the model is increased, may be appropriate if the actions taken for cholera response and wider DRR purposes require better levels of accuracy. The format of the rainfall assessment could also be changed so that they can be integrated and interact with other data sources, enhancing the ability for them to be used in cholera response and DRR purposes.

7. Summary and Recommendations

7.1 Summary of validation results

UNICEF Yemen's use of rainfall forecasts and the CRM to influence their cholera prevention activity represents a relatively novel approach to cholera control. This paper looked at the reliability of these tools and considers how UNICEF's approach to acting earlier (and the tools themselves) can inform cholera response in other areas which are affected by the disease.

After gaining an understanding (from work by Carmacho et al) that heavy rainfall led to an increase in cholera cases, UNICEF started to use rainfall forecasts from the Met Office and Cholera Risk Model information from the University of Florida in April 2018. The districts which were most likely to receive heavy rain and were predicted by the CRM to have high risk of the disease were considered against data on the districts with the highest levels of cholera. This was used to prioritise where interventions needed to be intensified. The interventions themselves, (mainly WASH awareness raising) were classed as 'no regrets'; if the surge in cholera cases did not entail, the communities would still be better informed and prepared in how they could avoid it. The CRM was used initially but at the current time, UNICEF are only using the rainfall assessments to inform their decision-making.

To assess the reliability of the CRM, its predictions in 2017, 2018 and 2019 were compared to recorded cases of cholera in Yemen. In the most populous governorates (comprising about 80% of the population), the CRM's predictions were accurate 60% of the time. Assessments of the CRM's performance in other countries (by UF) also supports this finding. In addition, using the BHC framework, analysis of sensitivity, specificity, accuracy and precision, and negative predictive value, suggests that the CRM was able to capture weekly changes in the number of cholera cases in all of the governorates. The CRM had the highest accuracy in 2017 and UF suggest this may be due to the preventative risk-based interventions being taken by UNICEF in 2018 onwards. However, more cholera surveillance data would be needed to test this. Cholera has occurred consistently in Yemen each week from 2017 to 2019 and this suggests cholera is becoming endemic.

The forecast models used in the Weekly Rainfall Assessments were analysed by comparing them to satellite observation data, due to the lack of in-situ rain gauge data. Forecasts for light rain were, typically, within 11km of where light rain was observed. The forecasts for heavy rain differed in location, to the observation data, by at least 160km. These results are comparable to forecasts from other national weather centres. Rainfall forecasts were found to be weakly correlated with the weekly number of new cholera cases. The statistical modelling work suggested that targeted interventions based on the rainfall forecasts may have reduced the number of cholera cases, however more data would be needed to validate this.

7.2 Introduction to Recommendations

The evaluation of UNICEF's response to Yemen's cholera outbreaks in 2016 and 2017 suggests that timely control measures could have limited its spread. Given the prevailing risk factors and vulnerabilities in Yemen and the ongoing cholera outbreak, it describes a system-wide lack of anticipation and preparation for a major epidemic, which left UNICEF and other actors 'chasing the epidemic'. Getting ahead of the epidemic curve, the report suggests, can make a substantial difference during an outbreak. The need for early detection and response to contain outbreaks is also recognised in Axis 1 of the GTFCC's Roadmap to Ending Cholera, which calls for early surveillance systems (prepositioning and preparedness of WASH and health systems).

UNICEF's approach to using information on cholera risk to target their interventions represents a novel approach within the field of cholera control. In the Disaster Risk Reduction (DRR) field, risk informed early action is a now well-known and sophisticated concept. It should be recognised however that its evolution has been a long journey which started with individual pilots implemented by the Red Cross, World Food Programme, Food and Agriculture Organisation and others. These tested the theory that acting ahead of a climatological hazard (which is forecast to have high impacts) could save more lives and improve the effectiveness of interventions. Events like the Red Cross' Global and Regional Dialogue Platforms were (and still are) used to raise awareness of early action within the wider DRR community, share learning from pilots and to encourage discourse and ideas that could enhance the approach. Eventually, standardised methodologies and frameworks have been developed which provide a step-by-step approach to implementing an early action plan. In parallel, technological developments and financing mechanisms have emerged which act as enablers for practitioners to act ahead of a hazard.

It can be argued that the use of forecasts and cholera risk information in Yemen represents one of the first pilots/trials of their use and that the general concept of using risk information to inform cholera response is in its infancy. The results of the validation work described in this report are encouraging and indicate that the CRM and rainfall forecasts can add value to cholera decision making where the context or use case is similar to Yemen (i.e. cholera control and prevention interventions are already underway). Whilst both the CRM and rainfall forecasts have limitations, these should not discourage their uptake as waiting for the perfect tool risks missing opportunities to reduce the impact caused by cholera.

The recommendations drawn from this assessment are presented in the following way:

- Recommended use of CRM and rainfall forecasts in Yemen
- Recommended use of the CRM and rainfall forecasts elsewhere
- The need for further pilots
- Encouraging discourse and learning lessons from pilots
- Opportunities to enhance the CRM and rainfall forecasts in the short-term (already in progress or straightforward to implement
- Opportunities to enhance the CRM and rainfall forecasts in the medium to longer term
- Recommendations on data availability to support tools such as the CRM

For all of these recommendations, a caveat could be made that there are still many questions about the way that these tools should be applied which can only really be answered through more detailed engagement with potential users. The COVID pandemic has restricted routine exploration with users in Yemen during the course of this project.

7.3 Recommended use of the CRM in Yemen

Currently UNICEF are not using the CRM in their cholera decision making and have not been doing so for over a year and a half. During brief interactions with the team, the main reason for this was scepticism about its predictions and finding it hard to interpret. In light of the validation work undertaken for this Case Study, UF should demonstrate to UNICEF and partners how use of the CRM can add value to their planning and preparedness activity and how it complements the SOPS already in place for using rainfall data. The development of training materials on the CRM which describe the genesis and evidence base for the model, summarise validation work undertaken on its outputs and make it easy for users to understand its outputs, will facilitate uptake for users in Yemen (and elsewhere). Additionally, further analysis of the CRM's performance should be undertaken with new epidemiological data sets which have been obtained from UNICEF which provide more granular case data.

7.4 Recommended use of rainfall forecasts in Yemen

The rainfall assessment product which is provided to UNICEF each week needs to be reviewed and streamlined due to only the narrative summary of weather and rainfall by district tables being used. A review of the threshold used by UNICEF to "Take action" should also be undertaken to identify a threshold which means the forecast models have greater skill, following the performance analysis of the models used in the forecasts in this study.

Due to the absence of any other forecasts in the country, the rainfall assessments are used more widely than for just cholera decision-making (e.g. flood forecasting and camp management). As the product was not originally designed with this in mind, the suitability of the rainfall assessment in its current form (and considering the validation work undertaken), is under consideration with UNICEF. The outcome of these discussions will inform the details that will be presented in the forecasts. A way to secure the long-term funding for the ongoing provision of the rainfall assessments (once their new content and format has been agreed) is also recommended until the Yemen Met Service can resume issuing forecasts itself.

Longer-range information on the rainy season ahead, in terms of its likely onset and intensity could also be of use to UNICEF for longer term planning and preparation and it is recommended that the availability of seasonal forecasts for Yemen is explored with the World Meteorological Organization (WMO).

7.5 Recommended use of the CRM elsewhere

In areas where epidemic cholera is a frequent occurrence, preventative measures will already be underway in anticipation of an outbreak (often on a seasonal basis). In these contexts, the CRM and rainfall forecasts can be used to inform planning and preparation activities. The cholera risk information provided by the CRM should also be used to intensify early control measures such as surveillance and reporting, strengthening healthcare systems and community engagement. Using cholera risk information in this way can help to flatten the epidemic curve

It should be recognised that willingness of UNICEF to use risk information to intensify interventions was born out of the scale of the epidemic Yemen was experiencing; any information which could be used to prioritise control activity (that was better than taking chances) was helpful. In other contexts where the nature of the epidemic is not as severe, there may be less willingness to use risk information to target control measures. The way the CRM and rainfall forecasts are used alongside other tools and models would therefore need careful consideration with potential users. Funding to encourage early action based on risk information could also help to overcome financial barriers to acting before an epidemic is declared.

The range of interventions that are taken for cholera treatment, control and prevention, whilst adapted to local contexts, are broadly similar around the world. It is therefore recommended that these are explored with a range of operational and strategic cholera stakeholders so that they can be used as a basis for SOPs if and when further pilots take place.

A joint OCHA/GTFCC/FCDO workshop took place in Spring 2021 to share the concept of early action for cholera with a wide range of stakeholders. One of the three break-out sessions provided an opportunity to initially explore which actions/interventions had the best potential to be informed by risk information. Even though participants came from varying backgrounds and operated in different contexts, there was a reassuring level of consistency in the actions they felt could be brought forward if cholera risk data were available to them.

7.6 Recommended use of the rainfall forecasts elsewhere

When used alongside the CRM, forecasts can provide another layer of risk information and provide a higher level of confidence to a potential user due to their dynamic nature (the forecasts are updated hourly and short range forecasts for days to a week ahead are associated with higher accuracy).

There may be localised differences in how rainfall and cholera are connected. Understanding the nature of this relationship, by area, in a country would indicate where use of rainfall forecasts has most value in cholera decision making. A weather sensitivity analysis (which can assess the relationship between rainfall and cholera, by area) could be used to establish this.

If forecasts are used, the spatial accuracy of forecasts versus the size of administrative areas for which decisions are taken would need to be considered in individual use cases, as this could represent a limitation to their application in smaller countries. Thresholds for action would need to be considered (as in Yemen).

Short-range forecast products to support cholera response should be co-produced between the user (e.g. UNICEF) and the National Meteorological Service (NMS) of the country in question. If the NMS is not functioning (as is the case in Yemen) or is unable to provide the required information, the local regional centre, through the WMO cascade system might be able to assist. Failing that any global producing centre such as the Met Office can work with the user to design a rainfall forecast service. A full understanding of needs and potential uses of these forecasts is required at this stage to ensure the forecast is suitable.

As described in Section 2, numerous environmental factors are associated with cholera including temperature, relative humidity, sea surface temperature and social risk factors (e.g. sanitation conditions) so there should be caution in using rainfall alone as a determinant of cholera. Rather, short-range rainfall forecasts should be used alongside tools such as the CRM.

Longer-range weather forecasts may also be valuable in preparing for cholera response. Rainy seasons are thought to amplify outbreaks either through contamination of surface water and open sources or through population movements related to the rains (evaluation of UNICEF cholera response 2018). Outlooks for the upcoming rainy seasons are developed at a series of Regional Climate Outlook Forums (COFs) around the world (coordinated by the WMO). Where they are functioning, the COFs also provide impact assessments though collaborative workshops with regional personnel from a range of institutes and agencies. The seasonal forecasts give a likelihood prediction of whether the month will be above or below average (precipitation). The exact timing of the onset of a rainy season may be inferred from these, but due to the probabilistic nature of seasonal forecasts, there will be a large uncertainty in these predictions. The appetite for inclusion of cholera stakeholders in these fora should be assessed. Furthermore, Moore et al. (2017) suggest that there may be a link between El Niño and the coastal environmental conditions that lead to cholera, in Africa. Further analysis would be required to determine whether the impact of El Niño, and other inter-annual climate events, is observed on cholera outbreaks elsewhere in the world, including Yemen. If a correlation is found this could provide an earlier warning for the purposes of even longer-term strategic planning.

The diagram below summarises recommendations on how the CRM and rainfall forecasts could be used now to inform cholera response. It is likely that interventions profiled would change according to practitioner and context in which they are applied, so it is recommended this approach is consulted on with a range of cholera stakeholders.

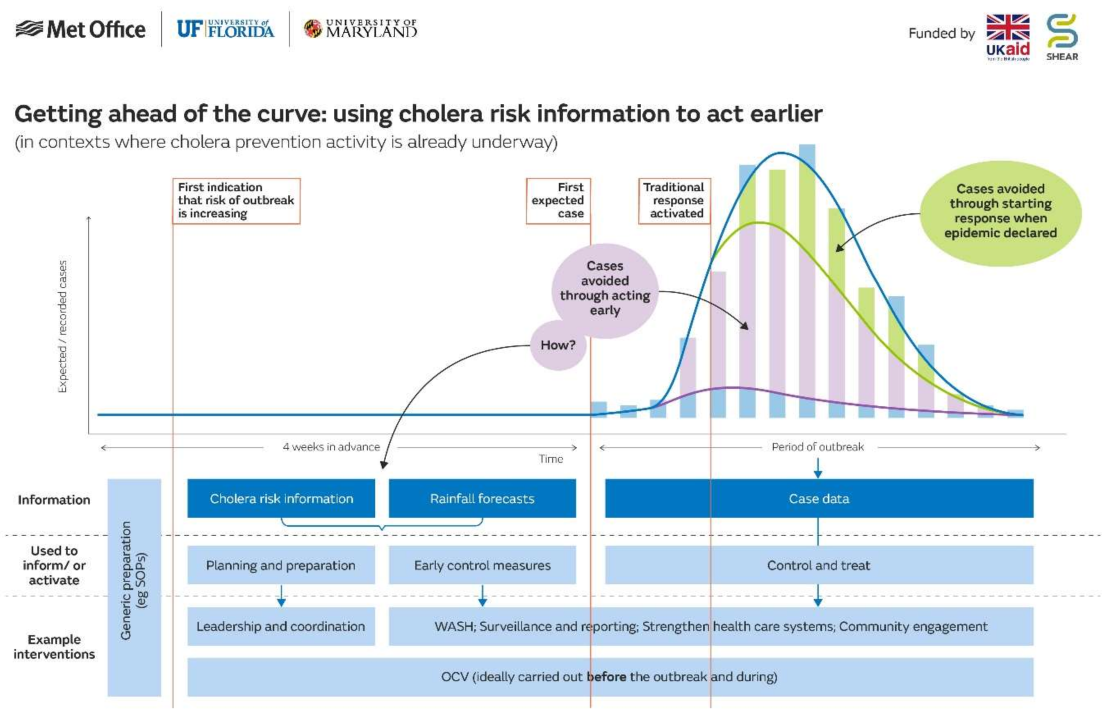


Figure 7a: Proposed use of the CRM and rainfall forecasts to inform cholera decision making and the impact this can have on the epidemic curve.

7.7 The need for further pilots using risk information to inform cholera response

For those unaccustomed to the anticipatory action approach, risk informing tools like the CRM and rainfall forecasts may be hard to value within a more generalised response framework; further pilots are recommended to inform the evidence-base for the merit of such tools.

Yemen is, so far, one of the only examples of using risk information in decision making around cholera (and this was predominantly based on using the rainfall forecasts). Whilst it provides a useful basis for discussion, more examples of use of the information in practice are required to shape this approach.

The design of these pilots could be informed by the Met Office's People Led Services Approach (PLSA). This was developed in recognition that successful early action can only be achieved through structured engagement between the providers of risk information and the users. Working together, they can co-produce a process which pulls risk information through into action. The PLSA is based on (and advances) the Met Office's Impact Based Forecasting approach and draws on best practice of using risk information from the UK and around the world. The PLSA also compliments the Red Cross Climate Centre's Trigger Methodology for Forecast Based Financing.

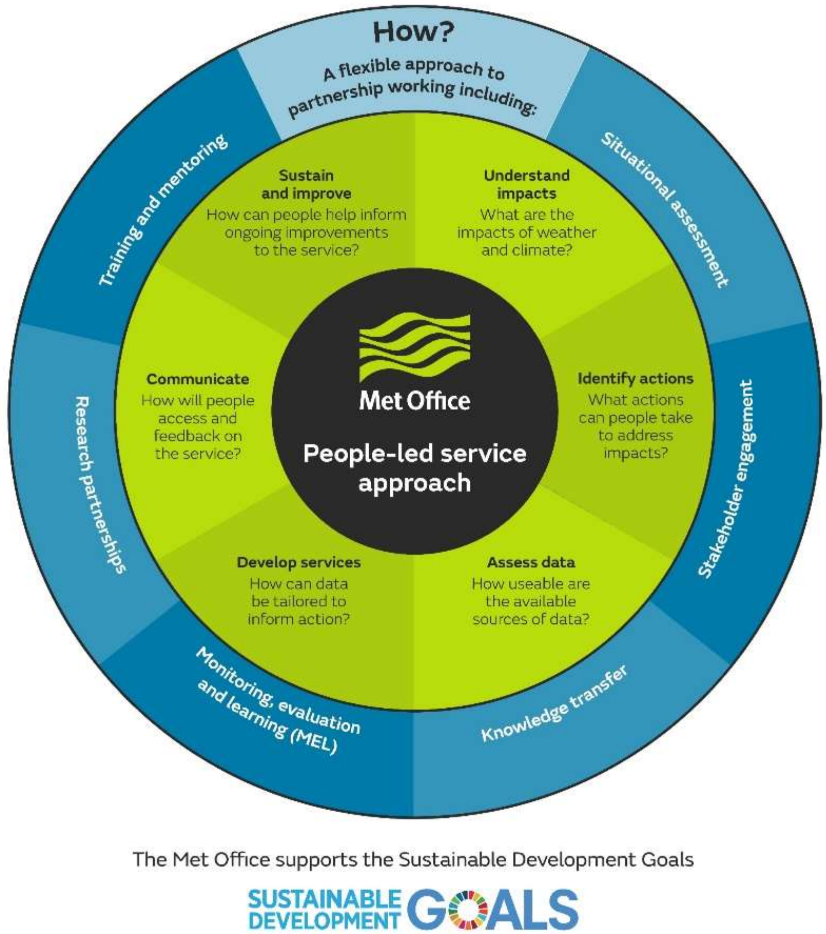


Figure 7b: The Met Office's People Led Service Approach: Ensuring science is pulled through to action.

Whilst primarily used in DRR and humanitarian contexts, we recommend the approach is also appropriate in the cholera context, since the elements involved are similar. If the approach was followed to inform the design any pilots which take place, the questions in Box 7 could be considered.

A workshop (described in 7.5) took place in April 2021. Hosted by UN OCHA, REAP, FCDO and the GTFCC, the concept and coordination of the workshop drew substantially on the experiences and lessons gained from this Yemen Early-action for Cholera pilot. The objectives of the workshop were to raise awareness of the concept of early action for cholera and to explore how the CERF fund could be used to support further testing of this. There was a general sense of openness (from the wide range of stakeholders participating) that using risk information could represent a way to reduce the cholera burden during an epidemic. It is essential however that the concept of anticipatory action in the cholera domain is not seen to obviate or undermine the need for longer term investment in WASH and health infrastructure, as described in the GTFCC's Roadmap for Ending Cholera.

OCHA have now committed up to $10 million to working with the GTFCC to supporting pilots which will test how using risk information can support earlier response. These pilots could provide an opportunity to take forward some of the recommendations presented in this case study.

1. **Understand impacts:** Status of cholera in the country; seasonality of cholera or monthly distribution; relationship between climatic parameters and cholera bacteria in human population.
2. **Identify actions in anticipatory space:** How is cholera response organised and coordinated; what are the cholera interventions that are generally taken, by whom; how are they triggered; what is the lead time between deciding to take action and triggering action; what is the planning and preparation process for the interventions; how long does this take; what is the cost of interventions and how are they funded; what are barriers to interventions; what are the local considerations (cultural/social/political/economic); which interventions are considered "no regrets"; what is the impact of 'acting in vain'.
3. **Assess data:** Availability of epi data/WASH data/population data; how is this currently used to inform interventions; how are interventions paid for; are there cholera tools/models being used already/what evaluations have been undertaken on cholera response/what lessons have been learned; does the CRM cover the country in question; what rainfall forecast data is available from national meteorology service – and what is their capacity for delivering tailored rainfall products; are there known barriers up uptake of information.
4. **Develop service**: Identify where risk information could inform and enhance cholera decision making; develop SOPs; explore what the financing considerations around earlier response/action; agree how the effectiveness of acting early can be measured – what is the impact on the curve of the epidemic; what are the consequences of acting earlier and (potentially) flattering the curve with regards to fatigue/scepticism. Ways to evaluate risk based interventions and the performance of the tools used to inform these should be directly informed by the availability of data identified in step 3.
5. **Communicate:** Who will risk information be used by; what format should information be provided in (email, web portal etc); ensure training materials and user guides are available; develop service level agreements which include feedback mechanisms
6. **Sustain and improve**: Evaluate and adapt approach based on learnings; knowledge management. The target being to avoid cholera becoming endemic in that particular region.

Box 7: Questions that could inform the structure of the pilot design.

7.8 Encouraging discourse and lesson sharing from pilots

If pilots which test the use of risk information to inform cholera go ahead, it is recommended that opportunities to share insight and experience from these are used to build awareness and momentum around the approach (following the trajectory of early action in the DRR context). There is known to be scepticism in the cholera community about the use of environmental data to inform cholera risk and this could come to represent a significant barrier to uptake. Discourse could be used to explore this and evidence from pilots may go some way in helping to overcome this.

Opportunities for such discourse could include:

Anticipatory Action Fora

- **Risk Informed Early Action Partnership (REAP):** REAP aims to encourage a systemic shift towards acting early. Whilst primarily focused on earlier emergency response, REAP has now formed a working group on health and cholera can be included in this.

- **Dialogue Platforms (Global and Regional):** Set up by the Red Cross to share best practice of pilots for forecast based early action for humanitarian response, these are now widely attended and consider a range of issues around anticipatory action. There may also be scope to include a health session in the Red Cross' Dialogue Platforms to consider early action for diseases or health conditions which have a predictable environmental element (e.g. Dengue fever, cholera, heat stress and others).
- **Understanding Risk** is a global community of experts and practitioners with interests in the field of disaster risk identification, specifically risk assessments and risk communication. Annual events are held to bring the community together.

Cholera/health specific fora:

- **A working group on the use of climate and weather information for predicting and preparing for cholera and vector-borne diseases** has already been established by the WMO, WHO and others to discuss scientific developments in this area.
- **Joint Operational Framework (JOF) set up between WASH and Health Clusters of the WHO to improve the coordinated and integrated preparedness and response to cholera in countries in humanitarian crisis.** The JOF was informed by findings from a joint global health and global WASH review and was created in consultation with partners working on the cholera response across different humanitarian contexts. It promotes a set of key tasks in the critical areas of leadership, coordination and integrated response to increase the efficiency and effectiveness of the cholera preparedness and response efforts.
- **GTFCC related fora.** The pilots will be co-designed between OCHA and the GTFCC and there will therefore be opportunities to share insight from these within the Task Force. If the pilots can demonstrate the value of using risk-data there may also be scope to explore how its use could also inform the implementation of the Roadmap for Ending Cholera. It may be particularly relevant to the 'Early Detection and Surveillance' pillar of the strategy.

7.9 Opportunities to enhance the CRM and rainfall forecasts in the short-term (already in progress or straightforward to implement)

Further pilots and associated discourse/learnings will help to develop and enhance the *approach/methodology* to early action in cholera response. There are also specific recommendations which could enhance the CRM and rainfall assessment tools themselves. Recommendations are categorised as either scientific or technical development where discreet activity can enhance the tools or 'research' where further insight or understanding is required.

The short/medium development activity recommendations include:

7.9.1 Explore if forecast rainfall information can be used to introduce 1 week lead-time into the CRM (scientific development)

The operational CRM uses historic rainfall data sets from the NASA GPM and is valid 4 weeks from issue. The potential enhanced predictability of the CRM would gain by ingesting forecast data needs further investigation by UF, but it is proposed that this could give the CRM a 1 week lead-time into its 4 week validity period.

Relevance of recommendation in Yemen: Currently UNICEF are not using the CRM and have not been doing so for over a year and a half. The reasons for this were described as finding it hard to understand and being sceptical about its predictions. It is hoped that by sharing this validation work with UNICEF Yemen, their confidence in the CRM and appetite to use it will increase so the extended lead time (if this is achieved through forecast data ingestion) could be helpful.

Relevance of recommendation elsewhere: Longer lead times are likely to be more helpful / relevant to potential users of the CRM.

7.9.2. Improving access to the CRM (technical development)

Output from the CRM is currently displayed in a PDF document which is sent as an attachment to an email at the start of each week. Accessing the CRM through a web-portal (which is in development) will improve access to the tool and enable users to interact with its outputs more flexibly. The web-based tool could also act as a dashboard which enables decision makers to see an overview of cholera risk and a regional/global level to identify potential hotspots. Development of a smartphone app (only for Android devices at present) to display cholera risk assessments is also underway.

Relevance of recommendation in Yemen: If UNICEF start to use the CRM again, then accessing it online maybe useful. Confirmation of this would need to be sought since there are likely to be issues with the internet in Yemen, which may make it preferable to retain capability to provide a PDF which can be downloaded, printed, and discussed.

Relevance of recommendation elsewhere: Providing a range of options for how the CRM is accessed (email, web, app) will help to tailor the service to the needs of individual users.

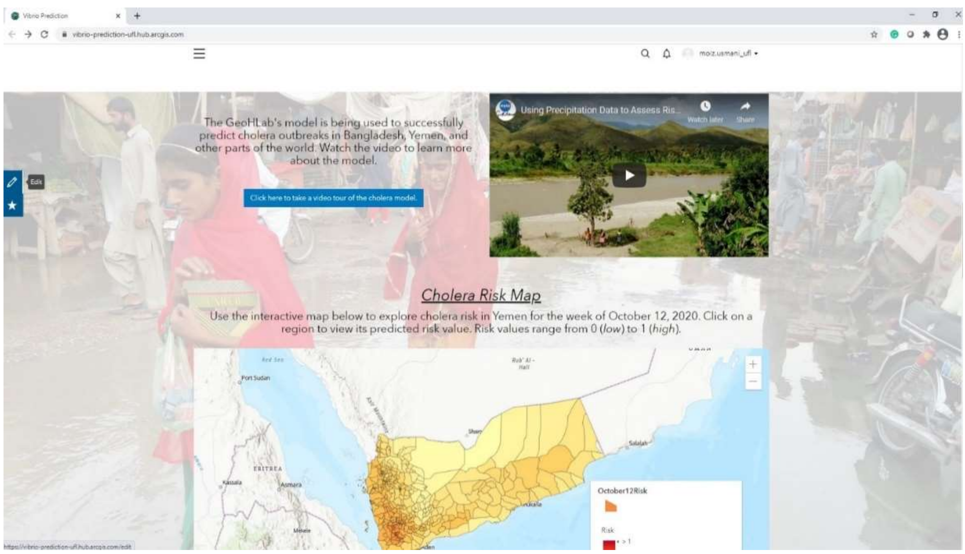


Figure 7b: CRM output on web-based portal (showing Yemen as an example)

7.9.3 Inclusion of the CRM in inter-comparison study of cholera models (research)

Identification of how the CRM compares with other cholera models will help practitioners to understand which tools are most appropriate for them in any given context. It is therefore recommended that the CRM is included in the cohort of models to be evaluated under the newly formed WHO Infections Disease (ID) modelling inter-comparison pilot study. This pilot study will independently evaluate multiple ID models using common, and agreed, metrics, including alternative cholera models, such as those of Medécins Sans Frontiers. The pilot emerged as an outcome of the WHO Expert Meeting on Infectious Disease Modelling in June 2020 and the terms of reference are being agreed at the time of writing.

Relevance of recommendation in Yemen: Understanding how the CRM compares with other tools will enable UNICEF to identify how it can add value to the tools that are already (or could be) available to them.

Relevance of recommendation elsewhere: Potential users would be able to identify the most appropriate tool to use to strengthen their decision making around cholera response.

7.9.4 Development of communication materials

Insight from users suggest that output from the CRM can be hard to interpret and thereby represents a barrier to its use. This finding supports the ongoing development of user guides and training video(s) which will inform decision makers of the properties of the CRM and rainfall forecasts and how to interpret them. The limitations of tools will also be described. The communications materials would also set out an evidence-based justification for the use of environmental factors in modelling cholera (described in section 2.2).

The way the CRM is used in individual contexts and by different types of stakeholders needs to be considered for each use case. Decision support tools (like the example shown in Annex 11) could be used to structure this dialogue.

Relevance of recommendation in Yemen: UNICEF have suggested that understanding of the output of the CRM is hard to understand. Team members have changed since the service started being provided and handovers may not have taken place on the CRM which could explain why it has fallen out of use. Training materials and user guides which can be easily understood by new users may therefore encourage its uptake if confidence in the CRM is increased.

Relevance of recommendation elsewhere: Training materials and user guides (which can be easily translated into different languages) would enable any new potential users to quickly learn about the limitations, benefits and how to use the tool.

7.10 Enhancements to the CRM and rainfall forecasts in the medium to longer-term

Whilst the CRM and rainfall forecasts are operationally ready in their current form and will be enhanced through the short- term recommendations described above, medium to longer term enhancements have the potential to improve confidence in the tools and enhance their usability.

7.10.1 Undertake further analysis of CRM performance in Yemen using newly acquired data sets (research recommendation)

More recently (after the analysis documented in this study had already been undertaken), UF was provided with data on water and sanitation, as well as epidemiological datasets on cholera cases at a district scale. This data should be interpreted and analysed for determining scale issues (i.e., what is the appropriate spatial and temporal scale at which CRM predictions are most valid?). A baseline study, similar to the one here, should then be designed to better understand the accuracy of the CRM at more local levels.

Relevance of recommendation elsewhere: Undertaking further analysis in Yemen using richer spatial epi data would determine at what scale the CRM should best be used at.

7.10.2 Validating the CRM through water sampling (scientific development)

The ability to test water for the *Vibrio Cholerae* bacteria in areas where the CRM has predicted cholera could provide a ‘ground truthing’ of the model’s predictions and increase confidence in the tool’s accuracy. A simple and cheap methodology for doing this in the field has been developed by University of Maryland (Dr Colwell) and is described in Annex 12. The idea of water sampling is important because it will provide further justification for anticipatory action and decision-making abilities for stakeholders. Development of a decision support system will be needed to assess susceptibility to outbreaks of cholera and to enable appropriate action once monitoring confirms the pathogenic bacteria is in a water system.

Undertaking water sampling in areas where the CRM predicts a risk could enable alerts to be issued as soon as cholera bacteria is detected in the water. This enhanced surveillance differs from traditional surveillance methods in which alerts are issued when cholera is confirmed in the human population (and at which point, avoidable transmission may have already occurred).

Research has been undertaken into doing this in Lake Tanganyika in East Africa with encouraging initial results (yet to be published so not referenced, Colwell et al, 2021). It is recommended that a further pilot (described in section 7.3.5) is used to test this in an operational context.

Relevance of recommendation in Yemen: If UNICEF start to use the CRM again, undertaking water sampling could provide them with higher confidence cholera risk information which could strengthen their current cholera response strategies.

Relevance of recommendation elsewhere: Confirmation of the presence of *Vibrio Cholerae* bacteria would help to provide greater confidence in cholera risk predictions and limit the chance of 'acting in vain' (where cholera risk is predicted to be high but actual cases don't occur). This maybe particularly relevant for users who are less willing to take 'no regrets' actions on the basis of predicted risk and require more certainty before planning/preparation activities and early response interventions are activated.

7.10.3 Development of the Transmission Mode of the CRM (scientific development)

It is suggested that continual cases of cholera seen in Yemen from 2017 onwards is an indication that the disease is becoming endemic in the country. This indicates that cholera will remain circulating in the human population in Yemen for some time. An example can be drawn from Haiti where a cholera outbreak lasted for about 10 years. The Transmission mode of the CRM (which is more relevant for epidemic cholera) is currently in development at UF and Yemen provides a unique opportunity to validate this, using the three years of datasets that have been obtained.

Relevance of recommendation in Yemen: Whilst the trigger mode of the CRM has been provided to UNCIEF in Yemen through this work, the transmission mode of the model may be more applicable for decision making in future.

Relevance of recommendation elsewhere: Development of the trigger mode and validation of its accuracy would enable the CRM to be used in countries with endemic cholera. Once the model has been validated, it can be used in conjunction with the trigger module to provide assessment on "given trigger risk, X, the corresponding transmission risk is likely to Y."

7.10.4 Validation of the CRM in other areas and contexts (research)

While the association of precipitation and temperature on cholera is established, thresholds on the degree of anomalous values of precipitation and temperature are expected to be a function of available water and sanitation infrastructure, as well as sociological perception of cholera disease in human population. It is also expected that some regions may be more resilient to an outbreak of cholera (due to existing behavioural norms - washing hands, boiling water etc). Therefore, it is suggested that the CRM should be validated in several regions where continuous time series of epidemiological data are available. This will shed further insights on region-wide impacts of temperature and precipitation on cholera. As an example, UF and UN OCHA have attempted validation of the trigger module on sporadic epidemiological data, harvested from WHO reports. .

Relevance of recommendation in Yemen: Continued analysis of the CRM's performance in Yemen will indicate if it continues to be a tool which can add value to decision-making.

Relevance of recommendation elsewhere: Validation of the CRM in a range of countries affected by cholera will indicate where the model's performance has the best potential to inform cholera response.

7.10.5 Improved understanding of the influence of climate on cholera (research)

Improving understanding of the impact that inter-annual climate events such as El Niño and other seasonal patterns may have on cholera could enable even earlier indications of where cholera may be an issue.

Relevance of recommendation in Yemen: Paz (2019) has suggested that whilst the cause of the Yemeni cholera outbreak in 2017 is unclear, a combination of the impact of the strong El Niño of 2015-16 on cholera incidence in Somalia, followed by south-western winds over the Gulf of Aden throughout the summer of 2016, contributed to the disease spreading from the Horn of Africa to Yemen. If further research finds a significant correlation between inter-annual events like El Niño and cholera, early predictions of El Niño would be relevant to UNICEF and could be used for longer-term decision planning and resource allocation.

Relevance of recommendation elsewhere: Improved understanding of inter-annual events on cholera could inform whether predictions for such events would be relevant for cholera responders.

7.11 Recommendations on data availability to support tools like the CRM

7.11.1 A repository of cholera time series.

Development of a dedicated repository of cholera time series in various parts of the globe is needed to understand patterns of emergence of this disease. Currently, WHO hosts a database of annual time series for various countries. However, such data are inadequate to provide sufficient confidence in the model and associated performance. For example, an experiment was conducted using very limited reconstructed time series from WHO database of cholera cases in Zimbabwe. UF used this time series to explore CRM risk predictions but the outputs remained inconclusive due to a lack of reliable/consistent epidemiological datasets. The development of data mining tools and a GIS based repository for cholera datasets could enhance future validation work on the CRM (and similar tools).

7.11.2 Assessment of intervention and mitigation activities

Assessing the impact of intervention activities which are informed by risk based information on cholera was particularly challenging during this study. The datasets on intervention and mitigation activities were not collected or routinely shared. Therefore, it is recommended that prototype reporting of ground intervention activities must be initiated once cholera risk is predicted to be high or if an outbreak is already occurring in communities. Data should be ideally be available in a GIS format and available at near real time in order to enable good analysis of the impact of acting earlier based on risk information. This type of analysis could also inform the evaluation of OCHA's pilots.

7.12 Consultation on the Case Study Recommendations

The recommendations described above regarding how the CRM and rainfall assessments could be used, are based on very limited insight from practitioners due to COVID and related factors. These will need to be further reviewed by practitioners in the cholera response community. It is therefore recommended that a consultation on the recommendations takes place with a representative group of operational and strategic practitioners.

7.13 Conclusions

Cholera represents an ongoing threat in areas where access to clean water and sanitation is poor (either due to poverty or natural or anthropogenic disasters) and environmental conditions favour growth and spread of the *Vibrio Cholerae* bacteria.

The use of the CRM and rainfall assessments in Yemen helped UNICEF to target their cholera prevention interventions and a reduction in cholera cases was observed. Whilst the drop in cases cannot be directly attributable to decisions that were informed by these tools, UNICEF's approach represents a novel example of responding earlier to fight cholera than may otherwise have been possible.

Validation of these tools indicates that in areas where epidemic cholera is a frequent occurrence, and preventative measures are already underway, the CRM and rainfall forecasts can be used to inform planning and preparation activities. The cholera risk information provided by the CRM can also be used to intensify early control measures such as surveillance and reporting, strengthening healthcare systems and community engagement.

To assess the use of risk-based tools and the impact that their use has on cholera levels, further pilots which test their use are required. Raising awareness of these pilots, sharing learnings, and drawing on the wider field of anticipatory action could enable such risk-based approaches to develop and add value to cholera interventions. If it can be applied at scale it could become a valuable tool to realise elements of the GTFCC's roadmap to end cholera by 2030 by "containing outbreaks—wherever they may occur—through early detection and rapid response".

7.13 Summary of recommendations

The recommendations presented in this paper are summarised below. Those shaded in green are already underway though the EACH project (ending in May 2021).

Recommendation	Description	Objective
Recommended continued use of the CRM and rainfall forecasts in Yemen		
Overcome barriers to uptake of CRM with UNICEF (see 7.3)	Share validation work and provide UNICEF with user guide and training materials on the CRM (in progress through EACH project)	Enable team to understand why and how to use the CRM.
Enhance use of rainfall forecasts (see 7.4)	Share validation work with UNICEF and explore thresholds at which they take action. Explore how current weekly forecast can be reformatted to meet needs for cholera response and DRR purposes. Identify how ongoing operational costs of forecasts can be met when EACH has finished in May 2021 (in progress through EACH project).	Update weekly forecast so it can support cholera and DRR decision making and can continue once EACH has closed.
Recommended continued use of the CRM and rainfall forecasts elsewhere		
Use of the CRM (see 7.5)	Use the model to plan and prepare for response.	Support planning and preparation and earlier control measures.
Use of rainfall forecasts (see 7.6)	Co-design of a forecast should take place between national meteorological	Use rainfall forecasts (in areas where there is a significant relationship between rainfall and

Recommendation	Description	Objective
	service and the cholera practitioner.	cholera (alongside the CRM) to support earlier planning and control.
The need for further pilots (see 7.7)		
Run pilots which test use of CRM and rainfall forecasts in other contexts (see 7.7)	Support further trials of the operational use of CRM and rainfall forecasts. Build strong evaluation metrics to assess value of acting/responding early into the design of these pilots.	Test /refine the approach for use of these tools in cholera decision making.
Encouraging discourse and learning lessons from pilots (see 7.8)		
Use relevant fora to further discourse on early action in cholera (see 7.8)	Use Anticipatory Action and health fora to share insight and learnings from the pilots.	Raise awareness of the use of risk-based tools to support early action for cholera. Encourage discourse, share best practice and build momentum for the concept.
Opportunities to enhance the CRM and rainfall forecasts in the short-term (already in progress or straightforward to implement) (see 7.9)		
Integrate rainfall forecast data into the CRM (see 7.9.1)	Ingesting rainfall forecast data sets into the CRM has the potential to introduce a lead-time of 1 week to when the model is valid, which would still be a 4-week period, i.e. the model would be valid for a period of 4 weeks, from 1 week after it is issued.	Introducing a lead-time of 1 week to the CRM's validity period would give users more time to plan/prepare their cholera response.
Development of web-based hub for cholera (see 7.9.2)	Development of a web-based platform and an app will mean users can access CRM data at any time, instead of waiting for a weekly email, containing the information in a PDF (underway: web-EACH, app-NASA).	Make it easier to access cholera risk information from the CRM.
Inclusion of CRM in an inter-comparison study of cholera models (see 7.9.3)	A WHO infectious disease modelling inter-comparison study is being set up to look at available cholera models and	Help users to understand where the CRM complements and adds value to

Recommendation	Description	Objective
	the CRM will be included in this. It is recommended that the needs of cholera response practitioners help to inform and shape this study.	other cholera models available to them.
Development of communications materials (see 7.9.4)	User guides and training materials will be developed which explain the evolution of the CRM and how to interpret and apply its output (underway EACH).	Enable any user of the CRM to be able to understand how to use it without face-to-face training or handovers.
Opportunities to enhance the CRM and rainfall forecasts in the medium-longer term (see 7.10)		
Further analysis of the CRM in Yemen (see 7.10.1)	Using more granular data that has recently been provided by UNCIEF, reanalyse the CRM's accuracy.	Improve understanding of the scale the CRM should be used at.
Validation of the CRM through performing water-sampling for cholera bacteria (see 7.10.2)	A simple and cheap method of testing water for the *Vibrio Cholerae* bacteria in the water of areas predicted to be high risk has been developed. It is recommended that this is undertaken when the CRM is used to inform cholera decision making.	Improve confidence in the CRM's predictions.
Development of the Transmission mode of the CRM (see 7.10.3)	Understand more about anomalous values of rainfall and temperature and available water and sanitation infrastructure and social perception of cholera to inform development of trigger mode of CRM. Validate the CRM in regions where this data (and continual times-series of cholera data) is available.	Providing a cholera-risk tool for areas with endemic cholera which provides information on how the disease will spread in the population.
Further validation work (see 7.10.4)	Validation in several regions where there is continuous time series of epi data.	Improve understanding of the impacts of temperature and precipitation on cholera.

Recommendation	Description	Objective
Improving understanding of climate events and cholera (see 7.10.5)	Look at inter-annual climate events such as El Niño and other seasonal patterns may have on cholera.	Understand if there is scope to provide longer-range (low confidence) cholera risk information.
Recommendations on data availability to support tools such as the CRM (see 7.11)		
Development of a repository of cholera time-series data (see 7.11.1)	Access to epi data is challenging but is crucial to understanding cholera. A dedicated repository would enhance access to data.	Improving access to epi data on cholera as this will enhance the development of tools like the CRM to enhance the ability and to validate their accuracy.
Assessment of cholera intervention and mitigation activities (see 7.11.2)	Datasets on cholera intervention activities are not routinely collected or shared so assessing the impact of these on cholera rates is challenging and makes the case for taking risk-informed action hard to establish.	Enabling assessment of the impact of risk-based interventions vs non risk based interventions to build the evidence base for early action for cholera.

References

Akanda, A. S., Griffiths, J. K., Colwell, R., Islam, S., & Jutla, A. S. (2011). Warming Oceans, Phytoplankton, and River Discharge: Implications for Cholera Outbreaks. *The American Journal of Tropical Medicine and Hygiene*, *85*(2), 303–308. https://doi.org/10.4269/ajtmh.2011.11-0181

Alam, M., Hasan, N. A., Sadique, A., Bhuiyan, N. A., Ahmed, K. U., Nusrin, S., Nair, G. B., Siddique, A. K., Sack, R. B., Sack, D. A., Huq, A., & Colwell, R. R. (2006). Seasonal Cholera Caused by Vibrio cholerae Serogroups O1 and O139 in the Coastal Aquatic Environment of Bangladesh. *Applied and Environmental Microbiology*, *72*(6), 4096–4104. https://doi.org/10.1128/AEM.00066-06

Ali, M., Nelson, A. R., Lopez, A. L., & Sack, D. A. (2015). Updated Global Burden of Cholera in Endemic Countries. *PLOS Neglected Tropical Diseases*, *9*(6), e0003832. https://doi.org/10.1371/journal.pntd.0003832

Andersson, N., & Mitchell, S. (2006). Epidemiological geomatics in evaluation of mine risk education in Afghanistan: Introducing population weighted raster maps. *International Journal of Health Geographics*, *5*, 1. https://doi.org/10.1186/1476-072X-5-1

Aziz, S. Pakhtigian, E, Akanda, A. Antarpreet. J, Huq. Anwar, Alam. M, Ashan.G and Colwell.R. Does improved risk information increase the value of cholera prevention? An analysis of stated vaccine demand in slum areas of urban Bangladesh (2021) *Journal of Social Science and Medicine*

Baker-Austin C, Trinanes J, Gonzalez-Escalona N, Martinez-Urtaza J. Non-Cholera Vibrios: The Microbial Barometer of Climate Change. Trends Microbiol. 2017 Jan;25(1):76-84. doi: 10.1016/j.tim.2016.09.008. Epub 2016 Nov 12. PMID: 27843109.

Bik, E. M., Bunschoten, A. E., Gouw, R. D., & Mooi, F. R. (1995). Genesis of the novel epidemic Vibrio cholerae O139 strain: Evidence for horizontal transfer of genes involved in polysaccharide synthesis. *The EMBO Journal*, *14*(2), 209–216. https://doi.org/10.1002/j.1460-2075.1995.tb06993.x

Blokesch, M., Bertuzzo, E., Murray, M., Casagrandi, R., Gatto, M., Mari, L., Rodriguez-Iturbe, I., Vesenbeckh, S. M., Rinaldo, A., & Righetto, L. (2012). Reassessment of the 2010-2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proceedings of the National Academy of Sciences*, *109*(17), 6602–6607. https://doi.org/10.1073/pnas.1203333109

Bookhagen, Bodo. n.d. “TRMM Tropical Rainfall Measurement Mission.” High Resolution Spatiotemporal Distribution of Rainfall Seasonality and Extreme Events Based on a 12-Year TRMM Time Series, in Review. Accessed June 3, 2020. http://www.geog.ucsb.edu/~bodo/TRMM/.

Bougoudogo, F. (1998). [Cholera in Africa and in Latin America]. *Bulletin De La Societe De Pathologie Exotique (1990)*, *91*(5 Pt 1-2), 406.

Brumfield, K.D., Hasan, N.A., Leddy, M.B., Cotruvo, J., Rashed, S.M., Colwell, R.R., Huq, A., 2020. A Comparative Analysis of Drinking Water Employing Metagenomics. PLoS One.

Cabot Venton, C.; Shitarek, T.; Coulter, L.; Dooley, O. The Economics of Early Response and Resilience: Lessons from Ethiopia. (2013) 46 pp

Camacho, Anton, Malika Bouhenia, Reema Alyusfi, Abdulhakeem Alkohlani, Munna Abdulla Mohammed Naji, Xavier de Radiguès, Abdinasir M. Abubakar, et al. 2018. “Cholera Epidemic in Yemen, 2016–18: An Analysis of Surveillance Data.” *The Lancet Global Health* 6 (6): e680–90. https://doi.org/10.1016/S2214-109X(18)30230-4.

Camacho, Anton, Malika Bouhenia, Andrew S. Azman, Marc Poncin, Nevio Zagaria, and Francisco J. Luquero. 2018. “Cholera Epidemic in Yemen – Author’s Reply.” *The Lancet Global Health*. Elsevier Ltd. https://doi.org/10.1016/S2214-109X(18)30395-4.

Cash, R. A., Music, S. I., Libonati, J. P., Snyder, M. J., Wenzel, R. P., & Hornick, R. B. (1974). Response of Man to Infection with Vibrio cholerae. I. Clinical, Serologic, and Bacteriologic Responses to a Known Inoculum. *Journal of Infectious Diseases*, *129*(1), 45–52. https://doi.org/10.1093/infdis/129.1.45

CDC, 2020. Guidance on the Inactivation or Removal of Select Agents and Toxins for Future Use [WWW Document].

Clemens, J. D., Nair, G. B., Ahmed, T., Qadri, F., & Holmgren, J. (2017). Cholera. *The Lancet*, *390*(10101), 1539–1549. https://doi.org/10.1016/S0140-6736(17)30559-7

Codeco, C. T., Lele, S., Pascual, M., Bouma, M., & Ko, A. I. (2008). A stochastic model for ecological systems with strong nonlinear response to environmental drivers: Application to two water-borne diseases. *Journal of The Royal Society Interface*, *5*(19), 247–252. https://doi.org/10.1098/rsif.2007.1135

Colwell, R. R. (1996). Global Climate and Infectious Disease: The Cholera Paradigm. *Science*, *274*(5295), 2025–2031. https://doi.org/10.1126/science.274.5295.2025

Colwell, R. R., Huq, A., Islam, M. S., Aziz, K. M. A., Yunus, M., Khan, N. H., Mahmud, A., Sack, R. B., Nair, G. B., Chakraborty, J., Sack, D. A., & Russek-Cohen, E. (2003). Reduction of cholera in Bangladeshi villages by simple filtration. *Proceedings of the National Academy of Sciences*, *100*(3), 1051–1055. https://doi.org/10.1073/pnas.0237386100

Colwell, Rita R. (2000). Viable but nonculturable bacteria: A survival strategy. *Journal of Infection and Chemotherapy*, *6*(2), 121–125. https://doi.org/10.1007/PL00012151

Colwell, Rita R., & Huq, A. (1994). Environmental Reservoir of Vibrio cholerae The Causative Agent of Cholera. *Annals of the New York Academy of Sciences*, *740*(1 Disease in Ev), 44–54. https://doi.org/10.1111/j.1749-6632.1994.tb19852.x

Colwell, R.R., & Huq, A. (1998). Global microbial ecology: Biogeography and diversity of Vibrios as a model. *Journal of Applied Microbiology*, *85*(S1), 134S-137S. https://doi.org/10.1111/j.1365-2672.1998.tb05292.x

Connelly, S., Fanelli, B., Hasan, N.A., Kaleko, M., Colwell, R.R., 2019. Oral Metallo-Beta-Lactamase Protects the Gut Microbiome From Carbapenem-Mediated Damage and Reduces Propagation of Antibiotic Resistance in Pigs. Front. Microbiol. 10, 1–12. https://doi.org/10.3389/fmicb.2019.00101

Conner, J. G., Teschler, J. K., Jones, C. J., & Yildiz, F. H. (2016). Staying Alive: Vibrio cholerae’s Cycle of Environmental Survival, Transmission, and Dissemination. *Microbiology Spectrum*, *4*(2). https://doi.org/10.1128/microbiolspec.VMBF-0015-2015

Constantin de Magny, G., Murtugudde, R., Sapiano, M. R. P., Nizam, A., Brown, C. W., Busalacchi, A. J., Yunus, M., Nair, G. B., Gil, A. I., Lanata, C. F., Calkins, J., Manna, B., Rajendran, K., Bhattacharya, M. K., Huq, A., Sack, R. B., & Colwell, R. R. (2008). Environmental signatures associated with cholera epidemics. *Proceedings of the National Academy of Sciences*, *105*(46), 17676–17681. https://doi.org/10.1073/pnas.0809654105

de Magny, G. C., Mozumder, P. K., Grim, C. J., Hasan, N. A., Naser, M. N., Alam, M., Sack, R. B., Huq, A., & Colwell, R. R. (2011). Role of Zooplankton Diversity in Vibrio cholerae Population

Dynamics and in the Incidence of Cholera in the Bangladesh Sundarbans. *Applied and Environmental Microbiology*, *77*(17), 6125–6132. https://doi.org/10.1128/AEM.01472-10

Dobson, Jerome E., Edward A. Bright, Phillip R. Coleman, Richard C. Durfee, and Brian A. Worley. 2000. "LandScan: A Global Population Database for Estimating Populations at Risk." *Photogrammetric Engineering and Remote Sensing*.

Eisenberg, M.C., Kujbida, G., Tuite, A.R., Fishman, D.N., Tien, J.H (2013)
Examining rainfall and cholera dynamics in Haiti using statistical and dynamic modeling approaches Epidemics

Eppinger, M., Pearson, T., Koenig, S. S. K., Pearson, O., Hicks, N., Agrawal, S., Sanjar, F., Galens, K., Daugherty, S., Crabtree, J., Hendriksen, R. S., Price, L. B., Upadhyay, B. P., Shakya, G., Fraser, C. M., Ravel, J., & Keim, P. S. (2014). Genomic Epidemiology of the Haitian Cholera Outbreak: A Single Introduction Followed by Rapid, Extensive, and Continued Spread Characterized the Onset of the Epidemic. *MBio*, *5*(6), e01721-14. https://doi.org/10.1128/mBio.01721-14

Farquharson, F. A.K., D. T. Plinston, and J. V. Sutcliffe. 1996. "Précipitation et Écoulement Au Yémen." *Hydrological Sciences Journal* 41 (5): 797–811. https://doi.org/10.1080/02626669609491546.

Fedak, K. M., Bernal, A., Capshaw, Z. A., & Gross, S. (2015). Applying the Bradford Hill criteria in the 21st century: How data integration has changed causal inference in molecular epidemiology. *Emerging Themes in Epidemiology*, *12*(1), 14. https://doi.org/10.1186/s12982-015-0037-4

Federspiel, F., & Ali, M. (2018). The cholera outbreak in Yemen: Lessons learned and way forward. *BMC Public Health*, *18*(1), 1338. https://doi.org/10.1186/s12889-018-6227-6

Frerichs, R. R., Keim, P. S., Barrais, R., & Piarroux, R. (2012). Nepalese origin of cholera epidemic in Haiti. *Clinical Microbiology and Infection*, *18*(6), E158–E163. https://doi.org/10.1111/j.1469-0691.2012.03841.x

Grad, Y. H., Miller, J. C., & Lipsitch, M. (2012). Cholera Modeling: Challenges to Quantitative Analysis and Predicting the Impact of Interventions. *Epidemiology*, *23*(4), 523–530. https://doi.org/10.1097/EDE.0b013e3182572581

Grim, C. J., Zo, Y.-G., Hasan, N. A., Ali, A., Chowdhury, W. B., Islam, A., Rashid, M. H., Alam, M., Morris, J. G., Huq, A., & Colwell, R. R. (2009). RNA Colony Blot Hybridization Method for Enumeration of Culturable Vibrio cholerae and Vibrio mimicus Bacteria. *Applied and Environmental Microbiology*, *75*(17), 5439–5444. https://doi.org/10.1128/AEM.02007-08

GTFCC, Overview of Ending Cholera, A Global Roadmap to 203o.

Hamner, S., Brown, B.L., Hasan, N.A., Franklin, M.J., Doyle, J., Eggers, M.J., Colwell, R.R., Ford, T.E., 2019. Metagenomic profiling of microbial pathogens in the little bighorn river, Montana. Int. J. Environ. Res. Public Health 16. https://doi.org/10.3390/ijerph16071097

Hasan, N. A., Ceccarelli, D., Grim, C. J., Taviani, E., Choi, J., Sadique, A., Alam, M., Siddique, A. K., Sack, R. B., Huq, A., & Colwell, R. R. (2013). Distribution of Virulence Genes in Clinical and Environmental Vibrio cholerae Strains in Bangladesh. *Applied and Environmental Microbiology*, *79*(18), 5782–5785. https://doi.org/10.1128/AEM.01113-13

Hashizume, M., Armstrong, B., Hajat, S., Wagatsuma, Y., Faruque, A. S. G., Hayashi, T., & Sack, D. A. (2008). The effect of rainfall on the incidence of cholera in Bangladesh. *Epidemiology*, *19*(1), 103–110. https://doi.org/10.1097/EDE.0b013e31815c09ea

Heltberg, R., P.B. Siegel, S.L. Jorgensen (2009) "Addressing Human Vulnerability to Climate Change: Toward a "˜No Regrets' Approach." *Global Environmental Change*. 19(1): 89-99.

Hood, M. A., & Winter, P. A. (2006). Attachment of Vibrio cholerae under various environmental conditions and to selected substrates. *FEMS Microbiology Ecology*, *22*(3), 215–223. https://doi.org/10.1111/j.1574-6941.1997.tb00373.x

Hourigan, S.K., Subramanian, P., Hasan, N.A., Ta, A., Klein, E., Chettout, N., Huddleston, K., Deopujari, V., Levy, S., Baveja, R., Clemency, N.C., Baker, R.L., Niederhuber, J.E., Colwell, R.R., 2018. Comparison of infant gut and skin microbiota, resistome and virulome between neonatal intensive care unit (NICU) environments. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.01361

Huffman, G., D. Bolvin, K. Braithwaite, K. Hsu, R. Joyce, and P. Xie. 2014. “Integrated Multi-SatellitE Retrievals for GPM (IMERG), Version 4.4. NASA’s Precipitation Processing Center.” 2014. ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/.

Huq, A., Small, E. B., West, P. A., Huq, M. I., Rahman, R., & Colwell, R. R. (1983). Ecological relationships between Vibrio cholerae and planktonic crustacean copepods. *Applied and Environmental Microbiology*, *45*(1), 275–283. https://doi.org/10.1128/AEM.45.1.275-283.1983

Huq, A., Xu, B., Chowdhury, M. A., Islam, M. S., Montilla, R., & Colwell, R. R. (1996). A simple filtration method to remove plankton-associated Vibrio cholerae in raw water supplies in developing countries. *Applied and Environmental Microbiology*, *62*(7), 2508–2512. https://doi.org/10.1128/AEM.62.7.2508-2512.1996

Huq, A., Haley, B.J., Taviani, E., Chen, A., Hasan, N.A., Colwell, R.R., 2012. Detection, isolation, and identification of Vibrio cholerae from the environment. Curr. Protoc. Microbiol. Chapter 6, Unit6A.5. https://doi.org/10.1002/9780471729259.mc06a05s26

Huq, Anwar, Anwar, R., Colwell, R., McDonald, M. D., Khan, R., Jutla, A., & Akanda, S. (2017a). Assessment of Risk of Cholera in Haiti following Hurricane Matthew. *The American Journal of Tropical Medicine and Hygiene*, *97*(3), 896–903. https://doi.org/10.4269/ajtmh.17-0048

Huq, Anwar, Hasan, N., Akanda, A., Whitcombe, E., Colwell, R., Haley, B., Alam, M., Jutla, A., & Sack, R. B. (2013). Environmental Factors Influencing Epidemic Cholera. *The American Journal of Tropical Medicine and Hygiene*, *89*(3), 597–607. https://doi.org/10.4269/ajtmh.12-0721

Huq, Anwar, Sack, R. B., Nizam, A., Longini, I. M., Nair, G. B., Ali, A., Morris, J. G., Khan, M. N. H., Siddique, A. K., Yunus, M., Albert, M. J., Sack, D. A., & Colwell, R. R. (2005a). Critical factors influencing the occurrence of Vibrio cholerae in the environment of Bangladesh. *Applied and Environmental Microbiology*, *71*(8), 4645–4654. https://doi.org/10.1128/AEM.71.8.4645-4654.2005

Huq, Anwar, Sack, R. B., Nizam, A., Longini, I. M., Nair, G. B., Ali, A., Morris, J. G., Khan, M. N. H., Siddique, A. K., Yunus, M., Albert, M. J., Sack, D. A., & Colwell, R. R. (2005b). Critical Factors Influencing the Occurrence of Vibrio cholerae in the Environment of Bangladesh. *Applied and Environmental Microbiology*, *71*(8), 4645–4654. https://doi.org/10.1128/AEM.71.8.4645-4654.2005

Jolliffe, I. T., and David B. Stephenson. 2012. *Forecast Verification : A Practitioner’s Guide in Atmospheric Science*. John Wiley & Sons.

Jutla AS, Akanda AS, Griffiths JK, Colwell R, Islam S. Warming oceans, phytoplankton, and river discharge: implications for cholera outbreaks. Am J Trop Med Hyg. 2011 Aug;85(2):303-8. doi: 10.4269/ajtmh.2011.11-0181. Erratum in: Am J Trop Med Hyg. 2011 Sep;85(3):582. PMID: 21813852; PMCID: PMC3144830.

Jutla, A. S., Akanda, A. S., & Islam, S. (2010a). Tracking Cholera in Coastal Regions Using Satellite Observations [1]: T RACKING C HOLERA IN C OASTAL R EGIONS U SING S ATELLITE O BSERVATIONS. *JAWRA Journal of the American Water Resources Association*, *46*(4), 651–662. https://doi.org/10.1111/j.1752-1688.2010.00448.x

Jutla, A. S., Akanda, A. S., & Islam, S. (2010b). Tracking Cholera in Coastal Regions Using Satellite Observations [1]: T RACKING C HOLERA IN C OASTAL R EGIONS U SING S ATELLITE O BSERVATIONS. *JAWRA Journal of the American Water Resources Association*, *46*(4), 651–662. https://doi.org/10.1111/j.1752-1688.2010.00448.x

Jutla, A., Whitcombe, E., Hasan, N., Haley, B., Akanda, A., Huq, A., Alam, M., Sack, R. B., & Colwell, R. (2013). Environmental Factors Influencing Epidemic Cholera. *American Journal of Tropical Medicine and Hygiene*, *89*(3), 597–607. https://doi.org/10.4269/ajtmh.12-0721

Jutla, Antarpreet, Aldaach, H., Akanda, A. S., Huq, A., & Colwell, R. R. (2015). Satellite Based Assessment of Hydroclimatic Conditions Related to Cholera in Zimbabwe. *PLOS-One*. https://doi.org/10.1371/journal.pone.0137828

Jutla, Antarpreet, Khan, R., & Colwell, R. (2017a). Natural Disasters and Cholera Outbreaks: Current Understanding and Future Outlook. *Current Environmental Health Reports*. https://doi.org/10.1007/s40572-017-0132-5

Jutla, Antarpreet, Khan, R., & Colwell, R. (2017b). Natural Disasters and Cholera Outbreaks: Current Understanding and Future Outlook. *Current Environmental Health Reports*, *4*(1), 99–107. https://doi.org/10.1007/s40572-017-0132-5

Kalan, L.R., Meisel, J.S., Loesche, M.A., Horwinski, J., Soaita, I., Chen, X., Uberoi, A., Gardner, S.E., Grice, E.A., 2019. Strain- and Species-Level Variation in the Microbiome of Diabetic Wounds Is Associated with Clinical Outcomes and Therapeutic Efficacy. Cell Host Microbe 25, 641-655.e5. https://doi.org/10.1016/j.chom.2019.03.006

Kaneko, T., & Colwell, R. R. (1975). Adsorption of Vibrio parahaemolyticus onto Chitin and Copepods. *Applied Microbiology*, *29*(2), 269–274. https://doi.org/10.1128/AM.29.2.269-274.1975

Kaper, J., Lockman, H., Colwell, R. R., & Joseph, S. W. (1979). Ecology, serology, and enterotoxin production of Vibrio cholerae in Chesapeake Bay. *Applied and Environmental Microbiology*, *37*(1), 91–103. https://doi.org/10.1128/AEM.37.1.91-103.1979

Khan, Rakib, Anwar, R., Akanda, S., McDonald, M. D., Huq, A., Jutla, A., & Colwell, R. (2017). Assessment of risk of cholera in Haiti following Hurricane Matthew. *American Journal of Tropical Medicine and Hygiene*, *97*(3), 896–903. https://doi.org/10.4269/ajtmh.17-0048

Khan, Rakibul, Nguyen, T. H., Shisler, J., Lin, L.-S., Jutla, A., & Colwell, R. (2018a). Evaluation of Risk of Cholera after a Natural Disaster: Lessons Learned from the 2015 Nepal Earthquake. *Journal of Water Resources Planning and Management*, *144*(8), 04018044. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000929

Khan, Rakibul, Nguyen, T. H., Shisler, J., Lin, L.-S., Jutla, A., & Colwell, R. (2018b). Evaluation of Risk of Cholera after a Natural Disaster: Lessons Learned from the 2015 Nepal Earthquake. *Journal of Water Resources Planning and Management*, *144*(8), 04018044. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000929

Khan, Rakibul, Usmani, M., Akanda, A., Palash, W., Gao, Y., Huq, A., Colwell, R., & Jutla, A. (2019). Long-Range River Discharge Forecasting Using the Gravity Recovery and Climate Experiment. *Journal of Water Resources Planning and Management*, *145*(7), 06019005. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001072

Kim, Kiyoung, Jongmin Park, Jongjin Baik, and Minha Choi. 2017. "Evaluation of Topographical and Seasonal Feature Using GPM IMERG and TRMM 3B42 over Far-East Asia." *Atmospheric Research*. https://doi.org/10.1016/j.atmosres.2016.12.007.

Koelle, K., Rodó, X., Pascual, M., Yunus, Md., & Mostafa, G. (2005). Refractory periods and climate forcing in cholera dynamics. *Nature*, *436*(7051), 696–700. https://doi.org/10.1038/nature03820

Lax, S., Smith, D.P., Hampton-marcell, J., Owens, S.M., Shogan, B.D., Weiss, S., Metcalf, J.L., Ursell, L.K., 2012. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science (80-.). 12702, 1048–1052

Lipp, E. K., Huq, A., & Colwell, R. R. (2002). Effects of Global Climate on Infectious Disease: The Cholera Model. *Clinical Microbiology Reviews*, *15*(4), 757–770. https://doi.org/10.1128/CMR.15.4.757-770.2002

Lobitz, B., Beck, L., Huq, A., Wood, B., Fuchs, G., Faruque, A. S. G., & Colwell, R. (2000). Climate and infectious disease: Use of remote sensing for detection of Vibrio cholerae by indirect measurement. *Proceedings of the National Academy of Sciences*, *97*(4), 1438–1443. https://doi.org/10.1073/pnas.97.4.1438

Louis, V. R., Russek-Cohen, E., Choopun, N., Rivera, I. N. G., Gangle, B., Jiang, S. C., Rubin, A., Patz, J. A., Huq, A., & Colwell, R. R. (2003). Predictability of Vibrio cholerae in Chesapeake Bay. *Applied and Environmental Microbiology*, *69*(5), 2773–2785. https://doi.org/10.1128/AEM.69.5.2773-2785.2003

Lucas, R. M., & McMichael, A. J. (2005). Association or causation: Evaluating links between "environment and disease." *Bulletin of the World Health Organization*, 4.

Luque Fernández, M. Á., Bauernfeind, A., Jiménez, J. D., Gil, C. L., Omeiri, N. E., & Guibert, D. H. (2009). Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003–2006: Analysis of a time series. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *103*(2), 137–143. https://doi.org/10.1016/j.trstmh.2008.07.017

Lyons, K. (2017, October 12). Yemen's cholera outbreak now the worst in history as millionth case looms. *The Guardian*.

Martinez, R. M., Megli, C. J., & Taylor, R. K. (2010). Growth and Laboratory Maintenance of *Vibrio cholerae*. *Current Protocols in Microbiology*, *17*(1), 6A.1.1-6A.1.7. https://doi.org/10.1002/9780471729259.mc06a01s17

Mata, L. (1994). Cholera El Tor in Latin America, 1991-1993. *Annals of the New York Academy of Sciences*, *740*, 55–68. https://doi.org/10.1111/j.1749-6632.1994.tb19853.x

Miller, ChristopherJ., Drasar, BohumilS., & Feachem, RichardG. (1982). CHOLERA AND ESTUARINE SALINITY IN CALCUTTA AND LONDON. *The Lancet*, *319*(8283), 1216–1218. https://doi.org/10.1016/S0140-6736(82)92340-6

Mishra, A., Taneja, N., Sharma, R. K., Kumar, R., Sharma, N. C., & Sharma, M. (2011). Amplified fragment length polymorphism of clinical and environmental Vibrio cholerae from a freshwater environment in a cholera-endemic area, India. *BMC Infectious Diseases*, *11*(1), 249. https://doi.org/10.1186/1471-2334-11-249

Mitchell, Todd. 2014. "Elevation Data in NetCDF." 2014. http://research.jisao.washington.edu/data_sets/elevation/.

Mittermaier, M., & Roberts, N. (2010). Intercomparison of Spatial Forecast Verification Methods: Identifying Skillful Spatial Scales Using the Fractions Skill Score, *Weather and Forecasting*, *25*(1), 343-354. Retrieved Jan 14, 2021, from https://journals.ametsoc.org/view/journals/wefo/25/1/2009waf2222260_1.xm

Mohammed, Safa, Mohamed Hamouda, Mohammed Mahmoud, and Mohamed Mohamed. 2020. "Performance of GPM-IMERG Precipitation Products under Diverse Topographical Features and Multiple-Intensity Rainfall in an Arid Region." *Hydrology and Earth System Sciences Discussions*. https://doi.org/10.5194/hess-2019-547.

Mukerjee, S., Basu, S., & Bhattacharya, P. (1965). A new trend in cholera epidemiology. *BMJ*, *2*(5466), 837–839. https://doi.org/10.1136/bmj.2.5466.837

Nasr-Azadani, F., Khan, R., Rahimikollu, J., Unnikrishnan, A., Akanda, A., Alam, M., Huq, A., Jutla, A., & Colwell, R. (2016). Hydroclimatic sustainability assessment of changing climate on cholera in the Ganges-Brahmaputra basin. *Advances in Water Resources*. https://doi.org/10.1016/j.advwatres.2016.11.018

Neogi, S. B., Lara, R., Alam, M., Harder, J., Yamasaki, S., & Colwell, R. R. (2018). Environmental and hydroclimatic factors influencing Vibrio populations in the estuarine zone of the Bengal delta. *Environmental Monitoring and Assessment*, *190*(10). https://doi.org/10.1007/s10661-018-6925-7

Oliver, J. D. (2010). Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiology Reviews*, *34*(4), 415–425. https://doi.org/10.1111/j.1574-6976.2009.00200.x

Paz, S. (2009). Impact of Temperature Variability on Cholera Incidence in Southeastern Africa, 1971–2006. *EcoHealth*, *6*(3), 340–345. https://doi.org/10.1007/s10393-009-0264-7

Paz S. The cholera epidemic in Yemen - How did it start? The role of El Niño conditions followed by regional winds. Environ Res. 2019 Sep;176:108571. doi: 10.1016/j.envres.2019.108571. Epub 2019 Jul 2. PMID: 31288197.

Pollitzer, R. (1954). Cholera studies. 1. History of the disease. *Bulletin of the World Health Organization*, *10*(3), 421–461.

Pollitzer, R., Swaroop, S., & Burrows, W. (1959). *Cholera*. World Health Organization.

Ponnusamy, D., Kozlova, E. V, Sha, J., Erova, T.E., Azar, S.R., Fitts, E.C., 2016. Cross-talk among flesh-eating Aeromonas hydrophila strains in mixed infection leading to necrotizing fasciitis. Proc. Natl. Acad. Sci. 113, 722–727. https://doi.org/10.1073/pnas.1523817113

Prakash, Satya, Ashis K. Mitra, Amir AghaKouchak, Zhong Liu, Hamidreza Norouzi, and D. S. Pai. 2018. "A Preliminary Assessment of GPM-Based Multi-Satellite Precipitation Estimates over a Monsoon Dominated Region." *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2016.01.029.

Qadri, F., Islam, T and Clemens, J. (2017) Cholera in Yemen — An Old Foe Rearing Its Ugly Head. *The New England Journal of Medecine*

Rebaudet S, Sudre B, Faucher B, Piarroux R. Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. *J Infect Dis.* 2013 Nov 1;208 Suppl 1:S46-54. doi: 10.1093/infdis/jit195. PMID: 24101645.

Rivera, I. G., Chowdhury, M. A. R., Sanchez, P. S., Sato, M. I., Huq, A., Colwell, R. R., & Martins, M. T. (1995). Detection of cholera (ctx) and zonula occludens (zot) toxin genes in Vibrio cholerae O1, O139 and non-O1 strains. *World Journal of Microbiology & Biotechnology*, *11*(5), 572–577. https://doi.org/10.1007/BF00286376

Roberts, Nigel M., and Humphrey W. Lean. 2008. "Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events." *Monthly Weather Review*. https://doi.org/10.1175/2007MWR2123.1.

Roszak, D. B., & Colwell, R. R. (1987). Survival strategies of bacteria in the natural environment. *Microbiological Reviews*, *51*(3), 365–379.

Roy, M.A., Arnaud, J.M., Jasmin, P.M., Hamner, S., Hasan, N.A., Colwell, R.R., Ford, T.E., 2018. A Metagenomic Approach to Evaluating Surface Water Quality in Haiti. Int. J. Environ. Res. Public Health 15, 2211. https://doi.org/10.3390/ijerph15102211

Sasaki, S., Suzuki, H., Fujino, Y., Kimura, Y., & Cheelo, M. (2009). Impact of Drainage Networks on Cholera Outbreaks in Lusaka, Zambia. *American Journal of Public Health*, *99*(11), 1982–1987. https://doi.org/10.2105/AJPH.2008.151076

Schwartz, B. S., Harris, J. B., Khan, A. I., Larocque, R. C., Sack, D. A., Malek, M. A., Faruque, A. S. G., Qadri, F., Calderwood, S. B., Luby, S. P., & Ryan, E. T. (2006). Diarrheal epidemics in Dhaka, Bangladesh, during three consecutive floods: 1988, 1998, and 2004. *The American Journal of Tropical Medicine and Hygiene*, *74*(6), 1067–1073.

Sharp, J., & Salaam-Blyther, T. (2017). *Yemen: Cholera Outbreak*. 3.

Singleton, F. L., Attwell, R. W., Jangi, M. S., & Colwell, R. R. (1982). Influence of salinity and organic nutrient concentration on survival and growth of Vibrio cholerae in aquatic microcosms. *Applied and Environmental Microbiology*, *43*(5), 1080–1085.

Skok, Gregor, and Nigel Roberts. 2016. "Analysis of Fractions Skill Score Properties for Random Precipitation Fields and ECMWF Forecasts." *Quarterly Journal of the Royal Meteorological Society*. https://doi.org/10.1002/qj.2849.

Stamps, B.W., Leddy, M.B., Plumlee, M.H., Hasan, N.A., Colwell, R.R., Spear, J.R., 2018. Characterization of the Microbiome at the World ' s Largest Potable Water Reuse Facility 9, 2435. https://doi.org/10.3389/fmicb.2018.02435

Tamplin, M. L. (n.d.). Attachment of Vibrio cholerae Serogroup 01 to Zooplankton and Phytoplankton of Bangladesh Waters. *APPL. ENVIRON. MICROBIOL.*, 4.

Taylor, D. L., Kahawita, T. M., Cairncross, S., & Ensink, J. H. J. (2015). The Impact of Water, Sanitation and Hygiene Interventions to Control Cholera: A Systematic Review. *PLOS ONE*, *10*(8), e0135676. https://doi.org/10.1371/journal.pone.0135676

Vezzulli, L., Grande, C., Reid, P. C., Hélaouët, P., Edwards, M., Höfle, M. G., Brettar, I., Colwell, R. R., & Pruzzo, C. (2016). Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proceedings of the National Academy of Sciences*, *113*(34), E5062–E5071. https://doi.org/10.1073/pnas.1609157113

UNICEF: Personal communications took place throughout the project through meetings and emails. The insight gained from these was not as extensive as planned to due COVID and workload of the Yemen UNICEF teem.

UNICEF publication, Evaluation of level 3 response to cholera epidemic in Yemen: A crises within a crises, 2018.
Evaluation_of_the_UNICEF_Level_3_response_to_the_cholera_epidemic_in_Yemen_HQEO-2018-001.pdf

UNICEF publication: Evaluation of UNICEF Level 3 response to the cholera epidemic in Yemen: A crisis within a crisis. Informal briefing the UNICEF Executive Board, Augst 2018 https://www.unicef.org/french/about/execboard/files/EB_presentation-Yemen_Cholera_eval-2018-08.22.pdf

USAID Yemen WASH Fact Sheet (last updated 2020). https://www.usaid.gov/yemen/fact-sheets/water-sanitation-and-hygiene-fact-sheet#:~:text=As%20a%20result%20of%20the,humanitarian%20crisis%20in%20the%20world.&text=As%20of%20December%202019%2C%20cholera,January%202018%2C%20including%201%2C513%20deaths.

Walters, David, Anthony J. Baran, Ian Boutle, Malcolm Brooks, Paul Earnshaw, John Edwards, Kalli Furtado, et al. 2019. "The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 Configurations." *Geoscientific Model Development*. https://doi.org/10.5194/gmd-12-1909-2019.

WHO. (2018). Weekly epidemiological record (Vol. 93). https://apps.who.int/iris/bitstream/handle/10665/274654/WER9338.pdf?ua=1

WHO Regional Office for the Eastern Mediterranean - web page: http://www.emro.who.int/pandemic-epidemic-diseases/cholera/outbreak-update-cholera-in-yemen-28-june-2020.html#:~:text=The%20cumulative%20total%20number%20of,the%20333%20districts%20of%20Yemen.

WHO, lessons learned from cholera response in Yemen (awaiting source)

Wilkinson, Emily., Weingärtne, Lena., Choularton, Richard., Bailey, Meghan., Todd, Martin., Kniveton, Dominic., and Cabot Venton, Courtenay (2018) Forecasting Hazards, Averting Disasters, Implementing Forecast Based Early Action at Scale. Report for the WISER programme.

ANNEXES

Annex 1: CRM report for Yemen from 11 January 2021

Yemen 011121.pdf

Annex 2: Rainfall forecasts for Yemen

Yemen Rainfall Assessment - 202101	Yemen Rainfall Guidance - 2021011	Yemen Rainfall Assessment - 202101

Annex 3: UNICEF Case Study on use of CRM and rainfall forecasts

Yemen Case Study on Rainfall data for Cholera Preparedness.pdf

Annex 4: Details on formulas and additional methodology for validation of CRM

The purpose of this effort is the use the existing model (without any improvements) and validate it using Yemen epidemiological data. Although the TM has been validated using historical data from Sudan, Bangladesh, Mozambique, Algeria, Cameroon, Haiti, and other parts of the world, Yemen provides a unique near real time opportunity to examine the validity of the trigger module of the CRM (Jutla et al., 2013; Jutla et al., 2015). Current version of CRM was developed with the purpose to provide the longest possible prediction lead time. This model was originally developed using cholera data collected from British India period and subsequently validated at several other location. In our previous analysis and published study (Jutla et al., 2013), we have reported “odds of occurrence of the disease during above and below normal rainfall and following the months of above average air temperature at the nine locations”. This analysis formed the key basis for development of the model for assessment of risk for following four weeks from the time the values of risk are computed. Subsequently, we tested our hypothesis on several location in the world. For example Figure 6 in the study published in PLOS (Jutla et al., 2015), shows that the hypothesis was true for several locations in Africa. We then used this hypothesis in Nepal, Sudan, Haiti, Mozambique and found it to be valid there. Hence, we converted it to a working model which was then used at locations such as Nepal (Khan et al., 2018a) and Haiti (Huq et al., 2017a) on a retrospective basis. The model appears to capture the trigger of cholera in those locations and hence the same model was used for Yemen. CRM is a hypothesis driven model and hence it does not require calibration unlike traditional disease simulation models.

There is an opportunity to check model forecast on 1, 2, 3 weekly basis, however, this is beyond the scope of current work. However, during assimilation of Met Office precipitation data into the modeling schema, such understanding will provide further insights on what to do with short term cholera risk. The intervention strategies for a short term (such as weekly scale) are going to be very different that the long term (monthly to seasonal). The current scope is to validate the model outputs in its current form.

The ideal version of prediction system has eight different routine. Figure A4a (below) shows basic structure of the model. The prediction system starts with identification of major disaster (man-made or natural) which is then fed to hydrological and climatological routine where large scale assessment of regional conditions supporting vibrios is conducted. Thereafter, environmental routine makes an evaluation on possible locations of the vibrios given hydroclimatological conditions. Vibrio survival routine was developed using about 40 years of data knowledge from University of Maryland. This is in form of a lookup table as a function of temperature and survival of vibrios under those conditions (Huq et al., 2005a). The WASH routine implies availability of safe water and sanitation infrastructure at the time prior and after disaster, where such information is routed to population routine. Population routine implies density of humans and locations of settlements in a region. This dataset is derived from LandScan data managed by Oak Ridge National Lab in the USA. Water security routine has information on where and when any intervention is being done so that outbreak of cholera can be stopped. Finally, our in-house developed algorithm combines the entire information and produces a risk value. The advantage of using risk values instead of the actual case (prevalence or incidence) values is that the algorithm becomes independent from the epidemiological data requirements. In other words, the algorithm do not require epidemiological data to produce the risk values. However, we do validate our model outputs once the epidemiological data becomes available.

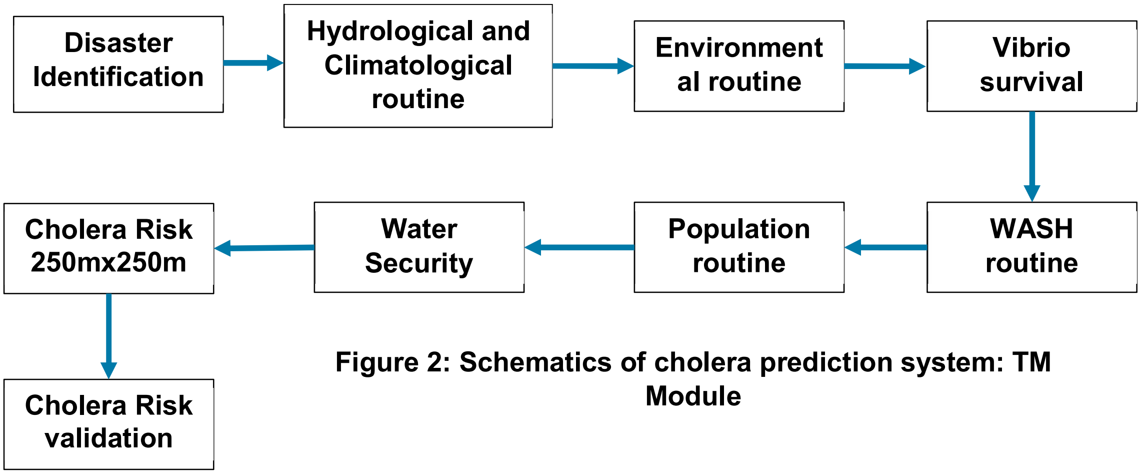


Figure 2: Schematics of cholera prediction system: TM Module

Figure A4a: Schematics of cholera prediction system

However, different regions will have different datasets available, therefore, depending on availability of data we will modify the utility of a particular routine. For example, in Yemen, information on water security, and WASH routines are were not used since such data are not available.

The analysis was started with the entire datasets and then continued to separately to all three years (2017, 2018 and 2019). The individual year analysis was conducted to identify any changes in dynamics of cholera in the region. A separate year wise analysis may provide an indication as to whether cholera has transitioned from epidemic to endemic modes in the country.

Pearson correlation coefficients (including 95% confidence) were computed between CRM risk model output and epidemiological time series of cholera cases. The analysis was conducted in the following way (for each governorate): Obtain weekly cholera risk values for all the pixels in a governorate from 2017 to 2019.

- Take an average of the risk value to develop one time series for a governorate.
- Correlate Week 1 risk value of CRM with total number of cholera cases in the next four weeks (week 1 to week 4). Correlate Week 2 value of CRM with average of week 2 to week 5 values of cholera cases. (Sample plot between cholera cases and CRM risk values is shown in Figure A4b).
- Statistical significance was computed at 95% confidence interval. Details on Pearson correlation coefficient and associated significance here https://en.wikipedia.org/wiki/Pearson_correlation_coefficient)

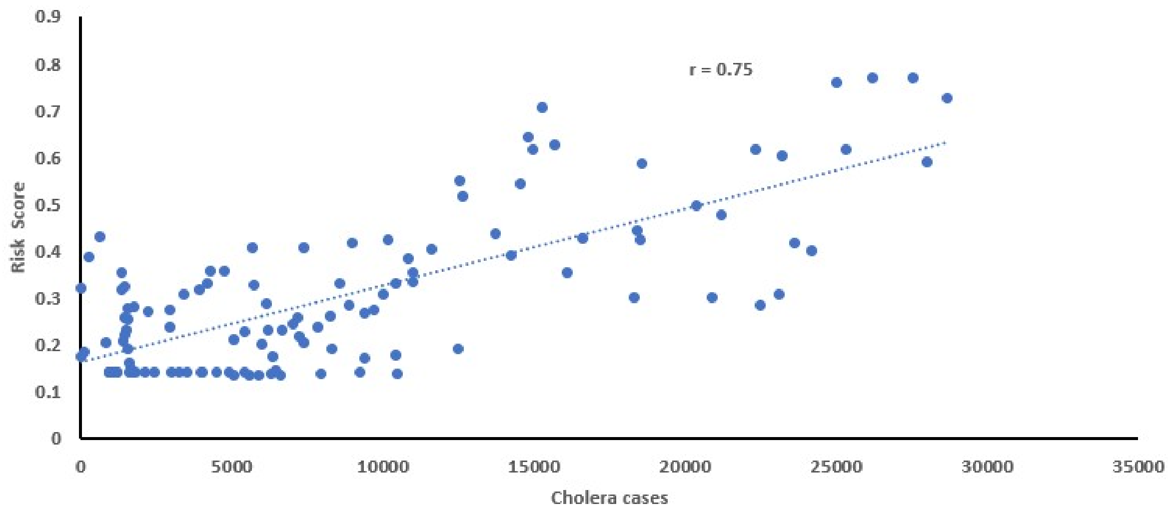


Figure A4b: Weekly correlation between risk scores and cholera cases in Al-Hudaydah from 2017-2019.

Pearson correlation coefficient "measures linear correlation between two variables *X* and *Y*. It has a value between +1 and −1. A value of +1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation." (https://en.wikipedia.org/wiki/Pearson_correlation_coefficient). This is a parametric value of association between two variables. The standard student t-test (https://en.wikipedia.org/wiki/Student%27s_t-test) was used to determine statistical significance of the correlation values and has been used by scientists all over the world for over several decades. As a complementary analysis, we have used a non-parametric Kendall tau rank correlation coefficient which is a statistic "used to measure the ordinal association between two measured quantities" (https://en.wikipedia.org/wiki/Kendall_rank_correlation_coefficient). The purpose here is to determine if CRM model outputs are statistically related to epidemiological data. If these are, then we deem that the model is performing satisfactorily and will stop further analysis.

CRM uses daily rainfall data obtained from the National Aeronautics and Space Administration (NASA). Rainfall data, at a resolution of 0.25°X0.25° from the Tropical Rainfall Measuring Mission (TRMM), was employed to compute the long-term average (four weekly average of rainfall). Daily rainfall data at a spatial resolution of 0.1°X0.1° from the Global Precipitation Mission (GPM 1998-2020) was averaged over four week time scale. Following procedure was followed:

1. Obtain long term average (1998-2018), on four week scale at governorate level, from TRMM.
2. Obtain four week precipitation averaged over a governorate for a particular four weeks in question.
3. Subtract step 2 from step 1 to get a four weekly deviation, or, $Precipitation\ anomaly = Monthly\ value_{Precipitation} - Long\ term\ average\ value_{Precipitation}$

Daily air temperature data on the surface, at a spatial resolution of 0.5°X0.625°, were obtained from NASA's Modern-Era Retrospective analysis Research and Application, Version 2 (NASA-MERRA 2) (hydrological and climatological routine). A similar methodology to compute deviations from long term averages was used. LandScan population data at a spatial resolution of 1kmX1km from Oak Ridge National Laboratory were acquired and employed for this model (population routine). The model output was resampled at 0.1°X0.1° (native resolution to GPM data) and is valid for the next four weeks(Anwar Huq et al., 2017a; Antarpreet Jutla et al., 2015). Weighted raster overlay algorithm (Andersson & Mitchell, 2006) was used to compute risk maps for cholera. Weighted raster overlay is a technique for applying a common measurement scale of values to diverse and dissimilar inputs to create an integrated output with attributable outcomes (e.g., high risk to low risk). A population density layer along with two month's lagged monthly mean air temperature anomaly and one month lagged monthly total precipitation anomaly layers were used to produce the hydroclimatic risk map for likelihood of occurrence of cholera. Thereafter, information on regional water resources (WASH infrastructure) was added to the risk assessment. When applying the weighted raster overlay algorithm (Andersson & Mitchell, 2006), all input raster layers must have an assigned integer value, or it must be converted to an integer. Each input raster is assigned a new value based on an evaluation scale which is shown in our previously published manuscript (Rakibul Khan et al., 2018a). The new values were deemed to be "reclassification" of the original input raster values. The evaluation scale was determined based on the range of all raster layers for the variable under consideration. For example, the air temperature anomaly evaluation scale was determined based on maximum and minimum values of the raster layers for all May, June, July anomalies. Every input raster was weighted according to importance (in terms of percent influence) and was converted to relative percentage; total being 100. Changing evaluation scales or percent influence can change the results in the final risk map. Different weights were computed while determining risk of cholera under various scenarios e.g. hydroclimatic and WASH based risk assessment. The relative weight of each variable was assigned a risk level, e.g. very high, high, moderate high, low, or very low. Hydroclimatological departure from normal conditions was assumed to the strongest contributor to risk of cholera. Using each pair of precipitation and temperature anomalies, along with population density, three composite maps of spatial cholera risk for July, August and September were generated. Figure A4c shows the flowchart for implementation of the weighted raster overlay algorithm. The blue dotted box at the left represents layers used to generate the hydro-climatic risk map and the solid red box incorporates the WASH infrastructure into the risk computation. Figure 5b shows sample risk scores and cholera time series for one of the governorates. A sample temperature and preciptiation long term averages is shown in Appendix A.

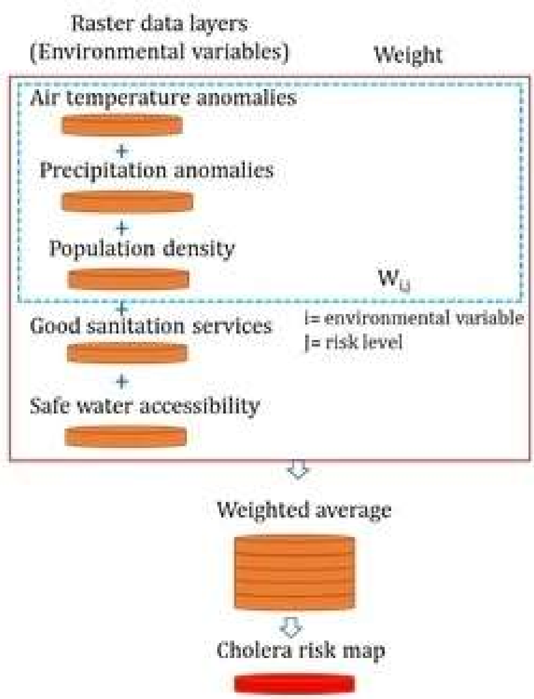


Figure A4c: Weighted raster overlay flow chart. The dotted box represents the layers used to generate hydroclimatic risk map and the solid box to incorporate water resources into the risk calculations.

Annex 5: Limitations of satellite data

A limitation of using the IMERG satellite product over Yemen is that the product will not be able to use precipitation gauges to assist in the bias correction process. This means that if there are biases between satellite coverage and true precipitation coverage it may not be possible to correct for these biases.

A study of IMERG performance across the Kingdom of Saudi Arabia, which is on Yemen's Northern border, has found that the product performs well in detecting light rain (0-10 mm) precipitation owing to its 13-channel microwave imager and dual frequency precipitation radar (Mohammed et al. 2020). They also found that the detection errors in the product increased as the intensity of precipitation increased and that this error tended to come in the form of an underestimation by the satellite. A topographical analysis done in the same study finds that the IMERG product is more biased at higher altitudes (more than 1000 m). These results are presented in the context of other studies that have found IMERG products to have higher errors in mountainous and in coastal regions. A study by Prakash et al. 2018 found that there were increased errors and bias over mountainous regions in India. Another study finds increased errors in both coastal and mountainous regions in Far-East Asia (Kim et al. 2017). Overall, Mohammed et al. 2020 concluded that there is generally a low correlation coefficient between ground observation data and IMERG precipitation over Saudi Arabia.

Although Yemen and Saudi Arabia do not share identical climatology, owing mainly to the Mediterranean influence on northern Saudi Arabia and the extent to which the ITCZ affects the regions differently, the study by Mohammed et al. 2020 highlights the limitations of validating forecasts using a satellite product alone. It is particularly important to recognise this as a limitation to this analysis given that the mountainous areas in the west of Yemen are the regions that experience the most precipitation but are also the most populated.

Annex 6: Further details of the contingency table analysis

The classification of the satellite and forecast data into separate binary fields allows the categorisation of the data into 2 x 2 contingency tables. Each cell in the table, as illustrated in Figure A6a, represents one possible outcome from comparing the satellite to the forecast data.

Dataset	Satellite data above threshold	Satellite data below threshold	
Precipitation forecast above threshold	Hits (a)	False Alarms (b)	Forecast rate (a + b) / N
Precipitation forecast below threshold	Misses (c)	Correct Rejections (d)	c + d
	Base rate (a + c) / N	b + d	N = a + b + c + d

Figure A6a: Classification of cells in a 2 x 2 contingency table.

Each day of the month was classified into one of the four contingency categories described above, so the total number of records, N, will correspond to the number of days in the selected month. For a day to be classified as a Hit, Miss or False Alarm day, it is not necessary for the comparison to be at the grid cell level in the satellite and precipitation Yemen domain. Each grid cell is compared for the same location in the satellite and precipitation field and it is a comparison of the precipitation events aggregated to the Yemen domain level. As a result, the contingency tables used in this report provide a preliminary investigation of the Yemen domain. However, they provide little information about the skill of the forecast compared to the satellite data at smaller spatial resolutions than that of the Yemen domain.

Many statistics can be derived from a contingency table (see Jolliffe and Stephenson 2012). In this report, the base rate and forecast rate are used. The base rate is the proportion of events observed ((a+c)/N, using the nomenclature defined), and the proportion of events forecast is called the forecast rate ((a+b)/N). The ratio of the forecast rate to the base rate provides the frequency bias. A bias of one is obtained when the precipitation forecast rate is the same as the base rate, the forecast is then well calibrated. However, a bias of one may not give the best result, if the objective is to minimise the impact of the forecast on the users. This is especially the case if the economic or societal impact of a False Alarm is small compared to the impact of a Miss. In such a situation, users may tolerate more False Alarms compared to Misses and as a result a bias greater than one may be more acceptable from an impacts perspective. Care should be taken when considering the frequency bias in this analysis. Normally, a forecast bias occurs when there are systematic differences between forecasts and observations and the implication is that the forecasts tend to be too high or too low. However, in this analysis, ground observations are not available and satellite data are used as a proxy for these missing ground observations. However, the satellite data may themselves be biased compared to the unknown ground observation. As a result, any deviation of the frequency bias from the value of 1, cannot be solely attributed to error in the forecast.

The proportion of Hits, Misses, False alarms and Correct Rejections for the 24-hour precipitation accumulation for day 2 (top row), day 3 (middle row) and day 6 (bottom row), from January 2017 to August 2019 for the Yemen domain is shown in Figure 6b. The Figure shows the following:

1. A clear seasonal cycle is noted:

a. Typically, between May and October the number of Hits and Correct Rejections is proportionally higher compared to November to April.
b. Between November and March there are fewer satellite precipitation events (Misses + Hits - base rate) compared to the rest of the year.

2. The accuracy of the GM day 3 forecasts for the 10 mm threshold generally exceed those of the day 6 GM forecasts, throughout the wet season. The day 2 forecasts from the CAM are generally more accurate in terms of the percentage of correct forecasts (number of Hits and Correct Rejections) than the day 3 GM forecasts. However, it is noticeable that the day 2 forecasts have a higher percentage of Misses compared to the day 3 forecasts, which have a higher percentage of False Alarms.
3. The accuracy of the GM forecasts improves over time for the period investigated here, especially between 2017 and 2018. This could be due to improvements in the model, but it is also possible that different weather processes could affect the region in 2017 compared to 2018 and 2019. However, the results from the CAM are more mixed, with some improvements in the CAM performance between 2017 and 2018, particularly in the first half of the year. However, the percentage of Hits and Correct Rejections falls for all months (except January and August) when 2018 is compared to 2019

These results generally tie in with the known climatology of the region, which is influenced by the ITCZ and the RSCZ. The transition periods between the dry season (December to March) and the wet season (June to September) are the months of April/May and October/November. In April 2018 and 2019, the Hit rate across day 2, 3, and 6 exceeded 65 %, whereas the Hit rate for April 2017, was low at approximately 20 % and 15 % for day 3 and 6 rising to 50 % for day 2. The difference in the April Hit rates could show the variable nature of the transition from the dry to the wet season with the transition from dry to wet season arriving later in 2017 compared to 2018 and 2019. However, it may also show the improvements in the weather forecast over this period.

The base rate looks at the proportion of precipitation events in the satellite data and will not be affected by forecast model updates. This data displays a similar pattern: higher base rates for April 2018 and 2019 compared to April 2017. This suggests that the timing of the transition to wet season can be variable. Camacho et al. (2018) also observed a link between the start of the wet season and the outbreak of cholera waves. Therefore, the results will focus on the period April to November, which covers the transition to and from the wet season (Section 3.1).

In the wet season, the proportion of Misses is large compared to the number of False Alarms for day 2 but for day 3 and day 6 the number Misses is small compared to the number of False Alarms. This could be attributed to the different forecast models, different thresholds, the re-gridding of the CAM forecast data to the coarser satellite resolution (Annex 5), or timing differences as the lead time, from when the forecast is issued to when it is valid, increases. The higher resolution of the CAM, compared to the GM, allows for improved modelling of convective precipitation, and combined with the better resolved topography due to the smaller grid cell size would suggest using a higher “take action” threshold compared to the GM. The threshold chosen for the “take action category” was 20 mm for day 1 and 2 compared to 10 mm for subsequent days.

To allow for a comparison of the model data with the satellite data necessitated a re-gridding of the forecast data to that of the satellite data, as discussed in Annex 5. The grid resolution for the satellite data and the GM data were very similar. However, the resolution of the satellite data (11 km) are approximately 3 times greater than that of the CAM model (4 km). The re-gridding will inevitably result in a smoothing of the forecast precipitation fields and this could explain the increase in the number of Misses (and decrease in False Alarms) observed compared to the GM forecasts. Ideally, in future work, the CAM should either be compared to satellite data at a similar resolution, or the 20 mm threshold adjusted to a lower threshold to take into consideration the spatial averaging that happened as part of the re-gridding. Hence, the interpretation of any results from this verification analysis from the CAM should be treated with caution, and further analysis that investigates the sensitivity of the threshold to the re-gridding of the CAM data to the satellite scale should be undertaken prior to any analysis of the Day 2 forecasts.

The next step is to analyse the monthly scatter plots of the daily counts of precipitation, where both the satellite and the forecast data have at least one grid cell “take action threshold”, for the Yemen domain. The 24-hour precipitation for day 3 and day 6 are shown in Figure A6b and Figure A6c, respectively. Daily data from different years are plotted using different shapes: 2019 with diamonds, 2018 with squares, 2017 with circles. This is the equivalent of producing a scatter plot of the number of cells for the forecast and satellite data that have been classified as a Hit event. Counts reported in the satellite (forecast) data are not required to be in (or near) the same cell location in the forecast (satellite) data. The black dotted line represents the one-one correspondence line, which is where all the points would lie if the forecast data was in perfect agreement with the satellite data.

May, for both day 3 and day 6 precipitation events, tends to be the most variable month across the wet season years, containing both the largest count of satellite precipitation grid squares and forecast grid squares, not necessarily on the same day, with values close to 3000 which is approximately 20% of the total grid cells in the Yemen domain. Across the early wet season months (April to July), the spread of the counts associated with the Hit category appears to be smaller in 2017. As described earlier whether this is due to changes in the global model or different weather process affecting the Yemen region in 2017 compared to 2018 and 2019 cannot be ascertained from these data sources. For both day 3 and day 6, for the wet season months, there is a slight tendency for the forecast data to predict more precipitation events compared to the satellite data (more noticeable in August).

Mohammed et al (2020), found that for precipitation events that exceeded 10 mm the satellite data underestimated the precipitation compared to observations. Given that there are no observations available for Yemen, the fact that the satellite data underestimates precipitation events compared to the forecast data (whilst not providing robust, statistical significant evidence) may suggest that at least a small component of the bias could be due to the satellite data under reporting precipitation events.

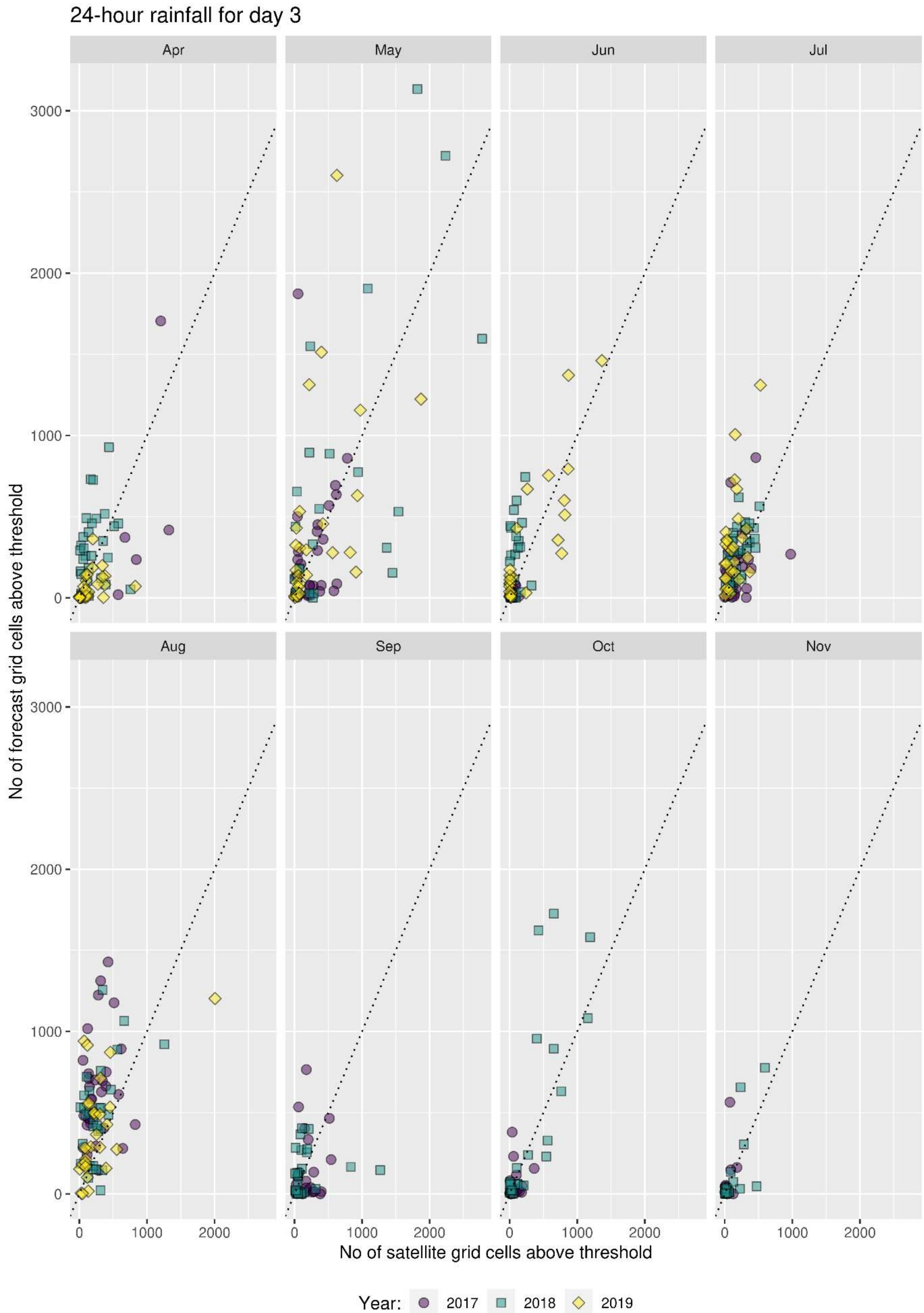


Figure A6b: Scatter plot of daily counts where both forecast and satellite data have cells above "active threshold" for day 3. Each point represents one day, and days in different years have different shapes (2017 circle, 2018 square, 2019 diamond). The Yemen domain has 14 700 grid cells in total.

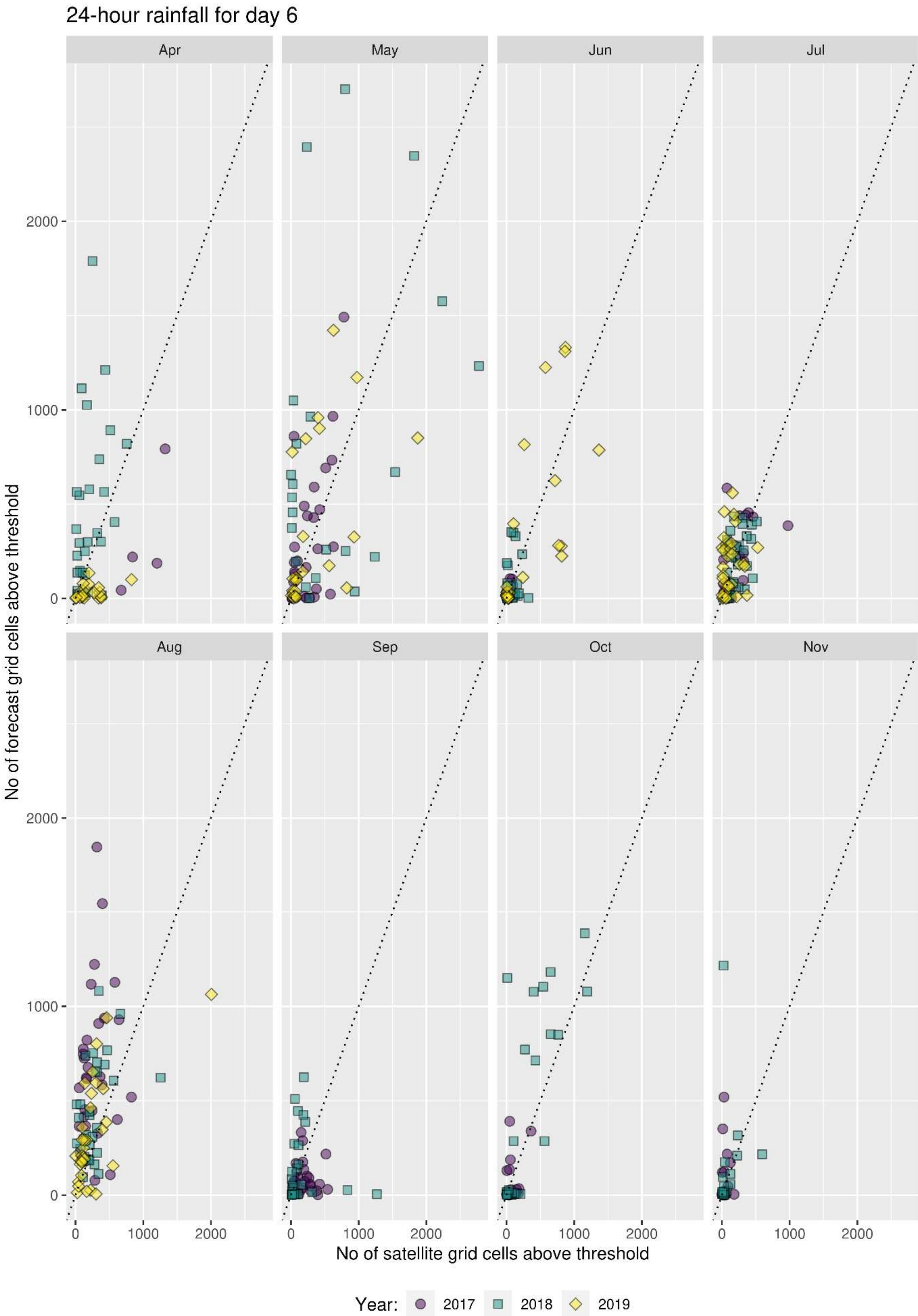


Figure A6c: Scatter plot of daily counts where both forecast and satellite data have cells above "active threshold" for day 6. Each point represents one day, and days in different years have different shapes (2017 circle, 2018 square, 2019 diamond). The Yemen domain has 14 700 grid cells in total.

Annex 7: Calculation of Fractions Skill Score (FSS)

The calculation of the FSS can be considered in three main steps:

1. Re-grid the forecast model data, so that it is on the same spatial resolution as the satellite data,
2. Convert both the satellite and forecast fields to binary data,
3. Decide upon the extent of the spatial area around each grid cell. For each spatial area, calculate in both the satellite and forecast data the fraction of precipitation events,
4. Calculation of the Fractions Skill Scores.

To assess how similar the forecast and satellite fields are spatially, the Fractions Brier Score (FBS) was calculated. The FBS for a forecast that matches (in location) the satellite data perfectly is 0. However, the scores obtained for non-perfect forecasts are dependent on the frequency of the precipitation event. As a result, Roberts and Lean (2008) proposed deriving the Fractions Skill Score (FSS) from the FBS. A perfect forecast would have a FSS of 1; scores close to 0 imply very little, or no skill in the ability of the forecast to replicate the satellite fractions.

One of the aims of this analysis is to investigate the extent of any spatial skill, not just at the grid cell level, but for a region surrounding the grid cell. Figure A7a is a schematic representation of the domain, the outer domain, and the location of a single grid cell (row=i, col=j). The domain represents the satellite and forecasts fields that were extracted for the Yemen domain as defined in Section 6.1. Every grid cell in the domain (represented by an x in Figure A7a), will have a value of either one or zero, depending on whether the rain amount was above a specified threshold or not. Values outside of the domain are assumed to have a value of 0 regardless of what was forecast or observed. Each grid cell within the domain has an associated neighbourhood. In the example in Figure A7a, the cell i,j (highlighted in yellow) has a neighbourhood consisting of a square made up of 25 grid cells (blue shaded cells), with the cell i,j located at the centre of the 5 x 5 grid square. Darker blue cells represent a rain event above the threshold, the fraction of rain events in the associated neighbourhood of cell i,j is 5/25 or 20 %. A field of precipitation fractions can be derived, if the proportion of precipitation events is calculated for each 5 x 5 neighbourhood that has a centre that falls within the domain. Separate precipitation fractions fields were derived for the satellite and the forecast data.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	**x** i	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0

											, j																		
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	X	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure A7a: Schematic scheme of outer domain, domain and neighbourhood.

It is not necessary to restrict the neighbourhood to 5 x 5 grid squares, any grid square (provided its length is odd, so that the centre of the neighbourhood sits directly on a grid square in the domain) may be considered. Typically, as the size of the length of the grid square increases from 1 up to a maximum of twice the longest side of the domain minus 1, the FSS increases until it reaches an asymptote. If the number of events exceeding the threshold for the satellite and forecast fields are the same, this asymptote will have a value of 1. However, should a bias be present then the FSS will reach its asymptote at values less than one. It is not possible, using this method, to determine whether the bias is in the forecast data, the satellite data, or both.

A plot of FSS plotted against the length of the neighbourhood square, can be used to identify the spatial scale at which the forecast becomes useful, FSS_{ufc}. Provided that the frequency of precipitation events in the domain is small (less than 20 %), the spatial scale at which FSS becomes greater than 0.5 can be used (Skok and Roberts 2016). Typically, when estimating the length of the neighbourhood that corresponds to the FSS_{ufc}, quantile thresholds are used to remove the influence of any precipitation bias between the forecast model and the observational data set used. However, the current analysis used fixed values, e.g. “take action” category (Section 6.1) as these could trigger (dependent on other non-meteorological risks) on the ground targeted action. Using absolute values allows for an assessment of the impact of False Alarms (when an event was forecast but didn’t occur) and Misses (when an event occurred but this was not forecast).

In this analysis any grid cell with its centre laying within the area shown in Figure A7a, is said to lie within the Yemen domain.

The equations used in this study to calculate FSS:

1. Use a threshold to convert precipitation fields, f, to binary fields

$$I_{field} = \begin{cases} 1 & f \geq q_{th} \\ 0 & f < q_{th} \end{cases}$$

2. Calculate the fraction of precipitation events, for each neighbourhood of size n, in the Yemen domain.

$$f(n)_{(i,j)} = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_{field} \left[i + k - 1 - \frac{n-1}{2}, j + l - 1 - \frac{n-1}{2} \right]$$

3. Calculate the Fractions Brier Score

$$\mathrm{FBS}_{(\mathrm{n})} = \frac{1}{\mathrm{N_x N_y}} \sum_{\mathrm{i=1}}^{\mathrm{N_x}} \sum_{\mathrm{j=1}}^{\mathrm{N_y}} \left[\mathrm{f(n)}_{(\mathrm{satellite,i,j})} - \mathrm{f(n)}_{(\mathrm{forecast,i,i})} \right]^2$$

4. Calculate the low skill FBS

$$\mathrm{FBS}_{(n,low)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} f(n)^2_{(satellit\ ,i,j)} + \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} f(n)^2_{(forecast,i,j)}$$

5. Calculate the FSS

$$FSS_{(n)} = 1 - \frac{FBS_{(n)}}{FBS_{(n,low)}}$$

Table 8 shows the number of days that were included in the FSS analysis, each month, after dry days were excluded.

Obs	2017	2018	2019
Neighbourhood size at which the specified bounds cross the $\mathbf{FSS_{ufc}}$	Count	Count	Count
Apr	5	20	17
May	19	19	16
Jun	7	11	11
Jul	21	28	17
Aug	28	26	26
Sep	21	15	
Oct	9	14	
Nov	4	9	

Table 8: The number of days included in the FSS analysis after dry days were excluded.

Annex 8: Further analysis of the FSS results

Plots were produced for all months between April and November (the wet season), and the results are summarised in Table 9 and Table 10 for Day 3 and 6, respectively. When the average FSS curve does not cross the FSS_{ufc} line, which is when the forecast could be considered useful, this is represented by a dash in the tables.

Table 9 and Table 10 show that the models have the greatest skill for July and August, i.e. they have the smallest neighbourhood sizes compared to the rest of the year. For day 3 accumulations, the forecasts become skilful when the monthly mean neighbourhood square has a length that is between 300 km (90 000 km2) and 800 km (640 000 km^2). To put this into perspective, Yemen has an estimated area of approximately 427 900 km^2. However, this comparison should be treated cautiously as the land mass of Yemen does not conform to a square, unlike the domain of this analysis. The effect of the domain size on the results is discussed earlier. By day 6, August has the most consistent skill with the length ranging between 300 km and 650 km. The difference in length at which the upper and lower bounds cross the FSS_{ufc} for July and August varies between 180 km and 1000 km.

Day 3	2017			2018			2019		
Neighbourhood size at which the specified bounds cross the FSS_{ufc}	upper 95 % CI bound	mean curve	lower 95 % CI bound	upper 95 % CI bound	mean curve	lower 95 % CI bound	upper 95 % CI bound	mean curve	lower 95 % CI bound
Apr	29 (319)	101 (1111)	-	21 (231)	45 (495)	83 (913)	39 (429)	95 (1045)	-
May	37 (407)	69 (759)	-	33 (383)	113 (1243)	-	35 (385)	55 (605)	121 (1331)
Jun	89 (979)	-	-	81 (891)	-	-	11 (121)	77 (847)	133 (1463)
Jul	29 (319)	75 (825)	-	19 (209)	31 (341)	47 (517)	21 (231)	41 (451)	69 (759)
Aug	33 (363)	49 (539)	71 (781)	19 (209)	35 (385)	65 (715)	41 (451)	57 (627)	81 (891)
Sep	-	-	-	71 (781)	-	-	-	-	-
Oct	135 (1485)	-	-	9 (99)	27 (297)	61 (671)	-	-	-
Nov	3 (33)	87 (957)	-	39 (429)	109 (1199)	-	-	-	-

Table 9: The useful Fractions Skill Score neighbourhood size in grid cells (km in brackets) for the mean curve, and the lower and upper bounds of the 95% confidence intervals for 2017, 2018 and 2019, for 24-hour accumulation for day 3 precipitation, excluding dry days.

Day 6	2017			2018			2019		
Neighbourhood size at which specified the specified bounds cross the FSS_{ufc}	upper 95 % CI bound	mean curve	lower 95 % CI bound	upper 95 % CI bound	mean curve	lower 95 % CI bound	upper 95 % CI bound	mean curve	lower 95 % CI bound
Apr	17 (187)	-	-	43 (473)	73 (803)	-	-	-	-
May	69 (759)	127 (1397)	-	-	-	-	29 (319)	61 (671)	-
Jun	75 (825)	-	-	125 (1375)	-	-	79 (869)	117 (1287)	-
Jul	71 (781)	-	-	21 (231)	47 (517)	109 (1199)	21 (231)	47 (517)	107 (1177)
Aug	39 (429)	57 (627)	99 (1089)	17 (187)	29 (319)	51 (561)	39 (429)	59 (649)	87 (957)
Sep	51 (561)	-	-	81 (891)	-	-	-	-	-
Oct	-	-	-	19 (209)	45 (495)	-	-	-	-
Nov	23 (253)	77 (847)	-	27 (297)	-	-	-	-	-

Table 10: The useful Fractional Skill Score neighbourhood size in grid cells (km in brackets) for the mean curve, and the lower and upper bounds of the 95 % confidence intervals for 2017, 2018 and 2019, for 24-hour accumulation for day 6 precipitation, excluding dry days.

Annex 9: Epidemiological statistical model

The WRA issued by the Met Office since April 2018 are used to anticipate locations where high daily precipitation accummulations may lead to increases in cholera levels, thus warranting targeted intervention. The objective of this epidemiological analysis is to answer two questions:

1. Are the forecast models in the WRA good predictors of the cholera risk in 2017? In 2017, the WRA were not provided to UNICEF and thus this provides a baseline for this analysis.
2. Were the actions taken based on the WRA effective at reducing cholera risk in 2018 and 2019?

To answer these questions, it is necessary to characterise the increases in cholera levels that warrant intervention. Thresholds in the WRA used to trigger action were: 20 mm/day for the CAM, and 10 mm/day for the GM (Section 6.1). Due to the lack of contextual information, it is not known at what cholera count or epidemiological threshold warrants action, independently of the WRA. It is therefore necessary to link the forecast precipitation to the cholera data available for 2017 to infer an epidemiological threshold that warrants a WASH intervention.

The derivation of the epidemiological threshold that warrants a WASH intervention is the objective of the first part of this analysis. To derive the epidemiological threshold, a mathematical formulation is proposed that models the relationship between the weekly number of new cholera cases and the precipitation forecasts. This mathematical model can then be used to calculate the expected number of weekly cholera counts for a given precipitation amount. These expected counts are compared to the actual values and the accuracy of the model is assessed using RMSE and bias statistics. The mathematical model is derived from cholera counts and precipitation forecasts using data from 2017.

The results from this analysis help provide an answer to question 1 above. To answer question 2, the mathematical model derived from 2017 is again used but with precipitation forecasts from 2018 and 2019. RMSE and biases are calculated for the 2018 and 2019 data and compared to the 2017 RMSE and biases.

In Yemen the precipitation forecasts are issued as an indicator of the risk of rising numbers of cholera cases. In 2017, prior to the WRA being issued, the number of cholera cases are assumed not to be affected by targeted WASH interventions, i.e. WASH interventions prior to 2018 will have been reactionary and based on epidemiological counts only. In 2018, it is assumed that the number of cholera cases will be reduced with the targeted WASH interventions due to the provision of the WRA allowing for preventative action.

The epidemiological data for 2017 is examined to determine whether the cholera and precipitation data are correlated. The investigation into whether there is a statistical link between the WRA and cholera data is limited to the day 1 forecasts, which is from the CAM. It is assumed that the statistical relationship does not, within the confidence intervals of the fitted model, depend on the choice of model (CAM or GM) or which forecast day is used.

To test whether the two datasets are correlated, the daily precipitation data are transformed into:

- weekly mean precipitation (measured in mm per day), to provide a measure of total precipitation over the week, calculated from accumulated daily forecasts (7 model runs);
- weekly maximum precipitation, i.e. the highest of the 7 daily precipitation values. This was investigated because the forecasts were used by examining the daily precipitation for the next 7 days and acting when any one of these 7 forecasts exceeded the threshold.

Quantile-Quantile (Q-Q) plots (not shown, created using the Statistical Package for Social Sciences (SPSS) software package) are used to determine if the variables, cholera count, mean and max precipitation, follow normal distributions. This was achieved by plotting the expected normal value against the observed value. If the data follow the diagonal (expected normal value equals observed value), then the variable follows a normal distribution. The Q-Q plots (not shown) show that neither cholera count nor precipitation are normally distributed and therefore the correlation between cholera count and precipitation is assessed using Spearman's (rank order) correlation. An investigation into the lag between precipitation and cholera count is presented in later.

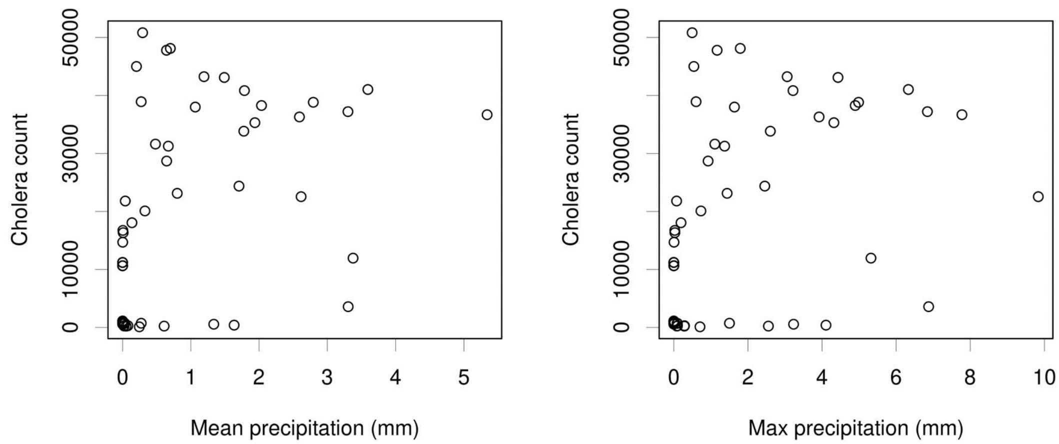


Figure A9a: Scatter plots of cholera count against mean (left) and max (right) precipitation.

In 2018 and 2019 preventive action was taken if the daily accumulated precipitation exceeded 20 mm for the CAM. As the epidemic threshold used prior to the provision of the CAM data is not known, the epidemic threshold will be derived from the relationship between cholera cases and precipitation in 2017. This relationship should take into account that a threshold of 20 mm of daily accumulated precipitation was used as a limit over which the number of cholera cases were expected to be of epidemic concern, due to using the CAM to investigate the statistical link between precipitation and cholera. No set definition for what would constitute an epidemic threshold was available prior to this study. Therefore, a model is sought that estimates the number of new cholera cases for an epidemic threshold, given the chosen precipitation threshold. This is limited by only have 1 year's epidemiological data to develop this model, ideally there would be at least 3 years of epidemiological data, but this was not available. With the wide scatter shown in the scatter plots (Figure A9a), the choice of model is only to (a) provide the means of estimating an epidemic threshold and (b) adjust for the non-normal distribution of precipitation. Furthermore, a model that transforms the precipitation into a linear risk index lying between 0 and 1 would be convenient. In consideration that:

- the distribution of precipitation is highly skewed (of the log-normal type);
- a threshold is used on daily precipitation to trigger action;
- this threshold could be used to define significant epidemic levels in terms of a cholera count threshold;
- both cholera count and precipitation variables have a natural lower bound at 0;
- it would be useful to transform precipitation into a risk index between 0 and 1 to relate to the counts and to help set the cholera count threshold;

the following non-linear model is proposed for estimating weekly cholera counts, $\hat{c}$:

$$\hat{c} = \frac{4E}{\pi} \tan^{-1} \frac{p}{P}$$

p is the maximum daily precipitation in a week,

P is the threshold of the maximum daily precipitation in a week ($P = 20$ mm has been used for the CAM),

E is the estimated cholera epidemic threshold.

The model presented here has the property that, if $p = P$, the CAM precipitation threshold, then $\hat{c} = E$ the epidemic threshold. The cholera risk index C can be defined as:

$$C = \frac{2}{\pi}\tan^{-1}\frac{p}{P}$$

This means that the cholera count estimate $\hat{c} = 2EC$ is a linear function of the risk index C with an epidemic threshold E that can be determined by linear regression. Also, it can be noted that $0 \leq C < 1$.

The parameter E is estimated by nonlinear regression (using SPSS's nonlinear regression module), with the constraint $E \geq 0$. A regression is carried out with the fixed precipitation threshold $P = 20$ mm to estimate the epidemic threshold E for the 20 mm threshold that was used for the CAM. This is equivalent to a simple linear regression of cholera counts against the risk index C (with the constraint that $E \geq 0$).

The performance of the precipitation forecasts with respect to their predicting cholera cases is assessed, through the nonlinear model, by computing for each year the root mean square error (RMSE) and the bias. The bias is derived from the weekly deviations of the predicted cholera counts ($\hat{c}=2EC$, where C is the cholera index) from the observed cholera count (c), $\hat{c}-c=2EC-c$, averaged over the year. The RMSE is derived from taking the square root of the mean of the square of the weekly deviations.

Both mean and max precipitation, shown in Figure A9b, are most strongly correlated with the cholera count with no lag, shown in Figure A9a, with coefficients of 0.51 for mean and 0.45 for max precipitation, although these are still considered weak correlations. The coefficients were found to be significant at the < 0.01 level (2-tailed test) with up to 3 weeks lag for mean precipitation and up to 1 week lag for max precipitation (from complementary analyses using SPSS). Furthermore, the gradual waning of correlation with increasing lag suggests strong temporal autocorrelation with a possible seasonal trend.

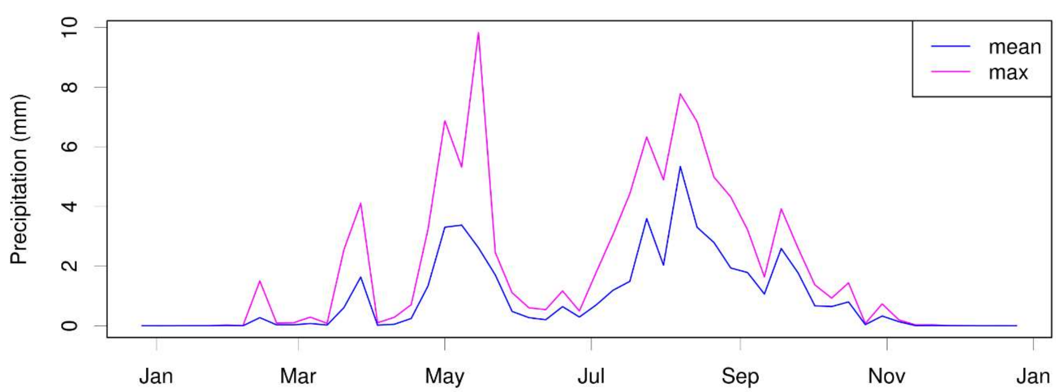


Figure A9b: Weekly mean and maximum precipitation for Yemen in 2017.

Due to using weekly country-level data, with only 53 data points over a single one-year period, it is not possible to examine seasonality, although it is known that there is a seasonality to the precipitation in Yemen. This limits the analysis, however, for this study the weekly data are assumed to be independent. This assumption could be addressed by having epidemiological data prior to 2017. Allowing for temporal and seasonal dependency could result in a higher correlation coefficient but would most likely reduce the statistical significance of the result, result due to the decrease in the number of data points. Also, the WRA may have been used in conjunction with additional 'local' knowledge from the WASH teams, such as air temperature, population density, known areas of conflict (thus potential damage to sanitation infrastructure) and previous cholera intervention, which cannot be taken into account in this analysis. Therefore, the temporal dependency and seasonality are not examined here as they form a small part of the assumptions of the relationship between precipitation and cholera incidence. The epidemiological data was not analysed at a governorate level due to time constraints.

Having established a statistically significant correlation between cholera counts and precipitation, with the strongest temporal correlation occurring between the precipitation in the same week as that of the new cholera cases, the relationship between counts and precipitation with no temporal lag is examined.

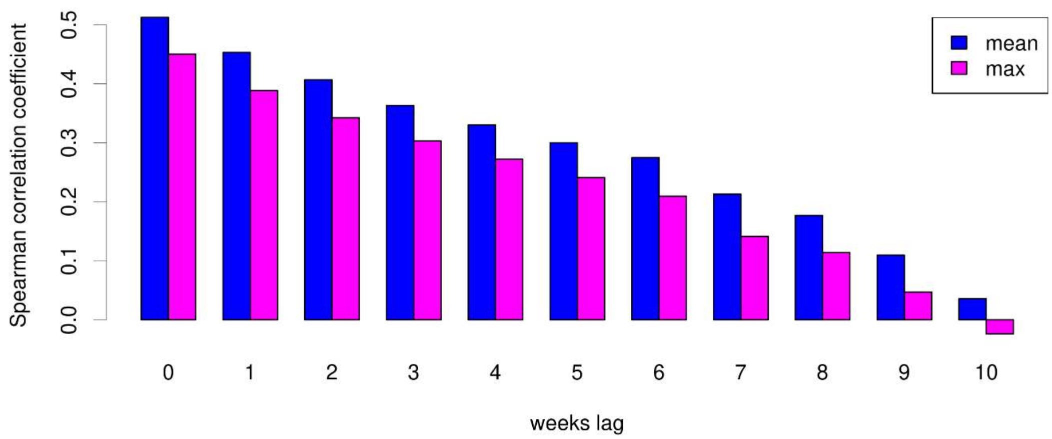


Figure A9c: Spearman's (rank order) correlation coefficients for country-level cholera counts against weekly mean and max precipitation with lags of 0 to 10 weeks. Coefficients are significant at the < 0.01 level up to 3 weeks and 1week lag for mean and max precipitation, respectively (additional analysis using SPSS).

By setting a precipitation threshold of P = 20 mm, the estimate for the epidemic threshold is E = 91 881 with bootstrap estimate of its 95 % trimmed range of [63 096;132 470]. The Pearson correlation coefficient is 0.39 (95 % Confidence Interval of 0.13 to 0.60) with a 0.004 significance level, obtained by examining the cholera counts against the computed cholera risk index C. The epidemic threshold of 91 881 is well above the highest cholera case count of ~ 50 000. This is consistent with the precipitation threshold of 20 mm that is in excess of the highest max precipitation value of ~ 10 mm. With such a precipitation threshold, for the whole country of Yemen at least, at no stage in 2017 did the weekly count of new cholera cases exceed the estimated epidemic threshold. Similar questions regarding the choice of the threshold are also relevant here, however the discussion presented earlier is also valid here.

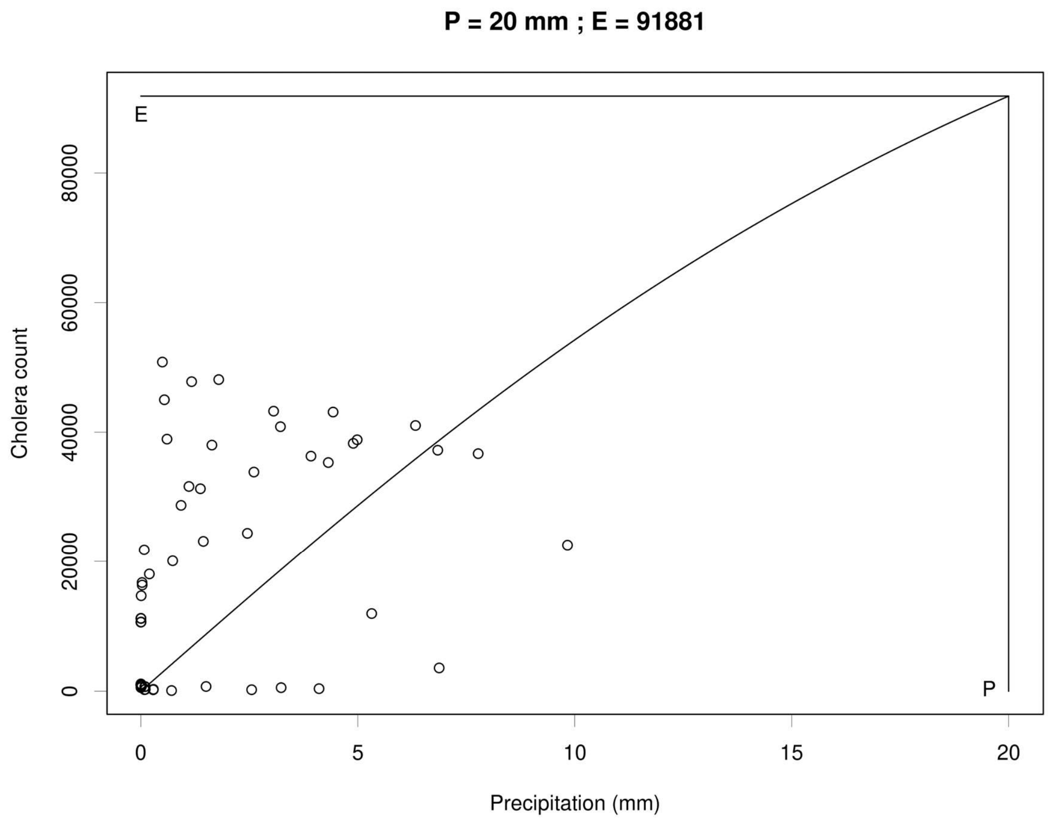


Figure A9d: Scatter plots with regression lines with precipitation threshold P = 20 mm. Vertical line indicates the precipitation threshold P; horizontal line indicates the epidemic threshold E.

Following the cholera epidemic in 2017 the number of weekly new cholera cases rose in the second half of 2018 to 15 000 cases, peaked again in March/April 2019 at over 30 000 cases and remained high with two secondary peaks at two-month intervals. The precipitation-based cholera risk index was computed for each week by taking the maximum risk index value from the individual forecast day risk values and using the forecast model dependent 20 mm and 10 mm thresholds. The threshold is represented by the risk index line C = 0.5 and the cholera risk index only exceeds this on two occasions in 2018, consistent with the previous observation that precipitation averaged over the whole of Yemen rarely reaches the threshold values used.

For 2017, the year before the forecasts were being used to help manage the epidemic in Yemen, the RMSE is relatively high at 20 000, shown in Figure A9e, and the regression has a Pearson correlation coefficients of 0.39 previously calculated from the nonlinear regression.

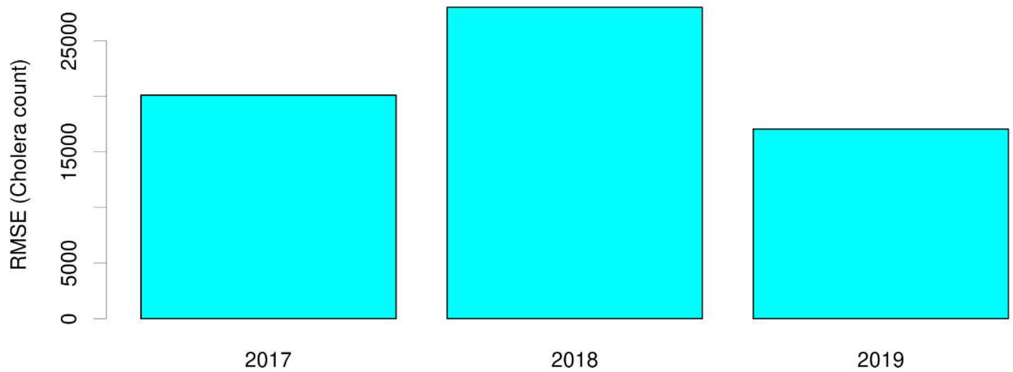


Figure A9e: RMSE in cholera count using fixed precipitation thresholds for 2017 to 2019.

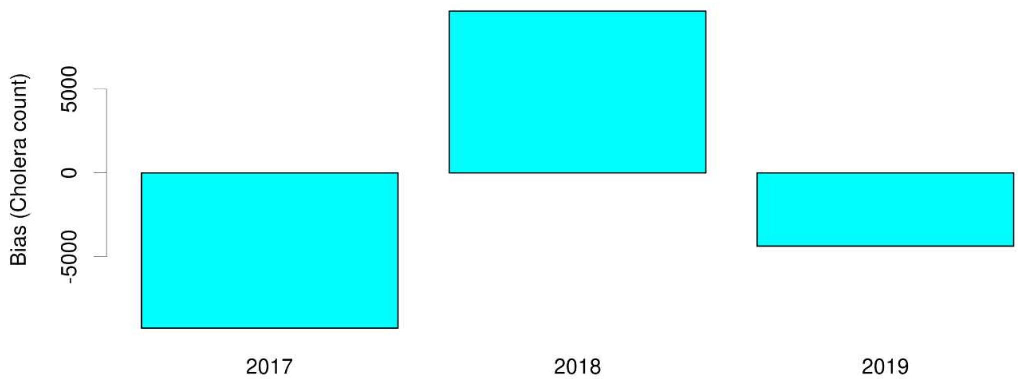


Figure A9f: Bias in cholera count using fixed precipitation thresholds for 2017 to 2019.

For 2017, the negative bias, shown in Figure A9f, means that new cholera cases are underestimated by up to 10 000 (on average) when using fixed 20 mm and 10 mm precipitation thresholds.

For 2018 and 2019, the results show a higher RMSE in 2018 with positive biases of up to 10 000. These results could be explained by a number of different factors that could either act independently or in combination:

- the precipitation forecasts were used effectively to prevent increases in number of cholera cases;
- an increase in the number of cholera cases was prevented due to factors other than the precipitation forecasts, e.g. through monitoring water supply microbiology;
- there is significant scatter in the risk index to cholera count relationship and the results are due to a random effect, natural to this type of data;
- the annual seasonality which has not been accounted for (only one year, 2017, before the forecasts was available).

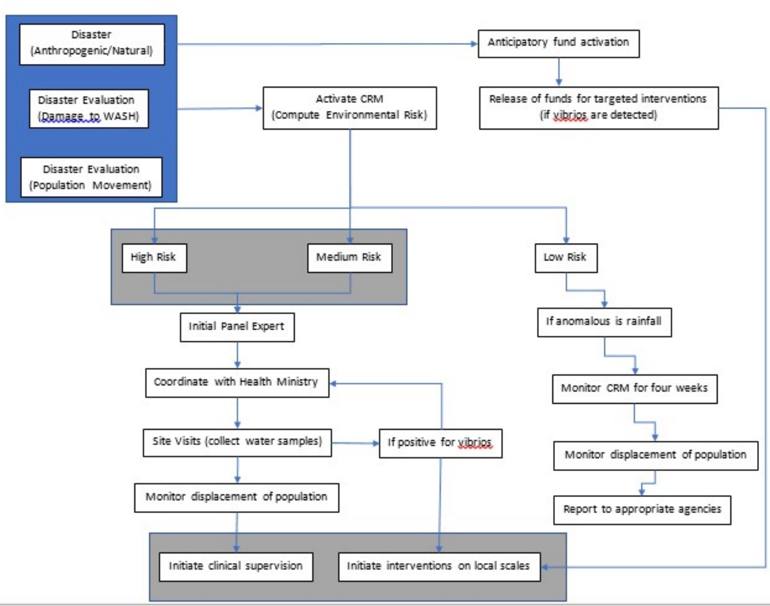
Having established a statistically significant spearman's correlation between cholera counts and precipitation (< 0.01 level of significance), the strongest correlation was found when there is no lag between precipitation and cholera counts. Thus, the relationship between counts and precipitation is examined assuming there is no lag. A positive bias implies that there are more predicted cases from the fitted models than were actual realised. If the reason for this positive bias is solely down to effective disease prevention through targeted interventions due to the WRA then a positive bias should be seen in 2018. Based on the RMSE results, the precipitation forecasts appear to be better at predicting cholera risk in 2019 (lowest RMSE) than in any other year.

Annex 10: Alternative approach to modelling cholera counts

The methodology and results described above come from just one approach to modelling the cholera counts. A different approach which draws upon statistical theory is to fit a generalized additive model (GAM). A GAM is an extension of the generalised linear model framework which is itself an extension of the linear model framework. A linear model assumes that the response variable, in this case cholera counts, is normally distributed and can be described by the sum of a set of linear predictor variables. However, typically, count data are not normally distributed. They either follow a Poisson distribution (where the mean is equal to the variance) or the negative binomial distribution. A generalised linear model assumes that the response data has a distribution which is a member of the exponential family (both Poisson and negative binomial distributions belong to the exponential family). The means of the response variable are then linked to a smoothed monotonic function of the predictor variables via a link function, (the default link function for the negative binomial distribution is and the Poisson distribution is the log function).

Results from fitting a Poisson distribution to the count data suggest that for the cholera data the mean is not equal to the variance, and so a negative binomial distribution would be arguably better. Preliminary investigations with such a model, analysis not shown, suggested that the response variables are not temporally independent. A GAM allows the count data to be modelled by the sum of smoothed functions of the predictor variables, rather than by the smoothed monotonic functions of the linear predictors as for the generalised linear model. As a result the seasonality in the data, such as that shown in the time series plots, can be incorporated via a dependency on a temporal variable, such as the week of the year. Whilst this GAM gave a better correlation to the 2017 cholera count, its predictions of the cholera count for 2018 were still poor. It is possible that this type of model could be improved upon, and one recommendation for further work would be to investigate the GAM approach further.

Annex 11: Example of a decision support tool for using the CRM



Annex 12: Water sampling protocols

University of Maryland (UMD) has previously developed and optimized an end-to-end protocol for sample collection, metagenomic sequencing, and bioinformatics. Accordingly, UMD will provide a training video and assist virtually with all aspects of sample collection and processing. At each sampling event, 1 L grab samples will be collected from surface water using virgin wide mouth bottles. Water samples will be concentrated via syringe filtration by passing 250 mL of water through a 0.22 μm pore size filter membrane, in triplicate, totalling three concentrations per sampling event. Each membrane will be placed in a separate 1 mL flip cap tube containing DNA/RNA Shield™ Lysis Buffer—this method of pathogen inactivation is compliant with Center for Disease Control guidelines for inactivation of infectious agents, including viruses, bacteria, fungi, and parasites (CDC, 2020). Lysed samples will be stored free of direct light at ambient temperature for up to one month or at <-20 °C for long term storage. Following international shipping regulations for liquids, samples will be packaged and shipped to UMD at ambient temperature. UMD will isolate DNA from each filter membrane, prepare DNA libraries, and perform whole-metagenome sequencing coupled with cloud-based bioinformatics, using techniques proven in both clinical and environmental settings (Brumfield et al., 2020; Connelly et al., 2019; Hamner et al., 2019; Hourigan et al., 2018; Kalan et al., 2019; Lax et al., 2012; Ponnusamy et al., 2016; Roy et al., 2018; Stamps et al., 2018) to achieve microbial (bacterial, archaeal, viral, fungal, and protozoan) identification to species, subspecies, and/or strain level and quantification of relative abundance. Analogously, antibiotic resistance and virulence-associated genes present in each sample will be identified. UMD will also employ conventional, direct molecular detection techniques as described previously (Huq et al., 2012).