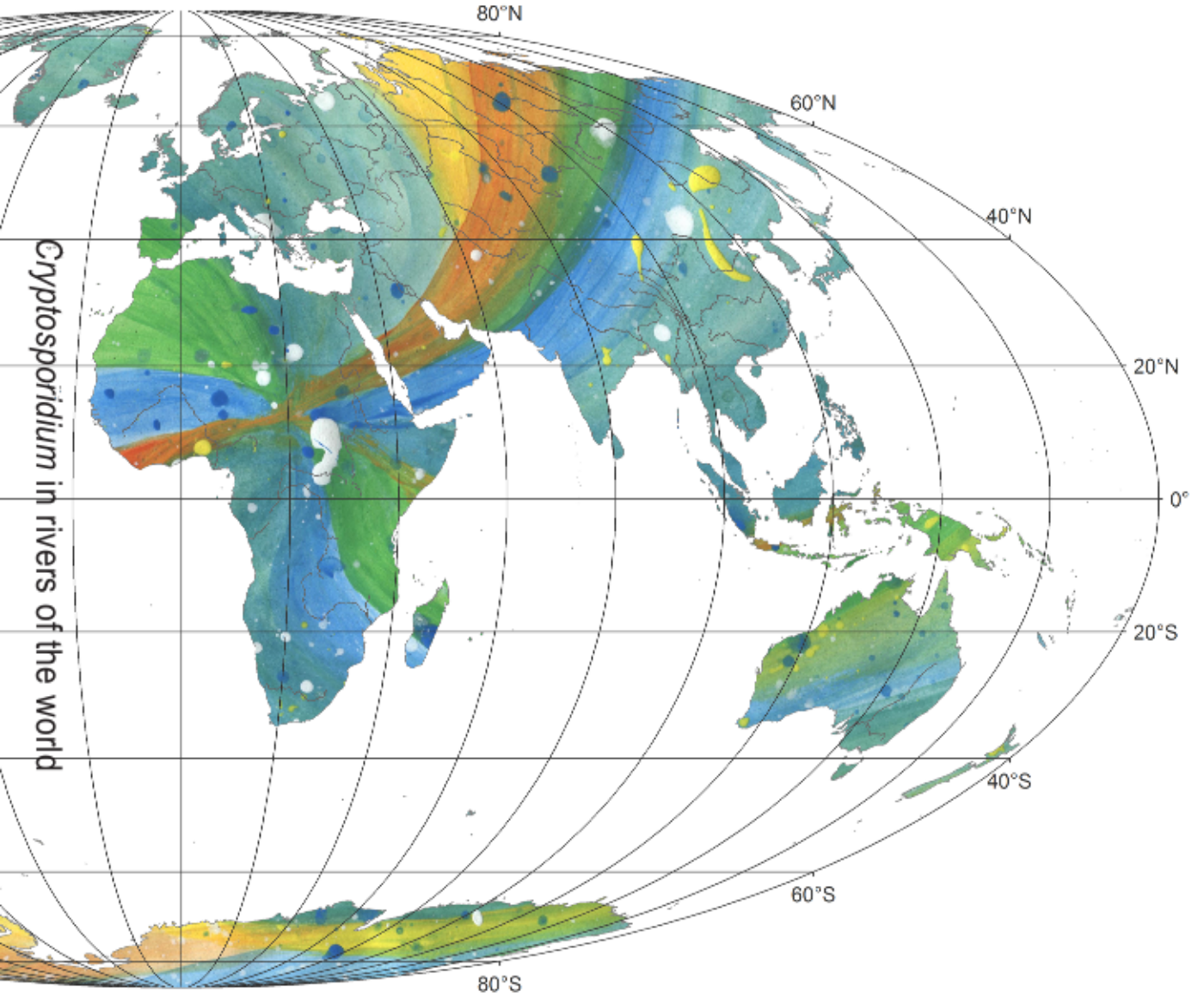


Cryptosporidium in rivers of the world the GloWPa-Crypto model

Lucie C. Vermeulen

Lucie C. Vermeulen

Propositions

1. For studying water pollutants at the global scale, models are better than measurements.
   (this thesis)

2. To combat diarrhoeal disease, wastewater treatment must be improved.
   (this thesis)

3. Process-based modelling and scenario analysis should get more attention in the study of environment-related health risks.

4. Cartography is a form of art: many different interpretations can be generated from the same raw material.

5. To meet the Sustainable Development Goals studying trade-offs and synergies between the goals is indispensable.

6. Life is like modelling: we aim to minimize uncertainty, but will never fully succeed.

7. Research should be approached like music, recognizing that a symphony is more than the sum of the parts.

Propositions belonging to the thesis, entitled:
"*Cryptosporidium* in rivers of the world - the GloWPa-Crypto model"

Lucie C. Vermeulen

Wageningen, 16th February 2018

Cryptosporidium in rivers of the world

the GloWPa-Crypto model

Lucie C. Vermeulen

Thesis committee

Promotors
Prof. Dr C. Kroeze
Professor of Water Systems and Global Change
Wageningen University & Research

Prof. Dr G.J. Medema
Chair on Water and Health
Delft University of Technology
Chief Science Officer
KWR Watercycle Research Institute, Nieuwegein

Co-promotor
Dr N. Hofstra
Assistant professor, Environmental Systems Analysis Group
Wageningen University & Research

Other members
Dr M. Flörke, University of Kassel, Germany
Prof. Dr A.A. Koelmans, Wageningen University & Research
Prof. Dr A.M. de Roda Husman, Utrecht University and RIVM, Bilthoven
Prof. Dr J.B. Rose, Michigan State University, East Lansing, USA

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE).

Cryptosporidium in rivers of the world

the GloWPa-Crypto model

Lucie C. Vermeulen

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 16 February 2018
at 1.30 p.m. in the Aula.

Lucie C. Vermeulen
Cryptosporidium in rivers of the world - the GloWPa-Crypto model,
238 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2018)
With references, with summary in English

ISBN 978-94-6343-820-9
DOI 10.18174/426782

Table of contents

Chapters 2, 3, 4 and 6 have been published as peer-reviewed scientific articles. Chapters 5 and 7 will be submitted for publication soon. The text, figures and tables of the published articles have been adjusted to the PhD thesis format. Editorial changes were made for reasons of uniformity of presentation. References should be made to the original articles.

Chapter 1

Introduction

1.1 Background

1.1.1 Diarrhoeal disease and water quality

Diarrhoeal disease is very common around the world. Diarrhoea is still the fourth leading cause of death among children under five years worldwide (United Nations Inter-agency Group for Child Mortality Estimation, 2015) and ranks sixth in global burden of disease among all ages (GBD 2013 DALYs and HALE Collaborators, 2015). Diarrhoeal disease can be caused by many different enteric pathogens, such as viruses (e.g. rotavirus, adenovirus), bacteria (e.g. certain strains of *Escherichia coli*, *Shigella* spp., *Campylobacter* spp.), or parasites (e.g. *Cryptosporidium* spp., *Giardia* spp.) (Lanata et al., 2013; Liu et al., 2016). Some of these pathogens do not only cause diarrhoea, but can also have more severe and sometimes chronic sequelae (Garg et al., 2006; Kirk et al., 2015). Enteric pathogens spread via the faecal-oral route, meaning that people can get infected when they ingest pathogens, commonly via water or food. The contamination of surface waters with faeces is an important pollution problem, and improvements in water, sanitation and hygiene (WASH) are needed to reduce the burden of diarrhoeal disease (Cairncross et al., 2010; Clasen et al., 2015). Knowing more about the global burden of disease caused by enteric waterborne pathogens and about the geographical distribution of pollution is important for human health decision making and water and sanitation planning.

This thesis specifically focusses on the protozoan parasite *Cryptosporidium*, for a number of reasons: 1) *Cryptosporidium* is one of the most common and important diarrhoeal pathogens worldwide, causing a large disease burden and mortality especially among children in developing countries and among immunocompromised people such as HIV/AIDS patients (Lanata et al., 2013; Liu et al., 2016; Squire and Ryan, 2017). It is associated with malnutrition and growth faltering (Squire and Ryan, 2017). 2) *Cryptosporidium* is a leading cause for waterborne disease outbreaks in industrialized countries (Efstratiou et al., 2017b; Karanis et al., 2007). 3) *Cryptosporidium* is considered a reference pathogen for enteric protozoan pathogens (Medema et al., 2009). Reference pathogens are chosen based on their disease burden and the difficulty to remove them via treatment, meaning that control of the reference pathogen would also indicate that other, similar pathogens are sufficiently controlled (Medema et al., 2009). 4) Cryptosporidiosis is

substantially underreported due to a lack of simple inexpensive diagnostic tools (Checkley et al., 2015; Shirley et al., 2012). 5) There is a lack of effective treatment for cryptosporidiosis, and currently no vaccine exists (Checkley et al., 2015; Shirley et al., 2012). 6) Oocysts, the robust survival stage of *Cryptosporidium*, can survive in the environment for months and are resistant to disinfection (Nasser, 2016; Peng et al., 2008). 7) *Cryptosporidium* is zoonotic, meaning that it can infect humans and a variety of animal hosts, including important livestock species such as cattle (Robertson et al., 2014; Ryan et al., 2014). Environmental transmission between humans and animals is common (Squire and Ryan, 2017). 8) Knowledge on sources, fate and transport of *Cryptosporidium* in the environment is relatively good (see section 1.1.4 for an overview), making it a suitable target for a first model at the global scale (Hofstra et al., 2013).

The global burden of disease of cryptosporidiosis is difficult to estimate and very uncertain, due to the aforementioned lack of diagnostics and underreporting (Checkley et al., 2015; Shirley et al., 2012). Global burden of disease estimates mostly rely on national public health statistics, large scale surveys and epidemiological data analysis (GBD 2013 Mortality and Causes of Death Collaborators, 2015). Quantitative microbial risk assessment (QMRA) is a useful tool to calculate burden of disease based on environmental exposure, such as waterborne or foodborne exposure, rather than based on health statistics. QMRA calculates risk of infection from exposure to a certain dose of pathogens and an organism-specific dose-response curve (Haas et al., 1999). From infection risk the resulting burden of disease can be calculated. QMRA can also help inform what level of water treatment is needed to meet a health-based target, and it can aid in decision making on investments in the water system (Howard et al., 2006; World Health Organization, 2017a). Infection with enteric pathogens can occur via direct contact with contaminated faeces (Zambrano et al., 2014), but often transmission occurs via an environmental route, in which rivers play an important role (Cacciò and Widmer, 2014). In the case of waterborne transmission, the exposure depends on the concentration of the pathogen in the ingested water.

Data on the concentration of pathogens in rivers are very scarce, as monitoring is time-consuming and expensive, with for *Cryptosporidium* a typical cost of $250-$400 US per 10L sample (Efstratiou et al., 2017a). Often faecal indicator organisms (FIOs), such as the bacterium *E. coli*, are monitored as a proxy for pathogens, but FIOs and pathogens do not necessarily correlate well (Committee on Indicators for Waterborne Pathogens National Research Council, 2004; Payment and Locas, 2011). Several industrialized countries do perform extensive waterborne pathogen monitoring for a period of time, mostly at source water intakes for drinking water production, an example is the *Cryptosporidium* monitoring as part of the Long Term 2 Enhanced Surface Water Treatment Rule (LT2) in the USA (Ongerth, 2013). However, typically, monitoring data are collected by water

authorities, health or environmental laboratories or drinking water utilities, and are not publically available. Furthermore, as nearly all available monitoring data are from industrialized countries, a large knowledge gap for the developing world exists (Efstratiou et al., 2017a). As the developing world is also most severely affected by diarrhoeal disease, this is an important shortcoming.

Some initiatives do exist that aim to cluster data and knowledge, such as the Global Environment Monitoring System (GEMS/Water) initiative of UN Environment, but most microbiological data in the GEMS database are faecal indicator data. Another initiative is the new Global Water Pathogen Project (http://www.waterpathogens.org/) that aims to create a data and knowledge hub and update the current benchmark reference work of Feachem et al. (1983). None of these initiatives currently provide accessible comprehensive quantifications of waterborne pathogens around the world. It is essential that such quantifications of pathogen pollution in surface waters become available to be able to evaluate health risk, and guide water and sanitation planning to the most efficient measures to control the health burden. Large scale water quality modelling can be a helpful tool to fill the gaps of regions with poor data, and help to identify where and when monitoring would be most needed. Modelling is based on understanding of the sources, pathways and processes of pathogens to and in rivers, and understanding of the effect of control measures, such as wastewater or manure treatment. Such understanding helps to design effective control strategies.

1.1.2 Water quality in the Sustainable Development Goals

The control of faecal water pollution is receiving increasing attention around the world. For example, water, sanitation and hygiene are strongly embedded in the Sustainable Development Goals (SDGs). The SDGs are 17 goals divided in 169 targets addressing different aspects of development that the world aims to reach in 2030 (United Nations General Assembly, 2015). Especially goal 6 ‘Ensure availability and sustainable management of water and sanitation for all’, but also other goals and targets contain elements that link to water quality (listed in Box 1).

The Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP) published SDG baseline data on the current status of drinking water, sanitation and hygiene, showing the extent of the challenge that the world faces (WHO/UNICEF JMP, 2017). The JMP drinking water SDG baseline (target 6.1) found that in 2015, 29% of the world population did not have access to a safely managed drinking water source, and an estimated 2% drink surface water directly. Service levels are generally better in urban areas than in rural areas. In 22 countries over 10% of the population drink surface water directly, these are mainly located in sub-Saharan Africa (WHO/UNICEF JMP, 2017). The

Box 1: Waterborne pathogens and the SDGs

Goal 6 of the SDGs is 'Ensure availability and sustainable management of water and sanitation for all'. Targets belonging to this goal with the indicators used to assess progress include:

- Target 6.1: 'By 2030, achieve universal and equitable access to safe and affordable drinking water for all', with indicator 6.1.1: 'Proportion of population using safely managed drinking water services'.
- Target 6.2: 'By 2030, achieve access to adequate and equitable sanitation and hygiene for all and end open defecation, paying special attention to the needs of women and girls and those in vulnerable situations', with indicator 6.2.1: 'Proportion of population using safely managed sanitation services, including a hand-washing facility with soap and water'.
- Target 6.3: 'By 2030, improve water quality by reducing pollution, eliminating dumping and minimizing release of hazardous chemicals and materials, halving the proportion of untreated wastewater and substantially increasing recycling and safe reuse globally', with indicator 6.3.1 'Proportion of wastewater safely treated' and indicator 6.3.2 'Proportion of bodies of water with good ambient water quality'.

Other targets relevant to waterborne pathogens include target 1.4 on universal access to basic services, target 2.1 on safe food and food security, target 3.3 which mentions waterborne diseases, target 3.9 on the reduction of the disease burden from inadequate water, sanitation and hygiene (WASH), and target 4a on access to basic WASH in schools (WHO/UNICEF JMP, 2017).

JMP sanitation SDG baseline (target 6.2) found that in 2015, 61% of the world population did not have access to safely managed sanitation, and 12% still practiced open defecation (WHO/UNICEF JMP, 2017). Sanitation coverage is worst and open defecation most practiced in sub-Saharan Africa, Oceania (excluding Australia and New Zealand) and Central and Southern Asia. Marked differences exist between rich and poor populations, even within individual countries. Between 2000 and 2015 the number of people practicing open defecation has decreased, but in order to reach the SDG targets of ending open defecation, universal access to basic sanitation, and halving the proportion of untreated wastewater by 2030, progress will need to speed up considerably (WHO/UNICEF JMP, 2017). Improvements in WASH are required to effectively reduce the burden of diarrheal diseases (Cairncross et al., 2010; Clasen et al., 2015).

SDG indicator 6.3.2 'Proportion of bodies of water with good ambient water quality' is the indicator most directly related to surface water quality. To judge progress on this target, a large amount of data on different water quality variables around the world is required. The UN Environment Global Environment Monitoring System (GEMS/Water) is the responsible agency to assess a baseline and progress on indicator 6.3.2, a process which is currently underway. Unfortunately, no microbial variable was included in the baseline monitoring set for indicator 6.3.2, meaning that this initiative will likely not increase data availability on the occurrence of microbial pollution in the environment, which would have been relevant for combating diarrhoeal disease. Large scale spatially explicit modelling could be a key ingredient to assess progress on SDG indicator 6.3.2, for variables that have been included in the baseline set, but also to extrapolate to other variables, such as pathogens. Links between water quality and the SDGs can go in different directions; efforts to meet the SDGs can improve or deteriorate water quality, and water quality can affect progress towards certain targets. Modelling can be a tool to explore these links, and help to identify possible trade-offs and synergies. For example, if open defecation is ended (6.2) and all of these people are connected to sewage systems transporting the waste to rivers but only half of the wastewater is treated (6.3.1), what will be the effect on ambient water quality (6.3.2) and the resulting effect on waterborne disease occurrence (3.9) and food security (2.1)? Large scale water quality modelling can be a tool for such explorations.

1.1.3 Environmental transmission of enteric pathogens

As previously indicated, enteric pathogens like *Cryptosporidium* spread via the faecal-oral route. In this section the different sources and environmental transmission routes are described in more detail.

Cryptosporidium is in fact a genus consisting of over 25 species, of which 20 have been reported in humans (Ryan et al., 2014). Not all are equally infectious, the majority of human infections are caused by *C. hominis* (which infects predominantly humans) and *C. parvum* (which infects a variety of mammals, especially ruminant livestock is an important reservoir) (Ryan et al., 2014). Genotyping studies of *Cryptosporidium* spp. found in humans and animals indicate that environmental transmission is common (Squire and Ryan, 2017). Knowing more about environmental transmission of cryptosporidiosis is thus important for human health, but also for the livestock production sector, as it can cause diminished production and loss of income (Shafiq et al., 2015; Sweeny et al., 2011).

Oocysts are shed in faeces of infected humans and animals in high numbers (e.g. up to 10^9 oocysts per ill person and up to 10^9 oocysts/gram manure have been observed) (Chappell et al., 1996; de Waele et al., 2011). Sources of water pollution are commonly categorized

as point or nonpoint sources, the latter are also referred to as diffuse sources. For *Cryptosporidium,* point sources include the discharge of (treated and untreated) wastewater to rivers, and a typical example of a nonpoint source is animal manure on fields that can be transported to rivers with runoff (Atwill et al., 2006b; Davies et al., 2004). Human faeces on land, from the population practicing open defecation, is also considered a nonpoint source. On-site sanitation systems, such as septic tanks and pit latrines, could also contribute to the environmental load when these systems overflow or their contents are discharged unsafely (Berendes et al., 2017). Wastewater treatment can reduce oocyst numbers, but not all types of treatment are equally effective (Nasser, 2016). Oocysts from human faeces and animal manure can thus end up in rivers through a number of different pathways. The relative importance of the different sources will differ per location, and is often not known. Figure 1.1 shows a schematic representation of the main sources and environmental transmission of *Cryptosporidium.*

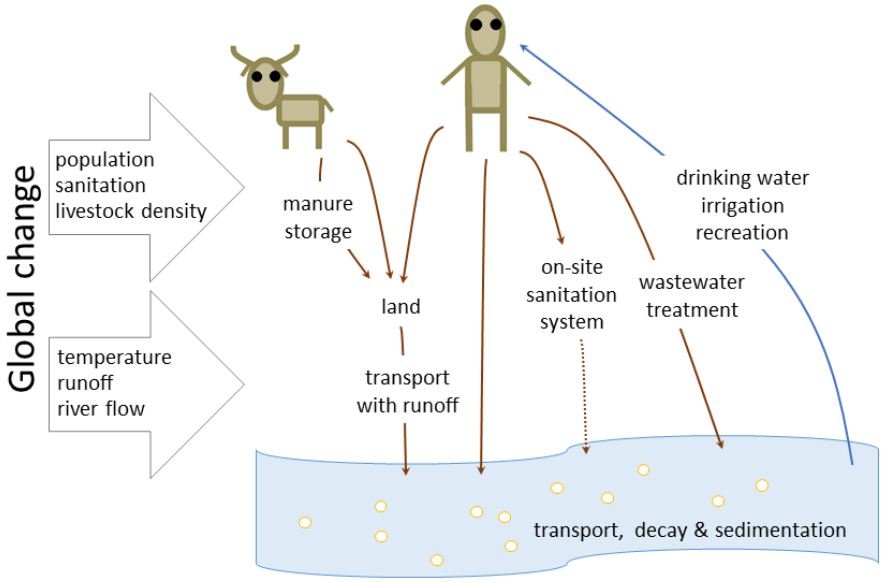


Figure 1.1: Schematic representation of the sources and environmental transmission of *Cryptosporidium* and the impacts of global change on this system. Brown arrows show faecal waste streams from animals and humans. Animal manure can be excreted directly on land during grazing or applied to land after storage, and from there be transported with runoff to rivers. Human faeces can be excreted on land during open defecation, can go directly to rivers (for example from hanging toilets), on-site sanitation systems (such as septic tanks and pit latrines) may also contribute to the environmental load when these systems overflow or their contents are discharged unsafely, and lastly human faeces is transported to rivers via sewers (treated or untreated). The blue arrow indicates the ingestion routes that can lead to infection. The large arrows on the left indicate impacts of global change.

Cryptosporidium is a very persistent pathogen that can survive for months in the environment (Monis et al., 2014; Peng et al., 2008). Nevertheless, oocysts decay does take place in the environment; survival is influenced by a number of different factors, such as ambient temperature and solar radiation (Peng et al., 2008). Furthermore, processes like sedimentation and resuspension from sediments can influence oocysts concentrations in rivers (Monis et al., 2014). *Cryptosporidium* cannot reproduce outside of a human or animal host, so growth in the environment does not take place.

Infection with *Cryptosporidium* can occur when contaminated water is ingested, for example during drinking, recreation or consumption of irrigated vegetables (Cacciò and Widmer, 2014; Dixon, 2016; Shirley et al., 2012). Ingestion of a dose as low as 10-30 oocysts is reported to cause infection in 20-66% of healthy adults (Chappell et al., 2006; DuPont et al., 1995). Comparing this to the high numbers that can be found in faeces of infected individuals makes it clear that for effective control of cryptosporidiosis, it is important that the different sources of oocysts and the transmission via rivers are studied.

1.1.4 Effects of global change on pathogens in rivers

Global change processes are likely to impact enteric waterborne pathogens in various ways. Figure 1.1 schematic depicts main sources and environmental transmission of *Cryptosporidium* and the impacts of some important elements of global change on this system. Global change is an umbrella term that encompasses planetary-scale changes in the Earth System, both physical (e.g. climatic changes, changes in biogeochemical cycles) and societal (e.g. changes in population or resource use) (Steffen et al., 2004). Climate change can have multiple impacts on enteric waterborne pathogens (Hofstra, 2011; Rose et al., 2001) and it is expected that risk of water- and foodborne disease will increase under global change (Smith et al., 2014). Increased (extreme) precipitation and runoff may lead to increasing loads of pathogens to rivers, changes in river discharge influence the dilution of pathogen concentrations, and changes in air and water temperature can affect survival rates in the environment (Hofstra, 2011; Rose et al., 2001). Human population growth, urbanisation, and changes in sanitation system use and wastewater treatment will affect the loads of waterborne pathogens from human faeces. Changes in livestock production as a response to changes in food demand, and management of manure will affect the loads of zoonotic waterborne pathogens. Qualitatively, these effects are apparent, but little quantitative information is available (Hofstra, 2011). The system is even more complex when we also consider the water use side. Climate change may alter water use, for example increasing irrigation water demand and river recreation in regions with higher temperatures, thereby potentially increasing human exposure to pathogens in rivers. On the other hand, if drinking water treatment is improved, this will decrease risk of infection via this route. A complex interplay of climate effects on food security,

resulting nutritional status and susceptibility to disease has also been envisioned, especially for the world's poorest regions, such as sub-Saharan Africa (Squire and Ryan, 2017).

There is a need for systems approaches to planetary-scale environment-related health problems (Pongsiri et al., 2017). Tools are needed that can quantify relations and allow for scenario analysis to explore the future and different management options. Spatially explicit models at different scales and resolutions can be part of such tools (Hofstra, 2011). The resulting maps can be helpful for decision makers by providing a visual overview. Modelling at the global scale is especially relevant in the context of studying the effects of global change on pathogen pollution in rivers in different world regions. Global waterborne pathogen modelling is the application of 'systems thinking' to a human health problem, in a way that is common in the environmental sciences but relatively new to the water and health field, where traditionally epidemiology is the standard.

1.2 Knowledge gaps and existing approaches

From the above, the following knowledge gaps have been identified to be addressed in this thesis:

- A lack of quantifications of *Cryptosporidium* concentrations in rivers worldwide that can be input for understanding the health impact of water use (via QMRA and disease burden studies)
- A lack of insight in the magnitude and relative importance of different sources of *Cryptosporidium* spp. to rivers that can help design effective control strategies
- A lack of tools to study the impact of global environmental change and management strategies on enteric waterborne pathogens in rivers around the world
- A lack of informative maps on waterborne pathogens that are useful for water and sanitation planning and can inform progress towards the Sustainable Development Goals
- A lack of application of 'systems thinking' to enteric pathogens in rivers, as is common in the environmental sciences. Specifically, a lack of application of large scale modelling to waterborne pathogens, as is already done for other water pollutants.

Models of substances in water are useful tools to study sources, transport, and concentrations of pollutants, and the demand for models of pathogens and faecal indicator organisms (FIOs) to inform water quality decision making is increasing (Cho et al., 2016; Oliver et al., 2016). Most commonly, water quality models estimate pollutant loads to rivers and combine these with an existing hydrological model to simulate environmental transport, while accounting for in-stream losses. The level of understanding of the environmental behaviour of waterborne pathogens and FIOs is relatively low compared to other pollutants, such as nutrients (Kay et al., 2008). Existing models of waterborne pathogens and FIOs are mostly limited to the watershed scale, with the exception of the WorldQual model that has been applied for faecal coliforms at the continental scale (Reder et al., 2017, 2015). Global models have been constructed for various other water pollutants, for example nutrients (such as nitrogen and phosphorus) (Beusen et al., 2015b; Seitzinger et al., 2010), silica (Beusen et al., 2009), sediments (Syvitski et al., 2005), insecticides (Ippolito et al., 2015), mercury (Amos et al., 2014), radionuclides (Smith et al., 2004) and recently for plastics (Lebreton et al., 2017; Siegfried et al., 2017).

There is a clear opportunity for the construction of a process-based waterborne pathogen concentrations model at the global scale. Such a model can provide information on pathogen concentrations in data-sparse regions, identify 'hotspot' regions with high concentrations, identify the relative contribution of different sources, and allow for evaluating scenarios (for example, to investigate the impact of management strategies or global change). A process-based model can be defined as a model that mathematically describes processes occurring in a system, as opposed to an empirical model that statistically relates an output variable to one or more observed input variables. Often 'mechanistic' is used synonymously to process-based. The outcome of a global waterborne pathogen model may be used as input for QMRA, and as a basis for policies aimed at reducing risk. A model at the global scale will have a relatively coarse resolution, and is therefore not suitable to exactly reproduce local observations, but rather to get an overview, e.g. of hotspots and relative change.

In this thesis, I present the Global Waterborne Pathogen model for *Cryptosporidium* (GloWPa-Crypto), the first global model of waterborne pathogen emissions to and concentrations in rivers, implemented for *Cryptosporidium*. This model builds on an exploratory model of *Cryptosporidium* emissions to surface waters (Hofstra et al., 2013).

1.3 Objectives and outline of the thesis

The objective of this thesis is to increase our knowledge on the sources, fate and transport of *Cryptosporidium* in rivers worldwide using spatially explicit modelling. This objective is divided in five sub-objectives:

- Obj.1: To learn from existing watershed-scale waterborne pathogen models and global nutrient export models for the development of a global waterborne pathogen model. (Chapter 2)
- Obj. 2: To develop a spatially explicit model to study the impact of sanitation changes, urbanisation and population growth on human emissions of *Cryptosporidium* to surface waters. (Chapters 3 and 4)
- Obj. 3: To collate and summarise published data on prevalence of *Cryptosporidium* infection in common livestock species and oocyst concentrations in livestock manure from the literature and to analyse their distribution by age and region. (Chapter 5)
- Obj. 4: To quantify livestock *Cryptosporidium* spp. loads to land on a global scale using spatially explicit process-based modelling, and to explore the effect of manure storage and treatment on oocyst loads using scenario analysis. (Chapter 6)
- Obj. 5: To estimate *Cryptosporidium* oocyst concentrations in rivers worldwide. (Chapter 7)

The five objectives are addressed in six scientific papers, presented in Chapters 2-7. Figure 1.2 shows an overview of the research presented in this thesis.

To meet the first objective, the literature on existing models of waterborne pathogens and global nutrient export was reviewed (Chapter 2). The different approaches were summarised and desirable characteristics for a global waterborne pathogen model were identified. A conceptual global waterborne pathogen model is presented.

For the second objective, initially a case study model was developed for human *Cryptosporidium* emissions in Bangladesh and India (Chapter 3). This model does not only include emissions from sewage systems, as was the case in the exploratory model (Hofstra et al., 2013), but also considers other faecal waste streams such as from open defecation and hanging toilets. In a scenario analysis for 2050, 'business as usual' is contrasted with sanitation improvement. Model performance is assessed in a sensitivity analysis. Next, the human emissions model was further developed for the world: GloWPa-Crypto H1 (Chapter 4). A scenario analysis for 2050 was performed using the Shared Socioeconomic Pathways (SSPs) (O'Neill et al., 2017).

To address the third objective, the literature on prevalence of *Cryptosporidium* infection and oocyst concentrations in manure from 11 livestock species was summarised (Chapter 5). Data were collected from >300 scientific studies, and analysed by age groups and world region. The results are input for the livestock loads model.

To meet the fourth objective, a model of *Cryptosporidium* loads to land from livestock manure was developed, GloWPa-Crypto L1 (Chapter 6). Livestock density maps are combined with prevalence and oocyst concentration estimates from the literature inventory in Chapter 5. Temperature-dependent oocyst decay during manure storage is accounted for. In a management scenario analysis, the effect manure storage and treatment can have on oocyst loads to land is investigated. Model performance is assessed in a sensitivity analysis.

To address the fifth objective, a model of global *Cryptosporidium* concentrations in rivers was developed, GloWPa-Crypto C1 (Chapter 7). The output of the human and animal load models (GloWPa-Crypto H1 and L1) was combined with hydrological output from the Variable Infiltration Capacity (VIC) model. Oocyst are routed through the stream network, and temperature- and solar radiation-dependent decay and loss due to sedimentation are accounted for. Model performance is assessed using sensitivity analysis and by validation with observational data.

Lastly, the methods and results are discussed in context, including a first analysis of risk and burden of disease based on the GloWPa-Crypto model, and final conclusions are drawn in the synthesis chapter (Chapter 8).

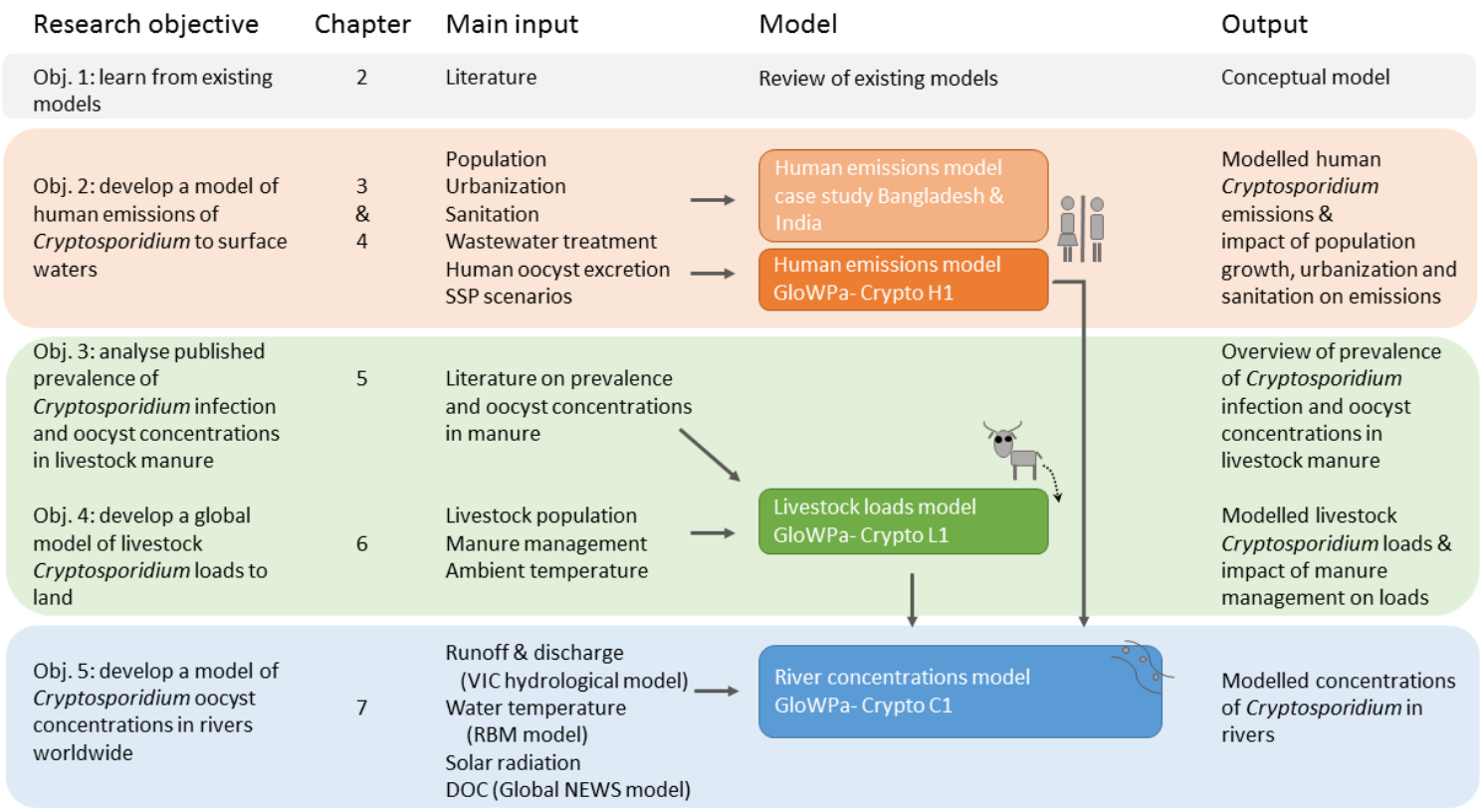


Figure 1.2: Schematic overview of the research presented in this thesis. This figure shows the research objectives, corresponding chapters, main input, the models that have been developed as part of this thesis and their main output.

Chapter 2

Advancing waterborne pathogen modelling: lessons from global nutrient export models

Lucie C. Vermeulen, Nynke Hofstra, Carolien Kroeze and Gertjan Medema

Abstract

Waterborne pathogens cause health problems worldwide. A global waterborne pathogen model could provide valuable new insights for data-sparse regions, by identifying pathogen hotspots and evaluating global change and risk management scenarios. Global waterborne pathogen modelling is not as advanced as global modelling of other water pollutants like nutrients. Our objective is to learn from existing watershed-scale waterborne pathogen models and global nutrient export models for the development of a global waterborne pathogen model. We describe how desirable characteristics of a global waterborne pathogen model are represented in current models of waterborne pathogens and river nutrient export. We present a conceptual global waterborne pathogen model, combining the existing knowledge on nutrient and pathogen models and identify data needs for such a model. Our review paves the way for an exciting new opportunity in waterborne pathogen modelling.

This chapter is based on: Vermeulen, L.C., Hofstra, N., Kroeze, C., Medema, G.J., 2015. Advancing waterborne pathogen modelling: lessons from global nutrient export models. Curr. Opin. Environ. Sustain. 14, 109–120. doi:10.1016/j.cosust.2015.05.003

2.1 Introduction

Global water quality models are useful tools to analyse sources, transport and concentrations of substances in surface water. Examples of such models include global models for river transport of nutrients such as nitrogen (N) and phosphorus (P) (Seitzinger et al., 2010), silica (Si) (Beusen et al., 2009), sediments (Syvitski et al., 2005), heavy metals (Amos et al., 2014; Smith et al., 2004) and persistent organic pollutants such as pesticides (Ippolito et al., 2015; Wania and Mackay, 1995). Waterborne pathogens are also transported by rivers and reduce the water quality. However, global waterborne pathogen models do not exist yet. Existing models of waterborne pathogens and faecal indicator organisms (FIOs) are mostly limited to specific watersheds or smaller areas (de Brauwere et al., 2014), with some exceptions at the level of states (Wickham et al., 2006), country (Medema and Schijven, 2001) and one recent model at the continental scale (Reder et al., 2015). Clearly, pathogen and FIO modelling is underdeveloped compared to that of other water pollutants, such as nutrients (Kay et al., 2008). Nutrients and pathogens in rivers have common sources, including sewage and livestock manure. Therefore, applying frameworks and methodologies from river nutrient export modelling can advance the field of waterborne pathogen modelling. Various models of global river nutrient export have been developed over the past decades (Kroeze et al., 2012), as well as numerous models at smaller scales (Grizzetti et al., 2003; Pärn et al., 2012).

Waterborne enteric pathogens are a major cause of diarrhoea, and diarrhoea is the second most important cause of infant deaths worldwide (UNICEF/WHO, 2009). Problems are largest in developing countries, but outbreaks of waterborne disease also regularly occur in industrialized countries (Karanis et al., 2007). Infection from contaminated water can be caused by a variety of organisms, including viruses, bacteria and protozoa (UNICEF/WHO, 2009). Some organisms, such as *Cryptosporidium*, are zoonotic, which means that animals can act as a reservoir of human pathogens (Fayer, 2004). Direct contact with faeces of infected humans or animals can lead to disease transmission, but often contaminated water is the transmission route. Contaminated water can be ingested as drinking water, during recreation or via irrigated crops (UNICEF/WHO, 2009).

The sources of pathogens in rivers are known, but poorly studied. Human waste is an important point source of pathogens in surface waters. It enters rivers through treated and untreated sewage. Nonpoint or diffuse sources of zoonotic waterborne pathogens include animal faeces on land (e.g. manure from livestock and wildlife excreta) that can enter surface water with runoff. Observational data of pathogens in surface waters are scarce or non-existing for many world regions, because monitoring is expensive. Modelling can, therefore, be useful for regions where problems are large, but data are lacking (Eisenberg et al., 2002; Ferguson and Kay, 2012).

Nutrient cycling is disturbed by human activities, leading to increased risks for eutrophication, hypoxia and harmful algal blooms (Bouwman et al., 2012; Seitzinger et al., 2010). The sources and flow pathways of nutrients (N and P) and waterborne pathogens overlap (Figure 2.1). Surface and subsurface runoff transport manure from fields to surface water, taking along both pathogens and nutrients. Geophysical factors (e.g. slope), land use (e.g. human settlement, livestock rearing) and climatic factors (e.g. precipitation, temperature) influence the amount of microorganisms and nutrients that end up in surface waters. Differences between pathogens and nutrients include, for instance, that pathogens do not cycle in the atmosphere as N does, nor weather from soils as P does, and are not in artificial fertilizers (Bouwman et al., 2012). On the other hand, nutrients do not die-off as pathogens do (Ramirez et al., 2004). Moreover, all people and animals worldwide excrete nutrients and their amount depends on their diet (Bouwman et al., 2012), while only infected people and animals excrete pathogens. The amount varies over the course of the infection and depends, among others, on host immunity.

A global waterborne pathogen model could provide valuable new insights by (1) providing quantitative information on pathogen levels in data-sparse regions, which are often developing countries, (2) describing the spatial distribution of pathogen loads and identifying pathogen hotspots (i.e. surface water with high pathogen concentrations), (3) enabling future concentration projections under global change scenarios (Ferguson and Kay, 2012; Hofstra, 2011) and (4) providing an evidence base to inform effective pathogen health risk management, for instance by determining the relative influence of contamination sources or studying effects of management strategies on pathogen concentrations in surface water. The objective of this review is to learn from existing watershed-scale waterborne pathogen models and global nutrient export models for the development of a global waterborne pathogen model. This review is performed in the context of the design of a global *Cryptosporidium* model, a first exploration of which has been published by Hofstra et al. (2013). We present desirable characteristics of a global waterborne pathogen model, present a conceptual global waterborne pathogen model and identify the data requirements for such a model.

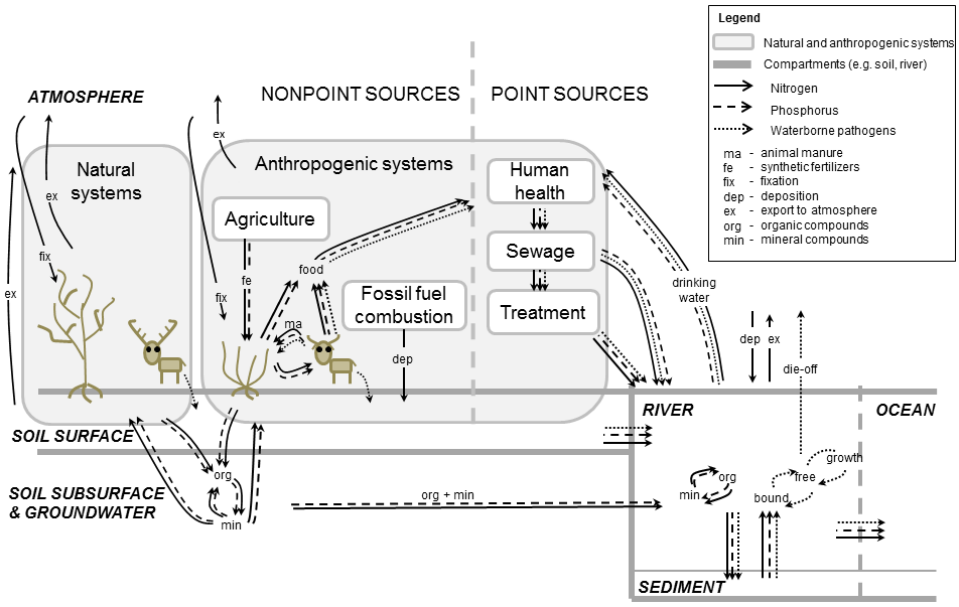


Figure 2.1: Overlapping environmental sources and pathways of waterborne pathogens and nutrients (N and P). This figure shows how point and nonpoint sources of pathogens and nutrients from natural and anthropogenic systems contribute to loading of surface waters. Pathogens and nutrients show overlap in sources (human and animal faeces) and pathways (overland flow, sewage discharge). Nitrogen differs from phosphorus and pathogens in that it has an atmospheric state and processes such as fixation and deposition as well as export play an important role in its global cycle. Pathogens differ from nutrients in that their prevalence is less static; die-off and growth processes play a role, and shedding rates of a population can change over time. For both pathogens and nutrients reversible transition processes within the surface water exist; for nutrients especially mineral to organic forms and for pathogens free or bound to particles are dominant processes. Both pathogens and nutrients can influence human health via food and drinking water.

2.2 Selection of models

2.2.1 Pathogen model selection

We included models of zoonotic waterborne pathogens that spread via the faecal-oral route and models of FIOs (hereafter together referred to as 'waterborne pathogen models'). The selected models were developed for the following organisms: *Escherichia coli, Cryptosporidium, Giardia, Campylobacter,* faecal indicator bacteria and faecal coliforms (Cho et al., 2010; Collins and Rutherford, 2004; Dorner et al., 2006; Ferguson et al., 2007; Fraser et al., 1998; Haydon and Deletic, 2006; Hipsey et al., 2008; Hofstra et al., 2013; Medema and Schijven, 2001; Petersen et al., 2009; Reder et al., 2015; Robles-Morua et al., 2012; Sadeghi and Arnold, 2002; Tang et al., 2011; Tian et al., 2002; Walker and Stedinger, 1999; Wu et al., 2009; Yeghiazarian et al., 2009). We excluded cholera (*Vibrio*

cholerae) modelling studies, because this pathogen does not solely follow the faecal oral route, it also grows in brackish water, and therefore the model dynamics are different (Epstein, 1993). We only included models at watershed or larger scale. We thus excluded approaches at smaller scales, such as laboratory, soil core, hill slope or farm scale, as these have been included in previous reviews (Oliver et al., 2009; Pachepsky et al., 2006). Most models calculated daily pathogen or FIO concentrations. We focus on fresh water streams, excluding lakes, beaches or estuaries (Nevers and Whitman, 2005; Olyphant and Whitman, 2004; Vijay et al., 2010). Furthermore, we focus on process-based models, and therefore also excluded empirical modelling studies aimed at statistically relating observed pathogen concentrations to environmental factors, such as climate variables, (Crowther et al., 2003; Kay et al., 2005; Kay and Mcdonald, 1983; Nevers and Whitman, 2005). We realize that our list of models is not exhaustive, but believe that for the purpose of this review including more models will not greatly increase information content or change conclusions. A full overview of the models and their general characteristics (e.g. pathogen species, spatial and temporal scale and steps, type of output) can be found in the supplementary material (Table S2.1).

2.2.2 Nutrient model selection

'Nutrients' in this review refers to N and P. In surface waters N and P can be found in dissolved organic, dissolved inorganic or particulate forms (Seitzinger et al., 2005). We focus on global models that simulate transport of nutrients from land to sea (Bouwman et al., 2009, 2005, 2012; Boyer et al., 2006; Caraco and Cole, 1999; Galloway et al., 2004; Green et al., 2004; Howarth et al., 1996; Mayorga et al., 2010; Morée et al., 2013; Seitzinger et al., 2010, 2005; Smith et al., 2003; van Drecht et al., 2009, 2003; Wollheim et al., 2008). Since we perform this review in the context of the design of a global waterborne pathogen model, all nutrient models in this review are global models, except for the model by Howarth et al. (1996), which covers the regions around the North Atlantic Ocean. This model was included because it was influential at the time. Most of these models include both process-based and empirical elements. We excluded studies that focus on modelling complete N and P cycles or the non-surface water related part of those cycles. We also excluded studies that provide global estimates that are not spatially explicit (Bennett et al., 2001; Fisher et al., 2010). The model output is generally the annual average nutrient load at a 0.5x0.5 degree grid, or for river basins. A full overview of the models and their main characteristics can be found in the supplementary material (Table S2.2).

2.3 Desirable characteristics of a global waterborne pathogen model

To evaluate existing models, we defined the characteristics of an ideal global waterborne pathogen model (see 1st column of Table 2.3). Ideally, the model calculates pathogen concentrations in surface water, since concentrations can be linked to risks associated with water consumption. Concentrations in surface waters can be calculated from point and nonpoint sources of pathogens. Ideally, the model accounts for differences between industrialized and developing countries. The model needs to incorporate all relevant environmental processes, such as hydrological processes, sedimentation and resuspension, attachment to particles, growth (when relevant) and die-off. The model then describes fate and transport within rivers, for instance by coupling to a hydrological model.

The ideal spatial and temporal resolution depends on the intended use of the model. To study spatial distribution of pathogen loads and pinpoint hotspot areas, a model with a relatively coarse resolution (e.g. 0.5x0.5 degree grid) giving static annual or seasonal estimates may give a reasonable first indication. Ideally, a global waterborne pathogen model can be used for scenario analysis in the context of global change to assess how waterborne pathogen concentrations change due to future trends in population, climate and land use change. Also to this aim a model with a relatively coarse spatial resolution giving static annual estimates can be suitable depending on the questions asked. When climate change is studied a smaller time step may be required, such as seasonal, monthly or daily. Finally, to study disease risk and risk mitigation options, a finer spatial resolution and dynamic daily estimates may be preferable.

2.4 Characteristics of existing pathogen and nutrient models

We evaluated how the desirable characteristics are represented in existing waterborne pathogen models and global river nutrient export models (see Tables 2.1 and 2.2). Model characteristics are presented for: (1) sources, (2) environmental processes, (3) spatial model characteristics, and (4) temporal model characteristics.

Table 2.1: Processes taken into account by waterborne pathogen models, indicating the first author of the paper in which the model or model results are published, the publication year, and whether specific sources or processes are taken into account by a given model (grey= yes, white = no). Since nearly all pathogen models are coupled to an existing hydrological model, in the column 'Hydrology' the names of these hydrological models are given. The remaining columns indicate: point sources (sewage outlets discharging into surface water), sewage overflows (extreme precipitation conditions leading to overflows of untreated sewage to the surface water), nonpoint sources (livestock manure applied to fields directly or from animal grazing), die-off of microorganisms in the environment, microorganism growth in the environment, sedimentation (settling of pathogens in stream or lake sediments), resuspension (from stream or lake sediments into the water column), attachment to particles suspended in the water column and direct defecation by animals into streams (circumventing runoff processes by directly adding these loads to the surface water). The * with numbers indicate comments.

Author(s)	Year published	Hydrology	Dynamic	Point sources (sewage)	Sewage overflows	Nonpoint sources (manure)	Direct defecation by animals into streams	Die-off in environment	Microorganism growth in environment	Sedimentation	Resuspension	Attachment to particles
Cho et al.	2010	Adjusted from the model by Steets and Holden (2003)		*1								
Collins & Rutherford	2004	Watershed Assessment Model incorporating an agricultural runoff model (WAM/GLEAMS)										
Dorner et al.	2006	WATFLOOD model										
Ferguson et al.	2007	The non-linear loss module of the IHACRES rainfall-runoff model										
Fraser et al.	1998	Not taken into account, merely delivery from land to surface water								*10		*10
Haydon and Deletic	2006	SimHyd model		*2	*2							
Hipsey et al.	2008	CAEDYM model										
Hofstra et al.	2013	Not taken into account, merely delivery from land to surface water.					*7					

Medema & Schijven	2001	WATNAT model			*5								
Petersen et al.	2009	Spatial lay-out of the watershed is included, but in-stream routing not.				*6	*6	*8		*8		
Reder et al.	2015	WorldQual, part of WaterGAP3										
Robles-Morua et al	2012	tRIBS model										
Sadeghi and Arnold	2012	SWAT model		*3								
Tang	2011	SWAT2005										
Tian et al.	2002	Watershed Assessment Model (WAMView) incorporating an agricultural runoff model (GLEAMS)		*4								
Walker & Stedinger	1999	Generalized Watershed Loading Function (GWLF) model coupled to basic stream routing and reservoir retention modeling.								*9		
Wu et al.	2009	WATFLOOD/SPL9 model										
Yeghiazarian et al.	2009	Water Erosion Prediction Project (WEPP) model				*11						

*1 Included, but not explicitly. User sets initial concentration values in this model.
*2 Only septic tank overflows, that can be added to the subsurface pathogen stores.
*3 Facultative inclusion of point sources on daily, average monthly or average annual basis. Point source loads need to be specified by user.
*4 Not only human sewage, mainly discharges from meat processing plants and dairy sheds are modelled as point sources.
*5 Sewage overflows modelled as constant discharge of a small percentage of raw sewage (0.87%)
*6 Some included. Livestock and manure are not taken into account in the model. Only birds, dogs and unspecified 'other mammals' are included.
*7 Study includes sensitivity analysis on direct deposition of human emissions in streams.
*8 Die-off, settling in sediment and possible regrowth together with other unaccounted processes are simulated as a single net value
*9 Oocyst settling and degradation in a stratified water column in reservoirs is included in the model.
*10 Not included, since in-stream processes are not modelled. However, transport of bacteria with runoff is assumed to be associated with sediment transport.
*11 Initial microorganism loads need to be specified by user. Only spread and state-transitioning is modelled.

Table 2.2: Processes taken into account by global nutrient models, indicating the first author of the paper in which the model or model results are published, the publication year, and whether specific sources or processes are taken into account by a given model (grey= yes, white= no). All nutrient models in this study which include river basin delineation have used the Simulated Topological Network STN-30 (Fekete et al., 2001; C. J. Vörösmarty et al., 2000). The remaining columns indicate: point sources (sewage outlets discharging into surface water), nonpoint sources (livestock manure applied to fields directly or from animal grazing) and retention (nutrient losses on land or within water, this implicitly includes sedimentation, resuspension and denitrification). None of the reviewed models were dynamic, and none included sewage overflows or direct defecation by animals into streams, therefore these are omitted from the table.

Author(s)	Year published	Hydrology	Point sources (sewage)	Nonpoint sources (manure)	Retention
Bouwman et al.	2005				
Bouwman et al.	2009	*1			
Bouwman et al.	2013	*1			
Boyer et al.	2006	*2	*4	*4	
Caraco & Cole	1999	*3			
Galloway et al.	2004	*5	*4	*4	
Green et al.	2004				
Howarth et al.	1996	*2	*4	*4	
Mayorga et al.	2010				
Morée et al.	2013	*6			
Seitzinger et al.	2005				
Smith et al.	2003		*7		
Van Drecht et al.	2003				
Van Drecht et al.	2009	*6			
Wollheim et al.	2008				

*1 Not included, only soil balances are modelled
*2 Not explicitly modelled, empirical relations of N inputs on the landmass to riverine N export.
*3 Regression model with runoff as variable
*4 Not modelled, but implicitly included in the import and export of food to an area.
*5 Riverine export estimates are based on earlier work by Howarth and Boyer, but then modified and extended to the global scale.
*6 Not included, only sewage emissions to the surface water are modelled
*7 Regression model with population density as variable

2.4.1 Sources

Most waterborne pathogen models in this review consider either point sources (sewage) or nonpoint sources (animal manure on land) of pathogens to the surface water (see Table

2.1). This is because some were specifically developed for urban watersheds (Cho et al., 2010; Petersen et al., 2009), and others for rural watersheds (Collins and Rutherford, 2004; Fraser et al., 1998; Tang et al., 2011; Tian et al., 2002). Some models combine both. We identified the following major pathways of faeces to the surface water in the reviewed models: (1) discharge of (treated) sewage to the surface water, (2) overflows of untreated sewage, (3) overland and/or subsurface flow transporting manure from land to water, and (4) direct defecation of animals into streams (Tables 2.1 and 2.2, Figure 2.1).

Point sources in the pathogen models include discharges of treated and untreated sewage. Only a few models account for sewage overflows. As sewage overflows bypass treatment, large numbers of pathogens can enter the surface water during an overflow event (Rechenburg et al., 2006). Medema and Schijven (Medema and Schijven, 2001) accounted for overflows by modelling them as a constant discharge of a small percentage (<1% in 1994) of raw sewage. However, this approach fails to take into account the local and episodic nature of overflows. Ferguson et al. (2007) incorporate overflows by looking at flow volumes and storage capacity of sewage treatment plants; if more water enters than the storage can hold, it is assumed to leave as an untreated discharge. None of the reviewed nutrient models account for sewage overflows. Most pathogen models including nonpoint pathogen sources include manure deposition on land and subsequent transport via overland flow. A few models include direct defecation of animals into streams (Collins and Rutherford, 2004; Ferguson et al., 2007; Petersen et al., 2009), which can lead to high pathogen loads. None of the reviewed nutrient models account for direct defecation. The majority of pathogen models and all nutrient models only consider livestock manure as nonpoint source, but depending on the pathogen species faeces from pets and wildlife can also play a role (Petersen et al., 2009).

The pathogen models have considerably different approaches for estimating inputs that are related to the pathogen sources. For example, in several models the input loads of microorganisms are not calculated by the model, but are user defined (Cho et al., 2010; Sadeghi and Arnold, 2002; Yeghiazarian et al., 2009). Such models are developed to model in-stream processes such as die-off and sedimentation of pathogens. Other models, such as those developed by Fraser et al. (1998) and Hofstra et al. (2013) are typical substance flow models. These models do not explicitly analyse in-stream processes, but aim at quantifying pathogen transport and loads to the surface water. Nutrient models use a wide range of input data that are also useful for pathogen modelling, such as human and animal population numbers, connection to sewer systems and waste water treatment etc. Most waterborne pathogen models focus on industrialized countries and are based on or validated with merely one or two years of observational data (Table S2.1).

One of the largest challenges as well as largest opportunities in developing a global waterborne pathogen model is the inclusion of developing countries, where problems are largest (UNICEF/WHO, 2009). Sewer coverage is often very limited in developing countries; in many places alternative sanitation systems are used including toilets hanging directly above streams, transport with open drains, or open defecation (WHO/UNICEF JMP, 2014). Therefore, we suggest that modelling pathogen sources in developing countries requires two additional pathways: (5) deposition of human faeces directly into surface water and (6) overland and/or subsurface flow transporting human faeces from land to water (Figure 2.2). None of the reviewed pathogen models incorporate direct deposition and only Reder et al. (2015) incorporate open defecation, but several do include human faeces from septic tanks entering the environment (Ferguson et al., 2007; Haydon and Deletic, 2006; Petersen et al., 2009; Reder et al., 2015). Most global nutrient models only include faeces from people connected to a sewage system, although some also (partly) include nutrients from faeces of the non-connected population (Bouwman et al., 2009; Green et al., 2004; Morée et al., 2013). The six identified pathways show that point source modelling should consider sewage, but also direct defecation into streams by humans and animals. Locally, other point sources such as discharges from meat processing plants can also be important (Tian et al., 2002). Nonpoint source modelling should consider animal manure, but also urban and rural wash-off of human faeces, especially in developing countries.

2.4.2 Environmental processes

We identified the following processes to govern fate and transport of pathogens to and in surface waters: (1) hydrological processes, (2) growth in the water environment (relevant for some bacterial pathogens, but not for parasites and viruses), (3) die-off, (4) attachment to particles, and (5) sedimentation and resuspension. Tables 2.1 and 2.2 show whether or not these processes are accounted for in the reviewed models.

Pathogen transport from manure on land is usually modelled as a function of hydrological processes, such as rainfall or runoff, sometimes including a mobilization fraction or taking into account manure stores built up during the antecedent dry period (Ferguson et al., 2007). Various pathogen models specifically focus on wet and dry weather event modelling (Cho et al., 2010; Ferguson et al., 2007; Haydon and Deletic, 2006; Hipsey et al., 2008; Petersen et al., 2009; Wu et al., 2009; Yeghiazarian et al., 2009), and are generally designed to run for several days after the occurrence of a (hypothetical) weather event. Precipitation events lead to peak FIO and pathogen loads in surface waters shortly after the event (hours-days, depending on the catchment characteristics) (Ferguson et al., 1996; Tryland et al., 2011; Wu et al., 2011). Increased runoff from agricultural lands, combined sewer overflows and resuspension of microorganisms in sediments are the likely causes of

the observed increases during and after a wet weather event (Garzio-Hadzick et al., 2010; Gibson III et al., 1998; Kistemann et al., 2002; Stott et al., 2011), and many waterborne pathogen models take these processes along (Table 2.1). However, peak concentrations can also occur during low flow conditions, due to the reduction in dilution (Nichols et al., 2009). Nutrient peak concentrations may also occur shortly after wet weather events, yet these are not the nutrient models' focus, as most are interested in annual totals. Most nutrient models do take runoff into account, although some do not explicitly model it but use an empirical relation instead (Table 2.2).

Most waterborne pathogen models incorporate environmental losses in die-off processes (Table 2.1). Pathogens vary in their ability to survive in the environment. Some organisms have lifecycle stages, such as spores, cysts (*Giardia)* or oocysts *(Cryptosporidium),* that are highly resistant to environmental stresses (Olson et al., 1999). For example, *Cryptosporidium* oocysts can survive for months in the environment before deactivation occurs (Medema et al., 2009). Growth of microorganisms is not included in most models, since environmental growth does not occur for some organisms (e.g. *Cryptosporidium*), or only under very specific conditions (e.g. *E. coli*) (Vital et al., 2008; Wang et al., 1996). Oliver et al. (2010) found that including growth of *E. coli* in faecal pats on fields improved model predictions. Nutrient retention is included in about half of the reviewed models (Table 2.2), usually as a loss fraction, that reflects the net effect of processes such as denitrification, uptake in organic matter, sedimentation and water abstraction.

Other processes that can take place within water bodies include particle attachment, sedimentation and resuspension of pathogens (Table 2.1, Figure 2.1). Pachepsky and Shelton have written an extensive review on this topic for E.coli and faecal coliforms (Pachepsky and Shelton, 2011). Pathogens can exist either free floating, attached to suspended solids or settled in sediments in a water stream or lake. They are only slightly more dense than water, so sedimentation of freely suspended pathogens is low (Pachepsky and Shelton, 2011). The flow behaviour of pathogens attached to particles will be governed by the behaviour of the particles (Pachepsky and Shelton, 2011). Resuspension from sediments (sometimes termed entrainment) can be included in models as a simple power function of flow rate (Collins and Rutherford, 2004; Wu et al., 2009) or calculated by model calibration (Cho et al., 2010). Water velocity, water depth and longer residence times of water (e.g. behind dams) affect sedimentation and pathogen die-off processes (Brookes et al., 2004). Particle attachment, sedimentation and resuspension are not explicitly modelled in the reviewed nutrient models, although some do describe the occurrence of denitrification within sediments (Galloway et al., 2004; Howarth et al., 1996; Mayorga et al., 2010) or 'spiralling' (uptake in biomass followed by mineralization) (Webster, 1975; Wollheim et al., 2008). As nutrients are mostly dissolved constituents,

although particulate forms also exist (Seitzinger et al., 2010), their dynamics in water differ from those of pathogens.

2.4.3 Spatial model characteristics

The general spatial structure of both waterborne pathogen models and global nutrient export models is a combination of inputs determining loads to the surface water, functions determining in-stream processes and a hydrological model that describes flow.

Global nutrient export models typically use large, publicly available datasets. For example, Van Drecht et al. (2009) combine country statistics, compiled by the UN and WHO, on population and sanitation to compute national N and P emissions from human waste. These are then distributed over a 0.5x0.5 degree grid based on population density. The Global NEWS 2 models (Mayorga et al., 2010) use this gridded dataset as input to aggregate over river basins on the basis of the STN 30p dataset (C. Vörösmarty et al., 2000), that delineates river basins and gives flow direction worldwide on a 0.5x0.5 degree grid. This way they compute nutrient export to the river mouths. Waterborne pathogen models can follow a similar approach (Reder et al., 2015). However, most existing waterborne pathogen models operate at a much finer spatial resolution.

The reviewed waterborne pathogen models employ different hydrological models (Table 2.1). A wide variety of hydrological models at different scales (watershed – global) are available, and these can be run at different spatial resolutions. Consequently, the waterborne pathogen model output is in a different spatial format; some models give output lumped for the entire watershed, some are spatially explicit within the watershed and compute values for multiple sub-catchments or points along a stream (Cho et al., 2010; Ferguson et al., 2007), while others give gridded output (e.g. 0.5x0.5 degree longitude x latitude) (Hofstra et al., 2013; Reder et al., 2015) (Table S2.1). Most nutrient models do not explicitly model hydrology, but use input datasets provided by hydrological models. Several nutrient models use input data at the spatial level of detail of the aforementioned STN 30 dataset. Some models aggregate these inputs for river basins. The number of basins in global nutrient models varies from 35 large rivers (Caraco and Cole, 1999) to 6081 basins (Mayorga et al., 2010). This shows that a ‘global’ model can take many forms and be developed further over time.

2.4.4 Temporal model characteristics

Pathogen models are generally constructed with one or at most a few years of available data. Pathogen measurements are mostly available at intervals of a few weeks (Table S2.1). Both static (point in time) and dynamic (period in time) modelling approaches exist (Table 2.1). On the contrary, global nutrient export models generally aggregate results

over longer time periods and as annual averages (Table S2.2). These differences can be partly explained by different objectives of nutrient and pathogen models, and a lack of global data on pathogens and nutrient emissions and concentrations at finer spatial and temporal resolutions. The data constraint is not in hydrological and meteorological information, which is available at much finer temporal resolution; for example, the 11 hydrological models that participate in the EU WATCH model intercomparison project range in time step from daily to 15 minutes (Haddeland et al., 2011). However, the information needed to compute pathogen and nutrient emissions to the surface water at, for instance, daily time steps is only locally available; global datasets are lacking.

The potential impacts of global change on human health are increasingly gaining attention (Haines et al., 2006; Hofstra, 2011; McMichael et al., 2009; Patz et al., 2008; Rose et al., 2001; K. Smith et al., 2014). The value of scenario analysis is widely recognized in global change studies (Swart et al., 2004). However, scenario analyses focusing on human health are scarce; examples include malaria distribution (Rogers and Randolph, 2000) and cholera (Koelle et al., 2005). Scenario analysis for waterborne pathogens is still in its infancy. Some examples exist; wet weather event modelling could be considered a form of scenario analysis. The same holds for river flow modelling (Liu et al., 2011), and comparison of management options as in the estuarine creek model by Vijay et al. (2010) (not included in this review). In addition, studies exist that project changes in the incidence of disease under climate change, without focusing on the pathogens directly (Kolstad and Johansson, 2011). However, here with 'scenario analysis' we refer to the use of comprehensive scenarios describing possible future states of the world, as is done in global change studies. Scenario analysis is performed in nutrient modelling, for example, the Millennium Ecosystem Assessment scenarios are used as a basis to explore future trends in river export of nutrients, as affected by agriculture, human population and sanitation systems (Seitzinger et al., 2010). Such approaches could be followed in global pathogen modelling.

2.5 Synthesis and a conceptual pathogen model

Currently, existing pathogen models do not meet the identified desirable characteristics. In Table 2.3 we summarize lessons from global nutrient models, and in Figure 2.2 we present a conceptual global waterborne pathogen model incorporating these lessons.

In general, waterborne pathogen models focus on water quality, and they often aim to assess contamination levels or health risk in a specific watershed, or to study the different contamination sources or the environmental factors that influence pathogen and FIO concentrations in surface waters. Global nutrient export models, on the contrary, mostly focus on understanding global nutrient fluxes, human impacts on nutrient fluxes, as well as effects on coastal water quality. These differences in focus are the underlying reason

for many of the observed differences between the models described in this review. The focus of the global waterborne pathogen model that we plan to develop and input and validation data availability will determine the spatial and temporal scale and resolution for which the model is developed.

Ideally, a global waterborne pathogen model includes point and nonpoint sources. Our review shows that while global nutrient models generally include both sources, waterborne pathogen models often focus on either. Global nutrient models show that it is feasible to include both sources in one model. Moreover, the generic input data used in global nutrient models could be used in a global pathogen model (Figure 2.2).

Waterborne pathogen models do not exist for developing countries. Many developing countries have large populations that are not connected to sewage systems. We can learn from global nutrient models and estimate emissions of pathogens in developing countries from process knowledge, as is done for nutrient emissions. However, the nutrient models generally only include sewers and assume that the emissions from people that are not connected to sewers will not reach the surface water. Ideally a global waterborne pathogen model would include the population not connected to sewers, as is indicated in Figure 2.2.

Table 2.3: Desirable characteristics of a global waterborne pathogen model. This table summarizes to what extent these characteristics are present in existing waterborne pathogen models and global nutrient models.

Model features	Desirable characteristics of a global waterborne pathogen model	Existing waterborne pathogen models	Existing global nutrient export models
Sources	Combine point sources and nonpoint sources		
	Combine industrialized and developing countries		
Environmental processes	Incorporate all relevant environmental processes		
Spatial features	Output at spatial detail relevant for risk assessment		
	Incorporates transport with rivers		
Temporal features	Output at temporal detail relevant for risk assessment		
	Future projections or scenario analysis possible		

Legend	Present in most or all models	Present in some models	Not present

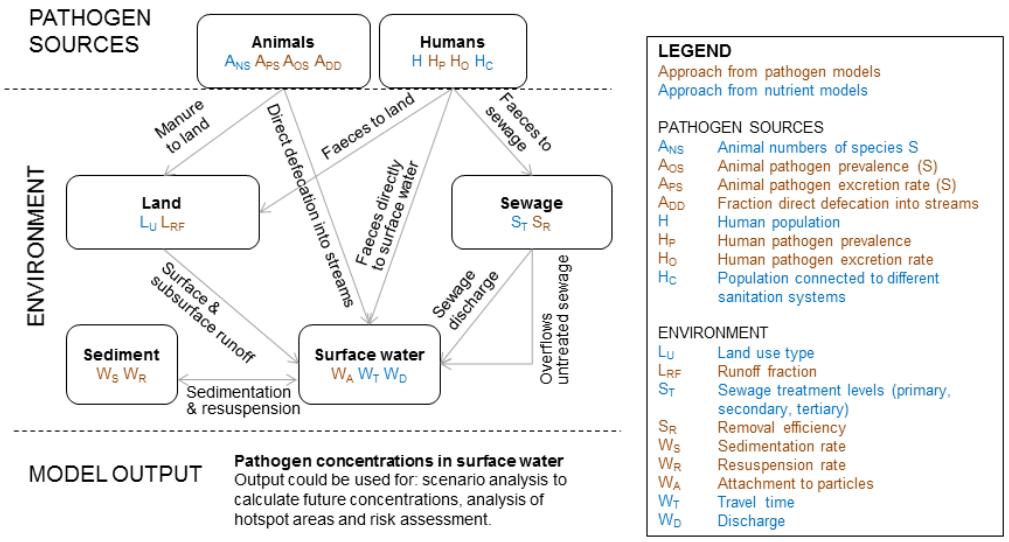


Figure 2.2: Conceptual model for a spatially explicit global waterborne pathogen model. Pathogens excreted in human and animal faeces enter and travel through the environment via different pathways (shown as arrows). The boxes show different compartments in which the pathogens can reside. The brown and blue letters show different model parameters. In the legend on the right these a full description of these is given. Brown and blue respectively indicate that the approach could be taken predominantly from existing pathogen models or from existing global nutrient export models (and associated datasets and hydrological models). Not shown in the figure are growth and die-off of pathogens, which could take place in all model compartments.

Pathogen modelling at daily time steps is desirable for risk assessment. This requires inclusion of specific environmental processes, such as inactivation and attachment to particles, sedimentation and resuspension, in a global waterborne pathogen model. We can adopt approaches from local waterborne pathogen models, that generally include many of these processes (Figure 2.2). These processes should, however, be applied at the global scale while maintaining a resolution relevant for risk assessment. Lack of input data may make it difficult to maintain such spatial and temporal detail. In order to assess hotspots and perform scenario analysis, the model would produce useful output at an annual or seasonal and 0.5 x 0.5 degree grid resolution. At this resolution some processes may prove to be unimportant as they average out over time.

Most pathogen and global nutrient models are coupled to hydrological models. The spatial and temporal characteristics of such hydrological models do not limit the development of a high resolution global pathogen model. We can learn from pathogen and global nutrient models as examples for coupling pathogen inputs and processes with hydrological models in a global pathogen model. Other substance transport models in rivers, such as sediment

transport (Syvitski et al., 2005), may also provide approaches and datasets relevant for the construction of the river transport part of a global waterborne pathogen model.

Global nutrient models provide great examples for using scenario analysis. In waterborne pathogen modelling, scenario analyses will open up new realms of knowledge, in particular in the context of global change (Ferguson and Kay, 2012; Hofstra, 2011). Scenario development for pathogen modelling could follow the lines of nutrient modelling scenarios, and could be further adapted to include projections for changes in, for example, prevalence of pathogenic organisms among human and animal populations, pathogen removal efficiencies by different sewage treatment types and the use of different sanitation systems worldwide. A model at 0.5x0.5 degree grid giving static annual or seasonal (when climate change is studied) estimates can be suitable for this.

The processes identified in this review lead to the design of a conceptual global waterborne pathogen model (Figure 2.2). This conceptual model takes the parameter relationships in the exploratory study of *Cryptosporidium* emissions by Hofstra et al. (2013) as a starting point. The conceptual model visualizes how knowledge from existing models of pathogens and nutrients could be used to create a global waterborne pathogen model. Depending on data availability the focus, spatial and temporal resolution could be adapted.

Concentration measurements availability for validation is critical for both pathogen and nutrient models. Therefore, compilation of available datasets is required, and continuous measurements at regular intervals over a longer period (multiple years) should be encouraged in measurement programs around the world.

2.6 Conclusion

A global waterborne pathogen model could provide valuable new insights for data-sparse regions, by identifying pathogen hotspots and evaluating global change and risk management scenarios. We identified desirable characteristics and present a conceptual global waterborne pathogen model that builds on knowledge from existing local waterborne pathogen models and global river nutrient export models. A global waterborne pathogen model needs to include pathogen-specific processes, approaches that can be taken from existing watershed scale waterborne pathogen models. With a spatial and temporal resolution such as currently employed in global nutrient export models, a global pathogen model could likely be used for hotspot analysis and assess future pathogen levels through scenario analysis. Global change scenarios could be adapted for future projections in waterborne pathogen modelling. For use in disease risk management, it is desirable for the model to run at a finer spatial and temporal

resolution. Our review paves the way for an exciting new opportunity in waterborne pathogen modelling.

Acknowledgements

Part of this research has been funded by the European Climate-KIC program, and the Gerbrand de Jong Fonds has kindly provided travel support. We thank colleagues at the Environmental Systems Analysis Group at Wageningen University for valuable discussions on this research, and we thank anonymous reviewers for their helpful comments

Chapter 3

Modelling the impact of sanitation, population growth and urbanisation on human emissions of *Cryptosporidium* to surface waters – a case study for Bangladesh and India

Lucie C. Vermeulen, Jelske de Kraker, Nynke Hofstra, Carolien Kroeze and Gertjan Medema

Abstract

Cryptosporidium is a protozoan parasite that can cause diarrhoea. Human faeces are an important source of *Cryptosporidium* in surface waters. We present a model to study the impact of sanitation, urbanisation and population growth on human emissions of *Cryptosporidium* to surface waters. We build on a global model by Hofstra et al. (2013) and zoom into Bangladesh and India as illustrative case studies. The model is most sensitive to changes in oocyst excretion and infection rate, and to assumptions on the share of faeces reaching the surface water for different sanitation types. We find urban centres to be hotspots of human *Cryptosporidium* emissions. We estimate that 53% (Bangladesh) and 91% (India) of total emissions come from urban areas. 50% of oocysts come from only 8% (Bangladesh) and 3% (India) of the country area. In future, population growth and urbanisation may further deteriorate water quality in Bangladesh and India, despite improved sanitation. Under our 'Business as usual' ('Sanitation improvements') scenario, oocyst emissions will increase by a factor 2.0 (1.2) for India and 2.9 (1.1) for Bangladesh between 2010 and 2050. Population growth, urbanisation and sanitation development are important processes to consider for large scale water quality modelling.

This chapter is based on: Vermeulen, L.C., De Kraker, J., Hofstra, N., Kroeze, C., Medema, G.J., 2015. Modelling the impact of sanitation, population and urbanisation estimates on human emissions of *Cryptosporidium* to surface waters – a case study for Bangladesh and India. Environ. Res. Lett. 10, 94017. doi:10.1088/ 1748-9326/ 10/ 9 / 094017

3.1 Introduction

Cryptosporidium is a protozoan intestinal parasite that causes diarrhoea in humans and animals worldwide. In the developing world diarrhoea is the third leading cause of death (WHO, 2008). But also in industrialized countries, *Cryptosporidium* outbreaks are regularly reported (Mackenzie et al., 1994). Through faeces of infected individuals *Cryptosporidium* oocysts – the robust survival stage of the pathogen - are excreted and spread in the environment. Surface water is an important mode of environmental transport, and a source of infection when ingested as drinking water, via irrigated crops or during recreation (Medema and Schijven, 2001). Ingestion of low numbers of oocysts causes a significant probability of infection (DuPont et al., 1995). Manure from livestock, and to a lesser extent from wildlife, is a diffuse source of oocysts to surface water via runoff (Cox et al., 2005; Thurston-Enriquez et al., 2005). Human faeces are a point source in various ways. Sewage systems can discharge faeces in to surface water, either treated or untreated. Several treatment types (primary, secondary and tertiary) are applied with different removal rates of *Cryptosporidium* (Hofstra et al., 2013). Faeces can also be deposited directly into the surface water, for example via hanging toilets (WHO/UNICEF JMP, 2014), or they can be deposited on land when people practice open defecation, where they can be a diffuse source.

Population growth, urbanisation and changes in sanitation are potentially important processes to consider for future water quality. Projections indicate that population will grow and urbanisation will increase, especially in the developing world (United Nations, 2014), but quantifications of the potential effects on water quality are limited. Sanitation improvements are lagging behind urbanisation rates, especially in the growing number of urban slums in developing countries (WHO/UNICEF JMP, 2000). The Sustainable Development Goals aim to halve the discharge of untreated sewage and end open defecation, to limit the spread of waterborne diseases. However, data on pathogens in surface water are scarce, especially in the developing world, as monitoring is time-consuming and costly.

Global modelling of pathogen emissions to surface water could contribute to adequate sanitation management to reduce the spread of waterborne diseases. The exploratory *Cryptosporidium* emission model (Hofstra et al., 2013) is a first global assessment of *Cryptosporidium* emissions to rivers in a spatially explicit way. This study accounts for people connected to a sewage system, but leaves out a large part of the global population, especially in developing countries, that is not connected to sewerage. This may lead to a significant underestimation of the actual situation, as a sensitivity analysis of the model showed that when assuming 20% of the faeces of the people not connected to sewer

systems would end up directly into the surface water, this would almost double human *Cryptosporidium* emissions (Hofstra et al., 2013).

Our aim is to develop a spatially explicit model to study the impact of sanitation, urbanisation and population growth on human emissions of *Cryptosporidium* to surface waters. We take the model by Hofstra et al. (2013) as starting point. We zoom into Bangladesh and India as illustrative case studies, and we apply the model in a scenario analysis to demonstrate the importance that population growth, urbanisation and changes in sanitation may have for future *Cryptosporidium* emissions to surface water.

3.2 Methods

3.2.1 Description of the original Hofstra model

A full model description can be found in Hofstra et al. (2013), here we give a short summary of the point sources sub-model, describing human emissions via sewage systems. The Hofstra model calculates the total annual emissions of *Cryptosporidium* oocysts to surface water on a 30 minute grid (0.5 x 0.5 degree) for the year 2000. The model is programmed in R. Firstly, country total annual emissions are calculated, and then these are distributed over grids. The oocyst excretion rate per ill person (O_i) is set at 10^9 oocysts per person per disease episode. Infection rates (fraction of the population experiencing a disease episode per year) have been set at 10% (I_d) and 5% (I_i) for developing and industrialized countries respectively. Countries having a score of 0.785 or lower on the United Nations Development Programme's (UNDP) Human Development Index (HDI) are classified as developing countries, Bangladesh and India fall into this category. Average excretion rate (O_p) in oocysts person^{-1} year^{-1} is then calculated as:

$$\mathrm{O_p} = \mathrm{I_d} \times \mathrm{O_i} \;\; if\; HDI \;\leq\; 0.785 \tag{3.1}$$

$$\mathrm{O_p} = \mathrm{I_i} \times \mathrm{O_i} \;\; if\; HDI > 0.785 \tag{3.2}$$

The fraction of oocysts removed by wastewater treatment (f_{rem}) depends on the treatment type. The Hofstra model defines four categories: no treatment (0% oocysts removed), mechanical (primary) treatment (f_p) (R_p =10%), biological (secondary) treatment (f_s) (R_s =50%) and advanced (tertiary) treatment (f_t) (R_t =95%). The fraction of sewage treatment that falls into each of the categories are country estimates based on data from WHO/UNICEF, and the fraction of oocysts removed is the weighted average of these efficiencies.

$$\mathrm{f_{rem}} = \mathrm{f_p} \times \mathrm{R_p} + \mathrm{f_s} \times \mathrm{R_s} + \mathrm{f_t} \times \mathrm{R_t} \tag{3.3}$$

The total human emissions per country (H) in oocysts year^{-1} are calculated by multiplying the country population (P) connected to a sewer system (f_c) with the average oocyst excretion rate per person (O_p) in oocysts person^{-1} year^{-1}. The fraction of oocysts removed in sewage treatment plants (STPs) is then subtracted (f_{rem}).

$$H = P \times f_c \times O_p \times (1 - f_{rem}) \quad (3.4)$$

The total human emissions per country (H) are then distributed over grid cells based on population density (taken from LandScan data maps) (Dobson et al., 2000). This is done under the assumption that places with the highest population densities are most likely to have a sewage system. Starting at the most densely populated grid cells, the total human emissions are allocated until all have been distributed.

3.2.2 An improved approach to account for sanitation types

The original Hofstra model divides the population in people that are either connected or not connected to a sewage system, and the emissions of the latter are ignored. No distinction is made between populations in urban and rural areas. We propose to divide the population in four emission categories: (1) people connected to sewage systems, (2) people as a direct source of pathogens in rivers, (3) people as a diffuse source and (4) people as non-source. Furthermore, we make a distinction between urban and rural populations. To this end, we reclassify the sanitation coverage data from the Demographic and Health Survey (DHS) Program to fit our emission categories (see Table 3.1). In classifying we assume that:

a. Faeces of people connected to a sewage system will reach the surface water (treated or untreated). We assume different oocyst removal per treatment type, similar to the original Hofstra model.
b. Faeces of people using septic tanks, pits or pit latrines or composting toilets will not reach the surface water. Soil passage effectively retains protozoan (oo)cysts (Ferguson et al., 2003). Furthermore, long storage time in septic tanks and latrines can cause oocyst die-off, and routes of disposal of contents are largely unknown. Therefore, we assume emissions to surface water of people using these systems to be zero.
c. Faeces of people using hanging toilets are a direct source of oocysts to the surface water. These are systems where a toilet facility is built above a stream or lake and faeces drop directly into the water.
d. Faeces of people without sanitation facilities are a direct source of oocysts in urban areas and a diffuse source in rural areas. In urban areas faeces of people without sanitation facilities likely end up in the surface water (e.g. via open

drains), due to the lack of space for open defecation on land. In rural areas, these emissions are likely to end up on the land and can form a diffuse source similar to animal manure (Thurston-Enriquez et al., 2005). Therefore, we classify the sanitation categories 'unknown', 'elsewhere' or 'no facilities, bush, field' to the diffuse sources for rural populations and to the direct sources for urban populations.

Table 3.1: Four emission categories are used in our model: (1) people connected to sewage systems, (2) people as a direct source of pathogens in rivers, (3) people as a diffuse source and (4) people as non-source. This table shows how we classify the types of sanitation (following the DHS Program) for urban and rural areas into these emission categories(International Institute for Population Sciences (IIPS) and Macro International, 2007; National Institute of Population Research and Training (NIPORT) et al., 2013).

Emission categories	Urban sanitation types based on DHS data	Rural sanitation types based on DHS data
Connected	To piped sewer system	To piped sewer system
Direct source	Hanging toilet, no facility, bush, field, unknown, elsewhere	Hanging toilet
Diffuse source	-	No facility, bush, field, unknown, elsewhere
Non-source	To septic tank, to pit, pit latrine, composting toilet	To septic tank, to pit, pit latrine, composting toilet

3.2.3 Estimating *Cryptosporidium* emissions in 2010

Using the assumptions above, we calculate human *Cryptosporidium* emissions to the surface water in Bangladesh and India for the year 2010. These countries were chosen as illustrative examples of developing countries with high population density and urbanisation rates where a variety of different sanitation types are used, and because Hofstra et al. (2013) indicated this region as one with emission hot-spots.

We use the formula by Hofstra et al. (2013) for the calculation of the human emissions via sewage systems, but calculate this for urban and rural areas separately. In addition, we calculate direct and diffuse emissions. This results in the following equations:

$$Connected\ emissions\ urban\ \ CE_u = P_u \times f_{cu} \times O_p \times (1 - f_{rem}) \quad (3.5)$$

$$Connected\ emissions\ rural\ CE_r = P_r \times f_{cr} \times O_p \times (1 - f_{rem}) \quad (3.6)$$

$$Direct\ emissions\ urban\ DE_u = P_u \times f_{du} \times O_p \quad (3.7)$$

$$Direct\ emissions\ rural\ \mathrm{DE_r} = \mathrm{P_r} \times \mathrm{f_{dr}} \times \mathrm{O_p} \quad (3.8)$$

$$Diffuse\ emissions\ rural\ \mathrm{DifE_r} = \mathrm{P_r} \times \mathrm{f_{difr}} \times \mathrm{O_p} \times \mathrm{f_{run}} \quad (3.9)$$

$$Total\ human\ emissions\ \mathrm{H} = \mathrm{CE_u} + \mathrm{CE_r} + \mathrm{DE_u} + \mathrm{DE_r} + \mathrm{DifE_r} \quad (3.10)$$

Where O_p is the average oocyst excretion (oocysts person^{-1} year^{-1}). It is calculated as described above (Equations 3.1 and 3.2), f_{rem} is the fraction of oocysts removed by sewage treatment. It is calculated as described above (Equation 3.3), P_u and P_r are the total urban and rural population of a country, respectively (Equations 3.5 –3.9), f_{cu} and f_{cr} are the fractions of the urban and rural populations that make use of sanitation that is connected to a sewer system (Equations 3.5 and 3.6), f_{du} and f_{dr} are the fractions of the urban and rural populations that make use of sanitation that is a direct source (Equations 3.7 and 3.8), f_{difr} is the fraction of the rural population that has no sanitation facilities and forms a diffuse source (Equation 3.9), and f_{run} is the fraction of faeces transported with runoff from land to surface water (Equation 3.9).

We have updated the baseline of the model to the year 2010. In Tables S3.1 and S3.2 in the Supplementary material all parameter values used in the calculations can be found. Oocyst excretion per ill person and oocyst removal efficiencies by different sewage treatment levels equal the original values as estimated by Hofstra et al. (2013). We assume that of the population connected to sewage systems, 20% receives treatment and the rest reaches the surface water untreated. This estimate is in line with the estimates for sewage treatment in Indian cities (Central Pollution Control Board: Government of India, 2015), although treatment levels are not specified here. Dhaka, the capital of Bangladesh, reportedly has only one STP with a capacity to treat only a third of the collected wastewater, and sewage overflows occur regularly (WASHplus project of USAID, 2010). Data on sewage treatment in other regions of Bangladesh is difficult to find, therefore we take the same value of 20% to represent Bangladesh also. We assume that currently only primary sewage treatment exists in India and Bangladesh.

We set the infection rate to 5%, lower than the 10% assumed by Hofstra et al. (2013) for developing countries. We based this on a short literature review on cryptosporidiosis prevalence, finding that 2.1-3.5% of diarrhoea cases in Bangladesh and India are caused by *Cryptosporidium* (Sitara Swarna Rao Ajjampur et al., 2010; Bhattacharya et al., 1997; Haque et al., 2009; Rahman et al., 1990). Ajjampur et al. (2010b) found that 40% of children experience multiple episodes of cryptosporidiosis, and more generally, children in developing countries are estimated to have on average 2.9 episodes of diarrhoea per year (Fischer Walker et al., 2012), meaning that an overall annual *Cryptosporidium* infection rate of 10% may be correct for children. However, for adults it is likely lower, Fischer

Walker et al. (2010) reported a median incidence of 0.299-0.675 episodes of diarrhoea per year for adults and children > 5 years in the South and South East Asia region. This means that a 10% overall annual infection rate is probably too high for Bangladesh and India. Therefore we decided to set the overall annual infection rate to 5%.

We run the model using sanitation input data per state instead of national averages, dividing Bangladesh in 7 and India in 35 regions (see Table S3.2 for an overview). The data on urban and rural populations per state are from the Bangladesh Bureau of Statistics and the Census of India (Bangladesh Bureau of Statistics, 2011; Census of India, 2011). For data on the usage of different sanitation types we take the most recent estimates of the DHS Program, this is for Bangladesh the year 2011 and for India 2005-2006 (International Institute for Population Sciences (IIPS) and Macro International, 2007; National Institute of Population Research and Training (NIPORT) et al., 2013). The runoff fraction is based on the median value for mobilisation of *Cryptosporidium* from animal manure by Ferguson et al. (2007). We study the effect of uncertainty in model input parameters in a sensitivity analysis (Section 3.3.3).

3.2.4 Spatial distribution of emissions in 2010

We estimate the spatial distribution of oocyst emissions to the surface water in Bangladesh and India. We spatially identify a country's urban and rural populations via density ranking, where the population in the most densely populated grid cells is defined as urban, based on a LandScan density map (Bright et al., 2011). We assume that among a population defined as urban, sanitation is distributed equally, proportional to the occurrence of different sanitation types.

3.2.5 Estimating *Cryptosporidium* emissions in 2050

We calculate the potential effect of urbanisation, population growth and sanitation changes on future human *Cryptosporidium* emissions for Bangladesh and India. We define two scenarios:

1. Business as usual: We assume that in 2050 the percentage of people connected to the different sanitation types in urban and rural areas is the same as today. Sewage treatment levels have also stayed the same (only primary treatment).
2. Sanitation improvements: In 2050 open defecation is no longer practiced, and hanging toilets are no longer used. In urban areas, the population that previously used either of these sanitation types are now mostly connected to the sewage system (75%) or use on-site systems such as septic tanks and latrines (25%). In rural areas it is the other way around, it is more likely people use on-site systems (75%) than sewer connections (25%). Sewage treatment levels have improved,

one third is primary treatment, one third secondary and one third tertiary treatment. Improvement of sanitation is in line with the current trends observed in the DHS data.

For both scenarios we use the population and urbanisation estimates for 2050 based on the Global Orchestration (GO) scenario of the Millennium Ecosystem Assessment. This scenario assumes globalisation and reactive environmental management, as opposed to regionalisation and proactive environmental management (Alcamo et al., 2006).

3.3 Results

3.3.1 Accounting for sanitation types

We divide the urban and rural populations of Bangladesh and India in different model emission categories (connected, direct source, diffuse source, non-source) based on sanitation types according to the DHS Program. Figure 3.1 shows that only accounting for people connected to a sewage system, as was done in the original Hofstra model, may lead to a large underestimation of human emissions in India and Bangladesh. The figure also shows large differences exist between sanitation types in urban and rural populations and between countries.

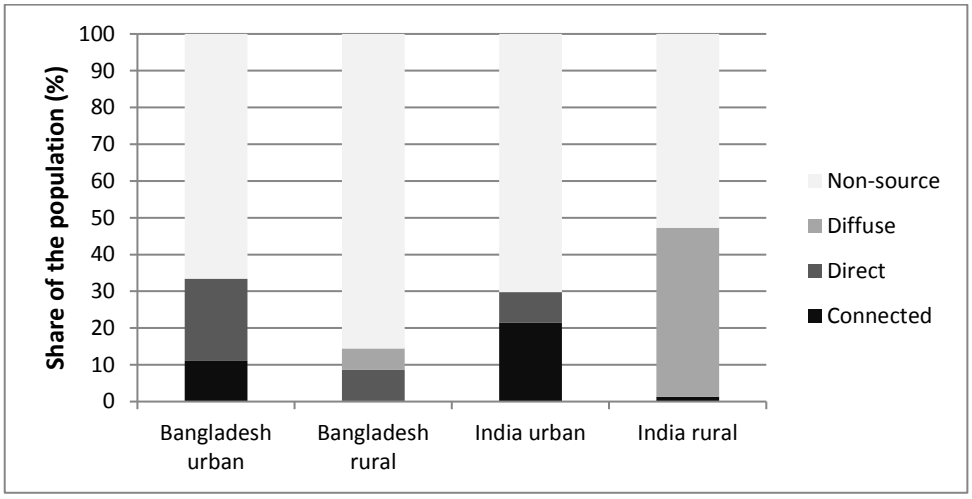


Figure 3.1: Share (%) of the population of Bangladesh and India in model emission categories, based on sanitation types. Data on sanitation types by the DHS Program were classified to fit the model emission categories 1) connected to a sewage system, 2) direct emission to the surface water (hanging toilets and for urban people the DHS categories 'unknown', 'elsewhere' or 'no facilities, bush, field'), 3) diffuse sources (for rural people the DHS categories 'unknown', 'elsewhere' or 'no facilities, bush, field') and 4) non-source (the DHS categories 'to septic tank', 'composting toilet' , 'to pit' and 'to pit latrine').

3.3.2 *Cryptosporidium* emissions in 2010

Total annual human *Cryptosporidium* emissions to the surface water in 2010 are for Bangladesh 1.0 x 10^{15} and for India 1.2 x 10^{16}, according to our model. In Bangladesh, of these total emissions 18% comes from the population connected to a sewage system, 81% from direct sources and 1% from diffuse sources. 53% of the total emissions of Bangladesh is urban. In India, of these total emissions 61% comes from the population connected to a sewage system, 32% from direct sources and 7% from diffuse sources. 91% of the total Indian emissions is urban. Compared to the Hofstra et al. (2013) estimate for 2000, our 2010 estimate for India is 1.5 times higher and for Bangladesh 8.6 times higher.

3.3.3 Sensitivity analysis

We studied the sensitivity of model output to changes in ten input parameters. Each parameter can take three different values in the sensitivity analysis, based on reasonable ranges the parameter can take (Table S3.1). We do the analysis in pairs of parameters, changing one or both at a time, as some parameters are strongly related to others. For example, the effect on model output of changing coverage of different sewage treatment levels (primary, secondary, tertiary) depends on the removal efficiencies assumed for these levels. Results of the sensitivity analysis are presented in Table S3.3. For both Bangladesh and India, the model was most sensitive to changes in the combination of oocyst excretion and number of infections. As these numbers are highly variable and uncertain (shown for 1 log unit change in oocyst excretion and halving or doubling of infection rate), the effect on model outcome is large, up to 20-fold increases or decreases. Uncertainty in whether faeces from septic tanks and pit latrines can reach the surface water can triple total oocyst emissions for Bangladesh. Uncertainty in the share of the faeces of people without sanitation facilities that reaches surface waters was a large contributor to uncertainty for India, mainly due to the large rural population without facilities.

3.3.4 Spatial distribution of emissions in 2010

Urban areas in both Bangladesh and India are hotspots of *Cryptosporidium* emissions to surface water. This is visualized in Figure 3.2, showing the spatial distribution of oocyst emissions over a 0.5 x 0.5 degree grid. Local differences in calculated oocyst emissions are large; in India 50% of oocysts originate from only 3% of grid cells, and 90% of oocysts originate from 12% of the grid cells. In Bangladesh the result is less extreme: 50% of oocysts originate from 8% of the grid cells and 90% of oocysts originate from 59% of the grid cells. This is because Bangladesh has a more evenly distributed population than India in the LandScan population density map.

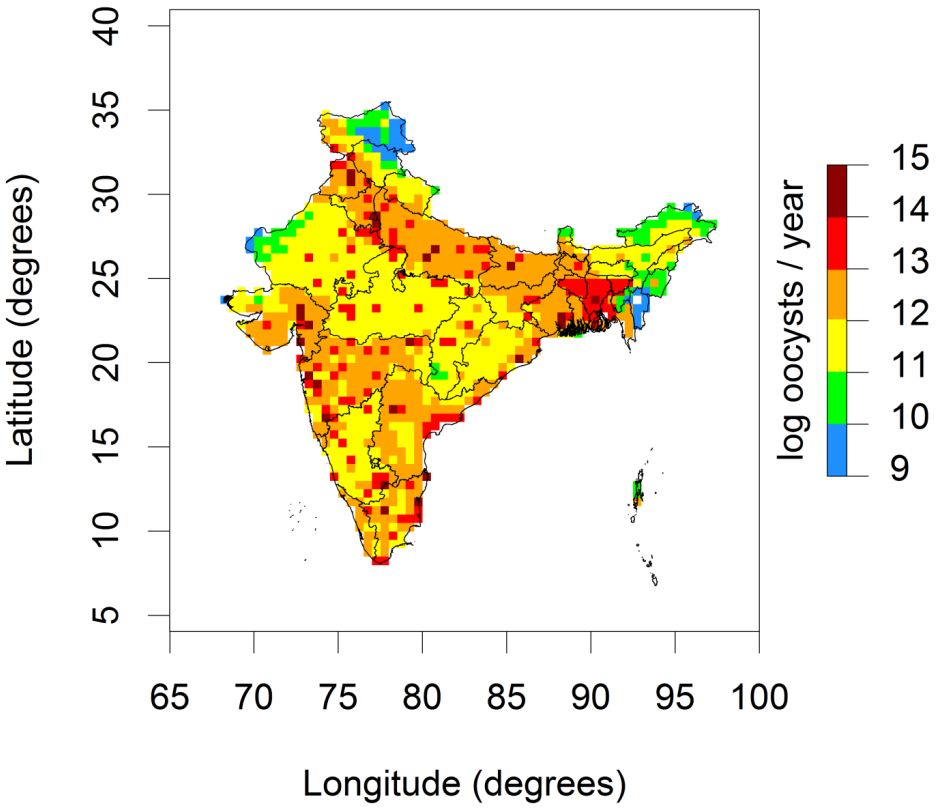


Figure 3.2: Spatial distribution of human oocyst emissions to the surface water in India and Bangladesh. This map shows log total annual human oocyst emissions to the surface water on a 0.5 x 0.5 degree grid, representing conditions for the year 2010. Emissions are distributed proportional to population size from a LandScan population density map (Bright et al., 2011). The results are calculated using population, urbanisation and sanitation data per state, dividing Bangladesh in 7 and India in 35 regions (borders shown on map).

3.3.5 Scenario analysis of future oocyst emissions

In the 'Business as usual' scenario the calculated total human emissions of oocysts to surface water for Bangladesh and India are higher in 2050 than in 2010 (Figure 3.3). According to our model oocyst emissions will increase by a factor 2.0 (India) and 2.9 (Bangladesh) in this period. This is mainly attributable to an expected increase in urban emissions of oocysts by a factor 2.1 (India) and 4.5 (Bangladesh). The calculated increase is solely due to population growth and urbanisation, as sanitation coverage is assumed to increase proportionally with population and urbanisation.

In the 'Sanitation improvements' scenario the calculated total human emissions for 2050 compared to 2010 are slightly higher for both Bangladesh (factor 1.1) and India (factor 1.2). For India, urban emissions go down slightly (factor 0.9) but rural emissions increase by a factor 4.0, because India had a very large population practicing open defecation that are now assumed to be partly connected to a sewage system.

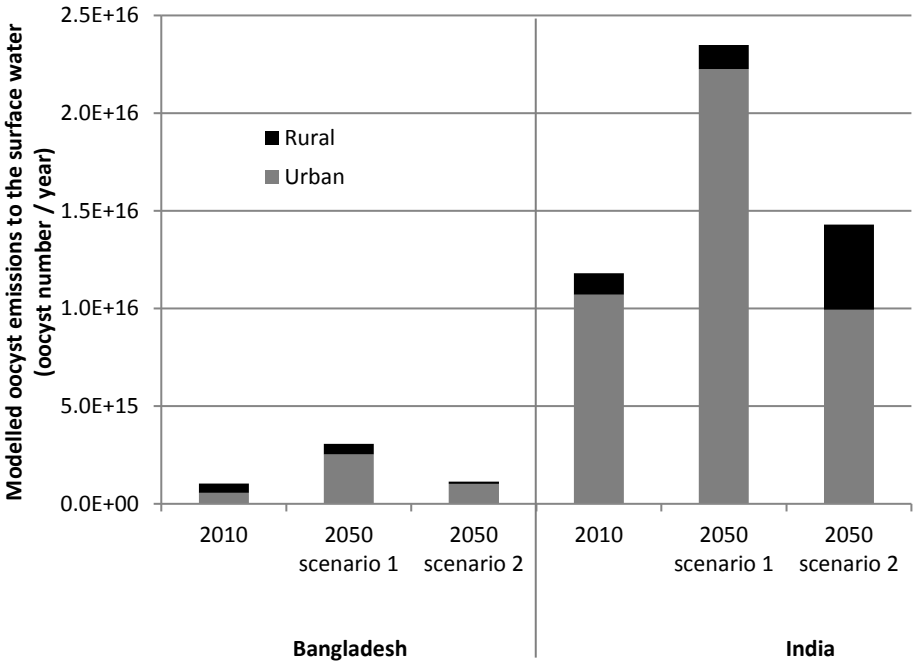


Figure 3.3: Total annual human *Cryptosporidium* oocyst emissions to surface water from urban and rural sources in Bangladesh and India in 2010 and for two scenarios in 2050, (1) 'Business as usual' and (2) 'Sanitation improvements'.

3.4 Discussion

To our knowledge, our model currently gives the only available estimate for *Cryptosporidium* oocyst emissions to surface water in Bangladesh and India. Our model can be used for pinpointing hotspot areas where problems may be largest, comparing pollution from different regions, and indicating dominant pollution sources now and in the future. The model can be informative for sanitation and water quality managers. The model in its current form is less appropriate for getting accurate emission values for specific locations within these countries, as uncertainties in this type of modelling

research are inevitably large. This is due to quality and availability of input data and model assumptions, among others. Oocyst excretion rates are highly uncertain. We assume a single annual oocyst excretion rate, as is done by Hofstra et al. (2013). This approach does not account for potential regional outbreaks of *Cryptosporidium* nor for variation in endemic cryptosporidiosis and shedding of oocysts. However, we did adjust the general infection rate estimate used in Hofstra et al. (2013) to be more representative for the situation in Bangladesh and India. Our categorisation of sanitation types assumes oocysts that end up in the soil through pit latrines and the like will not reach the surface water, but interflow may transport the faeces there (Davies et al., 2004), and in case of heavy rainfall these systems may flood. We ignore the contribution from septic tanks, it is assumed they are emptied in such a way the contents do not reach the surface water (e.g. in landfills) or after a long enough time for the oocysts to be inactivated. In addition, we classify the sanitation categories 'unknown', 'elsewhere' or 'no facilities, bush, field' to the diffuse sources for rural populations and to the direct sources for urban populations, based on the premise that due to lack of space in crowded urban areas it is likely that faeces will be disposed of towards the surface water. It is difficult to verify such assumptions, as regional or cultural variation in the use of sanitation can be large and the topic is often taboo (Dellström Rosenquist, 2005). In India 597 million people practice open defecation (WHO/UNICEF JMP, 2014). The uncertainty about what happens with these faeces is large, and can significantly influence model outcomes, as was shown in our sensitivity analysis. The fraction of faeces on land that is transported to surface water via runoff is dependent on geographical and climatic factors, such as slope and precipitation, which we have not taken into account in this study. Furthermore, measurement data to validate model outcomes are not available, to our knowledge. To improve our model, we would particularly require more data on oocyst excretion and the occurrence of cryptosporidiosis, data on the effect of different sanitation systems on pathogen survival, and measurements of pathogens in sewage and surface water for model validation.

Urban areas are hotspots of *Cryptosporidium* emissions; we estimate that in Bangladesh 53% and in India 91% of total emissions come from urban areas. The original Hofstra model spatially distributed point source sewer emissions over the most densely populated areas in a country only. By calculating at the state level, including a division between urban and rural populations, including direct and diffuse sources, and distributing proportionally to population size, we now produce a map that represents spatial distribution of oocyst emissions more accurately, as a larger share of the population is accounted for and the location of more population centres is represented.

The most problematic areas with regards to safe disposal of human faeces are likely to be urban slums. An urban slum can be defined as "an informal settlement in a city or town characterized by poor urban infrastructure, low water and sanitation service levels, high

population density and limited access for basic services" (Katukiza et al., 2013). It is questionable whether the DHS data on urban sanitation in Bangladesh and India also hold for slum areas, as actual surveys done in slums are scarce. According to UN Habitat, in India 29.4% and in Bangladesh 61.6% of urban population lived in a slum in 2009 (UN-HABITAT, 2012). A Bangladesh slum population survey identified 9,048 slum communities in the six major cities (Angeles et al., 2009). Hanchett et al. (2003) found that only 6-12% of households in slums of two Bangladesh cities has access to any form of latrines, septic tanks or sewerage, while the WHO/UNICEF JMP reports 50% overall urban access to improved sanitation in the year 2000 for Bangladesh (WHO/UNICEF JMP, 2014). Similarly, Agarwal et al. (2011) found that for urban areas in India, less than half of the poorest urban quartile had a flush toilet or pit latrine, while over 95% of the rest of the urban population did have this facility in 2005-2006. Furthermore, individual cities can differ in provision levels of basic services like sanitation, with smaller cities being generally underserved (Panel on Urban Population Dynamics, 2003). These examples highlight the inequalities between the richer and poorer urban populations, and the consequent difficulties for accurate spatial assessment of pollution originating from these areas. By categorizing the emissions of the urban population without sanitation access to the direct sources, we try to capture these slum populations in our model. It seems likely that part of the faeces from slums end up directly in surface waters (Nath, 2003; Nyenje et al., 2010), since sanitation coverage is low and there is little space for it to end up in the soil. In future, the problem will increase; it is expected that the number of people living in slums worldwide will have doubled by 2030 compared to 2000 (UN Millennium Project, 2005). The rate of urbanisation is so rapid that in developing countries planned urban expansion cannot keep up (UNEP/UN-HABITAT, 2010).

Sanitation coverage is improving in most world regions (WHO/UNICEF JMP, 2014). However, if more people are connected to sewers, this does not mean that adequate sewage treatment will also be installed (Baum et al., 2013; WHO/UNICEF JMP, 2000). A sewer connection without sewage treatment can cause faeces that now end up in the soil to reach the surface water untreated and affect water quality. More droughts and extreme rainfall events, which are expected with climate change, can cause problems for the existing STPs in the developing world, which are often old (UNEP/UN-HABITAT, 2010). Both in developing and developed countries rainfall events can cause sewer overflows, causing wastewater that was supposed to go to a STP to reach the surface water untreated. When flooding occurs, waste from open and inadequate sewers or other sanitation types will run off to lower-lying areas (UNEP/UN-HABITAT, 2010). This process is enhanced by urbanisation, because there will be more impervious surfaces inhibiting the water to infiltrate into the soil (Nyenje et al., 2010). Our current model does not incorporate rainfall event occurrence, as we are studying annual total emissions. Other

models studying waterborne pathogens at catchment scale at smaller time steps do incorporate this (e.g. Petersen et al. 2009).

The sensitivity analysis done by Hofstra et al. showed that the fate of faeces of people not connected to a sewage system may impact model outcomes considerably. In this study, we quantified this contribution for Bangladesh and India: we estimate that emissions from people not connected to a sewage system constitute 82% (Bangladesh) and 39% (India) of total emissions. Our scenario analysis highlights the importance of taking into account population growth, urbanisation and sanitation changes when predicting future water quality. We show that even with sanitation improvements (ending open defecation, improving sewage treatment levels) *Cryptosporidium* emissions to surface water are likely to increase.

3.5 Conclusion

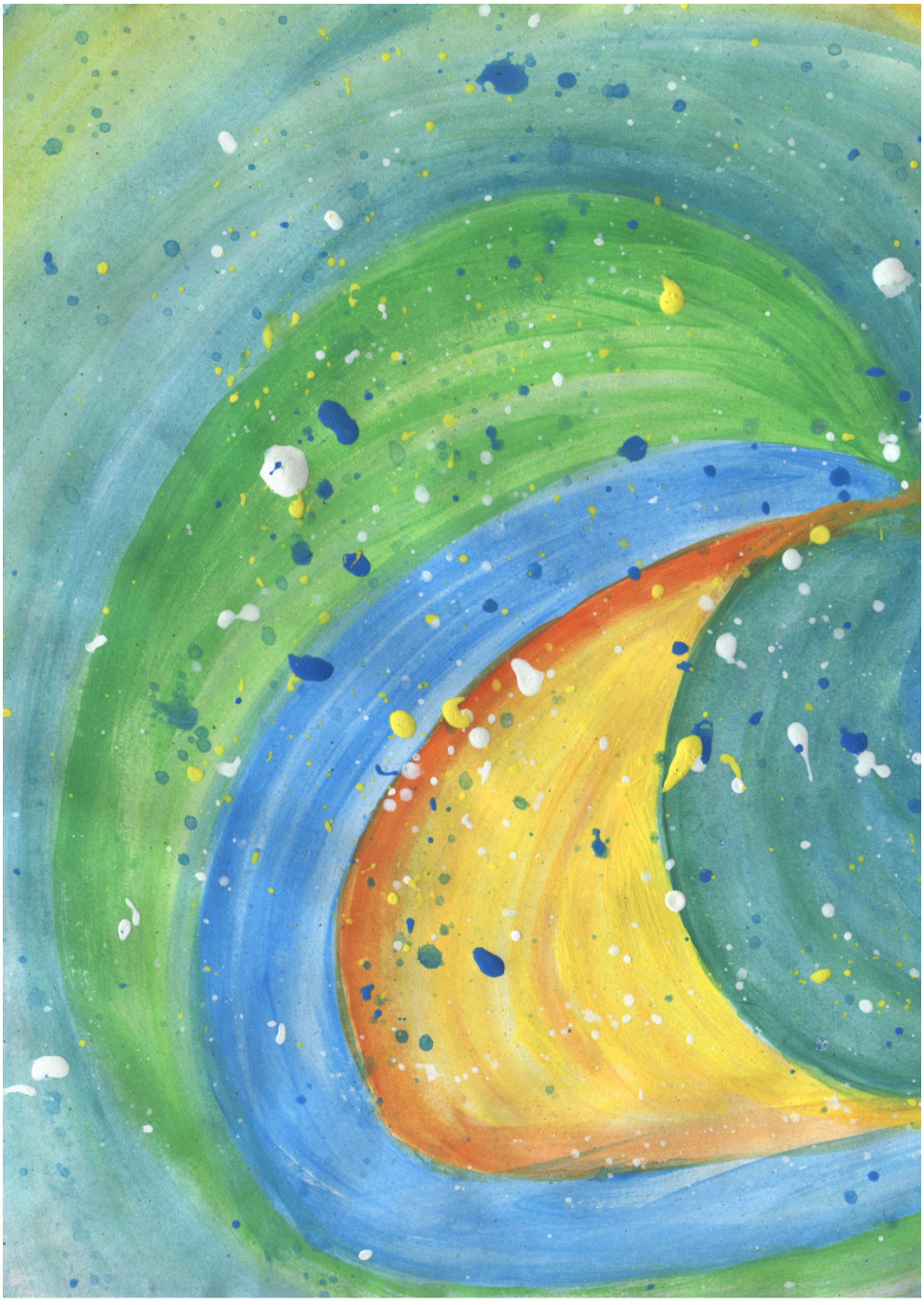
In this paper we propose a new method to calculate human emissions of *Cryptosporidium* to surface water, using a modified version of the Hofstra point sources sub-model. We modify the model by calculating human *Cryptosporidium* emissions using categorized DHS data on sanitation use per state, updating the baseline to 2010, adding direct and diffuse emissions, introducing a division between urban and rural populations, and creating a new spatial distribution.

Taking Bangladesh and India as case studies, we show that only accounting for people connected to a sewage system, as was done in the original Hofstra model, may lead to a large underestimation of human emissions in developing countries. Sewer connections with inadequate treatment, but also hanging toilets and open defecation negatively affect water quality. Urban centres are hotspots of human *Cryptosporidium* emissions in Bangladesh and India; we estimate that 53% (Bangladesh) and 91% (India) of total emissions come from urban areas. Our map indicates that 50% of oocysts originate from only 8% (Bangladesh) and 3% (India) of the country area. Future population growth and urbanisation are likely to lead to further deterioration of water quality in Bangladesh and India, in spite of efforts to improve sanitation. Under our 'Business as usual' scenario, oocyst emissions will increase by a factor 2.0 for India and 2.9 for Bangladesh between 2010 and 2050. Under our 'Sanitation improvements' scenario, oocysts emissions increase slightly for both Bangladesh (factor 1.1) and India (factor 1.2). The model is most sensitive to changes in oocyst excretion and infection rate, as well as assumptions on what share of faeces reaches the surface water for different sanitation types.

Population growth, urbanisation and the development of sanitation are important processes to consider for large scale modelling of current and future water quality related to human faeces. The new method proposed here is a first step for improved spatially explicit modelling of *Cryptosporidium*.

Acknowledgements

We thank anonymous reviewers for their comments and suggestions. Part of this work has been financially supported by the European Climate-KIC PhD programme.



Chapter 4

Impacts of population growth, urbanisation and sanitation changes on global human *Cryptosporidium* emissions to surface water

Nynke Hofstra and Lucie C. Vermeulen

Abstract

Cryptosporidium is a pathogenic protozoan parasite and is a leading cause of diarrhoea worldwide. The concentration of *Cryptosporidium* in the surface water is a determinant for probability of exposure and the risk of disease. Surface water concentrations are expected to change with population growth, urbanisation and changes in sanitation. The objective of this paper is to assess the importance of future changes in population, urbanisation and sanitation on global human emissions of *Cryptosporidium* to surface water. The GloWPa-Crypto H1 (the Global Waterborne Pathogen model for Human *Cryptosporidium* emissions version 1) model is presented and run for 2010 and with scenarios for 2050. The new scenarios are based on the Shared Socio-economic Pathways (SSPs) developed for the climate community. The scenarios comprise assumptions on sanitation changes in line with the storylines and population and urbanisation changes from the SSPs. In SSP1 population growth is limited, urbanisation large and sanitation and waste water treatment strongly improve. SSP1* is the same as SSP1, but waste water treatment does not improve. SSP3 sees large population growth, moderate urbanisation and sanitation and waste water treatment fractions that are the same as in 2010. Total global *Cryptosporidium* emissions to surface water for 2010 are estimated to be 1.6×10^{17} oocysts per year, with hotspots in the most urbanised parts of the world. In 2050 emissions are expected to decrease by 24% or increase by 52% and 70% for SSP1, SSP3 and SSP1* respectively. The emissions increase in all scenarios for countries in the Middle East and Africa (MAF) region, while emissions in large parts in Europe decrease in scenarios SSP1 and SSP3. Improving sanitation by connecting the population to sewers, should be combined with waste water treatment, otherwise (SSP1*) emissions in 2050 are expected to be much larger than in a situation with strong population growth and slow development of safe water and improved sanitation (SSP3). The results show that

population increase, urbanisation and changes in sanitation should be considered when water quality and resulting health risks are estimated by water managers or public health specialists.

This chapter is based on: Hofstra, N., Vermeulen, L.C., 2016. Impacts of population growth, urbanisation and sanitation changes on global human *Cryptosporidium* emissions to surface water. Int. J. Hyg. Environ. Health 219, 599–605. doi:10.1016/j.ijheh.2016.06.005

4.1 Introduction

Cryptosporidium is a pathogenic protozoan parasite and is a leading cause of diarrhoea worldwide (Shirley et al., 2012). *Cryptosporidium* transmission occurs via the faecal-oral route through direct contact or after ingestion of contaminated food or water. People exposed to a limited amount of contaminated drinking water, recreational water and food irrigated with contaminated water can be infected (e.g. DuPont et al., 1995). The global burden of diarrhoeal disease caused by *Cryptosporidium* is very uncertain (Checkley et al., 2015; Shirley et al., 2012). Reasons are a lack of diagnostic tools in developing countries (Shirley et al., 2012), but also substantial under-recognition and under-diagnosis in the developed world (Checkley et al., 2015). The global burden of disease is, however, important for health decision-making and planning processes (WHO, 2008). The disease burden could be estimated using quantitative microbial risk assessment (QMRA) (Howard et al., 2006). Risk of infection is estimated from the probability of exposure and dose-response relations. The probability of exposure is linked to the concentration in surface water (Medema et al., 2009), either directly (in the case of recreation) or indirectly (for drinking water and irrigated food). Therefore the concentration of the pathogens in surface water should be studied.

A global overview of *Cryptosporidium* concentrations in surface water does not exist. Observational data are sparse, particularly in developing countries (Medema et al., 2009). Monitoring programmes are expensive and measuring *Cryptosporidium* properly is difficult, for instance, because the procedure comprises several steps and recovery rates of *Cryptosporidium* are often low (Smith and Thompson, 2001). Modelling provides opportunities. For instance, modelling provides the opportunity to generate spatially continuous concentrations. A first exploration of global *Cryptosporidium* emissions to surface water showed that emissions are largest in big cities in China, India and Latin America (Hofstra et al., 2013). Here we present an improved version of the human emissions part of that model, the GloWPa-Crypto H1 model (the Global Waterborne Pathogen model for Human *Cryptosporidium* emissions version 1) that now also includes emissions from the population not connected to sewers. Although the model may not reproduce *Cryptosporidium* emissions to surface water exactly, it will help to identify hot-

spot regions, the main drivers and to develop strategies to avoid water pollution by *Cryptosporidium*.

Exposure to *Cryptosporidium* and concentrations of the pathogen in surface water will change with changing population size, urbanisation and sanitation. To get an understanding of the impact of these changes on the human *Cryptosporidium* emissions, we exploit another opportunity of modelling, which is that we can explore plausible futures under global change scenarios. In this paper we present the first application of Shared Socio-economic Pathways (SSPs), developed by the climate community to study climate change adaptation and mitigation, to estimate future (2050) emissions of *Cryptosporidium* to surface water. This approach provides ample opportunities to assess the importance of future changes in population, urbanisation and sanitation on global human emissions of *Cryptosporidium* to surface water.

4.2 Methods

4.2.1 GloWPa-Crypto H1 model

We apply the GloWPa-Crypto H1 model. The GloWPa-Crypto H1 model calculates human emissions of *Cryptosporidium* to the surface water. We identify three ways the emissions can reach the surface water. First of all, emissions by the population connected to a sewer reach the surface water after treatment, or directly. Secondly, emissions by the population using hanging toilets or practicing open defecation in urban areas are entered into the surface water directly. And finally, emissions by the population that practice open defecation in rural areas reach the surface water after runoff takes the faeces to the water, so these emissions can be considered diffuse emissions. The emissions by the population with access to on-site systems, such as latrines or septic tanks, are ignored, under the assumption that the faeces stay in the system or that the retention time is long enough that the oocysts will have decayed. This assumption may cause an under-estimation, as contents of septic tanks are regularly dumped on vacant land (e.g. AECOM and SANDEC/EAWAG, 2010).

This model builds on the human emissions part of the first explorations explained in Hofstra et al. (2013) and incorporates the improvements for sanitation types suggested in Chapter 3. The model we apply is described in detail for Rotavirus in Kiulia et al. (2015). Model variables specific for *Cryptosporidium* are the average oocyst excretion per person and the removal rates by the waste water treatment systems. Contrary to Kiulia et al. (2015), we do not specify excretion rates for different age categories. The excretion rates used are based on the literature, as explained in Hofstra et al. (2013) and are computed as the infection rate (10% in developing countries (Human Development Index (HDI) lower than 0.785) and 5% in developed countries) x the excretion rate per ill person (10^9). So for

developing countries the average overall excretion rate used is 1 x 10^8 oocysts/person/year, for developed countries the rate used is 5 x 10^7 oocysts/person/year. Treatment removal rates are also estimated from the literature as explained in Hofstra et al. (2013) and are defined as follows: primary treatment equals 10% removal; primary + secondary treatment equals 55% removal and primary + secondary + tertiary treatment equals 97.75% removal.

4.2.2 Scenario analysis

We use data for population and urbanisation for the year 2050 from the Shared Socio-Economic Pathways developed for the climate change community (from: https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about) and make rough assumptions for sanitation based on the scenario storylines presented by O'Neill et al. (2017). As a first exploration to use the model with scenarios, we select the scenarios SSP1 and SSP3 as two plausible futures. These scenarios are two extreme examples in terms of population growth, urbanisation, focus on the environment and equity and these are all factors that influence the *Cryptosporidium* emissions. Table 4.1 provides an overview of the changes in the scenarios.

SSP1 is entitled "Sustainability - Taking the green road" and focuses on sustainability, well-being and equity. Inequalities between and within countries are decreasing. Education, health, sanitation and safe water are improved and standards of living are overall high. Technological development is rapid and focusses on environmentally friendly processes that are available to all countries (O'Neill et al., 2017). In this scenario population growth is low (KC and Lutz, 2017) and urbanisation is high, but well managed (Jiang and O'Neill, 2017). We assume that access to sanitation and waste water treatment is strongly improved in this scenario. In 2050 the global population has access to a sewer or a pit latrine or septic tank. In urban areas the population that previously caused direct emissions is now connected to a sewage system (75%) or uses an on-site system (25%). In rural areas the population that previously caused direct or diffuse emissions is now connected to a sewage system (25%) or uses an on-site system (75%). Sewage treatment levels have improved. In the developed countries (HDI in 2010 > 0.785) waste water treatment is always tertiary. In developed countries one third of the waste water receives primary treatment, one third up to secondary treatment and one third up to tertiary treatment.

Table 4.1: Total (population) and mean (other variables) values for 2010 in the input variables population growth, urbanisation, sanitation and waste water treatment for scenarios SSP1, SSP1* and SSP3 for 2050 for the five different SSP-regions Asia, Latin America (LAM), Middle East and Africa (MAF), OECD and Reforming Economies (REF). Population for 2010 are aggregated data from LandScan 2010 (Bright et al., 2011), changes are calculated using data from the SSP database (https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about). Urban population and sanitation values are mean percentages over the countries in a region, not corrected for population.

	2010	2050 SSP1	2050 SSP1*	2050 SSP3
Population (number)				
ASIA	3.6E+09	4.1E+09	4.1E+09	4.9E+09
LAM	5.8E+08	6.7E+08	6.7E+08	8.5E+08
MAF	1.2E+09	2.1E+09	2.1E+09	2.7E+09
OECD	1.1E+09	1.3E+09	1.3E+09	1.1E+09
REF	2.8E+08	2.6E+08	2.6E+08	2.9E+08
Total	6.8E+09	8.4E+09	8.4E+09	9.8E+09
Urban population (%)				
ASIA	45	74	74	51
LAM	68	87	87	74
MAF	49	76	76	56
OECD	73	89	89	78
REF	53	80	80	59
Mean	57	81	81	63
Sanitation (%)				
Connected				
Urban				
ASIA	33	38	38	33
LAM	55	57	57	55
MAF	29	37	37	29
OECD	89	89	89	89
REF	66	67	67	66
Mean	50	55	55	50
Rural				
ASIA	8	13	13	8
LAM	12	15	15	12
MAF	7	15	15	7
OECD	35	35	35	35
REF	8	11	11	8
Mean	15	19	19	15
Direct				
Urban				
ASIA	6	0	0	6
LAM	2	0	0	2
MAF	10	0	0	10
OECD	1	0	0	1
REF	2	0	0	2
Mean	6	0	0	6
Rural				
ASIA	2	0	0	2
LAM	0	0	0	0
MAF	0	0	0	0
OECD	0	0	0	0
REF	1	0	0	1
Mean	0	0	0	0
Diffuse				

Rural				
ASIA	20	0	0	20
LAM	12	0	0	12
MAF	33	0	0	33
OECD	3	0	0	3
REF	14	0	0	14
Mean	18	0	0	18
Non-source				
Urban				
ASIA	60	62	62	60
LAM	43	43	43	43
MAF	61	63	63	61
OECD	8	8	8	8
REF	32	33	33	32
Mean	43	45	45	43
Rural				
ASIA	67	84	84	67
LAM	76	85	85	76
MAF	60	85	85	60
OECD	61	63	63	61
REF	77	89	89	77
Mean	66	80	80	66
Waste water treatment				
Primary				
ASIA	10	25	10	10
LAM	25	24	25	25
MAF	11	29	11	11
OECD	10	5	10	10
REF	17	31	17	17
Mean	13	20	13	13
Secondary				
ASIA	2	25	2	2
LAM	6	24	6	6
MAF	6	29	6	6
OECD	24	5	24	24
REF	16	31	16	16
Mean	10	20	10	10
Tertiary				
ASIA	0	49	0	0
LAM	0	52	0	0
MAF	1	41	1	1
OECD	36	90	36	36
REF	2	39	2	2
Mean	8	59	8	8
No treatment				
ASIA	88	0	88	88
LAM	69	0	69	69
MAF	82	0	82	82
OECD	30	0	30	30
REF	64	0	64	64
Mean	69	0	69	69

SSP3 is entitled “Regional rivalry – A rocky road” and focuses on regional progress, culture and national security. Little trade and international cooperation in combination with low investments in education and technology result in slow economic growth, widespread poverty, slow developments in technology, health care, safe water and improved sanitation, and low interest in environmental problems (O’Neill et al., 2017). In this scenario, the global population will grow rapidly, except for rich OECD countries (KC and Lutz, 2017) and urbanisation is slow and poorly managed (Jiang and O’Neill, 2017). We assume that the fractions of the population that have access to different sanitation types and wastewater treatment is the same as in 2010. However, the total number of people with access to the sanitation types will increase, because the population grows rapidly in most countries.

To analyse the importance of waste water treatment, we add a third scenario that is a variant on SSP1 (SSP1*). Population growth, urbanisation and access to sanitation remain the same, but the waste water treatment does not keep up with the population connected to sewers, as is currently the case (WHO/UNICEF JMP, 2000), so the percentage of waste water treated (primary, secondary or tertiary) is assumed to be the same as in 2010.

We present the human *Cryptosporidium* oocyst emissions to the surface water using input data for around the year 2010, so model output will represent approximately this year as well. The scenario results for SSP1, SSP1* and SSP3 are computed for the year 2050. We compare global and regional totals for the SSP-regions Asia, Latin America (LAM), Middle East and Africa (MAF), OECD and Reforming Economies (REF) (distribution of countries over the regions is available from: https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=about). We provide a map for *Cryptosporidium* emissions for 2010 at a 0.5 x 0.5 degree latitude x longitude (which equals approximately 50 x 50 km) grid and compare maps of future scenario results on a country by country basis (distribution maps of future population do, to our knowledge, not exist). We identify areas of high emissions (hotspots) and evaluate the most important emission sources.

4.3 Results

Figure 4.1 presents the total global emissions, the regional emissions and the types of emissions. For 2010, total global emissions are 1.6×10^{17} oocysts/year. The urban share of the emissions is 89%. Asia is responsible for 49% of the emissions, followed by LAM at 18%, MAF and OECD at 14%. Of the Asian emissions, 82% is caused by the population connected to a sewer and 15% are direct emissions. In LAM 95% and in OECD even 99% is attributable to the population connected to a sewer. MAF has the largest percentage of

direct emissions (20%). Diffuse emissions are low in all regions (0.03 – 2.3%), because we assume that only 2.5% of diffuse emissions reaches the surface water (based on Ferguson et al. (2007) as explained in Kiulia et al. (2015)) and only few countries have large fractions of diffuse emissions (determined by the rural population practicing open defecation).

Figure 4.2 presents the spatial distribution of the human *Cryptosporidium* oocyst emissions to the surface water for 2010. Hotspot areas are densely populated areas in China, Bangladesh, India, Nigeria, Africa along the Mediterranean coast, Europe, Mexico and the east and west coasts of Latin America. Highest average per capita emissions are 1.4 x 10^8 oocysts/person/year for Colombia and Fiji. In these countries a large share of the population is connected to a sewer, in both urban and rural areas, but only a limited fraction of the sewage is treated.

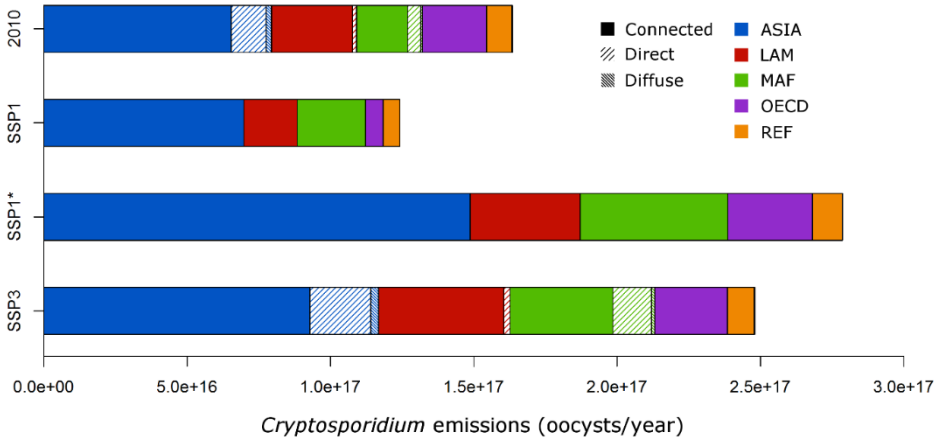


Figure 4.1: Global total *Cryptosporidium* emissions to surface water. Emissions have been specified for SSP-regions (Asia, Latin America (LAM), Middle East and Africa (MAF), OECD and Reforming Economies (REF), different colours) and for the different types of emissions (connected, direct and diffuse, different shadings).

Figure 4.1 also presents the estimated future total global emissions, the regions and the types of emissions. For SSP1, with moderate population growth, planned urbanisation and strong sanitation and water treatment improvements, the global total emissions decrease in 2050 and are expected to be 76% of the 2010 emissions. In this scenario the improved sanitation and waste water treatment weigh up to population growth. Direct and diffused emissions have disappeared, so all emissions are from the population connected to a sewer. For all regions, except for MAF, the emissions decrease (Table 4.2). The decrease is largest for OECD countries, because there everyone connected to a sewer obtains tertiary treatment. In MAF the increase is due to a strong increase in population and urbanisation

together with sanitation changes. The scenario results also show changes in the contribution of the different regions to total emissions. Asia is responsible for 56% of the emissions, an increase compared to the situation in 2010. MAF's share of the emissions also increased, from 14 to 19%. LAM's share of the emissions decreased from 18% to 15% and OECD's share of the emissions decreased from 14% to 5%. Globally, the urban share of emissions increases to 92%.

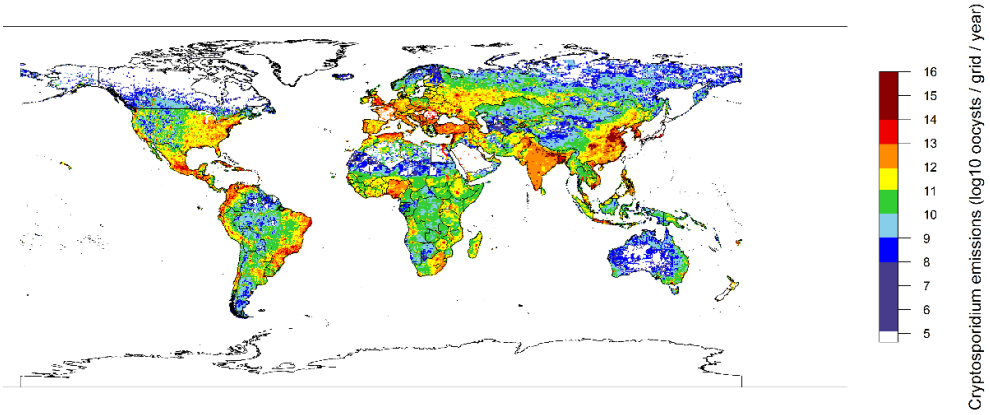


Figure 4.2: A map of *Cryptosporidium* emissions to surface water in oocysts/grid/year based on data for approximately the year 2010

Figure 4.3b&e and Table 4.2 paint the same picture for SSP1, with increases in MAF, both increases (particularly in India and Pakistan) and decreases in Asia, and decreases in other parts of the world. The strongest increase is found in Niger, Burundi, Rwanda and Uganda, where emissions nearly quadruple. In these countries the population strongly increases (in Niger and Uganda the population more than doubles) and the countries see strong urbanisation. In urban areas much more sewage reaches the surface water via sewers than in rural areas where people with pit latrines or septic tanks are assumed not to have any emissions and where of the faeces excreted by the population practicing open defecation only 2.5% are mobilised. So in these countries, population growth together with urbanisation do not weigh up to sanitation and waste water treatment improvements and emissions are expected to continue increasing for SSP1.

Argentina is one of the countries where emissions decrease the most. Argentina is classified as a developed country, but currently has only limited waste water treatment (in 2010 33% primary and 9% primary+secondary are assumed). In SSP1 all developed countries will obtain full coverage of tertiary treatment. So despite a population that is 111% of the population in 2010, the emissions in 2050 are strongly reduced to 2.9% of the 2010 emissions in Argentina for SSP1.

Table 4.2: Total emissions per region and for the world for 2010 and changes for SSP1, SSP1* and SSP3 for 2050

	2010 *Cryptosporidium* emissions (oocysts/year)	**2050 SSP1 Change in emissions (%)**	**2050 SSP1* Change in emissions (%)**	**2050 SSP3 Change in emissions (%)**
ASIA	7.9×10^{16}	-12	+87	+47
LAM	3.0×10^{16}	-38	+29	+54
MAF	2.3×10^{16}	+4	+125	+120
OECD	2.2×10^{16}	-73	+31	+13
REF	9.0×10^{15}	-34	+18	+06
TOTAL	1.6×10^{17}	-24	+70	+52

In the SSP1* scenario, population, urbanisation and sanitation changes are equal to SSP1, but waste water treatment does not keep up with efforts to connect the population to sewers. *Cryptosporidium* emissions for SSP1* are highest of all scenarios (Figure 4.1). Total emissions increase to 170% of the situation in 2010. Emissions increase strongest in Asia (+87%) and MAF (+125%) (Table 4.2).

Figure 4.3c&f show increases in nearly all countries for SSP1*. Only the Baltic States, Uruguay and Vietnam show slight improvements. Vietnam has large (13%) direct rural emissions. These are due to hanging toilets. While Vietnam's population becomes more urban, the remaining rural population gets access to mostly latrines and septic tanks that we assume to be a non-source, so the direct rural emissions disappear. The Baltic States' and Uruguay's population are decreasing, while their sanitation situation does not change much, resulting in reduced emissions compared to 2010. The strongest increase is found, again, in Niger. If population and urbanisation rates and the sanitation situation changes as explained for SSP1, but the waste water treatment does not improve, then the emissions in 2050 are expected to increase 8.6 fold for Niger. This result highlights the importance of waste water treatment. Only connecting the population to sewers is insufficient to improve the surface water quality.

The SSP3 scenario demonstrates for 2050 increases in *Cryptosporidium* emissions due to strong population growth (except for rich OECD countries), moderate urbanisation and sanitation usage proportional to the current situation (Figure 4.1). Total emissions are expected to increase to 2.5×10^{17}, an increase of 52%. The share of urban emissions is 91%. Asia is once again the strongest contributor at 47%, but the relative contribution by MAF has increased strongest, to 20% of total emissions. The change is mostly, but not only, driven by population growth. For instance in China, the population decreases, but

urbanisation causes an increase in emissions of 20%. In SSP3 the Chinese urban population is mostly connected to a sewer with limited treatment.

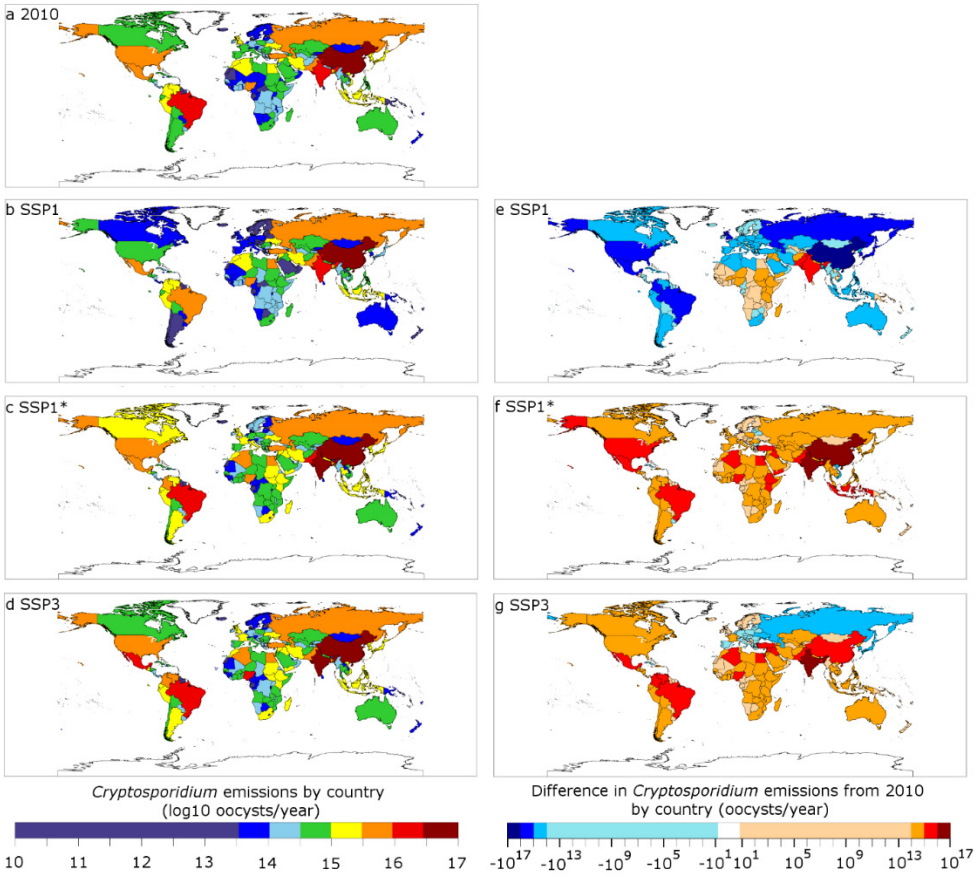


Figure 4.3: Maps of country-level *Cryptosporidium* emissions to surface water for 2010 (a) and 2050 for SSP1 (b), SSP1* (c) and SSP3 (d) and differences between 2050 and 2010 for SSP1 (e), SSP1* (f) and SSP3 (g).

In Asia and MAF for SSP3 the share of direct urban *Cryptosporidium* emissions increases for 2050 compared to 2010. For MAF this share increases from 20% to 26%. According to our assumptions, for SSP3 for 2050 3.8% of the world population produces direct emissions, compared to 2.8% in 2010. These direct emissions are 15% of total emissions for SSP3, compared to 11% in 2010. For all areas the share of the diffuse emissions remains the same. Figure 4.3d&g show the increases in *Cryptosporidium* emissions that are expected nearly everywhere for SSP3 in 2050. European countries that show a decrease in emissions have reducing populations. Rich European countries that do show an increase in emissions (France, UK, Scandinavia) have an increased population for SSP3.

Only Finland has a reduced population, but there urbanisation caused the increase. India shows the strongest absolute increase in emissions (a near doubling), while again Niger has the strongest relative increase (5-fold). Population growth and urbanisation in combination with the distribution of emission types over urban and rural population cause these increases.

4.4 Discussion

We have presented the results of the new GloWPa-Crypto H1 model for 2010 and for 3 scenarios for 2050. The GloWPa-Crypto H1 model is based on large existing datasets and assumptions on excretion rates and waste water removal rates taken from the literature. The main aim of this model is to get an understanding of patterns and changes due to changes in population, urbanisation and sanitation.

A sensitivity analysis for the model was performed in Chapter 3 for India and Bangladesh. The variable to which the model is by far the most sensitive is the excretion rate. A doubling in the number of infected people together with a 1log increase in the excretion by an infected person could lead to a 1900% increase in total emissions. We distinguish two different excretion rates for developed and developing countries, just like Shirley et al. (2012) do when they study the burden of cryptosporidiosis. Although excretion rates likely vary within and between countries and from year to year, the excretion rate mostly determines the order of magnitude, rather than spatial differences. More detailed information on the burden of cryptosporidiosis and excretion rates of infected people for different countries or areas, would improve the model. Overall, however, the patterns of hotspot areas that we see in our results would still be valid.

The model is also relatively sensitive to leakage of septic tanks and pit latrines, that is currently not included in our model. If half of the urban and rural emissions from the septic tanks and pit latrines reach the surface water, the emissions are 284% higher in case of Bangladesh and 85% higher in case of India. Septic tank and pit latrine leakage events are sporadic events, for instance in case of flooding (Tumwebaze et al., 2013), or illegal and undocumented, in case of illegal dumping of the waste (UN Water, 2015). Therefore, currently, there is no way to include these emissions in the model. Nevertheless, in many cases the assumption that waste from septic tanks and in particular pit latrines does not reach the water holds true. Even when sludge is disposed on the land, only a minor part will reach the surface water (we have assumed 2.5% of diffuse emissions reaching the surface water with runoff in this study) and *Cryptosporidium* may have decayed during the storage in the tank or latrine.

We have for the first time analysed changes in waterborne pathogen (in this case *Cryptosporidium*) concentrations in surface water with the new SSPs. We interpreted

values for sanitation and waste water treatment, because sanitation scenarios do, to our knowledge, not exist. Thus far, only some sanitation improvement scenarios have been explored at the local (e.g. Bao et al., 2013) and global scale (Haller et al., 2007). When developing scenarios, using internally consistent element combinations is essential (Schweizer and O'Neill, 2014). In our case, we interpreted improvements in sanitation in relation to technological development and focus on the environment. The current study should be seen as a first step towards the development of future sanitation scenarios, to show opportunities. A survey and workshops with experts in the field to determine crucial factors and their development, as was done to develop the SSPs (Schweizer and O'Neill, 2014), would benefit the quality of such scenarios.

One issue that could be important in the scenario analysis, is that the average per capita excretion rate could change in the future, particularly considering the sensitivity of the model to the excretion rate. In a world with lower emissions, the number of infections could decrease and therefore the excretion rate could decrease. Similarly, in a world with higher emissions, the number of infections could increase and the excretion rates as well. Also hygiene would have an influence. If, in the future, hygiene standards increase, the number of infections could decrease and the excretion rates as well. These effects are, however, very uncertain and the implementation of scenarios on epidemiology or behaviour were outside the scope of our study. A second issue that we did not account for is that technological advancement could increase the *Cryptosporidium* removal of the different waste water treatment types. The aim of the scenario analysis was to get an understanding of how the emissions change if population, urbanisation and sanitation change in the future and we have been able to do just that with the presented scenarios.

The GloWPa-Crypto H1 model currently only estimates emissions of the population to the surface water. We do not yet calculate concentrations in the surface water. Combining the model used in this paper with an animal emissions model, as *Cryptosporidium* is a zoonotic pathogen, and hydrology input will provide those concentrations. Moving to concentration estimates is essential for validation, but also to get an understanding of infection risk. The infection risk is related to concentrations in surface water through exposure and dose-response information (e.g. Teunis et al. 1997). In many regions more emissions will mean higher concentrations and therefore higher risk of infection to the population. Extrapolating that chain of thought would indicate that the risk of infection would decrease in the future for SSP1, except for the most vulnerable MAF region, and that the risk of infection would increase in the future in most parts of the world for SSP1* and SSP3. However, future changes in hydrology and exposure should also be accounted for. Exposure can be influenced by drinking water treatment and behaviour, among others.

4.5 Conclusion

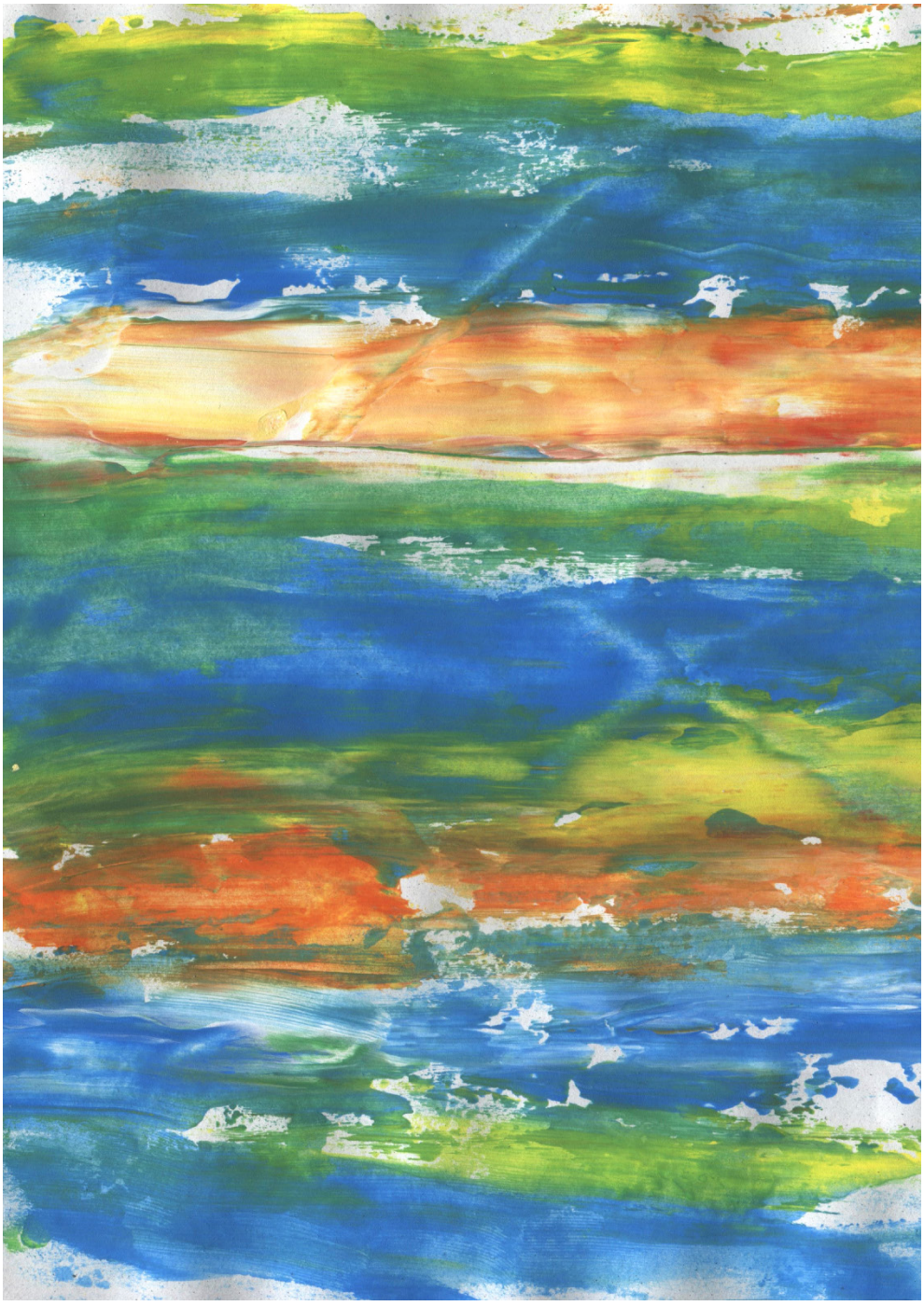
In this paper we presented for the first time a global application of the improved GloWPa-Crypto H1 model and we studied changes in *Cryptosporidium* emissions to the surface water for 2010 and for 2050 for the scenarios SSP1, SSP1* and SSP3. Total global emissions for 2010 are 1.6 x 10^{17} oocysts per year. Hotspot areas are densely populated areas in China, Bangladesh, India, Nigeria, Africa along the Mediterranean coast, Europe, Mexico and the east and west coasts of Latin America.

If population, urbanisation and sanitation change according to SSP1 in 2050, total global *Cryptosporidium* emissions to surface water will be reduced by 24%. In the Middle East and Africa region and India and Pakistan emissions will increase for SSP1 due to a combination of population growth and urbanisation. For SSP3 the total global emissions in 2050 are expected to increase by 52%. Only rich OECD countries see reduced emissions due to population decreases. For the rest of the world increases in emissions are expected. SSP1* provided the opportunity to study what would happen if population, urbanisation and sanitation would change according to SSP1, but waste water treatment would not be improved. In that case, emissions for nearly all countries are even higher than the emissions calculated for SSP3. Total global emissions in 2050 for SSP1* are expected to increase by 70%.

Changes in emissions of *Cryptosporidium* have consequences for the exposure of the population to *Cryptosporidium.* Increased emissions are expected to lead to increased concentrations of surface water, putting swimmers, but also people dependent on surface water for drinking at higher risk of infection. Our scenario analysis shows that population growth, urbanisation, sanitation and waste water treatment should be taken into account when water quality and resulting health risks are estimated by water managers or public health specialists.

Acknowledgement

Lucie Vermeulen is supported by the European Climate-KIC PhD programme.



Chapter 5

Prevalence of *Cryptosporidium* infection in livestock and oocyst concentrations in manure

Nynke Hofstra*, Lucie C. Vermeulen*, Jorien Benders and Gertjan Medema
*Shared first authors

Abstract

Prevalence of *Cryptosporidium* infection and *Cryptosporidium* oocyst concentration in livestock manure are important data to estimate environmental loads. This paper summarises the results of a systematic literature inventory for prevalence of *Cryptosporidium* infection in livestock and *Cryptosporidium* oocyst concentrations worldwide and analyses their distribution by age and region. 231 out of 457 retrieved papers published relevant *Cryptosporidium* infection prevalence data and 62 out of 394 papers provided useful data on oocyst concentrations. Average prevalence was found to be highest in pigs, sheep, cattle and buffaloes (19-22%). Oocyst concentrations in manure are strongly variable (0 – 10^{10} oocysts per gram); mean concentrations range from 2.3 to 4.8 ^{10}log oocysts / g manure. Overall young animals excrete higher concentrations than adults. Most evidence for varying regional prevalence exists for cattle. For adult cattle Africa has the highest mean prevalence at 17% and Oceania the lowest at 0.6%, and calves have the highest mean prevalence in Europe at 30.3% and the lowest in Latin America at 13.2%. Reported data most frequently reflect total *Cryptosporidium* and do not provide information about *Cryptosporidium* species. Influencing variables, such as hygiene standards and housing situation, together with sampling uncertainties (recovery efficiency and detection limit), may cause over- or underestimated prevalence and concentration. Future studies should report on influencing variables, sampling methods and recovery efficiency and detection limit in a more structured way to allow for meta-analysis. This study is the first to quantitatively summarise prevalence of *Cryptosporidium* infection and oocyst concentration in livestock manure and has already been used for a worldwide overview of environmental loads. Such loads are crucial to study environmental *Cryptosporidium* transmission and management options.

This chapter is based on: Hofstra, N., Vermeulen, L.C., Benders, J., Medema, G.J., 2017. Prevalence of *Cryptosporidium* infection in livestock and oocyst concentrations in manure. To be submitted.

5.1 Introduction

Cryptosporidiosis is a diarrhoeal disease in humans and livestock (Robertson et al., 2014; Ryan et al., 2014). Cryptosporidiosis is caused by the enteric protozoan parasite *Cryptosporidium* that comprises many species. Not all species are infectious for all livestock species and humans (Šlapeta, 2013). However, in particular *C. parvum* that is prevalent in ruminants is zoonotic. Evidence exists that young animals are infected with *Cryptosporidium* more often and shed oocysts in higher concentrations than adult animals (Robertson et al., 2014). Many publications discuss prevalence of *Cryptosporidium* infection and oocyst concentrations at farm or herd scale. *Cryptosporidium* has been found in livestock manure worldwide (Robertson et al., 2014), but quantitative information on the geographical occurrence of cryptosporidiosis is lacking. Thus far no overall quantitative overview of prevalence of *Cryptosporidium* infection in livestock and oocyst concentrations in livestock manure exists.

Quantitative data on prevalence of infection and oocyst concentrations in manure are highly relevant to estimate oocyst loads to the environment. Enteric pathogens such as *Cryptosporidium* spread via the faecal-oral route, and often transmission occurs via the environment, such as via surface waters contaminated with human faeces or animal manure (Cacciò and Widmer, 2014). Quantitative insight into environmental loads and the relative importance of different livestock species and age groups contributing to the environmental load is important to be able to estimate disease risk. Furthermore, quantitative insight into how the occurrence of cryptosporidiosis is distributed geographically is important, as this is relevant for human health decision making and manure management strategies. An example of an application of quantitative analysis is given in Chapter 6, where global, spatially explicit environmental *Cryptosporidium* loads are simulated. Two of the main variables in their load model were *Cryptosporidium* infection prevalence and oocyst concentration across the world. This paper provides the underlying literature inventory. The objective of this paper is: to collate and summarise published data on prevalence of *Cryptosporidium* infection in common livestock species and oocyst concentrations in livestock manure from the literature and to analyse their distribution by age and region. We add to the very detailed review of *Cryptosporidium* infection prevalence, association of infection with clinical disease, oocyst excretion and zoonotic transmission by Robertson et al. (2014), which is partly based on Santín and Trout (2007), by adding more studies and extracting the quantitative information about prevalence and concentrations in manure of different livestock species by age and region.

The literature inventory will be useful as an overview of age and spatial differences in prevalence and concentrations, as a basis for other modelling studies and to identify gaps in our understanding.

5.2 Methods

5.2.1 Systematic literature inventory

Prevalence: We did a systematic literature inventory to determine the prevalence of *Cryptosporidium* infection. We differentiated between eleven livestock species, young and adult animals and different world regions. On 03-05-2016 we used the following search string in the Scopus database: (cryptosporidiosis OR cryptosporidium) AND (prevalence OR "infection rate") AND (livestock OR cattle OR cows OR pigs OR sheep OR goats OR chickens OR ducks OR horses OR mules OR asses OR camels OR buffaloes). The results were limited to studies in English and published in the period 2001-2015. Of the 457 results 231 were deemed relevant and included in the analysis. In addition, 19 studies were added that were found in the literature search for concentration and also reported prevalence, leading to a total of 250 included studies reporting on over 150000 individual animals. We extracted data on prevalence in the general population, studies reporting only on prevalence among animals with diarrhoea were excluded. Studies that reported not finding any *Cryptosporidium* infection were also included, as this is relevant for calculating average prevalence in the population.

Concentration: We did a systematic literature inventory to determine the concentration in oocysts per gram manure for an infected animal. We differentiated between eleven livestock species, young and adult animals and possibly different world regions. On 14-4-2016 we used the following search string in the Scopus database: cryptosporidium AND (“infection intensity” OR concentration OR count OR shed* OR excretion OR OPG OR “per gram” OR {/g} OR “per g”) AND (livestock OR cattle OR cows OR pigs OR sheep OR goats OR chickens OR ducks OR horses OR mules OR asses OR camels OR buffaloes). The results were limited to studies in English and published in the period 2001-2015. We did not include inoculation studies, because the inoculation dose may affect the concentration (Zambriski et al., 2013), although that is not found in all studies (Zambriski et al., 2013). We did also not include semi-quantitative studies, because the bin ranges differ among studies and the highest category often has a very large range (e.g. $>10^4$ oocysts/visual range) that cannot be brought back to a concentration per gram of faeces. Of the 394 results, 62 were deemed relevant and included in the analysis. Another six studies identified in the results for prevalence were added, leading to a total of 68 included studies reporting on over 56000 individual animals. Studies that reported not finding any *Cryptosporidium* infection were excluded, as we intend to quantitatively estimate oocyst concentrations in manure from infected animals. Combining estimated prevalence of

infection in a population with the oocyst concentration in manure from infected animals can be the basis of environmental load calculations.

5.2.2 Analysis

Data from the literature were added to a database (Supplementary material S5). Reference, livestock species, age, world region, sample size, reported prevalence and minimum, maximum, median and mean concentration were reported where available. We focussed on generic *Cryptosporidium* rather than on the individual species because i) not all literature reports prevalence and concentrations for individual species and ii) for summaries insufficient prevalence and concentration data for all species are available. Therefore, we included studies reporting infection with any *Cryptosporidium* species and the results chapter provides summaries for *Cryptosporidium* spp.

To determine prevalence for a study we divided the number of animals testing positive for *Cryptosporidium* infection over the total number of animals tested in a study. Concentration was regularly reported in graphs. In those cases the graphs were read using software (http://arohatgi.info/WebPlotDigitizer/app) and the mean concentration in the graph was added to the database. We reported prevalence and concentration separately for young (<3 months) and adult animals and continent. Studies for which the age was not reported were labelled 'Age not specified'. In case the original data do not represent the exact age categories and more information was available in the paper, this information was used to obtain entries for the required age categories.

5.3 Results

Prevalence: To summarise the reported prevalence, we averaged over the prevalence per study, meaning that all studies get equal weight in determining the average. Our results show that cryptosporidiosis is ubiquitous in livestock around the world (Tables 5.1 and 5.2 and Figure 5.1 and 5.2). Average prevalence was found to be highest in pigs, sheep, cattle and buffaloes (19-22%), somewhat lower in goats, camels, ducks and chickens (11-15%) and lowest in horses (7.5%) and asses (0.9%). No studies were identified reporting cryptosporidiosis in mules. Prevalence varies with animal age. For cattle, buffaloes, sheep, goats, horses, pigs and ducks average prevalence in young animals was found to be higher than in adults. Only for chickens and asses average prevalence in young animals was found to be lower than in adults. However, for both these species sample size is too small to draw any meaningful conclusions on age-related prevalence. This may also be the case for other animal species. For example, for goats, horses and chickens prevalence in the group with unspecified age was found to be higher than prevalence for both young and adult animals, making it difficult to draw conclusions.

Table 5.1: Overview of sample sizes in the prevalence systematic literature inventory, specified for livestock species, adult (>3 months) and young (<3 months) animals, and world region.

			Number of animals sampled per age group (Number of studies per age group)			Number of animals sampled per world region (Number of studies per world region)						
Animal species	**Number of studies**	**Total number of animals sampled**	Adult	Young	Age not specified	Africa	Asia	Europe	Latin America	Middle East / North Africa	North America	Oceania
Buffaloes	14	5724	3163 (10)	2380 (11)	181 (3)	-	3730 (8)	-	-	1518 (5)	-	476 (1)
Camels	7	1203	874 (5)	-	329 (2)	-	-	-	-	1203 (7)	-	-
Cattle	153	93868	48221 (90)	27371 (91)	18094 (36)	8628 (20)	23216 (44)	26753 (38)	1535 (5)	7187 (18)	23722 (23)	2645 (5)
Chickens	9	3980	1039 (1)	1176 (2)	1765 (7)	90 (2)	2400 (2)	-	147 (1)	1290 (3)	53 (1)	-
Ducks	3	777	261 (1)	303 (1)	213 (2)	213 (2)	564 (1)	-	-	-	-	-
Goats	28	6665	3357 (13)	2544 (14)	764 (10)	669 (6)	2509 (8)	2452 (8)	105 (1)	919 4)	11 (1)	-
Horses	19	3600	2125 (13)	917 (7)	558 (6)	78 (1)	698 (2)	1378 (7)	486 (2)	521 (4)	439 (3)	-
Pigs	34	24264	8368 (20)	12155 (23)	3723 (9)	1077 (5)	12619 (15)	8951 (10)	-	-	664 (2)	935 (2)
Sheep	43	20086	11285 (22)	5620 (20)	3181 (14)	516 (5)	1196 (7)	5727 (14)	2520 (3)	3549 (6)	35 (1)	6543 (7)
Asses	1	124	116 (1)	8 (1)	-	-	-	-	-	124 (1)	-	-
Mules	-	-	-	-		-	-	-	-	-	-	-

Table 5.2: Overview of reported prevalence, specified for livestock species, adult (>3 months) and young (<3 months) animals, and world region.

	Prevalence (%)			Average* prevalence per age group (%)			Average* prevalence per world region (%)						
Animal species	Overall average*	Mini mum**	Maxi mum**	Adult	Young	Age not specified	Africa	Asia	Europe	Latin America	Middle East / North Africa	North America	Oceania
Buffaloes	21.1	0.0	50.6	14.4	27.0	23.8	-	19.2	-	-	25.4	-	15.6
Camels	12.2	0.0	37.9	16.7	-	1.0	-	-	-	-	12.2	-	-
Cattle	19.6	0.0	80.0	14.5	25.8	17.0	20.1	20.7	21.9	13.1	21.1	14.8	11.7
Chickens	11.6	0.0	34.4	8.9	6.7	13.5	18.2	9.2	-	12.9	13.1	0.0	-
Ducks	14.7	0.0	30.4	0.0	30.4	14.3	14.3	15.2	-	-	-	-	-
Goats	14.1	0.0	77.3	10.6	14.2	18.6	8.3	18.8	15.2	14.5	10.2	0.0	-
Horses	7.5	0.0	27.8	5.8	7.1	11.6	21.8	14.3	4.6	14.3	5.0	6.4	-
Pigs	21.8	0.0	80.0	21.0	25.4	14.1	30.7	19.8	23.2	-	-	15.9	15.8
Sheep	19.8	0.0	96.5	13.2	25.1	22.8	13.1	16.4	24.6	31.7	10.0	0.0	23.5
Asses	0.9	0.0	1.7	1.7	0.0	-	-	-	-	-	0.9	-	-
Mules	-	-	-	-	-	-	-	-	-	-	-	-	-

*Reported average prevalences are averages of the prevalences per study. Prevalence per study is defined as number of positive animals divided by number of animals sampled in that study.

**If a study reported cryptosporidiosis for both young and adult animals, these were recorded separately. Therefore, the reported minimum and maximum prevalence values can reflect a value reported for either young or adult animals and not the overall study prevalence.

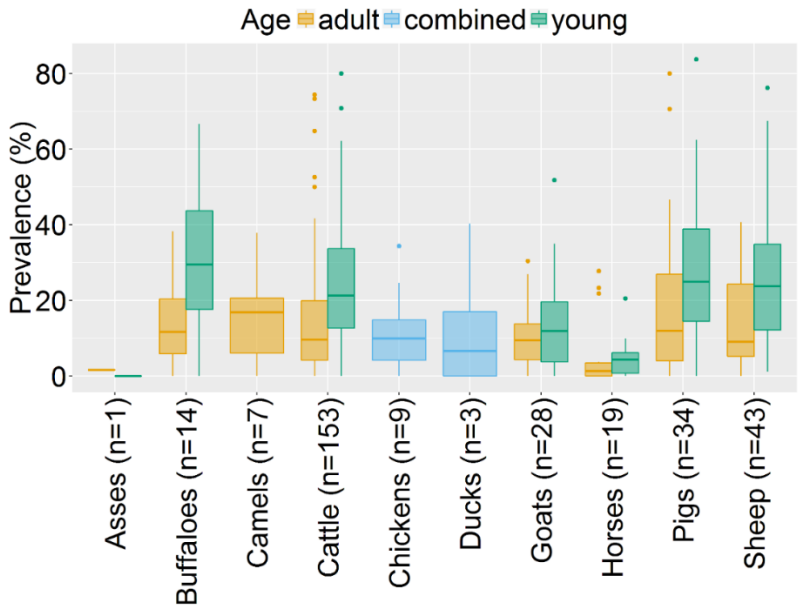


Figure 5.1: Reported prevalence of *Cryptosporidium* infection in livestock manure for 10 livestock species specified by young (< 3 months) and adult animals. Boxplots provide the 25th, 50th and 75th percentiles in the box and whiskers reach to 1.5 times the interquartile range (IQR 75th – 25th percentile). Values outside the IQR are plotted individually. n is the number of studies for each animal species.

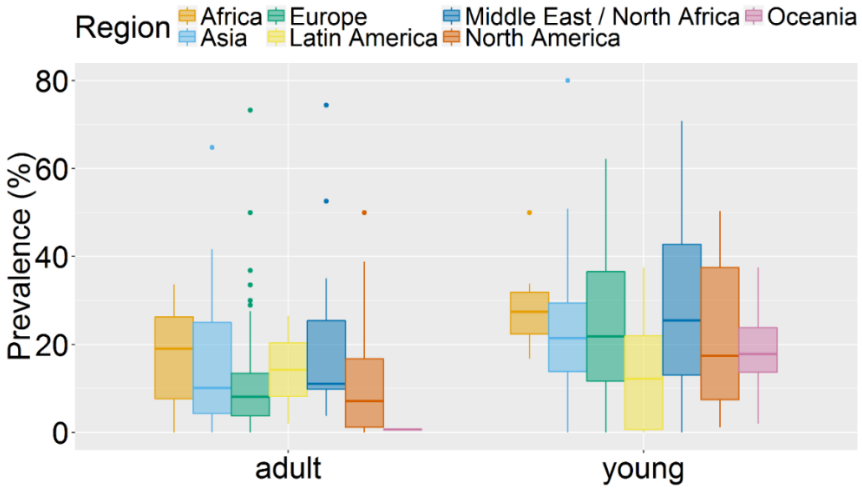


Figure 5.2: Reported prevalence of *Cryptosporidium* infection in cattle for young (< 3 months) and adult cattle specified by region. Boxplots provide the 25th, 50th and 75th percentiles in the box and whiskers reach to 1.5 times the interquartile range (IQR, 75th – 25th percentile). Values outside the IQR are plotted individually. Sample size for each boxplot is equivalent to Table 5.3.

Prevalence varies over world regions, and patterns are not consistent for the different animal species. For cattle, highest average prevalence was found in Europe. For pigs, horses and chickens, highest prevalence was found in Africa, for goats in Asia and for sheep in Latin America. However, the spatial coverage of studies is limited. With 153 studies, evidence is strongest for cattle (Tables 5.3 and 5.4). High prevalence (20-22%) among cattle is reported for Europe, Africa, Asia and the Middle East/North Africa region. Lower prevalence among cattle (11-15%) is reported for North America, Latin America and Oceania. In all world regions, calves are consistently found to have a higher prevalence of cryptosporidiosis than adult cattle.

Table 5.3: Overview of sample sizes for cattle, specified for world regions and age groups.

	Number of animals sampled (Number of studies)						
Cattle	Africa	Asia	Europe	Latin America	Middle East / North Africa	North America	Oceania
Overall	8628 (20)	23216 (44)	26753 (38)	1535 (5)	7187 (18)	23722 (23)	2645 (5)
Adult	3338 (10)	14066 (28)	13546 (25)	179 (1)	4350 (11)	12388 (14)	354 (1)
Young	1463 (7)	7014 (29)	9284 (22)	1356 (5)	2614 (13)	3808 (12)	1832 (3)
Age not specified	3827 (8)	2136 (7)	3923 (8)	-	223 (3)	7526 (8)	459 (2)

Table 5.4: Mean of reported prevalence for cattle, specified for world regions and age groups.

	Prevalence (%)			Average prevalence per world region (%)						
Cattle	Average	Minimum	Maximum	Africa	Asia	Europe	Latin America	Middle East North Africa	North America	Oceania
Overall	19.6	0.0	80.0	20.1	20.7	21.9	13.1	21.1	14.8	11.7
Adult	14.5	0.0	73.3	17.0	15.9	13.4	12.8	17.1	10.8	0.6
Young	25.8	0.0	80.0	28.9	24.7	30.3	13.2	26.1	25.1	17.9
Age not specified	17.0	0.0	68.4	16.4	23.0	26.4	-	13.5	6.3	8.0

Concentration: The concentration data reported in the literature vary widely. Most often the mean was reported, sometimes minimum and maximum or median only. To enable comparison and to summarise the data, we used the reported mean concentrations. In case no mean concentration was reported, we took an average of the minimum and maximum ((min+max)/2) and in case one or both of the minimum and maximum were not reported, we took the median concentration. We summarise medians of mean concentrations for cattle, pigs, sheep, goats and horses (Tables 5.5 and 5.6, and Figure 5.3). For other livestock species insufficient or no studies were available. Medians of the mean concentrations are available in Table 5.6. Figure 5.3 shows the spread in the data. Note that for pigs the 'age not specified' category is added to young and for goats to adults, because the median of mean values were for pigs larger than the young category and for goats smaller than the adults category, which means that most likely the 'age not specified' category mostly contained young pigs and adults goats. Figure 3 shows that concentrations range between 0 and nearly 10 ^{10}log oocysts per gram. Young animals excrete higher concentrations than adults. Medians of mean concentrations were highest for goat kids (5.3 ^{10}log OPG), although only 2 studies determine that value. Other young animal concentrations are in the same order of magnitude (4.4 – 4.9 ^{10}log OPG). Pigs over 3 months of age have highest median of mean concentrations (3.9 ^{10}log OPG) compared to other animals over 3 months of age (2.0 – 2.3 ^{10}log OPG). For horses there are insufficient data to identify specific values for animals under and over 3 months of age, so the overall median of mean concentrations is 3.1 ^{10}log OPG. There are insufficient data available to distinguish regional concentrations.

Table 5.5: Sample size for concentration, specified for livestock species, adult (>3 months) and young (<3 months) animals.

Animal species		Sample size*			
		Overall	Adult	Young	Age not specified
Overall	Total sampled	56709	-	-	-
	Total number of entries	159	-	-	-
	Total number of studies	68	-	-	-
Cattle	Total sampled	36738	14142	19542	3054
	Total number of entries	91	41	40	10
	Total number of studies	48	22	29	10
Goats	Total sampled	899	711	188	-
	Total number of entries	11	9	2	-
	Total number of studies	10	8	2	-
Horses	Total sampled	1278	-	-	-
	Total number of entries	4	-	-	-
	Total number of studies	4	-	-	-
Pigs	Total sampled	11336	2578	8758	-
	Total number of entries	37	11	26	-
	Total number of studies	12	4	13	-
Sheep	Total sampled	6458	3935	1736	787
	Total number of entries	16	7	4	5
	Total number of studies	12	5	4	5

* The sample size is not straightforward for concentration. In some studies the reported values include all animals tested in other studies the reported values only include the positive animals. As reported concentrations are high, reported values for positives of the total population are expected to be in the same order of magnitude, despite the presence of 0 values in the data. Some studies test different animal species for different *Cryptosporidium* species that are entered as different entries in our database, therefore the total number of entries exceeds the total number of studies of 68.

Table 5.6: Minimum, maximum and median of reported arithmetic mean concentrations, separated for age groups. As explained in the results, when arithmetic mean was unavailable (minimum+maximum)/2 was used or the median or the geometric mean. For goats and pigs the data in the category 'Age not specified' were added to the categories adult and young respectively.

Animal species	Minimum (^{10}log oocysts / g faeces)	Maximum (^{10}log oocysts / g faeces)	Median of mean concentration (^{10}log oocysts / g faeces)			
			Overall	Adult	Young	Age not specified
Cattle	0.5	9.3	3.2	2.0	4.9	2.5
Goats	0.0	5.4	2.3	2.2	5.3	-
Horses	0.4	5.5	3.1	-	-	-
Pigs	2.1	6.5	4.6	3.9	4.7	-
Sheep	1.2	6.3	3.4	2.3	4.4	2.7

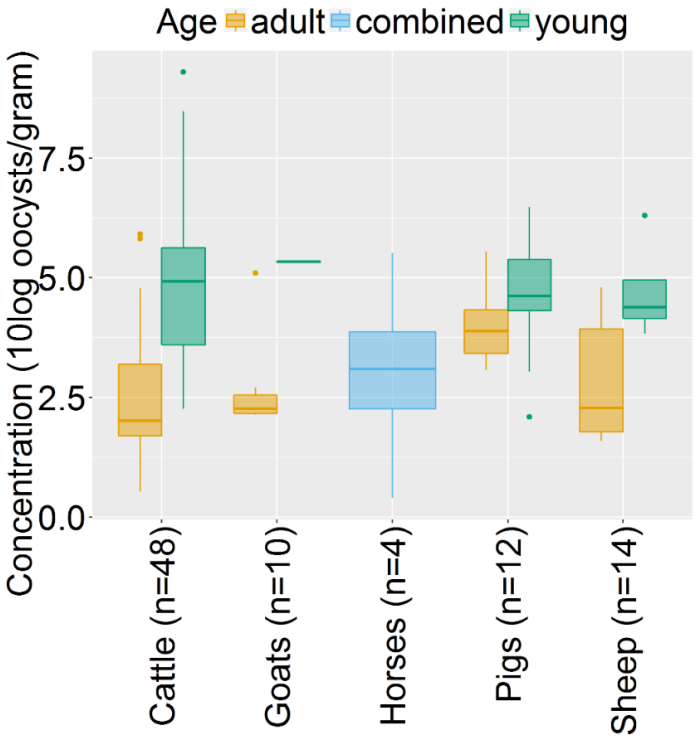


Figure 5.3: Reported *Cryptosporidium* concentration in livestock manure for 5 livestock species specified by young (< 3 months) and adult animals. Boxplots provide the 25th, 50th and 75th percentiles in the box and whiskers reach to 1.5 times the interquartile range (IQR 75th – 25th percentile). Values outside the IQR are plotted individually. n is the number of studies for each animal species.

5.4 Discussion

Cryptosporidium oocysts are omnipresent in livestock manure. This literature inventory shows that wide ranging prevalence and excretion rates are found in the literature. Differences are found between young and adult livestock animals and between world regions. Due to sparse data availability several specifications that are deemed important, such as specification by *Cryptosporidium* species and specification of concentration by region, were impossible. Also uncertainties related to, for instance, reporting of recovery efficiencies and detection limits in the literature are relevant to consider. The discussion focuses on these influencing factors and uncertainties.

5.4.1 Influencing factors

Many different *Cryptosporidium* species exist. The most prevalent *Cryptosporidium* species can change over the ages of the animal. Moreover, excreted concentrations can differ between species, although Björkman et al. (2015) find no statistically significant difference in concentrations of four different *Cryptosporidium* species in a study on beef calves in Sweden. Not all *Cryptosporidium* species will infect all livestock species or humans. Therefore, specification of prevalence and oocyst concentration for the different species would be ideal. However, the data found in our literature study do not justify reporting of prevalence and concentration by *Cryptosporidium* species. Many studies still used immunofluorescence microscopy to identify and count oocysts in manure. Microscopy does not allow for species specification (Efstratiou et al., 2017a). Moreover, an insufficient number of studies for each *Cryptosporidium* species exists to summarise the literature for each species separately. Robertson et al. (2014) qualitatively discuss the prevalence and concentration of the different *Cryptosporidium* species. They identify different *Cryptosporidium* species to be most relevant in different livestock species and at different ages of the livestock species. For example, *C. parvum* is found to be common in pre-weaned calves, relatively common in pre-weaned lambs and present, but less common in pigs. *C. bovis* and *C. ryanae* are common in young-post-weaned calves and *C. andersoni* is common in older post-weaned calves, yearlings and adults. *C. xiaoi* and *C. ubiquitum* are common in older lambs and sheep and *C. suis* and *C. scrofarum* are relatively common in pigs. As in particular *C. parvum* and to a lesser extent *C. ubiquitum* of the aforementioned *Cryptosporidium* species are highlighted for their zoonotic potential, other species could be ignored in loads studies that are used as a basis for risk assessment. Although confusion can arise for newly described species, such as *C. xiaoi* (Fayer and Santín, 2009), we recommend researchers to report the prevalence and concentrations for the different *Cryptosporidium* species separately, to enable use of their data in environmental loading and risk studies.

Cryptosporidium infection prevalence and the severity of the infection is influenced by a wide range of factors. The most commonly recognised factor is age of the livestock animals. Cattle prevalence and concentration peak around two weeks (de Waele et al., 2011; Follet et al., 2011; Kváč et al., 2003), pigs at a slightly higher age of five weeks to three months (Petersen et al., 2015; Ryan et al., 2003; Yui et al., 2014) and sheep around ten days to six weeks (Pritchard et al., 2007; Rieux et al., 2014). Other possibly influencing factors include farm management factors, such as hygiene conditions (Alonso-Fresán et al., 2009; Bomfim et al., 2005), herd size (Craig et al., 2007), housing situation (Duranti et al., 2009; Kváč et al., 2006) and organic versus conventional management (Silverlås et al., 2009; Silverlås and Blanco-Penedo, 2013); dairy versus beef cattle (Castro-Hermida et al., 2002; Duranti et al., 2009; Geurden et al., 2007, 2006; Kváč et al., 2006; Qi et al., 2015; Zhao et al., 2014); length of stay with the mother (Duranti et al., 2009; Silverlås et al., 2009) and preparturient versus periparturient versus post parturient livestock (Atwill et al., 2003; de Waele et al., 2012); latitude (de Waele et al., 2013); and season (Delafosse et al., 2006; Maurya et al., 2013; Noordeen et al., 2001; Starkey et al., 2005).

A range of studies develop statistical models to better understand the influence of the different factors on their observed prevalence and concentration. A meta-analysis would be interesting to better understand the factors that are important overall. However, that was not the objective of this literature inventory. Moreover, many experimental studies were not developed as input for a meta-analysis and do not provide details on many of the potentially influencing factors. The importance of the influencing factors is, though, expected to explain some of the wide range in prevalence and concentrations that was observed in this study. While Robertson et al. (2014) explain that “results from small farms in, for example, India cannot be extrapolated to large-scale farming in, for example, the United Kingdom”, we do put prevalence and concentration values from many different types of farm systems together and summarise these as a basis for load estimation. The data do, unfortunately, not justify further specification of the results. In many cases details of the farming system has not been explained in detail in the reviewed papers and even if they did, spreading in particular the concentration data over livestock animals, age groups, regions ánd farming system or other influencing factors would make the number of distinct groups very large and the number of studies within those groups very small. Our summary, that includes all the relevant literature that our systematic literature inventory selected, provides the most straightforward approach to quantify prevalence and concentration as a basis for load estimation.

5.4.2 Uncertainties from the literature studies

Data of prevalence of *Cryptosporidium* infection and oocyst concentration in manure are inherently variable and uncertain. *Cryptosporidium* sampling is not straightforward and

methods can be time consuming and costly. Already the sample collection can make a difference. Collecting fresh manure from the ground yields lower prevalence than collecting from the rectum directly (Atwill et al., 2006a). Throughout the literature used in this study many different sampling methods have been used. Efstratiou et al. (2017a) has summarised the sampling methods for the three phases particle concentration; selective concentration (purification/separation); and detection, identification and enumeration used to measure *Cryptosporidium* in water samples. For the sampling methods used for manure an equally wide variety of methods is used. The sensitivity of the methods differs. For example, Chalmers et al. (2011) compared different sampling methods regularly used in the UK for stool samples and found that microscopy using Ziehl-Neelsen staining had a significantly lower sensitivity than other microscopy based sampling methods. Moreover, PCR based methods are more sensitive than microscopic methods (Bruijnesteijn van Coppenraet et al., 2009). The uncertainty due to the sampling method is related to their sensitivity (the extent to which the method detects *Cryptosporidium* oocysts), specificity (the extent to which the method does not detect other oocysts for *Cryptosporidium* oocysts), recovery efficiency (the fraction of *Cryptosporidium* oocysts in the manure that is detected) and detection limits (the lowest concentration that can be detected). The latter two are discussed in more detail below.

Each sampling method has preparation and analysis steps during which oocysts can be lost. Each lab should therefore determine their recovery efficiency for the sampling method they use. Recovery efficiencies vary widely and can differ for manure from different livestock species (Kuczynska and Shelton, 1999). Recovery efficiencies can be as low as 0-9% (mean 4%), as provided by Oates et al. (2012) in their study of several livestock and wildlife species. Corrected for the recovery efficiency, their reported mean oocyst concentration of 1870.7 oocysts / g manure for beef cattle would be 46770 oocysts / g manure. That is a difference of 1.4 ^{10}log oocysts per gram. That is as much as half of the difference between adult and young cattle. Given this importance for quantification, recovery efficiencies should be reported with the data in publications. However, not for every study found in this literature inventory recovery efficiencies were mentioned. We therefore did not correct any of the values for the recovery efficiency. The reported concentrations are probably (considerably) lower than the actual concentrations.

Closely related to the recovery efficiency are the detection limits. Also detection limits vary between sampling methods (Kuczynska and Shelton, 1999). When a prevalence or concentration of 0 is reported, the actual prevalence or concentration may not be 0, but is simply below the detection limit. This could cause an underestimation of the prevalence, although mostly low-level shedders are missed that are not responsible for a large part of the environmental loads. While Atwill et al. (2003) reported lower environmental loads after including a more sensitive detection method, for concentration the detection limit is

expected to be less important, in particular because usually mean concentrations are reported. Recently, methods have been developed to detect low oocyst concentrations (Wells et al., 2016). This method lowered their detection limit from 100 oocysts / g to 5 oocysts / g manure. These methods were not used in the papers in this literature inventory, because we limited ourselves to the time period 2001 – 2015. When detection limits are reported, methods exist that estimate a distribution of the concentrations below the detection limit (Lee and Helsel, 2007). However, detection limits are often not reported and for concentration we therefore were unable to include studies that reported concentrations below the detection limit.

For estimating the prevalence, a final uncertainty is a possible positive publication bias. Only few studies have been published that report a prevalence of 0. In total 53 studies reported a prevalence of 0 for one or more age categories. These studies emerge across livestock species and regions. Although prevalence of 0 has been reported, the tendency may be to only publish positive prevalence. If that is indeed the case, our prevalence values may be overestimated. However, a positive publication bias cannot be proven, let alone quantified.

To summarise, specification for different *Cryptosporidium* species, influencing factors and uncertainties due to missing recovery efficiencies and detection limits and a possible positive publication bias introduce uncertainty into the mean prevalence and median of mean oocyst concentrations in this paper. However, the current summaries of the literature provide, for example, individual prevalence data (Robertson et al., 2014) and no overview of manure concentrations. For loads calculations, for which our quantitative outputs are established, assumptions will have to be made about mean prevalence and concentration in manure. That is exactly what this study has established using all the recent literature found with our search criteria.

5.5 Conclusion

Although the importance of livestock for *Cryptosporidium* infection is commonly understood and the literature has been casually reviewed for prevalence before (Robertson et al., 2014), this paper is the first to quantitatively summarise prevalence of *Cryptosporidium* infection and oocyst concentrations in livestock manure and to analyse their distribution spatially and by age category using a systematic literature inventory. We find that *Cryptosporidium* is ubiquitous in livestock manure. Average prevalence was found to be highest in pigs, sheep, cattle and buffaloes (19-22%), somewhat lower in goats, camels, ducks and chickens (11-15%) and lowest in horses (7.5%) and asses (0.9%). Median mean concentrations range from 2.3 to 4.8 ^{10}log oocysts / g manure. Young animals (< 3 months) have higher prevalence and oocyst concentrations than animals

older than 3 months of age. Prevalence varies over world regions, and patterns are not consistent for the different animal species. With 153 studies, evidence is strongest for cattle. High prevalence (20-22%) among cattle is reported for Europe, Africa, Asia and the Middle East/North Africa region. Lower prevalence among cattle (11-15%) is reported for North America, Latin America and Oceania. The data coverage was not sufficient to spatially distinguish oocyst concentrations. More data are required, in particular for prevalence in cattle in Latin America and Oceania and for other livestock species and concentrations worldwide. Moreover, concentration data were only found for five of the eleven livestock species included in our study and in particular chickens would need to be studied further.

Prevalence of *Cryptosporidium* infection and manure concentration data are inherently uncertain and depend on many influencing variables. These influencing variables and uncertainties may result in under- or overestimated mean prevalence and median concentrations. The literature does not allow for quantification of the rate of under- or overestimation. Future publications should consider including a structured description of their herd situation, such as hygiene conditions, housing situation and climatic zone, the recovery efficiency and detection limits of the sampling methods used, and should report prevalence and concentrations distinguishing for *Cryptosporidium* species to allow for inclusion in a meta-analysis.

The summarised prevalence and concentration data can be used and have been used (Chapter 6) to simulate oocyst loads to the environment, which is relevant to better understand concentrations in surface water and transmission from livestock to humans via the surface water. Such modelling efforts will help to find control mechanisms to reduce infection of humans and livestock via the surface water.

Acknowledgements

Lucie Vermeulen is supported by the EU Climate KIC PhD programme.

Chapter 6

Global *Cryptosporidium* loads from livestock manure

Lucie C Vermeulen, Jorien Benders, Gertjan Medema and Nynke Hofstra

Abstract

Understanding the environmental pathways of *Cryptosporidium* is essential for effective management of human and animal cryptosporidiosis. In this paper we aim to quantify livestock *Cryptosporidium* spp. loads to land on a global scale using spatially explicit process-based modelling, and to explore the effect of manure storage and treatment on oocyst loads using scenario analysis. Our model GloWPa-Crypto L1 calculates a total global *Cryptosporidium* spp. load from livestock manure of $3.2x10^{23}$ oocysts per year. Cattle, especially calves, are the largest contributors, followed by chickens and pigs. Spatial differences are linked to animal spatial distributions. North America, Europe and Oceania together account for nearly a quarter of the total oocyst load, meaning that the developing world accounts for the largest share. GloWPa-Crypto L1 is most sensitive to oocyst excretion rates, due to large variation reported in literature. We compared the current situation to four alternative management scenarios. We find that although manure storage halves oocyst loads, manure treatment, especially of cattle manure and particularly at elevated temperatures, has a larger load reduction potential than manure storage (up to 4.6 log units). Regions with high reduction potential include India, Bangladesh, Western Europe, China, several countries in Africa, and New Zealand.

This chapter is based on: Vermeulen, L.C., Benders, J., Medema, G.J., Hofstra, N., 2017. Global *Cryptosporidium* Loads from Livestock Manure. Environ. Sci. Technol. 51, 8663–8671. doi:10.1021/acs.est.7b00452

6.1 Introduction

Cryptosporidium is a protozoan parasite that is found all over the world and can cause diarrhoea in humans and animals (Robertson et al., 2014; Ryan et al., 2014; Šlapeta, 2013). Every year around 1.3 million people die of the consequences of diarrhoea (GBD 2013 Mortality and Causes of Death Collaborators, 2015). *Cryptosporidium* has been identified as one of the six major pathogens responsible for diarrhoea in children younger than 5 years in Africa and Asia (Liu et al., 2016). *Cryptosporidium* is transmitted via the faecal-oral route. Direct contact with faeces of humans or animals is a possible transmission route (Zambrano et al., 2014), but often transmission occurs via an environmental route, such as drinking of or recreation in contaminated water (Shirley et al., 2012) and eating of fresh produce that has been fertilized with manure or irrigated with contaminated water (Dixon, 2016). Livestock, particularly cattle, is considered to be an important reservoir of zoonotic *Cryptosporidium* (Robertson et al., 2014; Ryan et al., 2014). Cryptosporidiosis has been reported in many important livestock species, including cattle, buffaloes, pigs, goats, sheep, horses, camels, asses, chickens, and ducks, and it has been reported on all continents except Antarctica (Robertson et al., 2014; Ryan et al., 2014). Infection of livestock with *Cryptosporidium* can result in decreased production and loss of income for the livestock sector (Shafiq et al., 2015; Sweeny et al., 2011). Not all *Cryptosporidium* species in livestock are of public health significance; the majority of human infections are caused by *C. parvum* and *C. hominis.*

Studying the environmental transmission routes of *Cryptosporidium* is important for assessing and mitigating disease risk, yet observational data of *Cryptosporidium* in the environment are very scarce, as sampling is costly and time-consuming. Especially quantitative information about diffuse sources, such as loads from livestock, is rare. This is where process-based modelling can help; process knowledge can provide insights relevant for managing human and animal cryptosporidiosis when observational data are scarce. Process knowledge on *Cryptosporidium* from livestock manure includes the following. Oocysts, the robust survival stage of the pathogen, are excreted in manure of infected animals, depending on cryptosporidiosis prevalence and the excretion rate (concentration) of oocysts in manure that can vary between livestock species and age groups (Maddox-Hyttel et al., 2006; Maurya et al., 2013; R. P. Smith et al., 2014). Manure is either deposited directly on fields during grazing, stored or treated before it is applied to land, or it is used for other purposes such as burning for fuel. Oocysts can decay during storage and treatment (such as anaerobic digestion) of manure (Garcés et al., 2006; Hutchison et al., 2005; Jenkins et al., 2013). These are all factors that can be incorporated in a process-based model of *Cryptosporidium* loads from livestock manure. From the land, oocysts can spread further through the environment, they can, for example, be transported with runoff to surface waters.

In this paper we aim to quantify livestock *Cryptosporidium* spp. loads to land on a global scale using spatially explicit process-based modelling, and to explore the effect of manure storage and treatment on oocyst loads using scenario analysis. We present the model GloWPa-Crypto L1, a global spatially explicit model at a 0.5x0.5 degree grid that calculates total annual oocyst loads to land from manure of 11 livestock species.

6.2 Methods

GloWPa-Crypto L1 calculates livestock oocyst loads to land worldwide. We define 'load' as the annual total number of oocysts from livestock manure that end up on land. This accounts for all oocysts in manure that is dropped directly on land, and the proportion of oocysts that survive in manure that is stored before it is applied to land. GloWPa-Crypto L1 is programmed in R (R Development Core Team, 2016), it operates on a 0.5x0.5 degree grid and on an annual time step. The model is considered representative for approximately the year 2005. Figure 6.1 shows a schematic overview of major model components. The main input data for the model include: number of animals (from the Gridded Livestock of the World v2.0 (Robinson et al., 2014)), cryptosporidiosis prevalence and oocyst excretion rates (from the extensive literature review in Chapter 5), manure production and storage estimates (from IPCC and USEPA (IPCC, 2006; Safley et al., 1992)), intensive and extensive farming systems (from the IMAGE model (Bouwman et al., 2012)) and ambient temperature (from the WATCH forcing data (Weedon et al., 2011)). GloWPa-Crypto L1 is partly based on an exploratory global model of *Cryptosporidium* loads to surface water (Hofstra et al., 2013). In addition to the earlier work, GloWPa-Crypto L1 includes manure storage and oocyst decay, and prevalence and oocyst excretion are now based on an extensive literature review. GloWPa-Crypto L1 does not differentiate between different species of *Cryptosporidium*, as there is insufficient data available on prevalence and excretion rates of the different species in different livestock animals. Moreover, observational studies on oocysts in the environment usually do not distinguish between different species either, as the antibodies used for environmental surveillance are not specific to human or zoonotic *Cryptosporidium* species only.

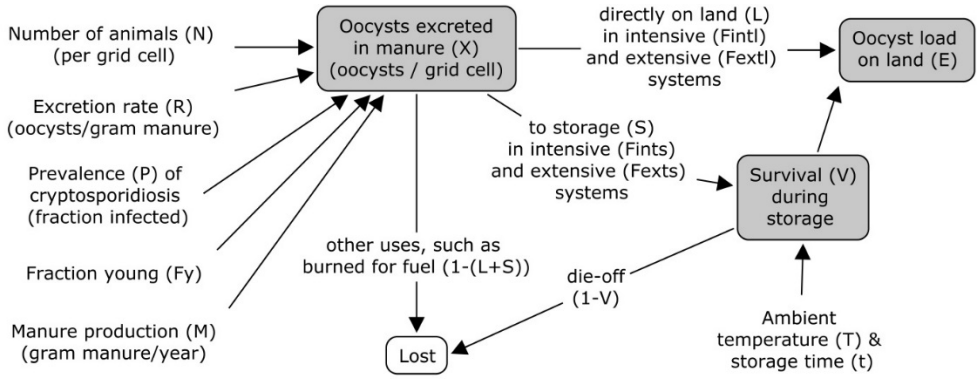


Figure 6.1: Schematic overview of major model components of GloWPa-Crypto L1. Grey boxes represent the major subcomponents that are calculated, the white box represents the oocysts that are lost, and the text without boxes are model inputs. The oocyst load to land (E) is the main model output.

6.2.1 Calculating oocyst excretion in livestock manure (X)

GloWPa-Crypto L1 includes 11 livestock species: cattle, buffaloes, pigs, sheep, goats, horses, camels, asses, mules, chickens and ducks. Oocyst excretion (X) per grid cell for each animal species (i) is calculated as follows:

$$X_i = N_i \times Fy_i \times My_i \times Py_i \times Ry_i + N_i \times (1 - Fy_i) \times Ma_i \times Pa_i \times Ra_i \quad (6.1)$$

Where X_i is the oocyst excretion in a grid cell (oocysts/year) for animal species i, N_i is the number of animals of species i in the grid cell, Fy_i is the fraction of animals of species i that is young (defined as under three months old), My_i and Ma_i are the manure production (M) per head for young (y) and adults (a) of species i (gram manure/year), Py_i and Pa_i are the prevalence (P) of cryptosporidiosis for young (y) and adults (a) of species i (fraction infected), and Ry_i and Ra_i are the excretion rates (R) of oocysts for an infected young (y) or adult (a) animal of species i (oocysts/gram manure).

The numbers of animals on a grid cell (Ni) are taken from the Gridded Livestock of the World v2.0 (Robinson et al., 2014) for six animal species: cattle, sheep, goats, pigs, chickens and ducks. These data were aggregated to a 0.5x0.5 degree grid. For the five other livestock species only country totals were available (FAO, 2013). We assumed that buffaloes are distributed over countries similar to cattle, and that horses, camels, mules and asses are distributed similar to sheep and goats, see the Supplementary material (Sm). Figures S6.1 and S6.2 visualize the distribution of livestock over the world. We define young animals as those under three months old, as prevalence and excretion rates for this group were found to differ from those of adult animals (Table S6.4 in the Sm). The fraction of animals that is young is estimated based on the fertility rates (FAO, 2016) and average

number of offspring per parturition of the different animal species (Table S6.3 in the Sm). Manure production (Ma_i and My_i) was calculated from average body mass of adult livestock, birth weight of young livestock, and manure production per 1000 kg mass (Table S6.1 and S6.2 in the Sm)(IPCC, 2006; Safley et al., 1992). Average prevalence of cryptosporidiosis (Py_i and Pa_i) and average oocyst excretion rate of infected animals (Ry_i and Ra_i) are based on an extensive systematic literature review (Chapter 5 and Table S6.4 in the Sm).

6.2.2 Calculating oocyst survival during manure storage (V)

For oocyst in manure that is dropped during grazing (L), we assume that everything ends up directly on land, according to the following equation:

$$L_i = X_i \times (Fintl_i + Fextl_i) \tag{6.2}$$

Where X_i is the oocyst excretion (oocysts/grid/year) for animal species i (see Equation 6.1), and $Fintl_i$ and $Fextl_i$ are the fraction of animals of species i that are kept in respectively intensive and extensive systems and drop manure on land during grazing.

However, if manure is stored before it is applied to land, oocysts will die off during the storage period. The following equation calculates the number of oocysts in manure that is stored and then spread on land (S):

$$S_{ij} = X_i \times (Fints_i + Fexts_i) \times FS_j \times \overline{V_j} \tag{6.3}$$

Where X_i is the oocyst excretion (oocysts/grid/year) for animal species i (Equation 6.1), $Fints_i$ and $Fexts_i$ are the fractions of animal species i that are kept in respectively intensive and extensive systems of which the manure goes to storage, FS_j is the fraction of stored manure kept in storage system j and $\overline{V}_j$ is the fraction of oocysts that survives during storage in system j, averaged over time (Equations 6.4 and 6.5).

Data on whether animals are kept in intensive or extensive farming systems ($Fintl_i$, $Fextl_i$, $Fints_i$, $Fexts_i$) were taken from the Integrated Model to Assess the Global Environment (IMAGE) according to Bouwman et al. (2012). Data on the use of different storage systems (FS_j) for the different animal species are from the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (IPCC, 2006) and underlying data from a USEPA report on Global methane emissions from livestock and poultry manure (Safley et al., 1992) (see Supplementary material S6.5). We assumed that all manure that is stored is applied to land in the same grid cell after storage. Manure trade was thus ignored.

Average oocyst survival (∇_j) in each storage system depends on storage time and temperature (Equations 6.4 and 6.5). We use an exponential decay function to calculate survival of oocysts over time (V):

$$V = e^{-K \times t} \quad (6.4)$$

Where t is the time (days) and K is a constant, that is dependent on temperature.

We derived a value for K for each grid cell, based on a relation between temperature (°C) and oocyst survival in livestock manure (measured as viability or infectivity) using data from seven studies (Collick et al., 2006; Garcés et al., 2006; Hutchison et al., 2005; Jenkins et al., 2013, 1999; Kinyua et al., 2016; Olson et al., 1999), see the Supplementary material S6.6 for more detail.

Data on the duration of manure storage worldwide are unavailable to our knowledge. Unless a short storage time was explicitly indicated (e.g. systems 'daily spread' and 'Pit storage shorter than one month', see Supplementary material S6.5) we assumed that manure is on average stored for 9 months (274 days), based on an assumed one or two harvests per year per location and spreading of manure at the start of the growing season. We integrated Equation 6.4 over the estimated storage time to get the average survival rate (∇_j) in every grid cell (Equation 6.5). We do an integration, because we assume that a manure storage system continually receives manure, as livestock produce manure every day. To clarify: when manure is accumulated for 9 months and then spread on a field, this manure will be between 1 day and 9 months old.

$$\overline{V_j} = \frac{\int_0^{t_j} V \, dt}{t_j} \quad (6.5)$$

We do not differentiate for other characteristics of manure storage systems because of a lack of data (see Table S6.7 in the Sm), although conditions under which manure is stored may differ per system, and these conditions might influence *Cryptosporidium* survival. Only anaerobic digesters are considered a special case, as they are a type of manure treatment rather than mere storage. Anaerobic digesters can be operated at medium or high temperatures (mesophilic or thermophilic digestion). Garcés et al. (2006) found that oocyst infectivity was reduced by 2 log units after mesophilic digestion, and by over 5 log units after thermophilic digestion. Data on the relative use of mesophilic and thermophilic digestion worldwide are not available, but mesophilic systems are reported to be more stable and most commercial-scale anaerobic digesters are operated at mesophilic temperatures (Labatut et al., 2014). Therefore, we took the conservative estimate that all digestion is mesophilic and assume a 2 log reduction ($\nabla_j = 0.01$).

6.2.3 Calculating the total oocyst load (E)

In GloWPa-Crypto L1, the total oocyst load (E) is defined as the number of *Cryptosporidium* oocysts in a grid cell ending up on land annually (Figure 6.1). The oocyst load in a grid cell is calculated as follows:

$$E = \sum_i L_i + \sum_{ij} S_{ij} \quad (6.6)$$

Where E is the total oocyst load in a grid cell (oocysts/year), L_i are the oocyst loads (oocysts/ year) in a grid cell from manure that is dropped directly on land by animal species i (Equation 6.2). These are summed for all animal species. S_{ij} are the oocyst loads (oocysts/ year) in a grid cell from manure of animal species i that has been stored in storage system j (Equation 6.3). These are summed for all animal species and storage systems.

6.2.4 Sensitivity analysis and scenario analysis

We test the sensitivity of our model to variation in the input variables in a nominal range sensitivity analysis. We change our input variables one at a time, based on the lower and upper end of a reasonable range the value can take. Tables S6.8-6.10 in the Supplementary materials show the input data of the sensitivity analysis.

We compare our baseline model to four alternative management scenarios, assuming that all manure goes directly to land (Scenario 1), all manure goes to storage (Scenario 2), all manure is treated by mesophilic anaerobic digestion (Scenario 3) and all manure is treated by thermophilic anaerobic digestion (Scenario 4). The difference between Scenario 1 and the baseline model represents the reduction in oocyst loads that is currently achieved by manure storage. The difference between the baseline model and Scenarios 2-4 represent the reduction potential. In all scenarios, the fraction of manure that is used for other purposes (e.g. burned for fuel) and leaves the system is unchanged. The assumptions on storage time and temperature in Scenario 2 are the same as in the baseline model. The assumption on the effect on oocyst survival of mesophilic anaerobic digestion (2 log reduction) and thermophilic anaerobic digestion (5 log reduction) are the same as in the baseline model.

6.3 Results & discussion

6.3.1 Oocyst loads to land

Our model calculates a total global oocyst load from animal manure of 3.2 x 10^{23} oocysts per year. Figure 6.2 shows how this is distributed over the world. The patterns are mostly

determined by the distribution of cattle, chickens and pigs in the Gridded Livestock of the World V2.0 data. North America, Europe and Oceania together only account for nearly a quarter of the total oocyst load, meaning that the developing world accounts for the largest share. This is likely mostly due to high animal numbers (Figure S6.1) and limited manure storage in developing countries, as cryptosporidiosis prevalence does not differ greatly between different world regions (Chapter 5 and Table S6.4).

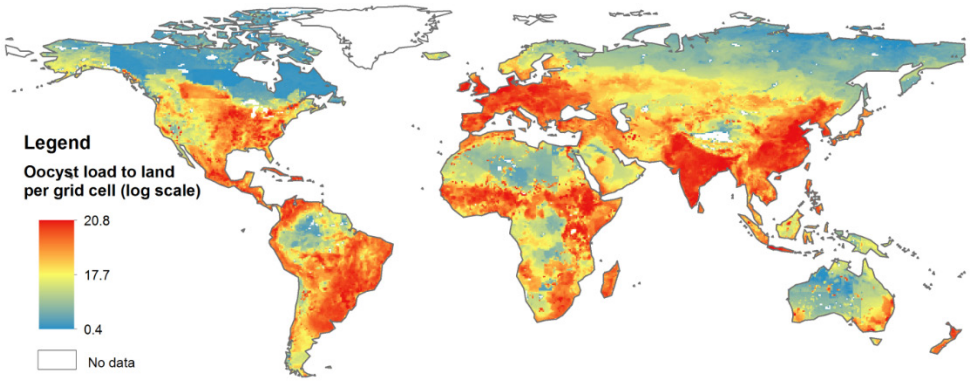


Figure 6.2: Oocyst loads to land (E) per grid cell per year. Grid cell size is 0.5x0.5 degree.

Compared to other studies calculating livestock oocyst loads to land, GloWPa-Crypto L1 model outcomes are in the same range. Atwill et al. (2006a) estimate a load of about $2.8x10^4$ to $1.4x10^5$ oocysts/animal/day for beef cattle from 22 feedlots in 7 states in the USA. With GloWPa-Crypto L1, we estimate for North American adult beef cattle a daily load of $2.9x10^5$ oocysts/animal/day. Starkey et al.(Starkey et al., 2007) estimate the daily *C. parvum*-like oocyst load from dairy cattle across all ages in the New York City Catskill/Delaware watershed to be $4.15x10^{10}$. They assume 258 herds within this watershed with an average of 125.3 animals per herd, hence, the oocyst load/animal/day is $1.28x10^6$. They estimate that preweaned calves (<2 months) produce 99.5% of this load. With GloWPa-Crypto L1, we estimate for North American cattle an average of $5.89x10^6$ oocysts/animal/day, of which 93% comes from young cattle (<3 months). In both cases our load estimate is somewhat higher, and our proportion attributed to calves is somewhat lower, but still in the same order of magnitude. All of these estimates are excluding die-off during storage.

6.3.2 Sources of oocyst loads

In Europe and North America most oocysts come from stored manure, in the other regions oocyst are predominantly excreted directly on land. Asia has the highest total oocyst load, followed by Africa, Latin America and Europe (Figure 6.3a) In Oceania the majority of

oocysts come from extensive systems, in Africa it is approximately equal, and in the other regions intensive systems dominate (Figure 6.3b). We did not assume different cryptosporidiosis prevalence for animals kept in different systems, although there is some evidence that in systems where large numbers of animals live closely together disease prevalence is higher, but the evidence was too weak and the input data too variable to provide a quantitative estimate (Castro-Hermida et al., 2002; Duranti et al., 2009; Lai et al., 2011; Maikai et al., 2011; Muhid et al., 2011; Silverlås and Blanco-Penedo, 2013).

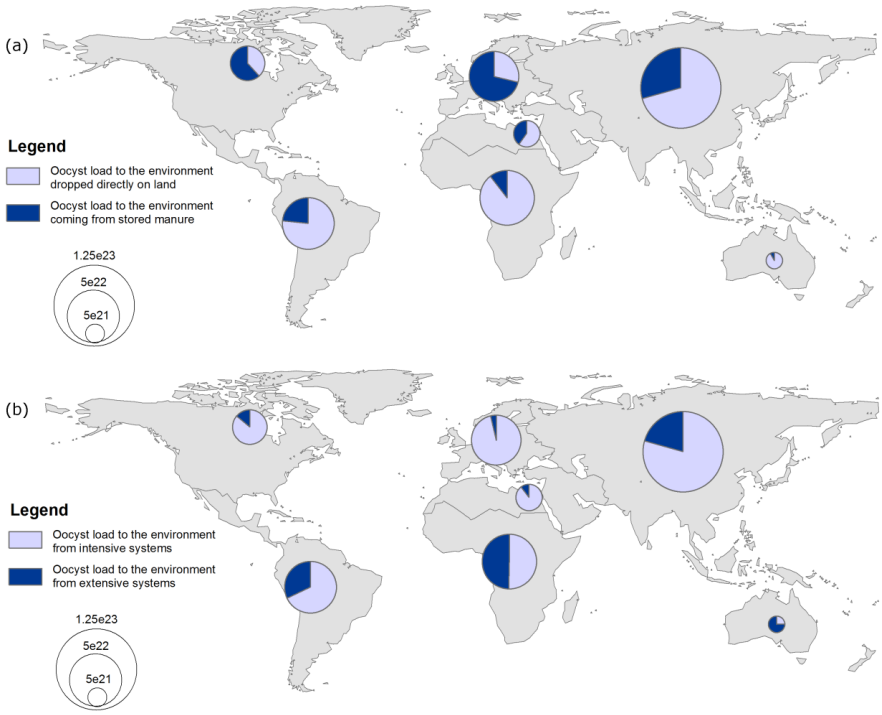


Figure 6.3: Oocyst load per world region, (a) from manure dropped directly on land and from stored manure, and (b) from intensive and extensive systems. Pie chart sizes are proportional to the size of the oocyst load. We distinguish seven world regions: Europe, Asia, Africa, North America, Latin America and Middle East – North Africa (MENA), see Supplementary material S6.1.

On a global scale, cattle are the dominant source of oocysts, followed by chickens and pigs (Figure 6.4). Intensive systems are the largest source of oocysts for most animal species, especially for pigs and chickens. Manure dropped directly on land is the largest source of oocysts for most animal species, except for chickens, pigs and ducks. For cattle, buffaloes, goats and sheep, young animals are the largest source of oocysts. For pigs, adults are the largest source of oocysts, although prevalence and oocyst excretion rates are higher for young pigs than for adults. The reason for this is that adult pigs produce much more manure than young pigs. It should be noted that the literature indicates that

cryptosporidiosis is more prevalent among dairy calves than among beef calves (Duranti et al., 2009; Geurden et al., 2007, 2006; Kváč et al., 2006; Qi et al., 2015), although the literature is not fully consistent on this (Castro-Hermida et al., 2002; Chuah et al., 2016; Dixon et al., 2011; Zhao et al., 2014). We did not make a distinction for dairy and beef cattle because the Gridded Livestock of the World v2.0 does not distinguish between these.

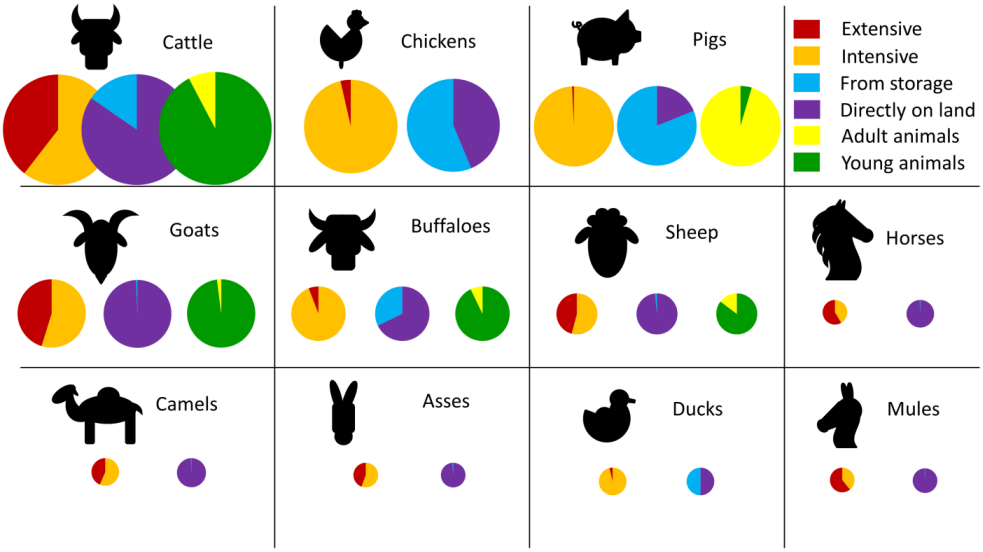


Figure 6.4: Oocyst load (E) per animal species attributed to intensive (Fint) or extensive (Fext) systems, coming from storage (S) or excreted directly on land (L), and coming from adult or young animals. Pie chart sizes are an indication of the total annual oocyst load per animal species.

6.3.3 Sensitivity analysis

Due to the lack of observational data of *Cryptosporidium* in the environment, a full model validation (in the meaning of comparing model outcomes to an independent set of observational data) is not possible for GloWPa-Crypto L1, as is the case for many large scale (ecological) models (Augusiak et al., 2014). Yet there are other ways to build trust in a model, that can be summarized in the process 'evaludation'(Augusiak et al., 2014), several of which we have incorporated in this study. These include transparency about model input data and assumptions (Section 6.2), comparing model outcomes to other studies calculating loads to land (Section 6.3.1), and doing a sensitivity analysis to study model performance.

The sensitivity analysis (Tables S6.8-6.10 in the SI) shows that the model is most sensitive to changes in the excretion rates, especially the excretion rates of young cattle (factor 27.9 or ^{10}log 1.4), young goats (factor 4.9), and young buffaloes (factor 3.8). This means that the absolute size of oocyst loads to land, the relative importance of the different animal species and the patterns on the maps should be interpreted with this in mind. Regarding prevalence, the model is most sensitive to changes in the prevalence among young cattle, adult pigs and chickens (factor 2.17, 1.45 and 1.41 respectively). For variables other than excretion rates and prevalence, the model is most sensitive to changes in the fraction of manure going to storage and to land (factor 1.96). The model is not very sensitive to changes in storage time and temperatures in different seasons, and to the excretion rates and prevalence among animal species that do not contribute much to the global total oocyst loads (e.g. mules, asses and ducks).

It is not surprising that the sensitivity analysis shows that the model is most sensitive to changes in the oocyst excretion rates, because the range over which they were varied in the sensitivity analysis is large, as excretion rates exhibit strong variation (over several orders of magnitude, see Chapter 5). Starkey et al. (2007) also report that their model to calculate oocyst loads is most sensitive to oocyst excretion. A source of uncertainty for excretion rates for all animal species is that recovery efficiencies of oocysts are often not determined, and this can affect faecal concentration estimates strongly (Kuczynska and Shelton, 1999). The literature review in Chapter 5 did not identify any studies for the excretion rates for mules, asses, camels, buffaloes, chickens and ducks. For mules, asses and camels the value for horses was taken as model input, and for buffaloes the value for cattle. For chickens and ducks, we determined model input values based on additional literature, all inoculation studies (i.e. not natural infection) with very young animals (see Table S6.5 in the Sm). This may lead to an overestimation of excretion rates, as inoculation with higher oocyst numbers can lead to higher shedding (Rhee et al., 1995; Zambriski et al., 2013), although not all studies observe this (Zambriski et al., 2013), and one study indicates that younger chickens shed more oocysts and for a longer period than older chickens (Sréter et al., 1995). Furthermore, nearly all studies in chickens were done with *C. baileyi*, and excretion rates can also differ for different *Cryptosporidium* species; one study found that *C. meleagridis* was shed by chickens in 2-3 times lower numbers than *C. baileyi* (Tůmová et al., 2002).

Observed prevalence is also uncertain, as different measurement methods can give different outcomes, in part due to the detection limit of the methods used that may cause 'low level shedders' to be missed (Ryan et al., 2005; Venu et al., 2013). Detection limits are often not discussed in studies measuring *Cryptosporidium* in animal faeces. Furthermore, we assumed that observed prevalence, usually measured at a point in time or over a short period, can be generalized to reflect the average prevalence throughout the year. If

observed prevalences are biased towards seasons, regions, or herds that experience higher than average levels of cryptosporidiosis infection, we may overestimate prevalence in our model.

We want to stress the need for more observational data of *Cryptosporidium*, in fresh faecal material but also in manure in different types of storage facilities and on the field, from animals of different age groups, and from different countries. Recovery rates and detection limits should be assessed and published with these data.

6.3.4 Scenario analysis: effect of manure storage and treatment

We compare the results for the four scenarios in Figures 6.5 and 6.6. The difference between the total oocyst load calculated in Scenario 1 and the baseline model was found to be a factor of 2.6. This represents the current reduction in oocyst loads due to manure storage. Around half of this current reduction is attributable to manure storage for chickens, around one third to pigs and approximately one tenth to cattle (Figure 6.5). Figure 6.6a is a map of the current reduction in oocyst load due to manure storage per country. High reduction is currently achieved in Europe, Bangladesh, China, countries in south-east Asia and the US. Low reduction takes place in Mongolia, Russia, parts of Africa and Australia.

The differences between Scenario 2-4 and the baseline model represent the oocyst load reduction potential of manure storage (2) and of manure treatment with mesophilic (3) and thermophilic (4) anaerobic digestion. If all manure would be stored under the conditions assumed in our model, the environmental oocyst load would be a factor 2 lower than in our baseline run. If all manure would be treated with mesophilic anaerobic digestion the load would be reduced by a factor 37, and for thermophilic anaerobic digestion by a factor of nearly 37000 (4.6 log units). Around half of the reduction potential in all three scenarios comes from cattle manure, after that chickens, pigs, goats and buffaloes are most important. Figure 6.6b-d are maps of the oocyst load reduction potential per country, resulting from manure going to storage, treated with mesophilic anaerobic digestion and thermophilic anaerobic digestion, respectively. The highest reduction potential can be found in India, Bangladesh, Western Europe, China, several countries in Africa and New Zealand. Low reduction potential is in Russia, Canada, and several countries in Africa. The spatial patterns in Figures 6.6b-d are similar, giving high values for countries with high livestock density and low current manure storage. The observation that especially western-Europe, Bangladesh and China have both a high current reduction (Figure 6.6a) and a high reduction potential (Figures 6.6b-d) follows from the high livestock density in these regions.

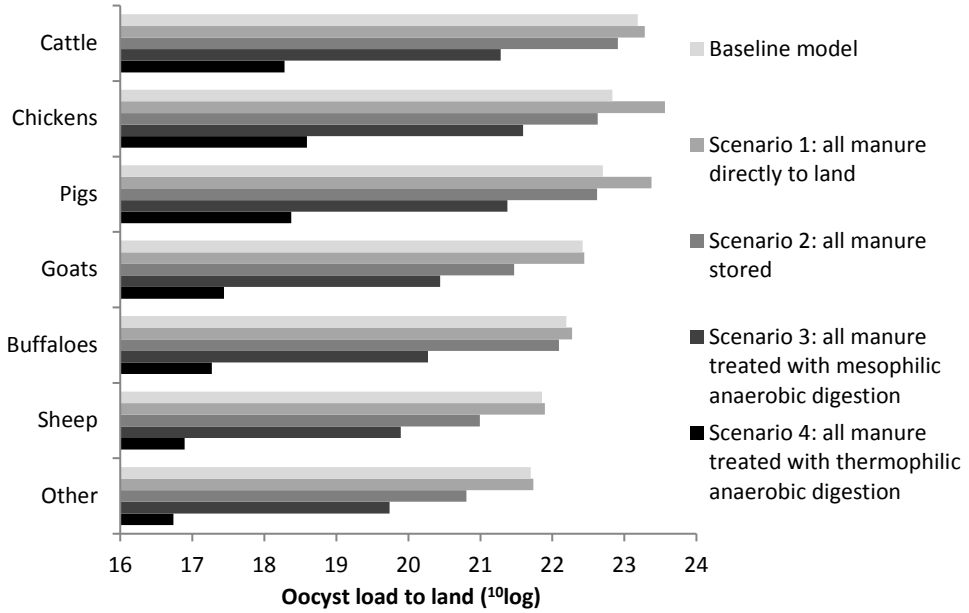


Figure 6.5: Oocyst load per animal species for the baseline model and four alternative management scenarios. The scenarios are assuming that all manure goes directly to land (Scenario 1), all manure goes to storage (Scenario 2), all manure is treated by mesophilic anaerobic digestion (Scenario 3) and all manure is treated by thermophilic anaerobic digestion (Scenario 4). The category 'Other' combines the results from horses, camels, asses, mules and ducks.

Manure storage alone is not a strategy by which oocyst loads to land are greatly reduced. In our model, we assumed continuous manure addition during the storage period, meaning that the manure going to land is ranging in age from fresh to old. Storing manure in batches, instead of with continuous addition of fresh manure, could improve oocyst load reduction (Hutchison et al., 2005). Manure treatment with mesophilic or thermophilic anaerobic digestion can have a much larger impact as oocyst loads could be reduced by several log units. Our estimate of oocyst survival during anaerobic digestion is based solely on the findings of Garcés et al. (2006), but Kinyua et al. (2016) confirm that treating manure by anaerobic digestion before it is applied to land can lower risk of cryptosporidiosis from contaminated crops and soils significantly. It would be worthwhile to further investigate the effects of different manure treatments (such as anaerobic digestion, but also other possible treatments) on oocyst survival.

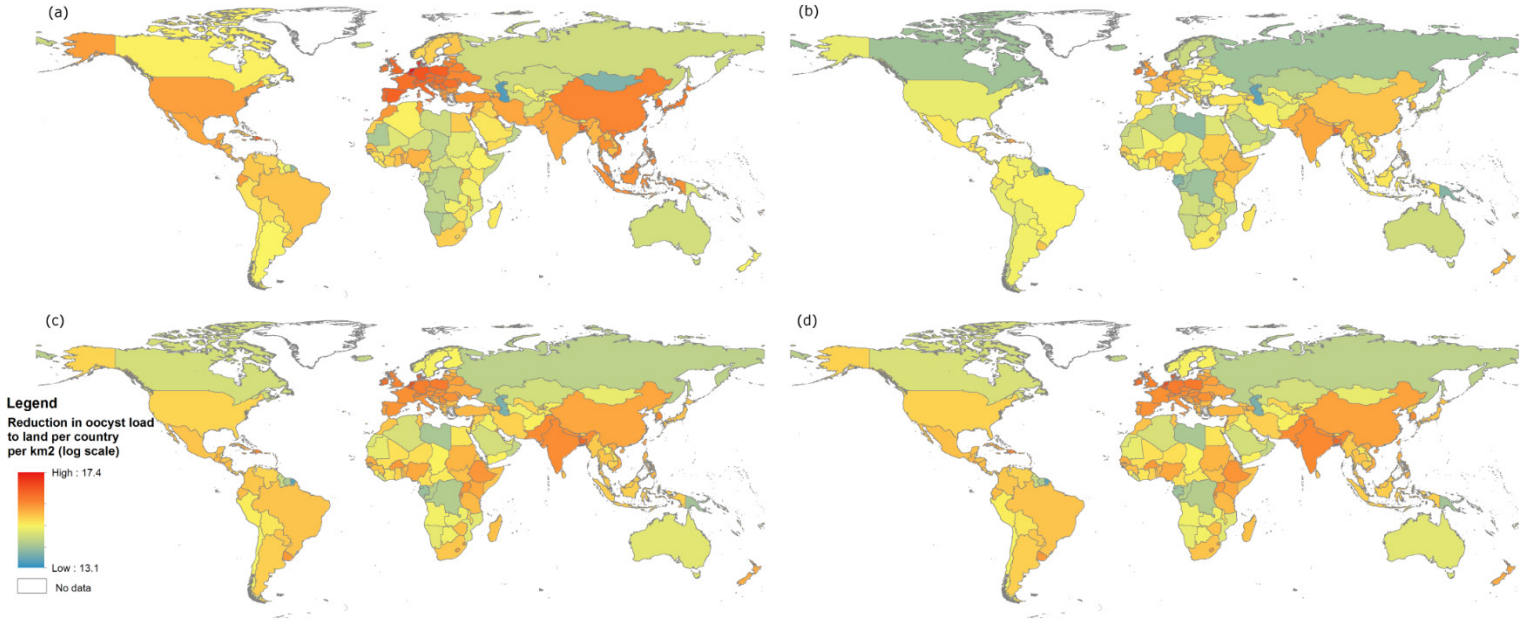


Figure 6.6: (a) Current reduction of oocyst load to land by storage (Scenario 1 minus baseline model run divided over country surface area in km^2). This map shows how much fewer oocysts end up on land because of current manure storage practices. High values indicate large reductions. (b-d) Reduction potential of oocyst load to land, showing respectively the baseline model run minus Scenario 2-4 divided over country surface area in km^2. These maps show how much fewer oocysts end up on land if all manure would go into storage (b), if all manure would be treated by mesophilic anaerobic digestion (c) and if all manure would be treated by thermophilic anaerobic digestion (d). High values indicate large reduction potential.

6.3.5 The importance of livestock *Cryptosporidium* for human disease

Not all *Cryptosporidium* species are infectious for humans. Livestock harbour many *Cryptosporidium* spp. that have not been implicated in human infection (Robertson et al., 2014; Ryan et al., 2014). The majority of human infections are caused by the species *C. hominis*, which predominantly infects humans, and *C. parvum*, which infects a variety of mammals. Ruminants are important reservoirs of *C. parvum* (Robertson et al., 2014; Ryan et al., 2014), especially (pre-weaned) calves (Fayer et al., 2007; Follet et al., 2011; Helmy et al., 2013; Khan et al., 2010; Kváč et al., 2006) and to a lesser extent adult cattle, lambs and goat kids (Castro-Hermida et al., 2011; Geurden et al., 2008; Koinari et al., 2014; Mi et al., 2014; Wang et al., 2014; Yang et al., 2009). Livestock can also harbour other *Cryptosporidium* spp. that have occasionally been reported in humans, examples are *C. meleagridis* from chickens, *C. andersoni* and *C. bovis* from cattle, *C. suis* and *C. scrofarum* from pigs, and *C. xiaoi* from sheep and goats (Robertson et al., 2014; Ryan et al., 2014).

GloWPa-Crypto L1 does not distinguish between *Cryptosporidium* species, for three main reasons: 1) comprehensive quantitative data on the relative occurrence of the various species in different livestock worldwide are not available, 2) *Cryptosporidium* species denomination has changed over the years and is subject to disagreement (Šlapeta, 2013), and 3) observational data on *Cryptosporidium* oocysts in the environment usually do not differentiate between species either, meaning that it would be near impossible to validate a species-specific model. However, if the outcome of GloWPa-Crypto L1 were to be used as input for risk assessment for human disease, data or assumptions are needed on the prevalence of *Cryptosporidium* spp. in livestock that are infectious for humans (mainly *C. parvum* and *C. hominis*) or only the input from the most relevant livestock species (cattle) should be incorporated.

GloWPa-Crypto H1 (Chapter 4) is the human counterpart of GloWPa-Crypto L1. It calculates global human *Cryptosporidium* emissions to surface water to be 1.6×10^{17} oocysts/year (Chapter 4). GloWPa-Crypto L1 calculates a much higher total global oocyst load from animal manure of 3.2×10^{23} oocysts/year. However, it should be noted that this is the load to land, not to surface water. Rainfall and subsequent runoff will transport only a small part of manure to surface waters, and in the meantime oocysts will also decay. Besides, as mentioned before, not all livestock *Cryptosporidium* is infectious for humans. A comparison of the relative importance of human and animal *Cryptosporidium* for waterborne disease is therefore, at this point, only speculative. However, our model suggests that the contribution from livestock should definitely not be ignored.

6.3.6 Outlook for *Cryptosporidium* modelling

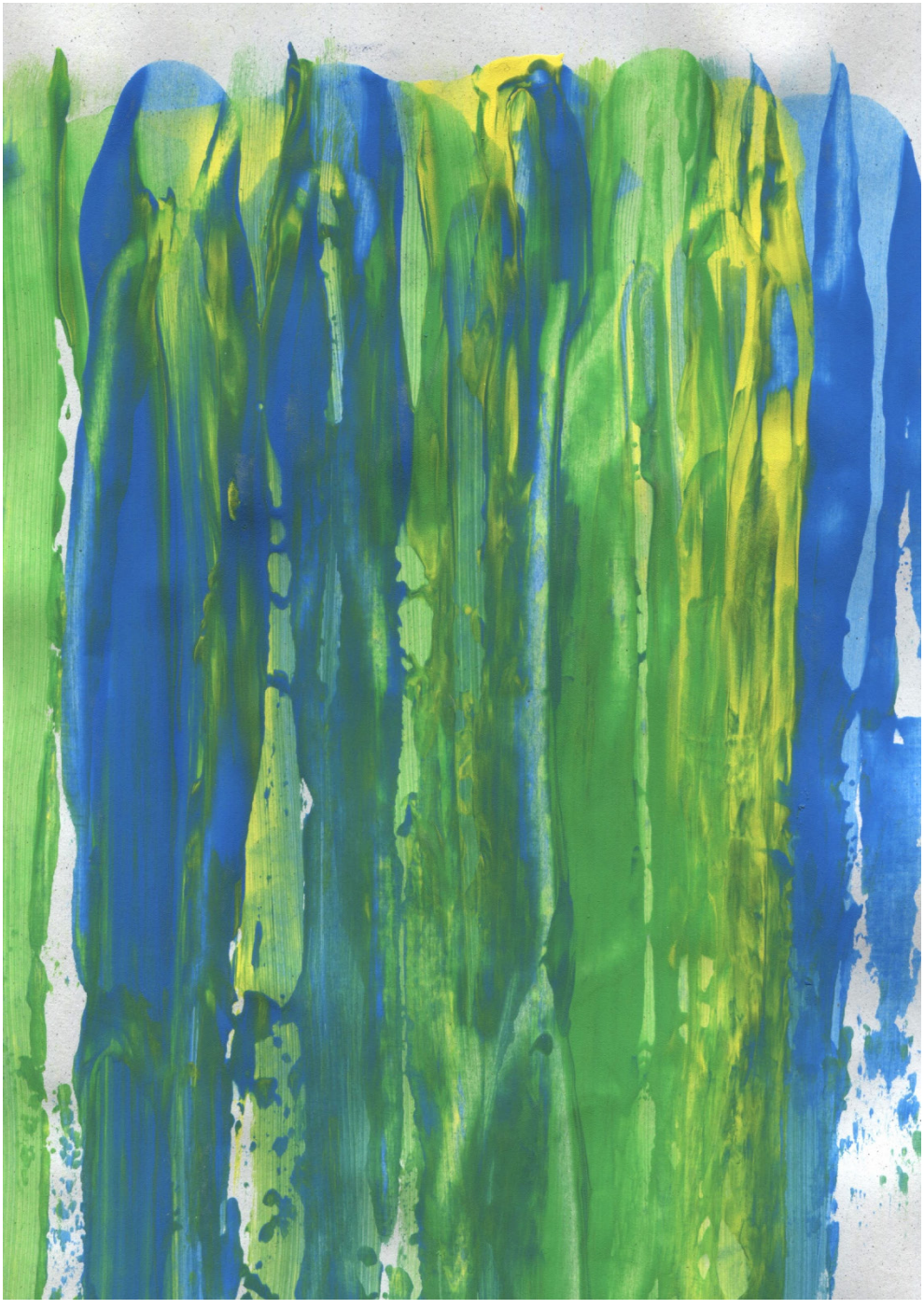
Gaining insight into the environmental pathways of *Cryptosporidium* is important in the context of managing human and animal cryptosporidiosis. Facing scarcity of observational data of *Cryptosporidium* in the environment, process-based modelling and scenario analysis can help to provide insight in handling options, such as the reduction potential from manure storage and treatment. More detailed scenario analyses could investigate the effects of different types of manure treatments, to answer specific management questions.

A next step is to go towards the exposure pathways that determine risk of contracting cryptosporidiosis, such as water- and foodborne pathways. A model of surface water oocyst concentrations can be constructed when the outcomes of GloWPa-Crypto L1 are combined with estimates on the survival of oocysts in manure on fields, transport with runoff to surface waters, hydrological information, and the outcomes of GloWPa-Crypto H1. Together with information on the share of *Cryptosporidium* spp. that are pathogenic for humans, such a model can provide a basis for risk assessments. In addition, GloWPa-Crypto L1 could be further refined to operate on a smaller time step or for specific regions. A model at a smaller time step could look into birthing seasons and herd structure development. This would require more detailed input datasets.

This paper provides a first spatially explicit assessment of *Cryptosporidium* spp. oocysts from livestock manure to land. The total global load is large (3.2×10^{23} oocysts per year) and should not be ignored in risk studies. Spatial differences are linked to animal spatial distributions. The GloWPa-Crypto L1 model is most sensitive to oocyst excretion rates, due to large variation reported in literature. Scenarios that include manure treatment (especially thermophilic anaerobic digestion) strongly reduce the loads to land (up to 4.6 log units). Manure treatment could be important to improve microbial environmental quality.

Acknowledgements

Part of this work has been financially supported by the European Climate-KIC PhD programme. We thank anonymous reviewers for their comments.



Chapter 7

Cryptosporidium concentrations in rivers worldwide

Lucie C. Vermeulen, Marijke van Hengel, Carolien Kroeze, Gertjan Medema, J. Emiel Spanier, Michelle T.H. van Vliet and Nynke Hofstra

Abstract

Cryptosporidium is a leading cause of diarrhoea and infant mortality worldwide. A better understanding of the sources, fate and transport of *Cryptosporidium* via rivers is important for effective management of waterborne transmission, especially in the developing world. We present GloWPa-Crypto C1, the first global, spatially explicit model that computes *Cryptosporidium* concentrations in rivers, implemented on a 0.5 x 0.5 degree grid and monthly time step. To this end, we first modelled *Cryptosporidium* inputs to rivers from human faeces and animal manure. Next, we use modelled hydrology from a grid-based macroscale hydrological model (the Variable Infiltration Capacity model). Oocysts transport through the river network is modelled using a routing model, accounting for temperature- and solar radiation-dependent decay and sedimentation along the way.

Monthly average oocyst concentrations are predicted to range from 10^{-6} to 10^{2} oocysts L^{-1} in most places. Critical regions ('hotspots') with high concentrations include densely populated areas in India, China, Pakistan and Bangladesh, Nigeria, Algeria and South Africa, Mexico, Venezuela and some coastal areas of Brazil, several countries in Western and Eastern Europe, and the Middle East. Point sources (human faeces) appears to be a more dominant source of pollution than diffuse sources (mainly animal manure) in most world regions.

Validation shows that GloWPa-Crypto medians are mostly within the range of observed concentrations. The model generally produces concentrations that are 1.5 – 2 log higher than the observations. This is likely predominantly due to the absence of recovery efficiency of the observations, which are therefore likely too low. Goodness of fit statistics are reasonable. Sensitivity analysis showed that the model is most sensitive to changes in input oocyst loads.

GloWPa-Crypto C1 paves the way for many new opportunities at the global scale, including scenario analysis to investigate the impact of global change and management

options on oocysts concentrations in rivers, and risk analysis to investigate human health risk.

This chapter is based on: Vermeulen, L.C., van Hengel, M., Kroeze, C., Medema, G.J., Spanier, E., van Vliet, M., Hofstra, N. 2017. *Cryptosporidium* concentrations in rivers worldwide. Submitted.

7.1 Introduction

Diarrhoea is still a major cause of death worldwide, especially in children younger than 5 years in developing countries (GBD 2013 Mortality and Causes of Death Collaborators, 2015). The zoonotic protozoan parasite *Cryptosporidium*, that is transmitted via the faecal-oral route, is an important cause of childhood diarrhoea and mortality (Liu et al., 2016). Infection can occur via direct contact with faeces of infected humans or animals (Zambrano et al., 2014), but also often occurs via a waterborne route through the environment. Examples are drinking of or recreation in rivers contaminated with faeces (Shirley et al., 2012), and consumption of fresh produce irrigated with contaminated water (Dixon, 2016). Oocysts, the robust survival stage of the pathogen, are excreted in faeces of infected humans and animals and can reach rivers either directly (point sources, such as sewer pipes) or indirectly (diffuse sources, such as manure transported with surface runoff). Oocysts are transported with rivers, and meanwhile decay and sedimentation decrease their viability and concentration.

Gaining insight in the transmission of *Cryptosporidium* via rivers is important for evaluating disease risk. Insight in the relative contribution of human versus animal sources, point versus diffuse sources and pathways and effect of control measures are important to design effective management strategies. But as sampling of *Cryptosporidium* (and or pathogen) concentrations is expensive and laborious, observational data are scarce, especially for the developing world. Modelling is a common approach to increase insight, for example by pinpointing hotspots of high concentrations or identifying the relative importance of pollutant sources.

Models of faecal microorganisms in rivers exist mostly at the catchment scale (Chapter 2). Most commonly, catchment scale waterborne pathogen models couple a pathogen loading estimate to an existing hydrological model. We are only aware of one other large scale concentrations model, WorldQual, which has been applied for faecal coliform bacteria for several continents (Reder et al., 2015; UNEP, 2016). Faecal coliforms are indicators of faecal pollution, but not pathogenic themselves. For other water quality variables, such as nutrients, large scale modelling is much more advanced (Chapter 2).

In this study we aim to estimate *Cryptosporidium* oocyst concentrations in rivers worldwide. We present the model GloWPa-Crypto C1 (the Global Waterborne Pathogen model for *Cryptosporidium* concentrations version 1), a global spatially explicit model at a 0.5°x0.5° longitude x latitude grid that calculates oocysts concentrations in rivers.

7.2 Methods

We present GloWPa-Crypto C1, a model to calculate mean monthly riverine oocyst concentrations. We use the output from the models GloWPa-Crypto H1 and L1 that calculate human and animal *Cryptosporidium* loads (Chapter 4 and 6) and couple this to output from the Variable Infiltration Capacity (VIC) model, a grid-based macroscale hydrological model (Liang et al., 1994), version 4.1.2. Oocysts are transported through the river network using a routing model, accounting for decay and sedimentation along the way. Figure 7.1 gives a schematic model representation. All calculations are performed on a 0.5x0.5° grid and monthly time step using the program R (R Development Core Team, 2016). We use the most recent available estimates for all input variables, and we consider the model representative for approximately the conditions around the years 2005-2010. The first two columns of Table S7.3 in the Supplementary materials give an overview of the variables that are included in the model.

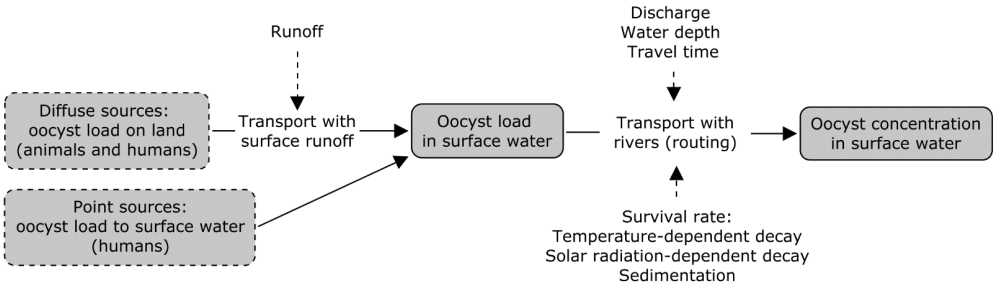


Figure 7.1: Flow chart overview of GloWPa-Crypto C1 model approach. This figure shows the 'flow' of oocysts from diffuse and point sources on the left to the concentrations in rivers on the right. In between, processes that are calculated are the transport with surface runoff, the transport with rivers (routing) and oocyst decay. Dotted arrows show the major influencing variables (runoff, discharge, water depth and travel time, and the survival rate as affected by temperature, solar radiation and sedimentation). The diffuse and point sources are presented in boxes with dotted lines, as they are not calculated in this current study, but in the GloWPa-Crypto H1 and L1 models.

7.2.1 Sources of *Cryptosporidium*

7.2.1.1 Human sources

Human *Cryptosporidium* loads to rivers are calculated in the GloWPa-Crypto H1 model (Chapter 4). In brief, prevalence of *Cryptosporidium* infection and oocyst excretion in

humans is estimated from literature, and gridded data of human population from LandScan (Bright et al., 2011) are divided in developed/ developing and urban/rural. Depending on the sanitation systems used by the different populations (from the WHO/JMP Joint Monitoring Program) and assumptions on the removal of oocysts by sewage treatment the final load to rivers is calculated. The model is programmed in R and operates on a 0.5x0.5° grid for the world at an annual time step, and is representative for conditions in approximately the year 2010. The annual oocyst loads are divided by 12 as an estimate of monthly oocyst loads. In Chapter 4 the load from the population that practices open defecation in rural areas (a diffuse source) was multiplied with a constant runoff fraction (0.025) to estimate the amount ending up in rivers. In the current study this was changed; the human diffuse sources were added to the animal diffuse sources and treated as described in Section 7.2.2.

7.2.1.2 Animal sources

Animal *Cryptosporidium* loads to land are calculated in the GloWPa-Crypto L1 model (Chapter 6). In brief, livestock population data are from the Gridded Livestock of the World v2.0 (Robinson et al., 2014), and prevalence and oocyst excretion by 11 livestock species (cattle, buffaloes, pigs, sheep, goats, horses, camels, donkeys, mules, chickens and ducks) are based on an extensive literature review (Chapter 5). The model differentiates between manure excreted directly on land during grazing and manure spread on land after storage, and temperature-dependent oocyst decay during storage is accounted for. The model is programmed in R and operates on a 0.5°x0.5° grid for the world at an annual time step, and is representative for conditions in approximately the year 2005. The annual oocyst loads are divided by 12 as an estimate of monthly oocyst loads. This means that the timing of the birthing season and application of stored manure on fields is not accounted for in the current version of the model, as comprehensive global data on this are not available.

7.2.2 Oocyst transport from land to rivers

Transport of oocysts from the land to rivers largely depends on surface runoff, as subsurface flow will generally transport few oocysts due to the filtering capacity of soils (Mawdsley et al., 1996; McLaughlin et al., 2013). Therefore, in GloWPa-Crypto C1 we only take into account transport with surface runoff. The VIC model (Liang et al., 1994) provides data of average surface runoff (mm day^{-1}) and average discharge (m^3 s^{-1}). These data are monthly averages of a baseline VIC run with WATCH forcing data (Weedon et al., 2011), averaged over the period 1970-2000 (0.5°x0.5° grid).

Given a certain surface runoff, we need to estimate the fraction of oocysts that is transported to rivers. Surface runoff will not transport all oocysts from manure to rivers; a

large part of the oocyst load will be retained within the faecal matrix or on the soil surface (Table 7.1), leading to a reduction of several log units of the oocyst load that will reach the river. We found 9 studies that quantify the amount of oocysts that are transported with runoff from faecal deposits on land, only one of which measured this under field conditions (Table 7.1).

Table 7.1: Overview of experimental studies on oocyst retention in the faecal matrix and on soils under rainfall.

Simulated rainfall or field conditions	Matrix and soil conditions	Observed oocyst retention	Studies
Field conditions	Oocysts in faecal matrix on vegetated soils	3.2 – 8.8 log	(Atwill et al., 2006b)
Field conditions	Oocysts in faecal matrix on bare soils	1.7 – 5 log	(Atwill et al., 2006b)
Simulated rainfall	Oocysts in faecal matrix on vegetated soils	0.7 - 6.8 log	(Davidson et al., 2014; Davies et al., 2004; Tate et al., 2004)
Simulated rainfall	Oocysts in faecal matrix on bare soils	0.4 - 2.5 log	(Davidson et al., 2014; Davies et al., 2004; Tate et al., 2004)
Simulated rainfall	Oocysts in faecal matrix in the lab	0.1 – 3 log	(Boyer et al., 2009; Bradford and Schijven, 2002)
Simulated rainfall	Oocysts in suspension on vegetated soils	0.6 – 3.1 log	(Atwill et al., 2002; Bhattarai et al., 2011; Trask et al., 2004)
Simulated rainfall	Oocysts in suspension on bare soils	0.0 – 1.6 log	(Atwill et al., 2002; Bhattarai et al., 2011; Trask et al., 2004)

From Table 7.1 we can see that both the faecal matrix and soils contribute to reducing the oocyst load in runoff, and that vegetated soils do so more effectively than bare soils. Under field conditions larger retention is observed than under simulated rainfall, which is likely partly due to decay, as Atwill et al. (2006b) measured oocysts in runoff over a period of one year. During this year, surface runoff varied from 0 – 20 mm. We therefore choose to linearly assign the observed log oocyst retention from vegetated plots under field conditions (3.2 – 8.8 log) to grid cells with runoff values ranging between 0 - 20 mm, where highest runoff will get lowest retention and vice versa. This approach is extrapolated and grid cells with higher runoff get assigned a retention between 1 – 3.2 logs. We assume that under field conditions, at least 1 log oocysts from a monthly load is retained, due to decay and landscape characteristics. In months that surface runoff is zero, oocyst retention is set to infinite. Oocyst stream input from diffuse sources (SD) is calculated by decreasing the oocyst loads on land (as calculated with GloWPa-Crypto L1) and the human diffuse oocyst loads (as calculated with GloWPa-Crypto H1) with the

assigned retention. The oocyst stream input from point sources (SP) is the total human load to rivers (as calculated in GloWPa-Crypto H1) minus the human diffuse sources.

Modelling the release of *Cryptosporidium* from manure to rivers is challenging, given the scarcity of observational data. We have also looked into the literature on existing modelling approaches, and conclude there currently is no single 'gold standard' approach to modelling the release of oocysts from manure and transport to rivers. Furthermore, existing modelling approaches require too detailed input data, in space and/or time, to directly apply them at a global scale on a monthly time step (see Supplementary Material S7.1 for a more detailed discussion).

7.2.3 Oocyst survival during transport with rivers

Oocyst survival during transport can be modelled following standard first order decay:

$$C_t = C_0 e^{-K \times t} \tag{7.1}$$

Where C_t is the oocyst concentration after time t (days), C_0 is the initial concentration, and K is the loss rate coefficient (day^{-1}). Writing this in a different way, the natural logarithm of the fraction that survives in month i (Fs_i) equals:

$$\ln(Fs_i) = -K_i \times t_i \tag{7.2}$$

The loss rate coefficient K_i (day^{-1}) can have several components (Thomann and Mueller, 1987):

$$K_i = K_{T,i} + K_{R,i} + K_{S,i} \tag{7.3}$$

$K_{T,i}$ is the temperature-dependent decay rate (day^{-1}), $K_{R,i}$ is the solar radiation-dependent decay rate (day^{-1}), and $K_{S,i}$ is the loss rate due to sedimentation (day^{-1}) in month i (see sections 7.2.3.1-7.2.3.3). Water residence time in a grid cell (t_i) is estimated as described in section 7.2.3.4. The Supplementary Material S7.3 contains some maps of the different loss components. It is important to note that decay and sedimentation have (partially) different effects on the oocysts. Sedimentation has an effect on the concentration (in the water phase) directly, while decay has an effect on the viability of oocysts. Loss of viability may lead to disintegration, thereby reducing the concentration.

7.2.3.1 Temperature-dependent survival (K_T)

To calculate temperature-dependent survival during transport with rivers, we use the approach by Peng et al. (2008), who compiled literature estimates on the survival of *Cryptosporidium* in raw waters. We apply the following relationship between K_T and water temperature (T_w):

$$K_{T,i} = K_4 e^{\lambda(T_{w,i}-4)} \quad (7.4)$$

Where K_4 and $K_{T,i}$ are respectively the decay rate coefficients (day^{-1}) at 4°C and water temperature $T_{w,i}$ in month i, and λ is a dimensionless constant. For T_w we use monthly water temperature as estimated by the VIC-RBM model framework which was applied globally (van Vliet et al., 2012). This is a coupled hydrological-water temperature model framework based on the VIC model (Liang et al., 1994) and RBM stream temperature model (Yearsley, 2009). For K_4 and λ we take the values reported for cell culture studies in raw waters by Peng et al., these are K_4 =0.0051 day^{-1} and λ=0.158 (Peng et al., 2008). In their review, Peng et al. distinguished between studies reporting *Cryptosporidium* survival for DAPI/PI or excystation studies, and studies that used cell culture to measure oocyst infectivity. For water temperatures between 4-20°C, observed K values were similar, but for temperatures over 25°C, K values from cell culture studies were notably higher (Peng et al., 2008), meaning that in the high temperature range DAPI/PI and excystation studies likely overestimate the survival of infectious oocysts. Therefore, we take the values reported for cell culture studies. For waters below 4°C, few studies have been done (Peng et al., 2008), we therefore assume that survival in waters below 4°C is the same as at 4°C. Other models of *Cryptosporidium* in rivers use similar K values for temperature-dependent decay (see Supplementary material S7.2).

7.2.3.2 Solar radiation-dependent survival (K_R)

For K_R we can follow Mancini (1978) and Thomann en Mueller (1987) to calculate water depth-averaged solar radiation-dependent decay:

$$K_{R,i} = \frac{k_I I_{A,i}}{k_e Z_i} \times (1 - e^{-k_e Z_i}) \quad (7.5)$$

Where $I_{A,i}$ is the average surface solar radiation ($kJ\ m^{-2}\ day^{-1}$) in month i, k_I is a proportionality constant ($m^2\ kJ^{-1}$), k_e is the attenuation coefficient (m^{-1}), and Z_i is the water depth (m) in month i. Surface solar radiation data are from the WATCH forcing data (Weedon et al., 2011). The ultraviolet (UV) part of solar radiation is most important in decreasing oocyst infectivity (Connelly et al., 2007). Attenuation of UV radiation in water is influenced by dissolved substances, an important one is dissolved organic carbon (Scully and Lean, 1994). We assume attenuation to be linearly dependent on dissolved organic carbon (DOC), according to Lambert Beer's law on substances in water:

$$k_e = k_d \times C_{DOC} \quad (7.6)$$

Where C_{DOC} is the DOC concentration ($mg\ L^{-1}$) and k_d is a proportionality constant ($L\ mg^{-1}\ m^{-1}$). The Global Nutrient Export from Watersheds (Global NEWS) model provides estimates of river export of DOC for the world (Harrison et al., 2005; Mayorga et al., 2010).

We divide total river DOC export over river discharge to obtain basin-averaged estimates of river DOC concentrations. We estimate the values for k_l and k_d by fitting the experimental data of oocyst survival in different water types under different solar conditions presented by King et al. (2008). We used a plot digitizer (http://arohatgi.info/WebPlotDigitizer/, accessed on 19 April 2017) to obtain the data from the infectivity-insolation plots presented in King et al. We can substitute eq. 7.5 and 7.6 in eq. 7.2 for solar radiation-dependent decay (temperature and sedimentation were controlled for in the experiment) gives:

$$\ln(Fs) = -\frac{k_l I_A}{k_d C_{DOC} Z} \times (1 - e^{-k_d C_{DOC} Z}) \times t \quad (7.7)$$

King et al. controlled for the effect of temperature-dependent decay and sedimentation in their experiments. We performed an optimization routine on eq. 7.7 minimizing the root mean squared error to find values for k_l and k_d, using the optim function of the stats package in R (R Development Core Team, 2016). We obtained the following values for the constants: k_l=0.0004798 and k_d=9.831. The model performs quite well in reproducing the observations by King et al. (Figure 7.2), and the outcome does not seem biased for different water types. King et al. looked at a combination of raw waters with different DOC content (pink dots) and tap water (blue dots), we took all of these together to have most data points. We performed a linear regression that yielded an adjusted R^2 of 0.626, with intercept of -0.66, slope of 0.84, and root mean squared deviation (RMSD) of 1.58.

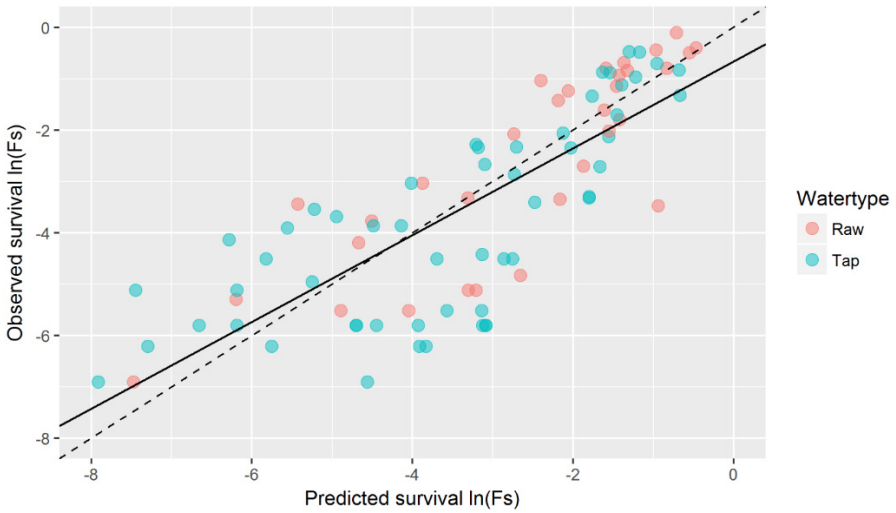


Figure 7.2: Observed versus predicted survival (given as ln(Fs)) for the data from King et al. (2008) for raw surface waters (pink dots) and tap water (blue dots). The dashed line indicates the 1-1 line, in case the model perfectly reproduces the observations all dots should be on this line. The solid line is a linear regression line of observed versus predicted values.

7.2.3.3 Sedimentation (K_S)

$K_{S,i}$, the loss due to sedimentation in month i, can be modelled following Thomann and Mueller (1987):

$$K_{S,i} = \frac{v}{Z_i} \tag{7.8}$$

Where v is the settling velocity (m day^{-1}) and Z_i is the river depth (m) in month i.

River depth Z_i is estimated according to section 7.2.3.4. We take v to be 0.1 m day^{-1}, based on the observations of oocyst settling velocity by Brookes et al. (2006) and Medema et al. (1998). The value found by Brookes et al. (0.1 m day^{-1}) is in between the values by Medema et al. for settling velocities of free oocysts (0.035 m day^{-1}) and oocysts attached to suspended particles (3.5 m day^{-1}). Resuspension of oocysts from sediments is ignored. Reder et al. (2015) apply a similar approach in their continental-scale model of faecal bacteria in rivers.

7.2.3.4 River geometry and water residence time

We use the river geometry equations by Leopold and Maddock (1953) to calculate river width, depth, and mean flow velocity from river discharge. The coefficients for these equations were empirically estimated by Allen et al. (1994), who used data from 674 stations across the USA. It was assumed that these coefficients can be applied globally, as these stations cover a wide range of hydro-climatic zones. This approach is taken from Van Vliet et al. (2012), who also applied it for the VIC-RBM model framework:

$$Z_i = 0.34\, Q_i^{0.341} \tag{7.9}$$

$$W_i = 1.22 Q_i^{0.557} \tag{7.10}$$

$$U_i = \frac{Q_i}{W_i \times Z_i} \tag{7.11}$$

where Z_i is river depth (m), Q_i is river discharge ($m^3 s^{-1}$), W_i is river width (m) and U_i is river flow velocity (m s^{-1}) in month i. The river discharge is naturalized discharge, meaning that the existence of dams and reservoirs are not taken into account.

The length of the river stretch in a grid cell is estimated by taking the distance between its midpoint and the midpoint of the cell it flows towards (based on the drainage direction map DDM30 (Döll and Lehner, 2002)). This was done using the function pointDistance from the raster package in R (Hijmans, 2016).The length of the river stretch divided by the flow velocity gives the residence time of water in a grid cell (t_i).

7.2.4 Oocyst transport with rivers – the routing model

Oocysts are routed through the river channel network based on the global flow direction map DDM30 (Döll and Lehner, 2002), which was also used for streamflow routing in VIC. The routing starts at the grid cells with the lowest flow accumulation (i.e. number of grid cells draining to that grid cell) and ends at the grid cells with highest flow accumulation. For each grid cell, monthly stream inputs from point (SP) and diffuse (SD) sources are added to the oocyst load that was already in the stream from the previous grid cell. The resulting oocyst load is decreased according to the calculated survival in the grid cell (Section 7.2.3).

$$L_{i,n} = \left(SD_i + SP_i + L_{i,n-1}\right) \times e^{-K_i \times t_i} \quad (7.12)$$

Where $L_{i.n}$ is the oocyst load (oocysts month^{-1}) in month i in a grid cell with flow accumulation number n, SD_i is the oocyst stream input (oocysts month^{-1}) from diffuse sources in month i, SP_i is the stream input (oocysts month^{-1}) from point sources in month i, $L_{i,n-1}$ is the oocyst load (oocysts month^{-1}) in month i from grid cells that drain into the current grid cell (n-1), K_i is the loss rate coefficient (day^{-1}) in month i and t_i is the water residence time (days) in the grid cell in month i.

The oocyst load is divided by the river discharge (converted to m^3 month^{-1}) to estimate average monthly oocyst concentration for the grid cell.

$$C_i = \frac{L_i}{Q_i} \quad (7.13)$$

Where C_i is the average oocyst concentration (oocysts m^{-3}), L_i is the total oocyst load (oocysts month^{-1}) and Q_i is the average river discharge (m^3 month^{-1}) in month i.

7.2.5 Model performance

7.2.5.1 Validation

We validate the model by comparing model outcomes with a set of observational data of *Cryptosporidium* in rivers around the world. We have gathered >4000 observations from 384 locations in 11 countries on 5 continents (Burnet et al., 2014, 2015; Chuah et al., 2016; Claßen et al., 2004; Ehsan et al., 2015; Kistemann et al., 2012; Lalancette et al., 2014; Rechenburg et al., 2009, 2006; Till et al., 2008). The Supplementary Material S7.5 provides more information on the sources of the data. For each observation, we selected the corresponding modelled value from the month and grid cell in which the observation was taken. These pairs of observed and predicted values were used for all further analysis. It should be noted that the observations are points in time and space, while GloWPa-Crypto uses climate and hydrological data averaged by month over a 30 year period on a

0.5x0.5 degree grid. This means that variation in predicted concentrations between months is captured to some degree (hydrologically, but not the calving or manure spreading season), but not between months in different years, while of course weather conditions might differ between years. Similarly, within-grid spatial variability is not captured by the model.

Oocyst concentrations in surface water are often around or below the detection limit of the most commonly employed methods (Efstratiou et al., 2017a). This means that zero values are very common and do not mean absence of the pathogen, but a concentration below the detection limit of the method used (Ongerth, 2016). In our data set, 68% of the observations is below the detection limit, ranging from 20-95% for the different countries. The reported detection limits range from 0.0028 - 1 oocysts L^{-1}. Simply ignoring observations below the detection limit in the estimation of, for instance, descriptive statistics is incorrect, as this will lead to overestimation of actual concentrations. Inserting a value for them, such as (half of) the detection limit, as is sometimes done is similarly incorrect (Haas and Scheff, 1990; Helsel, 2010a, 2010b). Therefore, we use statistical procedures that can work with these so-called 'censored data' by estimating a distribution for values below the detection limit, assuming lognormally distributed values (Helsel, 2010b; Lee and Helsel, 2005). These are available in the R package 'NADA' (Lee, 2017).

Where possible, we correct the observed concentrations for the recovery efficiency. A recovery efficiency can be determined by seeding a sample with a known amount of labelled oocysts and determining what percentage of these are recovered by the method (Efstratiou et al., 2017a). The actual observed oocyst concentration can then be adjusted accordingly. In our data set, 26% of the observations reported a recovery efficiency and are adjusted, the remainder are used 'as is'. Three countries reported a recovery efficiency for all observations (Belgium, Luxembourg and Thailand), three for a part of the observations, and four did not report a recovery efficiency. The recovery efficiency ranged between 0.026-0.89.

7.2.5.2 Sensitivity analysis

We perform a nominal range sensitivity analysis, changing one variable at a time based on a reasonable range the variable can take, usually both a decrease and an increase are tested. A description of the changes to the variables is provided in the Supplementary material S7.6.

7.3 Results and Discussion

7.3.1 Oocyst concentrations worldwide

Average oocyst concentrations are simulated to fall mostly in the range of 10^{-6} to 10^{2} oocysts L^{-1} worldwide (Table 7.2). Critical regions with high concentrations include densely populated areas in India, China, Pakistan and Bangladesh, Nigeria, Algeria and South Africa, Mexico, Venezuela and some coastal areas of Brazil, several countries in Western and Eastern Europe, and the Middle East. These hotspot regions align with the hotspot regions observed in the underlying human oocyst loads model GloWPa-Crypto H1 (Chapter 4), more than with the hotspot regions observed in the animal oocyst loads model GloWPa-Crypto L1 (Chapter 6). High values over 10^{3} oocysts L^{-1} (values typically found in untreated sewage (Nasser, 2016)) are found in various large cities in developing countries. Values over 10^{4} oocysts are only found in two grid cells in the month July in Bangladesh. Very low concentrations (10^{-13} – 10^{-6}) are found in very sparsely populated areas, such as the Arctic regions. The high concentrations in urban areas in developing countries are a concern, and the situation is expected to get worse in the future (Chapter 4).

Table 7.2: Descriptive statistics of oocyst concentrations grid cell values around the world. Q5 means that 5% of grid cells has a value below this, etcetera.

Stat	Concentration (oocysts/L)
Q5	3.0E-06
Q25	2.7E-03
Q50 (median)	4.9E-02
Q75	8.5E-01
Q95	1.5E+01
Mean	4.9E+00

Figures 7.3 and 7.4 show oocyst concentrations for the months January and July, both without and with a discharge mask. The masked plots show only large rivers (i.e. grid cells where average annual discharge is higher than 200 m^3/s). The figures show that for example India, parts of China, Mexico and Nigeria are expected to experience higher oocyst concentrations in January than in July. Grid cells with a monthly average discharge <1 m^3/s are excluded from the analysis altogether, as the model was found not to perform well for locations with very low discharge. The white areas in the plots are thus a result of either discharge below the threshold or zero oocyst loads.

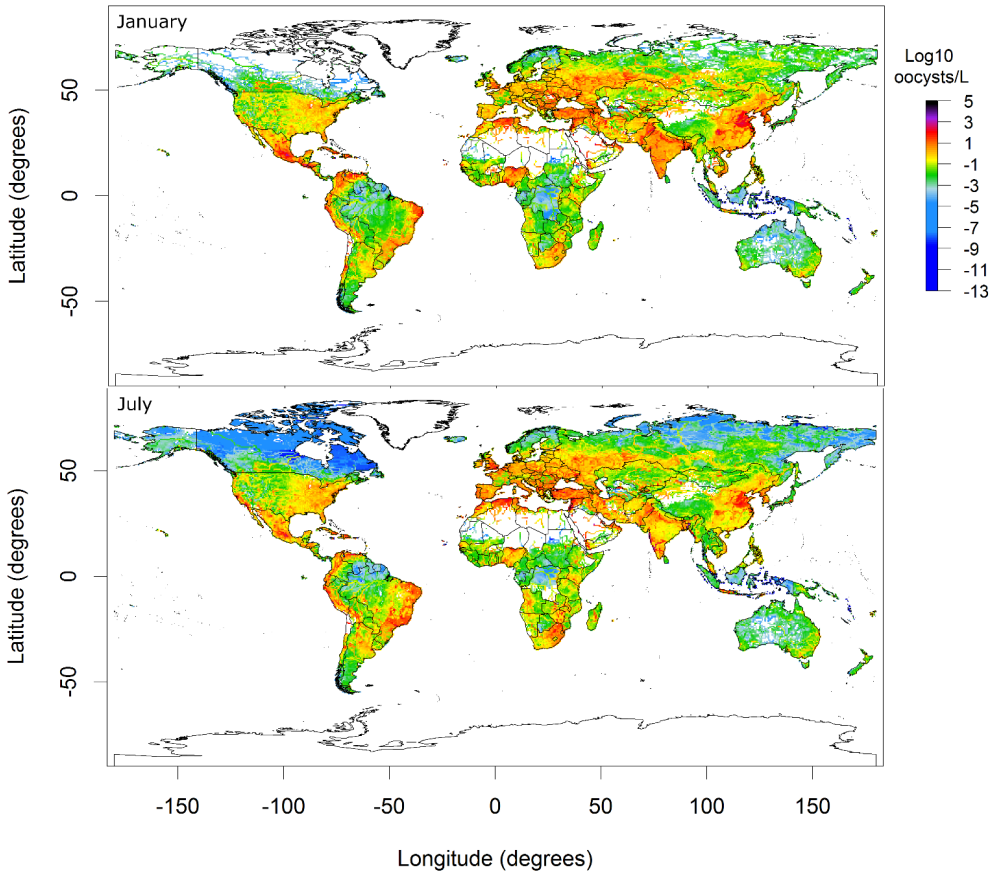


Figure 7.3: Oocyst concentrations in rivers worldwide for the months January and July (log10 oocysts/L)

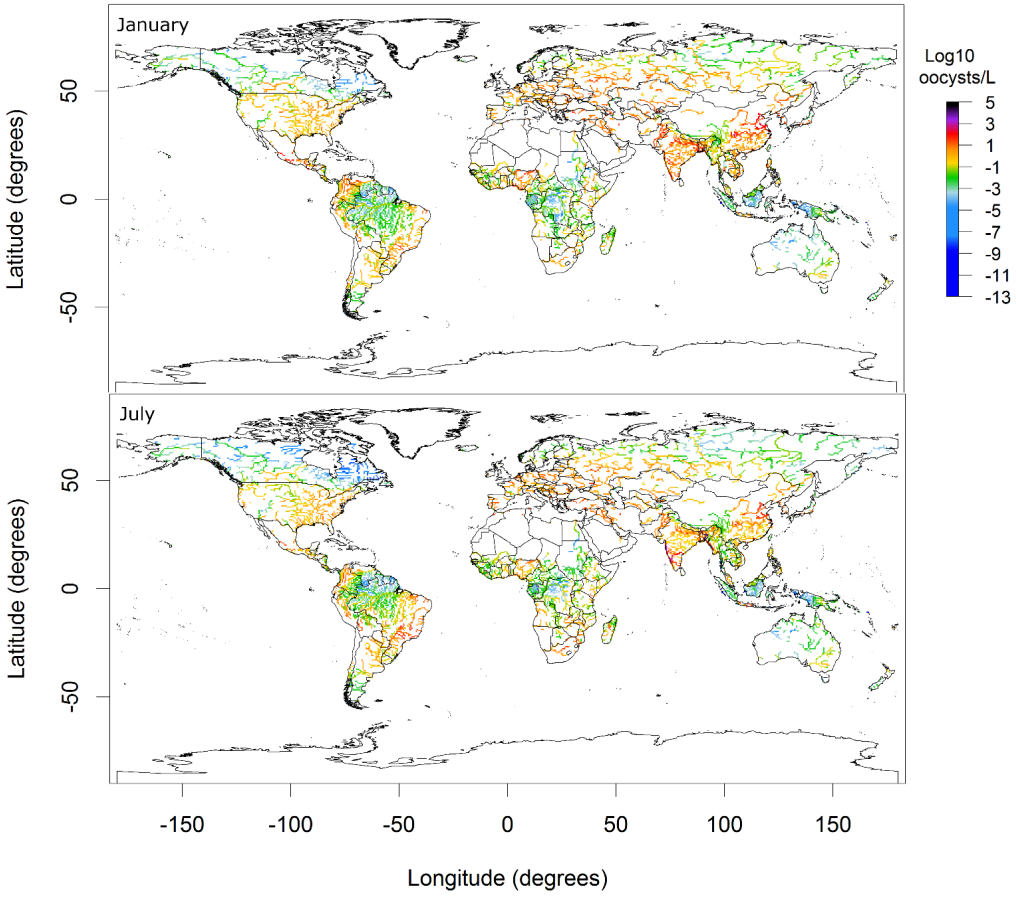


Figure 7.4: Oocyst concentrations in rivers worldwide for the months January and July (log10 oocysts/L), with a discharge mask. These plots show only large rivers (i.e. grid cells where average annual discharge is higher than 200 m^3/s).

A WHO report on risk assessment of *Cryptosporidium* (Medema et al., 2009) suggests a categorization of source waters for drinking water production in 6 categories (1= very pristine (~0.001 oocyst/L) to 6 = grossly polluted (~100 oocyst/L). Figure 7.5 shows what our results look like when these categories are assigned to the output concentrations. It can be seen that all 6 categories are covered by GloWPa-Crypto C1, meaning that model outputs are in a similar range with what is found in drinking water sources.

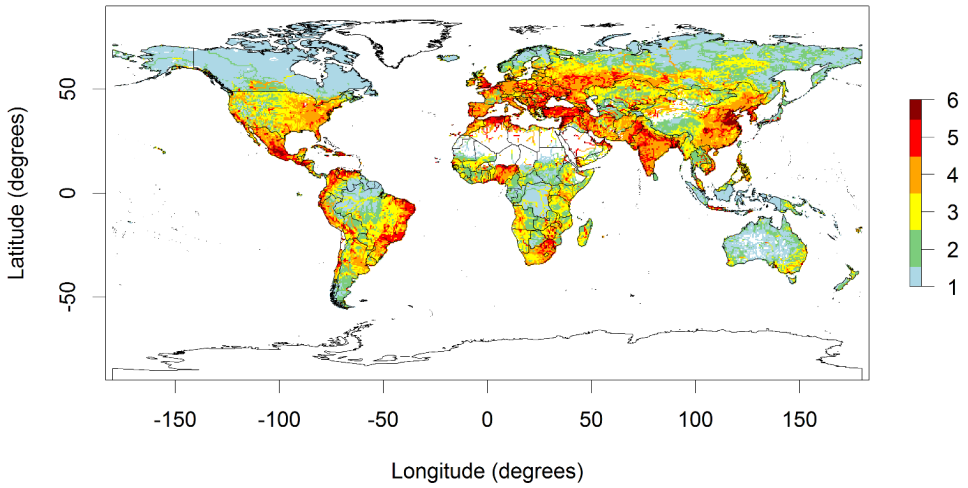


Figure 7.5: Annual mean oocyst concentration categorised to WHO pollution categories (1= very pristine (~0.001 oocyst/L) to 6 = grossly polluted (~100 oocyst/L) (Medema et al., 2009). Each category represents one log unit change in concentrations.

In the Supplementary material S7.4 we summarize our results over seven world regions: Africa, Asia, Europe, Latin America, Middle East – North Africa (MENA), North America and Oceania. We distinguish these regions as our input data for animal oocysts loads make the same distinction. Figure S7.2 shows boxplots of all concentration grid cell values by region and by month, and Figures S7.3-7.9 are concentration maps that zoom in on the different regions.

Figure 7.3 and 7.4 and the regional panels of Figure S7.2 in the Supplementary material show some effect of seasonality, but average monthly oocyst concentrations are predicted not to vary more than approximately one log unit throughout the year in most places. Regions with higher variability (2 log units) include parts of India and Southeast Asia, west Africa, Brazil and the west coast of North America. Of course, this is merely the variability in monthly average concentrations, which does not say anything about variability at shorter time steps. Many studies find that oocyst concentrations can respond strongly to peak runoff events (Atherholt et al., 1998; Kistemann et al., 2002) and may reach much higher levels than indicated by our model, for short time periods. Furthermore, GloWPa-Crypto operates under the assumption that manure is applied equally in all months, which is probably not actually the case in most cropping systems. GloWPa-Crypto should be used to look at 'the bigger picture', and not to make statements about actual levels of contamination in specific locations at specific times.

7.3.2 Source attribution

GloWPa-Crypto indicates that point sources are the dominant source of oocysts in most world regions (Figure 7.6). Point sources are human faeces, diffuse sources are predominantly livestock manure, plus the faeces of the human population that practices open defecation, but the latter is only a small part of the total. Blue regions in Figure 7.6 are grid cells where diffuse sources dominate, red regions are grid cells where point sources dominate. This map is based on average monthly oocyst stream input (variables SD and SP) per grid cell. It should be noted that GloWPa-Crypto L1 includes all *Cryptosporidium* spp., also those that are not infectious for humans (Chapter 6). This means that the risk of contracting cryptosporidiosis from animal or human sources is not equal, and these maps should not be used for assessing risk directly. Dominance of point sources of microbial pollution over diffuse sources of is also found in several other studies, for example in the Scheldt river for *E. coli* (Ouattara et al., 2013) and the European faecal coliform model (Reder et al., 2015). However, diffuse sources were found to dominate in a study on faecal indicators in Vietnam (Nguyen et al., 2016), and it is also observed that dominance can vary for different organisms and under epidemic conditions in Sweden (Sokolova et al., 2012).

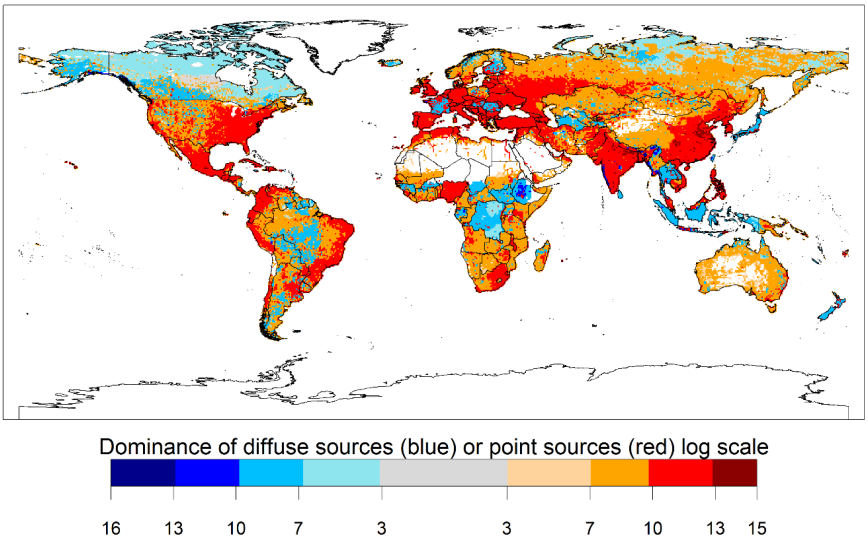


Figure 7.6: Dominance of diffuse sources (blue) or point sources (red). This map shows how much larger the average grid cell oocyst stream input from diffuse sources is than from point sources (log scale), and vice versa. E.g. the lightest shade of blue means that oocysts loads from diffuse sources are 3-7 log higher than oocyst loads from point sources in that area, etcetera.

7.3.3 Model performance

7.3.3.1 Validation

We compared model outcomes to a set of observational data, Table S7.2 and Figure S7.10 give an overview of the data. First we should note that in this set of validation data only 26% of observations could be corrected for recovery efficiency (mean 0.44, range 0.026-0.89). The remainder we used 'as is', but this means that these observed concentrations are almost certainly too low. This problem is also recognized in the literature. For instance, Efstratiou et al. (2017a) note that most published monitoring data do not specify the recovery efficiency. And when specified, it can vary widely, between <10% and >80%, depending on the method, the water characteristics (such as the turbidity) and the person performing the measurement (Efstratiou et al., 2017a). For this reason, we can expect the modelled concentrations to be higher than the observed concentrations, which is indeed the case (Figure 7.7, and Supplementary material S7.5).

The medians of the predicted concentrations mostly fall within the range of observed concentrations, but the model appears to generally produce concentrations that are around 1.5-2 log units higher than the observed concentrations (Figure 7.7, and Supplementary material S7.5). For the three countries that do have a recovery efficiency for all data (Belgium, Luxembourg and Thailand), observed and modelled concentrations are closer together than for the other countries (Figure 7.7), strengthening our notion that correcting for recovery efficiency is important. We recommend that monitoring programs for *Cryptosporidium* should always determine and report the recovery efficiency, preferably for each individual sample, in order for data to be useful for quantitative analysis (Efstratiou et al., 2017a; Ongerth, 2016).

Nevertheless, other factors likely also contribute to the observed difference between predicted and observed concentrations. These could include overestimation of, particularly human, oocyst excretion rates, and underestimation of oocyst losses. The GloWPa-Crypto H1 model uses a simple division between the developing and developed world, assigning each a single oocyst excretion rate, based on the available literature as assessed by Hofstra et al. (2013). Oocyst excretion rates are highly variable and depend on many factors (Chappell et al., 1996), and the model is sensitive to this (see section 7.3.3.2 and Chapter 3). Performing a thorough systematic literature inventory for prevalence of *Cryptosporidium* infection and oocyst concentrations in human faeces, similar to what we have done for livestock manure (Chapter 5), would be a good idea. Doing so was out of the scope of this thesis, but in light of this model validation this should be given priority in further research. The model may be underestimating oocyst losses, particularly because dams, lakes and reservoirs are not yet included. Increased water travel time will lead to higher modelled oocyst losses due to decay and sedimentation (see section 7.3.3.2).

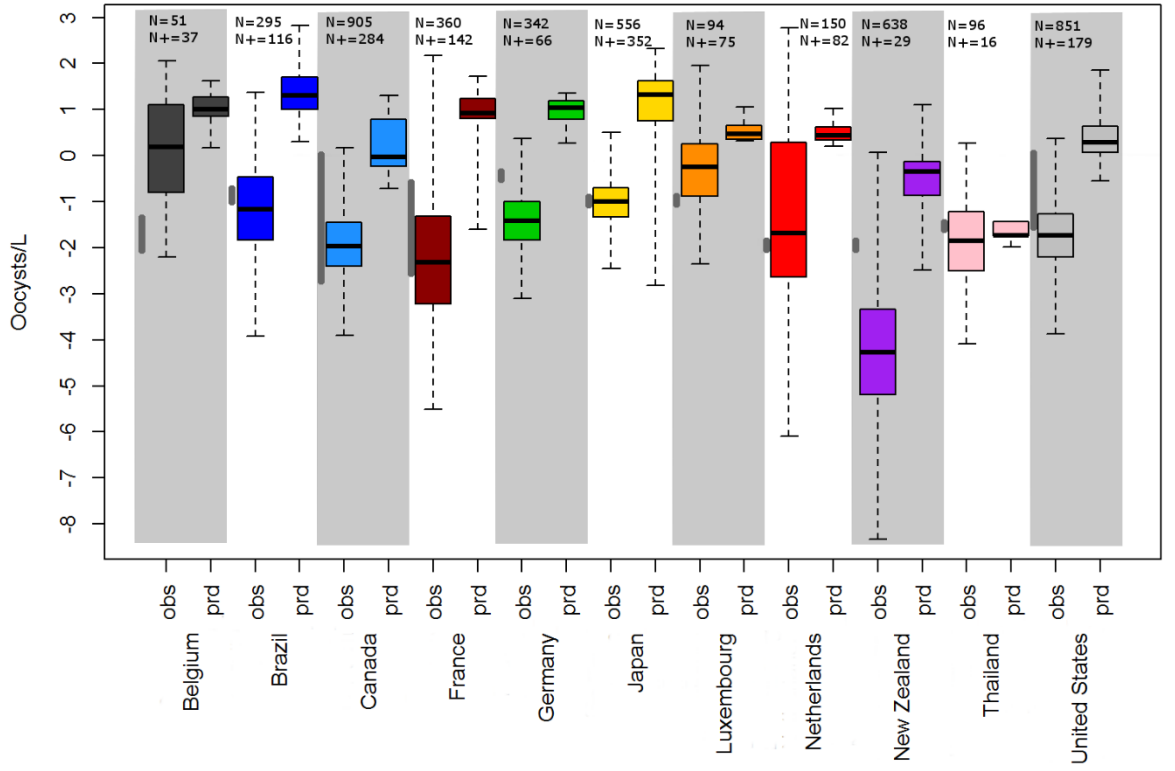


Figure 7.7: Boxplots of observed (obs) and predicted (prd) oocyst concentrations grouped by country. For the observational data, the function 'cenboxplot' from the R package NADA was used, which estimates the distribution of values below the detection limit, assuming the concentration is lognormally distributed (Lee, 2017; Lee and Helsel, 2005). Dark grey lines indicate the spread of detection limits for that country. All data below the detection limit are estimated by the cenboxplot function. N indicates the number of observations per country and N+ indicates the number of observations per country that are above the detection limit. The results are tenuous for countries where >80% of the data are below the detection limit (Lee and Helsel, 2005), which is the case for Germany (81%), Thailand (83%) and particularly New Zealand (95%). All observational data for Belgium, Luxembourg and Thailand and a part of the data for Brazil, Canada and the Netherlands were corrected for recovery, all other observational data were not corrected.

However, as the model is much more sensitive to oocyst excretion rates than to losses, losses alone could not explain all of the observed difference.

To assess model performance numerically, several goodness of fit statistics were calculated using the log transformed results. We calculated rank correlation coefficients, Spearman's rho: 0.42 and Kendall's tau: 0.30. This means there is a moderate positive correlation between observed and predicted values. Next, we applied the Index of Agreement, which compares the sum of the squared error to the potential error (Willmott et al., 1985). A value of 0 indicates no agreement, value 1 is best model performance. We obtained a value of 0.4, which is an average index and similar to what was found for the WorldQual continental faecal indicator model (UNEP, 2016). We found a root mean squared error of 1.87, which is quite large (as it is a log scale), but this is likely a result of the aforementioned differences between predicted and observed oocyst concentrations.

We also looked at the results for different months, but did not find obvious seasonality in observed oocyst concentrations (see Supplementary Material S7.5), consistent with our finding of limited seasonality in modelled concentrations (section 7.3.1). Correlation coefficients computed for the individual countries are worse than for the combined data, meaning that the model can reasonably predict differences between locations but does not perform well for predicting variability for a specific location. This is probably because variability in oocyst concentrations is inherently high. Figure 7.7 shows that the spread in observed concentrations is generally much larger than the spread in modelled concentrations. This is not surprising, as the model calculates monthly grid cell means, and the observations are points in time and space. Furthermore, observations from different years were taken together in the validation set, as the amount of data did not justify looking at trends over time, and because GloWPa-Crypto uses 30-years average climate and hydrological data as input.

7.3.3.2 Sensitivity analysis

The model is most sensitive to changes in input human oocyst loads (Table S7.3 in the Supplementary Material) and to a lesser extent to input animal loads, as human loads were found to dominate in most places. This is not surprising as for the sensitivity analysis this input was changed one log unit to reflect the variable and uncertain nature of oocyst excretion rates, even between infected individuals oocyst production can differ strongly (Chappell et al., 1996)(Chapter 5). Other models of pathogens or faecal indicators have similarly observed sensitivity to pathogen loading from humans and animals (R Coffey et al., 2010; Ferguson et al., 2007; Reder et al., 2017; Tian et al., 2002). To a lesser extent, the model was sensitive to oocyst retention, river length and water residence time in a grid cell, these affect the time during which decay and sedimentation take place.

We performed a low end and high end run, combining changes in all parameters, to investigate the combined effect on model output (see description in Supplementary material S7.6). The high and low end runs cause a 1.5 log change in median concentration in either direction. This 3 log difference yields very different WHO category maps of the low end and high end runs (Figure 7.8). While in the low end run the majority of the grid cells around the world falls in categories 1-3 (Very pristine, Pristine and Moderately polluted), in the high end run categories 4-6 (Polluted, Heavily polluted and Grossly polluted) are most dominant. Furthermore, the low end run shows that, even at the most optimistic end of model assumptions, there are some regions that come out as highly polluted in any case.

When comparing the difference between the high and low end run spatially, several points come to the attention: 1) the absolute difference between the high and low end runs (high minus low) is largest in areas where concentrations are highest, which are generally areas dominated by human sources, 2) the relative difference between the high and low end runs (high divided by low) is largest in the areas where the diffuse sources dominate (see Figure 7.6), which means the diffuse sources are relatively more uncertain (especially the fraction of oocysts in manure that is transported with runoff to surface water) but compared with the magnitude of the human sources this does not matter so much, and 3) the large rivers stand out, meaning that the uncertainty accumulates with the routing, which is a logical consequence of the way the model is constructed.

7.3.4 Model improvements and opportunities

Three important potential model improvements that we envision are: 1) refining human oocyst load input by the H1 model, by thoroughly analysing the literature on human oocyst excretion rates, 2) refining oocyst load input by the L1 model, for instance by incorporating monthly variation in manure input to the land, through including manure spreading regimes and birthing seasons, 3) including dams, lakes and reservoirs, instead of working with naturalized discharge, and related to this improving estimates of river flow path length and water residence times, and 4) refining the calculation of oocyst retention on land if more observational data become available.

GloWPa-Crypto C1 could be applied in scenario analysis to investigate the impact of global change and management options on oocyst concentrations in rivers. Furthermore, the model could serve as a basis for risk assessment studies to assess human health impacts, if the output were to be used as input for Quantitative Microbial Risk Assessment (QMRA). Especially for the latter application, it would be interesting to look into the option to make GloWPa-Crypto C1 a stochastic model, so that instead of calculating monthly average concentrations, it would provide output concentration distributions. Furthermore,

studying model output for the different continents in more detail would be interesting. GloWPa-Crypto C1 could also be downscaled to operate for a specific region with more detailed input data.

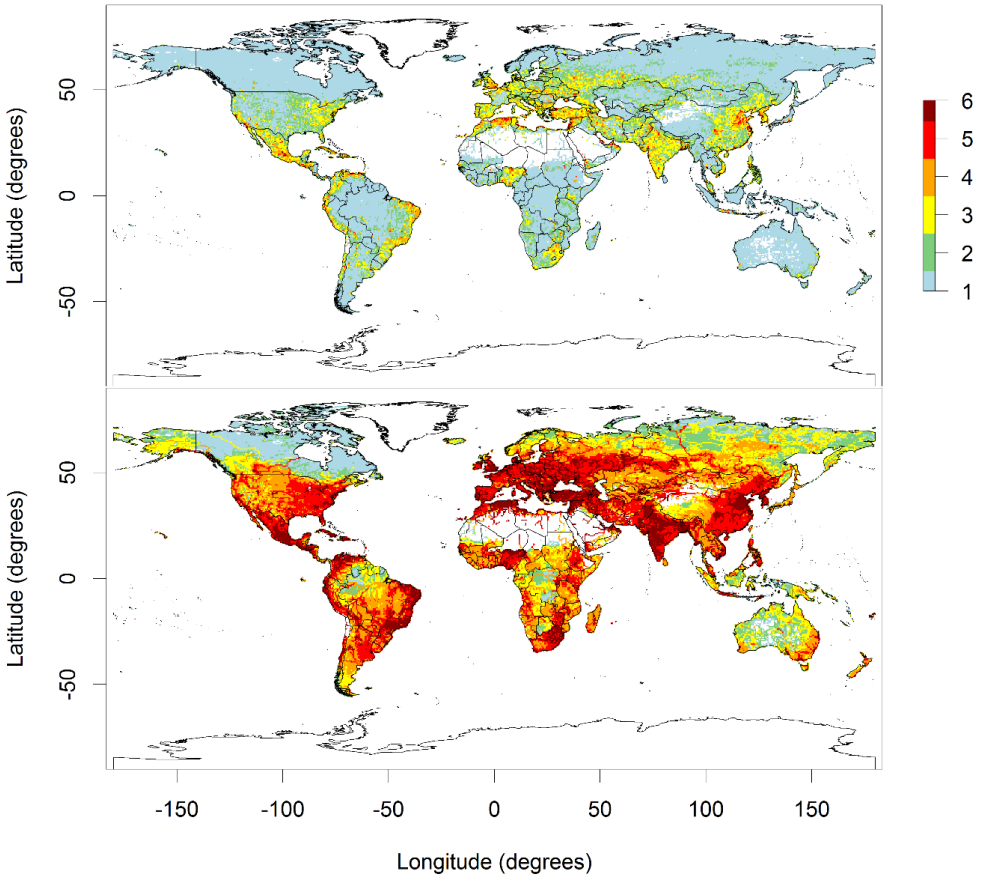


Figure 7.8: Low end (top) and high end (bottom) runs of annual mean oocyst concentration categorised to WHO pollution categories (1= very pristine (~0.001 oocyst/L) to 6 = grossly polluted (~100 oocyst/L). Each category represents one log unit change in concentrations.

7.4 Conclusion

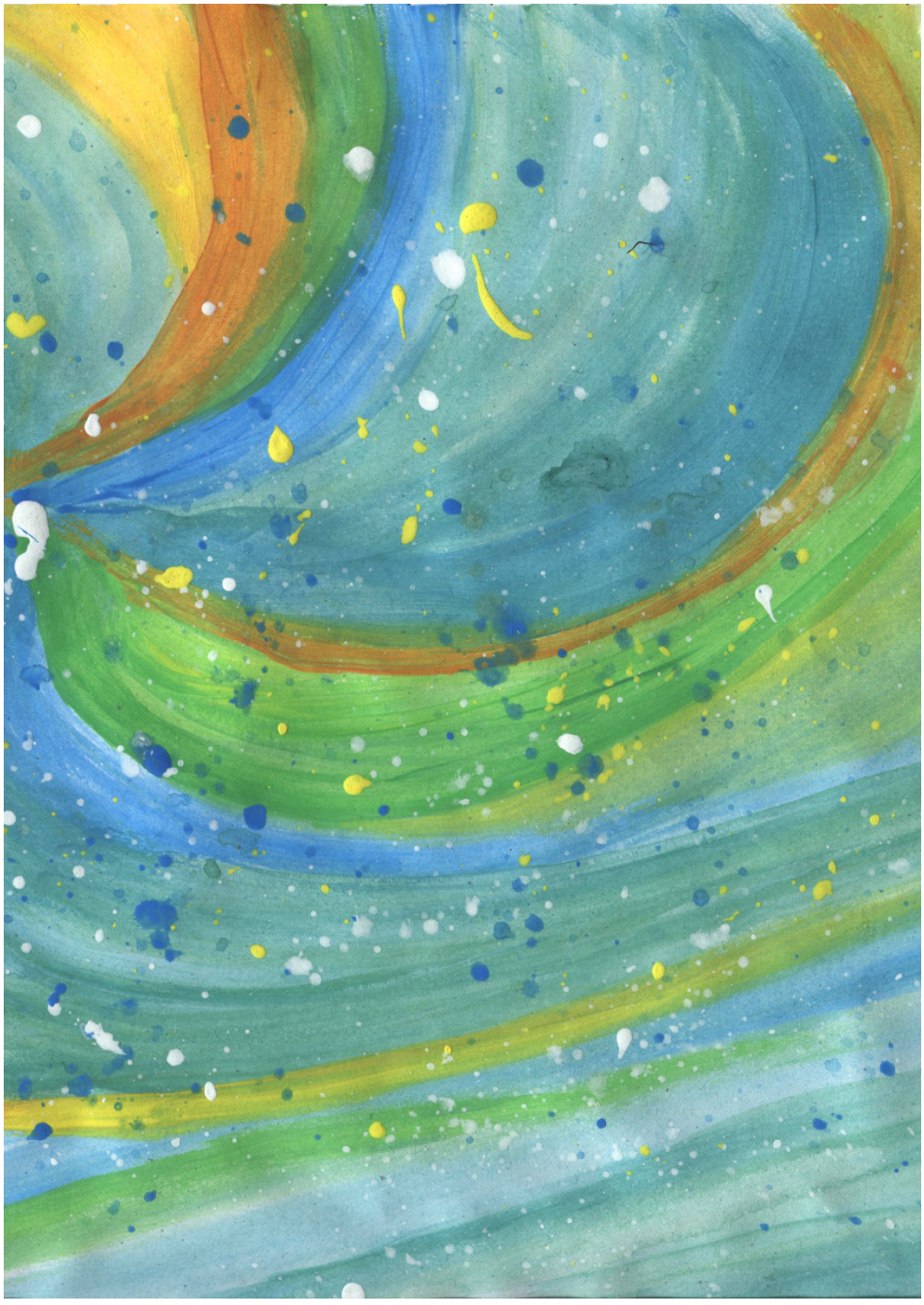
GloWPa-Crypto C1 is the first global model that computes pathogen concentrations in rivers. Monthly average oocyst concentrations are predicted to range from 10^{-6} to 10^{2} oocysts L^{-1} in most places. Hotspot regions with high concentrations include parts of India, China, Pakistan and Bangladesh, Nigeria, Algeria and South Africa, Mexico, Venezuela and some coastal areas of Brazil, several countries in Western and Eastern Europe, and the Middle East. Monthly average oocyst concentrations are predicted not to vary more than approximately one log unit throughout the year in most places. Point sources (human faeces) appears to be a more dominant source of pollution than diffuse sources (animal manure) in most world regions.

Due to the absence of recovery efficiency for most validation data, these are likely underestimating oocyst concentrations in rivers, and consequently modelled concentrations are expected to be higher. Model validation showed that GloWPa-Crypto medians are mostly within the range of observed concentrations, but indeed the model generally produces concentrations that are 1.5 – 2 log higher than the observations. Goodness of fit statistics are reasonable. We recommend that monitoring programs for *Cryptosporidium* should always determine and report the recovery efficiency and detection limit. Sensitivity analysis showed that the model is most sensitive to changes in input oocyst loads, river discharge and oocyst retention on land. Furthermore, river length and water residence time in a grid cell are quite influential on model outcome, these affect the time period under which decay and sedimentation take place.

GloWPa-Crypto C1 paves the way for many new opportunities at the global scale, including scenario analysis to investigate the impact of global change and management options on oocysts concentrations in rivers, and risk analysis to investigate human health risk.

Acknowledgements

Lucie Vermeulen was financially supported by the Gerbrand de Jong Fonds and the EU Climate KIC programme. We thank the many individuals who helped us by providing observational data for model validation: P. Berger, J.-B. Burnet, C.J. Chuah, E. Claerebout, S. Dorner, J. Enault, S. Itoh, M.W. LeChevallier, J.-F. Loret, G. McBride, M.T. Pepe Razzolini, M. Prévost, A. Rechenburg, B. Rempel, N. Ruecker, C. Schreiber, E. Sylvestre and M.I. Zanoli Sato. Their organisations are listed in the Supplementary Material S7.5. We thank J. van Geffen and M. van Weele of the Royal Netherlands Meteorological Institute (KNMI) for discussions on UV radiation. We thank B. King and P. Monis from S.A. Water and the University of Adelaide for discussions on their data of *Cryptosporidium* decay in surface waters.



Chapter 8

Synthesis

8.1 Objectives and main findings of the thesis

The objective of this thesis was to increase our knowledge on the sources, fate and transport of *Cryptosporidium* in rivers worldwide using spatially explicit modelling. This objective has been divided in five sub-objectives that were formulated as:

- Obj. 1: To learn from existing watershed-scale waterborne pathogen models and global nutrient export models for the development of a global waterborne pathogen model. (Chapter 2)
- Obj. 2: To develop a spatially explicit model to study the impact of sanitation changes, urbanisation and population growth on human emissions of *Cryptosporidium* to surface waters. (Chapters 3 and 4)
- Obj. 3: To collate and summarise published data on prevalence of *Cryptosporidium* infection in common livestock species and oocyst concentrations in livestock manure from the literature and to analyse their distribution by age and region. (Chapter 5)
- Obj. 4: To quantify livestock *Cryptosporidium* spp. loads to land on a global scale using spatially explicit process-based modelling, and to explore the effect of manure storage and treatment on oocyst loads using scenario analysis. (Chapter 6)
- Obj. 5: To estimate *Cryptosporidium* oocyst concentrations in rivers worldwide. (Chapter 7)

These objectives have been addressed in six scientific papers, presented in Chapters 2-7. Figure 8.1 recaps an overview of the research presented in this thesis, including the main findings from the different chapters. I found that urban areas in developing countries are hotspot regions, and human faeces appears to be a more dominant source of pollution than animal manure. Sanitation and treatment improvements are key to improve water quality now and in the future. *Cryptosporidium* infection is ubiquitous in livestock around the world, and cattle, especially calves, are expected to contribute most to the global load. This thesis has shown that waterborne pathogen modelling at the global scale is feasible and provides new opportunities: providing information on pathogen concentrations in data-sparse regions, identify hotspot regions with high concentrations, identify the relative contribution of different sources, and allow for scenario analysis. Examples of these opportunities have been demonstrated in this thesis.

In this synthesis chapter, a general discussion and conclusions are presented. Section 8.2 reflects on the methodology used in this thesis, focussing on methodological choices in model development (8.2.1), model performance, discussing validation, comparison with other modelling work and sensitivity analysis (8.2.2), and ending with an overview of research needs (8.2.3). Section 8.3 discusses the results in a broader context, reflecting on modelling opportunities and model applicability (8.3.1), and coming back to issues introduced in Chapter 1: global change (8.3.2), application in risk and burden of disease calculations (8.3.3) and the Sustainable Development Goals (8.3.4). Finally, in section 8.4 overall conclusions are drawn, summarizing the main findings of the thesis and thoughts from the discussions in this chapter.

8.2 Reflections on methodology

In this section, I will first reflect on methodological choices that were made in this thesis, focussing on why this particular type of modelling was chosen. Next, I will discuss model performance of GloWPa-Crypto, paying specific attention to validation, comparison with other modelling work and sensitivity analysis. Lastly, an overview of research needs is given.

8.2.1 Methodological choices in model development

When developing a model, many design choices have to be made. GloWPa-Crypto has been developed as a deterministic process-based model that contains some empirical parts. It is not calibrated, and it is static. It is spatially explicit, implemented at the global scale, with a spatial resolution of 0.5 x 0.5 degree grid, and a monthly time step. In this section I will reflect on these choices, coming back to the desirable characteristics of a global waterborne pathogen model as identified in Chapter 2.

All models are developed with a certain objective in mind, that guides the choices that are made. Furthermore, choices are often pragmatically based on the available input data and computation time. The objective of this thesis was to increase our knowledge on the sources, fate and transport of *Cryptosporidium* in rivers worldwide using spatially explicit modelling. Opportunities that we identified for a global waterborne pathogen model were to provide information on pathogen concentrations in data-sparse regions, identify hotspot regions with high concentrations, identify the relative contribution of different sources, and allow for scenario analysis. In Chapter 2, it was suggested that for these aims, developing a static, deterministic model with a spatial and temporal resolution such as currently applied in global nutrient export modelling (mostly static annual estimates on a 0.5 x 0.5 degree grid) would be a reasonable choice, and feasible to do at the global scale with the available input data.

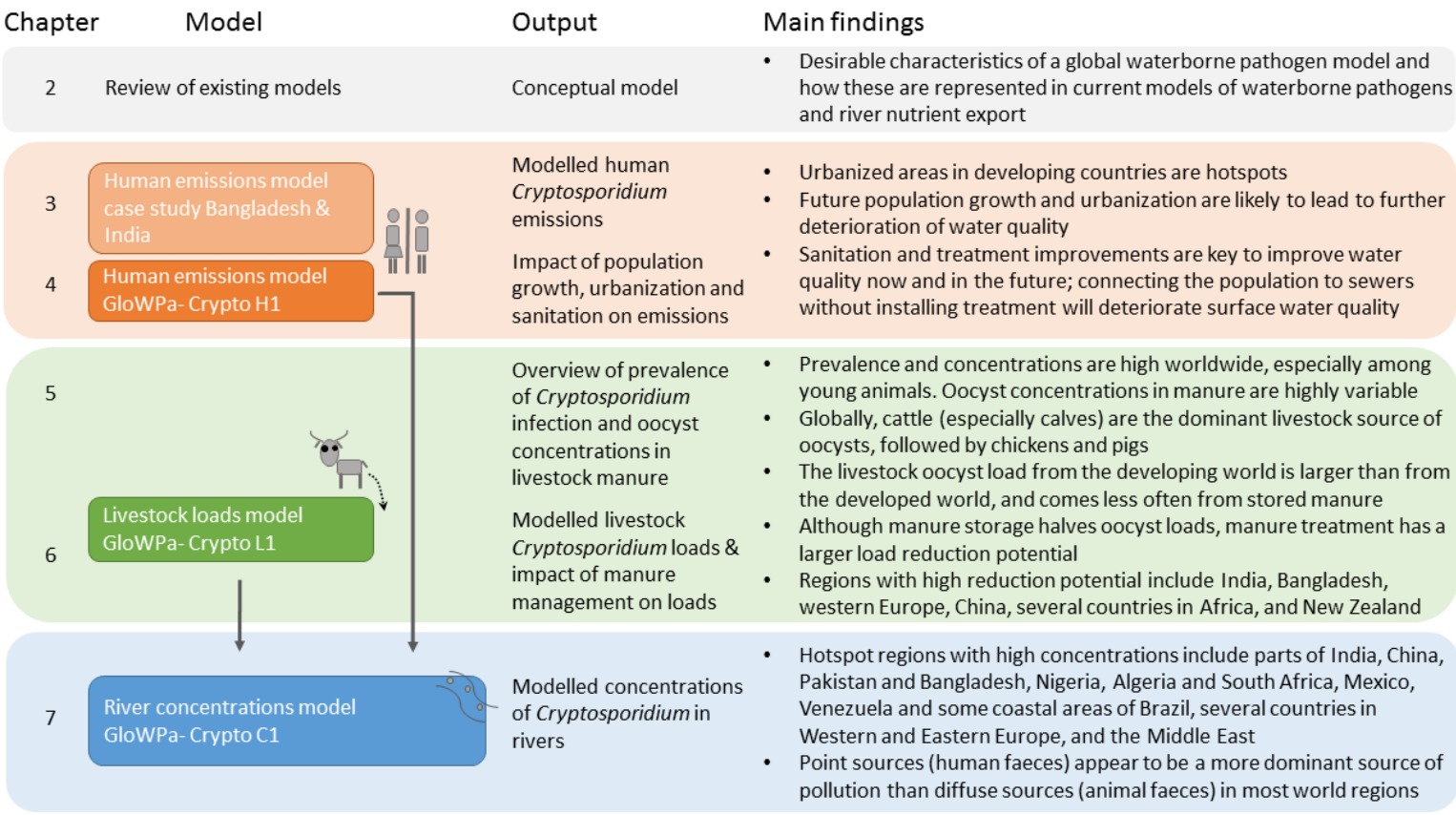


Figure 8.1: Schematic overview of the research and main findings of this thesis. This figure shows the thesis chapters, the models that have been developed therein, their main output and the main findings per chapter. See also Figure 1.2.

The global scale was chosen also because there was a clear opportunity for a waterborne pathogen model at this scale as this did not exist yet, and it is especially relevant since most knowledge is available for industrialized countries while problems are largest in the developing world. GloWPa-Crypto was therefore set up along these lines, global scale, static, deterministic, and on a 0.5 x 0.5 degree grid. It was also noted in Chapter 2 that for studying global change impacts and risk analysis, a finer spatial and temporal resolution may be more desirable. Therefore, GloWPa-Crypto was set up not on an annual but on a monthly time step, to facilitate going in this direction. The application of the model for the analysis of global change impacts and human health risk is further discussed in sections 8.3.2 and 8.3.4.

The choice for an uncalibrated process-based model was in part pragmatic, as there is simply insufficient observational data for *Cryptosporidium* in rivers to allow for developing a calibrated model at the global scale. Furthermore, the variability in observed oocyst concentrations is so high, that this may introduce additional uncertainties in the model. I consider it an accomplishment that this un-calibrated process-based model produces output that appears to be in a realistic range (Chapter 7, section 8.2.2). The advantage of process-based modelling is that everything is made explicit and the model is not a black box, which allows for exploring the impact of the different variables (de Brauwere et al., 2014). It should be noted that some parts of GloWPa-Crypto do have empirical elements: the decay functions in manure and in water, and the transport of oocysts with runoff from manure on land to rivers are based on experimentally determined relationships. This is done because explicitly modelling all processes that are taking place during decay or transport with runoff is 1) not possible with the available knowledge on these processes, and 2) not feasible at the global scale due to a lack of input data.

In the development of GloWPa-Crypto, like in the development of all other models, I needed to find a balance between complexity and simplicity. Making models more complex by introducing many variables and processes does not necessarily make them better, as this invariably also introduces additional uncertainties, while not necessarily providing a better answer to the questions the model is intended to answer. Choices for simplicity or complexity invariably bring about tensions and discussions. An example is our choice for grid-based modelling, while a part of the input data (e.g. urban/rural sanitation use and urbanisation statistics) are country statistics that have been downscaled to grid level. In this downscaling, grids were assigned to be either urban or rural based on their population density. The assumption was made that all urban grids within a country have the same proportion of sanitation systems in use, while it may very well be that in one urban grid everyone is connected to a sewage system while in the next grid they use latrines. This introduces uncertainties in the model, but as human and animal population, hydrology and climate data are available on a 0.5 x 0.5 degree grid, it would be a great loss

of information if we would have made GloWPa-Crypto provide only country-based output. And with country-based output, the model could not meet the specified aims of indicating hot spot areas, etc. Therefore, we chose to apply downscaling for the sanitation statistics. As the human and animal population densities mostly govern the spatial patterns in the output, I think this was a reasonable choice.

In the development of GloWPa-Crypto, I first looked at existing watershed-scale waterborne pathogen models and global nutrient export models (Chapter 2). GloWPa-Crypto is different from global nutrient export models in that calculating fresh water concentrations is most interesting, as these are most relevant for health risk. Global nutrient models, on the contrary, focus on pollution of coastal areas from river nutrient export. This also guided our choice for grid-based modelling, instead of aggregating for watersheds, as is done for example in Global NEWS (Mayorga et al., 2010). However, it is argued that global nutrient modelling should ideally also go in the direction of more distributed modelling (Kroeze et al., 2012). Watershed-scale waterborne pathogen models mostly operate on a finer spatial and temporal resolution than GloWPa-Crypto does, dynamic approaches are common and many models are calibrated (Chapter 2). Furthermore, more detailed processes are often included, for example in partitioning microorganisms to different states or considering different layers in the water. It makes a lot of sense to do this at the watershed level, but this level of detail is not feasible at the global scale. In Chapter 2 an inventory was made of processes included in models, leading to a conceptual global waterborne pathogen model (Figure 2.2).

GloWPa-Crypto has been developed largely following this conceptual model, but several changes were made along the way. Some processes were not included after all: direct defecation by animals into streams, overflows of untreated sewage, and resuspension from sediments. The first two were considered, but it was decided that too little data were available to make an adequate assumption that would not introduce too much uncertainty. Resuspension from sediments was not included for the same reason, and because dams, lakes and reservoirs have not yet been included in the model, meaning that we already know that we are doing this part wrong, and including resuspension in this version would still be wrong. This would be worthwhile to further investigate in a next version of the model. Compared to the conceptual model, extra processes were also added: the most important one is the addition of manure storage and its effect on oocyst decay. In Chapter 6 we calculated that this reduces the total livestock oocyst load estimate by a factor 2.6 and also has consequences for the importance of different livestock species, so this was deemed a relevant addition to the model.

8.2.2 Model performance

Multiple ways exist in which confidence in model performance can be enhanced, Augusiak et al. (2014) coined the term 'evaludation' for this process. Approaches to build confidence in model performance include: 1) transparency about model input data and assumptions, 2) expert judgment on model strengths and weaknesses (e.g. have all relevant processes been incorporated and does the model mimic the natural system as would logically be expected), 3) comparing model outcomes to observational data (validation), 4) comparing model outcomes to other modelling studies and 5) performing sensitivity analysis to judge the effect of changes in input variables (that can be based on variability and uncertainty) on model output. This thesis incorporates all five of these approaches where possible and appropriate.

The first and second approach are demonstrated in documenting the model in a clear, concise and comprehensive way, providing detail in Supplementary materials where possible. Furthermore, Chapters 2-7 are already published or are intended to be published in scientific journals, and as part of that undergo the process of peer review. Additionally, the results of several chapters have been presented and discussed at various international conferences and meetings, where further feedback from experts was invited. Approaches 3-5 are discussed in more detail in sections 8.2.2.1-8.2.2.3.

8.2.2.1 Validation

The third and perhaps most widely appreciated approach to enhance confidence in model performance is comparing model outcomes to an independent set of observational data. This is perhaps also the most challenging approach due to the scarcity of suitable data, which is recognized in many other modelling studies of pathogens and faecal indicators (Cho et al., 2016; Oliver et al., 2016). This also holds for other water quality variables such as nutrients, but data on pathogens and faecal indicators are relatively even more scarce, and generally not easily publically accessible (Cho et al., 2016; Monaghan et al., 2008). I have managed to compile a set of >4000 observational data to validate GloWPa-Crypto C1 with (Chapter 7).

Due to the absence of recovery efficiency for most validation data, these are likely underestimating oocyst concentrations in rivers, and consequently modelled concentrations are expected to be higher. Model validation showed that GloWPa-Crypto medians are mostly within the range of observed concentrations, but indeed the model generally produces concentrations that are 1.5 – 2 log higher than the observations. Goodness of fit statistics are reasonable: Spearman's rho: 0.42, Kendall's tau: 0.30 and Index of Agreement: 0.4 (Chapter 7). It appears that the model can reasonably predict differences between locations but does not perform well for predicting variability for a

specific location, since correlation coefficients computed for the individual countries are worse than for the combined data. It is not surprising that GloWPa-Crypto does not capture all variability, as the model resolution (0.5 x 0.5 degree grid and monthly time step) is rather coarse when comparing it to individual samples taken at a specific location at a point in time. Furthermore, monitoring data from different years were all put together in the validation set, as the amount of data did not justify looking at trends over time. GloWPa-Crypto uses climate and hydrological data averaged by month over a 30 year period. This means that variation between months is captured to some extent, but not between months in different years, while of course weather conditions might differ between years. Furthermore, GloWPa-Crypto is not (yet) designed to capture peak concentration events, which would be relevant for risk assessment (see Chapter 2 and section 8.3.3).

It can be argued that evaluation of model performance should be targeted towards model application rather than abstract accuracy of the model. Microbial water quality standards are typically standards of descriptive statistics of a set of monitoring data – such as means, medians or percentile values (Cho et al., 2016; Stott et al., 2011), because microbial water quality is inherently variable. Therefore, it makes sense to evaluate model performance by the accuracy of reproducing these descriptive statistics rather than point-to-point comparisons with monitoring data, and indeed examples exist of models that perform validation while doing some kind of 'lumping' (Cho et al., 2016; UNEP, 2016). Therefore, we also made the choice to present the validation results mostly in boxplots lumping different measurement locations and years together (Chapter 7).

Even though we made an effort to collect sufficient suitable data for validation, we ran into challenges finding enough suitable data, especially with reported recovery efficiency and preferably with sufficient sensitivity to limit the occurrence of censored data. A current challenge is therefore to develop effective monitoring schemes that help improve models but are also an efficient use of funds spent on sampling (Cho et al., 2016). Compilation of already available datasets is required, and observations at regular intervals over a longer period (multiple years) should be encouraged in monitoring programs around the world. We recommend that monitoring programs for *Cryptosporidium* should always determine and report the recovery efficiency and detection limit with their data (Chapter 7). Looking at the model hotspots, I suggest that monitoring programmes could be especially beneficial in data-sparse, urbanised, densely populated areas in developing countries, and to pick areas with differing hydrological and climatic conditions for comparison. Following this line of thought, interesting regions to focus monitoring programmes on include:

- India and Bangladesh: these countries harbour large numbers of both humans and cattle, section 7.3.2 suggests that both human and animal sources dominate in parts of the country, urbanisation and population growth are rapid and open defecation is still relatively common. The high climate variability and flood risk that comes with the monsoon rains is another feature that likely influences water quality throughout the year.
- Nigeria: the country that stands out most on the oocyst concentration map of Africa (Chapter 7). Nigeria is an emerging economy with high population density and growth, and it is a delta region where the Niger river drains into the ocean. This delta region has been observed to suffer from faecal pollution (MacPepple et al., 2017).
- The Middle-East North Africa (MENA) region: this region shows high concentrations on the map (Chapter 7), and is also interesting as it is a relatively dry region. Low flows can lead to high concentrations due to limited water availability for dilution of wastewater flows.
- Mexico City region: an example of a densely urbanised area with high expected concentrations.
- East China: a rapidly developing and urbanizing region with high population density and high expected concentrations. Nutrient modelling done for this region expects pollution increases for the future; it would be interesting to compare these estimates with an analysis for pathogens (Strokal et al., 2016). Furthermore, direct manure inputs into rivers have been identified as a major source of nutrients to water in this region (Strokal et al., 2016); as this has not been included in GloWPa-Crypto it would be interesting to study the potential relevance for pathogen pollution.

For all abovementioned regions it would be highly interesting and relevant if more water quality monitoring data were available, to inform decision makers and to evaluate the performance of GloWPa-Crypto in more detail. A relevant new initiative with respect to collecting data and knowledge is the Global Water Pathogen Project (GWPP), which is implemented by the International Hydrological Programme of the United Nations Educational, Scientific, and Cultural Organization (UNESCO), and Michigan State University (http://www.waterpathogens.org/). This initiative aims to be a state-of-the-art knowledge hub on waterborne pathogens, and it could be a platform on which data for model validation are compiled.

8.2.2.2 Comparison with other modelling work

The fourth approach to enhance confidence in model performance is to compare model outcomes to other modelling work. The only model truly comparable to GloWPa-Crypto in

scale and scope is the continental scale faecal coliform model WorldQual that is part of the WaterGAP framework (Reder et al., 2017, 2015; UNEP, 2016), as this is the only other large scale microbial water quality model. It should be noted that faecal coliforms are not pathogenic, but rather an indicator of faecal pollution, and faecal indicators and pathogens do not necessarily correlate well (Committee on Indicators for Waterborne Pathogens National Research Council, 2004; Lin and Ganesh, 2013; Payment and Locas, 2011). Faecal coliforms have other environmental sources than human faecal waste, such as the discharge of paper mills, and regrowth in the environment and drinking water distribution systems can also occur (Committee on Indicators for Waterborne Pathogens National Research Council, 2004; Gauthier and Archibald, 2001; Lin and Ganesh, 2013). This might lead to problems in model calibration and validation of WorldQual. Another difference between *Cryptosporidium* and faecal coliforms is the survival rate in water; oocysts are highly resistant to environmental stressors and could thus be expected to survive longer and be found farther downstream from a pollution source than faecal coliforms. For these reasons, GloWPa-Crypto was developed for pathogenic organisms directly. Moreover, this provides more relevant data in the context of risk assessment (section 8.3.3).

Nevertheless, I can compare our findings with regards to spatial patterns and hotspots to the findings of the WorldQual model. Figure 3.3 of the UNEP report shows faecal coliform concentrations for the month February in Latin America, Africa and the southern half of Asia (UNEP, 2016). When comparing this figure to a GloWPa-Crypto concentration map, it can be seen that similar patterns emerge (Figure 8.2). Both models find high concentrations in, for example, Mexico, along the west and east coasts of South America, the north coast of Africa, the Middle East, the Ganges basin, east China and Vietnam. It is not surprising that these patterns are similar, as they are mostly driven by population density of humans and to a lesser extent animals. The largest apparent discrepancies are Nigeria, that comes out as a hotspot in GloWPa-Crypto but not in WorldQual, and Sudan, that is a hotspot in WorldQual but not in GloWPa-Crypto. This finding of WorldQual is surprising to me, as the population density in Nigeria is nearly 10 times higher than in Sudan (United Nations DESA/Population Division, 2017). However, in their spatially explicit uncertainty analysis for Africa, Reder et al. (2017) note that the Sudan region has relatively high uncertainty. A possible reason may be that Sudan is a very dry country, and low dilution may lead to very high concentration estimates, an effect that we have also observed in GloWPa-Crypto but not so much for this particular region.

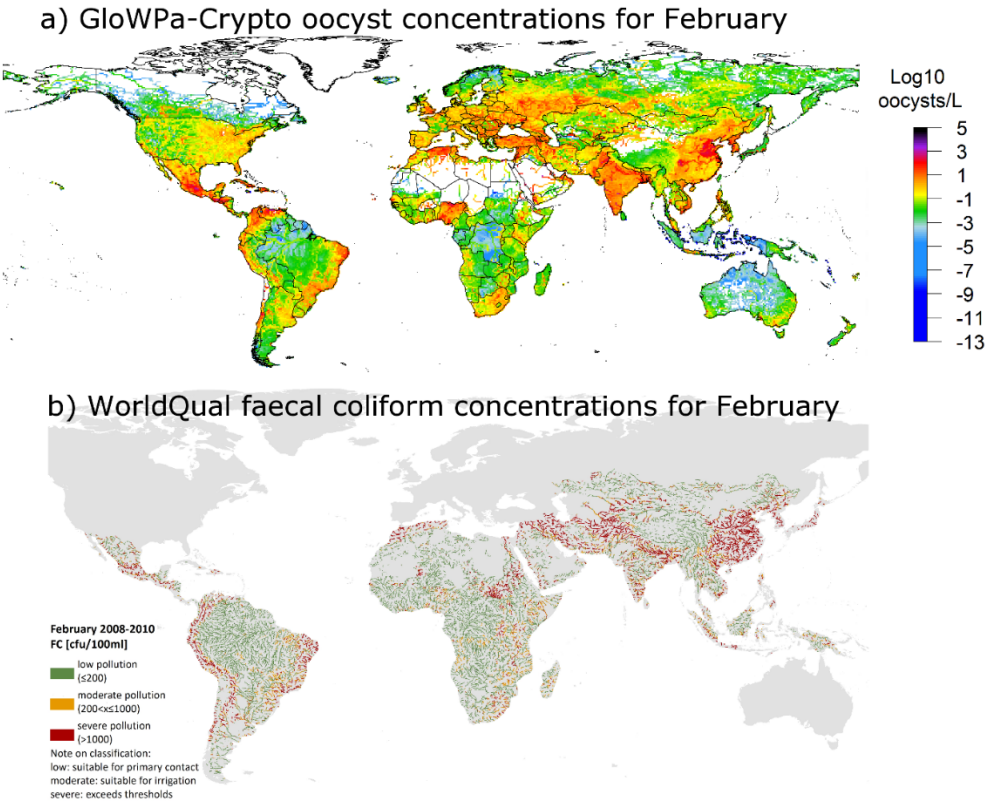


Figure 8.2: Comparison of GloWPa-Crypto oocyst concentration map (panel a) and WorldQual faecal coliform concentration map (panel b) for the month February. The WorldQual map has been produced by the Centre for Environmental Systems Research at the University of Kassel and was adapted from 'A Snapshot of the World's Water Quality: Towards a global assessment' (UNEP, 2016).

Notwithstanding the differences for Nigeria and Sudan, overall it is clear that the spatial patterns that are observed in GloWPa-Crypto and WorldQual are very similar, enhancing our confidence in the performance of the model. It would be worthwhile to conduct a more detailed model comparison for these two models, and perhaps also for some selected river basins for which basin-scale models of microbial water quality are available (Chapter 2). Plans to do so have been discussed at the workshop "Water Quality: a new challenge for global scale modelling" that took place 18-21 September 2017 in Wageningen. Model intercomparison projects have provided valuable knowledge to other fields, such as climate models and climate impact models (Taylor et al., 2012; Warszawski et al., 2014).

When comparing GloWPa-Crypto to large scale nutrient modelling, some findings are also similar: population density, connection to sanitation systems and treatment are important (especially for phosphorus) and agricultural management is important (especially for nitrogen) (Seitzinger et al., 2010), as are estimates on nutrient retention (i.e. losses within rivers) (Beusen et al., 2015a; Seitzinger et al., 2010). In the future, water pollution from nutrients and coastal eutrophication risk is expected to increase (Seitzinger et al., 2010; Strokal et al., 2016), as we also expect to happen for waterborne pathogens (Chapter 4). Differences also exist for nutrients and pathogens, in sources of pollution (for example, artificial fertilizers are an important source of nutrients to rivers and these do not contain any pathogens), in modelled processes (such as the processes governing in-stream losses and the uptake of nutrients by plants), and in impacts (such as the focus on coastal algal blooms in nutrient modelling). These similarities and differences have been reviewed in Chapter 2 of this thesis. Notwithstanding these differences, the general findings of similar important sources and trends in nutrient modelling enhance confidence in the performance of GloWPa-Crypto.

8.2.2.3 Sensitivity analysis

The fifth approach, sensitivity analysis, has been applied in various chapters (3, 6, 7). Sensitivity analysis is a procedure to determine how much of the variability and uncertainty in model outputs can be appointed to different model inputs. It is done to judge the robustness of model results. Sensitivity analysis can be done in multiple ways. A general distinction that can be made is between local (one-at-a-time) and global (all-at-once) approaches (Saltelli et al., 2004). In this thesis nominal range sensitivity analysis has been applied, a form of local sensitivity analysis. In nominal range sensitivity analysis, the parameters are changed one at a time over a reasonable range the parameter can take, while keeping all other parameters at their baseline (nominal) value. For global sensitivity analysis, a common approach is that all parameters are varied over their probability density functions at the same time in multiple runs (Monte Carlo sampling) (Oakley and O'Hagan, 2004; Saltelli et al., 2004). The advantage of such an approach is that the full range of outcomes is covered and possible interactions between parameters are accounted for. The disadvantage, however, is the high computational costs. Probably for this latter reason, local sensitivity analysis has been most common for microbial water quality modelling (Reder et al., 2017).

Our various nominal range sensitivity analyses (Chapters 3, 6, 7) indicate that sensitive parameters include oocyst excretion by humans and animals, the share of the population using different types of sanitation facilities, the proportion of manure going to storage and to land, oocyst retention in the faecal matrix and on soils, and parameters affecting the time during which loss (decay and sedimentation) during river transport takes place (river

length and residence time). It is not surprising that the oocyst excretion rate is the most sensitive variable, as oocyst excretion is inherently variable and uncertain. It can vary over orders of magnitude and even between infected individuals oocyst production can differ strongly (Chappell et al., 1996). For this reason we performed a systematic literature inventory to identify prevalence of *Cryptosporidium* infection and oocysts concentration in manure for the eleven livestock species that have been included in the model and specified these for age and world region where the data allowed (Chapter 5). For human oocyst excretion, I took the estimates from the exploratory model by Hofstra et al. (2013), who based it on available literature. It would be worthwhile to also perform a larger more systematic literature review summarizing human prevalence of *Cryptosporidium* infection and oocyst concentration in faeces by age and world region, to reduce model uncertainty. However, in the current model set up, the oocyst excretion determines the order of magnitude but not the spatial patterns of the result, so our conclusions on hotspot regions etc. will likely not be affected, unless regional differences in prevalence turn out to be substantial.

Reder et al. (2017) performed a global sensitivity analysis of their continental faecal indicator model WorldQual for Africa in a spatially explicit way. They found that the most sensitive parameters differ geographically, but also determined that only a small number of parameters dominate output uncertainty in most of the continent. Their overall conclusions on the importance of different parameters are not affected much by the spatial analysis, and the parameters they find are in fact similar to the parameters that we find to be most sensitive. They identified as sensitive parameters the human excretion rate, the population connected to different sanitation systems, the loss rate due to sedimentation and manufacturing effluent. This outcome is very similar to our findings; the first two were also identified by GloWPa-Crypto as highly sensitive parameters, the third to a somewhat lesser extent, and the fourth was not included by GloWPa-Crypto as manufacturing effluent is not likely to contain *Cryptosporidium*. We have not performed a full spatially explicit sensitivity analysis, but made a start using estimates of the low end and high end of model variables (section 7.3.3.2). Although it would certainly be interesting to do a global spatially explicit sensitivity analysis for GloWPa-Crypto, and defining probability distributions for all input parameters is something that is desirable when applying the model in risk analysis (section 8.3.3), we are confident that the results of our local sensitivity analyses here give a good general overview of model sensitivity. With the available knowledge and data GloWPa-Crypto is currently the best estimate for *Cryptosporidium* concentrations in rivers worldwide.

8.2.3 Research needs

Following from the conclusions of the research chapters in this thesis and the discussion in sections 8.2.1 on methodological choices in model development and 8.2.2 on model performance, main research needs can be identified that would improve the GloWPa-Crypto model. Here I list some of the main research needs, where I would give the top half of the list more priority than the lower half. It should be noted that this list is not exhaustive, but here I aimed to cover those research needs that would likely affect model outcomes most or reduce model uncertainty most, based on the various sensitivity analysis that have been performed.

- More data on oocyst excretion rates, especially of humans, and to a lesser extent of animals. Oocysts excretion is the most sensitive variable in the model, as it can easily vary over orders of magnitude. More data will not reduce this inherent variability, but for the purpose of going towards risk assessment it is very important to at least establish most likely orders of magnitude.
- More data on where human faeces end up, e.g. the use of different sanitation systems (as the JMP already collects) but also on subsequent waste management and wastewater treatment efficiency. An estimate on how much of the faeces of the population using on-site systems (septic tanks and pit latrines) reaches rivers should be included in GloWPa-Crypto. It is very difficult to estimate the occurrence of leakages and the proportion of safe faecal management, especially of on-site systems, as there is a lack of data (Berendes et al., 2017; WHO/UNICEF JMP, 2017). For this reason it has been ignored so far in the model under the assumption that residence time in the system is sufficiently long for most of the organisms to have decayed.
- Monitoring data of pathogens in surface waters that could be used for model validation and calibration, preferably with sufficient sensitivity to limit censored data and including determination of the recovery efficiency. Both newly collected data and compilation of existing data sets would be valuable. Data from developing countries are especially scarce, potentially interesting regions to focus monitoring on have been identified in section 8.2.2.1.
- More data on when and where animal manure ends up, e.g. manure spreading and grazing regimes, birthing seasons, geographical and age-related differences in the prevalence of *Cryptosporidium* infection, the use of manure storage and treatment systems and the decay rates in different systems and on fields.
- Field research should be done on the transport of oocysts from manure on land to rivers, hardly any field data are available on this.
- Research on the interactions between water flow velocity and oocyst losses from decay and sedimentation and possible resuspension from sediments. For

watershed scale faecal indicator models resuspension is also a challenging aspect to model, but evidence exists that it may be relevant (Cho et al., 2016).

- Linking to the former point, GloWPa-Crypto should be updated to include dams, lakes and reservoirs, in order to more accurately predict water residence times. For nutrients (N and P) it is estimated that reservoirs are responsible for ~23% of nutrient retention in rivers (Beusen et al., 2015a).
- Model performance in regions with extreme low flow conditions, both of the hydrological model and GloWPa, is currently poor and could be a focus of improvement. This problem is also recognized for watershed scale faecal indicator models (Cho et al., 2016). For this reason, we have excluded grid cells with a mean discharge of less than 1 $m^3 s^{-1}$ from the analysis, as is also done for hydrological studies, but this cut-off is arbitrary.

It should be noted that several model updates would be possible in the near future (i.e. including a first estimate of how much faeces from on-site systems reaches rivers, including dams, lakes and reservoirs, when the underlying VIC hydrology has been updated, and perhaps also a systematic literature review on human oocyst excretion would yield sufficient data to update the model for this). However, the other points involve extensive primary data collection or compilation of existing data that are not published in the scientific literature. GloWPa-Crypto could only be improved in case this kind of data would become available. Apart from research to improve the model in its current form, as listed here, it would also be possible to elaborate the model for additional new applications. This will be discussed in section 8.3.

8.3 Results in context

This thesis presents the first global model of waterborne pathogen concentrations in rivers, and as such it presents many new opportunities. The previous section reflected on the methods applied in this thesis and in this section I will discuss model results in context. First, modelling opportunities and model applicability are discussed (8.3.1). Then, the results are discussed in the context of several important issues that were touched upon in Chapter 1, namely global change and scenario analysis (8.3.2), risk assessment and burden of disease calculations (8.3.3) and the Sustainable Development Goals (8.3.4).

8.3.1 Modelling opportunities and model applicability

A global model of waterborne pathogens in rivers provides many opportunities, as has been indicated in various chapters in this thesis; it can provide information on pathogen pollution levels in data-sparse regions such as developing countries (Chapters 3 ,4, 6, 7), it

can identify hotspots of high loads and concentrations (Chapters 4, 6, 7), it can be a tool that enables scenario analysis, e.g. to study effects of global change (in Chapter 4 done for population, urbanisation and sanitation changes) or management options (in Chapter 6 done for livestock manure management), it can be a basis to inform studies of risk and disease burden (section 8.3.3), and it could be applied to help policy makers decide on the most efficient control strategies and follow progress (for example towards the Sustainable Development Goals, see section 8.3.4).

Furthermore, an important opportunity is to expand GloWPa-Crypto to other waterborne pathogens, antimicrobial-resistant micro-organisms and possibly also faecal indicator organisms and source tracking markers. A first application of the GloWPa human emissions model has been made for the pathogen rotavirus (Kiulia et al., 2015). Expanding to faecal indicator organisms would enhance the possibilities for validation, as much more data are available for indicators than for pathogens. However, this should be done with caution, as faecal indicators and pathogens do not necessarily correlate well, and some indicators have other environmental sources or experience environmental regrowth (Committee on Indicators for Waterborne Pathogens National Research Council, 2004; Lin and Ganesh, 2013; Payment and Locas, 2011). Expanding to source-tracking markers would help determine if the model is correctly representing different pollution sources. Another opportunity would be to incorporate GloWPa in multi-pollutant modelling (Kroeze et al., 2016), to do combined modelling of multiple water quality variables at once to enable a more complete understanding of water pollution and select control options that are most efficient to control the multitude of pollutants. Furthermore, developing GloWPa-Crypto from a deterministic into a stochastic model would be relevant for risk analysis (see section 8.3.3).

It is important to also keep in mind the limits of the applicability of global scale models. GloWPa-Crypto currently operates on a 0.5x0.5 degree grid and a monthly time step. This means the model is relatively coarse and it should not be used to give exact numbers for specific locations, but rather to see ‘the big picture’ and trends and for above-mentioned applications. GloWPa-Crypto does not capture peak concentrations, as it is designed to calculate average concentrations. Although application of GloWPa-Crypto in risk analysis and calculation of disease burden could be valuable (section 8.3.3), this model does not replace epidemiological studies of disease burden. However, it can serve as an addition that generates a better understanding of the sources and pathways that lead to this disease burden, and hence serve better to design effective control. Also, epidemiological studies on the prevalence of diseases in populations are needed to target large scale models on the most significant diseases/pathogens. Furthermore, GloWPa-Crypto complements both modelling at smaller scales and monitoring for pathogens. Modelling at, for instance, the watershed scale is highly relevant and necessary to study relevant

processes in detail and inform local water management (Cho et al., 2016; Oliver et al., 2016). Similarly, GloWPa-Crypto does not replace the need for observational data. In fact, it highlights the need for monitoring and we can use it to indicate interesting areas to conduct monitoring (section 8.2.2.1).

8.3.2 Global change and scenario analysis

Global change processes will likely affect future microbial water quality in multiple ways, as was mentioned in Chapter 1 in this thesis. Both physical (e.g. climatic) changes and socioeconomic development (e.g. changes in population, sanitation, water use) will have an effect on faecal pathogens in rivers (Hofstra, 2011; Rose et al., 2001).

This thesis has presented examples of how scenario analysis can be used to investigate effects of future developments. In Chapter 4 this was done to study effects of socioeconomic development (population, urbanisation and sanitation changes) using the Shared Socioeconomic Pathway (SSP) scenarios. In Chapter 6 this was done for management of livestock manure. The SSPs are socioeconomic development storylines on how the future may unfold, that have been developed for the global change community (O'Neill et al., 2017). Their counterparts are the Representative Concentration Pathways (RCPs), that describe potential future radiative forcing (from 2.6 to 8.5 W m^{-2}) and the greenhouse gas emissions leading to these, and consequent temperature and precipitation changes are calculated subsequently (van Vuuren et al., 2011). The SSPs and RCPs can be used together in a matrix structure to study future global change processes and impacts. For microbial pollution this has for the first time been done in studies at the watershed scale for Pakistan and Bangladesh (Iqbal, 2017; Islam, 2017). It will be highly interesting to perform a similar full scenario analysis with GloWPa-Crypto using a combination of RCPs and SSPs, to investigate the impacts of socioeconomic development and climatic changes together on pathogen concentrations in rivers. Doing so was out of the scope of this thesis, but it will be a next step.

For now, looking at the results of the modelling and the sensitivity analysis, my hypothesis is that we will find that, overall, socioeconomic development (especially sanitation and treatment) will matter more than climatic changes for average concentrations of enteric waterborne pathogens. Population growth in combination with connecting people to sewer systems (especially when the installation of adequate treatment lags behind) will likely increase concentrations much more than a few degrees of warming will affect decay and changes in river flow will affect dilution, a finding that is supported by literature (Islam, 2017; Sterk et al., 2016). Nevertheless, this will of course differ regionally and for different scenarios. For instance, Iqbal et al. (2017) found for the Pakistani Kabul river that under different future scenarios either point sources or nonpoint sources could become

relatively more important for microbial river pollution, and the magnitude of nonpoint sources is dependent on rainfall and runoff that are affected by climate change. Nevertheless, the fact that GloWPa-Crypto is very sensitive to the human oocyst excretion rates and sanitation changes supports the hypothesis that, overall, socioeconomic development will matter most (Chapters 3 and 4). Other models of pathogens or faecal indicators have similarly observed sensitivity to pathogen loading from humans (Coffey et al., 2010; Ferguson et al., 2007; Reder et al., 2017). Furthermore, hotspot regions were identified to be densely populated areas in developing countries, regions that also experience rapid changes in population and sanitation use (Chapters 3 and 4), another reason why socioeconomic developments are important. Developments in public health will also matter; for example, in case a vaccine would be developed for cryptosporidiosis and this vaccine would be widely administered, as is currently ongoing for rotavirus (World Health Organization, 2013), the prevalence and excretion rates will diminish. Consequently, emissions to surface water would also be expected to go down. However, although development is ongoing, there is no indication a vaccine for cryptosporidiosis will become available in the near future (Shirley et al., 2012).

It should be kept in mind that GloWPa-Crypto predicts average monthly concentrations, and both the frequency of occurrence and magnitude of peak concentrations might very well be more affected by climatic changes (Schijven and de Roda Husman, 2005), as they are influenced by the occurrence and magnitude of rainfall-runoff events (causing runoff of manure from land, and potential sewage and latrine overflows). In addition, peak concentrations are likely to strongly influence disease risk (Schijven and de Roda Husman, 2005) (see section 8.3.3) so this will certainly be an interesting theme to investigate. Continuing on the water use side: increased standard of living, better drinking water treatment and improved hygiene might increase population resilience and decrease the prevalence of cryptosporidiosis and other waterborne diseases. It is difficult to say which changes will have a larger impact, and in addition this will likely differ spatially, e.g. between world regions but also between rich and poor regions within individual countries. Scenario analysis is the tool that will enable us to investigate all of this quantitatively and pinpoint regions that are likely to be most at risk in the future.

8.3.3 Risk and burden of disease

One of the most interesting applications of GloWPa-Crypto is to use the concentration output as input for estimates of infection risk and burden of disease. Quantitative Microbial Risk Assessment (QMRA), which has been introduced in section 1.1.2, is the tool that enables making this link. In this section I present first examples of an application of QMRA for waterborne pathogens at the global scale, which has never been done before to my knowledge. This analysis is partly based on a first exploration performed in the MSc

thesis of Dennis de Raaij at Wageningen University, who calculated risk of infection and disease burden for January and July. I elaborated this to all other months and annual risk. First, I calculate risk of infection and disease burden for the case of drinking raw (untreated) surface water. Next, I produce maps of treatment targets for drinking water produced from surface water in order to comply with the WHO health target (World Health Organization, 2017a).

8.3.3.1 Dose

QMRA calculates risk of infection from the exposure (i.e. number of ingested oocysts) and the dose-response (i.e. the probability that a certain dose will cause an infection). Risk of infection is usually expressed as a probability (a number between 0-1), or a probability distribution for a population. The exposure, in the case of transmission via drinking water, is the oocyst concentration in the water that is drunk times the ingested water volume. For this first risk calculation, I look at the population drinking surface water directly, so without drinking water treatment. I assume that this population drinks 2 L of surface water per day. Using GloWPa-Crypto C1 monthly average concentration grids as input, the daily dose in a given month can be calculated as:

$$D_{d,i} = C_i \times V_d \tag{8.1}$$

Where $D_{d,i}$ is the daily oocyst dose in month i, C_i is the oocyst concentration in month i, and V_d is the daily ingested volume of water (here set to 2 L).

For the purpose of this risk calculation, GloWPa-Crypto C1 was adjusted to produce concentration maps that account for the proportion of *Cryptosporidium* spp. that are infectious for humans. The majority of human infections are caused by *C. hominis* (which infects predominantly humans) and *C. parvum* (which infects a variety of mammals, especially livestock is an important reservoir) (Ryan et al., 2014). *C. parvum* is mainly excreted by ruminants, especially calves, and to a lesser extent by adult cattle, lambs and goat kids (see Chapters 5 and 6) (Castro-Hermida et al., 2011; Fayer et al., 2007; Follet et al., 2011; Geurden et al., 2008; Helmy et al., 2013; Khan et al., 2010; Koinari et al., 2014; Kváč et al., 2006; Mi et al., 2014; Robertson et al., 2014; Ryan et al., 2014; Wang et al., 2014; Yang et al., 2009). For these first risk maps, 50% of oocysts excreted by calves (of cattle and buffaloes), 5% of oocysts excreted by adult cattle and buffaloes, and 20% of oocysts excreted by goat kids and lambs are assumed to be infectious for humans. All other oocysts excreted by livestock are assumed not to infect humans. GloWPa-Crypto-C1 was run with these adjustments to produce the concentration grids that were used to calculate the daily dose.

8.3.3.2 Risk

To go from ingested dose to risk, dose-response models are used. Dose-response relationships are mathematical models describing the probability of being infected given ingestion of a certain dose. These have been experimentally determined for most pathogenic organisms (by means of feeding studies with volunteers), and they are typically non-linear (Nilsen and Wyller, 2016). For *Cryptosporidium* several dose-response models exist, including exponential, hypergeometric and Beta-Poisson (Smeets et al., 2007; Teunis et al., 2002a, 2002b). The Beta-Poisson model was chosen for this first calculation, as an illustrative example.

The Beta-Poisson dose response model calculates the probability of infection per day in a given month as:

$$P_{d,i} = 1 - \left(1 + \frac{D_{d,i}}{\beta}\right)^{-\alpha} \quad (8.2)$$

Where $P_{d,i}$ is the probability of infection per day in month i, $D_{d,i}$ is the ingested dose per day in month i, β is set to 0.176 and α is set to 0.115 (Teunis et al., 2002b).

The monthly probability of infection in month i ($P_{m,i}$) is then calculated as:

$$P_{m,i} = 1 - \left(1 - P_{d,i}\right)^{30.42} \quad (8.3)$$

Where 30.42 is the average number of days in a month. The annual probability of infection P_y is then calculated as:

$$P_y = 1 - \left((1 - P_{m1}) \times (1 - P_{m2}) \times \ldots \times (1 - P_{m12})\right) \quad (8.4)$$

This QMRA model based on GloWPa-Crypto predicts that, in most world regions, the annual risk of getting infected with *Cryptosporidium* when drinking 2 L per day of raw surface water is 1, meaning that everyone would get infected at least once per year (Figure 8.3). The figure legend was made on a logarithmic scale to better visualize regions with lower risk, meaning that a risk of 1 is shown as 0 on this scale.

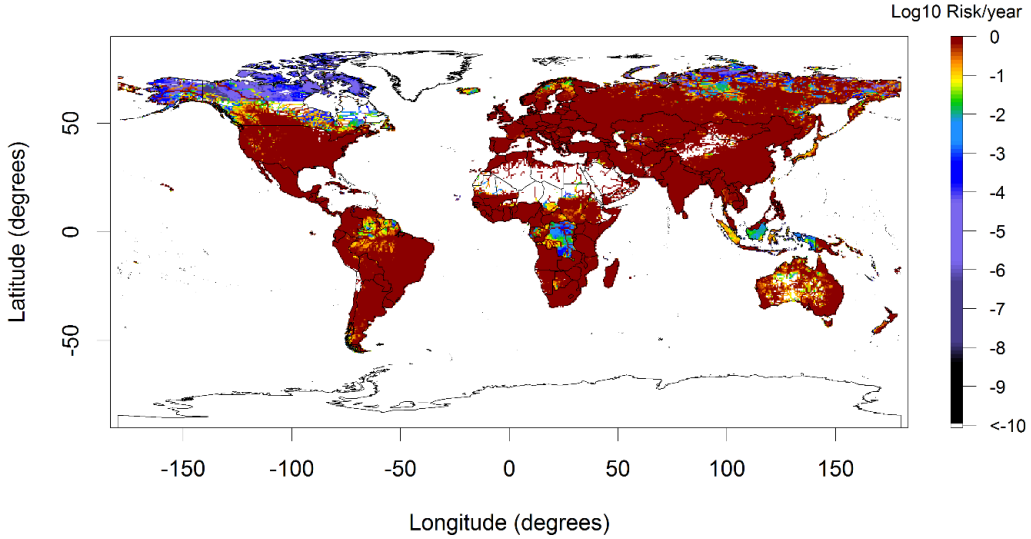


Figure 8.3: Theoretical annual risk of infection (P_y) when drinking raw surface water (2 L day^{-1} all year round).

8.3.3.3 Disease burden

Risk of infection can be translated into an estimate of disease burden, which is typically expressed in disability-adjusted life years (DALYs) per 1000 cases of the disease. First, not everyone who is infected with *Cryptosporidium* becomes ill. Therefore, the risk of infection has to be multiplied with the probability of infection leading to illness, for cryptosporidiosis commonly a value of 0.7 is applied (World Health Organization, 2017a). DALYs for a disease or other health condition are calculated as sum of the years of life lost (YLL) for the cases the disease causes death, plus years lost due to disability (YLD) for the period a person is affected by the disease. DALYs take into account a weighting for the severity of a disease. For cryptosporidiosis, a value of 1.47 DALYs per 1000 cases, or 0.00147 DALY per case, is a commonly used estimate of disease burden (Havelaar and Melse, 2003). It should be noted that this value is based on limited data from the Netherlands and the US and is therefore uncertain. Havelaar and Melse (2003) note that it is likely to be much higher in developing countries, but they do not provide an estimate of how much higher. This value is likely higher in developing countries for a large part due to the higher prevalence of HIV/AIDS, as the disease is more severe in immunocompromised people (Xiao et al., 2012). For now, I use 1.47 DALYs per 1000 cases for the entire world. The disease burden per year (B_y) can be calculated as:

$$B_y = P_y \times P_{ill} \times B_{case} \times Pop \tag{8.5}$$

Where B_y is the disease burden per year per grid cell, P_y is the probability of infection per year (eq. 8.4), P_{ill} is the probability infection causes illness (set to 0.7), B_{case} is the disease burden per case (set to 0.00147), and Pop is the human population per grid cell.

Equation 8.5 calculates the disease burden assuming everyone drinks raw (untreated) surface water, which is of course not the case. The WHO/UNICEF Joint Monitoring Program (JMP) estimates that 2.16% of the world population relied on raw surface water as their primary source of drinking water in 2015, or approximately 158 million people (WHO/ UNICEF Joint Monitoring Program, 2017). The ten countries with the highest proportion of the population drinking surface water in 2015 were Papua New Guinea (42%), Angola (24%), Kenya (23%), Eritrea (21%), Tajikistan (18%), Liberia (17%), Sierra Leone (16%), Madagascar (16%), Afghanistan (15%) and Swaziland (15%) (WHO/ UNICEF Joint Monitoring Program, 2017). Most people drinking raw surface water live in rural areas. By summing over all grid cells in a country and using data on the fraction of the population relying on raw surface water as downloaded from the JMP WASH database (WHO/ UNICEF Joint Monitoring Program, 2017), I can generate a country scale map of the expected disease burden from cryptosporidiosis from drinking raw surface water. This is calculated as:

$$B_{y,j} = \frac{\sum_j (B_y) \times f_{sw,j} \times 100{,}000}{Pop_j} \quad (8.6)$$

Where $B_{y,j}$ is the disease burden from drinking raw surface water per year per 100,000 population in country j, $f_{sw,j}$ is the fraction of the population drinking raw surface water in country j, and Pop_j is the human population in country j.

Figure 8.4 shows the result of this disease burden calculation for cryptosporidiosis contracted from drinking raw surface water on a country map, expressed in DALYs per 100,000 population per year. Countries that appear white on the map have no population depending on surface water for their drinking water, according to the JMP data. The disease burden is calculated to be largest in Papua New Guinea, sub-Saharan Africa, and to a lesser extent in Asia and Latin America. The values that I find appear to be in a plausible range; for example, for Brazil and China, respectively 0.52 and 0.83 DALYs per 100,000 people were found for drinking water transmission (Razzolini et al., 2016; Xiao et al., 2012). The WHO has estimated median (with 95% uncertainty interval) foodborne disease burden from cryptosporidiosis to be 13 (3-37) DALYs per 100,000 people for Africa, 0.6 (0.2-2) for the Americas, and 4 (0.4-20) for the Eastern Mediterranean region (Kirk et al., 2015). Unfortunately, waterborne transmission was not included in this study, but at least the relative order of magnitude for the different regions is comparable to my map. When comparing my map to the WHO Public Health and Environment (PHE) map of 'water,

sanitation and hygiene attributable burden of disease (low- and middle-income countries), 2012' (http://gamapserver.who.int/gho/interactive_charts/phe/wsh_mbd/atlas.html, accessed on 17-10-2017), the spatial patterns are similar, although East Africa stands out more in my map and West and Central Africa are more prominent on the WHO map. However, my map is only related to drinking raw surface water and East Africa has some countries with relatively high proportions of population drinking raw surface water. The WHO map, on the contrary, is all WASH-related disease, meaning that it combines different pathogens and exposure routes. Values on the WHO map are therefore also much higher, up to 5254 DALYs per 100,000 people per year for Angola, which gets a value of 19 on my map. The values on these maps can thus not directly be compared.

8.3.3.4 Treatment targets

The World Health Organization (WHO) has set the 'tolerable burden of disease' or, in other words, the target health outcome at 10^{-6} DALYs per person per year for waterborne exposure (World Health Organization, 2017a). Starting from this target, it is possible to back-calculate (using equations 8.1-8.5) to determine a tolerable oocyst concentration for drinking water and determine a drinking water treatment target to achieve this concentration. Calculating health-based treatment targets is a practice supported by the WHO (World Health Organization, 2017a), and it can help in planning investments in water infrastructure (Howard et al., 2006).

The tolerable infection risk can be found by dividing the health target 10^{-6} by the standard cryptosporidiosis DALYs of 0.00147 per ill person (Teunis et al., 2002b), yielding a tolerable illness risk of 6.80×10^{-4} per person per year. Dividing by 0.7 (probability of illness when infected) yields a tolerable infection risk of 9.72×10^{-4} per person per year. Continuing, I find a tolerable daily risk (daily probability of infection) of 2.66×10^{-6} and a tolerable daily dose of 4.07×10^{-6} oocysts. The tolerable average concentration when drinking 2 L per day is then 2.04×10^{-6} oocysts L^{-1}. The WHO guidelines for drinking water quality presents an example calculation that is very similar in Table 7.4 (World Health Organization, 2017a). However, as they use an exponential dose-response relationship where I used Beta-Poisson, and they assume 1 L per day water intake instead of 2 L, their resulting tolerable concentration in drinking water is somewhat higher at 1.3×10^{-5} oocysts L^{-1}. However, as I am looking at a population that is dependent on raw surface water for their drinking water and these people likely do not drink any other liquids, I consider 2 L more appropriate.

I compare this tolerable average concentration of 2.04×10^{-6} oocysts L^{-1} to a GloWPa-Crypto C1 concentration map to generate a map of what the level of treatment for drinking water production from surface water would need to be to meet the health target (Figure 8.5). This is the first time an analysis of this kind has been done using a global scale waterborne pathogen model as input, to my knowledge. This analysis should be seen as a

first, illustrative example of what would be possible.

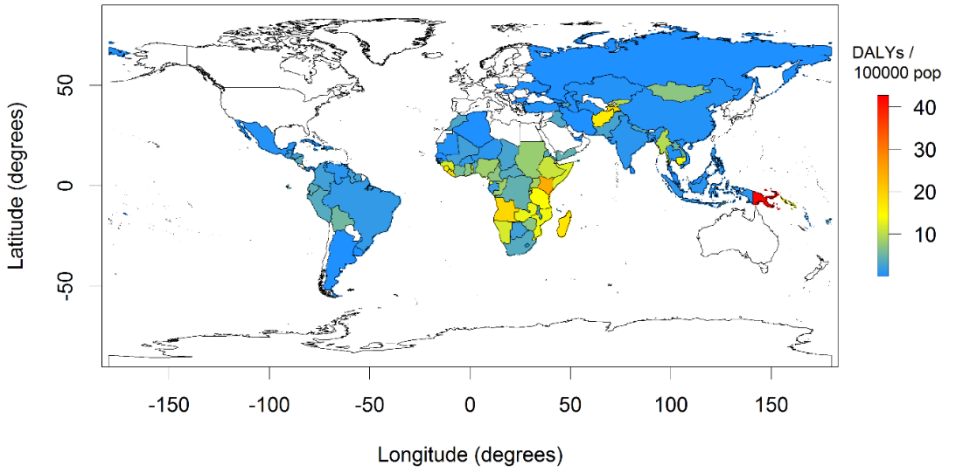


Figure 8.4: Disease burden (expressed in DALYs per 100,000 population per year ($B_{y,j}$)) for cryptosporidiosis contracted from drinking raw surface water (2 L day^{-1} all year). Countries that appear white on the map have no population depending on raw surface water for their drinking water, according to JMP data (WHO/ UNICEF Joint Monitoring Program, 2017).

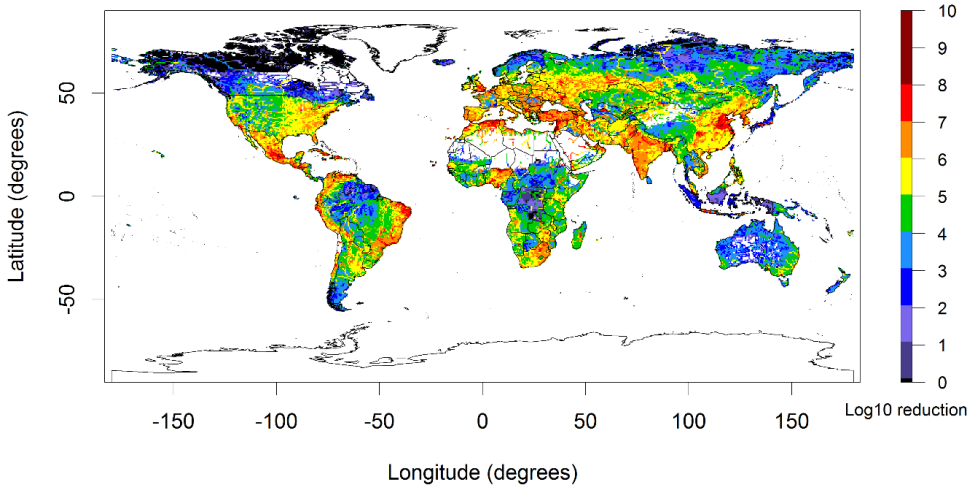


Figure 8.5: Log reduction in oocyst concentrations required to meet the health-based target of 10^{-6} DALYs per person per year, assuming everyone drinks 2 L per day of drinking water produced from surface water all year round.

To reach the WHO health-based target of 10^{-6} DALYs per person per year for the consumption of drinking water produced from surface water, drinking water treatment reaching 3-7 log removal of oocysts is needed for most world regions, and up to 9.4 log for

the most polluted regions (Figure 8.5). Values over 9 log are only found in 4 grid cells, around the Nile delta and Mexico City. Values between 8-9 logs are found in various densely populated urban areas in developing countries. Only some Arctic regions would require no treatment at all (Figure 8.5). If the exponential dose-response model and 1 L daily intake assumption would have been used, as the WHO did (World Health Organization, 2017a), the required reductions would be approximately one log unit lower. Typical oocyst removal efficiencies of drinking water treatment plants are not as high as the required reductions I calculate, more in the order of 2-5 log removal (Betancourt and Rose, 2004). As reference, this would indicate that in the most contaminated world regions oocysts concentrations in rivers are basically comparable to those found in raw sewage (see Chapter 7) (Nasser, 2016; World Health Organization, 2017b), which is not unthinkable, especially during periods of low river flow.

8.3.3.5 Discussion

The risk analysis done here should be seen as a very first exploration of QMRA using a global waterborne pathogen concentrations model, and it should be kept in mind that many issues need to be addressed in a future, more comprehensive QMRA based on GloWPa-Crypto. First, this risk calculation is based on consuming water with a certain average oocyst concentration, while it is suggested that actual risk is driven more by peak concentrations (Schijven and de Roda Husman, 2005), and thus by the frequency of occurrence and magnitude of peak concentrations. Nevertheless, average (arithmetic mean) concentrations are a better indicator for risk than, for instance, median or geometric mean concentrations, as the arithmetic mean is relatively sensitive to peaks (Haas, 1996). Secondly, for the purpose of this first risk calculation we have only looked at one single exposure route and ingested volume: drinking 2 L of water produced from surface water. A full QMRA should account for variation in ingested volumes in a population and combine multiple possible exposure routes, e.g. also taking into account drinking water produced from other sources, different drinking water consumption patterns, exposure from bathing, vegetables irrigated with contaminated water, etcetera. This would yield risk distributions instead of single values per grid cell. In this context should also be noted that GloWPa-Crypto currently focuses on river water while the consumption data consider all surface water sources. The conclusions probably do not directly apply to drinking water production from lakes or reservoirs, as sedimentation likely causes concentrations to be lower there (Chapter 7). Thirdly, I applied the Beta-Poisson dose-response model, applying another dose-response model gives a different result, as was seen in the comparison with the WHO calculation for health-based treatment targets. A full QMRA could include multiple dose-response models for comparison. Furthermore, the dose-response for sensitive populations (children, immunocompromised people) is likely to differ from the dose-response model I used, that

is based on feeding studies of healthy volunteers, and the DALY estimate of 1.47 is also based on a population with little HIV/AIDS infections (Havelaar and Melse, 2003; Teunis et al., 2002a, 2002b; Xiao et al., 2012). For sensitive populations the treatment target should likely be more stringent to achieve the target health outcome. Lastly, there is some evidence that immunity development against cryptosporidiosis occurs if the same person is infected multiple times, although frequent reinfections in children are also observed (Checkley et al., 2015; Teunis et al., 2002a). The potential protective effect of immunity is ignored in the current calculation, as is commonly done in QMRA.

It would be interesting to develop GloWPa-Crypto into a stochastic model, especially with the potential importance of peak concentrations in mind and their likely sensitivity to climate change (section 8.3.2). This would entail that for all model inputs probability distributions are estimated, instead of single values. The model would then be run many times using Monte Carlo sampling, and its output would be a concentration probability distribution for each grid cell. This approach has also been recognized and applied for catchment scale microbial modelling (Cho et al., 2016; Starkey et al., 2007). These concentration distributions can then be used as input for QMRA. Developing GloWPa-Crypto into a stochastic model was out of the scope of this thesis, but a first step in this direction was made by estimating the low end and high end ranges of model variables, as estimated for the sensitivity analysis in Chapter 7. A next step could be to take these high and low end estimates of the different variables and include these as input for risk assessment, to get an idea of potential variability. A downside of developing a stochastic model would be the much higher computational cost. Therefore, a careful judgment should be made whether it is necessary for answering particular research questions, or whether it makes the model more complex but not better. With current developments in computing power, this will likely become less of an issue in the future.

8.3.4 Meeting the Sustainable Development Goals

In Chapter 1 the Sustainable Development Goals (SDGs) were introduced. These are 17 goals subdivided in targets that the world aims to meet in 2030 (United Nations General Assembly, 2015). Especially Goal 6, but also other goals and targets contain elements that link to water quality (listed in Chapter 1 Box 1). On the one hand, water quality modelling can benefit from the data collected to assess SDG baselines and progress by using these as model input. On the other hand, there are several ways in which large scale water quality modelling could help meeting the SDGs; global water quality modelling can 1) complement collected monitoring data and thereby help assess progress towards SDG indicator 6.3.2 ‘Proportion of bodies of water with good ambient water quality’, 2) be a tool to calculate water quality impacts of meeting the SDGs, and 3) help formulate effective policy options that account for feedbacks in the system and trade-offs and

synergies that follow from these feedbacks. These four points will be explained in some more detail in this section.

First of all, water quality modelling benefits from the data collected to assess baselines and progress towards the SDGs, because it is valuable model input. For example, the Joint Monitoring Programme (that tracks the progress towards meeting the water, sanitation and hygiene (WASH) related SDG targets around the world) provides the GloWPa-Crypto model essential input data on the use of different sanitation systems around the world. JMP 'ladders' show the classes in which service levels are categorized. In 2017, updated JMP ladders were presented (WHO/UNICEF JMP, 2017); for sanitation these are now Open defecation, Unimproved, Limited, Basic and Safely managed. Previously, the ladder consisted of Open Defecation, Unimproved, Shared and Improved. This update is important, as the previous ladder did not explicitly include safe disposal or treatment of human waste. Safely managed is defined as 'Use of improved facilities that are not shared with other households and where excreta are safely disposed of in situ or transported and treated offsite'. The GloWPa model currently bases estimates of wastewater treatment on Van Drecht et al. (2009), and it would be worthwhile to look into updating this part of the model to be consistent with the newer JMP data. Looking into on-site sanitation systems is also relevant, as these are often poorly managed in developing regions, as the costs and responsibilities are often borne directly by the users themselves (Berendes et al., 2017). More JMP data on the safe and unsafe management of faecal waste from on-site systems would therefore be valuable input for the GloWPa-Crypto model. Furthermore, for water quality variables that are monitored by GEMS/Water, the collected monitoring data will be a valuable resource for model calibration and validation.

Then, as was suggested in Chapter 1, large scale water quality modelling could be key in assessing progress toward SDG indicator 6.3.2. The currently available monitoring data alone will likely not tell the whole story of the status of this indicator, for a number of reasons: 1) the coverage of observational data is limited, as supplying data to GEMS/Water is done on a voluntary basis and monitoring capacity is limited in many countries, 2) the SDG baseline set of water quality variables for which monitoring is required is limited, and does not contain any microbial water quality variable, 3) reporting is organized on a country basis, while many watersheds are transboundary systems, so good international cooperation is needed to accurately judge water quality in these systems and 4) observational data do not necessarily provide insight into pollution sources. Facing these limitations, it is apparent that water quality modelling could be a helpful complementary source of information. Specifically, modelled water quality could 1) complement the monitoring data of the baseline set to countries and regions where no measurements are taken, i.e. provide better spatial coverage, 2) expand to variables not included in the baseline set, 3) provide insight into transboundary pollution by allowing

analysis by watershed, and 4) provide insight in the relative magnitude of different sources of pollution, and thus give an indication of what policies would be effective to reduce pollution.

Next, modelling can be a tool to calculate impacts of meeting the SDGs. Modelling can be used to answer questions such as 'What will be the effect on water quality if open defecation is ended (target 6.2) while the population grows according to SSP3 and with different assumptions on wastewater treatment?'. Performing these kinds of model exercises can be useful to evaluate the targets and assess if they are stringent enough to have the desired impact. An example of this kind of analysis is given in Chapter 4 of this thesis, where the impact of connecting people to sewers without improving wastewater treatment is calculated. We conclude here that connecting the population to sewers without installing treatment will deteriorate river water quality. Another example would be to use the model to calculate wastewater treatment targets, as was done for drinking water treatment in section 8.3.4. SDG indicator 6.3.1 is the 'Proportion of wastewater safely treated'. GloWPa-Crypto could be applied to calculate what level of treatment would be considered 'safely treated', to reach a certain surface water quality target in different world regions.

The final point is to bring these type of calculations even further. Effective policies for sustainable development account for feedbacks within systems, and assess trade-offs and synergies of policy options that follow from these feedbacks. Global scale models can be part of the suitable tools to help evaluate policy options, as they are constructed to describe a system and can make feedbacks explicit. Trade-offs and synergies exist because links between the SDGs and water quality can go in different directions, i.e. efforts to meet the SDGs can improve or deteriorate water quality, and changes in water quality can affect progress towards certain targets. For example, if open defecation is ended (SDG target 6.2) and all people are connected to sewage system transporting the waste to rivers but only half of the wastewater is treated (6.3.1), what will be the effect on ambient water quality (6.3.2)? And what will be the resulting effect on waterborne disease occurrence (3.9) and food security (2.1)? Does this change when assuming socioeconomic development and climate change given certain SSP-RCP combinations? Can we design policies that maximize synergies and minimize trade-offs? Large scale water quality models, for example applied in scenario analysis, would be excellent tools for such explorations. This was also one of the conclusions of a the workshop "Water Quality: a new challenge for global scale modelling" that took place in Wageningen, 18-21 September 2017 (http://www.wur.nl/en/Expertise-Services/Chair-groups/Environmental-Sciences/Environmental-Systems-Analysis-Group/Workshop.htm, accessed on 23-10-2017).

8.4 Conclusions

This thesis presents the first global model of waterborne pathogen emissions to and concentrations in rivers, GloWPa-Crypto. The objective of this thesis was to increase our knowledge on the sources, fate and transport of *Cryptosporidium* in rivers worldwide using spatially explicit modelling, and this objective was addressed in six scientific papers (Chapters 2-7). Here I summarize the conclusions that can be drawn from this thesis.

Methodology

GloWPa-Crypto has been developed largely following the conceptual model presented in Chapter 2, which is based on a review of existing models of global nutrient export and watershed-scale pathogen and faecal indicator models. GloWPa-Crypto has been developed as a deterministic process-based model that contains some empirical parts. It is uncalibrated, static, spatially explicit, implemented at the global scale, with a spatial resolution of 0.5 x 0.5 degree grid, and a monthly time step. These choices were made with the model objectives in mind and based on pragmatic reasons, for instance regarding data availability. Validation with observational data, comparison with other modelling studies and sensitivity analysis have been applied to enhance confidence in model performance. Observational data are needed to inform and validate models, and modelling can fill the gaps and highlight where observational data are most needed.

Human *Cryptosporidium* emissions

We found that urbanised areas in developing countries are hotspots of *Cryptosporidium* emissions, this was observed in a case study for Bangladesh and India (Chapter 3), as well as in the global model for human *Cryptosporidium* emissions GloWPa-Crypto H1 (Chapter 4). Future population growth and urbanisation are likely to lead to further deterioration of water quality, we conclude from our scenario analysis using the SSP scenarios for 2050. If population, urbanisation and sanitation change according to SSP1 in 2050, total global *Cryptosporidium* emissions to surface water are expected to reduce by 24%. However, in the Middle East and Africa region, India and Pakistan emissions are expected to increase for SSP1 due to a combination of population growth and urbanisation. For SSP3 the total global emissions in 2050 are expected to increase by 52%, only rich OECD countries see reduced emissions due to population decreases. Sanitation and treatment improvements are key to improve water quality now and in the future. If population, urbanisation and sanitation change according to SSP1 but wastewater treatment would not be improved, emissions for nearly all countries are even higher than the emissions calculated for SSP3. We conclude that connecting more people to sewers without improving wastewater treatment will deteriorate river water quality.

Livestock *Cryptosporidium* emissions

Our systematic literature inventory showed that cryptosporidiosis is prevalent in livestock around the world, especially in young animals (Chapter 5). Oocyst concentrations in manure are highly variable, but are also generally higher for young animals than adults. These findings were applied in the global model for livestock *Cryptosporidium* loads to land GloWPa-Crypto L1 (Chapter 6). We calculate that globally, cattle (especially calves) are the dominant livestock source of oocysts, followed by chickens and pigs. The livestock oocyst load from the developing world is larger than from the developed world, and comes less often from stored manure. We find that although manure storage halves oocyst loads, manure treatment, especially of cattle manure and particularly at elevated temperatures, has a larger load reduction potential than manure storage. Regions with high reduction potential include India, Bangladesh, Western Europe, China, several countries in Africa, and New Zealand.

Oocyst concentrations in rivers

Monthly average oocyst concentrations are predicted to range from 10^{-6} to 10^{2} oocysts L^{-1} in most places, as calculated by GloWPa-Crypto C1 (Chapter 7). Hotspot regions with high concentrations include parts of India, China, Pakistan and Bangladesh, Nigeria, Algeria and South Africa, Mexico, Venezuela and some coastal areas of Brazil, several countries in Western and Eastern Europe, and the Middle East. Point sources (human faeces) appear to be a more dominant source of pollution than diffuse sources (animal manure) in most world regions.

Global change

This thesis presents examples of how scenario analysis can be used to investigate effects of future developments. In Chapter 4 this was done to study effects of socioeconomic development (population, urbanisation and sanitation changes) using the Shared Socioeconomic Pathway (SSP) scenarios. In Chapter 6 this was done for management of livestock manure. Future average oocyst concentrations in rivers are likely influenced more by changes in population, sanitation and treatment than by climatic changes. Scenario analysis is the tool that will enable us to investigate linkages quantitatively and pinpoint regions that are likely to be most at risk in the future.

Risk and burden of disease

This thesis presents a first exploration of worldwide risk and disease burden of cryptosporidiosis from drinking water produced from surface water, based on modelled oocyst concentrations from GloWPa-Crypto C1. The risk calculation indicates that when drinking 2 L per day of raw surface water, the annual risk of getting cryptosporidiosis is 1 in most world regions, meaning that everyone would get infected at least once per year.

The resulting cryptosporidiosis disease burden, when taking into account the population actually depending on raw surface water for their drinking water, is calculated to be largest in Papua New Guinea, sub-Saharan Africa, and to a lesser extent in Asia and Latin America. To reach the WHO health-based target of 10^{-6} DALYs per person per year for the consumption of drinking water produced from surface water, our model suggests that drinking water treatment reaching 4-7 log removal of oocysts is needed for most world regions, and up to 9.4 log for the most polluted regions.

SDGs

The world faces a tremendous challenge to meet the Sustainable Development Goals and ensure safe water for all. I have shown the contribution that global waterborne pathogen modelling could make in confronting this challenge. Facing the limitations of observational data, it is apparent that water quality modelling could be a helpful complementary source of information, to help assess progress towards SDG indicator 6.3.2 'Proportion of bodies of water with good ambient water quality'. Furthermore, models can be a tool to calculate water quality impacts of meeting the SDGs. Models can help formulate effective policy options that account for feedbacks in the system and trade-offs and synergies that follow from these feedbacks. Looking at it from another side, model development can also benefit from the data generated in assessing a baseline and progress towards SDG indicators.

In summary, waterborne pathogen modelling at the global scale is feasible and provides new opportunities: to provide information on pathogen concentrations in data-sparse regions, identify hotspot regions with high concentrations, identify the relative contribution of different sources, and allow for scenario analysis. Examples of these opportunities have been demonstrated in the research chapters of this thesis. More opportunities have been identified in this synthesis chapter, including application of the model in the analysis of risk, burden of disease and health-based treatment targets. Global water quality modelling could make a valuable contribution in meeting the Sustainable Development Goals.

Literature

AECOM, SANDEC/EAWAG, 2010. A Rapid Assessment of Septage Management in Asia: Policies and Practices in India, Indonesia, Malaysia, the Philippines, Sri Lanka, Thailand, and Vietnam. doi:10.13140/RG.2.1.4868.3602

Agarwal, S., 2011. The state of urban health in India; comparing the poorest quartile to the rest of the urban population in selected states and cities. Environ. Urban. 23, 13–28. doi:10.1177/0956247811398589

Ajjampur, S.S.R., Liakath, F.B., Kannan, A., Rajendran, P., Sarkar, R., Moses, P.D., Simon, A., Agarwal, I., Mathew, A., O'Connor, R., Ward, H., Kang, G., 2010. Multisite study of cryptosporidiosis in children with diarrhea in India. J. Clin. Microbiol. 48, 2075–2081. doi:10.1128/JCM.02509-09

Ajjampur, S.S.R., Sarkar, R., Sankaran, P., Kannan, A., Menon, V.K., Muliyil, J., Ward, H., Kang, G., 2010. Symptomatic and asymptomatic Cryptosporidium infections in children in a semi-urban slum community in southern India. Am. J. Trop. Med. Hyg. 83, 1110–1115. doi:10.4269/ajtmh.2010.09-0644

Alcamo, J., Van Vuuren, D., Cramer, W., 2006. Changes in ecosystem services and their drivers across the scenarios, in: Carpenter, S. (Ed.), Ecosystems and Human Well-Being: Scenarios. Island Press, Washington, D. C., pp. 279–354.

Allen, P.M., Arnold, J.G., Byars, B.W., 1994. Downstream channel geometry for use in planning-level models. JAWRA J. Am. Water Resour. Assoc. 30, 663–671. doi:DOI: 10.1111/j.1752-1688.1994.tb03321.x

Alonso-Fresán, M.U., Vázquez-Chagoyán, J.C., Velázquez-Ordoñez, V., Pescador-Salas, N., Saltijeral-Oaxaca, J., 2009. Sheep management and cryptosporidiosis in central Mexico. Trop. Anim. Health Prod. 41, 431–436. doi:10.1007/s11250-008-9208-4

Amos, H.M., Jacob, D.J., Kocman, D., Horowitz, H.M., Zhang, Y., Dutkiewicz, S., Horvat, M., Corbitt, E.S., Krabbenhoft, D.P., Sunderland, E.M., 2014. Global biogeochemical implications of mercury discharges from rivers and sediment burial. Environ. Sci. Technol. 48, 9514–22. doi:10.1021/es502134t

Angeles, G., Lance, P., Barden-O'Fallon, J., Islam, N., Mahbub, a Q.M., Nazem, N.I., 2009. The 2005 census and mapping of slums in Bangladesh: design, select results and application. Int. J. Health Geogr. 8, 32. doi:10.1186/1476-072X-8-32

Atherholt, T.B., LeChevallier, M.W., Norton, W.D., Rosen, J.S., 1998. Effect of rainfall on Giardia and crypto. J. / Am. Water Work. Assoc. 90, 66–80.

Atwill, E.R., Das Gracas C Pereira, M., Alonso, L.H., Elmi, C., Epperson, W.B., Smith, R., Riggs, W., Carpenter, L. V, Dargatz, D.A., Hoar, B., 2006a. Environmental load of Cryptosporidium parvum oocysts from cattle manure in feedlots from the central and western United States. J. Environ. Qual. 35, 200–206. doi:10.2134/jeq2005.0099

Atwill, E.R., Hoar, B., Das Gracas Cabral Pereira, M., Tate, K.W., Rulofson, F., Nader, G., 2003. Improved quantitative estimates of low environmental loading and sporadic periparturient shedding of Cryptosporidium parvum in adult beef cattle. Appl. Environ. Microbiol. 69, 4604–4610. doi:10.1128/AEM.69.8.4604-4610.2003

Atwill, E.R., Hou, L., Karle, B.M., Harter, T., Tate, K.W., Dahlgren, R.A., 2002. Transport of Cryptosporidium parvum oocysts through vegetated buffer strips and estimated filtration efficiency. Appl. Environ. Microbiol. 68, 5517–5527. doi:10.1128/AEM.68.11.5517

Atwill, E.R., Tate, K.W., Pereira, M.D.G.C., Bartolome, J., Nader, G., 2006b. Efficacy of natural grassland buffers for removal of Cryptosporidium parvum in rangeland runoff. J. Food Prot. 69, 177–184.

Augusiak, J., Van den Brink, P.J., Grimm, V., 2014. Merging validation and evaluation of ecological models to "evaludation": A review of terminology and a practical approach. Ecol. Modell. 280, 117–128. doi:10.1016/j.ecolmodel.2013.11.009

Bangladesh Bureau of Statistics, 2011. Population & Housing Census 2011 (Community report). Dhaka, Bangladesh.

Bao, P.N., Aramaki, T., Hanaki, K., 2013. Assessment of stakeholders' preferences towards sustainable sanitation scenarios. Water Environ. J. 27, 58–70. doi:10.1111/j.1747-6593.2012.00327.x

Baum, R., Luh, J., Bartram, J., 2013. Sanitation: A global estimate of sewerage connections without treatment and the resulting impact on MDG progress. Environ. Sci. Technol. 47, 1994–2000. doi:10.1021/es304284f

Bennett, E.M., Carpenter, S.R., Caraco, N.F., 2001. Human impact on erodable phosphorus and eutrophication: a global perspective. Bioscience 51, 227–234.

Berendes, D.M., Sumner, T.A., Brown, J.M., 2017. Safely managed sanitation for all means fecal sludge management for at least 1.8 billion of the world's poorest people. Environ. Sci. Technol. 51, 3074–3083. doi:10.1021/acs.est.6b06019

Betancourt, W.Q., Rose, J.B., 2004. Drinking water treatment processes for removal of Cryptosporidium and Giardia. Vet. Parasitol. 126, 219–234. doi:10.1016/j.vetpar.2004.09.002

Beusen, A.H.W., Bouwman, A.F., Dürr, H.H., Dekkers, A.L.M., Hartmann, J., 2009. Global patterns of dissolved silica export to the coastal zone: Results from a spatially explicit global model. Global Biogeochem. Cycles 23, n/a-n/a. doi:10.1029/2008GB003281

Beusen, A.H.W., Bouwman, A.F., Van Beek, L.P.H., Mogollón, J.M., Middelburg, J.J., 2015a. Global riverine N and P transport to ocean increased during the twentieth century despite increased retention along the aquatic continuum. Biogeosciences Discuss. 12, 20123–20148. doi:10.5194/bgd-12-20123-2015

Beusen, A.H.W., Van Beek, L.P.H., Bouwman, A.F., Mogollón, J.M., Middelburg, J.J., 2015b. Coupling global models for hydrology and nutrient loading to simulate nitrogen and phosphorus retention in surface water - Description of IMAGE-GNM and analysis of performance. Geosci. Model Dev. 8, 4045–4067. doi:10.5194/gmd-8-4045-2015

Bhattacharya, M.K., Teka, T., Faruque, a S., Fuchs, G.J., 1997. Cryptosporidium infection in children in urban Bangladesh. J. Trop. Pediatr. 43, 282–286. doi:DOI 10.1093/tropej/43.5.282

Bhattarai, R., Kalita, P., Trask, J., Kuhlenschmidt, M.S.S., 2011. Development of a physically-based model for transport of Cryptosporidium parvum in overland flow. Environ. Model. Softw. 26, 1289–1297. doi:10.1016/j.envsoft.2011.05.011

Björkman, C., Lindström, L., Oweson, C., Ahola, H., Troell, K., Axén, C., 2015. Cryptosporidium infections in suckler herd beef calves. Parasitology 142, 1108–1114. doi:10.1017/S0031182015000426

Bomfim, T.C.B., Huber, F., Gomes, R.S., Alves, L.L., 2005. Natural infection by Giardia sp. and Cryptosporidium sp. in dairy goats, associated with possible risk factors of the studied properties. Vet. Parasitol. 134, 9–13. doi:10.1016/j.vetpar.2005.05.067

Bouwman, A.F., Beusen, A.H.W., Billen, G., 2009. Human alteration of the global nitrogen and phosphorus soil balances for the period 1970-2050. Global Biogeochem. Cycles 23, GB0A04. doi:10.1029/2009gb003576

Bouwman, A.F., Van Drecht, G., Knoop, J.M., Beusen, A.H.W., Meinardi, C.R., 2005. Exploring changes in river nitrogen export to the world's oceans. Global Biogeochem. Cycles 19, n/a-n/a. doi:10.1029/2004GB002314

Bouwman, L., Klein Goldewijk, K., Van der Hoek, K.W., Beusen, A.H.W., Van Vuuren, D.P., Willems, J., Rufino, M.C., Stehfest, E., Ru, M.C., Stehfest, E., 2012. Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900–2050 period. Proc. Natl. Acad. Sci. 110, 20882–20887. doi:10.1073/pnas.1206191109

Boyer, D.G., Kuczynska, E., Fayer, R., 2009. Transport, fate, and infectivity of Cryptosporidium parvum oocysts released from manure and leached through macroporous soil. Environ. Geol. 58, 1011–1019. doi:10.1007/s00254-008-1580-x

Boyer, E.W., Howarth, R.W., Galloway, J.N., Dentener, F.J., Green, P.A., Vörösmarty, C.J., 2006. Riverine nitrogen export from the continents to the coasts. Global Biogeochem. Cycles 20. doi:10.1029/2005GB002537

Bradford, S.A., Schijven, J., 2002. Release of Cryptosporidium and Giardia from dairy calf manure: Impact of solution salinity. Environ. Sci. Technol. 36, 3916–3923. doi:10.1021/es025573l

Bright, E.A., Coleman, P.R., Rose, A.N., Urban, M.L., 2011. LandScan 2010. Oak Ridge, TN.

Brookes, J.D., Antenucci, J.P., Hipsey, M.R., Burch, M.D., Ashbolt, N.J., Ferguson, C.M., 2004. Fate and transport of pathogens in lakes and reservoirs. Environ. Int. 30, 741–759. doi:10.1016/j.envint.2003.11.006

Brookes, J.D., Davies, C.M., Hipsey, M.R., Antenucci, J.P., 2006. Association of Cryptosporidium with bovine faecal particles and implications for risk reduction by settling within water supply reservoirs. J. Water Health 4, 87–98. doi:10.2166/wh.2005.065

Bruijnesteijn van Coppenraet, L.E.S., Wallinga, J.A., Ruijs, G.J.H.M., Bruins, M.J., Verweij, J.J., 2009. Parasitological diagnosis combining an internally controlled real-time PCR assay for the detection of four protozoa in stool samples with a testing algorithm for microscopy. Clin. Microbiol. Infect. 15, 869–874. doi:10.1111/j.1469-0691.2009.02894.x

Burnet, J.-B., Penny, C., Ogorzaly, L., Cauchie, H.-M., 2014. Spatial and temporal distribution of Cryptosporidium and Giardia in a drinking water resource: Implications for monitoring and risk assessment. Sci. Total Environ. 472C, 1023–1035. doi:10.1016/j.scitotenv.2013.10.083

Burnet, J.B., Ogorzaly, L., Penny, C., Cauchie, H.M., 2015. Fine-scale spatial heterogeneity in the distribution of waterborne protozoa in a drinking water reservoir. Int. J. Environ. Res. Public Health 12, 11910–11928. doi:10.3390/ijerph120911910

Cacciò, S.M., Widmer, G., 2014. Cryptosporidium : parasite and disease, Cryptosporidium: Parasite and Disease. Springer. doi:10.1007/978-3-7091-1562-6

Cairncross, S., Hunt, C., Boisson, S., Bostoen, K., Curtis, V., Fung, I.C.H., Schmidt, W.P., 2010. Water, sanitation and hygiene for the prevention of diarrhoea. Int. J. Epidemiol. 39, i193–i205. doi:10.1093/ije/dyq035

Caraco, N.F., Cole, J.J., 1999. Human impact on nitrate export: an analysis using major world rivers. Ambio 28, 167–170.

Castro-Hermida, J.A., García-Presedo, I., Almeida, A., González-Warleta, M., Da Costa, J.M.C., Mezo, M., 2011. Cryptosporidium spp. And Giardia duodenalis in two areas of Galicia (NW Spain). Sci. Total Environ. 409, 2451–2459. doi:10.1016/j.scitotenv.2011.03.010

Castro-Hermida, J.A., González-Losada, Y.A., Ares-Mazás, E., 2002. Prevalence of and risk factors involved in the spread of neonatal bovine cryptosporidiosis in Galicia (NW Spain). Vet. Parasitol. 106, 1–10. doi:10.1016/S0304-4017(02)00036-5

Census of India, 2011. Provisional population totals - India - rural-urban distribution. New Delhi, India.

Central Pollution Control Board: Government of India, 2015. Status of STPs [WWW Document]. URL http://www.cpcb.nic.in/statusSTP.php (accessed 7.30.15).

Chalmers, R.M., Campbell, B.M., Crouch, N., Charlett, A., Davies, A.P., 2011. Comparison of diagnostic sensitivity and

specificity of seven Cryptosporidium assays used in the UK. J. Med. Microbiol. 60, 1598–1604. doi:10.1099/jmm.0.034181-0
Chappell, C.L., Okhuysen, P.C., Langer-curry, R., Widmer, G., Akiyoshi, D.E., Tanriverdi, S., Tzipori, S., 2006. Cryptosporidium hominis: experimental challenge of healthy adults 75, 851–857.
Chappell, C.L., Okhuysen, P.C., Sterling, C.R., Dupont, H.L., 1996. Cryptosporidium parvum: Intensity of Infection and Oocyst Excretion Patterns in Healthy Volunteers. J. Infect. Dis. 173, 232–236.
Checkley, W., White, A.C., Jaganath, D., Arrowood, M.J., Chalmers, R.M., Chen, X.M., Fayer, R., Griffiths, J.K., Guerrant, R.L., Hedstrom, L., Huston, C.D., Kotloff, K.L., Kang, G., Mead, J.R., Miller, M., Petri, W.A., Priest, J.W., Roos, D.S., Striepen, B., Thompson, R.C.A., Ward, H.D., Van Voorhis, W.A., Xiao, L., Zhu, G., Houpt, E.R., 2015. A review of the global burden, novel diagnostics, therapeutics, and vaccine targets for cryptosporidium. Lancet Infect. Dis. 15, 85–94. doi:10.1016/S1473-3099(14)70772-8
Cho, K.H., Cha, S.M., Kang, J.-H., Lee, S.W., Park, Y., Kim, J.H.J.-W., Kim, J.H.J.-W., 2010. Meteorological effects on the levels of fecal indicator bacteria in an urban stream: A modeling approach. Water Res. 44, 2189–2202. doi:10.1016/j.watres.2009.12.051
Cho, K.H., Pachepsky, Y.A., Oliver, D.M., Muirhead, R.W., Park, Y., Quilliam, R.S., Shelton, D., 2016. Modeling fate and transport of fecally-derived microorganisms at the watershed scale: State of the science and future opportunities. Water Res. 100, 38–56. doi:10.1016/j.watres.2016.04.064
Chuah, C.J., Mukhaidin, N., Choy, S.H., Smith, G.J.D., Mendenhall, I.H., Lim, Y.A.L., Ziegler, A.D., 2016. Prevalence of Cryptosporidium and Giardia in the water resources of the Kuang River catchment, Northern Thailand. Sci. Total Environ. 562, 701–713. doi:10.1016/j.scitotenv.2016.03.247
Clasen, T.F., Alexander, K.T., Sinclair, D., Boisson, S., Peletz, R., Chang, H.H., Majorin, F., Cairncross, S., 2015. Interventions to improve water quality for preventing diarrhoea. Cochrane database Syst. Rev. CD004794. doi:10.1002/14651858.CD004794.pub3
Claßen, T., Koch, C., Rechenburg, A., Christoffels, E., Kistemann, T., 2004. Hygienisch-mikrobielle Fließgewässerbelastung durch Regenwasserentlastungsanlagen am Beispiel der Swist. Umweltmedizin Forsch. und Prax. 9, 185–186.
Coffey, R., Cummins, E., Bhreathnach, N., Flaherty, V.O., Cormican, M., 2010. Development of a pathogen transport model for Irish catchments using SWAT. Agric. Water Manag. 97, 101–111.
Coffey, R., Cummins, E., Flaherty, V.O., Cormican, M., 2010. Analysis of the soil and water assessment tool (SWAT) to model Cryptosporidium in surface water sources. Biosyst. Eng. 106, 303–314. doi:10.1016/j.biosystemseng.2010.04.003
Collick, A.S., Fogarty, E.A., Ziegler, P.E., Walter, M.T., Bowman, D.D., Steenhuis, T.S., 2006. Survival of Cryptosporidium parvum oocysts in calf housing facilities in the New York City watersheds. J. Environ. Qual. 35, 680–7. doi:10.2134/jeq2005.0273
Collins, R., Rutherford, K., 2004. Modelling bacterial water quality in streams draining pastoral land. Water Res. 38, 700–712. doi:10.1016/j.watres.2003.10.045
Committee on Indicators for Waterborne Pathogens National Research Council, 2004. Indicators for Waterborne Pathogens. The National Academies Press, Washington, D. C. doi:10.17226/11010
Connelly, S.J., Wolyniak, E.A., Williamson, C.E., Jellison, K.L., 2007. Artificial UV-B and solar radiation reduce in vitro infectivity of the human pathogen Cryptosporidium parvum. Environ. Sci. Technol. 41, 7101–7106. doi:10.1021/es071324r
Côté, C., Massé, D.I., Quessy, S., 2006. Reduction of indicator and pathogenic microorganisms by psychrophilic anaerobic digestion in swine slurries. Bioresour. Technol. 97, 686–691. doi:10.1016/j.biortech.2005.03.024
Cox, P., Griffith, M., Angles, M., Deere, D., Ferguson, C.M., 2005. Concentrations of Pathogens and Indicators in Animal Feces in the Sydney Watershed. Appl. Environ. Microbiol. 71, 5929–5934. doi:10.1128/aem.71.10.5929-5934.2005
Craig, B.H., Pilkington, J.G., Kruuk, L.E.B., Pemberton, J.M., 2007. Epidemiology of parasitic protozoan infections in Soay sheep (Ovis aries L.) on St Kilda. Parasitology 134, 9–21. doi:10.1017/S0031182006001144
Crowther, J., Wyer, M.D., Bradford, M., Kay, D., Francis, C.A., Knisel, W.G., 2003. Modelling faecal indicator concentrations in large rural catchments using land use and topographic data. J. Appl. Microbiol. 94, 962–973.
Davidson, P.C., Kuhlenschmidt, T.B., Bhattarai, R., Kalita, P.K., Kuhlenschmidt, M.S., 2014. Effects of soil type and cover condition on Cryptosporidium parvum transport in overland flow. Water. Air. Soil Pollut. 225. doi:10.1007/s11270-014-1882-4
Davies, C.M., Ferguson, C.M., Kaucner, C., Krogh, M., Altavilla, N., Deere, D.A., Ashbolt, N.J., 2004. Dispersion and transport of Cryptosporidium oocysts from fecal pats under simulated rainfall events. Appl. Environ. Microbiol. 70, 1151–1159. doi:10.1128/AEM.70.2.1151
de Brauwere, A., Ouattara, N.K., Servais, P., 2014. Modeling fecal indicator bacteria concentrations in natural surface waters: a review. Crit. Rev. Environ. Sci. Technol. 140225123512007. doi:10.1080/10643389.2013.829978
de Waele, V., Berzano, M., Berkvens, D., Speybroeck, N., Lowery, C., Mulcahy, G.M., Murphy, T.M., 2011. Age-stratified Bayesian analysis to estimate sensitivity and specificity of four diagnostic tests for detection of Cryptosporidium oocysts in neonatal calves. J. Clin. Microbiol. 49, 76–84. doi:10.1128/JCM.01424-10

de Waele, V., Berzano, M., Speybroeck, N., Berkvens, D., Mulcahy, G.M., Murphy, T.M., 2012. Peri-parturient rise of Cryptosporidium oocysts in cows: New insights provided by duplex quantitative real-time PCR. Vet. Parasitol. 189, 366–368. doi:10.1016/j.vetpar.2012.05.002

de Waele, V., Van den Broeck, F., Huyse, T., McGrath, G., Higgins, I., Speybroeck, N., Berzano, M., Raleigh, P., Mulcahy, G.M., Murphya, T.M., 2013. Panmictic structure of the cryptosporidium parvum population in irish calves: Influence of prevalence and host movement. Appl. Environ. Microbiol. 79, 2534–2541. doi:10.1128/AEM.03613-12

Delafosse, A., Castro-Hermida, J.A., Baudry, C., Ares-Mazás, E., Chartier, C., 2006. Herd-level risk factors for Cryptosporidium infection in dairy-goat kids in western France. Prev. Vet. Med. 77, 109–121. doi:10.1016/j.prevetmed.2006.07.001

Dellström Rosenquist, L.E., 2005. A psychosocial analysis of the human-sanitation nexus. J. Environ. Psychol. 25, 335–346. doi:10.1016/j.jenvp.2005.07.003

Dixon, B., Parrington, L., Cook, A., Pintar, K., Pollari, F., Kelton, D., Farber, J., 2011. The potential for zoonotic transmission of Giardia duodenalis and Cryptosporidium spp. from beef and dairy cattle in Ontario, Canada. Vet. Parasitol. 175, 20–26. doi:10.1016/j.vetpar.2010.09.032

Dixon, B.R., 2016. Parasitic illnesses associated with the consumption of fresh produce — an emerging issue in developed countries. Curr. Opin. Food Sci. 8, 104–109. doi:10.1016/j.cofs.2016.04.009

Dobson, J.E., Bright, E.A., Coleman, P.R., Durfee, R.C., Worley, B.A., 2000. LandScan: A global population database for estimating populations at risk. Photogramm. Eng. Remote Sensing 66, 849–857.

Döll, P., Lehner, B., 2002. Validation of a new global 30-min drainage direction map. J. Hydrol. 258, 214–231. doi:10.1016/S0022-1694(01)00565-0

Dorner, S.M., Anderson, W.B., Slawson, R.M., Kouwen, N., Huck, P.M., 2006. Hydrologic Modeling of Pathogen Fate and Transport. Environ. Sci. Technol. 40, 4746–4753. doi:10.1021/es060426z

DuPont, H.L., Chappell, C.L., Sterling, C.R., Okhuysen, P.C., Rose, J.B., Jakubowski, W., 1995. The Infectivity of Cryptosporidium parvum in Healthy Volunteers. N. Engl. J. Med. 332, 855–859. doi:doi:10.1056/NEJM199503303321304

Duranti, A., SM, C., E, P., Di Egidio, A., De Curtis, M., Battisti, A., Scaramozzino, P., 2009. Risk Factors Associated with Cryptosporidium parvum Infection in Cattle. Zoonoses Public Heal. 56, 176–182. doi:10.1111

Efstratiou, A., Ongerth, J., Karanis, P., 2017a. Evolution of monitoring for Giardia and Cryptosporidium in water. Water Res. 123, 96–112. doi:10.1016/j.watres.2017.06.042

Efstratiou, A., Ongerth, J.E., Karanis, P., 2017b. Waterborne transmission of protozoan parasites: Review of worldwide outbreaks - An update 2011–2016. Water Res. 114, 14–22. doi:10.1016/j.watres.2017.01.036

Ehsan, A., Geurden, T., Casaert, S., Paulussen, J., De Coster, L., Schoemaker, T., Chalmers, R., Grit, G., Vercruysse, J., Claerebout, E., 2015. Occurrence and potential health risk of Cryptosporidium and Giardia in different water catchments in Belgium. Environ. Monit. Assess. 187. doi:10.1007/s10661-014-4157-z

Eisenberg, J.N.S., Brookhart, M.A., Rice, G., Brown, M., Colford Jr., J.M., 2002. Disease transmission models for public health decision making: analysis of epidemic and endemic conditions caused by waterborne pathogens. Environ. Health Perspect. 110, 783–790.

Epstein, P.R., 1993. Algal blooms in the spread and persistence of cholera. Biosystems. 31, 209–221.

FAO, 2016. Global Livestock Environmental Assessment Model - Model description Version 1.0.

FAO, 2013. FAOSTAT database collections [WWW Document]. URL http://faostat.fao.org (accessed 8.11.15).

Fayer, R., 2004. Cryptosporidium: a water-borne zoonotic parasite. Vet. Parasitol. 126, 37–56. doi:10.1016/j.vetpar.2004.09.004

Fayer, R., Santín, M., 2009. Cryptosporidium xiaoi n. sp. (Apicomplexa: Cryptosporidiidae) in sheep (Ovis aries). Vet. Parasitol. 164, 192–200. doi:10.1016/j.vetpar.2009.05.011

Fayer, R., Santin, M., Trout, J.M., 2007. Prevalence of Cryptosporidium species and genotypes in mature dairy cattle on farms in eastern United States compared with younger cattle from the same locations. Vet. Parasitol. 145, 260–266. doi:10.1016/j.vetpar.2006.12.009

Feachem, R.G., Bradley, D.J., Garelick, H., Mara, D.D., 1983. Sanitation and Disease - Health Aspects of Excreta and Wastewater Management, World Bank studies in water supply and sanitation ; no. 3. John Wiley & Sons, New York. doi:10.1111/j.1752-1688.1984.tb04765.x

Fekete, B.M., Vörösmarty, C.J., Lammers, R.B., 2001. Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error. Water Resour. Res. 37, 1955–1967.

Ferguson, C.M., Coote, B.G., Ashbolt, N.J., Stevenson, I.M., 1996. Relationships between indicators, pathogens and water quality in an estuarine system. Water Res. 30, 2045–2054.

Ferguson, C.M., de Roda Husman, A.M., Altavilla, N., Deere, D., Ashbolt, N., 2003. Fate and Transport of Surface Water Pathogens in Watersheds. Crit. Rev. Environ. Sci. Technol. 33, 299–361. doi:10.1080/10643380390814497

Ferguson, C.M., Kay, D., 2012. Transport of microbial pollution in catchment systems, in: Dufour, A., Bartram, J., Bos, R., Gannon, V. (Eds.), Animal Waste, Water Quality and Human Health. World Health Organization, IWA Publishing, London, UK, pp. 157–192.

Ferguson, M.F., Barry, F.W.C., Peter, J.B., Nicholas, J.A., Daniel, A.D., Ferguson, C.M., Croke, B.F.W., Beatson, P.J., Ashbolt, N.J., Deere, D.A., Ferguson, M.F., Barry, F.W.C., Peter, J.B., Nicholas, J.A., Daniel, A.D., 2007. Development of a process-based model to predict pathogen budgets for the Sydney drinking water catchment. J. Water Health 5, 187–208. doi:10.2166/wh.2007.013

Fischer Walker, C.L., Black, R.E., 2010. Diarrhoea morbidity and mortality in older children, adolescents, and adults. Epidemiol. Infect. 138, 1215–1226. doi:10.1017/S0950268810000592

Fischer Walker, C.L., Perin, J., Aryee, M.J., Boschi-Pinto, C., Black, R.E., 2012. Diarrhea incidence in low- and middle-income countries in 1990 and 2010: a systematic review. BMC Public Health 12, 220. doi:10.1186/1471-2458-12-220

Fisher, J.B., Sitch, S., Malhi, Y., Fisher, R.A., Huntingford, C., Tan, S.-Y., 2010. Carbon cost of plant nitrogen acquisition: A mechanistic, globally applicable model of plant nitrogen uptake, retranslocation, and fixation. Global Biogeochem. Cycles 24. doi:10.1029/2009GB003621

Follet, J., Guyot, K., Leruste, H., Follet-Dumoulin, A., Hammouma-Ghelboun, O., Certad, G., Dei-Cas, E., Halama, P., 2011. Cryptosporidium infection in a veal calf cohort in France: Molecular characterization of species in a longitudinal study. Vet. Res. 42, 116. doi:10.1186/1297-9716-42-116

Fraser, R.H., Barten, P.K., Pinney, D.A.K., 1998. Predicting Stream Pathogen Loading from Livestock using a Geographical Information System-Based Delivery Model. J. Environ. Qual. 27, 935–945. doi:10.2134/jeq1998.00472425002700040030x

Galloway, J., Dentener, F., Capone, D., 2004. Nitrogen cycles: past, present, and future. Biogeochemistry 70, 153–226.

Garcés, G., Effenberger, M., Najdrowski, M., Wackwitz, C., Gronauer, A., Wilderer, P.A., Lebuhn, M., 2006. Quantification of Cryptosporidium parvum in anaerobic digesters treating manure by (reverse-transcription) quantitative real-time PCR, infectivity and excystation tests. Water Sci. Technol. 53, 195–202. doi:10.2166/wst.2006.250

Garg, A.X., Marshall, J., Salvadori, M., Thiessen-Philbrook, H.R., Macnab, J., Suri, R.S., Haynes, R.B., Pope, J., Clark, W., Collins, S., Howard, J., Mahon, J., Matsell, D., Moist, L., Ray, J., Rosas-Arellano, P., 2006. A gradient of acute gastroenteritis was characterized, to assess risk of long-term health sequelae after drinking bacterial-contaminated water. J. Clin. Epidemiol. 59, 421–428. doi:10.1016/j.jclinepi.2005.08.014

Garzio-Hadzick, A., Shelton, D.R., Hill, R.L., Pachepsky, Y.A., Guber, A.K., Rowland, R., 2010. Survival of manure-borne E. coli in streambed sediment: Effects of temperature and sediment properties. Water Res. 44, 2753–2762.

Gauthier, F., Archibald, F., 2001. The ecology of “fecal indicator” bacteria commonly found in pulp and paper mill water systems. Water Res. 35, 2207–2218. doi:10.1016/S0043-1354(00)00506-6

GBD 2013 DALYs and HALE Collaborators, 2015. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990-2013: Quantifying the epidemiological transition. Lancet 386, 2145–2191. doi:10.1016/S0140-6736(15)61340-X

GBD 2013 Mortality and Causes of Death Collaborators, 2015. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. Lancet 385, 117–171. doi:10.1016/S0140-6736(14)61682-2

Geurden, T., Berkvens, D., Martens, C., Casaert, S., Vercruysse, J., Claerebout, E., 2007. Molecular epidemiology with subtype analysis of Cryptosporidium in calves in Belgium. Parasitology 134, 1981–1987. doi:10.1017/S0031182007003460

Geurden, T., Goma, F.Y., Siwila, J., Phiri, I.G.K., Mwanza, A.M., Gabriel, S., Claerebout, E., Vercruysse, J., 2006. Prevalence and genotyping of Cryptosporidium in three cattle husbandry systems in Zambia. Vet. Parasitol. 138, 217–222. doi:10.1016/j.vetpar.2006.02.009

Geurden, T., Thomas, P., Casaert, S., Vercruysse, J., Claerebout, E., 2008. Prevalence and molecular characterisation of Cryptosporidium and Giardia in lambs and goat kids in Belgium. Vet. Parasitol. 155, 142–145. doi:10.1016/j.vetpar.2008.05.002

Gibson III, C.J., Stadterman, K.L., States, S., Sykora, J., 1998. Combined sewer overflows: A source of Cryptosporidium and Giardia? Water Sci. Technol. 38, 67–72.

Green, P.A., Vörösmarty, C.J., Meybeck, M., Galloway, J.N., Peterson, B.J., Boyer, E.W., 2004. Pre-industrial and contemporary fluxes of nitrogen through rivers: a global assessment based on typology. Biogeochemistry 68, 71–105. doi:10.1023/B:BIOG.0000025742.82155.92

Grizzetti, B., Bouraoui, F., Granlund, K., Rekolainen, S., Bidoglio, G., 2003. Modelling diffuse emission and retention of nutrients in the Vantaanjoki watershed (Finland) using the SWAT model. Ecol. Modell. 169, 25–38. doi:10.1016/S0304-3800(03)00198-4

Guber, A.K., Shelton, D.R., Pachepsky, Y.A., Sadeghi, A.M., Sikora, L.J., 2006. Rainfall-induced release of fecal coliforms and other manure constituents: Comparison and modeling. Appl. Environ. Microbiol. 72, 7531–7539. doi:10.1128/AEM.01121-06

Haas, C., Scheff, P., 1990. Estimation of averages in truncated samples. Environ. Sci. Technol. 24, 912–919. doi:10.1021/es00076a021

Haas, C.N., 1996. How to average microbial densities to characterize risk. Water Res. 30, 1036–1038. doi:10.1016/0043-

1354(95)00228-6
Haas, C.N., Rose, J.B., Gerba, C.P., 1999. Quantitative microbial risk assessment. John Wiley & Sons, Inc., New York.
Haddeland, I., Clark, D.B., Franssen, W., Ludwig, F., Vosz, F., Arnell, N.W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S.N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G.P., Yeh, P., 2011. Multimodel estimate of the global terrestrial water balance: setup and first results. J. Hydrometeorol. 12, 869–884. doi:10.1175/2011JHM1324.1
Haines, A., Kovats, R.S., Campbell-Lendrum, D., Corvalan, C., 2006. Climate change and human health: impacts, vulnerability, and mitigation. Public Health 120, 585–596. doi:10.1016/j.puhe.2006.01.002
Haller, L., Hutton, G., Bartram, J., 2007. Estimating the costs and health benefits of water and sanitation improvements at global level. J. Water Health 5, 467–480. doi:10.2166/wh.2007.008
Hanchett, S., Akhter, S., Khan, M.H., Mezulianik, S., Blagbrough, V., 2003. Water, sanitation and hygiene in Bangladeshi slums: an evaluation of the WaterAid- Bangladesh urban programme. Environ. Urban. 15, 43–56. doi:10.1177/095624780301500219
Haque, R., Mondal, D., Karim, A., Molla, I.H., Rahim, A., Faruque, A.S.G., Ahmad, N., Kirkpatrick, B.D., Houpt, E., Snider, C., Petri, W. a, 2009. Prospective case-control study of the association between common enteric protozoal parasites and diarrhea in Bangladesh. Clin. Infect. Dis. 48, 1191–1197. doi:10.1086/597580
Harrison, J.A., Caraco, N., Seitzinger, S.P., 2005. Global patterns and sources of dissolved organic matter export to the coastal zone: Results from a spatially explicit, global model. Global Biogeochem. Cycles 19. doi:10.1029/2005GB002480
Hartley, D.M., 1987. Simplified Process Model for Water Sediment Yield From Single Storms Part I - Model Formulation. Trans. Am. Soc. Agric. Eng. 30, 710–717. doi:10.13031/2013.30464
Havelaar, A.H., Melse, J.M., 2003. Quantifying public health risk in the WHO Guidelines for Drinking Water Quality: a burden of disease approach. RIVM report 734301022. Bilthoven, the Netherlands. doi:10.1007/s12594-013-0053-3
Haydon, S., Deletic, A., 2006. Development of a coupled pathogen-hydrologic catchment model. J. Hydrol. 328, 467–480. doi:DOI 10.1016/j.jhydrol.2005.12.033
Helmy, Y.A., Krücken, J., Nöckler, K., von Samson-Himmelstjerna, G., Zessin, K.H., 2013. Molecular epidemiology of Cryptosporidium in livestock animals and humans in the Ismailia province of Egypt. Vet. Parasitol. 193, 15–24. doi:10.1016/j.vetpar.2012.12.015
Helsel, D., 2010a. Much ado about next to nothing: Incorporating nondetects in science. Ann. Occup. Hyg. 54, 257–262. doi:10.1093/annhyg/mep092
Helsel, D., 2010b. Summing Nondetects: Incorporating Low-Level Contaminants in Risk Assessment. Integr. Environ. Assess. Manag. 6, 361–366. doi:10.1002/IEAM.31
Hijmans, R.J., 2016. raster: Geographic Data Analysis and Modeling.
Hipsey, M.R., Antenucci, J.P., Brookes, J.D., 2008. A generic, process-based model of microbial pollution in aquatic systems. Water Resour. Res. 44, W07408. doi:10.1029/2007wr006395
Hofstra, N., 2011. Quantifying the impact of climate change on enteric waterborne pathogen concentrations in surface water. Curr. Opin. Environ. Sustain. 3, 471–479. doi:10.1016/j.cosust.2011.10.006
Hofstra, N., Bouwman, A.F., Beusen, A.H.W., Medema, G.J., 2013. Exploring global Cryptosporidium emissions to surface water. Sci. Total Environ. 442, 10–19. doi:http://dx.doi.org/10.1016/j.scitotenv.2012.10.013
Hornok, S., Bitay, Z., Széll, Z., Varga, I., 1998a. Assessment of maternal immunity to Cryptosporidium baileyi in chickens. Vet. Parasitol. 79, 203–212. doi:10.1016/S0304-4017(98)00170-8
Hornok, S., Heijmans, J.F., Békési, L., Peek, H.W., Dobos-Kovács, M., Drén, C.N., Varga, I., 1998b. Interaction of chicken anaemia virus and Cryptosporidium baileyi in experimentally infected chickens. Vet. Parasitol. 76, 43–55. doi:10.1016/S0304-4017(97)00046-0
Hornok, S., Széll, Z., Shibalova, T., Varga, I., 1999. Study on the course of Cryptosporidium baileyi infection in chickens treated with interleukin-1 or indomethacin. Acta Vet. Hung. 47, 207–216.
Hornok, S., Széll, Z., Sréter, T., Kovács, A., Varga, I., 2000. Influence of in ovo administered Cryptosporidium baileyi oocyst extract on the course of homologous infection. Vet. Parasitol. 89, 313–319. doi:10.1016/S0304-4017(00)00217-X
Howard, G., Pedley, S., Tibatemwa, S., 2006. Quantitative microbial risk assessment to estimate health risks attributable to water supply: Can the technique be applied in developing countries with limited data? J. Water Health 4, 49–65. doi:10.2166/wh.2005.058
Howarth, R., Billen, G., Swaney, D., Townsend, A., Jaworski, N., Lajtha, K., Downing, J., Elmgren, R., Caraco, N., Jordan, T., Berendse, F., Freney, J., Kudeyarov, V., Murdoch, P., Zhao-Liang, Z., 1996. Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. Biogeochemistry 35, 75–139. doi:10.1007/BF02179825
Hoyer, A.B., Schladow, S.G., Rueda, F.J., 2015. A hydrodynamics-based approach to evaluating the risk of waterborne pathogens entering drinking water intakes in a large, stratified lake. Water Res. 83, 227–236. doi:10.1016/j.watres.2015.06.014
Hutchison, M.L., Walters, L.D., Moore, A., Avery, S.M., 2005. Declines of zoonotic agents in liquid livestock wastes stored

in batches on-farm. J. Appl. Microbiol. 99, 58–65. doi:10.1111/j.1365-2672.2005.02585.x
International Institute for Population Sciences (IIPS), Macro International, 2007. National Family Health Survey (NFHS-3), 2005–06: India. Mumbai.
IPCC, 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, vol. 4 ch. 10.
Ippolito, A., Kattwinkel, M., Rasmussen, J.J., Schäfer, R.B., Fornaroli, R., Liess, M., 2015. Modeling global distribution of agricultural insecticides in surface waters. Environ. Pollut. 198, 54–60. doi:10.1016/j.envpol.2014.12.016
Iqbal, M.S., 2017. Quantifying The Impact of Socioeconomic Development and Climate Change on Escherichia coli Concentrations in The Pakistani Kabul River. Wageningen University. doi:10.18174/416179
Islam, M.M.M., 2017. Assessing the Impact of Socio-economic Development and Climate Change on Faecal Indicator Bacteria in the Betna River, Bangladesh. Wageningen University. doi:10.18174/419642
Jenkins, M.B., Liotta, J.L., Bowman, D.D., 2013. Inactivation kinetics of Cryptosporidium parvum oocysts in a swine waste lagoon and spray field. J. Parasitol. 99, 337–42. doi:10.1645/GE-3193.1
Jenkins, M.B., Walker, M.J., Bowman, D.D., Anthony, L.C., Ghiorse, W.C., 1999. Use of a sentinel system for field measurements of Cryptosporidium parvum oocyst inactivation in soil and animal waste. Appl. Environ. Microbiol. 65, 1998–2005.
Jiang, L., O'Neill, B.C., 2017. Global urbanization projections for the Shared Socioeconomic Pathways. Glob. Environ. Chang. 42, 193–199. doi:10.1016/j.gloenvcha.2015.03.008
Karanis, P., Kourenti, C., Smith, H., 2007. Waterborne transmission of protozoan parasites: A worldwide review of outbreaks and lessons learnt. J. Water Health 5, 1–38. doi:Doi 10.2166/Wh.2006.002
Katukiza, A.Y., Ronteltap, M., van der Steen, P., Foppen, J.W.A., Lens, P.N.L., 2013. Quantification of microbial risks to human health caused by waterborne viruses and bacteria in an urban slum. J. Appl. Microbiol. 116, 447–463. doi:10.1111/jam.12368
Kay, D., Crowther, J., Fewtrell, L., Francis, C.A., Hopkins, M., Kay, C., McDonald, A.T., Stapleton, C.M., Watkins, J., Wilkinson, J., Wyer, M.D., 2008. Quantification and control of microbial pollution from agriculture: a new policy challenge? Environ. Sci. Policy 11, 171–184. doi:10.1016/j.envsci.2007.10.009
Kay, D., Mcdonald, A., 1983. Predicting Coliform Concentrations in Upland Impoundments - Design and Calibration of a Multivariate Model. Appl. Environ. Microbiol. 46, 611–618.
Kay, D., Wyer, M., Crowther, J., Stapleton, C., Bradford, M., McDonald, A., Greaves, J., Francis, C., Watkins, J., 2005. Predicting faecal indicator fluxes using digital land use data in the UK's sentinel Water Framework Directive catchment: The Ribble study. Water Res. 39, 3967–3981. doi:DOI 10.1016/j.watres.2005.07.006
KC, S., Lutz, W., 2017. The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100. Glob. Environ. Chang. 42, 181–192. doi:10.1016/j.gloenvcha.2014.06.004
Khan, S.M., Debnath, C., Pramanik, A.K., Xiao, L., Nozaki, T., Ganguly, S., 2010. Molecular characterization and assessment of zoonotic transmission of Cryptosporidium from dairy cattle in West Bengal, India. Vet. Parasitol. 171, 41–47. doi:10.1016/j.vetpar.2010.03.008
King, B.J., Hoefel, D., Daminato, D.P., Fanok, S., Monis, P.T., 2008. Solar UV reduces Cryptosporidium parvum oocyst infectivity in environmental waters. J. Appl. Microbiol. 104, 1311–1323. doi:10.1111/j.1365-2672.2007.03658.x
Kinyua, M.N., Trimmer, J., Izurieta, R., Cunningham, J., Ergas, S.J., 2016. Viability and fate of Cryptosporidium parvum and Giardia lamblia in tubular anaerobic digesters. Sci. Total Environ. 554–555, 167–177. doi:10.1016/j.scitotenv.2016.02.170
Kirk, M.D., Pires, S.M., Black, R.E., Caipo, M., Crump, J.A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C.L., Hald, T., Hall, A.J., Keddy, K.H., Lake, R.J., Lanata, C.F., Torgerson, P.R., Havelaar, A.H., Angulo, F.J., 2015. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLoS Med. 12, 1–21. doi:10.1371/journal.pmed.1001921
Kistemann, T., Classen, T., Koch, C., Dangendorf, F., Fischeder, R., Gebel, J., Vacata, V., Exner, M., 2002. Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. Appl Env. Microbiol 68, 2188–2197.
Kistemann, T., Rind, E., Koch, C., Claßen, T., Lengen, C., Exner, M., Rechenburg, A., 2012. Effect of sewage treatment plants and diffuse pollution on the occurrence of protozoal parasites in the course of a small river. Int. J. Hyg. Environ. Health 215, 577–583. doi:10.1016/j.ijheh.2011.12.008
Kiulia, N.M., Hofstra, N., Vermeulen, L.C., Obara, M.A., Medema, G.J., Rose, J.B., 2015. Global occurrence and emission of rotaviruses to surface waters. Pathogens 4, 229–55. doi:10.3390/pathogens4020229
Koelle, K., Pascual, M., Yunus, M., 2005. Pathogen adaptation to seasonal forcing and climate change. Proc. R. Soc. Biol. Sci. 272, 971–977. doi:10.1098/rspb.2004.3043
Koinari, M., Lymbery, A.J., Ryan, U.M., 2014. Cryptosporidium species in sheep and goats from Papua New Guinea. Exp. Parasitol. 141, 134–137. doi:10.1016/j.exppara.2014.03.021
Kolstad, E.W., Johansson, K.A., 2011. Uncertainties associated with quantifying climate change impacts on human health: a case study for diarrhea. Environ. Health Perspect. 119, 299–305. doi:10.1289/ehp.1002060
Kroeze, C., Bouwman, L., Seitzinger, S., 2012. Modeling global nutrient export from watersheds. Curr. Opin. Environ. Sustain. 4, 195–202. doi:10.1016/j.cosust.2012.01.009

Kroeze, C., Gabbert, S., Hofstra, N., Koelmans, A.A., Li, A., Löhr, A., Ludwig, F., Strokal, M., Verburg, C., Vermeulen, L., van Vliet, M.T., de Vries, W., Wang, M., van Wijnen, J., 2016. Global modelling of surface water quality: a multi-pollutant approach. Curr. Opin. Environ. Sustain. 23, 35–45. doi:10.1016/j.cosust.2016.11.014

Kuczynska, E., Shelton, D.R., 1999. Method for Detection and Enumeration of Cryptosporidium parvum Oocysts in Feces, Manures, and Soils. Appl. Environ. Microbiol. 65, 2820–2826.

Kváč, M., Kouba, M., Vítovec, J., 2006. Age-related and housing-dependence of Cryptosporidium infection of calves from dairy and beef herds in South Bohemia, Czech Republic. Vet. Parasitol. 137, 202–209. doi:10.1016/j.vetpar.2006.01.027

Kváč, M., Květoňová, D., Půžová, G., Ditrich, O., 2003. Comparison of Selected Diagnostic Methods for Identification of Cryptosporidium parvum and Cryptosporidium andersoni in Routine Examination of Faeces. J. Vet. Med. Ser. B Infect. Dis. Vet. Public Heal. 50, 405–411. doi:10.1046/j.1439-0450.2003.00694.x

Labatut, R.A., Angenent, L.T., Scott, N.R., 2014. Conventional mesophilic vs. thermophilic anaerobic digestion: a trade-off between performance and stability? Water Res. 53, 249–58. doi:10.1016/j.watres.2014.01.035

Lai, M., Zhou, R.Q., Huang, H.C., Hu, S.J., 2011. Prevalence and risk factors associated with intestinal parasites in pigs in Chongqing, China. Res. Vet. Sci. 91, 121–124. doi:10.1016/j.rvsc.2011.01.025

Lalancette, C., Papineau, I., Payment, P., Dorner, S., Servais, P., Barbeau, B., Di Giovanni, G.D., Prévost, M., Di, G.D., Bruxelles, L. De, Plaine, C. De, Di Giovanni, G.D., Prévost, M., 2014. Changes in Escherichia coli to Cryptosporidium ratios for various fecal pollution sources and drinking water intakes. Water Res. 55, 150–61. doi:10.1016/j.watres.2014.01.050

Lanata, C.F., Fischer-Walker, C.L., Olascoaga, A.C., Torres, C.X., Aryee, M.J., Black, R.E., 2013. Global Causes of Diarrheal Disease Mortality in Children <5 Years of Age: A Systematic Review. PLoS One 8. doi:10.1371/journal.pone.0072788

Lebreton, L.C.M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., Reisser, J., 2017. River plastic emissions to the world's oceans. Nat. Commun. 8, 15611. doi:10.1038/ncomms15611

Lee, L., 2017. NADA: Nondetects and Data Analysis for Environmental Data. R package version 1.6-1.

Lee, L., Helsel, D., 2007. Statistical analysis of water-quality data containing multiple detection limits II: S-language software for nonparametric distribution modeling and hypothesis testing. Comput. Geosci. 33, 696–704. doi:10.1016/j.cageo.2006.09.006

Lee, L., Helsel, D., 2005. Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. Comput. Geosci. 31, 1241–1248. doi:10.1016/j.cageo.2005.03.012

Leopold, L.B., Maddock, T.J., 1953. The Hydraulic Geomtry of Stream Channels and Some Physiographic Implications. Geol. Surv. Prof. Pap. 252 57.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs. J. Geophys. Res. 99(D7), 14,414-415,428.

Lin, J., Ganesh, A., 2013. Water quality indicators: Bacteria, coliphages, enteric viruses. Int. J. Environ. Health Res. 23, 484–506. doi:10.1080/09603123.2013.769201

Liu, J., Platts-Mills, J.A., Juma, J., Kabir, F., Nkeze, J., Okoi, C., Operario, D.J., Uddin, J., et al., 2016. Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. Lancet 388, 1291–1301. doi:10.1016/S0140-6736(16)31529-X

Liu, Z.J., Hashim, N.B., Kingery, W.L., Huddleston, D.H., 2011. Fecal coliform modeling under two flow scenarios in St. Louis Bay of Mississippi. J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng. 45, 282–291. doi:10.1080/10934520903467949

Mackenzie, W.R., Hoxie, N.J., Proctor, M.E., Gradus, M.S., Blair, K.A., Peterson, D.E., Kazmierczak, J.J., Addiss, D.G., Fox, K.R., Rose, J.B., Davis, J.P., 1994. A Massive Outbreak in Milwaukee of Cryptosporidium Infection Transmitted through the Public Water-Supply. N. Engl. J. Med. 331, 161–167.

MacPepple, E., Okereke, N., Egwuonwu, C., H. Igoni, A., 2017. Effect of Raw Sewage on Surface Water Quality in Opobo Town, Nigeria. Environ. Ecol. Res. 5, 220–226. doi:10.13189/eer.2017.050307

Maddox-Hyttel, C., Langkjær, R.B., Enemark, H.L., Vigre, H., 2006. Cryptosporidium and Giardia in different age groups of Danish cattle and pigs-Occurrence and management associated risk factors. Vet. Parasitol. 141, 48–59. doi:10.1016/j.vetpar.2006.04.032

Maikai, B. V., Umoh, J.U., Kwaga, J.K.P., Lawal, I.A., Maikai, V.A., Cama, V., Xiao, L., 2011. Molecular characterization of Cryptosporidium spp. in native breeds of cattle in Kaduna State, Nigeria. Vet. Parasitol. 178, 241–245. doi:10.1016/j.vetpar.2010.12.048

Mancini, J.L., 1978. Numerical Estimates of Coliform Mortality Rates under Various Conditions. J. (Water Pollut. Control Fed. 50, 2477–2484. doi:10.2307/25040179

Maurya, P.S., Rakesh, R.L., Pradeep, B., Kumar, S., Kundu, K., Garg, R., Ram, H., Kumar, A., Banerjee, P.S., 2013. Prevalence and risk factors associated with Cryptosporidium spp. infection in young domestic livestock in India. Trop. Anim. Health Prod. 45, 941–946. doi:10.1007/s11250-012-0311-1

Mawdsley, J.L., Brooks, a E., Merry, R.J., 1996. Movement of the protozoan pathogen Cryptosporidium parvum through

three contrasting soil types. Biol. Fertil. Soils 21, 30–36. doi:10.1007/BF00335990
Mayorga, E., Seitzinger, S.P., Harrison, J.A., Dumont, E., Beusen, A.H.W., Bouwman, A.F., Fekete, B.M., Kroeze, C., Van Drecht, G., 2010. Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. Environ. Model. Softw. 25, 837–853. doi:10.1016/j.envsoft.2010.01.007
McLaughlin, S.J., Kalita, P.K., Kuhlenschmidt, M.S., 2013. Fate of Cryptosporidium parvum oocysts within soil, water, and Plant environment. J. Environ. Manage. 131, 121–8. doi:10.1016/j.jenvman.2013.09.017
McMichael, A.J., Neira, M., Bertollini, R., Campbell-Lendrum, D., Hales, S., 2009. Climate change: a time of need and opportunity for the health sector. Lancet 374, 2123–2125. doi:10.1016/S0140-6736(09)62031-6
Medema, G.J., Schets, F.M., Teunis, P.F.M., Havelaar, A.H., 1998. Sedimentation of free and attached Cryptosporidium oocysts and Giardia cysts in water. Appl. Environ. Microbiol. 64, 4460–4466.
Medema, G.J., Schijven, J.F., 2001. Modelling the sewage discharge and dispersion of Cryptosporidium and Giardia in surface water. Water Res. 35, 4307–4316.
Medema, G.J., Teunis, P., Blokker, M., Deere, D., Davison, A., Charles, P., Loret, J.F., 2009. Risk assessment of Cryptosporidium in drinking water. WHO, Geneva, Switzerland.
Mi, R., Wang, X., Huang, Y., Zhou, P., Liu, Y., Chen, Y., Chen, J., Zhu, W., Chen, Z., 2014. Prevalence and molecular characterization of Cryptosporidium in goats across four provincial level areas in China. PLoS One 9, 1–7. doi:10.1371/journal.pone.0111164
Monaghan, R.M., de Klein, C.A.M., Muirhead, R.W., 2008. Prioritisation of farm scale remediation efforts for reducing losses of nutrients and faecal indicator organisms to waterways: A case study of New Zealand dairy farming. J. Environ. Manage. 87, 609–622. doi:10.1016/j.jenvman.2006.07.017
Monis, P., King, B., Keegan, A., 2014. Chapter 13 - Removal and Inactivation of Cryptosporidium from Water, in: Cacciò, S.M., Widmer, G. (Eds.), Cryptosporidium: Parasite and Disease. Springer, pp. 515–552. doi:10.1007/978-3-7091-1562-6
Morée, A.L., Beusen, A.H.W., Bouwman, A.F., Willems, W.J., 2013. Exploring global nitrogen and phosphorus flows in urban wastes during the twentieth century. Global Biogeochem. Cycles 27, 836–846. doi:10.1002/gbc.20072
Muhid, A., Robertson, I., Ng, J., Ryan, U., 2011. Prevalence of and management factors contributing to Cryptosporidium sp. infection in pre-weaned and post-weaned calves in Johor, Malaysia. Exp. Parasitol. 127, 534–538. doi:10.1016/j.exppara.2010.10.015
Nasser, A.M., 2016. Removal of Cryptosporidium by wastewater treatment processes: A review. J. Water Health 14, 1–13. doi:10.2166/wh.2015.131
Nath, K.J., 2003. Home hygiene and environmental sanitation: a country situation analysis for India. Int. J. Environ. Health Res. 13 Suppl 1, S19-28. doi:10.1080/0960312031000102778
National Institute of Population Research and Training (NIPORT), Mitra and Associates, ICF International, 2013. Bangladesh Demographic and Health Survey 2011. Dhaka, Bangladesh and Calverton, Maryland, USA.
Nevers, M.B., Whitman, R.L., 2005. Nowcast modeling of Escherichia coli concentrations at multiple urban beaches of southern Lake Michigan. Water Res. 39, 5250–5260. doi:10.1016/j.watres.2005.10.012
Nguyen, H.T.M., Billen, G., Garnier, J., Rochelle-Newall, E., Ribolzi, O., Servais, P., Le, Q.T.P., 2016. Modelling of faecal indicator bacteria (FIB) in the Red River basin (Vietnam). Environ. Monit. Assess. 188. doi:10.1007/s10661-016-5528-4
Nichols, G., Lane, C., Asgari, N., Verlander, N.Q., Charlett, A., 2009. Rainfall and outbreaks of drinking water related disease and in England and Wales. J. Water Health 7, 1–8. doi:Doi 10.2166/Wh.2009.143
Nilsen, V., Wyller, J., 2016. QMRA for Drinking Water: 1. Revisiting the Mathematical Structure of Single-Hit Dose-Response Models. Risk Anal. 36, 145–162. doi:10.1111/risa.12389
Noordeen, F., Faizal, A.C.M., Rajapakse, R.P.V.J., Horadagoda, N.U., Arulkanthan, A., 2001. Excretion of Cryptosporidium oocysts by goats in relation to age and season in the dry zone of Sri Lanka. Vet. Parasitol. 99, 79–85. doi:10.1016/S0304-4017(01)00449-6
Nyenje, P.M., Foppen, J.W., Uhlenbrook, S., Kulabako, R., Muwanga, a, 2010. Eutrophication and nutrient release in urban areas of sub-Saharan Africa--a review. Sci. Total Environ. 408, 447–55. doi:10.1016/j.scitotenv.2009.10.020
O'Neill, B.C., Kriegler, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., van Ruijven, B.J., van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M., Solecki, W., 2017. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Glob. Environ. Chang. 42, 169–180. doi:10.1016/j.gloenvcha.2015.01.004
Oakley, J.E., O'Hagan, A., 2004. Probabilistic Sensitivity Analysis of Complex Models : A Bayesian Approach Author (s): Jeremy E . Oakley and Anthony O â€™ Hagan Published by : Wiley for the Royal Statistical Society Stable URL : http://www.jstor.org/stable/3647504 Probabilistic sensitiv 66, 751–769.
Oates, S.C., Miller, M.A., Hardin, D., Conrad, P.A., Melli, A., Jessup, D.A., Dominik, C., Roug, A., Tinker, M.T., Miller, W.A., 2012. Prevalence, environmental loading, and molecular characterization of Cryptosporidium and Giardia isolates from domestic and wild animals along the central California coast. Appl. Environ. Microbiol. 78, 8762–8772. doi:10.1128/AEM.02422-12

Oliver, D.M., Heathwaite, A.L., Fish, R.D., Chadwick, D.R., Hodgson, C.J., Winter, M., Butler, A.J., 2009. Scale appropriate modelling of diffuse microbial pollution from agriculture. Prog. Phys. Geogr. 33, 358–377. doi:10.1177/0309133309342647

Oliver, D.M., Page, T., Heathwaite, A.L., Haygarth, P.M., 2010. Re-shaping models of E. coli population dynamics in livestock faeces: Increased bacterial risk to humans? Environ. Int. 36, 1–7. doi:10.1016/j.envint.2009.08.006

Oliver, D.M., Porter, K.D.H., Pachepsky, Y.A., Muirhead, R.W., Reaney, S.M., Coffey, R., Kay, D., Milledge, D.G., Hong, E., Anthony, S.G., Page, T., Bloodworth, J.W., Mellander, P.E., Carbonneau, P.E., McGrane, S.J., Quilliam, R.S., 2016. Predicting microbial water quality with models: Over-arching questions for managing risk in agricultural catchments. Sci. Total Environ. 544, 39–47. doi:10.1016/j.scitotenv.2015.11.086

Olson, M.E., Goh, J., Phillips, M., Guselle, N., McAllister, T.A., 1999. Giardia Cyst and Cryptosporidium Oocyst Survival in Water, Soil, and Cattle Feces. J. Environ. Qual. 28, 1991. doi:10.2134/jeq1999.00472425002800060040x

Olyphant, G.A., Whitman, R.L., 2004. Elements of a predictive model for determining beach closures on a real time basis: The case of 63rd Street Beach Chicago. Environ. Monit. Assess. 98, 175–190. doi:10.1023/B:EMAS.0000038185.79137.b9

Ongerth, J.E., 2016. Cryptosporidium and Giardia in Water : Reassessment of Occurrence and Significance. J. Environ. Eng. 143, 1–8. doi:10.1061/(ASCE)EE.1943-7870.0001161.

Ongerth, J.E., 2013. LT2 Cryptosporidium data: What do they tell us about cryptosporidium in surface water in the United States? Environ. Sci. Technol. 47, 4029–4038. doi:10.1021/es4006509

Ouattara, N.K., de Brauwere, A., Billen, G., Servais, P., 2013. Modelling faecal contamination in the Scheldt drainage network. J. Mar. Syst. 128, 77–88. doi:10.1016/j.jmarsys.2012.05.004

Pachepsky, Y.A., Sadeghi, A.M., Bradford, S.A., Shelton, D.R., Guber, A.K., Dao, T., 2006. Transport and fate of manure-borne pathogens: Modeling perspective. Agric. Water Manag. 86, 81–92. doi:10.1016/j.agwat.2006.06.010

Pachepsky, Y.A., Shelton, D.R., 2011. Escherichia Coli and Fecal Coliforms in Freshwater and Estuarine Sediments. Crit. Rev. Environ. Sci. Technol. 41, 1067–1110. doi:10.1080/10643380903392718

Panel on Urban Population Dynamics, 2003. Cities transformed: Demographic change and its implications in the developing world. The National Academies Press, Washington, D. C.

Pärn, J., Pinay, G., Mander, Ü., 2012. Indicators of nutrients transport from agricultural catchments under temperate climate: A review. Ecol. Indic. 22, 4–15. doi:10.1016/j.ecolind.2011.10.002

Patz, J.A., Olson, S.H., Uejio, C.K., Gibbs, H.K., 2008. Disease Emergence from Global Climate and Land Use Change. Med. Clin. North Am. 92, 1473–1491. doi:DOI 10.1016/j.mcna.2008.07.007

Payment, P., Locas, A., 2011. Pathogens in Water: Value and Limits of Correlation with Microbial Indicators. Ground Water 49, 4–11. doi:10.1111/j.1745-6584.2010.00710.x

Peng, X., Murphy, T., Holden, N.M., 2008. Evaluation of the effect of temperature on the die-off rate for Cryptosporidium parvum oocysts in water, soils, and feces. Appl. Environ. Microbiol. 74, 7101–7107. doi:10.1128/AEM.01442-08

Petersen, C.M., Rifai, H.S., Stien, R., Stein, R., 2009. Bacteria load estimator spreadsheet tool for modeling spatial Escherichia coli loads to an urban Bayou. J. Environ. Eng. 135, 203–217. doi:10.1061/(ASCE)0733-9372(2009)135:4(203)

Petersen, H.H., Jianmin, W., Katakam, K.K., Mejer, H., Thamsborg, S.M., Dalsgaard, A., Olsen, A., Enemark, H.L., 2015. Cryptosporidium and Giardia in Danish organic pig farms: Seasonal and age-related variation in prevalence, infection intensity and species/genotypes. Vet. Parasitol. 214, 29–39. doi:10.1016/j.vetpar.2015.09.020

Pongsiri, M.J., Gatzweiler, F.W., Bassi, A.M., Haines, A., Demassieux, F., 2017. The need for a systems approach to planetary health. Tha Lancet Planet. Heal. 1, 257–259. doi:10.1016/S2542-5196(17)30116-X

Pritchard, G.C., Marshall, J.A., Giles, M., Chalmers, R.M., Marshall, R.N., 2007. Cryptosporidium parvum infection in orphan lambs on a farm open to the public. Vet. Rec. 161, 11–14.

Qi, M.Z., Fang, Y.Q., Wang, X.T., Du, S.Z., Guo, Y.X., Jia, Y.Q., Yao, L., Liu, Q.D., Zhao, G.H., Zhang, L.X., Wang, R.J., 2015. Molecular characterization of Cryptosporidium spp. in pre-weaned calves in Shaanxi Province, North-western China. J. Med. Microbiol. 64, 111–116. doi:10.1099/jmm.0.079327-0

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. R Found. Stat. Comput., R Foundation for Statistical Computing. doi:10.1007/978-3-540-74686-7

Rahman, M., Shahid, N.S., Rahman, H., Sack, D.A., Rahman, N., Hossain, S., 1990. Cryptosporidiosis: a cause of diarrhea in Bangladesh. Am. J. Trop. Med. Hyg. 42, 127–130.

Ramirez, N.E., Ward, L.A., Sreevatsan, S., 2004. A review of the biology and epidemiology of cryptosporidiosis in humans and animals. Microbes Infect. 6, 773–785. doi:DOI 10.1016/j.micinf.2004.02.021

Rankinen, K., Butterfield, D., Faneca Sànchez, M., Grizzetti, B., Whitehead, P., Pitkänen, T., Uusi-Kämppä, J., Leckie, H., 2016. The INCA-Pathogens model: An application to the Loimijoki River basin in Finland. Sci. Total Environ. 572, 1611–1621. doi:10.1016/j.scitotenv.2016.05.043

Razzolini, M.T.P., Lauretto, M. de S., Hachich, E.M., Sato, M.I.Z., Nardocci, A.C., 2016. Giardia and Cryptosporidium infection risk by simultaneous exposure to drinking water. Microb. Risk Anal. 4, 1–6. doi:10.1016/j.mran.2016.01.001

Rechenburg, A., Koch, C., Claßen, T., Kistemann, T., 2006. Impact of sewage treatment plants and combined sewer overflow basins on the microbiological quality of surface water. Water Sci. Technol. 54, 95. doi:10.2166/wst.2006.454

Rechenburg, A., Willkomm, M., Christoffels, E., Schreiber, C., Koch, C., Exner, M., Kistemann, T., 2009. Mikrobielle Fließgewässerbelastung aus Punkt- und diffusen Quellen in der Swist, in: Gewässerschutz Und Gewässergestaltung - von Der Idee Zur Konkreten Maßnahme. 27th Bochumer Workshop Bochum. Gesellschaft zur Förderung des Lehrstuhls für Siedlungswasserwirtschaft und Umwelttechnik an der Ruhr-Universität Bochum, Bochum, pp. 81–92.

Reder, K., Alcamo, J., Flörke, M., 2017. A sensitivity and uncertainty analysis of a continental-scale water quality model of pathogen pollution in African rivers. Ecol. Modell. 351, 129–139. doi:10.1016/j.ecolmodel.2017.02.008

Reder, K., Flörke, M., Alcamo, J., 2015. Modeling historical fecal coliform loadings to large European rivers and resulting in-stream concentrations. Environ. Model. Softw. 63, 251–263. doi:10.1016/j.envsoft.2014.10.001

Rhee, J.K., Kim, H.C., Park, B.K., 1998. Effect of Cryptosporidium baileyi infection on antibody response to sRBC in chickens. Korean J. Parasitol. 36, 33–36.

Rhee, J.K., Yook, S.Y., Park, B.K., 1995. Oocyst production and immunogenicity of Cryptosporidium baileyi in chickens and mallards. Korean J. Parasitol. 33, 45–54.

Rieux, A., Paraud, C., Pors, I., Chartier, C., 2014. Molecular characterization of Cryptosporidium isolates from beef calves under one month of age over three successive years in one herd in western France. Vet. Parasitol. 202, 171–179. doi:10.1016/j.vetpar.2014.03.004

Robertson, L.J., Björkman, C., Axén, C., Fayer, R., 2014. Cryptosporidiosis in Farmed Animals, in: Cacciò, S.M., Widmer, G. (Eds.), Cryptosporidium: Parasite and Disease. Springer, pp. 149–235. doi:10.1007/978-3-7091-1562-6

Robertson, L.J., Campbell, A.T., Smith, H. V, 1992. Survival of Cryptosporidium parvum oocysts under various environmental pressures. Appl. Environ. Microbiol. 58, 3494–3500.

Robinson, T.P., Wint, G.R.W., Conchedda, G., Van Boeckel, T.P., Ercoli, V., Palamara, E., Cinardi, G., D'Aietti, L., Hay, S.I., Gilbert, M., 2014. Mapping the global distribution of livestock. PLoS One 9, e96084. doi:10.1371/journal.pone.0096084

Robles-Morua, A., Mayer, A.S., Auer, M.T., Vivoni, E.R., 2012. Modeling riverine pathogen fate and transport in Mexican rural communities and associated public health implications. J. Environ. Manage. 113, 61–70. doi:10.1016/j.jenvman.2012.08.035

Rogers, D.J., Randolph, S.E., 2000. The global spread of malaria in a future, warmer world. Science (80-.). 289, 1763–1765. doi:10.1126/science.289.5485.1763

Rose, J.B., Epstein, P.R., Lipp, E.K., Sherman, B.H., Bernard, S.M., Patz, J.A., 2001. Climate variability and change in the United States: Potential impacts on water- and foodborne diseases caused by microbiologic agents. Environ. Health Perspect. 109, 211–221.

Ryan, U., Fayer, R., Xiao, L., 2014. Cryptosporidium species in humans and animals: current understanding and research needs. Parasitology 141, 1667–1685. doi:10.1017/S0031182014001085

Ryan, U.M., Bath, C., Robertson, I., Read, C., Elliot, A., Mcinnes, L., Traub, R., Besier, B., 2005. Sheep may not be an important zoonotic reservoir for Cryptosporidium and Giardia parasites. Appl. Environ. Microbiol. 71, 4992–4997. doi:10.1128/AEM.71.9.4992-4997.2005

Ryan, U.M., Samarasinghe, B., Read, C., Buddle, J.R., Robertson, I.D., Thompson, R.C.A., 2003. Identification of a novel Cryptosporidium genotype in pigs. Appl. Environ. Microbiol. 69, 3970–3974. doi:10.1128/AEM.69.7.3970-3974.2003

Sadeghi, A.M., Arnold, J.G., 2002. A SWAT/microbial sub-model for predicting pathogen loadings in surface and groundwater at watershed and basin scales, in: Total Maximum Daily Load (TMDL) Environmental Regulations: Proceedings of the March 11-13, 2002 Conference. pp. 56–63.

Safley, L.M., Casada, M.E., Woodbury, J.W., Roos, K.F., 1992. Global methane emissions from livestock and poultry manure. USEPA Washington, D. C.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity analysis in practice: a guide to assessing scientific models. John Wiley and Sons, Chichester, England.

Santín, M., Trout, J.M., 2007. No Title, in: Fayer, R., Xiao, L. (Eds.), Cryptosporidium and Cryptosporidiosis. CRC press - Taylor &Francis Group, p. 576.

Sattar, S., Chauret, C., Springthorpe, V., 1999. Giardia Cyst and Cryptosporidium Oocyst Survival in Watersheds and Factors Affecting Inactivation. AWWA Research Foundation & American Water Works Association.

Schijven, J., Bouwknegt, M., de Roda Husman, A.M., Rutjes, S., Sudre, B., Suk, J.E., Semenza, J.C., 2013. A Decision Support Tool to Compare Waterborne and Foodborne Infection and/or Illness Risks Associated with Climate Change. Risk Anal. 33, 2154–2167. doi:10.1111/risa.12077

Schijven, J., Derx, J., de Roda Husman, A.M., Blaschke, A.P., Farnleitner, A.H., 2015. QMRAcatch: Microbial Quality Simulation of Water Resources including Infection Risk Assessment. J. Environ. Qual. 0, 0. doi:10.2134/jeq2015.01.0048

Schijven, J.F., de Roda Husman, A.M., 2005. Effect of climate changes on waterborne disease in The Netherlands. Water Sci. Technol. 51, 79–87.
Schweizer, V.J., O'Neill, B.C., 2014. Systematic construction of global socioeconomic pathways using internally consistent element combinations. Clim. Change 122, 431–445. doi:10.1007/s10584-013-0908-z
Scully, N., Lean, D., 1994. The attenuation of ultrviolet radiation in temperate lakes. Arch. Hydrobiol. Beih. Ergebn. Limnol. 43, 135–144.
Seitzinger, S.P., Harrison, J.A., Dumont, E., Beusen, A.H.W., Bouwman, A.F., 2005. Sources and delivery of carbon, nitrogen, and phosphorus to the coastal zone: An overview of Global Nutrient Export from Watersheds (NEWS) models and their application. Global Biogeochem. Cycles 19. doi:10.1029/2005GB002606
Seitzinger, S.P., Mayorga, E., Bouwman, a. F., Kroeze, C., Beusen, a. H.W., Billen, G., Van Drecht, G., Dumont, E., Fekete, B.M., Garnier, J., Harrison, J. a., 2010. Global river nutrient export: A scenario analysis of past and future trends. Global Biogeochem. Cycles 24. doi:10.1029/2009GB003587
Seré, C., Steinfeld, H., 1996. World livestock production systems: Current status, issues and trends, Animal Production and Health Paper 127. Rome, Italy.
Shafiq, M.A.B., Maqbool, A., Khan, U.J., Lateef, M., Ijaz, M., 2015. Prevalence , Water Borne Transmission and Chemotherapy of Cryptosporidiosis in Small Ruminants. Pak. J. Zool. 47, 1715–1721.
Shirley, D.-A.T., Moonah, S.N., Kotloff, K.L., 2012. Burden of disease from cryptosporidiosis. Curr. Opin. Infect. Dis. 25, 555–63. doi:10.1097/QCO.0b013e328357e569
Siegfried, M., Koelmans, A.A., Besseling, E., Kroeze, C., 2017. Export of microplastics from land to sea. A modelling approach. Water Res.
Silverlås, C., Blanco-Penedo, I., 2013. Cryptosporidium spp. in calves and cows from organic and conventional dairy herds. Epidemiol. Infect. 141, 529–539. doi:10.1017/S0950268812000830
Silverlås, C., Emanuelson, U., de Verdier, K., Björkman, C., 2009. Prevalence and associated management factors of Cryptosporidium shedding in 50 Swedish dairy herds. Prev. Vet. Med. 90, 242–253. doi:10.1016/j.prevetmed.2009.04.006
Šlapeta, J., 2013. Cryptosporidiosis and Cryptosporidium species in animals and humans: A thirty colour rainbow? Int. J. Parasitol. 43, 957–970. doi:10.1016/j.ijpara.2013.07.005
Smeets, P.W.M.H., van Dijk, J.C., Stanfield, G., Rietveld, L.C., Medema, G.J., 2007. How can the UK statutory Cryptosporidium monitoring be used for Quantitative Risk Assessment of Cryptosporidium in drinking water? J. Water Health 5, 107–118. doi:Doi 10.2166/Wh.2007.140
Smith, J.T., Wright, S.M., Cross, M. a., Monte, L., Kudelsky, A. V., Saxén, R., Vakulovsky, S.M., Timms, D.N., 2004. Global Analysis of the Riverine Transport of 90Sr and 137Cs. Environ. Sci. Technol. 38, 850–857. doi:10.1021/es0300463
Smith, K., Woodward, A., Campbell-Lendrum, D., Chadee, D., Honda, Y., Liu, Q., Olwoch, J., Revich, B., Sauerborn, R., 2014. Human health: impacts, adaptation, and co-benefits, in: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 709–754.
Smith, M., Thompson, K., 2001. Cryptosporidium. The analytical challenge. Royal Society of Chemistry, Cambridge, UK.
Smith, R.P., Clifton-Hadley, F.A., Cheney, T., Giles, M., 2014. Prevalence and molecular typing of Cryptosporidium in dairy cattle in England and Wales and examination of potential on-farm transmission routes. Vet. Parasitol. 204, 111–9. doi:10.1016/j.vetpar.2014.05.022
Smith, S. V, Swaney, D.P., Talaue-McManus, L., Bartley, J.D., Sandhei, P.T., Mclaughlin, C.J., Dupra, V.C., Crossland, C.J., Buddemeier, R.W., Maxwell, B.A., Wulff, F., 2003. Humans, hydrology, and the distribution of inorganic nutrient loading to the ocean. Bioscience 53, 235–245.
Sokolova, E., Åström, J., Pettersson, T.J.R., Bergstedt, O., Hermansson, M., 2012. Estimation of pathogen concentrations in a drinking water source using hydrodynamic modelling and microbial source tracking. J. Water Health 10, 358–370. doi:10.2166/wh.2012.183
Squire, S.A., Ryan, U., 2017. Cryptosporidium and Giardia in Africa: current and future challenges. Parasit. Vectors 10, 195. doi:10.1186/s13071-017-2111-y
Sréter, T., Varga, I., Békési, L., 1995. Age-Dependent Resistance to Cryptosporidium baileyi Infection in Chickens. J. Parasitol. 81, 827–829.
Starkey, S.R., Wade, S.E., Schaaf, S., Mohammed, H.O., 2005. Incidence of Cryptosporidium parvum in the dairy cattle population in a New York City Watershed. Vet. Parasitol. 131, 197–205. doi:10.1016/j.vetpar.2005.04.040
Starkey, S.R., White, M.E., Mohammed, H.O., 2007. Cryptosporidium and dairy cattle in the Catskill/Delaware watershed: A quantitative risk assessment. Risk Anal. 27, 1469–1485. doi:10.1111/j.1539-6924.2007.00982.x
Steffen, W., Sanderson, A., Tyson, P., Jäger, J., Matson, P., Moore, B., Oldfield, F., Richardson, K., Schellnhuber, H., Turner, B., Wasson, R., 2004. Global Change and the Earth System: A Planet Under Pressure - Executive Summary. IGBP Secretariat, Royal Swedish Academy of Sciences.
Sterk, A., Schijven, J., Maria de Roda Husman, A., de Nijs, T., de Roda Husman, A.M., de Nijs, T., 2016. Effect of climate

change on runoff of Campylobacter and Cryptosporidium from land to surface water. Water Res. 95, 90–102. doi:10.1016/j.watres.2016.03.005
Stott, R., Davies-Colley, R., Nagels, J., Donnison, A., Ross, C., Muirhead, R., 2011. Differential behaviour of Escherichia coli and Campylobacter spp. in a stream draining dairy pasture. J. Water Health 9, 59–69. doi:10.2166/wh.2010.061
Strokal, M., Kroeze, C., Wang, M., Bai, Z., Ma, L., 2016. The MARINA model (Model to Assess River Inputs of Nutrients to seAs): Model description and results for China. Sci. Total Environ. 562, 869–888. doi:10.1016/j.scitotenv.2016.04.071
Swart, R.J., Raskin, P., Robinson, J., 2004. The problem of the future: sustainability science and scenario analysis. Glob. Environ. Chang. 14, 137–146. doi:10.1016/j.gloenvcha.2003.10.002
Sweeny, J.P.A., Ryan, U.M., Robertson, I.D., Jacobson, C., 2011. Cryptosporidium and Giardia associated with reduced lamb carcase productivity. Vet. Parasitol. 182, 127–139. doi:10.1016/j.vetpar.2011.05.050
Syvitski, J.P.M., Vörösmarty, C.J., Kettner, A.J., Green, P., 2005. Impact of humans on the flux of terrestrial sediment to the global coastal ocean. Science (80-.). 308, 376–380. doi:10.1126/science.1109454
Tang, J., McDonald, S., Peng, X., Samadder, S.R., Murphy, T.M., Holden, N.M., 2011. Modelling Cryptosporidium oocysts transport in small ungauged agricultural catchments. Water Res. 45, 3665–3680. doi:10.1016/j.watres.2011.04.013
Tate, K.W., Pereira, M.D.G.C., Atwill, E.R., 2004. Efficacy of vegetated buffer strips for retaining Cryptosporidium parvum. J. Environ. Qual. 33, 2243–2251. doi:10.2134/jeq2004.2243
Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An overview of CMIP5 and the experiment design. Bull. Am. Meteorol. Soc. 93, 485–498. doi:10.1175/BAMS-D-11-00094.1
Teunis, P.F.M., Chappell, C.L., Okhuysen, P.C., 2002a. Cryptosporidium dose-response studies: Variation between hosts. Risk Anal. 22, 475–485. doi:10.1111/0272-4332.00046
Teunis, P.F.M., Chappell, C.L., Okhuysen, P.C., 2002b. Cryptosporidium Dose Response Studies: Variation Between Isolates. Risk Anal. 22, 175–185. doi:10.1111/0272-4332.00014
Teunis, P.F.M., Medema, G.J., Kruidenier, L., Havelaar, A.H., 1997. Assessment of the risk of infection by Cryptosporidium or Giardia in drinking water from a surface water source. Water Res. 31, 1333–1346.
Thomann, R. V., Mueller, J.A., 1987. Principles of surface water quality modeling and control. Harper & Row Publishers, New York.
Thurston-Enriquez, J.A., Gilley, J.E., Eghball, B., 2005. Microbial quality of runoff following land application of cattle manure and swine slurry. J. Water Health 3, 157–171.
Tian, Y.Q., Gong, P., Radke, J.D., Scarborough, J., 2002. Spatial and temporal modeling of microbial contaminants on grazing farmlands. J. Environ. Qual. 31, 860–869.
Till, D., McBride, G., Ball, A., Taylor, K., Pyle, E., 2008. Large-scale freshwater microbiological study: Rationale, results and risks. J. Water Health 6, 443–460. doi:10.2166/wh.2008.071
Trask, J.R., Kalita, P.K., Kuhlenschmidt, M.S., Smith, R.D., Funk, T.L., 2004. Overland and Near-Surface Transport of Cryptosporidium parvum from Vegetated and Nonvegetated Surfaces. J. Environ. Qual. 33, 984–993.
Tryland, I., Robertson, L., Blankenberg, A.G.B., Lindholm, M., Rohrlack, T., Liltved, H., 2011. Impact of rainfall on microbial contamination of surface water. Int. J. Clim. Chang. Strateg. Manag. 3, 361–373. doi:Doi 10.1108/17568691111175650
Tůmová, E., Skřivan, M., Marounek, M., Pavlásek, I., Ledvinka, Z., 2002. Performance and oocyst shedding in broiler chickens orally infected with Cryptosporidium baileyi and Cryptosporidium meleagridis. Avian Dis. 46, 203–207. doi:Doi 10.1637/0005-2086(2002)046[0203:Paosib]2.0.Co;2
Tumwebaze, I.K., Orach, C.G., Niwagaba, C., Luthi, C., Mosler, H.-J., 2013. Sanitation facilities in Kampala slums, Uganda: users' satisfaction and determinant factors. Int. J. Environ. Health Res. 23, 191–204. doi:10.1080/09603123.2012.713095
UN-HABITAT, 2012. Global Urban Observatory - Urban Info Database.
UN Millennium Project, 2005. A home in the city: UN Millennium Project Task Force on improving the lives of slum dwellers.
UN Water, 2015. Wastewater Management - A UN-Water Analytical Brief.
UNEP, 2016. A Snapshot of the World's Water Quality: Towards a global assessment.
UNEP/UN-HABITAT, 2010. Sick Water? The central role of wastewater management in sustainable development. United Nations Environment Programme, UN-HABITAT, Norway.
UNICEF/WHO, 2009. Diarrhoea: why children are still dying and what can be done.
United Nations, 2014. World Urbanization Prospects: the 2014 revision. New York.
United Nations DESA/Population Division, 2017. World Population Prospects 2017 [WWW Document]. URL https://esa.un.org/unpd/wpp/ (accessed 9.26.17).
United Nations General Assembly, 2015. Transforming our World: The 2030 Agenda for Sustainable Development, A/RES/70/1. doi:10.1007/s13398-014-0173-7.2
United Nations Inter-agency Group for Child Mortality Estimation, 2015. Levels & Trends in Child Mortality - Report 2015.

New York. doi:10.1371/journal.pone.0144443
van Drecht, G., Bouwman, A.F., Harrison, J., Knoop, J.M., 2009. Global nitrogen and phosphate in urban wastewater for the period 1970 to 2050. Global Biogeochem. Cycles 23, GB0A03. doi:10.1029/2009gb003458
van Drecht, G., Bouwman, A.F., Knoop, J.M., Beusen, A.H.W., Meinardi, C.R., 2003. Global modeling of the fate of nitrogen from point and nonpoint sources in soils, groundwater, and surface water. Glob. Biogeochem. Cycles 17, 1115. doi:10.1029/2003gb002060
van Herk, F.H., McAllister, T. a., Cockwill, C.L., Guselle, N., Larney, F.J., Miller, J.J., Olson, M.E., 2004. Inactivation of Giardia Cysts and Cryptosporidium Oocysts in Beef Feedlot Manure By Thermophilic Windrow Composting. Compost Sci. Util. 12, 235–241. doi:10.1080/1065657X.2004.10702188
van Vliet, M.T.H., Yearsley, J.R., Franssen, W.H.P., Ludwig, F., Haddeland, I., Lettenmaier, D.P., Kabat, P., 2012. Coupled daily streamflow and water temperature modelling in large river basins. Hydrol. Earth Syst. Sci. 16, 4303–4321. doi:10.5194/hess-16-4303-2012
van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J., Rose, S.K., 2011. The representative concentration pathways: an overview. Clim. Change 109, 5–31. doi:10.1007/s10584-011-0148-z
Varga, I., Sréter, T., Békési, L., 1995. Quantitative method to assess Cryptosporidium oocyst shedding in the chicken model. Parasitol. Res. 81, 262–264.
Venu, R., Latha, B.R., Basith, S.A., Sreekumar, C., Raj, G.D., Raman, M., 2013. Comparative evaluation of conventional diagnostic methods for detection of bovine cryptosporidiosis. Indian J. Anim. Sci. 83, 124–128.
Vijay, R., Sardar, V.K., Dhage, S.S., Kelkar, P.S., Gupta, A., 2010. Hydrodynamic assessment of sewage impact on water quality of Malad Creek, Mumbai, India. Environ. Monit. Assess. 165, 559–571. doi:10.1007/s10661-009-0967-9
Vital, M., Hammes, F., Egli, T., 2008. Escherichia coli O157 can grow in natural freshwater at low carbon concentrations. Environ. Microbiol. 10, 2387–2396.
Vörösmarty, C., Fekete, B., Meybeck, M., Lammers, R., 2000. Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. Global Biogeochem. Cycles 14, 599–621.
Vörösmarty, C.J., Fekete, B.M., Meybeck, M., Lammers, R.B., 2000. Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution. J. Hydrol. 237, 17–39.
Walker, F.R., Stedinger, J.R., 1999. Fate and transport model of Cryptosporidium. J. Environ. Eng. 124, 325–333.
Wang, G., Zhao, T., Doyle, M.P., 1996. Fate of enterohemorrhagic Escherichia coli O157:H7 in bovine feces. Appl. Environ. Microbiol. 62, 2567–2570.
Wang, R., Li, G., Cui, B., Huang, J., Cui, Z., Zhang, S., Dong, H., Yue, D., Zhang, L., Ning, C., Wang, M., 2014. Prevalence, molecular characterization and zoonotic potential of Cryptosporidium spp. in goats in Henan and Chongqing, China. Exp. Parasitol. 142, 11–16. doi:10.1016/j.exppara.2014.04.001
Wania, F., Mackay, D., 1995. A global distribution model for persistent organic chemicals. Sci. Total Environ. 161, 211–232. doi:10.1016/0048-9697(95)04358-8
Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., Schewe, J., 2014. The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework. Proc. Natl. Acad. Sci. 111, 3228–3232. doi:10.1073/pnas.1312330110
WASHplus project of USAID, 2010. Bangladesh - sanitation deserves highest priority [WWW Document]. URL https://sanitationupdates.wordpress.com/2010/01/08/bangladesh-sanitation-deserves-highest-priority/ (accessed 7.30.15).
Webster, J., 1975. Analysis of potassium and calcium dynamics in stream ecosystems on three southern Appalachian watersheds of contrasting vegetation. University of Georgia, Athens.
Weedon, G.P., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J.C., Bellouin, N., Boucher, O., Best, M., 2011. Creation of the WATCH Forcing Data and Its Use to Assess Global and Regional Reference Crop Evaporation over Land during the Twentieth Century. J. Hydrometeorol. 12, 823–848. doi:10.1175/2011JHM1369.1
Wells, B., Thomson, S., Ensor, H., Innes, E.A., Katzer, F., 2016. Development of a sensitive method to extract and detect low numbers of Cryptosporidium oocysts from adult cattle faecal samples. Vet. Parasitol. 227, 26–29. doi:10.1016/j.vetpar.2016.07.018
WHO, 2008. The global burden of disease: 2004 update. WHO (World Health Organisation), Geneva.
WHO/ UNICEF Joint Monitoring Program, 2017. WASH database [WWW Document]. URL https://washdata.org/ (accessed 10.17.17).
WHO/UNICEF JMP, 2017. Progress on Drinking Water, Sanitation and Hygiene 2017 Update and SDG Baselines. doi:10.1111 / tmi.12329
WHO/UNICEF JMP, 2014. Progress on Drinking Water and Sanitation: 2014 update.
WHO/UNICEF JMP, 2000. Global Water Supply and Sanitation Assessment 2000 Report.
Wickham, J.D., Nash, M.S., Wade, T.G., Currey, L., 2006. Statewide empirical modeling of bacterial contamination of surface waters. J. Am. Water Resour. Assoc. 42, 582–591. doi:10.1111/j.1752-1688.2006.tb04477.x

Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res. 90, 8995. doi:10.1029/JC090iC05p08995

Wollheim, W.M., Vörösmarty, C.J., Bouwman, A.F., Green, P., Harrison, J., Linder, E., Peterson, B.J., Seitzinger, S.P., Syvitski, J.P.M., 2008. Global N removal by freshwater aquatic systems using a spatially distributed, within-basin approach. Global Biogeochem. Cycles 22. doi:10.1029/2007GB002963

World Health Organization, 2017a. Guidelines for drinking-water quality - fourth edition incorporating the first addendum. Geneva.

World Health Organization, 2017b. Potable Reuse: Guidance for producing safe drinking-water. World Health Organization, Geneva, Switzerland.

World Health Organization, 2013. Rotavirus vaccines WHO position paper - January 2013. Wkly. Epidemiol. Rec. 88, 49–64.

Wu, J., Rees, P., Dorner, S., 2011. Variability of E. coli density and sources in an urban watershed. J. Water Health 9, 94–106. doi:10.2166/wh.2010.063

Wu, J., Rees, P., Storrer, S., Alderisio, K., Dorner, S.M., 2009. Fate and transport modeling of potential pathogens: the contribution from sediments. J. Am. Water Resour. Assoc. 45, 35–44.

Xiao, S., An, W., Chen, Z., Zhang, D., Yu, J., Yang, M., 2012. The burden of drinking water-associated cryptosporidiosis in China: The large contribution of the immunodeficient population identified by quantitative microbial risk assessment. Water Res. 46, 4272–4280. doi:10.1016/j.watres.2012.05.012

Yang, R., Jacobson, C., Gordon, C., Ryan, U., 2009. Prevalence and molecular characterisation of Cryptosporidium and Giardia species in pre-weaned sheep in Australia. Vet. Parasitol. 161, 19–24. doi:10.1016/j.vetpar.2008.12.021

Yearsley, J.R., 2009. A semi-Lagrangian water temperature model for advection-dominated river systems. Water Resour. Res. 45, 1–19. doi:10.1029/2008WR007629

Yeghiazarian, L., Samorodnitsky, G., Montemagno, C.D., 2009. A Poisson random field model of pathogen transport in surface water. Water Resour. Res. 45. doi:Artn W11415Doi 10.1029/2009wr007896

Yui, T., Shibahara, T., Kon, M., Yamamoto, N., Kameda, M., Taniyama, H., 2014. Epidemiological studies on intestinal protozoa in pigs in Saitama, Japan. Japan Agric. Res. Q. 48, 87–93. doi:10.6090/jarq.48.87

Zambrano, L.D., Levy, K., Menezes, N.P., Freeman, M.C., 2014. Human diarrhea infections associated with domestic animal husbandry: A systematic review and meta-analysis. Trans. R. Soc. Trop. Med. Hyg. 108, 313–325. doi:10.1093/trstmh/tru056

Zambriski, J.A., Nydam, D. V., Bowman, D.D., Bellosa, M.L., Burton, A.J., Linden, T.C., Liotta, J.L., Ollivett, T.L., Tondello-Martins, L., Mohammed, H.O., 2013. Description of fecal shedding of Cryptosporidium parvum oocysts in experimentally challenged dairy calves. Parasitol. Res. 112, 1247–1254. doi:10.1007/s00436-012-3258-2

Zambriski, J.A., Nydam, D. V, Wilcox, Z.J., Bowman, D.D., Mohammed, H.O., Liotta, J.L., 2013. Cryptosporidium parvum: Determination of ID50 and the dose-response relationship in experimentally challenged dairy calves. Vet. Parasitol. 197, 104–112. doi:10.1016/j.vetpar.2013.04.022

Zhao, W., Wang, R., Zhang, W., Liu, A., Cao, J., Shen, Y., Yang, F., Zhang, L., 2014. MLST subtypes and population genetic structure of Cryptosporidium andersoni from dairy cattle and beef cattle in Northeastern China's Heilongjiang Province. PLoS One 9. doi:10.1371/journal.pone.0102006

Supplementary materials

This section contains supplementary materials for Chapters 2, 3, 5, 6 and 7.

Supplementary material for Chapter 2

Tables S2.1 and S2.2

Table S2.1 Overview of pathogen models included in the review and their main characteristics

Author(s)	Year published	Pathogen spp.	Region	Spatial scale	Spatial step	Temporal scale	Temporal step	Static vs dynamic	Output	Comments
Cho et al.	2010	Fecal indicator bacteria	Gwangju Creek, Korea	Watershed	Several points along stream	Weather event modelling. Data from wet and dry events 2006-2009.	Hourly	Dynamic	Fecal indicator concentrations over time on different locations in the stream during dry and wet weather events.	Model is mainly intended to investigate meteorological effects on fecal indicator bacteria concentrations.
Collins and Rutherford	2004	E. coli	Upper Managaotama catchment, Whatawhata, New Zealand	Watershed	Gridded, 10x10m. 37 subcatchments	Data from one year, November 2000 to October 2001	Daily	Dynamic	Mean daily E. coli concentrations over time in streams draining pastoral land.	Based on Tian et al. (2002). Mass balance formulation, specifically designed for application to steep grazed hill-country catchments.
Dorner et al.	2006	E. coli, Cryptosporidium spp.,Giardia spp.,Campylobacter spp, and E. coli O157:H7	Canagagigue Creek Watershed, Ontario, Canada	Watershed	Gridded, 1x1 km	Data from two years, June 2002 - May 2004.	Continuous	Dynamic	Pathogen concentrations in surface water over time	The model considers transport as a result of overland flow, subsurface flow to tile drainage systems, and in-stream routing. Contains a probabilistic pathogen loading model.
Ferguson et al.	2007	Cryptosporidium, Giardia and E. coli	Wingecarribee catchment, Sydney, Australia	Watershed	52 subcatchments.	Weather event modelling, wet, dry and intermediate.	Not specified	Static	Total microorganism loads generated and total loads downstream (export) per subcatchment for different weather event scenarios	Mass-balance approach
Fraser et al.	1998	Fecal coliforms	Saw Kill watershed, United States	Watershed	Gridded, 30x30m. 12 subcatchments.	Data from two months, June - July 1996.	Static averages for a given time period	Static	Fecal coliform delivery from agricultural lands to stream	GIS-based transport model (SEDMOD), models delivery from land to the surface water. Stream routing and climate not included, steady state conditions assumed.

Haydon and Deletic	2006	Generic, tested for E. coli	Three catchments in Australia (O'Shannassy, Myponga and Aldgate)	Watershed	Lumped output	Weather event modelling. Data from the year 2002.	Daily or hourly timestep possible	Dynamic	E. coli concentrations over time, total event load and peak concentration estimates	Two models. EG model: surface & subsurface flow coupled to hydrological model (SimHyd). ASP model: only surface flow coupled to a stormflow-baseflow separation model, no in-stream routing.
Hipsey et al.	2008	Generic, applicable to different species (protozoan, viral, bacterial)	Generic, tested for two Australian and one Brazilian reservoir	Reservoir	Several points in reservoir	Weather event modelling. A few days in the year 2003	Continuous	Dynamic	Concentrations over time after an inflow event	Uses an empirically determined or assumed input load, models dispersion.
Hofstra et al.	2013	Cryptosporidium	Global	Global	Gridded, 0.5x0.5 degree	Simulation for one year, 2000.	Static averages for a given time period	Static	Cryptosporidium emission to surface waters for the year 2002	First exploration of a global scale waterborne pathogen model. Model is not calibrated with observational data, but some comparison with literature is made. Only emissions, no concentrations in surface water.
Medema and Schijven	2001	Cryptosporidium and Giardia	The Netherlands	Country	13 sites spread over the country	Data from one year, May 1994 - May 1995	Daily	Dynamic	Emission model: annual load emitted per person. Dispersion model: daily estimates of average, min or max conc. in surface water	Based on two models developed for chemical polllutants, one for emission (PROMISE) and one for dispersion (WATNAT)
Peterse n et al.	2009	Fecal indicator bacteria, specifically E. coli	Buffalo Bayou watershed in Houston, Texas, United States	Watershed	Subwatersheds	One day, wet, dry or intermediate weather	Static averages for a given time period	Static	Spatial E. coli loads per day to an urban watershed in dry, wet or intermediate weather.	The model is intended for application in a predominantly urban area.

Reder et al.	2015	Fecal coliforms	Europe	Continent	Gridded, 5x5 arc minutes	1990-1995	Monthly	Static	Loads and in-stream concentrations of fecal coliforms in European rivers	The model is an application of the large scale waterquality model WorldQual, that has previously been applied to biological oxygen demand (BOD), total N, total P and total dissolved solids.
Robles-Morua et al.	2012	E. coli	Sonora River, Mexico	Watershed	River reaches of approx. 0.2 km	Daily	0.04 hour	Dynamic	E. coli concentration as function of distance downstream of wastewater discharge and compliance with norms	Uses empirically determined input load from 2 wastewater point sources. Employs EPA's QUAL2K water quality software in combination with hydological model tRIBS.
Sadeghi and Arnold	2012	Fecal coliform bacteria, subdivided in persistent and less persistent spp.	Has been applied to various watersheds worldwide.	Watershed	Subcatchments & hydrological response units	Can be applied for modeling continuous long time series.	Daily	Dynamic	Bacterial concentrations over time.	SWAT. Used for many purposes, microbial modeling only a minor application. Latest version 2012. Applied all over the world, several modified versions exist.
Tang et al.	2011	Cryptosporidium	North Leinster and South Leinster, catchments Ireland	Watershed	Subcatchments & hydrological response units	Data from March 2007 to June 2008.	Daily	Dynamic	Cryptosporidium concentrations over time per catchment.	SWAT2005
Tian et al.	2002	E. coli	Mangatoama catchment, Whatawhata, New Zealand	Watershed	Gridded, 400 m2 grids	Data from one year, November 1998 - November 1999	Daily	Dynamic	Mean daily E. coli concentrations over time in streams	Based on a mass balance formulation. Applied to a hilly catchment used exclusively for animal grazing.
Walker and Stedinger	1999	Cryptosporidium	Catskill-Delaware watershed, New York	Watershed	6 subwatershed-reservoi	Simulation of 32 years, 1960-1991	Daily timestep of some	Dynamic	Average monthly oocyst loads from dairy farming and sewage for each of the 12 calender months, and	Model is not calibrated with observational data, although some comparison with literature is made.

			City, United States		r systems		input, output per month		oocyst concentration in water used for drinking water.	
Wu et al.	2009	E. coli, extension to Cryptosporidium, Giardia, Campylobacter, E. coli O157:H7	Blackstone River Watershed, Massachusetts, United States	Watershed	Gridded, 1x1 km	Data from wet and dry events in 2005 -2006.	Continuous	Dynamic	E. coli concentrations over time	Based on the model by Dorner et al. (2006), but upgraded to include sedimentation/ resuspension processes.
Yeghiazarian et al.	2009	Calibrated for C. parvum, extension to other pathogens possible	Applied to a theoretical hillslope	Field or watershed	x meters from upstream start point	Weather event modelling	Continuous	Dynamic	Concentrations over time of microorganisms in different states (attached vs free), as well as risk of exceeding a maximum conc.	Stochastic process-based model, risk assessment is included.

Table S2.2 Overview of global nutrient models included in the review and their main characteristics

Author(s)	Year published	Nutrient type(s)	Region	Spatial information	Temporal scale	Static vs dynamic	Output	Comments
Bouwman et al.	2005	N	Global	Gridded, 0.5 x 0.5 degree and country data	Three years: 1970, 1995 and 2030	Static	Global river N flux to coastal marine systems. Results are mostly presented for world regions or per river basin.	Partly process-based.
Bouwman et al.	2009	N and P	Global	Gridded, 0.5 x 0.5 degree and country data	Four years: 1970, 2000,2030 and 2050	Static	N and P soil balances by world region, for 1970 and 2000 and for different scenarios in 2030 and 2050, mostly presented at continental scale.	Uses the four Millennium Ecosystem Assessment scenarios for scenario analysis.
Bouwman et al.	2013	N and P	Global	Gridded, 0.5 x 0.5 degree and country data	Four years: 1900, 1950, 2000 and 2050	Static	Changes in N and P soil budgets for 1900, 1950, 2000, and projections of N and P soil budgets under different scenarios in 2050, mostly presented at continental scale.	Uses the International Assessment of Agricultural Knowledge, Science and Technology for Development IAASTD baseline scenario for the 2000-2050 period, with a number of modifications for different scenarios
Boyer et al.	2006	N (reactive N)	Global	Data from 39 watersheds. Gridded data, country data and aggregated empirical data	Representing conditions during mid 1990s	Static	Annual riverine N loads to coastal zones and inland waters from the continents representing conditions during the mid 1990s.	Follows the approach of Howarth et al. (1996) but extended to the global scale.
Caraco & Cole	1999	Nitrate (NO3)	Global	Data from 35 large rivers.	Based on estimates from the 1980-1990s	Static	Mean annual NO3 export for 35 large rivers	A simple model that relates NO3 export to point and nonpoint source N loads from chemical fertilizers and deposition. Computes export as fraction of average loading per area of different sources. No extrapolation to other rivers.
Galloway et al.	2004	N	Global	Combines gridded data, country data and some empirical data.	Three years: 1860, 1990s, 2050	Static	Global nitrogen budgets for 1860, the early 1990s and 2050. Export estimates are also presented per continent.	Global N budgets. Based on Boyer et al. 2006 and Howarth et al. 1996.

Green et al.	2004	Reactive N (total, dissolved inorganic and total organic)	Global	Gridded, 0.5 x 0.5 degree	Pre-industrial to contemporary (mid 1990s).	Static	Reactive N loading to the landmass and subsequent riverine N fluxes under anthropogenic disturbance	Semi-empirical. Mass balance model for N loading to the landmass.
Howarth et al.	1996	N and P	North Atlantic Ocean	Data from 14 regions around the North Atlantic, combining most large river basins.	Based on estimates from the 1980-1990s	Static	Estimates of average annual total N and total P fluxes in rivers to the North Atlantic ocean. Output is aggregated per region.	Lumped export model, focussing on different anthropogenic N sources.
Mayorga et al.	2010	N, P and C in multiple forms. Also dissolved silica (Beusen et al. 2009)	Global	Gridded, 0.5 x 0.5 degree	Four years: 1970, 2000, 2030, 2050,	Static	Annual average nutrient yield per basin and global river nutrient export estimates for 1970, 2000, 2030 and 2050 . Output for different elements, forms and sources of nutrients, for each of the included basins, aggregated by region, continent or globally.	Semi-empirical. NEWS 2. More unified formulation compared to NEWS 1. For the scenario analysis with NEWS2, see the paper by Seitzinger et al. (2010). Builds on Van Drecht et al. 2009 and Bouwman et al. 2009. Scenario analysis with Millennium Ecosystem Assessment Scenarios.
Morée et al.	2013	N and P	Global	Gridded, 0.5 x 0.5 degree and country data	1900-2000, yearly timestep	Static	Global country-scale urban N and P flows from humans, animals, and industries and their waste discharges to surface water, and urban waste recycling in agriculture, mostly presented at continental scale.	Only urban emissions, taking into account nutrient flows from human waste, industrial and animal emissions (only equidae: horses, mules, donkeys). Flows have three possible fates: surface water, agricultural land and other.
Seitzinger et al.	2005	N, P and C in multiple forms: dissolved inorganic, dissolved organic and particulate	Global	Gridded, 0.5 x 0.5 degree	Representing conditions around 1995.	Static	Annual average nutrient yield per basin and global river nutrient export estimates for 1995. Output for different elements, forms and sources of nutrients, for each of the included basins, aggregated by region, continent or globally.	Global NEWS models, first version. Estimates export from 5761 watersheds globally as a function of land use, nutrient inputs, hydrology and other factors. It is a set of several models used for different nutrient forms. NEWS-PNU is multiple regression, all others are semi-process-based.

Smith et al.	2003	N and P (dissolved inorganic N and dissolved inorganic P)	Global	Data from 165 large rivers. Some data gridded, 0.5 x 0.5 degree.	Representing conditions during mid 1990s	Static	Average annual DIN and DIP fluxes, distributed globally around the world coastlines using data from 165 sites.	Spatially explicit regression model. Three variables proved to be strongly related to DIN and DIP loads: runoff, basin area and population. Model extrapolation to predictions for the global coastline.
Van Drecht et al.	2003	N	Global	Gridded, 0.5 x 0.5 degree and country data	Reference year 1995	Static	Global river N flux to coastal marine systems. Results are mostly presented for world regions or per river basin.	Partly process-based.
Van Drecht et al.	2009	N and P	Global	Gridded, 0.5 x 0.5 degree and country data	Four years: 1970, 2000,2030 and 2050	Static	Global N and P emissions from sewage for the years 1970 and 2000, and for different scenarios in 2030 and 2050, mostly presented at continental scale.	Uses the four Millennium Ecosystem Assessment scenarios for scenario analysis.
Wollheim et al.	2008	N	Global	Gridded, 0.5 x 0.5 degree	Representing conditions during mid 1990s	Static	Mean annual aquatic total N removal for the mid-1990s period	Uses the terrestrial model by Bouwman (2005) and Van Drecht (2003) as input for Total N river loading

Supplementary material for Chapter 3

Tables S3.1 to S3.3.

Table S3.1. Model parameter values for the standard model run and the sensitivity analysis for 2010. The sensitivity analysis is done with pairs of parameters that can take three different values each (see Table S3.2 for results). 'Standard value' indicates that the same value as in the standard model run is taken. Estimates for oocyst excretion and the removal efficiencies of oocysts by sewage treatment are based on a literature study by Hofstra et al. (2013). The infection rate and the percentage of the connected population that receives treatment were estimated from literature (Section 3.2.3).

	Model parameter value	Sensitivity		
		Value1	Value 2	Value 3
Oocyst excretion (oocysts / infected person)	1.00E+09	low (-1log) 1.00E+08	standard value	high (+1log) 1.00E+10
Infection rate (% of population infected)	5	low (half) 2.5	standard value	high (double) 10
Oocyst removal efficiencies sewage treatment (%)		low	standard value	high
Primary treatment	10	4	standard value	90
Secondary treatment	50	28	standard value	100
Tertiary treatment	95	40	standard value	99
Percentage of connected population with primary treatment (%)	20	standard value	0	0
Percentage of connected population with secondary treatment (%)	0	standard value	20	standard value
Percentage of connected population with tertiary treatment (%)	0	standard value	standard value	20
Urban no facilities that reaches surface water (%)	100	standard value	75	50
Rural no facilities that reaches surface water (%)	2.5	0	25	50
Urban septic tank and pit latrine content that reaches surface water (%)	0	standard value	25	50
Rural septic tank and pit latrine content that reaches surface water (%)	0	standard value	25	50

Table S3.2. Model parameter values per state. The data on urban and rural populations per state are from the Bangladesh Bureau of Statistics and the Census of India (Bangladesh Bureau of Statistics, 2011; Census of India, 2011). The sanitation numbers in this table are fractions of the population that use sanitation that fall into the emission categories 'connected', 'direct', 'diffuse' and 'non-source', as described in Section 3.2.2. Data on the usage of different sanitation types are based on the most recent estimates of the DHS Program, this is for Bangladesh the year 2011 and for India 2005-2006 (International Institute for Population Sciences (IIPS) and Macro International, 2007; National Institute of Population Research and Training (NIPORT) et al., 2013). Indian regions for which no sanitation data were available have for modelling purposes been assigned the population-weighted average Indian sanitation.

Region	Administrative level	Population			Sanitation						
					Urban			Rural			
		Total	Urban	Rural	Connected	Direct	Non-source	Connected	Direct	Diffuse	Non-source
INDIA TOTAL	country	1.21E+09	3.77E+08	8.33E+08	0.21	0.08	0.70	0.01	0.00	0.46	0.53
Andaman and Nicobar Islands	Union territory	379944	135533	244411	NA	NA	NA	NA	NA	NA	NA
Andhra Pradesh	state	84665533	28353745	56311788	0.48	0.15	0.37	0.02	0.00	0.73	0.25
Arunachal Pradesh	state	1382611	313446	1069165	0.09	0.12	0.80	0.03	0.00	0.32	0.65
Assam	state	31169272	4388756	26780516	0.02	0.03	0.95	0.00	0.00	0.29	0.71
Bihar	state	1.04E+08	11729609	92075028	0.08	0.28	0.64	0.01	0.00	0.84	0.16
Chandigarh	Union territory	1054686	1025682	29004	NA	NA	NA	NA	NA	NA	NA
Chhattisgarh	state	25540196	5936538	19603658	0.03	0.36	0.61	0.00	0.00	0.94	0.06
Dadra and Nagar Haveli	Union territory	342853	159829	183024	NA	NA	NA	NA	NA	NA	NA
Daman and Diu	Union territory	242911	182580	60331	NA	NA	NA	NA	NA	NA	NA
Delhi	Union territory	16753235	16333916	419319	0.67	0.23	0.10	0.17	0.00	0.35	0.49
Goa	state	1457723	906309	551414	0.02	0.17	0.81	0.02	0.00	0.40	0.57
Gujarat	state	60383628	25712811	34670817	0.61	0.14	0.25	0.05	0.00	0.70	0.25

Haryana	state	25353081	8821588	16531493	0.52	0.12	0.36	0.01	0.00	0.65	0.34
Himachal Pradesh	state	6856509	688704	6167805	0.26	0.11	0.63	0.01	0.00	0.60	0.39
Jammu and Kashmir	state	12548926	3414106	9134820	0.14	0.26	0.59	0.02	0.00	0.51	0.47
Jharkhand	state	32966238	7929292	25036946	0.06	0.29	0.64	0.00	0.00	0.95	0.05
Karnataka	state	61130704	23578175	37552529	0.23	0.20	0.57	0.00	0.00	0.79	0.21
Kerala	state	33387677	15932171	17455506	0.01	0.02	0.97	0.02	0.00	0.05	0.93
Lakshadweep	Union territory	64429	50308	14121	NA	NA	NA	NA	NA	NA	NA
Madhya Pradesh	state	72597565	20059666	52537899	0.37	0.20	0.43	0.00	0.00	0.90	0.09
Maharashtra	state	1.12E+08	50827531	61545441	0.72	0.12	0.17	0.02	0.00	0.80	0.19
Manipur	state	2721756	822132	1899624	0.01	0.02	0.98	0.00	0.00	0.09	0.91
Meghalaya	state	2964007	595036	2368971	0.01	0.03	0.96	0.00	0.00	0.40	0.59
Mizoram	state	1091014	561977	529017	0.00	0.00	1.00	0.00	0.00	0.05	0.95
Nagaland	state	1980602	573741	1406861	0.01	0.06	0.93	0.01	0.00	0.23	0.76
Odisha	state	41947358	6996124	34951234	0.09	0.42	0.49	0.00	0.00	0.89	0.11
Puducherry	Union territory	1244464	850123	394341	NA	NA	NA	NA	NA	NA	NA
Punjab	state	27704236	10387436	17316800	0.71	0.08	0.22	0.01	0.00	0.44	0.54
Rajasthan	state	68621012	17080776	51540236	0.18	0.20	0.63	0.00	0.00	0.92	0.08
Sikkim	state	607688	151726	455962	0.35	0.01	0.64	0.02	0.00	0.16	0.83
Tamil Nadu	state	72138958	34949729	37189229	0.26	0.50	0.24	0.00	0.00	0.83	0.16
Tripura	state	3671032	960981	2710051	0.00	0.02	0.97	0.00	0.00	0.04	0.96
Uttar Pradesh	state	2E+08	44470455	1.55E+08	0.21	0.26	0.53	0.00	0.00	0.84	0.16
Uttarakhand	state	10116752	3091169	7025583	0.43	0.08	0.49	0.02	0.00	0.58	0.41
West Bengal	state	91347736	29134060	62213676	0.33	0.05	0.62	0.00	0.00	0.55	0.45

BANGLADESH TOTAL	country	1.44E+08	33563183	1.1E+08	0.11	0.22	0.67	0.00	0.09	0.06	0.86
Barisal	division	8325666	1361943	6963723	0.02	0.05	0.94	0.00	0.09	0.02	0.89
Chittagong	division	28423019	6905480	21517539	0.02	0.26	0.72	0.00	0.06	0.03	0.91
Dhaka	division	47424418	15584835	31839583	0.21	0.32	0.47	0.00	0.12	0.06	0.81
Khulna	division	15687759	2822121	12865638	0.03	0.03	0.95	0.00	0.04	0.03	0.93
Rajshahi	division	18484858	3317022	15167836	0.01	0.08	0.91	0.00	0.09	0.05	0.86
Rangpur	division	15787758	2109071	13678687	0.00	0.11	0.88	0.00	0.04	0.14	0.81
Sylhet	division	9910219	1462711	8447508	0.07	0.06	0.87	0.00	0.14	0.05	0.81

Table S3.3. Sensitivity analysis results. We test the sensitivity of our modelled *Cryptosporidium* human emissions for Bangladesh and India to changes in a number of model parameters. The numbers shown are percentage change in oocyst emissions compared to the standard run (0). Positive numbers are therefore increases, negative numbers are decreases. We do the sensitivity analysis in pairs of parameters, changing one or both at a time, as some parameters are strongly related to one another. Parameters can take three different values each (Table S3.1). The first series shows changes in sewage treatment (all currently connected people switch to secondary or tertiary treatment) and removal efficiencies of *Cryptosporidium* during treatment (based on the ranges found by Hofstra et al. 2013: primary 4-90%, secondary 28-100%, tertiary 40-99%). The second series shows changes in whether or not the 'no facilities' category emissions reach the surface water. The third series shows changes in whether or not the 'septic tanks and pit latrines' category emissions reach the surface water. The fourth series shows changes in country population and urban population sizes (10% lower or higher). The fifth series shows changes in *Cryptosporidium* oocyst excretion (1 log unit lower or higher) and the number of infections (half or double). Each series contains 9 individual model runs, leading to a total of 45 runs per country.

		Bangladesh			**India**		
		Removal efficiencies					
		Low	*Middle*	*High*	*Low*	*Middle*	*High*
Sewage treatment	*Current situation (primary only)*	0	0	-3	1	0	-10
	All switch to secondary	-1	-1	-3	-2	-5	-11
	All switch to tertiary	-1	-3	-3	-4	-11	-11

		Share of faeces from rural population without facilities that reaches surface waters					
Share of faeces from urban population without facilities that reaches surface waters		*0%*	*25%*	*50%*	*0%*	*25%*	*50%*
	100%	-1	7	14	-8	68	143
	75%	-10	-2	5	-16	60	136
	50%	-19	-11	-4	-23	52	128
		Share of rural septic tank and pit latrine content that reaches surface waters					
Share of urban septic tank and pit latrine content that reaches surface waters		*0%*	*25%*	*50%*	*0%*	*25%*	*50%*
	0%	0	115	229	0	26	51
	25%	27	142	256	17	42	68
	50%	54	169	284	33	59	85
		Urbanisation					
Population size		*Low (-10%)*	*Middle*	*High (+10%)*	*Low (-10%)*	*Middle*	*High (+10%)*
	Low (-10%)	-22	-10	2	-34	-10	14
	Middle	-13	0	13	-27	0	27
	High (+10%)	-4	10	24	-20	10	40
		Number of infections					
Oocyst excretion per infected individual		*Low (half)*	*Middle*	*High (double)*	*Low (half)*	*Middle*	*High (double)*
	Low (-1log)	-95	-90	-80	-95	-90	-80
	Middle	-50	0	100	-50	0	100
	High (+1log)	400	900	1900	400	900	1900

Supplementary material for Chapter 5

Supplementary material S5 is digitally available on request. It is a large Excel file that could not be included in the printed version of the thesis.

Supplementary material for Chapter 6

S6.1. Distribution of livestock over the world

The number of animals in a grid cell (Ni) are from the Gridded Livestock of the World v2.0 (Robinson et al., 2014) for six animal species: cattle, sheep, goats, pigs, chickens and ducks. We made assumptions on the spatial distribution for five other livestock species (buffaloes, horses, mules, asses, camels) for which only country totals were available (from the FAO (2013)). We assumed that buffaloes are distributed over grid cells within countries proportionally to the distribution of cattle, and that horses, camels, mules and asses are distributed proportionally to the distribution of sheep plus goats. We chose cattle because we assume that these are most likely to be kept in similar circumstances as buffaloes, as they are both used for dairy and meat production. And the same holds for sheep plus goats and horses, mules, assess and camels, which are all animal species that are often kept free range. The way the country totals for these animals are distributed over the grid cells does not influence our calculation of the total oocyst load, it only influences where on the map this load is assigned too. However, as buffaloes, horses, mules, asses and camels are relatively little in number compared to the other livestock species (Figures S6.1 and S6.2) the appearance of the final oocyst load map is hardly influenced either. Figures S6.1 and S6.2 shows the distribution of livestock over the world, aggregated over seven world regions: Europe, Asia, Africa, North America, Latin America, Middle East - North Africa (MENA) and Oceania. We distinguish these seven regions as our input data for prevalence make the same distinction. The MENA region was taken separately from Africa and Asia because its characteristics differ from the surrounding areas, for example in relative occurrence of different livestock species (Figure S6.1 and S6.2). On the Livestock Geowiki website, detailed maps displaying the Gridded Livestock of the World v2.0 data can be found (http://livestock.geo-wiki.org/home-2/).

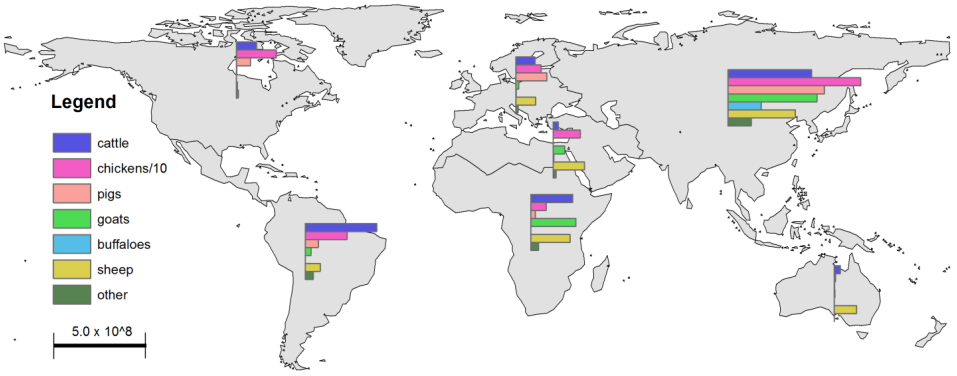


Figure S6.1: Distribution of livestock over the world. This figure shows the numbers of livestock summed per world region. Horses, camels, mules, asses and ducks were taken together in the category 'other'. Chicken and duck totals were divided by 10 for visual comparison purposes.

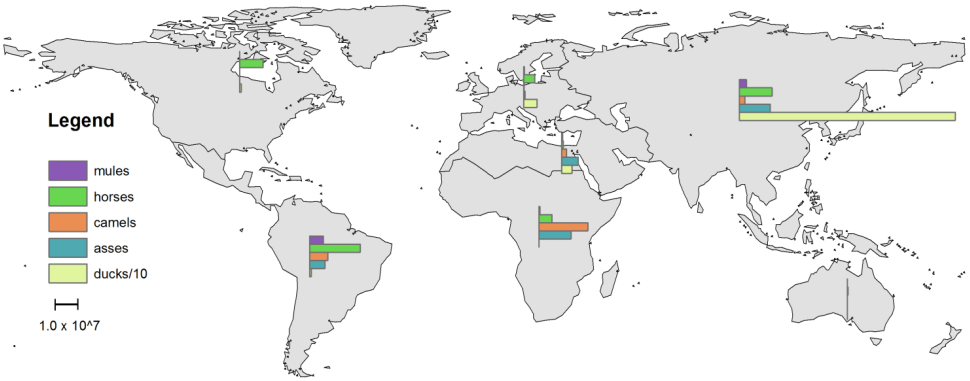


Figure S6.2: Distribution of livestock over the world. This figure shows the numbers of livestock summed per world region for horses, camels, mules, asses and ducks. Duck totals were divided by 10 for visual comparison purposes.

S6.2. Manure production

Manure production rates are calculated from average adult animal body mass (Table S6.1)(IPCC, 2006) and manure production per day per 1000 kg live animal mass (Table S6.2)(Safley et al., 1992), similar to what is done in manure production estimates for greenhouse gas emission inventories. For animals under three months old, we take the following birth weights for our calculations: cattle and buffaloes 40 kg, pigs 1.5 kg, sheep and goats 4.5 kg. Oocyst production per animal likely varies much more strongly with other factors than body mass, such as immune status. This is, for example, reflected by the fact that for many animal species it has been observed that younger (and thus smaller) animals excrete more oocysts than adults (Table S6.4). However, as there is no extensive research relating livestock factors to oocyst production, this is the best possible estimate with the available data. Using birth weight to calculate manure production of animals under three months might give an underestimation, as the animals grow and their manure production increases during this period. However, as oocyst shedding peaks shortly after birth in most livestock species (Robertson et al., 2014), we believe we made reasonable estimates for our purpose of calculating annual oocyst loads. No body weight estimates for chickens and ducks are needed, as the oocyst excretion rate values for these animals are in oocysts / infected animal / day, instead of excretion in oocysts / gram manure.

Table S6.1. Average adult body mass (kg) for the different animal species. Based on data from the IPCC guidelines for National Greenhouse Gas inventories (IPCC, 2006).

	Cattle	Buffaloes	Pigs	Sheep	Goats	Camels	Horses	Mules	Asses
Africa	250	400	28	28	30	217	238	130	130
Asia	350	400	28	28	30	217	238	130	130
Europe	550	400	190	48.5	38.5	217	377	130	130
Latin America	400	400	28	28	30	217	238	130	130
Middle East /North Africa	250	400	28	28	30	217	238	130	130
North America	550	400	190	48.5	38.5	217	377	130	130
Oceania	450	400	190	48.5	38.5	217	377	130	130

Table S6.2. Manure production per day per 1000 kg live animal mass. These data are based on the USEPA report (Safley et al., 1992). For cattle, we take the average of the reported values for beef cattle and dairy cattle, as the Gridded Livestock of the World dataset does not distinguish between beef and dairy cattle. For mules, we take the same values as for horses, asses and camels.

	Kg manure /day /1000 kg body mass
Beef cattle & buffaloes	58
Dairy cattle	86
Pigs	85
Sheep	40
Goats	41
Horses	51
Asses	51
Camels	51

S6.3. Fraction young animals

To estimate the fraction of animals that is young (defined as under three months old) (Table S6.3), we made assumptions based on the fertility rates and average age at first parturition as described in the Global Livestock Environmental Assessment Model (GLEAM)(FAO, 2016).

Table S6.3 Estimating the fraction of animals that is young (defined as under three months old)

	Average fertility rate (fraction of females having young per year)*	Average age at first parturition* (years)	Typical number of offspring per parturition	Fraction of animals that is under 1 year old (our estimate)	Fraction of animals that is under three months (our estimate)
Cattle	0.8	2.7	1	0.33	0.08
Buffaloes	0.7	3.2	1	0.33	0.08
Pigs	2	1.3	>10	0.90	0.22
Sheep	0.85	1.7	2	0.60	0.15
Goats	0.85	1.5	2	0.60	0.15

*Numbers are based on the Global Livestock Environmental Assessment Model (GLEAM)(FAO, 2016)

S6.4. Prevalence and excretion rates

S6.4.1 Systematic literature review

Average prevalence of cryptosporidiosis (Py_i and Pa_i) and average oocyst excretion rate of infected animals (Ry_i and Ra_i) are based on an extensive systematic literature review (Chapter 5). Table S6.4 shows the model input derived from this review. For all livestock except cattle we combined the prevalence data from studies from different parts of the world, as there is not enough data to reliably deduce regional prevalences. For cattle, buffaloes, pigs, sheep and goats we make a distinction in prevalence and excretion rates between young and adult animals. We do not distinguish between age groups for the other species due to a lack of data.

Table S6.4. Model input for cryptosporidiosis prevalences (Py and Pa) and oocyst excretion rates (Ry and Ra) for the different livestock species. These numbers are based on a systematic literature review (Chapter 5).

Livestock species		Average prevalence (%)				Average excretion (^{10}log oocysts / gram faeces)			
		All ages combined (P)	Aged over 3 months (Pa)	Aged under 3 months (Py)	Number of studies	All ages combined (R)	Aged over 3 months (Ra)	Aged under 3 months (Ry)	Number of studies
Cattle	*All regions combined*				153		2	4.9	48
	Africa		16.7	28.9	20				
	Asia		15.9	24.7	44				
	Europe		13.4	30.3	38				
	Latin America		12.8	13.2	5				
	Middle East / North Africa		16.4	26.1	18				
	North America		9.2	25.1	23				
	Oceania		0.6	17.9	5				
Buffaloes			14.4	27.0	14		2[b]	4.9[b]	
Pigs			19.8	25.4	34		3.9	4.7	12
Sheep			13.2	25.1	43		2.3	4.4	14

Goats		10.6	15.8	28		2.2	5.3	10
Horses	7.5			19	3.1			4
Camels	12.2			7	3.1[c]			
Chickens	11.6			9	6.8[d]			
Ducks	14.7			3	4.7[d]			
Asses	0.9			1	3.1[c]			
Mules	0.9[e]				3.1[c]			

[a] The literature review also distinguished 'unknown age'. Where the mean of the 'unknown age' group was not in between the means for adults and young, these data were included with the group it was closest to (this was the case for prevalence for goats, pigs, and for cattle in Africa, Latin America and the MENA region). [b] For buffaloes we take the same values as for cattle. [c] For camels, asses and mules we take the same value as for horses. [d] For chickens and ducks, the number represents ^{10}log oocysts per infected animal per day (see section S6.4.2). [e] For mules we take the same value as for asses.

S6.4.2 Excretion rates for chickens and ducks

The literature review on which the cryptosporidiosis prevalence and excretion rates are based did not yield any results for oocyst excretion rates for chickens and ducks (Chapter 5). This literature review looks specifically at studies published in the period 2001-2015, that concern natural *Cryptosporidium* infections (no inoculation), and give an estimate of oocysts excretion per gram faeces. We broadened this search with older studies, inoculation experiments and studies that estimate oocyst excretion per animal per day, in order to determine suitable model input for chicken and duck excretion rates.

Table S6.5 shows the eight studies for chickens we found, one of which also concerns ducks (Hornok et al., 2000, 1999, 1998a, 1998b, Rhee et al., 1998, 1995; Sréter et al., 1995; Varga et al., 1995). Some of these studies were experiments with a clinical treatment, in these cases we only looked at the control groups, to stay as close as possible to the situation in regular poultry. In most cases, we used a plot digitizer to obtain the data of interest from graphs as accurately as possible (http://arohatgi.info/WebPlotDigitizer/, accessed in July 2016).

One the basis of the data presented in Table S6.5, we used 6.8 log oocysts per day as model input for chickens, this is the ^{10}log of the average excretion rate observed in the eight studies. For ducks we used the single observation as model input. None of these studies concern natural *Cryptosporidium* infections; all are inoculation experiments, where the poultry were fed oocysts and their response was then measured. This might lead to an overestimation of the excretion rates, as it has been found that inoculation with higher oocyst numbers leads to higher oocyst shedding (Rhee et al., 1995; Zambriski et al., 2013), although another study on calves does not find this (Zambriski et al., 2013). All eight studies concern very young poultry, mostly in their first week of life. Often, the inoculation studies are done with very young chickens (<1 week old), but lower excretion rates are

observed in among somewhat older chickens (9 weeks)(Sréter et al., 1995). Furthermore, chickens inoculated at a younger age were found to shed for a much longer period (Sréter et al., 1995). Also in many other animal species it is found that young animals shed more oocysts than adults (Chapter 5), so this might also lead to an overestimation in our model. Despite these concerns, we believe that basing our model input on these available data is the best possible assumption.

Table S6.5. Excretion rates for chickens and ducks. In all studies the reported unit is oocysts per animal per day, and the *Cryptosporidium* species studied is in all cases *C. baileyi*. This table shows the general characteristics of the studies, the log10 of the mean of the observed oocyst count in faeces, the amount of oocysts with which the animals were inoculated prior to faecal measurements, and the age of the animals during the experiment.

Author	Publication year	Animal species	^{10}log of mean oocyst count	Inoculation amount of oocysts	Animal age during experiment
Hornok	1998	chickens	6.98	8.E+05	1 day
Hornok	1998	chickens	6.52	8.E+05	7 days, 28 days
Hornok	1999	chickens	6.93	8.E+05	7 days
Hornok	2000	chickens	6.78	8.E+05	10 days
Rhee	1998	chickens	5.54	5.E+05	2 days
Sreter	1995	chickens	6.16	6.E+05	7 days, 63 days
Varga	1995	chickens	7.23	6.E+05	7 days
Rhee	1995	chickens	5.43	2.E+02 to 2.E+06	2 days
Rhee	1995	ducks	4.72	2.E+02 to 2.E+06	2 days

S6.5. Data on intensive and extensive farming and manure storage systems

Data on whether animals are kept in intensive or extensive farming systems ($Fintl_i$, $Fextl_i$, $Fints_i$, $Fexts_i$) were taken from the Integrated Model to Assess the Global Environment (IMAGE) according to Bouwman et al. (2012). The classification is based on an FAO report that distinguishes pastoral systems (extensive) and mixed and landless systems (taken together as intensive)(Seré and Steinfeld, 1996). In both intensive and extensive systems, data are provided for the fraction of manure dropped directly on land during grazing, the fraction going to storage, and the fraction used for other purposes (such as burning for fuel) that leaves the system (Bouwman et al., 2012). The IMAGE model categories ‘dairy cattle’ and ‘beef cattle’ were averaged to represent cattle, as the Gridded Livestock of the

World v2.0 does not distinguish between beef and dairy cattle. The IMAGE model category 'sheep & goats' was considered representative for both our categories sheep and goats, and 'poultry' for ducks and chickens.

Table S6.6. Model input for manure storage (FS_j)

Manure management system categories (IPCC 2006)	Manure management system categories (Safley et al. 1992)	Assumption on survival (∇_j)	Assumption on storage time (t_j)
Daily spread	Daily spread	All survive (∇_j= 1)	0 days
Lagoon	Anaerobic lagoon	∇_j according to eq. 6.5	9 months
Liquid/slurry	Liquid systems (includes liquid/slurry storage and pit storage)	∇_j according to eq. 6.5	9 months
Pit < 1 month		∇_j according to eq. 6.5	1 month
Pit > 1 month		∇_j according to eq. 6.5	9 months
Solid storage	Solid storage & dry lot	∇_j according to eq. 6.5	9 months
Dry lot		∇_j according to eq. 6.5	9 months
Other	Other (includes deep pit stacks, litter, and other)	∇_j according to eq. 6.5	9 months
Anaerobic digester		2 log reduction in infectivity (∇_j= 0.01)	NA
Burned for fuel	Used for fuel [a]	Leaves the system (not storage, not to land)	NA
Pasture/range/ paddock	Pasture/range/paddock	To land (not storage)	NA

[a] According to the description, this category includes both manure going to anaerobic digesters and manure that is burned for fuel. We therefore assume that half of the manure in this category leaves the system, and the other half ends up on land after a 2 log reduction in infectivity.

Data on the use of different storage systems (FS_j) for the different animal species are from the 2006 IPCC Guidelines for National Greenhouse Gas Inventories (IPCC, 2006) and underlying data from a USEPA report on Global methane emissions from livestock and poultry manure (Safley et al., 1992) (Table S6.6). The 2006 IPCC guidelines provide data on the manure management systems used for cattle, buffaloes and pigs (IPCC, 2006). For the other animal species the USEPA data were used (Safley et al., 1992). The IPCC data are continent-specific, while the USEPA data are country-specific. Data are provided for an extensive list of countries. For countries not on the list, regional averages were used. For the manure management system category 'burned for fuel' we assumed manure leaves the system and does not get stored and spread on land. We assumed the category 'pasture/ range/ paddock' covers the manure dropped on land directly, as we account for in equation 6.2 in the main article.

S6.6. Temperature-dependent survival during manure storage

We derived the value of K based on seven studies that reported a relation between temperature (°C) and oocyst survival (measured as viability or infectivity) in livestock manure (Collick et al., 2006; Garcés et al., 2006; Hutchison et al., 2005; Jenkins et al., 2013, 1999; Kinyua et al., 2016; Olson et al., 1999) (Table S6.7). Since the data points of the studies using infectivity show similar inactivation as the data points from studies using viability assays, the data points from all of these studies were pooled (Figure S6.3). We excluded two studies that measured only presence of oocysts and not viability or infectivity, as these were found not to fit well with the other data (Côté et al., 2006; van Herk et al., 2004). The studies reported their findings in different metrics (e.g. percentage viability decline, t_{99}, times reduction, decrease in infectivity). We calculated each metric back to a daily survival rate and from there to a t_{90} value (time in days until a 1log reduction).

Figure S6.3 shows the data points from the seven studies and the linear relation between temperature and t_{90}. Based on this linear trend we use the following equation in GloWPa-Crypto L1 for the relation between ambient temperature and *Cryptosporidium* survival in manure (expressed as t_{90} value):

$$t_{90} = -2.5586 \times T + 119.63 \quad \text{(S6.1)}$$

In our model, T is the average annual air temperature in a grid cell, it is the average over 1970-2000 derived from the WATCH forcing data (Weedon et al., 2011). Average air temperature was used, as we are not interested in the conditions in a specific year, that might be unusually warm or cold in certain locations. We determined a t_{90} value for each grid cell using equation S6.1. We then substituted t in equation 6.4 in the main article with t_{90}, and V with 0.1 (1 log reduction), to determine a value for constant K for each grid cell.

In this analysis, we pooled data from different types of manure storage systems. Nevertheless, conditions under which manure is stored (e.g. moisture content, mixing with other materials) may differ per system, and these conditions might influence *Cryptosporidium* survival. For example, the two highest points in Figure S6.3 correspond to the study by Hutchison et al. (2005). We can only speculate as to why the observed survival in this study is so much higher than in the other studies. It could have to do with the storage type (liquid tank storage) or the *Cryptosporidium* strain, the researchers might have accomplished better oocyst recovery, or it may simply be natural variation. More research into this topic would be worthwhile, to get more insight into the effect of different manure management practices on oocyst survival. Furthermore, all seven studies concern cattle or pig manure. Information on the relation between temperature and oocyst survival in manure from other livestock species is not available, to our knowledge.

Nevertheless, the data in Figure S6.3 show a reasonable fit that shows a clear decrease in oocyst survival in manure with increasing temperatures. Therefore, we decided to take ambient temperature into account in our model.

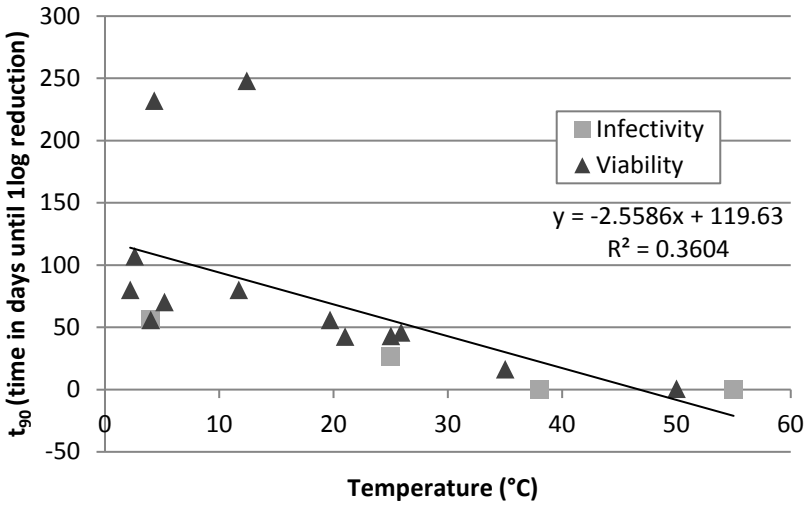


Figure S6.3: Oocyst survival in manure (t_{90}) under different ambient temperatures. We extracted the temperature and t_{90} from seven studies on survival of oocysts in manure under different temperature conditions. These were plotted and a linear trend line was fitted using Microsoft Excel. Since the data points of the studies using infectivity show similar inactivation as the data points from studies using viability assays, the data points from all of these studies were taken together for fitting the linear trend line. Reported temperature is ambient temperature, except for systems in which the storage system itself reached a certain temperature (for example during composting).

Temperatures below zero are modelled as an extrapolation of the trend line from Figure S6.3. Freezing has been found to reduce oocyst viability more strongly than temperatures just above zero, but the conditions under which oocysts are frozen matter. Snap-freezing oocysts was found to lead to 100% death, while slow freezing lead to higher survival rates, with a small proportion of oocysts still viable even after 750 hours at -22°C (Robertson et al., 1992). Sattar et al. (1999) found that oocysts in calf manure that were frozen at -20°C for at least 24 hours still had excystation rates between 39-72%. We did not consider freezing temperatures a special case in computing oocyst survival during manure storage, for three main reasons. Firstly, the relation between below-zero temperatures and oocyst survival is not straightforward, as Robertson et al. and Sattar et al. show. Secondly, we use ambient temperature (average annual air temperature), and when ambient temperatures

are below zero, this does not mean that the temperature within a manure storage system will also be below zero. And thirdly, if the average annual air temperature is below zero, it could very well be that part of the year the temperature is above zero and *Cryptosporidium* survives for a long time. Thus assuming that only few oocysts survive when the temperature is below zero could lead to underestimation of the actual situation. Therefore, we chose to model temperatures below zero as an extrapolation of the relation found in Figure S6.3.

Table S6.7. Studies that report a relation between temperature and survival of *Cryptosporidium* oocysts in manure. This table shows the general information about the seven studies that were found, with the metric they reported (changes in viability or infectivity). From these we calculated a daily survival rate and from that a t_{90} value.

First author	Publication year	Animal species	Measured viability or infectivity	System description	storage time	Season	Temperature (°C)	Viability (%)	Infectivity	Reduction (times)	Daily survival rate (calculated)	t_{90} (calculated)
Collick	2006	cattle	viability	Bedding	8 weeks	winter	2.6	30		3.33	0.98	107.1
						spring	2.2	20		5	0.97	80.1
						summer	11.7	20		5	0.97	80.1
						late summer	19.7	10		10	0.96	56.0
Hutchison	2005	pig and cattle	viability	Liquid tank storage	3 months	summer	12.4	From 90 to 60%			0.99	248.0
						winter	4.3				0.99	232.0
Jenkins	1999	cattle	viability	Composting	82 days		35			5log reduction	0.87	16.4
					5 days		50				0.02	0.6
Jenkins	2013	pigs	viability	Lagoon	t_{99} of 91.7 days	summer	25.9				0.95	45.9
					t_{99} of 140.7 days	winter	5.2				0.97	70.4
Kinyua	2016	pigs	viability	Anaerobic digestion			21			die-off 0.054/day	0.95	42.6
Olson	1999	cattle	viability	Wet faeces in the dark in lab	12 weeks		4	10		10	0.96	56.0
					12 weeks		25	5		20	0.95	43.0
			infectivity		12 weeks		4		10	10	0.96	56.0
					5 weeks		25		5	20	0.92	26.9
Garcés	2006	cattle	infectivity	Anaerobic digestion	1 hrs		38		Reduced by 2 log		1E-48	0.0
					4 hrs		55		Reduced by 5 log		1E-30	0.0

S6.7. Sensitivity analysis

We performed a nominal range sensitivity analysis changing one parameter at a time based on a reasonable range the parameter can take, usually both a decrease and an increase were taken. Table S6.8 shows the changed parameters and the outcomes of the sensitivity model runs for all parameters except prevalence and excretion rates, Table S6.9 shows it for prevalence, and Table S6.10 for excretion rates. Here we discuss the changes to the parameters and the results for each of these tables.

S6.7.1 Sensitivity analysis for all parameters except prevalence and excretion rates

Animal numbers per grid cell, animal mass and manure production were increased and decreased by 20% in the sensitivity analysis. Animal numbers per grid cell are aggregated from the Gridded Livestock of the World data, which have a much finer spatial resolution (3 arc minutes grid) and have been extensively validated (Robinson et al., 2014). At our 30 min grid we expect these not to vary more than 20%. Average animal mass and manure production are expected not to vary more than 20% due to physiological constraints. The fraction of animals that is under three months old and the mass of young animals were halved and doubled, as we think these are more uncertain. The fractions of manure going to storage and to land were increased and decreased by 50%, and the storage time (average 9 months) was changed to 6 months and 1 year. This is based on an assumed 1 or 2 crop yields per year and spreading of manure at the start of the growing season. The average storage time for the category 'Pit storage < 1 month' was varied from two weeks to two months. The constants in equation S6.1, that describes the relation between temperature and t_{90}, were halved and doubled, as this approximately covers the spread of the points in Figure S6.3. Instead of average annual air temperature, in the sensitivity analysis we took the average air temperature in January and July to reflect the natural variation between the seasons.

The last column of Table S6.8 shows that changing these parameters leads to a factor 0.66-1.97 change in the total oocyst load. The model is most sensitive to changing the fractions of manure that is going to storage and to land. Furthermore, the model is quite sensitive to the fraction of animals that is young and the mass of young animals. The model is not so sensitive to changes in storage time and temperatures in different seasons.

Table S6.8. Sensitivity analysis for all parameters except prevalence and excretion rates. Column 1 gives the variable name, column 2 the changed value, column 3 the baseline model value and column 4 explains the change in words. Column 5 and 6 show the results of the model run with that particular change. Column 5 shows the total global oocyst load to land, and column 6 shows

the total oocyst load divided by the total oocyst load of the baseline model run, so the relative effect of the change.

Variable name	Variable changed	Variable baseline	Explanation	Total oocyst load	Total divided by baseline run
animal numbers (N)	0.8	1	20% decrease	2.60E+23	0.80
animal numbers (N)	1.2	1	20% increase	3.89E+23	1.20
fraction young (Fy)	0.5	varies per animal	halve	2.37E+23	0.73
fraction young (Fy)	2	varies per animal	double	4.99E+23	1.54
mass all	0.8	varies per animal	20% decrease	2.73E+23	0.84
mass all	1.2	varies per animal	20% increase	3.76E+23	1.16
mass young	0.5	varies per animal	halve	2.30E+23	0.71
mass young	2	varies per animal	double	5.14E+23	1.58
manure per1000kgmass	0.8	varies per animal	20% decrease	2.73E+23	0.84
manure per1000kgmass	1.2	varies per animal	20% increase	3.76E+23	1.16
fraction to storage (Fints, Fexts)	0.5	0	Storage +50%, land - 50%	2.14E+23	0.66
fraction to land (Fintl, Fextl)	0.5	0	Land +50%, storage - 50%	6.38E+23	1.97
storage time (t_j)	183	274 days (9 months)	6 months	3.61E+23	1.11
storage time (t_j)	365	274 days (9 months)	1 year	3.06E+23	0.94
storage time low (t_j)	15	30 days (1 month)	2 weeks	3.26E+23	1.00
storage time low (t_j)	60	30 days (1 month)	2 months	3.22E+23	0.99
temp par1	-1.2793	-2.5586	less steep	3.43E+23	1.06
temp par2	239.26	119.63	higher	4.30E+23	1.33
temp par1 & temp par2	both doubled		higher & steeper	3.98E+23	1.23
temp par1 & temp par2	both halved		lower & less steep	2.84E+23	0.88
Temperature (T)	average January T	average annual temperature (°C)	Seasonal storage	3.46E+23	1.07
Temperature (T)	average July T	average annual temperature (°C)	Seasonal storage	3.05E+23	0.94

S6.7.2 Sensitivity analysis for prevalence

The prevalence values used in this model are based on the mean cryptosporidiosis prevalence per animal species and age group found in the literature review in Chapter 5. For this sensitivity analysis, we take the highest prevalence found per animal species and age group in this same literature review (Table S6.9, column 2). We did not do a sensitivity run for lowest observed prevalence, as this was zero in most cases and therefore not so meaningful.

Table S6.5 shows that the model is most sensitive to changes in prevalence among young cattle (defined as under three months old); the total oocyst load to land changes with a factor 2.17. After young cattle, the model is most sensitive to the prevalence among adult pigs, chickens, young goats and adult cattle. The effect size for these animal species is similar to the effect size observed in Table S6.8. Changing the prevalence to the upper end of the range for all animal species at the same time (last row of Table S6.9) changes the total oocyst load to land with a factor 3.69.

Table S6.9. Sensitivity analysis for prevalence (Pa and Py). Column 1 gives the variable name, column 2 the changed value and column 3 the baseline value (shown as %). Column 4 and 5 show the results of the model run with that particular change. Column 4 shows the total global oocyst load to land, and column 5 shows the total oocyst load divided by the total oocyst load of the baseline model run, so the relative effect of the change. First prevalence was changed per animal species, and the last row shows the effect when prevalence is changed for all animal species at the same time.

Variable name	Variable changed	Variable baseline	Total oocyst load	Total divided by baseline run
prev cattle young	80	varies per region, average 25.8	7.05E+23	2.17
prev cattle adult	73.3	varies per region, average 14.5	3.77E+23	1.16
prev buffaloes young	50.6	27	3.37E+23	1.04
prev buffaloes adult	50.6	14.4	3.27E+23	1.01
prev pigs young	80	25.4	3.29E+23	1.02
prev pigs adult	80	19.8	4.70E+23	1.45
prev sheep young	96.5	25.1	3.42E+23	1.05
prev sheep adult	96.5	13.2	3.31E+23	1.02
prev goats young	77.3	15.8	4.25E+23	1.31
prev goats adult	77.3	10.6	3.28E+23	1.01
prev horses	27.8	7.5	3.32E+23	1.02
prev camels	37.9	12.2	3.29E+23	1.01
prev chickens	34.4	11.6	4.58E+23	1.41
prev ducks	30.4	14.7	3.25E+23	1.00
prev asses	1.7	0.9	3.25E+23	1.00
prev mules	1.7	0.9	3.24E+23	1.00
prev all high	varies per animal	varies per animal	1.20E+24	3.69

S6.7.3 Sensitivity analysis for excretion rates

The excretion rates used in this model are based on the mean cryptosporidiosis excretion per animal species as found in a literature review in Chapter 5. For each study a mean excretion rate was determined, and as model input we use the median of these means. For this sensitivity analysis, we change the excretion rate by one geometric baseline deviation (SD) of the observed mean excretion rates of that animal species from this same literature analysis. For buffaloes, the same value as for cattle is used. For camels, assess and mules, the same value as for horses is used. For ducks, the geometric SD of the chicken studies was used (see Table S6.5).

Table S6.10 shows that the model is most sensitive to the excretion rate of young cattle. Adding one SD to the excretion rate leads to an increase in the total oocyst load to land by a factor 27.9. After young cattle, the model is most sensitive to changes in the excretion rate of young goats, young buffaloes, adult pigs, horses, chickens, adult cattle and camels. Decreasing the excretion rate with one SD for all animal species at once changes oocyst loads to land with a factor 0.07. Increasing the excretion rate with one SD for all animal species at once changes oocyst loads to land with a factor 39.8.

It can be seen from Tables S6.8-6.10 that the model is more sensitive to changes in the oocyst excretion rates than it is to changes in prevalence or other variables. This is not surprising, as oocyst excretion rates can vary over several orders of magnitude (Chapter 5). This means that the absolute size of oocyst loads to land, the relative importance of the different animal species and the patterns on maps should be interpreted with this fact in mind.

Table S6.10. Sensitivity analysis for excretion rates (Ra and Ry). Column 1 gives the variable name, column 2 the changed value, column 3 the baseline value and column 4 an explanation in words. The values in column 2 and 3 are in ^{10}log oocysts / gram manure, except for chickens and ducks, the value is in ^{10}log oocysts / infected animal / day. Column 5 and 6 show the results of the model run with that particular change. Column 5 shows the total global oocyst load to land, and column 6 shows the total oocyst load divided by the total oocyst load of the baseline model run, so the relative effect of the change. First the excretion rate was changed per animal species, and the last two rows show the effect when excretion rates are changed for all animal species at the same time.

Variable name	Variable changed	Variable baseline	Explanation	Total oocyst load	Total divided by baseline run
exc cattle young	3.1	4.9	1 SD lower	1.86E+23	0.57
exc cattle adult	0.7	2	1 SD lower	3.13E+23	0.97
exc buffaloes young	3.1	4.9	1 SD lower	3.10E+23	0.96
exc buffaloes adult	0.7	2	1 SD lower	3.23E+23	1.00
exc pigs young	3.9	4.7	1 SD lower	3.23E+23	0.99
exc pigs adult	2.9	3.9	1 SD lower	2.81E+23	0.87
exc sheep young	3.3	4.4	1 SD lower	3.19E+23	0.98
exc sheep adult	1	2.3	1 SD lower	3.23E+23	1.00
exc goats young	3.6	5.3	1 SD lower	2.99E+23	0.92
exc goats adult	1	2.2	1 SD lower	3.24E+23	1.00
exc horses	1	3.1	1 SD lower	3.22E+23	0.99
exc camels	1	3.1	1 SD lower	3.22E+23	0.99
exc asses	1	3.1	1 SD lower	3.24E+23	1.00
exc mules	1	3.1	1 SD lower	3.24E+23	1.00
exc chickens day	6.1	6.8	1 SD lower	2.71E+23	0.84
exc ducks day	4	4.7	1 SD lower	3.24E+23	1.00
exc cattle young	6.7	4.9	1 SD higher	9.05E+24	27.90
exc cattle adult	3.3	2	1 SD higher	5.47E+23	1.69
exc buffaloes young	6.7	4.9	1 SD higher	1.22E+24	3.77
exc buffaloes adult	3.3	2	1 SD higher	3.46E+23	1.07
exc pigs young	5.5	4.7	1 SD higher	3.37E+23	1.04
exc pigs adult	4.9	3.9	1 SD higher	7.54E+23	2.33
exc sheep young	5.5	4.4	1 SD higher	3.95E+23	1.22
exc sheep adult	3.6	2.3	1 SD higher	3.44E+23	1.06
exc goats young	7	5.3	1 SD higher	1.59E+24	4.90
exc goats adult	3.4	2.2	1 SD higher	3.32E+23	1.02
exc horses	5.2	3.1	1 SD higher	6.71E+23	2.07
exc camels	5.2	3.1	1 SD higher	5.78E+23	1.78
exc asses	5.2	3.1	1 SD higher	3.38E+23	1.04
exc mules	5.2	3.1	1 SD higher	3.28E+23	1.01
exc chickens day	7.5	6.8	1 SD higher	6.26E+23	1.93
exc ducks day	5.4	4.7	1 SD higher	3.25E+23	1.00
exc all low	varies per animal	varies per animal	1 SD lower	2.41E+22	0.07
exc all high	varies per animal	varies per animal	1 SD higher	1.29E+25	39.82

Supplementary material for Chapter 7

S7.1 Existing modelling approaches of oocyst transport from land to rivers

We have looked into the literature on existing modelling approaches of oocyst transport from manure on land to rivers, and conclude there currently is no single 'gold standard' approach.

An approach in several models is to model fluxes of pathogens between compartments. An example is the work of Yeghiazarian et al. (2009), who took a stochastic approach to modelling pathogen behaviour, especially focused on transitions between states (e.g. free, attached to particles, etc.). Sterk et al. (2016) use an equation relating pathogen release from manure to the duration and intensity of a rainfall event (Bradford and Schijven, 2002; Guber et al., 2006). QMRACatch is a tool for simulation of microbial water quality and associated risk of infection in catchments (Schijven et al., 2015) that simulates release rates of *Cryptosporidium* from manure in floodplains to be between 0 and 1, a simple approach not based observational data. Bhattarai et al. (2011) used their experimental results to calibrate a process-based model of *Cryptosporidium* in overland flow. The model uses a mass balance approach to calculate the concentrations of oocysts free in water and attached to mobile and immobile soil particles, these are then coupled to the output from the soil erosion model WEPP to calculate oocyst fluxes. Tang et al. (2011) applied the SWAT model to *Cryptosporidium* in two small agricultural catchments. They calibrated the decay and wash-off parameters with observational data. The model calculates an overall reduction (from oocyst input in manure to oocyst export in the two streams) of 3.5-4.2 log. Ferguson et al. (2007) do not model release of pathogens from faeces, but use a constant 'fraction of manure mobilized' in their process-based catchment scale model, meaning the fraction of manure on land that is transported to a stream during a large rainfall/runoff event. This constant varies between different animal species and is set between 0.005 and 0.05. Walker and Stedinger (1999) model the daily fraction of manure that is transported to streams as a function of the runoff, using a runoff threshold below which no manure is assumed to reach the stream. Two experiments using artificial precipitation in a laboratory setting were used to determine the upper bounds of the constants in their equation. The model by Dorner et al. (2006) predicts that the presence of agricultural soil drainage systems such as tile drainage can lead to increased pathogen transport. Overland flow in the this model is based on the sediment yield model by Hartley et al. (1987).

It can be seen from the above that a wide variety of approaches exist, there is no single method that is standard practice. All existing approaches described here require too detailed input data, in space and/or time, to directly apply them at a global scale on a monthly time step. Although crude, the approach we have chosen accounts for variability

in surface runoff and the magnitude of retention is based on observations under field conditions (Atwill et al., 2006b). More experimental data on oocyst release from manure and transport to rivers under various climatic and landscape conditions is needed to refine this approach.

S7.2 Temperature-dependent decay in other studies

Table S7.1 shows first order inactivation constants (K) of *Cryptosporidium* in surface waters as applied in other models (Rory Coffey et al., 2010; Dorner et al., 2006; Sterk et al., 2016; Tang et al., 2011). Our value for K_T is 0.0051 at 4°C and 0.064 at 20 °C, similar to these existing model approaches.

Several other models do not report comparable K values (Hoyer et al., 2015; Schijven et al., 2013; Walker and Stedinger, 1999), but are mostly based on the same experimental data as Peng et al. compiled (2008). Therefore, we conclude that modelling temperature-dependent decay in water according to Peng et al. is the current best estimate.

Table S7.1 First order inactivation constants (K) of *Cryptosporidium* in surface waters as applied in other models

Study	K value used in model	At temperature
Coffey 2010	0.01	4 – 15 °C
Dorner 2006	0.01	<4 – 15 °C
Dorner 2006	0.06	>20 °C
Sterk 2016	0.0018	4 °C
Tang 2011	0.05[a]	20 °C

[a] In soil solution, not water

S7.3 Loss due to decay and sedimentation

Figure S7.1 shows maps of the overall oocysts loss in January and July (top two panels), and below that the loss attributable to the different components: solar radiation-dependent decay, temperature-dependent decay, and sedimentation. The maps show oocyst loss on a log scale (log10(Fs)). Residence time of water in the grid cells is not taken into account in this figure, loss is expressed per day for comparison purposes.

Loss due to sedimentation is higher in shallower water, according to eq. 7.8 (Thomann and Mueller, 1987). It can be seen on the maps that dry areas (such as the Sahara region) have high loss rates and large river systems (such as the Congo and Amazon basins) have low loss rates due to sedimentation. Temperature-dependent decay is dependent on water temperature (van Vliet et al., 2012; Yearsley, 2009), which follows similar patterns as air temperature, with higher temperatures in the southern hemisphere in January and in the

northern hemisphere in July. Solar radiation-dependent decay is dependent on surface solar radiation (from the WATCH forcing data (Weedon et al., 2011)), water depth (calculated from discharge) and the dissolved organic carbon (DOC) concentration (from the Global NEWS model (Harrison et al., 2005; Mayorga et al., 2010)). The high decay that can be seen on the maps in the Nile basin and parts of the USA are due to the low DOC concentration there (see Figure 5A in Harrison et al. for a DOC yield map (kg C km^{-2} $year^{-1}$)), in combination with low discharge levels. We calculated DOC concentration from DOC export and river discharge, both at the river mouth, and assigned this value to all grid cells in the basin. This is a crude approach, but the current best estimate as no other global dataset of river DOC content is available, and it is relevant for determining solar radiation dependent-decay (King et al., 2008; Scully and Lean, 1994).

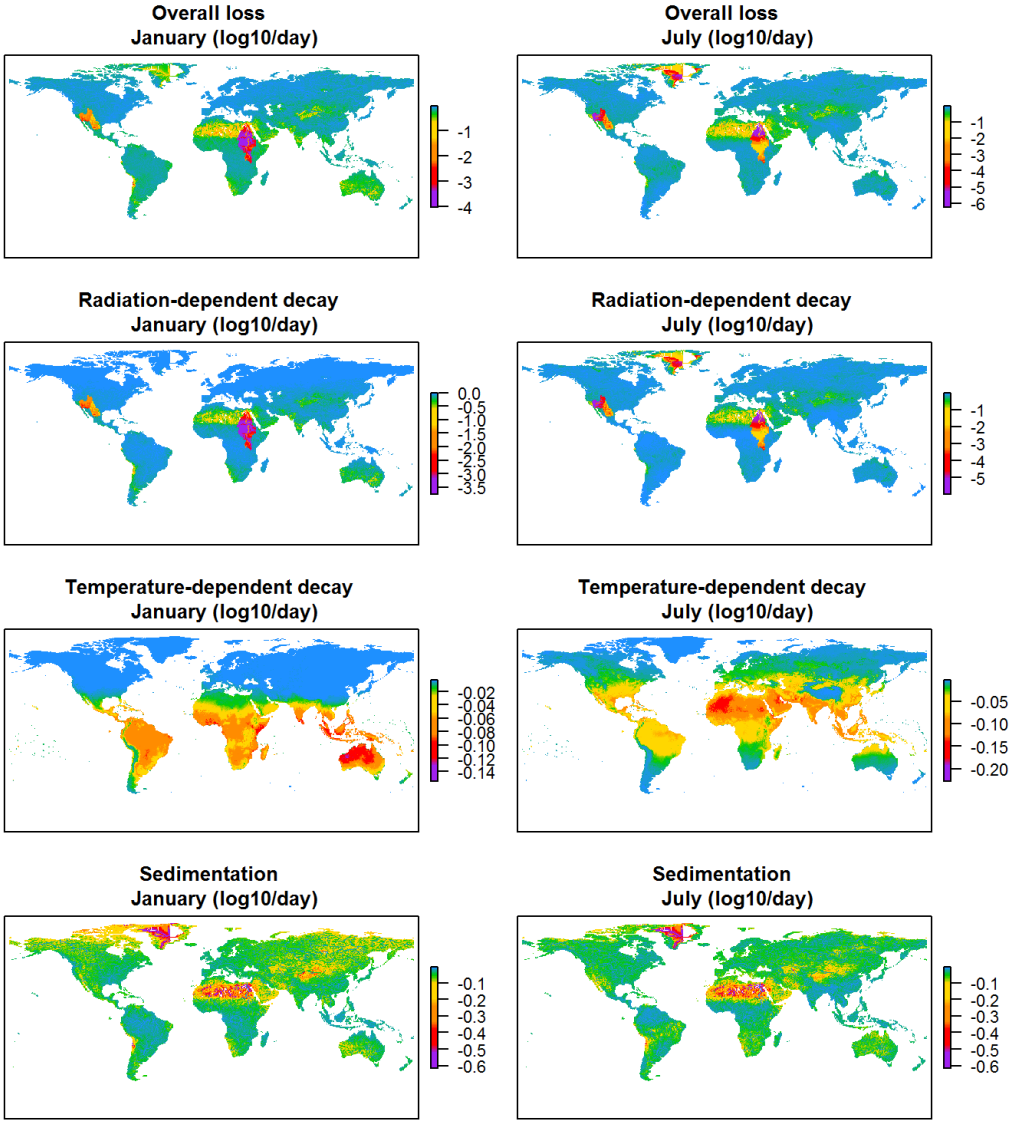


Figure S7.1 Maps of the overall oocysts losses in January and July (top two panels), and below that the loss attributable to the different components: solar radiation-dependent decay, temperature-dependent decay, and sedimentation. The maps show oocyst loss on a log scale (log10(Fs)). Residence time of water in the grid cells is not taken into account in this figure, loss is expressed per day for comparison purposes.

S7.4 Regional analysis

Figure S7.2 shows boxplots of all oocyst concentration grid cell values summarized by region and month. The top left panel indicates that the MENA region has overall highest concentration and North America overall lowest. However, it should be kept in mind that the many low values in North America are likely due to the large sparsely populated parts of Canada, that reduce median concentrations. On the contrary, the large sparsely populated Sahara region is mostly excluded from the analysis due to the low discharge levels, meaning that the MENA region has relatively few grid cells with low values. The boxplots should be interpreted in combination with the maps. The large LT2 dataset of *Cryptosporidium* observations in surface water in the United States shows median values from ca. 0.005 – 0.5 oocysts/L (Ongerth, 2013), covering approximately the upper half of the boxplot values for North America in Figure S7.2, indicating that the output range of our model is reasonable.

Figures S7.3-S7.9 show maps of modelled oocyst concentrations for the world seven regions, both with and without a discharge mask where grid cells with average annual discharge <200 m^3 s^{-1} are masked. Higher resolution plots are available on request.

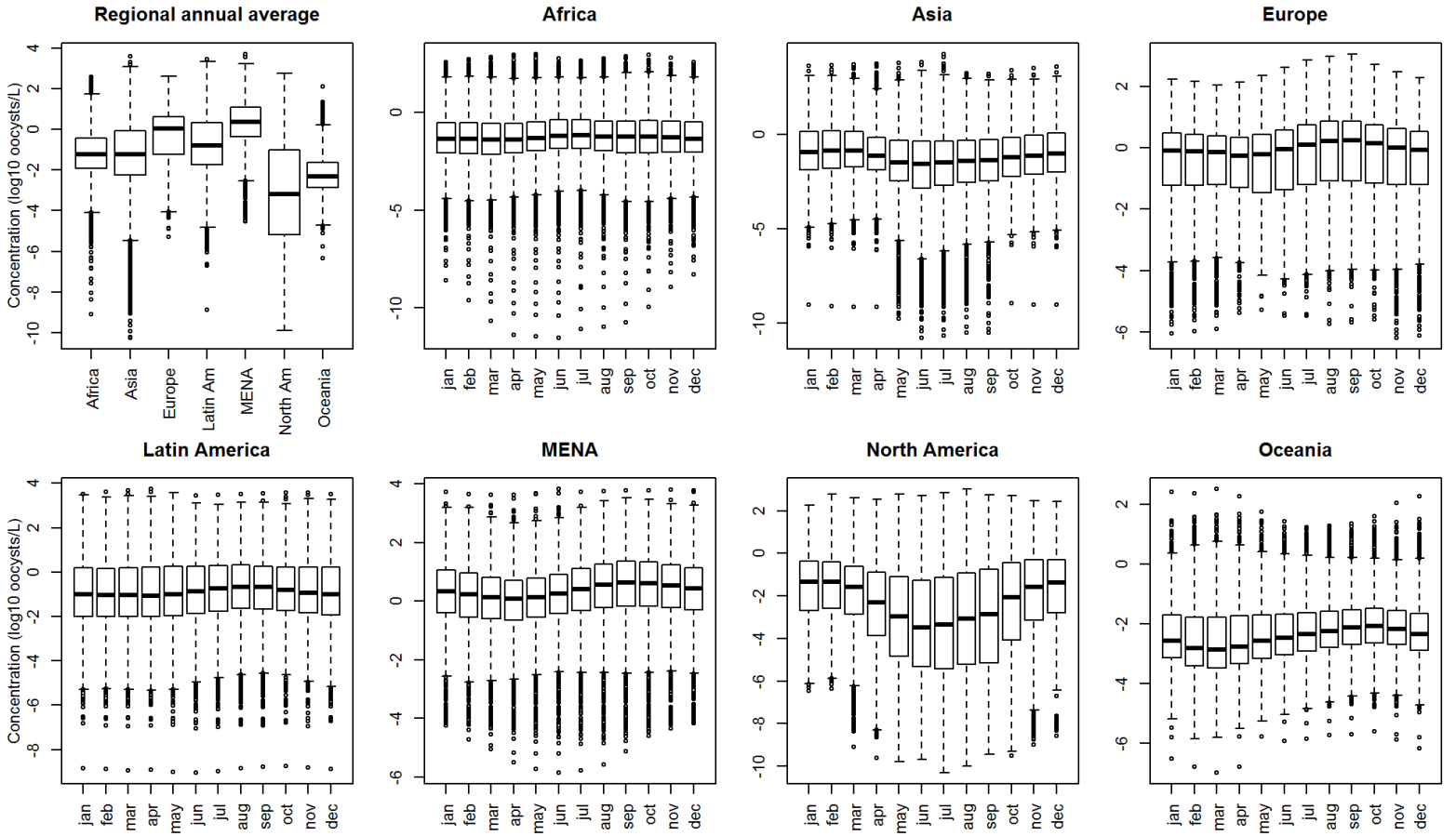


Figure S7.2: Boxplots of all oocyst concentration grid cell values (log10 oocysts/L) summarized by region and month.

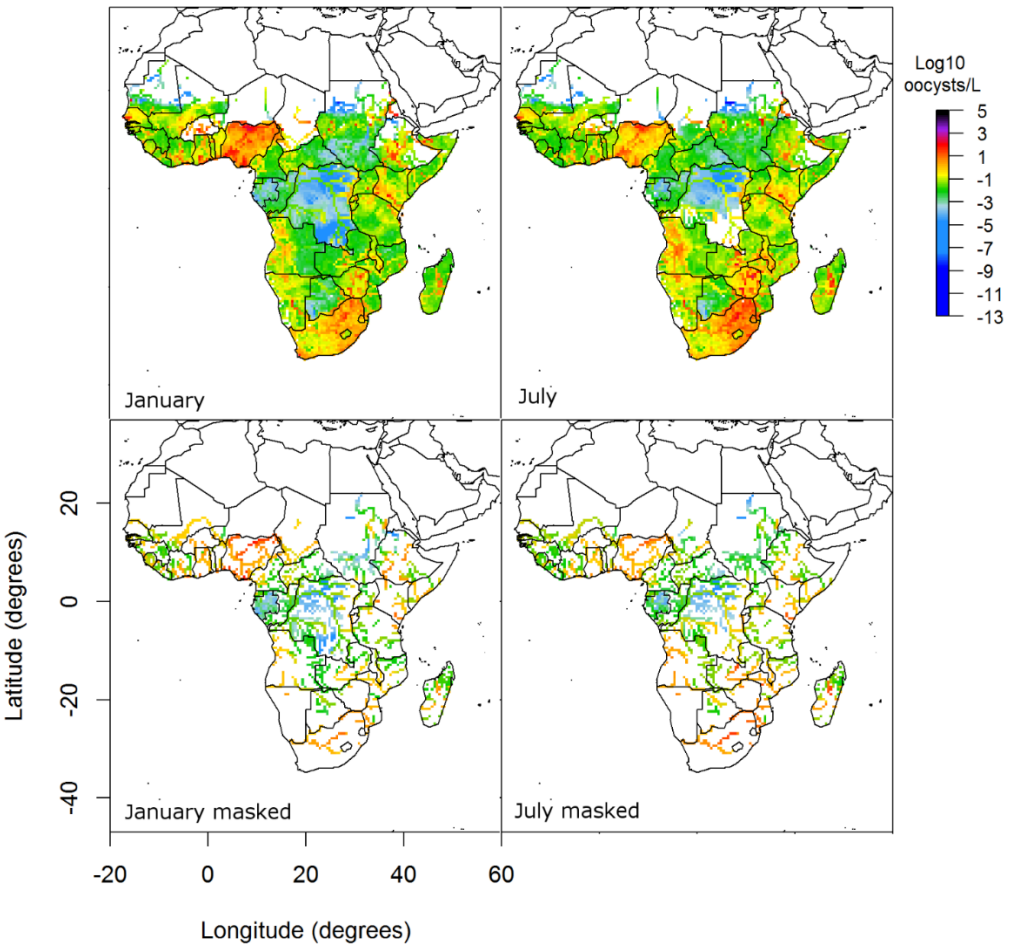


Figure S7.3 Modelled oocyst concentrations (log10 oocysts L^{-1}) in Africa for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

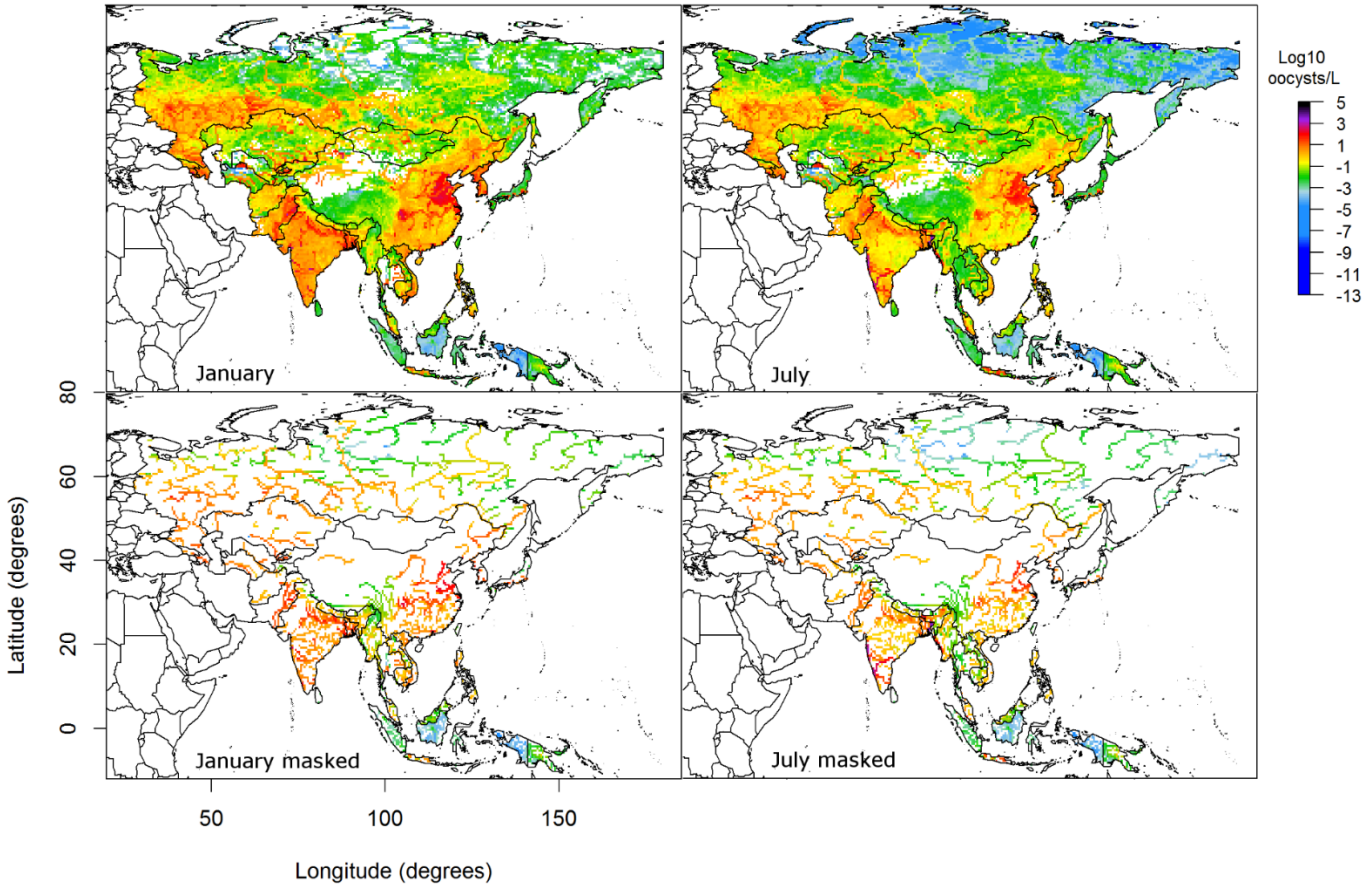


Figure S7.4 Modelled oocyst concentrations (log10 oocysts L^{-1}) in Asia for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

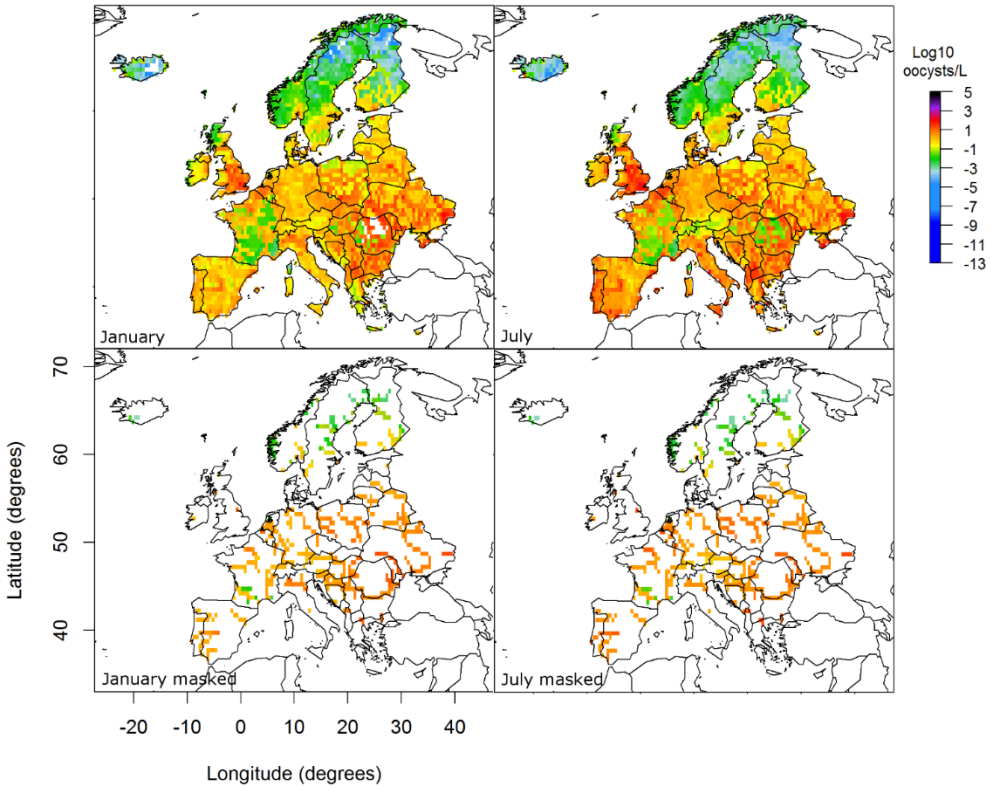


Figure S7.5 Modelled oocyst concentrations (log10 oocysts L^{-1}) in Europe for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

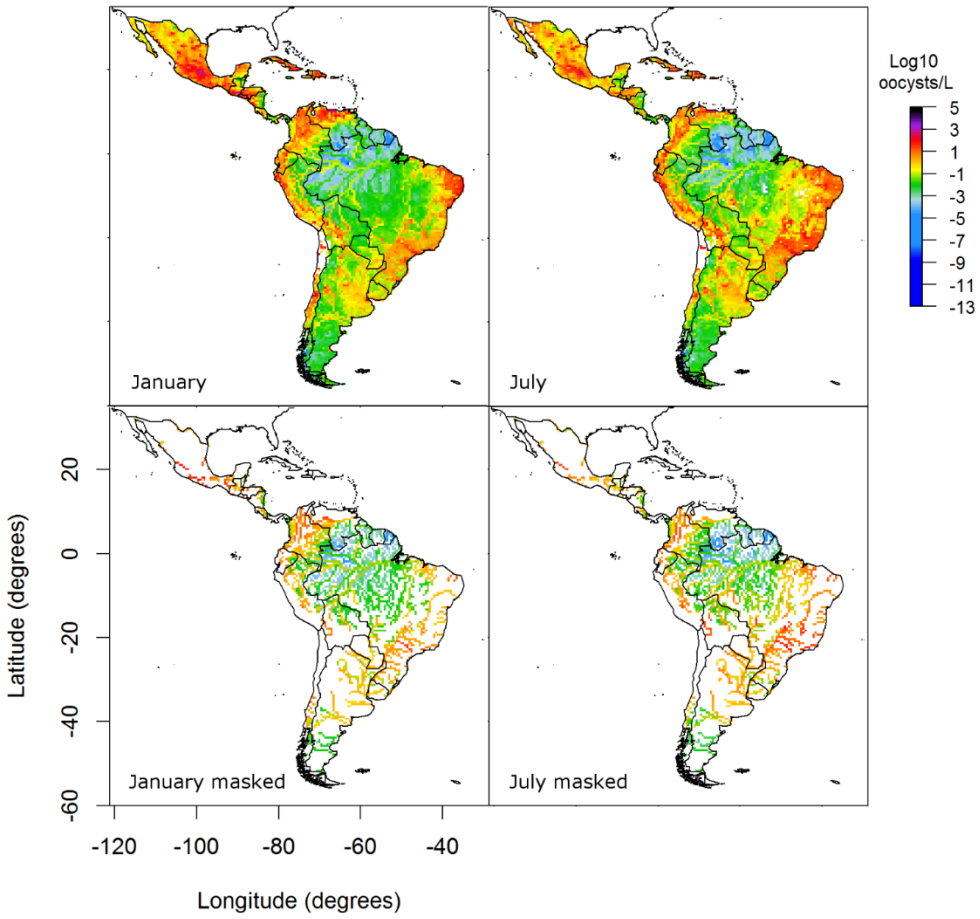


Figure S7.6 Modelled oocyst concentrations (log10 oocysts L^{-1}) in Latin America for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

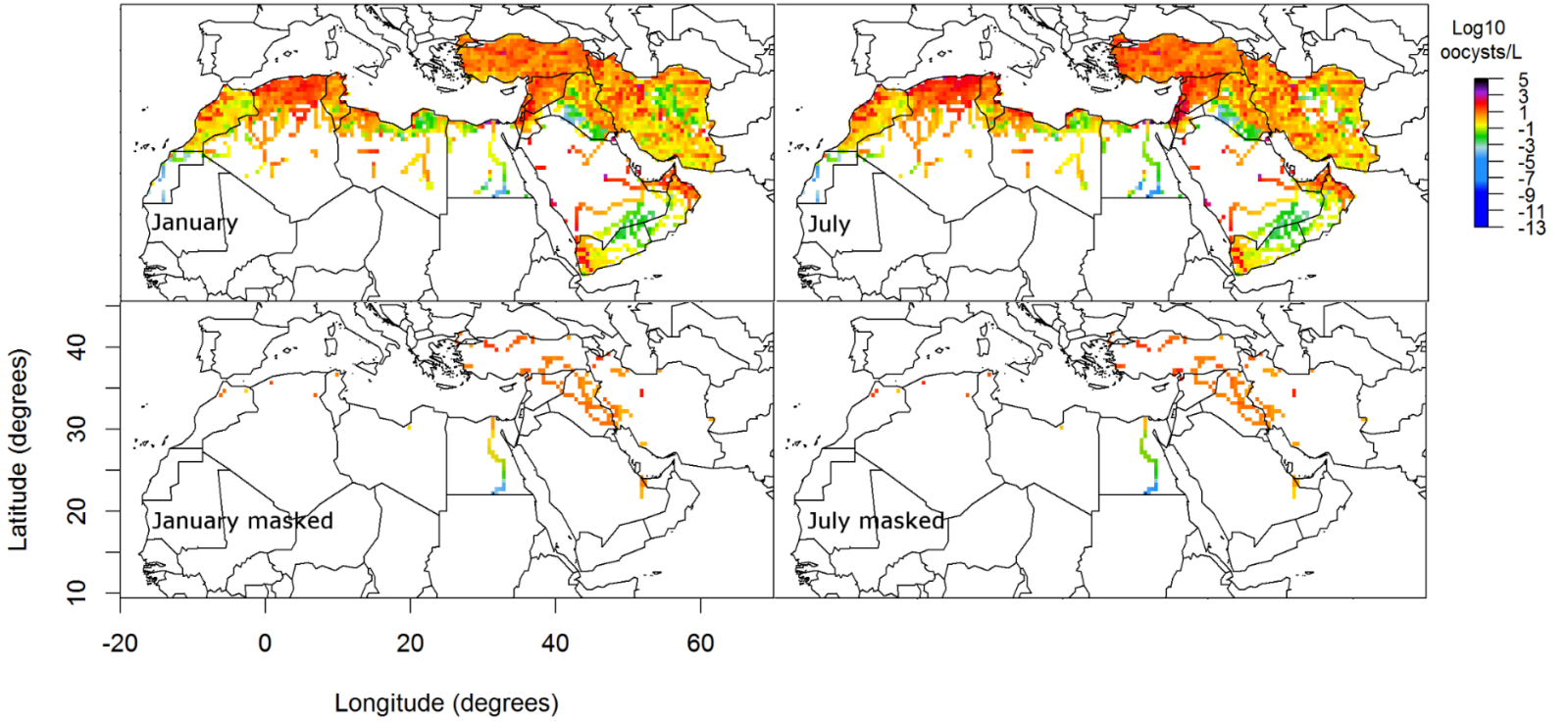


Figure S7.7 Modelled oocyst concentrations (log10 oocysts L^{-1}) the Middle East and North Africa for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

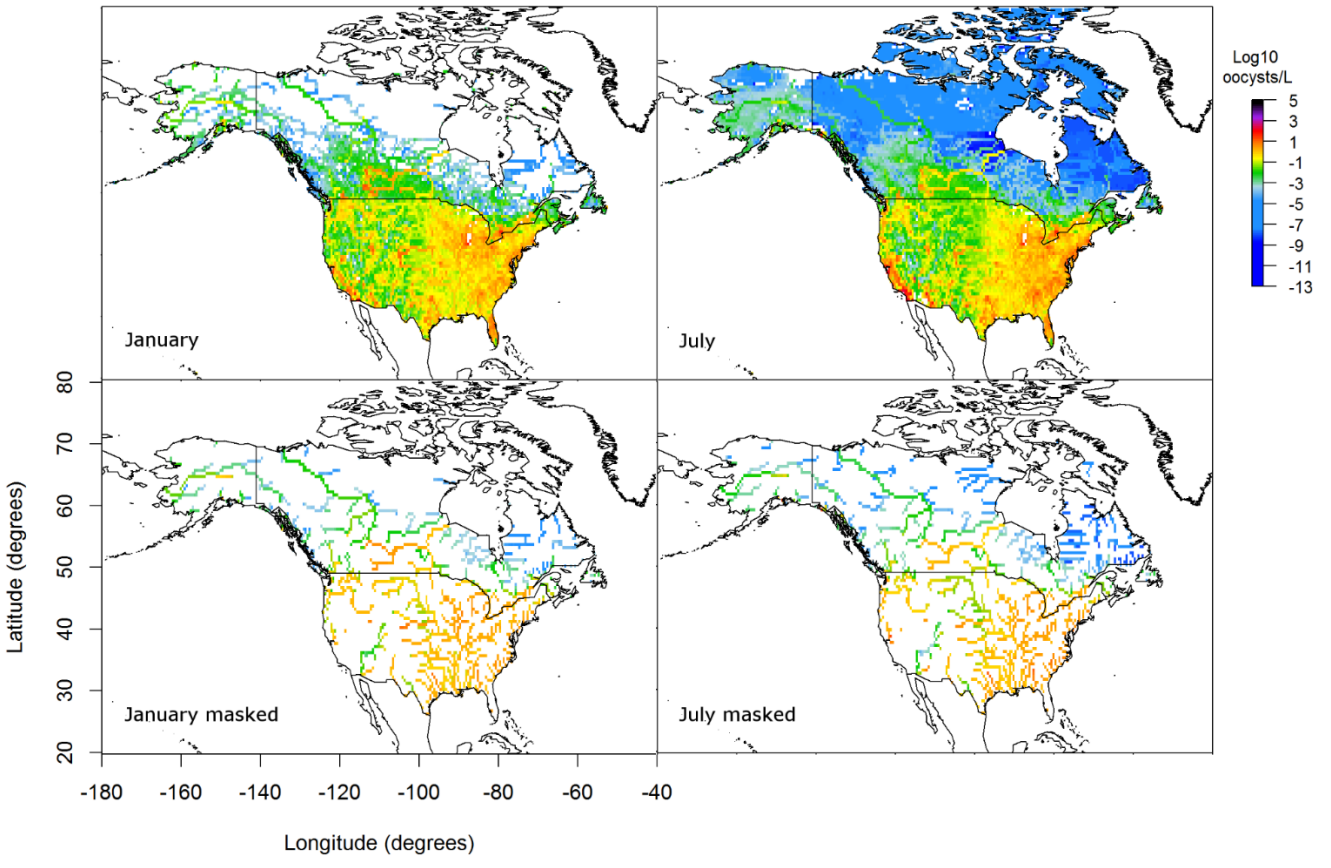


Figure S7.8 Modelled oocyst concentrations (log10 oocysts L^{-1}) in North America for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

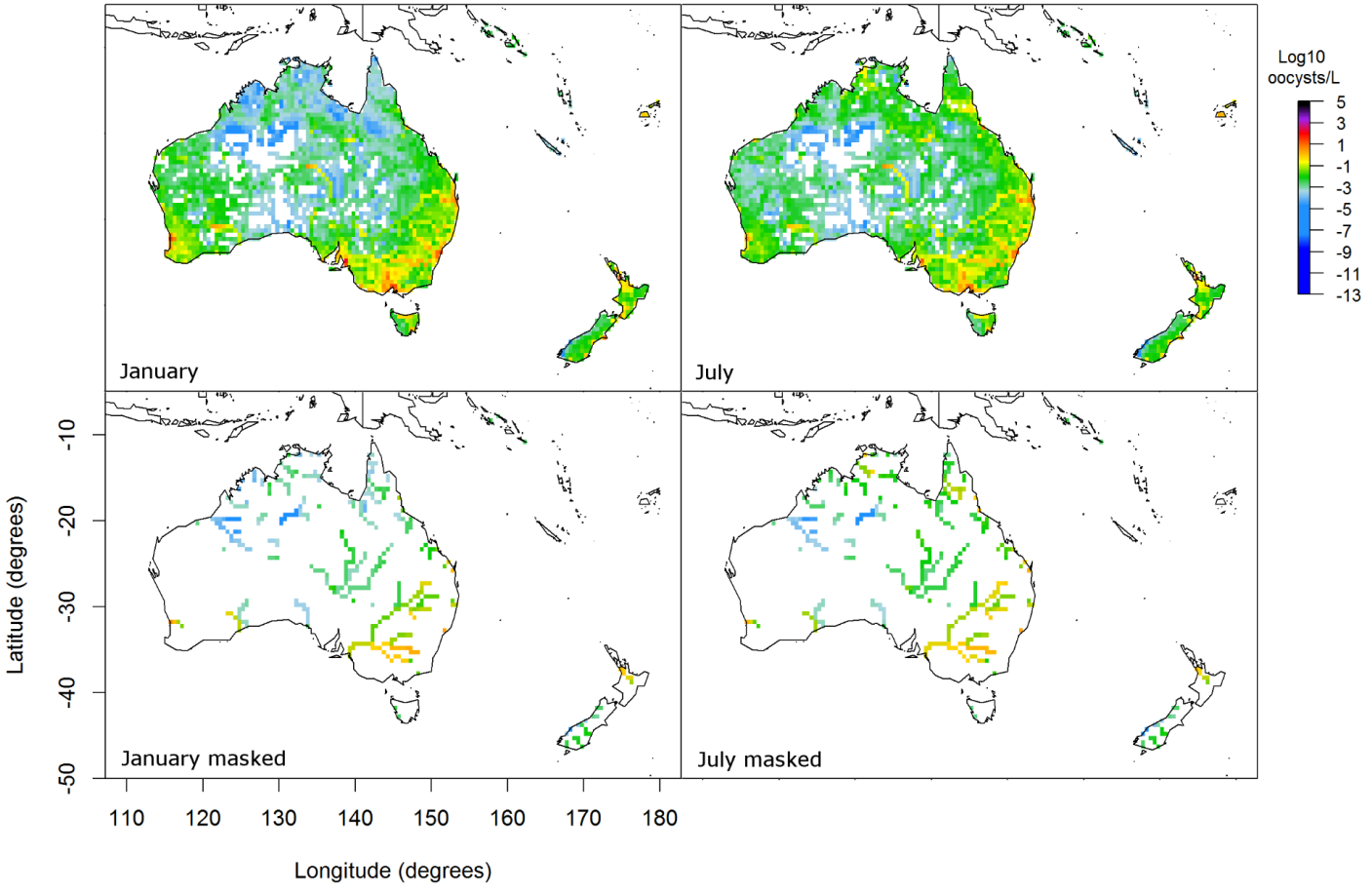


Figure S7.9 Modelled oocyst concentrations (log10 oocysts L^{-1}) in Oceania for January and July, both with and without a discharge mask (cells with average annual discharge <200 m^3 s^{-1} are masked).

S7.5 Model validation

We compare model outcomes to observational data of *Cryptosporidium* oocysts in rivers. Table S7.2 gives an overview of the data that were included, and Figure S7.10 provides a spatial visualisation of the locations. We included locations that concern flowing water, referred to by names as ‘river’, ‘canal’, ‘stream’ or ‘creek’, etc., and excluded locations with standing water, that were referred to as ‘lake’, ‘reservoir’ or ‘pond’, etc., as the model does not explicitly incorporate these.

The model predictions were generally found to be higher than the observational data. To analyse this somewhat further, we calculated for both the observed and modelled concentration pairs the WHO category to which they belong (1-6). We then calculated the difference between these categories for each pair, by subtracting the observed from the predicted category, and made a histogram of the result (Figure S7.11). This histogram shows that the difference in category appears to be normally distributed with a peak at 2, corresponding to the observed difference.

Table S7.2: Overview of the included validation data.

Country	N measurements	N locations	Main rivers/ water sources	Years	Organisation	Detection limit OR limit range in oocysts/L (% below detection limit)	Recovery efficiency (% for which recovery is specified)	Publication (if published)
Belgium	51	3	Meuse, Albertkanaal, Yser	1997-1998 2010-2012	Rijkswaterstaat, Ghent University	0.01-0.067 (27%)	0.026-0.44 (100%)	(Ehsan et al., 2015)
Brazil	295	29	19 rivers	2013-2016	Companhia Ambiental do Estado deSão Paulo, School of Public Health, University of São Paulo	0.1 (61%)	0.35-0.62 (71%)	
Canada	905	6	Milles Îles River, Ottawa River, Prairies River, Bow River	1994-2017	Polytechnique Montreal, City of Calgary	0.0028-1 (69%)	0.21-0.89 (70%)	(Lalancette et al., 2014)
France	360	16	10 rivers	1993-2017	Département Santé et Environnement	0.005-0.5 (61%)	NA	
Germany	342	41	Swist and tributaries	1999-2006	Klinikum der Universität Bonn, Institut für Hygiene und öffentliche Gesundheit	0.5 (81%)	NA	(Claßen et al., 2004; Kistemann et al., 2012; Rechenburg et al., 2009, 2006)
Japan	556	146	63 rivers and water sources	2008-2013	Kyoto University	1 (37%)	NA	
Luxembourg	94	10	Sûre and tributaries	2008-2012	Centre de Recherche Public - Gabriel Lippmann	0.1 (20%)	0.35 (100%)	(Burnet et al., 2014, 2015)
The Netherlands	150	8	Meuse, Rhine	1997-1998 2002-2010	Rijkswaterstaat	0.01 (45%)	0.026 (41%)	
New Zealand	638	22	20 rivers and water sources	1998-2000	Environment Bay of Plenty, Environment Canterbury, Environment Southland,	0.01 (95%)	NA	(Till et al., 2008)

					Environment Waikato, Hawkes Bay Regional Council, Horizons Regional Council, Marlborough District Council, National Institute of Water & Atmospheric Research, Northland Regional Council, Otago Regional Council, Wellington Regional Council			
Thailand	96	48	Kuang, Lai, Pong, San (8 rivers and water sources)	2014	National University of Singapore	0.05 (83%)	0.39 (100%)	(Chuah et al., 2016)
United States	851	55	40 rivers and water sources	2015-2017	Long Term 2 Enhanced Surface Water Treatment Rule	0.05-1 (79%)	NA	

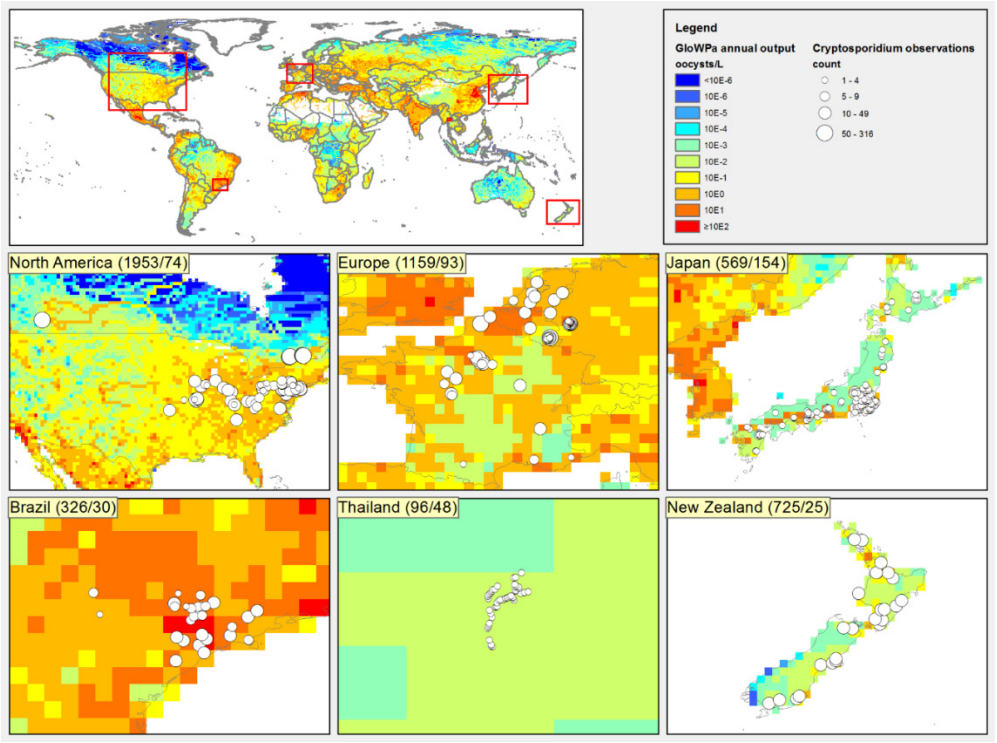


Figure S7.10: Spatial overview of the validation data set. Locations for which observational data were available are shown with white circles, the size indicates the number of observations in that location. The background map is the annual average oocyst concentration predicted by GloWPa-Crypto C1, as reference. The top left panel shows the entire world, the red squares indicate the regions for which data were available, these are shown zoomed in in the bottom six panels. The numbers behind the region names are the total number of observations and the number of locations (before exclusion of the locations with standing water).

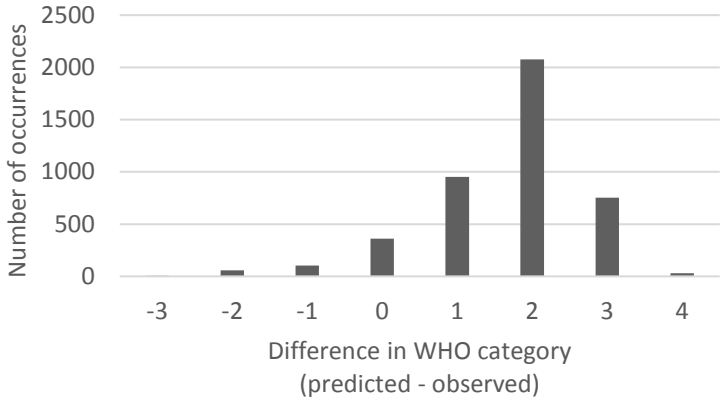


Figure S7.11: Difference in WHO category for predicted and observed oocysts concentrations.

We looked at the results separated for different months, to see if we could observe any seasonality. The following figures (S7.12-S7.14) are examples for Brazil, Japan and the USA of boxplots of the observed (top panel) and predicted (bottom panel) concentrations split by month. For the observational data, these were generated using the function 'cenboxplot' from the NADA package in R, that estimates the distribution for values below the detection limit, assuming a lognormal distribution (Lee, 2017). We conclude that no obvious seasonal patterns are present in the observational data. The model predicts slight seasonal variation, but not very much.

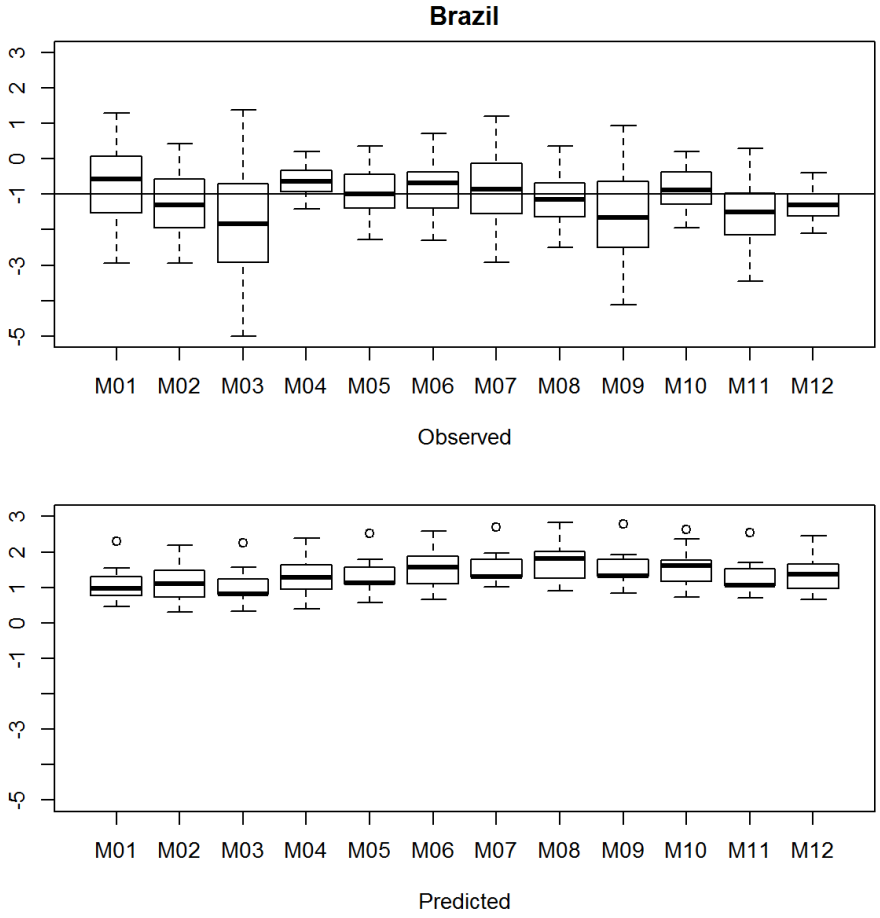


Figure S7.12: Boxplots of observed (top panel) and predicted (bottom panel) oocyst concentrations by month for Brazil.

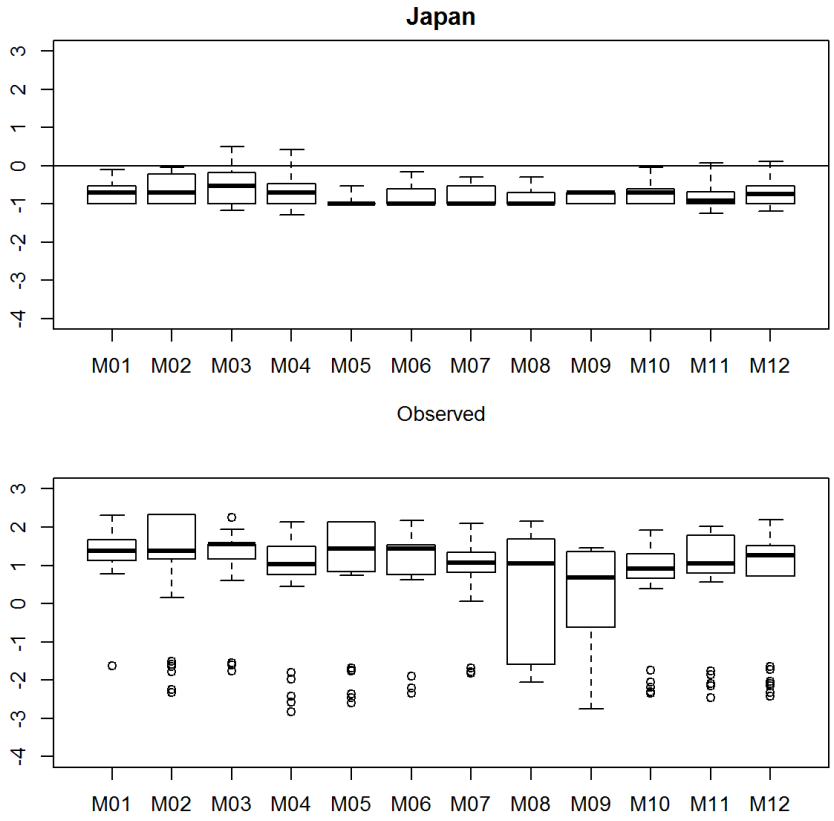


Figure S7.13: Boxplots of observed (top panel) and predicted (bottom panel) oocyst concentrations by month for Japan.

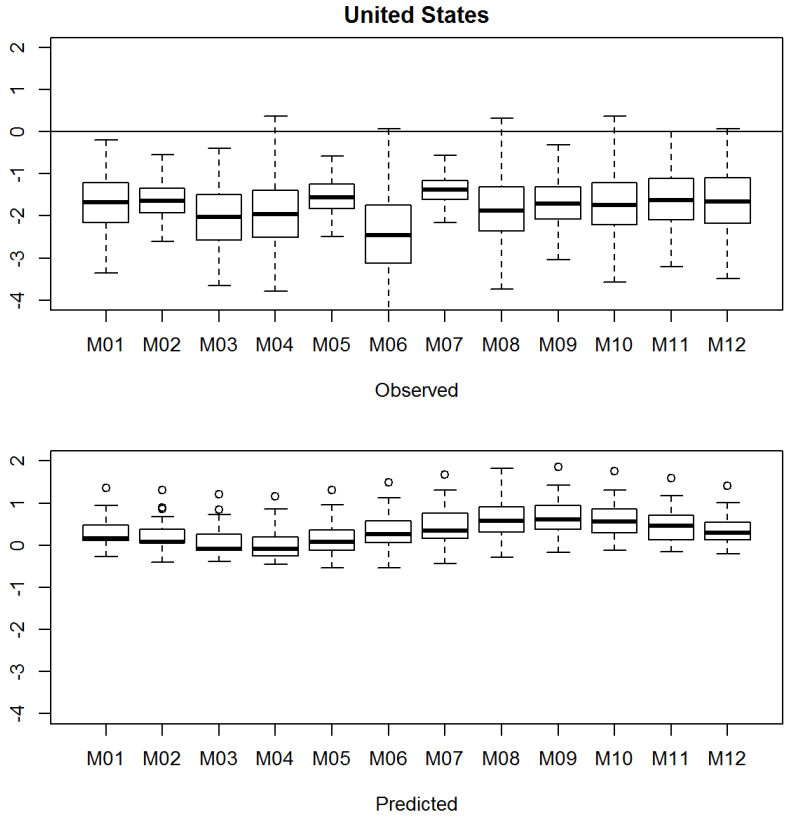


Figure S7.14: Boxplots of observed (top panel) and predicted (bottom panel) oocyst concentrations by month for the USA.

S7.6. Sensitivity analysis

S7.6.1 Methods

Human and animal oocyst loads are changed by 1 log unit, in the same range as was done in the GloWPa-Crypto H1 and L1 models (Chapters 4 and 6). Discharge, runoff, solar radiation and water temperature are varied by 20%. Retention of oocysts on land is more uncertain and is varied over several log units, as oocysts transport with runoff is influenced by many processes and observational data are scarce. The influence of oocyst survival is tested by varying the parameters K_4 and λ according to Peng et al. (2008), varying sedimentation velocity according to Medema et al. (1998), halving and doubling river depth and DOC concentrations, and leaving out components of the survival (solar radiation-dependent, temperature-dependent, sedimentation). Furthermore, river length and water residence time are increased 5 and 10 times. These are only increased upward as river length is the theoretical minimum flow distance (based on midpoint straight line

distance between grid cells), and residence time is computed from that. Moreover, the model is currently based on naturalized discharge, meaning that the presence of dams, lakes and reservoirs in rivers is not taken into account, so likely actual water residence times are longer. Lastly, low end and high end runs were performed, in which all changes were combined that have respectively increasing and decreasing effects on concentrations.

S7.6.2 Results and discussion

Table S7.3 shows the input changes and their effect on model output (median and mean oocyst concentrations, global annual averages). The model is most sensitive to changes in input human oocyst loads (Table S7.3) and to a lesser extent to input animal loads, as human loads were found to dominate in most places, this was discussed in the main text. To a lesser extent, the model was sensitive to oocyst retention, river length and water residence time in a grid cell, these affect the time during which decay and sedimentation take place. This will be discussed here in more detail.

Oocyst retention in the faecal matrix and on soils is a complex process that is likely influenced by many factors (such as vegetation, slope, precipitation intensity, manure characteristics, soil characteristics such as porosity and clay content, distance to the nearest water course, etc.), factors that are difficult if not impossible to capture at a 0.5x0.5° scale. Additionally, experimental evidence on oocyst transport with surface runoff is very thin, especially under field conditions. For this reason we chose the crude approach of linearly allocating a log retention value to a corresponding surface runoff value, ignoring all other landscape characteristics. River length and water residence time in a grid cell affect the time period under which decay and sedimentation take place. Due to the theoretical minimum flow distance and the fact that dams, lakes and reservoirs are not taken into account in the model, actual residence times are likely longer, meaning that decay could be more influential than is currently the case in the model. Increasing oocyst settling velocity also affects model outcomes relatively strongly, which has similarly been observed in the WorldQual model, that uses the same equation for sedimentation but with a different value for settling velocity (Reder et al., 2017). However, changes in K show that not sedimentation but solar radiation-dependent decay comes out as most influential. This is in line with literature suggesting sunlight may be the most potent environmental stressor influencing oocyst survival (Monis et al., 2014).

The model is also quite sensitive to changes in river discharge, but less to surface runoff. The impact of discharge and surface runoff on simulated concentrations are opposite in direction, increased discharge leads to decreased concentrations due to dilution, while increased runoff leads to increased concentrations due to increased oocyst release from

manure and transport. The dilution effect seems to be stronger than the increased release effect, as increasing both discharge and surface runoff leads to decreased concentrations. Other models have also observed that microbial concentrations are influenced more by dilution than release (Rankinen et al., 2016; Sterk et al., 2016). GloWPa-Crypto C1 is not very sensitive to changes in riverine DOC concentrations, solar radiation intensity, water depth, water temperature, K_4 and λ.

The low end and high end runs lead to 1 - 2 orders of magnitude change in outcomes in either direction. It should be noted that all of these observations are based on changes in global annual averages, while effects could vary in magnitude over time and space. For instance, which loss component (solar radiation, temperature or settling) is most influential is likely to vary depending on local conditions, e.g. in warm tropical waters with high DOC content, temperature might be much more influential than solar radiation. Reder et al. (2017) performed a spatially explicit uncertainty and sensitivity analysis of their continental faecal indicator model WorldQual for Africa, that shows that most sensitive parameters differ geographically. Nevertheless, they also note that only four parameters dominate output uncertainty in 93% of the continent, meaning that their overall conclusions on the importance of different parameters were not affected much by the geographical analysis. Although a spatially explicit sensitivity analysis would certainly be interesting to do with GloWPa-Crypto as well, we are confident that these first results give a general feel for the model sensitivity to changes in input parameters.

Table S7.3: Sensitivity analysis. This table shows the variable (column 1), its baseline value (column 2) and how it was changed in the sensitivity analysis (column3). The effect of the change is on model output is given for median (column 4) and mean (column 5) oocyst concentrations. The changed output was divided over the baseline output, meaning that a value of 1 indicates no change, a value above 1 an increase and a value below 1 a decrease.

Variable	Variable baseline value	Variable changed value	Median concentration change	Mean concentration change
Animal oocyst loads	Gridded, as in L1 model	+ 1 log	1.79	1.32
		- 1 log	0.86	0.97
Human oocyst loads	Gridded, as in H1 model	+ 1 log	8.57	9.68
		- 1 log	0.18	0.13
Both human and animal oocyst loads	Gridded, as in L1 & H1 model	+ 1 log	10.00	10.00
		- 1 log	0.10	0.10
Discharge (Q)	Gridded, from VIC	+ 20 %	0.84	0.84
		- 20 %	1.23	1.22
Runoff	Gridded, from VIC	+ 20 %	1.02	1.06
		- 20 %	0.98	0.98
Discharge (Q) & Runoff	Gridded, from VIC	+ 20 %	0.86	0.89
		- 20 %	1.21	1.20
	3.2-8.8 log	Wider (2-10 log)	0.87	1.13

Retention on land		Narrower (5-7 log) Lower (2.2-7.8 log) Higher (4.2-9.8 log)	3.80 1.79 0.86	1.09 1.12 0.98
K_4 and λ	0.0051 and 0.158	0.0029 and 0.118[a] 0.0093 and 0.095[b]	1.02 1.01	1.01 1.01
K	Gridded, calculated from K_T, K_R and K_S	Changed to 1, meaning no decay or sedimentation, all survive	1.15	1.10
K_T	Gridded, calculated from T_w	Only T-dependent decay (K_R and K_S changed to 1)	1.12	1.08
K_R	Gridded, calculated from DOC, solar radiation, and river depth	Only solar radiation-dependent decay (K_T and K_S changed to 1)	1.04	1.03
K_S	Gridded, calculated from v and river depth	Only sedimentation (K_R and K_T changed to 1)	1.12	1.07
Water temperature (T_w)	Gridded, RBM model output	+ 20 % - 20 %	0.98 1.01	0.99 1.01
River depth (Z)	Gridded, calculated from discharge	x 2 x 0.5	1.03 0.95	1.03 0.95
Length of river stretch (l)	Gridded, calculated from midpoint straight line distance between adjacent cells	x 10[c] x 5[c]	0.62 0.76	0.78 0.85
River residence time (t)	Gridded, calculated from river geometry	x 10[d] x 5[d]	0.62 0.76	0.78 0.85
DOC	River basin values, from Global NEWS	x 2 x 0.5	1.02 0.97	1.02 0.97
Solar radiation (I)	Gridded, WATCH forcing data shortwave radiation	+ 20 % - 20 %	0.99 1.01	0.99 1.01
Settling velocity (v)	0.1	0.035[e] 3.5[e]	1.01 0.61	1.01 0.79
High end		Changes above combined	25.76	16.24
Low end		Changes above combined	0.02	0.06

[a] Temperature-dependent decay as in sterilized water with DAPI/PI or excystation analysis (Peng et al., 2008). [b] Temperature-dependent decay as in river water with DAPI/PI or excystation analysis (Peng et al., 2008). [c] As this is the minimal theoretical flow distance, it was changed upward only [d] As lakes and reservoirs are not included and these likely have longer residence times, this was changed upward only. [e] Settling velocities of free oocysts and oocysts attached to suspended particles (Medema et al., 1998).

Summary

Diarrhoeal disease is very common around the world. Diarrhoea is still the fourth leading cause of death among children under five years worldwide and ranks sixth in global burden of disease among all ages. Enteric pathogens that cause diarrhoea spread via the faecal-oral route, and transmission through the environment is common. The contamination of surface waters with faeces is an important pollution problem, and improvements in water, sanitation and hygiene (WASH) are needed to reduce the burden of diarrhoeal disease. Knowing more about the global burden of disease caused by enteric waterborne pathogens and about the geographical distribution of pollution is important for human health decision making and water and sanitation planning in the context of meeting the Sustainable Development Goals (SDGs). *Cryptosporidium* is one of the most important and common diarrhoeal pathogens worldwide, and it is the main focus of this thesis.

The global burden of disease of cryptosporidiosis is difficult to estimate and very uncertain, and observational data of *Cryptosporidium* in the environment are scarce. Global modelling can be a helpful tool to fill the gaps of regions with poor data. It can also help to identify where and when monitoring would be most needed. Modelling is based on understanding of the sources, pathways and processes of pathogens to and in surface waters, and understanding of the effect of control measures, such as wastewater or manure treatment. Such understanding helps to design effective control strategies.

The objective of this thesis is to increase our knowledge on the sources, fate and transport of *Cryptosporidium* in rivers worldwide using spatially explicit modelling. In this thesis, I present the Global Waterborne Pathogen model for *Cryptosporidium* (GloWPa-Crypto), the first global model of waterborne pathogen emissions to and concentrations in rivers. For the development of this model lessons have been drawn from existing watershed-scale waterborne pathogen models and global nutrient export models (Chapter 2). Different approaches from literature have been summarised and desirable characteristics for a global waterborne pathogen model have been identified. GloWPa-Crypto is developed largely following the conceptual model presented in Chapter 2. GloWPa-Crypto is a deterministic process-based model that contains some empirical parts. It is uncalibrated, static, spatially explicit, implemented at the global scale, and has a spatial resolution of 0.5 x 0.5 degree grid and a monthly time step. These choices were made with the model objectives in mind and based on pragmatic reasons, for instance regarding data availability.

The focus in Chapters 3 and 4 is on human emissions of *Cryptosporidium* to surface waters, and on the impact of sanitation changes, urbanisation and population growth on these emissions. First the model was developed for Bangladesh and India as a case study (Chapter 3), then the model was implemented for the world (GloWPa-Crypto H1, Chapter 4). We find that urbanised areas in developing countries are pollution hotspots. We conclude from our scenario analysis using Shared Socioeconomic Pathway (SSP) scenarios for 2050 that future population growth and urbanisation are likely to lead to further deterioration of water quality. If population, urbanisation and sanitation change according to SSP1 in 2050, total global *Cryptosporidium* emissions to surface water are expected to reduce by 24%. However, in the Middle East and Africa region, India and Pakistan emissions are expected to increase for SSP1 due to a combination of population growth and urbanisation. For SSP3 the total global emissions in 2050 are expected to increase by 52%, only some developed countries see reduced emissions due to population decreases. Sanitation and treatment improvements are key to improve water quality now and in the future. If population, urbanisation and sanitation change according to SSP1 but wastewater treatment would not be improved, emissions for nearly all countries are even higher than the emissions calculated for SSP3. We conclude that connecting more people to sewers without improving wastewater treatment will deteriorate river water quality.

Livestock is the focus in Chapters 5 and 6. In Chapter 5, published prevalence of *Cryptosporidium* infection and oocyst concentrations in livestock manure were summarized from the literature and their distribution is analysed by livestock age and region. Prevalence and concentrations are found to be high worldwide, especially among young animals. Oocyst concentrations in manure are highly variable. The findings from this literature inventory were then implemented in the global scale model to calculate oocyst loads to land (GloWPa-Crypto L1, Chapter 6). Livestock density maps are combined with prevalence and oocyst concentration estimates from the literature review, and temperature-dependent oocyst decay during manure storage is accounted for. We find that globally, cattle (especially calves) are the dominant livestock source of oocysts, followed by chickens and pigs. The livestock oocyst load from the developing world is larger than from the developed world, and comes less often from stored manure. The effect of manure storage and treatment on oocyst loads is explored using scenario analysis. The scenario analysis shows that although manure storage halves oocyst loads, manure treatment, especially of cattle manure and particularly at elevated temperatures, has a larger load reduction potential than manure storage. Regions with high reduction potential include India, Bangladesh, Western Europe, China, several countries in Africa, and New Zealand.

Human and animal oocyst loads were combined with output from the VIC hydrological model, to estimate *Cryptosporidium* oocyst concentrations in rivers worldwide (GloWPa-

Crypto C1, Chapter 7). Transport of oocysts from manure and faeces on land with surface runoff is simulated, oocysts are routed through the stream network, and temperature- and solar radiation-dependent decay and loss due to sedimentation are accounted for. We find that monthly average oocyst concentrations are predicted to range from 10^{-6} to 10^{2} oocysts/L in most places. Hotspot regions with high concentrations include parts of India, China, Pakistan and Bangladesh, Nigeria, Algeria and South Africa, Mexico, Venezuela and some coastal areas of Brazil, several countries in Western and Eastern Europe, and the Middle East. Point sources (human faeces) appear to be a more dominant source of pollution than diffuse sources (animal manure) in most world regions. In Chapter 8 the results are discussed in the context of global change. Future average oocyst concentrations in rivers are likely influenced more by changes in population, sanitation and treatment than by climatic changes. Scenario analysis is the tool that will enable us to investigate linkages quantitatively and pinpoint regions that are likely to be most at risk in the future.

Validation with observational data, comparison with other modelling studies and sensitivity analysis are applied to enhance confidence in model performance. Observational data are needed to inform and validate models, and modelling can fill the gaps and highlight where observational data are most needed. Our various nominal range sensitivity analyses indicate that sensitive parameters include oocyst excretion by humans and to a lesser extent by animals, the share of the population using different types of sanitation facilities, the proportion of manure going to storage and to land, oocyst retention in the faecal matrix and on soils, and parameters affecting the time during which loss (decay and sedimentation) during river transport takes place (river length and residence time).

I performed a first exploration of risk, burden of disease and health-based treatment targets based on the GloWPa-Crypto model in Chapter 8. The risk calculation indicates that when drinking 2 L per day of raw surface water, the annual risk of getting cryptosporidiosis is 1 in most world regions, meaning that everyone would get infected at least once per year. The resulting cryptosporidiosis disease burden, when taking into account the population actually depending on raw surface water for their drinking water, is calculated to be largest in Papua New Guinea, sub-Saharan Africa, and to a lesser extent in Asia and Latin America. To reach the World Health Organization health-based target of 10^{-6} Disability-Adjusted Life Years (DALYs) per person per year for the consumption of drinking water produced from surface water, our model suggests that drinking water treatment reaching 4-7 log removal of oocysts is needed for most world regions, and up to 9.4 log for the most polluted regions.

The world faces a tremendous challenge to meet the Sustainable Development Goals and ensure safe water for all. I have shown the contribution that global waterborne pathogen modelling could make in confronting this challenge (Chapter 8). Facing the limitations of observational data, it is apparent that water quality modelling could be a complementary source of information, to help assess progress towards SDG indicator 6.3.2 'Proportion of bodies of water with good ambient water quality'. Furthermore, modelling can be a tool to calculate water quality impacts of meeting the SDGs. Models can help formulate effective policy options that account for feedbacks in the system and trade-offs and synergies that follow from these feedbacks. Looking at it from another side, model development can also benefit from the data generated in assessing a baseline and progress towards SDG indicators.

In summary, waterborne pathogen modelling at the global scale is feasible and provides new opportunities: to provide information on pathogen concentrations in data-sparse regions, identify hotspot regions with high concentrations, identify the relative contribution of different sources, and allow for scenario analysis. Examples of these opportunities have been demonstrated in the research chapters of this thesis. More opportunities were identified in the synthesis chapter, including application of the model in the analysis of risk, burden of disease and health-based treatment targets. Global water quality modelling could make a valuable contribution in meeting the Sustainable Development Goals.

Acknowledgements

This PhD has been an eventful journey. Many people have helped and supported me over the past years in many different ways, and of course they deserve thanks! The result would not have been the same without you.

First, my supervisor and co-promotor Nynke. Nynke, you are a fantastic supervisor. Your kindness, pragmatism, and sharp mind have greatly helped my PhD thesis (and before that, my MSc thesis) forward. You are the one who came up with the idea of waterborne pathogen modelling at the global scale, and I am glad I had the opportunity to help bring this idea to where it is now. I'm very happy with your optimism and openness, you make it easy to share. We always had fun as well, with dinners and even dancing during trips to conferences. I'm happy to have gotten to know you! Many thanks also to my promotors Carolien and Gertjan, the other two members of my wonderful supervision team. Carolien, you are very knowledgeable and strategic, but most of all I appreciated your kindness and warmth. You seem to radiate a calmness that comforts your students. Gertjan, I'm really happy you were involved in this project. Without your knowledge and experience the quality of my PhD thesis would not have been what it is now. You are very sharp, and your comments always triggered me to think more, explain better, do extra analyses ;) You too are a very kind person, stimulating and supportive.

Other people and organisations that have been important for the content of this thesis include: Michelle for discussions on hydrology, Emiel for processing the validation data, the many individuals and organisations who supplied data (listed in Chapter 7), graduate school WIMEK/SENSE and the WIMEK PhD council, the people from KNMI and SA Water, Jochem and Huub for discussions on decay and UV radiation, the EU Climate KIC programme for financial support and broadening my horizon, the Gerbrand de Jong Fonds for financial support, the Gezondheidsraad for the opportunity to do an interesting project on the side, Tetje, Merel, Huub and my dad for proof-reading sections of the thesis, Aritta for help with InDesign and last but not least Merel for the chapter page and cover art. During my PhD I have supervised several students for their BSc or MSc thesis. Some of them are now co-author of chapters in this book! So thanks to Jorien, Jelske, and Marijke.

The ESA group has been a nice working environment for the past years. Thanks to all colleagues for talks, coffee breaks, Sinterklaas celebrations, Christmas dinners, annual outings and New Year's drinks: Rik, Mathilde, Gerbert, Arnold, Wichertje, André, Karen, Lars, Dolf, Wim, Niklas, Bas, Sophie, Rob, Lenny and Marjolein. Special thanks to Rik, Arnold and Wichertje for giving me my first job at ESA/FSD (despite the spelling mistake in

the title of my try-out Natuurbericht). The current and former PhD students of the ESA group are a fun and supportive bunch of people, we shared international potluck dinners, drinks, parties, hikes, kubb games, we helped each other, and shared the bad moments also. Thank you all: Lena, Aritta, Sarahí, Eka, Jillian, Confidence, Leonardo, Gabriela, Jerry, Maddy, Halima, Alexey, Anna, Sofya, Siatwiinda, Dian, Saritha, Maryna, Mengru and Ang, and my ex-ESA friends: Alexander, Roy, Ingeborg, Matthias, Clara, Martine, Cheng, Shahid, Majed, and Julia. Special thanks go to colleague and friend Lena, who is willing to be paranymph during her maternity leave! Lena, I also want to thank you for your endless creativity and the fun we have had preparing the movie for Roy and other creative projects :)

Life is more than studying and working! Friends are needed for moral support and relaxation, and music is one of my favourite pastimes. The G104 housemates (Tess, René, Maarten, Lena and Annemarie) have been important over the past years, thanks for all the support, games, trips, food & drinks we shared! Robert and Jeroen, all the wine and coffee, here and in Italy, have been indispensable. Thanks to my friends from UCU (Annelli, Marjanne, Lidwien, Anneke, Jolien, Daan, Sophie, Joost, Roel, Barend) and high school (Sanne, Janine, Ilona) for the tea, dinners, games, in other words, the fun and moral support! Thanks to the ladies of my vocal group La Bomba (Wietske, Margo, Liza and Floor), we started from zero but have done some pretty nice performances over the last years! I learned a lot and made many friends while playing the trumpet at student orchestra de Ontzetting (Jochem, Alie, Sybren, Ina, Raymond and many others, thank you all). De Ontzetting shaped my experience of Wageningen as a student. LGBT+ organization SHOUT has been the host of many of the nicest parties I have been to in recent years. Being the education coordinator of SHOUT I find very meaningful, thanks also to all members of the education team for the fun and motivation!

Without my family, I would not have been where I am now. Pap en mam, you have been endlessly loving and supportive throughout my whole education, my hobbies, travel, everything. Huub, thanks for being my paranymph! Merel, thanks for your creativity and love. Thank you all for being there for me! During my PhD I have also gained some new family: thanks to my parents in law Thies and Hennie for making me feel welcome, and Willem & Janneke for all the fun! And Marleen and Freek, nice that you have joined the family!

Last but not least, my partner Tetje. You are the one who had to comfort me in the lows, and who I celebrated the highs with. Our relationship started one week before my PhD started, so you have been with me for this entire journey, from summer 2016 as my wife. I'm very happy we're going strong. Love you!

About the author

Lucie Catharina Vermeulen was born on 31 May 1988 in Zaanstad, the Netherlands. She attended primary school in Wormer, and high school St. Michael College in Zaandam, where she did a double profile (Nature & Technology and Nature & Health) and graduated cum laude in 2006.

She then followed a BSc program in Liberal Arts & Sciences at University College Utrecht, majoring in Life Sciences (mainly Biology), with some courses in Spanish, Development Studies and Music on the side. Part of this programme were an exchange to Macquarie University in Australia and an internship in Tanzania. She graduated summa cum laude in 2009.

In this same year, she moved to Wageningen and started her MSc Environmental Sciences. As part of this programme she did an internship at the United Nations University Institute for Water, Environment and Health (UNU-INWEH) in Canada, focussing on the impact of global environmental change on human health and coastal ecosystems. Her MSc thesis at the Environmental Systems Analysis Group focused on the impact of climate variables on the bacterium *Escherichia coli* in Dutch rivers. She graduated cum laude in 2012. Also in this year, Lucie worked at the ESA group as teaching assistant and at the Foundation for Sustainable Development doing work for Natuurbericht and Natuurkalender.

In January 2013 Lucie started her PhD on global waterborne pathogen modelling, also at the ESA group. During her PhD she took part in the EU Climate KIC programme and the SENSE research school, did a project for the Health Council of the Netherlands (Gezondheidsraad), presented her research at several international conferences and supervised multiple thesis students. In February 2018, Lucie started a new job at research institute RIKILT in Wageningen.

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

For specialised PhD training

The Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Lucie Catharina Vermeulen-Henstra

born on 31 May 1988 in Zaanstad, The Netherlands

has successfully fulfilled all requirements of the
Educational Programme of SENSE.

Wageningen, 16 February 2018

the Chairman of the SENSE board	the SENSE Director of Education
Prof. dr. Huub Rijnaarts	Dr. Ad van Dommelen

The SENSE Research School has been accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW)

KONINKLIJKE NEDERLANDSE
AKADEMIE VAN WETENSCHAPPEN

The SENSE Research School declares that **Ms Lucie Catharina Vermeulen-Henstra** has successfully fulfilled all requirements of the Educational PhD Programme of SENSE with a work load of 51 EC, including the following activities:

SENSE PhD Courses

- iGIS A practical post-graduate GIS course (2013)
- Environmental research in context (2013)
- Research in context activity: 'Contributing analysis to reviewing of IPCC chapter under coordination of the Netherlands Environmental Assessment Agency (PBL)' (2013)

Other PhD and Advanced MSc Courses

- Techniques for writing and presenting a scientific paper, Wageningen University (2013)
- PhD competence assessment, Wageningen University (2013)
- Writing a data management plan, Wageningen University (2014)
- Summer school The Journey, Climate KIC (2014)
- Reviewing a scientific paper, Wageningen University (2015)
- Risk assessment of infectious agents, Institute for Risk Assessment Sciences (IRAS) Utrecht University (2015)
- Teaching and Supervising thesis students, Wageningen University (2016)
- Using GIS in disease control programs, Royal Tropical Institute (KIT) Amsterdam and International Institute for Geo-Information Science and Earth Observation (ICT) Enschede (2016)

External training at a foreign research institute

- Research visit to compare modelling approaches, Centre for Environmental Systems Research, Universität Kassel, Germany

Management and Didactic Skills Training

- Supervising BSc student with thesis entitled 'Improving point sources of Cryptosporidium in the Hofstra model by looking at sanitation worldwide now and in 2050' (2014)
- Supervising two MSc students with thesis entitled 'Climate effects on pathogen delivery and survival in rivers: a global modelling approach' (2015) and 'The influence of present and future livestock production on the Cryptosporidium emissions to land' (2016)

Oral Presentations

- *Global modelling of Cryptosporidium from livestock manure in surface water.* 18th International Symposium on Health-Related Water Microbiology, 14-18 September 2015, Lisbon, Portugal
- *Global waterborne pathogen modelling.* 7th International Water & Health seminar, 29 June-1 July 2015, Cannes, France

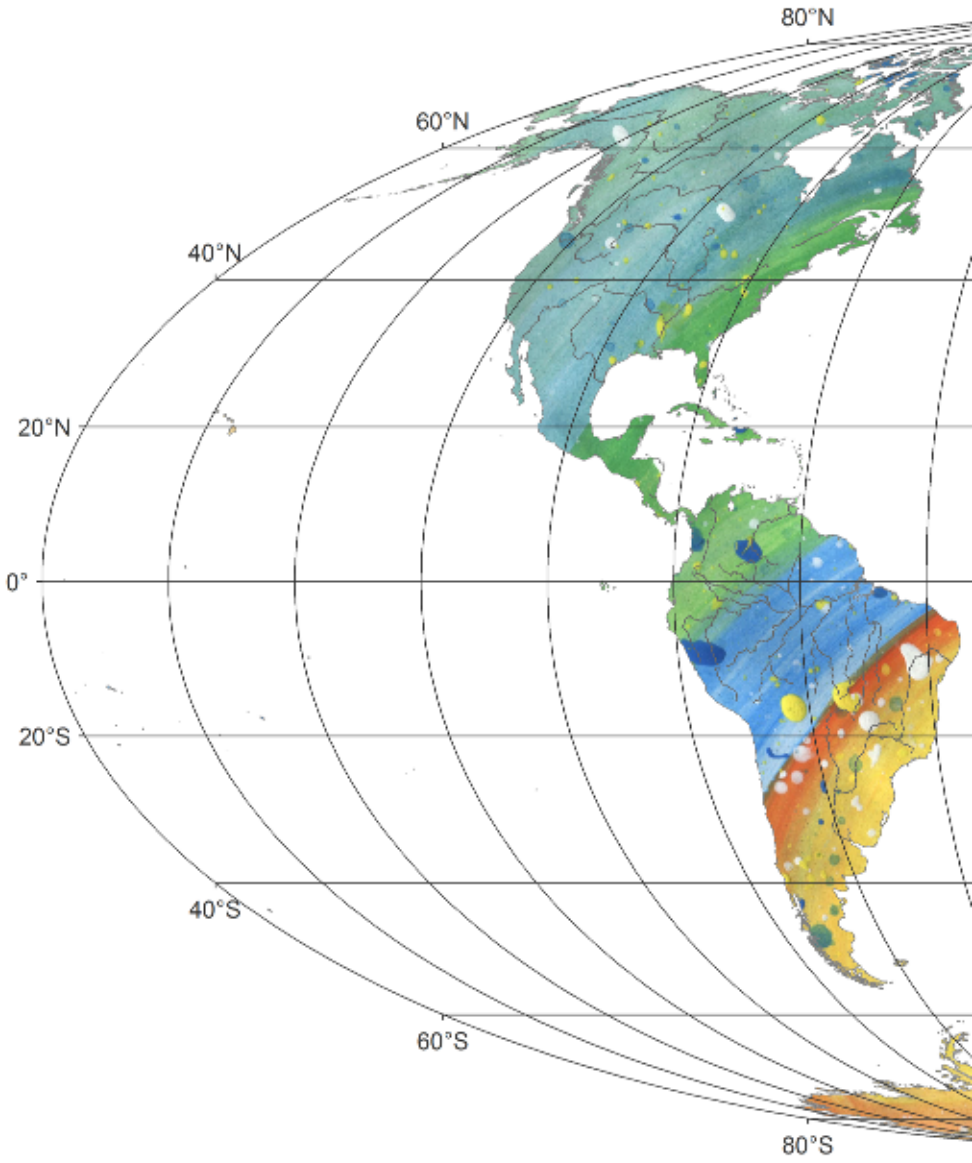
SENSE Coordinator PhD Education

Dr. Monique Gulickx

The research described in this thesis was financially supported by the EU Climate KIC program and the Gerbrand de Jong Fonds. Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Merel Vermeulen and Lucie Vermeulen

Printed by GVO drukkers & vormgevers B.V. on FSC-certified paper


80°N
60°N
40°N
20°N
0°
20°S
40°S
60°S
80°S