Contents lists available at ScienceDirect

Water Research

Cryptosporidium concentrations in rivers worldwide

Lucie C. Vermeulen [a, *], Marijke van Hengel [a], Carolien Kroeze [b], Gertjan Medema [c, d], J. Emiel Spanier [b], Michelle T.H. van Vliet [b], Nynke Hofstra [a]

[a] *Environmental Systems Analysis Group, Wageningen University, P.O. Box 47, 6700 AA, Wageningen, the Netherlands*
[b] *Water Systems and Global Change Group, Wageningen University, P.O. Box 47, 6700 AA, Wageningen, the Netherlands*
[c] *KWR Watercycle Research Institute, P.O. Box 1072, 3430 BB, Nieuwegein, the Netherlands*
[d] *Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA, Delft, the Netherlands*

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 $0.5 \times 0.5°$ 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). Oocyst 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 (incl. The UK, Belgium and Macedonia), 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 log10 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.

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 river water 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.

* Corresponding author.
  E-mail addresses: lucievermeulen@gmail.com, lucie.vermeulen@rivm.nl (L.C. Vermeulen).

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 other 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.

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 that calculates mean monthly oocyst concentrations in rivers.

2. Materials and methods

GloWPa-Crypto C1 couples the output from the loading models GloWPa-Crypto H1 and L1, that calculate human and animal *Cryptosporidium* loads (Hofstra and Vermeulen, 2016; Vermeulen et al., 2017), to output from the Variable Infiltration Capacity (VIC) model, a grid-based macroscale hydrological model (Liang et al., 1994), version 4.1.2. Section 2.1 discusses *Cryptosporidium* sources, section 2.2 the transport of oocysts from land to rivers, section 2.3 oocyst survival during transport (e.g. decay and sedimentation processes), section 2.4 oocyst transport with rivers (the routing model), and section 2.5 describes how validation and sensitivity analysis are applied to assess model performance. Fig. 1 gives a schematic model representation. All calculations are performed on a $0.5 \times 0.5°$ grid and monthly time step using the software environment 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. Table S1 in the Supplementary materials list the variables that are included in the model.

2.1. Sources of Cryptosporidium

2.1.1. Human sources

Human *Cryptosporidium* loads to rivers are calculated in the GloWPa-Crypto H1 model (Hofstra and Vermeulen, 2016). In brief, GloWPa-Crypto H1 calculates oocyst loads by multiplying human population data with estimates of *Cryptosporidium* infection, associated oocyst excretion and removal depending on the sanitation system used. Prevalence of *Cryptosporidium* infection and oocyst excretion in humans are 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 $0.5 \times 0.5°$ 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.

2.1.2. Animal sources

Animal *Cryptosporidium* loads to land are calculated in the GloWPa-Crypto L1 model (Vermeulen et al., 2017). 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 (Vermeulen et al., 2017). 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 $0.5° \times 0.5°$ 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.

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). In Hofstra and Vermeulen (2016) 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. Using a constant runoff fraction does not account for the variability in runoff and associated transport. Therefore, in GloWPa-Crypto C1 we no longer use this constant runoff fraction, but instead make oocyst transport from land to rivers dependent on surface runoff, both for
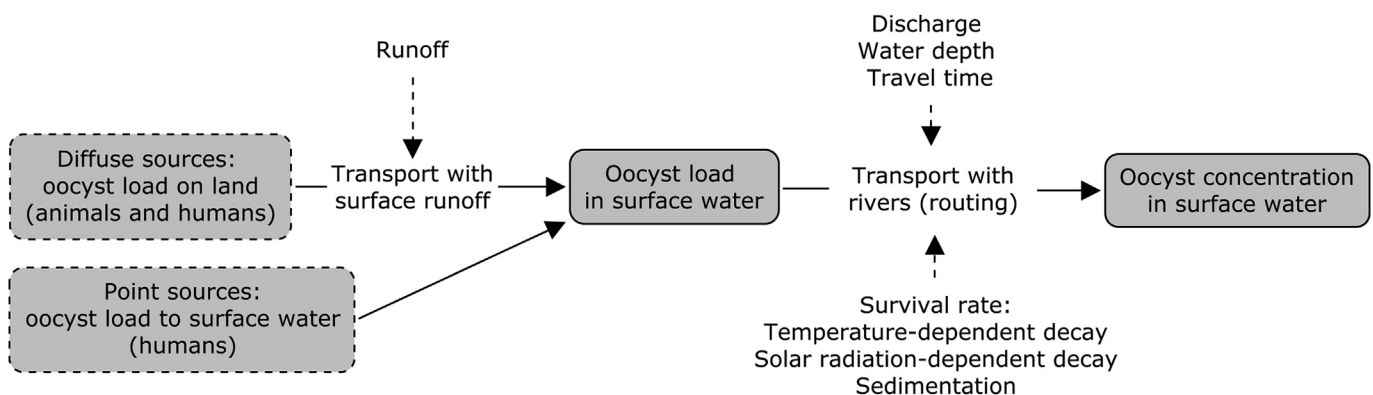

Fig. 1. Flow chart overview of GloWPa-Crypto C1 model approach. This figure shows the 'flow' of oocysts from diffuse and point sources on the left to the concentrations in rivers on the right. In between, processes that are calculated are the transport with surface runoff, the transport with rivers (routing) and oocyst survival. 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.

human diffuse sources and animal sources.

The VIC model (Liang et al., 1994) provides data of surface runoff (mm day^{-1}) and discharge (m^3 s^{-1}). Surface runoff and discharge are monthly mean averages of a baseline VIC run with WATCH forcing data (Weedon et al., 2011), averaged over the period 1970—2000 ($0.5° \times 0.5°$ 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 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 1).

From Table 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. (2006) measured oocysts in runoff over a period of one year. During this year, surface runoff varied from 0 to 20 mm. We therefore choose to linearly assign the observed log oocyst retention from vegetated plots under field conditions (3.2—8.8 log10) to grid cells with runoff values ranging between 0 and 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 and 3.2 log10. We assume that under field conditions, at least 1 log10 oocysts 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 animal 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 S2 for a more detailed discussion). Although crude, the approach we have chosen accounts for spatial and monthly variability in surface runoff and the magnitude of retention is based on observations

under field conditions (Atwill et al., 2006). More experimental data on oocyst release from manure and transport to rivers under various climatic and landscape conditions is needed to refine this approach.

2.3. Oocyst survival during transport with rivers

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

$$C_t = C_0 e^{-K \times t} \tag{1}$$

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:

$$\ln(Fs_i) = -K_i \times t_i \tag{2}$$

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

$$K_i = K_{T,i} + K_{R,i} + K_{S,i} \tag{3}$$

$K_{T,i}$ 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 2.3.1-2.3.3). Water residence time in a grid cell (t_i) is estimated as described in section 2.3.4. The Supplementary Material S4 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 and infectivity of oocysts. Loss of viability and infectivity may lead to disintegration, thereby reducing the concentration.

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):

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

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 mean water temperature as estimated by the VIC-RBM model framework which

Table 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 (log10)	Studies
Field conditions	Oocysts in faecal matrix on vegetated soils	3.2—8.8	Atwill et al. (2006)
Field conditions	Oocysts in faecal matrix on bare soils	1.7—5	Atwill et al. (2006)
Simulated rainfall	Oocysts in faecal matrix on vegetated soils	0.7 - 6.8	(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	(Davidson et al., 2014; Davies et al., 2004; Tate et al., 2004)
Simulated rainfall	Oocysts in faecal matrix in the lab	0.1—3	(Boyer et al., 2009; Bradford and Schijven, 2002)
Simulated rainfall	Oocysts in suspension on vegetated soils	0.6—3.1	(Atwill et al., 2002; Bhattarai et al., 2011; Trask et al., 2004)
Simulated rainfall	Oocysts in suspension on bare soils	0.0—1.6	(Atwill et al., 2002; Bhattarai et al., 2011; Trask et al., 2004)

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). Monthly mean water temperature is calculated as a 24-h mean extended to 30 days. 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 $\lambda = 0.158$ (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 and 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 S3).

2.3.2. Solar radiation-dependent survival (K_R)

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

$$K_{R,i} = \frac{k_l\, I_{A,i}}{k_e\, Z_i} \times \left(1 - e^{-k_e\, Z_i}\right) \tag{5}$$

Where $I_{A,i}$ is the average surface solar radiation (kJ m^{-2} day^{-1}) in month i, k_l 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 (DOC) (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:

$$k_e = k_d \times C_{DOC} \tag{6}$$

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 eqs. (5) and (6) in eq. (2) for solar radiation-dependent decay (temperature and sedimentation were controlled for in the experiment) gives:

$$\ln(Fs) = -\frac{k_l\, I_A}{k_d\, C_{DOC}\, Z} \times \left(1 - e^{-k_d\, C_{DOC} Z}\right) \times t \tag{7}$$

King et al. controlled for the effect of temperature-dependent decay and sedimentation in their experiments. We performed an optimization routine on eq. (7) 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. (Fig. 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.

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):

$$K_{S,i} = \frac{v}{Z_i} \tag{8}$$

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 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. It should be noted that filtration (attachment or straining) in permeable sediments could also remove oocysts from water (Harter et al., 2000). However, filtration was not included in the model, as its contribution to oocyst removal is probably lower than decay and sedimentation, and it difficult to estimate as data on the characteristics of streambed sediments are not available at the global scale.

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:

$$Z_i = 0.34\, Q_i^{0.341} \tag{9}$$

$$W_i = 1.22 Q_i^{0.557} \tag{10}$$

$$U_i = \frac{Q_i}{W_i \times Z_i} \tag{11}$$

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}), all 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 point-Distance 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).
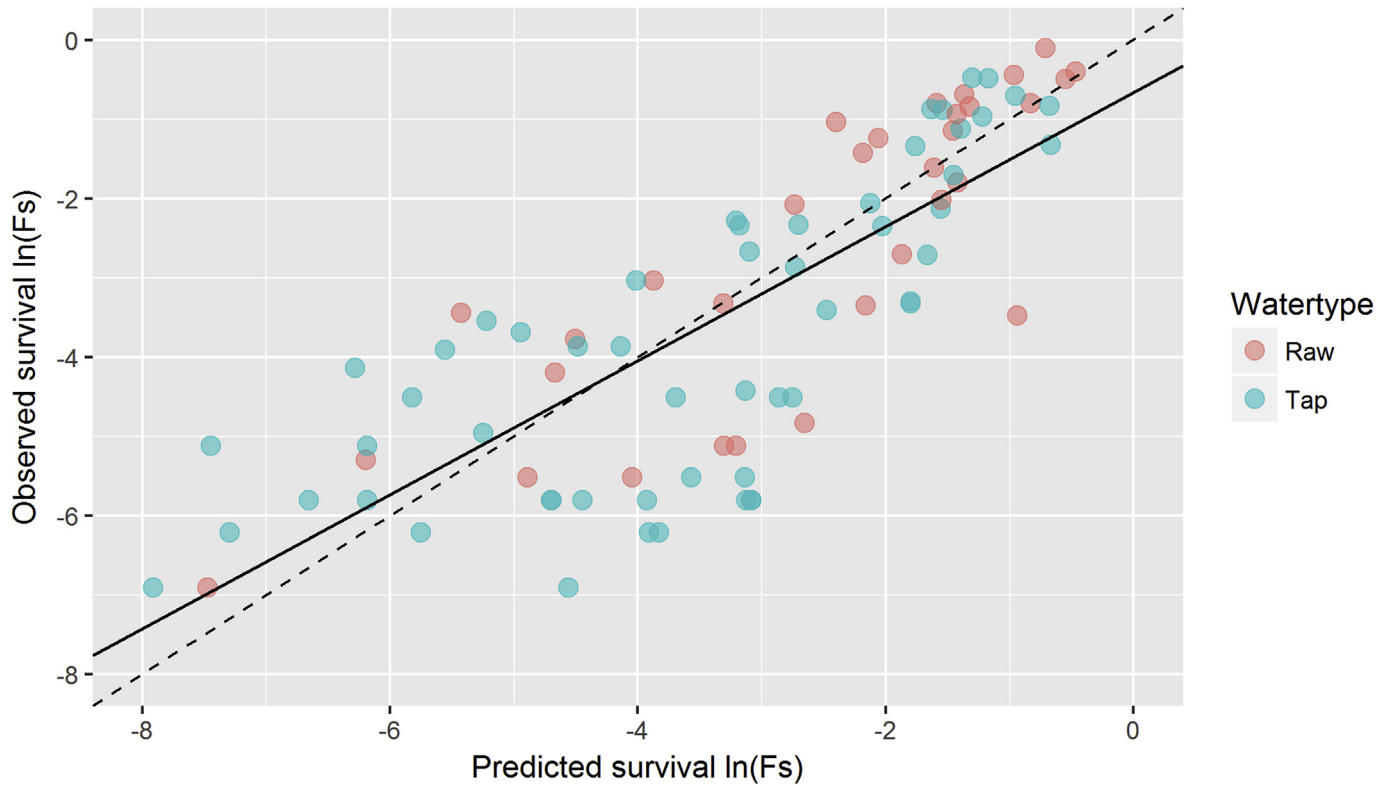
Fig. 2. Observed versus predicted survival (given as ln(*Fs*)) for the data from King et al. (2008) for raw surface waters (pink dots) and tap water (blue dots). The dashed line indicates the 1-1 line, in case the model perfectly reproduces the observations all dots should be on this line. The solid line is a linear regression line of observed versus predicted values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

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 2.3).

$$L_{i,\ n} = (SD_i + SP_i + L_{i,n-1}) \times e^{-K_i \times t_i} \tag{12}$$

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.

$$C_i = \frac{L_i}{Q_i} \tag{13}$$

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*.

2.5. Model performance

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 346 locations in 11 countries on 5 continents (Burnet et al., 2014, 2015; Chuah et al., 2016; Claβen et al., 2004; Ehsan et al., 2015; Kistemann et al., 2012; Lalancette et al., 2014; Medema et al., 2001; Rechenburg et al., 2009, 2006; Till et al., 2008). The Supplementary Material S6 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 0.5 × 0.5° 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., 2017). 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). For Japan, the dataset is incomplete, as observed

nondetects were not reported (personal communication). Therefore these Japanese data are only shown in the figures, but excluded for all calculations. In our complete data set excluding Japan, 73% of the observations is below the detection limit, ranging from 20 to 95% for the different countries. The reported detection limits range from 0.0028 to 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., 2017). The actual observed oocyst concentration can then be adjusted accordingly. In our data set, 28% 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 and 0.89.

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.

3. Results

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 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 (incl. The UK, Belgium and Macedonia), and the Middle East. These hotspot regions align with the hotspot regions observed in the underlying human oocyst loads model GloWPa-Crypto H1 (Hofstra and Vermeulen, 2016), more than with the hotspot regions observed in the animal oocyst loads model GloWPa-Crypto L1 (Vermeulen et al., 2017). 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.

Table 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

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.

Figs. 3 and 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 data presented in Figs. 3 and 4 are exactly the same, the discharge mask is merely a visual aid to better see large rivers only. 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 no data on oocyst loads. It should be noted that oocyst loading from wildlife is ignored in the model. Globally, livestock outnumber wildlife more than 20 to 1 (Smil, 2011), but for some natural areas, this might cause an underestimation of oocyst loads.

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)). Fig. 5 shows what our results look like when these categories are assigned to the modelled 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.

In the Supplementary material S5 we present regional concentrations maps for seven world regions: Africa, Asia, Europe, Latin America, Middle East − North Africa (MENA), North America and Oceania (Figs. S2−8). We distinguish these regions as our input data for animal oocysts loads make the same distinction. Furthermore, Figs. S11−13 show oocyst concentrations by month for selected locations for which validation data were available. Figs. 3 and 4, and Figures S2-8 and S11-13, show some effect of seasonality, but average monthly oocyst concentrations are predicted not to vary more than approximately one log10 unit throughout the year in most places. Regions with higher variability (2 log10) include parts of India and Southeast Asia, west Africa, Brazil and the west coast of North America.

3.2. Source attribution

GloWPa-Crypto indicates that point sources are the dominant source of oocysts in most world regions (Fig. 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 Fig. 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.

3.3. Model performance

3.3.1. Validation

We compared model outcomes to a set of observational data, Table S3 and Fig. S9 give an overview of the data. First we should note that in this set of validation data only 28% 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. (2017) note that most published monitoring data do not
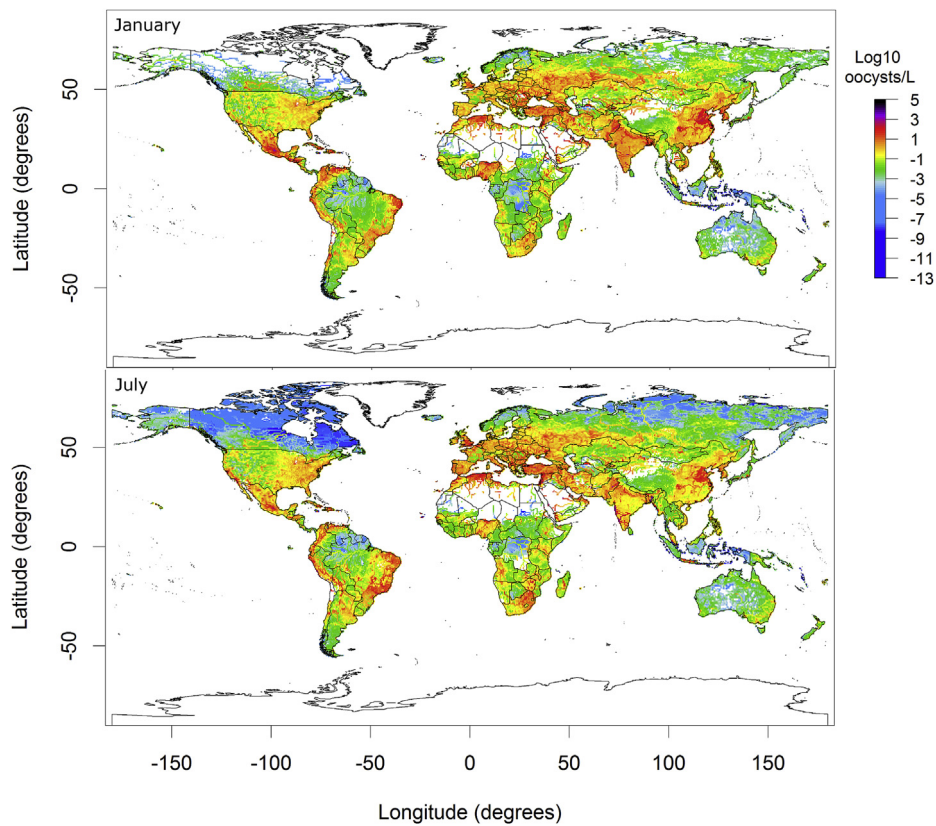
Fig. 3. Oocyst concentrations in rivers worldwide for the months January and July (log10 oocysts/L).
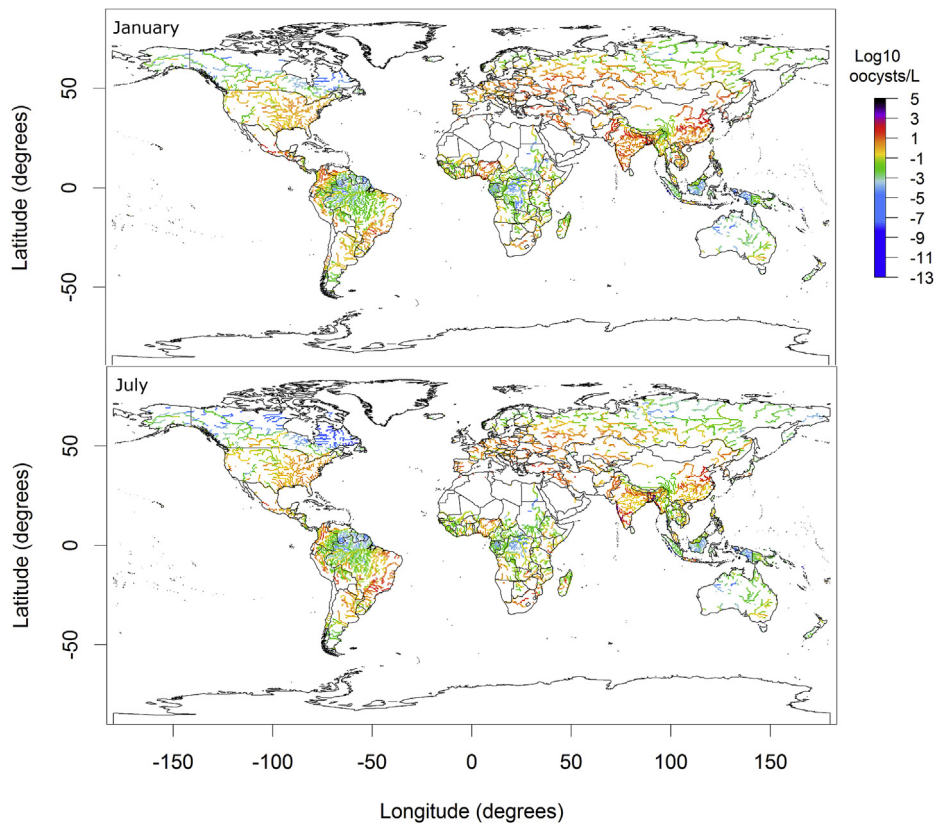

Fig. 4. Oocyst concentrations in rivers worldwide for the months January and July (log10 oocysts/L), with a discharge mask. These plots show only large rivers (i.e. grid cells where average annual discharge is higher than 200 m³/s).
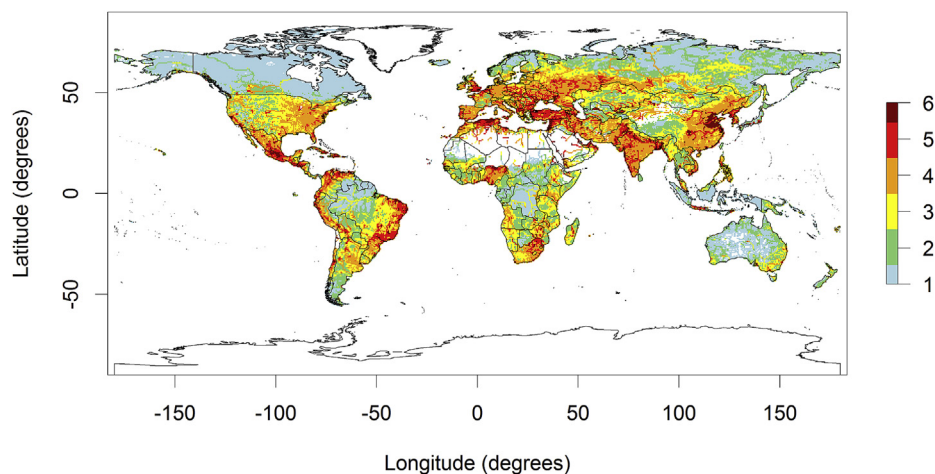
Fig. 5. Annual mean oocyst concentration categorised to WHO pollution categories (1 = very pristine (~0.001 oocyst/L) to 6 = grossly polluted (~100 oocyst/L)) (Medema et al., 2009). Each category represents one log10 unit change in concentrations.

Fig. 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 log10 higher than oocyst loads from point sources in that area, etcetera. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

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., 2017). For this reason, we can expect the modelled concentrations to be higher than the observed concentrations, which is indeed the case (Fig. 7, and Supplementary material S6).

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 log10 units higher than the observed concentrations (Fig. 7, and Supplementary material S6). 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 (Fig. 7), strengthening our notion that correcting for recovery efficiency is important.

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 3.3.2 and Vermeulen et al. (2015a)). 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 3.3.2). 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 between observed and predicted oocyst concentrations.

To assess model performance numerically, several goodness of fit statistics were calculated using the log transformed results. Japan was excluded in these calculations, as this data set did not
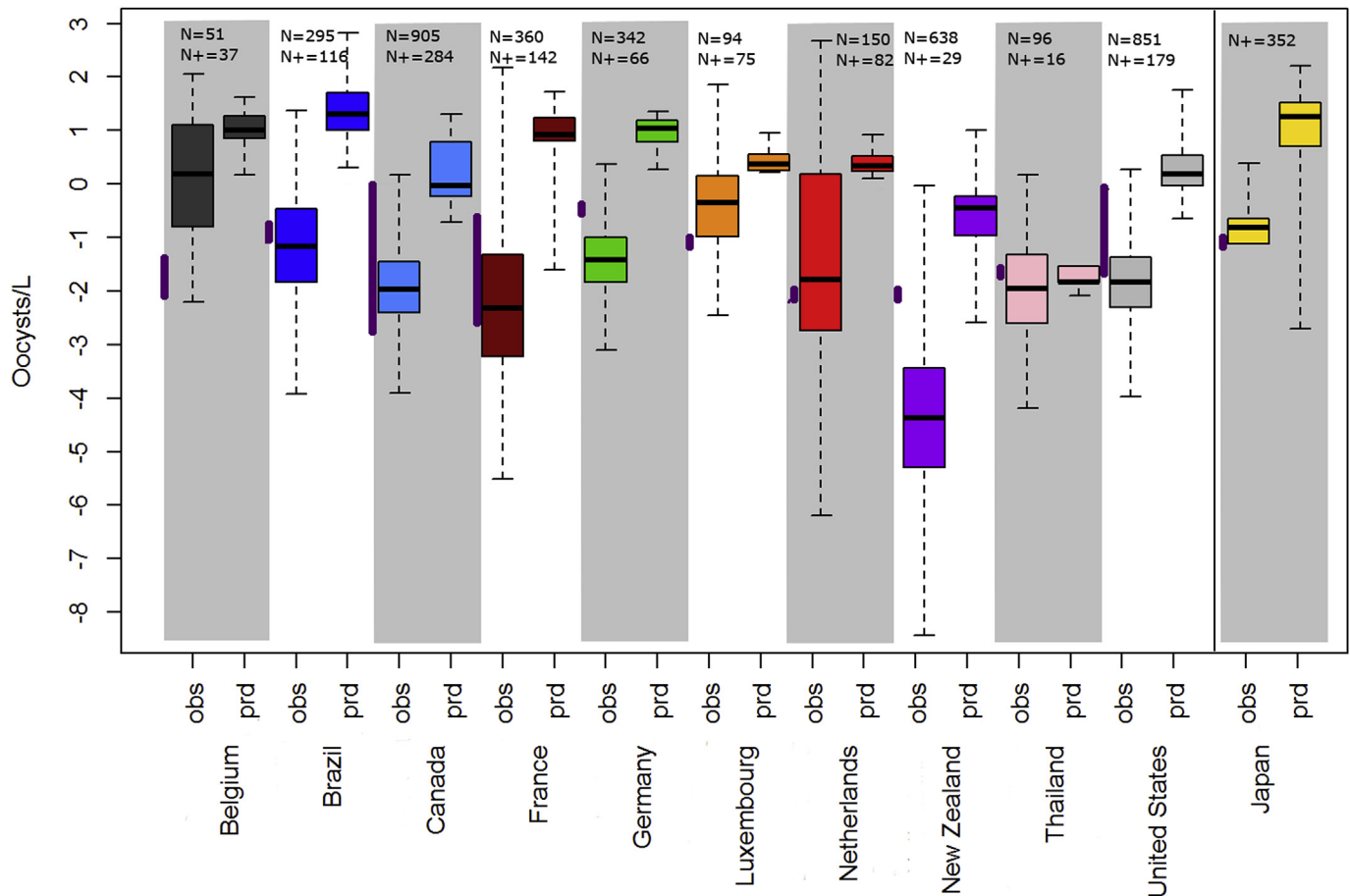
Fig. 7. Boxplots of observed (obs) and predicted (prd) oocyst concentrations grouped by country. For the observational data, the function 'cenboxplot' from the R package NADA was used, which estimates the distribution of values below the detection limit, assuming the concentration is lognormally distributed (Lee, 2017; Lee and Helsel, 2005). Dark purple 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. Japan is shown separately, as this data set did not report the nondetects. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

report the nondetects. We calculated rank correlation coefficients, Spearman's rho: 0.39 and Kendall's tau: 0.28. 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 0.40, 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 1.82, 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 S6), consistent with our finding of limited seasonality in modelled concentrations (section 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. Fig. 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.

3.3.2. Sensitivity analysis

The model is most sensitive to changes in input human oocyst loads (Table S4 in the Supplementary Material) and to a lesser extent to input animal loads, as human loads were found to dominate in most places. 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). The high and low end runs cause a 1.5 log10 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 (Fig. 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),
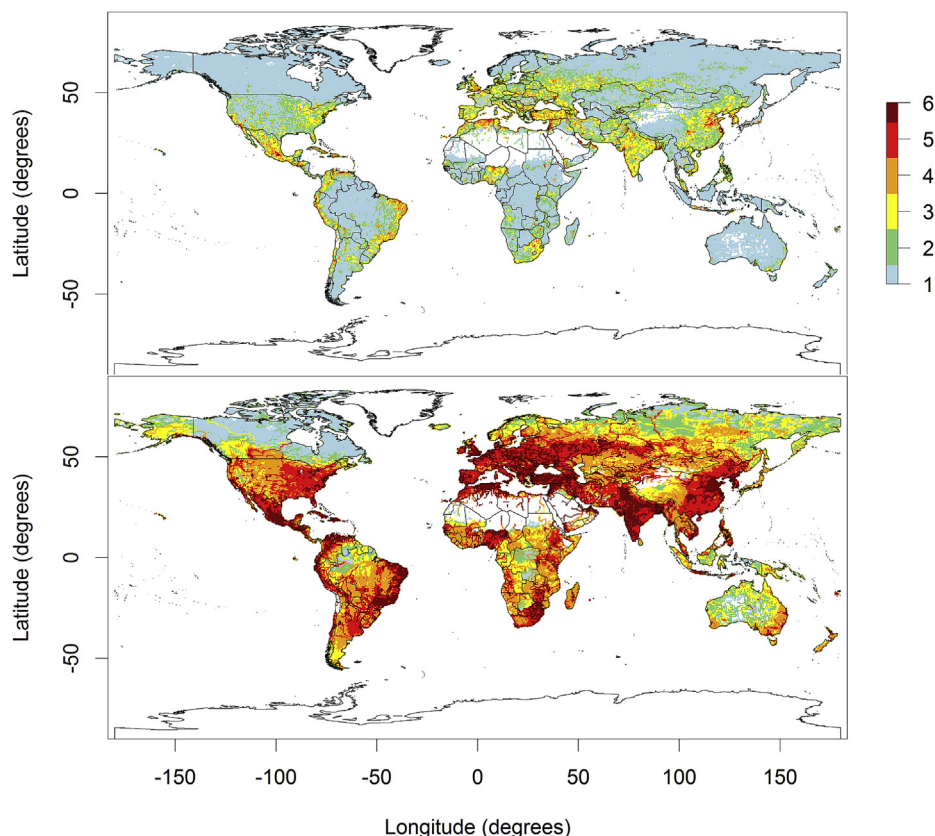
Fig. 8. Low end (top) and high end (bottom) runs of annual mean oocyst concentration categorised to WHO pollution categories (1 = very pristine (~0.001 oocyst/L) to 6 = grossly polluted (~100 oocyst/L)). Each category represents one log10 unit change in concentrations.

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 Fig. 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.

4. Discussion

Models of faecal microorganisms in rivers exist mostly at the catchment scale (Vermeulen et al., 2015b). 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 (Vermeulen et al., 2015b). Therefore, the development of GloWPa-Crypto provides important new opportunities, such as estimating hotspot regions with high concentrations (section 3.1) and assessing source attribution (section 3.2). Another interesting application would be to use GloWPa-Crypto for scenario analysis. Global change processes, such as population growth, socioeconomic development and climate changes, may influence pathogen concentrations in surface water in future (e.g. Fearnley et al., 2010; Hofstra, 2011; Schijven and de Roda Husman, 2005). The high concentrations in urban areas in developing countries are a concern, and the situation in these regions is expected to get worse in the future due to high population growth and installing sewer connections without sewage treatment keeping up (Hofstra and Vermeulen, 2016). Climatic changes potentially affect the runoff from land, transport with rivers and temperature-dependent oocyst decay. However, the current model predicts limited effects of seasonal variation in these climatic variables on monthly average concentrations (section 3.1). Yet 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. More detailed input data, for example on manure application worldwide, would be needed to operate GloWPa-Crypto on a finer temporal resolution.

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. Such big picture approaches are essential to address large scale transboundary issues like water pollution, for instance in the context of the Sustainable Development Goals that the world aims to reach in 2030 (United Nations General Assembly, 2015). Goal 6 is to 'ensure availability and sustainable management of water and sanitation for all'; large scale water quality models such as GloWPa-Crypto could be a tool help assess progress and identify policy options to reach this goal.

It should be noted that GloWPa-Crypto L1 includes all *Cryptosporidium* spp., also those that are not infectious for humans (Vermeulen et al., 2017). 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).

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 (Vermeulen et al., 2017), was out of the scope of this study. However, in light of the model validation and sensitivity analysis this should be given priority in further research. Oocyst excretion rates are variable and uncertain, even between infected individuals oocyst production can differ strongly (Chappell et al., 1996; Vermeulen et al., 2017). Other models of pathogens or faecal indicators have similarly observed sensitivity to pathogen loading from humans and animals (Coffey et al., 2010; Ferguson et al., 2007; Reder et al., 2017; Tian et al., 2002).We recommend that monitoring programs for *Cryptosporidium* in rivers should always determine and report the recovery efficiency, preferably for each individual sample, in order for data to be useful for quantitative analysis and model validation (Efstratiou et al., 2017; Ongerth, 2016).

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. It should be noted that more data should become available in order to be able to implement these improvements effectively, particularly number 2 and 4. Opportunities for GloWPa-Crypto C1 include that the model 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), a first exploration of this application has already been done (Vermeulen, 2018). Especially for application in risk assessment, 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.

5. Conclusion

GloWPa-Crypto C1 is the first global model that computes pathogen concentrations in rivers. Modelling the transport of *Cryptosporidium* through the environment is a helpful tool to get insight in the predicted order of magnitude of oocyst concentrations worldwide, to pinpoint hotspot regions of high concentrations that can be interesting places to focus further study on, and to gain insight in the relative importance of different pollution sources. Main findings from this study are:

- 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 (incl. The UK, Belgium and Macedonia), and the Middle East.
- Point sources (human faeces) appears to be a more dominant source of pollution than diffuse sources (animal manure) in most world regions.

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2018.10.069.

References

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. https://doi.org/10.1111/j.1752-1688.1994.tb03321.x. https://doi.org/DOI.

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., 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. https://doi.org/10.1128/AEM.68.11.5517.

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

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. Software 26, 1289–1297. https://doi.org/10.1016/j.envsoft.

2011.05.011.

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. https://doi.org/10.1007/s00254-008-1580-x.

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. https://doi.org/10.1021/es025573l.

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

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. https://doi.org/10.2166/wh.2005.065.

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. https://doi.org/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. Publ. Health 12, 11910—11928. https://doi.org/10.3390/ijerph120911910.

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.

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. https://doi.org/10.1016/j.scitotenv.2016.03.247.

Claßen, T., Koch, C., Rechenburg, A., Christoffels, E., Kistemann, T., 2004. Hygienisch-mikrobielle Fließgewässerbelastung durch Regenwasserentlastungsanlagen am Beispiel der Swist. Umweltmed. Forsch. 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.

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. https://doi.org/10.1021/es071324r.

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. https://doi.org/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. https://doi.org/10.1128/AEM.70.2.1151.

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. https://doi.org/10.1016/j.cofs.2016.04.009.

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

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

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. https://doi.org/10.1007/s10661-014-4157-z.

Fearnley, E., Weinstein, P., Dodson, J., 2010. Climate change, societal transitions and changing infectious disease burdens. In: Changing Climates, Earth Systems and Society. Springer, Netherlands, pp. 189—199. https://doi.org/10.1007/978-90-481-8716-4_9.

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. https://doi.org/10.2166/wh.2007.013.

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. https://doi.org/10.1016/S0140-6736(14)61682-2.

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

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. https://doi.org/10.1029/2005GB002480.

Harter, T., Wagner, S., Atwill, E.R., 2000. Colloid transport and filtration of Cryptosporidium parvum in sandy soils and aquifer sediments. Environ. Sci. Technol. 34, 62—70. https://doi.org/10.1021/es990132w.

Helsel, D., 2010a. Much ado about next to nothing: incorporating nondetects in

science. Ann. Occup. Hyg. 54, 257—262. https://doi.org/10.1093/annhyg/mep092.

Helsel, D., 2010b. Summing nondetects: incorporating low-level contaminants in risk assessment. Integr. Environ. Assess. Manag. 6, 361—366. https://doi.org/10.1002/IEAM.31.

Hijmans, R.J., 2016. Raster: Geographic Data Analysis and Modeling.

Hofstra, N., 2011. Quantifying the impact of climate change on enteric waterborne pathogen concentrations in surface water. Curr. Opin. Environ. Sustain. 3, 471—479. https://doi.org/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. https://doi.org/https://doi.org/10.1016/j.scitotenv.2012.10.013.

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. https://doi.org/10.1016/j.ijheh.2016.06.005.

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. https://doi.org/10.1111/j.1365-2672.2007.03658.x.

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. Environ. 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. https://doi.org/10.1016/j.ijheh.2011.12.008.

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—161. https://doi.org/10.1016/j.watres.2014.01.050.

Lee, L., 2017. NADA: Nondetects and Data Analysis for Environmental Data. R Package Version 1, 6-1.

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. https://doi.org/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.

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. https://doi.org/10.1016/S0140-6736(16)31529-X.

Mancini, J.L., 1978. Numerical estimates of coliform mortality rates under various conditions. J. Water Pollut. Control Fed. 50, 2477—2484. https://doi.org/10.2307/25040179.

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. https://doi.org/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. Software 25, 837—853. https://doi.org/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. Manag. 131, 121—128. https://doi.org/10.1016/j.jenvman.2013.09.017.

Medema, G.J., Ketelaars, H.A.M., Hoogenboezem, W., 2001. Cryptosporidium en Giardia: voorkomen in rioolwater, mest en oppervlaktewater. Vereniging van Rivierwaterbedrijven - RIWA.

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., Teunis, P., Blokker, M., Deere, D., Davison, A., Charles, P., Loret, J.F., 2009. Risk Assessment of Cryptosporidium in Drinking Water. WHO, Geneva, Switzerland.

Nasser, A.M., 2016. Removal of Cryptosporidium by wastewater treatment processes: a review. J. Water Health 14, 1—13. https://doi.org/10.2166/wh.2015.131.

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. https://doi.org/10.1007/s10661-016-5528-4.

Ongerth, J.E., 2016. Cryptosporidium and giardia in Water : reassessment of occurrence and significance. J. Environ. Eng. 143, 1—8. https://doi.org/10.1061/(ASCE)EE.1943-7870.0001161.

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. https://doi.org/10.1016/j.jmarsys.2012.05.004.

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. https://doi.org/10.1128/AEM.01442-08.

R Development Core Team, 2016. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing. R Found. Stat. Comput.

https://doi.org/10.1007/978-3-540-74686-7.

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. https://doi.org/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. Model. 351, 129—139. https://doi.org/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. Software 63, 251—263. https://doi.org/10.1016/j.envsoft.2014.10.001.

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. https://doi.org/10.1371/journal.pone.0096084.

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.

Scully, N., Lean, D., 1994. The attenuation of ultrviolet radiation in temperate lakes. Arch. Hydrobiol. Beih. Ergebn. Limnol. 43, 135—144.

Shirley, D.-A.T., Moonah, S.N., Kotloff, K.L., 2012. Burden of disease from cryptosporidiosis. Curr. Opin. Infect. Dis. 25, 555—563. https://doi.org/10.1097/QCO.0b013e328357e569.

Smil, V., 2011. Harvesting the biosphere: the human impact. Popul. Dev. Rev. 37, 613—636. https://doi.org/10.1111/j.1728-4457.2011.00450.x.

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. https://doi.org/10.2166/wh.2012.183.

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. https://doi.org/10.2134/jeq2004.2243.

Thomann, R.V., Mueller, J.A., 1987. Principles of Surface Water Quality Modeling and Control. Harper & Row Publishers, New York.

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. https://doi.org/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.

UNEP, 2016. A Snapshot of the World's Water Quality: towards a Global Assessment.

United Nations General Assembly, 2015. Transforming Our World: the 2030 Agenda for Sustainable Development, A/RES/70/1. https://doi.org/10.1007/s13398-014-0173-7.2.

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. https://doi.org/10.5194/hess-16-4303-2012.

Vermeulen, L.C., 2018. Cryptosporidium in Rivers of the World: the GloWPa-crypto Model. Wageningen University. https://doi.org/10.18174/426782.

Vermeulen, L.C., Benders, J., Medema, G.J., Hofstra, N., 2017. Global Cryptosporidium loads from livestock manure. Environ. Sci. Technol. 51, 8663—8671. https://doi.org/10.1021/acs.est.7b00452.

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

Vermeulen, L.C., Hofstra, N., Kroeze, C., Medema, G.J., 2015b. Advancing waterborne pathogen modelling: lessons from global nutrient export models. Curr. Opin. Environ. Sustain. 14, 109—120. https://doi.org/10.1016/j.cosust.2015.05.003.

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. https://doi.org/10.1175/2011JHM1369.1.

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. https://doi.org/10.1029/JC090iC05p08995.

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

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. https://doi.org/10.1093/trstmh/tru056.