Estimation of pathogen concentrations in a drinking water source using hydrodynamic modelling and microbial source tracking
Ekaterina Sokolova, Johan Åström, Thomas J. R. Pettersson, Olof Bergstedt and Malte Hermansson

ABSTRACT
The faecal contamination of drinking water sources can lead to waterborne disease outbreaks. To estimate a potential risk for waterborne infections caused by faecal contamination of drinking water sources, knowledge of the pathogen concentrations in raw water is required. We suggest a novel approach to estimate pathogen concentrations in a drinking water source by using microbial source tracking data and fate and transport modelling. First, the pathogen (norovirus, Cryptosporidium, Escherichia coli O157/H7) concentrations in faecal contamination sources around the drinking water source Lake Rådasjön in Sweden were estimated for endemic and epidemic conditions using measured concentrations of faecal indicators (E. coli and Bacteroidales genetic markers). Afterwards, the fate and transport of pathogens within the lake were simulated using a three-dimensional coupled hydrodynamic and microbiological model. This approach provided information on the contribution from different contamination sources to the pathogen concentrations at the water intake of a drinking water treatment plant. This approach addresses the limitations of monitoring and provides data for quantitative microbial risk assessment (QMRA) and risk management in the context of faecal contamination of surface drinking water sources.

Key words | Bacteroidales markers, Cryptosporidium, E. coli O157/H7, faecal contamination, norovirus, QMRA

INTRODUCTION
The faecal contamination of drinking water sources can lead to waterborne disease outbreaks (MacKenzie et al. 1994; Hrudey & Hrudey 2004). During two recent outbreaks (2010–2011) caused by faecal contamination of drinking water sources, tens of thousands of people in Sweden in the cities Östersund and Skellefteå suffered from waterborne Cryptosporidium infections (Skellefteå Municipality 2011; SMI 2011a). To estimate the potential risk for waterborne disease infections caused by faecal contamination of drinking water sources, a quantitative microbial risk assessment (QMRA) is often used (e.g. Westrell et al. 2003; De Roda Husman et al. 2009). QMRA involves hazard identification, exposure assessment, dose-response assessment and risk characterisation (Haas et al. 1999). Generic QMRA tools have been developed (Pettersson & Stenström 2007; Schijven et al. 2011) to perform QMRA for specific water treatment plants (WTPs) with different treatment trains and varying pathogen levels in the raw water.

To perform QMRA, data on pathogen concentrations at the water intake to the WTP are needed. However, pathogen concentrations in water sources are rarely measured due to the complexity and cost of the analytical tests (e.g. Brookes et al. 2005; Field & Samadpour 2007). Instead of pathogens, faecal indicators are used by water producers to monitor microbial water quality. However, concentrations of faecal indicators are not directly related to the concentrations of
pathogens in the environment (Ashbolt et al. 2001; Brookes et al. 2005; Harwood et al. 2005). Pathogen concentrations at the water intake to the WTP depend on the pathogen load that enters the water body from different contamination sources (e.g. wastewater treatment plants, on-site sewer systems, emergency sewer overflows as well as surface runoff from urban and agricultural areas) and on the pathogen fate and transport within the water body.

The pathogen load to the water body from different contamination sources varies strongly with time, often due to the prevalence of the disease in the population. Under epidemic conditions, pathogens are excreted from many more human or animal hosts than under endemic conditions. An increased pathogen load, which enters the water source with wastewater discharges or surface runoff, implies higher pressure on water treatment and, as a result, increased risk of water-borne infections. The pathogen load to the water body under both endemic and epidemic conditions can be estimated based on the assumption that the ratio of pathogens to faecal indicators in fresh faecal matter is the same as in the contaminated discharges to the water source.

The estimation of pathogen load to the water body can be performed using traditional faecal indicators (such as *Escherichia coli*) or indicators that are applied in microbial source tracking to determine the origin of faecal contamination, such as *Bacteroidales* genetic markers (Field & Samadpour 2007). Several *Bacteroidales* assays have been proposed for the identification of host-specific faecal matter from humans and cattle (Bernhard & Field 2000; Layton et al. 2006; Reischer et al. 2006, 2007; Kildare et al. 2007; Stricker et al. 2008; Converse et al. 2009). While *E. coli* bacteria represent the total faecal contamination from all warm-blooded animals (humans, livestock, domestic pets, wild animals) and birds, *Bacteroidales* genetic markers indicate the contribution from the pathogen host groups of interest, i.e. humans and cattle. It is therefore posited that the estimation of pathogen concentrations using *Bacteroidales* genetic markers is more accurate than the estimation using *E. coli*.

The fate and transport of microbial contamination within the water source can be simulated using a coupled hydrodynamic and microbiological modelling approach. This approach has been widely applied to simulate the fate and transport of traditional faecal indicators (e.g. Sanders et al. 2005; Kashefpour et al. 2006; Liu et al. 2006; Riou et al. 2007; Thupaki et al. 2010). However, relatively few studies have focused on modelling the fate and transport of the actual pathogens within the water source (e.g. Hipsey et al. 2004; Brookes et al. 2006; Hipsey et al. 2008), mainly due to the lack of data on pathogen concentrations.

In this study we have estimated concentrations of *Cryptosporidium*, norovirus and *E. coli* O157/H7 in contamination sources around a drinking water source, Lake Rådasjön, Sweden and simulated fate and transport of these pathogens within the lake. The concentrations in the contamination sources under endemic and epidemic conditions were estimated using the ratio of pathogens to faecal indicators (*E. coli*, human and ruminant *Bacteroidales* genetic markers) in fresh faecal matter. The spread of the pathogens within the lake from the contamination sources to the point of the WTP water intake was simulated using a three-dimensional hydrodynamic and microbiological model that takes into account pathogen inactivation in the environment. We suggest a novel approach to provide input data on pathogen concentrations in a drinking water source for QMRA by combining microbial source tracking data with hydrodynamic and microbiological modelling. Moreover, this approach provides information on the contribution from different contamination sources to the pathogen concentrations at the WTP water intake and enables prioritisation of mitigation measures.

**METHODS**

**Study area**

Lake Rådasjön (Figure 1) is located on the west coast of Sweden and constitutes the main water source for the city of Mölndal (60,000 consumers) and is a reserve water supply for the city of Gothenburg (500,000 consumers). The surface area of the lake is approximately 2.0 km², the catchment area of the lake is 14.9 km² and the maximum water depth is 23 m. The main inflow to the lake is the river Mölndalsån with a water flow in the range 1–20 m³ s⁻¹. The river Mölndalsån enters the lake in the southeast and drains the lake in the west to Lake Stensjön. The water intake for the city of Mölndal is located in the north-western part of the lake at 15 m depth (Figure 1).
Lake Rådasjön is subjected to faecal contamination from various sources located in the catchment of the lake. In terms of the risks of waterborne disease outbreaks, the sources of human and ruminant faecal contamination are assumed to be the most relevant. Human faecal contamination at the water intake may originate from on-site sewers located to the north of the lake. On-site sewers release partly treated effluents into streams that enter the lake close to the water intake (Figure 1, sites 3 and 7). Another source of human faecal contamination is an emergency sewer overflow at a pumping station in a separate sewer system located south of the lake (Figure 1, site P). Discharges of untreated wastewater from this source occur several times a year during periods of heavy rainfall after intrusion of stormwater into the sewer network. In addition, human faecal contamination can potentially enter the lake with untreated stormwater runoff from an urban area located to the east of the lake (Figure 1, site 18). Faecal matter from the grazing area may reach the lake after short transport in a small stream (Figure 1, site 17). Concentrations of faecal indicators in the contamination sources were measured during a monitoring campaign performed in 2008 (Sokolova et al. 2012).

Norovirus can enter Lake Rådasjön from the sources of human faecal contamination (sites 3, 7, 18 and P), while *E. coli* O157/H7 can enter the lake from the source of ruminant faecal contamination (site 17; WHO 2008). Furthermore, *Cryptosporidium* can enter the lake from the sources of both human and ruminant faecal contamination (WHO 2008), i.e. all the sources mentioned above.

**Estimation of norovirus, Cryptosporidium and E. coli O157/H7 concentrations in the contamination sources**

To estimate the concentrations of norovirus, *Cryptosporidium* and *E. coli* O157/H7 in the discharges from contamination sources to Lake Rådasjön (Figure 1), the ratio of pathogens to faecal indicators was defined according
to Equation (1):
\[
\frac{C_P}{C_I}^{\text{source}} = p \frac{C_P}{C_I}^{\text{faecal matter}}
\]  

(1)

where \(C_P\) is the pathogen concentration; \(C_I\) is the faecal indicator concentration; and \(p\) is the prevalence of a disease among people and animals present in the catchment of the lake. Prevalence of a disease was defined according to Equation (2):

\[
p = \frac{I \times t}{365}
\]  

(2)

where \(I\) is the incidence of disease (the number of new cases within a specified time period divided by the size of the population initially at risk) and \(t\) is the duration of excretion (days).

The parameters in Equation (1) were described by probability distributions in order to take into account their variability. Faecal indicator concentrations in the contamination sources were described by lognormal distributions with location, 50 and 95 percentiles (hereafter denoted as %). The faecal indicator and pathogen concentrations in faecal matter, as well as the data for incidence and duration of excretion, were obtained from the literature and described by probability distributions (Table 1). The estimation of the pathogen concentrations in the contamination sources was performed for endemic and epidemic conditions. It was assumed that under epidemic conditions on average half of the population of humans or animals excrete pathogens, i.e. the prevalence of a disease under epidemic conditions was described by a Beta distribution with \(\text{min} = 0\); \(\text{5}\% = 0.3\); \(\text{50}\% = 0.5\); \(\text{max} = 1\).

The concentrations of pathogens in the contaminated discharges to the lake were calculated according to Equation (1) using Monte Carlo simulations (10 000 iterations) and presented as percentiles (5, 50 and 95%).

Table 1 Literature data used for estimation of norovirus, Cryptosporidium and E. coli O157/H7 concentrations in the contamination sources

<table>
<thead>
<tr>
<th>Pathogen/faecal indicator</th>
<th>Type of data (unit)</th>
<th>Distribution (values)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norovirus</td>
<td>Duration of excretion (log days)</td>
<td>Lognormal 1 (0.30; 0.60; 1.48)</td>
<td>Atmar et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Incidence (^a)(^(-))</td>
<td>Beta (0.00; 1.00; 0.05; 0.19)</td>
<td>Hedlund (pers. comm. 2011)</td>
</tr>
<tr>
<td></td>
<td>Concentration in faecal matter (log gene copy numbers/g)</td>
<td>Lognormal 1 (3.30; 7.08; 9.46)</td>
<td>Nordgren et al. (2009)</td>
</tr>
<tr>
<td>Cryptosporidium human</td>
<td>Duration of excretion (log days)</td>
<td>Normal (1.48; 0.17)</td>
<td>Stehr-Green et al. (1987)</td>
</tr>
<tr>
<td></td>
<td>Incidence (^a)(^(-))</td>
<td>Beta (0.00; 1.00; 0.00005)</td>
<td>SMI (2010b)</td>
</tr>
<tr>
<td></td>
<td>Concentration in faecal matter (log oocysts/g)</td>
<td>Normal (7.00; 1.00)</td>
<td>Girdwood &amp; Smith (1999)</td>
</tr>
<tr>
<td>Cryptosporidium ruminant</td>
<td>Prevalence (^a)(^(-))</td>
<td>Beta (0.00; 1.00; 0.10; 0.17)</td>
<td>Enemark et al. (2002) and samples(^c)</td>
</tr>
<tr>
<td></td>
<td>Concentration in faecal matter (log oocysts/g)</td>
<td>Lognormal 2 (0.00; 2.00; 4.66)</td>
<td>Silverlås et al. (2010)</td>
</tr>
<tr>
<td>E. coli O157/H7</td>
<td>Prevalence (^a)(^(-))</td>
<td>Beta (0.00; 1.00; 0.10; 0.11)</td>
<td>Kistemann et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>Concentration in faecal matter (log CFU/g)</td>
<td>Lognormal 2 (0.00; 6.46; 8.41)</td>
<td>Hutchison et al. (2004)</td>
</tr>
<tr>
<td>Human marker</td>
<td>Concentration in faecal matter (log ME(^d)/g)</td>
<td>Triangular (9.82; 10.50; 10.94)</td>
<td>Reischer et al. (2007)</td>
</tr>
<tr>
<td>Ruminant marker</td>
<td>Concentration in faecal matter (log ME(^d)/g)</td>
<td>Triangular (9.48; 9.57; 9.63)</td>
<td>Reischer et al. (2006)</td>
</tr>
<tr>
<td>E. coli human</td>
<td>Concentration in faecal matter (log CFU/g)</td>
<td>Triangular (5.30; 6.80; 8.88)</td>
<td>Reischer et al. (2007)</td>
</tr>
<tr>
<td>E. coli ruminant</td>
<td>Concentration in faecal matter (log CFU/g)</td>
<td>Triangular (6.60; 7.30; 7.89)</td>
<td>Reischer et al. (2006)</td>
</tr>
</tbody>
</table>

\(^a\)Endemic conditions.

\(^b\)Parameters for distributions: Lognormal 1 (5; 50; 95%), Lognormal 2 (location; 5; 95%), Beta (min; max; 5; 50%), Normal (mean; std. dev.), Triangular (min; likeliest; max).

\(^c\)Samples of faecal matter from a herd represented at site 17.

\(^d\)Marker equivalent.
sources (sites 3, 7, 18 and P) were estimated using human *Bacteroidales* genetic markers (BacH), while in the ruminant faecal contamination source (site 17) the pathogen concentrations were estimated using ruminant *Bacteroidales* genetic markers (BacR) as faecal indicators. For comparison, the pathogen concentrations in all faecal contamination sources were estimated using *E. coli* as a faecal indicator.

**Coupled hydrodynamic and microbiological modelling**

The coupled hydrodynamic and microbiological modelling was used to estimate the contributions from the above-mentioned contamination sources (Figure 1) to the norovirus, *Cryptosporidium* and *E. coli* O157/H7 concentrations at the water intake of the city of Mölndal under endemic and epidemic conditions. For this purpose the pathogen concentrations in the contamination sources estimated using BacH and BacR markers were used. For every pathogen the simulations were performed using 5, 50 and 95% values of estimated pathogen concentrations in the contamination sources around the lake for endemic and epidemic conditions as input data. In addition, the simulations were performed to estimate the contributions to the pathogen concentrations at the water intake from the emergency sewer overflow (site P), the untreated stormwater source (site 18) and the cattle grazing area (site 17) for a hypothetical scenario in which these sources were located in the vicinity of the water intake on the north-western shore of the lake.

**Hydrodynamic model of Lake Rådasjön**

To simulate the water flows in Lake Rådasjön, the three-dimensional time-dependent hydrodynamic model MIKE 3 FM (DHI 2009) was used. The hydrodynamic model is based on the numerical solution of three-dimensional incompressible Reynolds-averaged Navier–Stokes equations invoking the assumptions of Boussinesq and of hydrostatic pressure (DHI 2009). The model consists of continuity, momentum, temperature, salinity and density equations and is closed by a turbulent closure scheme (DHI 2009).

The modelling domain was approximated with prisms (triangles in horizontal plane) using a flexible mesh approach (Figure 2). The length of the sides of the triangles varies from approximately 40 to 80 m, and was adjusted to describe the shoreline and bathymetry. The mesh resolution is finer in the narrow parts of the lake and where large
velocity gradients are expected, as well as in the vicinity of the emission points, to provide a better description of pollution spread. Vertically, the lake was divided into 27 layers. The thickness of the two uppermost layers can vary depending on the water level in the lake (sigma-layers), while the other layers have a fixed thickness (z-layers). In an undisturbed state the thickness of the two uppermost layers is 0.5 m each; these layers are followed by eight layers each of 1 m thickness down to a depth of 9 m. The range of depth from 9 to 16 m, where the thermocline is usually located, is divided into 14 layers each of 0.5 m thickness; below there are three layers with thicknesses of 1, 2 and 5 m.

The model was set up to simulate the hydrodynamic situation in Lake Rådasjön in the year 2008. The simulation was performed for the period January–December 2008. The output of the hydrodynamic model was then utilised in a microbiological model in order to simulate the transport of microbial contamination in the lake and estimate the pathogen concentrations at the water intake.

The initial rest conditions in the lake were defined by the constant surface elevation and the flow velocity was set to zero. The initial temperature field was described as homogeneous in the domain (2.1 °C) based on the assumption that in January the lake is well mixed and horizontal and vertical gradients of temperature at this time of year can be neglected.

The open boundary conditions (Figure 1, arrows) were defined by the discharge into the lake from the river Mölndalsån and by the water level in the downstream Lake Stensjön (Table 2). The land boundary was defined by zero normal velocity. The temperature on the open boundaries was described as zero gradients.

The model was set up to account for the hydrometeorological conditions (wind and precipitation on the lake surface), the withdrawal of water from the lake by WTPs, the inflow to the lake from small streams and to simulate the heat exchange between the atmosphere and the lake (Table 2). The water density was formulated as a function of temperature. The horizontal and vertical eddy viscosities were simulated using Smagorinsky and k-epsilon formulations, respectively. The bed resistance was described by a constant roughness height of 0.05 m.

<table>
<thead>
<tr>
<th>Type of data</th>
<th>Time resolution</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind speed</td>
<td>3 hours</td>
<td>SMHI*</td>
</tr>
<tr>
<td>Wind direction</td>
<td>3 hours</td>
<td>SMHI*</td>
</tr>
<tr>
<td>Precipitation</td>
<td>1 day</td>
<td>Härryda municipality</td>
</tr>
<tr>
<td>Air temperature</td>
<td>3 hours</td>
<td>SMHI*</td>
</tr>
<tr>
<td>Relative air humidity</td>
<td>3 hours</td>
<td>SMHI*</td>
</tr>
<tr>
<td>Cleanness coefficient</td>
<td>3 hours</td>
<td>SMHI*</td>
</tr>
<tr>
<td>Withdrawal of water by WTPs</td>
<td>1 hour</td>
<td>Gothenburg Water, City of Mölndal</td>
</tr>
<tr>
<td>Discharge in the river Mölndalsån</td>
<td>1 day</td>
<td>City of Mölndal</td>
</tr>
<tr>
<td>Water level in Lake Stensjön</td>
<td>1 day</td>
<td>City of Mölndal</td>
</tr>
<tr>
<td>Inflow from small streams</td>
<td>1 day</td>
<td>Estimated*b</td>
</tr>
</tbody>
</table>

*aSwedish Meteorological and Hydrological Institute.
*bThe inflow to the lake from the small creeks was estimated using precipitation data for the catchment area of the respective creek and surface runoff coefficients, which were assigned based on the rate of exploitation and slope.

**Microbiological model for norovirus, E. coli O157/H7 and Cryptosporidium**

The fate and transport of pathogens in Lake Rådasjön were simulated using a microbiological water quality model ECO Lab (DHI 2004) which was coupled to the hydrodynamic model of the lake. ECO Lab uses flow fields from the hydrodynamic model to calculate the concentrations of pathogens or faecal indicators in the lake.

Inactivation of pathogens was described in the microbiological model according to Equation (3):

\[
\frac{dC}{dt} = -kC
\]

where \( k \) is the decay rate of pathogens in the water, \( t \) is the time and \( C \) is the pathogen concentration.

The decay coefficient for inactivation of norovirus and E. coli O157/H7 in the lake was described by Equation (4) (Mancini 1978; Sokolova et al. 2012):

\[
k = k_0 \theta_I^{20} \theta_T^{(Temp-20)}
\]

where \( k_0 \) (day\(^{-1}\)) is the decay rate at 20 °C for a salinity of 0‰ and darkness; \( \theta_I \) is the light coefficient; \( \text{Int} \) (kW m\(^{-2}\))
is the light intensity integrated over depth; $\theta_T$ is the temperature coefficient; and Temp (°C) is the water temperature.

It was assumed that the inactivation of norovirus and E. coli O157/H7 is well represented by the inactivation of faecal indicators somatic coliphages and E. coli, respectively (WHO 2008). The coefficients in Equation (4) for somatic coliphages and E. coli were experimentally determined in the microcosm trials performed during different seasons for the conditions of Lake Rådasjön (Sokolova et al. 2012).

The decay coefficient for the inactivation of Cryptosporidium in the lake is described by Equation (5) (Walker Jr & Stedinger 1999):

$$k = 10^{(0.058 \times \text{Temp} - 2.68)}$$

(5)

Sedimentation of Cryptosporidium was simulated according to Equation (6):

$$\frac{dC}{dt} = - \nu_s \frac{dz}{dC}$$

(6)

where $dz$ is the thickness of the layer and $\nu_s$ is the settling velocity. The settling velocity was specified as 0.03 m day$^{-1}$, which is the settling velocity previously suggested for free oocysts (Medema et al. 1998). It was conservatively assumed that Cryptosporidium oocysts released into the lake were not attached to particles (Brookes et al. 2006).

**RESULTS**

Estimated concentrations of norovirus, Cryptosporidium and E. coli O157/H7 in the contamination sources

The concentrations of norovirus, Cryptosporidium and E. coli O157/H7 in the contamination sources around the lake were estimated for endemic and epidemic conditions (Figure 3). The estimated pathogen concentrations in the contamination sources were higher under epidemic conditions: the difference between epidemic and endemic conditions (comparing 50% values calculated using Bacteroidales markers) was approximately 2 log units for norovirus, 5 log units for Cryptosporidium, and 2 log units for E. coli O157/H7 (Figure 3). The difference between the 50 and 95% values of the estimated (using Bacteroidales markers as faecal indicators) pathogen concentrations in the contamination sources was between 2 and 3 log units for all sources and pathogens under both endemic and epidemic conditions (Figure 3).

Comparing the 50% values calculated using Bacteroidales markers as faecal indicators, the highest norovirus concentrations in the contamination sources under both endemic and epidemic conditions were estimated for the emergency sewer overflow (site P), followed by the on-site sewers (sites 3 and 7) and the untreated stormwater source (site 18; Figure 3). In the same way, the highest Cryptosporidium concentrations in the contamination sources under endemic conditions were estimated for the emergency sewer overflow (site P) followed by the cattle grazing area...
(site 17), the on-site sewers (3 and 7) and the untreated stormwater source (site 18); under epidemic conditions the order was the same with the exception of the cattle grazing area (site 17), for which the lowest concentrations were estimated (Figure 3).

The comparison of the 50% values of pathogen concentrations showed that the differences of more than 1 log unit between the estimations using *E. coli* and *Bacteroidales* markers were observed for on-site sewers (site 7) and untreated stormwater (site 18; Figure 3).

**Simulated concentrations of norovirus, *Cryptosporidium* and *E. coli* O157/H7 at the water intake**

The output results from the coupled hydrodynamic and microbiological modelling were the time series of pathogen concentrations at the water intake of the city of Mölndal during the year 2008 (e.g. Figure 4). These time series demonstrated that the contributions from different contamination sources to the pathogen concentrations at the water intake can vary strongly over time, up to 4 log units in case of norovirus (Figure 4). To simplify the analysis, the results were also presented using the arithmetic average values of the time series (Figure 5).

The modelling results illustrated that, under both endemic and epidemic conditions, on-site sewers (site 3) contributed the most to the norovirus concentration at the water intake followed by the emergency sewer overflow (site P), the on-site sewers (site 7) and the untreated stormwater source (site 18; Figure 5). In the case of *Cryptosporidium*, under endemic conditions the source that contributed most to the concentration at the water intake was the cattle grazing area (site 17) followed by the on-site sewers (site 3), the emergency sewer overflow (site P), the on-site sewers (site 7) and the untreated stormwater source (site 18). Under epidemic conditions the order was the same with the exception of the cattle grazing area (site 17), which contributed the least (Figure 5). According to the simulations, under endemic conditions the contributions from the studied contamination sources to the norovirus, *Cryptosporidium* and *E. coli* O157/H7 concentrations at the water intake were below 1 No. L\(^{-1}\), while under epidemic conditions the pathogen concentrations at the water intake were expected to be much higher (Figure 5).

The simulations demonstrated that if the emergency sewer overflow (site P), the untreated stormwater source (site 18) and the cattle grazing area (site 17) were located in the vicinity of the water intake on the north-western shore of the lake, the contributions from these sources to the pathogen concentrations at the intake would be up to 1 log unit greater (Figure 5) than the contributions from their actual locations (Figure 1).
DISCUSSION

In this paper, we have suggested the combination of microbial source tracking with coupled hydrodynamic and microbiological modelling in order to obtain data on pathogen concentrations in a drinking water source. To the best of our knowledge, this is the first study where the microbial source tracking data were used to calculate pathogen concentrations in the contamination sources to provide input data for coupled hydrodynamic and microbiological modelling of pathogen fate and transport within a drinking water source.

The estimations of the pathogen concentrations in the microbial contamination sources were performed using BacH and BacR genetic markers as well as *E. coli* as faecal indicators. For the on-site sewers (site 7) and untreated stormwater (site 18) sources, the pathogen concentrations estimated using *E. coli* were higher than the estimations using BacH. A probable explanation for this is that the stream which transports discharges from the on-site sewers (site 7, Figure 1) also drains a pasture area where horses graze, and untreated stormwater (site 18, Figure 1) can possibly transport faecal contamination from cats and dogs present in this urban area. Faecal contamination from horses, cats and dogs may have contributed to the total *E. coli* concentrations measured in these contamination sources. This might have resulted in an overestimation of pathogen concentrations since the only pathogen hosts of interest in case of these contamination sources were humans. Since BacH concentrations represent contamination from humans only, the estimations based on this indicator were considered to be more accurate. This example illustrates the advantages of using microbial source tracking data to estimate pathogen concentrations in contamination sources.

The *Cryptosporidium* concentrations were measured in the discharges from on-site sewers (site 3) on eight different occasions during 2010–2011 and varied from 0 to 3 oocysts/10 L (data from Gothenburg Water). These values can be compared to the *Cryptosporidium* concentrations in discharges from on-site sewers (site 3) estimated for endemic conditions using BacH markers: 0.02 and 8.5 oocysts/10 L as 50 and 95% values, respectively (Figure 3). The *Cryptosporidium* concentrations were measured at the WTP water intake on 26 different occasions during 2006–2011 and varied from 0 to 14 oocysts/10 L (data from Gothenburg Water). The concentration 14 oocysts/10 L, which was measured during the winter period, corresponds to the simulated *Cryptosporidium* concentrations for epidemic conditions (Figure 5).

The modelling results have demonstrated that the source with the highest estimated pathogen concentrations is not necessarily the main contributor to the pathogen concentrations at the water intake (Figures 3 and 5). Furthermore, the pathogen concentrations at the water intake can vary strongly over time (Figure 4). This illustrates that the risks for the drinking water supply are dependent on factors such as the magnitude of discharge from the contamination source and the pathogen fate and transport within the water source (Ashbolt *et al.* 2010). This emphasises the value of performing hydrodynamic and microbiological modelling in order to estimate the impact of various contamination sources on the water quality in a water source (e.g. Liu...
et al. 2006; Hipsey et al. 2008; Thupaki et al. 2010; Schijven & De Roda Husman 2011).

In this study, we have combined microbial source tracking with fate and transport modelling in order to determine the dominant faecal sources and to provide input data for QMRA. In a study by Reischer et al. (2011), the dominant faecal sources were identified by linking microbial source tracking to catchment information. A faecal pollution source profile was used to formulate a hypothesis about the dominant sources of pollution in the catchment; this hypothesis was then tested using multi-parametric microbial source tracking. Schoen et al. (2011) estimated the contribution made by each source to the total infection risk in a water body, which was assumed to be impacted by a mixture of secondary-treated disinfected municipal wastewater and untreated sewage, by applying QMRA using norovirus as the reference pathogen.

The method suggested in this article provides the data needed to evaluate the current and the future safety of the drinking water supply in terms of infection risks. The greatest risks for waterborne disease outbreaks can be expected under epidemic conditions due to the increased load of pathogens to the water source. We have therefore suggested a method to investigate the contributions from the microbial contamination sources to the pathogen concentrations at the water intake under both endemic and epidemic conditions. Data from actual epidemiological studies can be utilised as input data for this method. Low pathogen concentrations and complexity of the microbial analyses often limit monitoring of pathogen concentrations at the water intake, and different surrogates have been suggested to overcome these limitations (e.g. Brookes et al. 2005; Field & Samadpour 2007). However, monitoring of pathogens or surrogates cannot provide information on the possible future risks (Ashbolt et al. 2010). The suggested method can therefore be used to address the limitations of monitoring. It has been demonstrated that this method can be used to identify, select and prioritise between sources of contamination and, consequently, to provide decision-support data on the development of the water source catchment area as well as to facilitate the prioritisation of mitigation measures (Bambic et al. 2011; Schijven & De Roda Husman 2011).

During the winter/spring period of 2010–2011, the presence of Cryptosporidium oocysts in the drinking water sources of Lake Storsjön and the river Skellefteälven (Sweden) in combination with insufficient treatment at the WTPs resulted in large waterborne disease outbreaks in the cities Östersund and Skellefteå, respectively. Most likely, Cryptosporidium oocysts entered the water sources with wastewater discharges and were transported to the intakes of the WTPs. Hydrodynamic and microbiological modelling integrated with microbial source tracking data would be appropriate to estimate the possible pathogen concentrations at the intakes to the WTPs in these two areas. Furthermore, the estimated pathogen concentrations at the intakes to the WTPs could be used as input data for QMRA to indicate the health risks for the consumers. In the case of the outbreak in Skellefteå, preliminary modelling results confirmed that wastewater discharges could have reached the raw water intake in the river Skellefteälven (Skellefteå Municipality 2011). The outbreak in Östersund occurred shortly after autumn destratification of Lake Storsjön, i.e. during the period when there were mixed conditions in the lake and the faecal contamination could reach the raw water intake located deep in the lake. These examples illustrate the importance of understanding the hydrodynamic processes in drinking water sources and the need for fate and transport modelling.

CONCLUSIONS

By combining microbial source tracking data with coupled hydrodynamic and microbiological modelling, we have estimated the pathogen concentrations in different contamination sources around a drinking water source and evaluated the contributions from these contamination sources to the pathogen concentrations at the WTP water intake. The suggested approach provides input data for QMRA and decision support data for risk management and the prioritisation of mitigation measures.

The concentrations of norovirus, Cryptosporidium and E. coli O157/H7 in different contamination sources around Lake Rådasjön were estimated for endemic and epidemic conditions based on measurements of E. coli and Bacteroidales markers in these contamination sources. The estimated pathogen concentrations in the contamination sources under epidemic conditions were up to 5 log units
higher than under endemic conditions. The estimations based on Bacteroidales markers were considered to be source specific and thus more accurate than the estimations based on E. coli. From the results of coupled hydrodynamic and microbiological modelling, we can conclude that under endemic conditions the on-site sewers (site 3) were the source that contributed the most to the norovirus concentrations at the WTP water intake in Lake Rådasjön; the cattle grazing area (site 17) was the main contributor to the Cryptosporidium concentrations. Under epidemic conditions the on-site sewers (site 3) contributed the most to both norovirus and Cryptosporidium concentrations at the water intake. This modelling approach has proved to be a useful tool for simulating the pathogen fate and transport in a water source.

It can be concluded that integrating microbial source tracking with hydrodynamic and microbiological modelling yields more information than can be gained from applying these methods separately. The approach reported in this article can be used to prevent waterborne disease outbreaks caused by the faecal contamination of surface drinking water sources.

ACKNOWLEDGEMENTS

This research was funded by the Graduate School on Environment and Health (Miljö och Hälsa Forskarskolan) at Chalmers University of Technology and the University of Gothenburg, Sweden; the Swedish Water and Wastewater Association (Svenskt Vatten); and the EU project VISK (Interreg IV A program). The authors are also grateful to the Rådasjön project (SVU-project 29-122) management group for valuable discussions as well as to Gothenburg Water, the Härryda municipality, the city of Mölndal and the Swedish Meteorological and Hydrological Institute for providing data for this research.

REFERENCES


Silverlås, C., De Verdier, K., Emanuelson, U., Mattsson, J. G. & Björkman, C. 2010 Cryptosporidium infection in herds with


Sokolova, E., Borell Lövstedt, C. & Pettersson, T. J. R. 2011 Fate and transport modelling of microbial pollution in a lake used as a drinking water source. In *Proceedings of the 34th IAHR World Congress*, Brisbane, Australia, 26 June–1 July.


