Title	Bayesian modeling of virus removal efficiency in wastewater treatment processes
Author(s)	Ito, T.; Kato, T.; Takagishi, K.; Okabe, S.; Sano, D.
Citation	Water Science and Technology, 72(10), 1789-1795 https://doi.org/10.2166/wst.2015.402
Issue Date	2015-11
Doc URL	http://hdl.handle.net/2115/61970
Rights	©IWA Publishing 2015. The definitive peer-reviewed and edited version of this article is published in Water Science and Technology 72 (10) 1789-1795 2015 DOI: 10.2166/wst.2015.402 and is available at www.iwapublishing.com .
Type	article (author version)
File Information	Ito2014WST_rev_v2.pdf

[Instructions for use](#)

Bayesian modeling of virus removal efficiency in wastewater treatment processes

T Ito*, T Kato**, K. Takagishi***, S Okabe****, D Sano*****

*Division of Environmental Engineering, Faculty of Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan, E-mail: s26135027e@ec.hokudai.ac.jp

** Department of Computer Science, Graduate School of Engineering, Gunma University, Tenjinmachi 1-5-1, Kiryu, Gunma 376-8515, Japan, E-mail: katotsu@cs.gunma-u.ac.jp

*** Division of Electronics and Informatics, Faculty of Science and Technology, Gunma University, , Tenjinmachi 1-5-1, Kiryu, Gunma 376-8515, Japan.

**** Division of Environmental Engineering, Faculty of Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan, E-mail: sokabe@eng.hokudai.ac.jp

***** Division of Environmental Engineering, Faculty of Engineering, Hokkaido University, North 13, West 8, Kita-ku, Sapporo, Hokkaido 060-8628, Japan, E-mail: dsano@eng.hokudai.ac.jp

Abstract: Left-censored datasets of virus density in wastewater samples make it difficult to evaluate the virus removal efficiency in wastewater treatment processes. In the present study, we modelled the probabilistic distribution of virus removal efficiency in a wastewater treatment process with Bayesian approach, and investigated how many detect samples in influent and effluent are necessary for the accurate estimation. One hundred left-censored data of virus density in wastewater (influent and effluent) were artificially generated based on assumed log-normal distributions and the posterior predictive distribution of virus density and the log-ratio distribution were estimated. The estimation accuracy of distributions was quantified by Bhattacharyya coefficient. When it is assumed that the accurate estimation of posterior predictive distributions is possible when 100% positive rate is obtained for 12 pairs of influent and effluent, 11 out of 144, 60 out of 324, and 201 out of 576 combinations of detect samples gave the accurate estimation at the significant level of 0.01 in Kruskal-Wallis test when the total sample number was 12, 18, and 24, respectively. The combinations with the minimum number of detect samples were (12, 9), (16, 10), and (21, 8) when the total sample number was 12, 18, and 24, respectively.

Keywords: Bayesian model, left-censored data, truncated log-normal distribution, virus removal efficiency, wastewater treatment

Introduction

Enteric viruses, including human noroviruses, are causing outbreaks of waterborne diseases worldwide (Bosch et al., 2008). Since extremely high numbers of these pathogenic viruses are shed in feces of infected or asymptomatic individuals (Lee et al., 2007; Garcia et al., 2006), it is necessary to pay great attention to the virus removal efficiency in wastewater treatment process for preventing viral contamination of water environments receiving effluent from wastewater treatment plants. If the virus removal efficiency of each element process of wastewater treatment is assessed prior to the operation, it is possible to estimate the total virus removal efficiency in a wastewater treatment process and to calculate infection risks arising from the usage of treated wastewater for various purposes such as irrigation by quantitative microbial risk assessment (QMRA) (Shuval et al., 1997).

However, the evaluation of virus removal efficiency in wastewater treatment process can be a challenging task. Although the virus removal efficiency of wastewater treatment process is usually evaluated as a ratio of virus density in effluent to that in influent of a wastewater treatment process (Ottoson et al., 2006), non-detect samples in which the quantity of enteric viruses is below the analytical quantification limit make it difficult to obtain a precise ratio of the virus density before and after a treatment step. Particularly, treated wastewater is prone to yielding left-censored datasets, which have high numbers of non-detects (Helsel, 2006).

22

23 To address this problem of left-censored data, several statistical approaches have been proposed, in
 24 which the density of object substances is expressed by probabilistic density functions (PDF)
 25 (Kennedy and Hart, 2009; Kennedy, 2010; EFSA, 2010). Paulo et al proposed a Bayesian approach
 26 adapted for left-censored data of residual pesticide concentrations in food (Paulo et al., 2005). In
 27 our previous study, the Paulo model was applied to the artificially created enteric virus density in
 28 wastewater with a slight modification, in which the occurrence of the real zero of virus density is
 29 not assumed (Kato et al., 2013). Our previous study concluded that eight or more detect samples in
 30 a dataset (up to the total sample number of 48) are required to accurately estimate the posterior
 31 predictive distributions of virus density and the log-ratio posterior distributions as a virus removal
 32 efficiency, when 100% of untreated wastewater samples (influent) give positive results of the
 33 presence of enteric viruses (Kato et al., 2013). However, the prerequisite condition of 100% positive
 34 rate in influent samples does not coincide with reality, and non-detects also frequently appear in
 35 influent samples. It is necessary to confirm what level of accuracy in the estimation of posterior
 36 predictive distribution is obtained for a given sample size and number of non-detect.

37

38 The aim of this study is to clarify the minimum number of detect samples for the estimation of the
 39 probabilistic distribution of virus removal efficiency. One hundred paired datasets of virus density
 40 in influent and effluent were generated artificially from log-normal distributions for each
 41 combination of a sample size (12, 18, or 24) and the number of detected sample (from 1 to 24) in
 42 influent and effluent. The left-censored datasets of influent and effluent were created by setting a
 43 limit value of analytical quantification to obtain the assumed number of detect samples. Then, the
 44 modified Paulo model was applied to the generated left-censored datasets, in order to estimate
 45 posterior predictive distribution of virus density in influent and effluent. A log-ratio posterior
 46 distribution, regarded as virus removal efficiency was also obtained by dividing two posterior
 47 predictive distributions of virus density in influent and effluent. Finally, the difference between the
 48 true distribution and log-ratio posterior distribution was evaluated by Bhattacharyya coefficient.

49

Material and Methods

50 Estimation of the predictive distribution of enteric virus density and the log-ratio distribution

51 In this study, a parametric probabilistic model called the truncated log-normal distribution is
 52 employed to represent the underlying distribution of enteric virus concentrations. Each of
 53 observations in a given dataset is written by a tuple (x,y), where x is the numerical value of the
 54 observed enteric virus concentration and y is a Bernoulli variable indicating the presence of the
 55 detect; y = 1 if the observation is detected, and y = 0, otherwise. The value of x is undefined
 56 if y = 0. Each tuple has an additional variable θ representing the limit of quantification. In the
 57 truncated log-normal model, the concentration x is assumed to be drawn from the following PDF:
 58
$$TLN(x; \mu, \beta^{-1}, \theta) := \frac{1}{Z(\mu, \beta^{-1}, \theta)x} \exp\left(-\frac{\beta}{2}(\mu - \log_{10} x)^2\right) \quad \text{where} \quad Z(\mu, \beta^{-1}, \theta) := \sqrt{2\pi} \ln(10) \cdot$$

 59
$$\left(1 - \varphi\left(\sqrt{\beta}(\theta - \mu)\right)\right) \beta^{-1} \text{ and } \varphi \text{ is the cumulative density function of the standard normal}$$

 60 distribution. The probability of the Bernoulli variable is given by $\left(1 - \varphi\left(\sqrt{\beta}(\theta - \mu)\right)\right)$ for y = 1,
 61 and by $\varphi\left(\sqrt{\beta}(\theta - \mu)\right)$ for y = 0. This probabilistic model thus contains two model parameters, μ
 62 and β , to be inferred.

63

The likelihood function is required to infer the model parameters. Let us denote a dataset either of influent or of effluent by $X = \{(x_i, y_i)\}_{i=1}^n$, where the total number of samples is n in the dataset. Let θ_i be the quantification limit for i -th sample (x_i, y_i) . The likelihood function can then be written as $p(X|\mu, \beta) = \prod_{i=1}^n \left(\varphi\left(\sqrt{\beta}(\theta_i - \mu)\right) \right)^{1-y_i} \left(\left(1 - \varphi\left(\sqrt{\beta}(\theta_i - \mu)\right)\right) \text{TLN}(x_i; \mu, \beta^{-1}, \theta_i) \right)^{y_i}$. To obtain the inference of the model parameters in the form of a posterior distribution, say $p(\mu, \beta|X)$, this study adopted the Paulo et al.'s prior $\mu \sim N(0, 100)$ and $\beta \sim \text{Gam}(0.01, 0.01)$ where $N(m, v)$ and $\text{Gam}(a, b)$ (Paulo et al., 2005), respectively, denote the normal distribution with mean m and variance v and the Gamma distribution with shape parameter a and rate parameter b .

The posterior predictive distributions of an unknown concentration can be obtained by applying Bayesian inference to the posterior distribution of the model parameters. Given a dataset X , the probabilistic density at the common logarithm of a concentration $x_{\log}$ is written as $p_{\text{pred}}(x_{\log}|X) = \int N(x_{\log}; \mu, \beta^{-1}) p(\mu, \beta|X) d\mu d\beta$. It would be ideal if $p_{\text{pred}}(x_{\log}|X)$ was close to $N(x_{\log}; \mu_*, \beta_*^{-1})$ where (μ_*, β_*^{-1}) is the true value of the model parameter. When both a dataset of influent X_{inf} and a dataset of effluent X_{eff} are given, the probabilistic distribution of the log-ratio between two respective concentrations can be computed from the corresponding posterior predictive distributions. The probabilistic distribution of the log ratio is simply referred to as the log-ratio distribution, and denoted by $p(r|X_{\text{inf}}, X_{\text{eff}})$.

The software implementing the algorithm developed for inferring posterior predictive distribution of virus removal efficiency is available upon request to the corresponding author.

Bhattacharyya coefficient

Provided that the two datasets, X_{inf} and X_{eff} , are artificially generated, the true distribution of the log ratio is known, allowing us to assess the accuracy of the inference by comparing the inferred distribution with the true distribution. In this study, the Bhattacharyya coefficient (Bhattacharyya, 1943; Derpanis, 2008; Kazama et al., 2010) is employed for assessment of inferred distributions. Denoting by $p_{\text{true}}(r)$ the true distribution, the Bhattacharyya coefficient is defined by $\text{BC} = \int p(r|X_{\text{inf}}, X_{\text{eff}}) p_{\text{true}}(r) dr$. The coefficient takes a value between zero and one. Better inference gets a higher Bhattacharyya coefficient. $\text{BC} = 1$ implies the exact inference.

Statistical treatment

We assumed that the accurate estimation of posterior predictive distributions is possible when 100% positive rate is obtained for 12 sample pairs of influent and effluent. This assumption can be translated that the estimation is regarded to be accurate when Bhattacharyya coefficient values of certain combination of detect samples are not statistically different from those at 100% positive rate for 12 pairs of influent and effluent. For this purpose, Kruskal-Wallis test at a significant level of 0.01 was performed under the null hypothesis that the median value of Bhattacharyya coefficient is identical with that at 100% positive rate for 12 pairs of influent and effluent. Before performing the Kruskal-Wallis test, outliers in the Bhattacharyya coefficient values were detected by using interquartile range (IQR) between first (25%tile) and third (75%tile) quartiles, in which any Bhattacharyya coefficient values at a greater distance from first or third quartiles than 1.5 times IQR were regarded as outliers. The H statistic in Kruskal-Wallis test is defined by $H = \frac{12}{N(N+1)} \sum_{i=1}^c \frac{R_i^2}{n_i} -$

107 3(N + 1), where N is the total number of data and C is the number of detected sample combination.
 108 When the median values of three combinations of detect samples (ex., (12, 12), (12, 9), and (9, 12))
 109 that have 100 Bhattacharyya coefficient each were compared in Kruskal-Wallis test, the number of
 110 detected sample combination C = 3 and the total number of data N = 300. The mean of ranks in
 111 each group of the Bhattacharyya coefficient values is found by dividing the sum of ranked scores
 112 ($\sum R_i$) by n_i observations in the i-th group. In the matter of the correction factor for ties, it should be
 113 denoted by $H_0 = \frac{H}{1 - \sum(t_i^3 - t_i)/(N^3 - N)}$, where t_i is the number of tied value in the i-th rank. The H_0
 114 statistic (chi-square value) has been shown to be distributed approximately as a chi-square
 115 distribution ($df = C - 1$).
 116

Results and Discussion

117 Accuracy of the estimated distribution of virus removal efficiency when the total sample 118 number is 12 each for influent and effluent

119 When the total sample number is 12 each for influent and effluent, there are 144 combinations of
 120 detect samples from (1, 1) to (12, 12). One hundred pairs of virus density data in influent and
 121 effluent were generated for each combination of detect samples, based on two assumed (true) log-
 122 normal distributions ((μ , sigma) = (4, 1) and (1, 1)), respectively. Then, 100 each of posterior
 123 predictive distributions of virus density in influent and effluent were estimated, and 100 log-ratio
 124 posterior distributions for each combination of detect samples were subsequently obtained. These
 125 100 log-ratio posterior distributions were compared with the log-ratio distribution derived from the
 126 true log-normal distributions which were used to generate simulated datasets for evaluating the
 127 estimation accuracy, by using Bhattacharyya coefficient. One Bhattacharyya coefficient value was
 128 obtained for one inferred log-ratio posterior distribution, which means that 100 values of
 129 Bhattacharyya coefficients were calculated for each combination of detected sample.
 130

131 Percentiles (5, 25, 50, or 75%tile) of Bhattacharyya coefficient values when the total sample
 132 number is 12 each for influent and effluent are indicated in Fig. 1. The x-axis and y-axis are the
 133 number of detect samples of influent and effluent, respectively, and the z-axis is the value of
 134 Bhattacharyya coefficient ranging from 0 to 1. The 25 and 75%tile values of Bhattacharyya
 135 coefficient were used for the rejection of outliers using IQR. As intuitively, the Bhattacharyya
 136 coefficient values are increased with the number of detect samples, although how large
 137 Bhattacharyya coefficient value is enough for assuring the estimation accuracy is not clear. This is
 138 why we investigated the relative accuracy of the estimation by comparing with the Bhattacharyya
 139 coefficient values at 100% positive rate for 12 pairs of influent and effluent.
 140

141 Bhattacharyya coefficient values without outliers are shown in Fig. 2(a). The result of Kruskal-
 142 Wallis test indicated that 11 out of 144 combinations of the number of detect samples gave
 143 comparative estimation accuracy to that at 100% positive rate for 12 pairs of influent and effluent
 144 (Fig. 2(b)). The minimum number of detect samples of 9 was acceptable, when 12 of the coupled
 145 samples were all positive in the virus quantification. In our previous study, 8 or more detect samples
 146 in treated wastewater (effluent) were required for the accurate estimation, when the positive rate in
 147 influent was 100% and the total sample number is 12 each for influent and effluent (Kato et al.,
 148 2013). While the present study indicated that more than 11 detect samples under the same condition
 149 (100% of influent samples are positive). This difference is attributable to the employment of
 150 different parameters for evaluating the estimation accuracy. Kullback-Leibler divergence was used
 151 in the previous study, and Bhattacharyya coefficient was used in the present study. Although these
 152 two parameters are able to give us the information of estimation accuracy, we had better employ the
 153 larger value of 9 detect samples out of total sample number of 12 from a conservative viewpoint.

155
156 **Figure 1.** Percentile values of Bhattacharyya coefficient between the true distribution of virus
157 removal efficiency and estimated log-ratio posterior distribution when the total sample number is 12
158 each for influent and effluent. (a) 5%tile, (b) 25%tile, (c) 50%tile, and (d) 75%tile. The x-axis and
159 y-axis are the number of detect samples in influent and effluent, respectively. The z-axis is the value
160 of Bhattacharyya coefficient (ranging from 0 to 1). Level lines on the bottom panel are indicating
161 Bhattacharyya coefficient values of 0.9, 0.8 and 0.7.

164
165 **Figure 2.** Estimation accuracy of log-ratio distribution depending on combinations of detect
166 samples in influent and effluent. (a) Box plot of Bhattacharyya coefficient values without outliers
167 when the total sample number is 12 each for influent and effluent. Maximum, minimum and
168 quartile values are indicated. (b) White cell indicates that the combination of detect samples gives
169 accurate estimation at the significant level of 0.01 in Kruskal-Wallis test, while grey cell does not.

Accuracy of the estimated distribution of virus removal efficiency when the total sample number is 18 each for influent and effluent

When the total sample number is 18 each for influent and effluent, there are 324 combinations of detect samples from (1, 1) to (18, 18). One hundred paired datasets of virus density in influent and effluent were generated for each combination of detect samples as well as the total sample number was 12 each. One hundred Bhattacharyya coefficient values between the true distribution and the log-ratio posterior distribution were computed, and outliers are excluded as well. Bhattacharyya coefficient values without outliers are shown in Fig. 3(a). As a result, 60 out of 324 combinations of the number of detect samples gave comparative estimation accuracy to that at 100% positive rate for 12 pairs of influent and effluent (Fig. 3(b)). The minimum number of detect samples was 10, when the companion number of detect samples is larger than 15 (Fig. 3(b)).

Figure 3. Estimation accuracy of log-ratio distribution depending on combinations of detect samples in influent and effluent. (a) Box plot of Bhattacharyya coefficient values without outliers when the total sample number is 18 each for influent and effluent. Maximum, minimum and quartile values are indicated. The asterisk indicated the Bhattacharyya coefficient values without outliers at 100% positive rate for 12 pairs of influent and effluent. (b) The white cell indicates that the combination of detect samples gives accurate estimation at the significant level of 0.01 in Kruskal-Wallis test, while the grey cell does not.

Accuracy of the estimated distribution of virus removal efficiency when the total sample number is 24 each for influent and effluent

When the total sample number is 24 each for influent and effluent, there are 576 combination of detect samples from (1, 1) to (24, 24). One hundred paired datasets of virus density in influent and effluent were generated for each combination of detect samples as well as the total sample number was 12 or 18 each. One hundred Bhattacharyya coefficient values between the true distribution and the log-ratio posterior distribution were computed, and outliers are excluded as well. Bhattacharyya coefficient values without outliers are shown in Fig. 4(a). As a result, 201 out of 576 combinations of the number of detect samples gave comparative estimation accuracy to that at 100% positive rate for 12 pairs of influent and effluent (Fig. 4(b)). The minimum number of detect samples was 8, when the companion number of detect samples is larger than 20 (Fig. 4(b)).

210 (a)

(b)

211 **Figure 4.** Estimation accuracy of log-ratio distribution depending on combinations of detect
212 samples in influent and effluent. (a) Box plot of Bhattacharyya coefficient values without outliers
213 when the total sample number is 24 each for influent and effluent. Maximum, minimum and
214 quartile values are indicated. The asterisk indicated the Bhattacharyya coefficient values without
215 outliers at 100% positive rate for 12 pairs of influent and effluent. (b) The white cell indicates that
216 the combination of detect samples gives accurate estimation at the significant level of 0.01 in
217 Kruskal-Wallis test, while the grey cell does not.

218
219
220
221 These results indicate that the estimation accuracy depends on the number of detect samples, but the
222 dependency was not clear for the positive rate of virus quantification. In the case of other occasions,
223 such as the total sample number larger than 24 and unequal sample numbers between influent and
224 effluent, it is better to estimate the estimation accuracy using artificially generated data as
225 performed in this study. Furthermore, we assumed a log-normal distribution of enteric virus density
226 in wastewater samples, but the other distributions such as gamma distribution should be also tested
227 in the further study.

228
229 The virus removal mechanism is complex, and its efficiency depends on the types of wastewater
230 treatment and species of enteric viruses (Miura et al., 2015). The difference in virus removal
231 tendency is reflected by different estimation results in the proposed approach, and the number of
232 positive samples needed for the estimation is not affected by the particularities of wastewater
233 treatment and enteric virus species theoretically. The methodology for virus quantification also does
234 not affect the estimation results, but the algorithm users have to use the identical methodology for
235 both influent and effluent, although the quantification limit values are not necessary to be identical
236 between influent and effluent samples.

237
238 The algorithm developed in this study for inferring posterior predictive distribution of virus removal
239 efficiency has distinguished advantages compared to already-published ones in terms of the
240 treatment of left-censored data, but its availability should be compared with other approaches in the
241 further study.

242 Conclusions

243 The estimation accuracy of log-ratio distributions as the probabilistic distributions of virus removal

244 efficiency in wastewater treatment processes was dependent on the number of detect samples, rather
245 than positive rate of virus quantification. When it is assumed that the accurate estimation of
246 posterior predictive distributions is possible when 100% positive rate is obtained for 12 pairs of
247 influent and effluent, 11 out of 144, 60 out of 324, and 201 out of 576 combinations of detect
248 samples gave the accurate estimation when the total sample number was 12, 18, and 24,
249 respectively. The combinations with the minimum number of detect samples were (12, 9), (16, 10),
250 and (21, 8) when the total sample number was 12, 18, and 24, respectively.

References

- 251 Bhattacharyya, A. (1943) On a measure of divergence between two statistical populations defined
252 by their probability distribution. *Bulletin of the Calcutta Mathematical Society*, **35**, 99–110.
- 253 Bosch A., Guix S., Sano D. and Pinto R. (2008) New tools for the study and direct surveillance of
254 viral pathogens in water. *Current Opinion in Biotechnology*, **19**, 295-301.
- 255 Derpanis, K. G. (2008). The bhattacharyya measure. *Mendeley Computer*, **1**(4), 1990-1992.
- 256 European Food Safety Authority. (2010) Management of left-censored data in dietary exposure
257 assessment of chemical substances. *EFSA Journal*, **8**(3), 1557.
- 258 García C., DuPont H. L., Long K. Z., Santos J. I. and Ko, G. (2006) Asymptomatic norovirus
259 infection in Mexican children. *Journal of Clinical Microbiology*, **44**(8), 2997-3000.
- 260 Helsel D. R. (2006) Fabricating data: How substituting values for nondetects can ruin results, and
261 what can be done about it. *Chemosphere*, **65**(11), 2434-2439.
- 262 Kato T., Miura T., Okabe S. and Sano D. (2013) Bayesian modeling of enteric virus density in
263 wastewater using left-censored data. *Food and Environmental Virology*, **5**(4)185-193.
- 264 Kazama, J. I., De Saeger, S., Kuroda, K., Murata, M., and Torisawa, K. (2010). A Bayesian method
265 for robust estimation of distributional similarities. In Proceedings of the 48th Annual Meeting
266 of the Association for Computational Linguistics, 247-256 Association for Computational
267 Linguistics.
- 268 Kennedy M. C. (2010) Bayesian modeling of long-term dietary intakes from multiple sources. *Food
269 and Chemical Toxicology*, **48**, 250-263.
- 270 Kennedy M. and Hart A. (2009) Bayesian modeling of measurement errors and pesticide
271 concentration in dietary risk assessments. *Risk Analysis*, **29**(10), 1427-1442.
- 272 Lee N., Chan M. C. W., Wong B., Choi K. W., Sin W., Lui G., Chan P. K. S., Lai R. W. M.,
273 Cockram C. S., Sung J. J. Y. and Leung W. K. (2007) Fecal viral concentration and diarrhea in
274 norovirus gastroenteritis. *Emerging Infectious Diseases*, **13**(9), 1399-1401.
- 275 Miura T, Okabe S, Nakahara Y. and Sano D. (2015) Removal properties of human enteric viruses in
276 a pilot-scale membrane bioreactor (MBR) process. *Water Research*, **76**, 33-42.
- 277 Ottoson J., Hansen A., Bjorlenius B., Norder H. and Stenstrom T. A. (2006) Removal of viruses,
278 parasitic protozoa and microbial indicators in conventional and membrane processes in a
279 wastewater pilot plant. *Water Research*, **40**, 1449-1457.
- 280 Paulo M. J., van der Voet H., Jansen M. J. W., ter Braak C. J. F. and van Klaveren J. D. (2005) Risk
281 assessment of dietary exposure to pesticides using a Bayesian method. *Pest Management
282 Science*, **61**(8), 759-766.
- 283 Sano D., Ueki Y., Watanabe T. and Omura, T. (2006) Membrane separation of indigenous
284 noroviruses from sewage sludge and treated wastewater. *Water Science and Technology*, **54**(3),
285 77-82.
- 286 Shuval H., Lampert Y., Fattal B. (1997). Development of a risk assessment approach for evaluating
287 wastewater reuse standards for agriculture. *Water Science and Technology*, **35**(11), 15-20.
- 288 USEPA. (2012). Guidelines for water reuse. *USEPA and US Agency for International Development*,
289 Washington, DC.