VATTEN – Journal of Water Management and Research 68: 43–51. Lund 2012

ON SELECTION OF THRESHOLD VALUES IN ALARM-SYSTEMS FOR DRINKING WATER MONITORING

Om val av larmvärde vid dricksvattenkontroll i realtid

by JING LI[1] and KENNETH M PERSSON[2]
1 Predect i Stockholm AB, Invernessvägen 6–8, 182 76 Stocksund; jing.li@predect.se
2 Water Resources Engineering, PO Box 118, 221 00 LUND; Kenneth.persson@tvrl.lth.se,
Sydvatten AB, Skeppsgatan 19, 211 19 Malmö; Kenneth.m.persson@sydvatten.se

Abstract

During a one year long research project at Borensberg Waterworks, Motala, Sweden, water quality with respect to microscopic particle counts both in raw water and drinking water have been monitored by an online water quality monitoring system. Microscopic particle counts were documented with absolute values in real-time. In this paper, a discussion is introduced on how to analyze large data sets from this kind of monitoring and how to assess threshold values in practice for particle counts in water samples. Statistical tools were used for data sorting and arrangement. Log Reduction is used for comparison of two water qualities and visualization of the treatment process within the waterworks. Numerical analysis of GEV (Generalized Extreme Value) followed by recurrence curve analysis is suggested to be applied for threshold value definition. Results from threshold value definition show that the starting value should be determined by a combination of the theoretical analysis and the practical situation. A calibration of values for specific waterworks should be taken into account when choosing relevant threshold values for water quality monitoring.

Key words – Online Water Quality Monitoring System, microbiological water quality, microbiological contaminants, Borensberg Waterworks, Log Reduction, GEV (Generalized Extreme Value), threshold value definition

Sammanfattning

Med hjälp av ett automatiserat mätsystem för partikelhalter i vatten har vattenkvaliteten vid Borensbergs vattenverk i Motala kommun undersökts under ett års tid med avseende på mikroskopiska partiklar i dricksvatten och råvatten. Antalet mikroskopiska partiklar registrerades som absoluta värden i realtid. I denna artikel diskuteras hur larmgränser skall definieras i praktiken för partikelhalter och hur stora datamängder kan utvärderas för detta slags studier. Statistiska analysverktyg användes för att sortera och gruppera mätdata. Log-reduktion av antal partiklar användes för att jämföra olika slags vattenkvaliteter och för att tydliggöra beredningseffekten i vattenverket. En numerisk analys av generaliserade extremvärden följt av kurvanalys för återkomstfrekvensen föreslås tillämpas för att definiera lämpliga larmgränser. Resultaten av larmgränsdefinitioner visar att ursprungsvärdet bör bestämmas genom en kombination av teoretisk analys och erfarenheter från den faktiska anläggningen. En kalibrering behöver också göras i det specifika vattenverket för att bestämma lämpligast larmvärde i det enskilda fallet.

Introduction

Clean drinking water is one of the most essential and valuable resources for humanity. Yet it is becoming a highly threatened resource for human beings, due to population growth and lack of proper sanitation. It was estimated by WHO that in a global scale, about 900 million people lack access to clean water resources and nearly 2/5 of the world's population (2.6 billion people) lack access to improved sanitation services (WHO, 2010). Furthermore, there is evidence that diseases related to water, sanitation and hygiene account for around 2,213,000 deaths per year according to Disability Adjusted Life Years (DALYs) (Prüss et al., 2002).

Water is not only life-sustaining to humans but also for the survival of all organisms. Water-borne diseases caused by drinking water contaminated by human or animal faeces containing pathogenic microorganisms are transmitted directly when the water is consumed. Pathogens such as Cryptosporidium, Giardia and E-coli bacteria are the sources of frightening diseases such as Anemia, Cholera, Giardiasis and Diarrhoea, etc. (WHO, 2010).

One measure to limit problems with water-borne diseases can be to apply online monitoring to indicate presence of microscopic particles in waterworks and water mains. Online water quality monitoring systems based on sensor technologies have been developed, allowing water utilities to continuously monitor the quality in terms of physical and chemical indicators through the drinking water distribution systems in real time. Several suppliers have developed technologies to support the on-line monitoring, yet a thorough discussion of how to interpret monitoring results in terms of microbial risks is needed.

This paper aims to address the applicability of on-line monitoring and to evaluate how statistical analysis of particle distributions may be used to visualize the particle counts variation in a study period. Similarities and differences between raw water and drinking water are compared and a numerical method (generalized extreme value, GEV) is introduced to define threshold values for water quality monitoring.

Specified requirements for alerting or devices that warn when errors occur in the water are commented in Guidelines for drinking water from the Swedish Food Authority. (Vägledning, Swedish Food Authority, 2001). There are clear guidelines and requirements about the alarming system on what should be monitored. It defines the alarming system as the one which detects and records the data at the point (time and space) in which errors can arise or trigger a warning in the form of a clearly audible and/or visual signal when a numeric measurement value (alarm limit) is reached.

The pH adjustment, disinfection and monitoring of turbidity are parameters where errors must be avoided and water quality must be monitored on-line. In such cases, detection, alert and warning should take place continuously and automatically (Vägledning, 2001).

The objectives for the study were to use statistical analysis to assort particle measurement data and demonstrate the particle distribution through the study period. Based on these data, the similarities and difference between raw water and drinking water particle counts can be evaluated. Through maximum daily values, a numerical method to define the Threshold value (THV) for the monitoring system is introduced and the implications for practical use of THV discussed.

Methodology

Generally, there are three types of particles: inert organic, viable organic and inert inorganic. From a microbiological point of view, only viable organic particle pose a threat for health. Actual value of particle counts (Particle counts/ml) is used to demonstrate the particle number variation both for raw water and drinking water through the whole study period. Statistical methods were used for data sorting and arrangement and comparison between raw water and drinking water.

Log Reduction has been used as a measure to show the efficiency of water treatment. It is reported as a good and simple way to interpret waterworks' efficiency by calculating and visualizing the particle reduction before and after treatment, in logarithmic scale. The definition of Log Reduction:

$$\text{Log Reduction} = -\log\left(\text{Drinking water particle content} / \text{Raw water particle content}\right)$$

For surface water treatment, a 99.9 % (3 logs) removal or inactivation of Giardia and a 99.99 % (4 logs) removal or inactivation of entire viruses has been suggested (Hatukai et al., 1997). Furthermore, to be considered as a safe barrier, the Log Reduction for such a treatment step should be at least 3 logs in term of viable organic particles.

As a basis of the evaluation in this study, particle counts (1–25 microns) in raw water (from Boren Lake) and treated drinking water in Borensberg Waterworks were studied from October of 2009 to September 2010. Two different continuous water quality monitoring systems from Predect AB were installed (see details in Persson et al., 2011). The first, P-100, monitored the micro particle counts in the raw water, while the second, P-300, monitored the micro particle counts in treated drinking water. Both can work as an early warning system for water quality monitoring, continuously recording the particle counts every minutes (24 h/day, 7 days/week).

The operator is alerted through a text message which is sent when the particle count exceeds a defined THV. The differences between P-100 and P-300 are that P-100 monitors two different fractions (1–2 and 2–25 µm) while P-300 registers four different micro particle fractions from 1 to 25 Microns (1–2 µm, 2–7 µm, 7–15 µm and 15–25 µm). This is based on potential probability of presence of specific bacteria and parasite. Cryptosporidium and Giardia are generally in the size of 2–7 μm. Data related to the particle counts for different fractions are stored and visualized in tabular form for further distribution analysis. At the same time, the real time distribution curve is visualized on the computer screen within the system. In addition, the P-300 can sample water automatically if a defined THV is exceeded, which P-100 cannot.

Numerical analysis

A pure numerical analysis is used for comments about definition of threshold. Estimation of extreme values is used for the probability of events that are more extreme than any that has already been observed and could be used for any kind of extreme values analysis. GEV Distribution is determined by Equation 1.

$$g\,(z,\,\mu,\,\sigma,\,\xi) =$$

$$\frac{1}{\sigma}\left[1+\frac{\xi(z-\mu)}{\sigma}\right]^{1-1/\xi}\exp\left\{-\left[1+\frac{\xi(z-\mu)}{\sigma}\right]^{-1/\xi}\right\} \qquad \text{(Eq. 1)}$$

Where z: is the value of your variable, σ is the scale parameter, μ is the location parameter, and ξ is the shape parameter. The shape parameter, as indicated in the name, gives the GEV distribution different shape with different characteristics. Nominally, three types of shape parameters can be identified, depending on the size of the shape distribution:

Type I: $\xi \to 0$ (Gumbel distribution)
Type II: $\xi > 0$ (Fréchet distribution)
Type III: $\xi < 0$ (Weibull distribution)

Figure 1 gives a clear profile of these three types of GEV distributions with the following characteristics; Gumbel: light upper tail, Fréchet: heavy upper tail, and Weibull: bounded upper tail.

The distribution type is defined by the shape parameter ξ. The Gumbel distribution is commonly used in probability theory and statistical analysis to model the various distributions of the maximum/minimum of a number of samples. In hydrology, it can be used to present hydrological extreme events, such as the distribution of the maximum level of a river in a particular year based on observed maximum values for the past ten years. It is helpful to predict the risk that an extreme earthquake, flood or other natural disaster will arise.

The Probability Density Function and Cumulative Distribution Function and the Inverse Function of Interval Time are described in Eq. 2, Eq. 3 and Eq. 4, respectively.

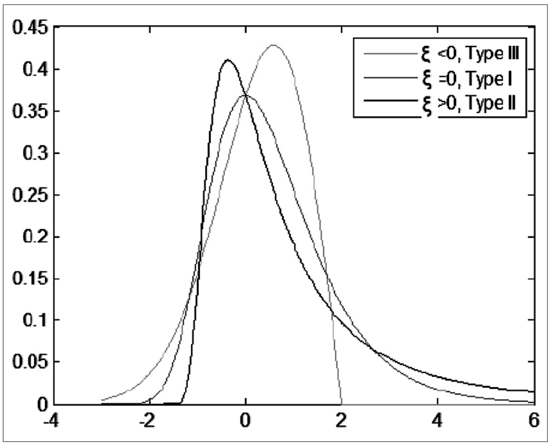

Figure 1. *The profiles of GEV distributions (Gumbel type I, Fréchet type II and Weibull type III Probability density functions).*

$$g\,(z,\,\mu,\,\sigma) = \frac{1}{\sigma}\exp\left\{-\exp\left[-\frac{(z-\mu)}{\sigma}\right]-\frac{(z-\mu)}{\sigma}\right\} \qquad (2)$$

$$G\,(Z) = \exp\left\{-\exp\left[-\frac{(z-\mu)}{\sigma}\right]\right\} \qquad (3)$$

$\mu = E(Z)$ can be estimated by Z mean

$$\sigma = \frac{\mathrm{std}(Z)6^{1/2}}{\pi} \qquad \text{(3A)}$$

Converting Equation 3 to 4 will provide the values of variables by giving different interval time.

$$Z(t) = \mu - \sigma \mathrm{I}n\left(-\ln\left(1-\frac{1}{t}\right)\right) \qquad (4)$$

Where Z(t) is the return level (the return value that you are looking for) and t is the return period (Z is expected to be exceeded once every t days). There is a probability $P= 1-\frac{1}{t}$ that z NOT will be exceeded by the maximum in any particular day.

Results and Discussion

Particle size distribution in raw water during the study period

By assorting the particle counts to the daily average value, the particle size distribution during the study period is clearly demonstrated.

Figure 2 shows the particle counts in raw water in terms of daily average value from October 2009 to July

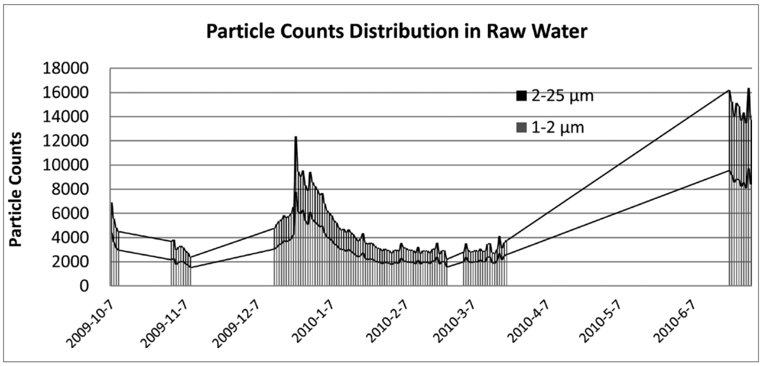

Figure 2. *Particle counts variation in raw water based on daily average value (before peak value appeared in July and August 2010).*

2010 (from Persson et al., 2011). There was no data recorded from April to June due to clogging of the equipment caused by contamination in raw water. High abundance of particle occurred in late July 2010, when the particle fraction size 1–2 µm was exceptionally high. In relative terms, the particle fraction size 2–25 µm was low compared with the values in other months yet still high in absolute terms.

Generally all particle fractions followed the same variation pattern. A small decrease was noted from October to November 2009. From late autumn it rose up to the first peak in early winter following a drop to low values in February. The values started to increase again when temperature rose. The smaller particle fraction is generally 3-fold higher than larger fraction, except for the summer 2010. Practically, the summer values could be explained by algae blooming. The change in particle fractions due to temperature has been reported by others (Scheifhacken et al., 2010), who found that autotrophic pico plankton 0.2–2 µm increased in number during blooming and were present in one order of magnitude higher concentrations than larger plankton. The high growth of algae is also the main reason for the technical problems with the operation of the monitoring system in Borensberg. After installing a proper pre-filtration system in June 2010, the particle counter was operating stable again.

Particle composition and variation in raw water

From a monthly average value of particle counts, one can easily find the variation of constituent percentage of two particle size fractions (Fig. 3) within total particle number. Except for July and August, particle fraction 1–2 µm in general composes 63 % of total counts. During algae growth in July and August, 2010, the small particle dominates totally.

Analysis of particle counts in drinking water

The drinking water treatment in the Borensberg Waterworks reduces the particle content significantly. In Fig. 4, the particle count distribution in drinking water is presented for the study period 2009–2010. The variation in particle count in drinking water differs also from the raw water. Instead of decreasing during autumn and winter, the total count increases and peaks in February and April. Also when the raw water was high in particles due to algae in the summer 2010, the total particle count in drinking water was low and even decreasing suggesting a treatment effect due to temperature. In the late

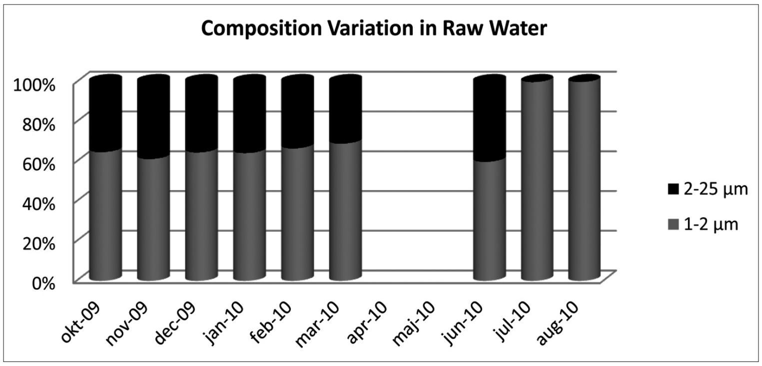

Figure 3. *General composition in raw water during the study period shows majority of the particle counts in faction 1–2 µm with instant high value in July and August.*

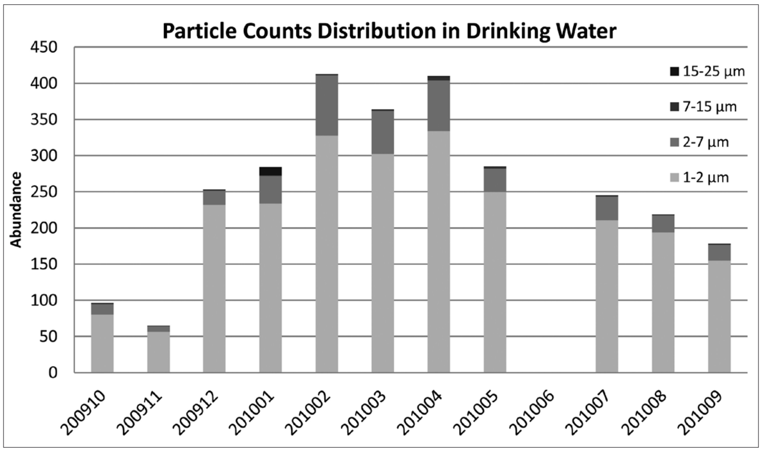

Figure 4. *Particle counts in drinking water based on monthly average value which shows that high abundance of particles occurred from Dec. 2009 to May 2010.*

winter and spring, particle abundance in drinking water still remained at high levels. A trend towards decreasing values can be found for the summer period, see Figure 4. Particle counts in drinking water during algae blooming season (Fig. 3) is undergoing a decreasing phase which shows high efficient treatment process within the waterworks during this season.

Composition variation of four particle size fractions in drinking water

The composition of total particle counts in drinking water and their variation is illustrated in Fig. 5, based on monthly average value. In Figure 5, the relative occurrence of the fractions 1–2 µm, 2–7 µm, 7–15 µm and 15–25 µm is presented. The treatment process reduces larger particles to a higher extent. So in drinking water, the particle fraction of 1–2 µm represents almost 85 % of the total particle counts. Also from the distribution curve, one can see that the particle fraction of 2–25 µm is slightly higher in winter and early spring (February, March and April) and that 2–7 µm dominates.

Comparison between raw water and drinking water

The relation between particle size fractions in raw water and drinking water was studied by analysis the selected time period when both data of particle counts in raw water and drinking water were available. The log-reduction of the number of particles could be calculated and the possible correlation between particle fractions in raw water and drinking water investigated.

Log Reduction can be used as a method for risk assessment of possible microbial and parasitic contamination of drinking water (Chowdhury, 2003). He suggested Log Reduction to be used as a tool to demonstrate the removal of particle number in water treatment process as well as the efficiency of specific waterworks. However, particles are not solely microorganisms.

By using the Log Reduction method, the total reduction of particle counts from raw water to drinking water can be evaluated, see Fig. 6. In general, the Log Reduction is 1.05 while in late July it was 2.82 logs due to higher content of particles due to algae bloom in the raw

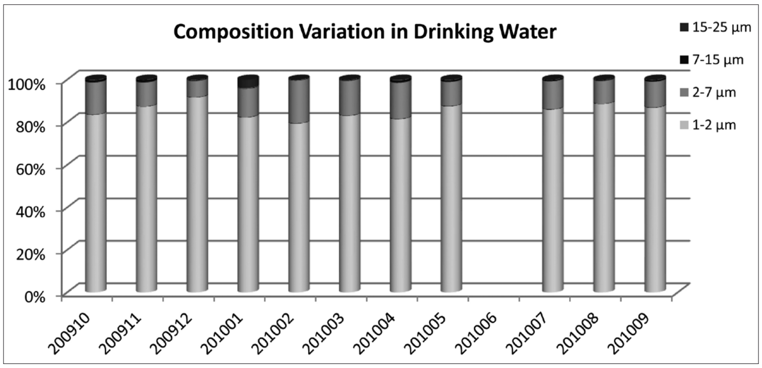

Figure 5. *Composition of particles in drinking water based on monthly average value.*

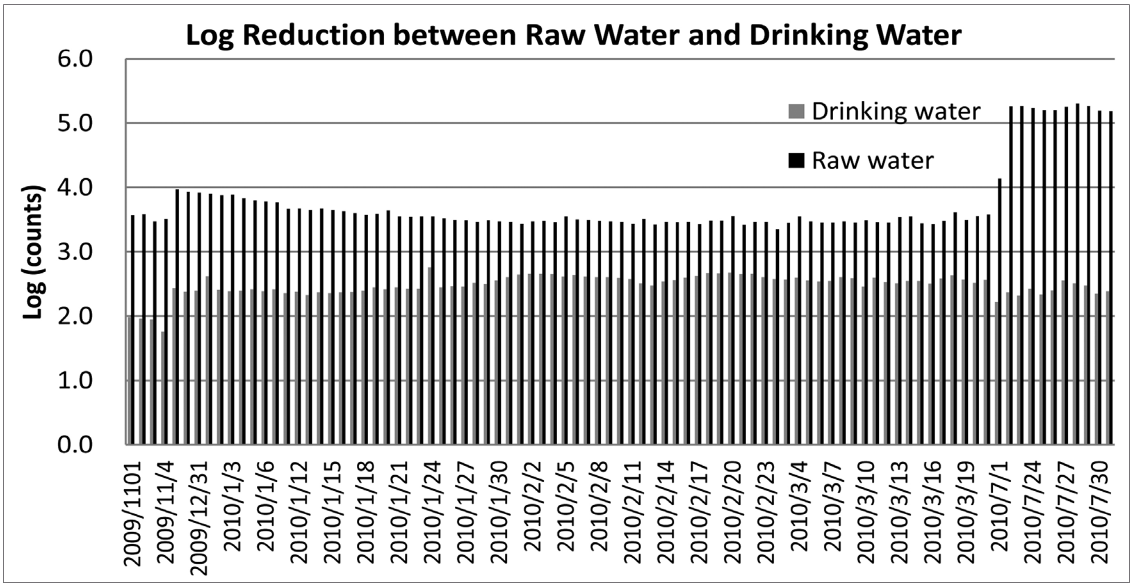

Figure 6. *Log Particle counts during the study period for the Borensberg Waterworks.*

water. Particle count in drinking water varied little during the study period.

Results demonstrate that water treatment within the waterworks reduces the particle count in drinking water, but that the total count also in drinking water is significant. Borensberg's waterworks has a downstream ultraviolet disinfection treatment.

To define a threshold value

In this study, a significant amount of data has been collected. The general idea of using particle size monitoring for on-line control is to better understand the raw water quality variation, the performance of the different treatment steps in the waterworks and to assess whether monitoring can be used for assessing increased risk of microbial contamination in drinking water. For the latter issue, the Generalized Extreme Values (GEV) distribution was introduced to define suitable threshold levels for the plant. The GEV aims to provide a proper method for threshold set-up and to discuss the possibility to apply auto-sampling. In this section, the process of analysis of defining THV for particle fraction 1–2 µm in drinking water is described in details.

The first step is to identify the daily peak values through the whole study period. Parmhat=gevfit(x) in Matlab is applied to define the shape parameter (ξ), scale parameter (σ) and location parameter (μ). For the fraction 1–2 µm in drinking water, there are 208 values in total used for extreme value analysis. The parameters of shape, scale and location are 0.29, 112.44 and 177.66 respectively. Since 0.29 is approaching 0, the Gumbel distribution is applied.

Results for Borensberg of probability Density Curve and Cumulative Frequency Curve are displayed in Fig. 7 and Fig. 8. From a cumulative probability curve (Fig. 8),

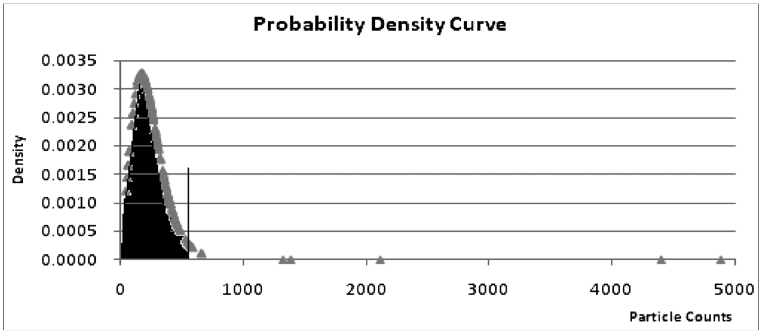

Figure 7. *Gumbel probability distribution of particle fraction 1–2 µm in drinking water, Borensberg Waterworks. Example of possibility of a value less than 500 to occur.*

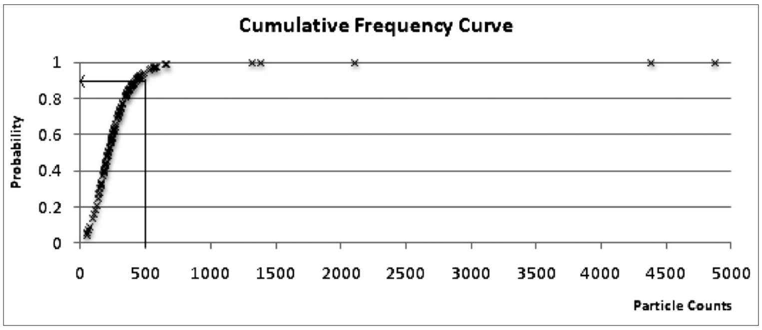

Figure 8. *Cumulative Frequency Curve of particle fraction 1–2 µm in drinking water, Borensberg Waterworks. Example of possibility of a value less than 500 to occur (Cumulative probability).*

we can estimate the probability and the related number of particle counts. Fig. 8 is the result of cumulative probability for any event less than the defined one to happen. The probability for it to occur equals to the size of the dark area shown in Figure 7. As an example, by reading from Fig. 8 the possibility for a value less than 500 to occur is 90 %, which represents the whole area within the density curve (Fig. 7) and the vertical line of the number 500.

Equation 5 is applied to study the relation between particle count and the interval time and the possibility for any particle count to exceed a given value, as well as its frequency. The results from the calculation are illustrated in the graph shown in Fig. 9 for analysis of interval period for a given specific value to present. Fig. 9 is functional to realize the particle abundance recurrence time and its frequency.

Numerically, from the Density Function Curve shown in Fig. 7, the particle counts of 700 is at the +2σ range edge where 98 % of the series is. It means that the chance for a number over 700 to occur is around 2 % which is also the same probability as presented in Fig. 9. Meanwhile, the frequency interval for 700 is 60 days reoccurrence, read from Fig. 9. Those events occurred are regarded as stochastic events in our case.

As a general point of view, the higher the particle concentration, the more potential microbiological pollution risks it is. From Figures 7 and 8, a set-up value of 460 can be assessed as a proper starting point of auto sampling suggestion, because the abundance which is over

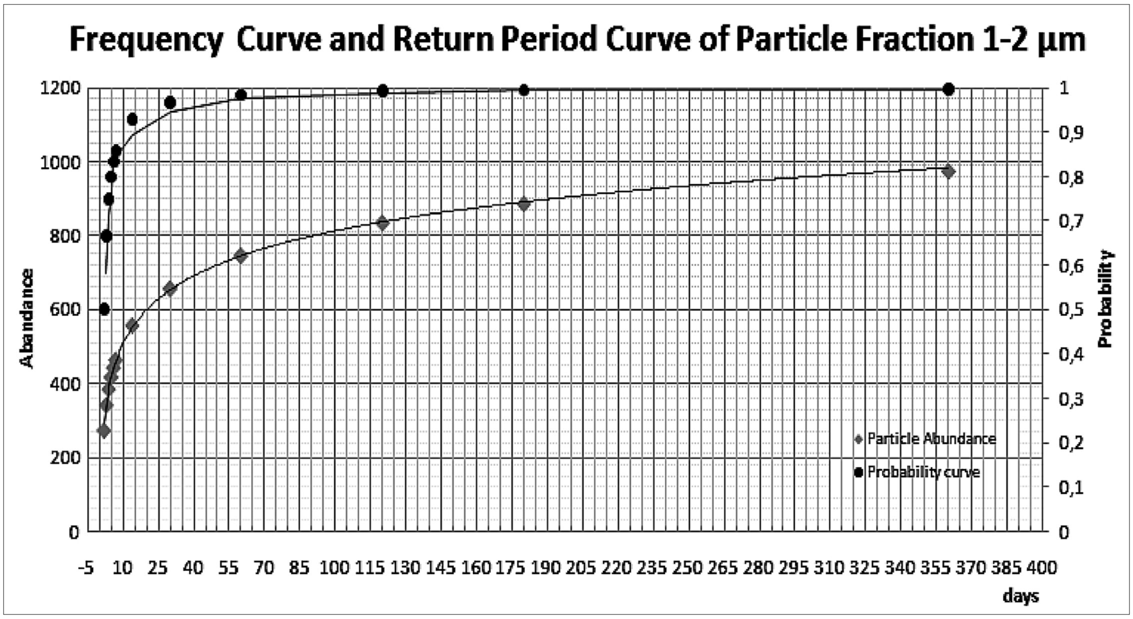

Figure 9. *Frequency curve and return period curve for particle threshold value definition of Fraction 1–2 µm in drinking water, Borensberg Waterworks.*

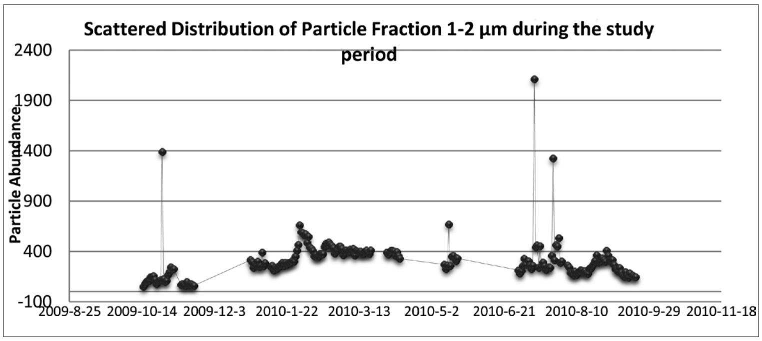

Figure 10. *Scattered distribution of particle fraction 1–2 µm in drinking water, Borensberg Waterworks.*

this amount has 10 % incidence to happen. High risk occasions should be considered for further laboratory analysis. This would probably include the seasonal circulation influence and its possible effect on the microbiological population. According to the scattered particle distribution curve in Fig. 10, it can also be seen that during spring time, the particle count is often higher than 460. Also during summer time, a reminiscence of the algae blooming affected raw water can be noticed in drinking water. So to use 460 as a static level would lead to an increased auto-sampling frequency during these periods or, to phrase it differently, high particle counts should be investigated more thoroughly to improve the understanding of possible microbial contamination.

It may hence be more practical to adjust the threshold during different seasons since particle counts are unevenly distributed over the year. When studying the deviation from the average seasonal value, the suggested boundary level shows that this threshold value covers winter, spring and summer seasons, since the suggested threshold value of 460 is about 77 %, 53 % and 119 % higher than the average value in winter (260), spring

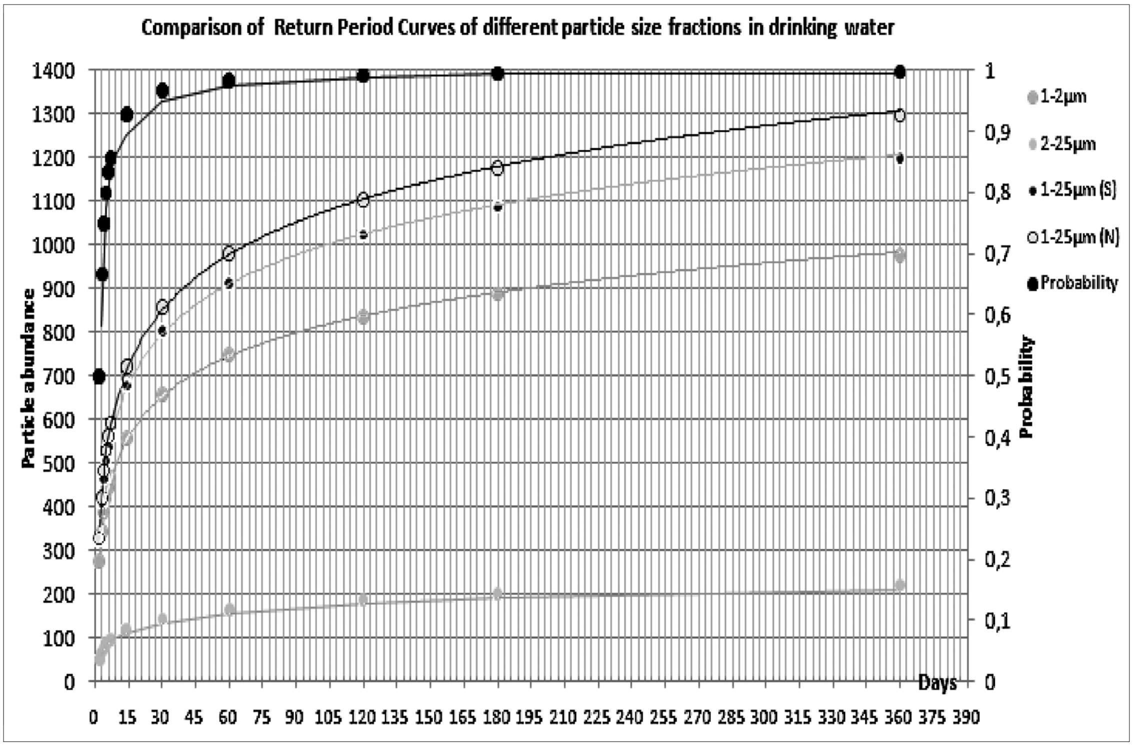

Figure 11. *Comparison of Return Period Curve of different particle size fractions in drinking water, (N) refers to the particle number calculated through numerical method and (S) are the sum of the observed particle size fractions.*

(300) and summer (210) respectively. But for autumn, adjustment is needed. For the autumn, when the average value for the particle is about 70, a threshold value of 460 would be too high and needed to be reduced, maybe down to 250, in order to not omit incidences of possible increased microbial pollution in the drinking water.

As a conclusion of extreme events analysis for particle fraction size 1–2 μm, a comparatively proper value for Borensberg Waterworks is suggested as 460 for a trial test. Adjustment should be done with practical laboratory analysis especially in autumn when there are less particle counts present.

Figure 11 shows the result of the return period and their probability for both two particle size fractions 1–2 μm and 2–25 μm and sum of them. It also shows that there is a deviation of the frequency curve of the summarized counts from modeled one.

Conclusion

The proper choice of a threshold/alarm level of particle counts in water when automatic sampling is done needs to be made after a break-in period for the current water supply. The threshold level may have to be adjusted according to actual water quality for site and local conditions. A proper start-up time is 1 year to gather information for understanding the water quality variation. Particle counts indicate the risk for microbial contamination, but are not equal to it. Particle counts are valuable for understanding the efficiency of treatment processes in a waterworks.

Acknowledgements

The technical and operational help with the work from Motala municipality (Anders Rönnmark, Anders Edholm, Anna-Karin Pålsson and Sven Larsson) as well as Predect AB is acknowledged. The project was financed by Svenskt vatten utveckling, Predect AB, Sweco AB and Motala municipality.

References

Bernier, J.-L. T., Maheux, A.F., Boissinot, M., Picard, F.J., Bissonnette, L., Martin, D., Dewailly, E. and Bergeron, M.G. (2009) Onsite Microbiological Quality Monitoring of Raw Source Water in Cree Community of Mistissini, Water Qual. Res. J. Can., 44 (4): 345–354.

Chowdhury S. (2003) Partikelmätning– ett nytt sätt att kontrollera vattenkvalitet, VA – FORSKRAPPORT Nr 2003-48.

Hatukai S. Ben-Tzur Y. and Rebhun M. (1997) Particle Counts and Size Distribution in System Design for Removal of Turbidity by Granular Deep Bed Filtration, Water Science and Technology, 36 (4): 225–230.

Hoekstra, E.J., Aertgeerts, R, Bonadonna, L., Cortvriend, J., Drury, D., Goossens, R., Jiggins, P., Lucentini,L., Mendel, B., Rasmussen, S., Tsvetanova, Z., Versteegh, A. and Weil, M. (2008) The advice of the Ad-Hoc Working Group on Sampling and Monitoring to the Standing Committee on Drinking Water concerning sampling and monitoring, for the revision of the Council Directive 98/83/EC, 2008. JRC European Commission EUR 23374, Brussels. Available at: http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/1331/1/reqno_jrc45484_final%20report%5B2%5D.pdf

Livsmedelsverket (2001) Vägledning till Livsmedelsverkets föreskrifter (SLVFS 2001:30) om dricksvatten, Sweden.

LeChevallier. M.W. and Norton. W.D. (1992) Examining relationships between particle counts and Giardia, Cryptosporidium and turbidity, Jour. AWWA. 84 (12). 54–62.

Persson, K.M., Chowdhury. S., Rönnholm, A. and Li, J. (2011) Realtidsmätning av vattenkvalitet och automatisk provtagning vid befarad mikrobiell förorening, Svenskt vatten utveckling 2011-06 Available at:
http://www.svensktvatten.se/Documents/Kategorier/SVU/Rapporter/Rapport_2011-06.pdf

Prüss, A., Kay, D., Fewtrell, L., and Bartram, J. (2002) Estimating the burden of disease from water, sanitation, and hygiene at a global level. Environ Health Perspect. 2002 110(5): 537–542.

Scheifhacken, N., Horn, H. and Paul,. L. (2010) Comparing in situ particle monitoring to microscopic counts of plankton in a drinking water reservoir, Water Research, 44(11), 3496–3510

WHO (2010) UN-Water Global Annual Assessment of Sanitation and Drinking-Water (GLAAS), Geneva, Switzerland.