Automatic Chromatogram Analysis

Jani Tomperi[*], **Esko K. Juuso***, **Kauko Leiviskä***

[*]*University of Oulu, Control Engineering Laboratory,*
P.O.Box 4300, FIN-90014 University of Oulu, Finland.
(e-mail: Jani.Tomperi | Esko.Juuso | Kauko.Leiviska @oulu.fi)

Abstract: Chromatography has been found an efficient tool for analysing the water quality, especially natural organic matter (NOM) content, in water treatment processes. Information of the fast and relatively easy chromatography measurements can be used, by itself or together with other process measurements, for process control and improving the water quality. Unfortunately, some chromatography measurement results are available only as graphical chromatograms or in software specific data file and cannot be transmitted to other software for more detailed mathematical analysis, or to the automation system of the water treatment plant. In this paper, the automatic chromatogram analysis method is developed using Matlab. Graphical chromatograms are converted to numerical data for easier and more efficient processing and analysing. The automatic analysis method calculates the significant values (e.g. peak maximums, peak areas) of the chromatograms, and provides an effective and fast way to monitor the water quality.

Keywords: Drinking Water; High-Pressure Size Exclusion Chromatography; Matlab; Natural Organic Matter; Water quality.

1. INTRODUCTION

Water quality consists of biological, chemical and physical characteristics of water. The quality of drinking water is affected by many factors, for example raw water source, water treatment techniques (sedimentation, filtration, disinfection), distribution network material and biofilm, etc. Natural organic matter (NOM) has become a significant issue in drinking water treatment processes since the negative health effects were discovered. NOM causes colour, taste and odour problems to drinking water, leads to the formation of disinfection by-products and increase chlorine demand, and promotes bacterial re-growth. The NOM affects also the water treatment operations: coagulant and disinfectants are dosed based on the amount of NOM, it blocks active carbon pores and is one of the major membrane foulants. The amount of NOM depends on climate, water source and season of the year. (Teixeira and Nunes 2011, Vuorio et al. 1998)

The most important water quality indicators pH, colour, taste and odour, dissolved metals and organics, salts and microorganisms, are measured frequently from a water treatment process. With basic measurements the general condition of the water treatment process and water quality can be achieved, but using chromatography techniques have been found to give more detailed information. Chromatography technique can be used to analyse the qualitative changes in the water treatment process. Changes in molecular size fractions can be measured from raw water, after different purification steps and from drinking water. In addition to analysing the quality of the water treatment process, chromatography technique can be used to assist the process control.

Chromatography is an analytical technique which can be used for separating chemical substances into its components for analysis or identification. There are e.g. paper, liquid, gas and ion-exchange chromatography techniques which all function in the same way: mobile phase (a liquid or a gas) flows through the stationary phase (solid, or liquid supported on solid) and carries the components of the mixture with it. Different components travel at different rates and the time that a compound travels to the output of the chromatography column is called a retention time. Different components of the analysed mixture are shown as separate peaks or patterns on the visual output, the chromatogram. Concentrations of substance components are measured as the peak area and peak height.

Fast, easy and practical High-Pressure Size Exclusion Chromatography (HPSEC) has been found to be a highly effective and precise analytical method of evaluating the molecular size and molecular weight (MW) distributions of dissolved humic substances in drinking water (Conte and Piccolo 1999, Matilainen et al. 2002, Teixeira and Nunes 2011, Zhou et al. 2000). The HPSEC method provides useful information about the quality of water and the amount of NOM. Correlation among HPSEC and traditional analytic methods like TOC and UV absorbance have been found but in comparison with traditional

analytical methods, HPSEC yields additional information on the removal and transformation of NOM.

In HPSEC analysis, humic molecules are separated on the basis of the different molecular size so that the largest molecules elute first and the smallest molecules elute last in the column (Matilainen et al. 2002). The amount of NOM in one molecular size fraction can be calculated as the height of the peak or the peak area. The peak area is less sensitive to the influence of peak dispersion mechanisms and thus is more reliable and corresponds better to the quantity of the component. Due to dispersion, the peaks may become shorter, broader and more asymmetrical but the total area under the peak remains nearly equal. Sometimes in chromatographic analysis the area under the individual peak curve cannot be resolved. The width of the base, wide of the half of the peak and the height of the peak are also often studied indicators. The total amount of the NOM in a measured water sample is calculated as the sum of all peak areas or peak heights.

Usually from five to seven, even eight different peaks (molecular size fractions) can be found in the chromatography measurement of the Finnish surface water sample. Peaks can be separated to high molecular weight (HMW), which has the lowest retention time, intermediate molecular weight (IMW) and low molecular weight (LMW), which has the highest retention time. Raw water consists mostly of large molecules and generally, high molecular weight matter is easier to remove than low molecular weight matter. (Matilainen et al. 2002, Vuorio et al. 1998)

However, usually chromatography measurement results are stored in software specific data and cannot be transferred or used in other software or systems e.g. Matlab or in the automation system of the water treatment plant. In the worst case only graphical chromatograms are available. In this study, the automatic chromatogram analysis method which converts images to numerical data and analyses the data is created. HPSEC measurements of water samples taken from one Finnish water treatment plant are used to demonstrate the functionality of the developed method. Water samples are collected from several stages of the water treatment process, from raw water to the distribution system. The automatic chromatography analysis method is programmed in Matlab and it consists of image processing of graphical chromatograms and analysis of the resulting data.

2. IMAGE DIGITIZATION

The image digitization is done using familiar methods and Matlab functions. In the first step of image processing all images containing the graphical chromatograms are loaded automatically from a specific folder using Matlab function 'imread'. Image data is converted to class 'double' and converted to gray-scale intensity image by eliminating the hue and saturation information while retaining the luminance. Same processing is done to the background image which is subtracted from the original image. After subtraction all values in image data greater than 1 are converted to value 1 and all values lower than 0 are converted to 0 to ensure the binary black and white image. Pixels with value 0 are displayed as black and pixels with value 1 are displayed as white. In the final step, the contrast of the image is adjusted using 'imadjust'-function and enhanced using 'adapthisteq'-function, and the intensity of the image is converted to double precision. (Matlab 2012)

After removing the background and processing the original image, numerical data of the image is stored to a mat-file where every pixel has its specific value, in this case between 0 and 1. The spectrum can now be found by reading the pixel values. After image processing some sporadic values of the spectrum may be missing, but they can be replaced by linear interpolation. Y-values of the spectrum are scaled to real values by reading the minimum and maximum values of the original spectrum from a separate file. For relative comparison, scaling can be also done, for instance, between 0 and 1. The main steps of image digitization are shown in Figure 1.

Locations of studied spikes are case specific and determined beforehand. The time/pixel -relation is used to select the specific spike for area and height calculation and analysis. The area of the spike is calculated using Matlab's 'trapz' integration method, which computes an approximation of the integral of the spectrum via the trapezoidal method. In the analysis, a single spike can be studied and the height of peak(s) or the total area of all peaks can be calculated. Chromatogram can also be divided into, for example, three molecular size fraction areas, which are studied.

3. RESULTS AND DISCUSSION

In this study water samples are taken from eight different process stages of a Finnish water treatment plant and measured using the HPSEC method. Block diagram of the water treatment plant is shown in Figure 2. Water samples are taken from raw water (#1), after adding flocculation chemical (#2), at slow mixing stage (#3), after clarification by flotation (#4), after active carbon filtration (#5), after disinfection using UV radiation (#6) and chlorine (#7) and from the distribution system (#8). Eight peaks are found and analysed in every measured spectrum.
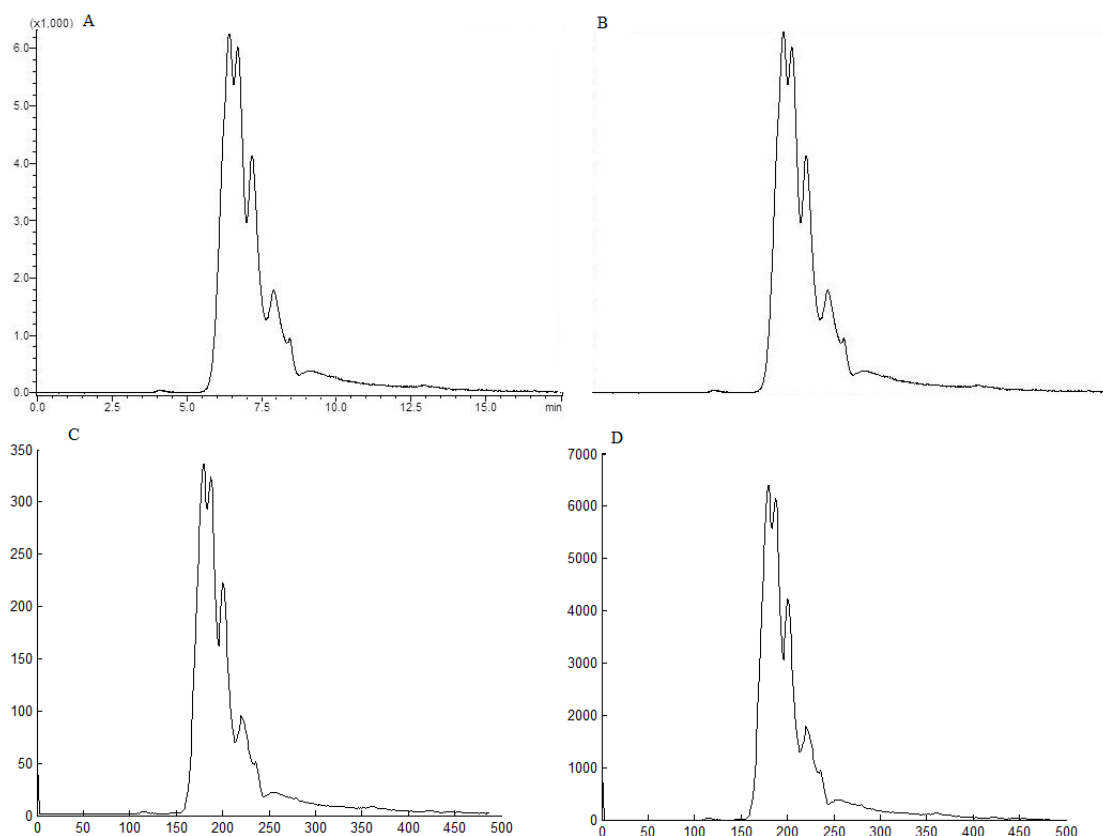
Figure 1. Digitization of a chromatograph: (A) The original chromatograph, (B) Chromatograph after the background is removed, (C) Image has been processed, stored to mat-file and after interpolation plotted, (D) Spectrum values have been scaled to real values.
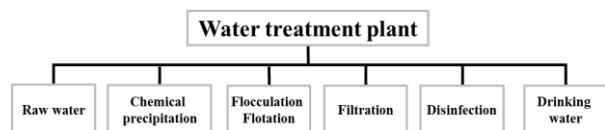

Figure 2. Block diagram of the Finnish water treatment plant.

The automatic chromatogram analysis can process hundreds of images in relatively short time. All results can be listed in numerical form or plotted to various figures. Analysis can be done by comparing different samples at the same process stage, samples from different (or consecutive) process stages, separate peaks in one spectrum etc. Examples of calculated peak heights and areas of raw water and drinking water samples in the numerical form are shown in Table 1. Eight different molecular size fractions are listed from high (#1) to low (#8). Using numerical data, the exact comparison between samples, process stages or peaks is more easy and precise than comparing several plotted spectra.

Table 1. Height and area of peaks analysed from raw water and drinking water samples.

Peak	Raw water		Drinking water	
	Peak height	Peak area	Peak height	Peak area
1	6400	129121	1051	12686
2	4226	54957	1600	17798
3	1798	22394	831	9392
4	947	5850	349	1822
5	382	18081	119	5502
6	134	2983	1072	10975
7	134	2706	57	1232
8	76	1210	81	1327

For visual examination, different figures can be plotted. In Figure 3 peaks are separated to HMW, IMW and LMW organic matter areas. Principles of separation can differ, but in this case the first two peaks are selected as HMW, the next three as IMW and last three as LMW. Figure shows that raw water (process stage #1) consists mostly of large molecules which are removed efficiently, but low molecular weight matter is harder to remove.
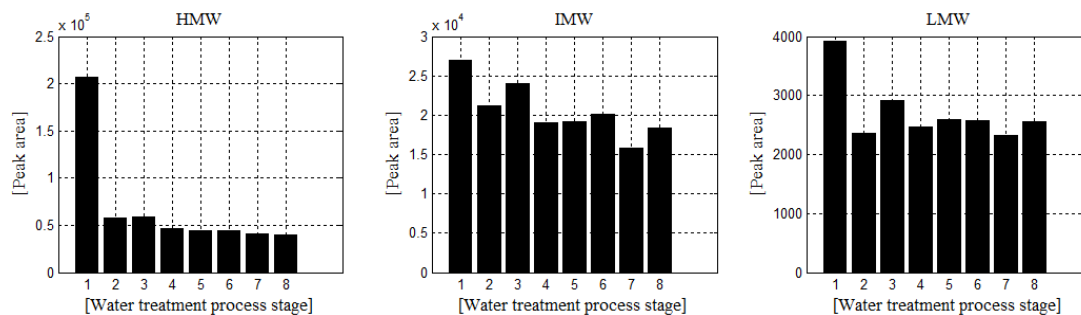
Figure 3. Molecular size fractions are examined as high, intermediate and low molecular weight areas.

In Figure 4, total areas of all peaks in every stage of the water treatment process are presented as bars. Total peak area is the highest in raw water (#1) and decreases towards the sample taken from the distribution system (#8). Thus the water treatment process is efficient and the amount of NOM diminishes.
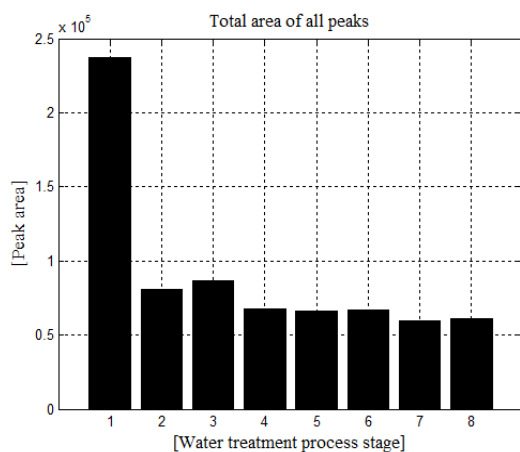

Figure 4. Total peak area of the water samples.

4. CONCLUSIONS

In this paper, a method for digitizing graphical chromatograms describing the operation of the water treatment plant has been developed. Methods familiar from image processing have been used. The digitized data is used in calculating peak heights and areas in the chromatogram that correspond the molecular size distribution of the flow in question.

The paper shows an example, how to use the calculated values in analysing how the water purification proceed in different stages of the plant and what happens to the water quality and NOM in different stages of water treatment. The example comes from one Finnish drinking water purification plant. In this case, NOM is divided into low, intermediate and high molecular weight areas and changes in these are followed through the eight stages of the plant. The results are logical compared with the practical experience.

In the future, the calculations of essential indicators, inter alia the peak area, are improved and the automatic chromatogram analysis method is used for processing and analysing multiple chromatography measurement series from Finnish water treatment processes. Results are compared to each other and with other process measurements to analyse the water quality and functionality of the water treatment process.

REFERENCES

Conte P. and Piccolo A. (1999). High Pressure Size Exclusion Chromatography (HPSEC) of humic substances: molecular sizes, analytical parameters, and column performance. *Chemosphere*, Vol. 38, No 3, pp. 517-528.

Matilainen A., Lindqvist N., Korhonen S. and Tuhkanen T. (2002). Removal of NOM in the diffrent stages of the water treatment process. *Environment International*, Vol. 28, pp. 457 – 465.

MATLAB. (2012). Image Processing Toolbox User's Guide. *Matlab MathWorks Inc*. web document.

Teixeira M.R. and Nunes L.M. (2011). The impact of natural organic matter seasonal variations in drinking water quality. *Desalination and Water Treatment*, Vol 36, pp. 344 – 353.

Vuorio E., Vahala R., Rintala J. and Laukkanen R. (1998). The evaluation of drinking water treatment performed with HPSEC. *Environment International*, Vol. 24, No 5/6, pp. 617 – 623.

Zhou Q., Cabaniss S. E. and Maurice P.A. (2000). Considerations in the use of High-Pressure Size Exclusion Chromatography (HPSEC) for determining molecular weights of aquatic humic substances. *Water Research*, Vol. 34, No 14, pp. 3505-3514.