Data clustering analysis in the assessment of wastes using in the sewage filtration

Dawid BEDLA1) ABDE, Ewa DACEWICZ2) BCDF

1) orcid.org/0000-0003-0500-144; University of Agriculture in Krakow, Faculty of Environmental Engineering and Land Surveying, al. Mickiewicza 24/28, 30-059, Cracow, Poland; e-mail: d.bedla@ur.krakow.pl
2) orcid.org/0000-0003-2167-8091; University of Agriculture in Krakow, Faculty of Environmental Engineering and Land Surveying, Poland; e-mail: ewa.wasik@ur.edu.pl

For citation: Bedla D., Dacewicz E. 2019. Data clustering analysis in the assessment of wastes using in the sewage filtration. Journal of Water and Land Development. No. 41 (IV–VI) p. 31–36. DOI: 10.2478/jwld-2019-0024.

Abstract

The paper discusses the use of multiclustering statistical analysis in the assessment of domestic wastewater filtration effectiveness. Calculations included data collected over four months of experiments with using waste as filling material of vertical flow filters for domestic sewage treatment. The effectiveness of pollutants removal was analysed in case of mechanically shredded waste in the form of PET flakes, PUR foam trims, shredded rubber tires and wadding. The organic compounds (CODcr, BODs) removal, suspend solids, biogens (as NH4+, PO43– ions) and oxygen saturation changing compared with sand filling was analysed. Multiclustering statistical analysis allowed to divide pollutants removal efficiency of analysed materials into 3 clusters, depending on the hydraulic loading. The first group consisted in quality parameters of treated sewage: the highest reduction of BOD5 and NH4-N. It included the values of quality parameters and indicators for the filtrates obtained at the lowest hydraulic load from columns filled with 60 cm of rubber tires or sand. The second group comprised the results for fillings containing foam, PET and rubber tires (the other hydraulic loads). It featured the highest reduction of total suspended solids and PO43–. Removal of easily biodegradable organic compounds was at a similar level in both cluster groups. The filter filled with polyester waste (wadding), which was as effective as 30 cm layer of sand, and the filters filled with 60 cm of sand working at the highest hydraulic load. Third group showed the lowest values of parameters and indicators for analysed filtrates.

Key words: cluster analysis, PET, polyester, polyurethane foam, rubber tires, sewage filtration

INTRODUCTION

Circular economy aims at treating the waste as recyclables that may be processed and used again. Quality of recycling, which depends on the methods of waste collecting and sorting, also plays an important role. Limited waste production and saving resources help to improve energy efficiency of the resource’s life cycle and to reduce atmosphere emissions of carbon dioxide. Researchers from Denkstatt GmbH in Austria found out that producing PET bottles from a regranulate instead of an original polymer reduced CO2 emissions by 80% [EURACTIV 2017]. Materials obtained in this way may also be used in the manufacture of fibers, fleece fabrics, food packaging or automotive foams [BEDELL et al. 2018] or as elements of thermal insulation [GUTIÉRREZ-GONZÁLEZ et al. 2012]. Shredded or chipped tires waste could be used in new rubber production [AOUDIA et al. 2017] and also added into rubber mixtures for thermal insulations or drainage layers [New Hampshire Department of Environmental Services 2011].

In the processes of wastewater treatment more popular become methods based on solid waste materials. An interesting research was conducted by MANGESH and CHARPE [2015]. In laboratory scale multimedia filter model consisting of three reactors placed in series was developed for treatment of domestic wastewater. In case their study different types of material was testing. One of those reactors was packed with media size 16–20 mm (plastic scrubbers).
During all processes of wastewater treatment the removal efficiency for BOD was on the level of 70%, for COD was 62% and for total suspended solids (TS) was 87% for 24 hours of detention time was observed. The multimedia filter process also gave best results for TS removal.

During the process of organizing the source describing the proper classification of wastewater treatment plays an important role. The linear methods (like regression) [KACZOR et al. 2017], as well as methods based on classification trees [DACEWICZ, CHMIELOWSKI in press], neural networks [WASIK et al. 2018] or data clustering could be used [DACEWICZ 2018]. Hierarchical or non-hierarchical methods can be used in data clustering analysis. Using the hierarchical method, agglomerates are determined by linking obtained from the earlier steps of the algorithm minor clusters [StatSoft 2018]. The results are a dendrogram which illustrate the structure of the source data and enable preliminary assessment of the number of clusters. This number could be used in the non-hierarchical analysis by the k-means method. The parameters of this agglomeration procedure include objects from which the algorithm will start to construct the clusters (starting centers), and the criterion which determines the way to stop the algorithm (stop criterion) [StatSoft 2018]. The algorithm of the k-means method of data clustering is based on the Euclidean distances matrix between objects and is corrected while performing the agglomeration procedure by transferring data between them.

This paper analyzes the possibility of using mechanically shredded waste, i.e. flakes of PET, elastic polyurethane foams, shredded rubber tires and wadding (polyethylene) for purification of domestic sewage. The efficiency of vertical flow filters in removing pollutants such as organic biogenic compounds and suspended solids using a multidimensional statistical analysis, i.e. principal component analysis and cluster analysis was assessed.

STUDY METHODS

Semi-technical models with identical PVC columns were built, to assess the performance of vertical flow filters filled with various types of waste materials (Photo 1). The model consisted of six identical PVC columns, 150 mm in diameter and 1000 mm high. The filling materials included four types of waste: fragmented polyurethane foam sponge (column 1), wadding from polyester foam (column 3), shredded PET flakes (column 4) and shredded rubber tires (column 5). As the reference columns with 60 cm sand (column 2) and 30 cm sand (column 6) was used.

The pretreated sewage collected in septic tank was analyzed. This sewage supplied the six columns at hourly interval per day. The three various hydraulic load was considered: 31.8, 76.4 and 114.6 mm$\cdot$day$^{-1}$ per measurement series. The physical and chemical analysis of sewage samples was carried out. The following indicators of pollution were measured: pH, temperature, oxygen saturation, concentrations of organic compounds (BOD$_5$ and COD$_{Cr}$), ammonium nitrogen, nitrates and orthophosphates, as well as total organic solids. Physical and chemical tests were performed for a mean of 10 measurement series. Samples of effluents was collected from the each columns at the intervals of 24 h.

Statistical analyses were conducted using Statistica 12.5 package. Statistical analysis of the filtrate quality from individual columns included few stages. The first stage provided basic descriptive statistics such as: an average value, minimum and maximum values, median and coefficient of variation. At the next step the types of support materials used for domestic wastewater treatment were grouped. Then we performed a non-parametric analysis of variance to reduced indicators describing sewage collected from individual model columns. The non-parametric Kruskal–Wallis test for variables lacking normal distribution and showing heterogeneity of variance was used to indicate the most significant rank differences between the groups of columns. A difference in means for a given pair of groups was considered significant when the probability value ($p$-value) was lower than 0.05. As interpretation of the results was not straightforward, performed multidimensional cluster analysis was made. It this cases a simple hierarchical method was adopted in which agglomerates were determined by linking minor clusters obtained from the earlier steps of the algorithm [StatSoft 2018]. Ward’s method was assumed as agglomeration method and the Euclidean distance was the distance measure. In the final step the correctness of classification into clusters on the basis of the agglomeration method, and regrouped the columns using the k-means method was verified.
RESULTS AND DISCUSSION

Figure 1 presents granulometric curves for shredded waste, i.e. polyethylene terephthalate flakes, flexible strips of polyurethane foam and shredded rubber tires. Because of its shape, the wadding (polyester) could not be subjected to the process of determining the granulation.

In the pretreated sewage high variability of oxygen saturation (CV = 1.16) was observed. In case of BOD₅ the value ranged from 50.5 mg O₂·dm⁻³ to 500.0 mg O₂·dm⁻³, with an average value reaching 172.6 mg O₂·dm⁻³. These values of BOD₅ are similar to presented by JÓŹWIAKOWSKI [2017a]. The average CODCr values for sewage from septic tank reached 280.9 mg O₂·dm⁻³. These results indicated a significant quality variation of supplied sewage, with a spread on the level 450.0 mg O₂·dm⁻³. Besides, the range concentration for suspended solids fluctuated from 29.7 mg·dm⁻³ to 243.7 mg·dm⁻³, with the average value on the level 92.0 mg·dm⁻³, which was low in compared with raw sewage. The sewage supplied the experimental model from the storage tank showed a low variation of biogenic indicators. The values achieved for NH₄-N ranged from 118.6 to 210.3 mg·dm⁻³, with an average value of 160.1 mg·dm⁻³ (Tab. 1). Achieved values of NH₄-N are at least two times higher than presented by JÓŹWIAKOWSKI [2017b]. The concentrations of PO₄³⁻ fell between 27.3 mg·dm⁻³ and 61.2 mg O₂·dm⁻³, with an average value of 43.3 mg O₂·dm⁻³.

Table 2. Quality characteristics of the treated sewage (effluents from columns 1–6) – average values

| Parameter/indicator | Value in column 8 | Value in column 4 | Value in column 5 | Value in column 6 |
|---------------------|-------------------|-------------------|-------------------|-------------------|
| pH                  | 7.40              | 7.81              | 7.29              | 7.54              |
| Oxygen saturation (%) | 41.50             | 54.10             | 37.89             | 65.46             |
| CODCr (mg O₂·dm⁻³)  | 121.86            | 103.93            | 111.89            | 236.71            |
| BOD₅ (mg O₂·dm⁻³)   | 16.25             | 9.00              | 28.50             | 66.85             |
| Suspended solids (mg·dm⁻³) | 15.99         | 12.13             | 16.54             | 44.49             |
| NH₄-N (mg·dm⁻³)     | 59.54             | 25.34             | 87.79             | 83.33             |
| PO₄³⁻ (mg·dm⁻³)     | 38.21             | 2.69              | 24.57             | 13.43             |
| Hydraulic load 31.8 mm·d⁻¹ |                 |                   |                   |                   |
| pH                  | 7.20              | 7.95              | 7.85              | 7.76              |
| Oxygen saturation (%) | 34.04             | 46.81             | 24.39             | 58.34             |
| CODCr (mg O₂·dm⁻³)  | 97.33             | 93.88             | 83.44             | 57.86             |
| BOD₅ (mg O₂·dm⁻³)   | 17.00             | 39.90             | 88.00             | 39.60             |
| Suspended solids (mg·dm⁻³) | 21.50          | 15.96             | 11.72             | 32.98             |
| NH₄-N (mg·dm⁻³)     | 45.38             | 46.40             | 109.79            | 50.42             |
| PO₄³⁻ (mg·dm⁻³)     | 46.17             | 43.33             | 48.89             | 33.96             |
| Hydraulic load 76.4 mm·d⁻¹ |                 |                   |                   |                   |
| pH                  | 7.89              | 7.24              | 7.79              | 7.91              |
| Oxygen saturation (%) | 21.93             | 41.56             | 24.47             | 56.17             |
| CODCr (mg O₂·dm⁻³)  | 94.67             | 86.43             | 66.67             | 82.78             |
| BOD₅ (mg O₂·dm⁻³)   | 35.00             | 59.20             | 83.80             | 14.00             |
| Suspended solids (mg·dm⁻³) | 17.10          | 12.47             | 23.15             | 15.75             |
| NH₄-N (mg·dm⁻³)     | 101.07            | 71.02             | 109.13            | 92.98             |
| PO₄³⁻ (mg·dm⁻³)     | 32.33             | 36.57             | 38.07             | 40.53             |

Table 1. Quality characteristics of the pretreated sewage supplied to the model columns

| Indicator | Descriptive statistics (number of samples = 30) |
|----------|-----------------------------------------------|
|          | average value | median | minimum value | maximum value | coefficient of variation |
| pH       | 7.88          | 7.87   | 7.20          | 8.46          | 0.04                     |
| Oxygen saturation (%) | 1.03          | 0.67   | 0.03          | 4.67          | 1.16                     |
| CODCr (mg O₂·dm⁻³) | 280.9         | 223.3  | 116.0         | 773.3         | 0.58                     |
| BOD₅ (mg O₂·dm⁻³)  | 172.6         | 164.2  | 50.5          | 500.0         | 0.63                     |
| Suspended solids (mg·dm⁻³) | 92.0          | 74.4   | 29.7          | 243.7         | 0.62                     |
| NH₄-N (mg·dm⁻³)     | 160.1         | 155.5  | 118.6         | 210.3         | 0.15                     |
| PO₄³⁻ (mg·dm⁻³)     | 43.3          | 44.0   | 27.3          | 61.2          | 0.20                     |

Source: own study.

In the pretreated sewage high variability of oxygen saturation (CV = 1.16) was observed. In case of BOD₅ the value ranged from 50.5 mg O₂·dm⁻³ to 500.0 mg O₂·dm⁻³, with an average value reaching 172.6 mg O₂·dm⁻³. These values of BOD₅ are similar to presented by JÓŹWIAKOWSKI [2017a]. The average CODCr values for sewage from septic tank reached 280.9 mg O₂·dm⁻³. These results indicated a significant quality variation of supplied sewage, with a spread on the level 450.0 mg O₂·dm⁻³. Besides, the range...
The analysis of variance was carried out to determine whether the type of waste used as a filler significantly differentiated the values of oxygen saturation, reduction of BOD₅, suspended solids, ammonium nitrogen and orthophosphates in the treated sewage collected from individual columns of the model. Outcomes of the non-parametric Kruskal–Wallis test (Tab. 3) confirmed considerable effects on the filling type on the reduction of the above-mentioned pollution indicators over four months of the study.

Table 3. Kruskal–Wallis test results based on pollutant elimination rate

| Pollution indicator elimination          | Results of Kruskal–Wallis test¹ |
|------------------------------------------|----------------------------------|
| Oxygen saturation (%)                    | 76.7390; p = 0.0000 [1 and 2, 3, 5, 2, 3 and 4; 4 and 5, 6] |
| BOD₅ reduction (%)                       | 18.0572; p = 0.0029 [2 and 3] |
| Suspended solids reduction (%)           | 33.9873; p = 0.0000 [1 and 5; 2, 3 and 4, 5; 5 and 6] |
| Ammonium nitrogen reduction (%)          | 44.11887; p = 0.0000 [1, 2 and 3, 3 and 4, 5, 6] |
| Orthophosphate reduction (%)             | 73.2049; p = 0.0000 [1 and 2, 3, 5; 2 and 4, 6; 3 and 4; 4 and 5] |

¹ Column pairs listed in parenthesis show significant differences (fulfil of columns: 1 foam, 2 sand 60 cm, 3 polyester, 4 PET, 5 rubber, 6 sand 30 cm).

Source: own study.

In analysed columns the least significant differences for elimination of easily biodegradable substances denoted as BOD₅ was found. It could be a result of created unfavourable conditions of filling material for colonization by heterotrophic bacteria.

Interpretation of the resulting column pairs did not provide a definite way of their grouping. For this reason the next step of investigation process was cluster analysis. Ward’s agglomeration method was used and the Euclidean distance measure was considered.

Results of Ward’s cluster analysis allowed to distinguish three groups of cases (clusters) that differed for all investigated columns over four months of the study (Fig. 2). The linkage distance was assessed to 700. From this point the rest of result interpretations were made based on this three clusters.

Figure 3 shows the average values for each of three clusters (groups) representing the value level for each parameter (filtrate quality indicator).

The first cluster consisted of cases of purified sewage with the highest oxygen saturation (an average value 63%) and reduction of BOD₅ and NH₄-N (an average value 72% and 93% respectively). It included the filtrate quality parameters obtained at the lowest hydraulic loads (hydraulic load 31.8 and 76.4 mm·d⁻¹) for the column filled with 60 cm of sand and for flow rate of hydraulic load 31.8 mm·d⁻¹ for rubber filled column. Crushed rubber filling containing large fractions ineffectively eliminated suspended solids (an average value 71%). The same high removal value of suspended solids (an average value 80%) obtained DACEWICZ and CHMIEŁOWSKI [in press]. The filtrates combined into the second cluster represented by the fillings made of foam, PET and rubber (the other hydraulic loads 76.4 mm·d⁻¹ and 114.6 mm·d⁻¹) featured the highest PO₄³⁻ reduction (an average value 45%). Ammonium nitrogen removal at the level 66% was observed, what was at similar level (67%) in compare with PET flakes using as a filters material [WAŚIK, CHMIEŁOWSKI 2017]. CHMIEŁOWSKI et al. [2018] indicated that NH₄-N ions were reduced in bio-filters filled with polyurethane foam and PET at a similar level (up 64% and up 69% respectively). Removal of easily biodegradable organic compounds was on comparable level as in the cluster 1 and was an average 90%. The obtained results was in the similar level like in case of column with PET flakes fulfil, and slightly highest than in research presented by WAŚIK and CHMIEŁOWSKI [2017]. Moreover CHMIEŁOWSKI et al. [2018] indicates, that the use of polyurethane foam allowed to remove BOD₅ by an average up 95%. MANGESH and CHARPE [2015] investigated the plastic scrubbers in the process of wastewater treatment and found the removal efficiency for BOD on the level of 70%, and for COD on the level 62%.
The third cluster included parameters describing the filtrates coming from the column filled with polyester and sand (30 cm), beside the filtrates obtained at the greatest hydraulic load 114.6 mm·d⁻¹ for 60 cm sand column. At the average value of oxygen saturation 47%, these columns removed easily biodegradable organic compounds on the average level 50%, and biogenic compounds in the case of NH₄-N, on the level 42%, and 31% for PO₄³⁻. SPYCHAŁA and LUCYK [2015] obtained satisfying nutrients removal from wastewater by heterotrophic bacteria inhabiting synthetic non-woven filters from polyester of various thickness. The average effectiveness of removing of organic compounds, designated as COD ranged from 48 to 64%.

The variance analysis for the three considered clusters (Tab. 4) indicated that statistically significant differences at a significance level of 0.05 were not observed only in the case of suspended solids reduction.

The variance analysis for the three considered clusters (Tab. 4) indicated that statistically significant differences at a significance level of 0.05 were not observed only in the case of suspended solids reduction.

### Table 4. Analysis of variance: 3 clusters

| Variable                        | Analysis of variance for 3 clusters | F | significance p |
|---------------------------------|-------------------------------------|---|----------------|
| Oxygen saturation               | 59.8882                             | 0.000000 |
| BOD₅ reduction                  | 101.4622                            | 0.000000 |
| Suspended solids reduction      | 1.8621                              | 0.161701 |
| NH₄-N reduction                 | 24.8268                             | 0.000000 |
| PO₄³⁻ reduction                 | 59.8215                             | 0.000000 |

Source: own study.

### CONCLUSIONS

1. In the processes of domestic sewage treatment shredded wastes in the form of PET flakes, trims of polyurethane foam, rubber tires and polyester (wadding) could be an alternative filling material for sand.

2. Multiclustering statistical analysis is a useful tool for classification of various waste materials and their efficient during sewage filtration process and should be considered as a support to reduction numbers of indicators describing process of domestic sewage treatment.

3. The first cluster grouped sewage quality parameters like: reduction of BOD₅ and NH₄-N. In the second cluster the best removal of total suspended solids and PO₄³⁻ it was found. Removal of easily biodegradable organic compounds was at a similar level in two received cluster groups (clusters 1 and 2). Third cluster showed the lowest level efficiency of pollutants removal for analysed filtrates. These results demonstrated some differences in the effectiveness of the pollutants removal between analysed waste materials.

4. Further studies to test waste materials in different combinations should be considered, especially in case of vertical flow filters without additional aeration.

### ACKNOWLEDGEMENT

This research was financed by the Ministry of Science and Higher Education of the Republic of Poland (DS-3337/KEKiOP).

### REFERENCES

Aoudia K., Azem S., Hocine N. A., Gratton M., Pettarin V., Seghar S. 2017. Recycling of waste tire rubber: Microwave devulcanization and incorporation in a thermoset resin. Waste Management. Vol. 60 p. 471–481.

Bedelli M., Brown M., Kiziltas A., Mielewski D., Mukerjee S., Tabor R. 2018. A case for closed-loop recycling of post-consumer PET for automotive foams. Waste Management. Vol. 71 p. 97–108.

Chmielewski K., Dacewicz E., Bedla D., Mazur R. 2018. Zastosowanie odpadów z tworzyw sztucznych w biofiltrach do oczyszczania ścieków bytowych [The use of plastic waste in biofilters for domestic sewage treatment]. Przemysł Chemiczny, Nr 9 p. 1456–1459.

Dacewicz E. 2018. Application of selective and porous materials for the removal of biogenic compounds and indicator bacteria from domestic wastewater, Acta Scientiarum Polonicae Formatio Circumiectus. Vol. 17. Iss. 2 p. 47–55.

Dacewicz E., Chmielewski K. in press. Decision support systems in the assessment of wastes using in the domestic sewage treatment process. Separation and Purification Technology [submitted in 2019].

EURACTIV 2017. New recycling levels [online]. [Access 10.10.2018]. Available at: https://www.euractiv.pl/section/energia-i-srodowisko/news/nowe-poziomy-recyklingu-zaakceptowane-unie-europejska-circular-economy

Gutiérrez-González S., Gadea J., Rodríguez A., Junco C., Calderón V. 2012. Lightweight plastic materials with enhanced thermal properties made with polyurethane foam wastes. Construction and Building Materials. Vol. 28. Iss. 1 p. 653–658.

Jóźwiakowski K. 2017a. Efficiency of organic substance removal in a hybrid sand filter with horizontal flow. Journal of Water and Land Development. No. 35 (X–XII) p. 95–100. DOI 10.1515/jwld-2017-0072.

Jóźwiakowski K. 2017b. Skuteczność usuwania związków biogenicznych w filtrach piaskowych z poziomym przepływem ścieków [Efficiency of biogenic compounds removal in a sand filters with horizontal flow]. Woda-Środowisko-Obszary Wiejskie. T. 17. Z. 4 p. 37–48.

Kaczmor B.G., Chmielewski K., Bugański P. 2017. Influence of accidental waters on the quality and loads of pollutants in wastewater discharged into the treatment plant. Journal of Water and Land Development. Vol. 33 (IV–VI) p. 73–78. DOI 10.1515/jwld-2017-0019.
STRESZCZENIE

W pracy omówiono wykorzystanie wieloklastrowej analizy statystycznej do oceny skuteczności filtracji ścieków bytowych. Obliczenia obejmowały dane zebrane w ciągu czterech miesięcy badań dotyczących wykorzystania odpadów jako materiału wypełniającego filtry z przepływem pionowym. Skuteczność usuwania zanieczyszczeń została przeanalizowana na przykładzie odpadów mechanicznie rozdrobnionych w postaci płatków PET, skrawków pianki PUR, rozdrobnionych opon gumowych i owaty. Analizowano związki organiczne (COD, BOD), zawiesiny ogólne, biogeny (jako jony NH₄⁺ i PO₄³⁻) oraz stopień nasycenia tlenu w przypadku filtracji z zastosowaniem do tego procesu wymienionych odpadów i porównano te wyniki z uzyskanymi w filtrach z wypełnieniem piaskowym. Wieloklastrowa analiza statystyczna dała podstawy do podziału skuteczności usuwania zanieczyszczeń przez materiały filtracyjne w zależności od obciążenia hydraulicznego na trzy grupy. Pierwsza z nich obejmowała parametry jakościowe oczyszczonych ścieków – największą redukcję BZT₅ i N-NH₄⁺. Znalazły się w niej wartości parametrów jakościowych i wskaźników dla filtratów otrzymanych w warunkach najmniejszego obciążenia hydraulicznego z kolumn wypełnionych odpadami gumowymi oraz piaskiem o miąższości 60 cm. Druga grupa obejmowała wyniki odnoszące się do wypełnień zawierających piankę poliuretanową, PET i gumę (dla pozostałych obciążeń hydraulicznych). Grupa ta charakteryzowała się największą redukcją zawiesin ogólnych i PO₄³⁻. Uzupełnienie łatwo ulegających biodegradowalnym związkom organicznych było na podobnym poziomie w obu grupach klastrów. Filtr wypełniony odpadami poliestrowymi (owatą) był tak samo skuteczny, jak filtry wypełnione 30-centymetrową warstwą piasku i filtry wypełnione piaskiem o miąższości 60 cm (w warunkach największego obciążenia hydraulicznego). W trzeciej grupie znalazły się najmniejsze wartości wszystkich analizowanych parametrów i wskaźników dla przedmiotowych filtratów.

Słowa kluczowe: analiza skupień, filtracja ścieków, opony gumowe, PET, pianka poliuretanowa, poliester