Henryk SOBCZUK[1] and Grzegorz ŁAGÓD[1]

SOME DISTANCE MEASURES FOR ESTIMATING THE EFFICIENCY OF WASTEWATER TREATMENT IN ACTIVATED SLUDGE SYSTEMS BY BIOMONITORING

POMIARY ODLEGŁOŚCI I PODOBIEŃSTWA TAKSOCENOZ W ASPEKCIE BIOINDYKACJI SYSTEMÓW OCZYSZCZANIA ŚCIEKÓW OSADEM CZYNNYM

Abstract: The results of bioindication, together with the physical and chemical parameters, may be used in wastewater treatment plant process management. The measurement of saprobiontic microorganisms' population is a relatively easy and cheap method for assessment of wastewater quality in sewage system. It can be also applied at the inflow of the wastewater treatment plant, at the following devices of technological line, as well as at outflow to receiver. Some methods allow to distinguish between distributions of saprobes population on the basis of species abundances. Splitting of a whole population of saprobes onto morphological-functional groups allows to simplify measurement and thus allow to make it *in situ*. The main idea of this type of measurement is to study the biocenosis entire structure. Because it is hardly possible, the approximately identified taksocenosis may be utilized for the needs of bioindication. An open question is how to compare the communities with different abundance of saprobes group. However using the different methods of biodiversity assessment the more or less visible differences among studied objects should be expected. In this paper the authors apply some existing distance measures based on entropy definitions to distinguish distributions of morphological-functional groups measured. The presented measures allow to tell how far - in terms of given measure of distance - are two distributions measured *in situ*, and differentiate them.

Keywords: bioindication, saprobes community, entropy measure

The results of bioindication, apart from the other physical and chemical parameters, may be used in wastewater treatment plant process management. Using the different methods of biodiversity assessment the more or less visible differences among studied objects should be expected [1-3]. The measurement of saprobiontic microorganisms' population is a relatively easy and cheap method for assessment of wastewater quality in sewer system, and can be applied at the inflow to the wastewater treatment plant as well as at outflow [1, 3, 4]. The main idea of this type of measurement is to study the biocenosis entire structure. Because it is hardly possible, the approximately identified taksocenosis may be utilized for the needs of bioindication.

The method described in paper prepared by Lagod et al [2] allows to distinguish between distributions of saprobes population on the basis of group of species abundances. Basically this method allows to draw some conclusions about the current status of wastewater but also on the history of the sewage quality change. If there is an event of microbiologically active pollutant discharge that seriously influences microbial activity at certain moment, the saprobe microorganism population will tend afterwards to the equilibrium state for a certain time. In order to discover such pollutants discharge in the past, the microorganisms distribution measurement and a comparison tool for different distributions are required.

[1] Faculty of Environmental Engineering, Lublin University of Technology, ul. Nadbystrzycka 40B, 20-618 Lublin, phone 81 538 44 81, email: H.Sobczuk@wis.pol.lublin.pl

The partitioning of a whole community onto morphological-functional groups allows to simplify measurement thus allow to make it even *in situ*. Trained employee can do it basically with an optical microscope [5]. An open question is how to compare populations with different abundance of group of species. The significance of differences amongst measurement results may be determined according to the popular statistical procedures, eg T-Student for distribution comparison. These tests require the specific strategy of the experiment - passive or active, determination of required number of samples for each studied level, rejection of doubtful results and the other actions standard for this type of tests. In case of WWTP (*Wastewater Treatment Plant*) sub-biocenosis object structures measurement, being in principle one of a kind, the single measurement result occurs. Thus, the use of usually applied mathematical tests is difficult.

The idea of entropy was introduced because many measures of this phenomenon may be applied to describe the differences among the described structures. The models similar to the one used in information theory [6] or likewise the others models describing the measure of entropy [7-10] may be useful in quantitative descriptions of different aggregations including sub-biocenosis such as populations of microorganisms settled inside WWTP devices.

The aim of this paper is to select, among the available descriptions of communities diversification, the one which would distinguish the WWTP sub-biocenosis communities the best.

Methods

Used method relies on the application of several available manners of taksocenosis diversification measure to the presented material. In this paper the authors apply some existing distance measures based on entropy definitions to distinguish between distributions of morphological-functional groups measured. Presented measures allow to tell how far - in terms of given measure of distance - are two distributions measured *in situ*, and differentiate them.

Each distance measure between two objects should fulfill three conditions: it should be positively defined ie the distance must be positive for the different objects and equal to zero for the identical objects; it should be symmetrical ie the distance from object A to B should be equal the distance from object B to A; and it should fulfill the triangle inequality. One of the simplest definitions of a distance function is an Euclidean distance function.

For the assessment of distance amongst random objects, like measured distributions of abundance of morphological-functional groups, the entropy-based measures are often applied. These measures usually lack of positive definition.

In mathematical theory of communication [6] there is a general entropy function of order α introduced in a form [8]:

$$H_\alpha = (1-\alpha)^{-1} \ln \sum_i p_i^\alpha \tag{1}$$

for α = 1 this formula reduces to well known Shanon formula:

$$H = -\sum_i p_i \ln p_i \tag{2}$$

where for $\alpha = 2$ it leads to another metric applied for biological diversity assessment [11]:

$$H = -\ln \sum_i p_i^2 \tag{3}$$

In [8] a mathematical background of application of entropy measures connected with generalized replicator equation was given. We will apply presented theory concerning state of the genetic system to the higher level - state of saprobes population on the basis of the abundances of species groups:

$$\dot{p}_i = h(p) f_i(p_i) \left\{ \sum_{j=1}^{n} w_{ij} f_j(p_j) - \theta^{-1}(p) \sum_{j,k=1}^{n} w_{kj} f_k(p_k) f_j(p_j) \right\}, \; i = 1, \ldots, n \tag{4}$$

where: $\dot{p}_i$ is a time derivative of probability p_i, $h(p)$ is the function determined by application, $f_i(p_i)$ are response functions with positive slope and starting from zero, $W = (w_{ij})$ is the matrix of interactions of microbial groups under consideration, and:

$$\theta(p) = \sum_{j=1}^{n} f_j(p_j) \tag{5}$$

The system of equations (4) describes the dynamics of the population. As shown in [8] on the basis of the equation (4) there is a possibility to calculate a response function for existing entropy measure and construct an entropy and distance function from response functions. The distance function generated in this way usually does not fulfill all the three requirements for distance function, for population comparison can be however applied.

As an example one can show that linear response function $f_i(p_i) = p_i$ generates distance function of the form:

$$H_{\hat{p}} = \sum \hat{p}_i \ln\left(\frac{p_i}{\hat{p}_i}\right) \tag{6}$$

which is equivalent to relative entropy, whereas logistic response function:

$$f_i(p_i) = 1 \big/ \left(b + c e^{-\alpha p_i}\right) \tag{7}$$

causes weighted logistic entropy:

$$H_{\hat{p}} = \sum \hat{f}_i \ln\left(\frac{1 - e^{-\alpha p_i}}{1 - e^{-\alpha \hat{p}_i}}\right) \tag{8}$$

to be an appropriate measure of distance. In this paper we will compare some measures of the distance among measured abundances of microorganisms, and also the distance given by the Euclidean norm.

Results and discussion

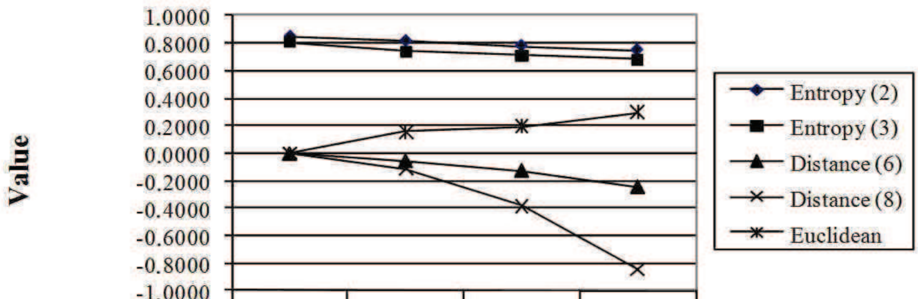
In the calculation we apply data set taken from publication [12], also analyzed in our previous papers [1, 2, 13]. The most important data are presented in Table 1 and used to calculate values of entropy according to equation (2) and (3) and distances to the first data set according to equations (6) and (8).

Table 1

Species numbers and species abundances in specified classes of wastewater purity [1, 2]

BOD_5 / organisms	**0÷10 g O_2/m^3**	**11÷20 g O_2/m^3**	**21÷30 g O_2/m^3**	**> 30 g O_2/m^3**
swimming ciliates	27	25	18	12
	1969	**3712**	**1477**	**1392**
attached ciliates	35	29	26	18
	5862	**5284**	**5198**	**1606**
crawling ciliates	14	14	12	7
	1449	**1535**	**1263**	**1142**
all ciliates	76	68	56	37
	9280	**10531**	**7938**	**4140**
rotifers	45	40	35	24
	6354	**2318**	**503**	**962**
flagellates	13	15	19	26
	4785	**4199**	**4463**	**6399**
amoebas	14	14	14	7
	3914	**3507**	**4385**	**4850**
others	8	9	1	1
	4214	**1328**	**1100**	**50**
all	152	146	125	95
	28547	**21883**	**18389**	**16401**

In rows: S - species richness (number of species), $\mathbf{n_T}$ - total number of individuals in cm^3 of wastewater (after multiplication by 10^3 in dm^3), others - the rest of organisms representing *Metazoa:* nematodes, oligochaetes, tardigrades, gastrotriches, arachnids, copepods, cladocers, turbellarians



	1	2	3	4
Entropy (2)	0.8481	0.8158	0.7803	0.7502
Entropy (3)	0.8077	0.7381	0.7123	0.6790
Distance (6)	0.0000	-0.0527	-0.1248	-0.2464
Distance (8)	0.0000	-0.1114	-0.3863	-0.8477
Euclidean	0.0000	0.1571	0.1947	0.2945

Fig. 1. Comparison of different distance measures for data distributions

Results are presented in Figure 1. It is clearly visible that distance measures presented here much more differentiate the data set than entropy measure. As such, the presented models can be used more effectively to parameterize measurement data distributions than entropy measures. Due to the possibility to generate measures according to the functional form of the response function in (4) one can find the best distance measure for the certain purpose. For the presented data set the measure (8) is differentiating data sets best of all applied methods.

Conclusion

Based on the calculation presented here the following conclusions are offered:

- As a result of the presented method the data sets used in our studies were successfully distinguished, as it was shown at Figure 1.
- Evidently, the most significant differences in microorganisms communities are visible when model described by equation (8) was applied.
- Entropy and distance measures should be appropriately chosen for complex data sets comparison. Equation (4) with accompanying theory allows to generate distance measures according to the data set character. It shows up that the distance measure (8) is best fitted to recognize differences in microbial distribution functions described here.
- In the situation of the necessity of subtle differences in compared taksocenosis structures perception, basing on the distance determination method, the most suitable method seems to be the model presented in equation (8).

Acknowledgement

This work was supported by the Ministry of Science and Higher Education of Poland, No. 4949/B/T02/2008/34.

References

[1] Montusiewicz A., Malicki J., Łagód G. and Chomczyńska M.: *Estimating the efficiency of wastewater treatment in activated sludge systems by biomonitoring*. [In:] Environmental Engineering, eds. L. Pawłowski, M. Dudzińska, A. Pawłowski. Taylor and Francis Group, London 2007, 47-54.

[2] Łagód G., Malicki J., Chomczyńska M. and Montusiewicz A.: *Interpretation of the results of wastewater quality biomonitoring using saprobes*. Environ. Eng. Sci., 2007, **24**(7), 873-879.

[3] Chomczyńska M., Montusiewicz A., Malicki J. and Łagód G.: *Application of saprobes for bioindication of wastewater quality*. Environ. Eng. Sci., 2009, **26**(2), 289-295.

[4] Lydy M.J., Crawford C.G. and Frey J.W.: *A comparison of selected diversity, similarity, and biotic indices for detecting changes in benthic-invertebrate community structure and stream quality*. Arch. Environ. Contamin. Toxicol., 2000, **39**, 469-479.

[5] Łagód G. and Sobczuk H.: *The number and size of samples required to measure the saprobe population at various pollutant concentrations in sewage*. Arch. Environ. Protect., 2008, **34**(3), 281-285.

[6] Shannon C.E. and Weaver W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana 1949.

[7] Pykh Y.A.: *Lyapunov functions as a measure of biodiversity: theoretical background*. Ecol. Indicat., 2002, **2**, 123-133.

[8] Pykh Y.A. and Malkina-Pykh I.G.: Replicator equations, response functions and entropy measures in science: mathematical background. [In:] Ecosystems and Sustainable Development V, WITpress, 2005.

[9] Naterer G.F. and Camberos J.A.: Entropy-Based Design and Analysis of Fluids Engineering Systems. CRC Press, Taylor and Francis Group, Boca Raton, London, New York 2008.

[10] Bogaert J., Farina A. and Ceulemans R.: *Entropy increase of fragmented habitats: A sign of human impact?* Ecol. Indicat., 2005, **5**, 207-212.
[11] Pielou E.C.: Ecological Diversity. John Wiley and Sons, New York 1975.
[12] Klimowicz H.: Znaczenie mikrofauny przy oczyszczaniu ścieków osadem czynnym. Zakł. Wyd. Instytutu Kształtowania Środowiska, Warszawa 1983.
[13] Łagód G., Sobczuk H. and Suchorab Z.: *Application of a saprobiontic microorganisms community analysis in the calibration of a model description of sewage self-purification in sewer systems*. Ecol. Chem. Eng., 2006, **13**(3-4), 265-275.

POMIARY ODLEGŁOŚCI I PODOBIEŃSTWA TAKSOCENOZ W ASPEKCIE BIOINDYKACJI SYSTEMÓW OCZYSZCZANIA ŚCIEKÓW OSADEM CZYNNYM

Wydział Inżynierii Środowiska, Politechnika Lubelska

Abstrakt: Rezultaty pomiarów bioindykacyjnych mogą stanowić przydatne uzupełnienie, a niekiedy nawet alternatywę dla pomiarów fizykochemicznych wykonywanych w celach kontroli i sterowania procesami oczyszczania ścieków. Pomiary liczebności odpowiednio dobranych populacji organizmów saprobiontycznych zasiedlających poszczególne urządzenia oczyszczalni ścieków mogą być relatywnie proste i tanie. Stąd też wykorzystujące je metody pomiarowe są możliwe do zastosowania zarówno do oceny jakości ścieków dopływających do oczyszczalni, oczyszczanych w kolejnych urządzeniach ciągu technologicznego, jak i odprowadzanych po oczyszczeniu do odbiornika. Niektóre ze wspomnianych metod pozwalają oceniać kompozycję zbiorowisk saprobów na podstawie liczebności poszczególnych gatunków. Jednakże podział całej populacji saprobiontów na grupy funkcjonalno-morfologiczne umożliwia znaczne uproszczenie pomiarów oraz wykonywanie ich *in situ*. Ze względu na fakt, iż pomiar struktury całej biocenozy saprobów jest trudny do realizacji, identyfikacja wybranych taksocenoz może być wykorzystana do celów bioindykacyjnych. Ciągle otwartą kwestią pozostaje, jak najlepiej porównywać pomiędzy sobą zbiorowiska saprobów złożone z grup o różnej liczebności. Używając istniejących metod oceny bioróżnorodności, uzyskać można wartości liczbowe indeksów i wskaźników mniej lub bardziej różnicujące porównywane między sobą zgrupowania organizmów. W niniejszym opracowaniu autorzy wykorzystują sposoby opisu odległości korzystające z definicji entropii w celu jak najlepszego zróżnicowania zbiorowisk, złożonych z grup morfologiczno-funkcjonalnych, do których sklasyfikowano organizmy bytujące w ściekach o różnym stopniu zanieczyszczenia. Prezentowane procedury pozwalają określić, jak daleko - w kategoriach przyjętych miar odległości - znajdują się od siebie rozkłady analizowanych zbiorów i odpowiednio je zróżnicować.

Słowa kluczowe: bioindykacja, zbiorowiska saprobów, pomiary entropii