The hydrological and hydrochemical division of the surface waters in the Gulf of Gdańsk

OCEANOLOGIA, 40 (3), 1998.
pp. 261–272.

KEYWORDS
Gulf of Gdańsk
Surface water
Temperature
Salinity
Nutrients
Surface layer division
Cluster analysis
Ward's method

Jacek Nowacki
Institute of Oceanography,
Gdańsk University,
Al. Marsz. Piłsudskiego 46, 81–378 Gdynia, Poland; e-mail: ocejn@univ.gda.pl

Ewa Jarosz
Louisiana State University,
Baton Rouge, LA 70803, USA

Manuscript received May 8, 1998, in final form July 15, 1998.

Abstract

The objective of this work was to define regions in the surface layer of the Gulf of Gdańsk that are characterised by different natural (temperature, salinity, total water depth) and anthropogenic (concentrations of major nutrients) factors. Additionally, approximate boundaries between the regions distinguished were delineated. To find regions of different hydrological and hydrochemical properties, Ward's method, an agglomerative hierarchical cluster technique, was applied to a data set containing measurements of temperature, salinity, total depths and nutrient concentrations.

Cluster analysis divided the surface waters of the Gulf of Gdańsk into seven regions. The surface waters in the shallow water part of the Gulf of Gdańsk, where six out of the seven regions were located, displayed a distinct diversity which was probably associated with the morphometry of the Gulf, nutrient sources and freshwater input. The last region encompassed the deep-water part of the Gulf of Gdańsk and Puck Bay.

1. Introduction

In the marine environment, relationships between anthropogenic factors, mostly pollutants, and natural conditions are readily detectable. As a result, the hydrological and hydrochemical characteristics of the waters in the Gulf of Gdańsk display considerable spatial and temporal variability. This is essentially due to interactions of marine and fresh waters, the latter being introduced to the system primarily by the river Vistula, anthropogenic diversity and the morphometry of the Gulf itself.

Divisions of the waters of the Gulf of Gdańsk for different purposes have already been attempted on a number of occasions. Even though their principal objective was not to determine zones with different hydrochemical and hydrological properties, the divisions of the Gulf thereby obtained have been used as an additional tool in describing other environmental properties in the study area. Morphometry has often been a criterion of such divisions (Trzosińska, 1978; Nowacki, 1993a), with the consequent hydrological (Nowacki, 1993b) or hydrochemical characterisation (Falkowska *et al.*, 1993) of the Gulf. Such a division could be of benefit in detailed analyses of regions with high concentrations and accumulations of pollutants, in descriptions of marine organisms in connection with the hydrological and hydrochemical properties of water, and in studies of the formation and behaviour of hydrochemical and hydrological fronts in the Gulf of Gdańsk. In the authors' opinion, however, existing knowledge about the diversity of Gulf waters is insufficient and a more accurate division of this area is required. The objective of the present work was therefore to define regions in the surface layer of the Gulf of Gdańsk characterised by different natural (temperature, salinity, total water depth) and anthropogenic (concentrations of major nutrients) factors. Additionally, approximate boundaries between the regions distinguished were determined. In order to find the regions with different hydrological and hydrochemical properties, an objective method – cluster analysis – was applied to a data set containing measurements of temperature, salinity, total depths and nutrient concentrations.

2. The data set and the methodology applied

To determine the division of the surface waters of the Gulf of Gdańsk down to 5 m into separate regions, a data set containing readings taken during monitoring cruises conducted in the study area by the Institute of Oceanography of the University of Gdańsk in 1981–1985 was analysed. The cruises had been planned to take place once a month, but owing to inclement weather, only 50 cruises actually took place. During each cruise the same station grid was sampled (Fig. 1). At each station, temperature (°C), salinity

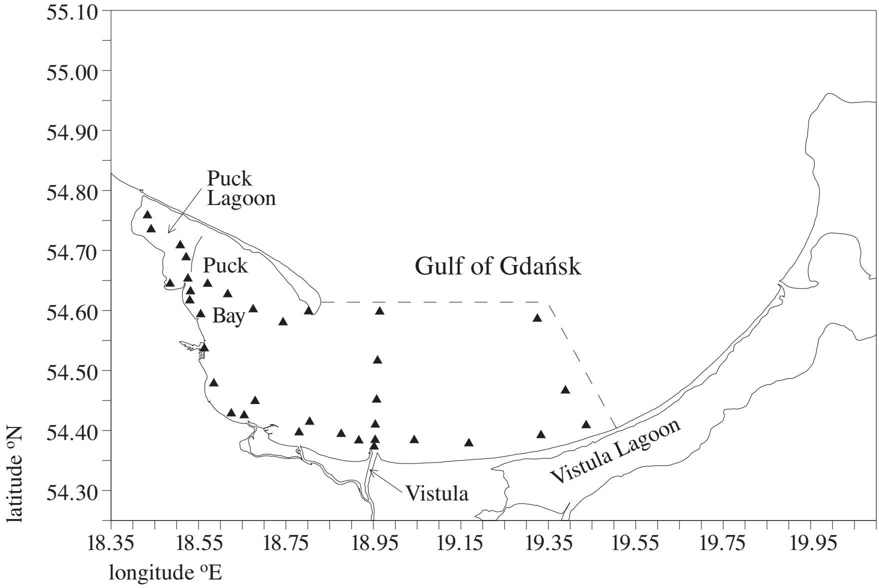


Fig. 1. Area of study and location of stations examined in the Gulf of Gdańsk in 1981–1985

(PSU) and the concentrations of dissolved phosphate (μmol dm^{-3}), silicate (μmol dm^{-3}), nitrate (μmol dm^{-3}), nitrite (μmol dm^{-3}) and ammonia (μmol dm^{-3}) were measured. However, for a variety of reasons, the time series obtained were sometimes shorter than expected, so only the stations with at least 36 measurements of each parameter listed above were taken into account in the analysis.

The data set was subjected to cluster analysis. This objective statistical method assumes only that objects (entities) characterised by similar properties do not belong to different groups (clusters). Cluster analysis encompasses a number of different classification algorithms. For the purpose of this paper, Ward's method (Ward, 1963) was chosen. This aims to find at each stage of grouping those two clusters, the merger of which gives the minimum increase in the total within the group error of squares. Ward's method not only ensures homogeneity (variance minimum or minimum of the total within the group error of squares) within a single group, but also heterogeneity (variance maximum) among clusters.

This method is an example of agglomerative hierarchical cluster techniques, which are generally defined as a series of successive fusions of N entities into groups. These methods ultimately reduce the data to a single cluster containing all the objects. The results of agglomerative hierarchical methods may be presented in the form of a dendrogram, which is a two-dimensional diagram illustrating the fusions that have been made at each successive level.

The basic procedure with all these agglomerative hierarchical techniques is similar. They begin with the computation of a similarity or distance matrix between the objects, which are considered as one-element clusters at the beginning of the analysis. Subsequently, the procedure generally consists of three major steps:

1. searching out in the distance or similarity matrix the closest or most similar pair of objects (s_i, s_j);
2. grouping these two objects (s_i, s_j) in one new cluster (s_a);
3. changing the distance or similarity matrix after the new cluster has been created by calculating distances or similarities between s_a and other objects and reducing the matrix order by one.

These three steps are repeated until there is only one group containing all the entities. This recurrent part of the procedure can be expressed by the formula (Lance and Williams, 1966, 1967; Marek, 1989)

$$d(s_a, s_k) = \alpha_i d(s_i, s_k) + \alpha_j d(s_j, s_k) + \beta d(s_i, s_j) + \\ + \gamma |d(s_i, s_k) - d(s_j, s_k)|, \quad (1)$$

where

$d(..., ...)$ – the distance or similarity between objects,
s_i, s_j, s_k – entities,
s_a – the new cluster obtained by grouping s_i and s_j,
$\alpha_i, \alpha_j, \beta, \gamma$ – coefficients depending on the clustering method chosen.

By substituting for α_i, α_j, β, and γ in (1), one can obtain a formula describing Ward's method (Marek, 1989):

$$d(s_a, s_k) = \frac{(n_i + n_k)d(s_i, s_k) + (n_j + n_k)d(s_j, s_k) - n_k d(s_i, s_j)}{(n_i + n_j + n_k)}, \quad (2)$$

where

n_i, n_j, n_k – the numbers of objects belonging to s_i, s_j, s_k.

Ward's method with the distance matrix was used to group the stations shown in Fig. 2. The Euclidean distance was selected as the measure of distance. To cluster the stations, eight comparative characteristics (variables) – temperature, salinity and the concentrations of dissolved phosphate, silicate, nitrate, nitrite, ammonia, and the total depth (m) of the station – were taken into consideration. Since a division of the surface waters usually characterised by the highest concentration of anthropogenic substances (Nowacki, 1992) was being sought, only the surface measurements of the first seven variables were taken into account.

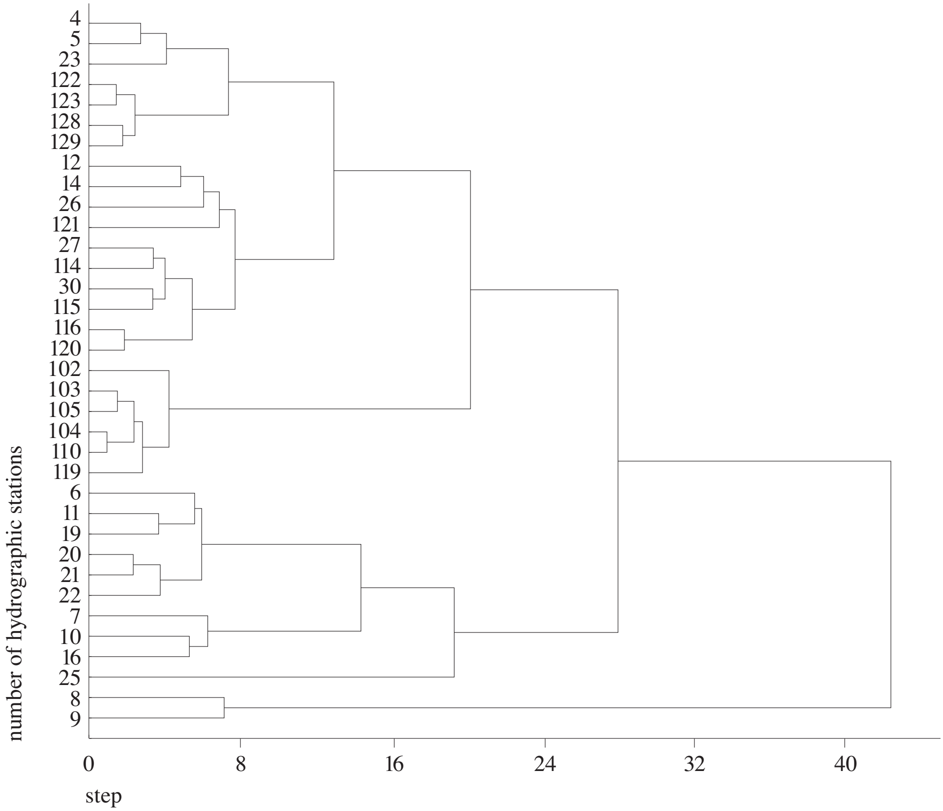


Fig. 2. Ward's method dendrogram for grouping 36 variables

The comparative variables were divided into two groups. The first one contained the annual averages of temperature, salinity and nutrient concentrations, as well as the total depth (36 comparative characteristics). The second group contained 61 comparative parameters, because the averages of the nutrients from the cold season (October, November, December, January, February, March) were included. This group of variables was analysed in order to include the annual nutrient cycle, which is characterised by maximum concentrations in the cold season and minimum ones associated with intensive phytoplankton growth (Trzosińska, 1990; Falkowska *et al.*, 1993) during the remaining part of the year.

3. Results

The analyses of the results obtained using Ward's method (Figs. 2 and 3) for the first group of comparative variables shows that there are seven major clusters. These occurred for almost the same value of the distance

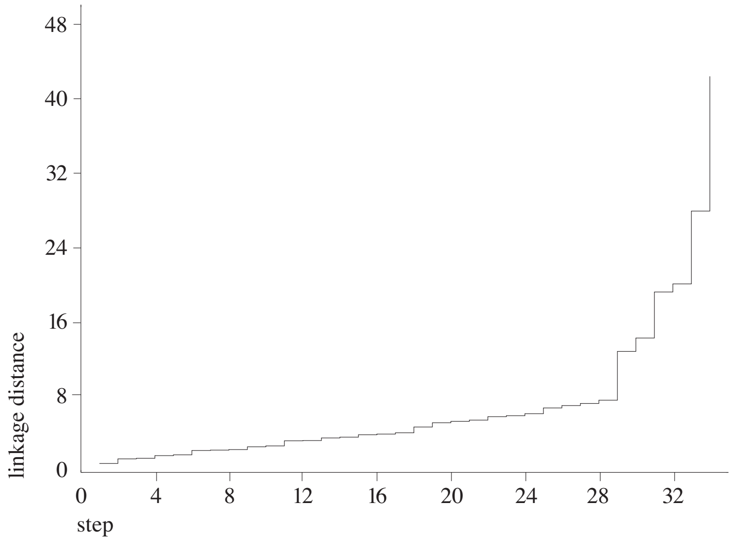


Fig. 3. Linkage distances across steps for 36 grouping variables

Table 1. Major clusters and stations belonging to these clusters – 36 comparative variables

Cluster number (region number)	Objects (station numbers) belonging to a cluster
I	8, 9
II	7, 10, 16
III	6, 11, 19, 20, 21, 22
IV	25
V	12, 14, 26, 27, 30, 114, 115, 116, 120, 121
VI	102, 103, 104, 105, 110, 119
VII	4, 5, 23, 122, 123, 128, 129

function (~ 8). Tab. 1 provides information about these clusters and the stations belonging to them.

The results obtained for the second group of comparative parameters were quite similar to those obtained for the first group of variables. For the distance function, in this case equal to 10, there were also seven major clusters. Among the 35 objects, only two were assigned to different clusters when the higher weight was attributed to the nutrients. This may suggest that the addition of 25 comparative characteristics did not cause any significant changes to appear in the clusters created.

To confirm the cluster analysis results and, at the same time, to prove the homogeneity of the clusters, the Kruskal-Wallis test for comparing more

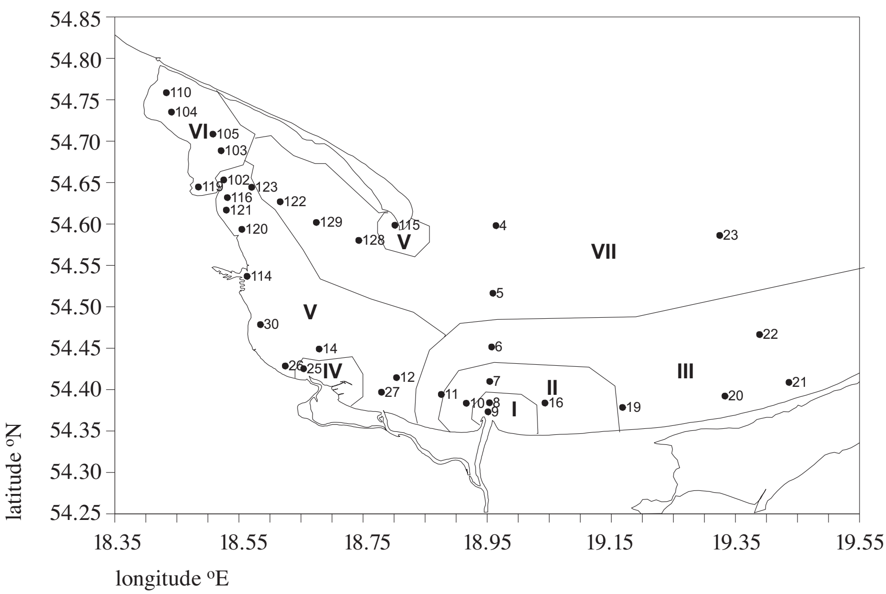
Table 2. Basic statistics computed for seven comparative variables for each cluster

Variable (component environment)	Statistics	Cluster number (regions numbers) I	II	III	IV	V	VI	VII
PO_4	$\bar{x}$	1.68	1.05	0.66	7.42	1.17	0.61	0.62
[μmol dm^{-3}]	s	1.52	1.34	0.74	8.16	1.44	0.55	0.79
	min	0.04	0.05	0.01	0.04	0.02	0.04	0.01
	max	8.51	12.19	5.79	44.30	10.42	2.95	10.22
	N	94	139	257	42	408	240	292
SiO_2	$\bar{x}$	57.30	30.32	16.34	18.59	12.12	12.29	11.37
[μmol dm^{-3}]	s	60.06	42.55	23.22	17.78	12.88	9.75	12.83
	min	2.20	0.60	0.50	2.10	0.90	0.50	0.30
	max	224.10	249.90	212.20	96.50	126.40	58.60	117.90
	N	98	143	265	46	434	260	324
NO_3	$\bar{x}$	28.65	16.31	8.33	7.34	3.90	2.36	3.12
[μmol dm^{-3}]	s	26.14	20.78	13.49	14.84	7.92	3.05	5.36
	min	0.14	0.14	0.04	0.14	0.04	0.04	0.05
	max	96.24	99.60	89.17	89.85	74.69	25.41	53.50
	N	92	132	253	44	401	236	287
NO_2	$\bar{x}$	0.87	0.57	0.32	0.37	0.25	0.21	0.22
[μmol dm^{-3}]	s	0.66	0.51	0.32	0.27	0.28	0.24	0.28
	min	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	max	2.83	2.32	1.91	1.14	2.56	0.91	0.20
	N	98	140	269	45	426	254	321
NH_4	$\bar{x}$	12.12	6.81	3.00	5.33	2.69	2.18	1.65
[μmol dm^{-3}]	s	19.02	12.18	5.23	5.03	4.12	2.59	1.90
	min	0.16	0.06	0.03	0.06	0.06	0.06	0.05
	max	90.93	73.58	44.63	20.99	49.66	17.64	23.80
	N	94	140	271	43	421	256	315
Temperature	$\bar{x}$	10.89	10.65	10.83	10.32	10.47	11.57	10.13
[°C]	s	7.06	3.46	6.40	6.27	6.32	6.82	6.47
	min	–0.26	0.16	0.03	–0.09	–0.28	0.01	–0.13
	max	21.62	21.64	22.31	19.33	23.45	24.11	23.80
	N	98	145	278	46	441	261	330
Salinity	$\bar{x}$	3.97	5.84	6.88	7.18	7.53	7.34	7.60
[PSU]	s	2.15	1.77	1.24	0.66	0.67	0.47	0.62
	min	0.33	1.37	1.14	5.12	2.00	5.14	4.45
	max	7.86	8.27	8.95	8.16	8.70	8.45	9.14
	N	98	145	278	45	439	267	331

$\bar{x}$ – mean, s – standard deviation, min – minimum value, max – maximum value, N – number of measurements.

than two populations was used. A null hypothesis, which was tested, stated, for instance, that population distributions of the concentration of dissolved phosphate at stations belonging to the same cluster were identical. This hypothesis was tested for all nutrients, salinity and temperature in the clusters containing at least two stations. In all the cases analysed, the null hypotheses were accepted at the significance level $\alpha = 0.05$. Acceptance of these hypotheses also permitted the conclusion that in each cluster, each variable had a certain distribution that could be characterised by statistics computed from the sample measurements. Since the actual distributions were not known, basic statistics such as the arithmetic mean, standard deviation, and maximum and minimum values were calculated for each variable except station depth (Tab. 2).

Thus on the basis of the cluster analysis result, the surface waters of the Gulf of Gdańsk are divisible into seven regions with different hydrological and hydrochemical characteristics (Fig. 4, Tab. 2). However, owing to the small number of measurement stations and their irregular locations, delineation of precise boundaries between them is not really possible. These are therefore drawn in such a way as to pass between stations belonging to different clusters. Fig. 4 shows that six of the regions – the majority of them – are located in the shallower part of the Gulf (maximum depth 20–30 m).



Roman numerals – regions; Arabic numerals – stations

Fig. 4. Division of the surface waters of the Gulf of Gdańsk based on the data collected in 1981–1985

Such a diversity in the shallow zone of the Gulf is probably associated with the large number of nutrient sources and the freshwater input.

The surface waters to the north of the Vistula mouth display a distinct diversity and can be divided into three different regions. The first region (I) lies just north of the river mouth and is stongly influenced by the Vistula discharge. This high impact of the riverine waters is reflected by the high concentrations of nutrients such as silicate, nitrate, nitrite and ammonia, as well as by the very low salinity. The second region (II) is farther north, where the nutrient concentrations are *ca* 50% lower and the salinity much higher than in the first zone. The third region (III), which may be treated as a transition region as far as hydrological and hydrochemical conditions are concerned, straddles the shallow and the deeper part of the Gulf of Gdańsk (depth *ca* 50 m). While high salinity waters ($\geq$ 6 PSU) are dominant there (80% of salinity measurements), waters of salinity $<$ 6 PSU have also been recorded there (20%). Moreover, the concentrations of dissolved nutrients decrease further in this area: they are *ca* 50% lower than in the second region.

The next region distinguished by the cluster analysis (VII) covers a major section of the deep part of the Gulf of Gdańsk and the deep part of Puck Bay. The average surface water salinity – 7.60 PSU – is the highest of all the regions. By contrast, the nutrient concentrations are much lower than those in the regions described previously. However, the decrease in these concentrations between this region and region (III) is not as abrupt as between (II) and (III). It is worth noting that the decrease in nutrient concentrations is associated with increasing salinity. Such behaviour is very apparent in regions (I), (II), (III) and (VII) and may suggest that there is a strong correlation between nutrient concentrations and salinity.

Another area (V) with a distinct hydrology and hydrochemistry occupied the western, shallow-water part of the Gulf of Gdańsk, the part of Puck Bay extending as far as the Tern Sandbank and a small area around the tip of Hel Peninsula. The average salinity there is 7.34 PSU and the average concentration of phosphate is 1.17 μmol dm^{-3}, silicate – 12.12 μmol dm^{-3}, nitrate – 3.9 μmol dm^{-3}, nitrite – 0.25 μmol dm^{-3} and ammonia – 2.69 μmol dm^{-3}. Within this region, there is a small area (IV) with the highest concentration of dissolved phosphate (7.42 μmol dm^{-3}). The concentrations of the other nutrients (except nitrate) are lower only than those measured in regions (I) and (II). Such high concentrations are probably associated with the municipal and industrial sewage flowing into the Gdynia port channel and then discharged into the Gulf of Gdańsk.

The last region (VI) is located in the shallowest part of Puck Bay (Puck Lagoon), north of the Tern Sandbank. The highest average temperature was recorded there. The average nutrient concentrations are similar to those measured in regions (V) and (VII). The salinity of the surface waters is on average 7.34 PSU. To the north-east of this region lies an area that could not be allocated to any of the other regions, because there were no stations in this region; this area has therefore been excluded from the cluster analysis.

4. Conclusions

Ward's method, which enabled different comparative hydrological and hydrochemical parameters to be accounted for, showed that the surface waters of the Gulf of Gdańsk could be divided into seven different regions. This method of classification appeared to yield results more objective than those emerging from the earlier studies because it employed a diversity of environmental parameters, but also was not influenced by attaching more importance to some of these characteristics. The statistical test and basic statistics calculated from the data also confirmed the homogeneity of the regions identified by the cluster method.

A distinct diversity of the surface waters was found in the shallow-water part of the Gulf of Gdańsk, where six out of the seven regions were located. This diversity was probably associated with the morphometry of the Gulf, nutrient sources and the freshwater input. The significant impact of the Vistula on the surface waters was much in evidence north and east of the river mouth. In this area, three regions with different hydrological and hydrochemical characteristics were found. The first zone (I), with the lowest salinity and highest concentrations of silicate, nitrate, nitrite and ammonia was located just north of the river mouth. Farther north and east, there were two more regions (II and III) characterised by increased salinity and reduced nutrient concentrations in comparison with the first zone. It was in these zones that the hydrological and hydrochemical fronts were probably located. The location and shape of the three regions result from the main direction of the spread of Vistula waters. The spatial extent of the freshwater plume was usually the greatest along the shore in an easterly direction (Matciak and Nowacki, 1995). The first region (I) is the area immediately affected by the stream of fresh waters flowing out with the Vistula, and is bordered by with isohaline 2 PSU (Matciak and Nowacki, 1996). The second (II) covers the area adjacent to the river mouth, bordered by the hydrological front of the Vistula, coincides with the positions of the 4–5 PSU isohalines (Matciak and Nowacki, 1996). The third one (III) covers the zone of Vistula water transformation, and is delineated by the mean positions of the 6–7 PSU isohalines (Cyberska, 1990).

West of the Vistula mouth there is a region (V) occupying the western, shallow water part of the Gulf, displaying relatively low nutrient concentrations and a higher salinity than those above. This region interacts with the littoral zone which is subject to considerable inflows of saline surface waters from the Gdańsk Basin. This might suggest that the riverine waters had little or no influence on the surface waters of this particular region. Within this area, the small zone (IV) was distinguished probably as a result of the high input of municipal and industrial sewage from the port channel. Farther north, the shallow part of Puck Bay was classified as a separate area with different hydrological and hydrochemical parameters. Finally, the last zone was found in the deep part of the Gulf of Gdańsk and Puck Bay. The highest salinity and the lowest nutrient concentrations were characteristic of the surface waters of this region.

The division obtained is also applicable to the whole water layer down to 15–20 m, except for the area of the freshened water plume near the Vistula mouth, where a significant change in all parameters occurs at a depth of 1–2 m. Such a conclusion is confirmed by the preliminary investigation of the vertical structure of the water column (Cyberska 1990; Nowacki, 1992; Nowacki and Matciak, 1996). The water division obtained in this work appears to result essentially from the variability in the hydrological conditions (T, S). In spite of the fact that the data were collected in 1981–1985, the number of zones and their spatial extent are not expected to have changed, so they should be detectable today as well. The division of the surface waters of the Gulf of Gdańsk presented here might be useful for a greater understanding of the concentration and accumulation of other substances such as pollutants, which may be introduced into the marine system with the discharge from rivers and port channels. Such an understanding may be helpful in evaluating the environmental conditions in the Gulf.

For a precise determination of the location of the zones, the number of stations in the study area would have to be increased. This is very important in the regions where the properties of the water are highly diverse, for example near the Vistula mouth (Nowacki and Matciak, 1996). Finally, for a more complete classification of the aquatic environment of the Gulf of Gdańsk, further investigations based on a greater number of comparative variables including physical, chemical and biological parameters, is required.

References

Cyberska B., 1990, *Water salinity in the Gdańsk Basin*, [in:] *The Gulf of Gdańsk*, A. Majewski (ed.), Wyd. Geol., 237–255, (in Polish).

Falkowska L., Bolałek J., Nowacki J., 1993, *Nutrient and oxygen in the Gulf of Gdańsk*, Stud. i Mater. Oceanol., 64, 131–162.

Lance G. N., Williams W. T., 1966, *A generalised sorting strategy for computer classifications*, Nature, 212, p. 218.

Lance G. N., Williams W. T., 1967, *A general theory of classificatory sorting strategies hierarchical systems*, Comp. J., 9, 373–380.

Marek T., 1989, *Cluster analysis in empirical research: Methods SAHN*, PWN, Warszawa, 170 pp., (in Polish).

Matciak M., Nowacki J., 1995, *The Vistula river discharge front – surface observations*, Oceanologia, 37 (1), 75–88.

Matciak M., Nowacki J., 1996, *Hydrological conditions in the area enclosed by the Vistula river discharge front*, Prz. Geofiz., 41 (4), 275–285, (in Polish).

Nowacki J., 1992, *Konzentration der N- und P-Verbindungen im Küstengebiet der Danziger Bucht. Die Belastung der Danziger Bucht – ein europäisches Problem globaler Bedeutung*, Bibl. Informationssys. Univ. Oldenburg, Oldenburg, 115–123.

Nowacki J., 1993a, *Gulf morphometry*, [in:] *Puck Bay*, K. Korzeniewski (ed.), Fund. Rozw. Uniw. Gdańs, Gdańsk, 71–78, (in Polish).

Nowacki J., 1993b, *Temperature, salinity and density of water*, [in:] *Puck Bay*, K. Korzeniewski (ed.), Fund. Rozw. Uniw. Gdańsk, Gdańsk, 79–111, (in Polish).

Trzosińska A., 1978, *Calculation of water volume of the Gdańsk Basin for a chemical budget of the Baltic Sea*, Prz. Geofiz. 23 (31), 7–13, (in Polish).

Trzosńska A., 1990, *Nitrogen and phosphorus compounds*, [in:] *The Gulf of Gdańsk*, A. Majewski (ed.), Wyd. Geol., 275–291, (in Polish).

Ward J. H., 1963, *Hierarchical grouping to optimise and objective function*, J. Am. Statist. Assoc., 58, 301, 236–244.