ICES Journal of Marine Science, 60: 1–10. 2003
doi:10.1006/jmsc.2002.1320

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Distribution fields for aquatic ecosystem components: method of optimization of correlation zones

I. Kalikhman, Y. Z. Yacobi, and M. Gophen

Kalikhman, I., Yacobi, Y. Z., and Gophen, M. 2003. Distribution fields for aquatic ecosystem components: method of optimization of correlation zones. – ICES Journal of Marine Science, 60: 1–10.

A method was initially suggested for the identification of zones where there are strong correlations between two distribution fields. The method was based on outlining areas where the gradients of the two fields are in the same or opposite directions. In the present study, correlation zones were optimized by varying the range of angles between gradients, simultaneously changing the correlation value and zone area. The major role was given to the correlation value; the zone area was maintained so that the optimized correlation zone was significant. The method was applied in an attempt to establish the interaction of spatial distribution of a grazer (Cladocera) and its food (nanophytoplankton) in Lake Kinneret (Israel), using data from multidisciplinary surveys carried out on 26 December 1991 and 13 May 1993.

Keywords: correlation zone, gradient, optimization, patchiness, significance.

Received 17 December 2001; accepted 20 August 2002.

I. Kalikhman, Y. Z. Yacobi, and M. Gophen: Yigal Allon Kinneret Limnological Laboratory, Israel Oceanographic and Limnological Research Ltd, PO Box 8030, Haifa 31080, Israel. Correspondence to I. Kalikhman: tel: 972 4 851 5202; fax: 972 4 851 1911; e-mail: igal@ocean.org.il

Introduction

Patchiness is a fundamental attribute of ecosystems (Steele, 1976). The state of the ecosystem in different zones can be controlled by various dominant factors causing full or partial overlap of patches of different components. Until recently, there was no quantitative method for elucidating relationships between patchy distribution fields. As a result, the underlying causes for the distribution pattern observed could not be revealed.

To help fill this gap, Kalikhman *et al.* (1995) elaborated a method of identifying zones where there are strong correlations between two distribution fields. The method enabled the evaluation of the fish-zooplankton interaction, and assessment, for the first time, of a threshold level of food source concentration from data obtained in nature. A threshold means the following: if the prey density is lower than the threshold, the predator's behaviour is aimed mainly at searching for food; if the prey density is higher than the threshold, the feeding dominates (Nikolsky, 1963). Similar methods were used by Ellison *et al.* (2000) in testing a hypothesis that mangroves appear in Bangladesh in discrete zones depending on elevation. An improvement of the method (determining the zone significance) allowed the estimation of the extent of heterogeneity of phytoplankton distribution (Kalikhman, 1999).

By definition, the aim of identifying a correlation zone is to obtain the strongest correlations between ecosystem components (Kalikhman *et al.*, 1995). At the same time, development of the method considered in the present study, has shown that stronger correlations are observed in smaller zones; in contrast, weaker correlations are associated with larger zones. The significance of a correlation zone depends on both these factors. The real value of the significance cannot be determined as, for significant correlations, the correlation should *exceed* the critical value of a correlation coefficient, no matter for how many (Sokal and Rohlf, 1969). Therefore, during optimization, the major role is given to the correlation value; the zone area is maintained so that the indicated condition of significance is satisfied. The aim of the present study is to demonstrate a method of optimization of correlation zones.

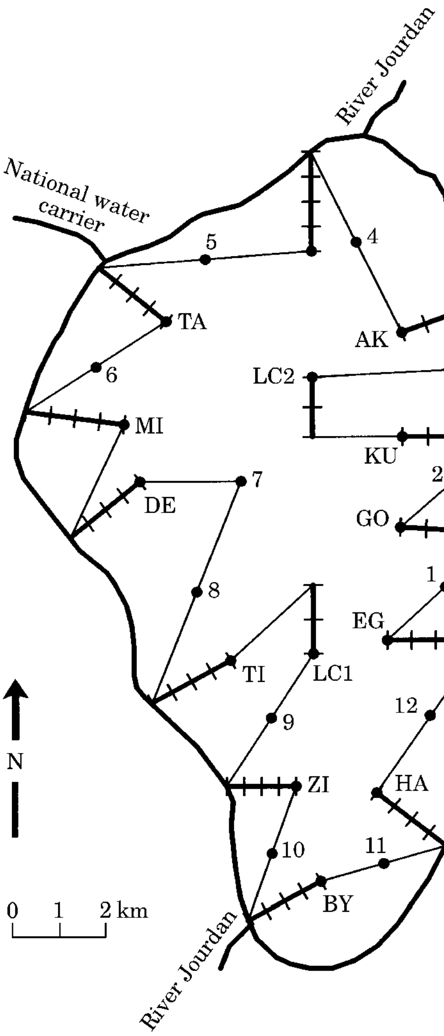

Figure 1. Map of the multidisciplinary survey coded as follows: thick lines – acoustical transects; perpendicular marks – sampling distance unit; thin lines – connecting tracks; points – sampling stations.

Method and material

Our example is based on measurements carried out in the subtropical Lake Kinneret, Israel. The distribution fields were reconstructed on the basis of the data from multidisciplinary surveys carried out on 26 December 1991 and 13 May 1993. Each survey consisted of an acoustic survey of fish concentrations conducted together with the sampling of plankton (Figure 1). Densities of zoo- and phytoplankton were measured at fixed stations in integrated mixed samples within the 0–10 m layer (Kalikhman *et al.*, 1995).

The biomass of zooplankton is dominated by three major groups: Cladocera – 58%, Copepoda – 35% and Rotifera – 7% (Gophen, 1978). Distribution fields for Cladocera, the main food source for fish (Gophen, 1978), and the main food source for Cladocera – nanophytoplankton (Gophen, 1973; Serruya *et al.*, 1980), were analysed together. The term nanophytoplankton includes all phytoplanktonic components which pass through a 20 μm mesh net (Sieburth *et al.*, 1978). Most algal species in Lake Kinneret are nanophytoplanktonic and belong to the following classes: Chlorophyta, Bacillariophyta, Pyrrhophyta, Cryptophyta and Cyanophyta (Pollingher, 1986).

The density of nanophytoplankton varies throughout the annual cycle in Lake Kinneret (Pollingher and Berman, 1982), and the relative abundance mainly depends on density of the pyrrhophyte *Peridinium gatunense* (Nygaard). *Peridinium* is a large species with a cell diameter surpassing 40 μm. This species forms dense blooms from February through June, with almost a regular periodicity. During the bloom period, *Peridinium* comprises 90 to 99% of the algal biomass and, on an annual basis, the biomass ranges from 65 to 95% (Berman *et al.*, 1992).

Although throughout the paper we use the term nanophytoplankton, we actually measured chlorophyll *a* in nanophytoplankton as an index of the amount of nanophytoplankton. In order to determine the contribution of nanophytoplankton to the total bulk of chlorophyll *a*, the water samples were filtered through a net of 20 μm. The particles that passed the 20 μm net were collected on to glass fibre. The subsequent procedure was identical to that used for estimation of chlorophyll *a* in the total phytoplankton.

Identification of correlation zones

A method of identifying correlation zones is given in Kalikhman *et al.* (1995), and that of determining the zone significance in Kalikhman (1999). If, at a certain point, the gradients of two fields are co-phased (the angle between them is acute enough), a positive correlation may be expected between the fields in proximity to that point (Figure 2, top). In contrast, if at a certain point, the gradients of two fields are anti-phased (the angle between them is obtuse enough), a negative correlation may be expected between the fields in proximity to that point (Figure 2, bottom). If the gradients are out of phase (the angle between them is close to right), no correlation is expected.

Initially, the range of angles between the gradients (α) indicative for the presence of correlations was chosen constant based upon the results of experimentation (Kalikhman *et al.*, 1995); for example, $0° < \alpha < 90°$ (for positive correlations) or $90° < \alpha < 180°$ (for negative ones). The optimized range will be considered later in section "Optimization of Correlation Zones".

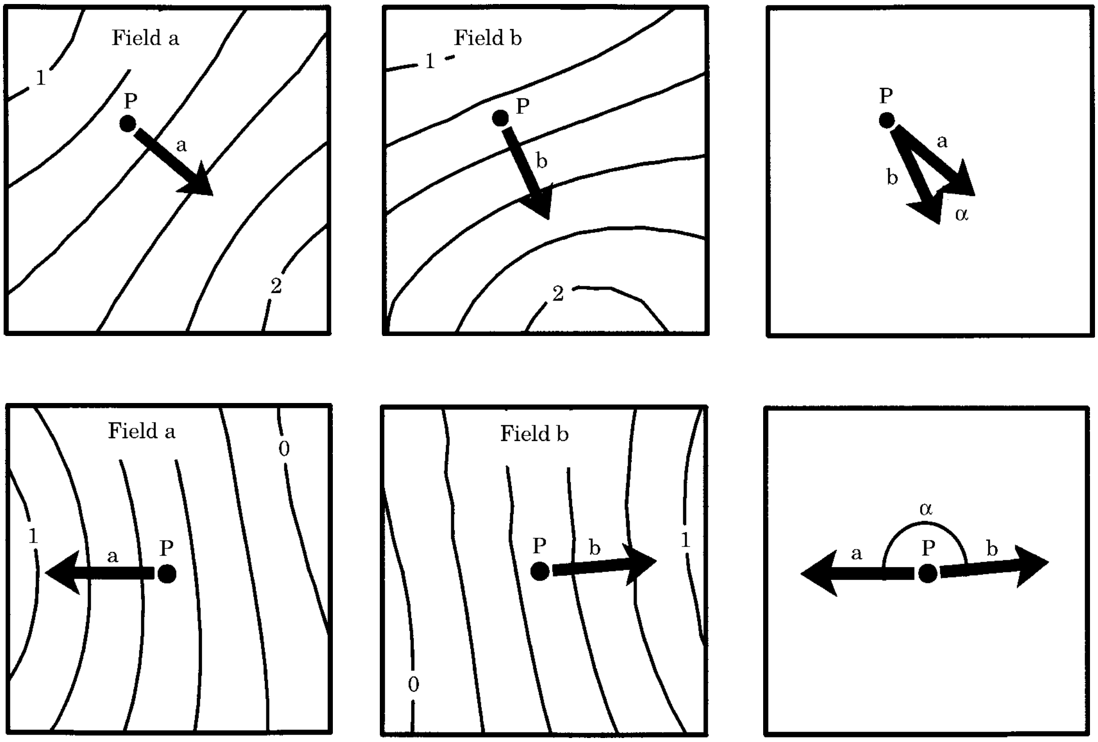

Figure 2. A pair of distribution fields in proximity to point P: top – the co-phased gradients; bottom – the anti-phased gradients.

To grid the data in constructing the distribution fields, the *Kriging* method was used. This algorithm assumes an underlying variogram, which is a measure of how quickly things change on the average (Cressie, 1991). To identify the correlation zones, the two pairs of distribution fields (obtained in December 1991 and May 1993) were sampled at nodes of a regular grid. The maximally admissible distance between nodes was estimated on the basis of the analysis of the autocorrelation for the fields (Sokal and Wartenberg, 1981; Legendre and Fortin, 1989; McArdle and Blackwell, 1989). Based on the analysis, the distance between the nodes was chosen equal to 1 km (Kalikhman *et al.*, 1992, 1995).

For each pair of the fields, the angles between the gradients at the same grid node were calculated (Figure 3a and b, top). If the corresponding gradients are co-phased at a number of adjacent nodes, a positive correlation of the fields may be expected in the area containing these nodes (Figure 3b, top). In contrast, if the corresponding gradients are anti-phased at a number of adjacent nodes, a negative correlation of the fields may be expected in the area containing these nodes (Figure 3a, top). The points where a positive or negative correlation was expected were grouped; the areas containing each group of points were outlined (Figure 3a and b, bottom).

After outlining the potential correlation zones, the correlation values were calculated for each zone. As different variables were measured at the same points, there are two possibilities for calculating the correlations: the first one – at the measurement points; the second one – at the regular grid nodes (Kalikhman, 1999). The distribution fields were sampled at the measurement points and regular grid nodes, located in each zone outlined (it should be emphasized that the original or interpolated field values were taken, not the gradients). Correlations were calculated for the corresponding measured or interpolated field values belonging to each zone (Figure 3a and b, bottom). Calculation of the correlation with or without interpolation use different values of variables (Figure 3a and b, bottom left and right); this fact explains the small difference of correlation values obtained (within the limit of 2.1% for the considered cases). Further on (Figure 4), the lowest correlation values are taken into consideration.

The significance of a correlation zone was estimated on the basis of the Fisher z-criterion according to the correlation value found, the number of measurement points at which the correlation was determined, and the admissible error probability (Sokal and Rohlf, 1969). To estimate the significance of correlation zones, only the number of measurement points can be used (Figure 3a

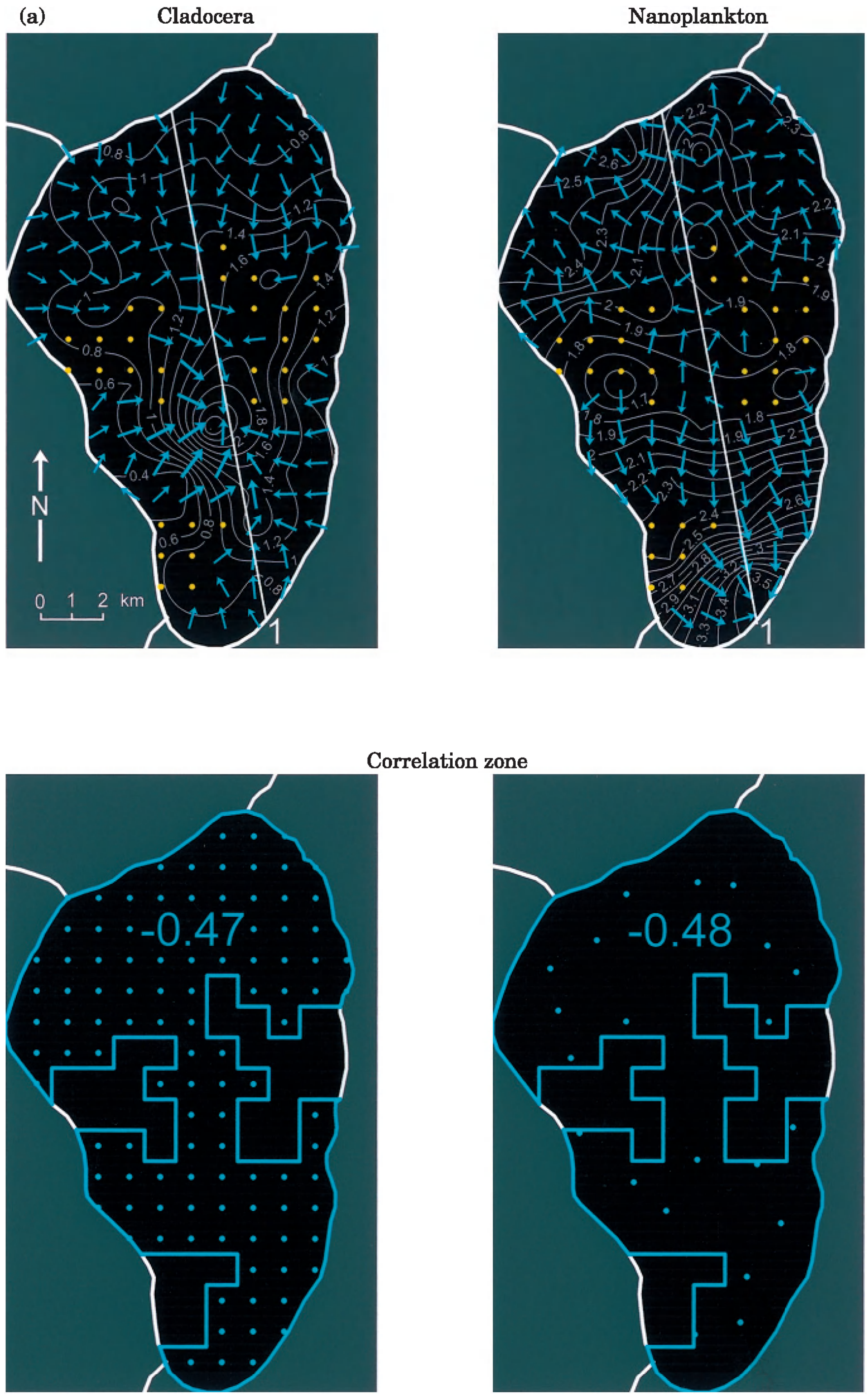

Figure 3(a).

and b, bottom right), even though the correlation was calculated at nodes of a regular grid (Figure 3a and b, bottom left). This is explained by the fact that a regular grid may contain any, as great as one desires, number of such nodes; however, the significance should be independent of this number (Kalikhman, 1999). The Fisher z-criterion can only be applied if the points of measurements are located randomly in space, i.e., they are not

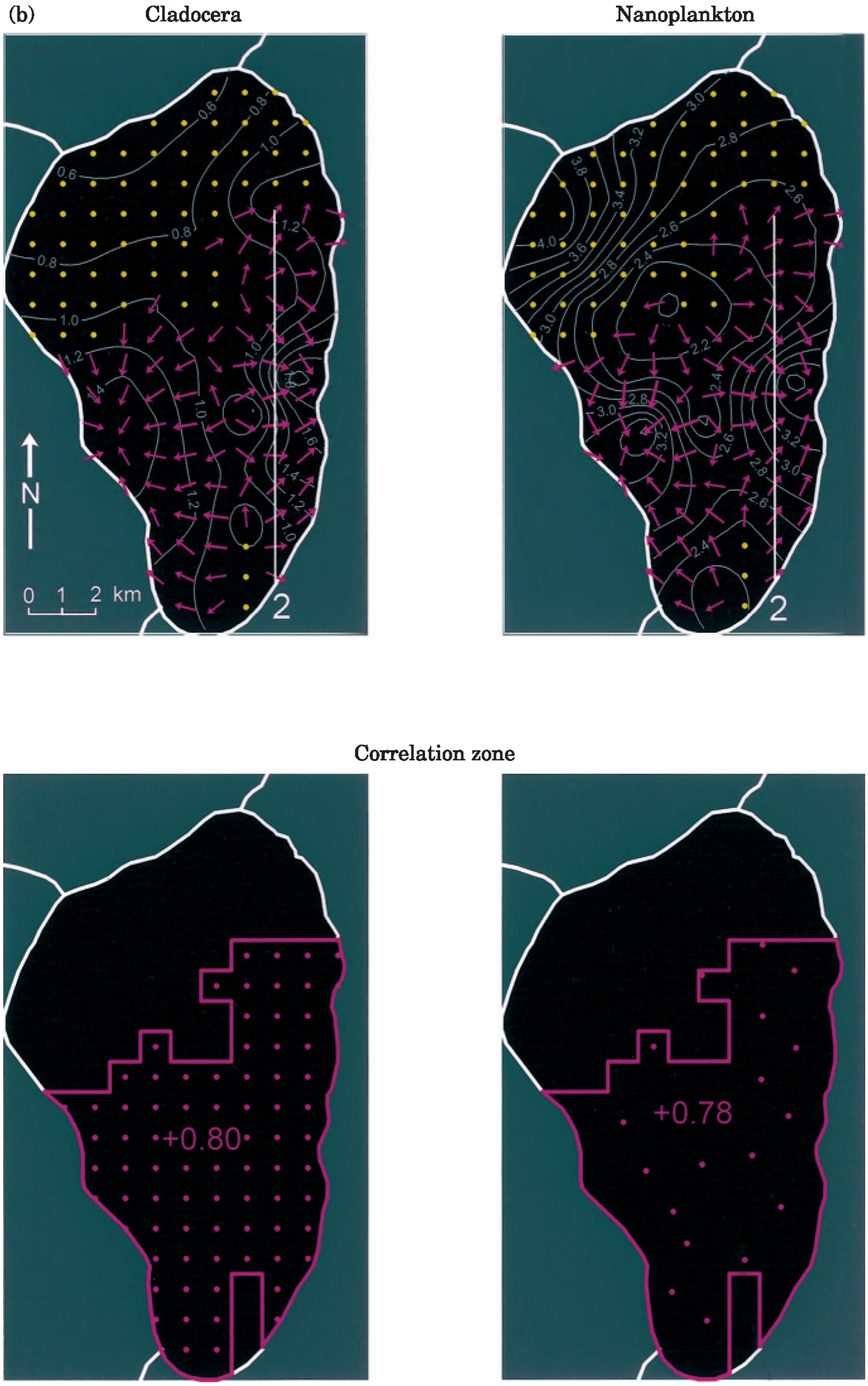

Figure 3(b).

situated on any known configuration (or any part of it). Therefore, the increase (or decrease) in the area of a correlation zone means the corresponding increase (or decrease) of the number of such points within the zone.

Optimization of correlation zones

As mentioned earlier, the aim of the optimization is to obtain the strongest correlation between ecosystem

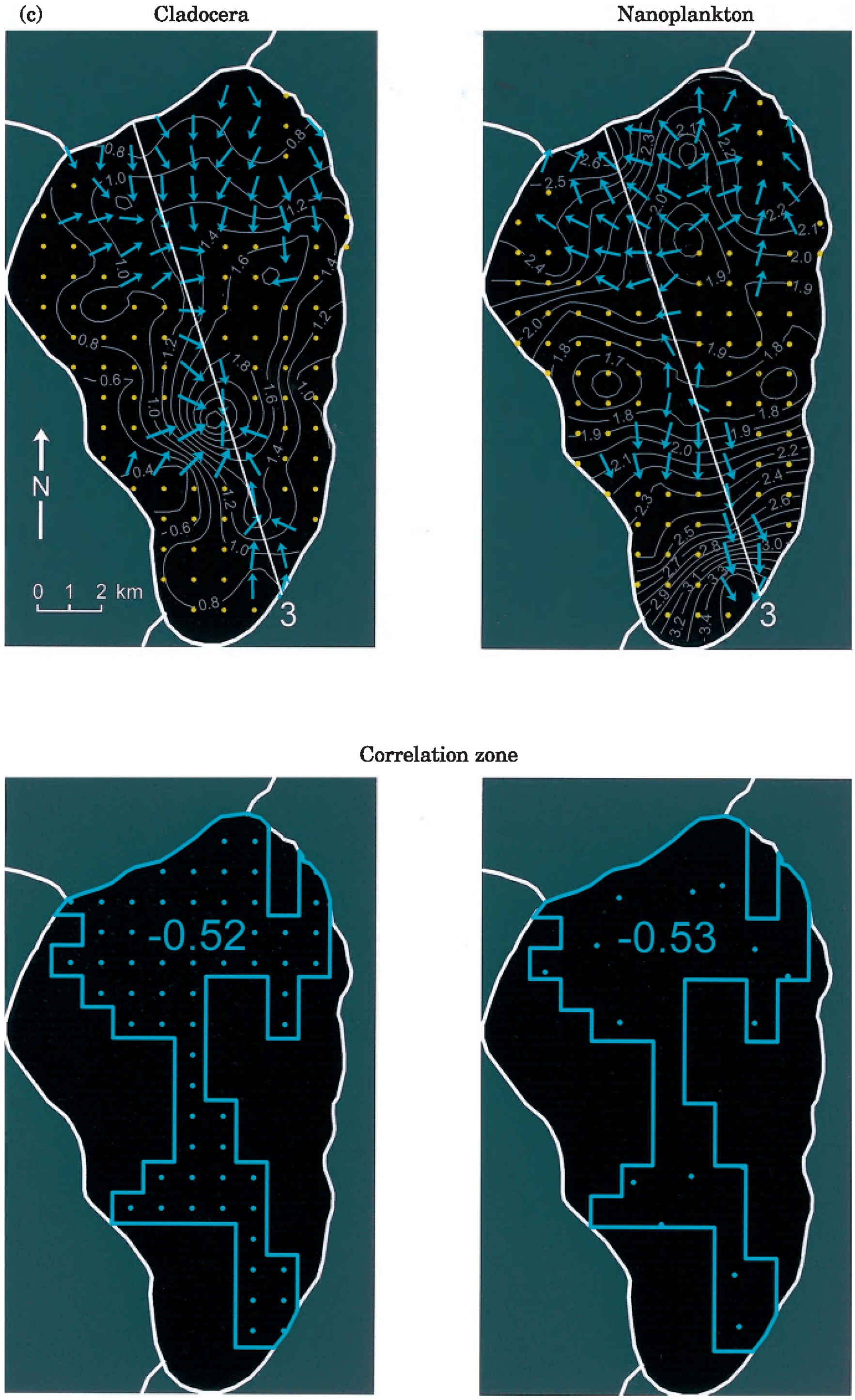

Figure 3(c).

components. Calculations have shown that the following regularities are typical for correlation zones: as a result of narrowing the range of angles between the gradients, the correlation values are increased and the number of measurement points is decreased; with smaller number of points, the curve of critical values of the correlation coefficient rises more abruptly than the correlation value of a zone (Figure 4). Therefore, to obtain the strongest

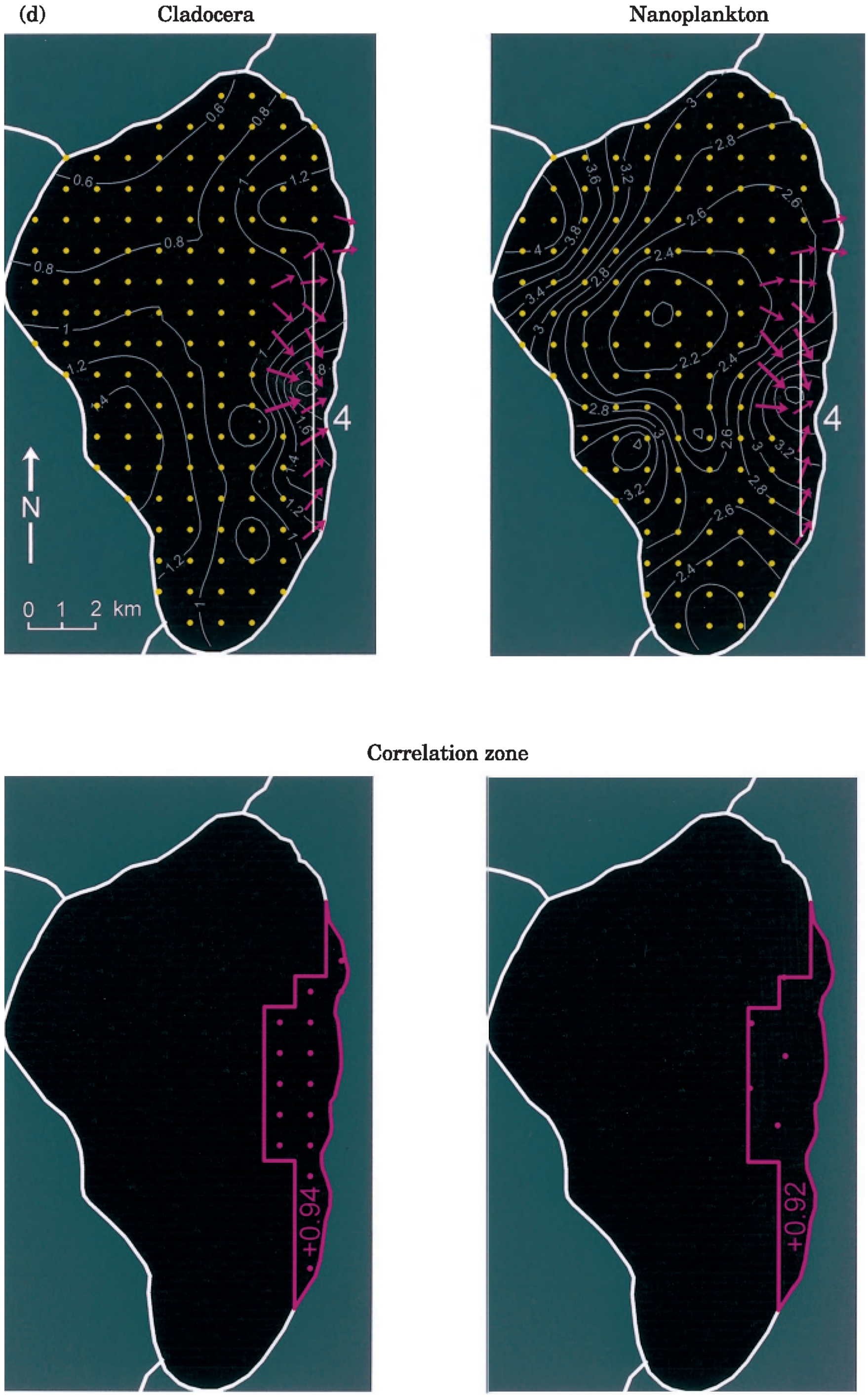

Figure 3(d).

Figure 3. Maps demonstrating the distribution fields and correlation zones. The top panel represents the pairs of distribution fields for Cladocera density (mg/l) and nanophytoplanktonic chlorophyll density (µg/l) on 26 December 1991 (a and c) and 13 May 1993 (b and d). The gradients are directed from minimal to maximal values. The gradients at the same node are considered to indicate the positive correlation (shown by crimson arrows) or the negative one (shown by blue arrows). The range of angles between the gradients (α) is chosen as follows: $90° < \alpha < 180°$ (a); $0° < \alpha < 90°$ (b); $118° < \alpha < 180°$ (c); $0° < \alpha < 35°$ (d). The gradients not meeting the indicated conditions are not shown; instead, the yellow points are given. The bottom panel corresponds to the same distribution fields and represents regular grid nodes (left) or measurement points (right) for sampling the pair of distribution fields to calculate the positive correlation (shown by crimson points) or the negative one (shown by blue points). The correlation zone is displayed on the same panel by the corresponding colour.

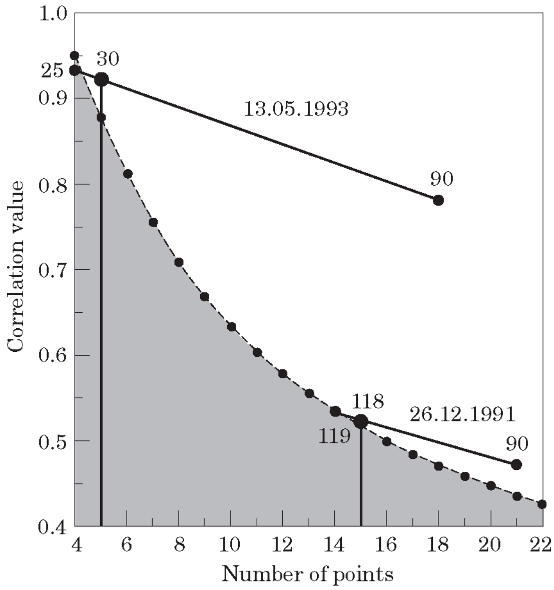

Figure 4. Dependencies of the correlation value and number of measurement points while varying the range of angles between the gradients. The number indicates the limit (for a positive correlation zone – upper; for a negative one – lower) of the range of angles between the gradients. The dashed curve shows the critical values r_α of the correlation coefficient (r is insignificant with $r<r_\alpha$ and significant with $r>r_\alpha$) corresponding to various numbers of measurement points and the error probability p=0.05 (Sokal and Rohlf, 1969).

correlation, it is necessary to find the minimal number of measurement points ensuring the zone significance.

Let us consider the negative correlation zone observed on 26 December 1991. With the range of angles between the gradients 90°<α<180°, 21 measurement points fall into the zone and the correlation value is −0.47 (Figure 3a). When the range of angles is 118°<α<180°, 15 points fall into the zone and the correlation value is −0.52 (Figure 3c). With a further decrease of the range of angles between the gradients (119°<α<180°), the number of points is 14 and the correlation value is −0.53, i.e., already below the curve of the critical values of the correlation coefficient.* Therefore, the correlation zone containing 15 measurement points is taken to be the optimal one (in Figure 4 shown by a large point). As this point is located above the curve of critical values of the correlation coefficient, the correlation zone is significant.

Let us consider the positive correlation zone observed on 13 May 1993. With the range of angles between the gradients 0°<α<90°, 18 measurement points fall into the zone and the correlation value is +0.78 (Figure 3b). When the range of angles is 0°<α<30°, 5 points fall into the zone and the correlation value is +0.92 (Figure 3d). With a further decrease of the range of angles between the gradients (0°<α<25°), the number of points is 4 and the correlation value is +0.94, i.e., already below the curve of the critical values of the correlation coefficient. Therefore, the correlation zone containing 5 measurement points is taken to be the optimal one (in Figure 4 shown by a large point). As this point is located above the curve of critical values of the correlation coefficient, the correlation zone is significant.

*Regarding 15 or 14 points, the critical values of the correlation coefficient are equal respectively to 0.514 or 0.532 (Sokal and Rohlf, 1969).

Ecological application

Cladocera and nanophytoplankton densities in the negative correlation zone, in December 1991 (Figure 3a and c), and those in the positive correlation zone, in May 1993 (Figure 3b and d), were approximately equal to each other. To elucidate the conditions inside the correlation zones in more detail, transects were selected across the negative correlation zone (Figure 3a, transect 1; Figure 3c, transect 3) and the positive one (Figure 3b, transect 2; Figure 3d, transect 4). The correlations of Cladocera and nanophytoplankton densities on transect 1 and 2 were lower than on transect 3 and 4 (Figure 5). As follows from Figure 5, in the case of non-optimized correlation zones, the association between Cladocera and nanophytoplankton densities is less obvious. The result obtained confirms the necessity of optimizing correlation zones.

The concentration range of both variables along these transects was fairly similar and it was within the variability found all over the lake in both surveys (Figure 5). Transects were chosen to maximize the negative and positive correlations between spatial distributions of Cladocera and nanophytoplankton. Consequently, the correlation values along the transects were higher than in the full correlation zones. No positive and negative correlation zones were simultaneously observed.

We assume that the association found between Cladocera and nanophytoplankton densities, in both surveys, reflects the status of phytoplanktonic components in Lake Kinneret in different seasons. In December, algal food supply is much below the zooplankton demand (Gophen *et al.*, 1999). In fact, in December 1991, 79.6% of all chlorophyll originated from nano phytoplankton.* At this time, Cladocera's requirement for food could not be satisfied by the nanophytoplankton alone and should be supplemented by other sources such as protozoa, detritus and bacteria (Serruya *et al.*, 1980). The negative correlation between the distributions of Cladocera and nanophytoplankton is a result of the intensive grazing of a predator on the most abundant food resource available (Figure 5, bottom left). In

*Unpublished data obtained in the present study.

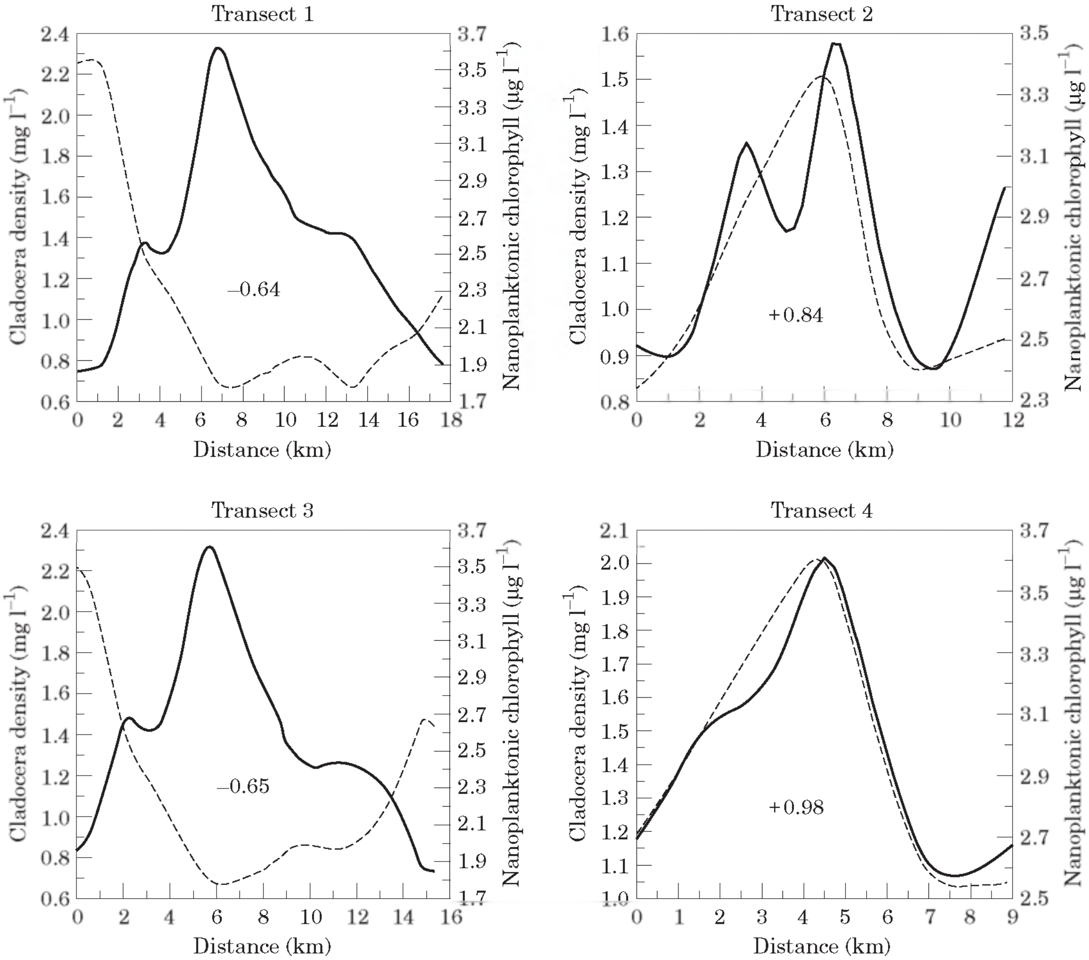

Figure 5. Interpolated values of various distribution fields on numbered transects (1 and 3 – 26 December 1991; 2 and 4 – 13 May 1993): Cladocera density (solid line) and nanophytoplanktonic chlorophyll density (dashed line). The number indicates the correlation on the transect. The location of different transects numbered (1–4) is shown in Figures 3.

contrast, in May, algal food supply is abundant (Gophen *et al.*, 1999). In fact, in May 1993, only 8.5% of all chlorophyll originated from nanophytoplankton.* At that time, most of the phytoplankton was composed of *Peridinium gatunense* which is considerably larger than nanophytoplanktonic algae and cannot be consumed by Cladocera during the *Peridinium* life, but decomposes rapidly within the water column after the cell death (Zohary *et al.*, 1998). The decomposition of *Peridinium* cells may be a stimulant for the development of algae, because of the enrichment of an otherwise nutrient-deficient aquatic environment. Zooplankton feeding on detrital particles is a further factor that enhances nutrient (phosphorus, nitrogen) recycling, and consequently contributes to the augmentation of algal reproduction. Therefore, the positive correlation between Cladocera and nanophytoplankton densities is indicated (Figure 5, bottom right). It is clear that no threshold level was found in the Cladocera-nanophytoplankton interaction, as the ranges of nanophytoplankton densities in both surveys were similar to each other and both components were influenced by other interactions, independent of the variables measured in the present study.

Conclusions

A method of identifying and optimizing correlation zones represents a new mathematical tool for the analysis of interrelationships between patchily distributed variables. Its application to ecological studies allows us to formulate hypotheses on dominant factors controlling the patchiness under various conditions. The example outlined above demonstrates the capability of correlation zone determination to pinpoint interactions

*Unpublished data obtained in the present study.

within an ecosystem. This may be a preliminary technique in an attempt to understand processes occurring in an ecosystem. The method is a powerful tool for analyzing correlations between distribution fields in any systems (e.g., terrestrial, aquatic, atmospheric, etc.), including investigations of relationships between distributions of atmospheric factors, characteristic of soils, abundance of pests, harvest of specific crops, vegetation of plants, composition of rocks or minerals, as well as studies based on satellite information.

Acknowledgements

We wish to thank the two anonymous reviewers for their constructive criticism, which allowed us to greatly improve the manuscript. We are grateful to the Israeli Ministry of Absorption, the Israeli Ministry of Science and Technology, the Rich Foundation, AID and BSF, whose financial support made this study possible.

References

Berman, T., Yacobi, Y. Z., and Pollingher, U. 1992. Lake Kinneret phytoplankton: stability and variability during twenty years (1970–1989). Aquatic Sciences, 54: 104–127.

Cressie, N. A. C. 1991. Statistics for Spatial Data. John Wiley & Sons, Inc., New York. 900 pp.

Ellison, A. M., Mukherjee, B. B., and Karim, B. 2000. Testing patterns of zonation in mangroves: scale dependence and environmental correlates in the Sundarbans of Bangladesh. Journal of Ecology, 88(5): 813–824.

Gophen, M. 1973. Zooplankton in Lake Kinneret. *In* Israel Scientific Conference on the Interaction between Hydrographic, Physical and Biological Factors in Warm Lakes, pp. 61–67. Ed. by T. Berman. Israel Council for Research and Development, Jerusalem.

Gophen, M. 1978. Zooplankton. *In* Lake Kinneret, pp. 299–311. Ed. by C. Serruya. Monographiae Biologicae, Junk, The Hague.

Gophen, M., Walline, P., Ostrovsky, I., Azoulay, B., and Easton, J. 1999. Water quality and fishery management in Lake Kinneret, Israel. *In* Theoretical Reservoir Ecology and its Applications, pp. 493–503. Ed. by J. G. Tundidsi, and M. Straskraba. International Institute of Ecology, Brazilian Academy of Sciences and Backhuys Publishers.

Kalikhman, I., Walline, P., and Gophen, M. 1992. Simultaneous patterns of temperature, oxygen, zooplankton and fish distribution in Lake Kinneret, Israel. Freshwater Biology, 28: 337–347.

Kalikhman, I., Ostrovsky, I., Walline, P., Gophen, M., and Yacobi, Y. Z. 1995. Distribution fields for aquatic ecosystem components: method of identification of correlation zones. Freshwater Biology, 34: 317–328.

Kalikhman, I. 1999. Distribution fields for aquatic ecosystem components: determination of significance of correlation zones. Hydrobiologia, 400: 1–11.

Legendre, P., and Fortin, M. J. 1989. Spatial pattern and ecological analysis. Vegetatio, 80: 107–138.

McArdle, B.H., and Blackwell, R. G. 1989. Measurement of density variability in the bivalve *Chione stutcchburyi* using spatial autocorrelation. Marine Ecology Progress Series, 52: 245–252.

Nikolsky, G. V. 1963. The Ecology of Fishes. Academic Press, London. Translated from the Russian by L. Birkett.

Pollingher, U. 1986. Phytoplankton periodicity in a subtropical lake (Lake Kinneret, Israel). Hydrobiologia, 138: 127–138.

Pollingher, U., and Berman, T. 1982. Relative contributions of net and nanophytoplankton to primary production in Lake Kinneret. Archiv fur Hydrobiologie, 96: 33–46.

Serruya, C., Gophen, M., and Pollinhger, U. 1980. Lake Kinneret: carbon flow patterns and ecosystem management. Archiv fur Hydrobiologie, 88: 265–302.

Sieburth, M. N., Smetacek, V., and Lenz, J. 1978. Pelagic ecosystems: heterotrophic compartments of plankton and their relationship to plankton size fractions. Limnology and Oceanography, 23: 1256–1263.

Sokal, R. R., and Rohlf, F. J. 1969. Biometry. The Principles and Practice of Statistics in Biological Research. W. H. Freeman and Company, San Francisco.

Sokal, R. R., and Wartenberg, D. E. 1981. Space and population structure. *In* Dynamic Spatial Model, pp. 186–213. Ed. by D. A. Griffith, and R. D. Mackinnon. Alphen aan den Rijn, Sijthoff, Netherlands.

Steele, J. H. 1976. Patchiness. *In* The Ecology of the Seas, pp. 98–115. Ed. by D. H. Cushing, and J. J. Walsh. Blackwell Scientific Publications, London.

Zohary, T., Pollingher, U., Hadas, O., and Hambright, K. D. 1998. Bloom dynamics and sedimentation of Peridinium gatunense in Lake Kinneret. Limnology and Oceanography, 43: 175–186.