ELSEVIER

Ecological Modelling 87 (1996) 285–294

ECOLOGICAL MODELLING

Modeling spatial distributional patterns of benthic meiofauna species by Thomas and related processes

D. Pfeifer [a,*], H.-P. Bäumer [b], H. Ortleb [c], G. Sach [d], U. Schleier [c]

[a] *Institut für Mathematische Stochastik, Universität Hamburg, Bundesstr. 55, D-20146 Hamburg, Germany*
[b] *Carl von Ossietzky Universität Oldenburg, HRZ – Angewandte Statistik, Postfach 25 03, D-26111 Oldenburg, Germany*
[c] *Carl von Ossietzky Universität Oldenburg, Fachbereich Mathematik, Postfach 25 03, D-26111 Oldenburg, Germany*
[d] *Christian-Albrechts-Universität Kiel, Zoologisches Institut, Olshausenstr. 40, D-24098 Kiel, Germany*

Received 8 February 1995; accepted 2 October 1995

Abstract

Spatial distributional patterns of benthic meiofauna in the presence of certain macrofauna species are of particular interest in the ecological research of tidal flats. The promotion of benthic meiofauna by macrofauna species is frequently observed in field experiments and is in part responsible for the spatial patchiness of such distributions. It is the aim of this paper to show how Thomas and related cluster processes can be used for a stochastic modeling of such patterns when only aggregate abundances of individuals from quadrat counts are available. In particular, three data samples of *Harpacticus obscurus* (Crustacea: Copepoda) from sites where the polychaete *Lanice conchilega* is predominant are analyzed.

Keywords: Cluster analysis; Marine ecosystems; Spatial patterns; Zooplankton

1. Introduction

Promotion of benthic meiofauna by macrofauna species is a frequently observed phenomenon in marine sediments and has been studied in various field experiments by several authors (see Reise, 1985 and references therein). A series of such experiments was recently carried out within the research project "Ökosystemforschung Niedersächsisches Wattenmeer" near the island of Spiekeroog in areas with a high spatial density of the polychaete species *Lanice conchilega*. Sediment samples were taken by means of a multiple-tube corer consisting of 25 square tubes each measuring 4 cm^2 (thus covering an area of 100 cm^2). Within three samples a high frequency of the copepod *Harpacticus obscurus* was observed. Table 1 contains the observed abundances within each of the individual tubes (below: sample average μ and index-of-dispersion D (cf. Diggle, 1983, Chapter 2.5; Richter and Söndgerath, 1990, Chapter 3.1.5.1).

The outstanding values of the index-of-dispersion obtained for each sample show clearly that an extremely high degree of spatial patchiness is present in each of the locations. Correspondingly, the index-of-dispersion test rejects the (null) hypothesis of a random spatial distribution for this species at any reasonable level of significance, for

* Corresponding author. Fax: (+49-40) 4123 4924.

all sites. General biological explanations for this kind of aggregation have recently been investigated by Ekschmitt (1993). In the context of benthic fauna, Reise (1985, Chapter 11.2) discusses similar observations around the feeding pockets of the lugworm *Arenicola marina* (involving nematodes and further meiofauna species), among others.

For the purpose of mathematical modeling, the use of Thomas and related processes seems to be an appropriate tool here (see Diggle, 1981, p. 54; Richter and Söndgerath, 1990, p. 65ff; Stoyan and Stoyan, 1992, p. 337; Pfeifer et al., 1992). Recall that a Thomas process is a two-phase stochastic point process where the so-called *parent points* are generated according to a homogeneous spatial Poisson process with intensity $\lambda_P > 0$, say, and the so-called *daughter points* according to non-homogeneous spatial Poisson processes clustering around each of the parent points with intensity $\lambda_D > 0$, say. The spatial distribution of individual daughter points is assumed to be bivariate radially symmetric normal with variance $\sigma^2 > 0$ for each component.

In the situation above the parent process corresponds to the locations of *Lanice conchilega*, while the daughter processes correspond to the locations of *Harpacticus obscurus*.

The main difficulty here is of course the fact that due to the sampling device the locations of the daughter individuals are not precisely known. Thus neither the shape nor the number of clusters can be seen immediately from the data. Therefore a reliable statistical analysis is very cumbersome unless more specific tools are developed for the particular situation under consideration. This is the aim of the subsequent section.

2. The statistical approach

Unfortunately there is only a very limited number of papers dealing with the statistical analysis of quadrat count data for Thomas-like processes; see, e.g., Pielou (1957) or Gleeson and Douglas (1975). Both papers essentially use the method of *empty quadrats* for estimating the intensity of the underlying Poisson parent process; a procedure which is usually extremely biased. The authors thus come to the conclusion that the method is not appropriate in many cases, especially when the daughter process has a larger variability so that points are spreading out in quadrats which were originally free of parent points. Moreover, the method is not applicable at all if there are no empty quadrats, as is, e.g., the case of the data from site 18.

In this paper we propose to proceed in two steps: Firstly, to estimate the number L of quadrats free of parent individuals by comparing the number of daughter individuals for each quadrat with a suitable threshold S by

$$\hat{L} = \text{number quadrats with at most } S \text{ daughter individuals;} \tag{1}$$

secondly, to estimate the intensity λ_P of the parent process by

$$\hat{\lambda}_P = -\frac{n}{m(A)} \ln\left(\frac{\hat{L}}{n}\right) \tag{2}$$

Table 1
Observed abundances of *Harpacticus obscurus* in areas of high density of *Lanice conchilega*, near Spiekeroog (May 1992)

Site 6					Site 8					Site 18				
95	1	3	0	42	165	22	1	94	68	245	142	326	52	293
1	1	1	4	2	11	82	111	97	153	222	368	84	18	67
0	5	8	81	24	0	0	24	13	15	239	25	477	213	204
11	1	6	71	116	31	1	46	22	11	18	570	183	494	47
1	5	116	2	10	2	0	5	8	6	238	119	126	591	20
$\mu = 24.28$					$\mu = 39.52$					$\mu = 215.24$				
$D = 1464.21$					$D = 1490.13$					$D = 3377.47$				

where n is the total number of quadrats and $m(A)$ is the total area covered by the sampling device; and thirdly, to estimate the intensity λ_D of the daughter Poisson process by

$$\hat{\lambda}_D = \frac{N}{\hat{\lambda}_P}, \tag{3}$$

where N is the total number of daughter individuals. In the sequel we call quadrats with less than S daughter points *weakly occupied*. Such quadrats are considered to be free of parent points with high probability, replacing the notion of "empty quadrats" in a suitable way for estimation purposes.

The motivation for Eqs. 2 and 3 is as follows. In the Poisson case, the probability p_0 of a quadrat to contain no parent individual is given by

$$p_0 = e^{-\lambda_0} = \exp\left(-\lambda_P \frac{m(A)}{n}\right),$$

where $\lambda_0 = \lambda_P m(A)/n$ is the expected number of points in an area of size $m(A)/n$, i.e. the size of each of the n sampling quadrats. On the other hand, L/n represents the empirical fraction of empty quadrats which by the law of large numbers is approximately equal to p_0, if the total number of quadrats, n, is large enough. Hence $\hat{L}/n$ is a statistical estimate for p_0; solving the resulting equation

$$\hat{L}/n = \exp\left(-\hat{\lambda}_P \frac{m(A)}{n}\right)$$

for $\hat{\lambda}_P$ gives Eq. 2 (see also Stoyan et al., 1989, Chapter 2.7, p. 54). Further, for sufficiently large observation area sizes $m(A)$, by the independence of the number of daughter and parent points, we have for the expectation $E(N)$ of the total number N of daughter individuals

$$E(N) \approx \lambda_P \lambda_D m(A).$$

This is due to the fact that $\lambda_P \lambda_D$ corresponds to the theoretical intensity of the point process consisting of *all* daughter points in the plane, per unit area. Replacing λ_P by its estimate $\hat{\lambda}_P$ and solving for λ_D gives Eq. 3.

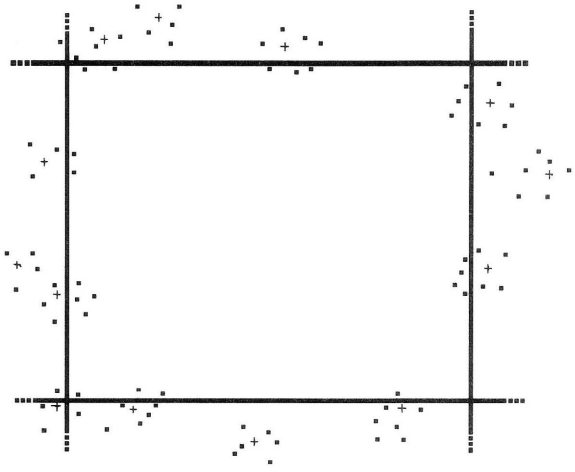


Fig. 1. Graph of a weakly occupied quadrat, surrounded by clusters of daughter points. Here + denotes the position of parent points while ■ corresponds to the location of daughter points.

In order to choose an appropriate threshold S we suggest to use

$$S = \frac{N}{3(n-l)}$$

where again N is the total number of daughter individuals and l denotes the number of quadrats being free of *daughter* points. This recommendation is based on the following heuristic arguments.

Considering first a parent point being "close" to the edge of a quadrat which was originally empty, we see that about one half of the daughter points is spreading out into this quadrat, if the projection of the point to the quadrat boundary is roughly in the middle between the two adjacent vertices (Fig. 1). If the point is located "close" to a vertex itself, only about one quarter of daughter points is spreading out into the quadrat. Due to the homogeneity of the parent Poisson process, we see that, averaging over the length of an edge between vertices, about one third of the daughter points spreads out into the quadrat, if the centering parent point is "close" to the boundary of the quadrat. On the other hand, $n-l$ is the total number of quadrats being occupied by daughter points; on average we thus have roughly $N/(n-l)$

daughter points in such a quadrat. This explains Eq. 4. Note that there is no simple closed formula available for the exact average number of daughter points spreading out into a weakly occupied quadrat.

3. Adaptive bias reduction and simulations

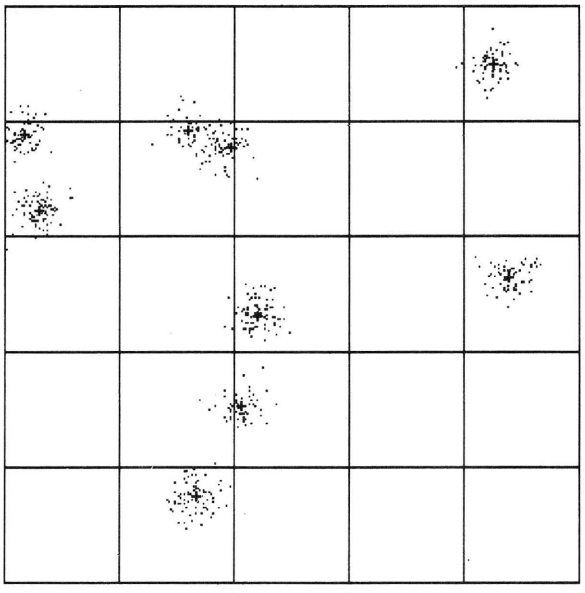
Although the threshold approach provides a certain bias reduction over the methods proposed in Pielou (1957) or Gleeson and Douglas (1975), the resulting estimates in Eqs. 2 and 3 are still not completely satisfactory. We therefore suggest to process the obtained estimates through the following adaptive bias reduction procedure:

1. simulate a certain number of spatial patterns with input parameters given by the estimates of Eqs. 2 and 3 from the original data;
2. compare the mean values of the corresponding parameter estimates obtained by simulation with the input parameters and adjust by appropriate scaling;
3. repeat the procedure until a sufficient result is obtained.

Table 2 contains the result of a three-fold bias reduction from 1000 simulations each, with respect to the parent intensity λ_P, whereby the total area size of 100 cm^2 was for practical reasons set to 1 [unit]. The standard deviation σ for the bivariate normal distribution was chosen to be 2.33 mm; this is on the one hand a reasonable

Table 2
Simulation results from bias reduction

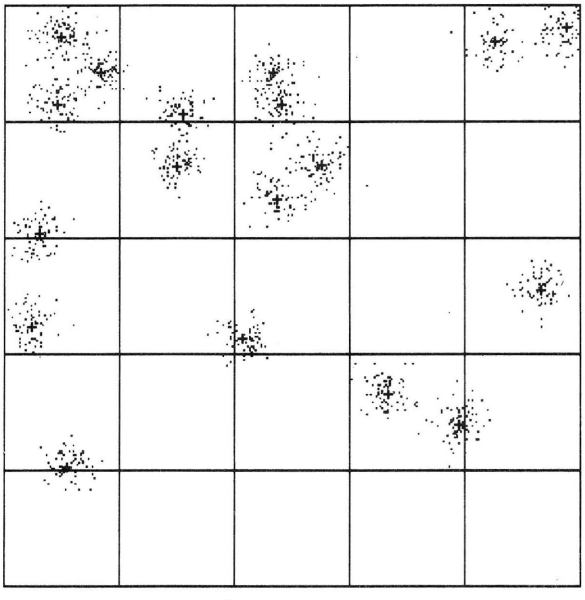
Simulation No.	Parameter	Site 6	Site 8	Site 18
1	input $\hat{\lambda}_P$	11.15	18.34	31.82
	input $\hat{\lambda}_D$	54.40	53.84	169.08
	D	1464.21	1490.13	3377.47
	μ	24.28	39.52	215.24
Results	average $\hat{\lambda}_P$	14.14	22.66	36.56
	st. deviation	4.87	6.79	9.49
	average $\hat{\lambda}_D$	43.30	43.33	145.52
	st. deviation	9.42	9.04	30.99
	average D	1025.70	1010.45	3086.76
	st. deviation	298.83	290.79	878.43
	f	0.7885	0.8093	0.8703
2	input $\hat{\lambda}_P$	8.79	14.84	27.69
	input $\hat{\lambda}_D$	69.05	66.57	194.30
Results	average $\hat{\lambda}_P$	11.64	18.80	32.82
	st. deviation	4.59	6.00	8.45
	average $\hat{\lambda}_D$	53.11	52.25	163.46
	st. deviation	12.66	11.04	33.29
	average D	1298.86	1246.32	3576.84
	st. deviation	385.89	362.46	1034.43
	f	0.9579	0.9755	0.9695
3	input $\hat{\lambda}_P$	**8.42**	**14.47**	**26.85**
	input $\hat{\lambda}_D$	**72.09**	**68.28**	**200.41**
Results	average $\hat{\lambda}_P$	11.04	18.32	31.77
	st. deviation	4.39	5.74	8.36
	average $\hat{\lambda}_D$	54.92	53.94	168.26
	st. deviation	12.93	11.68	34.84
	average D	1339.28	1286.62	3725.20
	st. deviation	390.27	378.63	1058.18



9	10	0	2	67
129	89	20	0	0
2	3	77	0	64
0	15	48	0	0
0	75	0	0	0

Fig. 2. Simulation for site 6. $\hat{\lambda}_P = 11.15$; $\hat{\lambda}_D = 54.67$; $D = 1350.16$. Position of parent points, +; position of daughter points, ■.

magnitude from biological considerations, on the other hand this choice provides dispersion indices which are in the order of magnitude of the corresponding dispersion indices for the original data set. The reduction factor f is the ratio between the original input value for λ_P estimated from the data set, and the resulting empirical average of simulated estimates for λ_P. With each repetition of the procedure, the new input value for λ_P was obtained from the preceding one by multiplica-



229	53	113	1	117
47	85	137	1	0
95	14	45	0	64
47	4	4	118	29
18	0	0	0	0

Fig. 3. Simulation for site 8. $\hat{\lambda}_P = 18.34$; $\hat{\lambda}_D = 66.54$; $D = 1698.43$. Position of parent points, +; position of daughter points, ■.

Table 3
Estimates from bias reduction

Parameter	Site 6	Site 8	Site 18
$\hat{\lambda}_P$	8.42	14.47	26.85
$\hat{\lambda}_D$	72.09	68.28	200.41

tion with the actual value of f, while the input value for the daughter intensity was obtained via Eq. 3, with N being counted from the original data set.

Additionally, also empirical averages of the simulated index-of-dispersion and of the simulated estimates for λ_D are given in Table 2, together with their empirical standard deviations.

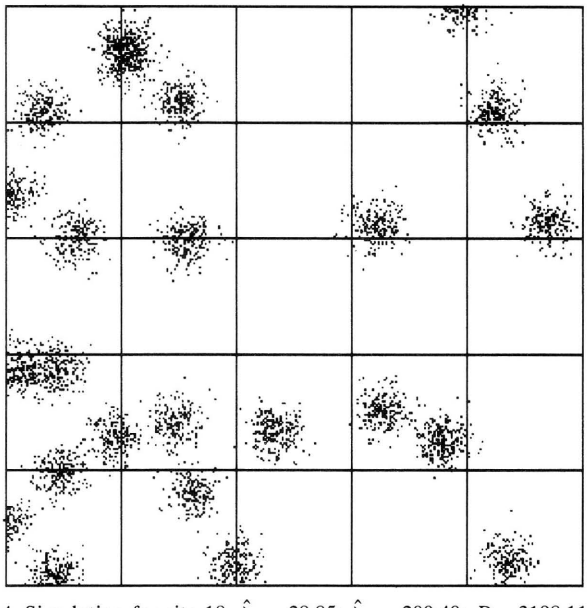
A comparison of the underlined values in Table 2 shows that the third simulation already provides a very good coincidence between the original input parameters and simulated estimates, so that (cf. the bold values above) the values in Table 3 can be regarded as appropriate estimates for λ_P and λ_D. Note that the empirical standard deviations for $\hat{\lambda}_P$ cannot essentially be improved by any other method since the number of parent points in the whole area A is Poisson-distributed with mean $\lambda_P m(A) = \lambda_P$ here, and standard deviation $\sqrt{\lambda_P}$; thus, with the estimates from the original data set, $\sqrt{\hat{\lambda}_P} = 3.33$, 4.28 and 5.64, respectively, come close to the theoretically minimal standard deviation.

4. Visualization of simulation results

The above three figures (Fig. 2, Fig. 3 and 4) show visualizations of simulations for Thomas processes with the input parameters from Table 3. On the left hand side, the sampling quadrats are shown, together with parent points (+) and daughter points (■). On the right hand side, the corresponding quadrat counts for daughter points are given. In the legend of each figure, the resulting estimations for λ_P and λ_D from the particular picture are shown, as well as the corresponding index-of-dispersion D.

5. Adjustments of the model for benthic meiofauna

A closer look to the results obtained above shows that the model reproduces the structure of



302	457	0	55	202
288	109	8	161	203
159	110	1	41	31
526	262	215	431	11
382	264	85	2	182

Fig. 4. Simulation for site 18. $\hat{\lambda}_P = 28.85$; $\hat{\lambda}_D = 200.40$; $D = 3198.11$. Position of parent points, +; position of daughter points, ■.

the given data quite well (in a statistical sense) except for the fact that seemingly too many empty quadrats occur in the simulations. This is most probably due to the fact that individuals of copepoda species are not exclusively clustering around meiofauna funnels but are also moving in the sediment. If a one-sided standard Wilcoxon rank sum test is performed with the original data set and the count results from the above three simulations to test whether the original data come from a stochastically larger distribution compared with the model under consideration, we obtain the following result (Table 4).

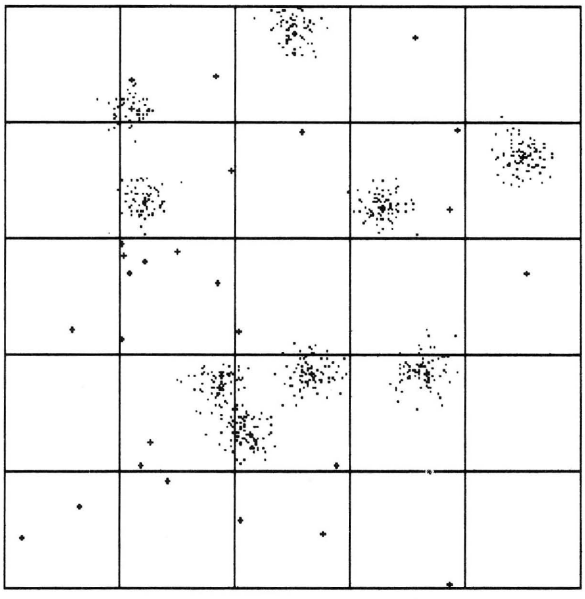
Table 4
Wilcoxon test for the standard Thomas process

Location	Site 6	Site 8	Site 18
Rank sum	693	633	682
Approx. critical value, $\alpha = 0.05$	723	723	723
Approx. critical value, $\alpha = 0.10$	704	704	704
Approx. critical value, $\alpha = 0.20$	681	681	681

We see that at a significance level of 80%, the null hypothesis of equal quadrat count distributions is rejected, at least for sites 6 and 18. The approximate critical values were computed using the normal approximation for the rank sum, having mean 637.5 and variance 2656.25 under the null hypothesis (see e.g. Hartung et al., 1993, p. 514ff). Note that the test procedure is applicable since by the construction principle of Thomas's processes, given the parent points, the pattern of

Table 5
Modified simulation results from bias reduction

Simulation No.	Parameter	Site 6	Site 8	Site 18
1	input $\hat{\lambda}_P$	11.15	18.34	31.82
	input $\hat{\lambda}_D$	51.68	51.15	160.63
	D	1464.21	1490.13	3377.47
	μ	24.28	39.52	215.24
Results	average $\hat{\lambda}_P$	15.49	24.40	37.89
	st. deviation	5.21	7.01	9.40
	average $\hat{\lambda}_D$	36.79	38.33	134.39
	st. deviation	8.27	7.72	27.36
	average D	926.87	908.25	2859.13
	st. deviation	274.16	263.88	801.62
	f	0.7198	0.7516	0.8398
2	input $\hat{\lambda}_P$	8.02	13.78	26.72
	input $\hat{\lambda}_D$	71.85	68.09	191.30
Results	average $\hat{\lambda}_P$	11.97	19.28	33.32
	st. deviation	4.40	5.60	8.61
	average $\hat{\lambda}_D$	49.02	48.96	153.55
	st. deviation	12.24	10.03	329.92
	average D	1287.58	1216.87	3371.52
	st. deviation	392.12	362.01	971.87
	f	0.9315	0.9512	0.9550
3	input $\hat{\lambda}_P$	**7.47**	**13.11**	**25.52**
	input $\hat{\lambda}_D$	**77.19**	**71.59**	**200.33**
Results	average $\hat{\lambda}_P$	11.27	18.33	32.36
	st. deviation	4.15	5.56	8.30
	average $\hat{\lambda}_D$	50.95	50.69	157.83
	st. deviation	13.10	10.80	31.12
	average D	1337.67	1263.11	3471.29
	st. deviation	377.96	367.33	1010.16

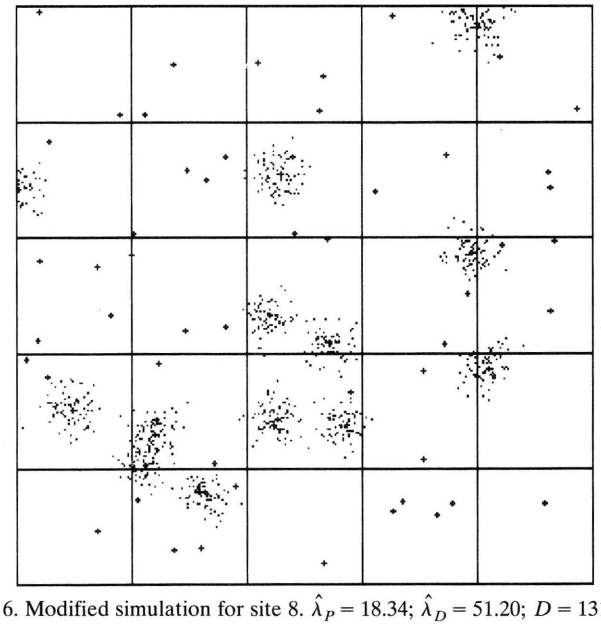


8	40	72	1	1
4	69	2	81	84
1	7	12	10	1
0	78	163	73	0
2	1	2	1	0

Fig. 5. Modified simulation for site 6. $\hat{\lambda}_P = 11.15$; $\hat{\lambda}_D = 60.70$; $D = 1514.73$. Position of noise points, +; position of daughter points, ■.

daughter points forms an inhomogeneous Poisson process, hence in each simulation, the individual quadrat counts are (conditionally) independent.

A seemingly better description of the situation is obtained if we assume that a fixed proportion p of the individuals is scattered *randomly* in the



2	2	3	36	46
41	6	80	6	5
4	4	119	45	36
94	113	169	27	47
4	94	1	4	1

Fig. 6. Modified simulation for site 8. $\hat{\lambda}_P = 18.34$; $\hat{\lambda}_D = 51.20$; $D = 1313.70$. Position of noise points, +; position of daughter points, ■.

sediment. This means that the Thomas process is superimposed by an independent homogeneous Poisson noise process of low intensity $\lambda_S = p\lambda_{tot}$, where

$$\lambda_{tot} = \lambda_P \lambda_D + \lambda_S \tag{5}$$

denotes the overall intensity of the process ($\approx$ mean number of individuals per unit square, up to edge effects). From this formula we obtain $\lambda_P \lambda_D = (1-p)\lambda_{tot}$, so that the estimation for λ_D should be based on the quantity $(1-p)N$ instead of N (cf. Eq. 3):

$$\hat{\lambda}_{D,mod} = \frac{(1-p)N}{\hat{\lambda}_P} = (1-p)\hat{\lambda}_D \tag{6}$$

Likewise, the threshold value S should be increased to

$$S_{mod} = S/(1-p). \tag{7}$$

Similarly as above, an adaptive bias reduction is possible. Table 5 contains computations analogous to those in Table 2. The resulting parameter estimates were again used for a visualization of three simulations with the above modifications (Fig. 5, Fig. 6 and 7); for the *Harpacticus obscurus* data, $p = 0.05$ turned out to be a good choice for the noise process.

Table 6
Wilcoxon test for the modified Thomas process

Location	Site 6	Site 8	Site 18
Rank sum	643	629	613
Approx. critical value, $\alpha = 0.05$	723	723	723
Approx. critical value, $\alpha = 0.10$	704	704	704
Approx. critical value, $\alpha = 0.20$	681	681	681

The Wilcoxon rank sum test applied to the latter simulated count data here provides the following results (Table 6). In this case, the hypothesis of equal quadrat count distributions cannot be rejected at significance levels of 80% or higher.

6. Conclusions

The above simulation studies show that the modified Thomas process model is in very good coincidence with the data sets of sites 6, 8, and 18. To find evidence for the conjecture that the

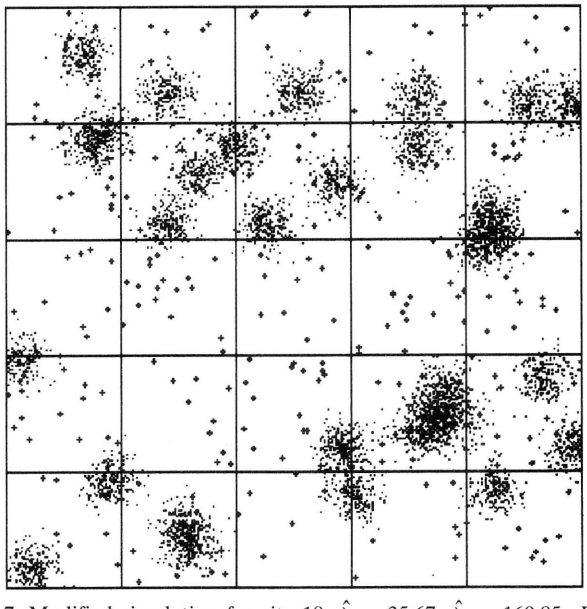


Fig. 7. Modified simulation for site 18. $\hat{\lambda}_P = 35.67$; $\hat{\lambda}_D = 169.85$; $D = 3406.58$. Position of noise points, +; position of daughter points, ■.

Table 7
Table of estimated and true densities for locations 6, 8 and 18

Location	Site 6	Site 8	Site 18
$\hat{\lambda}_P$	7.47	13.11	25.52
d	9.00	15.00	24.00

individuals of the copepod species *Harpacticus obscurus* are spatially clustering around *Lanice conchilega* funnels (at least in the observed area) we have to compare the estimated intensities $\hat{\lambda}_P$ for the parent process with density estimates d for *Lanice conchilega* funnels. Such data were fortunately independently obtained within the same regions of observation. Table 7 shows the six corresponding figures, with d denoting the number of funnels per 100 cm^2 each.

A comparison of the corresponding values indicates very clearly that the conjecture of the promotion of meiofauna by benthic macrofauna can be assured locally for the above-mentioned species.

Programs for the estimation of parameters, simulation and adaptive bias reduction for quadrat counts in a Thomas process of the above type are available upon request from the authors.

Acknowledgements

This research was carried out during phase B of the *Ökosystemforschung Niedersächsisches Wattenmeer* (ELAWAT) financed by the Bundesministerium für Wissenschaft und Forschung, Förderkennzeichen 03F0112A. The contribution of H. Ortleb was, in particular, financed by the Forschungszentrum TERRAMARE e.V., Wilhelmshaven.

References

Diggle, P.J., 1983. Statistical Analysis of Spatial Point Processes. Academic Press, New York.

Ekschmitt, K., 1993. Über die räumliche Verteilung von Bodentieren. Zur ökologischen Interpretation der Aggregation und zur Probenstatistik. Dissertation, Universität Bremen.

Gleeson, A.C. and Douglas, J.B., 1975. Quadrat sampling and the estimation of Neyman type A and Thomas distributional parameters. Aust. J. Stat., 17: 103–113.

Hartung, J., Elpelt, B. and Klösener, K.-H., 1993. Statistik. Lehr- und Handbuch der angewandten Statistik, 9th ed. Oldenbourg, Munich.

Pfeifer, D., Bäumer, H.-P. and Albrecht, M., 1992. Spatial point processes and their applications to biology and ecology. Model. Geo-Biosphere Processes, 1: 145–161.

Pielou, E.C., 1957. The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas's distributions. J. Ecol., 45: 31–47.

Reise, K., 1985. Tidal Flat Ecology. An Experimental Approach to Species Interactions. Ecological Studies 54. Springer, New York.

Richter, O. and Söndgerath, D., 1990. Parameter Estimation in Ecology. The Link between Data and Models. VCH, Weinheim.

Stoyan, D. and Stoyan, H., 1992. Fraktale – Formen – Punktfelder. Methoden der Geometrie-Statistik. Akademie-Verlag, Berlin.

Stoyan, D., Kendall, W.S. and Mecke, J., 1989. Stochastic Geometry and Its Applications. Wiley, New York.