

Beiträge der Angewandten Statistik zur Bearbeitung von Maßstabsfragen und zur Versuchsplanung für die Untersuchung räumlicher Strukturen und dynamischer Vorgänge im Watt

Projekt B3:

Dietmar Pfeifer¹, Hans-Peter Bäumer² und Matthias Albrecht¹

Carl von Ossietzky Universität Oldenburg, Fachbereich Mathematik¹ und HRZ-Angewandte Statistik²

Projektziel

In der Pilotphase der Ökosystemforschung Niedersächsisches Wattenmeer konzentrieren sich die Forschungsarbeiten in den Einzelprojekten darauf, sowohl raum- als auch zeitbezogene Maßstabsfragen zu klären. Dazu sind in der Regel umfangreiche Freilandarbeiten erforderlich. Diese Feldversuche, die mit erheblichem Aufwand verbunden sind und teilweise unter erschwerten Bedingungen stattfinden, sind in Bezug auf ihre Auswertung effizient zu planen und durchzuführen. Daher werden Methoden und Verfahren der Angewandten Statistik so eingesetzt, daß die Relationen von Versuchsplanung, Versuchsdurchführung und Versuchsauswertung berücksichtigt werden, um eine möglichst effiziente Probenahme und eine aussagekräftige Analyse der erhobenen Daten zu gewährleisten.

Projektansatz

Um diese Projektziele zu realisieren wird wissenschaftliches Consulting auf dem Gebiet der Angewandten Statistik angeboten, werden Workshops zu ausgewählten Themen organisiert und durchgeführt, wird geeignete Software ausgewählt, beschafft und installiert. Darüberhinaus werden Lücken in der Verfügbarkeit von Software für spezifische Aufgaben durch die Eigenentwicklung von Programmen geschlossen.

Projektdurchführung

Das Gesamtspektrum statistischer Verfahren und Methoden, die bisher in Bezug auf Versuchsplanung, -durchführung und -auswertung empfohlen und eingesetzt worden sind, läßt sich in diesem Rahmen nicht annähernd

ausschöpfen. Daher werden ausgewählte Methoden und Techniken zur Analyse räumlicher Muster im Anwendungskontext vorgestellt.

Als ein Basismodell ist der räumliche Poisson-Prozeß anzusehen. Das in Abbildung 1 dargestellte Punktmuster gibt eine Simulation von Ansiedlungsorten von *Arenicola marina* wieder. Vom Standpunkt des Substanzwissenschaftlers ist insbesondere die Frage von Bedeutung, ob ein solches Punktmuster durch einen homogenen Poisson-Prozeß erzeugt worden ist. Um diese Forschungshypothese zu prüfen, stellt die Angewandte Statistik den Dispersionsindex-Test zur Verfügung. Bei großem Stichprobenumfang N ist die Teststatistik D_N mit $D_N = (N-1)s^2/x_{am}$ annähernd chiquadratverteilt mit $(N-1)$ Freiheitsgraden. Dabei bezeichnet x_{am} das arithmetische Mittel und s^2 den gebräuchlichen Schätzer für die Varianz.

Das Untersuchungsgebiet wird in N Beobachtungsfenster zerlegt, die sich nicht überlappen und bei dieser Variante des Dispersionsindex-Tests von gleichem Flächeninhalt sind. Für jedes der N Beobachtungsfenster wird die Anzahl der Objekte im Fenster bestimmt, im Beispiel die Anzahl von *Arenicola marina*.

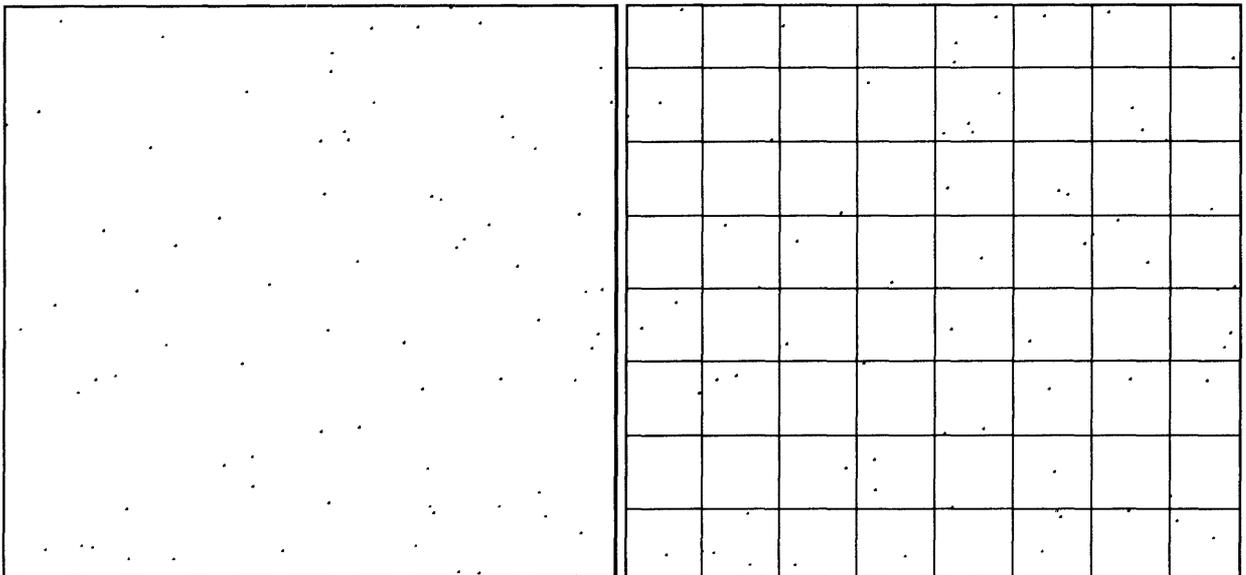


Abb. 1: Ansiedlungsorte *Arenicola marina*

Abb. 2: räumliche Aggregation in 64 Beobachtungsfenstern

Der Wert der Teststatistik D_{64} ist 48,89. Wird eine Irrtumswahrscheinlichkeit von 0,1 gewählt, dann ist die Nullhypothese nicht zu verwerfen, daß die Ansiedlungsorte von *Arenicola marina* von einem homogenen Poisson-Prozeß erzeugt worden sind. Die Daten liefern demnach keinen Hinweis darauf, daß die Ansiedlungsorte nicht "rein zufällig" sind.

Die Simulation wurde mit dem Programm POISSON und das Testverfahren mit dem Programm DISPEX durchgeführt. Die Anwendungssoftware ist im Projekt entwickelt worden und unter MS-DOS lauffähig.

Station 6				
95	1	3	0	42
1	1	1	4	2
0	5	8	81	24
11	1	6	71	116
1	5	116	2	10

Station 8				
165	22	1	94	68
11	82	111	97	153
0	0	24	13	15
31	1	46	22	11
2	0	5	8	6

Bei den Daten handelt es sich um Anzahlen von Individuen der Art *Harpacticus obscurus* in den 25 Kammern eines Multicorers, mit dem jeweils ein Gebiet des Flächeninhalts von insgesamt 400 cm² beprobt wird.

Der Wert der Teststatistik D_{25} ist 1464,20 im Fall von Station 6 und 1490,14 im Fall von Station 8. Für jede vernünftige Wahl der Irrtumswahrscheinlichkeit ist die Nullhypothese zu verwerfen. Die beobachteten Werte für die Teststatistik liefern einen Hinweis darauf, daß Clusterbildung vorliegt. In einem solchen Fall ist es naheliegend, mit geostatistischen Verfahren die Struktur des räumlichen Zusammenhangs zu analysieren. Als Hilfsmittel dazu dient das Variogramm oder das Semivariogramm.

Für die Stationen 6 und 8 werden die folgenden Semivariogramme geschätzt.

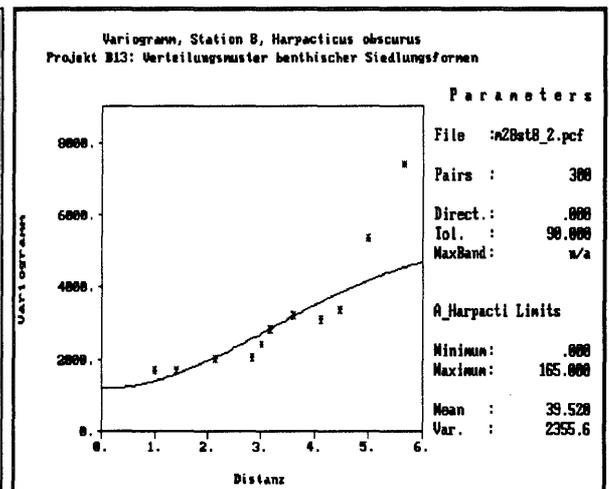
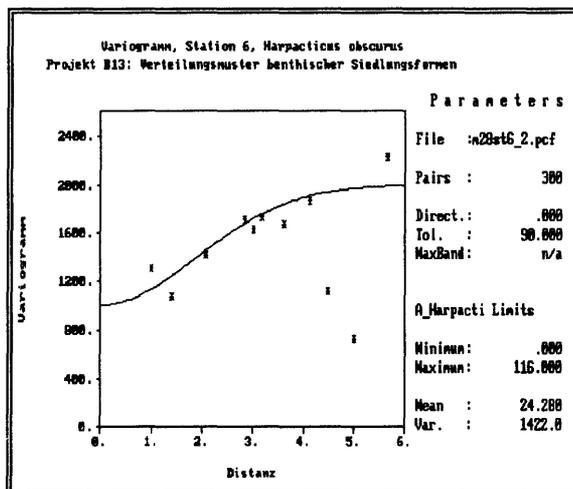


Abb. 3: geschätztes und modelliertes Semivariogramm, Station 6

Abb. 4: geschätztes und modelliertes Semivariogramm, Station 8

Für Station 6 liefert das geschätzte Semivariogramm und seine Modellierung, insbesondere der Wert des Modellparameters zur Reichweite, einen Hinweis darauf, daß der gewählte räumliche Maßstab ausreicht, um eine Ansiedlung von *Harpacticus obscurus* zu erfassen. Dagegen liefert im Fall von Station 8 das geschätzte Semivariogramm und seine Modellierung einen Hinweis darauf, daß der gewählte räumliche Maßstab nicht ausreicht, um eine Ansiedlung von *Harpacticus obscurus* zu erfassen. Das Semivariogramm für Station 6 deutet ferner auf die Auswirkungen eines Hole-Effekts hin.

Die Semivariogramme wurden mit dem Programmsystem GEOEAS (*GEO*statistical *E*nvironmental *A*ssessment *S*oftware), Version 1.2.1, geschätzt und modelliert. Unter GEOEAS wird das Semivariogramm als Variogramm bezeichnet.

Auf der Grundlage der Modellierung des geschätzten Variogramms oder Semivariogramms ist eine Vorhersage der Abundanzen über dem Gebiet der Probenahme möglich. Dazu dienen Varianten des Verfahrens von Krige, das auch kurz als Kriging bezeichnet wird. Liegen für die Gitterpunkte eines regelmäßigen Rasters über dem Untersuchungsgebiet die vorhergesagten Werte für die Abundanz vor, so lassen sich mit einem numerischen Verfahren Isolinien erzeugen.

Für die Stationen 6 und 8 resultieren dann die folgenden zweidimensionalen Darstellungen. Damit wird nur eine Möglichkeit angegeben, die beobachteten Besiedlungsstrukturen von Benthosorganismen zu visualisieren.

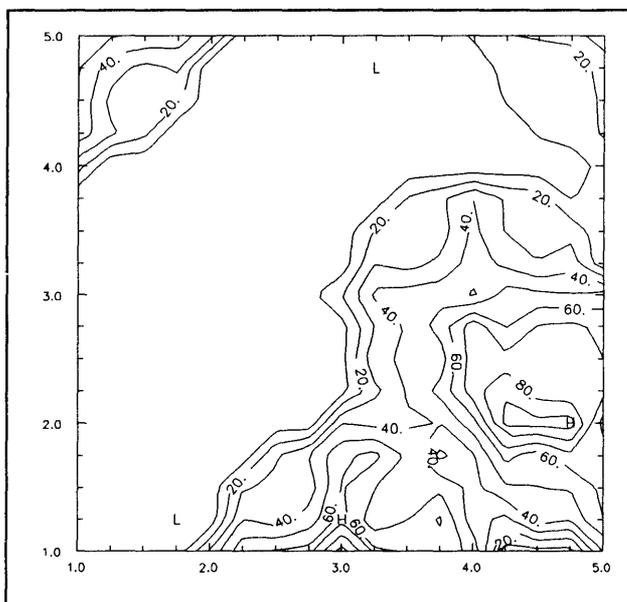


Abb. 5: Kriging; Linien gleicher Abundanz, Station 6

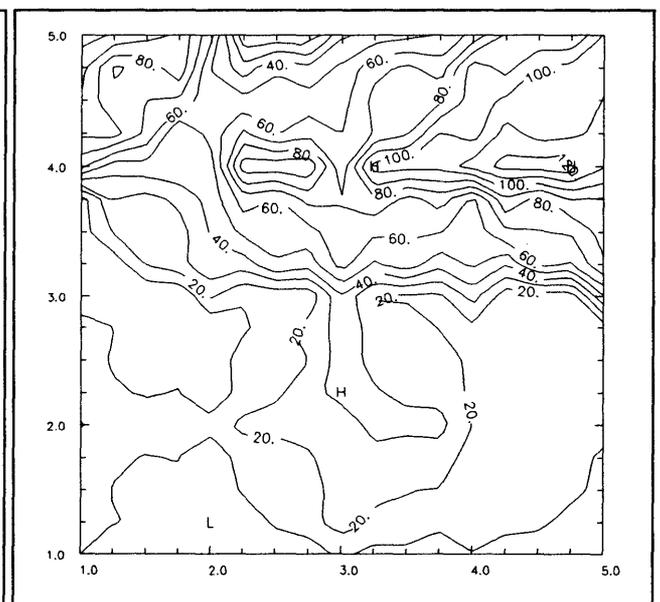


Abb. 6: Kriging; Linien gleicher Abundanz, Station 8

Kriging und numerische Interpolation erfolgten mit dem Programmsystem GEOEAS.

Punktprozesse dienen auch als Basis für komplexere Modelle, wie zufällige abgeschlossene Mengen, mit denen auch kompliziertere räumliche Strukturen modellierbar sind. Bei dem Booleschen Modell, einer speziellen zufälligen abgeschlossenen Menge, werden beispielsweise die Schwerpunkte von geometrischen Objekten, die Keime, als Punkte aufgefaßt, die von einem homogenen Poisson-Prozeß erzeugt worden sind. Die geometrischen Objekte werden auch als Körner bezeichnet und sind im zweidimensionalen Raum häufig Kreise von festem oder zufälligem Radius. In Abbildung 7 ist die Realisierung eines Booleschen Modells mit Kreisen von festem Radius dargestellt.

Neben der Modellierung haben zufällige Mengen auch Bedeutung im Zusammenhang mit der Bestimmung des Bedeckungsgrads. Um den Bedeckungsgrad zu schätzen, gibt es Techniken, wie das Gitterpunkt- und das Linealverfahren. Für das in Abbildung 8 dargestellte Beispiel der Ansiedlung von Muschelbänken auf der Lütetsburger Plate wurde nach dem Gitterpunktverfahren ein Bedeckungsgrad von 13,4% und nach dem Linealverfahren von 11,5% geschätzt. Im Fall eines Booleschen Modells läßt sich für das Linealverfahren auch die Standardabweichung schätzen. Es gibt zwar kein Testverfahren zur Nullhypothese, daß ein Boolesches Modell vorliegt, aber es gibt Eigenschaften, die sich als Nullhypothese formulieren lassen und die als Hinweis darauf zu werten sind, ob anwendungsbezogen von einem Booleschen Modell ausgegangen werden kann. In Bezug auf die vorliegende Anwendung des Linealverfahrens läßt sich die Standardabweichung des Schätzers mit 3,7% abschätzen. Der zeitliche Aufwand, um den Bedeckungsgrad nach dem Gitterpunktverfahren zu schätzen, ist gering und betrug im Fall der vorliegenden Anwendung bei insgesamt 1497 Gitterpunkten weniger als 2 Stunden. Die Anwendung des Linealverfahrens ist zeitaufwendiger, insbesondere wenn die Standardabweichung geschätzt werden soll.

Zur Auswertung von Daten aus der Anwendung des Gitterpunkt- und Linealverfahrens sowie zur Durchführung des zugehörigen Testverfahrens ist inzwischen im Projekt auch Anwendungssoftware entwickelt worden, die unter MS-DOS lauffähig ist.

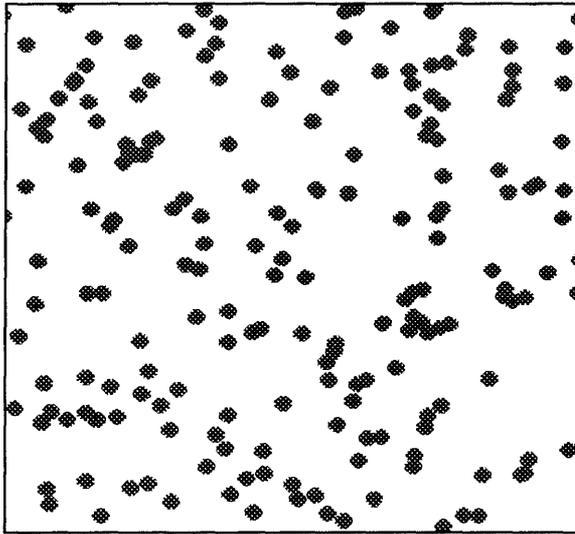


Abb. 7: Boolesches Modell, Kreise mit festem Radius als Körner

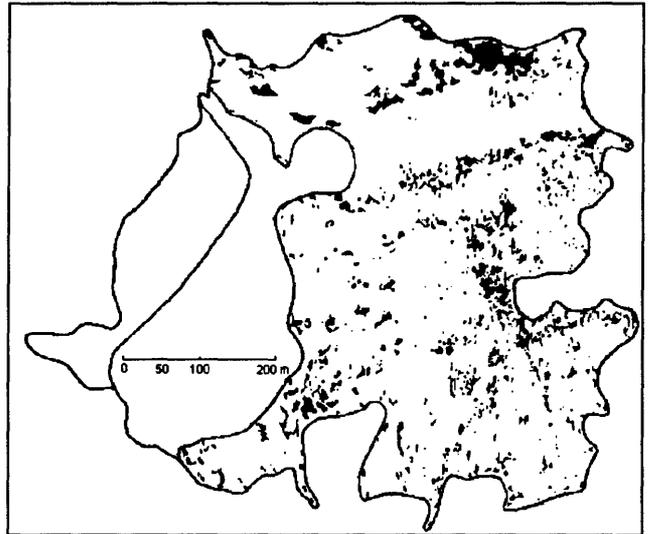


Abb. 8: Muschelbänke Lütetsburger Plate

Die Daten zur Abundanz von *Harpacticus obscurus* wurden freundlicherweise von K. H. van Bernem, Projekt B13, und die Karte zur Ansiedlung von *Mytilus edulis* auf der Lütetsburger Plate von M. Herlyn, Projekt A3.3, zur Verfügung gestellt.

Literatur

Auf eine Zusammenstellung von Literatur wird verzichtet. Die folgenden Veröffentlichungen sind als Sonderdruck verfügbar und geben Hinweise auf relevante Literatur:

Pfeifer, D., Bäumer, H.-P., and Albrecht, M. (1992): Spatial Point Processes and their Applications to Biology and Ecology. *Modeling of Geo-Biosphere Processes*, 1, 145-161

Pfeifer, D., Bäumer, H.-P., and Albrecht, M. (1993): Moving Point Patterns: the Poisson Case. (pp. 248-257) In Opitz, O., Lausen, B., and Klar, R. (eds.): *Information and Classification. Concepts, Methods, and Applications. Proceedings of the 16th Annual Conference of the Gesellschaft für Klassifikation e.V.* Dortmund, April 1-3, 1992. *Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin etc.: Springer