

BERICHTE UND MITTEILUNGEN

RÄUMLICHE SIMULATION IN DER DIFFUSIONSFORSCHUNG
Am Beispiel der Simulationsmodellierung der räumlichen Diffusion von Agrarinnovationen
in der polnischen Landwirtschaft

Mit 4 Abbildungen (z. T. als Beilage V) und 4 Tabellen

ZDZISZAW KAMIŃSKI

Summary: Spatial simulation in diffusion research. The example of simulation modelling of the spatial diffusion of innovations in Polish agriculture

The study presented describes the result arrived at by the author in his research on the spatial diffusion of agricultural innovations in Poland. The subject under examination is the diffusion of modern livestock buildings among peasant farmers in Southern Wielkopolska between 1967–1975. The study aims at a reconstruction of this process and at an understanding of its mechanism in a spatio-temporal approach. This aim has been realized within the framework of the wave-model of innovation diffusion using the Monte Carlo simulation model. The article presents the successive stages of simulation modelling, from the identification of spatial dependences of the observed diffusion process, through the formulation of the rules of a stochastic model and a simulation of the spatial patterns of diffusion for two model approaches, to the determination of the adequacy and the evaluation of the simulation experiments carried out.

Vorbemerkungen

Das Ziel des vorliegenden Aufsatzes ist der Versuch, eine Klasse von Raum-Zeit-Modellen, die auf der Theorie der stochastischen Prozesse beruhen, auf die Simulationsmodellierung der räumlichen Diffusion von Agrarinnovationen anzuwenden. Diese Modelle sind stochastische Monte-Carlo-Simulationen, die prognostische Funktionen erfüllen und daher nicht nur für die geographische Forschung, sondern auch für die räumliche Planung von Bedeutung sind.

Die Entwicklungsphase der sog. „Quantitativen Geographie“ der sechziger und siebziger Jahre ging auf eine verstärkte Orientierung an den exakten Wissenschaften und ihren Methoden zurück. Dies hat eine Expansion statistischer Verfahren und mathematischer Modelle hervorgerufen, um sowohl die Struktur der untersuchten Wirklichkeit zu be-

schreiben als auch bestimmte Phänomene und Prozesse zu erklären und zu prognostizieren. Dank der Fortschritte in der elektronischen Datenverarbeitung ist auch die numerische Simulationsmodellierung zunehmend populärer geworden. Sie findet vor allem Anwendung beim Experimentieren mit Erscheinungen und Prozessen, deren direkte Untersuchung zu teuer oder sogar undurchführbar wäre. Die Effektivität der numerischen Simulation erhöht sich, wenn die zu analysierenden Prozesse einen diskreten, nicht-linearen oder stochastischen Charakter tragen bzw. wenn exakte Lösungen dieser Prozesse nicht möglich sind. Mit derartigen Phänomenen, die sich als Resultat von sehr komplexen Systemen ergeben, haben es Geographen häufig zu tun.

In Polen befindet sich die Entwicklung der Simulationsmodellierung räumlicher Prozesse noch in den Anfängen. Dennoch gibt es in der geographischen Literatur eine Reihe von Arbeiten, die man als Beispiele ihrer Anwendung auch für Bedürfnisse der Raumplanung erwähnen kann. Das erste polnische Simulationsmodell über das Wachstum von Städten stammt von ZIPSER (1972) und baut auf dem sog. *intervening opportunities*-Modell (STOUFFER 1940) auf. Des weiteren ist die Arbeit von MAZURKIEWICZ (1977) zu nennen, in der das Lowry-Typ-Modell (LOWRY 1965) als Grundlage der Simulation der raum-funktionalen Strukturveränderungen einer Stadt verwendet wurde. Ein neuer Vorschlag im Bereich der Simulationsmodellierung des ländlichen Siedlungssystems geht auf DOMAŃSKI und WIERZBICKI (1981) zurück. Die einzigen polnischen Beispiele der Anwendung stochastischer Simulationsmodelle in der Diffusionsforschung bilden die Arbeiten von DRAMOWICZ (1975) und KAMIŃSKI (1982). Die von DRAMOWICZ erarbeiteten Modelle dienen zur Abbildung der Urbanisierungsprozesse im ländlichen Raum. Die Untersuchungen von KAMIŃSKI werden im folgenden Teil des Aufsatzes ausführlich dargestellt.

Beobachtung des Diffusionsprozesses

Forschungsgegenstand ist der Prozeß der räumlichen Ausbreitung von Agrarinnovationen. Im folgenden Beispiel handelt es sich um moderne Wirtschaftsgebäude und Ställe, die in den Jahren 1967–1975 im südlichen Gebiet von Wielkopolska (Großpolen) im privaten Sektor der Landwirtschaft neu errichtet oder modernisiert worden sind. Mit der Untersuchung wird angestrebt, den beobachteten Diffusionsprozeß zu rekonstruieren und seinen raum-zeitlichen Ablauf zu erkennen.

Die Kartenserie in Abb. 1 (Beilage V) stellt die räumliche Diffusion der erwähnten Agrarinnovationen im untersuchten Gebiet (rund 6000 km²) dar. Der Ausbreitungsprozeß setzt im Jahre 1967 im westlichen und südlichen Gebiet von Großpolen ein. Die Lokalisation der ersten Adoptoren, d. h. der 23 Bauern, die die Innovation zuerst angenommen haben, bestimmt den Beginn der Diffusion zum Zeitpunkt t_0 und bildet die Ausgangssituation für die weitere Entwicklung dieses Prozesses. Schon im Jahre 1968 (t_1) hat sich die kumulative Zahl von Adoptoren auf 80 Personen erhöht und sich bis 1969 nochmals verdoppelt. Kennzeichnend für diese Phase des Ausbreitungsprozesses ist die Häufung der neuen Adoptoren um die ersten Neuerer herum. Im Jahr 1975 (t_2) beträgt die kumulative Zahl von Adoptoren 1681 Personen, und fast das ganze Gebiet von Großpolen ist von diesem Diffusionsprozeß erfaßt. Lediglich im nördlichen Teil des Untersuchungs-

raumes hat sich ein kleines Gebiet gebildet, in dem sich die Diffusion nicht ausbreiten konnte.

Räumliche Abhängigkeiten des Diffusionsprozesses

In der Simulationsmodellierung räumlicher Diffusionsprozesse spielt die Weitergabe von Informationen eine sehr wichtige Rolle. Mit anderen Worten, man fragt danach, wie sich die Ausbreitung von Informationen innerhalb der Bevölkerung eines Gebietes raum-zeitlich vollzieht. Zu diesem Zweck geht man von der Annahme aus, daß die gesamte Kommunikation innerhalb der Bevölkerung auf persönlichen Kontakten beruht. Diese vereinfachende Annahme findet unmittelbare Anwendung bei der Bearbeitung des sog. *Mittleren Informationsfeldes (Mean Information Field - MIF)*, das die Wahrscheinlichkeit der nachbarschaftlichen Kontakte bei der Informationsweitergabe ausdrückt, die aufgrund der empirisch-bestimmten *Distanzabnahmefunktionen (distance decay function)* formuliert werden kann (HÄGERSTRAND 1967).

Es ist klar, daß ein solches Mittleres Informationsfeld den konkreten lokalen Bedingungen eines sozio-ökonomischen Systems, in dem der Diffusionsprozeß stattfindet, entsprechen sollte. Das hier vorgestellte Verfahren beruht darauf, daß die Bestimmung des MIF aufgrund der Beteiligung an den landwirtschaftlichen Märkten und damit über ein sog. *Beteiligungsfeld* erfolgt. Jedes „Beteiligungsfeld“ hat seine be-

Tabelle 1: Reale Verteilung und theoretische Verteilungen der Beteiligungsdichte an den Märkten

Actual distribution and theoretical distributions of journey to market densities

| Entfernung gegenüber Ringmitte (in km) | Empirische Beteiligungsdichte (Personen/km ²) | Distanzabnahmefunktionen | | |
|---|---|--------------------------|--------------------------|-------------------------------------|
| | | Pareto | Exponential- Funktion | Pareto- Exponential- Funktion |
| 1 | 4,1400 | 10,2415 | 2,3445 | 4,1675 |
| 3 | 1,8842 | 1,7050 | 1,6449 | 1,8083 |
| 5 | 1,1942 | 0,7408 | 1,1540 | 1,0710 |
| 7 | 0,4663 | 0,4278 | 0,8096 | 0,6956 |
| 9 | 0,5484 | 0,2838 | 0,5680 | 0,4725 |
| 11 | 0,5138 | 0,2046 | 0,3985 | 0,3297 |
| 13 | 0,1408 | 0,1557 | 0,2796 | 0,2342 |
| 15 | 0,1539 | 0,1233 | 0,1961 | 0,1684 |
| 17 | 0,1498 | 0,1005 | 0,1376 | 0,1223 |
| 19 | 0,1425 | 0,0838 | 0,0965 | 0,0894 |
| 21 | 0,0645 | 0,0712 | 0,0677 | 0,0658 |
| 23 | 0,0243 | 0,0613 | 0,0475 | 0,0487 |
| 25 | 0,0319 | 0,0535 | 0,0333 | 0,0361 |
| 27 | 0,0620 | 0,0472 | 0,0233 | 0,0268 |
| 29 | 0,0138 | 0,0420 | 0,0164 | 0,0200 |
| | | $r^2 = 0,89$ | $r^2 = 0,91$ | $r^2 = 0,99$ |

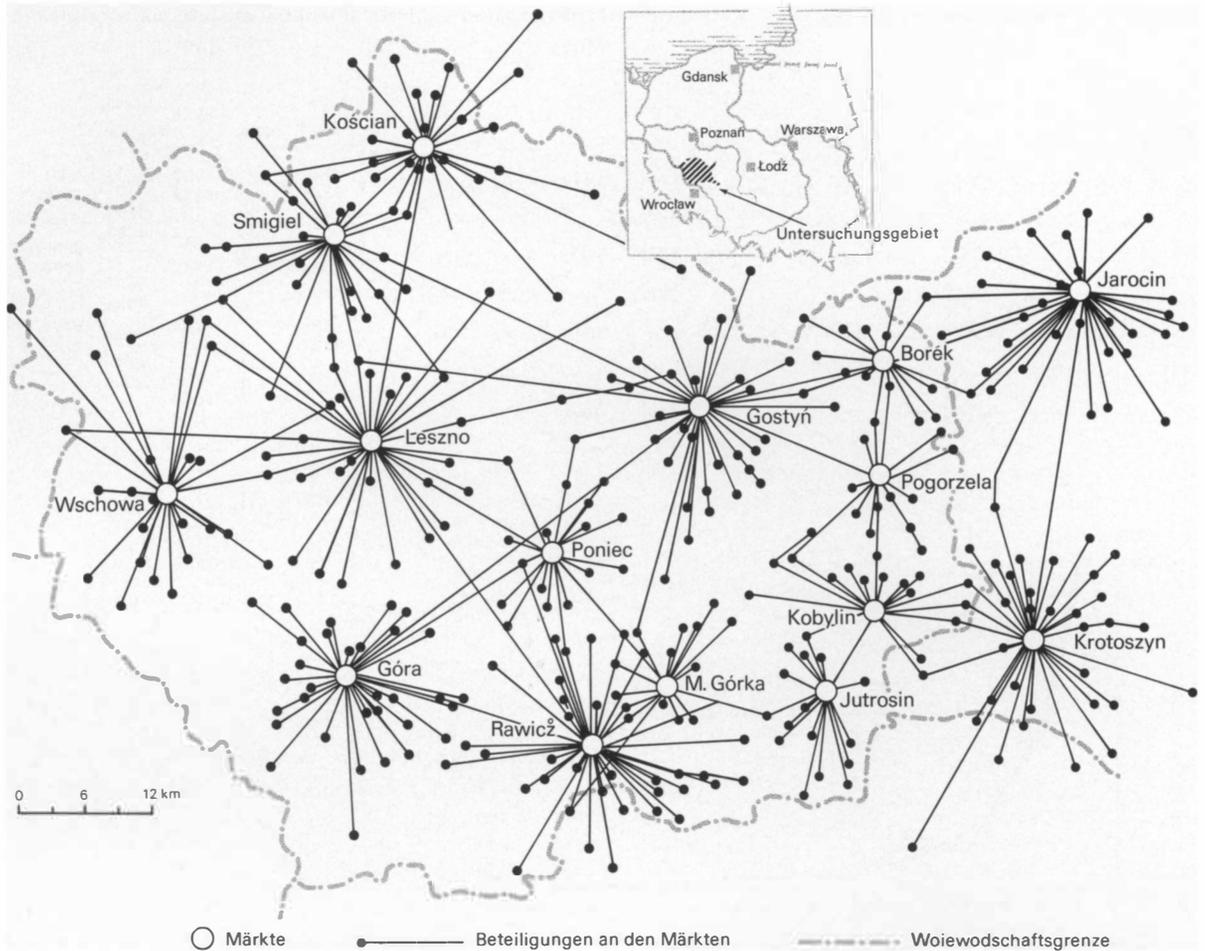


Abb. 2: Räumliche Ausbreitung der Beteiligungen an den Märkten im südlichen Gebiet von Wielkopolska (Großpolen) im Jahre 1976

Spatial distribution of journeys to markets in southern Wielkopolska, 1976

stimmte räumliche Struktur, und die Intensität der persönlichen Kontakte verringert sich mit wachsender Entfernung zum Zentrum des jeweiligen Feldes (vgl. Abb. 2 und Tab. 1). Es erweist sich, daß zwischen den beiden Variablen, d. h. der Teilnahme der Bauern an den Märkten und der Entfernung zu den Märkten, ein recht enger negativer Zusammenhang besteht (Korrelationskoeffizient $r = -0,720$). Die empirische Verteilung der Beteiligungsichte in Abhängigkeit von der Entfernung ist in Tab. 1 dargestellt. Zur mathematischen Beschreibung dieses beobachteten Zusammenhanges werden drei Distanzabnahmefunktionen verwendet, und zwar die Pareto-, die Exponential- und die Pareto-Exponential-Funktion, deren Parameter mit Hilfe der „Methode der kleinsten Quadrate“ geschätzt worden sind. Die Bestimmung der Anpassungsgüte der verwendeten Funktionen an die empirische Verteilung wird mit Hilfe des Bestimmtheitsmaßes (r^2) durchgeführt.

Aus dem Vergleich der einzelnen Bestimmtheitsmaße (vgl. Tab. 1) geht hervor, daß die Pareto-Exponential-Funktion die beste Anpassung zeigt. Diese Funktion verbindet zufälliges mit zweckmäßigem Verhalten der Menschen im geographischen Raum (vgl. MORRILL u. PITTS 1967).

Aufbau des Mittleren Informationsfeldes

Abb. 3 zeigt das Verfahren beim Aufbau des Mittleren Informationsfeldes (vgl. MORRILL u. PITTS 1967 sowie GOULD 1969). Den Anfang bildet das sog. Netz von erwarteten Kontakten (*Expected Floating Grid - EFG*), das aus 19 hexagonalen Raumeinheiten (cells) besteht. Die Fläche einer Zelle beträgt 25 km², die Gesamtfläche des EFG ist 475 km². Als Radius des EFG werden ca. 11 km angenommen, weil anschließend die Beteiligung an den Märkten stark abnimmt (vgl.

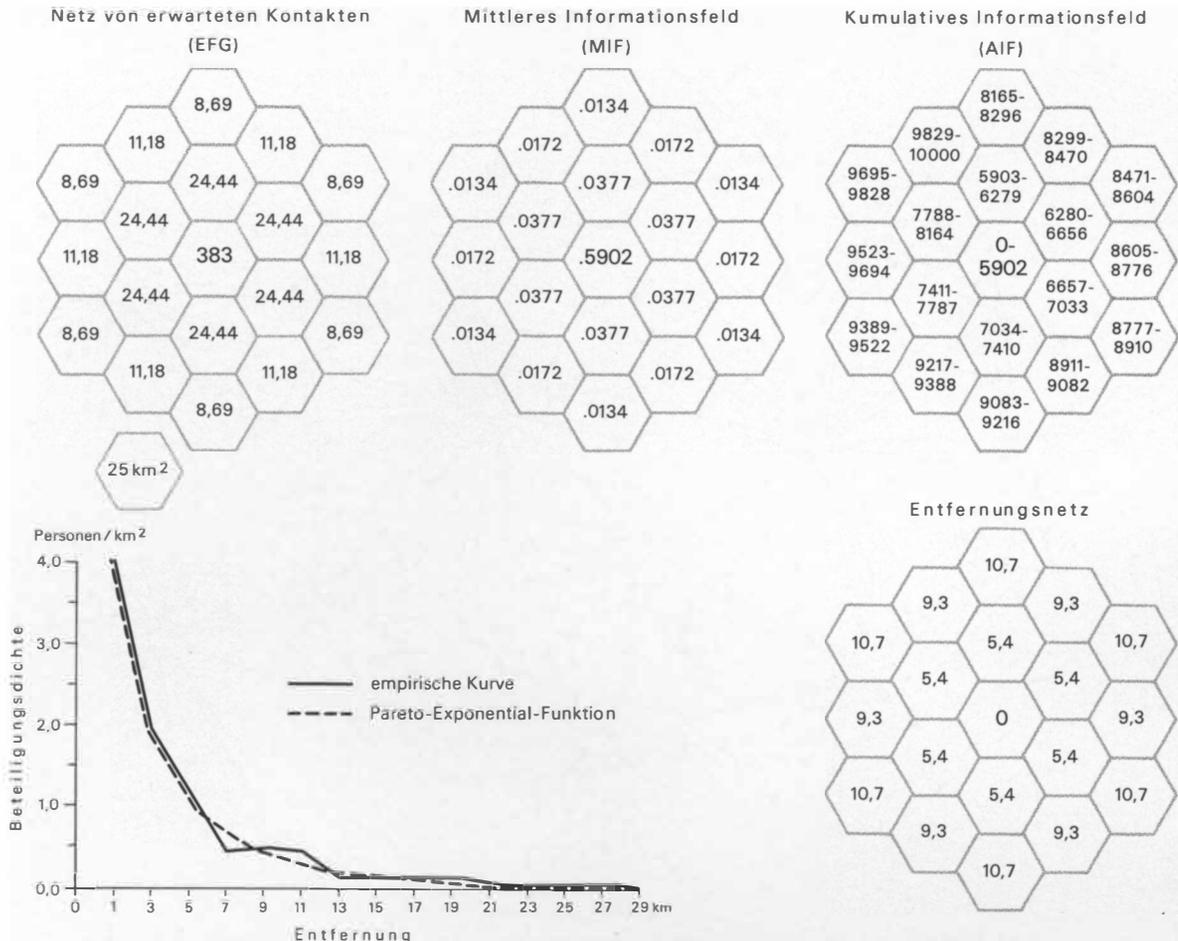


Abb. 3: Konstruktion des Mittleren Informationsfeldes (MIF) und des Kumulativen Informationsfeldes (AIF)
Construction of the mean information field (MIF) and the cumulative information field (AIF)

Abb. 3). Die Berechnung der Werte für jede Zelle des EFG wird in Anlehnung an die oben abgeschätzte Pareto-Exponential-Funktion durchgeführt. In der Zentralzelle steht die Zahl von Personen, die in den jeweiligen Marktorten wohnen (im Beispiel 388). Das MIF enthält eine zu EFG analoge Aussage, jetzt jedoch ausgedrückt als Wahrscheinlichkeit von Kontakten (Summe aller Zellen = 1,0). Abb. 3 stellt die Etappen beim Aufbau des MIF und des sog. *Accumulated Information Field (AIF)* dar. Dieses kumulative Informationsfeld ist eine Matrix, die Intervalle von Zufallszahlen enthält. Diese Zufallszahlenintervalle entstehen durch Kumulation der Zufallszahlen aus dem MIF im Uhrzeigersinn.

Da das oben dargestellte Verfahren bei der Konstruktion des MIF vollständig isotrop ist, d. h. keine Richtungsunterschiede zulässt, und damit zu wenig realistisch ist, sollte man die Abschätzung der Wahrscheinlichkeitswerte besser aufgrund der Zahlen potentieller Adoptoren durchführen. Dabei geht man

von der Annahme aus, daß die potentiellen Adoptoren der untersuchten Agrarinnovationen solche Bauern sind, die eine Feldfläche von über 10 ha besitzen. In jede Zelle des Simulationsnetzes, das 281 hexagonale Raumeinheiten umfaßt, wird die Zahl solcher potentieller Innovationsübernehmer eingetragen.

Aus dem Obigen geht hervor, daß sich die Wahrscheinlichkeit P_i^* , daß ein Adoptor (teller), der die zentrale Zelle von MIF bewohnt, und ein potentieller Innovationsübernehmer (receiver), der eine andere Zelle i von MIF bewohnt, zur Zeit t zusammen treffen, wie folgt abschätzen läßt:

$$P_i^* (t_j) = \frac{Y_i \cdot L_i(t_j)}{\sum_{i=1}^{19} Y_i \cdot L_i(t_j)}$$

wobei Y_i = Wahrscheinlichkeit der persönlichen Kontakte entsprechend dem Isotropenmodell des MIF,

- L_i = Zahl von potentiellen Innovationsübernehmern in der i -ten Zelle,
 t_j = Annahmezeit von Innovationen ($t_0 = 1967 (G_0)$, $t_1 = 1968 (G_1)$, $t_2 = 1969 (G_2)$,
 \dots , $t_8 = 1975 (G_8)$),
 i = Zelle.

Das bedeutet, daß jede neue Verteilung der Wahrscheinlichkeitswerte sowohl von der Entfernung vom Diffusionszentrum als auch von der Zahl der potentiellen Innovationsübernehmer in jeder Zelle des Simulationsnetzes und in jeder Generation des Simulationsmodells ($G_1, G_2, G_3, \dots, G_8$) abhängig ist.

Prinzipien des Simulationsmodells

Die Simulation des untersuchten Diffusionsprozesses stützt sich auf folgende Prinzipien (siehe KAMIŃSKI 1982):

1. Das Gebiet, in dem der beobachtete Ausbreitungsprozeß der Agrarinnovation stattfindet, ist in hexagonale Raumeinheiten, sog. Zellen, mit gleichgroßen Flächen (25 km^2) unterteilt.
2. Die Zeitabfolge des Diffusionsprozesses wird in einzelne Zeitspannen, die sog. Generationen des Simulationsmodells, gegliedert. Bekannt ist also die Zeit $t_0 = 1967 (G_0)$, die über den Anfang des beobachteten Prozesses entscheidet.
3. Die Kommunikation innerhalb der Bevölkerung beruht auf persönlichen Kontakten, wobei man in diesem Bereich zwei alternative Prozeduren unterscheidet:
 - a) die Prozedur von HÄGERSTRAND, in der man davon ausgeht, daß sich jeder Adopter in einer Generation nur mit einem Nicht-Adopter treffen kann und
 - b) die Prozedur, in der in jeder Generation des Modells eine Angleichung der Zahlen von simulierten und realen Adoptoren erfolgt.
4. Bekannt ist die Lokalisation der ersten Innovationsadoptoren zur Zeit $t_0 (G_0)$ und die Zahl potentieller Innovationsübernehmer innerhalb jeder Zelle des Simulationsnetzes für jede Generation des Simulationsmodells.
5. Bekannt ist weiterhin das sog. Mean Information Field, das von der Distanzabnahmefunktion und von der Zahl der potentiellen Innovationsübernehmer abhängig ist.
6. Die Übernahme der Neuerung durch einen Nicht-Adopter ist von bestimmten Kategorien der psychischen Hemmung (psychological resistance) gegenüber der Innovation abhängig. Um dem Rechnung zu tragen, wird angenommen, daß die psychische Hemmung gegen die Innovation normalverteilt ist.

Nicht berücksichtigt werden physische Barrieren, weil es in dem untersuchten Gebiet keine solchen

Effekte gibt. Desweiteren wird unterstellt, daß außerhalb der untersuchten Region keine Möglichkeiten zur Übernahme der Innovation bestehen, also $P_i^*(t) = 0$ für i außerhalb der Untersuchungsregion gilt.

Die Simulation des Diffusionsprozesses geht auf Monte-Carlo-Operationen zurück, mit denen Pseudozufallszahlen erzeugt werden. Auf diese Weise bestimmt man

- die Lokalisation der Zelle, in der ein neuer Innovationsübernehmer wohnt,
- eine konkrete Person, die die Innovation annehmen könnte,
- die Kategorien der psychischen Hemmung.

Simulationsprozeduren

Die überwiegende Mehrheit von Forschern, die sich mit der Problematik der räumlichen Diffusion beschäftigen, lehnt sich dabei an das Modell von HÄGERSTRAND (1967) an, bei dem davon ausgegangen wird, daß die Übernahme der Neuerung direkt nach Erhalt der Information erfolgt. Es kann jedoch nicht genug darauf hingewiesen werden, daß dieser Mechanismus wenig realistisch ist. Diese Schwierigkeiten kann man überwinden, indem man die erwähnte HÄGERSTRAND-Prozedur modifiziert. Diese Modifizierung beruht darauf, für jede Generation die Zahl der simulierten neuen Adoptoren der Zahl der realen neuen Adoptoren anzugleichen. Die Zweckmäßigkeit dieser Auffassung haben schon früher unter anderem YAPA (1976), BIVAND (1978) und KAMIŃSKI (1982) betont.

Man nimmt dabei folgendes an:

1. T bezeichnet die Zahl von Zeitspannen, Generationen,
2. $xx[t]$, $t = 1, 2, \dots, T$ bezeichnet die realen Summen von Adoptoren,
3. $yy[t]$ bezeichnet die Zahl von Adoptoren im Simulationsmodell und
4. $y[i]$, $i = 1, 2, \dots, N$ bezeichnet das Register von Adoptoren.

In der klassischen HÄGERSTRAND-Prozedur werden $m \leq xx[t]$ potentielle Innovationsübernehmer zu Adoptoren, weil nur $xx[t]$ Informationen übermittelt werden. Ein Teil davon kann das Untersuchungsgebiet verlassen oder infolge der Wirkung der psychischen Hemmung nicht zum Tragen kommen. Die Modifizierung der HÄGERSTRAND-Prozedur beruht auf der Notwendigkeit der Erzeugung von $m = xx[t + 1] - xx[t]$ neuen Adoptoren in der Generation t . Dies bedeutet, daß eine Verlängerung der Simulation notwendig wird, bis $yy[t] = xx[t + 1]$ gilt. Damit die Wahrscheinlichkeit der Informationsübermittlung im weiteren Verlauf gleich bleibt, muß man solange Personen aus dem Register der aktiven Adoptoren

auslösen und sie als neue Adoptoren registrieren, bis $yy = xx[t + 1]$ ist.

Angenommen, daß die Eingangsdaten des Simulationsmodells bekannt sind, d. h. die Zahl von potentiellen Innovationsübernehmern $z[i]$, $i = 1, 2, \dots, N$, die Zahl der aktiven Adoptoren $x[i]$, $i = 1, 2, \dots, N$, $xx[t]$, $t = 1, 2, \dots, T$, sowie die Wahrscheinlichkeitswerte des MIF, dann kann man die oben besprochene modifizierte Prozedur der Informationsübermittlung während des räumlichen Diffusionsprozesses in folgenden Schritten darstellen:

1. Zum Ausgangszeitpunkt $t = 1$ stimmt die Zahl der Adoptoren mit der im Register überein ($yy[1] = xx[1]$).
2. Man löst den Informationsübermittler aus dem Register von aktiven Adoptoren und legt die Zelle k , zu welcher der Übermittler gehört, fest.
3. Durch das abzuschätzende Informationsfeld über k wird die Zelle des ausgewählten potentiellen Übernehmers bestimmt.
4. Falls eine auf diese Weise ausgewählte Person ein Adoptor, aber kein Übermittler ist oder für sie eine psychische Hemmung zutrifft oder die Zelle außerhalb des Untersuchungsgebietes liegt, dann muß man zu Punkt 2, falls aber diese Person ein Übermittler ist, zu Punkt 3 zurückgehen. Ansonsten kann man das Verfahren fortsetzen.
5. Die Zahl der Adoptoren im Simulationsmodell ist um 1 zu erhöhen, und wenn sie kleiner als die reale Summe von Adoptoren ist, dann muß man zu Punkt 2 zurückspringen.
6. Die Zahl der aktiven Adoptoren entspricht jetzt den Adoptoren im Register.
7. Man erhöht t um 1 und beginnt wieder bei Punkt 1. Falls $t + 1 = T$ ist, endet die Prozedur.

Simulation des räumlichen Ablaufs des Diffusionsprozesses

Im untersuchten Beispiel wird sowohl die bekannte Prozedur von HÄGERSTRAND (Version 1 des Modells) als auch die oben dargestellte modifizierte Prozedur (Version 2) angewandt. Jedes Verfahren beruht auf der Erzeugung von 99 Serien der Simulation, wobei jede Serie 8 Generationen des Simulationsmodells umfaßt ($G_1 = 1968, \dots, G_8 = 1975$). Insgesamt werden 198 Simulationsserien des Diffusionsprozesses erarbeitet.

Die Monte-Carlo-Simulation des Diffusionsprozesses besteht in der Erzeugung der Pseudo-Zufallszahlen, die aus dem sog. multiplikativen Generator von Zufallszahlen nach folgender Formel erfolgt:

$$X_{i+1} = \text{frac}(\pi \cdot x_i)$$

Die Simulation der effektiven Informationsanweisung (Innovationsübernahme) ist eine zweistufige Prozedur. Die erste Zufallszahl bestimmt die Lokali-

sation einer Zelle, die ein Informationsadressat bewohnt, die zweite identifiziert die konkrete Person aus den potentiellen Adoptoren. Der Effekt der psychischen Hemmung wird für jeden potentiellen Annehmer noch vor der Durchführung des Simulationsmodells bestimmt. Diese Vorgehensweise ist Ausdruck des stochastischen Wesens des räumlichen Diffusionsprozesses.

Testen der Simulationsmodelle

Das Testen der Simulationsmodelle gehört bis heute zu den noch wenig angewandten Verfahren innerhalb der geographischen Simulationsforschung. Im vorliegenden Fall basiert das Testen des Simulationsmodells auf der HOPE-Typ-Prozedur (HOPE 1968). Diese Prozedur wurde zum ersten Mal zur Überprüfung des HÄGERSTRAND-Modells von CLIFF und ORD (1973) verwendet. Sie besteht aus folgenden Schritten:

1. Man erzeugt m unabhängige Realisationen des Simulationsmodells $x[i,k]$, $i = 1, \dots, N$, $k = 1, \dots, m$.
2. Dann berechnet man eine mittlere Karte $S[i]$ aus den m Realisationen und aus der Karte des beobachteten Prozesses $y[i]$

$$S[i] = \left\{ \sum_{k=1}^m x[i,k] + y[i] \right\} / (m + 1), \quad i = 1, 2, \dots, N.$$

3. Man ermittelt für jede Karte ein Anpassungsmaß.
4. Man ordnet die $(m + 1)$ Ergebnisse nach der Größe der Standardnormalabweichungen und lehnt die Nullhypothese ab, falls keine signifikanten Unterschiede zwischen einer mittleren Karte und der Karte des beobachteten Prozesses bestehen.

Im 3. Schritt der HOPE-Typ-Prozedur ist eine Version des I-Autokorrelationstestes von MORAN (vgl. CLIFF u. ORD 1973) für die Unterschiede zwischen den Karten verwendet worden, d. h. zwischen der mittleren Karte (2. Schritt) und den Simulationskarten (m Realisationen) als auch der Karte des beobachteten Prozesses $y[i]$. In dieser Version des I-Testes nimmt man an, daß die Unterschiede zwischen den Karten nicht gleich Null sind, daß also

$$\sum_{i=1}^n n_i \neq 0 \text{ gilt.}$$

Die angewandte Variante des I-Testes von MORAN ist wie folgt definiert:

$$I_k = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} z_{ik} z_{jk}}{w \sum_{i=1}^n z_{ik}^2},$$

wobei $w_{ij} = w_{ij}^* / \sum_{j=1}^n w_{ij}^*$ mit

$$w_{ij}^* = \begin{cases} 1 & \text{wenn } i \text{ ein Nachbar von } j \text{ ist} \\ 0 & \text{in übrigen Fällen} \end{cases}$$

und $w = \sum_{i,j} w_{ij}^*$,

$$z_{ik} = x_{ik} - \bar{x}_i, \quad \bar{x}_i = \frac{1}{m} \sum_{k=1}^m x_{ik}, \quad k = 1, 2, \dots, m,$$

wobei \bar{x}_i ein Wert der Zelle i der mittleren Karte und m die Zahl der Karten zum Vergleich mit der mittleren Karte ist.

Die Standardnormalabweichung der oben besprochenen I-Statistik ist $SND(I) = [I - E_R(I)] [Var_R(I)]^{-1/2}$.

Im 4. Schritt der HOPE-Typ-Prozedur ordnet man die Standardnormalabweichungen vom kleinsten (Rang 1) bis zum größten Wert (Rang $m + 1$). Die am besten angepaßte Karte ist diejenige, für die die Standardnormalabweichung $SND(I)$ dem Wert Null möglichst nahe kommt. Unter Beachtung der Position der Karte des beobachteten Prozesses in der geordneten Reihe der $SND(I)$ kann man die Anpassungsstufe des Modells bestimmen. In diesem Fall testet man die Hypothese H_0 , ob die Unterschiede zwischen der mittleren Karte und der Karte des beobachteten Prozesses, bei einem angenommenen Signifikanzniveau ($\alpha = 0,05$), nicht signifikant sind.

Im Fall der Verwendung der HÄGERSTRAND-Prozedur werden sechs von acht Modellgenerationen

getestet ($G_1, G_2, G_4, G_6, G_7, G_8$). Dabei wird angenommen, daß das Testen einer Generation in dem Moment beginnt, in dem die Globalzahl der simulierten Adoptoren $yy[t + 1]$ größer oder gleich der Zahl realer Adoptoren $xx[t + 1]$ ist. Wenn diese Voraussetzung nicht erfüllt ist, d. h. wenn $yy[t + 1] < xx[t + 1]$ ist, dann ist eine Fortsetzung des Simulationsmodells notwendig. Die Testergebnisse der beiden Modellierungsverfahren sind in Tab. 2 dargestellt.

Bewertung des Simulationsmodells

Die wichtigste Eigenschaft eines gut angepaßten Simulationsmodells besteht darin, daß die Unterschiede zwischen den Simulationskarten und der Karte des realen Prozesses nicht signifikant sind. Zur Überprüfung wurde der räumliche I-Autokorrelationstest von MORAN herangezogen (s. oben), und zwar für jede zu testende Generation des Modells (Tab. 2). Daraus läßt sich nicht nur ablesen, ob das angewandte Monte-Carlo-Simulationsmodell den zu erforschenden Diffusionsprozeß gut widerspiegelt, sondern man kann auch diejenigen der zwei Versionen des Simulationsmodells auswählen, die diesen Prozeß besser beschreibt.

Diese Auswahl ist anhand der Analyse der räumlichen Autokorrelation der Differenzen zwischen den Simulationsergebnissen und den realen Werten des beobachteten Prozesses möglich. Die Resultate dieser Analyse stellt ebenfalls Tab. 2 dar. Da die I-Statistik

Tabelle 2: Test des Simulationsmodells
Testing the simulation model

| Generation | Nummer der Serie, die am besten zur durchschnittlichen Karte angepaßt ist (nach $SND(I)$ Werten) | $SND(I)$ für die bestangepaßte Serie | Kartenposition des beobachteten Prozesses in der geordneten Reihe $SND(I)$ | $SND(I)$ für die Karte des beobachteten Prozesses | Signifikanzniveau für die Karte des beobachteten Prozesses nach HOPE-Prozedur |
|------------------|--|--------------------------------------|--|---|---|
| VERSION 1 | | | | | |
| G_1 | 23 | -0,1575 | 87 | -0,8270 | 0,14 |
| G_2 | 29 | -0,0030 | 84 | -0,8248 | 0,17 |
| G_4 | 68 | -0,0576 | 68 | -0,9371 | 0,33 |
| G_6 | 16 | 0,0254 | 60 | -0,9354 | 0,41 |
| G_7 | 1 | -0,0325 | 58 | -0,6933 | 0,43 |
| G_8 | 27 | 0,0135 | 25 | -0,9394 | 0,26 |
| VERSION 2 | | | | | |
| G_1 | 63 | 0,0035 | 100 | 3,1700 | 0,01 |
| G_2 | 89 | 0,0056 | 93 | 2,5828 | 0,08 |
| G_3 | 99 | 0,0475 | 68 | 1,6424 | 0,33 |
| G_4 | 63 | -0,0004 | 74 | 2,4427 | 0,27 |
| G_5 | 63 | -0,0140 | 74 | 3,0533 | 0,27 |
| G_6 | 91 | 0,4018 | 62 | 3,6111 | 0,39 |
| G_7 | 63 | -0,0565 | 70 | 5,0358 | 0,31 |
| G_8 | 63 | 0,5971 | 81 | 6,9063 | 0,20 |

Tabelle 3: Test der theoretischen Verteilungen mit Hilfe des χ^2 -Anpassungstests
Testing the theoretical distributions using the χ^2 -test

| Werte des Tests | Theoretische Verteilungen | | | | | |
|-------------------------------------|---------------------------|-----------------------|------------------|-----------------------|------------------|-----------------------|
| | Negativ-Binomial | | Polya-Aeppli | | NEYMAN-Typ A | |
| | realer Prozeß | simulierter Prozeß | realer Prozeß | simulierter Prozeß | realer Prozeß | simulierter Prozeß |
| χ^2 | 23,62 | 14,48 | 11,63 | 35,68 | 47,17 | 757,33 |
| Anzahl der Freiheitsgrade (FG) | 14 | 12 | 14 | 12 | 14 | 12 |
| Signifikanzniveau (α) | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 | 0,05 |
| Kritischer Wert (χ^2_α) | 23,69 | 21,03 | 23,69 | 21,03 | 23,69 | 21,03 |

von MORAN asymptotisch normalverteilt ist, ist anzunehmen, daß eine signifikante räumliche Autokorrelation der Unterschiede (auf dem 5%-Niveau signifikant) nur im Fall der zweiten, modifizierten Version des Modells vorhanden ist. Kennzeichnend ist dabei die Tatsache, daß in dieser Version nur positive Werte der räumlichen Autokorrelation auftreten (vgl. Tab. 2). Dies bedeutet, daß das Vorhandensein eines Sachverhaltes in einer Raumeinheit die Wahrscheinlichkeit vergrößert, daß dieser Sachverhalt auch in den benachbarten Raumeinheiten vorhanden ist. Damit hat sich die zweite, modifizierte Version des Simulationsmodells als besser erwiesen.

Zur Bewertung der Simulationsexperimente stehen noch andere Methoden zur Verfügung. In der hier durchgeführten Untersuchung wird zusätzlich die Quadratanalyse, ein weniger kompliziertes Verfahren, herangezogen (vgl. hierzu BAHRENBURG, GIESE, NIPPER, SCHICKHOFF, STREIT 1978; HARVEY 1966). Die Quadratanalyse ermöglicht, die Anordnung der untersuchten Punkte und somit die räumliche Entwicklungstendenz von beobachteten Diffusionsprozessen zu untersuchen.

Bei der Anwendung der Quadratanalyse geht man ebenfalls von dem über die zugrundeliegende Fläche gelegten Gitternetz mit 281 hexagonalen Zellen aus. Um die Diffusion von Neuerungen zu beschreiben und die Ergebnisse der Simulationsmodellierung (Version 2) mit den realen Werten vergleichen zu können, werden insgesamt vier theoretische Verteilungen eingesetzt, und zwar: Poissonverteilung, Negativ-Binomialverteilung, Polya-Aeppli-Verteilung und NEYMAN-Typ A-Verteilung. Als Anpassungstest wird der χ^2 -Test gewählt, der zu den häufigsten Tests bei der Quadratanalyse gehört (Tab. 3). In der Tabelle sind keine Werte für die Poissonverteilung aufgenommen, weil diese Verteilung, die eine zufällige Anordnung der Punkte beschreibt, sich als

völlig unbrauchbar erwiesen hat, da sowohl für den realen als auch für den simulierten Diffusionsprozeß die χ^2 -Werte sehr weit über den kritischen Werten χ^2_α liegen.

Der Vergleich der χ^2 - und χ^2_α -Werte der angewandten theoretischen Verteilungen zeigt, daß befriedigende Resultate des Anpassungstests nur in einigen Fällen eingetreten sind. Im Fall der NEYMAN-Typ A-Verteilung und im Fall der Polya-Aeppli-Verteilung für den simulierten Diffusionsprozeß muß H_0 abgelehnt werden. Die Negativ-Binomialverteilung hat sich als beste Verteilung erwiesen, sowohl für den realen als auch für den simulierten Prozeß. Diese Feststellung bestätigt ein Untersuchungsergebnis von HARVEY (1966), der sich mit der Anwendung der Quadratanalyse zum Testen des HÄGERSTRAND-Modells beschäftigt hat. Die Negativ-Binomialverteilung beschreibt am besten die Tendenz zur klumpenhaften Anordnung der Punkte. Diese Anordnung ist eine Konsequenz der wellenförmigen Ausbreitung von Neuerungen im Raum.

Um den Grad der Clusterung im Zeitverlauf zu erfassen, wird das Verhältnis von Varianz (m_2) und Mittelwert (m_1) der Häufigkeitsverteilungen für die einzelnen Generationen ermittelt. Genauer wird das sog. L-Maß angewandt, das wie folgt definiert ist:

$$L = \frac{N}{N-1} \cdot \frac{m_2}{m_1},$$

wobei m_1 = Mittelwert, m_2 = Varianz, N = Anzahl der Zellen ist.

Auf diese Weise lassen sich drei Punktmuster identifizieren, und zwar eine klumpenhafte ($L > 1$), eine zufällige ($L = 1$) und eine regelmäßige ($L < 1$) Verteilung. Dieses L-Maß, das in der polnischen geographischen Literatur als die „Anzahl von Lexis“ bezeichnet wird, wurde schon in den dreißiger Jahren

Tabella 4: Lexis-Maß nach der Zeitfolge des Diffusionsprozesses und nach den Generationen des Simulationsmodells

Lexis measure by the temporal sequence of diffusion and the simulation model generations

| Zeitfolge in Jahren (Generationen) | realer Prozeß | simulierter Prozeß |
|------------------------------------|---------------|--------------------|
| 1967 (G ₀) | 1,005 | 1,005 |
| 1968 (G ₁) | 1,324 | 1,684 |
| 1969 (G ₂) | 1,362 | 1,992 |
| 1970 (G ₃) | 1,407 | 2,091 |
| 1971 (G ₄) | 1,490 | 2,226 |
| 1972 (G ₅) | 1,526 | 2,485 |
| 1973 (G ₆) | 1,623 | 2,838 |
| 1974 (G ₇) | 1,797 | 3,164 |
| 1975 (G ₈) | 2,082 | 3,452 |

von STEINHAUS (1936), einem polnischen Mathematiker, verwendet, um die räumliche Verteilung eines Siedlungsystems zu beschreiben.

Die Anwendungsergebnisse dieses Ansatzes stellen Tab. 4 und Abb. 4 dar, in denen die verschiedenen Werte des Lexis-Maßes sowohl für den realen als auch für den simulierten Diffusionsprozeß für die verschiedenen Generationen des Simulationsmodells (Version 2) angegeben werden. Auf diese Weise läßt sich feststellen, daß der erforschte Diffusionsprozeß einem Entwicklungstrend folgt, der sich durch eine immer größere klumpenförmige Anordnung der Adoptoren auszeichnet. Dieser Trend läßt sich auch für Prognosezwecke benutzen. Es muß aber betont werden, daß diese Tendenz bei den simulierten Diffusionsprozessen stärker eingetreten ist. Diese Feststellung ermöglicht, eine noch bessere Übereinstim-

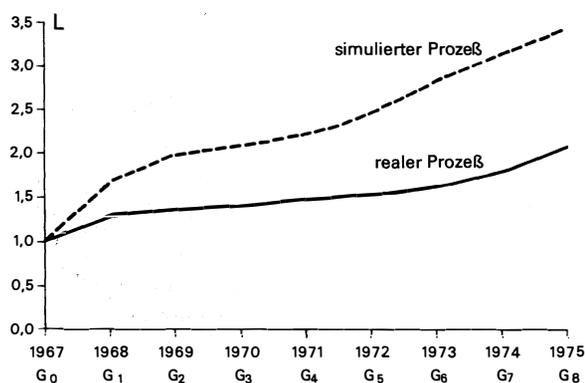


Abb. 4: Lexis-Maß für den simulierten und realen Diffusionsprozeß nach Generationen (1967-1975)

Lexis measure of the simulated and observed diffusion process according to generations (1967-1975)

mung zwischen dem angewandten Monte-Carlo-Simulationsmodell und dem untersuchten Diffusionsprozeß herzustellen und auf diese Weise zu einer Vorhersage des erforschten Phänomens überzugehen.

Literatur

- BAHRENBERG, G., GIESE, E., NIPPER, J., SCHICKHOFF, I. u. STREIT, U.: Methoden der quantitativen Geographie. Werkstattpapiere 5. Geographisches Institut der Justus Liebig-Universität Gießen. Gießen 1978.
- BIVAND, R.: Nowe podejście do problemu przekazu informacji w modelach dyfuzji. Sprawozdania PTPN za 1978 rok. Poznań 1978.
- CLIFF, A. D. a. ORD, J. K.: Spatial autocorrelation. London 1973.
- DOMAŃSKI, R. a. WIERZBICKI, A.: A simulation model for developing service centers in rural settlement network. Laxenburg: International Institute for Applied Systems Analysis, Working Paper WP-81-58, 1981.
- DRAMOWICZ, K.: Symulacja cyfrowa i analiza systemowa w badaniach procesów urbanizacji wsi. Prace Geograficzne IG i PZ PAN 112, 1975.
- GOULD, P. R.: Spatial diffusion. Resource Paper, No 4, Association of American Geographers, Washington 1969.
- HÄGERSTRAND, T.: Innovation diffusion as a spatial process. Chicago: University of Chicago Press, 1967.
- HARVEY, D. W.: Geographical processes and the analysis of point patterns: testing models of diffusion by quadrat sampling. In: Transactions Institute of British Geographers 40, 1966, 81-95.
- HOPE, A. C. A.: A simplified Monte Carlo significance test procedure. In: Journal of the Royal Statistical Society, Ser. B, 30, 1968, 582-598.
- KAMIŃSKI, Z.: Przestrzenna dyfuzja innowacji rolniczych. Uniwersytet im. A. Mickiewicza w Poznaniu, Seria Geografia 25. Poznań 1982.
- LOWRY, I.: A short course in model design. In: Journal American Institute of Planners 30, 1965, 158-166.
- MAZURKIEWICZ, L.: Zastosowanie metody symulacji w badaniach zmian przestrzennej struktury miasta (na przykładzie Wałbrzycha). Dokumentacja Geograficzna 5, 1977.
- MORRILL, R. L. a. PITTS, F. R.: Marriage, migration and the Mean Information Field. In: Annals of the Association of American Geographers 57, 1967, 401-422.
- STEINHAUS, H.: O charakterystyce skupienia osiedli. In: Czasopismo Geograficzne, H. 2-3, 1936, 288-297.
- STOUFFER, S. A.: Intervening opportunities: A theory relating mobility and distance. In: American Sociological Review 5, 1940, 845-867.
- YAPA, L. S.: On the statistical significance of the observed map in spatial diffusion. In: Geographical Analysis 8, 1976, 255-268.
- ZIPSER, T.: Modele symulacyjne wzrostu miast oparte na modelu procesu wyboru celów. In: Przegląd Geograficzny 44, 1972, 479-493.

Abb.1 Diffusion von Agrarinnovationen im südlichen Teil von Wielkopolska (Großpolen) 1967-1975

Diffusion of agricultural innovations in Southern Wielkopolska, 1967-1975

Beilage V zu ERDKUNDE 42,3 Beitrag Kaminski

