

INSTYTUT CYBERNETYKI TECHNICZNEJ

POLITECHNIKI WROCŁAWSKIEJ

Raport serii PREPRINTY Nr 104

WYKORZYSTANIE MODELI DYNAMIKI  
POPULACJI W PROGNOZOWANIU

/rozprawa doktorska/

Halina Kwaśnicka

Promotor:

prof. Zygmunt Szparkowski

Słowa kluczowe: dynamika populacji, modele matematyczne,  
symulacja, prognozowanie.

Wydawnictwa Politechniki Wrocławskiej

Wrocław 1979

S P I S T R E Ś C I

	str.
WSTĘP . . . . .	1
1. Przegląd metod prognozowania . . . . .	5
1.1. Próba charakterystyki procesów rozwoju . . . . .	5
1.2. Klasyfikacja metod prognozowania . . . . .	6
1.3. Modele substytucji w prognozowaniu . . . . .	10
2. Model pewnej klasy ewolucyjnych procesów rozwoju . . . . .	15
2.1. Wprowadzenie . . . . .	15
2.2. Model . . . . .	17
2.3. Model z reprodukcją losową o rozkładzie ciągłym . . . . .	22
2.4. Deterministyczna wersja modelu . . . . .	24
2.5. Przybliżenie liniowe wartości oczekiwanej liczby elementów w typach dla dowolnego czasu w modelu dyskretnym . . . . .	26
2.6. Dominacja jednego typu w populacji . . . . .	28
2.7. Wartość oczekiwana liczby elementów powstałych w wyniku reprodukcji jednego elementu, gdy jakości poszczególnych elementów są zmiennymi losowymi o rozkładzie gamma . . . . .	30
2.8. Zmiany w czasie udziałów w populacji elementów poszczególnych typów . . . . .	33
3. Metoda prognozowania wykorzystująca zaproponowany model ewolucyjnych procesów rozwoju . . . . .	36
3.1. Sformułowanie zadania prognozy . . . . .	36
3.2. Estymacja stanu populacji na podstawie stanu losowej próbki . . . . .	37

3.3.	Identyfikacja parametrów modelu dyskretnego . . . . .	41
3.4.	Identyfikacja parametrów modelu ciągłego. . . . .	42
3.5.	Określenie horyzontu prognozy . . . . .	43
3.6.	Wpływ niedokładności estymacji stanu populacji na prognozę . . . . .	45
3.7.	Wpływ zmian parametrów modelu na prognozę . . . . .	47
4.	Badania symulacyjne modelu. . . . .	50
4.1.	Typowe przebiegi otrzymane z modelu probabilistycznego. . . . .	50
4.2.	Szanse rozwoju elementów nowych typów . . . . .	53
4.3.	Przebiegi otrzymane przy założeniu znajomości losowej próbki z populacji . . . . .	56
4.4.	Przykładowe realizacje przy zmiennych w czasie jakościach typów . . . . .	59
4.5.	Wpływ pojemności środowiska i maksymalnego współczynnika reprodukcji na stan populacji w czasie . . . . .	60
5.	Algorytm prognozujący . . . . .	63
5.1.	Schemat blokowy algorytmu . . . . .	63
5.2.	Prognoza pewnego procesu rozwoju . . . . .	72
6.	Przykład prognozy . . . . .	76
WNIOSKI	. . . . .	85
LITERATURA	. . . . .	88

## STRESZCZENIE

W pracy przedstawiono probabilistyczny model pewnej klasy ewolucyjnych procesów rozwoju, będący wielotypowym procesem gałęzkowym. Model, budowany w oparciu o przesłanki biologiczne opisuje dynamikę populacji składającej się z elementów różnych typów, funkcjonujących w określonym środowisku zaspokajając w nim tę samą potrzebę. Elementy są różnie preferowane przez środowisko w zależności od stopnia zaspokajania potrzeby. Elementy funkcjonują w populacji przez pewien czas, podlegają losowej reprodukcji, po czym są eliminowane. Reprodukacja elementów wykazuje tendencję do zaspokojenia zapotrzebowania środowiska, limitowaną istniejącym maksymalnym współczynnikiem reprodukcji elementów. Równolegle występuje stopniowe wypieranie elementów "gorszych" typów przez elementy "lepszyc" typów. Przeprowadzono badania zaproponowanego modelu pod kątem sprawdzenia możliwości wykorzystania go do prognozowania technicznego. Zaproponowano algorytm prognozujący, który w oparciu o znaną historię zmian populacji /lub losowej próbki/ przeprowadza estymację parametrów modelu, prognozę zmian liczebności typów-estymowane wartości średnie i dyspersje, określa horyzont prognozy, umożliwiając ocenę wpływu elementów nowych typów. Pokazano przykład zastosowania algorytmu.

## W s t ę p .

Prognozowaniu rozwoju technologii, a w szczególności ilościowym metodom prognozowania poświęca się w ostatnich latach dość dużo uwagi, np. /2/, /18/, /23/, /28/. Rozwój technologii jest przy tym rozumiany jako ciągły proces dynamiczny, podczas którego istniejące formy zaspokajania potrzeb materialnych są zastępowane przez nowe, lepsze /21/. Do opisu i prognozowania rozwoju konkretnych technologii służą najczęściej modele substytucji /16/, /18/, /21/. Są one stosowane do opisu zmian procentowych udziałów na rynku produktów poszczególnych typów. Jak dotąd modele te budowane są zazwyczaj dla dwóch typów produktów w oparciu o dane dotyczące procesu wytwarzania /np. bilans kosztów/, co wymaga włączania do zbioru danych wejściowych prognozy parametrów ekonomicznych prognozowanego procesu.

Szczególne zapotrzebowanie /5/, /16/ istnieje na modele opisujące dynamikę rynku w przypadku, kiedy występuje kilka typów produktów /kilka technologii/ zaspokajających alternatywnie tę samą konkretną potrzebę konsumenta. Modele takie powinny umożliwić określenie, między innymi, które produkty i w jakim okresie zdominują rynek, a które i kiedy zostaną z niego wyparte.

Niejednokrotnie wskazuje się /2/, /19/, /30/ na analogie procesów rozwoju zachodzących w obrębie populacji biologicznych z procesami rozwoju zachodzącymi w sferze działalności ludzkiej. Sugeruje się, że modele opisujące konkurencję w populacji osobników tego samego gatunku można stosować np. do opisu konkurencji występującej na rynku między produktami zaspokajającymi tę samą potrzebę konsumenta.

W pracy proponuje się nową metodę prognozowania tego typu procesów rozwoju, wykorzystującą probabilistyczny model substytucji. Zdaniem autora, metoda może znaleźć zastosowanie w przypadkach, w których występuje wspomniana analogia między biologicznymi a technologicznymi procesami rozwoju, tzn. w podobny sposób preferowane jest lepsze "przystosowanie do potrzeb środowiska".

Procesy rozwoju będące przedmiotem zainteresowań autora można pokrótce scharakteryzować następująco:

Istnieje środowisko wykazujące zapotrzebowanie na elementy zaspokajające konkretną potrzebę i istnieje populacja elementów różnych typów, które tę potrzebę zaspokajają. Typ elementu charakteryzuje między innymi wskaźnik określający jakość "zaspokajania potrzeby". Elementy funkcjonują w populacji przez pewien czas, po czym są eliminowane; ich miejsce zajmują elementy uzyskane w procesie reprodukcji. Reprodukacja jest procesem losowym, w którym elementy o większym wskaźniku jakości są preferowane. Reprodukacja wykazuje tendencję do zaspokojenia zapotrzebowania środowiska, limitowaną istniejącym maksymalnym współczynnikiem reprodukcji elementów poszczególnych typów. Równolegle występuje tendencja do wypierania elementów gorszych typów przez elementy lepszych typów, prowadząca do zdominowania populacji przez typ o największym wskaźniku jakości.

Do opisu dynamiki takiej populacji proponuje się w niniejszej pracy probabilistyczny model matematyczny, będący wielotypowym procesem gałązkowym. Model ten, wstępnie naszkicowany w /10/ wykazuje pewne podobieństwo do modeli rozważanych w /9/, /14/, /21/. Podobny model, choć pod innym kątem rozważany jest też w pracy /15/ Model umożliwia określenie zmian stanu populacji /zdefiniowanego jako wektor o składowych równych liczbom elementów poszczegól-

nych typów w populacji/ w kolejnych chwilach czasu, przy znajomości stanu początkowego, relatywnych wartości wskaźników jakości i zapotrzebowania środowiska.

W pracy przeprowadzono badania - analityczne i symulacyjne - zaproponowanego modelu pod kątem sprawdzenia jego przydatności do prognozowania rozwoju technologii. Otrzymane rezultaty, zdaniem autora, potwierdzają taką możliwość.

Zasadniczym celem, było opracowanie metody prognozowania procesów rozwoju rozważanej klasy w oparciu o badania symulacyjne modelu.

Opracowano algorytm prognozujący na maszynę cyfrową w języku ALGOL 1900. Jako dane wejściowe do prognozowania wykorzystuje się dane o dotychczasowych zmianach stanu populacji /lub tylko zmianach stanu w losowej próbie z populacji/.

W rezultacie zastosowania algorytmu otrzymuje się dla kolejnych chwil czasu:

- przybliżenia liniowe wartości oczekiwanych liczby elementów poszczególnych typów w populacji,
- estymowane z zadanej krotności realizacji losowych średnie wartości liczby elementów poszczególnych typów i ich wariancje,
- horyzont prognozy, rozumiany jako czas, po którym wystąpią istotne zmiany prognozy /np. 10 % / wskutek pojawienia się w populacji elementów nowego typu, nie uwzględnionych przy obliczaniu prognozy.

Układ prezentowanej pracy jest następujący:

Rozdział pierwszy zawiera przegląd i krótkie charakterystyki znanych metod prognozowania. Bardziej szczegółowo omówiono kilka modeli substytucji technologicznej.

Rozdział drugi zawiera opis matematyczny, dyskusję założeń i prezentację niektórych własności modelu. Zaprezentowano trzy wersje modelu dla trzech różnych procesów reprodukcji elementów: reprodukcja losowa o rozkładzie dyskretnym /Poissona/, reprodukcja losowa o rozkładzie ciągłym /gamma/ i reprodukcja deterministyczna.

Rozdział trzeci poświęcono metodologii prognozowania z wykorzystaniem zaproponowanego modelu. Omówione są problemy związane z budową algorytmu: estymacji stanu populacji na podstawie losowej próbki, identyfikacji parametrów modelu, określenie /i zdefiniowanie/ horyzontu prognozy.

Rozdział czwarty prezentuje wybrane rezultaty badań symulacyjnych modelu.

Rozdział piąty zawiera opis opracowanego algorytmu prognozującego. Podano schemat blokowy algorytmu i sposób generowania zmiennych losowych o wymaganych rozkładach. Prosty przykład ilustruje działanie algorytmu.

Rozdział szósty zawiera przykład zastosowania algorytmu. Wykorzystano dane dotyczące zużycia energii na świecie w latach 1951-1974, do wytwarzania której służą różne surowce /różne typy energii/, /18/, /21/. Dane z lat 1951-1961 posłużyły do opracowania prognozy na lata 1961-2010.

Pracę kończą wnioski dotyczące możliwości udoskonalenia algorytmu prognozowania i ewentualnych zastosowań.

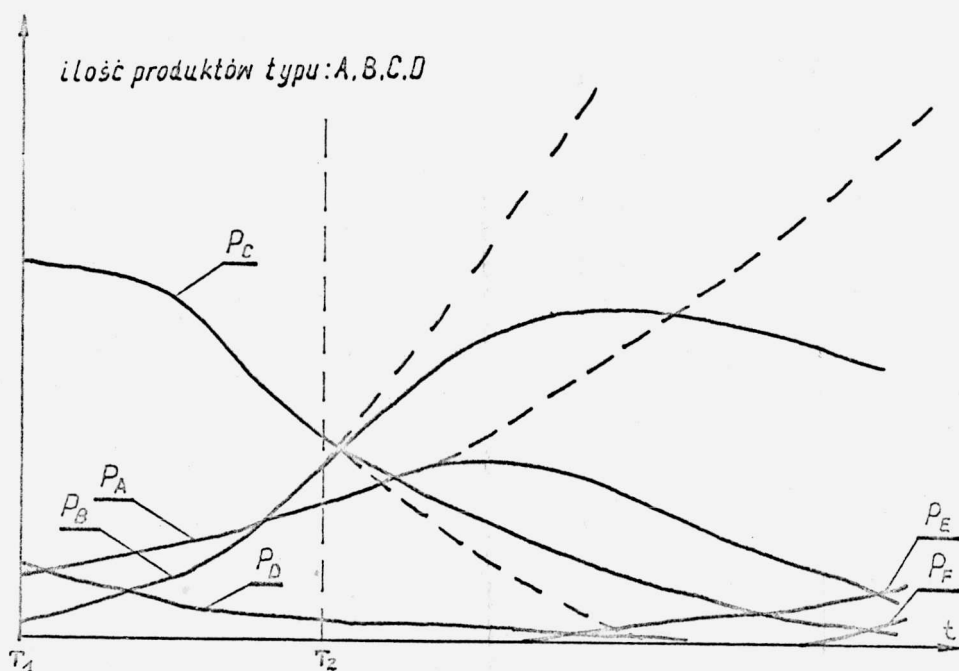


## 1. PRZEGLĄD METOD PROGNOZOWANIA

### 1.1. Próba charakterystyki procesów rozwoju.

W niniejszej pracy rozważane będzie prognozowanie zmian zachodzących w zbiorze produktów /populacji/ zaspokajających tę samą potrzebę konsumenta /środowiska/, otrzymywanych przy użyciu różnych technologii. Ilość produktów na rynku /w środowiska/, otrzymywanych różnymi technologiami /różnych typów/ będzie się zmieniać w zależności od ich preferencji przez środowisko i możliwości technologicznych.

Jak w każdym przypadku, źródłem danych wejściowych do prognozowania jest przeszłość. Przyjmijmy, że znamy historię zmian ilościowych populacji produktów  $P$ , otrzymywanych przy użyciu technologii  $A/P_A/$ , technologii  $B/P_B/$ ,  $C/P_C/$ , oraz  $D/P_D/$ , w przedziale czasu  $t \in [T_1, T_2]$ . Niech ilości produktów poszczególnych typów zmieniają się jak na rys. 1.1.



Rys. 1.1. Historia i możliwe prognozy ilościowych zmian produktów czterech typów.

Widać, że w przedziale czasu  $(T_1, T_2)$  poszczególne krzywe są typu:  $P_A, P_B$  - eksponencjalny wzrost,  $P_C$  - logistyczny spadek,  $P_D$  - eksponencjalny spadek. Aproksymacja krzywych dla  $t > T_2$  daje prognozę zaznaczoną na rysunku liniami przerywanymi. Maleje do zera ilość  $P_C$  i  $P_D$ , natomiast ilość produktów typu A i B rośnie do nieskończoności. Prognoza w takiej postaci jest nie do przyjęcia, gdyż ilość produktów P na rynku zwiększałaby się nieograniczenie.

Z danych rzeczywistych wynika, że najczęściej krzywe zmian ilościowych produktów poszczególnych typów mają charakter epidemiologiczny - linie ciągłe na rys. 1.1. /5/, /17/, /27/. Wzrost ilości produktów otrzymywanych przy użyciu danej technologii jest początkowo powolny, następnie, gdy zostaną usunięte trudności - szybko wzrasta, a wreszcie ulega nasyceniu, ponieważ zostaje osiągnięty górny limit praktycznych możliwości tej technologii. Po pewnym czasie technologia ta zostaje wypierana przez nową, która z kolei przechodzi przez te same etapy /na rys. 1.1. - krzywe  $P_E, P_F$ /.

Istnieje szereg metod prognozowania. Wydaje się, że większość z nich, szczególnie metody prognozowania nie jakościowego, a ilościowego są mało dopracowane. Progności często opracowują metody tylko na własny użytek, służące do rozwiązania jednego, konkretnego zadania /m.in. tak powstają modele substytucji/.

## 1.2. Klasyfikacja metod prognozowania.

Dotychczas nie istnieje końcowa, ogólnie zaakceptowana klasyfikacja metod prognozowania. W niniejszym rozdziale przedstawiony jest krótki przegląd metod prognozowania w oparciu o klasyfikację zaproponowaną przez J.M.Cetrona /5/:

1. Prognozy intuicyjne - tworzone w oparciu o opinię jednego lub grupy ekspertów.

a) Indywidualne - zasięgana jest opinia jednego eksperta.

b) Kolektywne - zasięgana jest opinia grupy ekspertów:

(1) przez głosowanie /Pools/ - opinia ekspertów uzgadniana jest przez głosowanie,

(2) panelowe /Panels/ - eksperci uzgadniają opinię w czasie bezpośredniej dyskusji,

(3) delphi /Delphi/ - przy zastosowaniu tej metody nie dopuszcza się do bezpośredniego kontaktu ekspertów, co pozwala wyeliminować takie mankamenty jak:

- dominacja silnych indywidualności w grupie,
- dążenie do konformizmu,
- niechęć do publicznej zmiany stanowiska, itp.

Metody z tej grupy są najprostszymi, a jednocześnie szeroko stosowanymi metodami tworzenia prognoz. Są to metody jakościowe, pozbawione obiektywizmu. Za najbardziej obiektywną z tej grupy metod uważana jest metoda delphi.

2. Ekstrapolacja trendów - dotyczy przewidywania rozwoju wyłącznie bieżącej technologii, oparta jest na zasadzie ciągłości /23/, tzn. jeżeli kolejne metody wytwarzania powodowały określone, regularne tempo wprowadzania ulepszeń, to tempo to będzie utrzymane.

a) Prosta - kontynuacja tego co było, tak jak było, bez uwzględnienia jakichkolwiek czynników. Jest metodą niedokładną. W literaturze podaje się jako jej zaletę - obiektywizm.

b) Krzywe trendu - jakościowa analiza krzywych trendu dla szczególnych przypadków. Uwzględnianie możliwych czynników ograniczających wzrost; w przypadku gdy te czynniki są znane - szacowanie wartości górnej granicy wzrostu.

c) Równania - matematyczny opis trendu przy pomocy równania liniowego, parabolicznego, eksponencjalnego lub logistycznego. Wybór typu równania, identyfikacja parametrów równania - najczęściej metodą najmniejszych kwadratów.

Podstawowa trudność przy stosowaniu metod tej grupy - problem górnej granicy trendu. Jeżeli znamy naturalną granicę /nasyćenie/ np. 100% wydajności, nie należy przedkładać trendu powyżej niej. Drugi poważny problem związany jest z wyborem parametrów do prognozowania. Podstawą prognozowania trendów technicznych jest najczęściej wykres określonego parametru technicznego względem czasu. Parametr powinien być /wg /2/ /:

- 1° operacyjny - dający się zmierzyć w mających obiektywne znaczenie kategoriach ilościowych,
- 2° musi rzeczywiście odzwierciedlać poziom możliwości funkcjonalnych prognozowanej techniki,
- 3° musi dać się zastosować do różnych technik /możliwości porównania wyrobów opartych na różnych technikach/,
- 4° taki, dla którego są dostępne odpowiednie dane historyczne.

Dla przykładu, złe parametry to: wzmocnienie lub szerokość pasma wzmacniacza elektronicznego, szybkość samolotu, itp. Dobre parametry /były dla nich przeprowadzane prognozy w USA /5/ /23/ /- tonomile/h i pasażeromile/h przy prognozowaniu zdolności przewozowych samolotów transportowych, lumen/wat - wydajność źródeł światła, może być parametrem iloczyn wzmocnienia i szerokości pasma przy prognozowaniu rozwoju wzmacniaczy elektronicznych.

3. Korelacja trendów - czasami trend złożonego parametru technicznego może być łatwiej wyrazić jako wynik współzależności pomiędzy dwoma lub większą ilością trendów /3/.

a) Poprzedzających zdarzenia - wyznaczenie korelacji trendów pomiędzy dwoma procesami rozwoju, z których jeden jest "poprzednikiem" drugiego /np. szybkość samolotów transportowych i wojskowych w [mile/h] /. Problem w dobraniu trendów, między którymi rzeczywiście istnieje współzależność.

b) Analiza korelacji - w praktycznych problemach często jest ważne znalezienie aktualnej relacji pomiędzy dwoma lub więcej trendami /niekoniecznie czasowo-zależnymi/. Dla więcej niż dwóch zmiennych niezależnych stosuje się wielowymiarową analizę regresji. W przypadku, kiedy może być znaleziona stała utrzymująca się korelacja, metody tej grupy dają obiektywne podejście do prognozowania. W tych metodach wykorzystuje się modelowanie matematyczne. W oparciu o analizę obiektu konstruuje się model, następnie dokonuje się estymacji aktualnej charakterystyki rozwoju /krzywe trendów/ i identyfikacji parametrów modelu. Do grupy tej można zaliczyć modele substytucji, które są szerzej omówione w rozdz. 1.3.

4. Analogie - metody wykorzystujące analogie postępu technicznego do innych procesów.

a) Wzrostu - oparte na analogii biologicznej /2/. Lenz zauważa, że równanie R.Pearla, opisujące wzrost populacji biologicznej jako funkcji czasu w ograniczonym środowisku /np. muszki w butelce, komórki drożdży w ustalonej pożywce, białe myszki na skończonej przestrzeni/ jest identyczne jak równanie krzywej logistycznej postępu technicznego, zaproponowane przez de Solla Price.

b) Historyczne - interesuje prognozę wpływ nowej technologii na istniejące /z tej samej dziedziny/. Ocenia ten wpływ w oparciu o podobne zajścia w historii, np. General Electric Company robiło prognozę do 2000 r. zmian udziałów w zużytej w USA energii otrzymywanej z różnych surowców, na skutek wprowadzenia paliw nuklearnych. Użyto tego samego typu krzywych wzrostu, które były obserwowane dla paliw kopalnych i mocy elektrowni wodnych w latach 1880 - 1960, /23/.

### 1.3. Modele substytucji w prognozowaniu.

Jednym z często spotykanych modeli substytucji dwóch "konkurentów" jest model Fisher-Pry'a /16/. Niech  $f$  - oznacza udział procentowy na rynku technologii wypierającej /  $f < 1$  /.  $(1-f)$ -oznacza udział procentowy technologii wypieranej.

Podstawę tego modelu stanowi równanie wyprowadzone w oparciu o hipotezę, że udział na rynku technologii wypierającej wzrasta proporcjonalnie do udziału jeszcze nie całkiem wyeliminowanej z rynku technologii wypieranej. Postulowana przez autorów relacja może być zapisana w postaci równania różniczkowego

$$\frac{1}{f} \cdot \frac{df}{dt} = k \cdot (1-f) \quad , \quad (1.3.1)$$

gdzie  $k$  - stała, definiowana jako roczny współczynnik wzrostu udziału nowej technologii podczas kilku pierwszych lat substytucji, /otrzymywana na podstawie danych historycznych/.

Oznaczając przez  $u$  - czas, dla którego  $f=0.5$ , otrzymuje się jedną z wersji krzywej logistycznej /krzywej typu " S " /:

$$f = \frac{1}{1 + \exp(k(u-t))} \quad . \quad (1.3.2)$$

W równaniu na  $f$  występują dwa parametry, które należy identyfikować:  $k, u$ .

Wadą tego modelu jest to, że uwzględnia on udziały tylko dwóch konkurujących technologii, procesów lub produktów. Spotyka się inne zarzuty pod adresem tego modelu /16/, np to, że przy jego budowie autorzy nie uwzględnili przesłanek ekonomicznych, takich jak koszty, użyteczność dla specyficznych celów, itp.

Bardzo podobnym do omawianego wyżej modelu jest model Ayres-Noble-Overly'ego, przy budowie którego uwzględniono mechanizm kosztów. Autorzy wprowadzają współczynniki reprezentujące

użyteczność jednostki kosztów obu konkurujących technologii /28/. Otrzymane przez autorów końcowe równanie na udział na rynku technologii wypierającej w zasadzie nie różni się w postaci ani w interpretacji parametrów od równania Fisher-Pry'a.

Stern /28/ opracował model substytucji, w którym nie wykorzystuje punktu  $f=0.5$ . Autor sugeruje, że model ten może również funkcjonować dla więcej niż dwóch konkurujących technologii. Stern podzielił cały rynek na części w oparciu o różne wykorzystania produktu. Włączone są również do modelu współczynniki użyteczności jednostki kosztów. W modelu Sterna uwzględnione są jeszcze takie parametry jak:

- liczba jednostek produkowanych w ciągu roku produktu I i II,
- cena rynkowa produktu I,
- czas pojawienia się produktu II,
- czas, po którym cena produktu II zrówna się z ceną produktu I /w okresie wprowadzania na rynek produktu II, jego cena będzie wyższa niż produktu I /,
- stała zanikania, równa średniej czasu życia /w latach/ całkowitego wyposażenia sektora przemysłu produktu I,
- współczynnik rocznego wzrostu produktu I /jak  $k$  w równ. 1.3.1/,
- "czas uczenia" = ilość lat, po których produkcja produktu II będzie na tyle opanowana, by mogła być łatwo wprowadzana.

Jak twierdzi sam autor /16/, użyteczność tego modelu jest znacznie ograniczona wymaganą dużą liczbą danych technicznych i ekonomicznych.

W pracy /24/ autorzy wskazują na potrzebę wykorzystania modelowania matematycznego do prognozowania, widząc w tym dwa główne cele: 1° opis dynamiki rynku w oparciu o produkt - technologię w przypadku, kiedy w tym samym okresie czasu jeden bądź więcej typów produktów

jest zastępowanych przez dowolną liczbę nowych typów produktów, które zaspokajają szczególną potrzebę konsumentów, 2° rozwijać systematycznie metodologię włączającą dynamikę i cechy adaptatywności do modeli ekstrapolacji trendów.

W pracy /21/ przedstawiony jest logistyczny model substytucji, zbudowany w oparciu o analizę bilansu kosztów. Autorzy zajmują się również udziałami na rynku produktów poszczególnych typów a nie ich bezwzględnymi wielkościami. Model uwzględnia dowolną liczbę konkurujących technologii.

Udział produktów i-tego typu wyraża się wzorem (wyr. 4.23 w /21/)

$$f_i(t) = \frac{1}{1 + \sum_{j \neq i} \frac{f_{j0}}{f_{i0}} \exp(-c_{ji}(t-t_0))} \quad , \quad (1.3.3)$$

gdzie:  $f_{j0}$  - udział na rynku produktu j-tego typu w  $t=t_0$ ,

$c_{ji}$  - współczynnik zależny od różnicy kosztów produktów j-tego i i-tego typu.

Autorzy analizują dalej model probabilistyczny, wykazując że model deterministyczny wyznacza medianę rozkładu prawdopodobieństwa dla  $f_i(t)$ . Autorzy przeprowadzają identyfikację parametrów modelu wykorzystując model probabilistyczny, natomiast do prognozowania stosują model deterministyczny. Jako wynik prognozy podają medianę udziałów  $f_i(t)$ . Do rezultatów otrzymanych przy wykorzystaniu tego modelu będzie nawiązane w rozdziale 5 niniejszej pracy, przy prezentowaniu przykładu.

Interesującym wydaje się fakt, że otrzymany przez autorów model deterministyczny, budowany w oparciu o bilans kosztów jest analogiczny do modeli występujących w biologii; modelu M. Eigena, dotyczącego ewolucji makromolekuł /7/, /29/ i modelu R.A.Fishera opisującego zmiany częstości występowania w populacji m alleli tego



samego genu, przy działaniu naturalnej selekcji /9/. Model Eigena zmian stężenia makromolekuł j-tego typu /przy założeniu, że wszystkie cząstki powielane są dokładnie/ ma postać (wyr.II.48 w/7/):

$$\dot{x}_j = k_0 (E_j - \bar{E}) \cdot x_j \quad (1.3.4)$$

W pracy /21/ model zmian udziału produktu j-tego typu ma postać (wyr. 4.13 w /21/):

$$\dot{f}_j = (\varphi(t) - c_j/\alpha) \cdot f_j \quad (1.3.5)$$

Porównując wyrażenia /1.3.4/ i /1.3.5/ widoczna jest odpowiedniość modeli, przedstawiona w tabeli 1.

TABELA 1.

Analogia modelu M. Eigena i V. Peterka-F. Flecka /21/.

Model M.Eigena		Model V.Peterka-F.Flecka	
zmienna	interpretacja	zmienna	interpretacja
$x_j$	Stężenie molekuł j-tego typu.	$f_j$	Udział na rynku produktów j-tego typu.
$E_j$	"Produktywność" cząstek typu j.	$c_j$	Koszt związany z produkcją jednostki produktu typu j.
$\bar{E}$	$\bar{E} = \left( \frac{\sum_{i=1}^N E_i \cdot x_i}{\sum_{i=1}^N x_i} \right) -$ średnia produktywność cząstek.	$\alpha \cdot \varphi(t)$	$\alpha \cdot \varphi(t) = \sum_{i=1}^N c_i \cdot f_i$ - średnie koszty produktów.

W równaniach /1.3.4/ i /1.3.5/ parametry  $E_j$  i  $c_j$  oraz  $\bar{E}$  i  $\alpha\varphi(t)$  mają przeciwne znaki. W modelu Eigena preferowane są cząstki o większej produktywności, natomiast w /1.3.5/ preferowane są produkty o mniejszych kosztach jednostkowych, stąd wspomniana różnica w znakach.

W  $t \rightarrow \infty$ , w modelu Eigena zdominują cząstki jednego typu - o największej produktywności; w modelu substytucji - zdominują produkty o najniższych kosztach jednostkowych. Mimo, że oba modele budowane były niezależnie od siebie, wykazują duże analogie między sobą.

Podobna analogia występuje pomiędzy przedstawianym modelem substytucji technologicznej a modelem Fishera. Powyższe analogie wydają się potwierdzać hipotezę, że modele dynamiki populacji biologicznych mogą być wykorzystywane do prognozowania technicznego.

## 2. MODEL PEWNEJ KLASY EWOLUCYJNYCH PROCESOW ROZWOJU

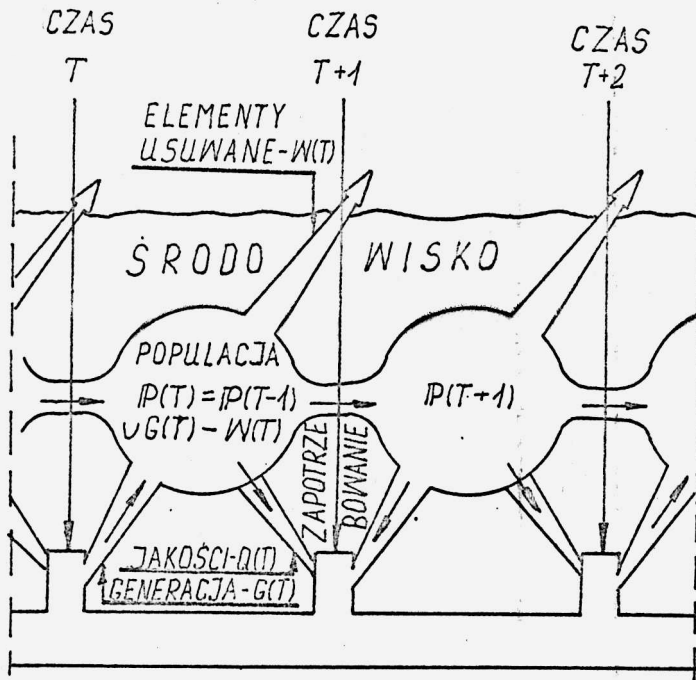
### 2.1. Wprowadzenie.

Rozpatrujemy populację składającą się z elementów o różnych typach. Elementy populacji funkcjonują w określonym środowisku, zaspokajając tę samą potrzebę. Elementy w różnym stopniu mogą zaspokajać daną potrzebę i w zależności od tego być różnie preferowane przez środowisko. Zakładamy specyficzne środowisko, w którym:

- ocena stopnia zaspokojenia potrzeby przez każdy element jest zmienną losową o jednakowym rozkładzie prawdopodobieństwa dla elementów tego samego typu, funkcjonujących w środowisku w tym samym czasie. Miarę stopnia zaspokojenia potrzeby będziemy nazywać jakością elementu,
- jakości poszczególnych elementów są zmiennymi losowymi, niezależnymi dla różnych elementów.

Będziemy nazywać środowisko bardziej jednorodnym, im mniejsza będzie wariancja jakości elementów.

Środowisko ma ograniczoną pojemność. Elementy są wprowadzane do populacji, funkcjonują w niej przez określony czas, po czym są eliminowane. Zapotrzebowanie środowiska jest określone przez różnicę między jego pojemnością a aktualną liczbą elementów należących do populacji /tzn. funkcjonujących w środowisku/. Rozpatrywany proces jest zilustrowany na rys. 2.1.1. Zakładamy dyskretny czas. W każdym przedziale czasu, w środowisku istnieje populacja. Składa się ona z tych elementów, które "przeżyły" z populacji wcześniejszej i z nowej generacji elementów, powstałych w wyniku procesu reprodukcji. Reprodukacja elementów jest procesem losowym, z tendencją do zaspokojenia zapotrzebowania środowiska



Rys. 2.1.1. Rozwój populacji.

oraz preferującym elementy o większej jakości. Elementy poszczególnych typów są reprodukowane niezależnie. Dla każdego typu istnieje maksymalny współczynnik reprodukcji, który limituje stopień zaspokojenia środowiska przy jego dużym zapotrzebowaniu. Kiedy maksymalny współczynnik reprodukcji nie odgrywa istotnej roli /przy niezbyt dużym zapotrzebowaniu środowiska na nowe elementy/, elementy są reprodukowane w taki sposób, że oczekiwana wartość ilości wszystkich nowych elementów jest równa zapotrzebowaniu środowiska oraz, dla poszczególnych typów, wartość oczekiwana współczynnika reprodukcji jest proporcjonalna do ich jakości. Interesować się będziemy zmianą stanu generacji w czasie, tzn. ilością wprowadzanych do populacji elementów poszczególnych typów w kolejnych przedziałach czasowych. Jeżeli "czas życia" elementów populacji wynosi jedną, przyjętą przez nas jednostkę czasu, wówczas możemy uważać, że przy zmianie czasu z  $t$  na  $(t+1)$  zachodzi proces reprodukcji, po czym populacja z czasu  $t$  "ginie", natomiast w  $(t+1)$  funkcjonują tylko nowo powstałe elementy. Zapotrzebowanie

środowiska na nowe elementy jest teraz równe jego pojemności. W tej sytuacji, tym co będzie nas interesować są zmiany stanu populacji w czasie, tzn. zmiany w czasie liczby elementów poszczególnych typów, wchodzących w skład populacji.

W dalszej części pracy zajmować się będziemy takimi właśnie populacjami, pamiętając jednakże, że słowo "populacja" będzie równoznaczne ze słowem "generacja" dla sytuacji omówionej wcześniej, zilustrowanej na rys. 2.1.1.

W większości przypadków środowisko, w którym funkcjonować będą interesujące nas populacje jest w pełni jednorodne, tzn. takie, w którym elementy tego samego typu są oceniane jednakowo. Ponadto wykazano, że w pewnych przypadkach losowość oceny jakości elementów może być nie do rozróżnienia od losowości reprodukcji elementów /rozdział 2.7 i 4.4/.

Omawiany proces jest wielotypowym procesem gałęzowym, z dyskretnym czasem. Pojęcia używane w niniejszej pracy, takie jak "potomek", "populacja" itp. używane są podobnie jak w teorii procesów gałęzowych /8/, /12/, /20/, /25/.

## 2.2. Model.

Wprowadzimy następujące oznaczenia:

$t = 0, 1, 2, \dots$  - dyskretny czas,

$M(t)$  - pojemność środowiska w chwili  $t$ ,

$P(t)$  - populacja w chwili  $t$ ,

$N_i(t)$  - liczba elementów  $i$ -tego typu, należących do populacji w chwili  $t$ ,

$e_{i_k}$  -  $k$ -ty element  $i$ -tego typu, należący do  $P(t)$ ,

$m$  - liczba typów wchodzących w skład populacji,

$N_{\Sigma}(t)$  - liczba elementów wszystkich typów, należących do  $P(t)$ ,

$q_i(t)$  - jakość elementów  $i$ -tego typu, należących do  $P(t)$ ,  $q_i(t) \in \mathbb{R}_+$ ,

$f(\eta_{ik}(t), E[\eta_{ik}(t)]; \eta_{ik}(t) \in \mathcal{N}$  - rozkład prawdopodobieństwa dla procesu reprodukcji k-tego elementu i-tego typu, należącego do  $\mathcal{P}(t)$ , gdzie:  $\forall t, i = \overline{1, m}, k = \overline{1, N_i(t)}$ ;  $\eta_{ik}(t)$  - zmienna losowa dyskretna, o rozkładzie  $f$ , określająca liczbę potomków elementu  $e_{ik}$ , należącego do populacji w chwili  $t$ ,

$r_i$  - maksymalny współczynnik reprodukcji dla elementów i-tego typu,

$S(t)$  - stan populacji w chwili  $t$

$$S(t) \stackrel{\text{def}}{=} (N_1(t), N_2(t), \dots, N_m(t)).$$

Przyjmujemy następujące założenia:

1. Środowisko jednorodne.
2. Dla dwóch dowolnych elementów  $e_{ik}$ ,  $e_{jr}$  należących do  $\mathcal{P}(t)$ , zmienne losowe  $\eta_{ik}(t)$ ,  $\eta_{jr}(t)$  są zmiennymi losowymi niezależnymi.
3. Warunkowa wartość oczekiwana ilości potomków każdego elementu  $e_{ik}$  należącego do  $\mathcal{P}(t)$  jest proporcjonalna do jego jakości:

$$\forall t; i = \overline{1, m}; k = \overline{1, N_i(t)}: E[\eta_{ik}(t) | S(t)] = c \cdot q_i(t),$$

gdzie  $c > 0$ , współczynnik proporcjonalności.

4. Warunkowa wartość oczekiwana wielkości populacji w  $(t+1)$  jest proporcjonalna do wielkości populacji w  $t$

$$E[N_\Sigma(t+1) | S(t)] = g \cdot N_\Sigma(t),$$

gdzie  $g$  - funkcja zależna od  $r_i$ , czasu i pojemności środowiska. W szczególności może również zależeć od  $N_\Sigma(t)$ ,

5.  $r_i = r$ ,  $\forall i = \overline{1, m}$ , jednakowy maksymalny współczynnik reprodukcji dla elementów wszystkich typów.

6. W interesującym nas przedziale czasu, pojemność środowiska jest wielkością stałą:  $M(t) = M$ .

7.  $q_i(t) = q_i$ ,  $\forall i = \overline{1, m}$ , stałe w czasie jakości elementów poszczególnych typów /założenie to nie będzie obowiązujące w rozdz. 2.7 i 4.4/.

8. Stan populacji w chwili początkowej jest wielkością zdeteminowaną.

Ponieważ  $N_{\Sigma}(t+1) = \sum_{e_{i_k} \in P(t)} \eta_{i_k}(t)$ , możemy wyznaczyć:

$$E[N_{\Sigma}(t+1)|S(t)] = \sum_{e_{i_k} \in P(t)} E[\eta_{i_k}(t)|S(t)] = c \cdot \sum_{i=1}^m N_i(t) \cdot q_i \quad (2.2.1)$$

Definiujemy średnią jakość populacji w sposób następujący:

$$q_{sr}(t) \stackrel{df}{=} \frac{\sum_{i=1}^m N_i(t) \cdot q_i}{N_{\Sigma}(t)} \quad (2.2.2)$$

Zatem

$$E[N_{\Sigma}(t+1)|S(t)] = c \cdot q_{sr}(t) \cdot N_{\Sigma}(t) \quad (2.2.3)$$

Porównując założenie 4 i powyższe równanie, możemy wyznaczyć współczynnik proporcjonalności, występujący w założeniu 3:

$$c = g/q_{sr}(t).$$

Założenie 3 przyjmuje więc następującą postać:

Zał. 3<sup>a</sup>.

$$E[\eta_{i_k}(t)|S(t)] = \frac{g}{q_{sr}(t)} \cdot q_i$$

Funkcja tworząca prawdopodobieństwa stanu populacji w chwili  $t$  jest równa z definicji

$$F(\bar{s}, t) = \sum_{S(t)} P[S(t)] \cdot \bar{s}^{S(t)}, \quad (2.2.4)$$

gdzie:  $P[S(t)]$  - prawdopodobieństwo tego, że w chwili  $t$  populacja będzie w stanie  $S(t)$ ,

$\bar{s}$  -  $m$ -wymiarowy wektor parametrów zespolonych,  $\bar{s} = (s_1, s_2, \dots, s_m)$ ,

$\sum_{S(t)}$  - suma po wszystkich możliwych stanach populacji,

$$\sum_{S(t)} \equiv \sum_{N_m(t)=0}^{\infty} \sum_{N_{m-1}(t)=0}^{\infty} \dots \sum_{N_1(t)=0}^{\infty}$$

$$(\bar{s})^{S(t)} \stackrel{\text{df}}{=} \prod_{i=1}^m (s_i)^{N_i(t)}$$

Funkcja tworząca prawdopodobieństwa stanu populacji w  $(t+1)$  jest

$$F(\bar{s}, t+1) = \sum_{S(t+1)} P[S(t+1)] \cdot \bar{s}^{S(t+1)} = \sum_{S(t+1)} \left( \sum_{S(t)} P[S(t)] \cdot P[S(t+1)|S(t)] \right) \cdot \bar{s}^{S(t+1)}$$

$$F(\bar{s}, t+1) = \sum_{S(t)} P[S(t)] \cdot G(\bar{s}, t+1) \quad , \quad (2.2.5)$$

gdzie:  $G(\bar{s}, t+1)$ -warunkowa funkcja tworząca prawdopodobieństwa stanu populacji.

$$G(\bar{s}, t+1) \stackrel{\text{df}}{=} \sum_{S(t+1)} P[S(t+1)|S(t)] \cdot \bar{s}^{S(t+1)}, \quad (2.2.6)$$

$$G(\bar{s}, t+1) = \sum_{S(t+1)} \prod_{i=1}^m P[N_i(t+1)|S(t)] \cdot (s_i)^{N_i(t+1)} = \prod_{i=1}^m G_i(s_i, t+1) \quad (2.2.7)$$

$G_i(s_i, t+1)$ -warunkowa funkcja tworząca prawdopodobieństwa  $N_i(t+1)$  elementów  $i$ -tego typu w chwili  $(t+1)$ .

Oznaczmy przez  $H_i(s_i, t+1)$  funkcję tworzącą prawdopodobieństwa liczby potomków jednego elementu  $i$ -tego typu /funkcja ta jest dla elementów tego samego typu jednakowa/. Ponieważ

$$N_i(t+1) = \sum_{k=1}^{N_i(t)} \eta_{ik}(t) \quad , \quad (2.2.8)$$

oraz reprodukcja elementów jest procesem niezależnym, możemy zapisać:

$$G(\bar{s}, t+1) = \prod_{i=1}^m (H_i(s_i, t+1))^{N_i(t)} \quad (2.2.9)$$

Charakter omawianego procesu wskazuje na to, iż jest on łańcu-



chem Markowa, w którym stanem procesu jest stan populacji.

Prawdopodobieństwo przejścia ze stanu  $S(t)$  do stanu  $S(t+1)$  jest równe współczynnikowi przy  $\bar{s}^{S(t)}$  we wzorze /2.2.9/.

Na podstawie wyrażenia /2.2.9/ można wyznaczyć warunkowe wartości oczekiwane i wariancje ilości elementów poszczególnych typów

$$E[N_i(t+1)|S(t)] = \left. \frac{\partial G(\bar{s}, t+1)}{\partial s_i} \right|_{\bar{s} = \mathbb{1}}$$

Ponieważ  $\left. \frac{\partial H_i(s_i, t+1)}{\partial s_i} \right|_{s_i=1} \stackrel{\text{zał. 3a}}{=} \frac{g}{q_{sr}(t)} \cdot q_i$ ,

różniczkując /2.2.9/ otrzymujemy:

$$E[N_i(t+1)|S(t)] = \frac{N_i(t) \cdot g \cdot q_i}{q_{sr}(t)}, \quad (2.2.10)$$

$$D^2[N_i(t+1)|S(t)] = \left[ \frac{\partial^2 G(\bar{s}, t+1)}{\partial s_i^2} + \frac{\partial G(\bar{s}, t+1)}{\partial s_i} - \left( \frac{\partial G(\bar{s}, t+1)}{\partial s_i} \right)^2 \right] \Bigg|_{\bar{s} = \mathbb{1}}$$

$$D^2[N_i(t+1)|S(t)] = N_i(t) \cdot D^2[\eta_{ik}(t)|S(t)], \quad (2.2.11)$$

gdzie:  $D^2[\eta_{ik}(t)|S(t)]$  - wariancja zmiennej losowej liczby potomków elementu  $e_{ik}$  należącego do  $\mathbb{P}(t)$ .

W niniejszej pracy przyjęto, że rozkład prawdopodobieństwa zmiennej losowej  $\eta_{ik}$  jest rozkładem Poissona, o wartości oczekiwanej, jak w założeniu 3a

$$E[\eta_{ik}(t)|S(t)] = \frac{g \cdot q_i}{q_{sr}(t)} = \lambda_i(t),$$

$$P(\eta_{ik}(t) = l) = \frac{[\lambda_i(t)]^l}{l!} \cdot e^{-\lambda_i(t)} = \frac{\left( \frac{g \cdot q_i}{q_{sr}(t)} \right)^l}{l!} \cdot e^{-\frac{g \cdot q_i}{q_{sr}(t)}} \quad (2.2.12)$$

Funkcja tworząca prawdopodobieństwa  $p(\lambda, l) = e^{-\lambda} \cdot \frac{\lambda^l}{l!}$  ma postać  $f(s, \lambda) = e^{-\lambda(1-s)}$  zatem

$$H_i(s_i, t+1) = \exp\left(-\frac{g \cdot a_i}{a_{br}(t)} \cdot (1-s_i)\right), \quad (2.2.13)$$

$$G_i(s_i, t+1) = \exp\left(-\frac{N_i(t) \cdot g \cdot a_i}{a_{br}(t)} \cdot (1-s_i)\right) \quad (2.2.14)$$

Z powyższego wzoru wynika, że zmienna losowa  $N_i(t+1)$  jest również zmienną losową o rozkładzie Poissona, z wartością oczekiwaną, daną przez /2.2.10/. Mamy więc dany rozkład prawdopodobieństwa dla procesów reprodukcji typów w populacji /zmiennie losowe  $N_i(t+1)$   $i=1, m$ /. Podobnie, zmienna losowa wielkości populacji  $N_Z(t+1)$  jako suma zm. losowych o rozkładzie Poissona jest zmienną losową o tym samym rozkładzie i wartości oczekiwanej, jak w zał. 4. Można wykazać, że jeżeli  $g=1$  /środowisko nasycone/, to zmiana liczby elementów w całej populacji w czasie jest tzw. krytycznym procesem Galton-Watson'a /12/. Wartość oczekiwana w tym procesie jest stała w czasie i wynosi  $N_Z(0)$ , a wariancja rośnie w czasie zgodnie z wyrażeniem:  $D^2[N_Z(t)|N_Z(0)] = t \cdot N_Z(0)$ .

### 2.3. Model z reprodukcją losową o rozkładzie ciągłym.

W niektórych procesach rzeczywistych, które wydają się być nieźle opisywane przy pomocy przedstawionego modelu trudno jest zdefiniować element, natomiast można mówić o jednostkach. Zasadnicza różnica polega na tym, że ilość jednostek w całej populacji /jak również w poszczególnych typach/ nie musi być wielkością całkowitą. W biologii procesem takim będzie np. proces ilościowych zmian biomasy fenotypów /11/, /9/, /14/; w procesach technologicznych - np. zmiany ilości wyprodukowanej energii przez elektrownie różnych typów /np. wodne, jądrowe itp./.

W takich procesach wygodniej jest mówić o reprodukcji typów w populacji. Dla ustalonego czasu  $t$ ,  $N_1(t), N_2(t), \dots, N_m(t)$  są zmiennymi losowymi typu ciągłego, mogą przyjmować wartości ze zbioru  $\mathcal{R}_+$ . Taki model będzie dalej nazywany modelem ciągłym. Będzie się on różnił od modelu zaproponowanego w rozdziale 2.3 /dyskretnego/ tylko rozkładem prawdopodobieństwa dla procesu reprodukcji typów. Model ciągły można opisać analogicznie do dyskretnego przy pomocy funkcji charakterystycznych. Poczynione w rozdziale 2.2 założenia są nadal aktualne z tym, że zał. 3 przyjmuje postać wyrażenia /2.2.10/.

Oznaczmy przez  $\phi_k(u_k, t+1)$  warunkową funkcję charakterystyczną liczebności  $k$ -tego typu w populacji w chwili  $(t+1)$ . Wtedy funkcja charakterystyczna stanu populacji w  $(t+1)$ ,  $\phi(\bar{u}, t+1)$  wynosi

$$\phi(\bar{u}, t+1) = \prod_{k=1}^m \phi_k(u_k, t+1) \quad , \quad (2.3.1)$$

gdzie  $\bar{u} = (u_1, u_2, \dots, u_m)$  - wektor parametrów zespolonych.

Podobnie:

$$E[N_j(t+1)|S(t)] = \left. \frac{\partial \phi(\bar{u}, t+1)}{\partial u_j} \cdot \frac{1}{i} \right|_{\bar{u}=\bar{0}}$$

$$D^2[N_j(t+1)|S(t)] = \left. \frac{\partial^2 \phi(\bar{u}, t+1)}{\partial u_j^2} \cdot \frac{1}{i^2} \right|_{\bar{u}=\bar{0}} - \left( \left. \frac{1}{i} \frac{\partial \phi(\bar{u}, t+1)}{\partial u_j} \right|_{\bar{u}=\bar{0}} \right)^2$$

Ponieważ zmienne losowe  $N_j(t+1)$   $j=\overline{1, m}$  mogą przyjmować wartości ze zbioru liczb rzeczywistych dodatnich, zatem przyjęty dla nich rozkład powinien być określony na dodatniej półosi i ciągły.

W dalszych rozważaniach przyjęto rozkład gamma:

$$f[N_i(t+1)|S(t)] = \frac{b^p}{\Gamma(p)} [N_i(t+1)]^{p-1} \cdot e^{-b \cdot N_i(t+1)} \quad (2.3.2)$$

$$E[N_j(t+1)|S(t)] = p/b = \frac{N_j(t) \cdot a_j \cdot g}{a_{j\tau}(t)} = \lambda_j(t) ;$$

$$D^2[N_j(t+1)|S(t)] = p/b^2 ,$$

$$\phi_k(u_k, t+1) = \frac{1}{(1 - i \cdot u_k / b)^p} = \frac{1}{(1 - i \cdot u_k / b)^{b \cdot \lambda_k(t)}}$$

Warunkowa funkcja charakterystyczna stanu populacji w (t+1) wynosi

$$\Phi(\bar{u}, t+1) = \prod_{k=1}^m \frac{1}{(1 - i u_k / b)^{b \cdot \lambda_k(t)}}$$

Warunkowa funkcja charakterystyczna jednowymiarowej zmiennej

losowej  $N_\Sigma(t+1) = \sum_{k=1}^m N_k(t+1)$  wynosi  $\Phi_\Sigma(u, t+1)$ :

$$\phi_\Sigma(u, t+1) = \prod_{k=1}^m \phi_k(u, t+1) = \frac{1}{(1 - i \cdot u / b)^{b \cdot N_\Sigma(t) \cdot g}} \quad (2.3.3)$$

$N_\Sigma(t+1)$  - jest zmienną losową o rozkładzie gamma, z wartością oczekiwaną  $N_\Sigma(t) \cdot g$  i wariancją  $N_\Sigma(t) \cdot g / b$ .

Zarówno w modelu ciągłym jak i dyskretnym, wyznaczenie rozkładów prawdopodobieństwa dla zmiennych losowych  $N_i(t)$ ,  $i = \overline{1, m}$  w dowolnej chwili czasu  $t$ , w zależności od stanu początkowego populacji  $S(0)$  jest zadaniem trudnym. Zmusza to do wyznaczania stanów populacji w kolejnych chwilach.

#### 2.4. Deterministyczna wersja modelu.

Przedstawiony w rozdziale 2.2. ogólny model probabilistyczny sugeruje jego wersję deterministyczną. Przyjmując, że proces reprodukcji elementów zachodzi w sposób deterministyczny, zgodnie z wartościami oczekiwanymi danymi równaniem /2.2.10/ otrzymuje się deterministyczny model zmian stanu populacji w kolejnych chwilach czasowych. Aby model był w pełni zdefiniowany należy

przyjąć postać  $g$  w założeniu 4. Zarówno w modelach populacji biologicznych /6/, /26/ jak i w procesach technologicznych /2/, /1/, /5/, /21/, /23/ przyjmowany jest wzrost logistyczny. W niniejszej pracy również przyjęto, że wielkość populacji zmienia się logistycznie, "dążąc" do zaspokojenia zapotrzebowania środowiska.

Jednym z możliwych sposobów przedstawienia równania zał. 4 w postaci logistycznej jest przyjęcie następującej funkcji  $g$  :

$$g(t) = 1 + A \cdot (1 - N_{\Sigma}(t)/M). \quad (2.4.1)$$

Liczba elementów w populacji w  $\{t+1\}$  będzie określona następującym równaniem

$$N_{\Sigma}(t+1) = [1 + A \cdot (1 - N_{\Sigma}(t)/M)] \cdot N_{\Sigma}(t), \quad (2.4.2)$$

w którym  $(1+A) = r$  - maksymalny współczynnik reprodukcji.

Można wykazać, że jeżeli maksymalny współczynnik reprodukcji w równaniu /2.4.2/ będzie mniejszy od 2, tzn.  $A < 1$ , to populacja zbliża się do nasycenia /stanu równowagi/ bez oscylacji; jeżeli  $2 < A+1 < 3$ , to występują oscylacje zbieżne (/26/, str.43,44).

Zgodnie z równaniem /2.2.10/ liczby elementów poszczególnych typów będą się zmieniać według następującego równania

$$\forall t; i = \overline{1, m} : N_i(t+1) = \frac{N_i(t) \cdot q_{ji}}{q_{br}(t)} \cdot [1 + A \cdot (1 - \frac{N_{\Sigma}(t)}{M})] \quad (2.4.3)$$

Powyższe równania w pełni definiują model deterministyczny.

Znając parametry modelu i początkowy stan populacji, można wyznaczyć stan populacji w dowolnej chwili czasowej. Można wykazać:

$$N_i(t) = \frac{N_i(0) \cdot (q_{ji})^t}{\sum_{j=1}^{m} N_j(0) \cdot (q_{ji})^t} \cdot N_{\Sigma}(t-1) \cdot [1 + A \cdot (1 - \frac{N_{\Sigma}(t-1)}{M})] \quad (2.4.4)$$

W modelu deterministycznym nie zachodzi konieczność wyznaczania stanów populacji w kolejnych chwilach czasowych. W powyższym

równaniu występuje wielkość populacji w chwili  $(t-1)$ .  $N_{\Sigma}$  można również wyznaczyć dla dowolnego  $t$  na podstawie znajomości  $N_{\Sigma}(0)$  zgodnie z następującym wyrażeniem:

$$N_{\Sigma}(t) = N_{\Sigma}(0) \cdot [1 + A \cdot (1 - \frac{N_{\Sigma}(0)}{M}) \cdot [1 + A \cdot (1 - \frac{N_{\Sigma}(0)}{M}) [1 + \dots \cdot [1 + A \cdot (1 - \frac{N_{\Sigma}(0)}{M})]^1]^2 \dots]^t]^{t-1}]^t \quad (2.4.5)$$

2.5. Przybliżenie liniowe wartości oczekiwanej liczby elementów w typach dla dowolnego czasu w modelu dyskretnym.

Zajmiemy się modelem probabilistycznym, w którym reprodukcja jest procesem losowym o rozkładzie Poissona.

Funkcja tworząca stanu populacji  $S(t+1)$  wyraża się wzorem

$$F(\bar{s}, t+1) = \sum_{S(t+1)} P[S(t+1)] \cdot \prod_{i=1}^m (\bar{s}_i)^{N_i(t+1)} \quad (2.5.1)$$

Korzystając z niezależności reprodukcji elementów many

$$P[S(t+1)] = \sum_{S(t)} \prod_{i=1}^m P[N_i(t+1)|S(t)] \cdot P[S(t)]$$

Zatem

$$F(\bar{s}, t+1) = \sum_{S(t)} P[S(t)] \cdot \prod_{i=1}^m G_i(\bar{s}_i, t+1) \quad (2.5.2)$$

Z rozdz. /2.2./:

$$G_i(\bar{s}_i, t+1) = [H_i(\bar{s}_i, t+1)]^{N_i(t)} = \exp[-N_i(t) \cdot \lambda_i(t) \cdot (1 - \bar{s}_i)]$$

gdzie  $\lambda_i(t)$  wartość oczekiwana  $\eta_{ik}$ ,  $k = \overline{1, N_i(t)}$ .

Podstawiając do /2.5.2/ mamy

$$F(\bar{s}, t+1) = \sum_{S(t)} P[S(t)] \cdot \prod_{i=1}^m \exp \left[ -N_i(t) \cdot \frac{q_i \cdot g(t) \cdot N_{\Sigma}(t)}{\sum_{j=1}^m N_j(t) \cdot q_j} \cdot (1 - \bar{s}_i) \right] \quad (2.5.3)$$

Oznaczmy przez

$$\gamma_i(t) = \frac{N_i(t) \cdot q_i \cdot g(t) \cdot N_{\Sigma}(t)}{\sum_{j=1}^m N_j(t) \cdot q_j} \quad (2.5.4)$$

$f_i(t)$  jest nieliniową funkcją  $N_i(t)$ . Rozwińmy ją w szereg Taylora w otoczeniu wybranego a priori punktu  $S_0(t) = [N_{01}(t), \dots, N_{0m}(t)]$ . Ograniczmy się do dwóch pierwszych wyrazów rozwinięcia

$$f_i(t) \approx c_i(t) + \sum_{j=1}^m d_{ij}(t) \cdot [N_j(t) - N_{0j}(t)] \quad (2.5.5)$$

gdzie:

$$c_i(t) = f_i(t) \Big|_{S_0(t)} = \frac{N_{0i}(t) \cdot g_i \cdot N_{0m}(t) \cdot g(t)}{\sum_{j=1}^m N_{0j}(t) \cdot g_j}, \quad d_{ij} = \frac{\partial f_i(t)}{\partial N_j(t)} \Big|_{S_0(t)}$$

Podstawiając do /2.5.3/ zamiast  $f_i(t)$  jej rozwinięcie w szereg, otrzymujemy

$$\begin{aligned} F(\bar{s}, t+1) &= \sum_{S(t)} P[S(t)] \cdot \prod_{i=1}^m \exp[-c_i(t) - \sum_{j=1}^m d_{ij}(t) [N_j(t) - N_{0j}(t)] (1-s_i)] = \\ &= \prod_{i=1}^m \exp[-(c_i(t) - \sum_{j=1}^m d_{ij}(t) \cdot N_{0j}(t)) (1-s_i)] \cdot \sum_{S(t)} P[S(t)] \cdot \\ &\quad \cdot \prod_{i=1}^m \exp[-\sum_{j=1}^m d_{ij}(t) \cdot N_j(t) (1-s_i)] \end{aligned} \quad (2.5.6)$$

Ostatni iloczyn w /2.5.6/ można zapisać w następującej postaci

$$\begin{aligned} \prod_{i=1}^m \exp[-\sum_{j=1}^m d_{ij}(t) \cdot N_j(t) (1-s_i)] &= \prod_{j=1}^m \left( \exp[-\sum_{i=1}^m d_{ij}(t) \cdot (1-s_i)] \right)^{N_j(t)} \\ F(\bar{s}, t+1) &= \prod_{i=1}^m \exp[-(c_i(t) - \sum_{j=1}^m d_{ij}(t) \cdot N_{0j}(t)) (1-s_i)] \cdot \\ &\quad \cdot \sum_{S(t)} P[S(t)] \cdot \prod_{j=1}^m \left( \exp[-\sum_{i=1}^m d_{ij}(t) \cdot (1-s_i)] \right)^{N_j(t)} \end{aligned} \quad (2.5.7)$$

Oznaczmy nowy wektor parametrów zespolonych

$$\bar{s}' = (s'_1, s'_2, \dots, s'_m), \quad \text{taki, że: } \forall_{j=1, \dots, m}; s'_j = \exp[-\sum_{i=1}^m d_{ij}(t) (1-s_i)] \quad (2.5.8)$$

Wtedy

$$F(\bar{s}, t+1) = \prod_{i=1}^m \exp[-(c_i(t) - \sum_{j=1}^m d_{ij}(t) N_{0j}(t)) (1-s_i)] \cdot \sum_{S(t)} P[S(t)] \cdot \prod_{j=1}^m (s'_j)^{N_j(t)} \quad (2.5.9)$$

Z definicji 
$$\sum_{S(t)} P[S(t)] \cdot \bar{x}^{S(t)} = F(\bar{x}, t) \quad (2.5.10)$$

Zatem

$$F(\bar{x}, t+1) = \prod_{i=1}^m \exp[-c_i(t) - \sum_{j=1}^m d_{ij}(t) \cdot N_{oj}(t)] \cdot (1-s_i) \cdot F(\bar{x}, t) \quad (2.5.11)$$

Różniczkując /2.5.11/ po  $s_k$  przy  $\bar{x} = 1$ :

$$E[N_k(t+1)] = c_k(t) - \sum_{j=1}^m d_{kj}(t) N_{oj}(t) + \sum_{j=1}^m d_{kj}(t) \cdot E[N_j(t)] \quad (2.5.12)$$

Przyjmijmy, że punkt  $S_o(t)$ , w otoczeniu którego  $\gamma_i(t)$  było rozwinięte w szereg Taylora wybrany był tak, że  $\forall i=1, m: N_{oi}(t) = E[N_i(t)]$ .

Wtedy:

$$E[N_k(t+1)] = c_k(t) = \frac{E[N_k(t)] \cdot q_k \cdot E[N_{\Sigma}(t)] \cdot g(t)}{\sum_{j=1}^m E[N_j(t)] \cdot q_j} \quad (2.5.13)$$

Powyższe wyrażenie jest zgodne z modelem deterministycznym, przedstawionym w rozdziale 2.4. W rozdziale 4 zaprezentowane będą eksperymenty wskazujące na zgodność modelu deterministycznego z estymowanymi wartościami średnimi zmian liczebności poszczególnych typów w modelu probabilistycznym, zarówno dyskretnym jak i ciągłym.

Badania symulacyjne wskazują na podobny typ zachowania się modeli deterministycznego i probabilistycznych, dlatego też wydaje się możliwe wnioskowanie o zachowaniu się modeli probabilistycznych na podstawie własności modelu deterministycznego.

## 2.6. Dominacja jednego typu w populacji.

W rozdziale 2.2. zdefiniowane było pojęcie średniej jakości populacji jako /wyr. 2.2.2/

$$q_{sr}(t) = \frac{\sum_{i=1}^m N_i(t) \cdot q_i}{N_{\Sigma}(t)}$$

Wykażemy, że w zdefiniowanym w rozdz. 2.4. modelu, średnia ja-



kość populacji jest rosnącą funkcją czasu, aż do osiągnięcia maksimum. W tym celu obliczymy różnicę  $q_{\text{śr}}(t+1) - q_{\text{śr}}(t)$

$$q_{\text{śr}}(t+1) - q_{\text{śr}}(t) = \frac{\sum_{i=1}^m N_i(t+1)q_i}{N_{\Sigma}(t+1)} - q_{\text{śr}}(t) = \frac{\sum_{i=1}^m \frac{N_i(t)}{N_{\Sigma}(t)} \cdot (q_i)^2}{q_{\text{śr}}(t)} - q_{\text{śr}}(t) \quad (2.6.1)$$

ale  $\frac{N_i(t)}{N_{\Sigma}(t)} = p_i(t)$  oznacza prawdopodobieństwo znalezienia w populacji w chwili  $t$  elementu o jakości  $q_i$ , zatem wariancja jakości w populacji, w czasie  $t$  wynosi:

$$\sigma^2(t) = \sum_{i=1}^m p_i(t) \cdot [q_i - q_{\text{śr}}(t)]^2 = \sum_{i=1}^m p_i(t) (q_i^2 - [q_{\text{śr}}(t)]^2) \quad (2.6.2)$$

Wyrażenie /2.6.1/ przyjmuje więc następującą postać:

$$q_{\text{śr}}(t+1) - q_{\text{śr}}(t) = \frac{\sigma^2(t) + [q_{\text{śr}}(t)]^2}{q_{\text{śr}}(t)} - q_{\text{śr}}(t) = \frac{\sigma^2(t)}{q_{\text{śr}}(t)} \geq 0, \forall t \quad (2.6.3)$$

Powyższe wyrażenie wykazuje, że średnia jakość populacji jest funkcją rosnącą, aż do momentu, w którym  $\sigma^2(t)$  będzie równa zeru. Wystąpi to wtedy, gdy wszystkie elementy w populacji będą miały tę samą jakość. Inaczej mówiąc, populacja będzie się składała tylko z elementów jednego typu.

Wiemy, że liczby elementów poszczególnych typów zmieniają się zgodnie z wyrażeniem /2.4.3/. Możemy policzyć granice  $N_i(t)$  dla  $i = \overline{1, m}$  przy  $t \rightarrow \infty$ . Okazuje się, że:

$$\text{dla takiego } k, \text{ że } \max_{j=\overline{1, m}} q_j = q_k, \quad \lim_{t \rightarrow \infty} N_k(t) = \lim_{t \rightarrow \infty} N_{\Sigma}(t) = M, \quad \text{oraz}$$

$$\text{dla } i = \overline{1, m}, \quad i \neq k: \quad \lim_{t \rightarrow \infty} N_i(t) = 0.$$

Wobec powyższego, po dostatecznie długim czasie populacja będzie składać się z elementów tego typu, którego jakość jest największa.

Maksymalna wartość średniej jakości populacji jest równa  $q_k$ :

$\max_t q_{\text{śr}}(t) = q_k$ . W przypadku gdy mamy środowisko nasycone /tzn.

$\forall t: N_{\Sigma}(t) = M$  i funkcja  $g(t) = 1$  /, równania /2.4.3/ przyj-

mują postać

$$\forall_{t,i=\overline{1,m}}; N_i(t+1) = \frac{N_i(t) \cdot q_i}{q_{sr}(t)}; N_i(t+1) - N_i(t) = N_i(t) \cdot [q_i/q_{sr}(t) - 1] \quad (2.6.4)$$

Liczba elementów danego typu rośnie w tym przedziale czasu, w którym  $q_i < q_{sr}$ , zaczyna maleć z chwilą zrównania się średniej jakości populacji z jego jakością. W środowisku nienasyconym:

$$N_i(t+1) - N_i(t) = N_i(t) \cdot \left[ \frac{q_i [1 + A(1 - N_\Sigma(t)/M)]}{q_{sr}(t)} - 1 \right] \quad (2.6.5)$$

Z powyższego równania wynika, że przy dostatecznie dużym niedosyceniu środowiska, liczba elementów nawet tych najgorszych /o najmniejszej jakości/ będzie się zwiększać. Jako pierwszy zaczyna być eliminowany z populacji typ o najmniejszej jakości, dla którego najwcześniej jest spełniony warunek

$$q_i [1 + A(1 - N_\Sigma(t)/M)] > q_{sr}(t)$$

Powyższe własności będą zilustrowane w rozdziale 4.1.

2.7. Wartość oczekiwana liczby elementów powstałych w wyniku reprodukcji jednego elementu, gdy jakości poszczególnych elementów są zmiennymi losowymi o rozkładzie gamma.

Zakładamy, że dla ustalonej chwili czasu  $t$ ,  $\forall e_{ik} \in P(t)$

$q_{e_{ik}}$  - oznacza jakość  $k$ -tego elementu  $i$ -tego typu, oraz,  $q_{e_{ik}}$  - zmienna losowa o rozkładzie  $\Gamma(p_i, b)$ , z wartością oczekiwaną jednakową dla wszystkich elementów tego samego typu, równą  $q_i$  / $p_i = q_i \cdot b$  / i wariancją  $p_i/b^2$  / $i=\overline{1,m}$ /. Zakładamy, że  $b$  - jest jednakowe dla wszystkich elementów w populacji.

Założenie 3<sup>a</sup> /rozd. 2.2./ mówi, że warunkowa wartość oczekiwana zmiennej losowej  $\eta_{ik}$  /liczby potomków  $e_{ik}$ / jest proporcjonalna do jakości elementów  $i$ -tego typu i odwrotnie proporcjonalna do średniej jakości populacji, przy czym jakości elementów

były deterministyczne. Przy założeniu losowości jakości elementów, założenie to przyjmuje postać:

$$E[\eta_{ik}(t) | S(t), Q] = \frac{q_{e_{ik}}}{\sum_{e_{ij} \in P(t)} q_{e_{ij}}} \cdot g(t) \cdot N_{\Sigma}(t) \quad (2.7.1)$$

gdzie:  $Q = N_{\Sigma}(t)$  - wymiarowy wektor jakości elementów  $e \in P(t)$ :

$$Q = [q_{e_{11}}, \dots, q_{e_{1, N_1(t)}}, \dots, q_{e_{m1}}, \dots, q_{e_{m, N_m(t)}}]$$

Chcemy wyznaczyć  $E[\eta_{ik}(t) | S(t)]$ . W tym celu zajmiemy się zmienną losową:

$$U = \frac{q_{e_{ik}}}{\sum_{e_{ij} \in P(t)} q_{e_{ij}}} \quad (2.7.2)$$

Wprowadźmy oznaczenia:

$X$  - zmienna losowa  $q_{e_{ik}}$ ,

$Y$  - zm. losowa równa sumie jakości wszystkich elementów, z wyjątkiem  $e_{ik}$ :

$$Y = \sum_{\substack{e_{ij} \in P(t) \\ e_{ij} \neq e_{ik}}} q_{e_{ij}} = \sum_{i=1}^m \sum_{j=1}^{N_i(t)} q_{e_{ij}} - q_{e_{ik}}$$

Zmienną  $U$  możemy przedstawić jako

$$U = \frac{X}{X+Y}$$

gdzie:

$X$  - ma rozkład  $\Gamma(p_i, b) = \Gamma(p_1, b)$ ,

$Y$  - ma rozkład  $\Gamma(\sum_{j=1}^m N_j \cdot p_j - p_i, b) = \Gamma(p_2, b)$ .

Znajdziemy rozkład zmiennej losowej  $U$ :

$$\begin{cases} U = \frac{X}{X+Y} \Rightarrow Y = X/U = X \cdot V \\ V = X \quad X = V \end{cases} \quad |J| = V/U^2,$$

$$f(x, y) = f_1(x) \cdot f_2(y)$$

$$\begin{aligned} h(u) &= \int_0^{\infty} \frac{b^{p_1}}{\Gamma(p_1)} v^{p_1-1} e^{-bv} \cdot \frac{b^{p_2}}{\Gamma(p_2)} \left( \frac{v \cdot (1-u)}{u} \right)^{p_2-1} \cdot e^{-b \cdot \frac{v(1-u)}{u}} \cdot \frac{v}{u^2} dv = \\ &= \frac{b^{p_1+p_2}}{\Gamma(p_1) \cdot \Gamma(p_2)} \cdot \left( \frac{1-u}{u} \right)^{p_2-1} \cdot \frac{1}{u^2} \cdot \int_0^{\infty} v^{p_1+p_2-1} \cdot e^{-v \frac{b}{u}} dv \end{aligned}$$

Podstawiając  $z=b \cdot v/u$  wyznaczymy całkę

$$\int_0^{\infty} v^{p_1+p_2-1} e^{-v \frac{b}{u}} dv = (u/b)^{p_1+p_2} \Gamma(p_1+p_2)$$

Zatem

$$h(u) = \frac{\Gamma(p_1+p_2)}{\Gamma(p_1) \cdot \Gamma(p_2)} \cdot \frac{u^{p_1+p_2} \cdot (1-u)^{p_2-1}}{u^2 \cdot u^{p_2-1}} = \frac{1}{\beta(p_1, p_2)} \cdot (1-u)^{p_2-1} \cdot u^{p_1-1} \quad (2.7.3)$$

Zmienna losowa  $U$  ma rozkład beta z parametrami  $p_1, p_2$ .

Dla rozkładu beta

$$m_1 = E[u] = p_1 / (p_1 + p_2) \quad (2.7.4)$$

Możemy teraz wyznaczyć warunkową wartość oczekiwaną zmiennej losowej  $\eta_{ii}(t) | S(t)$

$$\begin{aligned} E[\eta_{ii}(t) | S(t)] &= \sum_{k=0}^{\infty} k \cdot \int_0^1 \frac{[g(t) \cdot N_{\Sigma}(t) \cdot u]^k}{k!} e^{-g(t) \cdot N_{\Sigma}(t) \cdot u} \cdot f(u) du = \\ &= g(t) \cdot N_{\Sigma}(t) \cdot \int_0^1 u \cdot f(u) \cdot du = g(t) \cdot N_{\Sigma}(t) \cdot E[u], \end{aligned}$$

$$E[\eta_{ii}(t) | S(t)] = \frac{g(t) \cdot N_{\Sigma}(t) \cdot q_i}{\sum_{j=1}^m N_j(t) \cdot q_j} \quad (2.7.5)$$

Warunkowa wartość oczekiwana zmiennej losowej  $\eta_{ii}(t) | S(t)$  jest jednakowa dla  $i=1, \dots, N_1(t)$ , proporcjonalna do wartości oczekiwanej jakości elementów  $i$ -tego typu, a odwrotnie proporcjonalna do oczekiwanej wartości średniej jakości populacji.

Wyrażenie /2.7.5/ jest zgodne z zał. 3<sup>B</sup>. Wyznaczenie rozkładu prawdopodobieństwa dla  $\eta_{ii}(t)$  wymaga wyznaczenia skomplikowanych całek, czego nie udało się zrobić w niniejszej pracy.

2.8. Zmiany w czasie udziałów w populacji elementów poszczególnych typów.

Założmy, że interesuje nas, jak w czasie zmieniają się udziały w populacji elementów różnych typów, natomiast liczby elementów w całej populacji i w poszczególnych typach nie są ważne. Postawmy pytanie: czy zmiana pojemności środowiska oraz maksymalnego współczynnika reprodukcji /który jest jednakowy dla wszystkich elementów/ będzie miała wpływ na krzywe udziałów elementów różnych typów w populacji. Zajmiemy się modelem deterministycznym:

$$N_i(t) = \frac{N_i(0) \cdot (q_i)^t}{\sum_{j=1}^m N_j(0) \cdot (q_j)^t} \cdot N_{\Sigma}(t)$$

$u_i(t) = \frac{N_i(t)}{N_{\Sigma}(t)}$  udział elementów i-tego typu w  $P(t)$ ,

$$u_i(t) = \frac{N_i(0) \cdot (q_i)^t}{\sum_{j=1}^m N_j(0) \cdot (q_j)^t} = \frac{u_i(0) \cdot (q_i)^t}{\sum_{j=1}^m u_j(0) \cdot (q_j)^t}, \quad i = \overline{1, m} \quad (2.8.1)$$

$$\sum_{i=1}^m u_i(t) = 1$$

Zmiany w czasie udziałów elementów poszczególnych typów w populacji, w modelu deterministycznym nie są zależne od pojemności środowiska i maksymalnego współczynnika reprodukcji. Model deterministyczny opisujący zmiany udziałów typów w populacji zachowuje się tak, jak model /deterministyczny/ opisujący zmiany stanu populacji w przypadku, gdy wielkość populacji jest stała w czasie / $q(t) = 1$ /. Omówione własności tego modelu dają się przenieść na model /2.8.1/. Należy jeszcze sprawdzić, jak będzie w przypadku probabilistycznym. Czy rozkłady prawdopodobieństwa dla  $u_i(t+1)/\bar{u}(t)$ ,  $\bar{u}(t) = [u_1(t), u_2(t), \dots, u_m(t)]$  będą zależne od

pojemności środowiska i parametru  $A/a$  więc maksymalnego współczynnika reprodukcji/. Ze względu na stopień trudności przy wyznaczaniu rozkładów dla  $u_i(t+1)/\bar{u}(t)$  w modelu dyskretnym, ograniczymy się do modelu ciągłego, przedstawionego w rozdz. 2.3. Reprodukacja elementów jest więc procesem losowym o rozkładzie gamma.

$$\forall i=1, m : u_i(t+1) = \frac{N_i(t+1)}{N_\Sigma(t+1)} = \frac{N_i(t+1)}{N_i(t+1) + \sum_{\substack{j=1 \\ j \neq i}}^m N_j(t+1)}$$

Wiadomo /rów. 2.3.2/, że  $f[N_i(t+1) | S(t)]$  jest rozkładem  $\Gamma(p_1, b)$ ,

$$p_1 = b \cdot \frac{N_i(t) \cdot q_{i1}}{\sum_{j=1}^m N_j(t) \cdot q_{j1}} \cdot N_\Sigma(t) \cdot g(t) = \frac{b \cdot u_i(t) \cdot q_{i1}}{\sum_{j=1}^m u_j(t) \cdot q_{j1}} \cdot N_\Sigma(t) [1 + A \cdot (1 - N_\Sigma(t)/M)] \quad (2.8.2)$$

Warunkowy rozkład zmiennej losowej  $\sum_{\substack{j=1 \\ j \neq i}}^m N_j(t+1) | S(t)$  jest  $\Gamma(p_2, b)$ ,

$$p_2 = \sum_{\substack{j=1 \\ j \neq i}}^m p_j = b \cdot \frac{\sum_{\substack{k=1 \\ k \neq i}}^m u_k(t) \cdot q_{k1}}{\sum_{j=1}^m u_j(t) \cdot q_{j1}} \cdot N_\Sigma(t) [1 + A \cdot (1 - N_\Sigma(t)/M)] \quad (2.8.3)$$

Z rozdziału 2.7 wiadomo, że zmienna losowa  $(u_i(t+1)/\bar{u}(t), N_\Sigma(t))$  będzie mieć rozkład beta z parametrami  $p_1, p_2$  określonymi przez równania /2.8.2/ i /2.8.3/.

$$h(u_i(t+1) | \bar{u}(t), N_\Sigma(t)) = \frac{1}{\beta(p_1, p_2)} \cdot [1 - u_i(t+1)]^{p_2 - 1} \cdot [u_i(t+1)]^{p_1 - 1} \quad (2.8.4)$$

$$m_1 = \frac{p_1}{p_1 + p_2}, \quad \mu_2 = \frac{p_1 \cdot p_2}{(p_1 + p_2)^2 \cdot (p_1 + p_2 + 1)}$$

Aby wyznaczyć rozkład  $h(u_i(t+1)/\bar{u}(t))$  należałoby znać rozkład  $f(N_\Sigma(t)/N_\Sigma(0))$ . Znany jest rozkład  $N_\Sigma(t)/N_\Sigma(t-1)$ , natomiast rozkładu dla  $N_\Sigma(t)/N_\Sigma(0)$  autorowi nie udało się znaleźć.

Na podstawie wzoru /2.8.4/ możemy powiedzieć, że parametry rozkładu  $u_i(t+1)/\bar{u}(t)$  w istotny sposób zależą od wielkości populacji,

a więc i od pojemności środowiska i maksymalnego współczynnika reprodukcji. Jedyne na warunkową wartość oczekiwaną  $u_i(t+1)/\bar{u}(t), N_{\Sigma}(t)$ , pojemność środowiska i parametr  $A$  nie będą miały wpływu, natomiast wariancja i wyższe momenty zależą od  $N$  i od  $A$ .

$$E[u_i(t+1)|\bar{u}(t)] = \frac{u_i(t) \cdot q_i}{\sum_{j=1}^m u_j(t) q_j} \quad (2.8.5)$$

Można się spodziewać podobnych rezultatów przy analizie rozkładów  $u_i(t+1)/\bar{u}(t)$  w modelu dyskretnym.

### 3. METODA PROGNOZOWANIA WYKORZYSTUJĄCA ZAPROPONOWANY MODEL EWOLUCYJNYCH PROCESÓW ROZWOJU

#### 3.1. Sformułowanie zadania prognozy.

Przy opracowywaniu zaprezentowanej w niniejszym rozdziale metody prognozowania uwzględniono sytuację, w której dla osoby przeprowadzającej prognozę dostępne są dane dotyczące zmian zachodzących w losowej próbce z populacji. Na podstawie zmian stanu populacji w losowej próbce, proponowany algorytm powinien przeprowadzić estymację zmian stanu całej populacji w tym okresie czasu, a otrzymane rezultaty wykorzystać do prognozowania jako znaną historię.

Zakładamy, że wielkość całej populacji jest w tym przedziale czasu znana. W przeciwnym przypadku można ją oszacować stosując np. metodę podwójnego losowania (/8/, str.49).

Opracowano dwie wersje algorytmu prognozującego, różniące się tylko wykorzystanym modelem probabilistycznym: algorytm I wykorzystuje model dyskretny /z rozdz. 2.2/, algorytm II - model ciągły /przedstawiony w rozdz. 2.3/. Obie wersje algorytmu wykorzystują również model deterministyczny. Zadania stawiane przed algorytmem /dla I i II / są następujące:

- na podstawie losowej próbki przeprowadzić estymację zmian stanu populacji /wykonać, jeżeli znana jest historia próbki a nie populacji/,
- na podstawie zadanej historii zmian stanu populacji przeprowadzić identyfikację parametrów modelu /metodą największej wiarygodności /,



- wyznaczyć estymowane wartości oczekiwane i odchylenia standardowe wielkości populacji oraz liczby elementów poszczególnych typów w żądanym okresie prognozowania,
- wyznaczyć zmiany stanu populacji w żądanym okresie prognozy, przy wykorzystaniu modelu deterministycznego,
- określić horyzont prognozy, zdefiniowany jako czas, po którym liczebności poszczególnych typów zmienią się o 10 % swej wartości, wskutek wprowadzenia do populacji elementów nowego typu.

Liczbę elementów nowo wprowadzanego typu, jego jakość /w stosunku do najlepszego z istniejących/ oraz czas wprowadzenia go do populacji zadaje prognozujący jako dane wejściowe.

W prognozie uwzględnione będą tylko elementy tych typów, które istnieją w populacji w chwili przeprowadzania prognozy. Przedstawiony w rozdziale 2 model nie uwzględnia możliwości pojawienia się w populacji elementów nowych typów. Inna wersja tego modelu, uwzględniająca możliwość powstawania elementów nowych typów w procesie reprodukcji jest przedstawiona w /10/, /15/.

Problemy związane z zastosowaniem zaproponowanego modelu do prognozowania to:

1. Estymacja stanu populacji na podstawie stanu losowej próbki.
2. Identyfikacja parametrów modelu
  - a/ dyskretnego,
  - b/ ciągłego.
3. Określenie horyzontu prognozy.

### 3.2. Estymacja stanu populacji na podstawie stanu losowej próbki.

Znane są:

$N_2(t)$  w przedziale czasu, oznaczmy go umownie od 0 do T,

$R(t)$  dla  $t \in [0, T]$  - wielkość próbki losowej w chwili  $t$ ,  
 $k_i(t)$  dla  $t \in [0, T]$ ,  $i = \overline{1, m}$  - liczebność elementów  $i$ -tego typu  
w próbie liczącej  $R(t)$  elementów,

Szukane:  $N_i(t)$  dla  $t \in [0, T]$ ,  $i = \overline{1, m}$ .

Prawdopodobieństwo, że w próbie zawierającej  $R(t)$  elementów  
z populacji  $N_\Sigma(t)$ -elementowej, zawierającej:

$N_1(t)$  - elementów 1 typu,

$N_2(t)$  - elementów 2 typu,

⋮

$N_m(t)$  - elementów  $m$ -tego typu, będzie odpowiednio:

$k_1(t)$  - elementów 1 typu,

⋮

$k_m(t)$  - elementów typu  $m$ , dane jest rozkładem hipergeometrycznym  
/schemat losowania bez zwracania/

$$P(k_1, k_2, \dots, k_m) = \frac{\binom{N_1}{k_1} \binom{N_2}{k_2} \dots \binom{N_{m-1}}{k_{m-1}} \binom{N_\Sigma - \sum_{i=1}^{m-1} N_i}{R(t) - \sum_{i=1}^{m-1} k_i}}{\binom{N_\Sigma(t)}{R(t)}} \quad (3.2.1).$$

Przyjmujemy estymator największego prawdopodobieństwa /8/. W po-  
pulacji wyróżniamy tylko typ 1:

$$P(k_1, \sum_{i=2}^m k_i) = \frac{\binom{N_1}{k_1} \cdot \binom{N_\Sigma - N_1}{R - k_1}}{\binom{N_\Sigma}{R}} \quad (3.2.2)$$

Prawdopodobieństwo to jest funkcją posiadającą jedno maksimum  
ze względu na  $N_1$  /8/.

Wyznaczamy takie  $N_1(t)$ , dla którego

$$\frac{P(k_1, \sum_{i=2}^m k_i) \text{ dla } N_1}{P(k_1, \sum_{i=2}^m k_i) \text{ dla } (N_1-1)} \gg 1 \quad (3.2.3)$$

Po przekształceniach otrzymujemy:  $N_1(t) = \frac{k_1(t)}{R(t)} \cdot N_{\Sigma}(t)$ .

Przyjmujemy nowe:  $N_2'(t) = N_2(t) - N_1(t)$ ;  $R'(t) = R(t) - k_1(t)$

i wyróżniamy typ drugi. Otrzymujemy:

$$P(k_2, \sum_{i=3}^m k_i) = \frac{\binom{N_2}{k_2} \cdot \binom{N_2' - N_2}{R' - k_2}}{\binom{N_2'}{R'}} \xrightarrow{\max_{N_2(t)}} :$$

$$N_2(t) = \frac{k_2(t)}{R'(t)} \cdot N_2'(t) = \frac{k_2(t)}{R(t) - k_1} \cdot (N_{\Sigma}(t) - \frac{k_1 \cdot N_{\Sigma}(t)}{R(t)}) = \frac{k_2(t) \cdot N_{\Sigma}(t)}{R(t)} \quad (3.2.4)$$

Postępowanie takie przeprowadzamy aż do  $N_{m-1}(t)$ .

Następnie wyznaczamy  $N_m(t)$  jako różnicę:

$$N_m(t) = N_{\Sigma}(t) - \sum_{i=1}^{m-1} N_i(t) = \frac{k_m(t) \cdot N_{\Sigma}(t)}{R(t)}$$

Powyższą procedurę powtarzamy dla  $t=0, 1, 2, \dots, T$ . Ostatni stan populacji z tego przedziału czasu  $S(T)$  przyjmowany jest dalej za stan początkowy do przeprowadzania prognozy.

Wykażemy, że estymator  $N_i(t) = k_i(t) \cdot N_{\Sigma}(t) / R(t)$  jest nieobciążony i zgodny. Dla przykładu zajmiemy się typem pierwszym, w ustalonej chwili czasu  $t$ .

Określamy ciąg zmiennych losowych:  $U_n = \frac{N_{\Sigma}(t)}{R(t)} \cdot \frac{1}{n} \cdot \sum_{j=1}^n k_1^j(t)$ ,  
 $k_1^j$  - ilość elementów 1-szego typu w  $j$ -tej próbie z populacji w  $t$ .

$$E[k_1^j] = \frac{N_1 \cdot R}{N_{\Sigma}}, \quad E[k_1^j - E[k_1^j]]^2 = V$$

Nieobciążoność: Czy  $E[U_n] \stackrel{?}{=} N_1$

$$E[U_n] = E\left[\frac{N_{\Sigma}(t)}{R(t)} \cdot \frac{1}{n} \cdot \sum_{j=1}^n k_1^j\right] = \frac{N_{\Sigma}(t)}{R(t)} \cdot \frac{1}{n} \cdot \sum_{j=1}^n E[k_1^j],$$

$$E[U_n] = N_1(t) \quad (\text{cbdo}).$$

Zgodność:  $\lim_{n \rightarrow \infty} P(|U_n - N_1(t)| > \varepsilon) = 0, \varepsilon > 0.$

$$P(|U_n - N_1(t)| > \varepsilon) = P\left(\left|\frac{N_2(t)}{R(t)} \cdot \frac{1}{n} \cdot \sum_{j=1}^n k_1^j - N_1(t)\right| > \varepsilon\right) =$$

$$P\left(\left|\sum_{j=1}^n k_1^j - n \cdot \frac{N_1(t)}{N_2(t)} \cdot R(t)\right| > \frac{\varepsilon \cdot n \cdot R(t)}{N_2(t)}\right) = P\left(\left|\sum_{j=1}^n \left(k_1^j - \frac{R}{N_2} \cdot N_1(t)\right)\right| > \frac{\varepsilon \cdot n \cdot R(t)}{N_2(t)}\right).$$

Zmienna losowa  $X_n = \sum_{j=1}^n (k_1^j - R(t) \cdot N_1(t) / N_2(t))$  ma wartość oczekiwaną  $E[X_n] = 0$ , oraz wariancję  $E[X_n^2]$  równą n-krotnej wariancji  $k_1^j (= n \cdot V)$ .  $k_j^i, k_1^i$  są niezależne. Korzystając z nierówności Czebyszewa mamy

$$P(|X_n| > \frac{\varepsilon \cdot n \cdot R(t)}{N_2(t)}) < \frac{n \cdot V \cdot [N_2(t)]^2}{\varepsilon^2 \cdot n^2 \cdot [R(t)]^2} = \frac{V \cdot [N_2(t)]^2}{\varepsilon^2 \cdot [R(t)]^2} \cdot \frac{1}{n}$$

$$\lim_{n \rightarrow \infty} P(|U_n - N_1(t)| > \varepsilon) = \lim_{n \rightarrow \infty} \frac{V \cdot [N_2(t)]^2}{\varepsilon^2 \cdot [R(t)]^2} \cdot \frac{1}{n} = 0 \quad (\text{cbdo}).$$

Podobnie dla pozostałych typów,  $\forall t \in [0, T]$ .

Jako estymator największego prawdopodobieństwa jest też asymptotycznie najefektywniejszy /4/.

Istnieje niebezpieczeństwo, że w dostępnej próbie nie będzie uwzględniony typ liczący niewiele elementów. Może to być typ, który aktualnie jest eliminowany z populacji. W takim przypadku błąd prognozy, spowodowany nieuwzględnieniem tego typu nie będzie istotny. Natomiast istotny wpływ na przebieg prognozy będzie miało nieuwzględnienie typu, który został właśnie wprowadzony do populacji, ma stosunkowo wysoką jakość, ale jeszcze nie został rozpowszechniony w populacji. Błąd wynikający z tego faktu będzie analizowany w rozdz. 3.6.

Można wykazać, że aby prawdopodobieństwo tego, że w próbie będzie choć jeden reprezentant typu, którego liczebność wynosi jeden procent całej populacji było nie mniejsze niż 0.9, wielkość próby powinna wynosić nie mniej niż 300 elementów.

### 3.3. Identyfikacja parametrów modelu dyskretnego.

Znany jest stan populacji w pewnym przedziale czasu /historia/  $S(t)$  dla  $t \in [0, T]$ .

Szukane parametry modelu:  $A, M, q_1, q_2, \dots, q_m$ .

Przyjęto również estymator największego prawdopodobieństwa.

$P(S(T), S(T-1), \dots, S(1), S(0))$  – prawdopodobieństwo tego, że w przedziale czasu  $[0, T]$  populacja znajdowała się w stanach  $S(0), S(1), \dots, S(T)$ .

Szukamy takich wartości parametrów  $A^*, M^*, q_1^*, \dots, q_m^*$ , dla których powyższe prawdopodobieństwo przyjmuje wartość maksymalną.

$$\max_{A, M, q_i} P(S(T), \dots, S(1), S(0)) \longrightarrow A^*, M^*, q_1^*, \dots, q_m^* \quad (3.3.1)$$

$$P(S(T), \dots, S(0)) = \prod_{t=1}^T P(S(t)|S(t-1)) = \prod_{t=1}^T \prod_{i=1}^m P(N_i(t)|S(t-1))$$

Podstawiając za  $P(N_i(t)|S(t-1))$  i przekształcając, otrzymujemy następujące wyrażenie.

$$P(S(T), \dots, S(0)) = e^{-\sum_{t=1}^T N_2(t-1)[1+A \cdot (1-N_2(t-1)/M)]} \cdot \prod_{t=1}^T \prod_{i=1}^m \frac{1}{[N_i(t)]!}$$

$$\left\{ \frac{N_i(t-1) \cdot q_i \cdot N_2(t-1) \cdot [1+A \cdot (1-N_2(t-1)/M)]^{N_i(t)}}{\sum_{j=1}^m N_j(t-1) \cdot q_j} \right\} \quad (3.3.2)$$

Na podstawie powyższego wzoru można stwierdzić:

Jeżeli  $P(S(T), \dots, S(1), S(0))$  przyjmuje wartość maksymalną dla

$A^*, M^*, q_1^*, \dots, q_m^*$ , to przyjmuje ją również dla każdego wektora

$A^*, M^*, c \cdot q_1^*, c \cdot q_2^*, \dots, c \cdot q_m^*$ , gdzie  $c$  – dowolna stała dodatnia.

Należy przyjąć zatem jakoś dowolnie wybranego typu, która będzie wartością odniesienia dla jakości pozostałych typów. Liczba zmiennych optymalizacyjnych zmniejsza się o jedną i wynosi  $(m-1)+2=m+1$ .

Ze względów obliczeniowych wygodniej będzie użyć do celów opty-

malizacji zamiast wyrażenia /3.3.2/, jego logarytm

$$\begin{aligned} \ln P(S(T), \dots, S(0)) = & - \sum_{t=1}^T N_{\Sigma}(t-1) [1 + A \cdot (1 - N_{\Sigma}(t-1) / M)] + \\ & + \sum_{t=1}^T \sum_{i=1}^m \left\{ N_i(t) [\ln(N_i(t-1) q_i) + \ln(N_{\Sigma}(t-1) [1 + A \cdot (1 - N_{\Sigma}(t-1) / M)]) + \right. \\ & \left. - \ln \left( \sum_{j=1}^m N_j(t-1) q_j \right) \right] - \sum_{k=1}^{N_i(t)} \ln(k) \right\}. \end{aligned} \quad (3.3.3)$$

Ostatni składnik wyr. /3.3.3/ nie zależy od zmiennych optymalizacyjnych, można więc go pominąć.

Zadanie  $\max_{A, M, q_i > 0, i=1, \dots, m} (\ln P(S(T), \dots, S(0)))$ , równoważne zadaniu /3.3.1/

można rozwiązać wykorzystując jedną z metod optymalizacji statycznej. W niniejszej pracy wykorzystano metodę Rosenbrocka.

### 3.4. Identyfikacja parametrów modelu ciągłego.

Znane:  $S(t)$ ,  $t \in [0, T]$ .

Szukane: takie  $A^*$ ,  $M^*$ ,  $b^*$ ,  $q_1^*$ ,  $\dots$ ,  $q_m^*$ , które będą maksymalizować funkcję gęstości prawdopodobieństwa zaistnienia takiej sekwencji stanów populacji:

$$\max_{A, M, b, q_i > 0, i=1, \dots, m} f(S(T), S(T-1), \dots, S(0)) \quad (3.4.1)$$

$$f(S(T), \dots, S(0)) = \prod_{t=1}^T \prod_{i=1}^m f(N_i(t) | S(t-1)).$$

Zastosujemy przybliżenie:  $\Gamma(z) \approx \sqrt{\frac{2\pi}{z}} \cdot e^{-z} \cdot z^{-z}$ .

Po podstawieniu otrzymujemy

$$f(S(T), \dots, S(0)) = \prod_{t=1}^T \frac{b^{\sum_{i=1}^m \lambda_i(t-1)}}{\prod_{i=1}^m N_i(t)} \cdot e^{-b \cdot \sum_{i=1}^m N_i(t)}.$$

$$\prod_{i=1}^m \frac{[N_i(t)]^{b \cdot \lambda_i(t-1)} \cdot \sqrt{b \cdot \lambda_i(t-1)} \cdot e^{b \cdot \lambda_i(t-1)}}{\sqrt{2 \cdot \pi} \cdot [b \cdot \lambda_i(t-1)]^{b \cdot \lambda_i(t-1)}}. \quad (3.4.2)$$

Pomijając czynniki stałe /niezależne od zmiennych optymalizacyjnych/ i logarytmując powyższe wyrażenie otrzymujemy

$$\ln f' = \sum_{t=1}^T \left\{ b \cdot \sum_{i=1}^m \lambda_i (t-1) (\ln b + 1) - b N_{\Sigma}(t) + \sum_{i=1}^m [\ln (b \cdot \lambda_i (t-1)) \cdot (1/2 - b \cdot \lambda_i (t-1)) + b \cdot \lambda_i (t-1) \cdot \ln (N_i(t))] \right\}. \quad (3.4.3)$$

gdzie:

$$\lambda_i(t-1) = \frac{N_i(t-1) \cdot q_i}{q_{\Sigma}(t-1)} \cdot [1 + A \cdot (1 - N_{\Sigma}(t-1)/M)] \quad (3.4.4)$$

Podobnie jak w modelu dyskretnym, należy przyjąć jakoś jednego, dowolnie wybranego typu, a optymalizować po pozostałych parametrach. Liczba zmiennych optymalizacyjnych:  $(m-1)+3=m+2$ .

Zadanie  $\max_{A, M, b, q_i} (\ln f')$  zostało w pracy rozwiązane na m.c. z wykorzystaniem metody optymalizacji statycznej Rosenbrocka.

### 3.5. Określenie horyzontu prognozy.

Horyzont prognozy, zdefiniowany w rozdz. 3.1. określany będzie dla modelu deterministycznego, przedstawionego w 2.4.

Oznaczmy czas pojawienia się nowego,  $(m+1)$ -ego typu w populacji przez  $t_p$ . Ilość elementów  $i$ -tego typu w chwili  $t_p$  wynosi:

$$\forall i=1, m : N_i(t_p) = \frac{N_i(0) \cdot (q_i)^{t_p}}{\sum_{j=1}^m N_j(0) \cdot (q_j)^{t_p}} \cdot N_{\Sigma}(t_p-1) [1 + A \cdot (1 - N_{\Sigma}(t_p-1)/M)]. \quad (3.5.1)$$

Począwszy od  $t_p$  w populacji funkcjonują elementy  $(m+1)$  typów.

Zakładając, że w  $t=t_p$  pojawiło się w populacji  $N_{m+1}(t_p)$  elementów nowego typu, to dla czasów  $t > t_p$  możemy zapisać:

$$N_i'(t) = \frac{N_i'(t_p) \cdot (q_i)^{(t-t_p)}}{\sum_{j=1}^{m+1} N_j(t_p) \cdot (q_j)^{(t-t_p)}} \cdot N_{\Sigma}(t-1) [1 + A \cdot (1 - N_{\Sigma}(t-1)/M)]$$

Podstawiając do powyższego /3.5.1/ za  $N_i(t_p)$ ,  $i=1, m$  dochodzimy po przekształceniach do następującej postaci:

$$N_i'(t) = N_i(t) \cdot \frac{1}{1 + \frac{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^{t_p} \cdot N_{m+1}(t_p)}{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^t \cdot N_{\Sigma}(t_p-1) [1+A \cdot (1-N_{\Sigma}(t_p-1))/M]}} =$$

$$= N_i(t) \cdot \frac{1}{1 + \frac{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^{t_p} \cdot N_{m+1}(t_p)}{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^t \cdot N_{\Sigma}(t_p)}} \quad (3.5.2)$$

gdzie:

$N_i(t)$  - ilość elementów i-tego typu dla  $t > t_p$ , gdyby w  $t_p$  w populacji nie pojawiły się elementy (m+1)-ego typu,

$N_i'(t)$  - ilość elementów typu i dla  $t > t_p$  przy założeniu, że w  $t_p$  powstało  $N_{m+1}$  elementów typu (m+1)-ego.

Horyzont prognozy jest to czas, dla którego:  $\frac{N_i'(t_h) - N_i(t_h)}{N_i(t_h)} = 0,1$

tzn.

$$\frac{1}{1 + \frac{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^{t_p} \cdot N_{m+1}(t_p)}{\sum_{k=1}^m N_k(0) \cdot (a_{i,k}/a_{i,m+1})^{t_h} \cdot N_{\Sigma}(t_p)}} \leq 0,9 \quad (3.5.3)$$

$t_h$  - horyzont prognozy.

Dla zadanych wartości  $t_p$ ,  $a_{m+1}$ ,  $N_{m+1}(t_p)$ , algorytm prognozujący przedstawiony w rozdziale 5.1 wyznacza  $t_h$ .

Na podstawie wyr. /3.5.2/ można sformułować następujące wnioski:

1. Błąd procentowy ilości elementów poszczególnych typów, spowodowany wprowadzeniem do populacji elementów nowego typu jest /w każdej chwili/ jednakowy dla wszystkich typów. Jest on zawsze ujemny.

2. Wartość bezwzględna błędu procentowego rośnie ze wzrostem czasu. Gdy wprowadzony nowy typ ma jakość większą od wszystkich



istniejących w chwili wprowadzania, to dla  $t \rightarrow \infty$  wypiera on wszystkie typy z populacji, a sam dominuje.

### 3.6. Wpływ niedokładności estymacji stanu populacji na prognozę.

Niedokładna znajomość stanów populacji, służących jako dane wejściowe do przeprowadzenia prognozy, ma na nią dwojaki wpływ:

1. Niedokładne dane wejściowe do zadania identyfikacji parametrów modelu powoduje otrzymanie obciążonych błędem rezultatów. Wpływ niedokładności identyfikacji parametrów modelu będzie przedyskutowany w rozdz. 3.7.

2. Wpływ na prognozę obciążonego błędem stanu początkowego populacji. Można tu wyróżnić dwa czynniki:

a/ obciążone błędami  $N_i(0)$  dla  $i=\overline{1,m}$ . /Początek okresu prognozy będziemy oznaczać  $t_0=0$ . /.

b/ możliwość nieuwzględnienia w populacji wszystkich typów.

#### 3.6.1. Wpływ zmian $N_i(0), i=\overline{1,m}$ na $S(t)$ .

Rozpatrzmy wpływ zmian  $N_i(0)$  na  $N_i(t)$ ,  $i=\overline{1,m}$  w przypadku modelu deterministycznego. Wiadomo, że:

$$N_i(t) = \frac{N_i(0) \cdot (a_i)^t \cdot N_2(t)}{\sum_{j=1}^m N_j(0) \cdot (a_j)^t} \quad (3.6.1.1)$$

Założmy, że dla każdego  $i=\overline{1,m}$ ,  $N_i'(0) = N_i(0) + \Delta_i$ .

W chwili  $t$  ilość elementów  $i$ -tego typu w populacji, zamiast /3.6.1.1/ wyraża się:

$$N_i'(t) = \frac{(N_i(0) + \Delta_i) \cdot (a_i)^t \cdot N_2(t)}{\sum_{j=1}^m (N_j(0) + \Delta_j) \cdot (a_j)^t} \quad (3.6.1.2)$$

Przekształcając powyższe wyrażenie otrzymujemy:

$$N_i'(t) = N_i(t) \cdot \frac{1 + \sigma_i(0)}{1 + \frac{\sum_{j=1}^m \sigma_j(0) \cdot N_j(0) \cdot (q_{ij})^t}{\sum_{j=1}^m N_j(0) \cdot (q_{ij})^t}}, \quad \sigma_k(0) = \frac{\Delta_k}{N_k(0)}$$

$$N_i'(t) \approx N_i(t) \cdot \left[ 1 + \sigma_i(0) - \frac{\sum_{j=1}^m \sigma_j(0) \cdot N_j(0) \cdot (q_{ij})^t}{\sum_{j=1}^m N_j(0) \cdot (q_{ij})^t} \right], \quad (3.6.1.3)$$

gdzie  $\sigma_j(0)$  - względny błąd ilości elementów j-tego typu w  $t=0$ .

Z powyższego wynika, że

$$\sigma_i(t) = \sigma_i(0) - \frac{\sum_{j=1}^m \sigma_j(0) \cdot N_j(0) \cdot (q_{ij})^t}{\sum_{j=1}^m N_j(0) \cdot (q_{ij})^t}. \quad (3.6.1.4)$$

Utrzymując w dalszym ciągu, że  $\max q_j = q_k$  dla  $t \rightarrow \infty$  w populacji istnieją elementy tylko k-tego typu,

$$\lim_{t \rightarrow \infty} \sigma_k(t) = \sigma_k(0) - \sigma_k(0) = 0.$$

$\forall i=1, \overline{m}$  można oszacować  $\sigma_i(t)$ . Załóżmy, że  $\max_{\sigma_j > 0} \sigma_j(0) = \sigma_s(0)$ ,

$\min_{\sigma_j < 0} \sigma_j(0) = \sigma_r(0)$ . Otrzymujemy wówczas oszacowanie:

$$-|\sigma_r(0)| - \sigma_s(0) \ll \sigma_i(t) \ll \sigma_s(0) + |\sigma_r(0)|. \quad (3.6.1.5)$$

### 3.6.2. Wpływ nieuwzględnienia typu na $S(t)$ .

Nieuwzględniony w  $S(0)$  typ będzie miał taki sam wpływ na liczebności elementów poszczególnych typów w populacji, jaki miało pojawienie się nowego typu w populacji, analizowane przy określaniu horyzontu prognozy. W wyrażeniu 3.5.2 należy podstawić za czas pojawienia się typu w populacji  $t_p=0$  i otrzymujemy:

$$N_i'(t) = N_i(t) \cdot \frac{1}{1 + \frac{N_{m+1}(0) \cdot (a_{m+1})^t}{\sum_{j=1}^m N_j(0) \cdot (a_j)^t}} \quad (3.6.2.1)$$

Zatem procentowy błąd ilości elementów i-tego typu, spowodowany brakiem (m+1)-szego typu w  $S(0)$  wynosi:

$$\sigma_i(t) = - \frac{N_{m+1}(0)}{\sum_{j=1}^m N_j(0) \cdot (a_j/a_{m+1})^t} \cdot \frac{1}{1 - \frac{N_{m+1}(0)}{\sum_{j=1}^m N_j(0) \cdot (a_j/a_{m+1})^t}} = - \frac{N_{m+1}(t)}{N_{\Sigma}(t)} \quad (3.6.2.2)$$

Jak widać, procentowy błąd liczby elementów poszczególnych typów, spowodowany brakiem jednego typu w stanie początkowym jest ujemny i jednakowy dla wszystkich typów, co do modułu jest równy procentowemu udziałowi w populacji nieuwzględnionego typu. Można oszacować czas, dla którego udział elementów (m+1)-szego /nieuwzględnionego/ typu nie przekroczy określonej wartości  $c$ :

$$\frac{N_{m+1}(t)}{N_{\Sigma}(t)} \ll c \implies \frac{N_{m+1}(0) \cdot (a_{m+1})^t}{\sum_{j=1}^m N_j(0) \cdot (a_j)^t} \ll c \quad (3.6.2.3)$$

Najbardziej sprzyjające warunki do wzrostu liczebności elementów (m+1)-szego typu w populacji będą wtedy, gdy wszystkie  $q_j$  pod znakiem sumy w /3.6.2.3/ zastąpimy przez  $q_{\min}$ ,  $q_{\min} = \min_j q_j$ . Wtedy, przekształcając /3.6.2.3/ otrzymujemy:

$$t \ll \frac{\ln \frac{c}{1-c} + \ln \frac{N_{m+1}(0)}{N_{\Sigma}(0)}}{\ln \frac{q_{\min}}{q_{\max}}} \quad (3.6.2.4)$$

### 3.7. Wpływ zmian parametrów modelu na prognozę.

Wpływ niedokładności identyfikacji pojemności środowiska oraz maksymalnego współczynnika reprodukcji  $(1+A)$  na  $S(t)$  będzie

omówiony przy badaniach symulacyjnych modelu. W tym rozdziale ograniczymy się do przeanalizowania, jak w modelu deterministycznym wpływają na stan populacji  $S(t)$  zmiany jakości typów.

Zakładamy, że  $\forall i=\overline{1,m} \quad q'_i = q_i \cdot (1 + \sigma_{q_i})$ , wtedy

$$N'_i(t) = \frac{N_i(0) [q_i (1 + \sigma_{q_i})]^t \cdot N_{\Sigma}(t)}{\sum_{j=1}^m N_j(0) [q_j (1 + \sigma_{q_j})]^t} = N_i(t) \frac{(1 + \sigma_{q_i})^t \cdot \sum_{j=1}^m N_j(0) \cdot (q_j)^t}{\sum_{j=1}^m N_j(0) (q_j)^t \cdot (1 + \sigma_{q_j})^t} \quad (3.7.1)$$

Procentowy błąd liczby elementów  $i$ -tego typu w populacji w  $t$ :

$$\sigma'_i(t) = \frac{\sum_{j=1}^m N_j(0) \cdot (q_j)^t}{\sum_{j=1}^m N_j(0) \cdot (q_j)^t \left( \frac{1 + \sigma_{q_j}}{1 + \sigma_{q_i}} \right)^t} - 1 \quad (3.7.2)$$

Jeżeli procentowy błąd jakości wszystkich typów jest taki sam, tzn.  $\forall i=\overline{1,m} \quad \sigma_{q_i} = \sigma$ , to  $\sigma'_i(t) = 0$ ,  $\forall i=\overline{1,m}$ .

Błąd wynikający ze zmian jakości typów jest dość istotny, może doprowadzić do dominacji innego typu niż w rzeczywistej populacji, oraz wyeliminowania z populacji elementów typu faktycznie najlepszego.

Założmy, że  $\max_j q_j = q_k$  oraz  $\max_j q'_j = q_s$ . Przekształcając wyrażenie /3.7.2/ można uzyskać:

$$\sigma'_i(t) = \frac{\sum_{j=1}^m N_j(0) (q_j/q_i)^t}{\sum_{j=1}^m N_j(0) (q'_j/q'_i)^t} - 1 \quad (3.7.3)$$

Obliczając błędy  $\sigma'_k(t)$  i  $\sigma'_s(t)$  przy  $t \rightarrow \infty$ , otrzymujemy:

$$\lim_{t \rightarrow \infty} \sigma'_k(t) = 0 - 1 = -1,$$

$$\lim_{t \rightarrow \infty} \sigma'_s(t) = \infty - 1 = \infty.$$

Istotnie, w populacji w nieskończoności powinno być  $N_k(t) = N_{\Sigma}(t) = M$  elementów typu  $k$ , a jest ich zero, więc  $\sigma'_k(\infty) = \frac{0 - M}{M} = -1$ , natomiast elementów typu  $s$  powinno być zero, a jest ich  $M$  /bo

ten typ zdominował w populacji/, więc  $\tilde{o}_s(\infty) = \frac{M-0}{0} \rightarrow \infty$

Liczebności pozostałych typów w populacji dla  $i/k \wedge i/s$ , zachowują się w nieskończoności tak samo dla primowanych jak i nieprimowanych jakości - zerują się.

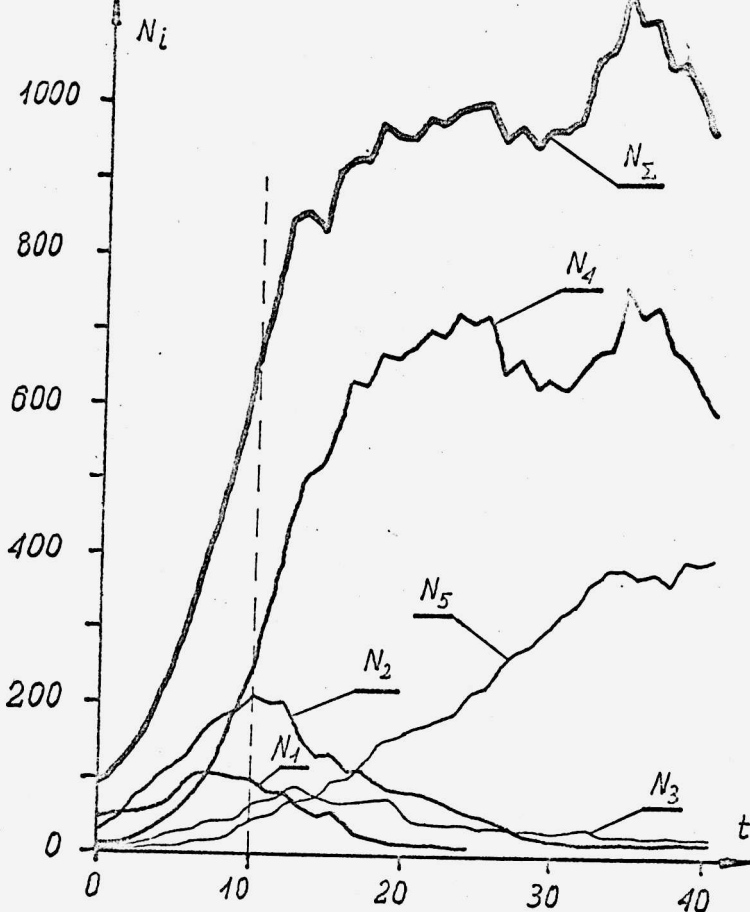
#### 4. BADANIA SYMULACYJNE MODELU

##### 4.1. Typowe przebiegi otrzymane z modelu probabilistycznego.

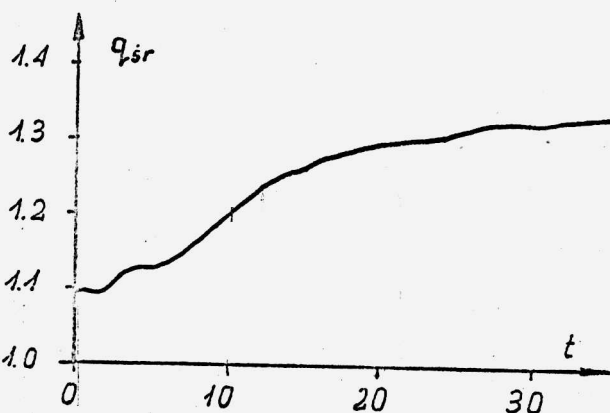
Rysunek 4.1.1 przedstawia zmianę liczby elementów różnych typów i wielkości populacji w czasie. Wynik ten otrzymano z modelu, w którym proces reprodukcji jest losowy o rozkładzie Poissona /model dyskretny/. Parametry modelu:

$$S(0) = [40, 30, 20, 8, 2], \quad A = 0.3, \quad M = 1000,$$

$$q_1=1, \quad q_2=1.1, \quad q_3=1.2, \quad q_4=1.3, \quad q_5=1.4,$$



Rys. 4.1.1. Zmiany stanu populacji w czasie.



Rys. 4.1.2. Zmiany średniej jakości w czasie.

Dla  $t < 10$  liczba elementów wszystkich typów zwiększa się. Wynika

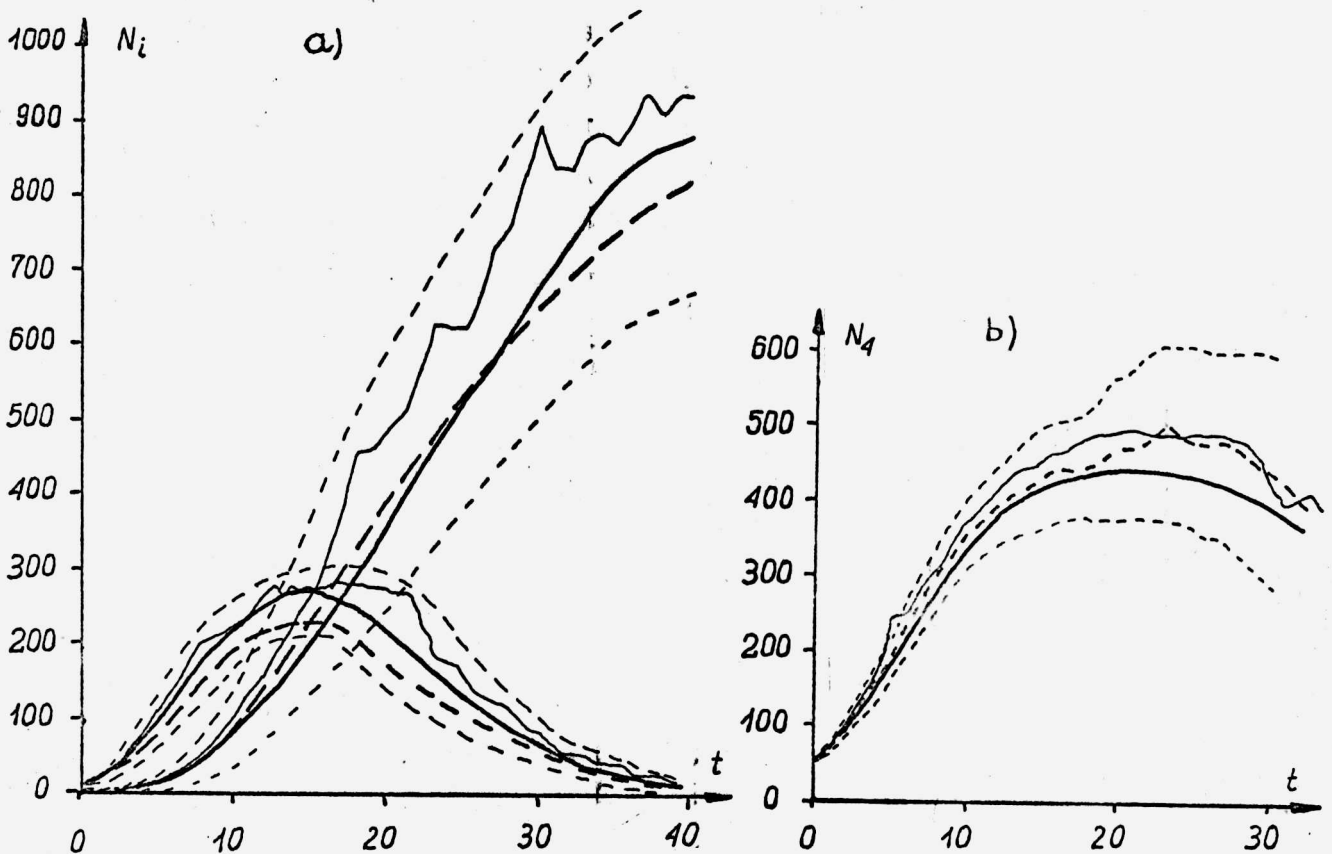
to z dużego niedosycenia środowiska. Dla  $t=10$ ,  $N_2(t)=0.7M$ . Od tego czasu zaczyna być stopniowo eliminowany z populacji typ 1 /naj-słabszy/. Dla  $t=12$  zaczyna być eliminowany również typ drugi. Stopniowo pozostałe typy są eliminowane, z wyjątkiem typu piątego /najlepszego w populacji/. W  $t=60$  w populacji udział typu piątego przekracza 90%. Rys. 4.1.2 przedstawia zmiany średniej jakości populacji w czasie, uzyskane w tym samym eksperymencie.  $q_{sr}(0)=1.102$ , więc  $\frac{q_1}{q_{sr}(0)} < 1$ . O wzroście liczebności pierwszego typu decyduje więc niedosycenie środowiska. Ilość elementów typu pierwszego zaczyna maleć gdy  $q_{sr}(t)=1.164$ .

Widać, że  $q_{sr}(t)$  jest rosnącą funkcją czasu aż do uzyskania maksimum równego 1.4, gdy w populacji funkcjonują elementy tylko piątego typu / $t=104$ /.

W rozdziale 1.2, przy przedstawianiu grupy metod ekstrapolacji trendów powiedziane było, że metodami tymi najczęściej przeprowadza się prognozę jednego, odpowiednio wybranego parametru, np. tonomile/h przy badaniu rozwoju samolotów transportowych. Załóżmy, że element w populacji odpowiada jednemu samolotowi transportowemu, a samolot jest tym lepszy, im jego zdolność przewozowa /mierzona w tonomilach/h/ będzie większa /jest to jego jakość/. Rys. 4.1.1 wskazywałby jak zmieniać się będzie ilość samolotów transportowych różnych typów /np. turboodrzutowe I generacji, II generacji i in./, natomiast rys. 4.1.2 byłby prognozą zmian średnich zdolności przewozowych samolotów transportowych. Przy założeniu, że w prognozowanym okresie nie powstanie nowy typ tych samolotów, byłaby to ta prognoza, którą przeprowadzono w USA metodą ekstrapolacji trendów /5/. Ten przykład nie mógł być w niniejszej pracy przetestowany ze względu na brak danych.

Przeprowadzono analogiczny eksperyment /jak na rys. 4.1.1/

z modelem ciągłym, w którym przyjęto parametr  $b=1$ . Uzyskane rezultaty są bardzo podobne do tych z rys. 4.1.1 i 4.1.2, i z tego względu nie zaprezentowano ich w niniejszym rozdziale. Rysunki 4.1.3<sup>a</sup> i 4.1.3<sup>b</sup> pokazują porównanie przebiegu deterministycznego z wartością średnią, estymowaną z 50 realizacji.



Rys. 4.1.3. Przykładowe przebiegi otrzymane z modelu deterministycznego i probabilistycznego, a) model dyskretny, b) model ciągły,

- - - - -wartość średnia estymowana,
- -przebieg deterministyczny,
- - - - -wartość średnia  $\pm$  odchylenie standardowe,
- -przykładowa realizacja,

$$S(0) = [68, 82, 85, 51, 18], A=0.3, M=1000,$$

$$q_1=1, q_2=1.1, q_3=1.2, q_4=1.3, q_5=1.4 .$$

W każdym z przeprowadzonych eksperymentów, przebieg deterministyczny dobrze przybliżał estymowaną wartość średnią, nawet jeżeli



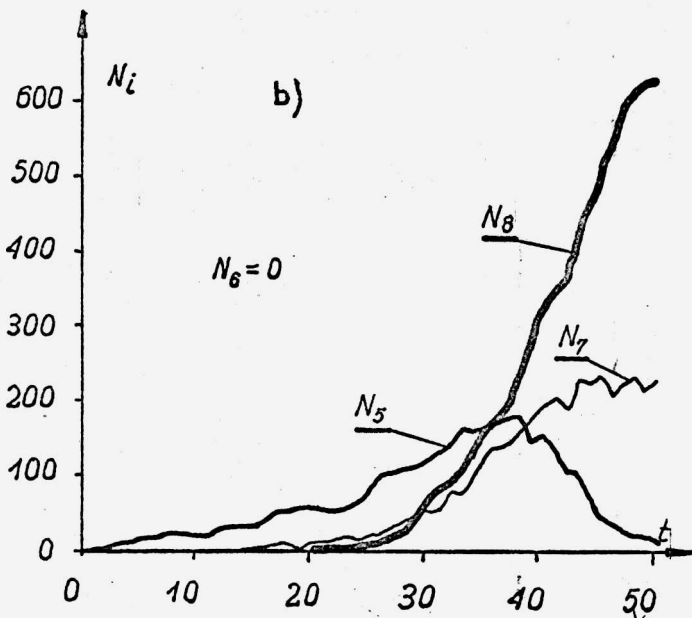
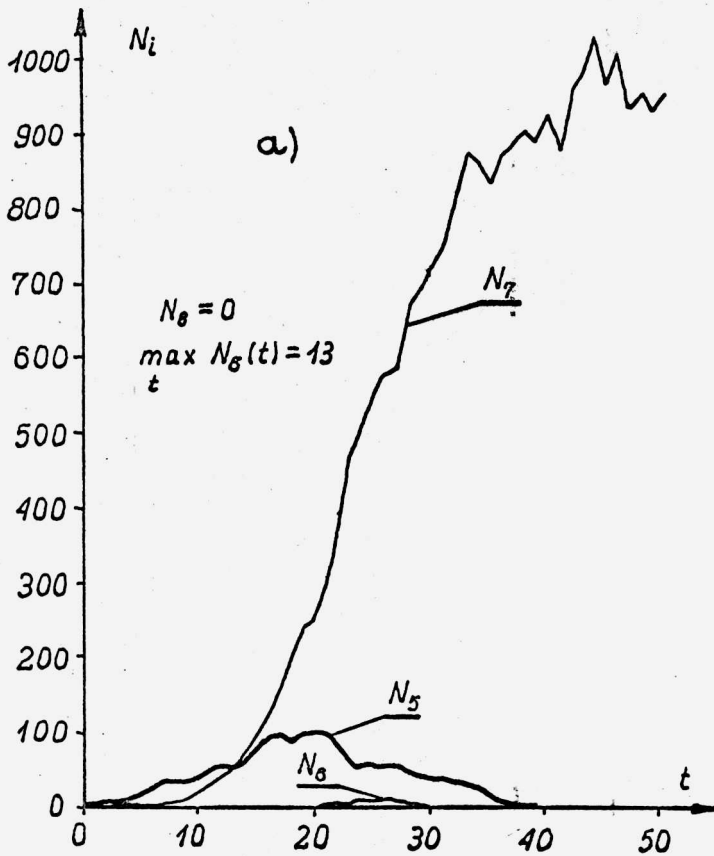
do estymacji wykorzystano tylko 10 realizacji. Tylko bardzo nieliczne realizacje nie mieściły się w przedziale  $E[N_1(t)] \pm \sqrt{D_1^2(t)}$ . Taką realizację pokazano na rys. 4.1.3<sup>b</sup> /t=5/.

Przeprowadzone eksperymenty potwierdzają fakt, że losowość jest bardziej uwidoczniła w procesach, w których występują stosunkowo niewielkie populacje. Zarówno w procesie Poissona jak i przy rozkładzie gamma, stosunek odchylenia standardowego do wartości oczekiwanej maleje ze wzrostem wartości oczekiwanej. Dla rozkładu Poissona  $\frac{\sqrt{D^2(x)}}{E(x)} = \frac{1}{\sqrt{\lambda}}$ , dla rozkładu gamma  $\frac{\sqrt{D^2(x)}}{E(x)} = \frac{1}{\sqrt{b}} \cdot \frac{1}{\sqrt{\lambda}}$ . Stąd bardziej gładkie krzywe, gdy większa jest liczebność populacji.

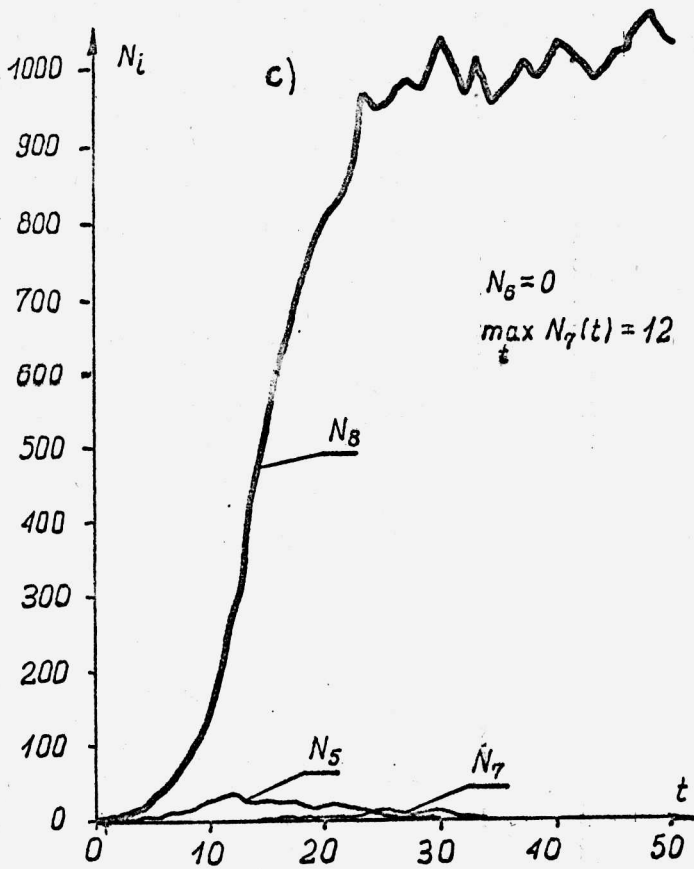
#### 4.2. Szanse rozwoju elementów nowych typów.

Przeprowadzono eksperymenty, w których mogły pojawiać się elementy nowych typów, o lepszych jakościach. W  $S(0)$  w populacji istniały elementy pięciu typów. Eksperyment był tak dobrany, że prawdopodobieństwo pojawienia się przynajmniej jednego elementu szóstego typu wynosiło 0.1, siódmego typu - 0.07, ósmego typu - 0.049. Jeżeli pojawił się w populacji choć jeden element któregoś z tych typów, rozpowszechniał się z czasem zgodnie z modelem. Jakości typów wynosiły odpowiednio: 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7. Pozostałe parametry modelu wynoszą:  $A=0.3$ ,  $M=1000$ ,  $S(0)=[300, 100, 80, 70, 3, 0, 0, 0]$ , model dyskretny / reprodukcja o rozkładzie Poissona/. Na rys. 4.2.1 pokazano kilka przykładowych realizacji. W każdym z pokazanych przykładów zwyciężył inny typ. W realizacji z rysunku 4.2.1<sup>a</sup> dla  $t=27$  istniało w populacji 13 elementów typu szóstego, ale już w następnej chwili /rys. 4.2.2a/  $q_{sr}$  przekroczyło wartość

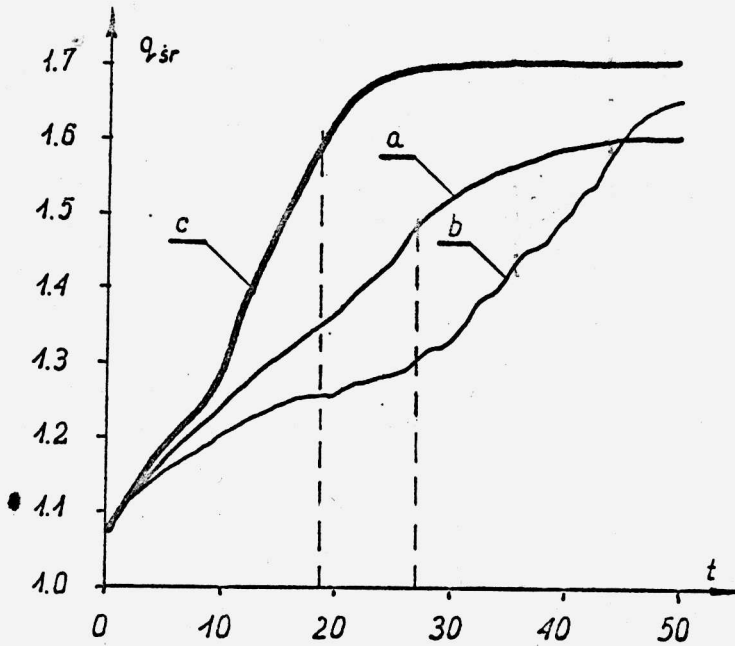
1.5 /ponieważ dość szybko w populacji rozpowszechnił się typ 7/, nasycenie środowiska było już osiągnięte, więc ten typ został szybko wyeliminowany z populacji. Typ ósmy pojawił się w  $t=8$  - jeden element, utrzymywał się przez pięć następnych chwil, jego ilość nie zwiększyła się w tym czasie /wpływ losowości/. Również od  $t=117$  do 121 było  $N_8(t)=1$ , ale również "bez powodzenia".



Rys. 4.2.1a,b. Przykładowe krzywe obrazujące zmiany  $N_5(t)$ ,  $N_6(t)$ ,  $N_7(t)$ ,  $N_8(t)$  w eksperymencie umożliwiającym powstawanie nowych typów.



Rys. 4.2.1c. Przykładowe krzywe obrazujące zmiany  $N_5(t)$ ,  $N_6(t)$ ,  $N_7(t)$ ,  $N_8(t)$  w eksperymencie umożliwiającym powstawanie nowych typów.



Rys. 4.2.2. Zmiany średniej jakości populacji w eksperymencie z rys. 4.2.1.

W tym przykładzie średnia jakość populacji osiągnęła maksymalną wartość równą  $q_7=1.6$ . W przykładzie b) typy 7 i 8 zostały rozpowszechnione w populacji; typ 6 pojawił się kilkakrotnie w pojedynczych ilościach, ale był eliminowany z populacji. Typ siódmy "rywalizował" z typem ósmym, dopiero w  $t=15$  w populacji całkowicie zdominowały elementy ósmego typu. Na rys. 4.2.1.c widać, że typ ósmy pojawił się w populacji w  $t=1$  i nie wyginął losowo, lecz zaczął dość szybko dominować. Spowodowało to bardzo szybki wzrost średniej jakości populacji /rys. 4.2.2, krzywa c/. Typ siódmy pojawił się w  $t=15$ , zaczął się stopniowo rozpowszechniać, ale w  $t=19$  /ze względu na szybki rozwój  $N_8 / q_{sr}$  wynosiło już 1.6 i wobec tego typ 7 po osiągnięciu 12 elementów został wyeliminowany.

Przeprowadzono również podobne eksperymenty z innymi parametrami modelu / $A, M, q_i, S(0)$ , model dyskretny i ciągły/. Wszystkie wskazują na to, że pojawienie się jednego elementu danego typu, nawet istotnie lepszego, nie oznacza że potrafi on zdominować w populacji. Gdy jego liczebność wynosi 1 lub 2 elementy, może zajść taka realizacja, w której wskutek losowości reprodukcji te pojedyncze elementy wyginą. W przeprowadzonych eksperymentach nie zdarzyło się, by losowo został wyeliminowany typ, który osiągnął liczebność  $\gg 3$  /przy jego jakości większej od  $q_{sr}(t)$  /.

#### 4.3. Przebiegi otrzymane przy założeniu znajomości losowej próbki z populacji.

Przeprowadzono szereg eksperymentów, w których zakładano znajomość losowej próbki z populacji w pewnym przedziale czasu. W niniejszym rozdziale zaprezentowano jeden z takich eksperymentów.

Jako populację rzeczywistą wykorzystano wyniki uzyskane po symulacji modelu deterministycznego, w którym:

$M=1000$ ,  $A=0.3$ ,  $q: 1, 1.1, 1.2, 1.3, 1.4$ ,  $S(0)=[40, 30, 20, 8, 2]$ .

Otrzymane krzywe przedstawione są na rys. 4.3.1 linią przerywaną.

Następnie przeprowadzono losowanie próbki z tej populacji, dla  $t \in [0, 5]$ . Opracowano algorytm na m.c. przy pomocy którego otrzymano losową próbkę w  $t=0, 1, 2, 3, 4, 5$ , zadając wielkości próbek.

W tabeli 2 zestawiono stany populacji "rzeczywistej", stany w losowej próbce i estymowane stany populacji.

TABELA 2.

Rezultaty estymacji stanu populacji na podstawie losowej próbki.

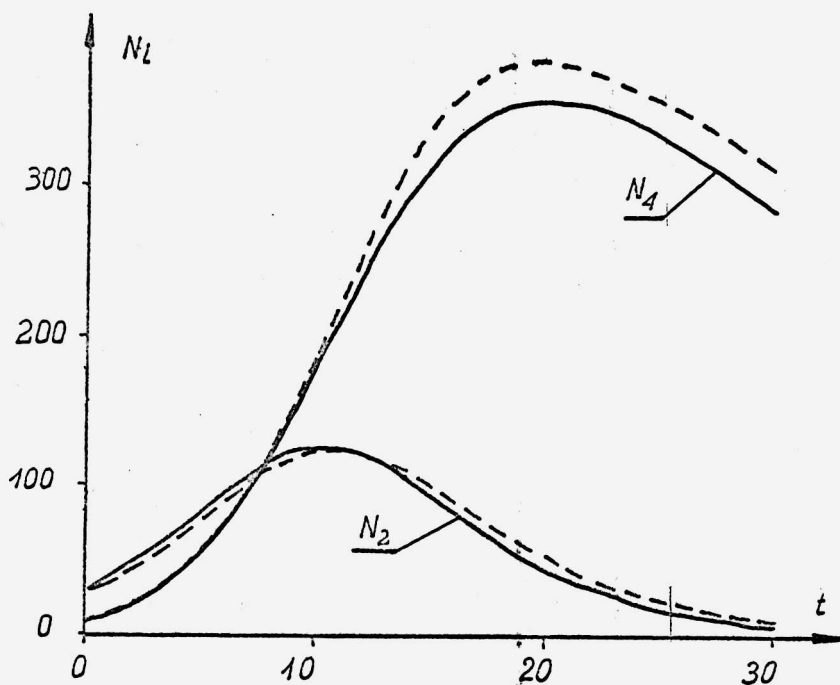
t	$N_1$	$N_2$	$N_3$	$N_4$	$N_5$	
0	40	30	20	8	2	Populacja rzeczywista
1	46	38	28	12	3	
2	52	47	38	18	5	
3	58	58	50	26	8	
4	64	70	66	36	12	
5	68	82	85	51	18	
0	20	15	10	4	1	Stan w los. próbce
1	28	24	18	8	2	
2	42	38	31	14	5	
3	43	43	38	20	6	
4	51	56	53	30	10	
5	45	54	56	34	11	
0	40	30	20	7	3	Populacja estymowana
1	45	39	28	13	2	
2	52	46	39	17	5	
3	58	59	51	27	8	
4	63	70	66	38	12	
5	68	82	86	52	19	

Wielkości próbek wynosiły:  $R(0)=50$ ,  $R(1)=80$ ,  $R(2)=130$ ,  $R(4)=200$ ,  
 $R(3)=150$ ,  $R(5)=200$ .

Estymowany stan populacji w  $t \in [0,5]$  wykorzystano do identyfikacji parametrów modelu dyskretnego, probabilistycznego. Zadana ilość iteracji w metodzie Rosenbrocka - 100. Po identyfikacji uzyskano następujące wartości parametrów modelu:

$A=0.29955$ ,  $M=1011.5$ ,  $q_1=1.0222$ ,  $q_2=1.1222$ ,  $q_3=1.2254$ ,  $q_4=1.3307$ ,  $q_5=1.4$ .

Na rysunku 4.3.1 linią ciągłą narysowano przebiegi uzyskane z modelu przy wykorzystaniu wyników estymacji i identyfikacji /są to estymowane wartości średnie z 50 realizacji/.



Rys. 4.3.1. Przebiegi  $N_2(t)$  i  $N_4(t)$  uzyskane przy założeniu znajomości losowej próbki z populacji, oraz przebiegi "rzeczywiste".

Podobne wyniki uzyskano przy zastosowaniu modelu z procesem reprodukcji o rozkładzie gamma. Im większy był stosunek wielkości próbki do wielkości populacji i im większa ilość iteracji przy identyfikacji parametrów modelu, tym bardziej pokrywały się krzywe

uzyskane na podstawie znajomości próbki z krzywymi uzyskanymi z modelu wyjściowego.

#### 4.4. Przykładowe realizacje przy zmiennych w czasie jakościach typów.

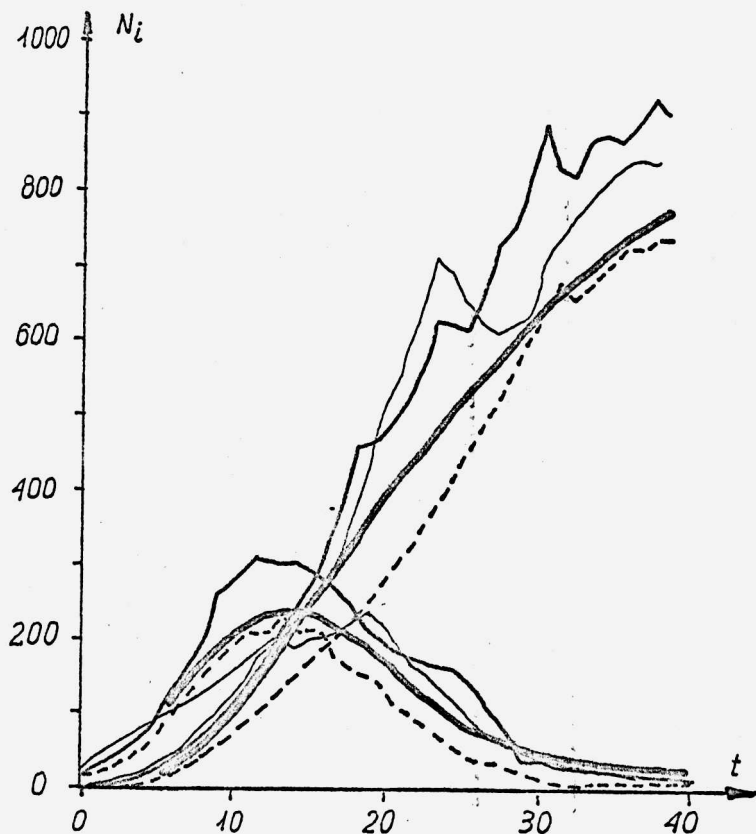
Przeprowadzono eksperymenty, w których jakości typów:

a/zmieniały się sinusoidalnie,

b/były zmiennymi losowymi o rozkładzie równomiernym.

Na rysunku 4.4.1 pokazano przykładowe realizacje zmian liczby elementów dla dwóch typów otrzymane z modelu, w którym:

$$S(0)=[40, 30, 20, 8, 2], \quad q_1^0=1.0, \quad q_2^0=1.1, \quad q_3^0=1.2, \quad q_4^0=1.3, \quad q_5^0=1.4.$$



Rys. 4.4.1. Zmiany liczebności typu 5 i 3 otrzymane przy różnych założeniach o jakościach typów i procesie reprodukcji.

—————  $q_i(t)$  - zmienna losowa o rozkładzie równomiernym na prze-

działe  $(q_i^0 - 0.15, q_i^0 + 0.15)$ ,  $i = \overline{1, m}$ , reprodukcja deterministyczna.

————  $q_i(t) = q_i^0 + 0.15 \sin\left(\frac{\pi}{5} \cdot t + \varphi_i\right)$ ,  $\varphi_1 = 0^\circ$ ,  $\varphi_2 = 90^\circ$ ,  $\varphi_3 = 0^\circ$ ,  
 $\varphi_4 = 270^\circ$ ,  $\varphi_5 = 180^\circ$ , reprodukcja losowa o rozkładzie Poissona.

————  $q_i(t) = q_i^0$ , reprodukcja deterministyczna.

-----  $q_i(t) = q_i^0$ , reprodukcja losowa o rozkładzie Poissona.

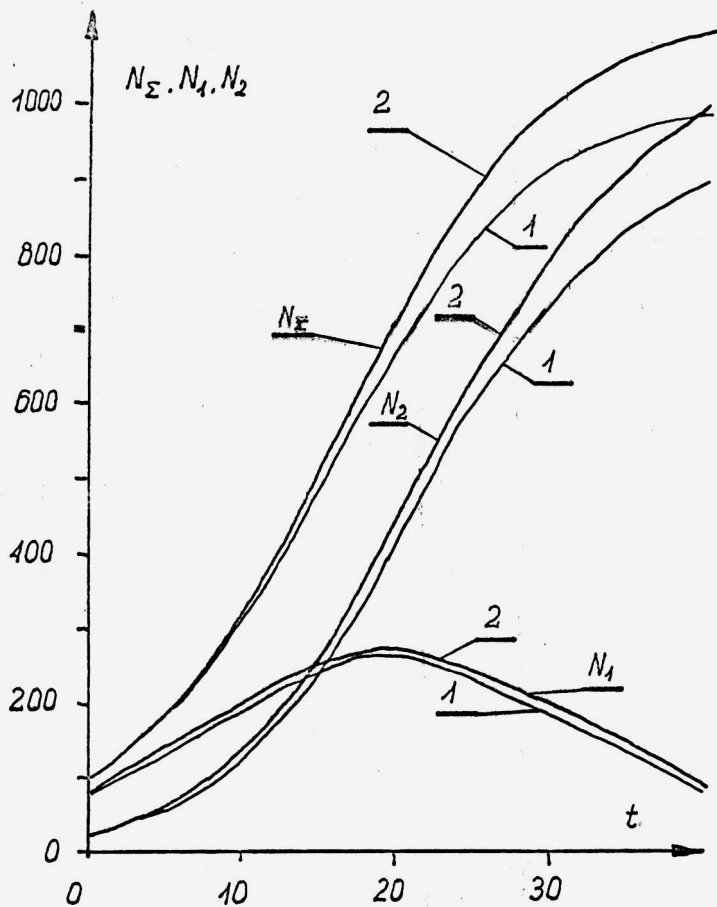
Biorąc pod uwagę stosunkowo mały rozmiar populacji  $N_2(0) = 100$ , przy którym istotny jest wpływ losowości można powiedzieć, że poszczególne realizacje nie różnią się od siebie więcej niż realizacje uzyskane przy tych samych parametrach modelu.

#### 4.5. Wpływ pojemności środowiska i maksymalnego współczynnika reprodukcji na stan populacji w czasie.

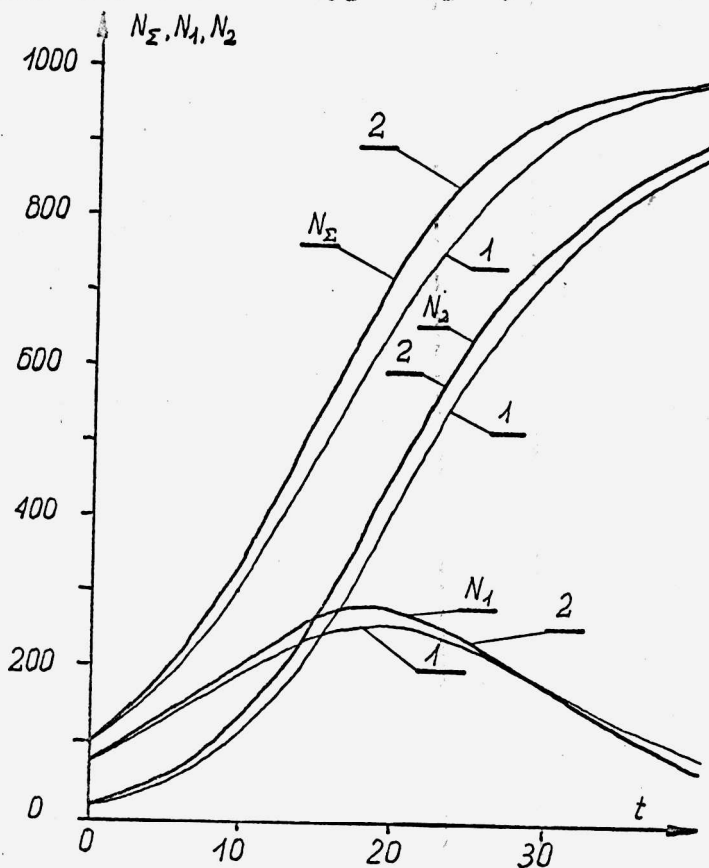
Na rysunku 4.5.1 przedstawiono zmianę wielkości populacji i liczby elementów dwóch typów /z których składa się badana populacja/ w czasie, dla dwóch wartości pojemności środowiska: krzywe oznaczone numerem 1 są uzyskane przy  $M_1 = 1000$ , krzywe o numerze 2 - przy  $M_2 = M_1 \cdot (1 + 0.1)$ . Przy niedosyceniu środowiska błąd  $N_2(t)$  i  $N_1(t)$  spowodowany dziesięcioprocentowym odchyleniem pojemności środowiska jest bardzo mały; po dostatecznie długim czasie /po dojściu do nasycenia/ błędy zwiększają się do błędu pojemności. Warunki eksperymentu:  $A = 0.15$ ,  $S(0) = [80, 20]$ ,  $q_1 = 1.0$ ,  $q_2 = 1.1$ ,  $M_1 = 1000$ ,  $M_2 = 1100$ . Błąd typu pierwszego nie przekracza kilku procent, błąd typu drugiego zwiększa się z czasem, osiągając 10% wtedy, gdy typ pierwszy jest wyeliminowany.

Na rysunku 4.5.2 przedstawione są zmiany wielkości populacji i wielkości typów, uzyskane przy dziesięcioprocentowym odchyleniu parametru A. Krzywe oznaczone numerem 1 uzyskano przy  $A_1 = 0.15$ , krzywe o numerze 2 - przy  $A_2 = A_1 \cdot (1 + 0.1)$ .





Rys. 4.5.1. Estymowane wartości średnie wielkości populacji i ilości elementów w typach przy dwóch wartościach pojemności.



Rys. 4.5.2. Estymowane wartości średnie wielkości populacji i ilości elementów w typach przy dwóch wartościach wsp. A.

Błędy wielkości całej populacji w początkowym okresie czasu rosną /przy niedosyceniu środowiska/ po czym zaczynają maleć, spadając do zera dla  $N_2(t)=M$ . W zaprezentowanym eksperymencie dla  $t=9, 10, \dots, 23$  błąd  $N_2(t)$  wynosił w przybliżeniu 10%. Dla  $t > 23$  błąd wielkości populacji malał, w  $t=46$  zarówno przy  $A_1$  jak i  $A_2$  osiągnięte było nasycenie środowiska, wobec czego nie było błędu  $N_2(t)$ .

Tempo wzrostu liczby elementów każdego typu należącego do populacji jest większe dla większej wartości współczynnika  $A$ , ale też i szybkość eliminowania typu z populacji jest większa.

Warunki eksperymentu:  $S(0)=[80, 20]$ ,  $M=1000$ ,  $q_1=1.0$ ,  $q_2=1.1$ ,  $A_1=0.15$ ,  $A_2=0.165$ , model z reprodukcją losową o rozkładzie Poissona.

## 5. ALGORYTM PROGNOZUJĄCY

### 5.1. Schemat blokowy algorytmu.

W pracy zaproponowano dwie wersje algorytmu prognozującego. Pierwsza zbudowana jest z wykorzystaniem modelu probabilistycznego, w którym reprodukcja elementów jest losowa o rozkładzie Poissona, druga - modelu, w którym reprodukcja jest losowa o rozkładzie gamma.

Program na maszynę cyfrową został napisany w języku ALGOL 1900. Jako dane wejściowe należy podać znaną historię zmian stanu populacji /lub zmiany stanu w losowej próbce z populacji/ oraz zmienne sterujące programem /nie wymagające żadnych dodatkowych informacji o prognozowanym procesie/. W tabeli 3 podano wszystkie zmienne wejściowe /z ich nazwami i znaczeniem/ w takiej kolejności, w jakiej należy je zadać do algorytmu.

TABELA 3.

Dane wejściowe do algorytmu prognozującego.

Lp.	Nazwa zmiennej	Znaczenie	Uwagi
1	M	Liczba typów w populacji	
2	TM	Czas, w którym znana jest historia zmian stanu populacji /lub próbki/	Czas mierzony jest od zera: $t=0, 1, 2, \dots, TM$
3	TP	Zadany okres prognozy	Prognoza przeprowadzana jest dla $t \in [0, TP]$ . Za $S(0)$ dla progn. podstawiany jest stan $S(TM)$
4	PROBA	Zmienna sterująca. Musi być równa 1 gdy dysponujemy losową próbą, dowolna, $\neq 1$ w przeciwnym przypadku	Spełnia rolę przełącznika w programie.

5	NPOP[J]	Wielkość populacji dla $t \in [0, TM]$ /historia/	Tablica jednowymiar. $(TM+1)$ - wymiarowa
6a	RP[J] K[I, J]	Wielkość próby w $t=J$ , $J \in [0, TM]$ Ilość elementów $I$ -tego typu w próbie RP[J], $I=1, M$ , $J \in [0, TM]$	Jeżeli PROBA=1 zadać te dane w kolejności: RP[0], K[1, 0], ..., K[M, 0], RP[1], K[1, 1], ..., K[M, 1]... RP[TM], K[1, TM], ..., K[M, TM]
6b	N[I, J]	Ilość elementów $I$ -tego typu w populacji w czasie J, $J \in [0, TM]$ , $I=1, M$	Te dane zadać tylko wówczas, gdy nie ma danych 6a. /PROBA≠1/. Kolejność: N[1, 0], N[2, 0], ..., N[M, 0], N[1, 1], ..., N[M, 1], ..., N[1, TM], ..., N[M, TM].
7	X[I] T[I]	Punkt startowy $I$ -tej zmiennej optymalizacyjnej przy identyfikacji parametrów modelu. Długość kroku w $I$ -tym kierunku przy identyfikacji /metoda Rosenbrocka/	$I=1, M+1$ - model dyskretny /I wersja alg./ $I=1, M+2$ - model ciągły /II wersja alg./ .Kolejność: X[1], T[1], X[2], T[2], ...
8	X[M+2] /lub X[M+3]/	Zadana wartość jakości dla $M$ -tego typu / $q_m$ - dowolna liczba dodatnia/	w I wersji alg. jest to X[M+2], w II wersji- X[M+3]. Różna ilość zm. optym. /patrz rozdz. 3.3, 3.4/
9	L	Liczba iteracji przy identyfikacji parametrów modelu /metoda Rosenbrocka/	
10	NORM[I]	Tablica dystrybuant rozkładu normalnego $\mathcal{N}(0,1)$ , dla przedziału $[0, 3]$ , z krokiem 0.1, $I=0, 1, \dots, 30$ .	Patrz sposób generowania zmiennych losowych o rozkładzie Poissona i gamma.
11	X	Zmienna losowa początkowa- dowolna liczba rzeczywista, dodatnia	
12	IP	Liczba przebiegów /realizacji/ na podstawie których estymowane są wartości średnie i odchylenia standardowe dla $N_2(t)$ , $N_i(t)$ , $i=1, M$ , $t \in [0, TP]$ .	
13	ILQ	Dla ilu różnych wartości jakości nowego typu ma być określany horyzont prognozy	
14	ILT	W ilu różnych chwilach ma być wprowadzany do populacji nowy typ w celu wyznaczenia horyzontu prognozy	

15	C1	Ile razy jakość nowo wprowadzanego typu jest większa od maksymalnej jakości, występującej w populacji.	Dla jednej wartości C1 zadać tyle wartości T1, ile wynosi zmienna ILT. Powyższą sekwencję danych powtórzyć tyle razy, ile wynosi zmienna ILQ.
	T1	Nowy typ jest wprowadzany do populacji w $t=T1$ .	

Jako wielkości wyjściowe z algorytmu otrzymujemy:

1. Jeżeli była losowa próbka, to wynik estymacji stanów populacji w przedziale czasu  $t \in [0, TM]$ .
2. Parametry modelu otrzymane w wyniku identyfikacji. Jeżeli algorytm optymalizacji został zakończony wcześniej /nie po L iteracjach/ wskutek braku polepszenia funkcji celu po wykonaniu kroków we wszystkich kierunkach, to również informację przy której iteracji to nastąpiło.
3. Wynik prognozy, czyli dla  $t \in [0, TP]$ :
  - dla  $i = \overline{1, M}$ ,  $N_i(t)$  oraz  $N_z(t)$  uzyskane z modelu deterministycznego,
  - dla  $i = \overline{1, M}$   $E[N_i(t)]$  i  $E[N_z(t)]$ , estymowane na podstawie zadanej ilości realizacji,
  - dla  $i = \overline{1, M}$   $\sqrt{D^2(N_i(t))}$  i  $\sqrt{D^2(N_z(t))}$ , estymowane na podstawie zadanej liczby realizacji.
4. Horyzont prognozy dla zadanych parametrów /jakości nowego typu i czasu wprowadzania go do populacji/.

Ze względu na to, że w chwili zadawania danych do algorytmu prognozującego nie jest wiadome ile wynosić będą jakości poszczególnych typów /prócz m-tego/ trudno jest podać wartość jakości nowo wprowadzanego do populacji typu. Dlatego poprzez zmienną C1 należy zadać pożądany stosunek jakości nowego typu do maksymalnej jakości jaka wystąpi w populacji / $C1 = q_{m+1} / q_{max}$ /. Dotychczas przez M oznaczana była pojemność środowiska, a przez

m - liczba typów w populacji. W programie, ze względu na konieczność stosowania dużych liter do nazywania zmiennych, przez M oznaczona jest liczba typów w populacji /patrz 1 w tabeli 3/, natomiast pojemność środowiska będzie w programie oznaczona przez S.

Tablica X /w tabeli 3 punkt 7/ jest punktem startowym do optymalizacji prawdopodobieństwa  $P\{S(TM), S(TM-1), \dots, S(0)\}$  - do identyfikacji parametrów modelu. Odpowiedniość pomiędzy elementami tablicy X a parametrami modelu jest następująca:

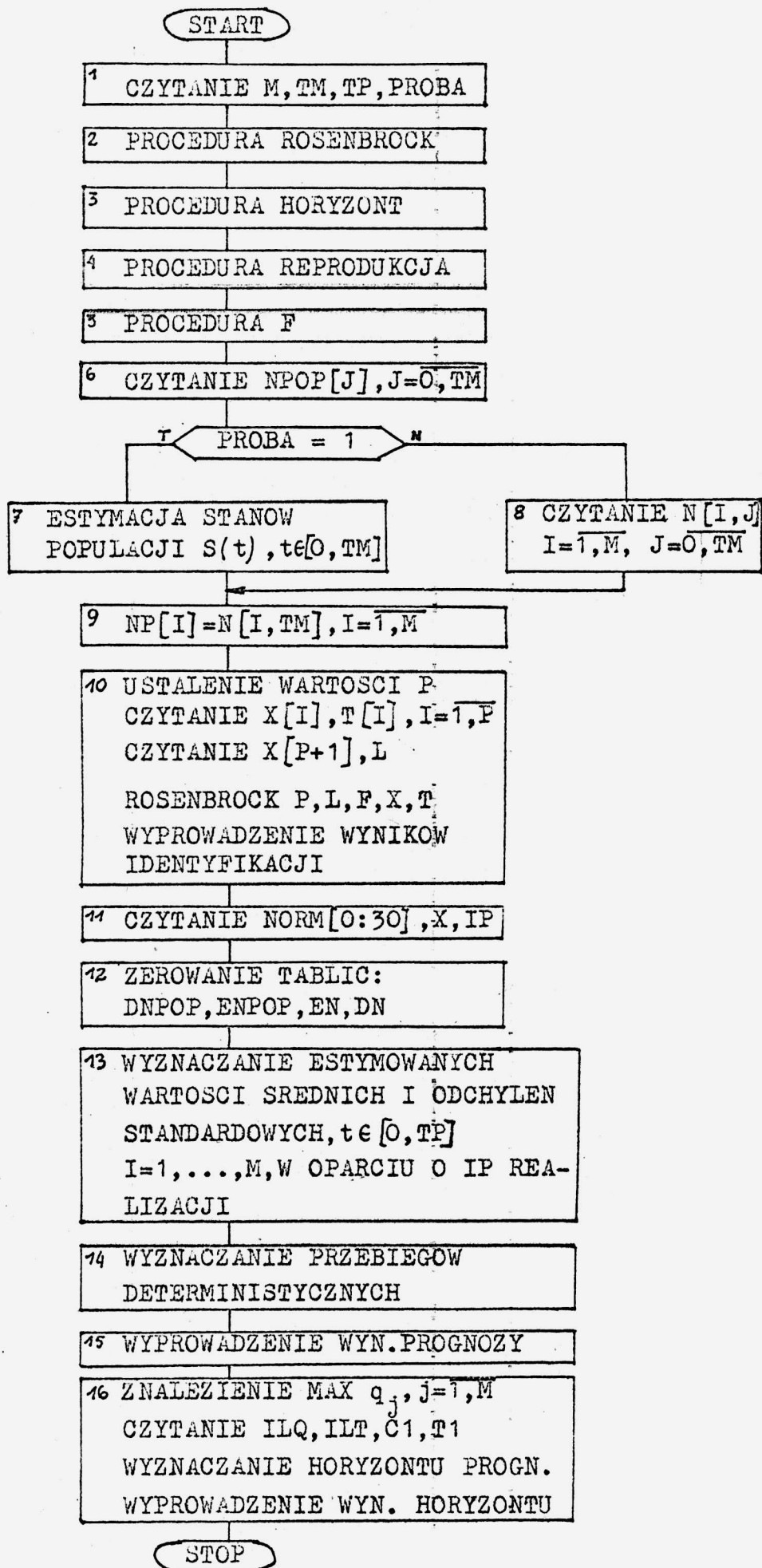
Dla wersji I /model dyskretny/ -  $X[1]=A$ ,  $X[2]=S/\text{pojemność środowiska}/$ ,  $X[3]=q_1$ ,  $X[4]=q_2, \dots, X[M+2]=q_m$ , dla wersji II /model ciągły/ -  $X[1]=A$ ,  $X[2]=S/\text{pojemność}/$ ,  $X[3]=B/\text{parametr } b \text{ w rozkładzie gamma } \Gamma(p, b)/$ ,  $X[4]=q_1, \dots, X[M+3]=q_m$ .

Ogólny schemat blokowy algorytmu prognozującego przedstawiony jest na rys. 5.1.1. Poszczególne bloki w schemacie ponumerowano w lewych górnych rogach.

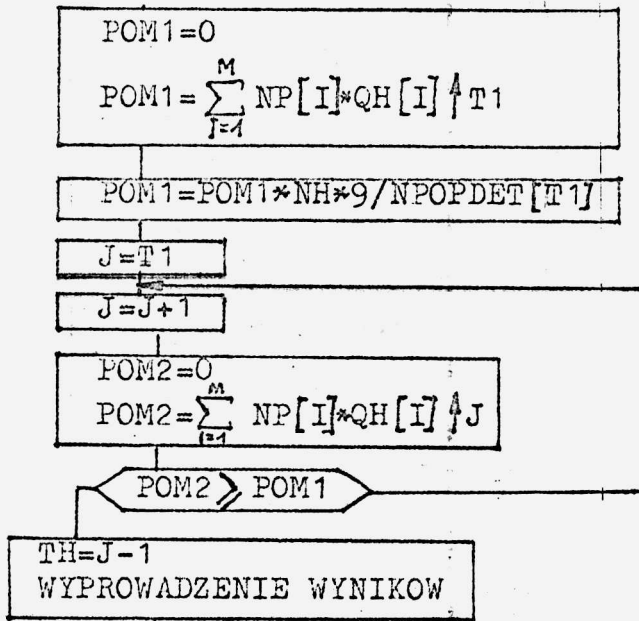
Blok nr 2 zawiera procedurę szukania minimum funkcji wielu zmiennych metodą Rosenbrocka. Procedurę tę nazwano ROSENBROCK. Powoduje ona znalezienie takich parametrów modelu, przy których funkcja celu  $F = -P\{S(TM), \dots, S(0)\}$  - przyjmuje minimum; czyli  $P\{S(TM), \dots, S(0)\}$  - maksimum.

Blok nr 3 - jest to procedura nazwana HORYZONT, wyznaczająca dla określonych warunków horyzont prognozy, zdefiniowany w rozdz. 3.1 i 3.5. Dokładniejszy schemat blokowy procedury HORYZONT przedstawiony jest na rys. 5.1.2. Procedura ta posiada dwa parametry formalne: QH i T1, gdzie QH jest tablicą M-wymiarową,  $QH[I]=q_i/q_{m+1}$ , T1 - czas wprowadzenia do populacji elementów (m+1)- ego typu.

Blok nr 4 jest procedurą losowej reprodukcji elementów poszczegól-



Rys. 5.1.1. Schemat blokowy algorytmu prognozującego.



Rys. 5.1.2. Schemat blokowy procedury HORYZONT,  
 NPOPDET =  $N_{\Sigma}$  liczone z modelu deterministycznego,  
 TH - horyzont prognozy.

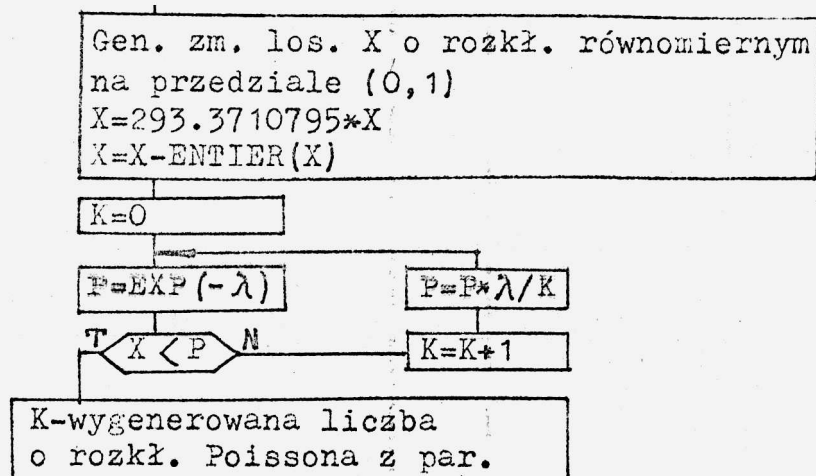
gólnych typów. Procedura REPRODUKCJA jest różna dla obu wersji algorytmu. Zanim zostanie omówiona budowa tej procedury, zaprezentowana zostanie zastosowana w algorytmie metodologia generowania zmiennych losowych o rozkładzie Poissona i gamma.

W modelu dyskretnym zachodzi konieczność generowania - dla każdego t z przedziału  $[0, T_F]$  - M zmiennych losowych o rozkładzie Poissona /liczby elementów poszczególnych typów/. Przy generowaniu tych zmiennych wykorzystano metodę odwracania dystrybuanty /13/, /31/. Schemat generowania zmiennej losowej o rozkładzie Poissona, z parametrem  $\lambda$  przedstawiono na rys. 5.1.3.

Zastosowany generator liczb losowych o rozkładzie równomiernym na przedz.  $(0, 1)$  został przebadany i wykazuje dobre własności statystyczne. Dla dużych wartości  $\lambda$ , rozkład  $Y = (K - \lambda) / \sqrt{\lambda}$

dąży do rozkładu normalnego  $\mathcal{N}(0, 1)$ . Można zatem /aby skrócić czas obliczeń na m.c./ wygenerować zmienną losową Y o rozkł.  $\mathcal{N}(0, 1)$





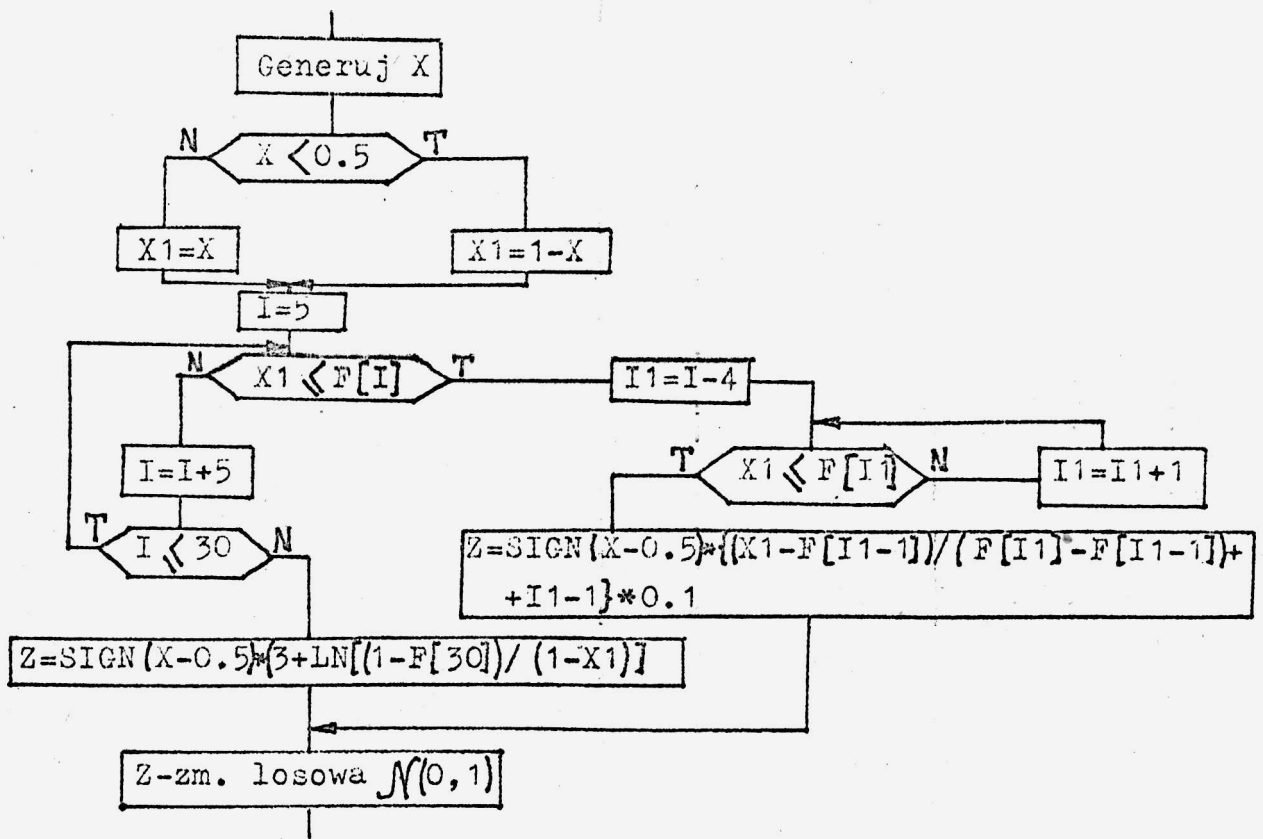
Rys. 5.1.3. Schemat generowania zmiennej losowej o rozkładzie Poissona z parametrem  $\lambda$ .

i wyznaczyć przybliżoną wartość zmiennej losowej o rozkładzie Poissona ze wzoru:

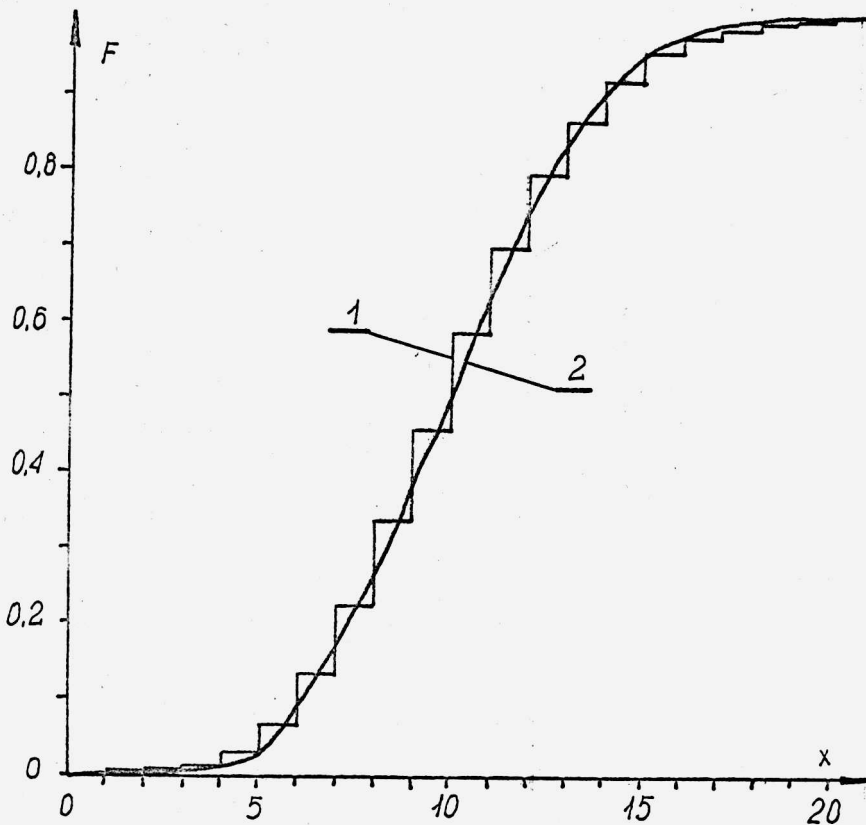
$$K = \text{ENTIER}(\sqrt{\lambda} * Y + \lambda + 0.5)$$

W algorytmie, przy generowaniu rozkładu  $\mathcal{N}(0,1)$  zastosowano również metodę odwracania dystrybuanty. Przybliżono dystrybuantę rozkładu normalnego : w przedziale  $(-3,3)$  - linią łamaną z krokiem 0.1, poza tym przedziałem - funkcją eksponencjalną. Ze względu na symetrię rozkładu normalnego, wystarczy pamiętać dodatnią "połówkę" rozkładu. Schemat blokowy generowania zmiennej losowej o rozkładzie  $\mathcal{N}(0,1)$  przedstawia rys. 5.1.4. Rysunek 5.1.5 pokazuje dokładność przybliżenia dystrybuanty rozkładu Poissona przy  $\lambda = 10$  rozkładem  $\mathcal{N}(10, \sqrt{10})$ .

W modelu ciągłym należy generować - w każdej chwili  $t \in [0, TP]$  -  $M$  zmiennych losowych o rozkładzie  $\Gamma(p, b)$  - wartości  $N_i(t)$ ,  $i = \overline{1, M}$ . Jeżeli zmienna losowa  $Z$  ma rozkład  $\Gamma(p, 1)$ ,  $Z' \sim \Gamma(p, b)$ , to  $Z = b \cdot Z'$  / 31/. Zmienną  $Z$  można przedstawić w postaci  $Z = Z_1 + Z_2 - Z_3$ , gdzie:  
 $Z_1 \sim \Gamma(p', 1)$ ,  $p' = \text{ENTIER}(p)$ ,  
 $Z_2 \sim \Gamma(1, 1)$  - rozkład wykładniczy z wartością oczekiwaną = 1,  
 $Z_3 \sim \beta(p_1, 1-p_1)$  rozkład beta,  $p_1 = p - p'$ .



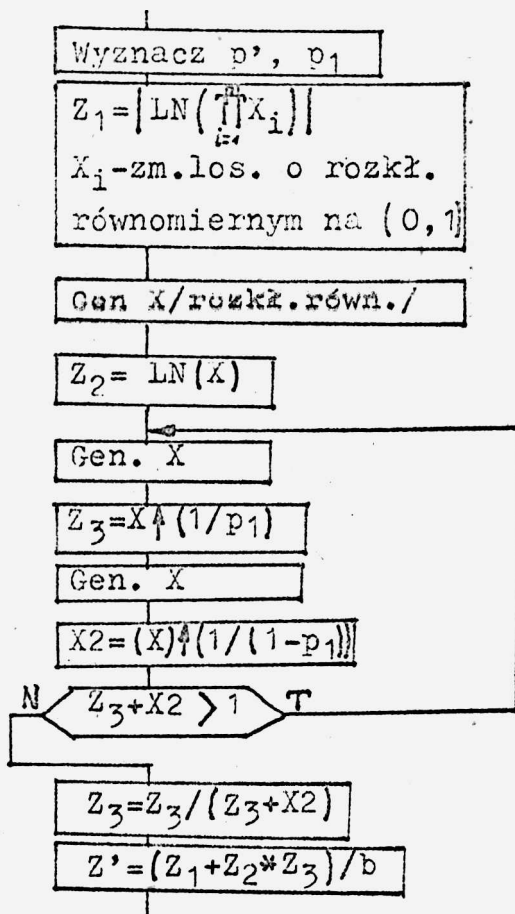
Rys. 5.1.4. Schemat generowania zmiennej losowej  $\mathcal{N}(0, 1)$ .



Rys. 5.1.5. Dystrybuanta rozkładu Poissona dla  $\lambda = 10$  i rozkładu normalnego  $\mathcal{N}(10, \sqrt{10})$ .

1 - dystrybuanta rozkładu Poissona,

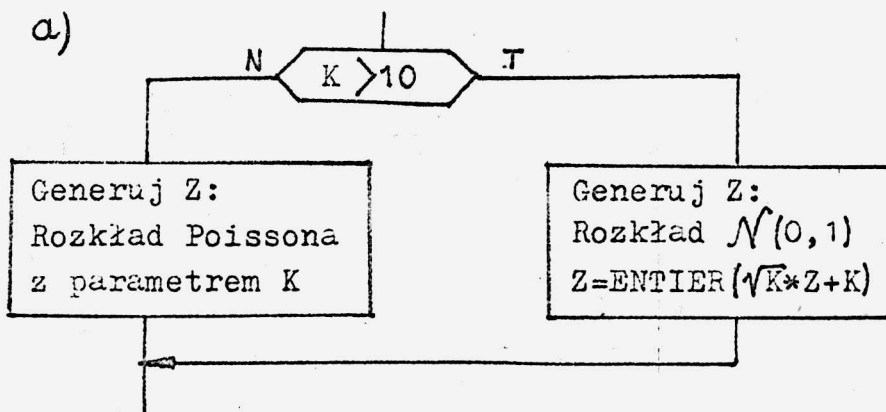
2 - dystrybuanta rozkładu normalnego.

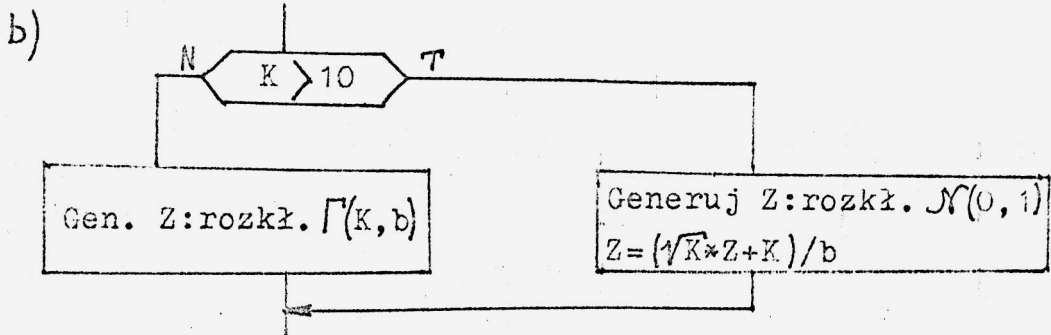


Rys. 5.1.6. Schemat generowania zmiennej losowej o rozkładzie  $\Gamma(p, b)$ .

Dokładny schemat generowania zmiennej losowej  $Z'$  /o rozkł.  $\Gamma(p, b)$ / przedstawiony jest na rys. 5.1.6.

Na rys. 5.1.7a pokazany jest blok nr 4 ze schematu 5.1.1, w przypadku modelu dyskretnego /wersja I/, na rys. 5.1.7b - w przypadku modelu ciągłego /wersja II/. Procedura REPRODUKCJA(K) ma jeden parametr formalny: K oznacza  $E[N_i(t)]$  w wersji I, i K - oznacza  $E[N_i(t)] * b = p$  w wersji II.





Rys. 5.1.7. Procedura REPRODUKCJA(K). a/ wersja I, b/ wersja II.

Blok nr 5 algorytmu prognozującego wyznacza wartości funkcji celu do optymalizacji. Wyznacza prawdopodobieństwo zajścia sekwencji stanów populacji  $S(0), S(1), \dots, S(TM)$  w obu wersjach. Dla wersji pierwszej wyznaczana jest wartość wyrażenia /3.3.3/ z rozdziału 3.3, natomiast dla wersji II algorytmu - wartość wyrażenia /3.4.3/.

Blok nr 7 algorytmu wykonuje estymację stanów populacji dla  $t \in [0, TM]$  zgodnie z estymatorem zaproponowanym w rozdziale 3.2.

W bloku nr 9 zapamiętywany jest stan początkowy populacji służący do wyznaczania kolejnych realizacji prognozy i przebiegu deterministycznego.

Blok 10 różni się w wersji I i II ustaleniem ilości zmiennych optymalizacyjnych /do identyfikacji parametrów modelu/: w wersji I -  $P=M+1$ , w wersji II -  $P=M+2$  /wymiar tablicy T/. Pozostałe bloki w algorytmie są jednakowe w obu wersjach /z wyjątkiem zadawanego parametru przy wywoływaniu procedury REPRODUKCJA(K)/, Algorytm wyznacza zmiany stanu populacji w czasie zgodnie z zaprezentowanym wcześniej modelem.

## 5.2. Prognoza pewnego procesu rozwoju.

Przeprowadzono szereg eksperymentów testujących zaproponowany algorytm. W niniejszym rozdziale przedstawiono jeden z takich eksperymentów. Jako rzeczywisty przebieg procesu rozwoju

przyjęto wynik eksperymentu na modelu uwzględniającym niezerowe prawdopodobieństwo zmiany typu w procesie reprodukcji, wskutek czego w populacji pojawiają się nowe typy, zarówno lepsze jak i gorsze /10/, /15/.

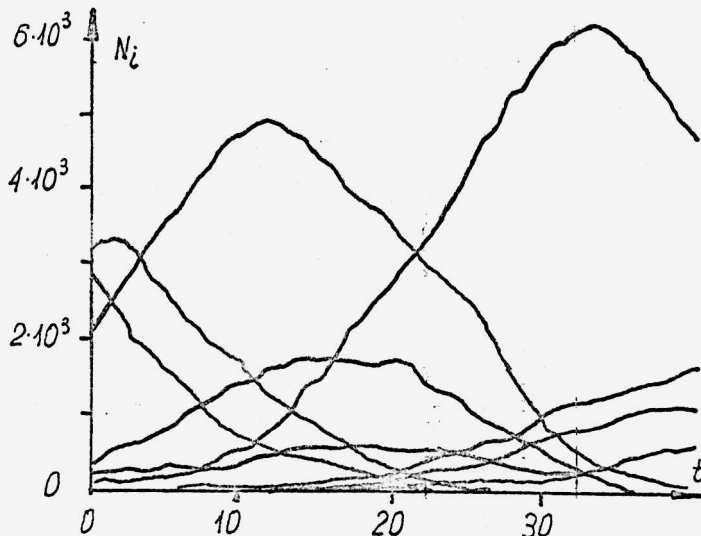
Stan populacji i jakości typów występujących w populacji w chwili początkowej dane są w tabeli 4.

TABELA 4.

Liczebności i jakości poszczególnych typów w

Typ	Ilość elem.	Jakość	Typ	Ilość elem.	Jakość
1	3	0.383	11	2	0.173
2	24	0.444	12	6	0.312
3	238	0.405	13	727	0.323
4	151	0.312	14	2764	0.301
5	1	0.220	15	300	0.258
6	4	0.383	16	2	0.193
7	77	0.405	17	16	0.249
8	2054	0.379	18	128	0.258
9	3293	0.323	19	116	0.234
10	54	0.249	20	1	0.193

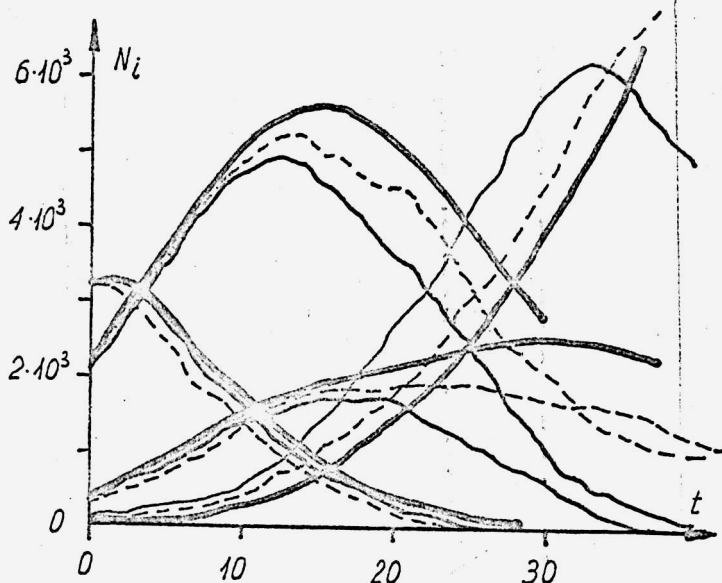
Środowisko było nasycone, więc funkcja  $g(t)=1$ , reprodukcja losowa o rozkładzie Poissona,  $E[N_2(t)] = N_2(0) = 10000$ .



Rys. 5.2.1. Liczby elementów wybranych typów z tabeli 4 w czasie.

Rysunek 5.2.1 pokazuje zmianę liczebności kilku typów z powyższej tabeli. Na rysunku pokazane są również niektóre nowo powstałe typy w procesie "rzeczywistym". Pierwszy nowy typ, lepszy od wymienionych w tabeli, pojawił się w  $t=7$ .

W pracy przyjęto, że znany jest stan w chwili początkowej  $S(0)$  i jakości typów występujących w  $S(0)$ . W oparciu o te dane przeprowadzono prognozę zmian liczebności typów, występujących w  $t=0$ . Następnie porównano otrzymane rezultaty z "rzeczywistą" realizacją przebiegu. Przykładowe krzywe pokazuje rysunek 5.2.2.



Rys. 5.2.2. Wynik prognozy dla typów o numerach 2, 3, 8, 9.

- realizacja prognozowanego procesu /"rzeczywista"/,
- — — deterministyczna prognoza,
- - - przykładowa realizacja otrzymana przy prognozowaniu wykorzystującym model probabilistyczny.

Biorąc pod uwagę fakt, że w  $t=30$  w populacji "rzeczywistej" istniało już 3 nowe typy, z których liczebność najlepszego wynosiła ponad 1000 elementów, prognoza wydaje się być całkiem niezła. Dla  $t=15$ , maksymalny błąd procentowy /tak przy deterministycznej jak i probabilistycznej prognozie/ nie przekraczał

dwadziestu procent.

Przeprowadzono podobne eksperymenty, w których jakości typów były identyfikowane. Ze względu na czas wykonywania obliczeń na maszynie cyfrowej ograniczono się wówczas do mniejszej liczby typów w prognozowanym procesie rozwoju niż w zaprezentowanym eksperymencie /mniejsza ilość zmiennych optymalizacyjnych/. Otrzymane rezultaty wydają się potwierdzać możliwość wykorzystania zaproponowanego algorytmu do prognozowania rzeczywistych procesów rozwoju.

## 6. PRZYKŁAD PROGNOZY

Przeprowadzono - jako przykład - prognozę zużycia energii: na świecie do 2011 roku /50 lat od 1961r./. Dane wejściowe uzyskano z materiałów opracowanych w 1978r. w International Institute for Applied Systems Analysis w Laxenbergu, w Austrii /18/,/21/. Jako znany okres /historię zmian stanu populacji/ potraktowano lata 1951 - 1961. W tych latach istniały trzy typy energii:

- typ 1 - jako surowiec do wytwarzania energii służy ropa naftowa,
- typ 2 - surowcem jest gaz naturalny,
- typ 3 - surowcem jest węgiel.

Energia jądrowa w tym czasie jeszcze nie była rozpowszechniona, autorzy uwzględniają ją dopiero od 1961r. Natomiast energia, do wytwarzania której używano drewna nie jest uwzględniana przez autorów, ponieważ jej udział jest już bardzo mały.

W tabeli 5 podano ilość zużytej na świecie energii wymienionych wyżej trzech typów.

TABELA 5

Ilość zużytej na świecie energii w latach 1951 - 1961 /3 typy/  
w [MLD TCE].

Rok	Typ 1	Typ 2	Typ 3	Ogółem
1951	710	320	1630	2660
1952	750	340	1635	2725
1953	800	360	1640	2800
1954	840	380	1640	2860
1955	950	400	1760	3110



TABELA 5 /cd./

Rok	Typ 1	Typ 2	Typ 3	Ogółem
1956	1030	430	1860	3320
1957	1070	460	1930	3460
1958	1140	500	2020	3660
1959	1220	570	2100	3890
1960	1320	620	2190	4130
1961	1400	670	2160	4230

Wykorzystano drugą wersję algorytmu, tzn. reprodukcja elementów była procesem losowym o rozkładzie gamma. Zadano L=200 iteracji do identyfikacji parametrów modelu. Po 126 iteracjach metoda szukania minimum funkcji Rosenbrocka została przerwana wskutek braku polepszenia przy wykonaniu kroków we wszystkich kierunkach.

Wyniki identyfikacji są następujące:

Pojemność środowiska  $S=2.0192 \cdot 10^9$ ,

Współczynnik max. repr.  $A=4.8148 \cdot 10^{-2}$ ,

Parametr rozkł. gamma  $B=1.4203 \cdot 10^{-1}$ ;

Jakość elementów typu 1 -  $q_1=103.7870$ ,

typu 2 -  $q_2=105.0364$

typu 3 -  $q_3=100.0000$  /wartość zadana/.

Przeprowadzono eksperyment, w którym za punkt startowy przyjęto rok 1951, tzn. pierwszy rok znanej historii. Celem tego eksperymentu było sprawdzenie, jak uzyskany po identyfikacji model przybliży krzywe wykorzystane do identyfikacji. W kolejnym eksperymencie, za punkt startowy przyjęto końcowy czas ze znanej historii, to jest 1961 rok. Przeprowadzono prognozę na okres 50 lat, tzn. do 2011 roku. Dostępne w /18/i/21/ dane rzeczywiste sięgały do 1974 roku. Do tego czasu możliwe jest porównanie otrzymanej prognozy z przebiegiem rzeczywistym.

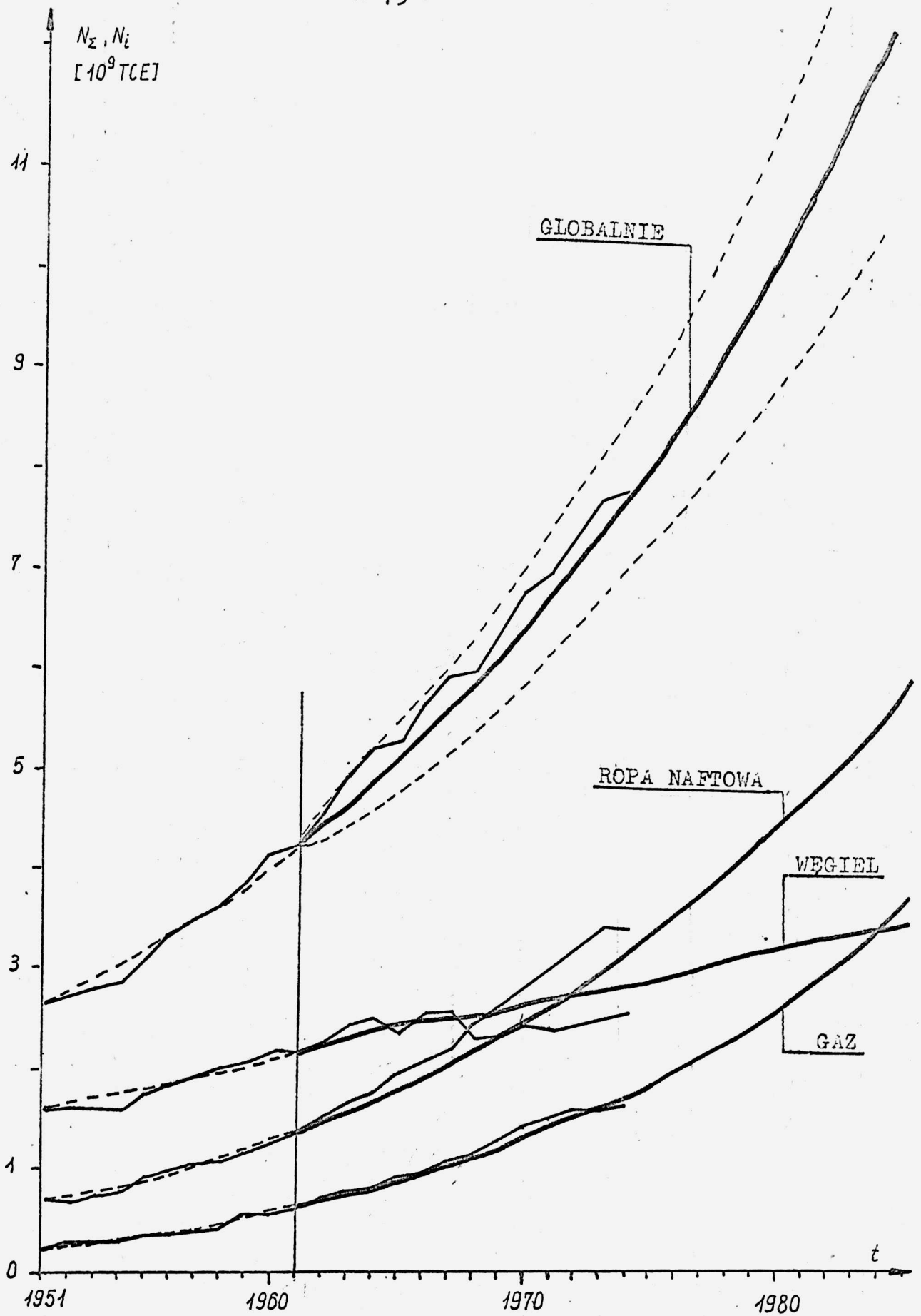
Na rysunku 6.1 przedstawione są krzywe zmian zużycia energii

do 1974 roku /linia ciągła, cienka/. Linia grubą przerywaną zaznaczono przybliżenie uzyskane z modelu krzywych, służących jako dane wejściowe do identyfikacji, Linia grubą zaznaczone są prognozy - estymowane wartości średnie  $E[N_i(t)]$ ,  $i=1,2,3$ . Linia cienką przerywaną zaznaczony jest przedział  $E[N_i(t)] \pm \sqrt{D^2[N_i(t)]}$ . Wartości średnie i odchylenia standardowe estymowane były na podstawie 50 realizacji. Na rysunkach 6.2.a,b,c pokazano prognozę - estymowane wartości średnie i przedziały  $E[N_i(t)] \pm \sqrt{D^2[N_i(t)]}$  odpowiednio dla typu 1,2,3. Na rys. 6.1 widoczne są dość duże wahania poszczególnych krzywych przebiegów rzeczywistych. Dlatego też występuje dość duża wariancja ilości elementów poszczególnych typów /rys. 6.2/. W roku 2000 dla energii produkowanej z węgla /typ 3/ odchylenie standardowe wynosi ok. 38 % /dla typu 1 i 2 ok. 27 %/.

Przeprowadzono ten sam eksperyment wykorzystując pierwszą wersję algorytmu /reprodukcja losowa o rozkładzie Poissona/. Po  $L=150$  iteracjach uzyskano następujące parametry modelu:

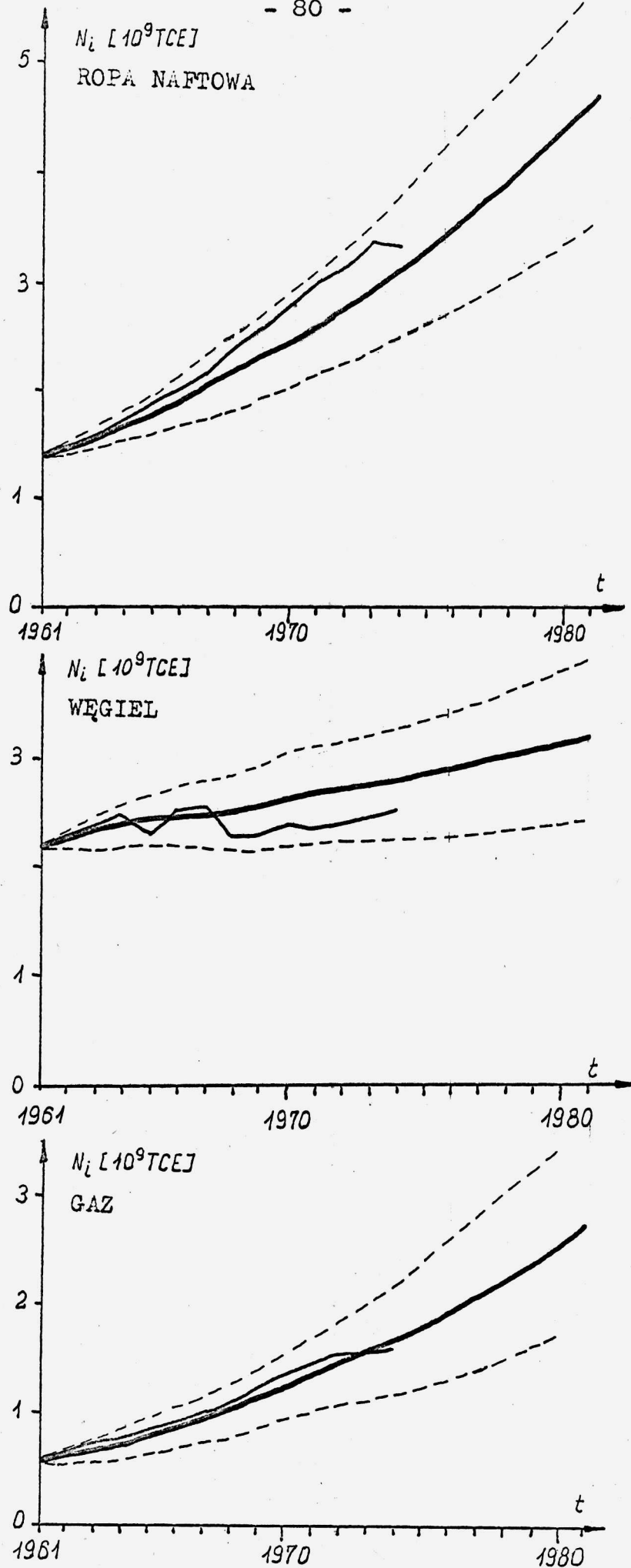
$$S=1.6227 \cdot 10^8, \quad A=4.8761 \cdot 10^{-2}, \quad q_1=104.0264, \quad q_2=104.9900, \\ q_3=100.0000.$$

W tabeli 6 zestawiono rezultaty prognozy - estymowane wartości średnie i odchylenia standardowe  $N_i(t)$ ,  $i=1,2,3$ , uzyskane przy wykorzystaniu algorytmu prognozującego w wersji I i wersji II. Widoczne są małe zmiany w wartościach średnich ilości elementów poszczególnych typów w populacji otrzymanych przy wykorzystaniu obu wersji. Porównując  $E[N_i(t)]$  z algorytmu I, z algorytmu II i  $N_i(t)$  otrzymane z modelu deterministycznego, można zauważyć, że różnice między nimi /dla ustalonych  $t$  i typu/ nie przekraczają kilku procent. Istnieją znaczne różnice w odchyleniach standardowych poszczególnych zmiennych losowych /tabela 6/. Przy zastosowaniu



Rys. 6.1. Dane rzeczywiste i prognoza zużycia energii na świecie.

- dane rzeczywiste.
- prognoza od roku 1961.
- - - przedział  $[N_{\Sigma}(t)] \pm \sqrt{D^2[N_{\Sigma}(t)]}$
- · · prognoza od roku 1951.



Rys. 6.2. Prognoza zużycia energii: estymowane wartości średnie i przedział  $E[N_i(t)] \pm \sqrt{D^2[N_i(t)]}$ ,  $i=1,2,3$ , na tło uwzględniające niepewności w danych podstawowych.

TABELA 6.

Rezultaty prognozy zużycia energii na świecie, uzyskane przy zastosowaniu wersji I i wersji II algorytmu prognozującego.

Rok	ALGORYTM Z MODELEM DYSKRETNYM						ALGORYTM Z MODELEM CIĄGŁYM					
	$E[N_1]$	$\sqrt{D^2[N_1]}$	$E[N_2]$	$\sqrt{D^2[N_2]}$	$E[N_3]$	$\sqrt{D^2[N_3]}$	$E[N_1]$	$\sqrt{D^2[N_1]}$	$E[N_2]$	$\sqrt{D^2[N_2]}$	$E[N_3]$	$\sqrt{D^2[N_3]}$
1962	1499	36	719	23	2226	44	1502	95	715	62	2240	115
1964	1700	77	838	58	2340	78	1687	205	844	152	2352	207
1966	1943	109	968	78	2451	109	1936	293	974	207	2450	287
1968	2201	146	1117	97	2558	134	2177	390	1121	253	2539	353
1970	2495	182	1296	110	2672	177	2461	489	1312	299	2647	465
1972	2814	212	1496	147	2798	180	2750	561	1527	405	2779	472
1974	3182	251	1717	185	2892	201	3106	696	1750	505	2833	522
1976	3610	292	1957	228	2988	225	3550	804	1976	614	2895	583
1978	4051	324	2241	268	3088	268	3952	838	2265	725	2970	688
1980	4566	351	2552	328	3198	297	4458	906	2569	888	3071	762
1990	7942	724	4933	649	3750	451	7578	1846	5071	1745	3624	1160
2000	13345	1270	9133	1163	4311	616	12418	3189	9446	3102	4302	1633
2010	21888	2104	16355	2057	4777	755	19955	5245	17000	5412	4835	1991

waniu rozkładu Poissona do procesu reprodukcji nie ma możliwości dopasowania wariancji rozkładu do wariancji prognozowanego procesu /jest ona równa wartości oczekiwanej/. Natomiast przy wykorzystaniu rozkładu gamma  $\Gamma(p, b)$  - wariancja ilości elementów poszczególnych typów jest identyfikowana w oparciu o dane historyczne /poprzez identyfikację parametru  $b$ /.

Z danych z /18/ wynika, że w 1961 roku zaczęła być produkowana energia jądrowa /w zauważalnej ilości/. Wyznaczono horyzont prognozy, dla następujących parametrów:

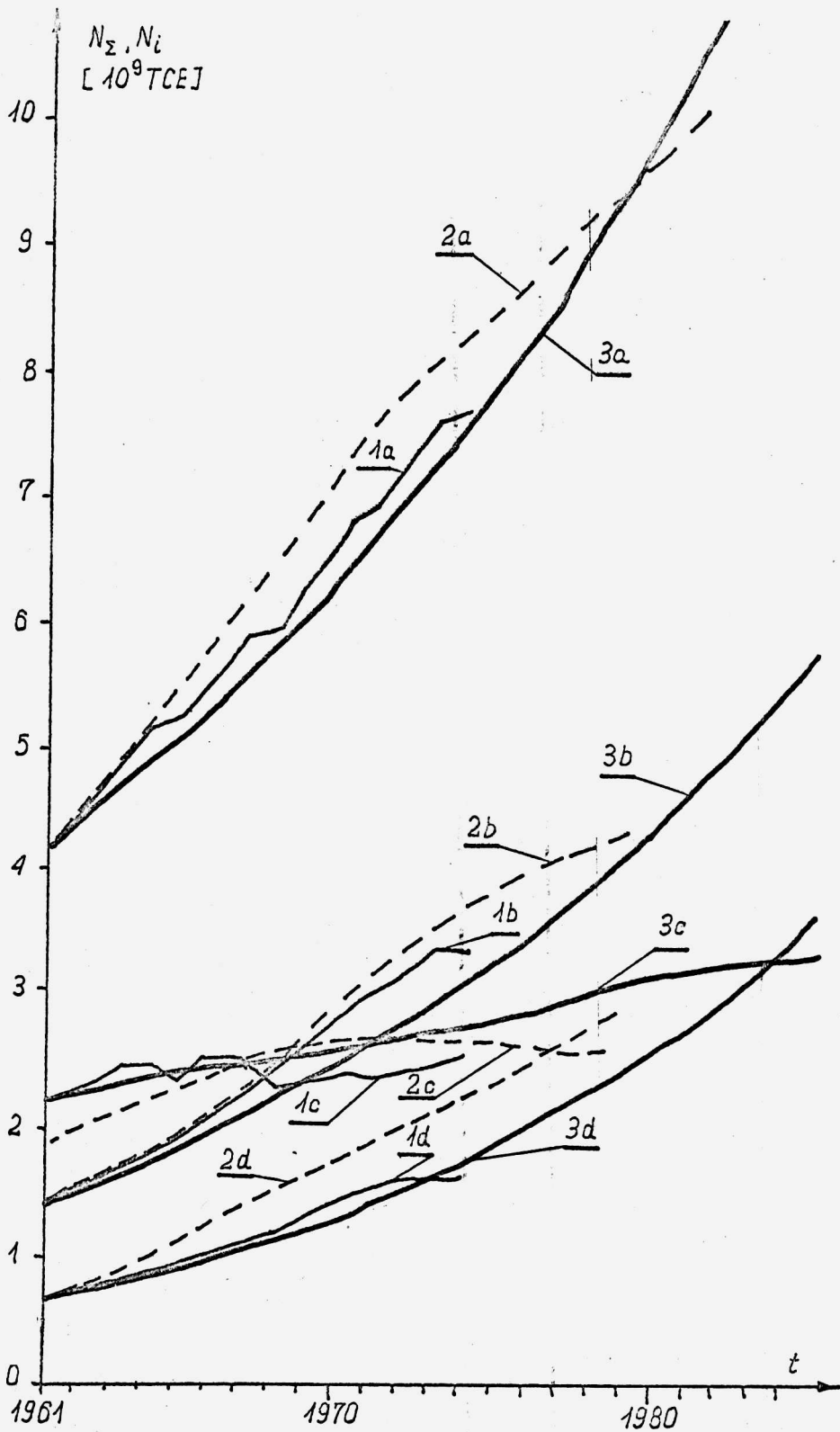
$$\begin{aligned} q_4 &= 1.2 * q_{\max} \quad / = 126.0437 /, \quad T_1 = 1 & t_h &= 31 \text{ lat,} \\ q_4 &= 1.5 * q_{\max} \quad / = 157.5546 /, \quad T_1 = 1 & t_h &= 15 \text{ lat,} \\ q_4 &= 1.8 * q_{\max} \quad / = 189.0655 /, \quad T_1 = 1 & t_h &= 11 \text{ lat,} \\ q_4 &= 1.2 * q_{\max} \quad / = 126.0437 /, \quad T_1 = 10 & t_h &= 43 \text{ lata.} \end{aligned}$$

Zakładając, że istotnie lepszy typ będzie wprowadzony do populacji już w następnym roku, tzn. 1962, przez kilkanaście lat nie uzyska on liczebności 10 % wielkości populacji.

Na rys. 6.3 przedstawiona jest prognoza zużycia energii przeprowadzona przez C. Marcetti 'ego i N. Nakicenowic'a w /18/ oraz prognoza zużycia energii otrzymana przy wykorzystaniu zaproponowanego w rozdz. 5 algorytmu, na tle przebiegów rzeczywistych.

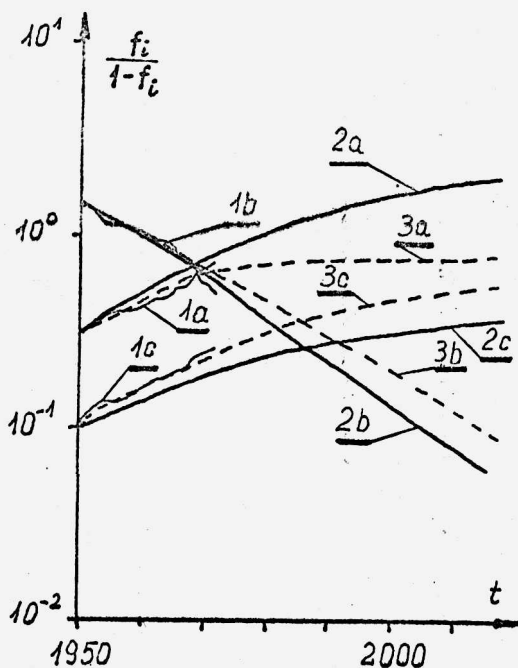
W pracy /21/ autorzy wykonali prognozę zużycia energii wykorzystując model substytucji przedstawiony krótko w rozdz. 1.3.

Jako wynik prognozy autorzy podali wykres w skali logarytmicznej udziałów energii poszczególnych typów /bez nuklearnej/. Fragment tego wykresu jest przerysowany na rysunku 6.4. Na ten rysunek naniesione są również krzywe udziałów różnych typów energii, otrzymane w niniejszej pracy.



Rys. 6.3. Prognozy: wyznaczona w pracy /18/ i w niniejszej pracy na tle przebiegów rzeczywistych.

- 1 - dane rzeczywiste zużycia energii,
- 2 - prognoza z /18/,
- 3 - prognoza otrzymana w niniejszej pracy.
- a - globalnie, b - ropa naftowa, c - węgiel, d - gaz



Rys. 6.4. Fragment prognozy zużycia energii na świecie z pracy /21/ rys. 8, str. 30 . Liniją przerywaną zaznaczona jest prognoza uzyskana przy zastosowaniu algorytmu przedstawionego w rozdz. 5.

- 1 - dane rzeczywiste,
- 2 - prognoza z /21/,
- 3 - prognoza z rys. 6.1.
- a - ropa naftowa
- b - węgiel
- c - gaz

$f_i$  - udział energii i-tego typu w zużyciu globalnym.



## W n i o s k i

Podjęta w pracy próba wykorzystania modelu dynamiki populacji do prognozowania technicznego wydaje się być obiecująca.

Zaproponowany w rozdziale 2 model zmian stanu populacji w czasie, zbudowany w oparciu o przesłanki biologiczne zdaje się odpowiadać zapotrzebowaniu wyrażonemu w /16/ na probabilistyczny model substytucji, uwzględniający konkurencję elementów wielu typów. Model ten opisuje klasę "spontanicznych" procesów rozwoju, w których preferencje dla poszczególnych typów, wyrażane przez relatywne wskaźniki jakości, są proporcjonalne do prawdopodobieństwa reprodukcji elementu danego typu. Założono ponadto, że preferencje te są stałe w czasie.

Zastosowany do budowy algorytmu model jest na tyle złożony, że autorowi nie udało się określić efektywnie rozkładów prawdopodobieństw liczby elementów każdego typu w populacji, w dowolnej chwili czasu  $t$ , w zależności od stanu początkowego populacji. Narzuciło to konieczność podejścia symulacyjnego i wyznaczania wartości oczekiwanych i odchyłeń standardowych poprzez estymację z zadanej krotności realizacji.

Uderzające jest podobieństwo charakterystyk dynamicznych rzeczywistych procesów rozwoju /5/, /23/ i charakterystyk otrzymanych w rezultacie badań symulacyjnych modelu /rozdział 4 /. W obu przypadkach występują logistyczne i epidemiologiczne krzywe wzrostu, podobny jest wpływ nienasycenia środowiska, itp.

Zastosowanie rozkładu gamma do opisu procesu reprodukcji elementów umożliwia dopasowanie wariacji rozkładów liczby elementów poszczególnych typów w populacji /poprzez identyfikację parametru

b w rozkładzie  $\Gamma(p, b)$  do wariacji obserwowanej w procesie rzeczywistym.

Opracowany algorytm prognozujący /rozdział 5/, oprócz przebiegów deterministycznych i estymowanych wartości średnich, dostarcza dodatkowo informacji o zmianie wiarygodności prognozy z upływem czasu. Wyznaczane są odchylenia standardowe ilości elementów poszczególnych typów i wielkości całej populacji, jak również horyzont prognozy, umożliwiający ocenę wpływu "nowych wynalazków".

Zaproponowany algorytm prognozujący nie wymaga danych ekonomicznych o prognozowanym procesie. Może to być z jednej strony wadą algorytmu - nie uwzględnia on parametrów ekonomicznych procesu, z drugiej zaś zaletą - dane takie są trudno dostępne /28/, ich brak uniemożliwia wykorzystanie algorytmu.

Można przyjąć, że wskaźniki jakości elementów syntetyzują w pewnym sensie parametry ekonomiczne procesów ich produkcji, dystrybucji i eksploatacji. Wskaźniki te nie muszą być dane explicite. Proponowany w rozdziale 5 algorytm dokonuje estymacji wskaźników jakości elementów na podstawie znanej historii zmian stanu populacji /lub losowej próbki z populacji/.

W rozdziale 6 zaprezentowano przykład wykorzystania zaproponowanego algorytmu do prognozowania rzeczywistych procesów rozwoju. Jest to prognoza zużycia energii na świecie, na lata 1961 - 2011, w oparciu o znaną historię zmian z lat 1951 - 1961. Porównując otrzymane rezultaty z danymi rzeczywistymi /do 1974 roku/ oraz z prognozą z prac /18/ i /21/, można powiedzieć, że są one zadowalające.

Zdaniem autora, zaproponowany algorytm może być wykorzystywany w różnych dziedzinach techniki, niekiedy dość odmiennych, bez

szczegółowego wnikania w ich specyfikę.

Autor widzi kilka możliwości lepszego przybliżenia zaproponowanego w pracy modelu do rzeczywistości. Jedną z nich jest uwzględnienie struktury wiekowej populacji, podobnie jak w pracy /22/. W kolejnych chwilach czasu elementy populacji podlegałyby dwóm procesom losowym: procesowi reprodukcji i procesowi przeżywania do kolejnej grupy wiekowej /"starzenia się"/. Prawdopodobieństwa reprodukcji elementów zależałyby wtedy od ich grupy wiekowej.

Inna możliwość udoskonalenia modelu wiązałaby się z dopuszczeniem zmiennego w czasie zapotrzebowania środowiska. Szczególnie interesującym byłoby też wykrywanie i uwzględnianie stałych trendów w zmianach wskaźników jakości elementów poszczególnych typów.

Przedstawiony w niniejszej pracy model opisuje zaspokajanie konkretnej potrzeby w jednorodnym środowisku. Producenta może interesować prognoza dotycząca produkowanych przez niego elementów, które mogą zaspokajać różne potrzeby w różnych środowiskach. W takim przypadku celowe byłoby zbudowanie modelu globalnego, obejmującego podmodele opisujące zmiany stanu w tych populacjach, w których funkcjonują elementy produkowane przez tego producenta.

Naturalnie, w modelu uwzględniającym powyższe sugestie znacznie powiększy się liczba parametrów. Pojawi się wtedy niebłachy problem wyznaczenia interesujących rozkładów prawdopodobieństwa oraz przeprowadzenia identyfikacji parametrów takiego modelu.

Autor składa podziękowania Dyrektorowi Ośrodka Badań Progностycznych PWr dr K.I. Pelcowi za udostępnienie niepublikowanych danych, umożliwiającą sprawdzenie zaproponowanego w pracy algorytmu i zaprezentowanie przykładu prognozy.

L I T E R A T U R A

- /1/ Athreya K.B., Ney P.E.: Branching Processes. Berlin, 1972.
- /2/ Ayres R.U.: Prognozowanie rozwoju techniki i planowanie długookresowe. Warszawa, 1973.
- /3/ Baworowski J.L.: Prognozowanie nauki i techniki. Wrocław, 1977
- /4/ Bubnicki Z.: Identyfikacja obiektów sterowania. Warszawa, 1974
- /5/ Cetron J.M.: Technological Forecasting. New York, London, Paris, 1969.
- /6/ Dinamičeskaja teorija biologičeskich populacji, p. red. R.A. Poluektowa. Moskwa, 1974.
- /7/ Eigen M.: Selforganization of Matter and Evolution of Biological Macromolecules. Die Naturwissenschaften 58, 465-523, 1971.
- /8/ Feller W.: Wstę do rachunku prawdopodobieństwa. Warszawa, 1966.
- /9/ Fisher R.A.: The Genetical Theory of Natural Selection. Dover, New York, 1958.
- /10/ Galar R., Kwaśnicka H., Kwaśnicki W.: Simulation of Some Processes of Development. IMACS Congress Simulation of Systems 79, Sorrento, September, 1979.
- /11/ Hansell R.J.C., Marchi E.: Aspects of Evolutionary Theory and the Theory of Games. Victoria Conference, 1973.
- /12/ Harris T.E.: The Theory of Branching Processes. Berlin, 1963.
- /13/ Jermakow S.M., Michailow G.A.: Kurs statističeskowo modelirovanija. Moskwa, 1976.
- /14/ Karlin S.: Equilibrium Behavior of Population Genetic. Models with Non-random Mating. J.Appl.Prob. 5, 487-566, 1968.

- /15/ Kwaśnicki W.: Symulacja pewnej klasy ewolucyjnych procesów rozwoju. Praca doktorska, Wrocław, 1979,
- /16/ Linstone H.A., Sahal D.: Technological Substitution. New York, Oxford, Amsterdam, 1976.
- /17/ Lofgren L.: Względne objaśnianie systemów, w "Ogólna teoria systemów" p. red. G.J.Klira. Warszawa, 1976, /327-389/.
- /18/ Marcetti C., Nakicenovic N.: The Dynamics of Energy Systems and the Logistic Substitution Model. V.1, Administrative Report, IIASA, Laxemburg, 1978.
- /19/ Milsum J.H.: Podstawa hierarchiczna dla systemów ogólnych żywych, w "Ogólna teoria systemów" p. red. G.J.Klira, Warszawa, 1976, /143-186/.
- /20/ Mode C.J.: Multitype Branching Processes. New York American Elsevier, 1971.
- /21/ Peterka V., Fleck F.: The Dynamics of Energy Systems and the Logistic Substitution Model. V.2, Administrative Report, IIASA, Laxemburg, 1978.
- /22/ Pollard J.H.: Mathematical Models for The Growth of Human Populations. Cambridge, 1973.
- /23/ Prognozowanie w technice, p.red. J.R.Brighta, M.E.Schoemana, Warszawa, 1978.
- /24/ Sharif M.N., Kabir C.: System Dynamics Modeling for Forecasting Multilevel Technological Substitution, in /16/.
- /25/ Siewastanow B.A.: Wietwiaszcziesja processy. Moskwa, 1971.
- /26/ Smith J.M.: Matematyka w biologii. Warszawa, 1974.
- /27/ Steindl J.: Technical Progress and Evolution, Referat na konferencję " Science Policy and R D Management", Berlin Zachodni, 6-10.12.1978.
- /28/ Stern M.O., Ayres R.U., Shapanka A.: A Model for Forecasting

- the Substitution of One Technology for Another, in /16/.
- /29/ Thompson C.J., McBride J.L.: On Eigen's Theory of The Self Organization of Matter and the Evolution of Biological Macromolecules. Mathematical Biosciences 21, 127-142, 1974.
- /30/ Waddington C.M.: Stabilization in Systems. Chreods and Epigenetic Landscapes. Futures, 139-146, 1977.
- /31/ Zieliński R.: Generatory liczb losowych. Warszawa, 1972.

Raport złożono w Redakcji  
Instytutu 09. 1979.

mgr inż. Halina Kwaśnicka.

Niniejszy raport otrzymują:

1. OINT - Biblioteka Międzyinstytutowa I-6,25	1
2. Biblioteka Główna PWr	1
3. Z-ca Dyrektora d/s Kształcenia Kadry Naukowej	2
4. Promotor - prof. Z. Szparkowski	1
5. Recenzenci	2
6. Wydział Informatyki Uniwersytetu Jagiellońskiego	1
7. Dyrektor OBP PWr dr K.I.Pelc	1
8. Dr inż. R.Galar	1
9. Autor	7
<hr/>	
RAZEM:	17 egz.