

Określenie zasadności modeli w naukach społecznych

Witold Kwaśnicki

Instytut Nauk Ekonomicznych, Uniwersytet Wrocławski

kwasnicki@prawo.uni.wroc.pl

<http://prawo.uni.wroc.pl/~kwasnicki>

Streszczenie

Przedstawiono problemy z określaniem stopnia podobieństwa zachowania się modeli procesów społeczno-gospodarczych do procesów rzeczywistych. Podkreślono konieczność wypracowania specyficznych metod dla tej klasy procesów i problemy z zastosowaniem znanych metod określania zasadności modeli w naukach technicznych do nauk społecznych. Wyrażono opinię, że jednym z możliwych podejść jest określenie podobieństwa na poziomie tzw. wyidealizowanych faktów oraz, że skutecznym może okazać się wykorzystanie analizy spektralnej.

Badanie procesów gospodarczych i społecznych dokonywane jest często przy wykorzystaniu podejścia symulacyjnego, zwłaszcza poprzez badanie wyidealizowanych modeli komputerowych tych procesów. Konieczność stosowania tej techniki wynika z dużej złożoności systemów społeczno-gospodarczych i konieczności opisu tych procesów w kategoriach modeli nieliniowych, trudnych, ze swej natury, dla badania analitycznego. Badanie symulacyjne jest specyficznym sposobem badania rzeczywistości i jako takie wymaga opracowania specyficznych metod badania zasadności modelu. Warto rozróżnić pojęcie *modelowania*, jako procesu odnoszącego się głównie do określenia zależności pomiędzy systemami rzeczywistymi i modelami, oraz *symulacji* związanej głównie z określeniem zależności pomiędzy modelem i jego implementacją komputerową. Przez *komputerową symulację systemów* rozumiemy specyficzną technikę rozwiązywania problemów, polegającą na śledzeniu w czasie zmian zachodzących w dynamicznym modelu systemu przy wykorzystaniu techniki komputerowej.

Model komputerowy jest to w istocie zbiór instrukcji zapisanych w odpowiednim języku komputerowym, służący wygenerowania danych o relacjach, wartościach zmiennych i zjawiskach zidentyfikowanych w opisywanej rzeczywistości. Choć nie zawsze widoczny, niejako w tle tak rozumianego modelu komputerowego istnieje zawsze model matematyczny rozumiany jako sformalizowany opis badanego systemu z uwzględnieniem celu tworzenia tego opisu.

Przez eksperyment, podobnie jak Gordon (1974, s. 317), rozumieć będziemy sprawdzenie zachowania się modelu komputerowego dla jednego zbioru warunków (rozumianych najczęściej jako jeden zestaw wartości parametrów modelu). Natomiast badanie symulacyjne modelu jest to zbór wszystkich eksperymentów wykonanych dla realizacji określonego celu.

Model rzeczywistego procesu (układu, systemu, obiektu, itp.) to zawsze uproszczony, przybliżony i wyidealizowany sposób przedstawienia (reprezentacji) wybranego procesu rzeczywistego, ukierunkowany na realizację określonych zastosowań, z uwzględnieniem wszelkich ograniczeń związanymi z samym procesem (np. dotyczącymi ograniczonych możliwości obserwacji procesu rzeczywistego).

Dokładność, czy też zasadność modelu, rozumiana jako jego wierność w stosunku do rzeczywistości, w decydujący sposób wpływa na wiarygodność wniosków uzyskanych z badań modelu a dotyczących bezpośrednio badanego procesu. Budowa adekwatnego modelu wiąże się zawsze z koniecznością pogodzenia wymogów prostoty modelu i istotności uwzględnienia w modelu maksymalnie dużej liczby cech procesu rzeczywistego.

Dopasowanie modelu do rzeczywistości wiąże się nieodłącznie z koniecznością identyfikacji, przez którą rozumiemy rozpoznanie lub też ustalenie charakterystycznych cech lub parametrów procesu, które są niezbędne do budowy modelu matematycznego na podstawie sygnałów wejściowych i wyjściowych oraz znajomości praw i mechanizmów kierujących zachowaniem się procesu rzeczywistego, dla zadanej klasy procesów i zadanej klasy sygnałów wejściowych (testujących).

Najogólniej rzecz ujmując chodzi o zbudowanie takiego wskaźnika jakości, najlepiej w formie skalaru, dającego podstawy do oceny różnicy zachowania się modelu w stosunku do procesu rzeczywistego. Schematycznie proces dopasowywania modelu do rzeczywistości przedstawić można jak na Rys. 1. Badanie zasadności modelu (realizowana po dokonaniu identyfikacji) rozumiana jest jako sprawdzenie czy rozwiązanie zadania identyfikacji doprowadziło do zbudowania sensownego modelu, mogącego stanowić podstawę do przeprowadzenia na nim badań, których wyniki stanowiłyby bazę do wnioskowania o rzeczywistym procesie. Tak określone podejście wydaje się być jasnym i precyzyjnym, w rzeczywistości jednak budujący modele napotykają wiele problemów przy określaniu dokładności modelu i podjęciu ostatecznej decyzji ustalających fakt zasadności modelu. Stosunkowo niewielkie problemy występują w sytuacji kiedy w badanym systemie daje się wyróżnić wyraźne trendy rozwojowe i one to są dominującym składnikiem zmiennych opisujących zachowanie się systemu, natomiast składowe stochastyczne są jedynie niewiele znaczącym 'szumem', który niejako może być traktowany jako tło dla dominującego trendu (patrz Rys. 2). W takiej sytuacji możliwe jest zastosowanie znanych metod identyfikacji parametrów modelu (np. Bubnicki, 1974, Mańczak, 1971) i w ten sposób osiągnąć maksymalnie dobre dopasowanie przebiegu otrzymanego w symulacji do przebiegu rzeczywistego (jako kryterium można przyjąć np. często stosowane kryterium średniokwadratowe, czy też współczynnik podobieństwa Theila).

W symulacji spotykamy się często z sytuacjami kiedy zmianach wartości zmiennych opisujących system dominują składniki stochastyczne, i to zarówno w przebiegach obserwowanych w procesie rzeczywistym jak i w przebiegach procesu symulowanego. Taką typową sytuację przedstawia Rys. 3. Oba procesy, rzeczywisty i symulowany, są z reguły procesami stochastycznymi. Powstaje pytanie, w jaki sposób określić czy oba te procesy są takie same (w sensie podobieństwa rozkładów prawdopodobieństwa)? I pytanie zasadnicze, jak dobrać parametry modelu tak aby na pewnym etapie badań symulacyjnych możliwe było stwierdzenie, że oba procesy są takie same, albo też wykazują daleko idące podobieństwo. Mniejszy problem powstawałby gdyby zmienne opisujące system mogłyby być traktowane jako procesy stochastyczne stacjonarne, lub też ergodyczne. Dla tego typu procesów opracowano wiele metod pozwalających na określenie czy oba procesy są takie same, a niektóre z nich nawet na określenie stopnia podobieństwa procesów stochastycznych. Niestety większość symulowanych procesów to procesy niestacjonarne, dla których niezmiernie trudno jest opracować efektywne metody oceny zasadności modelu. O pewnych przybliżonych metodach (jak np. określanie przedziałów ufności i metodach opartych na analizie spektralnej) piszemy w dalszej części opracowania.

Zatem w jaki sposób podejść do problemu planowania badań symulacyjnych i określania zasadności modelu? Tradycyjnie (np. Conway, 1963) zaplanowanie badania symulacyjnego podzielić można na dwa etapy:

- planowanie strategiczne, tzn. zaprojektowanie zbioru eksperymentów (określenie miary oceny systemu i sposobów przeprowadzania testów istotności tych miar), oraz
- planowanie taktyczne, rozumiane jako określenie sposobu prowadzenia eksperymentów (określenie sposobów dokonywania pomiarów dla każdego przebiegu i liczbę przebiegów dla każdego eksperymentu).

Niewiele zrobiono dotychczas w zakresie metod planowania strategicznego specjalnie opracowanych dla projektowania eksperymentów symulacyjnych. Większość metod klasycznych jest nieodpowiednia dla tego celu, ponieważ przyjmuje się w nich istnienie niezależnych danych, na których opiera się testy statystyczne, a dane z symulacji często nie spełniają tego warunku. Jak wspomnieliśmy, niektóre z nich, jak np. metoda analizy sekwencyjnej, czy też analizy spektralnej mogą się okazać się użyteczne w zastosowaniu do symulacji, jednakże zawsze wymagać będą ich dostosowania do specyfiki badań symulacyjnych.

Problem oceny zasadności modelu jest jednym z bardziej niejasnych, nie dających się sformalizować i ukonkretnić problemów symulacji komputerowej. Wbrew jednak stosunkowo niewielkiej liczbie publikacji poświęconych temu problemowi jest on jednym z ważniejszych elementów w procesie budowy modelu, a w szczególności symulacyjnego modelu komputerowego. Praktycznie nie sformułowano generalnych wskazówek jak oceniać model symulacyjny, jego zasadność oraz jak dobrać strukturę modelu i wartości jego parametrów by możliwe było stwierdzenie że model dostatecznie dobrze opisuje rzeczywistość. Praktycznie każdy z budujących modele symulacyjne opracowuje dla siebie na swój własny użytek sposoby oceny modelu i doboru jego parametrów. W dużym stopniu kierując się swoją intuicją i swoim doświadczeniem.

Zatem każdy z budujących model symulacyjny powinien określić zbiór kryteriów, w oparciu o które możliwe byłoby rozstrzygnięcie czy model jest „prawdziwy” czy nie. Świadomie napisaliśmy słowo „prawdziwy” w cudzysłowie, jako, że nie jest możliwe jednoznaczne określenie czy model jest dobry czy zły, zamiast tego lepiej jest mówić o stopniu zgodności modelu z rzeczywistością, pozwala nam to na ulepszanie modelu w tym sensie, że w miarę jego udoskonalania stopień zgodności z rzeczywistością rośnie. Jest to postulat zgodny z propozycją Poincare (1925, 1935) i Duhema (1883, 1894, 1914), a także Poppera (1977). Jak wyraził to Carnap (Carnap, 1963) „... zamiast o zasadności modelu możemy mówić o stopniowym wzroście zaufania do modelu.” A zaufanie to rośnie w miarę tego jak eksperymentujemy z nim i nie napotykamy na negatywne wyniki jego konfrontacji z rzeczywistością.

Filozofowie nauki lubią kategorie czyste, a takimi czystymi stanowiskami filozoficznymi odnoszącymi się do badania zasadności modelu są racjonalizm (model jest prawdziwy jeśli wysnuty jest z pewnej liczby syntetycznych sądów, których prawdziwość nie może być podważona), empiryzm (tylko doświadczenie i obserwacja są ostatecznym sędzią w ocenie prawdziwości modelu) oraz stanowisko pozytywistyczne (model jest na tyle prawdziwy na ile jest zdolny przewidywać przyszłe zachowania się systemu rzeczywistego). Niestety żadne z tych trzech metodologicznych stanowisk nie może być zastosowane w swej czystej formie przy podejściu symulacyjnym. W istocie do oceny modelu symulacyjnego (i prawdopodobnie w każdym innym modelu) musimy uwzględnić wszystkie trzy stanowiska, jakkolwiek filozofowie oponowaliby przeciwko takiemu podejściu, bo - jak wielu z nich twierdziłoby - „nie można pożenić ognia z wodą”.

Podejście symulacyjne stosowane jest najczęściej w sytuacjach gdy obserwacja procesów zachodzących w rzeczywistości jest bardzo kosztowna a obserwowany system jest na tyle złożonym, że niemożliwe jest opisanie go stosunkowo prostym układem równań dających się rozwiązać analitycznie. Już te dwie cechy wymuszają pewne specyficzne podejście do oceny zasadności modelu.

Ważnym etapem w ustalaniu zasadności modelu jest upewnienie się, czy zmienne wejściowe do modelu symulacyjnego (jak np. rozkłady prawdopodobieństwa wykorzystywane do generacji strumieni zgłoszeń) dobrze odzwierciedlają rzeczywiste dane wejściowe. I tutaj pomocne mogą być różnego rodzaju testy statystyczne pozwalające określić czy przyjęte przez nas rozkłady prawdopodobieństwa odpowiadają ‘rozkładom

rzeczywistym'. Często przy budowie modelu symulacyjnego, niejako bezwiednie, z rozpędu, przyjmujemy podobne rozkłady prawdopodobieństwa dla różnych elementów danej klasy, tymczasem przy bliższym przyjrzeniu się systemowi rzeczywistemu może się okazać, że należy przyjąć różne rozkłady dla różnych elementów - np. dla dwóch urzędników wykonujących podobną pracę przy obsłudze klientów w banku, czy dwóch lekarzy pracujących w tej samej przychodni. Nie zawsze takie rozróżnienie może być uznane za konieczne. Dobrym testem jest zbadanie wrażliwości modelu na niewielką zmianę parametru charakteryzującego zmienną wejściową (np. charakteryzującego rozkład prawdopodobieństwa), Jeśli taka niewielka zmiana spowoduje istotną zmianę wyników otrzymanych podczas symulacji to należy uznać to za sugestię bliższego przyjrzenia się temu elementowi modelu i dokładniejszego dopasowania go do danych otrzymanych z obserwacji systemu rzeczywistego.

Zatem jednym z postulatów jest by na pewnym etapie budowy modelu starać się go zdezagregować i próbować takie lokalne modele dopasować do obserwacji rzeczywistych. Na kolejnych etapach prac nad 'całościowym' modelem badanego systemu, kiedy takie 'lokalne' modele połączone zostały w jedną całość, należałoby się przyjrzeć miejscom szczególnie wrażliwym, tzn. takim gdzie niewielka zmiana powoduje radykalną zmianę charakteru zachowania się modelu. Takie elementy szczególnie wrażliwe należałoby szczególnie dokładnie opracować i dopasować do rzeczywistości.

Przed przystąpieniem do budowy modelu całościowego, jak również w trakcie prac nad nim, w miarę jak zdobywamy dokładniejsze i pełniejsze dane o samym procesie rzeczywistym, który modelujemy, budujący model powinien (na podstawie całego swojego dotychczasowego doświadczenia) wybrać pewną liczbę postulatów, służących mu jako podstawa do dalszej pracy. Wybór tych postulatów musi być dokonany na podstawie sądów *a priori*. Nie będąc dogmatykiem, badacz może zmieniać zbiór tych „*a priori* postulatów” w trakcie pracy nad modelem. Jest to często proces interakcyjny, z wieloma powrotami (iteracjami) do wcześniejszych etapów prac nad modelem. Propozycję formułowania takich postulatów Nicolas Kaldor (1961) ujął w formę tzw. wyidealizowanych faktów (*stylised facts*). Kaldor czynił często uwagi dotyczące *abstraction, tendencies* and *stylised facts*. Fragment poniżej jest typowym dla jego poglądów:

Jakakolwiek teoria musi bazować na różnego rodzaju abstrakcjach, jednakże rodzaj czynionych abstrakcji nie może być dokonywany w próżni: musi on odpowiadać charakterystycznym cechom procesu gospodarczego jakiego doświadczamy. Dlatego też teoretyk, kiedy podejmuje decyzję o wyborze szczególnego podejścia badawczego, powinien zacząć od zebrania wszelkiego rodzaju faktów odnoszących się do badanego procesu, które to fakty uznaje on za ważne i charakterystyczne dla tegoż procesu. Ponieważ fakty, w takiej formie jak przedstawiane są przez ludzi zajmujących się zbieraniem danych statystycznych, są zawsze obiektem przeróżnego rodzaju obróbek i wyborów, dlatego też najczęściej nie poddają się one podsumowaniu w formie jasnych, precyzyjnych sformułowań. Dlatego też teoretyk, w mojej opinii, powinien mieć swobodę w formułowaniu 'wyidealizowanych faktów' bez zbytniego dbania o dużą ich dokładność z danymi historycznymi, jak również też o ich pełną liczbę, tak by całkowicie charakteryzowały badany proces.” (Kaldor, 1961)

W procesie rozwoju gospodarczego takimi typowymi zachowaniami i zjawiskami (*stylised facts*) są np. (1) różnorodność cen produktów oferowanych do sprzedaży i brak jednej, wspólnej wszystkim podmiotom gospodarczym ceny równowagi, (2) wzrost średniej ceny produktów i średniego zysku firm funkcjonujących na danym rynku ze wzrostem koncentracji rynku (konkurencja doskonała, oligopol, duopol, monopol), (3) pojawienie się innowacji powoduje chwilową monopolistyczną pozycję firmy wprowadzającej tę innowację na rynek, dzięki czemu firma ta uzyskuje dodatkowy zysk, który z czasem maleje w miarę jak

konkurenci naśladowują firmę pionierską, (4) ścisła zależność pomiędzy wielkością ekonomii skali a stopniem koncentracji rynku (im większa ekonomia skali tym większa koncentracja rynku), (5) porównując gospodarki różnych krajów, różnorodność produktywności kapitału (tzn. stosunku Produkcji/Kapitał) jest znacznie mniejsza aniżeli różnorodność produktywności pracy ludzkiej (tzn. stosunku Produkcja/Praca)¹, (6) na wielu rynkach obserwujemy charakterystyczny rozkład wielkości firm funkcjonujących na rynku (zbliżony do rozkładu Pareto) o prawie stabilnym nachyleniu w przeciągu długiego okresu czasu, (7) proces rozwoju gospodarczego jest niepowtarzalnym procesem historycznym w którym istotną rolę odgrywają zdarzenia przypadkowe (ścieżki rozwojowe - *path-dependence*) oraz tzw. zjawiska kumulowania się zmian (*cumulative causation*).

Wybór takich 'wyidealizowanych faktów' zależy będzie od stopnia zaufania jakim je darzy badacz. A stopień ten zależy będzie od tego na ile dany postulat jest ogólny i stosowany był wielokrotnie z pozytywnym skutkiem przez innych, oraz na ile dany fakt potwierdzony może być przy zastosowaniu pewnych wypróbowanych narzędzi analizy, jak np. testów statystycznych.

Ocena modelu często wiąże się z wyborem modelu ze zbioru alternatywnych modeli opisujących daną rzeczywistość. Nigdy nie jest tak, że mamy model idealnie pasujący do rzeczywistości, bo jak powiedział William Edwards Deming: "Wszystkie modele są złe tylko niektóre są użyteczne". Oceniając jakikolwiek model nie oceniamy każdego pojedynczego założenia, formuły czy konkluzji wynikającej z zastosowania modelu. Z reguły budujemy swego rodzaju ogólny wskaźnik oceny danego modelu i ze zbioru alternatyw wybieramy ten o największym wskaźniku globalnym. By ułatwić sobie proces oceny modelu często posługujemy się sub-kryteriami które z pewnymi wagami wpływają na ten jeden globalny wskaźnik (mamy zatem tutaj typowy problem oceny wielokryterialnej). Wydaje się, że takimi najczęściej używanymi sub-kryteriami są:

1. *poprawność* – wyniki modelu powinny być możliwie najbliższe wyników obserwacji lub eksperymentów.
2. *spójność* – model powinien być nie tylko spójny wewnętrznie ale także zgodny (spójny) z ogólnie akceptowanymi teoriami, które odnoszą się do badanego problemu.
3. *uniwersalność* – konsekwencje danego modelu powinny odnosić się do szerszej grupy zjawisk niż ten pojedynczy problem, który był inspiracją budowy danego modelu.
4. *prostota* – model powinien przyczyniać się do tworzenia porządku w obrębie badanej klasy zjawisk; pewne subiektywne odczucia harmonii i piękna odgrywają często istotną rolę w ocenie stopnia prostoty modelu. Albert Einstein zwykł mawiać, że „Najlepsze wyjaśnienie jest to najprostsze ... ale nie prostsze”. Alfred North Whitehead proponował natomiast: "Szukaj prostoty ... a potem jej niedowierzaj!".
5. *plodność* – model powinien rzucić nowe światło na dobrze znane zjawiska, powinien stymulować nowe odkrycia.
6. *użyteczność* – to praktyczne kryterium dominuje w naukach stosowanych.

Podstawowym elementem oceny zasadności modelu jest zwykle konfrontacja wyników otrzymanych z symulacji modelu z obserwacjami systemu rzeczywistego - historycznymi danymi statystycznymi, czy też celowo zbieranymi danymi w ramach projektu budowy modelu wybranego systemu rzeczywistego.

¹ Kaldor (1985, s. 64) pisze, o „jednym z najbardziej udokumentowanych 'wyidealizowanych faktów', mianowicie podczas gdy stosunek Kapitał/Praca rośnie bardziej lub mniej proporcjonalnie do produktywności, i jest największy w państwach najbogatszych i najmniejszy w najbiedniejszych, to stosunek Kapitał/Produkcja jest prawie taki sam w krajach bogatych jak i biednych - nie jest wcale większy w Ameryce ... aniżeli w Indiach.”

Tak jak wspomnieliśmy, w zasadzie nie ma problemów z systemami wykazującymi duże regularności, dającymi się opisać za pomocą modeli deterministycznych (lub quasi-deterministycznych). W takich sytuacjach możliwy jest wybór pewnej klasy modeli mogących dostatecznie dobrze opisywać daną rzeczywistość i następnie wykorzystanie wielu dostępnych metod identyfikacji nieznanymi parametrów modelu, tak by opis był najlepszy w sensie przyjętego kryterium jakości podobieństwa zachowania się modelu do procesów obserwowanych w rzeczywistości (np. Bubnicki, 1974; Mańczak, 1971). Podobne podejście do identyfikacji możemy zastosować w sytuacji gdy znamy rozkłady prawdopodobieństwa z dokładnością do parametrów charakteryzujących te rozkłady. W tej sytuacji też można zastosować znane probabilistyczne metody identyfikacji parametrów tychże rozkładów.

Niestety wiele modeli symulacyjnych, budowanych często poprzez szukanie różnego rodzaju analogii i które ze swej natury wykazują daleko idące podobieństwa swej struktury i mechanizmów rozwoju do struktury i mechanizmów systemów rzeczywistych, uniemożliwiają stosowanie tego typu podejść do identyfikacji parametrów modelu. Problemem samym w sobie jest określenie miary podobieństwa wyników otrzymanych z symulacji modelu z danymi rzeczywistymi. Trudności tego rodzaju przyczyniają się do istnienia dużej różnorodności propozycji poradzenia sobie z tym problemem.

Jeśli jednym z celów budowy modelu jest, lub może być, możliwość przewidywania przy jego wykorzystaniu, to jednym z testów oceny zasadności modelu byłby test na zdolność przewidywania korzystając z tzw. predykcji retrospektywnej.

Dobrze jest sformułować operacyjne kryteria oceny zgodności zachowania się modelu symulacyjnego z danymi historycznymi, lub aktualnymi zmianami obserwowanymi w systemie rzeczywistym. Jest to pole niemalże dziewicze, wybór takich kryteriów jest prawie zawsze sądem subiektywnym samego badacza.

Przeoglądając literaturę dotyczącą badania stopnia zgodności modeli symulacyjnych z rzeczywistością można określić pewne typowe i najczęściej stosowane metody statystyczne w poradzeniu sobie z tym problemem, są to:

- **analiza wariancji**, tzn. testowanie hipotez, że średnia wartość (lub ogólnie jakaś charakterystyka szeregów czasowych otrzymanych w symulacji - traktowanych jako proces stochastyczny) jest równa średniej wartości (lub danej charakterystyce) szeregu otrzymanego w wyniku obserwacji procesu rzeczywistego. Wadą tej metody jest to, że w większości testów statystycznych zakłada się normalność rozkładu i niezależność stochastyczną. Stosowanie tej metody powinno być bardzo ostrożne przy podejściu symulacyjnym, jako, że bardzo często rozkłady prawdopodobieństwa zmiennych symulowanych odbiegają od rozkładu normalnego (i tu sam badacz powinien odpowiedzieć czy rozkład zmiennej symulowanej dostatecznie zbliża się do rozkładu normalnego, np. na podstawie twierdzenia granicznego), oraz czy spełniony jest postulat niezależności statystycznej (w symulacji najczęściej to co ma wydarzyć się w przyszłości zależy od stanu aktualnego procesu).

- **testy zgodności** (lub inaczej testy nieparametryczne) pozwalające określić czy hipoteza o zgodności rozkładów empirycznych jest zgodna z przyjętym rozkładem prawdopodobieństwa. Przy odpowiedniej modyfikacji tych testów, można pokusić się o przetestowanie hipotezy, że wyniki generowane podczas symulacji mogą być traktowane jako realizacja procesu stochastycznego o rozkładach prawdopodobieństwa wyliczanych na podstawie danych historycznych. Tutaj najczęściej stosowanymi testami są: test χ^2 , test zgodności lparty na twierdzeniu Kołmogorowa-Smirnowa, oraz testy Walda-Wolfowitza i Wilcoxon-Manna-Whitneya (Fisch, 1969, s. 467-473). Test Kołmogorowa-Smirnowa. pozwala ocenić stopień zgodności rozkładów generowanych na podstawie wyników symulacji i na podstawie danych rzeczywistych w ten sposób, że nie gubi informacji przez łączenie się sobą różnych kategorii, tak jak to jest czynione w teście χ^2 .

- **analiza regresji.** Metoda ta pozwala na ocenę regresji pomiędzy szeregami rzeczywistymi i generowanymi oraz testowanie, czy otrzymane równania regresji mają odcinki w których różnice pomiędzy danymi generowanymi w symulacji i rzeczywistymi można uznać za bliskie zeru.

- **analiza czynnikowa.** Zastosowanie analizy czynnikowej do zbioru wyników generowanych przez model i do zbioru danych uzyskanych z obserwacji, i porównanie czy wyniki analizy czynnikowej otrzymane w obu sytuacjach różnią się od siebie w istotny sposób - podejście to zaproponowane zostało przez Cohen'a i Cyert'a (1961).

- **współczynnik podobieństwa** (niekiedy nazywany współczynnikiem nierówności) Theila (Theil, 1961). Współczynnik podobieństwa Theila U mierzy stopień możliwości predykcyjnych (retrospektywnych) P_i i obserwowanych A_i .

$$U = \frac{\sqrt{\frac{1}{n} \sum (P_i - A_i)^2}}{\sqrt{\frac{1}{n} \sum P_i^2} + \sqrt{\frac{1}{n} \sum A_i^2}}$$

Współczynnik U przyjmuje wartości od zera do jedności, równa się zero kiedy przewidywanie jest doskonałe, natomiast równy jest 1 jeśli przewidywanie jest złe. W późniejszej swojej pracy Theil (1966) podaje inną definicję współczynnika nierówności, natomiast w pracy (Theil, 1964) przedstawił uogólnienie tej metody na przypadki wielu zmiennych.

- **określenie przedziału ufności,** dzięki czemu możliwe jest określenie stopnia podobieństwa pomiędzy zmianami symulowanymi a zmianami w systemie rzeczywistym.

- **analiza spektralna.** Zmiany wartości zmiennych otrzymywanych w trakcie symulacji (podobnie jak dane z obserwacji procesów rzeczywistych) wykazują duży stopień autokorelacji. Fakt ten powoduje niedoszacowanie wariancji z próby i nieskuteczność predykcji w klasycznych testach statystycznych np. podczas estymacji. W stosunku do analizy wariancji, estymacji średniej czy wariancji, analiza spektralna niesie w sobie znacznie więcej informacji o badanym procesie. Analiza spektralna, dzięki możliwości porównania estymowanych widm wyników symulacji z podobnym estymowanym widmem procesu rzeczywistego pozwala na wyciągnięcie stosunkowo obiektywnych i godnych zaufania wniosków o stopniu podobieństwa procesu symulowanego z procesem rzeczywistym.

Te dwa ostatnie podejścia (określanie przedziałów ufności i analiza spektralna) wydają się być najbardziej efektywnymi i najbardziej obiecującymi podejściami do badania stopnia podobieństwa modelu symulacyjnego do rzeczywistości.

Liczba testów statystycznych jakie proponuje się do testowania zgodności danych modelowych i rzeczywistych jest stosunkowo duża (patrz np. Shannon, 1975, s. 208, ale też większość książek ze statystyki matematycznej, np. Fisz, 1969). Metody te opracowano przy założeniu niezależności pomiędzy błędami statystycznymi, równości wariancji i rozkładu normalnego. Procesy stochastyczne leżące w sferze zainteresowania teoretyków to z reguły procesy stacjonarne i nie wykazującego skorelowania. Założenie to jest oczywiście nieprawdziwe w procesach leżących w sferze zainteresowań symulacji (z reguły niestacjonarne i o wysokim stopniu autokorelacji).

Nie zawsze jednak musimy badać zasadność modelu stosując wyrafinowane metody statystyki. Niekiedy, dzięki wykorzystaniu specyfiki badanego systemu, jak również zależnie od celu budowy modelu, możliwe jest opracowanie innych metod badania zasadności, uznanych przez użytkownika modelu za wystarczające. Law i Kelton (1982, s. 341) opisują interesujący przypadek zastosowania *Testu Turinga* do określania stopnia podobieństwa wyników otrzymanych z symulacji modelu i wyników rzeczywistych. Ludzie znający system rzeczywisty 'od podszewki' pytani są czy dane im przedstawione do oceny odzwierciedlają to

co znają z rzeczywistości. Przy czym nie wiedzą oni czy dane te są wynikami obserwacji systemu rzeczywistego, czy pochodzą z badania modelu. Eksperti nie oceniają wyników w kategoriach ‘dobre’ czy ‘złe’, ale podają dosyć szczegółową analizę przedstawianych im danych i wskazują co jest dobre a co im się nie podoba w dostarczonych im danych. Takie dosyć szczegółowe oceny wystawiane przez ekspertów z jednej strony służą modelującym do poprawiania samego modelu, ale jednocześnie służą do oceny stopnia zgodności modelu z rzeczywistością. Idea ta zastosowana została w modelu ISEM opisującego System Zatrudnienia sił powietrznych Stanów Zjednoczonych (*Air Force Manpower and Personnel System*), który to zbudowany został w celu określanie alternatywnych polityk zatrudnienia w siłach powietrznych. Wyniki modelu pokazywane były analitykom i decydentom z *Air Force*. Procedura ‘Testu Turinga’ powtarzana była wielokrotnie, w efekcie otrzymano zadowolający model zatrudnienia w *Air Force*.

Analiza spektralna

Wyniki symulacji są zazwyczaj silnie skorelowane a procesy, symulowany i rzeczywisty, nie są procesami ergodycznymi. Liczenie zatem średniej i wariancji względem zmiennej czasowej mija się z celem. Obejściem często stosowanym w takiej sytuacji jest obliczanie średniej i wariancji nie względem czasu i testowanie zgodności średniej z wyników symulacji i średniej z danych rzeczywistych, ale wielokrotne powtarzanie eksperymentu symulacyjnego i liczenie średniej dla każdej chwili w okresie symulacji i porównywanie tego wyniku z przebiegiem rzeczywistym.² Oczywiście, że w tej sytuacji wydłuża się czas badania symulacyjnego, rosną też koszty samego badania. Analiza spektralna, choć też nie jest wolna od wad, pozwala porównać przebiegi symulacyjny z rzeczywistym, bez wielokrotnego powtarzania eksperymentów. Analiza spektralna pozwala badać średni poziom działania systemu, odchylenia od tego średniego poziomu oraz długość ich trwania - a o to najczęściej chodzi w badaniach symulacyjnych. Umożliwia też budowanie pasm ufności i testowanie hipotez przy porównywaniu wyników symulacji w przypadku analizy alternatywnych polityk, czy też strategii gospodarczych. Rysunek 4 przedstawia typowy wykres gęstości mocy otrzymywany przy analizowaniu danych o rozwoju gospodarczym. Estymowany wykres, otrzymany po odfiltrowaniu trendu liniowego, wraz z 95% przedziałami ufności również pokazane są na rysunku. Wadą analizy spektralnej jest to, że musimy mieć stosunkowo długie szeregi czasowe, rzędu kilkuset a nawet kilku tysięcy danych, choć niekiedy można analizować szeregi składające się ok. 100 elementów.

Niech $\{x(t), t \in T\}$ oznacza proces stochastyczny z którego realizacją jest szereg czasowy $\{x(t), t = 1, 2, 3, \dots, n\}$. Generowany proces może być wstępnie opisany za pomocą momentów pierwszego i drugiego rzędu, wartości oczekiwanej μ_t , wariancji σ_t , oraz autokowariancji $\gamma(t, s)$.

$$\mu_t = E[x(t)]$$

$$\sigma_t^2 = E[(x(t) - \mu_t)^2]$$

$$\gamma(t, u) = E[(x(t) - \mu_t)(x(t+u) - \mu_{t+u})]$$

Dla procesów stacjonarnych momenty pierwszego i drugiego rzędu nie są funkcjami czasu, a autokowariancja jest tylko funkcją opóźnienia. Przekształcenie Fouriera funkcji autokowariancji (zwane również spektrum, gęstością widmową, widmową gęstością mocy) ma dla procesów stacjonarnych postać

² Często jednak działania takie wydają się być słabo uzasadnionymi (jak np. w sytuacji szerszej klasy procesów ewolucyjnych, kiedy to każda realizacja procesu ewolucyjnego jest procesem niepowtarzanym).

$$f(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} e^{-iu\lambda} \gamma(u)$$

Wczesne podejście do analizy spektralnej bazowało na dyskretnym przekształceniu Fouriera

$$w(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^T x(t) e^{-it\lambda}$$

Dla zbudowania estymatora spektrum $f(\lambda)$ w całym przedziale częstości $[0, \pi]$ $w(\lambda)$ jest liczone dla częstości podstawowych $\lambda = \lambda_j = 2\pi j/T, j = 1, 2, 3, \dots, [T/2]$. Bezpośrednie liczenie tej statystyki wymaga wykonania liczby operacji proporcjonalnych do T^2 . Współczesne algorytmy, oparte o algorytm tzw. Szybkiej Transformaty Fouriera (FFT) pozwalają liczbę tę zmniejszyć do rzędu $T \log_2 T$. Periodogramem albo próbkową gęstością widmową nazywa się wielkość równą

$$I(\lambda) = |w(\lambda)|^2$$

Jeżeli $f(\lambda)$ jest ciągłą to $I(\lambda)$ jest asymptotycznie zbieżnym estymatorem $f(\lambda)$.

Nie będziemy tutaj omawiać metod estymacji gęstości widmowej. Współcześnie dobre algorytmy liczenia tych estymatorów zawarte są w każdym pakiecie obliczeń statystycznych. Z reguły estymatory te opracowane są dla procesów stacjonarnych, co znów powoduje ograniczonosc stosowania tego podejścia do badań symulacyjnych i nakazuje daleko idącą ostrożność w stosowaniu analizy spektralnej w tego typu badaniach.

Pierwsze próby liczenia gęstości widmowych mocy dla procesów niestacjonarnych zostały pojęte stosunkowo niedawno (np. Trevino, 1982), jednakże nadal nie wyszły one poza fazę pewnych propozycji teoretycznych.

Osobnym problemem jest określenie podobieństwa widmowych gęstości mocy otrzymanego z badań symulacyjnych do analogicznego widma procesu rzeczywistego. Niewielu autorów pojęło się teoretycznego zbadania tego problemu (wyjątkiem jest tutaj Anderson, 1993) dlatego też praktycznie każdy z budujących modele symulacyjne i próbujący odpowiedzieć sobie na pytanie o stopień odpowiedniości modelu do rzeczywistości przy pomocy analizy spektralnej stara się sam opracować praktyczne testy pozwalające mu na ocenę stopnia podobieństwa tych dwóch widm.

Przykład próby zastosowania analizy spektralnej i wyidealizowanych faktów do określenia zasadności modelu przedstawili Silverberg i Verpsagen (1995) w ewolucyjnym modelu opisującym dynamikę rozwoju gospodarczego i pokazującym proces konwergencji gospodarek siedmiu krajów OECD (Belgii, Niemiec, Francji, Wielkiej Brytanii, Włoch, Szwecji i USA). Na Rys. 5a pokazano widmo gęstości mocy wariancji dochodu narodowego na osobę w obrębie badanych krajów w ostatnich 120 latach. Autorzy uzyskali podobny wykres w swoim modelu w różnych przebiegach symulacyjnych; na Rys. 5b pokazano gęstości mocy tego współczynnika uzyskanego w dwóch przebiegach symulacyjnych. Podobieństwo pomiędzy wynikami symulacyjnymi i rzeczywistymi jest rzeczywiście duże, autorzy jednak bardzo ostrożnie podchodzą do tych wyników, pisząc, że „podobieństwo pomiędzy empirycznymi i modelowymi wyidealizowanymi faktami jest na tyle bliskie, że powinno skłaniać do prowadzenia dalszych badań empirycznych jak i teoretycznych by potwierdzić naszą hipotezę o naturze obserwowanego procesu konwergencji.”

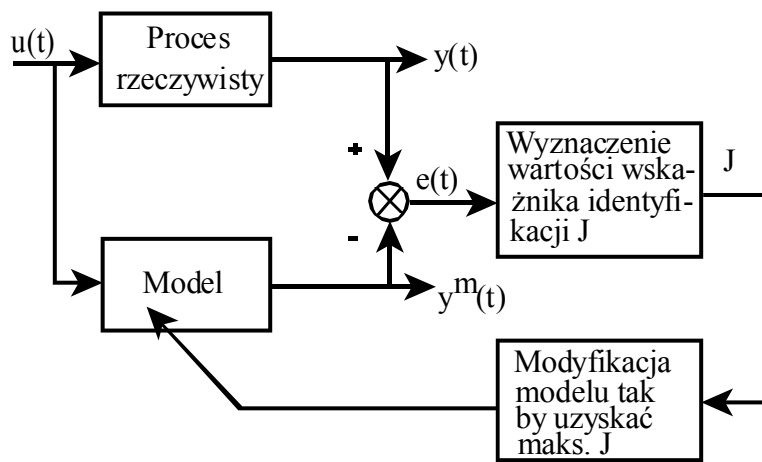
Podsumowanie

Dobrze znane i efektywne metody określenia zasadności modeli w naukach technicznych i w fizyce mają ograniczone możliwości zastosowania w przypadku modelowania procesów społeczno-gospodarczych. Podstawowymi przyczynami niemożności przeniesienia metod

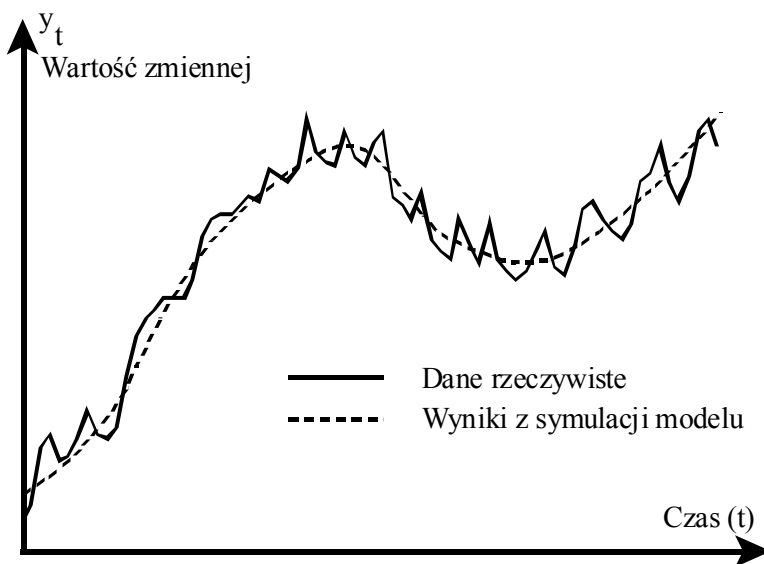
opracowanych w naukach technicznych i w fizyce jest niepowtarzalność tych procesów, duża ich złożoność i obecność niestacjonarnych procesów stochastycznych. Z metod typowo ilościowych obiecuje się być analiza spektralna. Natomiast na obecnym etapie rozwoju badań symulacyjnych efektywnymi metodami oceny dopasowania modeli symulacyjnych do rzeczywistości może być dokonane głównie na poziomie jakościowym poprzez określenie wyidealizowanych faktów obserwowanych w procesach rzeczywistych i próba ich odnalezienie modelu poprzez odpowiednio zaprojektowane eksperymenty symulacyjne.

Literatura

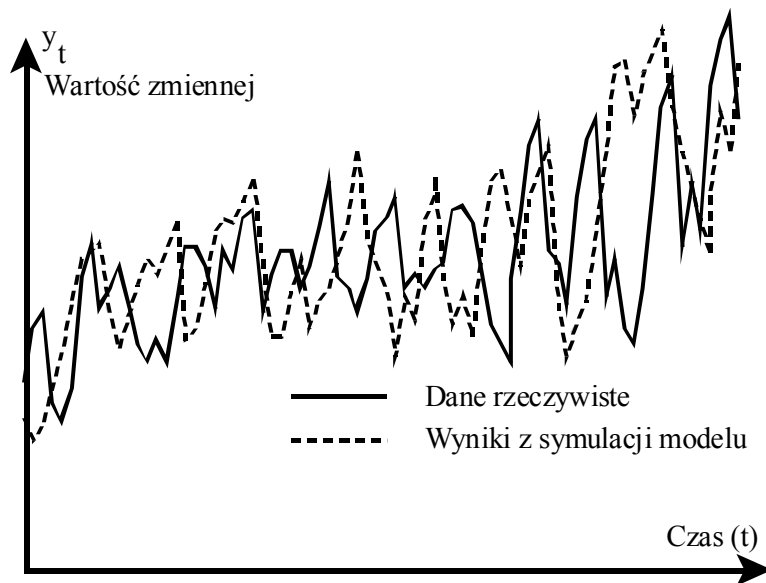
- Anderson T.W. (1993), Goodness of Fit Tests for Spectral Distributions, *The Annals of Statistics*, Vol. 21, No. 2, 830-847.
- Brillinger D.R., Krishnaiah P.R. (eds.), (1983), *Time Series in the Frequency Domain*, Amsterdam: North-Holland.
- Bubnicki Zdzisław (1974), *Identyfikacja obiektów sterowania*, Warszawa: PWN.
- Carnap R. (1963), Testability and Meaning, *Philosophy of Science*, III.
- Conway R.W. (1963), Some Tactical Problems in Digital Simulation, *Management Science*, X, No 1, Oct. pp. 47-61.
- Cohen Kalman J., Cyert Richard M. (1961), Computer Models in Dynamics Economics, *The Quarterly Journal of Economics*, LXXXV, February, ss. 112-127.
- Cyert Richard M (1966), A Description and Evaluation of some Firms Simulations, *Proceedings of the IBM Scientific Computing Symposium on Simulation Models and Gaming*, IBM, White Plains, N.Y.
- Duhem P. (1914), *La theorie physique, son objet, sa structure*, Paris.
- Duhem P. (1883), Physique at metaphysique, *Revue de Questions Scientifique*, serie 2, t.2.
- Duhem P. (1894), Quelques réflexions au sujet de la physique expérimentale, *Revue de Questions Scientifique*, serie 2, t. 3.
- Fisz Marek (1969), *Rachunek prawdopodobieństwa i statystyka matematyczna*, Warszawa: PWN.
- Gordon Geoffrey (1974), *Symulacja systemów*, Warszawa: Wydawnictwa Naukowo-Techniczne
- Kaldor N., 1961, Capital Accumulation and Economic Growth, in Lutz F. (ed.) *The Theory of Capital*, London, Macmillan.
- Kaldor N., 1985, *Economics Without Equilibrium*, University College Cardiff Press.
- Mańczak Kazimierz (1971), *Metody identyfikacji wielowymiarowych obiektów sterowania*, Warszawa: WN-T.
- Otnes Robert K., Enochson Loren, 1978, *Analiza numeryczna szeregów czasowych*, Warszawa: WN-T.
- Poincaré Henri (1925), *La science at l'hypothese*, Paris.
- Poincaré Henri (1935), *La valeur de la science*, Paris.
- Popper Karl (1977), *Logika odkrycia naukowego*, Warszawa.
- Shannon R.E., 1975, *Systems Simulation: The Art and Science*, Englewood Cliffs, N.J.: Prentice-Hall
- Theil H. (1961), *Economic Forecasts and Policy*, Amsterdam: North-Holland.
- Theil H. (1964), *Optimal Decision Rules for Government and Industry*, Chicago: Rand McNally.
- Theil H. (1966), *Applied Economic Forecasting*, Chicago: Rand McNally.
- Trevino George (1982), The Frequency Spectrum of Nonstationary Random Processes, w O.D. Anderson (ed.), *Time Series Analysis: Theory and Practice 2*, Amsterdam: North-Holland Publishing Company.
- Silverberg G. and B. Verspagen, (1995), 'Long Term Cyclical Variations of Catching Up and Falling Behind. An Evolutionary Model', *Journal of Evolutionary Economics*, vol 5, 209-227 (reprinted in Dopfer, K., 1996, *The Global Dimension of Economic Evolution*, Heidelberg: Physica Verlag, 29-48).



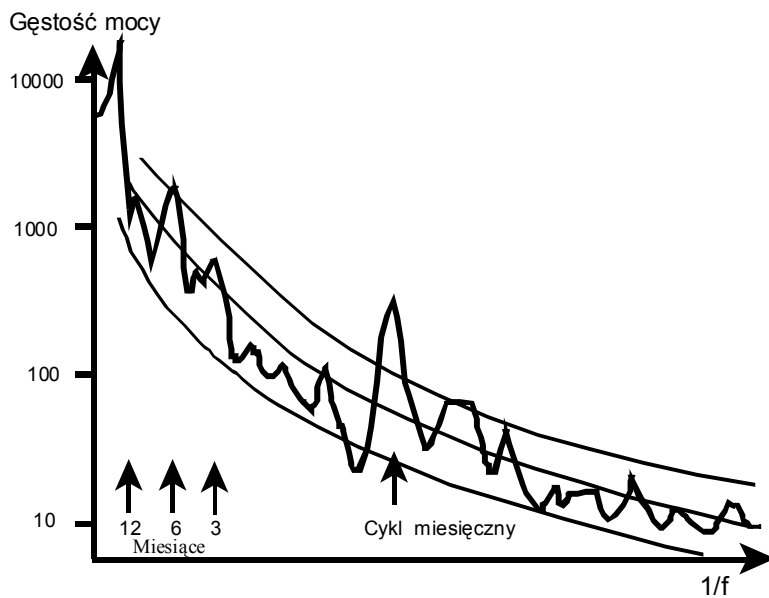
Rysunek 1. Model a rzeczywistość



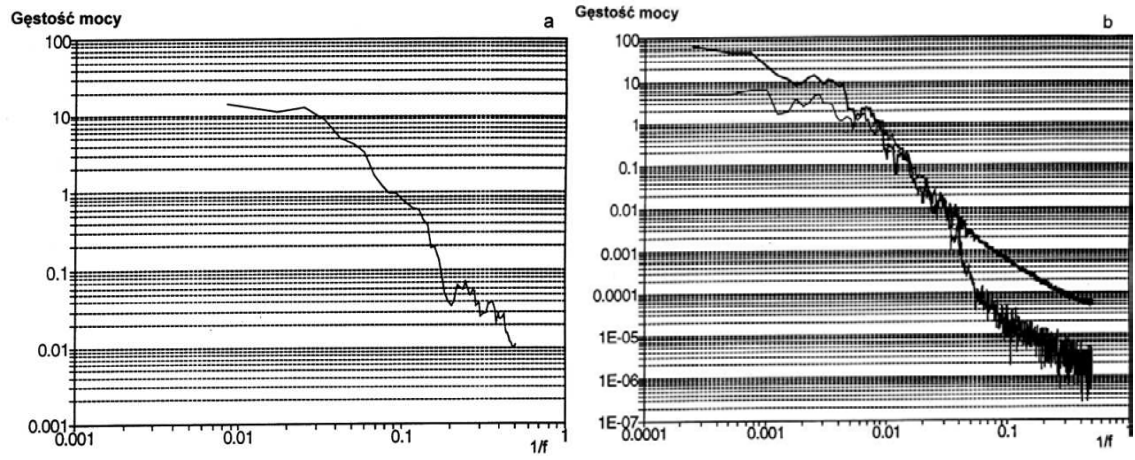
Rysunek 2. Trendy rozwojowe a rzeczywistość



Rysunek 3. Model stochastyczny a rzeczywistość



Rysunek 4. Przykładowy wykres gęstości mocy dla indeksu cen na giełdzie



Rysunek 5. Analiza spektralna wzrostu gospodarczego w wybranych krajach OECD