

Prof. Tomasz Lipniacki,  
Pracownia Modelowania w Biologii i Medycynie  
Instytut Podstawowych Problemów Techniki PAN

Obory, 31 maja 2022

RECENZJA ROZPRAWY DOKTORSKIEJ w *dyscyplinie Matematyka i Informatyka*

***Modelowanie matematyczne rozprzestrzeniania się gruźlicy w populacjach niejednorodnych***

wykonanej przez **mgr Marcina Choińskiego** pod kierunkiem prof. dr hab. Urszuli Forys (promotor) i dr inż. Mariusza Bodziocha (promotor pomocniczy).

**Charakterystyka pracy**

Praca doktorska mgr. Marcina Choińskiego zawiera matematyczną i numeryczną analizę układów dwóch albo czterech równań różniczkowych zwyczajnych zaproponowanych do opisu epidemii gruźlicy. Autor rozpatruje model susceptible – infected – susceptible (SI), który dla populacji jednorodnej opisywany jest dwoma równaniami, a dla populacji niejednorodnej (składającej się z osobników niskiego i dużego ryzyka) czterema równaniami. Istotnym elementem pracy jest dyskretyzacja czasu (dla obu modeli) przy użyciu albo algorytmu Eulera, albo zbliżonego algorytmu (również pierwszego rzędu) pozwalającego na większą zgodność z ODEs dla tego samego kroku czasowego.

Od strony formalnej praca, napisana po polsku, liczy 165 stron i składa się ze wstępu, trzech rozdziałów merytorycznych, podsumowania i czterech „rachunkowych” dodatków. Brakuje wyróżnionego abstraktu, choć przewodnik po rozprawie znajduje się na dwóch ostatnich stronach wstępu. Większość materiału rozprawy zawarta jest w pięciu publikacjach; Referencje 16, 27, 28, 29 i 30 – odpowiednio w *Discrete and Continuous Dynamical Systems - Series B*, IF=1.4; *Mathematical Biosciences and Engineering*, IF=2.1; Materiały konferencyjne; *Communications in Nonlinear Science and Numerical Simulation*, IF=4.3, *Mathematica Applicanda* (bez IF). Współautorami wszystkich prac są promotorzy rozprawy, przy czym w pierwszej z prac mgr. Choiński jest środkowym autorem a w pozostałych czterech pierwszym autorem. Prace ukazały się w latach 2019-2022 i aktualnie w bazie Google Scholar mają łącznie 4 cytowania, ale są to cytowania własne.

**Opinia o pracy**

Głównym celem pracy, w sformułowaniu autora, jest weryfikacja hipotezy stanowiącej, że modele krzyżowe (dwu-populacyjne) są konieczne do prawidłowego opisu epidemii w populacjach niejednorodnych. Moim zdaniem teza jest dość oczywista i intensywnie eksploatowana (choćby przez piszącego tę recenzję w kontekście COVID-19 i media <https://royalsociety.altmetric.com/details/91054089>), ale jej kolejne potwierdzenie ma pewną wartość. Od strony matematycznej i numerycznej praca wydaje mi się poprawna (z dokładnością do pewnej liczby nieprecyzyjnych sformułowań, choć nie podjąłem się przesłania ogromnej liczby przekształceń (podejrzewam, że przeprowadzonych lub zweryfikowanych w „Mathematica”). Ponadto wyniki zostały w większości opublikowane. W kontekście faktu, że praca rozpatrywana jest w dyscyplinie *Matematyka i Informatyka* przychyliam się do dopuszczenia jej do publicznej obrony. **Niemniej mam poważne zastrzeżenia do rozpatrywanych modeli epidemii, dyskretyzacji czasu oraz doboru przykładów do symulacji numerycznych – w dużej części zupełnie oderwanych od epidemiologicznej rzeczywistości.** Poniżej zamieszczam swoje krytyczne uwagi z nieśmiałą nadzieją,

że pomogą doktorantowi, a może także współautorom w dalszych zmaganiach w dziedzinie matematycznego modelowania epidemii.

### *Uwagi krytyczne*

W pracy brakuje odniesienia do faktu, że zdecydowana większość infekcji prątkami gruźlicy jest bezobjawowa. Ten fakt niekoniecznie musi być uwzględniony w zgrubnym modelowaniu, bo osoby bezobjawowe raczej nie zarażają, niemniej możliwy długi okres pomiędzy zarażeniem się, a przejściem do postaci zakaźnej może oznaczać konieczność uwzględnienia fazy „exposed” tj. osób już zarażonych a (jeszcze) niezarażających.

1. **Model 1.1 (oraz model 3.1)** Podstawowym problemem tego modelu jest założenie, że w przypadku braku infekcji wzrost populacji jest proporcjonalny do wielkości populacji ze współczynnikiem proporcjonalności  $A$ . W rezultacie układ (bez infekcji) nie posiada przyciągającego stanu stacjonarnego. Układ taki zupełnie nie nadaje się do analizy stanów stacjonarnych bądź cykli granicznych w wypadku występowania infekcji – a to jest przedmiotem pracy. W analizowanym modelu stabilny przyciągający stan stacjonarny pojawia się dopiero gdy występuje infekcja i związana z nią dodatkowa śmiertelność (Rys. 1.4 b). Jeśli gruźlica może ustabilizować wielkość populacji to dlaczego nie AIDS lub inna potencjalnie śmiertelna choroba zakaźna? Jasne jest, że tego typu model 1.1. nie ma sensu w kontekście pytania o stany stacjonarne.

Wydaje się, że Autor dochodzi sam do podobnego wniosku w Dyskusji na stronie 27, pisząc: „Przyjrzyjmy się modelowi 1.3 (*tożsamy z 1.1*). Stwierdzenie 1.2 wskazuje na zaistnienie przypadku gdy populacja rozrasta się nieograniczenie [...] Z tego powodu uznajemy model 1.3 za niepoprawny do modelowania dynamiki epidemii w długim czasie, [...]”.

Czytelnik dociera do tego stwierdzenia po straconych dwóch godzinach, a dodatkowo następnie okazuje że model dla niejednorodnej populacji (3.1) oparty jest na dokładnie tym samym założeniu proporcjonalnego wzrostu w ustalonym tempie. Założenie to nie jest bezsensowne, ale nie nadaje się do analizy epidemii w języku stanów stacjonarnych. Jeżeli model bez epidemii nie posiada niezdegenerowanego stanu stacjonarnego, to jasne jest że stany stacjonarne dla przypadku występowania epidemii będą artefaktami.

Być może dla modeli typu 1.1 można zadać pytanie o wartość parametrów, dla których stosunek  $I/S$  jest stały.

Człon  $AI$  w drugim równaniu modelu 1.1 sugeruje, że autor zakłada że osoby chore na gruźlicę rodzą chore dzieci. W rzeczywistości gruźlica wrodzona jest bardzo rzadka, bardziej prawdopodobne jest zarażenie się od rodziców, ale i tak gruźlica u dzieci poniżej 14 roku życia (przynajmniej w Polsce) jest rzadka, np.:

[https://wl.uwm.edu.pl/sites/default/files/download/201801/2017\\_tb\\_pl\\_dr\\_toczydlowski.pdf](https://wl.uwm.edu.pl/sites/default/files/download/201801/2017_tb_pl_dr_toczydlowski.pdf)

2. **Wprowadzenie modelu ciągłego w którym  $S$  i  $I$  są rzeczywiste.** Na stronie 14 Autor pisze „Rozważamy tutaj populację, której zagęszczenie (czyli liczba osobników na jednostkę powierzchni) w momencie  $t$  wynosi  $N(t)$ . Przyjęcie oznaczenia  $N(t)$  jako liczebności prowadzi do nierealistycznego warunku, że  $N(t) \in \mathbb{N}$  dla każdego  $t$ .” Jak dla mnie założenie, że liczba osobników jest naturalna jest realistyczne (!). To co autor próbuje niesłusznie zatuszować to fakt używania ciągłego opisu do procesu w którym zmienne zmieniają się w sposób dyskretny. Nawet jeśli  $N(t)$  będzie opisywać ilość osobników na hektar to też będzie zmieniało się w sposób dyskretny (choć w ogólności nie będzie liczbą całkowitą). Używanie opisu ciągłego do procesów dyskretnych jest standardem, choć ma swoje ograniczenia – jeśli liczba osobników (bądź molekuł) jest rzędu jednośc przybliżenie ciągłe jest obciążone dużym błędem.
3. **Wprowadzenie dyskretyzacji czasu, model 2.1.** Około połowy rozważań poświęcone jest modelom z czasem dyskretnym, w których używany jest albo schemat Eulera, albo podobny

schemat również pierwszego rzędu. Motywacja dyskretyzacji czasu jest dla mnie zupełnie nieprzekonywująca. We wstępie Autor pisze „Wielkości te [chodzi o liczby osób zarażonych] mają postać dyskretną, podano je w rocznikach statystycznych, zatem za długość kroku dyskretyzacji można przyjąć rok.” To uzasadnienie nie ma sensu. Autor sam pokazuje, (Rys. 2.2), że zachowanie układu zależy od wielkości kroku dyskretyzacji czasu  $h$ , powyżej pewnej wartości  $h$  pojawiają się oscylacje, a dla dużych wartości  $h$  rozwiązania ujemne (Rys. 2.8). Zachowanie układu nie może zależeć od tego jak często rocznik statystyczny publikuje dane – czy gdyby rocznik podawał dane z rozbiciem na miesiące trajektoria epidemii gruźlicy byłaby inna? Czy jeśli Zosia patrzy na termometr za oknem co godzinę, a Staś raz dziennie to pogoda Stasia będzie inna? Druga metoda dyskretyzacji, choć bliższa rozwiązaniom ODEs, również prowadzi do ujemnych rozwiązań.

Faktycznie, na stronie 78, Autor konstatuje, że „możliwość wystąpienia dwóch rodzajów bifurkacji pokazuje, że modele epidemiologiczne z użyciem EEM [schemat Eulera] nie obrazują odpowiednio dynamiki epidemii.” To samo dotyczy niestety drugiego schematu NSDM. Oznacza to, że całe rozważania dotyczące modeli dyskretnych nie mają nic wspólnego z modelowaniem epidemii, a są tylko, może wdzięcznym, ćwiczeniem matematycznym. Nie przeszkadza to autorowi w kontynuowaniu tych rozważań dla bardziej złożonego modelu w rozdziale 4.

Ostatecznie jedyny model który ma sens epidemiologiczny w kontekście badania stanów stacjonarnych to model 1.13, którego analizie Autor poświęca 4 strony – faktycznie prowadzi on od rozsądnych rozwiązań (Rys. 1.6), jak rozumiem nieciekawych z matematycznego punktu widzenia. Z pozostałych modeli, model typu 1.13 dla populacji niejednorodnej, tj. model 3.35 wydaje się najrozsądniejszy, ale trudno się o tym przekonać na podstawie zawartej w pracy analizy. Dla tego modelu, na Rys. 3.8 Autor pokazuje zgodność z danymi epidemiologicznymi trajektorii dla osobników zainfekowanych (z grupy małego ryzyka), nie podając jednakże trajektorii liczby osobników podatnych (czyli *de facto* wielkości populacji), która jak podejrzewam na podstawie Rys. 3.9 też istotnie zmienia się w czasie – ciekawe czy w sposób zgodny z wartościami dla województwa Warmińsko-Mazurskiego (skąd pochodzą dane) – chętnie zobaczę ten wykres na obronie.

O oddaleniu rozważań autora od rzeczywistości epidemicznej można przekonać się również przeglądając dobór przykładów w symulacjach numerycznych. Np. Rys. 2.2 pokazany jest stan stacjonarny w którym mamy ok. 12 osobników  $S$  (podatnych) i ok. 1010 osobników  $I$  (zainfekowanych) – jak rozumiem na jednostkę powierzchni (!). Gruźlica jest poważną chorobą, ale przykład w którym 99% jest zainfekowanych jest dziwny. Dla większego  $h$  dochodzi do bifurkacji podwojenia okresu i  $S$  oscyluje w przedziale od 10.1 do 12 a  $I$  pomiędzy 1006.5 i 1009.5. Na Rys. 3.5 a, b, pokazana jest trajektoria dla której liczba osobników (zainfekowanych i podatnych) z grupy małego ryzyka dąży do zera (punkt stacjonarny), liczba osobników zainfekowanych z grupy dużego ryzyka do 5, a liczba osobników podatnych (też z grupy wysokiego ryzyka) do 1. Czyli trajektoria odpowiada procesowi w którym wszyscy osobnicy z grupy małego ryzyka wymierają!

Nawet rozsądny model może mieć dziwne rozwiązania, problem w tym, że wszystkie modele poza modelem 1.13 dla dostatecznie długiego czasu wykazują głównie dziwne rozwiązania.

Reasumując, przedstawiona rozprawa doktorska, poparta publikacjami Autora świadczy o Jego biegłości w dziedzinie matematyki stosowanej. Mimo dużych obiekcji dotyczących konstrukcji modeli, dyskretyzacji czasu i dyskusji wyników wnoszę o dopuszczenie rozprawy doktorskiej do publicznej obrony i liczę na ustosunkowanie się Autora do postawionych zarzutów.

J. Lipiński