

4 Instrukcja do ćwiczenia nr 4:

Budowa modelu systemu dynamicznego w SIMULINKu.

4.1 Wstęp

Nawiązując do zakończenia ćwiczenia 7 (proszę nie zważać na numerację ćwiczeń).

Modelowanie matematyczne to coś więcej niż aproksymacja danych funkcją. Model budujemy na bazie mniej lub bardziej uproszczonych zależności między zmiennymi, zależności opisywanych znanymi (lepiej lub gorzej) prawami lub zasadami. Cytat z wikipediI: „... *model systemu to grupa funkcji wiążących ze sobą różne zmienne i w ten sposób opisujących powiązania między wielkościami w systemie*”. Koncentrujemy się na opisie matematycznym, a algorytmy aproksymacji używamy do kalibracji parametrów modelu – czasowe przebiegi zmiennych modelu (symulację procesu) „dopasowujemy” do danych uzyskanych z pomiarów wielkości w systemie.

W ćwiczeniu ze zbiornikiem i symulowaniem wpływu wody budowaliśmy prosty model zapisując go formułą matematyczną – równaniem różniczkowym zwyczajnym. W tym ćwiczeniu zrobimy dokładnie to samo – też zbudujemy model matematyczny. Pamiętając, że celem tego ćwiczenia nie jest zbudowanie najlepszego modelu, a raczej ćwiczenie użycia do tego celu jednego z narzędzi (SIMULINKa), pomińmy gotowe, profesjonalne modele epidemiologiczne¹. W ćwiczeniu 3 do symulacji używaliśmy MATLABa i zapisanych w nim algorytmów rozwiązywania równań różniczkowych. Teraz zastosujemy inne narzędzia – posługujące się symbolami graficznymi. Dlaczego efekt będzie ten sam? Bo pod tymi elementami graficznymi kryją się równania matematyczne i algorytmy ich rozwiązywania. I na to chciałbym zwrócić szczególną uwagę: SIMULINK to nie nowy algorytm, to jest tylko inna (często znacznie wygodniejsza dla nie-matematyków) forma zapisu równań.

Pominiemy – z braku czasu – przypadki równań algebraicznych i od razu przejdziemy do równań różniczkowych (którymi są opisywane systemy dynamiczne – pozostaniemy przy epidemii, a to jest proces dynamiczny). Alternatywą są równania różnicowe, o których w skrócie jest mowa w Dodatku.

4.2 Założenia o modelu epidemii – nie tylko w jej początkowej fazie jak w ćwiczeniu 7.

1. Wyodrębnienie 4 kategorii w populacji N ludzi:

O – **Odporni** – osoby zdrowe, które nabyły odporność po chorobie lub szczepieniu, Jeżeli przyjmujemy, że w chwili początkowej $O(0) = 0$ to uznamy, że nie ma odporności „samoistnej”, „naturalnej”.

P – **Podatni** – osoby zdrowe, które nie nabyły odporności. W chwili początkowej $P(0) = N - O(0)$.

¹Np. model epidemiologiczny typu SEIR <https://rmostowy.github.io/covid-19/prognoza-polska/>

Z – **Zarażeni** – osoby, które są zarażone i w czasie T powrotu do zdrowia zarażają innych. Nie rozróżniamy osób chorych z objawami od chorych bezobjawowych.

D – **Zdiagnozowani** – część osób zarażonych, u których przeprowadzono test diagnostyczny i dał on wynik pozytywny.

2. Założenia dotyczące zależności między kategoriami

a. Nie uwzględniamy przyrostu naturalnego (urodzin i zgonów) oraz migracji – zatem:

$$P = N - Z - O,$$

b. Przyrost liczby zarażonych jest proporcjonalny do liczby osób, z którymi każdy zarażony się spotyka. Wśród tych osób tylko część (P/N) jest podatna – i ta część może być zarażona. Równocześnie przyrost liczby zarażonych jest pomniejszany o przyrost liczby odpornych – tych, którzy opuszczają szpital, a T dni wcześniej zostali do szpitala przyjęci (podobnie jest z bezobjawowymi):

$$\frac{dZ(t)}{dt} = \beta \frac{P(t)}{N} Z(t) - \frac{dZ(t-T)}{dt}$$

Uwagi:

- Choroba (z objawami lub bez), rozumiana jako okres zarażania, trwa przez okres T dni.
 - Czy zarażeni zarażają? Przyjmijmy, że tak, choć w różnym stopniu (w szpitalach, na kwarantannie – mniej, chorzy bezobjawowi – bardziej), ale zamiast wprowadzania tu jakiegoś współczynnika podziału, można przyjąć, że efekt ten jest uwzględniony w ogólnym **współczynniku zarażania** β .
- c. O pewnej części (η) osób świeżo zarażonych (zarażonych w jednym dniu – jeżeli przyjmiemy krok symulacji równy jednostce czasu)² wiemy, że są nosicielami wirusa, bo zostali poddani testom (z pozytywnym wynikiem testu), pozostała część $1 - \eta$ „zwalcza swojego wirusa” nieświadomie – bezobjawowo:

$$\frac{dD(t)}{dt} = \eta Z(t),$$

η – **współczynnik diagnozowania**, sprawność diagnozowania.

d. Przyrost liczby osób odpornych na zakażenie jest równy (dopóki nie ma szczepionki) przyrostowi liczby osób, które wyzdrowiały – czyli zarażonych T dni wcześniej:

$$\frac{dO(t)}{dt} = \frac{dZ(t-T)}{dt}.$$

²O wyborze rodzaju modelu – ciągłego czy dyskretnego – jest mowa w Dodatku, ale jest to zaledwie mizerny wstęp do zagadnienia

3. Czas choroby T przyjmujemy dla uproszczenia stały i ten sam dla wszystkich, mimo że czas trwania choroby jest wielkością indywidualną.
4. Zakładamy, że współczynniki zarażania β i wykrywania zarażonych η m.in. przez testowanie są stałe, choć wiadomo, że zmieniają się w czasie (w 4. punkcie ćwiczenia „uzmiennimy” współczynnik β – bardzo prymitywnie, zastępując stałą wartość arbitralnie wybraną funkcją czasu).
5. Wartości początkowe zmiennych.
Symulację zaczynamy od 4 marca, gdy został zdiagnozowany pierwszy zarażony, czyli $D(0) = 1$. Jeżeli przyjmiemy, że sprawność diagnozowania jest równa η , to liczba zarażonych w tym dniu $Z(0) = 1/\eta$. O wartościach początkowych liczby odpornych O była mowa powyżej, przy definiowaniu tej zmiennej.
6. Dane pomiarowe pochodzą z oficjalnych komunikatów MZ $D_{pomiar}(t)$ – jak w ćwiczeniu 7, ale uzupełnione do 22 maja br.

4.3 Cele ćwiczenia

1. Jak posługiwać się SIMULINKiem.
2. Ilustracja budowy prostego modelu w SIMULNKu.
3. Budowa modelu rozwoju epidemii – uogólnionego w porównaniu do ćwiczenia 7.
4. Próba kalibracji modelu (tylko próba, bo mamy dane tylko z początkowego okresu epidemii).
5. Ilustracja możliwości ekstrapolacji wariantowego rozwoju epidemii – w zależności od zmienności współczynników (zależnych od podejmowanych decyzji i sterowania rozwojem epidemii).

4.4 Plan

1. Przegląd podstawowych elementów SIMULINKa (potrzebnych do budowy modelu epidemiologicznego):
 - a. bloki i ich parametry,
 - b. wspólna przestrzeń zmiennych (MATLAB Workspace),
 - c. zmienne (sygnały),
 - d. opis (komentarz) diagramu,
 - e. ustawianie parametrów symulacji (czas końcowy t_k , wybór solvera, maksymalny/minimalny/automatyczny krok czasowy).
 - f. uruchomienie symulacji z Command Window poleceniem `sim`.

2. Budowa modelu w SIMULINKu:

- a. diagram modelu,
- b. ustawienie wartości początkowych i ograniczeń,
- c. symulacja modelu dla różnych współczynników zarażania β ,
- d. porównanie wyników symulacji z danymi oficjalnymi,
- e. pętla kalibracji,
- f. przykład modelu z współczynnikiem zmiennym w czasie.

4.5 Przebieg ćwiczenia

1. Budujemy najprostszy model ciągły

Zadanie: zbudować model opisany równaniem $\frac{dZ}{dt} = \beta Z$.

Którego bloku użyjemy – różniczkującego, czy całkującego? I dlaczego całkującego?

Każdy blok, który ma wejście i wyjście, realizuje określoną funkcję przejścia: Wyjście zależy od wejścia, a w niektórych blokach także od stanu tego bloku. Np. integrator wyprowadza na wyjście całkę sygnału wejściowego – całkę oznaczoną od 0 do bieżącego czasu, z warunkiem (stanem) początkowym zapisanym w parametrach tego bloku.

Blok narzuca relację, ale do nas należy interpretacja wejścia i wyjścia. Jeżeli na wejście podamy sygnał $x(t)$, to na wyjściu będzie całka z $\int_{-\infty}^t x(\tau) d\tau$. Ale jeżeli przyjmiemy interpretację taką, że na wyjściu jest sygnał $x(t)$, to na wejście musimy zinterpretować jako $\frac{dx(t)}{dt}$. Przypominam, że całka oznaczona w granicach od $-\infty$ do t z pochodnej funkcji jest równa wartości tej funkcji w górnej granicy całkowania, przy czym wartość tej całki od $-\infty$ do 0 jest stanem (warunkiem) początkowym, a w przedziale $[0, t]$ jest obliczana w kolejnych krokach symulacji.

Ponieważ model jest opisany równaniem różniczkowym, a szukamy rozwiązania tego równania, tj. funkcji $Z(t)$, to wybieramy blok całkujący. Związek między jego wejściem, a wyjściem jest określony zadaniem równaniem $\frac{dZ}{dt} = \beta Z$.

Ten związek mamy przedstawić w diagramie. Zaczynamy od sprowadzenia do okienka edycyjnego bloku Integrator. Wejście interpretujemy jako pochodną $\frac{dZ(t)}{dt}$, zatem na wyjściu otrzymamy sygnał $Z(t)$. Aby przeprowadzić symulację, musimy „zaspokoić” wszystkie wejścia użytych bloków. Jak zatem skonstruować sygnał pochodnej (aby doprowadzić go do wejścia integratora)? O tym mówi równanie modelu: pochodna ma być równa iloczynowi sygnału $Z(t)$ i stałej β . Używamy więc bloku wzmacniacza (zawsze gdy mamy sygnał pomnożyć przez stałą), na jego wejście podajemy $Z(t)$ i na wyjściu wzmacniacza dostajemy potrzebną wielkość – $\beta Z(t)$, która ma być równa pochodnej. Doprowadzenie tego sygnału do wejścia bloku całkującego dopina pętlę – tak zostaje zapisane równanie. Ogólnie: każda pętla to jedno równanie. Model jest gotowy.

Aby obserwować przebieg symulacji używamy bloku Scope, ewentualnie (ale rzadko)

Display, a w przypadku, gdy wyniki chcemy przesłać do przestrzeni roboczej MATLABa – bloku To workspace.

- Z biblioteki sprowadzamy bloki: Integrator, Gain, Scope, To workspace, Display (rzadko stosowany w modelach ciągłych).
- Równanie zapisujemy w postaci pętli.
- Opisujemy sygnały i bloki i otrzymujemy diagram jak na rysunku 1.

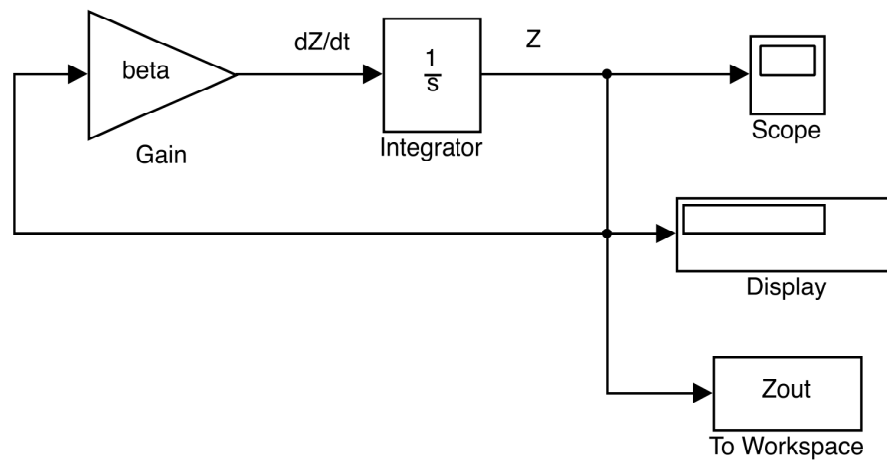


Figure 1: Diagram prostego modelu $\frac{dZ}{dt} = \beta Z$.

- Aby uruchomić symulację należy wcześniej:
 - Wybrać horyzont symulacji (czas końcowy – w okienku górnego menu) – wartość wstępna = 10). Tu pojawia się pytanie o jednostkę czasu. Najczęściej wybieramy jednostki układu SI (chyba, że pracujemy w np USA). W naszym modelu dogodniejszym wyborem jednostki będzie dzień – bo dane były zmierzone w odstępach jednego dnia,
 - Jeżeli rezultatem symulacji mają być dane obliczane też dla każdego dnia, to krok obliczeń powinien być równy 1 (to nie jest ogólna zasada). W parametrach symulacji (zębate kółko / śrubka) ustawiamy (tu są różnice między wersjami SIMULINKa, więc poniższy opis może być mylący):
Solver type: Fixed-step,
Step size: 1.
 - Ustawić warunek początkowy zmiennej Z – klik na integratorze i wypełniamy pole Initial condition. Możemy też zaznaczyć Limit output i zadać dolną i górną granicę sygnału wyjściowego.
Uwaga: Do ww. pól można wpisywać wartości stałe, ale zdecydowanie lepiej

używać tu nazw zmiennych, które przed symulacją zdefiniujemy i nadamy im wartości w Command Window MATLABa (lub uruchamiając skrypt zawierający te definicje).

- W bloku wzmacniacza Gain jest ustawiona wstępna wartość wzmocnienia = 1. Sugerowałbym zmianę tej wartości i obserwację zmian wykresu generowanego w bloku Scope.
- Uruchamiamy symulację i obserwujemy przebiegi zmiennej $Z(t)$.

2. Wstęp do modelu epidemiologicznego.

W modelu uwzględnimy nie tylko osoby zarażone.

- Dodajemy integrator – blok reprezentujący zbiór osób zdiagnozowanych dodatnio – wśród osób zakażonych, część jest zdiagnozowana pozytywnie (a tylko tę liczbę znamy z oficjalnych raportów i przebieg zmienności tej liczby może posłużyć do kalibracji modelu).
- Dodajemy blok opóźnienia **Trasport Delay** reprezentującego czas T trwania choroby – każda osoba po przejściu choroby staje się odporna. Blokowi **Trasport Delay** zmieniłem nazwę na **Okres choroby**. Po kliknięciu na nim należy wpisać wielkość opóźnienia T w przyjętych jednostkach – u nas – dniach. Jeżeli do pola opóźnienia wpisujemy nazwę zmiennej T , to musimy pamiętać, żeby taką zmienną zdefiniować (jak wszystkie inne).
- Uwaga: Nazwa pojawiająca się w ikoncie bloku (np. **Gain** jest nazwą zmiennej wpisanej do bloku jako współczynnik wzmocnienia. Zatem przed uruchomieniem symulacji należy w Command Window nadać wartości tym zmiennym – **beta**, **T**.

3. Rozbudowa modelu - uwzględnimy ograniczoną populację i wpływ wzrastającej liczby odpornych na zaraźliwość.

Liczba osób zarażanych jest tym mniejsza im – wśród osób kontaktujących się z zarażonym – jest więcej osób odpornych. Zmniejszenie zaraźliwości można wyrazić współczynnikiem Podatni/Wszyscy czyli P/N .

- Potrzebne bloki: **Constant**, **Sum**, **Product**.
- Kiedy używamy bloku **Gain**, a kiedy **Product**: **Gain** gdy sygnał (zmienną) mnożymy przez wyrażenie stałe, a **Product**, gdy konieczne jest obliczenie iloczynu 2 lub więcej zmiennych. **Gain** ma jedno wejście, a **Product** - dowolną liczbę wejść.
- Uwagi dotyczące nie tylko estetyki, ale także czytelności:
 - przy komplikacji diagramu pojawiają się skrzyżowania sygnałów – należy ich unikać,
 - jest możliwość obrotu lub odwrócenia dowolnego bloku (prawy klik na blok i z menu wybieramy **Format**),
 - proszę porównać czytelność diagramów z rysunku 3 i 4.

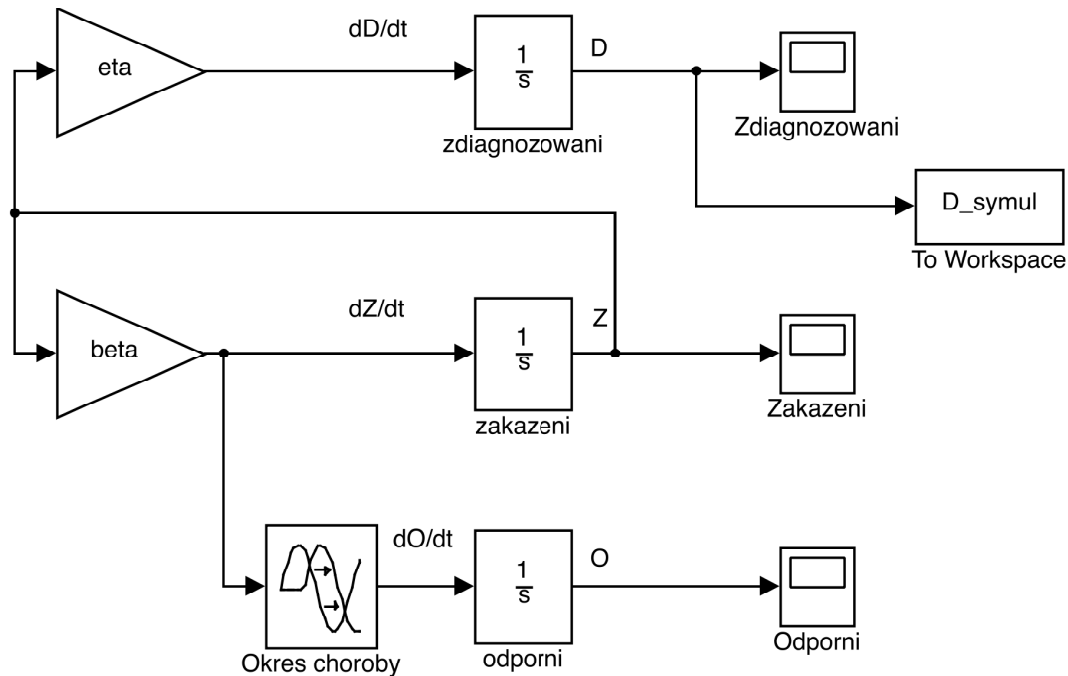


Figure 2: Diagram modelu epidemiologicznego - wersja 1.

- W skrypcie `glowny.m` są zdefiniowane zmienne użyte w modelu – należy go wywołać w Command Window MATLABa.
- Należy uruchomić symulację i obserwować wykresy bloków Scope (w okienku wykresu jest ikonka przeskalowania osi tak, aby był widoczny cały wykres),
- Należy kilkakrotnie uruchamiać symulację, za każdym razem niewiele zmieniając współczynnik β – to jest współczynnik „przekazywania” wirusa. Jaka jest wrażliwość wyników symulacji na wartość tego współczynnika?
- Symulacja wywołana z MATLABa. Proszę zdjąć komentarz z ostatnich linii skryptu `glowny.m` – za opisem: `Wywołanie symulacji`, zmienić w poleceniu `sim` nazwę zbudowanego modelu (w przykładzie jest `model_nr_1`) i wykonać go. W MATLABie powinien pojawić się wykres symulacji na tle danych oficjalnych.
- Porównanie z danymi oficjalnymi – prawdopodobnie zaobserwujemy wyraźne niedopasowanie.

4. Wstęp do kalibracji modelu

- Dysponujemy za **małą ilością pomiarów** (jesteśmy wciąż na początkowym etapie epidemii) aby kalibrować model w całym zakresie trwania epidemii.
- **Współczynnik diagnozowania** η mówiący jaką część osób zarażonych wykrywamy na drodze testowania możemy przyjąć na podstawie przeprowadzonego

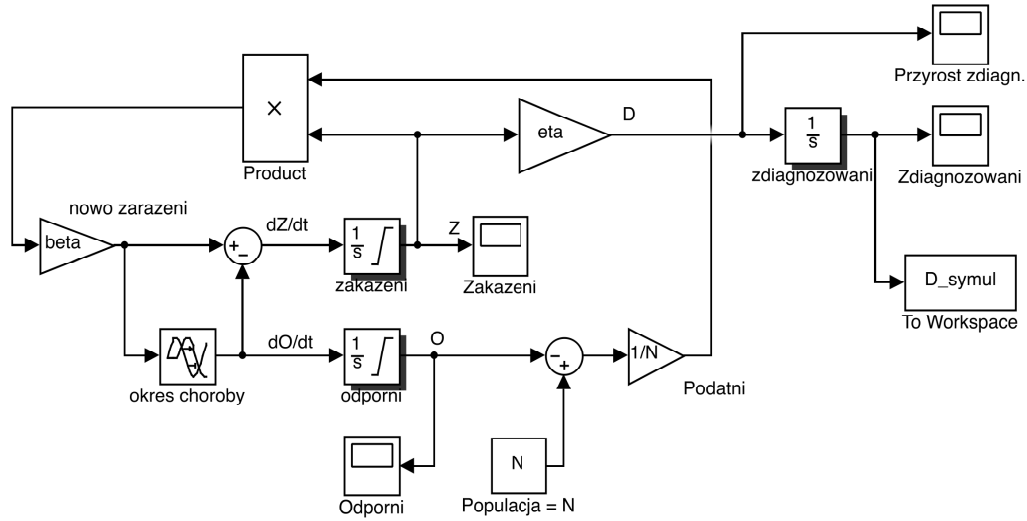


Figure 3: Diagram modelu epidemiologicznego ze stałym współczynnikiem zarażania β .

w maju w Krakowie badania. Wykazało ono, że na 1000 badanych osób bez żadnych objawów, 20 wyników testów jest pozytywnych. Stąd wynikałoby, że 2% populacji stanowią osoby zarażone. Wiemy, że w tym czasie osób zdiagnozowanych pozytywnie w województwie małopolskim było ok. 1000 osób, a wszystkich mieszkańców województwa ok. 3,5 mln. Zatem można przypuszczać, że zarażonych jest 2% tj. 70 000, co przy 1000 zdiagnozowanych daje sprawność wykrywania na poziomie $\eta = 1/70$.

- Szacunkową wartość **współczynnika zaraźliwości** β można określić na podstawie wyników ćwiczenia nr 7 (choć model był nie taki sam) i wstępnie przyjąć $\beta = 0.1$.
- Początkowa liczba osób zdiagnozowanych pozytywnie (w dniu 4.03.2020) $D(0) = 1$, a liczba osób zarażonych w tym dniu – jeżeli przyjąć stały współczynnik wykrywalności η , to $Z(0) = D(0)/\eta = 70$.
Początkowa liczba osób „z natury” odpornych nie jest znana (na razie) – możemy równie dobrze przyjąć $O(0) = 0$ jak i np. 20% populacji N . Jedyne liczebność populacji możemy określić stosunkowo dokładnie i przyjąć $N = 38 \cdot 10^6$.
- Symulacja modelu z wyznaczonymi po kalibracji współczynnikami nie da wiarygodnej odpowiedzi na pytanie o przyszły przebieg epidemii. Zatem do ekstrapolacji poza okres, dla którego mamy dane pomiarowe należy podchodzić bardzo sceptycznie – co najwyżej jako ekstrapolację aktualnej tendencji. Można ją wykorzystać także do prognozowania skutków alternatywnych strategii. Przebieg ten zależy od decyzji, które będą miały wpływ na przede wszystkim współczynnik zarażania β .

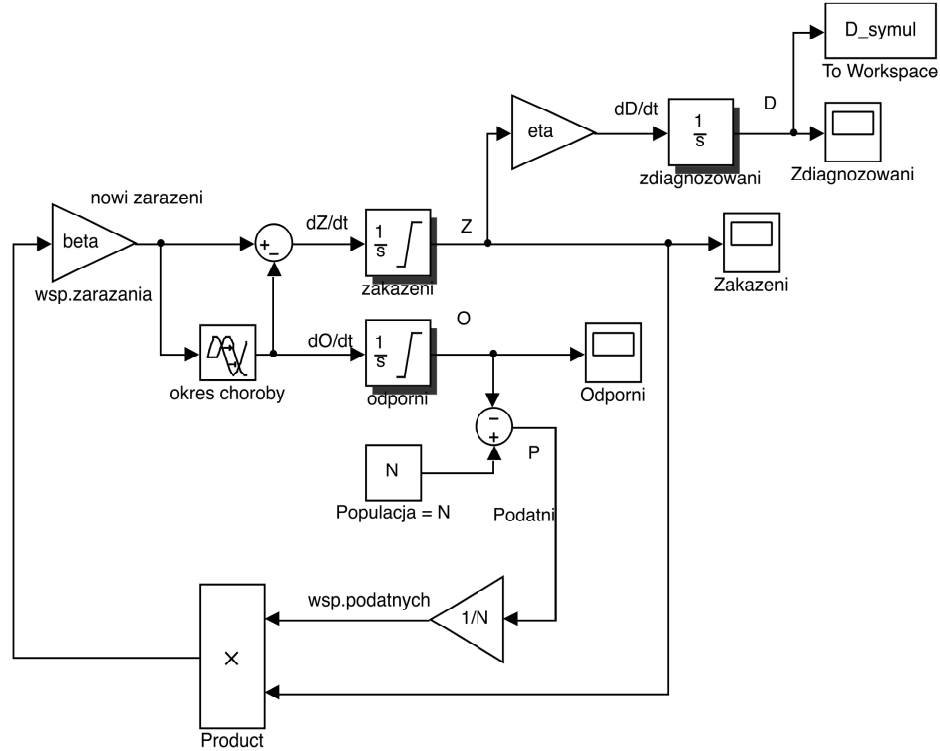


Figure 4: Inny diagram modelu epidemiologicznego ze stałym współczynnikiem zarażania β .

- Symulacja z dalszym horyzontem czasowym t_k wymaga uwzględnienia faktu, że wartości przyjętych współczynników nie są stałe. Na potrzeby tego ćwiczenia, podobnie jak w ćwiczeniu 7, możemy przyjąć, że zmienność współczynnika β będzie aproksymowana przeskalowaną funkcją arcus tangens (pamiętając, że to był tylko pomysł „pierwszy z brzegu”). W dalszej perspektywie należałoby prawdopodobnie uwzględnić dodatkowy składnik reprezentujący wychodzenie z „lock-down’u”, kolejny składnik po rozpoczęciu szczepień itd.
- Przykład „uzmiennienia” współczynnika zarażania β jest pokazany na rysunku 5. Wzmacniacz o stałym wzmocnieniu β został zastąpiony blokiem realizującym funkcję zdefiniowaną przez użytkownika. W tym przykładzie jest to (niewidoczna na diagramie) funkcja $\beta(u) = a * (\pi/2 - \arctan(b \cdot u))$, gdzie a i b to parametry modelu, a sygnałem wejściowym u jest czas.
Uwaga: Latex, w którym piszę tę instrukcję toleruje nazwę `arctan` dla arcusa tangensa, a w MATLABie należy użyć nazwy `atan`.
- skrypt `glowny_wsp_var.m` jest wstępem do kalibracji. Ustawia parametry modelu o przykładowej nazwie `model_epid_wsp_var` (rys. 4) i uruchamia dwa razy symulację (dla różnych wartości parametrów). Oblicza też odległość symulacji od danych (stosując funkcję `norm` obliczającą normę euklidesową).

Dla weryfikacji dotychczasowych działań umieściłem na rysunku 6 wykres przebiegu liczby osób zarażonych, jaki otrzymałem symulując mój model.

- Sądzę, że ten przykład nasunie Państwu pomysł jak zbudować algorytm, który wyliczy optymalne wartości parametrów a i b . Podobnie jak w ćwiczeniu 7.
Uwagi:
 - Przyznaję, że lepiej byłoby umieścić obydwie te parametry w jednym wektorze.
 - Współczynnik zarażania można uzależnić również od zmiennych modelu (nie tylko od czasu).
- Na eksperymenty z modelem dyskretnym (patrz – Dodatek) prawdopodobnie nie wystarczy czasu – w przeciwnym przypadku będziemy improwizować.

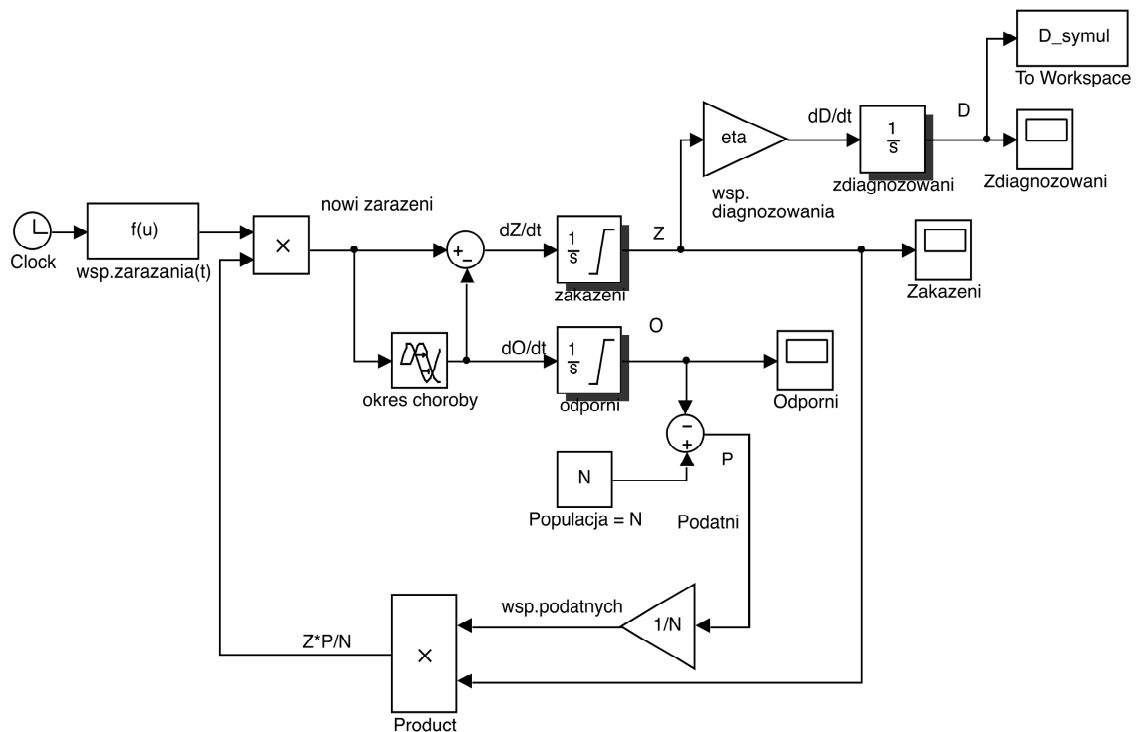


Figure 5: Diagram modelu ze zmiennym współczynnikiem zarażania.

To, co było na tym ćwiczeniu, to „robótki ręczne”. Trzeba pamiętać, że w SIMULINKu są specjalistyczne narzędzia np. usprawniające kalibrację modelu – np. w przyborniku Simulink Design Optimization.

„Dalej żegluj sam, żeglarzu...”

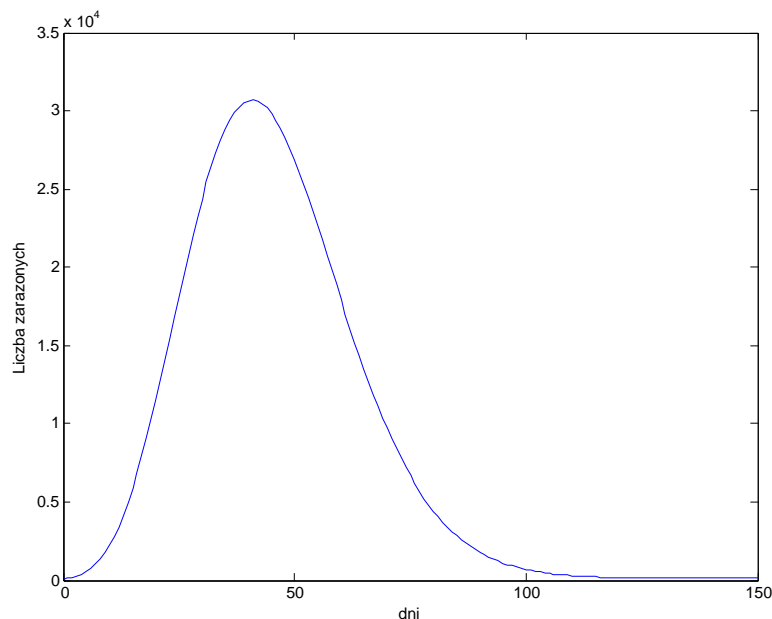


Figure 6: Wykres zmian liczby zarażonych.

4.6 Sprawozdanie

W sprawozdaniu należy umieścić:

1. Zbudowany model (w wersji stałego lub zmiennego współczynnika zarażania) – rysunek diagramu otrzymany po wybraniu z menu File / Print – może być w formacie .ps, chyba, że nowsza wersja SIMULINKa produkuje wygodniejsze formaty.
2. Wykres przebiegu liczby zakażonych z podaniem parametrów symulacji (solver, krok, wszystkie parametry modelu, wartości początkowe zmiennych).
3. Odpowiedź na pytanie: Ile jest pętli w diagramie na rysunku 4?
4. Nadobowiązkowo – dla pasjonatów – skrypt lub opis algorytmu automatycznie kalibrującego model.

Uwaga:

W modelach przedstawionych na rysunkach 3 i 4 jest błąd (niezgodność modelu z założeniami), który nie ma istotnego wpływu na wynik.

Jeżeli ktoś go znajdzie, to koniecznie proszę o nim napisać w sprawozdaniu.

5 Dodatek: Model ciągły czy dyskretny

- Blok integratora $1/s$ (dla modeli ciągłych) można zastąpić blokiem Unit Delay $1/z$ (dla modeli dyskretnych).
- W bloku $1/z$ nie ma możliwości wpisania stanu początkowego ani ograniczeń na sygnał wyjściowy (nasycenia) – trzeba zastosować blok nadania wartości początkowej sygnałowi oraz bloku Saturation.
- Wybór kroku ma inną „wagę”. W modelach ciągłych ma znaczenie jedynie dla dokładności rozwiązania równania różniczkowego, a w modelach dyskretnych krok trzeba dostosować do procesu zachodzącego w obiekcie.
- Proponuję porównanie modelu z rysunku 1 z odpowiednikiem dyskretnym. Należy wymienić blok $1/s$ na $1/z$, ustawić stały krok $= 1$ i porównać wyniki.

Zapewne dokładne relacje między opisem ciągłym a dyskretnym były omawiane na Systemach dynamicznych, więc tu odwołam się tylko do intuicji:

Rozważmy prosty model ciągły

$$\frac{dx(t)}{dt} = ax(t).$$

Prawa strona równania różniczkowego wyznacza **przyrost** dx zmiennej x przypadający na dt .

Dla tego samego obiektu budujemy model dyskretny

$$x[n] = bx[n-1]$$

Prawa strona równania różnicowego wyznacza **wartość** zmiennej x w następnej chwili czasu.

Dlatego, opisując ten sam obiekt raz modelem ciągłym, a raz dyskretnym, można przyjąć, że $b = 1 + a$ – wartość w chwili następnej to jest wartość z chwili poprzedniej + przyrost.