

بسم الله الرحمن الرحيم



پروژه درس شبیه سازی کامپیوتری

موضوع پروژه: شبیه سازی مونت کارلو برای مدل سازی شیوع کووید ۱۹ در طول زمان در کشور ایران

استاد: جناب آقای دکتر امیر حسین طباطبایی

گروه علوم کامپیوتر

دانشجویان

ایمان کیانیان

نوید جمالی پور نوکنده

مهراد دهقان

محمد حسن پور

شهاب گلزار

نیم سال اول سال تحصیلی ۱۴۰۰-۱۳۹۹

ویروس کرونا (COVID-19) هم اکنون در اکثر کشورهای جهان شیوع پیدا کرده که باعث خسارات جانی سنگین و آسیب رساندن به مواضع اقتصادی-اجتماعی شده است. با استفاده از یک روند تصادفی رویکرد مدل سازی، یک مدل شبیه سازی مونت کارلو برای نشان دادن پویا بودن شیوع COVID-19 توسعه داده شد.

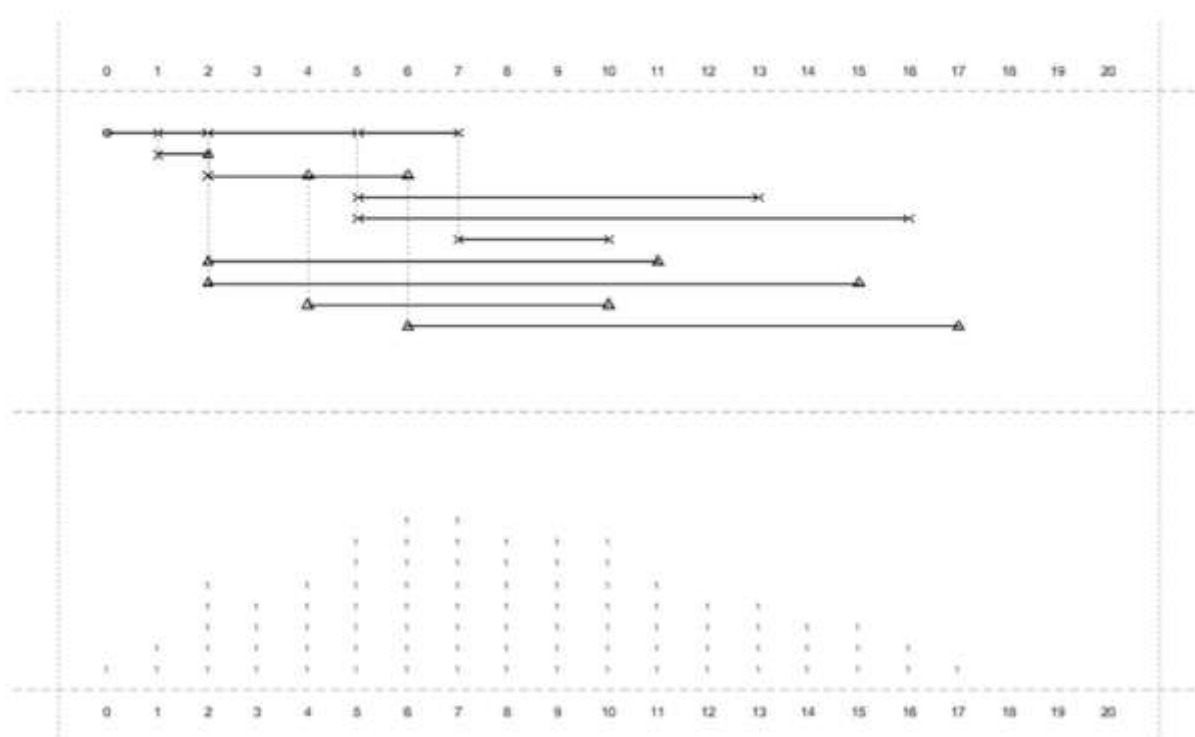
در ابتدا، عملکردهای مختلف مورد انتظار (خصوصیات نظری) را با فرض تعدادی از سناریوهایی که به صورت قراردادی تعریف شده اند بررسی کردیم. سپس مطالعات شبیه سازی پس از آن بر روی داده های واقعی گزارش شده COVID-19 (در طی دوره ۱۱ اسفند تا ۱۰ اردیبهشت) برای ایران انجام شد. با توجه به تعداد اولیه موارد آلوده به COVID-19 که در کشور حدود ۱۰ مورد بود، این مدل تخمین زده است که تعداد موارد فعال در تاریخ ۸ تیر در کشور ایران (۵۹۸۹ نفر) خواهد بود. در نهایت تعداد کل موارد تایید شده می تواند برای ایران در طی ۶۰ روز از تاریخ ۱۱ اردیبهشت تا ۸ تیر ۱۳۴۲۵۳ باشد. نتایج تخمین زده شده در رابطه با شیوع COVID-19 با آنچه که گزارش شده مطابقت داشت. این مدل شبیه سازی به عنوان یک تصمیم گیری موثر و سازگار در نظر گرفته شد، چه از لحاظ ابزار تجزیه و تحلیل برای نیاز مبارزه فوری با COVID-19 و چه از لحاظ مدل سازی سایر بیماری های عفونی در آینده.

همه گیری ویروس کرونا (COVID-19) در شهر ووهان چین در دسامبر سال ۲۰۱۹ آغاز شد. سازمان بهداشت جهانی در اوایل مارس ۲۰۲۰ به طور رسمی اعلام کرد که یک بیماری همه گیر بنام کووید ۱۹ در اکثر کشورهای جهان گسترش یافته که خسارات جانی سنگین و آسیب رساندن به مواضع اقتصادی-اجتماعی را در پی داشته است. شیوع کووید ۱۹ در واقع در حال حاضر تهدیدی می باشد که بهداشت عمومی منطقه ای را به یک بحران اجتماعی-اقتصادی جهانی تبدیل کرده است.

این مطالعه با هدف توسعه یک مدل شبیه سازی مونت کارلو مبتنی بر پویا بودن شیوع COVID-19 انجام شده است که می تواند به عنوان یک ابزار تصمیم گیری برای نیاز مبارزه فوری با COVID-19 و برای مدل سازی سایر بیماری های عفونی در آینده مورد استفاده قرار گیرد.

با این حال، این روش برای مدل سازی فرآیندهای دینامیکی انتقال ویروس/بیماری متفاوت است. در این مطالعه، پیشرفت COVID-19 با یک مدل شبیه سازی مونت کارلو مشخص شده که شبیه یک مدل فرآیند تصادفی می باشد. با در نظر گرفتن هر فرد به عنوان یک نقطه تصادفی در جمعیت، مدل جدید بسیار شهودی است و تفسیر آن آسان می باشد. خروجی های مدل جدید شامل تعداد تخمین های روزانه از موارد تازه تأیید شده و تعداد روزانه موارد فعال (به عنوان مثال، موارد جدید و موارد عفونی) در یک دوره مشاهده مشخص باشد. بنابراین خروجی های مدل حاوی اطلاعاتی برای پاسخگویی به مهمترین سوالات تصمیم گیری از قبیل این بود که اگر تعداد موارد فعال COVID-19 به اوج خود برسد، چه مدت طول می کشد تا تعداد موارد فعال به یک سطح مشخص کاهش یابد، تعداد کل موارد تأیید شده تا پایان شیوع COVID-19 چه مقدار خواهد بود و غیره. مهمترین پارامتر از مدل شبیه سازی، میانگین/تعداد موارد مورد انتظار یک فرد مبتلا به کووید ۱۹ است که در طی دوره ابتلای این فرد، فرد دیگری آلوده شده باشد. در این مقاله "Rt" اساساً به عنوان پارامتر میزان عفونت در نظر گرفته می شود.

در حقیقت تعیین میزان عفونت (R_t) در طی یک دوره مشاهده شده، به عوامل زیادی بستگی دارد، از جمله حداقل برآزش از تعداد موارد مشاهده شده که COVID-19 آن‌ها تأیید شده، تفسیر تأثیر موارد مداخلات/سیاست‌های دولت، ظرفیت سیستم مراقبت‌های بهداشتی و ارزیابی پتانسیل عواقب ناشی از واکنش‌های عمومی نسبت به سیاست‌های دولت و تغییر رفتار مردم با وادار کردنشان به تبعیت از این سیاست‌ها.



دوره مشاهده شده: روزها

شکل ۱. شمای مدل فرایند تصادفی با جمعیت فرضی ۱۰ نفر (۱۰ خطوط افقی در قسمت بالای شکل): انتقال ویروس از یک نفر آغاز می‌شود (که توسط دایره توخالی اولیه نشان داده شده است)؛ سپس پنج نفر را آلوده می‌کند (که توسط علامت ضربدر نشان داده شده)؛ دو نفر از پنج نفر ثانویه که آلوده بودند، توانستند گروه سوم افراد را آلوده کنند (که با مثلث نشان داده شده‌اند). قسمت پایین شکل الگویی از تعداد موارد مبتلا را نشان می‌دهد.

روش و جزئیات فنی مدل شبیه‌سازی

پویا بودن شیوع گسترش COVID-19 با چند پارامتر کلیدی در ساخت مدل شبیه‌سازی مشخص می‌شود. پارامترها و فرضیات کلیدی عبارتند از:

(۱) پارامتر میزان عفونت (R_t)

(۲) تعداد واقعی افراد که توسط هر یک از موارد مبتلا شده آلوده می‌شوند، از یک توزیع پواسون با میانگین R_t پیروی می‌کند. تعداد دقیق روزها برای کسی که مبتلا شده از توزیع دو جمله‌ای منفی با میانگین/زمان انتظار (μT) در روز پیروی می‌کند.

(۳) تعداد موارد محدود از افراد و نسبت مصونیت در یک جمعیت می‌تواند توسط کاربر از قبل تعیین شود.

(۴) سایر شرایط اولیه از جمله تعداد اولیه افراد مبتلا، دوره مشاهده (در روز) و دوره شبیه‌سازی (در روز) از یک مطالعه شبیه‌سازی.

مدل مفهومی فرآیند تصادفی با فرض یک جمعیت فرضی کوچک ۱۰ نفره به بهترین نحو در شکل ۱. نمایش داده شده است. دوره مشاهده ۲۰ روز بود. فرآیندهای شیوع COVID-19 از یک نفر شروع شد (که توسط یک دایره توخالی نشان داده شده بود) که پنج نفر را آلوده کرد (که با علامت ضربدر نشان داده شده بودند)، تعداد افراد مبتلا توسط توزیع پواسون با پارامتر R_t مدل‌سازی شد که می‌تواند در مراحل مختلف دوره مشاهده متفاوت باشد. حدس ما این است که حداقل یک روز برای یک فرد تازه مبتلا شده لازم است تا شروع انتقال COVID-19 به گروه بعدی صورت گیرد. زمان‌های ابتلا مستقل از یکدیگر هستند، به طوریکه بیش از یک نفر می‌تواند از فرد مبتلا شده در همان روز آلوده شوند. تعداد دقیق روزها برای فردی که مبتلا شده، از یک توزیع دو جمله‌ای منفی با میانگین/زمان انتظار μT در روز پیروی می‌کند. به محض آنکه توزیع پواسون که تعداد افراد مبتلا را ایجاد می‌کند، فرارسید، منبع فرد مبتلا شده از فرآیند انتقال خارج می‌شود (به عنوان مثال، با فرض اینکه منبع فرد مبتلا شده، بهبود یافته یا فوت شده یا حداقل در شرایطی نباشد که قادر به آلوده کردن افراد دیگر باشد، در این صورت دیگر در گروه موارد مبتلا شده باقی نمی‌ماند. اگر فرد آلوده شده، شخص دیگری برای انتقال ویروس نداشته باشد، این فرد به طور متوسط از گروه موارد مبتلا با میانگین μT در روز حذف می‌شود. طبق شکل ۱، قسمت بالای شکل به ما نشان می‌دهد که چه تعداد و چه زمانی موارد آلوده شده برای یک جمعیت کوچک ۱۰ نفره در یک دوره مشاهده ۲۰ روزه اتفاق می‌افتد. در قسمت پایینی شکل الگوی مربوط به تعداد موارد مبتلا را نمایش می‌دهد (یعنی تعداد افرادی که به طور روزانه از قبل آلوده بودند). مدل شبیه‌سازی برای استفاده در پایتون دارای قالب زیر می‌باشد:

Simulation (days=300, nd=30, $R_t=rr$, $\mu T=4$, sizeV=1, limit=80,000,000, pp=0.001, $n_0=1$)

پارامتر ورودی	تعریف پارامتر ورودی
Days	دوره مشاهده (در روز) از یک مطالعه شبیه‌سازی
nd	دوره شبیه‌سازی (در روز) از یک مطالعه شبیه‌سازی
Rt	میانگین تکثیر/ میزان عفونت (به عنوان مثال ، تعداد مورد ثانویه مورد انتظار که هر مورد عفونی موجود ایجاد خواهد کرد) ؛ Rt تابعی از زمان / روز است
muT	تعداد/ میانگین روزهای مورد انتظار برای یک فرد عفونی موجود برای آلوده کردن یک فرد مستعد در جمعیت؛ منظور داشتن پارامتر برای توزیع دو جمله ای منفی
sizeV	پارامتر پراکندگی برای توزیع دو جمله ای منفی به طوری که: $\text{variance} = \mu T + (\mu T)^2 / \text{sizeV}$
Limit	بررسی / اندازه جمعیت هدف
PP	نسبت افراد دارای مصونیت در جمعیت
n0	تعداد اولیه افراد عفونی قبل از دوره مشاهده / شبیه‌سازی

جدول ۱. تعاریف اصطلاحات برای تابع شبیه‌سازی(Simulation)

در این تابع هشت پارامتر / اصطلاح آرگومان وجود داشت و تعاریف این اصطلاحات آرگومان در جدول ۱ آورده شده است. با توجه به عدم اطمینان ارائه شده توسط توزیع احتمالات برای تعیین تعداد موارد عفونت و زمان وقوع چنین رویدادی ، ما باید دوره مشاهده ("روز") را نسبتاً طولانی تر از دوره شبیه‌سازی ('nd') بپذیریم. برای مدل‌سازی داده‌های COVID-19، مشخص شد که $nd = 60$ برای به دست آوردن نتایج شبیه‌سازی تثبیت شده کافی بود. طول Rt باید حداقل به اندازه مقدار nd باشد. این بهتر است از طریق یک مثال توضیح داده شود.

اکنون فرض کنید یک الگوی نرخ عفونت فرضی به صورت زیر داشتیم؛

$$rr=[2.8 \text{ for } i \text{ in range}(45)]$$

در این صورت طول Rt برابر ۴۵ است.

اگر سعی کنیم مدل شبیه‌سازی را به صورت $\text{Simulation}(nd = 60, Rt = rr)$ در پایتون اجرا کنیم، یک پیام خطا در پنجره کنسول پایتون برمی‌گردد: خطا در $\text{Simulation}(nd = 60, Rt = rr)$: طول Rt نباید از nd کوچکتر باشد. از آنجا که فرایندهای شبیه‌سازی مونت کارلو شامل تولید اعداد تصادفی است که به معنی شبیه‌سازی

می‌باشد، نتایج تجزیه و تحلیل به دلیل نقاط شروع مختلفی که ذاتاً توسط یک مقدار seed تصادفی تعریف می‌شوند، در معرض تغییر قرار می‌گیرند. بنابراین، وقتی سناریوهای مختلف انتقال بیماری را مقایسه می‌کنیم، با استفاده از مدل شبیه‌سازی، برای مقایسه صحیح هر سناریو لازم است همان seed تصادفی مشخص شود.

بهتر است در این کد هر مدل شبیه‌سازی ۱۰۰۰ بار اجرا شود و مقادیر میانه به عنوان محتمل‌ترین تخمین در نظر گرفته شود و سپس میتوان به صحت شبیه سازی اطمینان حاصل نمود.

خروجی‌های مدل شبیه‌سازی شامل سه بخش اطلاعات است: تعداد تخمینی موارد فعال روزانه، تعداد تخمینی موارد جدید روزانه و تعداد کل موارد تایید شده در دوره شبیه سازی. بنابراین، جمع بندی تعداد تخمینی تمام موارد تأیید شده روزانه باید دقیقاً برابر با کل موارد تایید شده باشد.

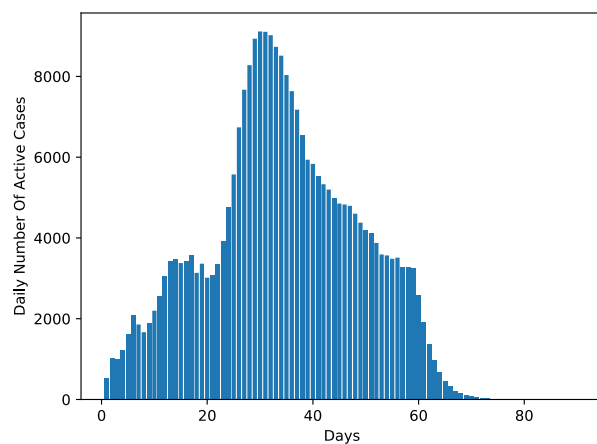
بررسی تاثیر مقادیر μT و $sizeV$

دو پارامتر (μT و $sizeV$) برای تعریف یک توزیع چند جمله‌ای منفی وجود دارد که برای توصیف فاصله زمانی برای اینکه فرد مستعد آلوده شود استفاده می‌شود. در این مطالعه، ما سناریوهای مختلف را با بررسی مقادیر مختلف μT و $sizeV$ مقایسه کردیم. سه تنظیم شبیه‌سازی مختلف که مورد بررسی قرار گرفتند به صورت زیر می‌باشند:

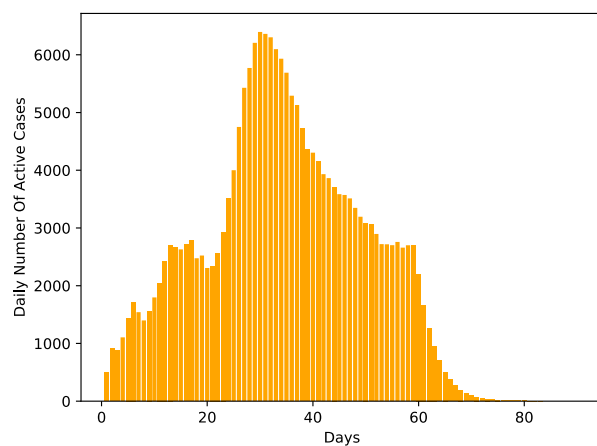
Simulation(nd=100, $\mu T=2.8$, $sizeV=1.3$) (حالت ۱)

Simulation(nd=100, $\mu T=3.3$, $sizeV=1.3$) (حالت ۲)

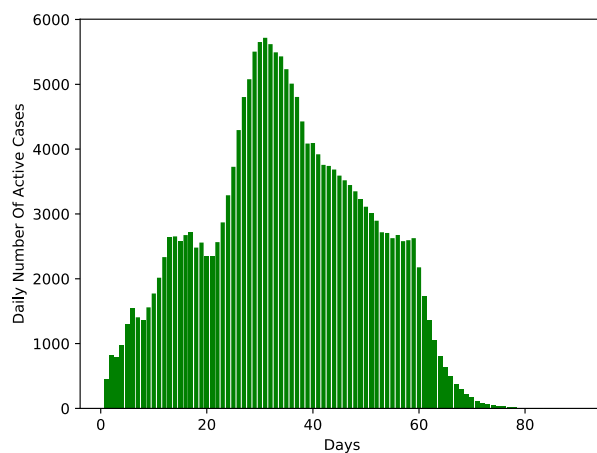
Simulation(nd=100, $\mu T=2.8$, $sizeV=0.9$) (حالت ۳)



شکل حالت ۱



شکل حالت ۲



شکل حالت ۳

در این مقایسه و با بررسی دقیق داده ها ، سطح اوج بیماران فعال کرونا با تغییر μT و $sizeV$ تغییر چشمگیری میکند. هر چقدر μT کمتر باشد ، نقطه اوج بالاتری داریم و سرعت نزول سریع تر است. یعنی μT بالاتر باعث رخداد نقطه اوج کمتر و سرعت نزول ملایم تر خواهد شد.

بلعکس ، با کاهش $sizeV$ نقطه اوج بیماران فعال کرونا کاهش پیدا میکند و سرعت نزول ملایم تر خواهد شد . پس اگر $sizeV$ افزایش پیدا کند نقطه اوج بالاتری داریم و سرعت نزول سریع تر است.

به عبارت دیگر مقایسه مقادیر μT (۲,۸ در مقابل ۳,۳ روز) سطح اوج را کاملاً افزایش می دهد. مقایسه مقادیر $sizeV$ نشان داد که کاهش مقادیر $sizeV$ (۰,۹ در مقابل ۱,۳) یعنی مقدار کوچکتر $sizeV$ به معنی واریانس بزرگتر می باشد. همچنین با ثابت نگه داشتن $sizeV$ ، μT با مقدار پایین تر، پیک بالاتری خواهد داشت. در مقابل، با ثابت نگه داشتن $sizeV$ ، μT با مقدار بالاتر، با روند آروم تری نزول می کند. بنابراین به طور خلاصه، این قسمت از مطالعه شبیه سازی نشان داد، همانطور که توسط تئوری اساسی انتظار می رود، یک Rt بالاتر یا یک μT کوچکتر با یکسان نگه داشتن سایر موارد باعث می شود تعداد بیشتری از افراد آلوده شوند.

نتایج مطالعات شبیه سازی بر روی داده های کشور ایران

مدل شبیه سازی با مجموعه ی واقعی دیتای کووید ۱۹ مورد آزمایش قرار گرفت: موارد تایید شده کووید ۱۹ گزارش شده برای ایران طی دوره ۱۱ اسفند سال ۱۳۹۹ تا ۱۰ اردیبهشت ۱۴۰۰.

الگوی میزان عفونت (Rt) برای ایران برای دوماه اول (۱۱ اسفند تا ۱۰ اردیبهشت) به صورت زیر تخمین زده شده است:

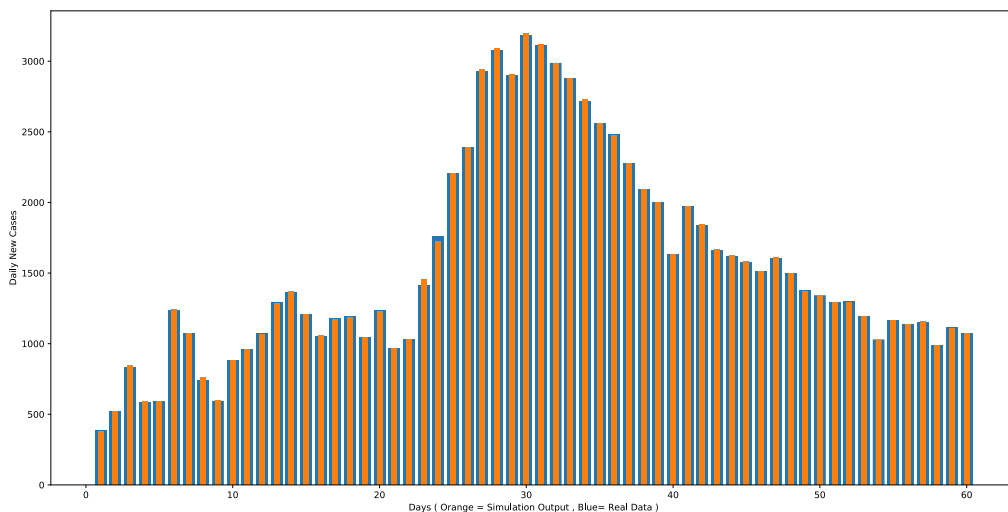
$rr_ir1 = [5.95, 2.25, 3.2, 0.03, 0.86, 4.15, 0.75, 0, 0.3, 2.48, 1.3, 1.3, 1.92, 1.2, 0.55, 0.58, 1.31, 1.2, 0.25, 1.8, 0.33, 1.3, 2, 1.6, 2, 1.24, 1.78, 1.15, 0.85, 1.3, 0.92, 0.8, 0.87, 0.8, 1, 0.76, 0.7, 0.84, 0.68, 0.6, 1.65, 0.65, 0.7, 0.98, 0.8, 0.9, 1.1, 0.85, 0.8, 0.83, 0.77, 1.2, 0.7, 0.585, 1.35, 0.95, 1.1, 0.65, 1.35, 0.92]$

پارامتر μT و $sizeV$ نیز در تنظیمات شبیه سازی تنظیم شده تا انعطاف بیشتری در تناسب با الگوی داده های مشاهده شده داشته باشد.

تنظیمات شبیه سازی برای دوره اول به این صورت بود:

$Simulation(nd = 60, \mu T = 2.8, sizeV = 1.3, n0 = 205)$

نمودار مقایسه داده های موارد جدید مبتلا شده در هر روز به صورت واقعی و شبیه سازی شده در دوره اول شبیه سازی (۲ ماه اول) به شرح زیر است:



که خطوط نارنجی مقدار شبیه سازی شده و خطوط آبی مقدار واقعی است. به وضوح روشن است که شبیه سازی به نتایج واقعی بسیار نزدیک هستند به اصطلاح fit هستند.

در مجموع ۵ مورد شبیه سازی برای بازه های ۲ ماهه مورد بررسی قرار گرفتند که به شرح زیر می باشند:

`simu=Simulation(nd=60,Rt=rr_ir1, muT=2.8,sizeV=0.9,limit=80000000,n0=205) (ir1)`

`simu=Simulation(nd=60,Rt=rr_ir2, muT=2.8,sizeV=1.3,limit=80000000,n0=1073) (ir2)`

`simu=Simulation(nd=60,Rt=rr_ir3, muT=2.8,sizeV=1.3,limit=80000000,n0=2489) (ir3)`

`simu=Simulation(nd=60,Rt=rr_ir4, muT=2.8,sizeV=1.3,limit=80000000,n0=2190) (ir4)`

`simu=Simulation(nd=60,Rt=rr_ir5, muT=2.8,sizeV=1.3,limit=80000000,n0=5960) (ir5)`

همچنین نتایج این شبیه سازی ها در فایل پایتون ضمیمه شده کاملاً قابل اجرا است.