

HAW HAMBURG

MODELLIERUNG DYNAMISCHER SYSTEME

---

# Ausbreitung von Tuberkulose

Ein Populationsmodell

Prof. Dr. Fohl

---

Matthias Nitsche

`matthias.nitsche@haw-hamburg.de`

Hamburg University of Applied Sciences, Department of Computer Science

Berliner Tor 7

20099 Hamburg, Germany

4. September 2017

## Kurzfassung

In dieser Arbeit wird eine Übersicht von mathematischen Analyseverfahren zur Übertragung von Tuberkulose präsentiert. Hierzu zählen die Compartment Modelle wie SEIR und agentenbasierte Modelle (ABM). Tuberkulose ist eine komplexe Erkrankung mit unterschiedlichen Krankheitsstadien. Das Ziel von derartigen Modellierungen ist das identifizieren von optimalen Parametern, sodass ein teil der Realität dargestellt werden kann. Dies führt idealerweise zu neuen Erkenntnissen wie z.B. durch Impfung, Tuberkulose langfristig verhindert werden kann. Auf dem Weg wird der Einfluss von Impfung, Immunerkrankungen, Alter und Lokalität untersucht. Modelle dieser Arbeit wurden mit hilfe von Matlab und Python Mesa realisiert.

# Inhaltsverzeichnis

<b>1</b>	<b>Domäne - Tuberkulose</b>	<b>3</b>
1.1	Krankheitsbild	3
1.2	Behandlung	4
1.3	Faktoren	5
1.4	Ausbreitung von Infektionskrankheiten	5
<b>2</b>	<b>Compartment Modelle zur Ausbreitung von Tuberkulose</b>	<b>6</b>
2.1	Matlab	8
2.2	SEEIR MDR-TB Modell	8
2.3	Clustered SELEIR Modell	11
<b>3</b>	<b>Agentenbasierte Modellierung zur Ausbreitung von Tuberkulose</b>	<b>13</b>
3.1	Modellierung von Tuberkulose in Mesa	14
3.2	Eigenschaften des Modells	17
3.3	Resultate	20
<b>4</b>	<b>Fazit</b>	<b>22</b>

# 1 DOMÄNE - TUBERKULOSE

Tuberkulose ist eine bakterielle Infektionserkrankung die besonders häufig die Lunge befällt. Die WHO gibt Tuberkulose als eine der top 10 Todesursachen der Welt an, mit einer Sterberate von 1.8 Millionen Menschen und 10.4 Millionen Neuinfektionen in 2015. Es ist ein aktives Sustainable Development Goal (SDG) das bis 2030 eingedämpft werden soll. Tuberkulose ist eine Krankheit die in Ländern mit schlechter medizinischer Versorgung und tropischem Klima (Sonne und starker Regen) gesondert auftritt. [World Health Organization, 2016]

## 1.1 KRANKHEITSBILD

Tuberkulose (TB) wird vor allem über die Luft übertragen. Das befallen der Lungen resultiert nicht zwangsweise in einer aktiven Infektion der Krankheit. In Abbildung 1.1 ist der Kreislauf der Tuberkulose Übertragung.

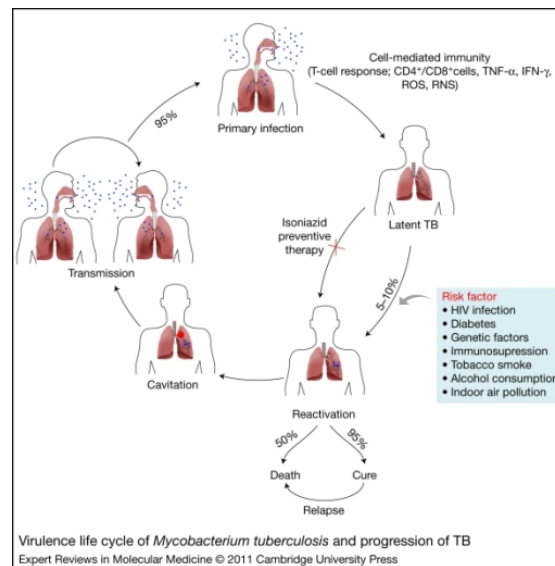


Abbildung 1.1: Tuberkulose Kreislauf in [Kumar et al., 2011]

In kurz, ein Mensch wird über direkten Kontakt von einem anderen über die Luft mit Tuberkulose latent infiziert, hat wesentlich geringere Wahrscheinlichkeit infiziert zu werden und eine hohe Wahrscheinlichkeit bei nicht Medikation zu sterben.

1. Die **Übertragung** findet über die Luft zwischen zwei Menschen statt. Häufig ist deshalb der Ort der Übertragung in Lebensgemeinschaften auf engem Raum, öffentlichen Verkehrsmitteln oder Arbeitsplätzen. Klimaanlage beschleunigen die Verbreitung von Bakterien in der Luft. Bei direktem Kontakt ist die Wahrscheinlichkeit auf Ansteckung bei 95% und es wird eine Hauptinfektion ausgelöst.

2. Die **Hauptinfektion** ist meist ein Befall der Lunge. Verschiedene Immunfaktoren beeinflussen die Geschwindigkeit des Fortschreitens.
3. Betroffene sind ab diesem Schritt **latent infiziert**. Während das Subjekt in der latenten TB Phase ist, tauchen nicht zwangsweise Symptome auf. Die meisten latent infizierten Träger von TB haben in ihrem ganzen Leben keinerlei Auswirkungen der Krankheit. In dem latenten Stadium können Subjekte niemanden anstecken. Einige Faktoren beeinflussen jedoch die Wahrscheinlichkeit auf eine aktive Infektion, sogenannte Risikofaktoren. Diese sind vor allem Immunbedingt wie z.B. HIV, Diabetes, genetische Faktoren, Immunsuppressiver und viele mehr. Subjekte mit derartigen Ausprägungen haben eine Wahrscheinlichkeit von 5-10% die Krankheit als aktive Infektion zu entwickeln. In den Fällen wo eine Präventivtherapie nicht hilft erfolgt meist eine aktive Krankheitsentwicklung.
4. Die **aktive Infektion/Reaktivierung** erfolgt irgendwann nachdem ein Subjekt latent infiziert wurde. Ab diesem Punkt gibt es nur zwei Möglichkeiten, Menschen die eine Heilung bekommen werden zu 95% wieder gesund. 50% die keine Heilung bekommen sterben und 50% bessern sich von selbst. Ein Rückfall führt meistens zum Tode. Infizierte sind ab jetzt Überträger der Krankheit. Menschen die von TB geheilt werden gehen zurück ins latente Stadium. Eine vollständige Heilung ist nicht mehr möglich.

An diesem Model wird relativ schnell klar warum TB so gefährlich ist. Betroffene können nie wieder von der Krankheit regenerieren und Immunfaktoren, die besonders im Alter schlechter werden, begünstigen den Ausbruch der Krankheit. Zudem kommt das TB in zwei Formen auftritt. TB als herkömmliche, behandelbare Erkrankung und Multi-Drug-Resistant TB (MDR-TB), bei der TB resistent gegenüber den derzeitig stärksten Medikationen ist. Tatsächlich ist nach der WHO die MDR-TB Erkrankungen am steigen.

## 1.2 BEHANDLUNG

Die Behandlung von TB kann vielschichtig erfolgen. Häufig eignet sich eine Impfung direkt bei der Geburt, aber auch später bei bisheriger nicht Erkrankung. Besondere Maßnahmen müssen bei TB und HIV berücksichtigt werden. Die Inzidenzrate von TB ist fast endemisch bei Trägern von HIV. Da HIV über Blut oder Körpersekrete weitergegeben wird, sind vor allem Länder betroffen wo es einen hohen Teil an HIV infizierten gibt, schlechte Aufklärung und Mittel zur Verhütung. In Regionen wo TB bereits als Epidemie ausgebrochen ist, muss die Bevölkerung aufgeklärt werden, sodass gegenschritte eingeleitet werden können. Zentren für Medikation, Behandlungen von Erkrankten und Isolation von der gesunden Bevölkerung. Isolation ist hier besonders wichtig, sowohl für nicht erkrankte als auch erkrankte. Betroffene Regionen müssen gewarnt werden sodass erkrankte selbstständig Hilfe suchen und nicht Betroffene nur für die wichtigsten Fälle in die Öffentlichkeit gehen. Individuen können außerdem durch Gesundheitsfaktoren wie Nahrung und geringen Konsum von Zigaretten oder Alkohol höhere Wahrscheinlichkeiten auf nicht Infektion haben. [World Health Organization, 2017]

### 1.3 FAKTOREN

Es gibt einige Faktoren die besonders gut die Ausbreitung von Tuberkulose beschreiben können. Dazu gehören

- Die Abdenkung von Impfung in der Bevölkerung
- Immunerkrankungen wie HIV oder Diabetes
- Alter
- Alkohol und Zigarettenkonsum
- Lebensraum in Quadratmeter pro Individuum
- Transport (Flugzeug oder Schiff) und Öffentliche Verkehrsmittel (Bus oder Bahn)
- Regen und Hitze (tropisches Klima)
- Gruppierungen von Menschen (z.B. durchschnittliche Größe und Menge an Unternehmen, durchschnittliche Familiengröße, öffentliche Plätze)
- TB resistent gegen Medikation (MDR-TB)

Häufig lassen sich diese Faktoren aus Berichten der WHO und weiteren extrahieren. So ist es möglich eine Bevölkerung anhand dieser Faktoren als Durchschnitt der Bevölkerung zu modellieren. Das erkennen und aufdecken dieser Variablen ist enorm wichtig. Die aufgelisteten Faktoren sind die wichtigsten und offensichtlichsten Faktoren. Es gibt dutzende weitere die auf den ersten Blick nicht erkenntlich sind und eventuell hervorragende Prädiktoren für TB sind. [World Health Organization, 2016]

### 1.4 AUSBREITUNG VON INFEKTIONSKRANKHEITEN

Die Modellierung der Ausbreitung von Tuberkulose gehört zu einer höheren Kategorie von Modellen der Ausbreitung von Infektionskrankheiten. Infektionskrankheiten sind nach der Epidemiologie, Erkrankungen die durch einen Erreger hervorgerufen werden, die meist parasitär in einem Wirt Leben und transient (kurze Dauer, hohe Infektion) oder persistent (lange Dauer, niedrige Infektion) auftreten. Sie gehen in der Regel mit starken septischen Symptomen einher wenn eine aktive Phase eintritt.

Bei der Modellierung der Ausbreitung von Infektionskrankheiten wird in drei Modellen unterschieden. Diese Modelle sind nicht zwingend diskret, da es viele Überlappungen geben kann.

1. **Compartment Modelle** sind solche in denen Individuen einer Bevölkerung in Klassen eingeordnet werden. Die Standardklassen beinhalten Susceptibles (Empfänger), Exposed (latent infizierte), Infected (Infiziert) und Recovered (Erholte).

2. **Statistische Modelle** schauen sich historische Daten von Bevölkerungen an wo TB aufgetreten ist. Mithilfe dieser Daten können z.B. Prädiktionsmodelle entwickelt werden, um vorhersagen, wo die nächste TB Epidemie ausbricht.
3. **Agentenbasierte Modellierung** modelliert jedes einzelne Individuum aufgrund von Bevölkerungsparametern und lässt diese auf Grids miteinander interagieren. Der Vorteil zu reinen Compartment Modellen ist das die Lokalität der einzelnen Agenten eine Rolle spielt. Der Nachteil, dass große Simulation enorme Rechenleistung erfordern.

Im Folgenden werden die Compartment Modelle und die agentenbasierte Modellierung am Beispiel von Tuberkulose beschrieben und erläutert.

## 2 COMPARTMENT MODELLE ZUR AUSBREITUNG VON TUBERKULOSE

Compartment Modelle werden über Differentialgleichungssysteme beschrieben. Das einfachste dieser Modelle ist das Susceptible-Infected-Recovered (SIR) Modell. Die Bevölkerungspopulation wird durch  $N = S + I + R$  zu einem Zeitpunkt  $t$  evaluiert. Diese Werte werden durch das Differentialgleichungssystem 2.1 zu einem Zeitpunkt  $t$ , über ein Zeitintervall mit einer fixen Schrittweite approximiert. [Hethcote, 2000]

$$\begin{aligned}\frac{dS}{dt} &= \mu N - \mu S - \beta \frac{I}{N} S \\ \frac{dI}{dt} &= \beta \frac{I}{N} S - (\gamma + \mu) I \\ \frac{dR}{dt} &= \gamma I - \mu R\end{aligned}\tag{2.1}$$

Die Parameter des SIR Modells sind  $\mu$  die Sterberate,  $\beta$  die Infektionsrate und  $\gamma$  die Erholungsrate. Die Sterberate ist die generelle Sterberate in der Bevölkerung. Für ein realeres, kontinuierliches Modell mit Zyklen, bräuchte dieses Modell auch einen entsprechenden Geburtenratenparameter. Das vorliegende einfache *SIR* Modell ist ein

- gewöhnliches Differentialgleichungssystem
- erster Ordnung
- welches nicht-linear ist, durch  $\frac{I}{N} S$ , da  $S$  nicht mit einer Konstanten multipliziert wird. Dadurch sind keine einfachen analytischen Lösungen möglich. Allerdings konnte [Harko et al., 2014] eine exakte Lösung für das *SIR* Modell präsentieren.

Das SIR Modell kann leicht durch latent infizierte (Exposed) zum SEIR Modell erweitert werden. Das Differentialgleichungssystem 2.2 bekommt eine zusätzliche Klasse  $E$  mit dem korrespondierenden Zunahmeparameter  $\alpha$ , wobei  $N = S + E + I + R$  gilt. Auch

hier nimmt  $E$  um  $(\alpha + \mu)$  in  $I$  ab und steigt um  $\gamma I$ , die Abnahme der infizierten, an. [Hethcote, 2000]

$$\begin{aligned}
\frac{dS}{dt} &= \mu N - \mu S - \beta \frac{I}{N} S \\
\frac{dE}{dt} &= \beta \frac{I}{N} S - (\alpha + \mu) E \\
\frac{dI}{dt} &= \alpha E - (\gamma + \mu) I \\
\frac{dR}{dt} &= \gamma I - \mu R
\end{aligned} \tag{2.2}$$

Die Übergangparameter mit den zugehörigen Compartments stellen Übergangswahrscheinlichkeiten zwischen den Compartments dar. Sollte es z.B. möglich sein direkt von  $S$  zu  $I$  zu laufen, kann ein entsprechender Zunahmeparameter eingefügt werden. Das letzte und vollständigste Differentialgleichungssystem MSEIRS in 2.3 stellt eine relativ nahe Modellierung von Tuberkulose dar.

$$\begin{aligned}
\frac{dM}{dt} &= N - \sigma M - \mu M \\
\frac{dS}{dt} &= \sigma M - \mu S - \rho(t) \beta \frac{I}{N} S \\
\frac{dE}{dt} &= \beta \frac{I}{N} S - \rho(t) (\alpha + \mu) E + \lambda R \\
\frac{dI}{dt} &= \rho(t) \alpha E - (\gamma + \mu) I \\
\frac{dR}{dt} &= \gamma I - (\mu + \lambda) R
\end{aligned} \tag{2.3}$$

In diesem Differentialgleichungssystem wurden drei zusätzliche Modifikationen unternommen.

1. Es wurde eine Klasse von geimpften  $M$  bei der Geburt eingefügt die mit einer Wahrscheinlichkeit von  $\sigma$  nach  $S$  gehen können.
2. Es gibt einen direkten Weg von  $R$  zu  $E$  mit einer Wahrscheinlichkeit  $\lambda$ .
3. Es wurde ein variierender Zeitparameter  $\rho(t)$  eingefügt um saisonale Schwankungen zu berücksichtigen. Dieser könnte z.B. 1/365 für 1 Tag pro Berechnung darstellen. Dies müsste im Fall der Tuberkulose, deutlich genauer modelliert werden, da diese vor allem in heißen regnerischen Monaten auftritt.

Für ein vollständiges modernes Modell fehlen noch altersbasierte Parameter, Einfluss von Immunerkrankungen, eine Wachstumsrate, der Einfluss von MDR-TB, Lokalität über z.B. Clusterings, der Einfluss von Therapien und einige mehr. [Hethcote, 2000]

## 2.1 MATLAB

Zur Modellierung von den folgenden Compartment Modellen wurde Matlab verwendet. Insbesondere der *ode45* solver in Kombination mit den Plotting Möglichkeiten.

```
intervall = 0:1:60;  
y0        = [S E L Es I R];  
opts      = odeset('RelTol', 1e-5);  
[t, y]     = ode45(@(t,y) model(t, y), intervall, y0, opts);
```

Wobei  $t$  der aktuelle Zeitschritt und  $y_0$  die zu integrierenden Werte zum Zeitpunkt  $t = 0$  der Differentialgleichungen sind. Im folgenden wurden Experimente mit Daten über Tuberkulose von dem US Ministerium für Gesundheit entnommen. Zum einen die Population  $N$  von 1953-2015 und zum anderen die Inzidenzraten  $I$  von 1953-2015.

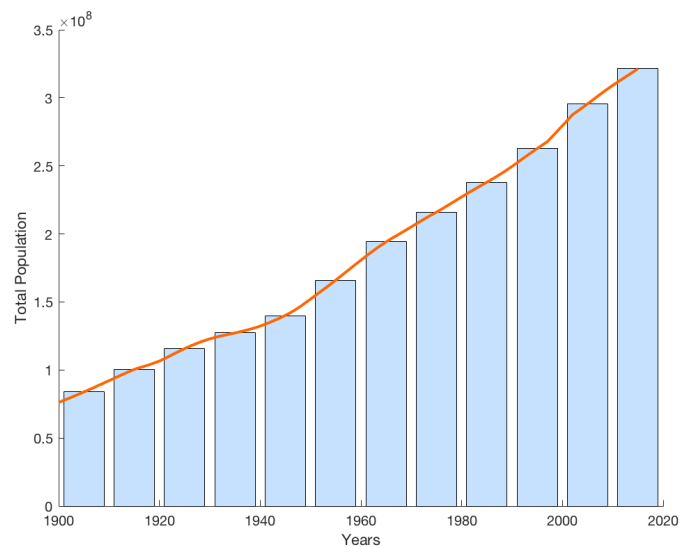


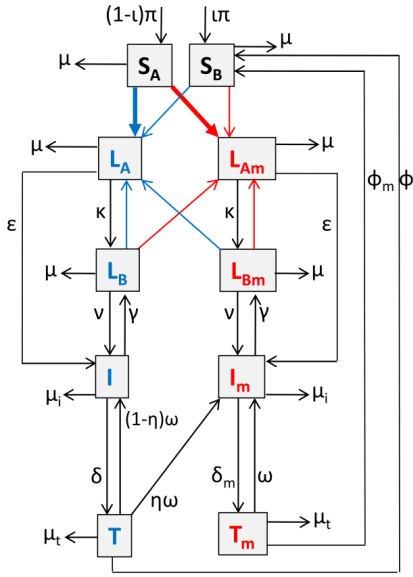
Abbildung 2.1: US Population zwischen 1953-2015

Die US Populationsentwicklung ist in Abbild 2.1 zu sehen. Die durchschnittliche Rate ist der orangenen Linie zu entnehmen. Es ist ein stetiges Wachstum auf ein level zugehend von 320 Millionen im Jahr 2015.

## 2.2 SEEIR MDR-TB MODELL

Im ersten Experiment wurde versucht das Modell von [Trauer et al., 2014] in Matlab zu implementieren. Das reproduzieren ist aufgrund der menge an Differentialgleichungen und Parametern nicht trivial. Das System inkorporiert MDR-TB und ist in zwei größere Gruppen eingeteilt. Das Modell von [Trauer et al., 2014] ist in Abbildung 2.2 zu sehen.





Das Modell ist in zwei Systeme aufgeteilt mit kleinen Übergängen zwischen diesen Systemen. Die Differentialgleichungssysteme beschreiben ein klassisches SEIR Modell mit einer latenten Phase mehr.  $L_A$  ist die erste latente Phase wo die Wahrscheinlichkeit auf Erkrankung geringer ist als bei der aktiveren Phase  $L_B$ . Die Autoren haben herausgefunden das eine Modellierung so besser abzubilden ist. Das Differentialgleichungssystem was aus  $S_A$  folgt stellt das Transfer Diagramm der herkömmlichen TB dar.  $S_B$  hingegen stellt die MDR-TB dar mit höheren Resistenzen. Es ist fest zu stellen das  $\mu$  die Sterberate ist.  $\phi$  hingegen beschreibt den Übergang von einer Genesung, direkt in die anfällige Gruppe der MDR-TB. Ferner gibt es einen Rückfallfaktor  $\eta\omega$  nach Genesung von  $T$  zu  $I_m$  direkt in die MDR-TB.

Das obige Transfer Diagramm wird als Differentialgleichungssystem realisiert und sieht wie folgt aus

$$\begin{aligned}
 \frac{dL_A}{dt} &= \lambda S_A + \lambda_d(S_B + L_B + L_{Nm}) - (\epsilon + \kappa + \mu)L_A \\
 \frac{dL_{Am}}{dt} &= \lambda_m S_A + \lambda_{dm}(S_B + L_B + L_{Nm}) - (\epsilon + \kappa + \mu)L_{Am} \\
 \frac{dL_B}{dt} &= \kappa L_A + \lambda I - (\lambda_d + \lambda_{dm} + \nu + \mu)L_B \\
 \frac{dL_{Bm}}{dt} &= \kappa L_{Am} + \lambda I_m - (\lambda_d + \lambda_{dm} + \nu + \mu)L_{Bm} \\
 \frac{dI}{dt} &= \epsilon L_A + \nu L_B + (1 - \eta)\omega T - (\gamma + \delta + \mu_i)I \\
 \frac{dI_m}{dt} &= \epsilon L_{Am} + \nu L_{Bm} + \eta\omega T + \omega T_m - (\gamma + \delta_m + \mu_i)I_m \\
 \frac{dT}{dt} &= \delta I - (\varphi + \omega + \mu_t)T \\
 \frac{dT_m}{dt} &= \delta_m I_m - (\varphi_m + \omega + \mu_t)T_m
 \end{aligned} \tag{2.4}$$

Wie zuvor auch ist die Populationszahl  $N = S_A + S_B + L_A + L_B + L_{Am} + L_{Bm} + I + I_m + T + T_m$  errechnet durch alle Variablen zum Zeitpunkt  $t$ . Die spezifischen  $\lambda$  Parameter kontrollieren die Übergänge zwischen den  $S$  und  $L$  Klassen, also quasi die latente Infektion.

$$\lambda = \beta\rho(I + oT))N$$

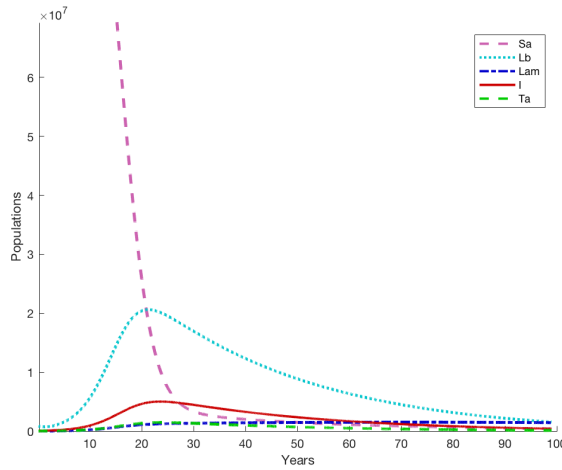
$$\lambda_d = \beta_m\rho(I_m + oT_m))/N$$

$$\lambda_{dm} = \chi\beta_m\rho(I_m + oT_m))/N$$

Die eingestellten Parameter für dieses Modell sind im folgenden aufgelistet und nahe an [Trauer et al., 2014] angelehnt.

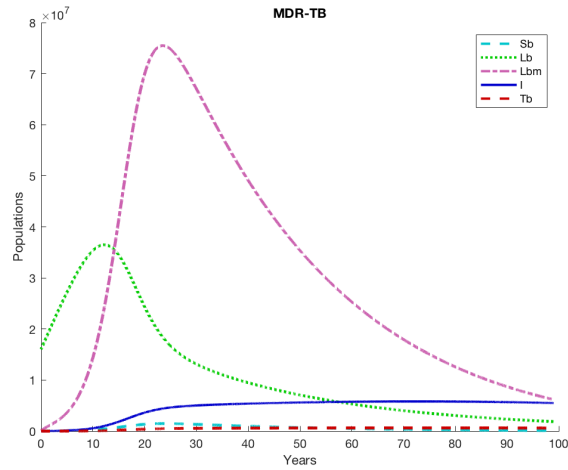
Variable	Wert	Beschreibung
Infektionsparameter		
$\epsilon$	0.129	Early progression - Diel et al. (2011)
$\kappa$	0.821	Transition to late latency
$\nu$	0.075	Reactivation - Blower et al. (1995)
$\gamma$	0.63	Spontaneous Recovery - Tiemersma et al. (2011)
$\mu_i$	0.37	TB death rate
$\mu_t$	$0.5 \cdot \mu_i$	Treated TB death rate - Harries et al. (2001)
$\eta$	0.035	Amplification - Cox et al. (2007)
$o$	0.21	Treatment modification of infectiousness - Fitzwater et al. (2010)
$\chi$	0.49	Partial immunity - Colditz et al. (1994)
$\phi$	2	Drug susceptible treatment rate per year - WHO (2010)
$\phi_m$	0.5	MDR-TB treatment rate per year
Populationsparameter		
$\pi$	0.025	Birth rate during sensivity Analysis - UNDESA
$\mu$	0.016	TB-free mortality - WHO (2013b)
$\rho$	0.35	Infectious propertion - WHO (2013c)
Kontrollparameter		
$\iota$	0.65	BCG vaccination rate - World Bank (2013)
$\delta$	0.65	Detection rate - WHO (2013c)
$\delta_m$	0	MDR-TB detection
$\omega$	0.25	Default rate - WHO (2013c)
$\beta$	0.25	Effective contact rate
$mdr$	0.06	proportion of cases MDR-TB 4-6%

Die Resultate des Modells sind unscheinbar. Sobald es zwei Populationen gibt mit TB und MDR-TB, überwiegt die Infektionsrate von MDR-TB. Das liegt daran, dass es mehr Eingänge zu dem System des MDR-TB gibt und aus der MDR-TB nur geringe Abläufe sind. Es ist zu beobachten das ein geringer Anteil der Population in einen Heilungsprozess kommt. Durch fehlende Geburtenrate sind zum Ende des berechneten Zeitraums fast alle Individuen durch die Mortalitätsrate  $\mu$  gestorben. Hier fehlt ein System das sich theoretisch auf einem Equilibrium einpendelt bzw. wo die Krankheitsausbreitungen deutlich geringer sind.



Alle Geheilten  $T$  sind direkt in die Klasse  $S_B$  gekommen. Es ist zu beobachten das die empfänglichen enorm schnell abnehmen und zeitgleich latent infizierte übermäßig stark zunehmen. Die Infizierten wiederum nehmen erst stark zu und dann ab. Die Heilung von Tuberkulose läuft gegen 0.

In Abbild 2.2 hingegen ist zu beobachten das die empfänglichen leicht zunehmen und latent infizierte extrem zunehmen. Diese kommen aus der Gruppe ohne MDR-TB. Vor allem die Gruppe der langfristig infizierten nimmt zu. Der Trend aller Kurven nach unten ist durch die Sterberate  $\mu$  zu begründen. Die infizierten wiederum sind deutlich geringer, nehmen aber über Zeit zu.



Durch diese Art der Modellierung kann durch Variationen der Parameter geschaut werden was effiziente Strategien wären um TB in dieser Simulation einzudämmen. Dies ist in diesem Fall leider nicht gelungen. Die Approximation von optimalen Parametern ist nicht analytisch lösbar und erfordert einiges an mathematischen Anstrengungen.

### 2.3 CLUSTERED SELEIR MODELL

Das Clustered SELEIR Modell aus [Aparicio and Castillo-Chavez, 2009] arbeitet mit zwei Populationen  $N_1$  und  $N_2$ . Das Transfer Diagramm 2.2 zeigt einige Ähnlichkeiten zu den bisherigen Compartment Modellen, allerdings deutlich komplexer. Um alles schwieriger zu machen sind  $U$  die Empfänger ( $S$ ),  $A$  sind die infizierten ( $I$ ) und  $L$  ist eine neue Klasse von latent infizierten Individuen. Wie zuvor gilt der Kreislauf  $U \rightarrow E \rightarrow A \rightarrow R$ . Hinzu kommt, dass die latente Phase fein granularer dargestellt ist. Anstelle von dem

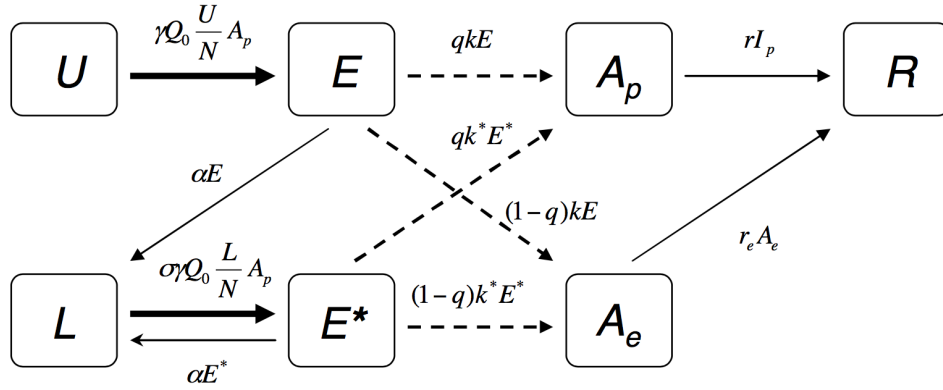


Abbildung 2.2: Clustered Tuberkulose in [Aparicio and Castillo-Chavez, 2009]

direkten Übergang von latent infizierten  $E$  zu infizierten gibt es drei Klassen.  $E$  kann standardmäßig zur Klasse der Infizierten wandern. Hier gibt es zwei Klassen, der Befall der Lunge “pulmonale TB”  $A$  oder der Befall anderer Organe “extrapulmonale TB”  $A_e$ .  $E$  kann außerdem in die Klasse  $L$ , lange latente Infizierte die eine wesentlich geringere Chance auf den Ausbruch haben. Diejenigen die  $L$  mit einer Wahrscheinlichkeit  $\sigma \gamma Q_0 \frac{L}{N} A_p$  verlassen sind in einer hochgradig exponierten Variante von  $E$ , die höhere Wahrscheinlichkeit auf aktiven Ausbruch hat. Wie zuvor werden die Cluster als zwei Populationen in zwei Differentialgleichungssystemen dargestellt. Das Transferdiagramm wird dupliziert und eine explizite Abhängigkeit aufgebaut zwischen diesen erzeugt. Eine generelle Population  $N_1$  bei der es zufällig mit einem sehr geringen Wert Neuinfektionen für TB gibt. Eine Population mit clustern von erkrankten  $N_2$  die bei jeder Neuinfektion in  $N_1$  dahin bewegt wird. Es gibt dann eine Kontaktwahrscheinlichkeit zwischen  $N_1$  und  $N_2$ . Das volle Differentialgleichungssystem kann in [Aparicio and Castillo-Chavez, 2009] nachvollzogen werden und würde den Rahmen hier sprengen.

In dem Experiment mit dem Clustered SELEIR Modell ist herauszufinden, wie gut die generellen Populationsparameter für Tuberkulose die Kurve von realen Daten approximieren kann. Das System wurde mit den Bevölkerungsparametern von Tuberkulose aus [Trauer et al., 2014] befüllt. Diese Parameter wurden kombiniert mit den Daten der amerikanischen Bevölkerung. In Abbildung 2.3 sehen wir die Vorhersage des Modells über den zeithorizont in Blau im Vergleich zu der realen Entwicklung in Grün.

Es wird relativ klar, dass der generelle Trend getroffen wird, die Vorhersagelinie aber nicht der realen folgt. Varianzen in den realen Daten führen nicht zu Varianzen in den syntetisch erzeugten Daten des Modells. Probleme die hier vor allem auftreten sind die Menge der Parameter. Diese unschärfe sorgt dafür dass es nur durch komplexe analytische Verfahren möglich ist zu bestimmen, inwiefern Parameter angepasst werden müssen. Manche reale

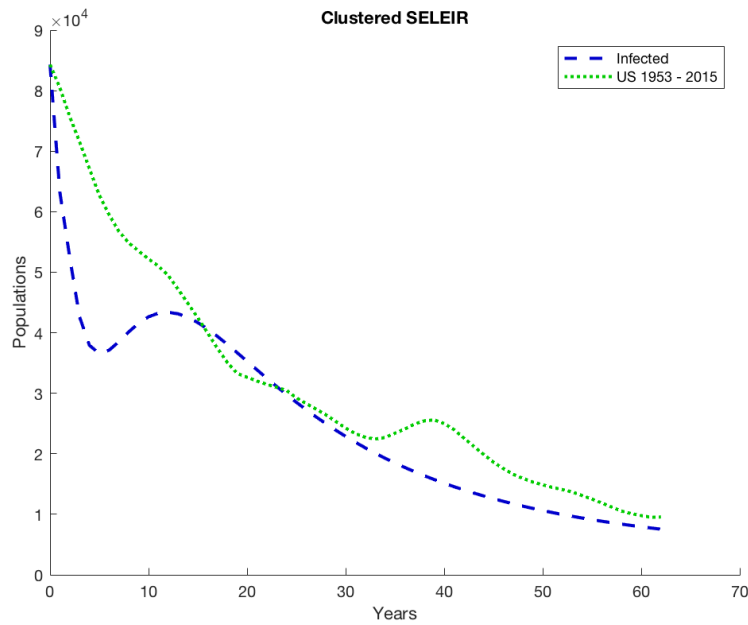


Abbildung 2.3: Clustered SELEIR Modell

Parameter sind zudem nicht bekannt und Parameter sind auch nur als Durchschnitt in diesem Zeitraum statisch.

### 3 AGENTENBASIERTE MODELLIERUNG ZUR AUSBREITUNG VON TUBERKULOSE

Mit der agentenbasierte Modellierung (ABM) werden Individuen oder Kollektive simuliert und deren Aktionen in einer bestimmten Sichtweise analysiert. Das Ziel ist es neue Erkenntnisse aus den Interaktionsmustern zu entwickeln. Im Vergleich zu den Compartment Modellen ist es bei ABM möglich auch die Lokalität eines Agenten einfließen zu lassen. Ferner lassen sich Agenten auf Schichten platzieren, wobei jede Schicht für unterschiedliche Modellierungsaspekte steht. Es kann z.B. eine Schicht für die physikalischen Interaktionsmuster und Wege existieren und eine andere um die soziale Zusammensetzung zu modellieren. Jede Abstraktionsebene kann getrennt von den anderen Schichten hinzugefügt werden, für die entsprechende Rechenleistung.

Im Folgenden wird ein Modell vorgestellt, dass die Ausbreitung von Tuberkulose als ABM realisiert. Hierbei laufen Agenten auf einem Grid per Random Walk. Nur bei direktem Kontakt auf demselben oder umliegenden Grid kann eine Infektion erfolgen. Verschiedenste Bevölkerungsparameter und Übergangswahrscheinlichkeiten wie in Compartment Modellen kommen zum Einsatz. Die Simulation wurde in Python Mesa, einem ABM

Framework, geschrieben.

### 3.1 MODELLIERUNG VON TUBERKULOSE IN MESA

Mesa ist ein ABM Framework in Python [Kazil, 2014]. Es existieren unterschiedliche Funktionalitäten um Agenten auf einem oder mehreren Feldern zu platzieren und mit diesen über einen Zeitraum zu planen. Im folgenden ein paar Erläuterungen zu dem Mesa Framework und der Modellierung von Tuberkulose.

Die *AgentDistribution* hält alle Populationsparameter. Jeder neue Agent wird aus dieser Klasse mit Stichprobenparametern befüllt. Unter anderem kommt hier das Konzept von den Compartment Modellen - *SLIR* - zum Einsatz um Agenten in einzelne Klassen einteilen zu können. Bei den Ansätzen wurde sich vor allem an dem Modell von [Perez and Dragicevic, 2009] orientiert.

---

```
class AgentDistribution():  
  
    def __init__(self, S=100., L=10., I=1., R=0.):  
        pass
```

---

Agenten müssen auf einem Grid platziert werden können. Zu aller erst wird eine neue Stichprobe generiert und ein neuer Agent mit diesen Daten erzeugt. Jeder Agent wird dem Mesa Scheduler hinzugefügt und auf dem Grid platziert. In jedem Update oder Tick wird die Simulation um eine Zeiteinheit  $t$  weitergeführt. Der scheduler geht Round Robin alle Agenten durch und kann optimal Populationsdaten einsammeln. Ein Agent bewegt sich aufgrund seiner zu implementierenden *step* Schnittstelle. Dieser *step* wird an alle Agenten propagiert, die dann ihre individuellen Handlungen durchführen.

---

```
from Mesa import Model  
from grid import Grid  
  
class DiseaseModel(Model, Grid):  
  
    def __init__(self, agent_dist, width, height):  
        Grid.__init__(self, width, height)  
  
        self.number_of_agents = agent_dist.N  
        self.time_unit = agent_dist.time_unit  
        self.agent_dist = agent_dist  
        self.schedule = RandomActivation(self)  
        for i in range(self.number_of_agents):  
            self.create_agent(unique_id=i)  
  
    def create_agent(self, unique_id=None):  
        attributes = self.agent_dist.generate_sample()  
        agent = Human(unique_id, self, attributes)  
        self.schedule.add(agent)  
        self.place_agent(agent)  
  
    def step(self):  
        self.schedule.step()
```

---

Der Mesa Agent besitzt alle nötigen Eigenschaften um mit anderen Agenten zu kommunizieren und wichtige Lokalkriterien abzufragen. Ein spezieller *AAgent* besitzt die Schnittstelle, um beliebige Interaktionen mit dem Mesa Grid und Scheduler zu machen wie z.B. der Random Walk. Der *TBAgent* besitzt dann alle Eigenschaften um speziellere Aktionen bezüglich der Tuberkulose aufzunehmen. Ein Mensch hat prinzipiell nichts mit TB zutun, weshalb dieser dann eine Unterkategorie des *TBAgent* darstellt. Der Mensch kann sich Reproduzieren, Sterben, Altern, TB entwickeln/anstecken, impfen, behandeln lassen und sich bewegen.

---

```
import tuberculosis as tb

from mesa import Agent

class AAgent(Agent):

    def __init__(self, unique_id, model):
        Agent.__init__(self, unique_id, model)

class TBAgent(AAgent):

    def __init__(self, unique_id, model):
        AAgent.__init__(self, unique_id, model)

    def susceptible(self):
        return tb.susceptible(self.state)

    def infected(self):
        return tb.infected(self.state)

class Human(TBAgent):

    def __init__(self, unique_id, model, attrs):
        TBAgent.__init__(self, unique_id, model)

        self.age = attrs['age']
        self.birth = attrs['birth']
        # ...

    def step(self):
        self.move()
        self.aging()
        self.reproduce()
        state = tb.next_state(self)
        self.update_state(state)
        self.chance_of_death()
```

---

Das Tuberkulose Modell beinhaltet alle Parameter zur Tuberkulose und die Berechnung der Infektions- und Übergangswahrscheinlichkeiten. Jeder *Human* Agent muss daher bestimmte Attribute wie Alter, HIV positiv/negativ, Impfung, weitere Immunerkrankungen, TB Behandlung/Therapie und die Lebenserwartung initialisiert bekommen. Es müs-

sen weitere Schnittstellen erfüllt sein um die Lokalität des Agenten und der herumliegenden Agenten abzufragen.

---

```
def next_state(h):
    state = h.state

    if h.vaccinated:
        return R

    aging_chance = aging(h.age, h.le)

    if state == S:
        if infect(h.inf_cm(), h.inf_nb()):
            state = SState(h.hiv, h.immune_diseases, aging_chance)

    elif state == L:
        state = LState(h.treatment, h.hiv, h.immune_diseases, aging_chance)

    elif state == I:
        state = IState(h.treatment, h.hiv, h.immune_diseases, aging_chance)

    elif state == R:
        if infect(h.inf_cm(), h.inf_nb()):
            state = RState(h.treatment, aging_chance)

    else:
        raise Exception("'state={}' must be in {}".format(state, ", ".join(
            STATES)))

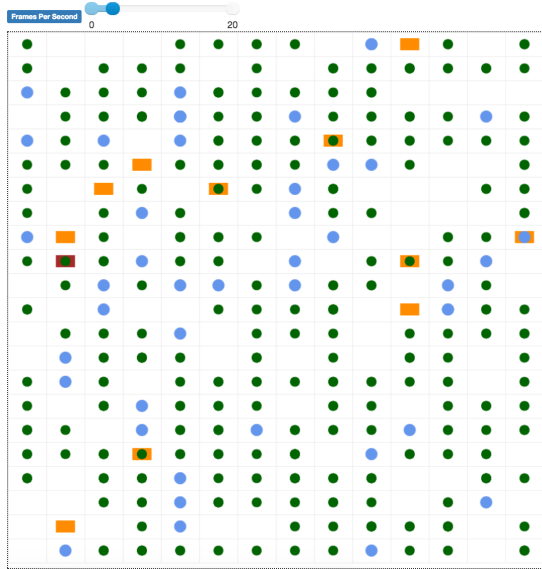
    return state
```

---

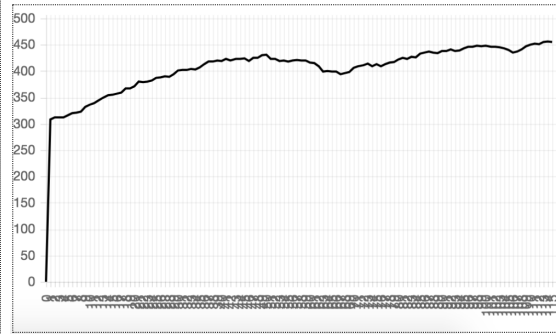
Impfung führt immer zu  $R$ , nur im  $S$  oder  $R$  Status kann ein Agent infiziert werden. Diese Wahrscheinlichkeit wird durch *infect* berechnet, wodurch überhaupt determiniert wird ob es eine Ansteckungsgefahr gibt. Die Ansteckungsgefahr wird durch die umliegenden, aktiv infizierten Agenten berechnet. Nicht jeder direkte Kontakt führt zu einem neuen Infekt. Danach wird berechnet, aufgrund der Agentenparameter, wie wahrscheinlich ein Agent in das latente oder aktive TB Stadium läuft.

Zu guter letzt kann die Simulation über ein Backend Server abgerufen werden. Hierbei bindet Mesa sich an eine konfigurierte IP, standardmäßig 127.0.0.1 auf einem festgelegten Port z.B. 9000. Über z.B. einen Browser kann dann über <http://127.0.0.1:9000> auf die Simulations GUI zugegriffen werden. In Abbildung 3.1a ist das Simulationsgrid zu erkennen, wobei  $S$ , rund und Blau,  $R$ , rund und Grün,  $L$ , rechteckig und Gelb und  $I$ , rechteckig und Rot, ist. In Abbildung 3.1b ist die Gesamtpopulation zu sehen. Dabei ist auf der Y-Achse die Bevölkerungsanzahl und auf der X-Achse der Zeitpunkt  $t$  zu sehen. Die Populationskomposition, daher  $N = S + L + I + R$  kann in Abbildung 3.2a, für  $L$  und  $I$ , sowie in Abbildung 3.2b, für  $S$  und  $R$  beobachtet werden. Die Simulation ist in Jahren und wurde auf der X-Achse über ca 100 Jahre geführt.

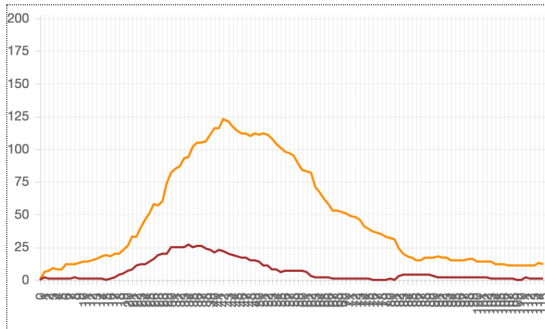




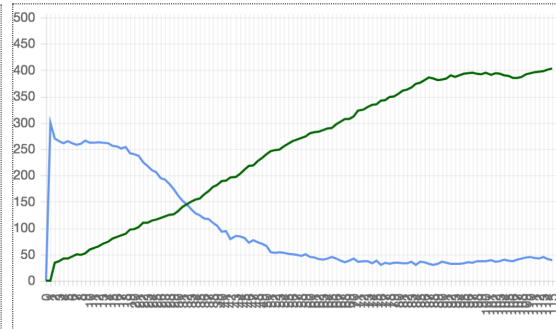
(a) Das Mesa Grid



(b) Die Gesamtpopulation  $N = S + L + I + R$



(a) Die Populationen  $L$  und  $I$



(b) Die Populationen  $S$  und  $R$

### 3.2 EIGENSCHAFTEN DES MODELLS

Das vorliegende ABM Modell in Mesa ist nahe angelehnt an Compartment Modelle. Dabei werden ein paar stochastische Eigenschaften von Distributionen ausgenutzt und mit Effekten kombiniert, die durch direkten Kontakt hervorgerufen werden. Des weiteren werden Parameter vorgestellt und ein paar Experimente beschrieben. Für das ziehen von Stichprobendaten wurden Funktionen aus Python *numpy* verwendet. Zum ziehen von binären Events wurde die Binominal Distribution mit einem wahrscheinlichkeitsparameter genutzt. Die formale Definition der Binominal Distribution ist

$$P(k; n, p) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Die Binominaldistribution eignet sich für Stichproben die aus einem diskreten Raum kommen. Es wird dann eine Distribution mit 0 und 1 aufgestellt, wobei 0 für keine Ausprägung und 1 für eine Ausprägung steht. Diese Ausprägungen werden dann über einen Parameter  $p$  mit Anteilen belegt. Die resultierende Binominaldistribution zieht dann Stichproben mit einer Wahrscheinlichkeit, dass der Wert 0 ist mit  $p$  und die Restwahrscheinlichkeit, dass der Wert 1 ist. Interessanter wird es mit der Gamma Distribution. Das Alter einer Bevölkerung lässt sich beispielsweise wie in Abbildung 3.3 initialisieren.

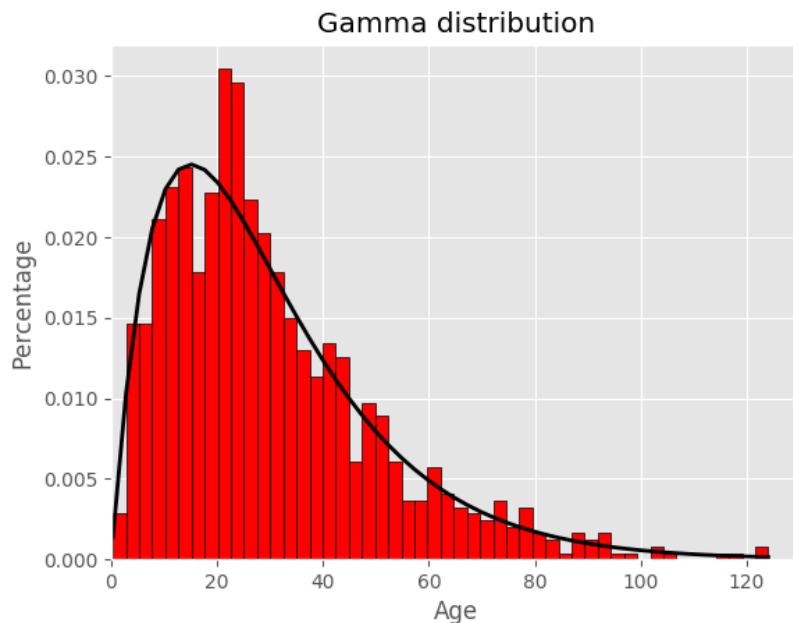


Abbildung 3.3: Gamma Distribution für Altersverteilungen

Die vorliegende Bevölkerung ist im Durchschnitt über 10 und unter 40 und die Lebenserwartung einzelner Individuen sehr gering bis ins hohe Alter. Für die Simulation wurden insgesamt folgende Parameter angenommen

- Alter - Stichproben aus einer Gamma Distribution
- Geburtenrate - in prozentualen Anteilen per Update (entscheidet darüber ob ein Agent, Kinder bekommt)
- Impfung - Ob ein Agent geimpft wurde

- Klassen - Die Klasse  $S$ ,  $L$ ,  $I$  oder  $R$
- Behandlung - Ob ein Agent behandelt wird
- HIV - HIV positiv/negativ eines Agenten
- Immunkrankheiten - Immunerkrankungen eines Agenten

Außerdem wurden folgende Effekte programatisch berücksichtigt

- Neugeborene - Die Geburtenrate entscheidet wieviele Individuen neu erzeugt werden pro Update. Neugeborene sind standardmäßig immer 0 Jahre alt bei der Geburt und bekommen HIV nur dann wenn die Eltern diese Krankheit auch haben. Ob die Eltern geimpft sind gibt einen Ausschlag darüber aus, ob Kinder ebenfalls geimpft werden.
- Tod - Individuen sterben am Alter, daher wenn sie die Lebenserwartung aus der Alterdistribution erreichen, steigt die Chance auf Tod. Ferner, gibt es Sterben durch Tuberkulose, sowie eine zufällige Sterberate die immer eintreten kann.
- Die Umgebung eines Individuums spielt eine Rolle, da mehr Kontakt zu anderen Infizierten das Risiko deutlich erhöht. Es gibt auf dem Grid einen Radius der Infektion, der bestimmt, über wieviele Felder die Infizierten andere anstecken können.
- Die Zeit, wenn ein Individuum in der Klasse  $E$  ist, wird aufgenommen. Die Wahrscheinlichkeit in Klasse  $I$  zu wandern steigt mit fortlaufender Zeit.
- Die Zeit, wenn ein Individuum in der Klasse  $I$  ist, wird aufgenommen. Die Wahrscheinlichkeit in Klasse  $R$  zu wandern oder zu Sterben steigt mit fortlaufender Zeit.

Die Übergangswahrscheinlichkeiten zwischen den Compartments werden anhand einer Markov Kette definiert. Jeder Agent hat nur seine aktuelle Klasse  $S$ ,  $L$ ,  $I$  oder  $R$ . Die Markov Eigenschaft verlangt das

$$\Pr(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \Pr(X_{n+1} = x \mid X_n = x_n) \\ \Pr(X_1 = x_1, \dots, X_n = x_n) > 0$$

Das heißt, der nächste Zustand  $x_{n+1}$ , daher z.B. die Übergangswahrscheinlichkeit das  $x_n = L$  in die Klasse  $x_{n+1} = I$  läuft ist nur bedingt durch den derzeitigen Zustand  $x_n$ . Das heißt konkret das  $x_{n-1}, x_{n-2}, \dots, x_1$  keine Rolle spielen. Nach der Bayesianischen Forderung wollen wir im Prinzip aus dem aktuellen Wissen “prior” und neuen Beobachtungen eine neue Erkenntnis “posterior” finden. Jeder Agent wird dadurch simuliert wieviele der anderen Agenten sich in den einzelnen Klassen befinden. Durch die individuellen Ausprägungen eines Agenten und durch die die Übergangswahrscheinlichkeiten

der oben beschriebenen Markov Kette, wird die Wahrscheinlichkeit vom Übergang in ein anderes Compartment berechnet.

$$M = \begin{bmatrix} .95 & .05 & .0 & .0 \\ .0 & .9 & .1 & .0 \\ .0 & .1 & .7 & .2 \\ .0 & .45 & .05 & .5 \end{bmatrix} \quad \begin{array}{l} \text{S - Empfänger} \\ \text{L - latent Infizierte} \\ \text{I - Infizierte} \\ \text{R - Geheilte} \end{array}$$

Die Markov Kette für Übergangswahrscheinlichkeiten in Kombination mit den folgenden Parametern stellt das komplette Modell dar.

Variable	Wert	Beschreibung
Distributionen und Initialisierung		
$S$	100	Empfänger
$L$	10	Latent Infizierte
$I$	1	Infizierte
$R$	0	Geheilte
Alter	$Gamma(k = 2.0, \theta = 15, N = N)$	Alter, Formparameter $k$ und Skalierungsparameter $\theta$ über die Population $N$
Lebenserwartung	80	Normalverteilung um 80 herum mit Varianz von 10
Geimpfte	0.1	Geimpfte Teile der Bevölkerung
Geburtenrate	0.013	Wahrscheinlichkeit auf Neugeburt pro Tick
HIV	0.1	Wahrscheinlichkeit auf HIV bei Initialisierung und Neugeburt
Behandlung	0.3	Wahrscheinlichkeit auf HIV Behandlung bei Erkrankung
Immunkrankheiten	0.2	Wahrscheinlichkeit auf Immunerkrankung bei Initialisierung oder Neugeburt
Lokalitätsparameter		
Infektionsradius	5	Reichweite der Ansteckungsgefahr sind 5 Felder
Bewegungsweite	1	Schritt pro Agent pro Tick.
Moore + Center	<i>True</i>	Diagonale zählen als Nachbarn + mehrere Agenten auf einem Platz
Infektionsrisiko Center	0.3	Wahrscheinlichkeit das eine Infektion auftritt wenn Infizierter auf demselben Feld
Infektionsrisiko Nachbar	0.1	Wahrscheinlichkeit das eine Infektion auftritt wenn Infizierter auf einem benachbarten Feld
Risikoparameter		
HIV Risiko	0.4	Wahrscheinlichkeit das eine Klasse in die nächste wandert ist ehört bei HIV
Immunkrankheiten Risiko	0.1	Wahrscheinlichkeit das eine Klasse in die nächste wandert ist ehört bei Immunkrankheiten
Behandlung Chance	0.4	Wahrscheinlichkeit das Infizierter geheilt wird ist erhöht
Alter Risiko	$\frac{\text{Alter}}{\text{Lebenserwartung}} \cdot 0.001$	Wahrscheinlichkeit das eine Klasse in die nächste wandert ist ehört bei höherem Alter

### 3.3 RESULTATE

Mit obigen Konfigurationsparametern ist es nun spannend zu beobachten, ob es eine Reihe von Faktoren, gibt die Tuberkulose eindämmen können. Es konnte beobachtet werden, wenn die Impfung bei Neugeborenen der Population deutlich erhöht wird, mit z.B. 60% aller Neugeborenen, verschwindet TB über einen langen Zeitraum vollständig. Da Infizierte keinerlei Chance auf Heilung besitzen, ist die sinnvollste Strategie bei der nächsten Generation anzufangen. Interessant ist auch das durch das einführen von MDR-TB, über einen langen Zeitraum alle Menschen gestorben sind, da diese durch die Behandlung von Medikamenten nicht verschwindet. Die Variation von den Lokalitätsparametern hat auch einen großen Einfluss. Es stellte sich heraus das nur größere Radien  $> 4$  Felder für eine TB Epidemie sorgen konnten. Kleinere Cluster wurden ebenfalls gebildet. Teile der Bevölkerung waren noch nicht infiziert, während andere Teile schon viele Infizierte vorweisen konnte. Langfristig kann ein Stabiles Maß gefunden werden, wenn allein die geimpften Eltern, ihre Kinder ebenfalls impfen wie in den unteren Abbildungen zu sehen ist.

Die Zeitachse ist in kleineren Schritten über 100 Jahre abgebildet. Die Geimpften füllen die  $R$  Klasse. Eine höhere Bevölkerung führt zu höheren Anteilen an latenter TB mit

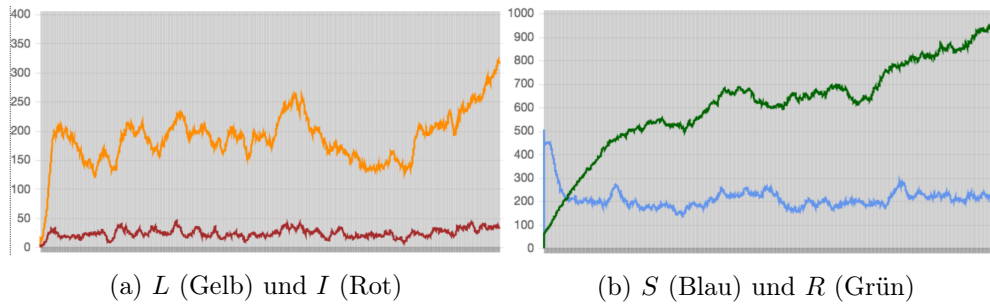


Abbildung 3.5: Gesamtpopulation  $N = S + L + I + R$

periodischen Kurven. Aktiv Infizierte sind relativ gering und gehen periodisch in Schritten rauf und runter.

## 4 FAZIT

In dieser Arbeit wurde gezeigt, wie agentenbasierte Modellierung und Compartment Modelle genutzt werden können, um die Ausbreitung von Infektionskrankheiten wie Tuberkulose zu simulieren. Idealerweise kommen neue Erkenntnisse ans Licht, wie die Übertragung von Tuberkulose verhindert werden kann.

Die beiden Compartment Modelle haben nur geringfügig brauchbare Ergebnisse erzeugt. Auch wenn interessante Effekte zu beobachten waren, können die Ergebnisse nur als plausible Indikatoren für die Ausbreitung von Tuberkulose herangezogen werden. Die agentenbasierte Simulation hat gezeigt dass unterschiedliche Strategien, zumindest auf synthetisierten Daten, einen signifikanten Einfluss zur Dämmung von Tuberkulose haben können. Nichts desto trotz ist es überaus schwierig relevante Ergebnisse aus diesen Modellen zu bekommen. Dies liegt unter anderem an der enormen Komplexität und der Menge an Variablen die so ein System beeinflussen.

Eine der vielversprechendsten Teile der Modellierung mit statistischen Prädiktionsmodellen über eine Menge beschreibender Variablen konnten nicht näher betrachtet werden. Diese Modelle stellen sinnvolle Erweiterungen zu rein theoretischen Modellen dar. Von realen Daten zu lernen beinhaltet den Vorteil, die unterliegenden Trends, deutlich präziser anzunähern. Voraussetzung ist allerdings, dass diese Daten existieren und auch tatsächlich das System von Interesse erklären können.

## LITERATUR

- [Aparicio and Castillo-Chavez, 2009] Aparicio, J. P. and Castillo-Chavez, C. (2009). Mathematical modelling of tuberculosis epidemics. *American Institute of Mathematical Sciences*, 6:209–237.
- [Harko et al., 2014] Harko, T., Lobo, F. S., and Mak, M. (2014). Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*, 236:184–194.
- [Hethcote, 2000] Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42:599–653.
- [Kazil, 2014] Kazil, J. (2014). Mesa: Agent-based modeling in python 3+.
- [Kumar et al., 2011] Kumar, A., Farhana, A., Guidry, L., Saini, V., Hondalus, M., and Steyn, A. J. (2011). Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Reviews in Molecular Medicine*, 13.
- [Perez and Dragicevic, 2009] Perez, L. and Dragicevic, S. (2009). An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8(1):50.
- [Trauer et al., 2014] Trauer, J. M., Denholm, J. T., and McBryde, E. S. (2014). Construction of a mathematical model for tuberculosis transmission in highly endemic regions of the asia-pacific. *Journal of Theoretical Biology*, 358:74–84.
- [World Health Organization, 2016] World Health Organization (2016). Global tuberculosis report 2016. Technical report, World Health Organization, Geneva, Switzerland.
- [World Health Organization, 2017] World Health Organization (2017). Treatment of tuberculosis guidelines. Technical report, World Health Organization, Geneva, Switzerland.