

Angewandte Mathematik und Optimierung Schriftenreihe
Applied Mathematics and Optimization Series
AMOS # 39(2015)

Robin Stanley Mathea

Modellierung von Infektionskrankheiten zur
Beurteilung von Syndromüberwachung

Herausgegeben von der
Professur für Angewandte Mathematik
Professor Dr. rer. nat. Armin Fügenschuh

Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg
Fachbereich Maschinenbau
Holstenhofweg 85
D-22043 Hamburg

Telefon: +49 (0)40 6541 3540
Fax: +49 (0)40 6541 3672

e-mail: appliedmath@hsu-hh.de
URL: <http://www.hsu-hh.de/am>

Angewandte Mathematik und Optimierung Schriftenreihe (AMOS), ISSN-Print 2199-1928
Angewandte Mathematik und Optimierung Schriftenreihe (AMOS), ISSN-Internet 2199-1936

Arbeit zum Erreichen des akademischen
Grades

Master Of Science



HELMUT SCHMIDT
UNIVERSITÄT

Universität der Bundeswehr Hamburg

Robin Stanley Mathea

Modellierung von Infektionskrankheiten zur Beurteilung von
Syndromüberwachung

Modelling of infectious diseases for evaluation of
syndrom surveillance

Fakultät für Maschinenbau

Professur für Angewandte Mathematik

Studiengang: Maschinenbau

Matr.-Nr. 858840

Übernahme: 30.05.2015

Abgabe: 28.09.2015

1. Prüfer: Univ.-Prof. Dr. Armin Fügenschuh

2. Prüfer: Univ.-Prof. Dr.-Ing. Bernd Niemeyer

Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst, keine anderen als die im Literaturverzeichnis genannten Quellen und Hilfsmittel, insbesondere keine dort nicht genannten Internetquellen benutzt, alle wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe und dass die auf dem elektronischen Speichermedium abgegebene Fassung der Arbeit der gedruckten entspricht.

Hamburg,

.....

(Datum)

(Unterschrift)

Inhaltsverzeichnis

Abbildungsverzeichnis	iv
Symbolverzeichnis	v
Zusammenfassung	vi
Abstract	vii
1 Einleitung und Hintergrund der Arbeit	1
2 Ziele der Arbeit	2
3 Methodik	4
3.1 Faktoren des Ebola-Ausbruchs 2014	4
3.2 Analyse bisher verwendeter mathematischer Modelle	8
3.2.1 SIR Modelle	8
3.2.2 WHO Modell	11
3.2.3 Yale Modell	15
3.3 Erstellung eines SEIR Modells mit vier Bevölkerungsgruppen	18
3.3.1 Datenaquise	18
3.3.2 Das multisoziale multiregionale SEIR Modell	20
3.4 Simulation eines syndrombasierten nahezu Echtzeit-Überwachungssystems .	28
4 Ergebnisse	32
4.1 Abgrenzung zu bisherigen SEIR Modellen	32
4.2 Simulation einer Syndromüberwachung	33
4.2.1 Kalibrierung des SEIR Modells	33
4.2.2 Vergleich von professionellem Medizinpersonal und CHWs	37

5 Diskussion und Ausblick	40
5.1 Zusammenfassung und kritische Prüfung der Ergebnisse	40
5.2 Grenzen der Arbeit	41
5.3 Ausblick	41
Literaturverzeichnis	45

Abbildungsverzeichnis

3.1	Darstellung des Yale SEIR Modells	16
3.2	Multisoziales multiregionales SEIR Modell mit vier Bevölkerungsgruppen .	24
3.3	Falldaten der CDC bis 30. Woche, mit gleitendem Mittelwert (gestrichelt).	27
3.4	Krankheitsverlauf bis 30. Woche gemäß SEIR Modell mit WHO Daten. . .	27
3.5	Krankheitsverlauf bis 60. Woche gemäß SEIR Modell mit WHO Daten. . .	28
4.1	Krankheitsverlauf bis 30. Woche gemäß kalibriertem SEIR Modell.	34
4.2	Krankheitsverlauf bis 60. Woche gemäß kalibriertem SEIR Modell.	35
4.3	Falldaten der CDC bis 60. Woche, mit gleitendem Mittelwert (gestrichelt).	36
4.4	Simulationsergebnis des Vergleichs von Medizinerinnen und CHWs.	38
4.5	Krankheitsverlauf bis 30. Woche gemäß kalibriertem SEIR Modell.	38

Symbolverzeichnis

N	Gesamtpopulation
S	Gesunde bzw. Infizierbare
E	latent Infizierende
I	Erkrankte bzw. Infizierende
R	nicht mehr Infizierte
\mathfrak{R}_0	Basisreproduktionszahl
\mathfrak{R}_t	Nettoreproduktionszahl
β	Ansteckungs- oder Infektionsrate
γ	Rate, mit der Personen die Klasse der Infizierenden verlassen
ϑ	Kontaktrate
ε	Rate, mit der infizierte Personen ansteckend werden
T_{sim}	in der Simulation betrachteter Zeitraum
δ	Sterbewahrscheinlichkeit
α	Wahrscheinlichkeit einer hygienischen Bestattung
p	arme Bevölkerung
r	reiche Bevölkerung
U	Stadtbevölkerung
R	Landbevölkerung

Zusammenfassung

In dieser Arbeit wurde ein multisoziales, multiregionales SEIR Modell mit vier Klassen zur Analyse der Auswirkungen des Einsatzes von Community Health Worker zusammen mit dem Einsatz eines syndrombasierten nahezu Echtzeit-Überwachungssystem erstellt. Modelliert wurde die Ebola-Epidemie in Liberia aus dem Jahr 2014. Die Berechnung erfolgte mit der Software Vensim von Ventana. Der Verlauf der Krankheit konnte gut wiedergegeben werden, allerdings waren die Modellparameter nicht optimal eingestellt, weshalb es zu einer einseitigen Verteilung der Krankenzahlen auf nur einen Teil der Bevölkerung kam. Dennoch konnte mit der abschließenden Simulation gezeigt werden, dass der Einsatz von Community Health Workern im Vergleich zum Einsatz von professionellem medizinischen Personal einen Vorteil bei der Früherkennung und Einschätzung von Epidemien bringen würde.

Abstract

In this thesis a multisocial and multiregional SEIR model with four categories is created in order to analyse the effects of using Community Health Workers together with a syndrome-based near-real time surveillance system. The Ebola epidemic disease of 2014 in Liberia forms the basis for modeling. For calculation the software Vensim by Ventana is used. The progression of disease could be reflected well but there was a one-sided distribution of the invalid data in relation to only a part of population because of incorrect model parameter settings. However a terminal simulation demonstrated that the use of Community Health Workers instead of professional medical staff was more effective concerning the early recognition and assessment of epidemic diseases.

1 Einleitung und Hintergrund der Arbeit

Die Ebola-Epidemie in Westafrika von 2014 forderte deutlich mehr Opfer, als alle vorherigen Ausbrüche dieser Krankheit [13, 25]. Seither gibt es viele Überlegungen, wie ein Ausbruch ähnlichen Ausmaßes besser eingedämmt werden kann. Ein wichtiger Faktor hierbei ist die Zeit, die vergeht, bis der Ausbruch als solcher erkannt wird, denn erst danach kann die internationale Hilfe anlaufen. Ein interessanter Ansatz dazu ist der Einsatz von nicht-professionellen freiwilligen Helfern, sog. Community Health Workers (CHWs), in Verbindung mit einem syndrombasierten nahezu Echtzeit-Überwachungssystem. Um den Nutzen eines solchen Einsatzes abzuschätzen wurde ein multisoziales, multiregionales SEIR Modell mit vier Klassen entwickelt (Abschnitt 3.4). Für die Wahl der Modellparameter wurde speziell auf die begünstigenden Faktoren der Epidemie und deren Quantifizierbarkeit eingegangen (Abschnitt 3.1). Nach eingehender Literaturrecherche wurde das Modell implementiert (Abschnitt 3.3) und mit den real erhobenen Daten kalibriert (Unterabschnitt 4.2.1). Die Berechnung der komplexen Differentialgleichungssysteme wurde mit der Software Vensim von Ventana durchgeführt. Abschließend wurde ein bereits existierendes statisches, stochastisches Modell in das entwickelte dynamische SEIR Modell implementiert, um abschließend den Einsatz von CHWs im Vergleich zum Einsatz mit professionellen Medizinern beurteilen zu können (Unterabschnitt 4.2.2).

2 Ziele der Arbeit

Das Ziel dieser Arbeit ist es zu untersuchen, ob der Einsatz von nicht professionellen, angeleiteten lokalen Community Health Workers (CHWs, Gemeinde-Gesundheitshelfer) im Zusammenspiel mit einem nahezu Echtzeit-Überwachungssystem für Krankheitssyndrome, die frühere Erkennung von Epidemien ermöglichen kann. Die Grundlage hierfür ist ein in Kapitel 3.3 erstelltes, mathematisches SEIR Modell zum Verlauf der Ebola-Epidemie in Liberia.

Dieses Gesamtziel der Arbeit wird systematisch in drei Unterziele aufgeteilt:

Ziel 1

Die Untersuchung bisheriger mathematischer Modelle unter Berücksichtigung der Genauigkeit ihrer Prognosen bezüglich künftiger Opferzahlen.

Der Zweck hiervon ist es, die Stärken und Schwächen mathematischer Modelle zu erkennen, gegebenenfalls auszubessern und gleichzeitig die Grenzen des eigenen Modells besser einschätzen zu können.

Ziel 2

Die Erstellung eines SEIR Modells mit vier unterschiedlichen Bevölkerungsgruppen.

Dieses Modell soll den Ebola-Ausbruch in Liberia 2014 möglichst genau beschreiben. Um die Auswirkungen eines Einsatzes von CHWs zu untersuchen, ist die Betrachtung verschiedener sozialer Bevölkerungsgruppen von Vorteil (siehe Abschnitt 3.3). Das Modell bildet die mathematische Grundlage zur Realisierung von Ziel 3:

Ziel 3

Der Vergleich von CHWs und professionellem Medizinpersonal bezüglich der Früherkennung von Epidemien unter der simulierten Verwendung eines nahezu Echtzeit-Überwachungssystems für Krankheitssyndrome.

Hiermit wird eine Alternative zum aktuellen Gesundheitssystem in Liberia untersucht mit dem Zweck, bei künftigen Epidemie-Ausbrüchen schneller reagieren zu können und somit die Opferzahlen zu reduzieren.

3 Methodik

Die in Kapitel 2 formulierten Ziele werden in diesem Kapitel systematisch bearbeitet. Zunächst wird in Abschnitt 3.1 auf die Faktoren der Ebola-Epidemie von 2014 eingegangen und unter deren Berücksichtigung werden in Abschnitt 3.2 die mathematischen Modelle von zwei bedeutenden Institutionen, der Weltgesundheitsorganisation (WHO) und der Yale Universität, auf Vorteile und Grenzen untersucht. Im Anschluss wird, entsprechend Ziel 2, mit diesen Erkenntnissen in Abschnitt 3.3 ein SEIR Modell erstellt, welches die Grundlage für die Simulation eines nahezu Echtzeit-Überwachungssystems für Krankheitssyndrome bildet. Mit dieser Simulation kann der Vergleich aus Ziel 3 erfolgen.

Die in dieser Arbeit erstellten Modelle und Simulationen beschränken sich auf Liberia. Dies hat mehrere Gründe. Zum einen ist die Beschränkung auf eines der drei hauptsächlich betroffenen Länder (Guinea, Sierra Leone und Liberia) deshalb sinnvoll, da jedes Land unterschiedlich bei der Eindämmung der Epidemie vorgegangen ist. Außerdem bestanden schon zum Beginn des Ausbruchs unterschiedliche Voraussetzungen. Von den drei Ländern wurde Liberia für die nähere Betrachtung gewählt, eine eingehendere Erklärung dazu wird in Abschnitt 3.3 geliefert.

3.1 Faktoren des Ebola-Ausbruchs 2014

Die Ebola-Epidemie in Westafrika begann bereits im Dezember 2013 [25]. Als sog. Indexfall, der erste Infizierte dieses Ausbruchs, erkrankte ein 18 Monate alter Junge in Guinea in der Grenzregion zu Sierra Leone und Liberia. Er starb am 26. Dezember, steckte jedoch noch einige Familienmitglieder, sowie medizinisches Personal des lokalen Krankenhauses an. Bis Mitte März 2014 konnte sich die Krankheit unentdeckt weiterverbreiten. Am 14. März wurde erstmals über ein seltsames Fieber in Guinea berichtet, am 19. März daraufhin dann Ebola als mögliche Ursache genannt. Am 23. März 2014 wurde die Weltgesundheitsorganisation über den ersten, per Laboruntersuchung bestätigten, Ebola-Fall informiert [24]. Die Krankheit verbreitete sich rapide und betraf bald auch die

Nachbarländer Sierra Leone und Liberia. Am 8. August 2014 rief die WHO für die Region den internationalen Gesundheitsnotstand aus. Mitte September lief daraufhin, 9 Monate nach dem ersten Auftreten der Krankheit, die internationale Hilfe an [25]. Am 9. Mai 2015 wurde Liberia als erstes der drei am meisten betroffenen Länder als frei von Ebola erklärt. Da sich jedoch kurze Zeit später ein neuer Ebola-Fall in dem Land ereignete, wurde dieser Status widerrufen. Am 3. September wurde der Status erneut von der WHO bestätigt, nach dem eine 42-tägige Periode ohne bestätigtem Auftreten der Krankheit verstrichen war [28].

Wie zu Beginn des Kapitels bereits vorweg genommen wurde, fokussiert sich diese Arbeit auf den Verlauf der Epidemie in Liberia. In dem Land tobte von 1989 bis 2003 ein blutiger Bürgerkrieg der Zehntausende Menschenleben forderte [23]. Neben diesen absoluten Verlusten wurde jedoch auch das Gesundheitssystem des Landes fast vollständig vernichtet. Zum Kriegsende waren lediglich 51 der ehemals 293 medizinischen Einrichtungen einsetzbar [12]. Anfang 2010 hatte das liberianische Gesundheitsministerium bereits 378 medizinischen Einrichtungen, wenn auch nur teilweise, in Betrieb genommen [8]. Durch die lange Zeit des Krieges und des Wiederaufbaus hatte sich neben der formellen eine informelle Gesundheitsversorgung etabliert. Zu dieser gehören u.a. Verkäufer von Medikamenten (sog. „black bagger“) und traditionelle Heiler und Hebammen. Eine Studie aus dem Jahr 2011 belegt, dass selbst acht Jahre nach Kriegsende in 10 von 13 Fällen auf die informelle Gesundheitsversorgung zurückgegriffen wurde. Der Aufbau eines funktionierenden Gesundheitssystems wurde von der Bevölkerung als wichtigster Regierungsauftrag beschrieben [27].

Die Funktionsfähigkeit bzw. Regierbarkeit eines Staates kann anhand des „fragile state index“ beurteilt werden [4]. Dieser Index wird aus zwölf Indikatoren zusammengerechnet, welche u.a. die sozialen, ökonomischen und politischen Gegebenheiten der jeweiligen Länder berücksichtigen. Die drei am meisten betroffenen Länder der Epidemie wurden 2014 mit einer Warnung („alert“) beurteilt [24]. Liberia war hierbei das am stärksten vom Zerfall gefährdete von 178 bewerteten Ländern [7]. Schon 2007 ordnete der renommierte britische Ökonom Paul Collier diese drei Ländern den 55 weiteren „bottom billion countries“ zu. Diese Länder bzw. Ökonomien warnte er davor, dass in Folge von Konflikten, wie z.B.

des liberianischen Bürgerkriegs, ein Rückfall in einen weiteren Konflikt wahrscheinlich ist. Im Nachhinein stellte sich dieser weitere Konflikt als medizinische Krise dar.

In jedem Fall konnte die Funktions- bzw. Handlungsfähigkeit der liberianischen Regierung und damit auch deren Gesundheitssystem bereits vor der Ebola-Epidemie mit berechtigten Zweifeln bedacht werden. Dadurch wurde die Überwachung und die Eindämmung der Epidemie von vornherein, im Vergleich zu stabileren Staaten, erschwert. Da die Krankheit in der Grenzregion der drei Staaten ausbrach, wurde somit auch deren Ausbreitung begünstigt [6, 25].

Somit lässt sich erklären, warum trotz der Bemühungen der Regierung, ein funktionierendes Gesundheitssystem aufzubauen, vor dem Ausbruch der Ebola-Epidemie auf 10.000 Einwohner lediglich 2,8 Beschäftigte im Gesundheitswesen kamen. Der Bevölkerung von insgesamt 4,29 Millionen Menschen standen nur 51 Ärzte zur Verfügung [2]. Die von der Regierung selbstgesteckten Ziele bezüglich der Bereitstellung von medizinischer Infrastruktur, welche sich an den Millenniums-Entwicklungsziele der UN orientierten, wurden auch fünf Jahre nach Kriegsende noch nicht erreicht [1]. Durch die mangelhafte Infrastruktur konnten natürlich nicht alle aufgetretene Krankheitsfälle unverzüglich von den Behörden erfasst werden. Dies ist bei der Betrachtung der Fallzahlen und dem Rückschließen auf charakteristische Zeitspannen, wie z.B. der durchschnittlichen Inkubationszeit oder Krankheitsdauer, zu berücksichtigen [25].

Inwieweit Zugang zu medizinischer Versorgung möglich ist, kann anhand der sog. „five A’s“ beurteilt werden [19]:

- Zugänglichkeit (englisch: **A**ccessibility): Ist es geografisch und physikalisch möglich, Zugang zu Versorgungseinrichtungen zu erhalten? Ist die Entfernung dahin zu weit, stehen Transportmöglichkeiten zur Verfügung?
- Erschwinglichkeit (englisch: **A**ffordability): Kann sich der Patient die nötige medizinische Versorgung leisten?
- Verfügbarkeit (englisch: **A**vailability): Entspricht die Qualität der angebotenen Versorgungsleistungen dem realen Bedarf?

- Anpassbarkeit (englisch: **Accommodation**): Wie gut reagiert das System auf Bedarfsänderungen? Kann sichergestellt werden, dass überall das Benötigte zur Verfügung steht? Ist ausreichend Personal und Aufnahmekapazität vorhanden?
- Akzeptanz (englisch: **Acceptability**): Nimmt die Bevölkerung die angebotenen Leistungen an?

Bereits 2008 wurde Liberias Gesundheitssystem unter Berücksichtigung dieser fünf Dimensionen im Rahmen einer Befragung beurteilt. Dabei wurde festgestellt, dass 97,5% der Befragten bereits mindestens eine Beschränkung beim Zugang zu medizinischer Versorgung erlebt hatten. Als Haupthinderungsgründe wurden mangelnde Bezahlbarkeit und die Entfernung zur benötigten Versorgung genannt [9].

Nun lässt sich nachvollziehen, warum in dieser Arbeit ausgerechnet CHWs als interessante Alternative zur bisher etablierten Gesundheitsversorgung genannt werden. Sie erreichen deutlich mehr Menschen, können diesen schneller helfen und dazu bei der Erhebung der Statistiken einen enormen quantitativen Vorteil bringen.

Die aufgezählten Faktoren der Ebola-Epidemie sind lediglich eine Auswahl an begünstigenden Umständen des Ausbruchs. Unzählige Studien haben sich mit dieser Thematik bereits befasst. Sie untersuchten u.a. die Auswirkungen der örtlichen Gebräuche, wie den traditionellen Beerdigungsritualen [13, 15] oder dem Einfluss von sozialpsychologischen Effekten auf den Verlauf der Epidemie [21]. Die hier genannten Faktoren geben jedoch schon einen guten Überblick über die Komplexität und das Zusammenspiel der verschiedenen Dimensionen dieser Epidemie. Es wird deutlich, dass eine mathematische Modellierung dieser Epidemie, so komplex sie auch gestaltet werden würde, niemals ein genaues Abbild der realen Gegebenheiten sein, geschweige denn eine exakte Zukunftsprognose erstellen kann. Mit dem Wissen über die nicht mögliche Realisierung einer genauen Erfassung und Evaluation aller Faktoren sollten die Ergebnisse aller Modelle kritisch beurteilt werden. Nichtsdestotrotz können ohne den Versuch einer Integration dieser nicht messbaren Faktoren bei der Modellierung ebenfalls keine exakten Ergebnisse erwartet werden. Deshalb gilt es, diese Einflüsse bestmöglich zu berücksichtigen, um damit die Genauigkeit des Modells zu steigern.

3.2 Analyse bisher verwendeter mathematischer Modelle

Zum Entwickeln eines eigenen Modells der Ebola-Epidemie ist es hilfreich, bereits verwendete Modelle zu Rate zu ziehen. Deshalb wurde zuerst untersucht, welche Art von Modellen für eine Epidemie dieses Ausmaßes geeignet ist [14]. In Unterabschnitt 3.2.1 werden die Erkenntnisse dieses Auswahlprozesses kurz zusammengefasst. Anschließend werden die bereits verwendeten Modelle von der WHO und der Yale Universität analysiert.

3.2.1 SIR Modelle

Die in den Abschnitten 3.2.3 und 3.3 vorgestellten SEIR Modelle basieren auf dem von William Ogilvy Kermack und Anderson Gray McKendrick entwickelten SIR Modell zur Beschreibung von Epidemien [10]. Ausgehend von diesem Basismodell wurde bereits eine Vielzahl von Krankheitsausbrüchen modelliert. In diesen Modellen wird die Population in verschiedene Klassen, welche den Krankheitszustand eines Individuums beschreiben, unterteilt.

Beim SEIR Modell werden folgende vier Klassen unterschieden:

- S - Gesunde bzw. Infizierbare (englisch: **susceptible**): Dieser Personenkreis kam mit der Krankheit noch nicht in Kontakt und ist nicht dagegen immun.
- E - Latent Infizierte (englisch: **exposed**): Diese sind zwar schon infiziert, können die Krankheit jedoch noch nicht weitergeben. Sie befinden sich in der Inkubationsphase, wo noch keine Symptome auftreten, weshalb das Erkennen der Erkrankung schwierig ist.
- I - Erkrankte bzw. Infizierende (englisch: **infective**): Diese können die Krankheit an andere Personen weitergeben und zeigen Symptome.
- R - Nicht mehr Infizierende (englisch: **removed** oder **recovered**): Diese Personen können die Krankheit nicht mehr weitergeben. Sie sind entweder gesund und immun oder tot und nicht mehr ansteckend.

Diese vier Klassen haben die Oberklasse:

- N - Gesamtpopulation: Diese ist die Summe aller anderen Klassen und daher konstant.

Jedes Individuum der Population kann nur einer Klasse gleichzeitig angehören, es kann jedoch seine Klasse wechseln. In jedem Modell gibt es für diese Wechsel ein festgelegtes Schema, welches abhängig von der Ausgangsklasse festlegt, wohin die Person wechseln kann. Die möglichen Klassenwechsel werden mit Übergangskoeffizienten beschrieben. In den hier behandelten Modellen gibt es drei Übergangskoeffizienten:

- β - Ansteckungs- oder Infektionsrate: Sie beschreibt das Verlassen der Klasse der Gesunden.
- ε - Rate, mit der infizierte Personen ansteckend werden. Die Inkubationszeit, die Zeit, die eine Person der Klasse E angehört, ist durchschnittlich $\frac{1}{\varepsilon}$.
- γ - Rate, mit der Personen die Klasse der Infizierenden verlassen, mit der sie entweder gesund oder sterben. Die durchschnittliche Zeit, die sich eine Person in der Klasse I befindet, beträgt $\frac{1}{\gamma}$. Wenn infektiöse Leichname die Krankheit weiterverbreiten können, ist die Einführung einer Zusatzklasse für diese Toten möglich. Dann muss γ entsprechend aufgeteilt werden.

Die Klassen stehen in den Modellen so miteinander in Beziehung, dass ihre Größen einerseits voneinander gegenseitig beeinflusst werden und zusätzlich von den entsprechenden Übergangskoeffizienten und der verstrichenen Zeit seit dem Anfangszustand abhängen. Damit sind die Modelle deterministisch und dynamisch. Eine wichtige Eigenschaft von dynamischen Systemen ist ihre Systemgrenze. Dadurch werden sie von ihrer Umwelt entkoppelt. Bei den SIR Modellen wird das betrachtete System innerhalb dieser Grenze mit der Gesamtpopulation N beschrieben. Was darin nicht erfasst wurde, hat deshalb keine Auswirkungen auf das Modell.

Elemente eines dynamischen Systems werden in der Regelungstechnik durch Blöcke dargestellt, die Beziehungen untereinander und die Wechsel zueinander werden durch Pfeile gekennzeichnet. Ein einfaches System aus zwei Elementen A und B wird dann wie folgt dargestellt: $\boxed{\text{Element } A} \xrightarrow{\text{Beziehung zwischen den Elementen}} \boxed{\text{Element } B}$. In dieser Arbeit werden deterministische Systeme behandelt, d.h. die Änderungsraten der Zustandsgrößen sind von den jeweiligen Zuständen und Umwelteinwirkungen abhängig und eindeutig bestimmbar. Zufälle individueller Schicksale werden durch empirisch ermittelte, deterministische Mittelwerte ersetzt, das Verhalten des globalen Realsystems kann dennoch gut angenähert werden.

Die SIR Modelle enthalten Populationsklassen, deren Größe sich abhängig von der, seit dem Anfangszeitpunkt vergangenen, Zeit und der momentanen Größe anderer Klassen ändert. Diese Änderung wird jeweils mit einer Differentialgleichung beschrieben. Somit ergibt sich für jedes Infektionsmodell ein System von Differentialgleichungen (DGLS). Da die Anfangsgröße jeder Klasse bekannt ist, kann über die beschriebenen Änderungen zu jedem gewünschten Zeitpunkt die Klassengröße errechnet werden.

Das SIR Modell ist das bekannteste Infektionsmodell. Hier wird von einer dauerhaften Immunisierung nach der Genesung bzw. dem Tod ausgegangen [20]. Es gibt nur eine Richtung, in die Klassenwechsel möglich sind. Und zwar von S nach I und von dort nach R . Aus R ist kein Wechsel möglich: $\boxed{S} \xrightarrow{\text{Ansteckungsrate}} \boxed{I} \xrightarrow{\text{Gesundungsrate}} \boxed{R}$. Die Wechsel zwischen den Klassen erfolgen gemäß folgender Gleichungen:

$$\frac{dS}{dt} = -\beta SI, \quad (3.1a)$$

$$\frac{dI}{dt} = \beta SI - \gamma I, \quad (3.1b)$$

$$\frac{dR}{dt} = \gamma I. \quad (3.1c)$$

Die Ansteckung ist abhängig von der Ansteckungsrate β und den momentanen Größen der Klassen S und I . Da die Personen, welche die Klasse S verlassen, nur in die Klasse I wechseln können, sind beide Änderungen betragsmäßig gleich und haben verschiedene Vorzeichen. Das Verlassen der Klasse wird mit negativem Vorzeichen, der Wechsel in eine Klasse entsprechend mit positivem Vorzeichen beschrieben. Die Genesung bzw. das

Sterben ist nur von der Rate γ und der Anzahl an Infizierten abhängig und geht positiv in die Bilanz der Klasse R und entsprechend negativ in die von Klasse I ein.

Das SEIR Modell [17] erweitert das zugrunde liegende SIR Modell um die Klasse der latent Erkrankten E : $\boxed{S} \xrightarrow{\text{Ansteckungsrate}} \boxed{E} \xrightarrow{\text{Inkubationsrate}} \boxed{I} \xrightarrow{\text{Gesundungsrate}} \boxed{R}$. Die Größe der Klasse E hat nur Auswirkungen auf das Wachstum von Klasse I und ist selber von deren Größe abhängig:

$$\frac{dS}{dt} = -\beta SI, \quad (3.2a)$$

$$\frac{dE}{dt} = \beta SI - \varepsilon E, \quad (3.2b)$$

$$\frac{dI}{dt} = \varepsilon E - \gamma I, \quad (3.2c)$$

$$\frac{dR}{dt} = \gamma I. \quad (3.2d)$$

Für die Charakterisierung einer Krankheit wurde von [10] die aussagekräftige Basisreproduktionszahl \mathfrak{R}_0 eingeführt, mit:

$$\mathfrak{R}_0 = \frac{\beta N}{\gamma}. \quad (3.3)$$

Sie gilt für den Fall, dass die gesamte Bevölkerung zur Klasse S gehört und nun eine Person infiziert wird, denn dann wird die Krankheit mit $\beta SI = \beta S \cdot 1 = \beta N$ übertragen. Diese Ansteckungsrate wird nun ins Verhältnis zur Genesungsrate gesetzt. Da $\frac{1}{\gamma}$ die durchschnittliche Infektionszeit war, gibt sie gleichfalls an, wie viele Menschen während der mittleren Infektionszeit von einem Kranken angesteckt werden, falls die Zahl der Infizierbaren S konstant gehalten wird [20]. $\mathfrak{R}_0 = 1$ ist der Schwellwert, ab dem eine Krankheit in der Bevölkerung erhalten bleibt. Dementsprechend entstehen Epidemien nur für $\mathfrak{R}_0 > 1$. Der Schwellwert hat den Vorteil, dass es sich dabei um einen relativen Wert handelt und damit verschiedene Ausbrüche mit jeweils unterschiedlichen Parametern und Anfangswerten vergleichbar sind, da absolute Zahlen nur Aussagen über den Einzelfall zulassen.

3.2.2 WHO Modell

In diesem Abschnitt wird das WHO Modell des WHO Ebola Response Teams [25] erläutert.

Zuerst wurden die erhobenen und gesammelten Daten kritisch geprüft und bereinigt. In den Statistiken werden die Krankheits- und Todesfälle gemäß einheitlicher Definition in drei Kategorien eingeteilt [18]. Es wird jeweils zwischen verdächtigen, wahrscheinlichen und bestätigten Fällen unterschieden. Die Patienten werden anhand verschiedener Merkmale erfasst, dazu zählen Alter, Geschlecht, Wohnort, Ort der Erfassung usw. Bei der Analyse der erhobenen Daten wurde bemerkt, dass z.T. bei Datumsangaben Tag und Monat vertauscht wurden. Konnte solch ein Fehler festgestellt werden, wurde er korrigiert. Die WHO geht davon aus, dass deshalb alle Daten vor März und ab September 2014 für Liberia korrekt sind. Für den dazwischen liegenden Zeitraum kann dieser Umstand nicht angenommen werden, obwohl genau die darin erhobenen Daten für die Erstellung des WHO Modells verwendet wurden.

Wurde bei der Datenauswertung festgestellt, dass eine unlogische Kausalkette vorliegt, wurden die betroffenen Fälle nicht berücksichtigt. Eine unlogische Kausalkette lag z.B. dann vor, wenn laut Datenaquise erst nach dem Tod Symptome aufgetreten sind oder der Patient vor der Einlieferung bereits aus dem Krankenhaus entlassen wurde.

Fehlten bei Fällen Angaben, wurden diese bestmöglich geschätzt. War beispielsweise das Datum des Auftretens der ersten Symptome unbekannt, wurde mit dem länderspezifisch errechneten Median von dem Tag der Datenerfassung rückwärts extrapoliert. Insgesamt wurden fünf wichtige Zeitabschnitte definiert und berechnet:

- Die Inkubationszeit: Sie verstreicht zwischen der Infizierung eines Patienten und dem Zeitpunkt, zu dem er ansteckend wird. Es wurde für diese Epidemie angenommen, dass dies zeitgleich mit dem Auftreten der ersten Symptome beim Patienten der Fall war.
- Die Zeitspanne vom Auftreten der ersten Symptome bis zur Aufnahme in einer medizinischen Versorgungseinrichtung.
- Die Zeitspanne zwischen der Aufnahme in einer medizinischen Versorgungseinrichtung bis zur Entlassung bzw. bis zum Tod des Patienten.
- Das sog. „serial interval“: Es beziffert die Zeitspanne vom Auftreten der Symptome bei einem Patienten bis zum Auftreten der Symptome bei einem, von ihm ange-

steckten, Patienten.

- Die Generationszeit: Sie liegt zwischen der Infizierung eines Patienten und der Infizierung eines von ihm angesteckten Patienten.

Bei der Berechnung der Inkubationszeit wurden nur wahrscheinliche und bestätigte Fälle einbezogen, die entweder Kontakt zu einem wahrscheinlichen oder bestätigten Ebola-Patienten hatten bzw. an einer traditionellen Beerdigung eines solchen Todesfalls teilnahmen. Die übrig gebliebenen Fälle wurden hinsichtlich der Zeitspanne ihrer Ansteckung entsprechend in eintägige und mehrtägige Exposition untergliedert. Die Inkubationszeit wurde abschließend mit einer log-likelihood Funktion an die ausgewählten realen Daten angepasst, dabei wurden die Expositionsdauer gammaverteilt aufgetragen [11].

Für das „serial interval“ wurden ausschließlich Fälle mit nur einem einzigen Kontakt zu einem Infizierten. Bei beiden Fällen musste der Tag des Auftretens der ersten Symptome angegeben worden sein.

Bei der Bestimmung der Generationszeit wurde die Annahme getroffen, dass mit dem Auftreten der ersten Krankheitssymptome der Patient gleichzeitig ansteckend wird. Dadurch sind Generationszeit und „serial interval“ identisch.

Ein weiterer charakteristischer Wert von Epidemien ist die Letalität. Sie gibt das Verhältnis von Todesfällen gemessen an der Gesamtzahl von Erkrankungen an. Diesen Wert während eines aktuellen Ausbruchs zu ermitteln, ist dabei nicht so trivial, wie es auf den ersten Blick scheint. Es darf nämlich nicht einfach das Verhältnis mit den aktuellen Gesamtzahlen von Todes- und Krankheitsfällen gebildet werden. Dafür gibt es den einleuchtenden Grund, dass ein gewisser Anteil der zum aktuellen Zeitpunkt gemeldeten Kranken, jedoch lebenden Patienten, zu einem späteren Zeitpunkt sterben wird. Dies kann jedoch zum Zeitpunkt der Erhebung nicht in die Berechnung einfließen. Damit wird die Letalität niedriger ausfallen, als sie es in Wahrheit ist. Deshalb wurden Fälle aussortiert, bei denen das Auftreten der ersten Krankheitssymptome so nahe am Zeitpunkt der Erhebung lag. Das Bewertungskriterium hierfür war das Konfidenzintervall der durchschnittlichen Zeitspanne zwischen dem ersten Auftreten von Krankheitssymptomen bis zum Tod.

Ein weiteres Problem bei der Berechnung ist die Dunkelziffer an Erkrankten, die sterben, ohne registriert worden zu sein. Zusätzlich sind in 54% der erfassten Krankheitsfälle keine

Angaben über den Ausgang der Krankheit beim betroffenen Patienten möglich. Darum wurden für die Berechnung nur diejenigen Fälle herangezogen, bei denen bekannt war, ob der Patient überlebt hat oder nicht. Die ermittelte Letalität lag unabhängig von der Nation und vom Zeitpunkt der Erhebung bei ca. 70%.

Eine wichtige Größe bei der Betrachtungen von Epidemien ist die Basisreproduktionszahl \mathfrak{R}_0 . Für sie ist ein Schwellwert definiert, der Auskunft darüber gibt, ob und wie stark sich eine Krankheit verbreitet. Darum wird sie während Ausbrüchen meistens erhoben. Die WHO ermittelte gleichzeitig auch die Nettoreproduktionszahl \mathfrak{R}_t . Beide Zahlen werden prinzipiell gleich berechnet. Der Unterschied liegt darin, dass sich die Krankheit anfangs exponentiell verbreitet, dies lässt sich mit \mathfrak{R}_0 beschreiben. \mathfrak{R}_0 ist für das Einbringen einer Krankheit in eine vollkommen infizierbare Population definiert. Steigt die Zahl an Krankheitsfällen, ist diese Voraussetzung auch bei großen Populationen irgendwann nicht mehr erfüllt. Nach einer gewissen Zeit ändert sich außerdem das Verhalten der Menschen, sie achten mehr auf Risikofaktoren und versuchen diese zu minimieren. Die Ausbreitungsgeschwindigkeit nimmt demzufolge ab. Sie wird nun mit \mathfrak{R}_t beschrieben, meist ist, wie auch bei dieser Epidemie, $\mathfrak{R}_0 > \mathfrak{R}_t$. \mathfrak{R}_t wurde anhand der wöchentlichen Fallzahlen in einem gleitenden 4-Wochen-Fenster berechnet. Durch dieses Fenster ist die Anzahl an betrachteten Fällen nicht erheblichen Schwankungen unterworfen worden. Dennoch konnten temporäre Trends aufgedeckt werden. Die ermittelten Werte wurden anschließend mit dem bei Stichproben üblichen Markov-Chain-Monte-Carlo-Verfahren [5] optimiert.

Die WHO stellt zwei Modelle zur Prognose des zukünftigen Krankheitsverlaufs vor. Beide Modelle verwenden \mathfrak{R}_t . Darum wurde für dieses für den Startzeitpunkt der Prognose berechnet. Hierfür wurde \mathfrak{R}_t anhand der Daten der zurückliegenden 6 Wochen ermittelt. Das erste Modell verwendete eine Regression. Dafür wurden aus den Daten der zurückliegenden 6 Wochen nur die wahrscheinlichen und bestätigten Fallzahlen einbezogen.

Das zweite Modell ist ein stochastisches und etwas komplexer. Es verwendet einen Poisson Prozess [22], der die Anzahl der Infizierten zu einem Zeitpunkt t berechnet. Für das Modell wurden die Wahrscheinlichkeiten für das Bekanntwerden von Krankheitsfällen und die Verzögerung zwischen Auftreten und Melden der Krankheit berücksichtigt. Die

entscheidenden Parameter der verwendeten Verteilung sind der ermittelte, zeitabhängige Parameter \mathfrak{R}_t der zurückliegenden 6 Wochen und das gammaverteilte, durchschnittliche „serial interval“, die Summe von durchschnittlicher Infektionsdauer und Inkubationszeit. Beide werden für jeden Kranken zu jedem betrachteten Zeitpunkt miteinander multipliziert. Damit wird beschrieben, dass nach dieser Zeit ein Infizierter eine entsprechend \mathfrak{R}_t große Anzahl an Personen angesteckt hat. Zieht man eine Parallele zu einem SIR Modell, wird mit dieser Multiplikation der Übergang von der Klasse I in die Klasse R beschrieben.

Beide Modelle kommen annähernd zu gleichen Ergebnissen, wobei das Modell mit der Poissonverteilung schmalere Konfidenzintervalle aufweist. Da die Modelle auf Messwerten basieren, schwanken die Ergebnisse je nach dem, wie viele Werte bereits bekannt sind. Ein schmaleres Konfidenzintervall bedeutet dann, dass der Erwartungsbereich der Schätzung kleiner ist und somit die Ergebnisse des Modells zuverlässiger sind [11].

3.2.3 Yale Modell

Mitte Oktober 2014 wurde in der renomierten Fachzeitschrift „The Lancet“ eine aktuelle Studie [15] der Yale Universität veröffentlicht. In dieser wurde ein erweitertes SEIR Modell der Ebola-Epidemie in Liberia vorgestellt. In diesem Abschnitt wird das Modell erläutert.

Grundsätzlich beinhaltet das Modell (siehe Abbildung 3.1) die vier üblichen Klassen S, E, I und R. Allerdings werden die Klassen I und R wiederum aufgeteilt. Entscheidend ist hier, ob der Patient am Ende sterben wird oder überlebt. Dementsprechend folgt er mit der Wahrscheinlichkeit δ bzw. $(1 - \delta)$ dem jeweiligen Pfad. Für die Überlebenden wird die Heilungsrate mit γ_R angegeben. Mit der Rate α wird ein Patient in eine Behandlungseinrichtung verbracht. Ist er dort erstmal aufgenommen worden, wird angenommen, dass er in Quarantäne ist und seine Umwelt nicht mehr anstecken kann. Als in Quarantäne befindlich, werden auch die Fälle bezeichnet, die in der regelmäßig von medizinischen Helfern betreut werden, da man davon ausgehen kann, dass sie im Todesfall entsprechend gefahrlos beerdigt werden. Wegen der Annahme einer perfekten Quarantäne in Behandlungseinrichtungen, wurden die Krankheits- und Todesfälle von medizinischen Helfern von den offiziellen Statistiken subtrahiert.

Es wurde, wie im WHO Modell, die Annahme getroffen, dass mit dem Ende der Inkubationszeit der Patient ansteckend wird und die ersten Symptome der Krankheit aufweist.

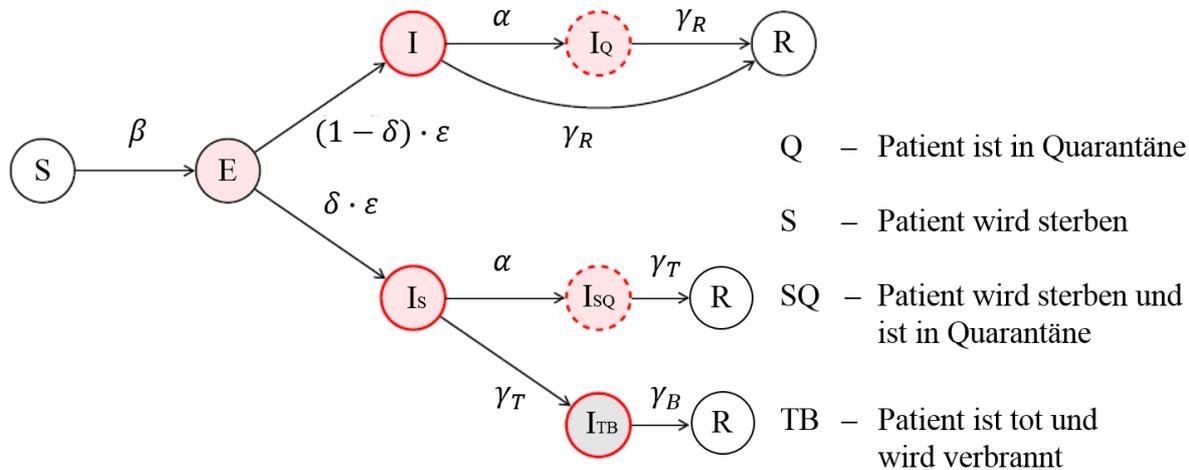


Abbildung 3.1: Darstellung des Yale SEIR Modells

Mit der Rate γ_T stirbt ein Infizierter, mit der Rate γ_B wird die Zeit bis zum Ende seiner traditionellen Beerdigung beschrieben. Hiernach wird der Leichnam als nicht mehr infizierend betrachtet. Die Wahrscheinlichkeit des Patiententodes wird mit der Letalitätsrate des WHO Modells beziffert. Ein Übergang von I_{SQ} nach I_Q oder umgekehrt ist nicht vorgesehen.

Es wird davon ausgegangen, dass die veröffentlichten Zahlen über Krankheits- und Todesfälle nicht das gesamte Ausmaß der Epidemie umfassen, es demnach eine Dunkelziffer an nicht erfassten Fällen gibt. Deshalb wird die Annahmen getroffen, dass in der Gemeinschaft Verstorbene nur teilweise in die offiziellen Statistiken eingeflossen sind. Diese nicht erfassten Fälle werde in der Berechnung der Ansteckungsrate berücksichtigt. Ebenso wird hier die Wahrscheinlichkeit einbezogen, dass nicht ausreichend Betten in Behandlungseinrichtungen zur Verfügung stehen. Auf Anraten des amerikanischen Centers for Disease Control and Prevention (CDC) verteilte das liberianische Gesundheitsministerium Schutz-Sets an Angehörige von Patienten, die diese zu Hause pflegten. Diese Sets boten allerdings nicht den gleichen Schutz vor Ansteckung, wie es die Ausrüstung der Helfer in den Behandlungseinrichtungen tat. Auch dieser Umstand floss in die Berechnungsformel der Ansteckungsrate mit ein. Damit waren alle benötigten Werte für die Berechnung der Basisreproduktionszahl \mathfrak{R}_0 bekannt. Anschließend wurde das Modell durch eine Anpas-

sung an die offiziellen Statistiken des liberianischen Gesundheitsministeriums kalibriert. Diese wurden unregelmäßig erhoben. Deshalb wurden die Fälle zur Kalibrierung wochenweise zusammengefasst. Das Modell diente der Berechnung der benötigten Krankbetten in den nächsten zwei Monaten im Montserrado County (Liberia). Deshalb wurden kurze Zeit vorher neu geschaffene Bettenkapazitäten und damit ein Anstieg an gemeldeten Fällen entsprechend in das Modell integriert. Zusätzlich wurden verschiedene Szenarien erstellt. Diese variierten hinsichtlich der Anfangszeitpunkte und unterschiedlich intensiven und effektiven Hilfsmaßnahmen im Rahmen der modellierten Intervalle.

Für die Startwerte des Modells wurden weitere Annahmen getroffen, da die Anzahl an Patienten in der Inkubationsphase nicht bekannt sein konnte. Das Modell beschränkt sich auf das Montserrado County, dort wurde jeweils ein Zensus in den Jahren 1984 und 2008 durchgeführt. Mit deren Bevölkerungszahlen wurde die Einwohnerzahl vor dem Ausbruch 2014 extrapoliert. Allerdings ist diese Zahl zumindest zu hinterfragen, da von 1989 bis 2003 in Liberia Bürgerkrieg herrschte und dieser sich auf die Bevölkerungszahl im betrachteten Gebiet ausgewirkt haben wird. Dies könnte sowohl durch Tötungen an sich als auch durch Migration über die Gebietsgrenzen hinweg geschehen sein. In Montserrado leben laut Schätzung 32% der liberianischen Bevölkerung. Von diesem Anteil leben über 90% in der Landeshauptstadt Monrovia. Gerade in diesem Gebiet ist deshalb die nationale und internationale Hilfe konzentriert gewesen. Es wird darum davon ausgegangen, dass die Schätzungen anderer Modelle bezüglich der Dunkelziffer der Epidemie für Montserrado lediglich eine pessimistische Obergrenze darstellen. Um diese dennoch für das Gebiet zu beurteilen, wurden die Übergangsraten zwischen den Klassen des Modells an den zeitlichen Verlauf der Zahlen der offiziellen Statistiken angepasst. Abschließend wurde, wie im WHO Modell, zur Optimierung des Modells das Markov-Chain-Monte-Carlo-Verfahren angewendet.

Entgegen der Annahme, dass die von der CDC geschätzte Dunkelziffer von 60% in Montserrado eher unwahrscheinlich ist, konvergierte die ermittelten Werte gegen diese obere Grenze. Deshalb wurde davon ausgegangen, dass die vom Modell prognostizierten Fallzahlen nur eine vorsichtige Schätzung des realen Verlaufs wäre.

3.3 Erstellung eines SEIR Modells mit vier Bevölkerungsgruppen

Bevor das Modell erstellt werden konnte, wurden zunächst einmal die benötigten Daten gesammelt und aufbereitet.

3.3.1 Datenaquise

Für das erstellte SEIR Modell wurde der Ausbruch in Liberia als zu modellierendes dynamisches System gewählt. Dies hat mehrere Gründe. Zum einen wurde sich gegen eine ganzheitliche Modellierung des Ausbruchs in Guinea, Sierra Leone und Liberia entschieden, da in jedem Land unterschiedliche Bedingungen vor und während der Epidemie herrschten (vgl. Abschnitt 3.1). Wie die Politik, die Bevölkerung oder die Weltgemeinschaft auf die Krise reagierten, war von großem Einfluss auf die Ausbreitung der Krankheit. Somit musste sich für ein Land entschieden werden. Hierbei wies Liberia einige Vorteile auf. Einerseits hat erst im Jahr 2008 ein Zensus im Land stattgefunden, deshalb sind die für benötigten Bevölkerungszahlen gut zu extrapolieren. Die aussagekräftige Basisreproduktionszahl \mathfrak{R}_0 stimmt in Liberia am längsten mit der tatsächlichen Nettoerproduktionszahl \mathfrak{R}_t überein. In Guinea gilt diese Annahme bis zum 30. März 2014, in Sierra Leone bis zum 6. Juli 2014 und in Liberia sogar bis zum 24. August 2014 [25]. Da \mathfrak{R}_t in dem SEIR Modell konstant bleibt, ist es vorteilhaft, wenn dieser Wert in der Realität auch möglichst lange konstant ist. $\mathfrak{R}_{t,Guinea}$ schwankt zwischen März und Juli um den Schwellwert $\mathfrak{R}_{t,Guinea} = 1$, steigt im August dann, weil in einer Region des Landes die Fallzahlen deutlich steigen. In Sierra Leone ist der Wert einerseits erst ab Mitte Juni berechnet worden, zum anderen sinkt er auch bis August, da sich die Lage einigermaßen stabilisierte. In Liberia galt zwischen März und August nahezu dauerhaft $\mathfrak{R}_{t,Liberia} > 1$. Dies spiegelt das stetige Ansteigen der Fallzahlen wieder [25].

Das erstellte Modell sollte natürlich möglichst nahe an dem realen Verlauf der Epidemie liegen. Als Datengrundlage wurden deshalb die von der US-Behörde Centers for Disease Control and Prevention (CDC) veröffentlichten Fallzahlen verwendet [3]. Hierbei wurden die Krankheits- bzw. Todesfälle aufsummiert. Diese wurden tageweise aufgelistet, häufig sind jedoch auch Tage ohne Fallzahlen enthalten. Dies betrifft mitunter nur einzelne Tage, geht jedoch bis zu einer Anreihung von 9 Tagen ohne aktuelle Daten. Da das erstellte SEIR Modell mit in konstanten Zeitschritten Daten ausgibt, mussten die realen

Rohdaten entsprechend geschätzt werden. Als Schrittweite wurde, wie bei den zwei WHO Modellen, ganze Wochen gewählt. Als Anfangszeitpunkt wurde Dienstag, der 25. März 2014, gewählt, da hier die ersten Fallzahlen von der CDC veröffentlicht wurden. Dieses Datum steht im Folgenden für den Wochenanfang der Woche 0. Die Abschätzung der unbekanntes Fallzahlen war nur dann nötig, wenn an einem Dienstag keine Daten veröffentlicht wurden. In diesem Fall wurden mit einer linearen Interpolation von den zwei benachbarten Randpunkten die Werte ermittelt.

Unstimmigkeiten in den Rohdaten erzwangen die Grenzen der simulierten Zeitspanne T_{sim} . Als untere Grenze wurde der 5. Juni, 10 Wochen nach dem Anfangszeitpunkt der Datenveröffentlichung, gewählt. Von der 1. bis zur 4. Woche steigt die Zahl der Krankheitsfälle von 8 auf 34. Eine Woche später sind nur noch 13 bekannt. Da die Krankheits- bzw. Todesfälle aufsummiert werden, ist ein Rückgang der Zahlen bei fortlaufender Zeit unlogisch. Einen solchen unlogischen Sprung gibt es nochmal zwischen der 6. und 7. und zwischen der 7. und 8. Woche. Bis zur 9. Woche bleiben die Zahlen konstant. In der 10. Woche gibt es dann einen weiteren Krankheitsfall, danach erfolgen längere Zeit keine weiteren unlogischen Sprünge. Somit bietet sich die 10. Woche als Startzeitpunkt an.

Die obere Grenze von T_{sim} wird ebenfalls durch einen solchen unlogischen Sprung nahe gelegt. Von der 30. auf die 31. Woche steigt die Zahl der Krankheitsfälle von 4584 auf 6068. Dieser Anstieg um 1483 Fälle in einer Woche ist außergewöhnlich, da vorher und nachher Anstiege zwischen 250 und 750 verzeichnet wurden. Zusätzlich reduziert sich die Zahl der Todesfälle von 2661 auf 2486 um 175. Dies ist ein weiterer unlogischer Sprung. Im Zeitraum zwischen der 10. und 30. Woche gibt es ansonsten keine augenscheinlichen Auffälligkeiten, darum wurde dies als Simulationszeitraum gewählt. Dennoch können die Ursachen der o.g. Unstimmigkeiten auch hier für Verfälschungen gesorgt haben. Dies zu untersuchen ist jedoch nicht möglich.

Eine mögliche Erklärung für die anfänglichen Ungereimtheiten zwischen der 4. und 8. Woche mag in der Unterteilung der verschiedenen Fälle liegen. Wenn anfangs ein Patient laut offizieller Definition [18] als möglicherweise mit Ebola infiziert galt, kann er in die Statistik eingegangen sein. Wenn sich der Verdacht im Nachhinein jedoch nicht bestätigte, müsste er wieder herausgerechnet werden. Wenn dies nur teilweise geschah, der Patient zwar in den zukünftigen Statistiken gestrichen wird, jedoch nicht in den vergangenen, wäre dies ein Grund für den unlogischen Sprung. Vor allem in der Anfangszeit des Ausbruchs

erscheint dieses Szenario, mit Blick auf die Umstände im Land (vgl. Kapitel 3.1) und die verbreitete Unwissenheit über die Krankheit, plausibel.

In der 19. Woche, am 8. August 2014, wurde der internationalen Gesundheitsnotstand ausgerufen. Mitte September, um die 25. Woche herum, lief die internationale Hilfe an [25]. Mit dem vermehrten Eintreffen von Experten und Helfern erscheint es plausibel, dass einen guten Monat später eine Überarbeitung der Statistiken durchgeführt wurde. Dies wäre eine mögliche Erklärung für die Unstimmigkeiten zwischen der 30. und 31. Woche. Mit dem Beginn der internationalen Hilfsmaßnahmen wurde Einfluss auf die Ausbreitung der Epidemie genommen. Aus humanitärer Sicht ist dies natürlich erfreulich. Für die mathematische Modellierung des Krankheitsverlaufs bedeutet das auf der anderen Seite, dass sich hier auch die Modellparameter ändern müssten. Aus diesem Grund ist die Wahl der oberen Grenze des Simulationszeitraums ebenfalls akzeptabel.

3.3.2 Das multisoziale multiregionale SEIR Modell

In Abschnitt 3.1 wurde beschrieben, dass es in Liberia unterschiedlich verfügbaren Zugang zu medizinischer Versorgung gab. Vor allem die Bezahlbarkeit und die Entfernung zur nächstgelegenen, entsprechenden Einrichtungen wurden als hauptsächliche Hinderungsgründe genannt [9]. Deshalb ist es sinnvoll, die Bevölkerung nicht als homogen zu betrachten, sondern verschiedene Bevölkerungsgruppen zu unterteilen entsprechend ihrer Möglichkeiten, medizinische Versorgung in Anspruch nehmen zu können. Aus diesem Grund wurden die Liberianer in vier Bevölkerungsgruppen unterteilt. Dabei wurde einmal gemäß der Zahlen der Afrikanischen Entwicklungsbank zwischen arm (englisch: **poor**) und reich (englisch: **rich**) unterschieden, wobei in dieser Arbeit als arm gilt, wer unter 10\$ pro Tag verdient [16]. Zum anderen wurde gemäß neuester UNICEF Studien zwischen Stadt- und Landbevölkerung (englisch: **Urban** und **Rural**) differenziert [26]. Die entstandenen vier Bevölkerungsgruppen sind in Tabelle 3.1 aufgeführt.

Tabelle 3.1: Einteilung der liberianischen Bevölkerung in vier Gruppen.

	arme Bevölkerung (p)	reiche Bevölkerung (r)
Stadtbevölkerung (U)	pU	rU
Landbevölkerung (R)	pR	rR

Die CDC geht zu Beginn der Epidemie von einer Gesamtbevölkerung von $N = 4.294.000$ Menschen aus. Mit den Statistiken der Afrikanischen Entwicklungsbank und von UNICEF entsteht die Anfangspopulation des Modells gemäß Tabelle 3.2.

Tabelle 3.2: Verteilung der Bevölkerung zu Beginn der Epidemie.

$N = 4.294.000$	arme Bevölkerung (p)	reiche Bevölkerung (r)	%
Stadtbevölkerung (U)	1.878.625 (43.75%)	203.965 (4.75%)	48,5%
Landbevölkerung (R)	2.007.445 (46.75%)	203.965 (4.75%)	51,5%
%	90,5%	9,5%	100%

Das der Ausbruch im ländlichen Grenzgebiet von Guinea, Liberia und Sierra Leone begann und sich schnell ganzheitlich in allen drei Ländern, auch in die Städte hinein, verbreitete, zeigt dass die Bevölkerungsgruppen sich gegenseitig ansteckten. Es fand eine Durchmischung untereinander statt. Diese Durchmischung kann mathematisch mit Hilfe der Kontaktrate angegeben werden. Die Bestimmung dieser Kontaktraten zwischen den einzelnen Gruppen ist daher entscheidend für das Modell.

Da der simulierte Zeitraum nur einige Wochen beträgt, haben die natürlichen Geburten- und Sterberaten keinen Einfluss auf die Anzahl der Gesamtbevölkerung. Ebenso wird die Migration über die Systemgrenze nicht berücksichtigt, um das Modell nicht noch komplexer zu gestalten. Obwohl die Gesamtbevölkerung durch die Einteilung in die verschiedenen Bevölkerungsgruppen als inhomogen betrachtet wird, werden die einzelnen Gruppen wiederum als homogen verteilt angenommen. Damit wird impliziert, dass jedes Individuum der Gruppe die selbe Wahrscheinlichkeit hat, infiziert und geheilt zu werden bzw. zu sterben, weil die Durchmischung in der Gruppe als gleichmäßig betrachtet wird. Die Inhomogenität der Gesamtbevölkerung trägt mit gruppenspezifischen Raten β_{xy} und γ_{xy} den unterschiedlichen Ansteckungsrisiken und Heilungschancen Rechnung.

Die Berechnungsformeln entsprechend größtenteils den bekannten Formeln eines gewöhnlichen SEIR Modells, wie in Unterabschnitt 3.2.1 beschrieben. Der wichtige Unterschied besteht nun darin, dass die vier Bevölkerungsgruppen mit ihrem jeweils eigenen SEIR System aufeinander Einfluss nehmen. Dies kann sich jedoch nur auf die Ansteckungsrate γ auswirken. Dieser Parameter wird lediglich zur Berechnung des Klassenwechsels

von S nach E verwendet. Ansonsten sind die Bevölkerungsgruppen quasi entkoppelt voneinander. Der Klassenwechsel spiegelt sich in \dot{S} und \dot{E} wieder. Wie in Unterabschnitt 3.2.1 erklärt, ist der jeweilige Summand betragsmäßig gleichgroß. Da \dot{S} nur aus diesem Summanden besteht, wird die Kontaktrate der Bevölkerungsgruppen daran exemplarisch erläutert.

Es wird analog zu Kapitel 12 in [20], wo ein SIS Modell mit $2n$ -Unterklassen, wie den verschiedenen Bevölkerungsgruppen, entworfen wurde, beim multisozialen, multiregionalen SEIR Modell vorgegangen. Es werden vier Bevölkerungsgruppen betrachtet, demnach handelt es sich bei dem Modell um ein solches $2n$ -Unterklassen Modell mit $n = 2$. Die Klassen an sich werden mit $x = 1, \dots, 2n$ indiziert. Für jede x -te Unterklasse gilt:

$$S_x + E_x + I_x + R_x = N_x > 0, \text{ mit } N_x = \text{konstant.} \quad (3.4)$$

Der Kontaktrate eines Gesunden der Klasse x mit einem Kranken der Klasse y wird mit γ_{xy} angegeben. Für \dot{S} gilt dann allgemein:

$$\dot{S}_x = \frac{dS_x}{dt} = - \sum_{y=1}^n \gamma_{xy} S_x I_y. \quad (3.5)$$

In dem Modell soll u.a. berücksichtigt werden, wie wahrscheinlich Personen aus einer der vier Bevölkerungsgruppen sich anstecken können. Dafür wird die allgemeine Kontaktrate γ_{xy} aufgeteilt in:

$$\gamma_{xy} = \gamma_x \cdot \vartheta_{xy}, \text{ mit der Transmissionsrate } \vartheta_{xy}. \quad (3.6)$$

γ_x gibt nun, wie bei den klassischen SIR Modellen üblich, die allgemeine Ansteckungsrate einer Person der Unterklasse x an, ϑ_{xy} beschreibt das Kontaktverhalten zu anderen Unterklassen. Dabei ist $\vartheta_{xx} = 1$, weil die eigene Unterklasse homogen durchmischt ist. Dadurch ergibt sich:

$$\begin{aligned}
\dot{S}_x &= \frac{dS_x}{dt} = - \sum_{y=1}^n \gamma_{xy} S_x I_y, \\
&= -\gamma_x S_x \sum_{y=1}^n \vartheta_{xy} I_y, \\
&= -\gamma_x S_x \cdot \left(I_x + \sum_{\substack{y=1 \\ x \neq y}}^n \vartheta_{xy} I_y \right).
\end{aligned} \tag{3.7}$$

In Abbildung 3.2 ist das multiregionale, multisoziale SEIR Modell abgebildet, wie es in der Software Vensim von Ventana implementiert wurde. Zur Berechnung wurde hier ein Runge-Kutta-Verfahren 4. Ordnung mit automatischer Schrittweitensteuerung verwendet (siehe Abschnitt 3.2 in [14]). Bei dieser Darstellung sind die 16 verschiedenen Unterklassen mit jeweils einem rechteckigen Rahmen dargestellt. Die vorgesehenen Klassenwechsel sind mit Folgepfeilen veranschaulicht worden. Die Differentialgleichungen, welche diese Klassenwechsel beschreiben, sind in der Mitte dieser Pfeile in den „get“-Funktionen implementiert worden. Parameter und Variablen, die in diesen Gleichungen enthalten sind, zeigen mit jeweils einem blauen Pfeil auf die entsprechende Funktion.

Das Modell berücksichtigt zwei Infektionsraten, dabei wurde jeweils die Stadt- und die Landbevölkerung zusammengefasst. Die in Abschnitt 3.1 ermittelten Unterschiede zwischen arm und reich betrafen vor allem den Zugang zu medizinischer Versorgung. Das Risiko der Ansteckung ist davon nicht beeinflusst. Hier spielt die räumliche Nähe zu Infizierten eine entscheidendere Rolle [15]. Deshalb reicht es, nur zwei verschiedene Infektionsraten anzugeben. Diese Ansteckungsraten sind in der Softwareimplementierung mit b bezeichnet, da griechische Buchstaben nicht verfügbar waren. Aus dem selben Grund wurde die Heilungs- und Sterberate mit a bezeichnet.

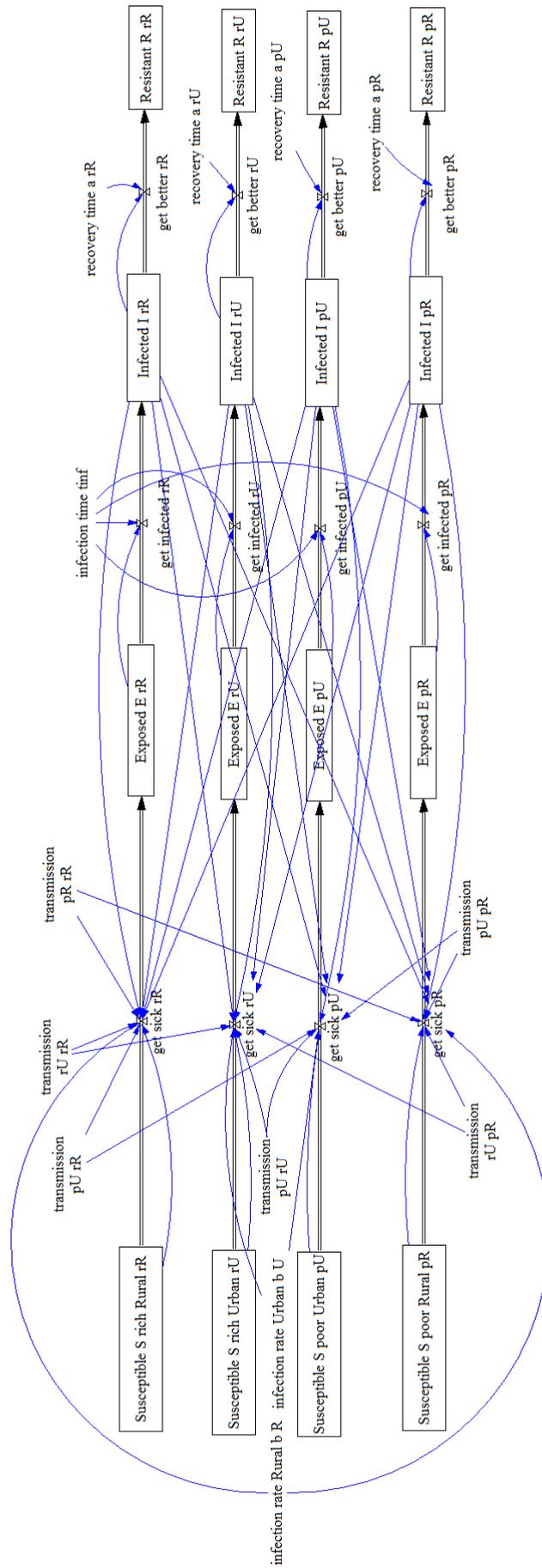


Abbildung 3.2: Multisoziales multiregionales SEIR Modell mit vier Bevölkerungsgruppen

Die Kontaktraten (englisch: transmission rate) konnten mit keiner Statistik quantifiziert werden, um das Modell nicht unnötig komplex zu machen, wurde deshalb für die Kontaktraten $\vartheta_{xy} = \vartheta_{yx}$ angenommen. Dementsprechend wurden sechs dieser Parameter implementiert.

Wie bereits in Unterabschnitt 3.3.1 beschrieben wurde, ist der Simulationszeitraum mit $10 < T_{sim} < 30$ gewählt worden. Als Zeitschritt des Runge-Kutta-Verfahrens wurde die kleinstmögliche Schrittweite von 0,0078125 gewählt, da es bei größeren Schrittweiten zu numerischen Integrationsfehlern kam. Als Zeiteinheit wurde „Wochen“ eingestellt. Ist ein Patient infiziert worden, nimmt die Krankheit keine Rücksicht auf seine soziale oder regionale Herkunft. Darum wurde nur eine globale Inkubationszeit verwendet.

Die Chance und Geschwindigkeit einer Heilung hingegen sind unterschiedlich, dementsprechend wurde für jede Bevölkerungsgruppe eine individuelle Heilungsrate implementiert. Wie in Unterabschnitt 3.3.1 beschrieben, sind die veröffentlichten Daten der ersten Wochen unstimmig. Deshalb geht dieses Modell von einem Infektionsbeginn zu Beginn der 10. Woche aus. Es wurde angenommen, dass die Krankheit von einem Infizierten über die Grenze nach Guinea eingeschleppt wurde. Diese Grenzregion kann als ländlich angesehen werden. Der Großteil der Landbevölkerung ist arm, deshalb wurde der Indexpatient der Unterklasse pR zugeordnet. Die Anfangswerte der gesunden Bevölkerung können Tabelle 3.2 entnommen werden. Es wurde angenommen, dass niemand in Liberia immun gegen Ebola war und der Indexpatient noch niemanden anstecken konnte. Deshalb sind die Anfangswerte aller Unterklassen für E , R und alle übrigen I auf Null gesetzt worden.

Die Systemparameter wurden aus den Daten der WHO übernommen, falls sie dort angegeben worden waren. Die restlichen Parameter wurden abgeschätzt. Von der WHO wurden gute Schätzungen zu folgenden Parameter veröffentlicht: Basisreproduktionsrate \mathfrak{R}_0 , durchschnittliche Inkubationszeit und die durchschnittliche Zeitspanne vom Melden des Krankheitsfall bis zur Heilung bzw. bis zum Tod. Diese stimmen auch einigermaßen mit Schätzungen von vorangegangenen Ebola-Ausbrüche überein. Damit fehlen die Parameter: Infektionsrate für Stadt- und Landbevölkerung und die sechs Kontaktraten zwischen den Bevölkerungsgruppen.

Wie in Unterabschnitt 3.2.1 beschrieben, kann mit dem Reziprokwert der durchschnittlichen Inkubationszeit (9,4 Tage) die Inkubationsrate berechnet werden. Mit der durchschnittlichen Zeitspanne vom Auftreten der Symptome bis zur Heilung (15,4 Tage) bzw.

bis zum Tod (7,9 Tage) wurde die durchschnittliche Heilungsrate ermittelt. Dafür wurde entsprechend der Letalität (67%) ein qualifizierter Mittelwert der beiden Zeitspannen gebildet (10,375 Tage). Es wurden die Zeitspannen vom Auftreten der Symptome an gewählt, weil von da an der Patient als ansteckend gilt und somit der Klasse I zuzuordnen ist. Erst danach konnten die entsprechenden Werte in das Modell eingepflegt werden. Hierbei musste noch beachtet werden, dass die Angaben der WHO in der Zeiteinheit „Tage“ veröffentlicht wurden, das Modell rechnet jedoch mit „Wochen“. Deshalb mussten die WHO Daten durch die Anzahl an Tagen pro Woche (sieben), geteilt werden. Die spezifischen Heilungsraten der Bevölkerungsgruppen wurden um diesen Wert herum variiert. Dabei wurde die Rate für die reichen Gruppen erhöht, weil sie sich eher die medizinische Versorgung leisten konnten und somit schneller geheilt werden konnten.

Auch die durchschnittliche Infektionsrate kann mit den vorhandenen Parametern leicht errechnet werden. Laut Abschnitt 3.2.1 ist:

$$\mathfrak{R}_0 = \frac{\beta N}{\gamma}, \text{ und es ergibt sich :} \quad (3.8a)$$

$$\beta = \frac{R}{N} \cdot \gamma. \quad (3.8b)$$

Diese durchschnittliche Infektionsrate wurde für die Stadtbevölkerung ein wenig erhöht und im gleichen Maße für die Landbevölkerung gesenkt, denn in der Stadt ist die Ausbreitung der Krankheit eher möglich [15]. Die Kontaktraten stellen gemäß Gleichung (3.6) einen Skalierungsfaktor dar. Diese Werte mussten frei gewählt werden. Lägen sie alle bei 1, würde ausschließlich die durchschnittliche Infektionsrate gelten. Deshalb wurden die Kontaktraten um den Wert 1 herum variiert. Hierbei wurde die Durchmischung innerhalb einer regionalen bzw. sozialen Gruppe als besonders hoch angenommen.

Nachdem diese Daten in das Modell implementiert worden sind, war der prognostizierte Krankheitsverlauf ernüchternd. Der Verlauf (siehe Abbildung 3.4) stimmte nicht annähernd mit den tatsächlichen Daten überein (siehe Abbildung 3.3).

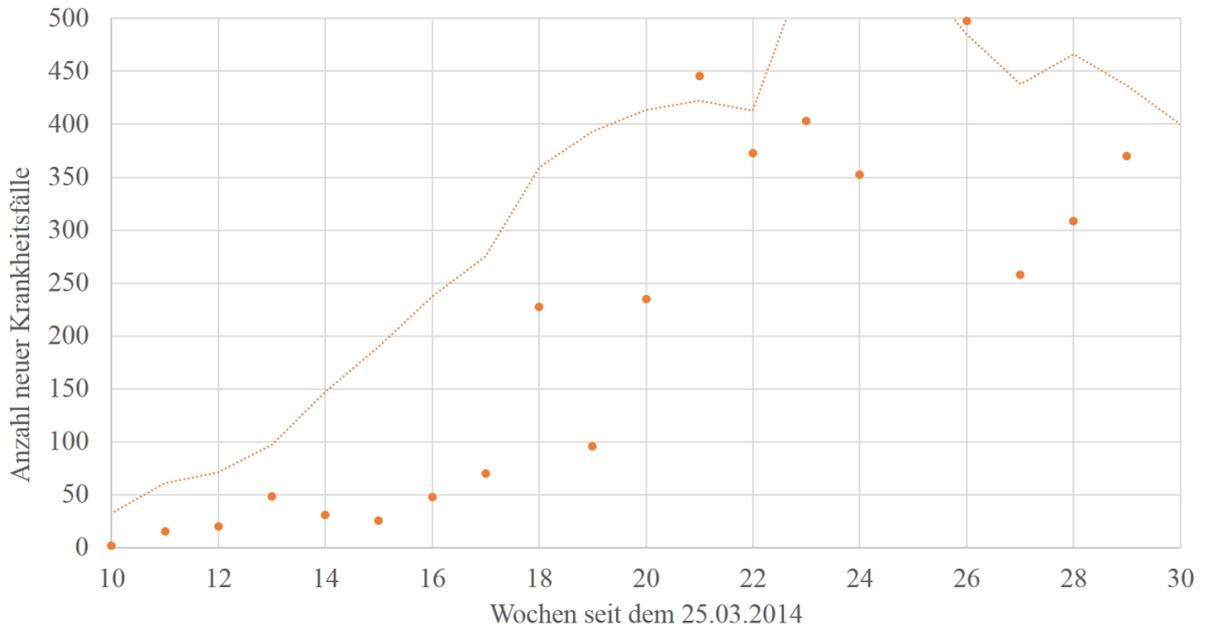


Abbildung 3.3: Falldaten der CDC bis 30. Woche, mit gleitendem Mittelwert (gestrichelt).

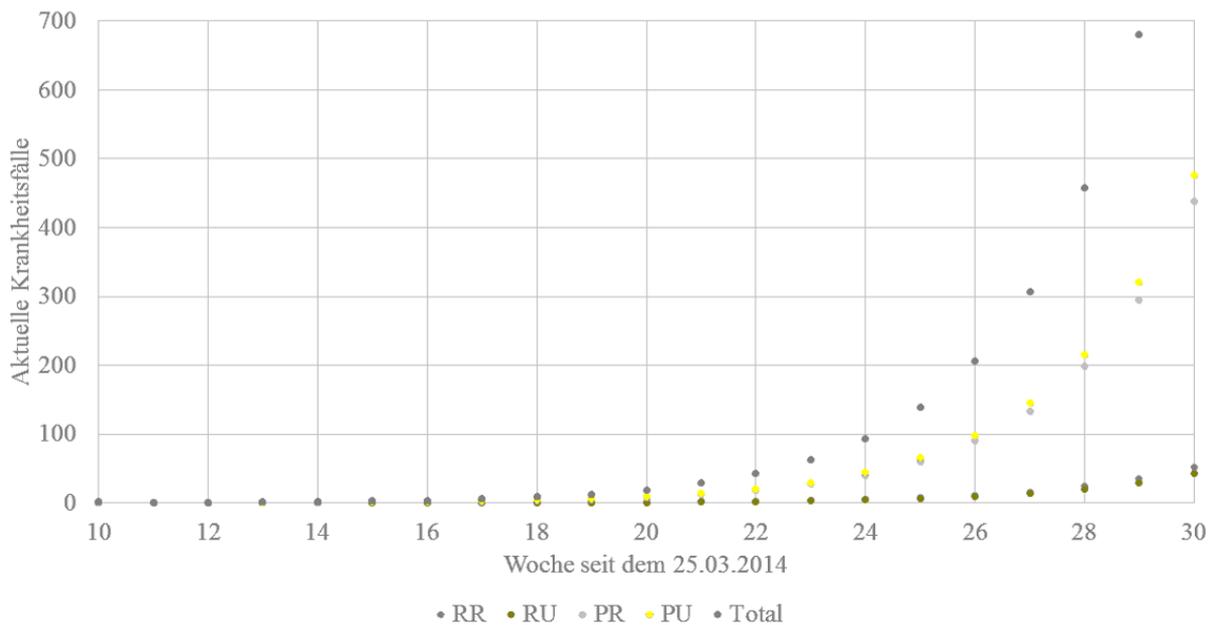


Abbildung 3.4: Krankheitsverlauf bis 30. Woche gemäß SEIR Modell mit WHO Daten.

Erst nachdem der Simulationszeitraum bis zur 60. Woche verlängert wurde (siehe Abbildung 3.5), konnte man erahnen, dass das Modell qualitativ eine Epidemie beschreiben soll. Allerdings stimmt die Dimension der Krankheitsfälle pro Woche nicht ansatzweise

mit den realen Daten überein. Der höchste Zuwachs wird in der 25. Woche mit 744 neuen Fällen beziffert. Zusätzlich gibt es noch den in Unterabschnitt 3.3.1 schon erwähnten, unlogischen Ausreißer mit 1483 neuen Krankheitsfällen in der 30. Woche. Die Beträge der Modelle unterscheiden sich ungefähr um den Faktor 1.000. Zusätzlich stimmen die zeitlichen Verläufe nicht überein. Auch ein leichtes Variieren der Modellparameter brachte kein besseres Ergebnis. Für die gewünschte Simulation in Abschnitt 3.4 mussten die Werte mit den realen Fallzahlen kalibriert werden. Diese Kalibrierung wird in Unterabschnitt 4.2.1 beschrieben.

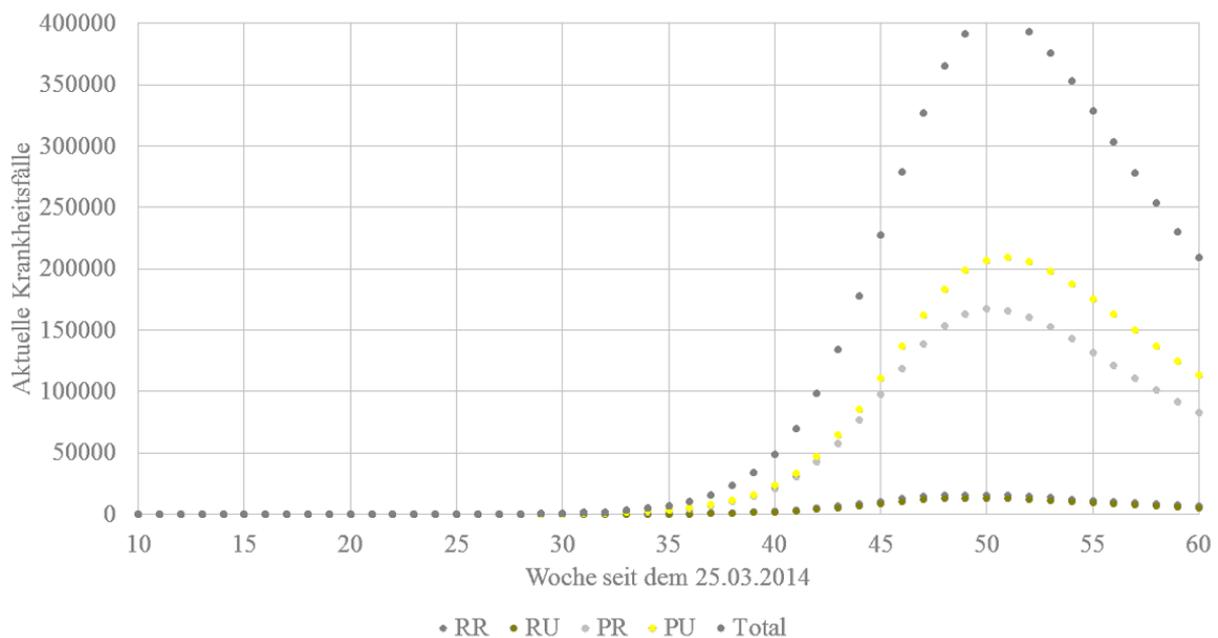


Abbildung 3.5: Krankheitsverlauf bis 60. Woche gemäß SEIR Modell mit WHO Daten.

3.4 Simulation eines syndrombasierten nahezu Echtzeit-Überwachungssystems

Das Ausmaß der Ebola-Epidemie hatte alle bisherigen Vorstellungen übertroffen. Die internationale Gemeinschaft begann erst ab dem 8. August 2014, viereinhalb Monate nach Beginn der Falldatenerhebungen, mit dem Ausrufen des internationalen Gesundheitsnotstands, die Krise ernst zu nehmen [25]. Für die Fragestellung, wie man das Ausmaß der Krankheit schneller hätte ermitteln können, ist die Einführung eines nahezu Echtzeit-Überwachungssystems eine plausible Lösung. Ein solches Überwachungssystem muss in der Lage sein, jede Inanspruchnahme von medizinischer Versorgung zu dokumentieren. Im

Idealfall werden die Daten nach der Eingabe am Endgerät in nahezu Echtzeit übermittelt und ausgewertet. Im DHSC NATO Military Medical Centre in München wird ein Prototyp (genannt ASTER) eines solchen Überwachungssystems betrieben. Dieses System arbeitet syndrombasiert, d.h. dass keine Differentialdiagnosen wie etwa Gelbfieber oder Magen-Darm-Grippe erfasst werden, sondern lediglich 69 verschiedene Symptome und Syndrome, wie zu hohe Körpertemperatur oder Durchfall. Der Vorteil hiervon ist, dass diese auch von nicht-professionellem Personal erfasst werden könnten. Die Dateneingabe für dieses System wird derzeit jedoch nur von medizinischem Fachpersonal getätigt. Das System summiert die verschiedenen, eingegebenen Syndrome nach Regionen und erzeugt automatisch eine Warnung, wenn ein speziell festgelegter Schwellwert überschritten wird. Infolge dieser Warnung muss dann ein entsprechend geschulter Arzt entscheiden, ob die Schwellwertüberschreitung eventuell saisonbedingt ist, oder ob eine Epidemie droht.

Die Medical Intelligence Einheit der Bundeswehr, ebenfalls in München stationiert, entwickelt gerade einer App (EMILIA), die die Dateneingabe für ein Überwachungssystem per Smartphone ermöglichen soll. In der vorliegenden Arbeit wird nun davon ausgegangen, dass beide Systeme kombiniert würden, d.h. dass die Dateneingabe für das nahezu Echtzeit-Überwachungssystem per Smartphone möglich wäre.

In Abschnitt 3.1 wurde jedoch bereits bemerkt, dass es in Liberia zu wenig medizinisches Personal gab. Das vorhandene Personal und auch internationale Helfer waren, wie in Unterabschnitt 3.3.1 beschrieben, voll damit beschäftigt, die Flut an Patienten zu versorgen. Die Einführung eines Überwachungssystems ist bei dem vorhandenen Gesundheitssystem zumindest kritisch zu hinterfragen.

Abhilfe könnte hier die Kombination eines solchen Überwachungssystems mit freiwilligen, nicht-professionellen Helfern, den CHWs, als Datenerheber leisten. Im Vorfeld dieser Arbeit wurde bereits, im Rahmen einer anderen Master Thesis [24], ein stochastisches Modell dieser innovativen Idee erstellt. Dieses CHW Modell ist statisch. Um seine Aussagekraft zu verbessern, wird es in diesem Abschnitt durch die Anwendung auf das multisoziale, multiregionale SEIR Modell in ein dynamisches gewandelt.

Das CHW Modell unterteilt die liberianische Bevölkerung in die selben vier Bevölkerungsgruppen wie das SEIR Modell. In Abschnitt 3.4 stand die Modellierung der Ebola-Ausbreitung im Vordergrund. Nun geht es darum, mit welchen Wahrscheinlichkeiten Patienten der verschiedenen Bevölkerungsgruppen an der medizinischen Versorgung und

Datenerhebung teilhaben können. In dem Modell wird dabei zwischen formeller und informeller Versorgung unterschieden. Unter Verwendung verschiedener Studien und Statistiken und mit dem Treffen einiger Annahmen konnten sowohl Anzahl als auch Art der Inanspruchnahme von medizinischer Hilfe durch die unterschiedlichen Bevölkerungsgruppen berechnet werden. Daraus wurde abgeleitet, mit welcher Wahrscheinlichkeiten ein Patient formelle oder informelle medizinischer Versorgung aufsucht. Dabei wurde vorausgesetzt, dass durch den Einsatz von CHWs diese als alleinige Anbieter der informellen Versorgung agieren. Das Ergebnis dieser Berechnungen ist in Tabelle 3.3 aufgeführt.

Tabelle 3.3: Wahrscheinlichkeiten für den Besuch von formeller oder informeller medizinischer Versorgung aufgeschlüsselt nach den vier Bevölkerungsgruppen.

	$p(\text{formelle Versorgung})$	$p(\text{CHW})$
arme Landbevölkerung pR	23%	77%
arme Stadtbevölkerung pU	21%	79%
reiche Landbevölkerung rR	76%	24%
reiche Stadtbevölkerung rU	100%	0%

Der flächendeckende Einsatz von CHWs zur Datenerhebung führt natürlich auch dazu, dass nicht nur Daten von der betrachteten Krankheit gemeldet werden, sondern nun vielmehr das gesamte Spektrum medizinischer Hilfeleistungen weitergemeldet werden würde. Dadurch entsteht ein Datenrauschen, aus dem heraus der Ausbruch einer Epidemie erkannt werden müsste. Das CHW Modell geht von einem wöchentlichen Rauschen von 2000 unterschiedlichen Dateneingaben aus.

Zusätzlich zu diesem Rauschen zieht der Einsatz von CHWs einen weiteren Aspekt mit sich mit, der unbedingt berücksichtigt werden muss: Die CHWs sind kein professionelles Medizinpersonal, sondern lediglich für ihre Aufgabe kurz geschult worden. Daher unterlaufen ihnen natürlich Fehler bei der Feststellung von Patientensyndromen, wodurch dem Überwachungssystem falsche Daten gemeldet werden. Die Genauigkeit bei der Diagnose von Symptomen wird in der Medizin durch die Sensitivität und Spezifität beschrieben, beides wurde bereits für beim Einsatz von CHWs erforscht [24]. Im Rahmen der Erstellung des CHW Modells wurden die Untersuchungsergebnisse zusammengefasst. Das

Resultat ist in Tabelle 3.4 aufgeführt. Wird ein kranker Patient von einem CHW untersucht, wird mit einer Wahrscheinlichkeit von 70,3% die Krankheit als solche erkannt. In 29,7% der Fälle bleibt sie jedoch unerkannt. Andersherum werden gesunde Patienten mit einer Wahrscheinlichkeit von 7,6% als krank eingestuft.

Tabelle 3.4: Sensitivität und Spezifität von CHWs bei der Datenerhebung.

Genauigkeit (s) von CHWs	Wahrer Krankheitszustand		
		Krank	Gesund
Test	Positiv	70,3%	7,6%
Resultate	Negativ	29,7%	92,4%

Im Verlauf der Epidemie werden sowohl die normalen Fälle des Datenrauschens, als auch die Krankheitsfälle der Epidemie untersucht. Dabei werden jeweils in einigen Fällen Fehler gemacht. Die Vergleichsgruppe zu den CHWs sind die professionellen Mediziner, denen hier unterstellt wird, dass sie jeden Fall richtig diagnostizieren. Für den Vergleich wird die Ebola-Epidemie von 2014 verwendet, wobei davon ausgegangen wird, dass das kalibrierte SEIR Modell aus Unterabschnitt 4.2.1 den wahren Krankheitsverlauf abbildet.

Im Folgenden wird angenommen, dass jeder Kranke auch medizinische Hilfe in Anspruch nimmt. Anhand des multisozialen, multiregionalen Modells kann festgestellt werden, wieviele Personen jeder Bevölkerungsgruppe in jeder Woche des betrachteten Zeitraums T_{sim} an Ebola erkrankt waren. Mit den Wahrscheinlichkeiten für das Aufsuchen von formeller oder informeller Hilfe ($p(k)$) abhängig von der Bevölkerungsgruppe k des Patienten und den Genauigkeiten der CHWs ($s(CHW)$) lässt sich so abschätzen, wie viele Fälle pro Woche gemeldet worden wären. Dafür gilt die Formel:

$$I_{Meldung, CHW} = \sum_{k=1}^x p(k) \cdot s(CHW) \cdot I_k, \quad (3.9)$$

bzw. für die professionellen Mediziner:

$$I_{Meldung, Med} = \sum_{k=1}^x p(k) \cdot I_k. \quad (3.10)$$

4 Ergebnisse

In diesem Kapitel wird die entwickelte Simulation ausgewertet. Zuerst werden die Unterschiede zu den zwei vorgestellten Modellen der Yale Universität und der WHO erläutert. Anschließend wird das Modell mit den realen Daten kalibriert und letztendlich im Rahmen der Simulation einer syndrombasierten nahezu Echtzeit-Überwachung der Einsatz von Community Health Workers (CHWs) dem bisherigen Gesundheitssystem gegenübergestellt.

4.1 Abgrenzung zu bisherigen SEIR Modellen

Der offensichtlichste Unterschied zu den Modellen der Yale Universität und der WHO ist die multisoziale und multiregionale Unterteilung der Bevölkerung. Obwohl die anfänglichen Simulationen, mit den aus der Literatur übernommenen Werten, einen deutlichen Dimensionsfehler aufwiesen, konnte das Modell nach der Kalibrierung mit den realen Fallzahlen einen plausiblen Verlauf der Krankheit berechnen.

Die WHO sagte aufgrund ihrer Modelle für den 2. November 9890 Krankheitsfälle voraus in Liberia voraus. In der Realität wurden bis zum 5. November 6535 Fälle verzeichnet. Die Modelle hatten den wahren Verlauf um ca. 50% überschätzt. Das Hauptmodell der WHO stützt sich auf einen Poisson-Prozess. Aus dieser Vorgehensweise war es nicht möglich, Erkenntnisse über die Ausbreitung der Krankheit in verschiedenen Bevölkerungsgruppen zu übernehmen. Dies könnte, wenn überhaupt, nur mit der Neudefinition verschiedener Zeitspannen, ähnlich dem verwendeten „serial interval“, umgesetzt werden. Dies modelliert ja indirekt den Klassenwechsel von I nach R . Wie in Abschnitt 3.1 beschrieben wurde, scheint das Berücksichtigen der verschiedenen Bevölkerungsgruppen jedoch angebracht zu sein. Eine solche Implementierung ist in einem SEIR Modell deutlich leichter realisierbar. Bei beiden WHO Modellen ist die Bestimmung des Peaks der Ausbreitung nicht vorgesehen gewesen. Vor allem das Regressionsmodell würde einen solchen Verlauf niemals vorhersagen können, da es von einem kontinuierlichen exponentiellen Wachstum ausgeht.

Das SEIR Modell der Yale Universität prognostizierte in seinem best-case Szenario, ausgehend von unverzüglicher und umfassender, effektiver Hilfsmaßnahmen für den 15. Dezember 2014, insgesamt 27.378 Krankheits- und 18.606 Todesfällen alleine im Montserado County. Am 17. Dezember 2014 meldete die CDC für das gesamte Liberia 7.797 Krankheits- und 3.290 Todesfälle. Trotz logischer Annahmen und sorgfältiger Prüfung, wurde das Ausmaß der Epidemie deutlich überschätzt. Somit wird deutlich, dass die SEIR Modelle sehr empfindlich bezüglich des Einstellens der Parameter sind.

4.2 Simulation einer Syndromüberwachung

In diesem Kapitel werden die Ergebnisse der Simulation erläutert. Zunächst wurde des erstellte SEIR Modell anhand der gemeldeten Daten kalibriert, anschließend wurde verglichen, ob CHWs oder professionelle Mediziner in Zusammenarbeit mit einem nahezu Echtzeit-Überwachungssystem den Ausbruch der Epidemie schneller melden können.

4.2.1 Kalibrierung des SEIR Modells

Mit der „Trial and Error“ wurden die Parameter zunehmend deutlicher verändert. Dafür wurde ein einfaches SEIR Modell erstellt, um zunächst die Größenordnung der Modellparameter richtig bestimmen zu können.

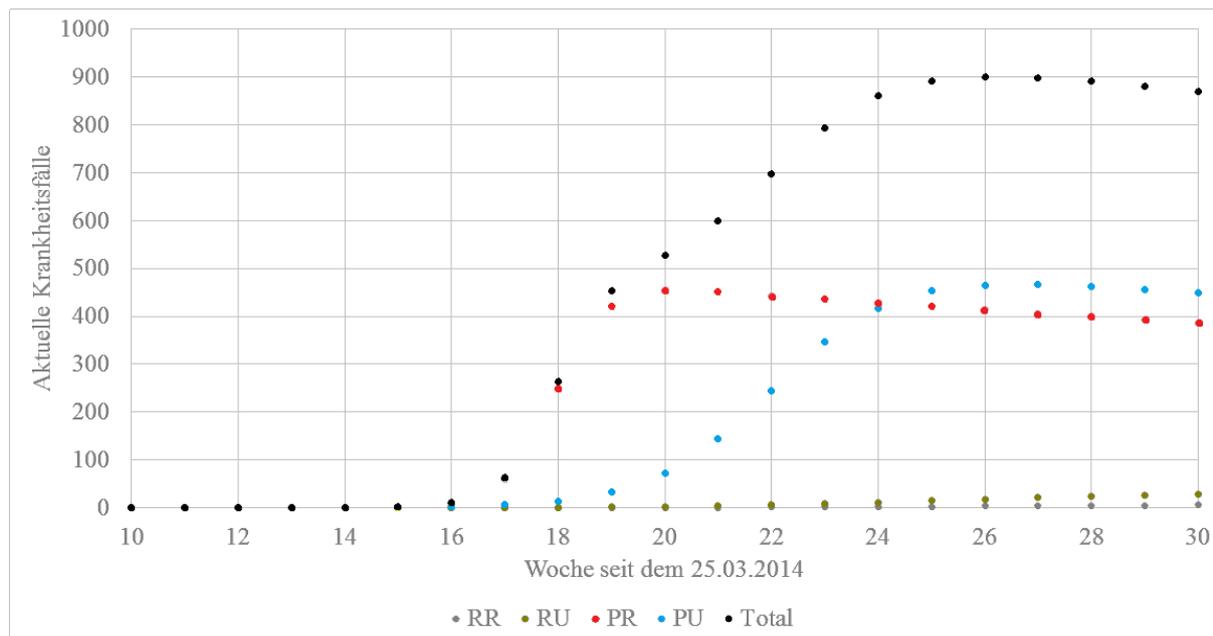


Abbildung 4.1: Krankheitsverlauf bis 30. Woche gemäß kalibriertem SEIR Modell.

Das erste SEIR Modell hatte zwei Probleme im Vergleich zu den realen Daten. Die Spitze des Ausbruchs lag zu spät auf der Zeitachse und war dabei auch um den Faktor 1.000 zu groß. Mit dem Erhöhen der Heilungsrate konnte er runterskaliert werden, weil die Infizierten schneller geheilt wurden bzw. starben und damit weniger Gesunde anstecken konnten. Allerdings verschob sich die Spitze dann auf der Zeitachse in die falsche Richtung nach rechts. Durch Erhöhen der Infektionsrate konnte die Spitze wieder nach links verschoben werden, weil die Menschen sich schneller gegenseitig ansteckten. Der Betrag der wöchentlichen Neuerkrankungen stieg dabei nur unwesentlich. Durch sukzessives Ändern der Parameter und Kalibrieren mit den realen Daten konnte schlussendlich ein befriedigendes Ergebnis erzielt werden (siehe Abbildungen 4.1 und 4.2).

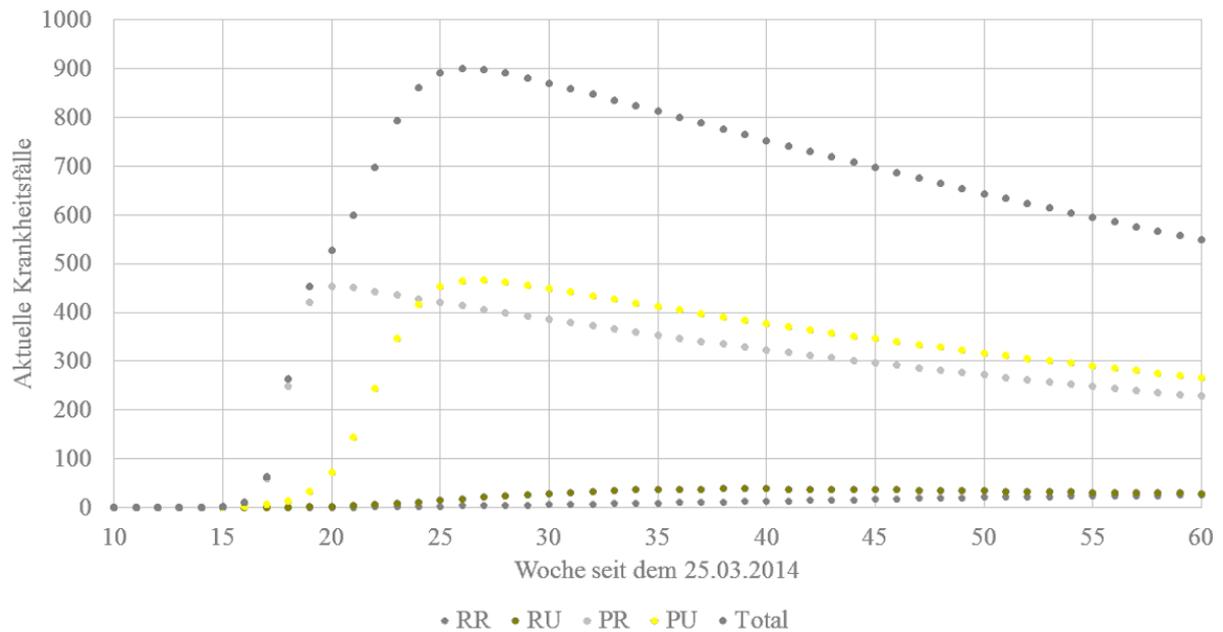


Abbildung 4.2: Krankheitsverlauf bis 60. Woche gemäß kalibriertem SEIR Modell.

Der resultierende Verlauf ist qualitativ mit dem der realen Daten (siehe Abbildung 4.3) nahezu übereinstimmend. Lediglich die Werte sind durchgehend ungefähr um den Faktor 2 größer. Die CDC geht davon aus, dass 60% der aufgetretenen Krankheitsfälle gar nicht erst bekannt geworden sind [13]. Dies entspräche einer Dunkelziffer von 2,5 Fällen pro einem gemeldeten Fall. Damit kann angenommen werden, dass das kalibrierte SEIR Modell dem wahren Verlauf der Epidemie ziemlich nahe gekommen ist.

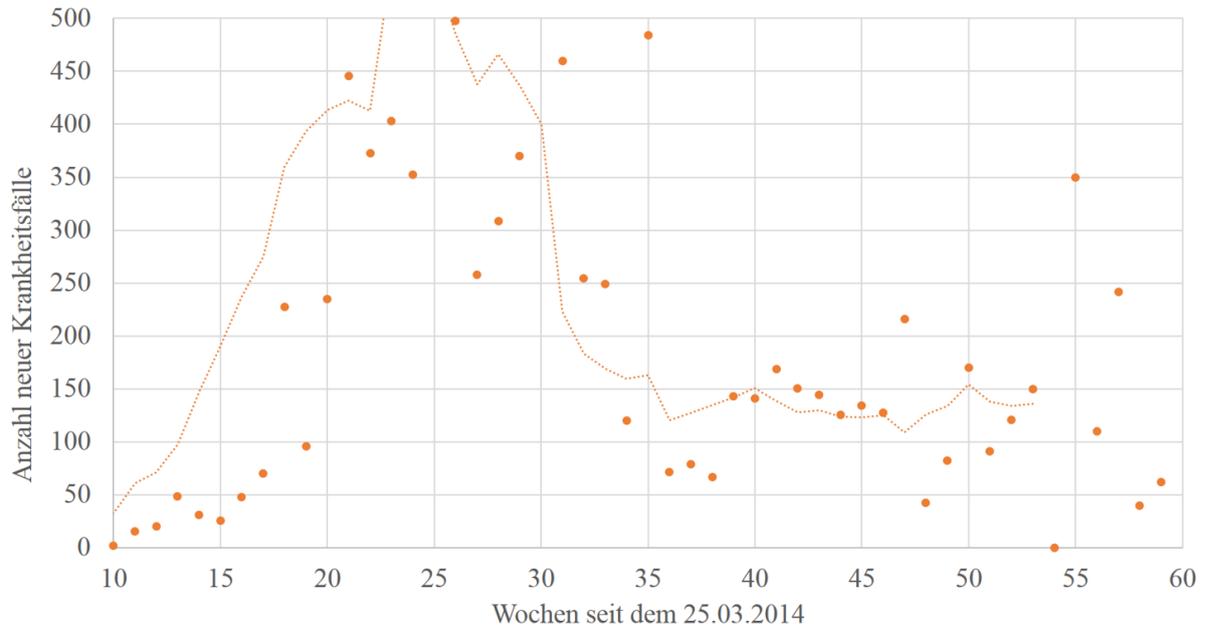


Abbildung 4.3: Falldaten der CDC bis 60. Woche, mit gleitendem Mittelwert (gestrichelt).

Die Parameter beider SEIR Modelle wurden in Tabelle 4.1 aufgelistet. Dabei wird deutlich, wie sehr sie von den ursprünglich ermittelten Werten abweichen.

Tabelle 4.1: Implementierte Parameter des SEIR Modells.

	mit WHO Daten	nach Kalibrierung
durchschnittliche Infektionsrate β	$1,2 \cdot 10^{-6}$	-
durchschnittliche Inkubationsrate ε	0,106	0,0175
Infektionsrate Stadtbevölkerung	$1,25 \cdot 10^{-6}$	0.0045
Infektionsrate Landbevölkerung	$1,15 \cdot 10^{-6}$	0.00175
Kontaktrate pU rR	0,9	0.008
Kontaktrate pU rU	1,5	0.01
Kontaktrate pU pR	0,7	0.007
Kontaktrate rU rR	1,3	0.0085
Kontaktrate rU pR	0,5	0.005
Kontaktrate pR rR	1,4	0.009
durchschnittliche Heilungsrate	0,675	-
Heilungsrate a pR	0,5	68
Heilungsrate a pU	0.6	69
Heilungsrate a rR	0,7	70
Heilungsrate a rU	0,85	71

4.2.2 Vergleich von professionellem Medizinpersonal und CHWs

Für den Vergleich von professionellem Medizinpersonal und CHWs wurde eine Simulation basierend auf unterschiedlichen Wahrscheinlichkeiten für den Besuch bei einem der beiden Anbieter von medizinischer Versorgung und den Genauigkeiten ihrer Diagnosen erstellt. Dabei wurde angenommen, dass beiden Gruppen ein syndrombasiertes, nahezu Echtzeit-Überwachungssystem zur Verfügung steht. Es sollte überprüft werden, welcher der beiden Anbieter mit einem solchen System den Ausbruch der Ebola-Epidemie hätte früher vorhergesagt werden können. Das Resultat der Simulation ist in Abbildung 4.4 dargestellt. Neben den berechneten Werten sind auch die von der CDC gemeldeten Fallzahlen aufgeführt. Überraschenderweise hätte laut Simulationsergebnis die Einführung eines nahezu Echtzeit-Überwachungssystem in Verbindung mit dem Einsatz von CHWs dafür gesorgt, dass später als in der Realität vor der Epidemie gewarnt worden wäre. Dies kann man

daran erkennen, dass die Zahl der von der CDC pro Woche gemeldeten Krankheitsfälle bis zur 25. Woche höher ist, als die von Medizinern oder CHWs gemeldeten Fälle. Am 8. August 2014 wurde der internationale Gesundheitsnotstand ausgerufen. Dieses Datum liegt in der 19. Woche.

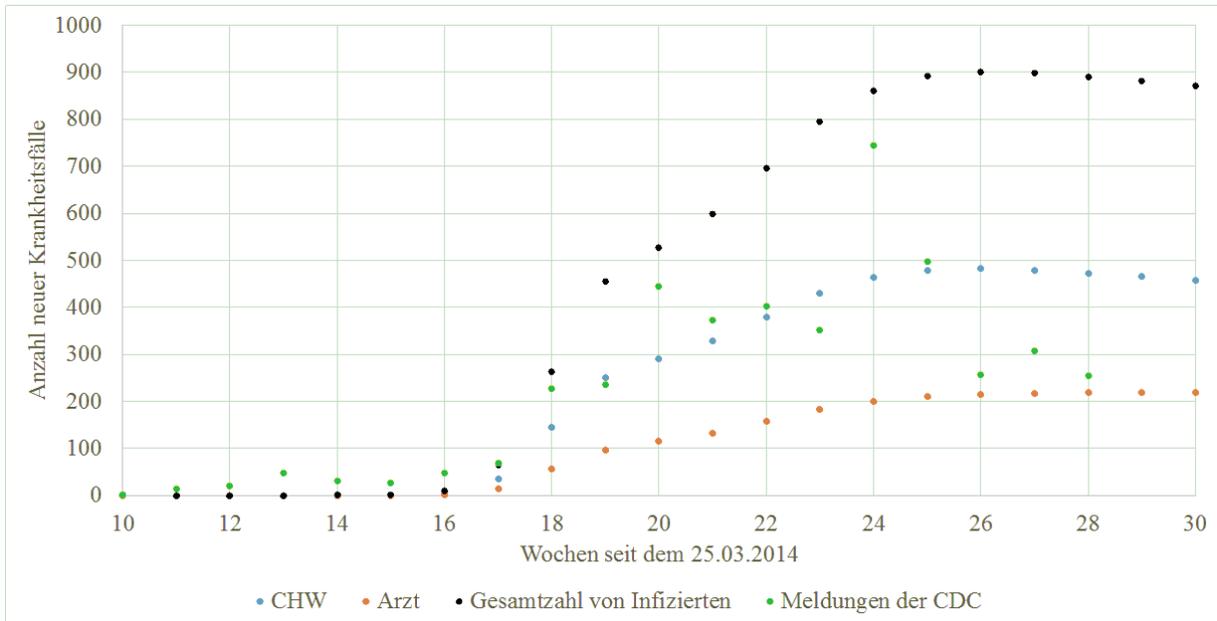


Abbildung 4.4: Simulationsergebnis des Vergleichs von Medizinern und CHWs.

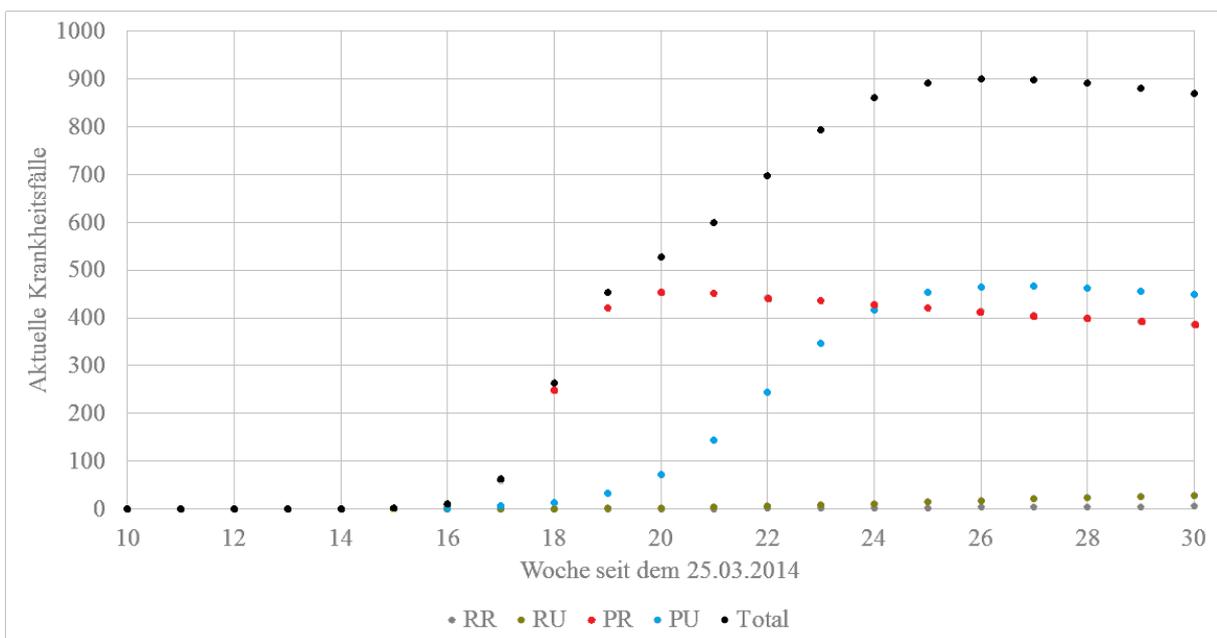


Abbildung 4.5: Krankheitsverlauf bis 30. Woche gemäß kalibriertem SEIR Modell.

Die plausibelste Erklärung dafür ist, dass das zugrundeliegende SEIR System nicht ausreichend gut genug kalibriert wurde. Dieser Verdacht wird dadurch erhärtet, dass der überwiegende Großteil der Infizierten aus der armen Bevölkerung besteht (siehe Abbildung 4.5). Die beiden anderen Bevölkerungsgruppen haben nur wenig Einfluss auf die Gesamtzahl der Fälle. Auch wenn deren Verlauf plausibel erschien, ist diese stark inhomogene Zusammensetzung der Krankheitsfälle unrealistisch. Außerdem ist auffällig, dass die CHWs nur ca. 50% der Krankheitsfälle melden würden, die Ärzte sogar nur ungefähr 25%. Dass deren simulierte Meldungen nicht annähernd mit ihren Meldungen aus der Realität übereinstimmen, ist der letzte und deutlichste Beweis, dass diese Simulation die Wahrheit nur ungenügend widerspiegelt.

Streicht man die wahren Meldungen der CDC aus dem Diagramm, kann abschließend dennoch die hauptsächliche Zielsetzung aus Kapitel 2 erfolgreich beantwortet werden. Denn laut dieser Simulation würden CHWs im Vergleich mit professionellen Mediziner deutlich früher und eindringlicher vor der Epidemie warnen können, da sie stets deutlich mehr Fälle pro Woche melden würden. Dieses Ergebnis erscheint plausibel, da CHWs deutlich mehr Kranke diagnostizieren könnten, als die professionellen Mediziner. Die Ungenauigkeiten bei der Diagnose würden zwar die wahren Fallzahlen etwas verschleiern, durch den Einsatz von CHWs im Zusammenspiel mit einem nahezu Echtzeit-Überwachungssystem würde das Ausmaß der Epidemie jedoch deutlich besser und schneller wiedergegeben werden können.

5 Diskussion und Ausblick

Abschließend sollen hier noch einmal die Ergebnisse dieser Arbeit zusammengefasst werden.

5.1 Zusammenfassung und kritische Prüfung der Ergebnisse

Die WHO kommt in ihrer Arbeit selbst zu der Erkenntnis, dass es eine viel zu große Ungewissheit bezüglich der Dunkelziffer der Fallzahlen gibt. Die dahingehenden Bemühungen, für ein besseres Lagebild zu sorgen, begannen sich erst auszuwirken, nachdem die internationale Hilfe angelaufen war. Das zeigt sich u.a. dadurch, dass die WHO in Liberia erst ab September 2014 von einer vollständig korrekten Datenerfassung ausgeht [25]. Dies wird damit begründet, dass die Patientenversorgung für die Helfer Vorrang vor der Falldokumentation hatte. Diese Probleme könnten durch die Implementierung von CHWs, mit entsprechender Soft- und Hardware ausgestattet, gelöst werden. Durch sie ständen deutlich mehr Helfer für die Patientenversorgung zur Verfügung und die Falldokumentation würde durch die Digitalisierung beschleunigt und erleichtert. Ein diesbezügliches Problem war, dass in 54% der bekannten Krankheitsfälle nicht bekannt ist, ob der Patient überlebt hat [25]. Diese Zahl wäre beim Einsatz von CHWs vermutlich deutlich geringer.

Die Erstellung eines multisozialen, multiregionalen SEIR Modell gestaltete sich schwierig. Ausgehend von den Daten aus dem WHO Bericht [25] wurden die fehlenden Modellparameter mit Formeln aus der einschlägigen Literatur oder durch nachvollziehbare Schätzungen bestmöglich ermittelt. Das Resultat war in vielerlei Hinsicht enttäuschend. Durch eine Kalibrierung mit den gemeldeten Fallzahlen durch Variation der Modellparameter der CDC konnte ein plausibler Krankheitsverlauf erzeugt werden. Bei näherer Betrachtung zeigte sich jedoch, dass die Aufteilung der Krankheitsfälle auf die verschiedenen Bevölkerungsgruppen sehr einseitig ausfiel und damit die Parameterwahl unplausibel ist.

Die Kernfrage aus Kapitel 2 konnte dennoch abschließend geklärt werden. Der Einsatz von CHWs zusammen mit der Bereitstellung eines nahezu Echtzeit-Überwachungssystems

könnte in Zukunft schneller und genauer vor einer möglichen Epidemie warnen.

5.2 Grenzen der Arbeit

Die Literaturrecherche zu dieser Arbeit wurde dadurch beeinträchtigt, dass einige interessante Arbeiten zum Thema mathematische Modellierung von Infektionskrankheiten, SIR Modelle, Berechnung der Basisreproduktionszahl uvm. ausschließlich in medizinischen Datenbanken wie z.B. PubMed veröffentlicht sind. Dazu war im Verlauf der Recherche leider kein Zugang möglich, deshalb konnte nur auf Arbeiten, welche in Google Scholar verfügbar sind, zugegriffen werden.

Die Schwächen des entwickelten multisoziale, multiregionale SEIR wurden bereits in Unterabschnitt 4.2.2 erläutert. Die Optimierung der Modellparameter hätte statt mit der „Trial-and-Error-Methode“ wahrscheinlich mit einer Monte-Carlo-Methode besser funktioniert. Als dieser Umstand festgestellt wurde, war die verbliebene Bearbeitungszeit jedoch leider zu kurz, um den entsprechende Algorithmus zu programmieren, in das Modell zu implementieren und die Ergebnisse erneut auszuwerten.

Das SEIR Modell wurde vor allem für den Beginn der Epidemie aufgestellt. Viele wichtige Faktoren (siehe Abschnitt 3.1) wie beispielsweise psychologische und soziokulturelle Effekte, konnten dafür nicht quantifiziert werden. Auch die beginnende Intervention der Weltgemeinschaft konnte in dem einmal initialisierten Modell nicht berücksichtigt werden, deshalb musste von vornherein ein Kompromiss zwischen Anfangs- und Endzustand des Simulationszeitraums getroffen werden.

Schlussendlich sind die verfügbaren Daten an sich eine weitere Grenze der Simulation. Alleine die Unterteilung der Krankheitsfälle in vermutete, wahrscheinliche und bestätigte Fälle zeigt, dass hier immer mit Unsicherheiten gearbeitet werden und die Daten genau geprüft werden müssen.

5.3 Ausblick

Obwohl die Simulation nicht ganz den wahren Verlauf der Krankheit widerspiegeln konnte, konnten wichtige Erkenntnisse erzielt werden. Außerdem offenbarten sich neue, spannende Forschungsfelder. Zunächst einmal kann das multisoziale, multiregionale SEIR Mo-

dell weiter verbessert werden, denn die Unterteilung der Bevölkerung ist nach wie vor sinnvoll. Hierfür bietet sich u.a. die Monte-Carlo-Methode an.

Weiterhin könnte untersucht werden, wie nach der Erkennung eines Ausbruchs ein schnellere und zuverlässigere Parameterbestimmung möglich ist. Am Rande dieser Arbeit wurden dazu erste Optimierungen mit der Software GAMS durchgeführt. Dadurch war es möglich, für einfach Simulationen ein worst-best-case Szenario zu erstellen. Hierfür wurde ein einfach SIR Modell eingepflegt und mit realen Daten kalibriert. Anschließend wurden die Modellparameter durch Minimierung der Quadratsummenfehler optimiert. Diese Verfahren erscheint deutlich besser geeignet, als die Regression des WHO Modells, da dieses das Überschreiten der maximalen Ausbreitungsgeschwindigkeit der Epidemie nicht berücksichtigen kann.

Literatur

- [1] African Development Bank. *Millennium development goals (MDGs) report 2014*. 2014.
- [2] Andrew S. Boozary, Paul E. Farmer und Ashish K. Jha. “The Ebola Outbreak, Fragile Health Systems, and Quality as a Cure”. In: *JAMA* 312.18 (2014), S. 1859.
- [3] Centers for Disease Control und Prevention. *Previous Case Counts*. URL: <http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/previous-case-counts.html> (besucht am 12.07.2015).
- [4] Lars Engberg-Pedersen und Finn Stepputat. *Fragile States: Definitions, Measurements and Processes*. 2008.
- [5] W. R. Gilks. “Markov Chain Monte Carlo”. In: *Encyclopedia of Biostatistics*. Hrsg. von Peter Armitage und Theodore Colton. Chichester und UK: John Wiley & Sons, Ltd, 2005.
- [6] C. Giorgetti. “International health emergencies in failed and failing states”. In: *Geophysical Journal International* 44 (2012).
- [7] J. J. Haken N.; Messner und K. Hendry et al. “Fund for peace; fragile states index 2014”. In: *Annual ranking report* July-August (2014).
- [8] Liberia Ministry of Health und Social Welfare. *Liberia Health Facility Accreditation Report 2009*. Monrovia, Liberia. 2009.
- [9] Kirsten Johnson. “Association of Combatant Status and Sexual Violence With Health and Mental Health Outcomes in Postconflict Liberia”. In: *JAMA* 300.6 (2008), S. 676.
- [10] W. Kermack und A. McKendrick. “Contributions to the mathematical theory of epidemics—I”. In: *Bulletin of Mathematical Biology* 53.1-2 (1991), S. 33–55.

-
- [11] Ulrich Krengel. *Einführung in die Wahrscheinlichkeitstheorie und Statistik: [für Studium, Berufspraxis und Lehramt]*. 6., verb. Aufl. Bd. 59 : Aufbaukurs Mathematik. Vieweg-Studium. Braunschweig und Wiesbaden: Vieweg, 2002.
- [12] National Transitional Government of Liberia. *Joint Needs Assessment*. 2004. URL: <http://apps.who.int/disasters/repo/12605.pdf> (besucht am 03.08.2015).
- [13] Martin I. Meltzer, Charisma Y. Atkins, Scott Santibanez et al. “Estimating the Future Number of Cases in the Ebola Epidemic — Liberia and Sierra Leone, 2014–2015”. In: *Morbidity and Mortality Weekly Report* 63 (2014). URL: <http://www.cdc.gov/mmwr/pdf/other/su6303.pdf>.
- [14] Robin S. Mathea. *Eine Übersicht über dynamische Modelle zur mathematischen Epidemiologie*. Techn. Ber. Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg, Fakultät für Maschinenbau, 2015.
- [15] Joseph A. Lewnard; Martial L. Ndeffo Mbah und Jorge A. Alfaro-Murillo et al. “Dynamics and control of Ebola virus transmission in Montserrado, Liberia: a mathematical modelling analysis”. In: *The Lancet Infectious Diseases* 14.12 (2014), S. 1189–1195.
- [16] L. Mubila M.; Lannes und M. S. Ben Aissa. “Briefing notes for AfDB’s long-term strategy. Income inequality in africa”. In: *African Development Bank Group Briefing Note No. 5* (2012).
- [17] E. A. Newton und P. Reiter. “A model of the transmission of dengue fever with an evaluation of the impact of ultra-low volume (ULV) insecticide applications on dengue epidemics”. In: *The American journal of tropical medicine and hygiene* 47.6 (1992), S. 709–720.
- [18] World Health Organization. *Technical Guidelines for Integrated Disease Surveillance and Response in the African Region*. 2010. URL: <http://t1p.de/7id6> (besucht am 24.08.2015).
- [19] R. Penchansky und J. W. Thomas. “The concept of access: Definition and relationship to consumer satisfaction”. In: *Medical Care* 19.2 (1981), S. 127–140.

- [20] Jan W. Prüss, Rico Zacher und Roland Schnaubelt. *Mathematische Modelle in der Biologie: Deterministische homogene Systeme*. Mathematik Kompakt. Basel: Birkhäuser Verlag, 2008.
- [21] Erik Pruyt, Willem L. Auping und Jan H. Kwakkel. “Ebola in West Africa: Model-Based Exploration of Social Psychological Effects and Interventions”. In: *Systems Research and Behavioral Science* 32.1 (2015), S. 2–14.
- [22] Sheldon M. Ross. *Stochastic processes*. 2nd ed. Wiley series in probability and statistics. Probability and statistics. New York: Wiley, 1996.
- [23] Liberia Istitute of Statistics und Geo-Information Services. *2008 National population and housing census final results*. URL: http://www.emansion.gov.lr/doc/Population_by_County.pdf (besucht am 13.09.2015).
- [24] Martijn Stijntjes. “Community Health Workers as data enterers in an early warning disease outbreak system”. Magisterarb. Utrecht University, 2015.
- [25] WHO Ebola Response Team. “Ebola Virus Disease in West Africa — The First 9 Months of the Epidemic and Forward Projections”. In: *New England Journal of Medicine* 371.16 (2014), S. 1481–1495.
- [26] UNICEF. *UNICEF, at a glance: Liberia. statistics*. 2013. URL: http://www.unicef.org/infobycountry/liberia_statistics.html (besucht am 27.08.2015).
- [27] Margaret E. Kruk; Peter C. Rockers; S. Tornorlah Varpilah und Rose Macauley. “Which Doctor? Determinants of Utilization of Formal and Informal Health Care in Postconflict Liberia”. In: *Medical Care* 49.6 (2011).
- [28] WHO. *Ebola transmission in Liberia over. Nation enters 90-day intensive surveillance period*. URL: <http://www.who.int/mediacentre/news/statements/2015/ebola-transmission-over-liberia/en/> (besucht am 03.09.2015).

