



universität
wien

MAGISTERARBEIT/MASTER'S THESIS

Titel der Magisterarbeit / Title of the Master's Thesis

Agentenbasierte Modellierung des Sars-CoV-2 Virus auf
Bezirksebene in Österreich am Beispiel des Bezirks Melk

verfasst von / submitted by
Daniel Karas, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the
degree of
Magister der Sozial- und Wirtschaftswissenschaften (Mag. rer. soc. oec.)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt / degree
programme code as it appears on the student
record sheet:

UA 066 951

Studienrichtung lt. Studienblatt / degree
programme as it appears on the student record
sheet:

Magisterstudium Statistik

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Immanuel Bomze

Abstract

Das Jahr 2020 stellte das Leben der gesamten Menschheit auf den Kopf. Innerhalb kürzester Zeit mussten wegen des *SARS-CoV-2* Virus einschneidende Maßnahmen, welche alle Lebensbereiche betrafen, eingesetzt werden, um die Verbreitung des Erregers zu verhindern. Aufgrund von Expertenwissen konnte die Politik auf verschiedenste valide Maßnahmen zur Bekämpfung des Virus zurückgreifen.

Diese Arbeit beschäftigt sich mit dem mathematischen und statistischen Hintergrund, mithilfe dessen die gesetzten Maßnahmen überprüft und Vorhersagen getätigt werden. Dabei werden die sogenannten *Agent Based Models* im Detail betrachtet. Diese werden nach einem generellen Überblick über die Epidemiologie des Virus im Allgemeinen behandelt, um eine gemeinsame Wissensgrundlage zu schaffen. In weiterer Folge werden konkrete Beispiele für solche Modelle anhand der Modellierung des *SARS-CoV-2* Virus präsentiert.

Der Kern dieser Arbeit bildet ein selbstentwickeltes *Agent Based Model*, welches zur Prognose des *SARS-CoV-2* Virus dient. Bei der Entwicklung des Modells wurde großer Wert auf die Einfachheit und Genauigkeit gelegt. Zusammenfassend wird durch diese Arbeit ein Überblick über die Methode der *Agent Based Models* gegeben und eine aktuelle Anwendung mittels eines selbstentwickelten Modells präsentiert.

The year 2020 turned the lives of all of humanity upside down. Due to the SARS-CoV-2 virus, drastic measures affecting all areas of life had to be implemented within a very short time in order to prevent the spread of the pathogen. Based on expert knowledge, politicians were able to draw on a wide variety of valid measures to combat the virus.

This thesis deals with the mathematical and statistical background with which the measures taken are checked and predictions are made. The so-called agent-based models are considered in detail. These will be addressed in general after an overview of the epidemiology of the virus to provide a common knowledge base. In the following, concrete examples of such models are presented using the modeling of the SARS-CoV-2 virus.

The core of this thesis is a self-developed agent-based model, which is used to predict the SARS-CoV-2 virus. When developing the model, great importance was attached to simplicity and accuracy. In summary, this thesis gives an overview of the method of agent-based models and presents a current application using a self-developed model.

Danksagung

An dieser Stelle möchte ich mich nun bei allen herzlich bedanken, die mich während des Verfassens dieser Magisterarbeit motiviert und unterstützt haben.

Beginnend gebührt mein Dank Herrn Univ.-Prof. Mag. Dr. Immanuel Bomze, der meine Magisterarbeit betreut und beurteilt hat. Ich möchte mich hiermit für die hilfreichen Anregungen und die konstruktive Kritik während der Erstellung des Modells und dieser Arbeit bedanken.

Zudem möchte ich mich bei allen bedanken, welche meine Magisterarbeit oder Teile davon Korrektur gelesen haben.

Abschließend möchte ich mich bei all meinen Familienmitgliedern bedanken, die mich durch mein Studium begleitet haben und immer ein offenes Ohr für meine Probleme hatten. Danke für eure Geduld, Unterstützung, Hilfsbereitschaft, Rückhalt und alles andere, was gar nicht in Worte zu fassen ist.

Inhaltsverzeichnis

1	Einleitung	1
2	Allgemeine Informationen zum SARS-CoV-2 Erreger	3
2.1	Chronologie der Pandemie	3
2.2	Epidemiologie des Virus	4
3	Agent Based Models	7
3.1	Agent Based Models im Detail	8
4	<i>Agent Based Models</i> zur Prognose des SARS-CoV-2 Virus	19
4.1	ABM-Simulation einer Stadt	19
4.2	Agent Based Model „COVID-19 ABM“	25
4.3	Agent Based Model für ganz Österreich	30
4.4	Vergleich der Modelle	35
5	Selbstentwickeltes agentenbasiertes Modell	37
5.1	Datengrundlage	38
5.2	Aufbau des Modells	41
5.3	Resultate	49
6	Resümee	63
	Literaturverzeichnis	65
	Abbildungsverzeichnis	67
A	Tabellen	68

1 Einleitung

Im Frühjahr 2020 entwickelte sich das *SARS-CoV-2* Virus, der zu Beginn nur in China sein Unwesen trieb, zu einer weltweiten Pandemie, durch die auf der ganzen Welt erhebliche Einschnitte im Alltag notwendig wurden. Dabei ist kein Land verschont geblieben, so wurde auch Österreich stark von der Pandemie getroffen. Durch diese Krise war die Gesundheitsversorgung am Rande eines Zusammenbruchs und konnte nur durch weitreichende nicht-medizinische Maßnahmen gerettet werden, welche es in dem Ausmaß bisher nicht gegeben hat. Es gab keine genauen Informationen über dieses Virus, weshalb die Forschung und Wissenschaft eine besondere Rolle während der Pandemie einnahm. Ohne dem Expertenwissen konnten keine Maßnahmen entwickelt, validiert und gesetzt werden. Die Wissenschaftler rückten folglich immer mehr in den Vordergrund und die politischen Würdenträger in den Hintergrund.

Neben den Epidemiologen und Biologen, welche für die Erforschung von Infektionskrankheiten bekannt sind, gerieten besonders die Mathematiker und Simulationsforscher aus einem ganz bestimmten Grund in ein neues Rampenlicht: der größte Wunsch jeder Person ist es, den Verlauf und die möglichen Auswirkungen der Pandemie zu prognostizieren und besser zu verstehen. Vor allem Modelle, welche geeignete Maßnahmen zur Eindämmung des Virus modellieren können, gewannen an Bedeutung. Diese gehörten zur Informationsgrundlage, auf welche sich die Entscheidung der Politiker über weitere Maßnahmen stützte. Darüberhinaus wurde mithilfe von mathematischen und statistischen Methoden versucht einzelne gesetzte Maßnahmen schon im Vorfeld zu evaluieren, um Aktionen zu finden, welche keinen nennenswerten Beitrag an der Eindämmung des Virus haben. Es gibt zwei verschiedene Ansätze, die eine geeignete Modellierung von Infektionskrankheiten ermöglichen.

Auf der einen Seite steht der klassische Ansatz der Differentialgleichungen. Diese mathematische Methode zur Modellierung existiert schon seit langer Zeit. Diese ermöglicht es, den Verlauf von Infektionskrankheiten mithilfe von mathematischen Gleichungen zu modellieren und zu prognostizieren. Das einfachste Modell dieser Art ist das *SIR-Modell*. Bei diesem werden mithilfe von Parametern wie beispielsweise der Infektionsrate die Verläufe folgender Gruppen modelliert: „susceptible“ (S), „infected“ (I) und „removed“ (R). Basierend auf diesem einfachen Ausgangsmodell werden komplizierte Simulationsverfahren entwickelt, welche mehr als drei Gruppen abbilden können. Obwohl diese Modellart seit Jahrzehnten erforscht und perfektioniert wurde, entstand in den letzten Jahren ein völlig neuer Ansatz aus den verschiedensten Gründen. Der Wesentlichste ist, dass die Methode der Differentialgleichungen sehr schwer zu verstehen und die Auswirkungen nicht-medizinischer Maßnahmen auf die Gleichungszusammensetzung schwer nachvollziehbar ist. Aus diesem Grund sollten verständlichere Modelle entwickelt werden und so rückten modernere Methoden in den Vordergrund, weil diese einerseits von den Politikern als

Grundlage herangezogen wurden. Andererseits musste die breite Öffentlichkeit ein gewisses Vertrauen zur Wissenschaft und den ausübenden Wissenschaftlern aufbauen, welches Hand in Hand mit dem besseren Verständnis geht.

Der moderne Ansatz sind sogenannte *Agent Based Models*, welche aus dem Bereich Data Science stammen und nicht derart mathematisch hochkomplex sind wie die zuvor erwähnte Methode der Differentialgleichungen. Die Entwicklung der *Agent Based Models* wurde erst in den letzten Jahren schneller vorangetrieben, weil in der nahen Vergangenheit die Rechenleistung der Rechner und die Datengrundlage immer größer wurden. Doch was sind *Agent Based Models*? Agentenbasierte Modelle sind einfache Algorithmen, die Abläufe, Fahrzeuge oder Menschen als *Agents* oder Einheiten betrachten, welche im Laufe einer Zeiteinheit bestimmte Verhaltensregeln und Abläufe beachten müssen. Dadurch wird es möglich, reale Abläufe, wie beispielsweise die Infektion einer Person über den Kontakt zu einer anderen, eins zu eins nachzumodellieren, ohne komplizierte mathematische Gleichungen verwenden zu müssen. Darüberhinaus konnten Maßnahmen aus der realen Welt einfach implementiert und deren Auswirkungen auf den Verlauf betrachtet und erforscht werden. Im Laufe der Pandemie entstanden deshalb sehr viele *Agent Based Models*, welche sich in Größe und Datengrundlage unterscheiden. Es gibt Modelle, welche nur für Städte entwickelt wurden, und andere, die für ganze Staaten erzeugt wurden. Obwohl die Größe so unterschiedlich und die Datengrundlage sehr verschieden sein kann, haben all diese Modelle etwas gemeinsam: die Daten müssen sehr detailreich und genau sein.

Diese Tatsache stellt die Grundlage dieser Arbeit dar, welche *Agent Based Models* behandelt sowie ein eigenständig entwickeltes Modell dieser Art für den Verlauf des *SARS-CoV-2* Virus auf Bezirksebene in Österreich präsentiert. Anfangs wird die Methode der *Agent Based Models* näher betrachtet und mit einigen Beispielen der aktuellen Modellierung der Coronaviruspandemie untermauert. Im Hauptteil folgt die Vorstellung des eigenen Modells sowie die Diskussion der Ergebnisse. Anhand dieser Überlegungen ergeben sich folgende Forschungsfragen: Ist es möglich, ein valides Modell für das Coronavirus auf einem handelsüblichen Laptop mit einer angemessenen Laufzeit zu entwickeln? Kann mithilfe von undetaillierten und geringen Daten ein vernünftiges Ergebnis erzielt werden, welches den Verlauf der Kennzahlen prognostizieren kann? Wie groß ist die Auswirkung der Maßnahmen zur Eindämmung des Virus? All diese Fragen werden am Beispiel des Bezirks Melk beantwortet.

2 Allgemeine Informationen zum SARS-CoV-2 Erreger

Dieses Kapitel soll eine Einführung zum *SARS-CoV-2* Virus und einen Überblick über die wichtigsten Zeitpunkte in der Chronologie der Pandemie geben. Dabei wird immer die Grundvariante beziehungsweise die erste aufgetretene Variante herangezogen und mit den Kennzahlen beschrieben.

2.1 Chronologie der Pandemie

Es war der Dezember 2019 als auf einem Großmarkt in der chinesischen Stadt Wuhan die ersten Fälle des *SARS-CoV-2* Virus auftraten. Anfänglich war die Ursache für die gehäufte Menge an Lungenentzündungen in dieser Region unbekannt. Kurz darauf folgten die ersten Todesfälle und das Virus verbreitete sich in der gesamten Welt. Anfang Februar wurde dasselbe Virus in verschiedensten Ländern der Welt nachgewiesen. Ungefähr einen Monat später wurde der Ausbruch des Coronavirus von der Weltgesundheitsbehörde (WHO) zu einer globalen Pandemie eingestuft (vgl. Rabadan 2020: 1 f.). Die Anzahl der infizierten Personen betrug am 31. März 2020 bereits 800 000 (vgl. Park, Cook, Lim, Sun & Dickens 2020: 1).

In Österreich begann die Pandemie eine Zeit später. Am 25. Februar 2020 wurden die ersten Fälle in Österreich entdeckt: ein junges Paar in Tirol, das zuvor in Italien war, hatte sich mit dem Coronavirus angesteckt. Zu dieser Zeit gab es nur die Regelung, dass erkrankte Personen für 14 Tage in Quarantäne müssen. Bis Ende Februar erhöhte sich die Anzahl auf acht Infizierte. Zu Beginn des Monats März rückten die Apres-Ski-Partys in Österreich, vor allem in Ischgl, in den Vordergrund und wurden zu Events, welche das Virus über die ganze Welt verteilten. Ab dem 10. März 2020 wurden in Österreich erste Maßnahmen zur Eindämmung des Virus eingeführt, welche vor allem die Veranstaltungsbranche betrafen. Zwei Tage später verstarb der erste Österreicher an einer Coronainfektion und vier Tage darauf trat der erste strenge Lockdown in Österreich in Kraft. Dabei wurde so gut wie alles geschlossen und die sozialen Kontakte wurden auf ein Minimum reduziert. Dieser blieb bis Mai 2020 in Kraft und wurde dann lediglich stufenweise gelockert (vgl. Dittler & Kreidl 2021: 1–11).

In den folgenden Monaten wiederholte sich die Entwicklungen vom Anfang immer wieder: nach den Lockerungen im Sommer begannen die Fallzahlen mit der kalten Jahreszeit stark zu steigen und es wurde abermals ein strenger Lockdown verhängt. Dies geschah sowohl im Herbst und Winter 2020 als auch zur selben Zeit im Jahr 2021. Ein weiterer Meilenstein war die Entwicklung einer Impfung gegen das Coronavirus, welche Ende Dezember 2020 in Österreich erstmals verabreicht wurde. Seit dem Auftreten der ersten Fälle mutierte und entwickelte sich das Virus rasant weiter, sodass es auch immer wieder zu Re-

infektionen kommen konnte. Im Laufe der Zeit wurden tausende von Varianten entdeckt, von denen sich nur wenige durchsetzen konnten, zum Beispiel die Delta- oder Omikron-variante. Im April 2022 wurden die meisten Maßnahmen gänzlich aufgehoben und das Leben gewinnt wieder an Normalität. Bis dahin gab es weltweit 509 805 166 Infektionen mit dem Coronavirus und über 6 000 000 Todesfälle. In Österreich gab es bis zum 25. April 2022 4 108 970 Infektionen und 18 054 Todesfälle (vgl. Johns Hopkins University 2022).

2.2 Epidemiologie des Virus

Um das Virus und damit auch den Verlauf und das Verhalten besser zu verstehen, werden neben der absoluten Zahl der Infizierten einige andere Kennzahlen benötigt. Im folgenden Teil der Arbeit soll eine kleine Einführung in die wichtigsten Parameter für das Coronavirus gegeben werden, sodass in weiterer Folge die mathematische statistische Modellierung besser zu verstehen ist.

Inkubationszeit

Die Inkubationszeit bezeichnet den Zeitraum zwischen einer Ansteckung mit dem Virus und dem Ausbruch, der das Auftreten der ersten Symptome markiert. Wichtig ist zu berücksichtigen, dass innerhalb dieser Zeit keine Symptome auftreten und es sehr schwer zu erkennen ist ob jemand bereits ansteckend ist. Ein weiterer wichtiger Punkt ist, dass dieser Zeitraum von Krankheit zu Krankheit und von Mensch zu Mensch sehr variieren kann. Beispielsweise beträgt die Inkubationszeit bei einer gewöhnlichen Erkältung nur ungefähr vier Tage und bei der Tollwut des Menschen ein bis zwei Monate. Im konkreten Fall des Coronavirus wurde die Inkubationszeit auf etwa fünf Tage geschätzt, wobei es in seltenen Fällen auch mehr als zwei Wochen sein können (vgl. Rabadan 2020: 14). Dabei ist zu berücksichtigen, dass dieser Wert je nach Studie zwischen vier und sechs Tagen liegt (vgl. Park et al. 2020: 6). Nach Hemmer, Geerdes-Fenge & Reisinger (2020: 894) geht aus einer Studie hervor, dass die Infektiosität einer infizierten Person ungefähr 2,3 Tage vor Symptombeginn einsetzt und etwa 0,7 Tage nach Start der Symptome am höchsten ist, wobei die durchschnittliche Infektiositätsdauer zwischen acht und neun Tagen beträgt.

Serielles Intervall

Das serielle Intervall ist eine sehr wichtige Kennzahl, weil sie die Zeit vom Eintreten der Symptome in einem ersten Fall bis zum Eintreten der Symptome bei einer Folgeinfektion angibt. Mithilfe dieser Kennzahl kann die Übertragung des Virus besser verstanden werden und die Zeit zwischen einer Infektionen und der Folgeinfektion eines weiteren Individuums besser geschätzt werden (vgl. Nishiura, Linton & Akhmetzhanov 2020: 1). Dieser Wert kann abermals stark von Studie zu Studie schwanken. Schätzungen reichen von fünf bis acht Tagen (vgl. Park et al. 2020: 6).

Basisreproduktionszahl R_0 und 7-Tages-Inzidenz pro 100 000 Einwohner

Die genannten Werte sollen dazu dienen, die Ansteckungsfähigkeit eines Erregers widerzuspiegeln. Die Basisreproduktionszahl R_0 beschreibt die geschätzten Neuinfektionen je infizierter Person. Nun wird zwischen zwei möglichen Szenarien unterschieden. Auf der einen Seite kann R_0 unter Eins liegen, sodass die Zahl der Neuinfektionen und die Zahl der Infektionen in der Population zurückgehen, da eine Person im Schnitt weniger als eine andere Person ansteckt. Auf der anderen Seite kann R_0 über Eins liegen. Wenn dies der Fall ist, steigen die Fallzahlen exponentiell an und sind auf längere Sicht nicht mehr kontrollierbar, weil eine Person mehr als eine andere Person infiziert. Diese Kennzahl ist ein Indikator für die Notwendigkeit von Maßnahmen zur Eindämmung von Krankheitserregern. Das Ziel ist es, diese Zahl konstant unter Eins zu halten, damit sich weniger Personen infizieren als es Infizierte gibt und es in weiterer Folge keine Neuinfektionen mehr gibt. Wie zuvor kann diese Kennzahl sehr unterschiedlich bei den verschiedensten Infektionskrankheiten sein. Beim *SARS-CoV-2* Virus beträgt die Basisreproduktionszahl knapp Drei, falls keine Maßnahmen getroffen werden. Im Vergleich dazu liegt sie bei der Grippe knapp über Eins (vgl. Rabadan 2020: 16–19). Park et al. (2020: 4) betrachten verschiedenste Studien, in denen die Basisreproduktionszahl des *SARS-CoV-2* Erregers geschätzt wurde. Dabei schwankte der Wert zwischen 1,9 und 6,5. Hier ist zu berücksichtigen, dass die Mehrheit der Schätzungen zwischen Zwei und Drei liegt (vgl. ebd.: 4 f.). Jedoch ist die Basisreproduktionszahl ungeeignet für Vergleiche der einzelnen Länder beziehungsweise einzelner regionaler Unterschiede, wie zum Beispiel von Gemeinde zu Gemeinde. Um dies besser in den Griff zu bekommen, entwickelte man die *7-Tages-Inzidenz* pro 100 000 Einwohner. Sie gibt die durchschnittliche Anzahl an Neuinfektionen in den letzten sieben Tagen gerechnet auf 100 000 Personen an (vgl. Hemmer et al. 2020: 893).

Faktoren für die Schwere der Krankheit

Einer der wichtigsten Faktoren für die Schwere des Krankheitsverlaufs ist das Alter des Infizierten. Betrachtet man zum Beispiel die Grippe, so erleiden Säuglinge und ältere Personen oftmals einen schweren Verlauf. Bei *Covid-19* ist dies sehr ähnlich. Die Fallsterblichkeit, welche der Anteil der Verstorbenen unter den Infizierten ist, nimmt mit zunehmendem Alter dramatisch zu. Liegt diese bei unter 40-Jährigen noch bei 0,2 Prozent, so ist sie bei über 80-Jährigen schon bei mehr als 15 Prozent. Des Weiteren dürfte es auch Unterschiede zwischen Männer und Frauen hinsichtlich der Schwere eines Verlaufs geben. Dies könnte sich einerseits auf die biologischen Eigenschaften und andererseits auf die Lebensweise zurückführen lassen. In vielen Ländern ist beispielsweise der Anteil der rauchenden Männern viel größer als der Anteil der Frauen, womit das Risiko eines schweren Verlaufs bei Männern bei Lungenkrankheiten höher ist (vgl. Rabadan 2020:

22–24). Ebenso schreibt Park et al. (2020: 7), dass bei älteren Individuen und jenen, die ein geschwächtes Immunsystem haben, ein größeres Risiko besteht einen schweren Verlauf durch *Covid-19* zu erleiden.

Herdenimmunität

Einer der wichtigsten Begriffe ist die Herdenimmunität. Aber warum genau ist dieser so bedeutend? Zu Beginn einer neuen Infektionskrankheit ist die gesamte Bevölkerung noch nicht mit dem Virus in Kontakt gekommen und anfällig für diesen. Dies resultiert anfangs in einen großen Anstieg der Fallzahlen. Sobald jedoch einige Personen der Population gegen das Virus immun werden, wird es für den Erreger schwieriger sich zu entfalten und die Ausbreitung wird verlangsamt. Wenn die Zahl der immunisierten Personen einen gewissen Schwellenwert, die Herdenimmunitätsschwelle, erreicht hat, kann sich das Virus nicht mehr in der Bevölkerung ausbreiten. Die Herdenimmunität hängt erheblich von der Basisreproduktionszahl ab, denn sie kann näherungsweise mit der Formel $1 - \frac{1}{R_0}$ berechnet werden (vgl. Rabadan 2020: 27–29). Am Beginn der Pandemie wurde eine Zahl zwischen 60 und 70 Prozent als Herdenimmunität genannt. Dies bedeutet, dass die Infektionskette erst zu einem Stillstand kommt, sobald 60 bis 70 Prozent der Bevölkerung immun sind (vgl. Hemmer et al. 2020: 893). Aufgrund der vielen Mutationen des *SARS-CoV-2* Virus ist diese Zahl nicht realistisch, da keine lange Immunisierung vorherrscht.

3 Agent Based Models

Das Kernstück eines *Agents Based Models*, welches in der restlichen Arbeit mit ABM abgekürzt wird, sind die sogenannten *Agents*. Diese sind Objekte, welche bestimmte Eigenschaften besitzen und gewisse Aktionen durchführen können (vgl. Wilensky & Rand 2015: 1). Bonabeau (2002: 7280) beschreibt sie als autonome Einheiten, die individuell Entscheidungen auf Basis einfacher Regeln treffen können. *Agents* können beispielsweise Menschen darstellen, die als Eigenschaften die demografische Daten eines bestimmten Landes erfüllen. Ein ABM ist ein rechnerisches Modell, welches mittels Interaktion zwischen den einzelnen *Agents* Phänome modelliert. Somit lassen sich komplexe Vorgänge aus den verschiedensten Fachgebieten nachvollziehbar darstellen (vgl. Wilensky & Rand 2015: 1). Darüberhinaus können ABMs neuronale Netze, evolutionäre Algorithmen oder andere Lernalgorithmen enthalten, um besser aus den Daten zu lernen (vgl. Bonabeau 2002: 7280).

Zur besseren Verständlichkeit existiert ein einfaches Beispiel: Um die Veränderung der Populationsgröße eines Landes über Jahre hinweg zu modellieren oder zu prognostizieren, werden *Agents* definiert, die die gesamte Bevölkerung mittels demografischer Daten widerspiegeln. Das heißt, dass die Objekte im Modell Eigenschaften wie Geschlecht, Alter und Herkunftsland besitzen. Diese *Agents* können Aktionen ausführen, beispielsweise zu sterben, ein Kind zu gebären oder auszuwandern. Im Modell interagieren diese Objekte miteinander und Aktionen werden aufgrund gewisser Auslöser durchgeführt. Somit kann die Populationsgröße eines Landes nach einem Jahr geschätzt werden (vgl. Bicher, Urach & Popper 2018: 2657).

Doch warum sind ABMs erst in der nahen Vergangenheit populär geworden? Aus welchem Grund werden sie nun häufiger angewandt als in der Vergangenheit? Diese Entwicklung geht Hand in Hand mit der immer stärkeren Rechenleistung der Computer. Einerseits macht es diese Entwicklung Wissenschaftlern möglich Berechnungen durchzuführen, die in der Vergangenheit nicht möglich waren. Andererseits gibt es nun die Gelegenheit riesige Datensätze zu generieren und zu analysieren. Diese sind die Grundlage für die Eigenschaften und Aktionen in einem *Agent Based Model*. Darüberhinaus ist die Rechenleistung maßgeblich dafür verantwortlich, dass Millionen von Objekten miteinander interagieren und gleichzeitig simuliert werden können. Ein weiterer wichtiger Punkt für die Popularität dieser Modelle ist die bessere Verständlichkeit und Nachvollziehbarkeit. Dies hat den Grund, dass ein ABM aus individuellen Objekten und deren Verhalten bestehen und nicht aus mathematischen Formeln und Gleichungen. Diese Art der Simulation verhält sich wie die intuitive Beschreibung eines Menschen von natürlichen Vorgängen. Das bedeutet, dass Menschen intuitiv die Abläufe in deren Umgebung mittels einzelner Einheiten und deren Verhalten versuchen zu erklären. Individuelle Erfahrungen werden in erster Linie durch Interaktionen zwischen Individuen beschrieben. Nach exakt dieser Art

und Weise gehen ABMs vor, womit die Beschreibung von *Agent Based Models* ähnlich der Vorgangsweise natürlicher Versuche ist. Der Einsatz dieser erleichtert das Verstehen und Beschreiben natürlicher Vorgänge (vgl. Wilensky & Rand 2015: 2).

3.1 Agent Based Models im Detail

Die Grundidee eines ABMs ist, dass sich nahezu alle realen Phänomene mittels *Agents*, einer *Umgebung* und einer Beschreibung der *Interaktionen* zwischen *Agent* und *Agent* oder *Agent* und *Umgebung* beschreiben lassen. Wie schon zuvor erwähnt sind das Herzstück dieser Modelle die *Agents*, die „autonomous individual or object with particular properties, actions, and possibly goals“ (ebd.: 32) sind. Als Umgebung kann sich eine Grundlage vorgestellt werden, auf der die Interaktion zwischen den Objekten sowie den Objekten und der Umgebung stattfinden. Diese kann auf geometrischer Natur (z.B. ein Raster, auf dem sich die Objekte fortbewegen), einem Netzwerk oder Realdaten basieren. Ein weiterer wichtiger Punkt sind die Interaktionen, welche sehr komplex werden können. Darüberhinaus kann sich das Kommunikationsverhalte der *Agents* über die Zeit verwandeln. Aufgrund verschiedenster Entscheidungen der *Agents* kann sich deren Status verändern und somit zusätzliche oder auch eine geringere Anzahl an Interaktionen bewirken (vgl. Wilensky & Rand 2015: 32; Macal & North 2010: 152). Aber es ist zu berücksichtigen, dass ein ABM keine eigenständige Technologie ist, sondern eine Art oder Möglichkeit Systeme auf Grundlage ihrer einzelnen Bausteine zu beschreiben. Deshalb kann *Agent Based Modelling* auch als mikroskopische Modellierung bezeichnet werden und steht im Gegensatz zu den makroskopischen Modellen (vgl. Bonabeau 2002: 7280). Die einzelnen Grundelemente werden im Folgenden näher beschrieben.

Zunächst gibt es einige wesentliche Charakteristika, die ein *Agent* erfüllen muss:

- „An agent is a self-contained, modular, and uniquely identifiable individual“ (Macal & North 2010: 153). Es ist klar definiert, welche Attribute zu einem *Agent* gehören und welche nicht. Darüberhinaus ist jedes Element über seine Attribute eindeutig unterscheidbar zu allen anderen und wird dadurch von jenen auch erkannt (vgl. ebd.: 153).
- Ein *Agent* soll autonom und selbstgesteuert handeln. Dabei wirkt er basierend auf einer begrenzten Anzahl an möglichen Aktionen in der Umgebung und in den Interaktionen mit anderen *Agents* unabhängig (vgl. ebd.: 153).
- Der Status eines *Agents* soll sich über die Zeit verändern können. Als Status werden die wesentlichen Variablen eines Elements bezeichnet, welche die aktuelle Situation beschreiben. Somit folgt daraus, dass sich der Status nicht immer auf alle Attribute eines *Agents* sondern auf eine Teilmenge derer bezieht (vgl. ebd.: 153).

- „An agent is *social* having dynamic interactions with other agents that influence its behaviour“ (vgl. Macal & North 2010: 153).

Außerdem kann ein *Agent* je nach Komplexität noch weitere Charakteristika haben, wie zum Beispiel Lernfähigkeit, Zielorientiertheit und Heterogenität (vgl. ebd.: 153).

Ein weiterer Grundstein sind die Interaktionen der *Agents* miteinander oder mit der Umgebung. Die Schaffung dieser beinhaltet: Einerseits deren Regeln und andererseits den Mechanismus dieser Interaktionen. Ein erster bedeutender Grundsatz ist dabei, dass es lediglich lokale Informationen gibt. Das heißt, dass nicht jeder *Agent* alle Informationen über jeden *Agent* hat, sondern nur jene seiner näheren Umgebung. Somit können nicht alle *Agents* mit allen anderen interagieren. Die erwähnte Umgebung wird Nachbarschaft eines *Agents* genannt. Diese verändert sich häufig im Laufe der Simulation und kann auf verschiedenste Weisen beschrieben werden sowie entstehen. Sie entwickelt sich aus Entfernungen, die die *Agents* zueinander haben, aus sozialen Netzwerken, sogenannte Freundeskreise, die jeder *Agent* haben kann, oder auf ganz andere Arten (vgl. ebd.: 154).

Zu guter Letzt ist die Umgebung zu erwähnen. Diese kann nicht nur einfache Informationen zur räumlichen Position der *Agents* beinhalten, sondern ebenso viele komplizierte Daten aus dem Bereich des geografischen Informationssystems, auch GIS bezeichnet (vgl. ebd.: 155).

Aufgrund der obigen Erklärungen wird deutlich, dass ein *Agent Based Model* viele verschiedene Facetten hat und es nicht eine einheitliche Definition geben kann. Deshalb ist es möglich, vier verschiedene Definitionen zu geben, die im Folgenden aufsteigend nach der Komplexität geordnet sind. Diese Definitionen sind informell und sollen lediglich eine Kommunikationsgrundlage schaffen (vgl. Macal 2016: 149).

Definition 3.1. „An *individual ABMS* [Agent-based modelling and simulation] is one in which the agents in the model are represented individually and have diverse characteristics“ (ebd.: 149).

In *individual ABMs* ist das Verhalten bereits vorgeschrieben und die Aktionen werden nicht aufgrund jeglicher Events, die als Auslöser dienen, hervorgerufen, sondern können anhand von einem vordefinierten Zeitplan ausgeführt werden. In dieser Art von ABMs braucht ein jeder *Agent* einen eigenen Status. Außerdem interagieren die *Agents* nicht miteinander oder mit der Umgebung (vgl. ebd.: 149).

Definition 3.2. „An *autonomous ABMS* is one in which the individual agents have internal behaviors that allow them to be autonomous, able to sense whatever condition occurs within the model at any time, and to act on the appropriate behavior in response“ (ebd.: 149).

Diese beiden Definitionen unterscheiden sich in nur einem Punkt. *Agents* können ihr Verhalten zu jeder Zeit ändern und haben keinen vorgeschriebenen Plan, hinsichtlich der

Abläufe ihrer Aktionen. Sie können auf verschiedenste Ereignisse autonom reagieren (vgl. Macal 2016: 149).

Definition 3.3. „An *interactive ABMS* is one in which autonomous agents interact with other agents and with the environment“ (ebd.: 150).

Hierbei kommen die Interaktionen der einzelnen *Agents* zueinander oder mit der Umgebung ins Spiel, wodurch sich komplexere Strukturen abbilden lassen (vgl. ebd.: 150). Das in dieser Arbeit entstandene Modell zur Vorhersage der Covid-19 Zahlen ist ein Musterbeispiel für diese Definition.

Definition 3.4. „An *adaptive ABMS* is one in which the interacting, autonomous agents change their behaviors during the simulation, as agents learn, encounter novel situations, or as populations adjust their composition to include larger proportions of agents who have successfully adapted“ (ebd.: 150).

Bei der letzten Definition wurde das Modell noch damit erweitert, dass die *Agents* ihr Verhalten aufgrund von Events oder ihrem Zustand anpassen und verändern. Sie können somit aus Vergangenen lernen (vgl. ebd.: 150).

Zusammenfassend ist zu erwähnen, dass diese Definitionen rein beschreibender Natur sind. Sie sollen eine Möglichkeit bieten, die Fülle an verschiedenen *Agent Based Models* in Kategorien zu unterteilen und eine gemeinsame Gesprächsgrundlage zu generieren. Es ist klar, dass damit keine eindeutige Abgrenzung von anderen Modellarten gelingt, dies ist aber auch nicht der Sinn dieser Definitionen.

<i>ABMS definition/ agent properties</i>	<i>Individuality</i>	<i>Behaviours</i>	<i>Interactions</i>	<i>Adaptability</i>	<i>Example</i>
Individual ABMS	Individual heterogeneous agents*	Prescribed, scripted [†]	Limited	None	Traffic model that has agents moving between origin–destination pairs according to a script
Autonomous ABMS	Individual heterogeneous agents*	Autonomous, dynamic [‡]	Limited	None	Taxation model in which agents choose occupations and places to work but do not interact with others
Interactive ABMS	Individual heterogeneous agents*	Autonomous, dynamic [‡]	Between other agents and the environment [§]	None	Infectious disease model in which agents transmit and are infected through contact and respond to their disease state according to prescribed behaviours
Adaptive ABMS	Individual heterogeneous agents*	Autonomous, dynamic [‡]	Between other agents and the environment	Agents change behaviours during the simulation	Healthcare model in which agents change their behaviours according to the state of their health

*Agents in the population have diverse set characteristics.
[†]Agent behaviour is exogenously provided and not based on endogenous events during the simulation.
[‡]Agent behaviour is endogenous based on the current agent state.
[§]Agent behaviours are based on the observed states and behaviours of other agents and the state of the environment.
^{||}Agents change behaviours during the simulation, agents learn, and/or populations adjust their composition.

Abbildung 1: Kurze Zusammenfassung der zuvor erklärten Definitionen (Macal 2016: 150)

Formaler Aufbau eines ABMs

Der formale Aufbau eines ABMs kann in drei grundlegende Punkte aufgeteilt werden: die Definition der *Agents* mit deren Eigenschaften, die Entwicklung von Regeln für das

Verhalten der *Agents* und die Implementierung der Interaktionen untereinander (vgl. De Marchi & Page 2014: 7–8).

1. Bevor mit der Entwicklung von Verhaltensregeln begonnen werden kann, müssen die *Agents* und deren Eigenschaften, auch Attribute genannt, definiert werden. Dabei ist es wichtig eine Menge für die einzelnen Attribute zu schaffen, damit die *Agents* anschließend über einen Vektor der Attribute definiert werden können. Sei M die Anzahl der Attribute. Dann gibt es für jede Eigenschaft eindeutige Ausprägungen $x_{i1}, x_{i2}, x_{i3} \dots$ für $i \in \{1, 2, \dots, M\}$, die eine abzählbare Menge an Möglichkeiten für jedes Attribut bilden. Somit gilt formal:

$$X = X_1 \times X_2 \times \dots \times X_M, \text{ wobei jedes } X_i \text{ abzählbar ist,}$$

ist der Raum der möglichen Eigenschaftsvektoren.

Außerdem ist anzumerken, dass es zwei Arten von Attributen gibt: einerseits fixe Eigenschaften, wie Geschlecht oder Körpergröße, und andererseits veränderbare Attribute, wie zum Beispiel Lokation oder Gesundheitstatus während einer Pandemie. Letztere sind entscheidend für den nächsten Punkt: Eigenschaften können sich über die Zeit hinweg verändern. Dies bedeutet, dass die Attribute der *Agents*, sowie die *Agents* selbst, von der Zeit abhängig sind. Sei N eine fixe Anzahl an *Agents*, so kann deren Status a_j^t zum Zeitpunkt t wie folgt definiert werden:

$$\forall j \in \{1, 2, \dots, N\} : a_j^t = (x_{j1}^t, x_{j2}^t, \dots, x_{jM}^t), \text{ wobei } x_{ji}^t \in X_i$$

Zu guter Letzt wird die Konfiguration oder der Stand des Modells A_t zum Zeitpunkt t als Matrix der Dimension $N \times M$ mit den Vektoren der Attribute der einzelnen *Agents* als Zeilen der Matrix beschrieben (vgl. ebd.: 7).

2. Im nächsten Punkt muss das Verhalten der *Agents* beschrieben und vorgegeben werden. Dies basiert hauptsächlich auf Regeln, wie sich die Einheiten verhalten sollen. Dabei gibt es verschiedenste Ansätze. Beginnend mit dem Einfachsten, können Regeln nur von der aktuellen Konfiguration A_t abhängig sein und das Verhalten kann sich aufgrund dieser anpassen. Geht man einen Schritt weiter, können die Regeln auch von vergangenen Konfigurationen $A_{t-1}, \dots, A_1$ abhängig sein und potentielle Aktionen anderer *Agents* berücksichtigen (vgl. ebd.: 8).
3. Im letzten Punkt werden die möglichen Interaktionen zwischen den *Agents* aufgegriffen. Hierbei ist zu berücksichtigen, dass alle Interaktionen beschrieben und vorgegeben werden müssen. Weiters ist darauf zu achten, dass die Interaktionen sowohl vom Status als auch von der zeitlichen Komponente abhängig sind. Das bedeutet, dass manche Aktionen erst dann gemacht werden, wenn eine andere Interaktion

davor geschehen ist. Darüberhinaus ist auch auf die Reihenfolge der Interaktionen zu achten. Es kann problematisch werden, wenn Aktionen gleichzeitig durchgeführt werden, vor allem wenn die *Agents* die Lokation verändern können. Es gibt einige Möglichkeiten dieses Problem zu umgehen. Einerseits kann eine zufällige Reihenfolge der Interaktionen durchgeführt werden. Andererseits ist es möglich, eine Reihenfolge zu bilden, indem die geografischen Daten hinzugenommen oder die Einheiten bevorzugt werden, die zu Beginn am vorteilhaftesten sind (vgl. De Marchi & Page 2014: 8).

Zusammenfassend ist zu erwähnen, dass dieser Aufbau lediglich eine Hilfestellung sein soll. Natürlich unterscheiden sich die einzelnen Modelle im Aufbau extrem voneinander, da nicht jedes Problem aus jeder Fachrichtung auf die gleiche Weise übersetzt werden kann. Diese Richtlinie soll vor allem ein besseres Verständnis für *Agent Based Models* geben und deren Nähe zur Vorstellung, wie Prozesse oder Systeme in der realen Welt ablaufen, darstellen.

Unterschiede zwischen ABMs und anderen Modellen

Es gibt einige Unterschiede zwischen ABMs und anderen Modellformen. Zunächst wird in diesem Vergleich nur zwischen dem *Equation Based Model*, welches in der Folge als EBM abgekürzt wird, und dem *Agent Based Model* differenziert, da die meisten Phänomene in irgendeiner Art und Weise mithilfe von Gleichungsmodellen beschrieben werden können (vgl. Wilensky & Rand 2015: 32).

- Ein wesentlicher Unterschied zwischen den Modelltypen liegt in der Grundpopulation. In ABMs kann die Population heterogen sein, weil man jedes einzelne Individuum simuliert und nicht die Gesamtheit. Gegensätzlich treffen EBMs typischerweise die Annahme, dass die Population homogen ist. Die Heterogenität der Population spielt jedoch eine wichtige Rolle bei vielen Modellen (vgl. ebd.: 32).
- Des Weiteren sind die Resultate und die Kommunikation zeitdiskret und nicht stetig, wenn man versucht Individuen, zu simulieren. Die meisten Gleichungsmodelle sind jedoch stetig. Die Verwendung eines EBMs bei *Population Dynamics* kann dazu führen, dass sich die Population stetig vergrößert, obwohl die Population nur diskrete Werte annehmen kann. Zum Beispiel hat ein Zehntel eines Wolfs eine Auswirkung auf die Vergrößerung eines Wolfsrudels, obwohl dieses gar nicht wahrhaftig existiert (vgl. ebd.: 32).
- Im Gegensatz zu den EBMs werden für *Agent Based Models* keine genauen Kenntnisse über das Gesamtphänomen benötigt. Bei EBMs ist es wichtig, das Gesamtverhalten zu verstehen, damit man eine Hypothese testen kann. Hingegen werden

bei ABMs Informationen und Verhalten von einzelnen Individuen benötigt. Diese interagieren mittels einfacher Regeln miteinander. Somit wird mithilfe von Simulationen von Individuen auf das Gesamtverhalten zurückgeschlossen. Es wird dadurch möglich, ein Modell ohne einer konkreten Hypothese des Gesamtverhaltens eines Problems zu erstellen und Resultate zu erhalten (vgl. Wilensky & Rand 2015: 33).

- *Agent Based Models* sind einfacher zu erklären, da sich diese Art der Modelle auf die Individuen konzentriert und nicht so sehr auf die Gesamtheit eines Problems. Somit sind diese Modelle sehr ähnlich zu realen Problemen und es wird einfacher, ein Modell über Expertenwissen zu validieren und kalibrieren, da jeder den natürlichen Aufbau des Modells verstehen kann (vgl. Wilensky & Rand 2015: 33; vgl. Bonabeau 2002: 7281; vgl. Macal 2016: 145). Macal (2016: 145) ergänzt, dass jedem die Perspektive der *Agenten* aus dem Alltag bekannt ist. Die Aktionen, die ein Mensch in der realen Welt durchführt, sind nämlich wie Aktionen, die ein *Agent* in seiner Umgebung durchführt. Dieses Verhalten kann sich auch über die Zeit verändern, somit ist jeder Mensch in einem gewissen Sinn ein *Agent Based Modeller*.
- Als letzter großer Punkt sind die detaillierteren Resultate von ABMs zu nennen. Dies lässt sich darin erkennen, dass diese Art der Modellierung sowohl die Gesamtergebnisse eines Problems als auch die einzelnen *Agents* gleichzeitig simuliert. Dieses Prinzip wird *bottom-up* genannt und steht im Gegensatz zum *top-down*-Prinzip, welches beispielsweise bei EBMen zur Anwendung kommt. Bei jenen Modellen sind Resultate über das Gesamtproblem vorhanden, jedoch nicht über die einzelnen Individuen (vgl. Wilensky & Rand 2015: 33). Ähnlich sehen dies De Marchi & Page (2014: 4), welche vor allem die Verhaltensregeln in den Vordergrund stellen, nach denen die *Agents* agieren. ABMs sind von diesen gewählten Regeln abhängig, aus denen sich anschließend durch die Interaktionen der einzelnen Einheiten eine Gesamteigenschaft erkennen lässt. Kleine Änderungen im Verhalten haben somit große Auswirkungen auf das Gesamtergebnis (vgl. ebd.: 4). Aus diesem Grund ist der *Agent* und dessen Verhalten das grundlegende Element für eine Modellierung mittels ABMs. Durch die Abhängigkeit der Einheiten vom Verhalten lässt sich neben Prozessen und Systemen, welche mit anderen Modellarten simuliert werden, noch vieles mehr betrachten (vgl. Macal 2016: 145).

Zu ergänzen ist, dass ABMs neue beziehungsweise einzigartige Einblicke in Systeme oder Prozesse liefern, die über andere Modellierungsarten nicht ersichtlich sind. ABMs sollen einen Einblick in den Ablauf und das Verhalten eines Systems geben, anstatt die genauesten Berechnungen durchzuführen. Damit können diese Modelle eine Basis für eine begründbare Erklärung der untersuchten Phänomene darstellen (vgl. ebd.: 145).

Die Zufälligkeit

Im Gegensatz zu den meisten deterministischen Modellformen ist es bei ABMs möglich der Zufälligkeit eine große Rolle zukommen zu lassen. Die einzelnen Entscheidungen oder Regeln, wie sich *Agents* verhalten sollen, beruhen auf Zufälligkeit (vgl. Wilensky & Rand 2015: 34; vgl. Bonabeau 2002: 7281). Die Zufälligkeit spielt somit eine große Rolle in ABMs, da sie einerseits eine Vereinfachung der realen Umstände ist und eine schnellere Modellbildung ermöglicht. Beispielsweise ist es mithilfe eines ABMs möglich einen Ameisenstamm zu simulieren, welcher sich auf Futtersuche befindet. Hierbei stellt die Fortbewegung einer Ameise ein großes Problem dar, weil es einen Unterschied macht, ob die Ameise über Stock und Stein oder über einen ebenen Weg wandert. Dies kann nicht mithilfe deterministischer Werte einfließen, weshalb hier die Zufälligkeit eine Rolle spielt. Der Weg einer Ameise zum Essen wird somit über den Zufall simuliert. Zusammenfassend kann die Zufälligkeit als Vereinfachung und Approximation der realen Umstände verstanden werden, welche die Erstellung eines Modells vereinfachen. Andererseits gibt es jedoch nicht genügend Informationen über den Ablauf eines komplexen Systems, damit alles deterministisch berechnet werden kann. Somit wird hierbei auf Modelle zurückgegriffen, welche mit einer gewissen Zufälligkeit umgehen können und Resultate liefern (vgl. Wilensky & Rand 2015: 34).

Kriterien für *Agent Based Models*

Auf welche Art ein Modell verwendet werden soll oder kann, hängt stark von der Situation ab. *Agent Based Models* sind nicht der heilige Gral der Modelle und sind beispielsweise nur unter gewissen Voraussetzungen sinnvoll einsetzbar. Im folgenden Teil der Arbeit werden gewisse Richtlinien vorgestellt, nach denen entschieden werden kann, ob ein ABM sinnvoll ist oder nicht (vgl. ebd.: 35).

- Erstens spielt die Anzahl der *Agents* eine große Rolle. Ein Problem mit vielen homogenen Objekten ist meist einfacher zu lösen, wenn auf andere Modelle zurückgegriffen wird. Es macht jedoch keinen Sinn auf ABMs zurückzugreifen, wenn nur eine handvoll *Agents* vorhanden sind. Hier ist es einfacher, die Interaktionen mittels Gleichungen zu beschreiben (vgl. ebd.: 35).
- Zweitens ist es von Vorteil, wenn die Individuen heterogen sind und die Heterogenität die Performance eines Systems beeinflusst. Warum ist das ein Argument für *Agent Based Models* (vgl. ebd.: 35)? Wenn dynamische Systeme auf ein Problem mit sehr vielen heterogenen Objekten angewandt werden, kann man folgende Problematik beobachten:

System dynamics modeling requires the creation of a separate “stock” for each group of agents with different properties, and, when the space of

properties is large, this becomes difficult to build, track, and integrate.
(Wilensky & Rand 2015: 35)

Im Gegensatz dazu müssen ABMs nicht alle möglichen Typen von *Agents* verfolgen. Es reicht, wenn definiert wird, wie die Eigenschaften von *Agents* sich bei Interaktionen verhalten (vgl. ebd.: 35). Außerdem ist es für ABMs möglich verschiedenste Daten, wie geografische oder soziale, und Netzwerke, wie Freundeskreis oder Regierungsbestimmungen, in einer unterschiedlich detaillierten Genauigkeit einzubauen. Dies ist essentiell, weil das Verhalten der *Agents* zu einem gewissen Zeitpunkt durch diese Eigenschaft beziehungsweise Bestimmungen auch beeinflusst werden kann und somit die Heterogenität zwischen den einzelnen Einheiten vergrößert werden kann (vgl. De Marchi & Page 2014: 5).

- *Agent Based Models* sind ideal, wenn ein Problem mit komplexen Interaktionen lernfähiger *Agents* vorliegt. Ein Grund dafür ist, dass wenige einfache Regeln benötigt werden, um die gesamten Interaktionen in einem ABM zu beschreiben. Darüberhinaus ist es in dieser Modellform möglich, dass Objekte ihr Verhalten aufgrund verschiedenster Interaktionen ändern können (vgl. Wilensky & Rand 2015: 36). Das komplexe individuelle Verhalten der Einheiten spielt ebenso eine große Rolle. Je komplizierter das Verhalten wird, desto schwieriger wird es, alles mit mathematischen Gleichungen zu beschreiben, weshalb die Komplexität der Gleichungen enorm wächst (vgl. Bonabeau 2002: 7281).
- Es sind nicht nur die Interaktionen zwischen *Agent* und *Agent* wichtig, sondern auch die komplexen Interaktionen zwischen *Agent* und der Umgebung. Die meisten Umgebungen können als stationäre *Agents* dargestellt werden und somit steht dies in direkter Verbindung mit dem vorherigen Argument (vgl. Wilensky & Rand 2015: 36).
- Zu guter Letzt gibt ein ABM wesentlich mehr Informationen zurück als andere Modellformen. Es ist vor allem nicht nur möglich das Endresultat eines System zu beobachten sondern auch den zeitlichen Verlauf. Dies passiert sowohl auf der aggregierten Ebene als auch auf der Individuenebene. Darüberhinaus können jene Interaktionen integriert werden, die nach bestimmten Interaktionen auftreten beziehungsweise nur temporär vorhanden sind (vgl. ebd.: 36).

Bonabeau (2002: 7280–7281) ergänzt zwei wesentliche Punkte. Einerseits sind ABMs sehr flexibel, da diese Modellform das Hinzufügen beziehungsweise die Wegnahme von *Agents* erleichtert. Des Weiteren ist es möglich, die Komplexität der *Agents* zu regulieren. Dies bedeutet, dass das Verhalten, die Regeln für Interaktionen oder die Lernfähigkeit veränderbar sind. Die verschiedensten Möglichkeiten der Aggregation auf Basis der *Agents*

verhelfen zu weiterer Flexibilität. Damit ist gemeint, dass man Einheiten zu Gruppen zusammenfassen oder alle individuell betrachten kann (vgl. Bonabeau 2002: 7281).

Auf der anderen Seite können *Agent Based Models* Phänomene, sogenannte *Emergent Phenomena*, beschreiben. Diese entstehen durch das Zusammenspiel der einzelnen Einheiten im System. Die Eigenschaften des Phänomens sind jedoch oftmals andere oder gegensätzlich zu jenen der Einzelbausteine. Das bedeutet, dass es nicht möglich ist, vom Verhalten der einzelnen Teile direkt auf das Gesamtverhalten zu schließen. Ein Beispiel dafür wäre ein Verkehrsstau. Dieser entsteht durch das Zusammenspiel einzelner *Agents*, den Verkehrsmitteln, sodass es zu einer Verzögerung kommt. Die einzelnen Einheiten bewegen sich aber noch immer verlangsamt vorwärts. Sobald jedoch das Gesamtbild betrachtet wird, bewegt sich der Stau in die entgegengesetzte Richtung fort. Somit kann man nicht vom einzelnen Verhalten der Verkehrsmittel auf das Gesamtverhalten schließen. Diese Eigenschaft macht es schwer, solche Phänomene zu beschreiben und zu verstehen. Das zuvor erwähnte *bottom-up*-Prinzip der ABMs ermöglicht es, solche Probleme zu simulieren und vorherzusagen (vgl. ebd.: 7280).

Ein weiteres Beispiel, warum und weshalb ABMs in solchen Fällen ideal für die Anwendung sind, wäre folgendes Szenario: Angenommen es befinden sich 20 Menschen in einem Raum. Es wird ihnen gesagt, dass sich jeder der Anwesenden zwei andere Personen als Person A und Person B merken soll. Anschließend ist es das Ziel sich im Raum herumzubewegen und immer darauf zu achten, dass sich Person A zwischen Person B und der Person selbst befindet. Sobald das Startzeichen ertönt, bewegen sich alle Personen wild im Raum herum. Bei einem zweiten Durchgang wird die Regel leicht abgeändert. Nun muss sich die Person selbst zwischen Person A und Person B befinden. Plötzlich verändert sich das Gesamtverhalten dahingehend, dass einzelne Knoten von Personen entstehen und sie nicht mehr wild herumlaufen. Das hier beschriebene Spiel ist ein Paradebeispiel für die zuvor genannten Phänomene und der Vorteil von ABMs ist deutlich erkennbar: Mit kleinen und einfachen Regeln werden große Auswirkungen auf das Gesamtverhalten bewirkt (vgl. ebd.: 7280).

De Marchi & Page (2014: 5) schreiben weiters, dass *Agent Based Models* die Fähigkeit haben, Eigenschaften aus mehreren Bereichen beziehungsweise Domänen gleichzeitig zu beschreiben und zu simulieren. Das bedeutet, dass es keinen Sinn macht mit einem Modell rein ökonomische oder politische Probleme zu betrachten, weil sich diese gegenseitig beeinflussen. Darüberhinaus kann nicht nur das Problem manipuliert werden sondern auch das Verhalten der *Agents* und die Interaktionen zwischen diesen. *Agent Based Models* ist es möglich, diese Wechselwirkung zwischen zwei Bereichen mithilfe der Interaktionen zwischen den einzelnen *Agents* und den aufgestellten Regeln zu simulieren und somit ein besseres Bild der realen Welt zu entwickeln (vgl. ebd.: 5). Dabei ist es wichtig, Folgendes zu beachten: Falls zu viele Bereiche gleichzeitig betrachtet oder inkludiert werden, geht die Erklärbarkeit des Modells immer weiter verloren. Es können zwar interessante und

korrekte Ergebnisse entstehen, doch die Frage nach dem Warum ist durch die gegenseitige Beeinflussung der einzelnen Bereiche nicht mehr erklärbar. Die beste Lösung ist daher der Mittelweg, wobei so viele Bereiche wie nötig und nicht so viele Bereiche wie möglich inkludiert werden (vgl. De Marchi & Page 2014: 6).

Nachteile und Herausforderungen von *Agent Based Models*

Wie jede andere Modellierungsart haben ABMs Nachteile, die es oftmals schwer machen mit ABMs zu arbeiten.

Das Hauptproblem ist, dass die Modellierung sehr rechen- und zeitintensiv ist. Somit wird eine große Rechenleistung gebraucht, sobald das Problem komplexer oder die Anzahl der Objekte größer wird. Diese Tatsache muss in Kauf genommen werden, wenn bis auf Individuenebene simuliert wird. Jedoch ist das nicht nur ein Problem von ABMs sondern von jeder Modellierung. Je mehr Details in einem Modell eingebaut werden, desto mehr Rechenleistung wird benötigt (vgl. Wilensky & Rand 2015: 37; Bonabeau 2002: 7287).

Ein weiterer Nachteil ist, dass es in ABMs mehr Parameter gibt als bei anderen Modellierungsarten, welche vom Modellierer gewählt werden können oder müssen. In weiterer Folge müssen diese Parameter geschätzt werden und an die realen Daten angepasst werden. Dies resultiert darin, dass noch mehr Rechenleistung benötigt und die Simulationszeit deutlich erhöht wird (vgl. Wilensky & Rand 2015: 37).

Zu guter Letzt kann kein *Agent Based Model* erstellt werden, ohne Kenntnis zu haben, wie der Vorgang auf Individuenebene abläuft. Dies ist der entscheidende Punkt für ABMs. Im Gegensatz dazu ist es nicht wichtig aber von Vorteil, das Problem auf einem aggregiertem Level zu verstehen, um ein *Agent Based Model* zu erstellen (vgl. ebd.: 38).

Macal (vgl. 2016: 152) geht nicht genauer auf die Nachteile von ABMs ein, sondern auf die Herausforderungen, die es in Zukunft zu lösen gibt. Er erwähnt vier Hauptaufgaben.

1. Zunächst muss das Vertrauen und die Glaubwürdigkeit von *Agent Based Models* erhöht werden. In vielen Bereichen gibt es seit Jahren oder Jahrzehnten dieselben anerkannten und akzeptierten Techniken und Modelle, um gewisse Phänomene zu beschreiben. Deshalb ist es schwierig, Anerkennung und Akzeptanz aufzubauen, wenn bereits eine Lösung vorhanden ist (vgl. ebd.: 152).
2. Die Transparenz der Modelle muss erweitert werden. Das bedeutet, dass eine exakte Beschreibung der Modelle vorhanden sein muss um diese nachvollziehbar replizieren zu können. Somit können die Resultate unabhängig validiert werden (vgl. ebd.: 152).
3. Als Nächstes muss das Wissen über *Agent Based Models* standardisiert und erweitert werden. Damit ist gemeint, dass eine einheitliche Sprache und Definitionen entwi-

ckelt werden müssen, um eine formale Kommunikationsbasis zu generieren (vgl. Macal 2016: 152).

4. Zu guter Letzt müssen mehr Tools zur Generierung von ABMs entwickelt werden, welche leicht anwendbar und einheitlich sind (vgl. ebd.: 152).

4 *Agent Based Models* zur Prognose des SARS-CoV-2 Virus

Im Anschluss an die im vorigen Kapitel gegebene Einführung zu *Agent Based Models* wird eine Anwendung derer betrachte, die den Hauptteil dieser Arbeit markiert und aktueller nicht sein könnte: *Agent Based Models* zur Simulation der Ausbreitung des *SARS-CoV-2* Virus. Darüberhinaus geht es nicht nur um die exakte Berechnung der Verbreitung oder Fallzahlen sondern meist um die Evaluation verschiedenster Maßnahmen, die getätigt wurden beziehungsweise möglich wären. Dabei handelt es sich meist um nichtmedizinische Methoden zur Eindämmung der Pandemie.

Da dieses Feld für ABMs sehr aktuell ist, gibt es sehr viele verschiedene Möglichkeiten und Ansätze zur agentenbasierten Modellbildung. Dieses Kapitel wird eine Auswahl an einzelnen Modellen geben und deren Aufbau beschreiben und vergleichen, bevor im nächsten Abschnitt das im Zuge dieser Arbeit selbst erstellte Modell vorgestellt und die dazugehörigen Resultate diskutiert werden.

4.1 ABM-Simulation einer Stadt

Um die Ausbreitung des *SARS-CoV-2* Virus innerhalb einer Stadt zu simulieren, erstellten Shamil, Farheen, Ibteahaz, Khan & Rahman (2021) dieses Modell. Dabei wird Ford County in den USA zum Validieren des Modells herangezogen. In weiterer Folge werden alle Experimente für eine verkleinerte Version von New York City, USA, durchgeführt. Im Folgenden wird zunächst der Aufbau beschrieben, bevor im Anschluss die Ergebnisse diskutiert werden (vgl. ebd.: 3).

Datengrundlage

In diesem Modell werden zwei Arten von Daten benötigt:

1. Erstens werden ortsspezifische Daten verwendet. Dabei handelt es sich auf der einen Seite um demografische Eigenschaften der Einwohner der Stadt und auf der anderen Seite um Daten über die Transportmöglichkeiten und die durchschnittliche Familiengröße. Darüberhinaus werden spezifische Daten bezüglich der COVID-19 Pandemie von der betrachteten Stadt benötigt. Diese sind die Anzahl der Infektionen in der Stadt und das Datum der Einschränkungen durch die lokale Regierung (vgl. ebd.: 3).
2. Zweitens sind in diesem Modell physiologische Daten von großer Bedeutung. Diese Daten werden wie folgt beschrieben:

The probability of a person coughing and sneezing, touching contaminated objects, coming into physical contact with others, or washing hands is also

an important parameter of our model, which would differ based on whether a person is at work, home, or hospitalized. (Shamil et al. 2021: 3)

Die Veränderung der Population über die Zeit durch Migration, Geburt oder Tod wird im vorliegenden Modell nicht berücksichtigt (vgl. ebd.: 3).

Aufbau der Agents

Jeder Einwohner der Stadt ist ein eigener *Agent*, der über die zuvor erwähnten Eigenschaften beschrieben wird. Jeder dieser *Agents* wird einer Familie zugeteilt und übt einen der vier Berufe, Mitarbeiter im Gesundheitswesen, Schüler, Arbeitnehmer oder Arbeitsloser, aus. Jeder dieser vier Typen verhält sich im Laufe einer Zeiteinheit sehr unterschiedlich. Somit gibt es für jeden eine eigene Menge an möglichem Verhalten oder Aufgaben pro Zeiteinheit (vgl. ebd.: 2).

- Ein Schüler kann folgendes Verhalten während einer Zeiteinheit durchführen: Zu hause bleiben, zur Schule gehen, lernen, nach Hause gehen oder eine Veranstaltung besuchen (vgl. ebd.: 2).
- Die Verhaltensregeln für Mitarbeiter im Gesundheitswesen unterscheiden sich natürlich von jenen eines Schülers, wie nachfolgend zu erkennen ist: Zu hause bleiben, zum Spital gehen, Patienten betreuen, nach Hause kommen oder eine Veranstaltung besuchen (vgl. ebd.: 2).
- Ein Arbeitnehmer hat einen ähnlichen Arbeitsablauf wie Mitarbeiter im Gesundheitswesen: Zu hause bleiben, zur Arbeit gehen, arbeiten, nach Hause gehen oder eine Veranstaltung besuchen (vgl. ebd.: 2).
- Zu guter Letzt unterscheidet sich das Verhalten eines Arbeitslosen nur dadurch, dass er nicht zur Arbeit geht: Zu hause bleiben, hinausgehen, eine Veranstaltung besuchen oder nach Hause gehen (vgl. ebd.: 2).

Bei Betrachtung der vier Arten und deren Verhalten fällt auf, dass sie alle die Möglichkeit haben, eine Veranstaltung zu besuchen. Dieses Verhalten soll die Freizeitkontakte mit anderen *Agents* widerspiegeln. Die Möglichkeiten des Verhaltens werden mithilfe anderer dynamischer Parameter unter Kontrolle gehalten. Das bedeutet, dass es zwei Parameter (Minimum und Maximum), die die Startzeit einer Aufgabe an diesem Tag fixieren, zwei Variablen (Minimum und Maximum), die die Dauer dieser Aufgabe beschreiben, und zwei Parameter für die Eintrittswahrscheinlichkeit dieser Aufgabe gibt. Neben dem berufsspezifischen Verhalten und Aufgaben gibt es noch gewisse Aktionen, die jeder *Agent* durchführen kann: niesen, Gegenstände kontaminieren, physischen Kontakt zu anderen *Agents* haben oder Hände waschen. All diese kleinen Aktionen, die jeder *Agent* durchführen kann, haben grundsätzlich einen negativen Einfluss auf die Übertragung des Virus

und betreffen eine andere Person. Eine Ausnahme stellt lediglich das Händewaschen dar, weil es einen positiven Einfluss für den *Agent* selbst hat. Diese Aktionen besitzen selbst wieder zehn Parameter, die sie regulieren. Somit haben sie jeweils zwei (Minimum und Maximum), um Folgendes zu kontrollieren (vgl. Shamil et al. 2021: 3):

- die mögliche Zeit zwischen zwei aufeinanderfolgenden Aktionen
- die Wahrscheinlichkeit über die Effektivität der Aktionen
- die Eintrittswahrscheinlichkeit dieser Aktion
- die Wahrscheinlichkeit für die Auswirkung auf andere *Agents* durch diese Aktion
- die Wahrscheinlichkeit über die eigene Beeinflussung durch diese Aktion

Des Weiteren hat jeder *Agent* eine Variable, *protection_level* genannt. Sie erhöht sich, wenn Maßnahmen wie das Tragen von Masken, welche die Übertragung des Virus verhindern sollen, eingeführt werden oder in Kraft treten. Natürlich kann dieser Wert von *Agent* zu *Agent* variieren und jeder benötigt einen anderen hohen Wert, damit er geschützt ist. Dies ist auch abhängig vom Beruf (vgl. ebd.: 3).

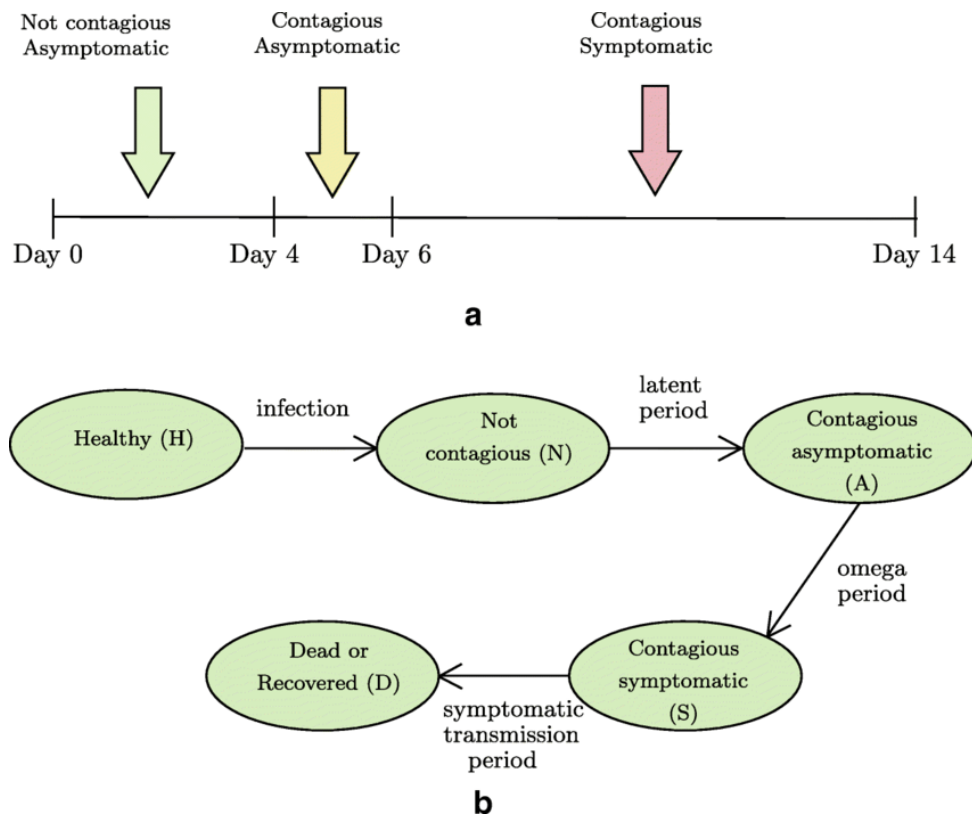


Abbildung 2: Länge und Verlauf der Zustände für die *Agents* (Shamil, Farheen, Ibte haz, Khan & Rahman 2021: 3)

Bevor die Beschreibung der Interaktionen und der Übertragung des Virus startet, müssen die möglichen Zustände der *Agents* erklärt werden. Zunächst ist wichtig festzuhalten,

dass jeder *Agent* mit *COVID-19* erkranken kann. Dabei durchleben alle die folgenden fünf Zustände, die in Abbildung 2 bildlich dargestellt sind:

- gesund oder nicht infiziert (H)
- infiziert, aber nicht ansteckend, asymptomatisch (N)
- infiziert, ansteckend und asymptomatisch (A)
- infiziert, ansteckend und symptomatisch (S)
- tot oder geheilt/immun (D)

Wie in Abbildung 2 zu erkennen ist, wird angenommen, dass es vier Tage dauert bis ein *Agent* nach einer Infektion ansteckend ist. Dies wird die Latenzzeit genannt. Nach weiteren zwei Tagen treten erste Symptome auf und nach insgesamt 14 Tagen hat ein *Agent* die Infektion überstanden (vgl. Shamil et al. 2021: 2).

Interaktionen zwischen *Agents* und die Übertragung des Virus

Bevor *Agents* in diesem Modell miteinander interagieren, werden sie pro Zeiteinheit in Gruppen auf Basis deren Verhaltens eingeteilt. Nach jeder Zeiteinheit werden diese Gruppen neu zugeteilt. Dabei unterscheiden Shamil et al. (ebd.) zwischen fünf Gruppen:

- zu Hause bleiben (F)
- zur Arbeit fahren (T)
- arbeiten oder zur Schule gehen (W)
- eine Veranstaltung besuchen (E)
- im Spital bleiben (H)

Für den öffentlichen Verkehr werden zunächst die freien Plätze auf Basis der Anzahl an Transportmitteln berechnet und einem neuen *Agent* wird automatisch ein freier Platz zugewiesen. Sobald eine Person einen Platz erhalten hat, wird dieser als besetzt markiert und steht nicht mehr zur Verfügung (vgl. ebd.: 4).

Innerhalb dieser Gruppen interagieren im Grunde alle miteinander. Das bedeutet, dass eine negative Aktion wie das Niesen einer infektiösen Person jeden *Agent* innerhalb dieser Gruppe beeinflusst. Jedoch gibt es Unterschiede basierend auf der Nähe der *Agents* zueinander. In diesem Modell wird jedem möglichen Paar innerhalb einer Gruppe ein Parameter für die Nähe zugeteilt, der zwischen Null und Eins sowie innerhalb vordefinierter Intervalle liegen muss (vgl. ebd.: 4).

Als Nächstes wird die Berechnung einer Infektion näher erklärt. Es ist wichtig, dass nur negative Aktionen wie das Niesen von Personen, die infektiös sind, einen Einfluss auf die Infektion einer Person haben. Dabei wird ein Kennwert namens *infection* wie folgt berechnet:

$$\text{infection} = (1 - \text{protection_level}) * E * \text{proximity}$$

Die Variable *protection_level* wurde bereits zuvor näher erklärt und *proximity* ist eben die Nähe, die im vorherigen Absatz genauer beschrieben wurde. Für E gilt: $0 \leq O_{min} \leq E \leq O_{max} \leq 1$. Dabei bezeichnen O_{min} und O_{max} die Wahrscheinlichkeitsgrenzen der Auswirkung einer Aktion auf andere *Agents*, welche im Teil über den Aufbau der *Agents* als dynamische und kontrollierende Parameter eingeführt wurden (vgl. Shamil et al. 2021: 4).

Im nächsten Schritt müssen die positiven Aktionen einer Person für sich selbst berücksichtigt werden. Dabei wird ähnlich zur zuvor verwendeten Formel vorgegangen:

$$\text{infection} = (-1) * (1 - \text{protection_level}) * E$$

Hier wird keine Nähe mehr benötigt, da diese Aktion Auswirkungen auf einen *Agent* selbst haben und nicht auf andere. Für E gilt: $-1 \leq S_{min} \leq E \leq S_{max} \leq 0$, wobei S_{min} und S_{max} die Grenzen der Auswirkung auf einen *Agent* selbst sind (vgl. ebd.: 4).

Nach diesen Berechnungen wurden jedem *Agent* Werte für *infection* pro Aktion zugewiesen. Doch wann ist ein *Agent* in diesem Modell infiziert? Zunächst wird ein Grenzwert, der *action_infection_threshold*, eingeführt. Sobald der Wert *infection* über diesem Grenzwert liegt, wird dieser zum *infection_level* einer suszipblen Person hinzugezählt (oder bei positiven Aktionen abgezogen). Dieser *infection_level* summiert sich über den Tag hinweg auf und wird am Ende des Tages mit einem weiteren Grenzwert verglichen, dem *infection_threshold*. Wenn der *infection_level* über dem *infection_threshold* liegt, kommt es zu einer Infektion dieses *Agents*. Die Einführung dieser Wahrscheinlichkeiten und der Grenzwerte hat den Hintergrund, dass nicht jede Aktion gleichbedeutend zu einer Infektion ist und dass auch die Summe aus mehreren Aktionen zu einer Infektion führen kann (vgl. ebd.: 4).

Die Umgebung

Zunächst ist anzumerken, dass sich die Simulation innerhalb einer Stadt abspielt. Dabei werden den *Agents* keine genauen Geoinformationsdaten für deren Zuhause zugewiesen. Sie werden einfach zu Familien und Gruppen, die von ihrem derzeitigem Verhalten abhängen, zugewiesen. Das bedeutet, dass in diesem Modell keine exakten Straßendaten, Gebäudedaten oder Ähnliches von einer Stadt benötigt werden.

Des Weiteren wird in diesem Modell eine Stadtregierung eingesetzt. Dieser ist es mög-

lich, Bestimmungen wie einen Lockdown, Maskenpflicht oder Schulschließungen durchzusetzen. Dabei werden in weiterer Folge eine Quarantäne und Contact Tracing eingeführt. Contact Tracing wird in diesem Modell aber nicht auf die klassische Art und Weise mittels Befragung der Personen simuliert, da dies auch in der Realität schwer durchführbar ist. Es wird daher angenommen, dass jeder Smartphonenuutzer eine Tracing-App installiert hat und somit automatisch in Quarantäne geschickt werden kann, sobald er sich mit einer infektiösen Person mehrere Tage hintereinander in derselben Gruppe aufgehalten hat (vgl. Shamil et al. 2021: 4).

Zusammenfassung der Resultate

Zusammenfassend ist zu erwähnen, dass das Modell die realen Coronadaten der Stadt Ford County sehr gut modelliert (vgl. ebd.: 7). Darüberhinaus wurden einige Experimente durchgeführt, die die verschiedensten Kombinationen von Lockdown und Contact Tracing simulieren. Diese resultieren darin, dass die effektive Reproduktionszahl innerhalb von drei Wochen unter Eins fällt, wenn 60 Prozent der Population über das Handy nachverfolgbar sind und gleichzeitig ein Lockdown herrscht. Sobald mehr als 75 Prozent nachverfolgbar sind, würde die effektive Reproduktionszahl schneller unter Eins fallen und die Anzahl der Neuinfektionen innerhalb von drei Monaten sogar bei Null sein (vgl. ebd.: 2).

4.2 Agent Based Model „COVID-19 ABM“

Dieses agentenbasierte Modell wurde für die Stadt Salzburg und deren umliegenden Gemeinden erstellt. Es ist aber nach Anpassung einiger Parameter leicht auf andere Städte anwendbar. Das Ziel dieses Modells ist es, die dynamische Ausbreitung des *SARS-CoV-2* Virus in der zuvor genannten Region zu modellieren und gewisse Szenarien von Interventionen der Regierung zu simulieren. Dabei greift es auf zwei existierende *Agent Based Models* zurück. Eines ist für die Aktivitäten und die Mobilität der *Agents* zuständig und das zweite übernimmt den Teil des Infektionsverhalten. Nachfolgend werden der Aufbau des Modells und die Resultate kurz diskutiert (vgl. Wallentin, Kazyieva & Reibersdorfer-Adelsberger 2020: 509 f.).

Datengrundlage

Für dieses Modell wurden vier verschiedene Datensätze zur Initialisierung verwendet.

1. Es werden anfänglich infizierte Personen in diesem Modell benötigt, die aus öffentlichen Daten vom Land Salzburg berechnet wurden. Dabei ist zu berücksichtigen, dass bei diesen infizierten *Agents* eine Exponentialverteilung für die Variable „Tage seit Infektion“ angenommen wird (vgl. ebd.: 511).
2. Als Nächstes werden Daten zur Bevölkerung benötigt. Diese sollten in einem 250 Meter mal 250 Meter Raster vorhanden sein. Im Punkt *Aufbau der Agents* wird näher auf diesen Raster eingegangen (vgl. ebd.: 511).
3. Darüberhinaus müssen Daten über Gebäude für die Aktivitäten vorhanden sein. Das bedeutet, dass Punktdaten für Arbeitsplatz, Schule, Einkaufsläden und vieles mehr benötigt werden (vgl. ebd.: 511).
4. Zum Schluss sind auch demografische Daten der Personen in dieser Region für eine gute Simulation erforderlich (vgl. ebd.: 511).

Die Daten der letzten drei Punkte werden hauptsächlich von der Homepage von *Statistik Austria* gewonnen.

Aufbau der Agents

In der untersuchten Region befinden sich ungefähr 200 000 Einwohner, welche als *Agents* modelliert werden. In dem Modell werden drei Gruppen unterschieden.

- Einereits werden die infizierten und potenziell infizierten *Agents* explizit als einzelne *Agents* simuliert (vgl. ebd.: 510).

- Die gesunden und suszeptiblen Personen werden in einem geografischen Raster dargestellt. Dieser hat eine Größe von 250 Meter mal 250 Meter (vgl. Wallentin et al. 2020: 510).
- Zu guter Letzt werden die genesenen und immunen *Agents* ebenso in einem Raster von 250 Meter mal 250 Meter dargestellt (vgl. ebd.: 510).

Agents interagieren innerhalb der Zelle miteinander, in der sie sich gerade befinden. Hier stellt sich die Frage, ob ein feinerer Raster notwendig wäre. Wallentin et al. (ebd.) beantwortet dies wie folgt: „Finer grids for demographic data are neither available due to data privacy restrictions, nor necessary to model human activity patterns in very high resolution.“ Weiters werden alle Aktivitäten, die sie im Laufe eines Tages ausführen, sowie die Mobilität der *Agents* über das Modell „Bicycle model for Salzburg City“ simuliert. Dabei werden die einzelnen Fahrradwege, die ein *Agent* zwischen zwei Aktivitäten zurücklegt, herausgenommen und nur die Aktivität innerhalb eines Tages betrachtet. Das bedeutet, dass der Weg zur Arbeit oder zur Schule nicht berücksichtigt wird. Detaillierter wird auf dieses Modell nicht eingegangen, da es den Rahmen dieser Arbeit sprengen würde (vgl. ebd.: 511).

Abschließend werden noch drei Annahmen zu den *Agents* gemacht.

1. Es wird angenommen, dass kein *Agent* ansteckend ist, solange er nicht Symptome entwickelt (vgl. ebd.: 510).
2. Weiters wird die Annahme getätigt, dass jeder *Agent* nach einer überstandenen Infektion für längere Zeit immun ist. Es war zur Zeit der Erstellung dieses Modells noch nicht klar, ob eine Immunität nach der Infektion herrscht und wie lange diese andauert (vgl. ebd.: 510).
3. Die *Agents* halten sich komplett an die verordnete Heimquarantäne und niemand verlässt sein Zuhause. Diese Annahme wurde nicht nur für *Agents* getätigt, die in Quarantäne geschickt werden, sondern auch für alle symptomatischen Personen (vgl. ebd.: 510).

Interaktionen zwischen Agents und die Übertragung des Virus

Um das Verhalten der *Agents* und deren Interaktionen besser zu verstehen, muss zunächst erwähnt beziehungsweise angenommen werden, dass als Kontakte, die relevant für das Infektionsgeschehen sind, nur solche herangezogen werden, die mindestens 15 Minuten andauern und bei denen ein Mindestabstand von zwei Metern unterschritten wird. Dies wurde ebenso in die „Stopp Corona“ App in Österreich implementiert. Aufgrund dessen beträgt die Zeiteinheit in diesem Modell 15 Minuten und das Modell unterteilt sich in zwei Phasen. Auf der einen Seite jene, die nur einmal pro Tag durchgeführt wird, und

auf der anderen Seite die Phase, die alle 15 Minuten wiederholt wird (siehe Abbildung 3) (vgl. Wallentin et al. 2020: 510).

Bevor eine Simulation überhaupt beginnen kann, müssen die *Agents* initialisiert und deren Infektionsstatus basierend auf realen Daten gewählt werden. Danach wird die Simulation gestartet und beginnt mit der Phase 1, welche die Abläufe darstellt, die nur einmal pro Tag durchgeführt werden. Darauf folgt dann der Tagesablauf in Zeitintervallen von 15 Minuten, der mit Phase 2 bezeichnet wird. Im Folgenden werden die einzelnen Phasen näher beschrieben (vgl. ebd.: 510).

- Bevor der Tag startet, wird immer der Infektionsstatus entsprechend angepasst. Falls hier ein *Agent* gesund wird, kommt er zurück in den Populationsraster von zuvor und wird nicht mehr als einzelner *Agent* simuliert. Um diesen Status zu erreichen, muss er zuvor, wie im vorherigen Modell, einige Gesundheitszustände durchleben. In diesem Modell sind die Abläufe basierend auf Expertenwissen wie folgt definiert: Sobald ein *Agent* infiziert wurde, ist er die ersten drei Tage infiziert, aber nicht infektiös. Am viertem Tag ist er infektiös und ohne Symptome, zwischen fünftem und 15. Tag ist er infektiös und hat Symptome und ab dem 16. Tag ist er genesen. Das bedeutet, dass er einen Tag ohne Symptome ansteckend ist und somit nur einen Tag Zeit hat, andere anzustecken (vgl. ebd.: 510).
- In der Phase 2 spielt sich das Hauptaugenmerk bei den infizierten *Agents* ab. Wie in Abbildung 3 erkennbar ist, werden nur die Kontakte jener *Agents* betrachtet, die auch infektiös sind. Zunächst wird geprüft, ob sich suszible Personen überhaupt an deren Ort befinden, bei denen es zu einer erfolgreichen Übertragung des Virus kommen kann. Dabei ist zu berücksichtigen, dass sobald ein *Agent* infiziert wird, er sich nicht mehr im Raster der Population befindet und ab diesen Zeitpunkt als einzelner *Agent* simuliert wird (vgl. ebd.: 510).

Wie funktioniert in diesem Modell eine erfolgreiche Übertragung des Virus auf einen anderen *Agent*? Es werden zwei Hauptparameter verwendet. Zunächst die Basisreproduktionszahl R_0 . Dieser Faktor beschreibt die Ausbreitungsgeschwindigkeit des Virus und die sozialen Interaktionen in einer Population. Aus diesem wird die Übertragungswahrscheinlichkeit berechnet. Dabei ist zu berücksichtigen, dass sich diese aufgrund der lokalen Immunitäten oder des Gebäudes, in dem sich der *Agent* aufhält, reduzieren kann (vgl. ebd.: 510 f.).

Weiters werden die durchschnittlichen Kontakte eines *Agents* pro Tag benötigt. Hierbei handelt es sich um Kontakte, die für das Infektionsgeschehen relevant sind, also länger als 15 Minuten andauern und der Abstand unter zwei Meter ist. Bei diesem Parameter ist zu beachten, dass er stark zwischen einzelnen Staaten und Kulturen variieren kann (vgl. ebd.: 511).



Abbildung 3: Detaillierter Tagesablauf innerhalb des Modells mit großem Augenmerk auf den Infektionsteil

(Wallentin, Kaziyeve & Reibersdorfer-Adelsberger 2020: 510)

Zusammenfassung der Resultate

Mithilfe dieses Modells werden vier verschiedene Szenarien für die Zeit nach dem ersten Lockdown in der Region um die Stadt Salzburg durchsimuliert:

- Ein verlängerter Lockdown, welcher für die ganze Population durchgeführt wird, bis eine Ausrottung des Virus zustandekommt (vgl. Wallentin et al. 2020: 511).
- Eine stufenweise Lockerung des Lockdowns wie es in Österreich tatsächlich passierte.
- Stufenweise Lockerung des Lockdowns mit parallelem Contact Tracing: Es wird das Szenario vom vorherigem Punkt um die Möglichkeit des Contact Tracings erweitert. Dabei wird zwischen drei verschiedenen Stärkegraden des Contact Tracings unterschieden (vgl. ebd.: 512).
- Als letztes Szenario wird wieder eine stufenweise Lockerung hergenommen, wobei

diese sehr genau überwacht wird. Sobald ein Wert über einen gewissen Grenzwert wächst, werden die Maßnahmen wieder aufgenommen (vgl. Wallentin et al. 2020: 512).

Wallentin et al. (ebd.) kamen bei deren Simulationen auf folgende Ergebnisse: Einerseits würde eine Verlängerung des Lockdowns relativ rasch eine Ausrottung des Virus bedeuten. Andererseits würde nach den Ergebnissen des Modells eine stufenweise Öffnung ohne beziehungsweise mit nur geringem Contact Tracing zu einem neuerlichen Anstieg der Infektionszahlen führen. Ein hohes Contact Tracing Management könnte jedoch eine stufenweise Öffnung ermöglichen und sogar bei sehr intensivem Contact Tracing zu einer Ausrottung des Virus führen (vgl. ebd.: 512–514).

4.3 Agent Based Model für ganz Österreich

Dieses Modell wurde aufgrund der *SARS-CoV-2* Pandemie in Österreich entwickelt. Mit Hilfe dessen wurden politische Entscheidungsträger unterstützt und gewisse Maßnahmen evaluiert und validiert. Die aktuellen Prognosen für Österreich stützen sich meist auf dieses Modell. Die Grundlage bildet ein zuvor entwickeltes *Agent Based Model*, das sogenannte GEPOC, das die Veränderung einer Population in einem Staat über die Jahre hinweg simulieren kann. In den nachfolgenden Unterkapiteln wird das Modell näher erklärt (vgl. Bicher, Rippinger, Brunmeir, Urach & Popper 2020: 1 f.).

Datengrundlage

In diesem Modell wird eine Vielzahl an verschiedenen Datenquellen in einer Einheit zusammengeführt. Dabei können diese Datenquellen in drei Gruppen unterteilt werden.

- Zunächst einmal sind Daten zu nennen, die die Bevölkerung von Österreich beschreiben und nachbilden. Dabei handelt es sich einerseits um die Bevölkerungsverteilung und deren demografische Daten und andererseits um Daten, die die Bevölkerung noch näher beschreiben, wie zum Beispiel Arbeitslosenzahlen, Pendlerzahlen und viele mehr (vgl. ebd.: 11).
- Des Weiteren werden viele Daten zur geografischen Lage in Österreich benötigt, wobei es sich in erster Linie um eine geografische Karte von Österreich mit allen Straßen handelt. Darüberhinaus sind aber auch die räumlichen Verteilungen der Krankenhäuser, Bildungseinrichtungen und Arbeitsstätten notwendig (vgl. ebd.: 11).
- Zu guter Letzt werden einige Datensätze bezüglich des Coronavirus benötigt. Es handelt sich vor allem um Verteilungen gewisser Parameter, die den Verlauf einer *SARS-CoV-2* Infektion bestimmen und einschätzen können. Diese sind oftmals abhängig von dem Alter der *Agents* oder anderen Einflüssen (vgl. ebd.: 16–19).

Aufbau der Agents

Jede Person eines Landes oder einer Region wird als ein einzelner *Agent* modelliert. Dieser hat verschiedenste Eigenschaften, die er erfüllen muss, wobei diese Attribute in vier Gruppen unterteilt werden: Populations-, Kontakt-, Krankheits- und Verordnungseigenschaften. Im folgenden Teil werden diese vier Gruppen und die darin enthaltenen Eigenschaften genauer beschrieben (vgl. ebd.: 1 f.).

1. *Population*: Diese Attribute beinhalten hauptsächlich demografische Daten der einzelnen *Agents*. Dabei handelt es sich zum Beispiel um das Geschlecht, das Alter und

die Position des Zuhauses, wobei das Zuhause mit Longitude und Latitude beschrieben wird. Somit wird dem Zuhause ein exakter Punkt im Land zugewiesen (vgl. Bicher, Rippinger et al. 2020: 1).

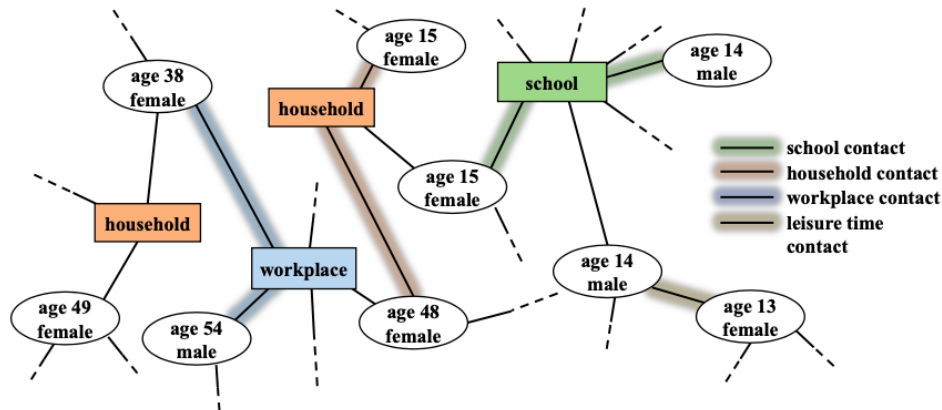


Abbildung 4: Kontaktnetzwerk der möglichen Kontakte von *Agents* in diesem Modell (Bicher, Rippinger, Brunmeir, Urach & Popper 2020: 2)

2. *Kontakt*: In diesem Modell sind die Kontakte essentiell für die Verbreitung des Virus. Dabei gibt es viele Möglichkeiten und Attribute eines *Agents*, die zu Kontakten führen. Zunächst wird jeder *Agent* einem Haushalt über die Position des Zuhauses zugewiesen. Darüberhinaus muss jeder zur Arbeit gehen, wobei zwischen Arbeitslosen, Arbeitern, Schülern und Pflegemitarbeitern unterschieden wird. Die Zuordnung zu einem Arbeitsplatz, egal ob Schule oder andere Arbeitsstätten, erfolgt immer über die Distanzen des Zuhauses zu den Arbeitsplätzen. Des Weiteren gibt es für jeden *Agent* eine gewisse Anzahl an Freizeitkontakten, die zufällig auf Basis der Abstände zwischen den einzelnen *Agents* gezogen werden. Insgesamt hat jeder *Agent* die Möglichkeit sich an drei verschiedenen Orten durch Kontakte zu anderen zu infizieren. Diese Orte sind als Knoten in einem Netzwerkgraphen zu verstehen, in denen sich die Personen treffen, wie in Abbildung 4 dargestellt ist (vgl. ebd.: 1 f.).
3. *Krankheit*: Jeder *Agent* durchläuft wie in den anderen Modellen verschiedene Status, die den Gesundheitszustand beschreiben. Folgende sind die Wichtigsten: „infected, infectious, symptoms, hospitalised, critical, confirmed, severe, asymptomatic, home-quarantined and recovered.“ (ebd.: 2) Dabei können all diese Eigenschaften nur den Wert *True* oder *False* annehmen. Des Weiteren ist es möglich, dass mehrere dieser Attribute zur gleichen Zeit den Wert *True* annehmen. Das bedeutet, dass das eine das andere nicht ausschließt und dass der Status wiederum die Aktionen der *Agents* beeinflusst (vgl. ebd.: 2).
4. *Verordnung*: Diese Eigenschaften betreffen vor allem die getroffenen Verordnungen beziehungsweise Einschränkungen. Dabei können diese auf die Arbeitsstätten ab-

zielen, die als offen oder geschlossen markiert sind, oder direkt das Verhalten der *Agents* beeinflussen. Ebenso wurde eine Variable, die sogenannte präventive Quarantäne, eingeführt, die die Personen beschreibt, die aufgrund des Contact Tracings in Quarantäne sind (vgl. Bicher, Rippinger et al. 2020: 2).

Wegen des komplexen Aufbaus ist eine der größten Herausforderung die Agenteninitialisierung, auf welche jedoch nicht näher eingegangen wird, da dies den Rahmen dieses Kapitels sprengen würde.

Interaktionen zwischen Agents und die Übertragung des Virus

Dieses Modell kann als eine hybride Version von *Agent Based Models* angesehen werden, da es sowohl ein zeitdiskretes als auch ein zeitstetiges Modell ist. Die Hauptsimulation wird in Tagen als Zeiteinheit aktualisiert. Innerhalb jedes Zeitschrittes teilt sich die Simulation in drei Teile. Teil 1 ist die Planungsphase, in welcher alle Aktionen geplant werden, die ein *Agent* während eines Tages ausführen soll. Danach startet Teil 2, wobei die zuvor geplanten Aktionen ausgeführt werden. Der letzte Teil ist die Datensammelungsphase, in der die aggregierten Werte der Zustände aufgezeichnet werden (vgl. ebd.: 3).

Neben der Übertragung des Virus sind die wichtigsten Aktionen, die ein *Agent* tätigen kann, die Kontakte zu anderen *Agents*. Wie schon zuvor erwähnt, kann der *Agent* Kontakte über den Haushalt, über die Arbeit und über die Freizeit generieren. Dabei wird die Anzahl der Kontakte über den Ort des Kontakts unterschieden. Zunächst ist klar, dass innerhalb eines Tages ein *Agent* zu jedem in seinem Haushalt lebenden *Agent* einmal Kontakt hat. Als Nächstes werden die Arbeitsplatzkontakte beschrieben. Hier wird eine Zufallszahl nach einer gewissen Verteilung gezogen, die die Anzahl der Kontakte vorgibt. Danach werden zufällig andere *Agents* von dieser Arbeitsstätte oder Schule für die zuvor definierte Anzahl gezogen, mit denen er Kontakt hat. Zu guter Letzt wird ebenso eine Zufallszahl für die Anzahl der Freizeitkontakte ermittelt und es werden wieder zufällig andere Personen ausgewählt. Hierbei wird eine Distanzmatrix berücksichtigt, die die Abstände der einzelnen *Agents* beschreibt und somit die Wahrscheinlichkeit für ein Freizeittreffen vorgibt. Die Summe all dieser Kontakte spiegeln die Interaktionen der einzelnen *Agents* innerhalb eines Tages. Diese Kontakte werden immer am Anfang des Tages ausgeführt (vgl. ebd.: 3). Da diese alle distanzabhängig sind, bezeichnen Bicher, Rippinger et al. (ebd.) jene als „spatially-continuous locally-biased contact network.“

Das Entscheidende in jedem dieser Modelle sind der Ablauf und der Verlauf einer Infektion. In diesem Modell wird die Übertragung nicht, wie in dem vorherigem Modell, über die Reproduktionszahl geregelt, sondern über eine kontaktspezifische Übertragungswahrscheinlichkeit. Dies hat den Vorteil, dass Parameter wie die Reproduktionszahl, die in traditionellen SIR Modellen als Voraussetzung benötigt werden, aus diesem Modell geschätzt werden können. Über die Infektionswahrscheinlichkeit wird geregelt, ob nun ein

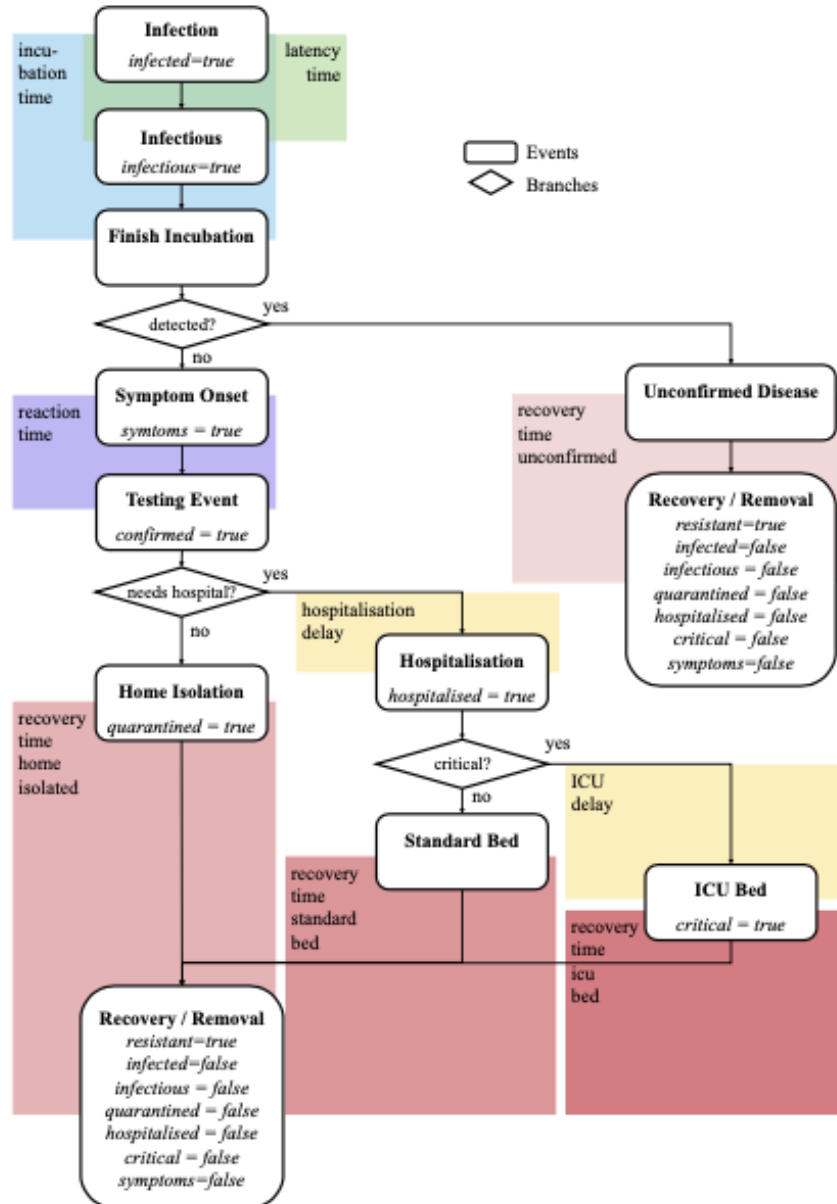


Abbildung 5: Verlauf des Gesundheitsstatus von *Agents* in diesem Modell (Bicher, Rippinger, Brunmeir, Urach & Popper 2020: 5)

infektiöser *Agent* bei einem Kontakt einen anderen *Agent* ansteckt. Dabei wird zwischen Kontakten innerhalb des Haushalts sowie innerhalb der Arbeit oder der Freizeit unterschieden und mit anderen Wahrscheinlichkeiten versehen. Sobald eine Infektion stattgefunden hat, beginnt der Gesundheitsverlauf, wie in Abbildung 5 beschrieben, und der *Agent* durchläuft die einzelnen Stationen der Infektion. Wie läuft dieser Gesundheitsverlauf nun ab? Nach der Infektion wird einem *Agent* eine Latenzzeit zugewiesen, die nach einer speziellen Verteilung berechnet wird. Nach dieser ändert sich sein Status zu „infektiös“. Sobald der Status auf „ansteckend“ geändert wird, wird eine Zufallszahl gezogen, welche über das Auftreten von Symptomen entscheidet. Dabei wird auf altersabhängige

Grenzen zurückgegriffen. Das bedeutet, dass ältere *Agents* eine größere Wahrscheinlichkeit haben als jüngere, Symptome zu entwickeln. Dies bildet die erste Abzweigung im Verlauf der Infektion. Der restliche Verlauf, wie er in Abbildung 5 abgebildet ist, passiert nach demselben Schema, wobei die Zeiten immer auf spezielle Verteilungen zurückgreifen (vgl. Bicher, Rippinger et al. 2020: 3 f.).

Zusammenfassung der Resultate

Die Resultate von diesem Modell sind wöchentlich sichtbar. Es wird zusammen mit zwei anderen Modellen für die Prognose der Entwicklung der Fallzahlen in Österreich verwendet, bildet oftmals die Grundlage für Diskussionen über Öffnungen oder Verschärfungen und dient als Entscheidungshilfe für die Politiker:innen Österreichs. Der Mehrwert und die Validierung wurden an den aktuellsten Daten Österreichs durchgeführt und die Prognosen haben in letzter Zeit als Grundlage für Entscheidungen gedient.

4.4 Vergleich der Modelle

Wie in den obigen Beschreibungen der Modelle ersichtlich ist, gibt es Gemeinsamkeiten und Unterschiede. Dies ist vor allem im Aufbau der *Agents* und in der Übertragung des Virus am besten zu erkennen. Im Folgenden werden die Gemeinsamkeiten und Unterschiede dieser beiden Teile beschrieben.

Agents

Bei den *Agents* gibt es einige Gemeinsamkeiten. Jedes Modell verwendet Ortsdaten, die das Zuhause und die Arbeitsstätte beziehungsweise Schule beschreiben. Der Unterschied liegt dabei vor allem in der Genauigkeit. Während in Modell 4.3 auf Longitude und Latitude der einzelnen Standorte zurückgegriffen wird, gibt es in Modell 4.1 keine genaue Angabe, wo sich diese befinden. Des Weiteren sind die Zustände, die ein *Agent* während einer Infektion durchlebt, nahezu identisch und variieren wiederum nur in der Genauigkeit. Die Zustände „infizierbar“, „infiziert“, „infektiös“ und „genesen“ durchleben die *Agents* in jedem Modell und diese bilden zugleich das Minimum an möglichen Zuständen der *Agents*. Ohne diese vier Zustände würde kein *Agent Based Model* für die Modellierung einer Infektionskrankheit funktionieren. Weiters werden auch in allen Modellen die einzelnen Zeiten zwischen den Zuständen berücksichtigt.

In diesem letzten Punkt steckt aber auch einer der größten Unterschiede. In den Modellen 4.1 und 4.2 werden die Zeiten, wie Latenzzeit oder Inkubationszeit, fixiert und variieren nicht von *Agent* zu *Agent*. Im Gegensatz dazu wird im Modell 4.3 für jede dieser Zeiten eine spezielle Verteilung verwendet, sodass diese über die *Agents* hinweg variieren können. Dies spiegelt die reale Welt besser wider, da man diese Zeiten nicht generalisieren kann. Ein weiterer wichtiger Punkt ist, dass die Agentensimulation zwischen den einzelnen Modellen unterschiedlich ist. Damit ist gemeint, dass bei den Modellen 4.1 und 4.3 jeder einzelne *Agent* als eine Einheit angesehen und modelliert wird. Im Modell 4.2 werden nur die Infizierten einzeln betrachtet und die *Agents*, die noch infiziert werden können, werden als eine große Gruppe simuliert. Mitunter der größte Unterschied liegt in den Aktionen und Interaktionen, die ein *Agent* tätigen kann. Auf der einen Seite werden nur Kontakte mit anderen *Agents* simuliert und aus diesen Kontakten weitere Infizierte generiert, wie in den Modellen 4.2 und 4.3 zu erkennen ist. Auf der anderen Seite steht das Modell 4.1. In diesem wird jede kleine Tätigkeit wie niesen und kontaminieren eines Gegenstandes von einem *Agent* simuliert. Somit gibt es nicht nur Kontakte, mit denen die *Agents* interagieren, sondern sie verkehren auch indirekt miteinander, indem Aktionen gesetzt werden, die die Anderen beeinflussen. Ebenso werden positive Vorgänge, wie Hände waschen, in diesem Modell berücksichtigt.

Infektion

Ein wesentlicher Unterschied aller Modelle ist die Herangehensweise in der Übertragung des Virus. Dabei werden drei verschiedene Ansätze verwendet. Zunächst verfolgen Wallentin et al. (2020) im Modell 4.2 einen klassischen Ansatz. Es wird auf die Reproduktionszahl zurückgegriffen, die auch bei Differentialgleichungsansätzen, wie bei SIR-Modellen, benötigt wird. Ein ähnlicher Ansatz wird im Modell 4.3 verwendet. Jedoch wird hier auf die Reproduktionszahl verzichtet und die Infektion über eine Übertragungswahrscheinlichkeit geregelt. Der Vorteil dabei ist, dass die Reproduktionszahl dadurch geschätzt werden kann. Ein komplett anderen Ansatz wird im Modell 4.1 herangezogen. Hier distanziert man sich von der Reproduktionszahl und führt Grenzwerte für eine Infektion ein. Diese entsteht nur dann, wenn die Summe der einzelnen Aktionen die Grenzwerte überschreitet. Dieser völlig andere Ansatz ermöglicht es Shamil et al. (2021), die Aktionen auf so detaillierter Ebene zu simulieren.

Zusammenfassend lässt sich sagen, dass der *Agentteil* in den einzelnen Modellen bis auf die Genauigkeit ähnlich ist. Der große Unterschied entsteht erst im Infektionsteil, genauer gesagt darin, wie ein *Agent* infiziert werden kann. Alle drei Modelle ergeben aber eine valide Simulation, die zu interessanten Resultaten geführt hat. Dies ist das Spezielle an *Agent Based Models*. Es wird dasselbe Phänomen aus drei verschiedenen Richtungen betrachtet, womit drei verschiedene Ansätze entstehen, die alle jedoch valide Ergebnisse liefern.

5 Selbstentwickeltes agentenbasiertes Modell

Der Hauptteil dieser Arbeit behandelt die Erstellung eines selbstentwickelten *Agent Based Model*, welches es ermöglicht, die Ausbreitung des *SARS-CoV-2* Virus zu modellieren und mögliche nichtmedizinische Sanktionen zu simulieren, um deren Auswirkungen zu studieren. Bei der Entwicklung dessen spielten mehrere Faktoren eine große Rolle, die die Modellierung beeinflussten.

Zunächst war es wichtig, dass dieses Modell auf einem herkömmlichen Computer in einer vertretbaren Zeit lauffähig ist. Deshalb beschränkt sich das Modell auf einen Bezirk in Österreich, da dabei die Anzahl der *Agents* überschaubar ist und somit die Laufzeit erheblich reduziert wird. Obwohl die Resultate sich auf den Bezirk Melk beschränken, ist dieses Modell auch für jeden anderen einsetzbar, solange die Kalibrierung der Infektionswahrscheinlichkeiten auf die Realdaten des gewählten Bezirks durchgeführt wird. Des Weiteren wurden manche Abläufe und der Infektionsverlauf vereinfacht. Auf diesen Punkt wird in einem späteren Kapitel noch näher eingegangen. Ein weiterer wichtiger Faktor war die Datenverfügbarkeit. Nicht alle Daten, die in den zuvor erklärten Modellen verwendet wurden, sind für die Öffentlichkeit zugänglich, wie zum Beispiel Mobilitätsdaten von Mobilfunkanbietern. Deshalb basiert dieses generierte Modell auf Daten, welche für jedermann zugänglich und öffentlich sind. Der Nachteil davon ist, dass das Modell an Genauigkeit verliert. Dies macht das Ganze sehr spannend. Es stellt sich nämlich die Frage, ob es möglich ist, mit weniger beziehungsweise weniger genauen Daten und einem weniger detaillierten Aufbau ein valides *Agent Based Model* zu erzeugen, welches für die Prognose und Validierung von Maßnahmen geeignet ist. Die Einschränkung auf einen Bezirk stellte weiters eine sehr große Herausforderung dar, da zu Beginn der Pandemie manche Bezirke noch nicht derart betroffen waren wie andere oder wie ganz Österreich.

In den folgenden Kapiteln wird zunächst auf die Datengrundlage und deren Probleme eingegangen, um danach den Aufbau des Modells genau zu beschreiben. Eine weitere wichtige Rolle spielte die Agenteninitialisierung, welche sich als sehr schwierig darstellte. Zu guter Letzt werden die Resultate dargestellt und in weiterer Folge diskutiert.

5.1 Datengrundlage

Die Datengrundlage ist im Vergleich zu den zuvor gezeigten Modellen um einiges geringer beziehungsweise aggregierter, weil es oftmals nur Daten auf Landesebene gibt. Deshalb wurde bei diesen Daten die Annahme getroffen, dass die Verteilung auf Landesebene in etwa der Verteilung in den einzelnen Bezirken entspricht. Generell können die Datensätze in drei Gruppen unterteilt werden:

- geografische Daten
- personenbezogene Daten
- Infektionsdaten

Geografische Daten

Hierbei handelt es sich um Daten, welche den Bezirk und seine Gemeinden mithilfe longitudinaler und latitudinaler Koordinaten widerspiegeln. Darüberhinaus wird das gesamte Straßennetz innerhalb eines Bezirks benötigt, weil in weiterer Folge einige Kontakt und die Zuweisung zu Haushalten, Schulen und Arbeitsstätten von der Distanz abhängen. Der erste Gedanke war, die Distanz der Mittelpunkte der einzelnen Gemeinden als Maß für den Abstand heranzuziehen, sozusagen die Luftlinie von einem Mittelpunkt zum anderen. Das ist jedoch nur bedingt richtig, da es immer auf die Straßenverbindung zwischen zwei Bezirken ankommt. Betrachtet man zum Beispiel Gemeinden, die durch einen Fluss voneinander getrennt sind und es keine Brücke über diesen innerhalb der Grenzen der beiden benachbarten Gemeinden gibt, müssen die Einwohner über eine andere Gemeinde fahren, um zur anderen zu kommen. Somit würde zwar die Luftlinie zeigen, dass der Nachbar beim Fluss der Nächste ist. Berücksichtigt man jedoch das Straßennetz, so wird der Abstand um einiges größer. Deswegen wird neben den Bezirks- und Gemeindegrenzen auch das gesamte Straßennetz innerhalb des Bezirks benötigt. In Abbildung 6 sind die Daten für den Bezirk Melk visualisiert.

Um diese Daten zu erhalten wurden die Grenzen und das Straßennetz von der Homepage *OpenStreetMap* für diesen Bezirk heruntergeladen. Es handelt sich hier aber nicht um einen einmaligen Download, der für jeden anderen Bezirk wiederholt werden muss, sondern es gibt die Möglichkeit, in der Analysesoftware R ein Paket zu installieren, welches es ermöglicht mittels einer API geografische Daten je nach gewähltem Bezirk zu laden.

Personenbezogene Daten

Wie die zuvor erwähnten Modelle benötigt dieses auch Daten, die die Bevölkerung innerhalb eines Bezirks widerspiegeln. Dabei gibt es aufgrund der Datenverfügbarkeit einige

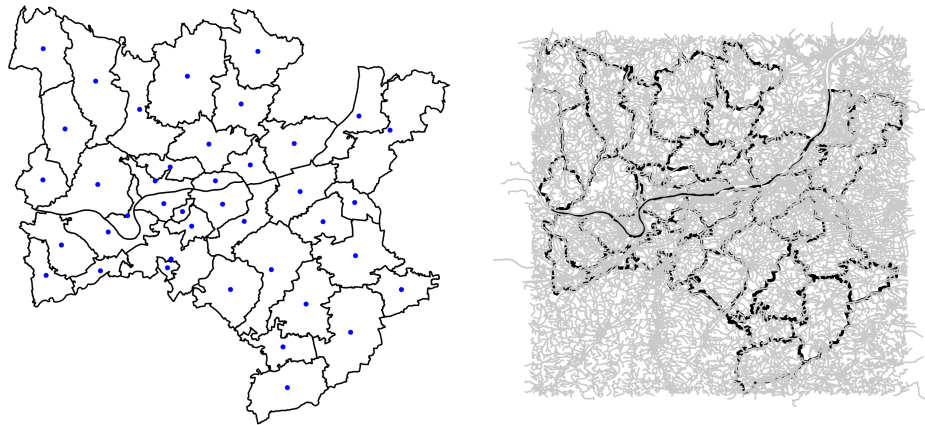


Abbildung 6: Die geografischen Daten am Beispiel des Bezirks Melk. Auf der linken Seite sieht man die Umrisse des Bezirks und der Gemeinden mit den Mittelpunkten in Blau. Die rechte Abbildung beinhaltet zusätzlich das gesamte Straßennetz in grau.
(Eigene Darstellung)

Herausforderungen. Zunächst sind die Gesamteinwohnerzahl und die Verteilung in Geschlecht und Alter erforderlich. Dabei stellte sich heraus, dass es auf Bezirksebene keine öffentlich zugängliche Datenquelle mit der Verteilung des Alters gibt. Deshalb wurde hier die Verteilung der Personen auf die Einzeljahre je nach Geschlecht innerhalb eines Bundeslandes verwendet. In weiterer Folge werden auch Daten bezüglich der Arbeitsplätze, Schulen und Haushalte erforderlich. Für Erstere gibt es einen Datensatz in *Statistik Austria*, welcher die Anzahl der Arbeitsstätten und die Beschäftigten in diesen Arbeitsstätten auf Gemeindeebene beinhaltet. Aus diesem lässt sich in weiterer Folge die Anzahl der Arbeitsplätze innerhalb einer Gemeinde und einer Arbeitsstätte berechnen. Bezüglich der Schulen gibt es einen idealen Datensatz, welcher die Anzahl der Schulen je Schultyp pro Gemeinde angibt. Zusätzlich wird auf einen Datensatz zurückgegriffen, der die Schüleranzahl je Schultyp pro Gemeinde beinhaltet. Aus der Kombination dieser beiden wird somit die Anzahl der Schulplätze berechnet. Neben den Schulen werden in diesem Modell auch die Kindergärten getrennt betrachtet, damit eine eventuelle Auswirkung der Schließung oder Nichtschließung beobachtet werden kann. Dazu gibt es wiederum einen Datensatz, welcher die Anzahl der Kindergärten und Kindergartenkinder auf Gemeindeebene beinhaltet. Als letzte personenbezogene Datengrundlage werden Informationen bezüglich der Haushalte benötigt. Da es abermals keine Daten auf Bezirksebene gibt, wird auf einen Datensatz zurückgegriffen, der die Verteilung der Haushalte aufgeteilt in Größe des Haushalts auf Landesebene beinhaltet.

Für all diese Datensätze wird *Statistik Austria* als Quelle herangezogen. Nähere Informationen, wie der Link zur Internetseite und der letzte Zugriff, ist in der Tabelle im Anhang A zu finden.

Infektionsdaten

Die letzte und zugleich wichtigste Datenquelle sind die Infektionszahlen. Die Granularität dieser Daten ist auf Bezirksebene und beinhaltet neben der Anzahl der Einwohner auch die Anzahl der täglichen Neuinfektionen, Toten und Geheilten, die Summe der Neuinfektionen, Toten und Geheilten und die *7-Tages-Inzidenz*. Diese Informationen stammen von *Open Data Österreich* und wurden einmalig heruntergeladen. Wie schon zuvor erwähnt, gibt es auch hierzu nähere Informationen in der Tabelle im Anhang A.

5.2 Aufbau des Modells

Der Modellaufbau stellt den wichtigsten Teil von diesem Kapitel dar. Zunächst ist anzumerken, dass dieses Modell im Bezug auf das Konzept sehr ähnlich zu jenem von Bicher, Rippinger et al. (2020) ist. Aufgrund der Datenbasis sind jedoch einige gravierende Unterschiede und Vereinfachungen entstanden. Die Zeiteinheit bei der Modellierung beträgt immer einen Tag. Folglich werden zunächst die *Agents* und deren Zustände näher beschrieben.

Agents

In diesem Modell wird jeder Einwohner eines Bezirks als ein einzelner *Agent* dargestellt und nicht zu Gruppen zusammengefasst. Die Zustände und Zeitverläufe der *Agents* sind angelehnt an das Modell in Kapitel 4.3, wobei jedoch eine wesentliche Anpassung getroffen wurde, welche eine große Auswirkung für dieses generierte Modell hat. Es wurde komplett auf die Hospitalisierungen verzichtet. Diese wurde stattdessen in den Quarantänebranch integriert. Das ist möglich, da Personen, die hospitalisiert sind, wie andere Personen in Quarantäne aus dem Kontaktnetzwerk herausgenommen werden. Somit ist der Ablauf derselbe wie für Personen in Quarantäne. Es wird jedoch ein zusätzlicher Aufwand benötigt, um dies umzusetzen. Der Nachteil ist, dass dadurch keine Hospitalisierungszahlen und Todeszahlen betrachtet werden können. Im Gegenzug wird bei der Modellierung viel Aufwand eingespart und somit die Simulationszeit verkürzt.

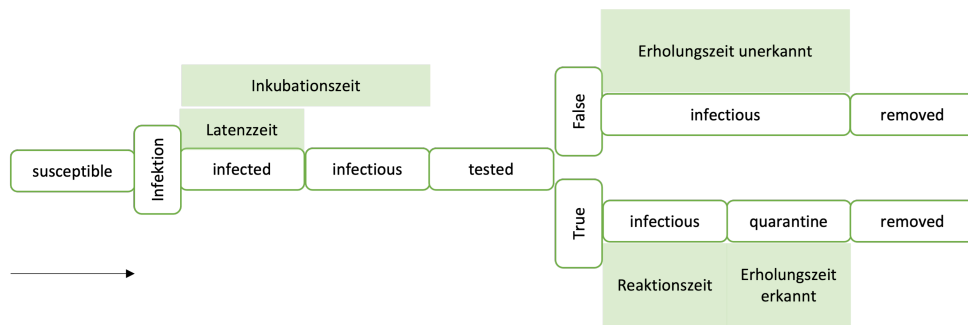


Abbildung 7: Zustandsverlauf in diesem Modell mit den einzelnen Zeiten (grün hinterlegt) (Eigene Darstellung)

Jeder *Agent* wird, wie in den anderen Modellen, zunächst initialisiert und bekommt seinen Ausgangszustand. Dieser Vorgang wird im nachfolgenden Unterkapitel diskutiert und erklärt. Wie in Abbildung 7 erkennbar ist, kann ein *Agent* im Laufe der Simulation sechs Zustände einnehmen: *susceptible*, *infected*, *infectious*, *tested*, *quarantine* und *removed*.

- *susceptible* und *removed*: Diese beiden Zustände stehen für den Anfang und das Ende einer Infektion. Jeder *Agent* ist, solange er noch nicht infiziert ist, empfänglich für

das Virus. Er hat daher den Status *susceptible*. Nachdem eine Person eine Infektion durchgemacht hat, ist der Betroffene in der realen Welt entweder gestorben oder genesen. Da in diesem Modell kein Unterschied zwischen gestorben und genesen gemacht wird, werden alle als *removed* bezeichnet, sobald die Infektion überstanden ist.

- *infected*: Dies ist der Startschuss des Krankheitsverlaufs. Nach einer „erfolgreichen“ Infektion wird der Status des *Agents* auf *infected* gesetzt. In diesem Zustand ist der *Agent* infiziert, aber nicht ansteckend. Diesen Status behält er, bis die Latenzzeit vorbei ist, welche mithilfe von einer Verteilung erzeugt wird, die auch im Modell im Kapitel 4.3 verwendet wird und im Anhang A zu finden ist.
- *infectious*: Nach der Latenzzeit ändert sich der Status des *Agents* auf *infectious*. Ab diesem Zeitpunkt ist der *Agent* ansteckend und kann das Virus an andere *Agents* weitergeben, falls diese den Status *susceptible* haben. Dieser Status bleibt auf jeden Fall aufrecht, bis die Inkubationszeit vorbei ist. Es lässt sich erkennen, dass die Zeit, die jemand auf jeden Fall ansteckend ist, die Inkubationszeit weniger der Latenzzeit ist. Die Berechnung der Inkubationszeit ist wiederum im Anhang A zu finden.
- *tested*: Dieser Zustand ist im engeren Sinn eigentlich ein Ereignis. Er gibt an, ob ein *Agent* durch einen Test als infiziert erkannt wird oder nicht. Dabei wird eine altersabhängige Verteilung (siehe Anhang A) verwendet, um die erkannten *Agents* zu bestimmen. Ab diesem Zeitpunkt teilt sich der Verlauf in zwei Zweige auf. Falls ein *Agent* nicht erkannt wird, bleibt er für eine gewisse Zeit, die wiederum über eine Verteilung bestimmt wird, infektiös und hat die Möglichkeit andere *Agents* anzustecken. Im zweiten Zweig bleibt der *Agent* noch die sogenannte Reaktionszeit, welche die Zeit zwischen Test und Testergebnis widerspiegeln soll, ansteckend, bevor er in Quarantäne geschickt wird.
- *quarantine*: Dies ist der letzte Zustand bevor der *Agent* als genesen gilt. Die Person wird in Quarantäne geschickt und alle Kontakte werden unterbunden. Vor allem wird die Annahme getroffen, dass auch keine Kontakte zum eigenen Haushalt mehr bestehen und dass sich die *Agents* komplett abschotten. Wie lange die Quarantäne dauert, wird wieder über eine Verteilung bestimmt (siehe Anhang A).

Sobald ein *Agent* eine Infektion durchgemacht hat und somit den Status *removed* hat, gilt dieser als genesen. Ebenso wird die Annahme getroffen, dass genesene Personen auch immun gegen das Virus sind. Dies ist zwar in der realen Welt nicht wahr, da nach einer gewissen Zeit der Schutz nachlässt und man wieder empfänglich für das Virus wird. Jedoch kann angenommen werden, dass es sicher über einen Monat dauert, bis der Schutz nachlässt. Diese Annahme kann getroffen werden, da eine Simulation meist nicht länger als

30 Tage dauert. Ein weiterer Punkt ist, dass die verwendeten Verteilungen für die Zeiten und die altersabhängige Wahrscheinlichkeitsaufteilung bei den Tests aus dem Modell von Bicher, Rippinger et al. (2020) herangezogen werden, weil sowohl die Berechnung dieser den Rahmen der Arbeit sprengen würde, als auch die benötigten Daten meistens nicht verfügbar sind.

Agenteninitialisierung

Die Agenteninitialisierung ist mitunter der wichtigste Punkt neben der eigentlichen Simulation, denn sie schafft die Grundlage für den weiteren Verlauf der Modellierung. Falls diese die reale Bevölkerung nicht entsprechend abbildet, ist die gesamte nachfolgende Modellierung zu hinterfragen.

Bevor aber die eigentlichen Daten für die einzelnen *Agents* erstellt werden, müssen zusätzliche Informationen aus bestimmten Datenquellen generiert werden. Dies betrifft vor allem die geografischen Daten aus Abschnitt 5.1. Die Initialisierungen bezüglich der Arbeit, der Schule oder des Kindergartens sind ortsabhängig und werden deshalb über eine Distanz zwischen den einzelnen Orten geregelt. Der Idealzustand wäre, dass für jede dieser Ortsvariablen eine longitudinale und eine latitudinale Koordinate vorhanden ist. Jedoch ist es schwierig dies umzusetzen und dazu öffentliche Daten zu finden, weshalb in diesem Modell auf eine Vereinfachung zurückgegriffen wird, indem sich der Ortsunterschied auf die unterschiedlichen Gemeinden innerhalb eines Bezirks beschränkt. Dies kann man sich so vorstellen, dass es pro Bezirk verschiedene Gemeinden und innerhalb dieser Gemeinden einen Mittelpunkt gibt, über den die Distanzen zwischen den Gemeinden berechnet werden. Zum besseren Verständnis wird die Variable Haushalt betrachtet. *Agents* werden aufgrund der Bevölkerungsverteilung der Gemeinden einer Heimatgemeinde und innerhalb dieser einem Haushalt zugewiesen. Im Laufe der Simulation treffen sich die *Agents* mit anderen. Offensichtlich ist es wahrscheinlicher, dass sich ein *Agent* mit einem anderen *Agent* innerhalb seiner Gemeinde beziehungsweise aus einer benachbarten Gemeinde trifft. Es genügt jedoch nicht mehr nur die Distanzen zu betrachten, sondern man muss auch die Bevölkerungszahl innerhalb einer Gemeinde mitberücksichtigen, denn je größer die Bevölkerung innerhalb einer Gemeinde ist, desto wahrscheinlicher ist es, sich mit jemanden aus dieser Gemeinde zu treffen. In diesem Modell wird daher die *Pendelwahrscheinlichkeit* eingeführt, mithilfe derer alle distanzabhängigen Variablen bestimmt werden. Die *Pendelwahrscheinlichkeit* wird folgendermaßen berechnet: Sei a_j die relative Populationsgröße der Gemeinde j in einem gewählten Bezirk und sei d_{ij} ein Aufwandsmaß für das Pendeln von Gemeinde i zu j . Dabei kann dieses Aufwandsmaß von der Distanz bis hin zur Reisezeit reichen. Für das Aufwandsmaß gilt: $0 < d_{ii} < d_{ij} \forall j \neq i$. Dann gilt

für die *Pendelwahrscheinlichkeit* p_{ij} von i nach j folgende Formel:

$$p_{ij} = \frac{1}{c_i} * \frac{a_j}{d_{ij}}$$

c_i wird dabei wie folgt festgelegt: $c_i = \sum_k \frac{a_k}{d_{ik}}$.

Mithilfe dieser Formel entsteht aus der Aufwandsmatrix eine Wahrscheinlichkeitsmatrix, mit der die ortsabhängigen Prozesse bestimmt werden können. In diesem Modell wurde für das Aufwandsmaß d_{ij} der Abstand zwischen den Mittelpunkten mittels des Straßennetzes der geografischen Daten berechnet. Dabei ist zu berücksichtigen, dass d_{ii} immer gleich Null ist. Um dies zu verhindern, werden die Aufwandsmaße d_{ij} für alle j um den durchschnittlichen Abstand des Mittelpunkts der Gemeinde i zum eigenen Rand erhöht. Somit gilt für alle i : $0 < d_{ii} < d_{ij} \forall j \neq i$. Dadurch ist die Voraussetzung für die Formel gewährleistet und es entsteht in der Folge eine Wahrscheinlichkeitsmatrix, wie in Abbildung 8 zu erkennen ist.

	centre_i	31502	31503	31504	31505
1	31502	0.067	0.024	0.008	0.019
2	31503	0.014	0.102	0.009	0.032
3	31504	0.01	0.02	0.076	0.017
4	31505	0.01	0.03	0.007	0.201

Abbildung 8: Ein Beispiel der Wahrscheinlichkeitsmatrix für vier Gemeinden mit den Kennzahlen 31502, 31503, 31504 und 31505 im Bezirk Melk
(Eigene Darstellung)

Nachdem die Wahrscheinlichkeiten für einen Bezirk berechnet wurden, kann die eigentliche Agentenbildung beginnen. Zunächst wird die Anzahl der Einwohner als *Agents* erstellt. Diesen *Agents* wird danach das Geschlecht gemäß der Verteilung in dem Bezirk aus den realen Daten zugewiesen. Um das Alter zu erhalten, muss als Nächstes auf den Datensatz der Einzeljahre auf Landesebene zurückgegriffen werden. Dabei wird je nach Geschlecht die Verteilung auf die Einzeljahre im jeweiligen Bundesland herangezogen. Somit haben die *Agents* ein Geschlecht und ein Alter, welches sich mit der realen Verteilung annähernd deckt.

Weiters werden die Haushalte zugewiesen. Wie zuvor erwähnt, gibt es hierbei nur Heimatgemeinden und keine exakten Koordinaten. Deshalb wird jedem *Agent* gemäß der Verteilung der Personen auf die realen Gemeinden eine Heimatgemeinde zugewiesen, indem jeder *Agent* die Gemeindekennzahl als eindeutigen Schlüssel zu seiner Gemeinde erhält. Daraufhin werden gemäß der realen Verteilung auf Länderebene solange Haushalte in einer Gemeinde erstellt, bis mindestens alle *Agents* einen Platz in einem Haushalt haben. Um nun jeden *Agent* einem eindeutigen Haushalt innerhalb seiner Heimatgemeinde zuzuweisen, wird in diesem Modell folgendermaßen vorgegangen: Zunächst wird jedem Haushalt ein *Agent* zugeteilt, der älter als 18 Jahre ist. Somit kann es auf keinen Fall passieren,

dass ein Kleinkind alleine in einem Haushalt ist oder nur Kinder zusammenleben. Danach werden alle anderen gemäß ihrer Heimatgemeinde einem Haushalt zugewiesen, indem sie eine eindeutige Nummer für deren Haushalt bekommen. Das heißt, dass jeder *Agent* einem eindeutigen Haushalt über die Gemeinde gemeinsam mit der Haushaltsnummer zugeteilt wird.

Im nächsten Schritt folgen die Arbeitsvariablen, genauer gesagt die Arbeits-, Schul- und Kindergartenzuweisung. Begonnen wird mit der Arbeitsplatzzuweisung. Bevor die *Agents* einem Arbeitsplatz zugewiesen werden, müssen die Arbeitsplätze erstellt werden. Dabei wurden auf Grundlage der Daten pro Gemeinde die Arbeitsstätten pro Gemeinde mittels einer eindeutigen Kennzahl erstellt. Die Arbeitsplätze wurden danach über die Anzahl der Beschäftigten in jeder Gemeinde ermittelt. Dies geschieht durch eine zufällige Ziehung mit Zurücklegen aus der Menge der Arbeitsstätten, bis die Anzahl der Beschäftigten erreicht wird. Dadurch entstehen Betriebe sowohl mit vielen als auch mit wenigen Arbeitsplätzen. Bevor die *Agents* nun einen Arbeitsplatz erhalten, werden sie in folgende Gruppen unterteilt:

- *Pensionisten*: Das sind *Agents*, die 65 Jahre oder älter sind. Sie werden keiner Arbeit zugewiesen.
- *Arbeiter*: Das sind *Agents*, die zwischen 18 und 65 Jahre alt sind. Sie werden einem Arbeitsplatz zugewiesen.
- *Arbeitslose*: Das sind *Agents*, die im arbeitendem Alter sind, aber arbeitslos sind und somit keiner Arbeit zugewiesen werden. Dies wird mittels einer Arbeitslosenquote in den einzelnen Gemeinden berechnet.
- *Schüler*: Das sind *Agents*, die zwischen sieben und 18 Jahren alt sind. Sie werden einem Schulplatz zugewiesen.
- *Kindergartenkinder*: Das sind *Agents*, die zwischen vier und sechs Jahre alt sind. Sie werden einem Kindergartenplatz zugewiesen.
- *Kinder*: Das sind *Agents*, die jünger als vier Jahre sind. Sie werden nicht zugewiesen und haben somit nur den Haushalt und die Freizeit als Interaktionsmöglichkeit.

Als Nächstes werden die Arbeiter einem Arbeitsplatz zugewiesen. Hierbei ist zu berücksichtigen, dass nicht jeder Arbeiter in seiner Gemeinde oder in seinem Bezirk bleibt und dort seine Arbeit verrichtet. Es gibt einen gewissen Anteil an Auspendlern. Mithilfe des Datensatzes „*Gemeindeergebnisse der Abgestimmten Erwerbsstatistik und Arbeitsstättenzählung 2019 (Gebietsstand 2019)*“ und der Tabelle 3 im Anhang A wird zunächst der Ort der Arbeit, dass heißt in der Heimatgemeinde, im Bezirk oder außerhalb des Bezirks,

der einzelnen *Agents* bestimmt. Danach wird als Erstes den *Agents*, welche in der Heimatgemeinde bleiben, ein Arbeitsplatz zugewiesen. Danach wird den *Agents*, die innerhalb des Bezirks bleiben, ein freier Arbeitsplatz mithilfe der zuvor berechneten *Pendelwahrscheinlichkeit* zugeordnet. Den restlichen *Agents* wird kein Arbeitsplatz zugewiesen, da sie außerhalb des Bezirks arbeiten.

Die initialisierung der anderen arbeitsspezifischen Ausprägungen funktioniert auf ähnliche Weise. Bei den Schülern in der Volksschule werden zunächst alle Schüler in die Volksschule deren Heimatgemeinden geschickt. Falls es zu wenig Platz gibt, werden sie nach der *Pendelwahrscheinlichkeit* in andere Gemeinden aufgeteilt, in denen noch ein Platz frei ist. Die anderen Schüler werden basierend auf der *Pendelwahrscheinlichkeit* den freien Schulplätzen in den einzelnen Gemeinden des Bezirks zugewiesen. All jene, die keinen Platz bekommen haben, werden außerhalb des Bezirks unterrichtet. Bei der Zuweisung zu Kindergartenplätzen wird sehr ähnlich vorgegangen. Zunächst werden anhand realer Daten alle Kindergartenplätze erstellt und danach den Kindern innerhalb einer Gemeinde zugewiesen. Gibt es innerhalb einer Heimatgemeinde zu wenig Plätze, werden die restlichen Kinder mittels der Pendelwahrscheinlichkeit anderen Gemeinden mit freien Plätzen zugewiesen. Falls es dann noch immer freie Plätze gibt, werden Kleinkinder im Alter von zwei oder drei Jahren den restlichen Plätzen zugeordnet.

Nach all diesen Vorgängen ist die erste Initialisierungsphase abgeschlossen. Den *Agents* wurden gewisse Eigenschaften basierend auf realen Daten zugewiesen. In weiterer Folge müssen nun die Anfangszustände jener berechnet werden. Dies wird in einem nachfolgendem Unterkapitel erklärt.

	Id_agent	sex	alter	Id_municipality	Id_household	type_of_work	Id_municipality_work	Id_workplace
1	19334	W	2	31533	1403	child	31533	0
2	24910	M	6	31549	1897	kindergarden	31549	1
3	754	M	80	31546	443	pension	31546	0
4	42409	M	10	31508	447	primary_school	31508	1
5	5645	W	17	31506	142	school	31525	1
6	35210	W	38	31515	103	unemployed	31515	0
7	21569	W	36	31520	851	work	31521	178

Abbildung 9: Hier sieht man einen kleinen Ausschnitt des Datensatzes, der die einzelnen *Agents* nach der Initialisierung beschreibt. Über die letzten drei Spalten ist ein *Agent* eindeutig einer Arbeitsstätte zugewiesen. Über die Spalten *Id_municipality* und *Id_household* ist der *Agent* einem eindeutigen Haushalt zugewiesen. Des Weiteren wird jedem *Agent* eine eindeutige ID über die Spalte *Id_agent* zugewiesen, damit jeder eindeutig abrufbar ist.

(Eigene Darstellung)

Interaktionen zwischen Agents und Ablauf einer Infektion

Es stellt sich nun die Frage, wie die *Agents* miteinander interagieren und wie es dabei zu einer Infektionen kommen kann. Zunächst wird die Interaktion zwischen den *Agents* näher beschrieben. Hierbei ist das Modell sehr ähnlich zu jenem aus Kapitel 4.3. Die

Interaktionen zwischen den *Agents* beruhen auf den Kontakten zwischen diesen. Jedoch werden die Kontakte je nach deren Art unterschieden.

- Eine Gruppe sind die Haushaltskontakte, welche jene Kontakte darstellen, die *Agents* zu deren Haushaltsmitgliedern haben. Innerhalb eines Tages hat ein *Agent* zu jedem Haushaltsmitglied einen Kontakt, sofern keiner der anderen oder er selbst in Quarantäne ist.
- Des Weiteren werden für jeden *Agent* Freizeitkontakte erstellt. Dabei wird mittels einer gewissen Verteilung (siehe Anhang A) die Anzahl der Kontakte pro Tag eines *Agents* berechnet. Danach werden mithilfe der *Pendelwahrscheinlichkeit* die *Kontaktagents* ausgewählt.
- Als Nächstes werden die Arbeitsplatzkontakte berechnet. Dies geschieht ähnlich wie zuvor. Jeder *Agent* bekommt über eine Wahrscheinlichkeitsverteilung eine Anzahl an Arbeitskontakten zugewiesen. Danach werden solange zufällig Mitarbeiter aus seiner Arbeitsstätte mit Zurücklegen gezogen, bis er die richtige Anzahl an Kontakten hat.
- Die weiteren Gruppen der Schul-, Volksschul- und Kindergartenkontakte basieren auf demselben Prinzip wie die Arbeitsplatzkontakte, nur mit anderen Verteilungen.

	Id_agent	type_of_contact	Id_contact
1	58224	high_school	72729
2	2505	household	11575
3	57170	kindergarden	32214
4	24257	leisure	28371
5	42740	primary_school	6775
6	25989	workplace	18966

Abbildung 10: Hier sieht man einen kleinen Ausschnitt aus der Kontakttable. Die *Id_agent*-Spalte gibt den *Agent* an, der den Kontakt zu dem *Id_contact* hat. (Eigene Darstellung)

Alle Kontakte werden zu Beginn eines jeden Tages erstellt und im Laufe des Tages ausgeführt. Ein Ausschnitt der entstehenden Kontakttable wird in Abbildung 10 dargestellt.

Damit sich ein *Agent* mit dem Virus ansteckt, muss der Kontaktagent infektiös sein. Eine Infektion über einen Kontakt wird mittels einer Infektionswahrscheinlichkeit simuliert, welche zwischen den einzelnen Kontakttypen unterschiedlich sein kann. Dabei wird

zunächst auf eine anfängliche Infektionswahrscheinlichkeit von Bicher, Rippinger et al. (2020) (siehe Anhang A) zurückgegriffen, die in weiterer Folge an die realen Daten angepasst wird. Somit kann am Ende des Tages über die Kontakte entschieden werden, ob sich ein *Agent* infiziert hat oder nicht und ob sich der Gesundheitsstatus ändert.

Initialisierung der Anfangszustände

Nun kommen wir zu einem der wichtigsten Punkte bei der Modellierung, der Initialisierung der Anfangszustände. Diese ist wesentlich für den weiteren Verlauf der Modellierung und entscheidend, ob diese überhaupt Sinn macht. Dabei wird zwischen zwei möglichen Ansätzen unterschieden.

Auf der einen Seite kann es eine zufällige Initialisierung nach der Vorgabe der Anzahl an Infizierten geben. Das bedeutet, dass nachdem die statischen Variablen, welche über die Zeit hinweg unverändert bleiben, der *Agents* erstellt wurden, zufällig welche ausgewählt werden können, die infiziert, infektiös oder in Quarantäne sind. Es kann somit eine gewisse Anzahl an *Agents* als infiziert, infektiös oder in Quarantäne zufällig gewählt werden. Dieser Vorgang macht dann Sinn, wenn allgemeine Ergebnisse betrachtet werden, wie zum Beispiel beim Vergleich der unterschiedlichen Maßnahmen, die getroffen werden können. Für die Vorhersage von Realdaten macht dies jedoch keinen Sinn, da es in den realen Daten die Anzahl der Neuinfektionen gibt. Das heißt, dass zwar diese Zahl der Infektiösen pro Tag aus dem Verkehr gezogen werden, indem sie in Quarantäne geschickt werden. Im Hintergrund gibt es aber viele Personen, die infektiös sind, jedoch keine Symptome haben. Diese sogenannte *Dunkelziffer* und der zeitliche Verlauf der Gesundheitszustände der *Agents* in dieser Dunkelziffer sind aber das Entscheidende für den Realdatenvergleich in diesem Modell. Deshalb wird dafür ein anderer Ansatz benötigt.

Um dieses zweite Vorgehen umzusetzen, muss ein Zeitpunkt t_0 in den realen Daten gewählt werden, ab dem simuliert werden soll. Bis zu diesem Zeitpunkt ist I_0 die Gesamtsumme an Infektionen in dem gewählten Bezirk. Nun wird die Simulation mit einer gewissen Anzahl an infektiösen *Agents* gestartet. Diese kann aufgrund der Clusteranzahl bestimmt oder auch manuell gewählt werden. Bei diesem Modell ist es einfacher die Anzahl manuell zu wählen, da es keine Daten bezüglich der Cluster auf Bezirksebene gibt und die Zahl I_0 auf Bezirksebene relativ klein ist. Die Simulation läuft jetzt so lange, bis die Zahl der Gesamtinfektionen im Modell größer oder gleich I_0 ist. Nachdem die Simulation gestoppt wurde, haben die *Agents* einen Status und der zeitliche Verlauf der Infektion ist bereits fortgeschritten. Dies kann als Anfangszustand herangezogen werden. Somit wird zum Zeitpunkt t_0 eine Dunkelziffer modelliert, mit welcher im weiteren Verlauf die Simulation durchgeführt wird und welche valide Ergebnisse liefert.

5.3 Resultate

Im folgendem Kapitel werden die Resultate der Modellierung am Beispiel des Bezirks Melk dargelegt. Dabei wird zu Beginn die technische Grundlage für die Simulation dargestellt, danach erste allgemeine Ergebnisse und in weiterer Folge ein Realdatenvergleich präsentiert.

Technische Grundlage

Die gesamte Modellierung und Simulation wird lokal auf einem handelsüblichen *MacBook Air* durchgeführt. Die Initialisierung und die Modellierung wurden in *R* (Version: 4.1.2 (2021-11-01)) und *RStudio* (Version: 2021.09.01) abgewickelt. Der gesamte Code für diese Arbeit ist unter folgendem Link in *GitHub* hinterlegt: https://github.com/karasdan/corona_modeling_austria.

Allgemeine Ergebnisse

Bevor näher auf die Ergebnisse eingegangen wird, muss ein technischer Aspekt betrachtet werden. Das Modell ist aufgrund der Infektionen mittels Wahrscheinlichkeiten sehr stark vom Zufall abhängig. Die Verwendung der *Monte-Carlo Simulation* kann dem entgegenwirken. Hierbei wird dieselbe Simulation mehrmals wiederholt und danach der Mittelwert der Ergebnisse als echtes Ergebnis herangezogen. Doch hier stellt sich die Frage, wie oft die Simulation durchgeführt werden muss, damit das Ergebnis stabil ist. Um diese Frage zu klären, wurde dasselbe Modell mit unterschiedlicher Anzahl an Durchläufen für die *Monte-Carlo Simulation* durchgeführt und die jeweiligen Ergebnisse betrachtet. Wie in Abbildung 11 zu erkennen ist, gibt es kaum Unterschiede zwischen den Änderungen der Durchläufe. Deshalb wurde eine konventionelle Annahme getroffen und es wurde 15 Wiederholungen als optimal gewählt. Das bedeutet, dass alle weiteren Ergebnisse immer mit einer *Monte-Carlo Simulation* von 15 Durchläufen durchgeführt wurden.

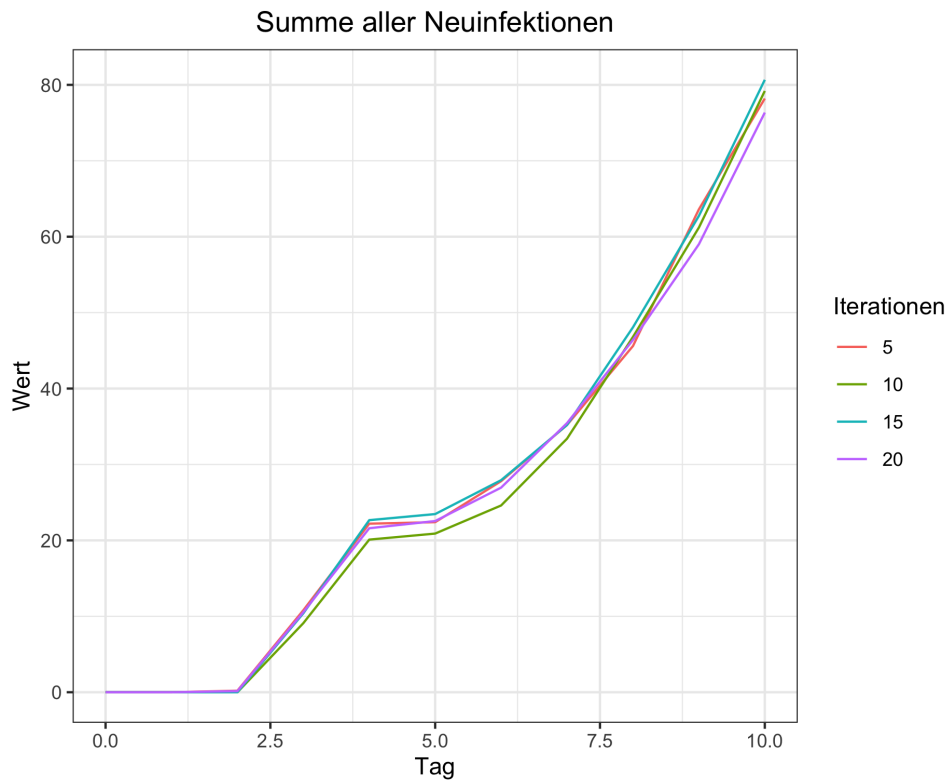


Abbildung 11: Hier sieht man die Vergleiche der Verläufe der einzelnen *Monte-Carlo Simulationen* mit fünf, zehn, 15 und 20 Iterationen bei der Variable der Summe der Neuinfektionen.

(Eigene Darstellung)

Alle nachfolgenden Experimente beruhen auf derselben Agenteninitialisierung. Dabei erhielten die *Agents* ihre Attribute gemäß der zuvor präsentierten Vorgehensweise. Im Anschluss wurden zufällig 100 Einheiten als infektiös markiert, um einen Ausbruch des Virus zu simulieren. Es wurden absichtlich zu Beginn sehr viele *Agents* infiziert, damit bald eine Veränderung zu erkennen ist. Somit konnte die Simulationslaufzeit ein wenig verringert werden. Des Weiteren wurde jedes Experiment auf eine Laufzeit von 50 Tagen beschränkt. Innerhalb dieser 50 Tage wurden fünf verschiedene Zusammensetzungen der Maßnahmen der Regierung ausgewählt und die Ergebnisse verglichen.

- Es gibt keine einzige gesetzte Maßnahme über die gesamte Laufzeit hinweg.
- Zwischen den Tagen 21 und 40 gibt es einen strengen Lockdown. Dabei werden die Freizeitkontakte auf ein Achtel reduziert und alle weiteren haushaltsfremden Kontakte komplett gestrichen. Somit bleiben für jeden *Agent* lediglich die Haushaltskontakte und sehr wenige Freizeitkontakte übrig. Durch diese geringe Anzahl an Freizeitkontakten sollen die Kontakte, die jede Person trotzdem hatte, wie beispielsweise Einkaufs- und Arbeitskontakte, welche nicht durch Homeoffice ersetzt werden konnten, simuliert werden.

- Zwischen den Tagen 21 und 40 gibt es einen „aufgeweichten“ Lockdown, bei dem das Prozedere gleich ist, wie beim zuvor genannten strengen Lockdown, außer dass die Kindergärten offen bleiben. Damit soll die Auswirkung der Kindertageseinrichtungsschließungen validiert werden.
- Als Nächstes wurden nur die Freizeitkontakte zwischen den Tagen 21 und 40 eingeschränkt. Damit soll herausgefunden werden, ob es möglich gewesen wäre, nur durch die Einschränkung der Freizeitkontakte einen Bruch in der Kurve der Anzahl der Neuinfektionen zu bewirken.
- Das letzte Experiment ist das Gegenteil vom vorherigen. Hier wird daher betrachtet, ob ein Bruch in der Kurve der Neuinfektionen möglich gewesen wäre, wenn alles bis auf die Freizeit eingeschränkt worden wäre.

Die Abbildungen der einzelnen Resultate sind immer gleich aufgebaut. Die erste Darstellung zeigt stets die Summe der entdeckten Infektionen, die zweite bildet die erkannten Infektionen, die Infizierten (aber noch nicht infektiös) und die Infektiösen ab und im letzten Graphen wird der Verlauf der täglichen Neuinfektionen betrachtet. Ebenso wurde neben dem Ergebnis von der *Monte-Carlo Simulation* auch ein 95-prozentiges Konfidenzintervall, um die einzelnen Kurven gebildet, um die Schwankungsbreite besser zu erkennen.

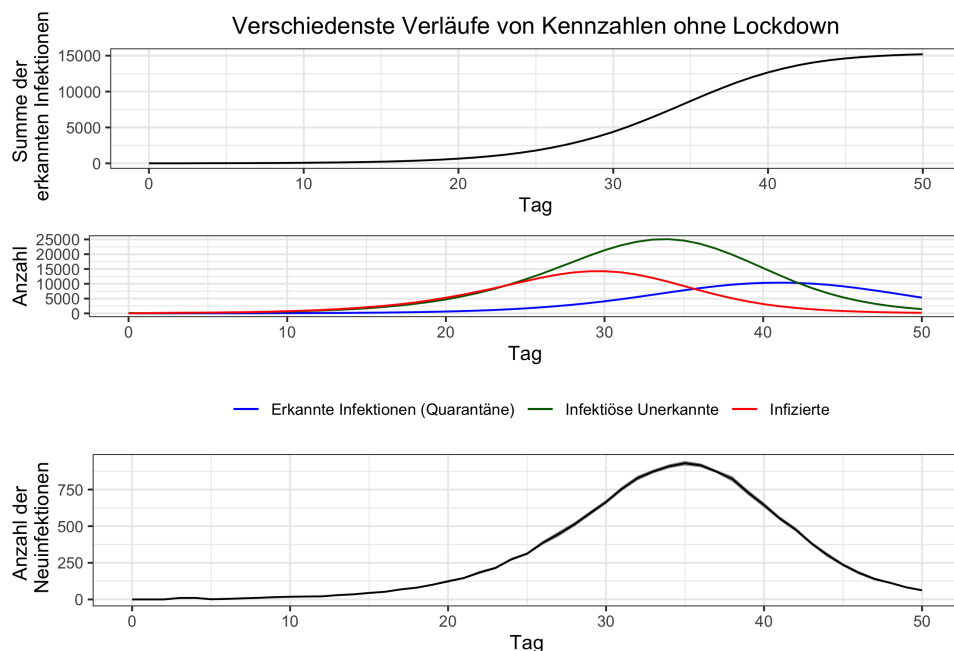


Abbildung 12: Hier kann man die Ergebnisse der Simulation von 50 Tagen mit 100 zufällig gewählten infektiösen *Agents* ohne Lockdown erkennen.
(Eigene Darstellung)

In Abbildung 12 kann der Verlauf einiger wichtiger Kennzahlen erkannt werden, falls kein Lockdown eingesetzt wird. Der obere Plot zeigt die Summe der erkannten Infektionen

in diesem Zeitraum durch dieses Modell. Hier kann man erkennen, dass nach ungefähr 35 Tagen der Höhepunkt erreicht ist und sich danach die Kurve abflacht. Dies spiegelt sich auch in der dritten Darstellung wider, in welcher die täglich erkannten Neuinfektionen gezeigt werden. Es ist eindeutig festzustellen, dass der Höhepunkt mit dem zuvor Erwähnten übereinstimmt. Dabei erreicht die Prognose eine Höhe von über 875 täglichen Neuinfektionen. In der mittleren Darstellung sieht man die einzelnen Phasen einer Infektion. Wie hier deutlich zu erkennen ist, ist die Dunkelziffer, also die unerkannt infektösen *Agents*, zu gewissen Zeiten viel höher als die wirklich identifizierte Anzahl an Infektionen. Ebenso liegt der Höhepunkt der Dunkelziffer ein bis zwei Wochen vor jenem der Anzahl der *Agents*, die sich in Quarantäne befinden.

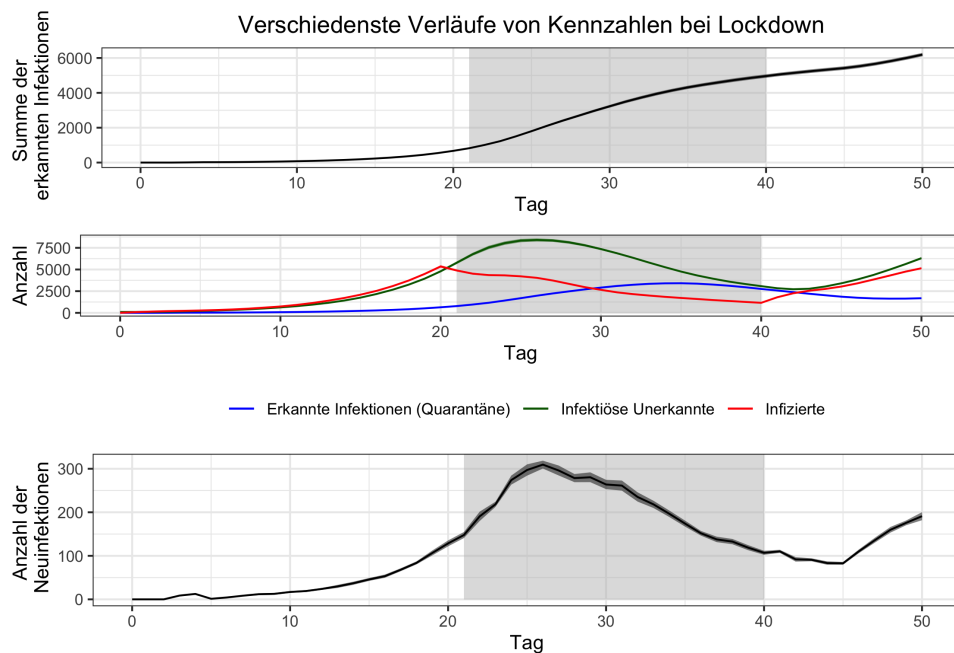


Abbildung 13: Hier kann man die Ergebnisse der Simulation von 50 Tagen mit 100 zufällig gewählten infektiösen *Agents* erkennen, wobei zwischen den Tagen 21 und 40 (grau hinterlegt) ein strenger Lockdown herrscht.
(Eigene Darstellung)

Nun folgen die Ergebnisse der mit einem Lockdown durchgeführten Simulation. Hier sind die Auswirkungen des Lockdowns sehr gut erkennbar. Betrachtet man die letzte Darstellung der Abbildung 13, so sieht man, dass ungefähr eine Woche nach dem Start des Lockdowns der Höhepunkt der Neuinfektionen erreicht ist. Danach fällt die Anzahl dieser rasch ab. Ebenso ist der Maximalwert der Neuinfektionen mit knapp 300 wesentlich geringer als ohne Lockdown. Darüberhinaus sinken die Neuinfektionen nach dem Lockdown noch knapp zehn Tage weiter, bis sie wieder steigen. Wenn nun das zweite Bild betrachtet wird, sticht sofort der Knick in den unerkannten Infizierten (rot) hervor. Dieser erklärt sich dadurch, dass ab dem 21. Tag plötzlich deutlich weniger Kontakte vorhanden sind. Daher infizieren sich weniger Personen innerhalb eines Tages als es Personen gibt, die

infektiös werden.

Vergleicht man die Werte zwischen den ersten zwei Experimenten, dann ist klar zu erkennen, dass ein Lockdown wirkungsvoll ist und somit die Fallzahlen in kurzer Zeit senken kann. Weiters wird deutlich, dass falls nach einem Lockdown keine entsprechenden Vorkehrungen getroffen werden, es sofort wieder zu einem Anstieg kommt und der Vorteil des durchgeführten Lockdowns sehr schnell verspielt ist.

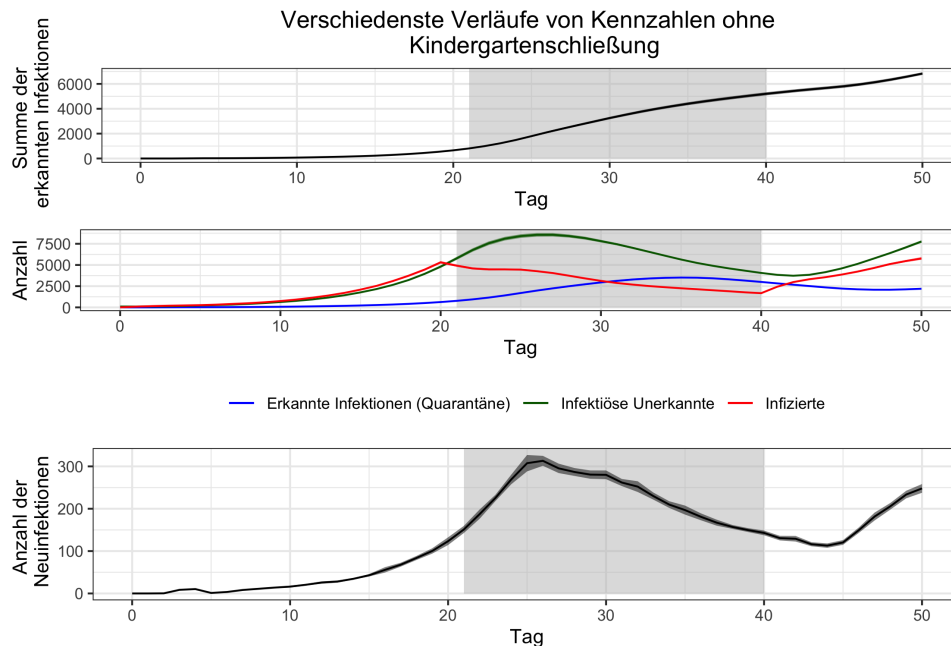


Abbildung 14: Hier kann man die Ergebnisse der Simulation von 50 Tagen mit 100 zufällig gewählten infektiösen *Agents* erkennen, wobei zwischen den Tagen 21 und 40 (grau hinterlegt) ein aufgeweichter Lockdown herrscht.

(Eigene Darstellung)

Nun wird ein weiteres Experiment präsentiert, da sich die Frage stellte, wie groß die Auswirkung einer Kindergartenschließung wäre beziehungsweise ob es überhaupt Sinn macht, diese extra zu berücksichtigen. Die Idee dahinter ist, ob es eventuell möglich gewesen wäre, den Kleinsten ihren Routinealltag während des Lockdowns zu ermöglichen. Eine weiterführende Überlegung wäre ein möglicher Schulbesuch der Volksschulkinder.

Bei der Diskussion dieser Gedanken wird Abbildung 14 betrachtet. Hier ist sehr gut zu erkennen, dass der Lockdown trotz geöffneter Kindergärten wirkungsvoll ist. Nach einer Dauer von sieben Tagen des Lockdowns zeichnet sich wiederum ein negativer Trend in der Kurve der Neuinfektionen ab. Des Weiteren ist diese Richtung auch nach knapp 14 Tagen im Verlauf der Gesamtsumme der Infizierten zu erkennen.

Ein großer Unterschied zur Abbildung 13 ist in der Dimension der einzelnen Verläufe festzustellen. Dies ist nicht überraschend, da es in dieser Konfiguration mehr Kontakte als in der vorherigen gibt. Jedoch ist der Unterschied ausschlaggebend. Vergleicht man die Anzahl der täglichen Neuinfektionen, sieht man, dass zum Ende des Lockdowns knapp 50

Prozent mehr Infektionen erkannt werden, als in Abbildung 13. Ebenso ist die Summe der Infizierten nach 50 Tagen um circa ein Sechstel höher als zuvor. Dieser große Unterschied würde sich in der weiteren Folge der Simulation fortführen. Darüberhinaus ist innerhalb eines Bezirks ein Unterschied von 50 Prozent mehr Neuinfektionen sehr gravierend, weshalb diese Modellierung für eine Kindergartenschließung spricht.

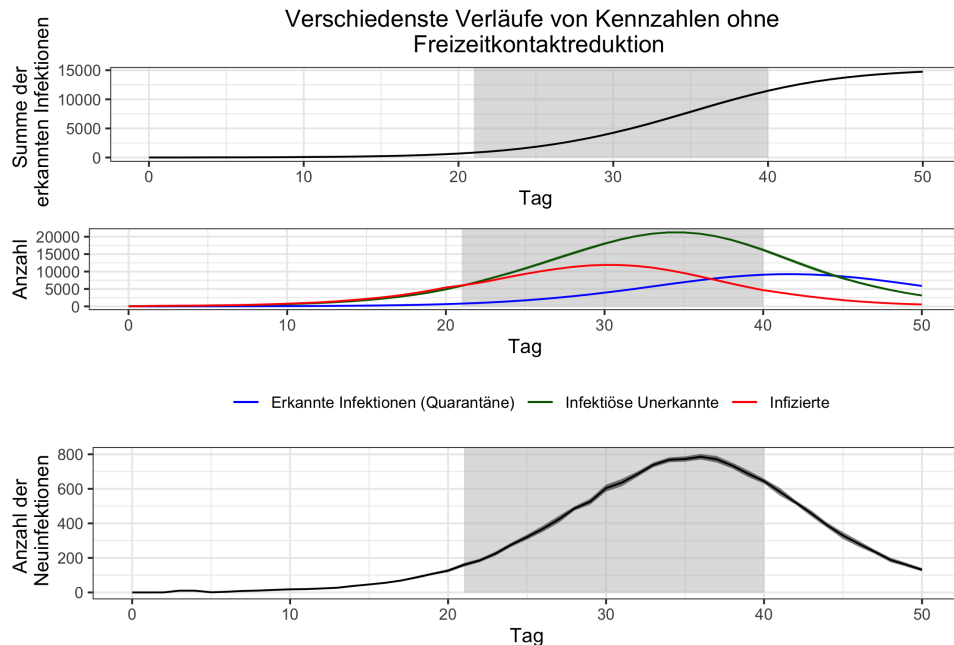


Abbildung 15: Hier kann man die Ergebnisse der Simulation von 50 Tagen mit 100 zufällig gewählten infektiösen *Agents* erkennen, wobei zwischen den Tagen 21 und 40 (grau hinterlegt) alle Regeln herrschten und die Freizeitkontakte nicht eingeschränkt wurden. (Eigene Darstellung)

Ist es möglich nur durch die Schließung von Schulen und Kindergärten sowie einer Homeoffice-Regelung einen Bruch des exponentiellen Anstiegs zu erzeugen? Könnten die Freizeitkontakte aufrecht bleiben, wenn alles andere geschlossen bleibt? Diese Fragen sind klar zu verneinen. Dies ist mithilfe der Abbildung 15 schnell erkennbar. Wenn die Abbildungen 12 (kein Lockdown) und 15 verglichen werden, ist kein nennenswerter Unterschied festzustellen. Der Höhepunkt wird um zwei bis drei Tage nach hinten verschoben und die Maximalwerte werden ein wenig geringer. Jedoch gibt es in Abbildung 15 kein Anzeichen eines negativen Trends wie in den Abbildungen 13 und 14. Das heißt, dass trotz Schulschließungen und Homeoffice-Regelung die Fallzahlen in etwa gleich bleiben und somit keinen großen Einfluss auf das Infektionsgeschehen haben. Der ausschlaggebende Antreiber eines großen Infektionsgeschehens sind die Freizeitkontakte. Dies spiegelt sich eindeutig in der obigen Abbildung wider.

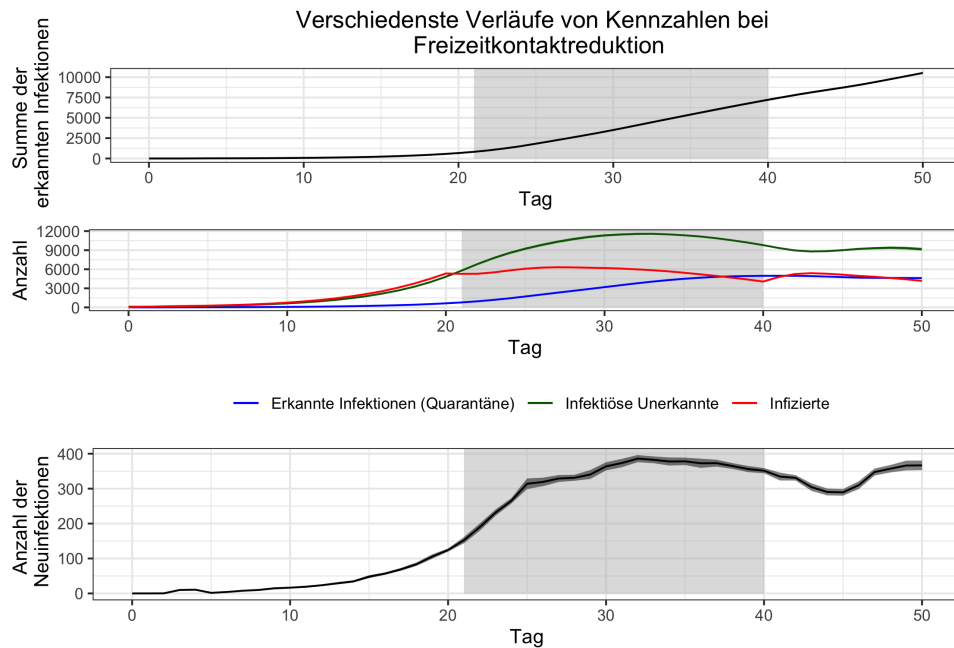


Abbildung 16: Hier kann man die Ergebnisse der Simulation von 50 Tagen mit 100 zufällig gewählten infektiösen *Agents* erkennen, wobei zwischen den Tagen 21 und 40 (grau hinterlegt) nur eine Freizeitkontaktreduktion herrscht.
(Eigene Darstellung)

In der Abbildung 16 werden die Resultate präsentiert, die entstehen, wenn das Gegenteil des vorherigen Experiments durchgeführt wird. Das heißt, dass alles geöffnet bleibt und nur die Freizeitkontakte eingeschränkt werden. Hier wird die Bedeutung der Freizeitkontakte noch besser sichtbar. Nur durch die Reduzierung der Freizeitkontakte wird ein Bruch des exponentiellen Wachstums bewirkt und die Neuinfektionen in der letzten Grafik befinden sich auf einem Tableau, welche zum Ende hin sogar zu fallen beginnen. Des Weiteren werden die Dimensionen der Zahlen wesentlich niedriger gehalten, als bei der Einführung keiner Regeln. Es zeigt, dass dies eventuell eine Möglichkeit wäre, die Fallzahlen auf einem stabilen Niveau zu halten ohne alles zusperren zu müssen.

Zusammenfassend ist nach diesem Unterkapitel klar, dass ein Lockdown das effektivste und schnellste Mittel ist, um die Fallzahlen zu senken. Darüberhinaus ist auch erkennbar, dass die treibende Kraft der Fallzahlen die Freizeitkontakte sind. Wenn diese zurückgeschraubt werden und auf einige Kontakte verzichtet werden, ist es sogar möglich die Fallzahlen auf einem stabilen Niveau zu halten, ohne weitere Schließungen durchführen zu müssen. Ein weiterer Hinweis dazu ist, dass all diese Experimente mit der ersten Variante des *SARS-CoV-2* Virus berechnet wurden, welche nicht so ansteckend war wie die *Delta*- oder *Omikron*-Variante. Diese Ergebnisse werden sich mit anderen Varianten verändern und nicht mehr gültig sein. Ebenso wurde keine verhaltensändernden Maßnahmen wie Hände waschen und Tragen eines Mund-Nasen-Schutzes eingeführt.

Realdatenvergleich

Bevor man schließlich einen Vergleich mit den Realdaten ziehen kann, muss zu guter Letzt ein Betrachtungszeitraum gewählt werden. Dabei ist zu berücksichtigen, dass zu Beginn ein exponentielles Wachstum vorherrschen muss und nach einiger Zeit Maßnahmen gegen den Anstieg ergriffen wurden, die nachmodelliert werden können. Im Rahmen dieser Arbeit liegt weiters der Fokus auf einer Zeitspanne in der Pandemie, in welcher die Maßnahmen, wie das Tragen eines Mund-Nasen-Schutzes, noch nicht so stark zum Einsatz gekommen sind. Es wurde deshalb die erste Phase der Pandemie kurz nach Ausbruch des Virus in Österreich betrachtet. Die Realdaten werden von *Open Data Österreich* heruntergeladen, wie im Anhang A ersichtlich ist. Im Bezirk Melk traten die ersten Fälle am 14. März 2020 auf. Ab dem 16. März 2020 wurde ein strenger Lockdown in Österreich verhängt, um einen Anstieg der Neuinfektionen zu verhindern. Es wurde stets die Gesamtsumme der erkannten Fälle betrachtet, um einen Vergleich mit dem erstellten Modell herzustellen.

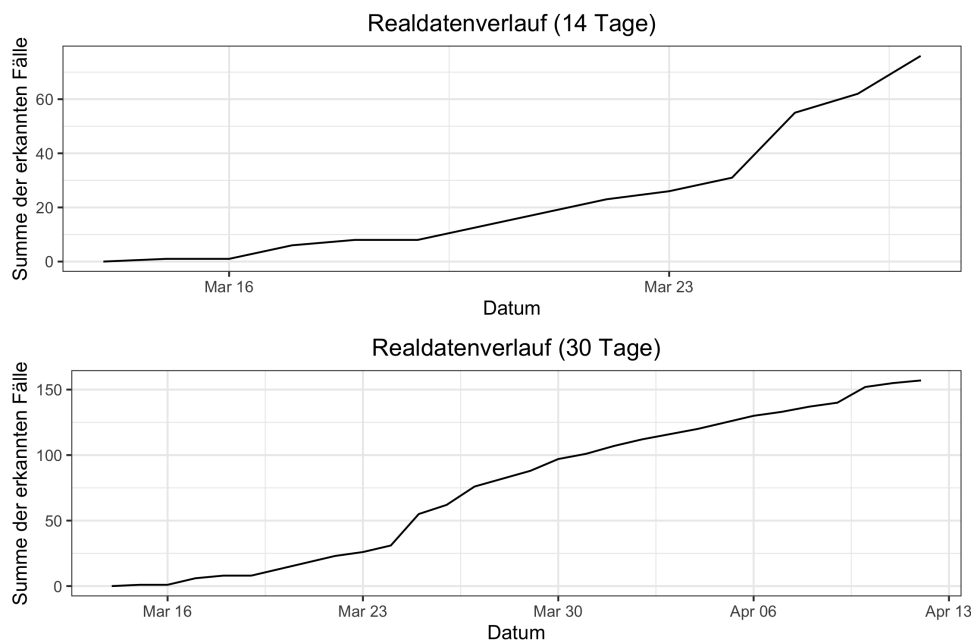


Abbildung 17: Die Abbildung zeigt einerseits den Verlauf in den ersten 14 Tagen und andererseits den Verlauf in den ersten 30 Tagen.
(Eigene Darstellung)

Betrachtet man Abbildung 17, erkennt man im unteren Bild, dass der Lockdown seine Wirkung erfüllt hat. Zu Beginn ist der Anstieg sehr groß und wird dann mithilfe des Lockdowns gebrochen, woraufhin die Kurve abflacht. Die obere Visualisierung stellt die ersten 14 Tage des Infektionsgeschehens im Bezirk Melk dar. Hier ist ein Anstieg gut zu erkennen und es lässt sich vermuten, dass ein exponentieller Anstieg vorherrscht. Um dies zu prüfen wurde mithilfe der Realdaten in den ersten 14 Tagen die Parameter einer Exponentialfunktion geschätzt. Die Ergebnisse werden in Abbildung 18 veranschaulicht.

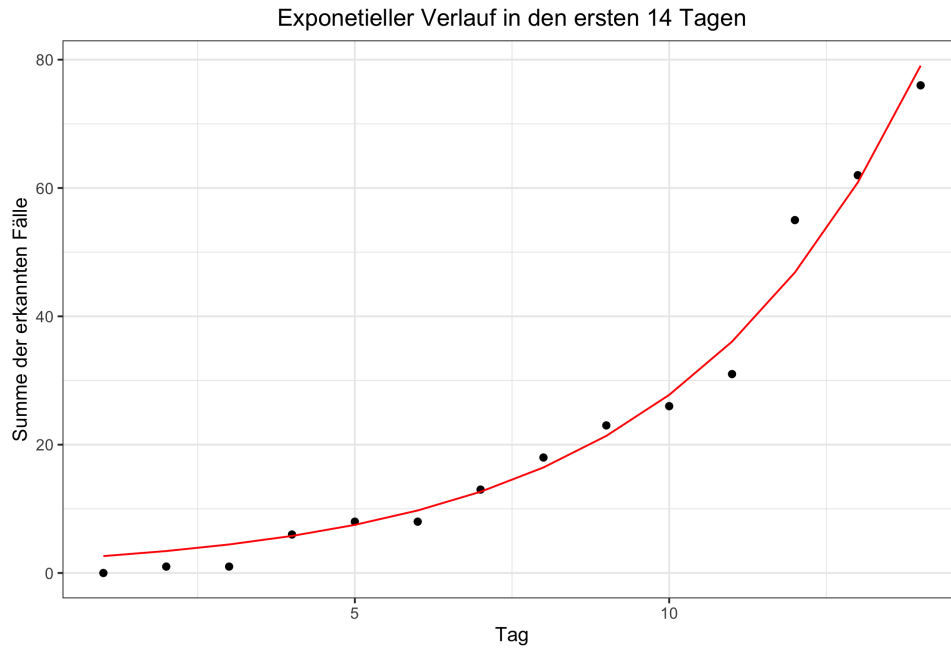


Abbildung 18: In dieser Abbildung ist der geschätzte exponentielle Verlauf (rot) in den ersten 14 Tagen zu erkennen.
(Eigene Darstellung)

In dieser Abbildung ist zu erkennen, dass die Schätzung der Exponentialfunktion nahezu perfekt zu den Realdaten passt. Somit herrscht ein exponentieller Anstieg in diesem Zeitraum vor, welcher in weiterer Folge essentiell für die Prognose Anpassung sein wird. Zusammenfassend kann erwähnt werden, dass der Zeitraum von 14. März 2020 bis 27. März 2020 zum Hyperparametertuning der einzelnen Parameter des Modells und die Tage vom 14. März 2020 bis zum 12. April 2020 für den Vergleich mit dem Modell ideal sind.

Wie schon zuvor erwähnt, müssen zuerst die Parameter des Modells an die Realdaten angepasst werden. Dabei werden zwei wesentliche Parameter betrachtet:

- die Anzahl A_0 jener *Agents*, die zu Beginn infektiös sind
- die Übertragungswahrscheinlichkeit p_{leisure} , die für die Übertragung des Virus bei Kontakten außerhalb des Haushalts zuständig ist

Die Übertragungswahrscheinlichkeit $p_{\text{household}}$ wird in diesem Modell nicht angepasst, weil angenommen wird, dass die Übertragungswahrscheinlichkeit innerhalb des Haushalts von Regierungsmaßnahmen nicht beeinflusst wird und somit gleich bleiben sollte.

Zunächst betrachten wir die Anpassung des Parameters A_0 . Um diesen anzupassen, wurde der zuvor gewählte Zeitraum, der zur Anpassung benötigt wird, auf die ersten zehn Tage (14.3.2020 – 23.3.2020) verkleinert. Es wurde ein kombiniertes Verfahren entwickelt, welches in folgende Schritte unterteilt werden kann:

1. **Schritt:** Wähle einen Suchbereich $\tilde{A} = \{A_1, A_2, \dots, A_n\}$ für die Anzahl A_0 aus.
2. **Schritt:** Führe nun eine Initialisierung der Anfangszustände, wie im vorherigen Kapitel beschrieben, durch. Dabei werden $A_1 \in \tilde{A}$ *Agents* zufällig als infektiös gewählt und keine Maßnahmen modelliert. Aus diesem Schritt entsteht eine Konfiguration der Gesundheitszustände aller *Agents* für den Beginn der Simulation.
3. **Schritt:** Nun beginnt die eigentliche Modellierung. Führe nun die Simulation mit der Ausgangskonfiguration aus Schritt 2 durch. Hier sind die Maßnahmen zu berücksichtigen, die in der Realität gesetzt wurden. Das bedeutet, dass im vorliegenden Fall ab dem zweiten Tag ein strenger Lockdown herrscht und dieser Lockdown bis zum Ende der Simulation anhält, da der 14. März 2020 in der Simulation der Tag Null ist, an dem die Ausgangskonfiguration angepasst wurde. Als Nächstes wird die Simulation mit der Ausgangskonfiguration und den Maßnahmen für neun weitere Tage (Tag Null = 14.3.2020) fünfmal durchgerechnet. Danach muss das Ergebnis validiert werden. Hierbei werden die Verdopplungszeiten der Summe der erkannten Fälle der Realdaten und der simulierten Daten verglichen. Es ist zu berücksichtigen, dass einerseits die Betrachtung der geschätzten Verdopplungszeit über den gesamten Beobachtungszeitraum nicht ausreicht und andererseits auch die ersten Tage eine große Unschärfe aufweisen, da der Wert stark davon abhängt, wie viele Neuinfektionen zu Beginn erkannt wurden. Um diese beiden Punkte zu beachten, wurde mit dem Vergleich erst ab dem vierten Tag begonnen und immer nur über vier Tagesschritten berechnet. Das bedeutet, dass ab dem vierten Tag immer die durchschnittliche Verdopplungszeit der nächsten vier Tage verglichen wurde, danach um einen Tag erhöht und wieder die nächsten vier Tage betrachtet wurden. Anschließend wird die Summe der quadratischen Abweichungen S_1 als Metrik für den Vergleich berechnet.
4. **Schritt:** Wiederhole Schritt 2 und Schritt 3 für alle $A_i \in \tilde{A}$. Somit entsteht die Menge $\tilde{S} = \{S_1, S_2, \dots, S_n\}$.
5. **Schritt:** Wähle nun $A_0 = A_{\tilde{i}}$, wobei $S_{\tilde{i}} = \min(\tilde{S})$.

Mithilfe dieses kurzen Algorithmus erhält man die optimale Anzahl an infektiösen *Agents* zu Beginn der Simulation und die optimale Anfangskonfiguration der Gesundheitszustände der *Agents*. Hier wurden im Gegensatz zu den anderen Simulationen nur fünf Durchläufe der *Monte-Carlo Simulation* durchgeführt, weil die Menge $\tilde{A}$ sehr groß sein kann und somit der Rechenaufwand reduziert werden muss, damit die Ergebnisse in einer vernünftigen Zeit verfügbar sind.

Der Parameter $p_{leisure}$ wird mit einer ähnlichen Vorgangsweise angepasst. Es gibt wiederum einen Suchraum $\tilde{P}$ über den der optimale Parameter gesucht wird. Mit dem zuvor

gefundenen Parameter A_0 und der Anfangskonfiguration werden wieder zehn Tage simuliert und mit den Realdaten auf die gleiche Art und Weise verglichen. Dabei entsteht abermals für jede Parametervariante eine Summe. All diese Summen werden verglichen und es wird wiederum das Minimum gewählt. Jener Parameter, der das Minimum der Summen erreicht, wird als $p_{leisure}$ bezeichnet.

In unserem Beispiel resultierten die beiden Algorithmen in folgender Parameterzusammensetzung:

$$A_0 = 42 \quad p_{leisure} = 6,5$$

Die Experimente und Vergleiche wurden nach der Parameteranpassung durchgeführt. Dabei wurde die *Monte-Carlo Simulation* wieder mit 15 Durchläufe durchgeführt und die Anfangskonfiguration der Gesundheitszustände sowie die zuvor berechneten Parameter verwendet.

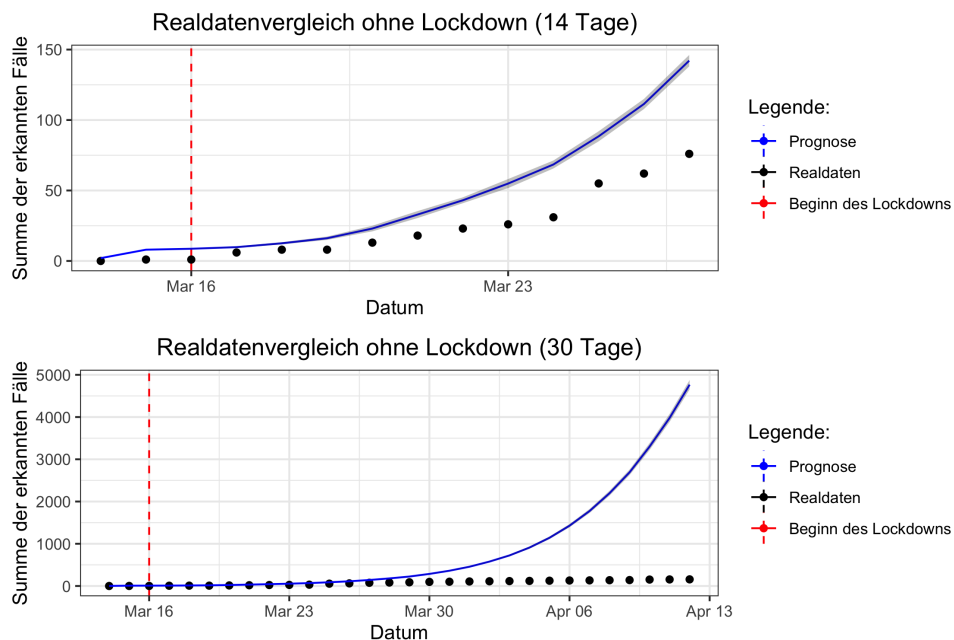


Abbildung 19: Hier sieht man den Vergleich der Realdaten mit der Prognose, wenn kein Lockdown simuliert wird. In der oberen Abbildung wurden 14 Tage und in der unteren 30 Tage simuliert.

(Eigene Darstellung)

In Abbildung 19 sieht man die Ergebnisse der zuvor erhaltenen Parameterzusammensetzung, wenn kein Lockdown in der Prognose herrscht. Wie im oberen Graphen zu erkennen ist, ist die Prognose in den ersten sieben Tagen noch sehr ähnlich zu den Realdaten. Danach wird der Unterschied zwischen den Realdaten und der Prognose immer größer. Betrachtet man den unteren Graphen, so kann man die verheerenden Auswirkungen sehen, falls kein Lockdown eingeführt worden wäre. Knapp 5 000 Menschen wären innerhalb eines Monats im Bezirk Melk positiv auf das *SARS-CoV-2* Virus getestet worden. Dies wären um ein Vielfaches mehr als mit einem Lockdown gewesen und hätte unser

Gesundheitssystem mit hoher Wahrscheinlichkeit zum Zusammenbruch gebracht.

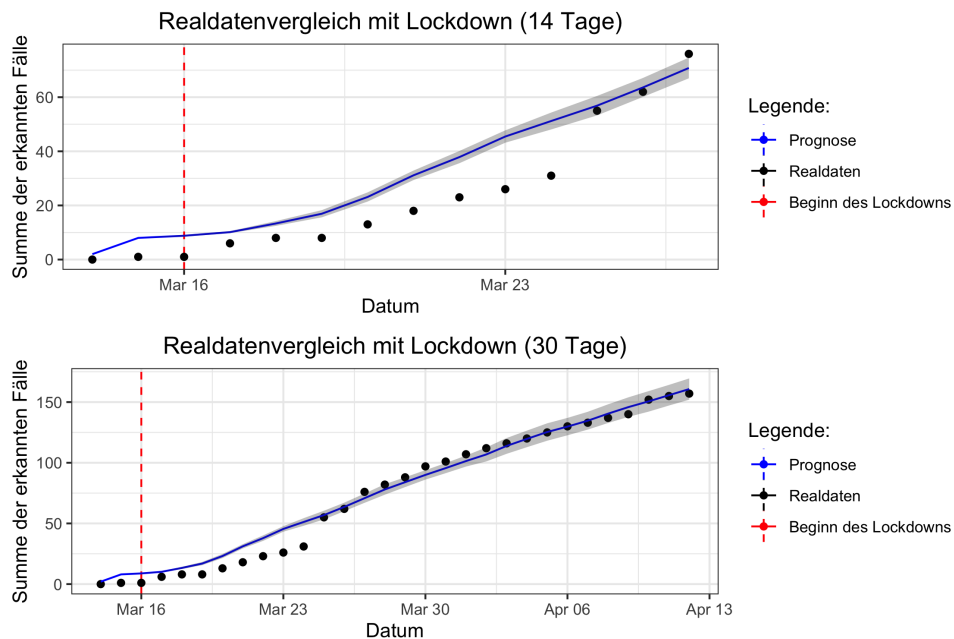


Abbildung 20: Hier sieht man den Vergleich der Realdaten mit der Prognose, wenn ein Lockdown simuliert wird. In der oberen Abbildung wurden 14 Tage und in der unteren 30 Tage simuliert.

(Eigene Darstellung)

Als Nächstes wurde eine Simulation mit Lockdown durchgeführt. Wie in Abbildung 20 zu erkennen ist, passt sich die Prognose sehr gut an die realen Daten an. Sowohl bei zwei als auch bei vier Wochen Simulation ist die Prognose nahezu perfekt. Zu Beginn ist die Prognose etwas zu hoch. Ein Grund dafür könnte eine zu hohe Anfangskonfiguration sein. Im Laufe der Simulation pendelt sich die Prognose jedoch gut ein und bleibt sehr nahe an den Realdaten. Vergleicht man die Abbildungen 19 und 20, so wird der enorme Unterschied zwischen Lockdown und keinem Lockdown sichtbar. Wenn ein Lockdown verhängt wird, steigt die Summe der erkannten Fälle nur auf circa 150 Personen, wohingegen in der ersten Abbildung die Zahl auf knapp 5 000 *Agents* klettert.

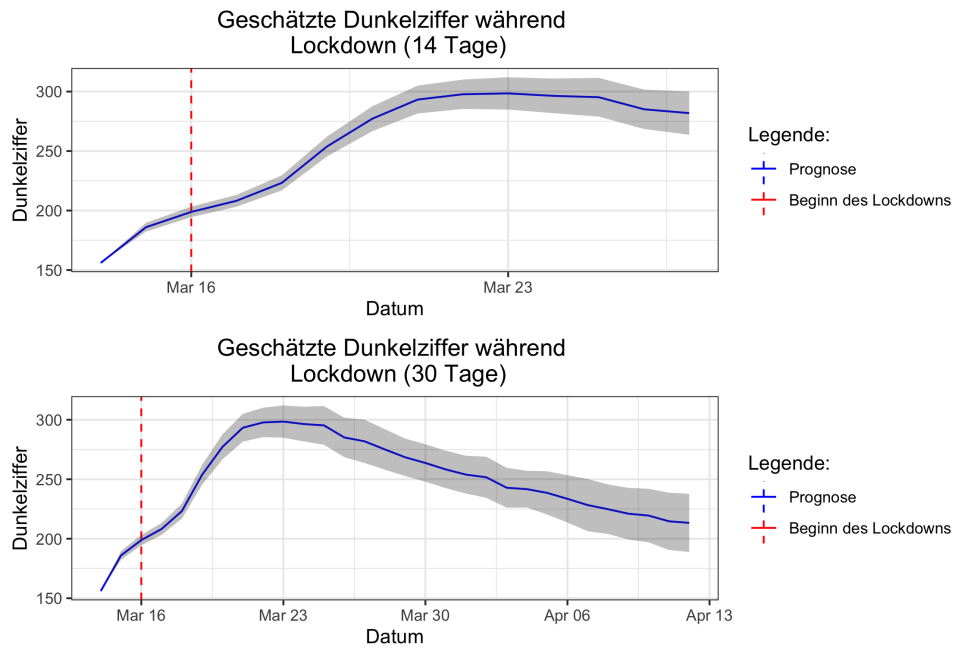


Abbildung 21: Durch das Modell geschätzte Dunkelziffer - Verlauf während Lockdowns (Eigene Darstellung)

Um die Vorteile des Modells und generell der *Agent Based Models* im Zusammenhang mit Infektionskrankheiten zu zeigen, betrachten wir in Abbildung 21 die geschätzte Dunkelziffer für den Lockdown. Aus den Realdaten und den Ergebnissen anderer Modelle wäre dies nicht möglich. Wie man hier gut erkennen kann, liegt der Höhepunkt der Dunkelziffer ungefähr eine Woche nach Beginn des Lockdowns, als knapp 300 Personen infiziert sind. Wenn dieser Wert mit den erkannten Fällen verglichen wird, sieht man, dass nur knapp 30 *Agents* erkannt wurden. Das heißt, dass am Höhepunkt die Anzahl der wirklich infizierten Personen um ein Zehnfaches höher war.

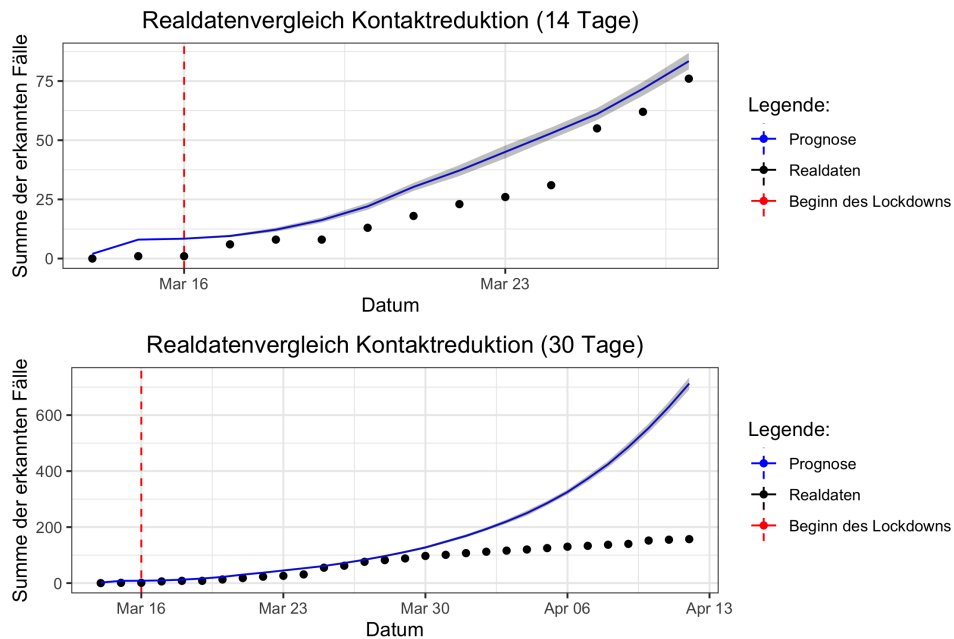


Abbildung 22: Hier sieht man den Vergleich der Realdaten zur Prognose, wenn eine Kontaktreduktion simuliert wird. In der oberen Abbildung wurden 14 Tage und in der unteren 30 Tage simuliert.
(Eigene Darstellung)

Abschließend wurde ein Experiment durchgeführt, welches bereits in den allgemeinen Resultaten durchaus interessant wirkte. Statt eines strengen Lockdowns wurde lediglich eine Kontaktreduktion eingeführt. Die zugehörigen Ergebnisse sind in Abbildung 22 veranschaulicht. Hier kann man erkennen, dass die Simulation für 14 Tage noch sehr vielversprechend ist. Die Prognose folgt den Realdaten und wirkt bei erster Betrachtung sogar besser als jene, bei der ein Lockdown eingeführt wurde. Nach den ersten 14 Tagen steigen die Zahlen jedoch enorm in die Höhe und nähern sich in großen Schritten den Zahlen, bei denen kein Lockdown eingeführt wurde. Die Gesamtsumme ist zwar nach einem Monat noch um ein Vielfaches geringer als jene ohne Lockdown, jedoch ist der Anstieg sehr groß und die Tendenz geht in Richtung der Zahlen in Abbildung 19. Somit lässt sich schlussfolgern, dass dies keine Option gewesen wäre, da ebenso diese Zahlen für einen einzelnen Bezirk viel zu hoch gewesen wären, um diese zu bewältigen.

Zusammenfassend ist festzustellen, dass dieses Modell nach dem Hyperparametertuning sehr gut zur Prognose der Summe der erkannten Fälle passt, was deutlich aus den Abbildungen hervorgeht. Darüberhinaus ist zu erwähnen, dass sich andere Maßnahmen, wie eine reine Kontaktreduktion, zu diesem frühen Zeitpunkt weniger gebracht hätten als ein strenger Lockdown.

6 Resümee

Das Jahr 2020 stellte sowohl Politiker:innen als auch Wissenschaftler:innen vor noch nie dargewesene Herausforderungen. Das *SARS-CoV-2* Virus und die damit verbundene Pandemie sind etwas völlig Neues für die gesamte Menschheit. Dadurch ergab sich die Gelegenheit zu raschen Veränderungen, welche nicht nur die politische und soziale Landschaft sondern auch die Wissenschaft betrafen. Diese Entwicklung ist in der Simulationsforschung beziehungsweise der Mathematik klar erkennbar.

Sogenannte *Agent Based Models*, welche seit langer Zeit existieren, sind durch die Pandemie populär geworden. Dafür gibt es mehrere Gründe. Zunächst steigt seit Jahren die Rechenleistung der Computer. Somit können kompliziertere Modelle in kürzerer Zeit erstellt und berechnet werden. Dies ist ein wesentlicher Punkt für ABMs. Vor 15 bis 20 Jahren wäre es unvorstellbar gewesen, einen ganzen Staat in einem *Agent Based Model* zu simulieren. Weiters ist der verständliche und intuitive Aufbau eines ABMs entscheidend, weshalb diese Methode auch zur Modellierung des *SARS-CoV-2* Virus herangezogen wurde. Die Akzeptanz der Bevölkerung, welche sehr eng mit dem Verständnis verbunden ist, ist entscheidend für die Ergebnisse. Somit haben ABMs einen Vorteil gegenüber des klassischen Ansatzes der Differentialgleichungen.

Während der Pandemie entstand eine Vielzahl an *Agent Based Models*, welche im Aufbau sehr unterschiedlich sind. Aber alle haben eines gemeinsam: Die Ergebnisse schneiden beim Realdatenvergleich sehr gut ab. Dies gilt auch für die Modelle, welche in dieser Arbeit präsentiert wurden. Der große Nachteil dieser liegt aber im Detail. Oftmals werden Daten verwendet, welche nicht öffentlich zugänglich sind und sehr detailliert sind. Dies benötigt einerseits die Berechtigung zum Zugriff und andererseits eine gewisse Rechenleistung, die ein herkömmlicher Computer nicht besitzt. Darüberhinaus reichen die Interaktionen dieser Modelle von einfachen Kontakten zwischen *Agents* bis hin zur Abbildung jeglicher Aktionen, die ein Individuum tätigen kann, um andere zu beeinflussen. All diese Argumente bewirken, dass die Modelle trotz ihres eigentlich intuitiven Aufbaus sehr kompliziert zu verstehen sind.

Hier setzt das selbstentwickelte *Agent Based Model* an. Dieses kommt mit einfachen Daten aus, welche für alle zugänglich sind. Der Aufbau, die Interaktionen und die Umgebung sind vereinfacht und ermöglichen die Erstellung dieses Modells auf einem handelsüblichen Rechner. Trotz dieser Einschränkungen und Vereinfachungen ist der Realdatenvergleich nach der Anpassung der Parameter sehr zufriedenstellend. Die Nachteile, welche aufgrund der Beschränkung auf einen Bezirk und der Nichtberücksichtigung der Hospitalisierungen entstehen, lassen sich durch einfache Anpassungen des Modells entfernen. Das Modell ist beliebig auf alle Bezirke oder auf ganz Österreich erweiterbar, da die Datengrundlage für alle Bezirke gleich und außerdem vorhanden ist. Um die Hospitalisierungen einzubauen sind kleine Veränderungen in der Logik des Krankheitsverlaufs vorzunehmen.

Des Weiteren könnte eine Erweiterung, welche in dieser Arbeit aufgrund des Umfangs und der Rechenleistung nicht durchführbar war, vorgenommen werden. Vor dem selbstentwickelten Modell, welches auf alle Bezirke Österreichs anwendbar ist, kann ein weiteres *Agent Based Model* entwickelt werden, das als *Agents* die einzelnen Bezirke in Österreich hat und somit Interaktionen zwischen diesen zulässt und eine gewisse Abhängigkeit dieser untereinander abbilden kann. Der Vorteil von diesem Ansatz wäre, dass die Verkehrskontrollen, welche zu einer gewissen Zeit in der Pandemie in Österreich ab einer definierten Inzidenz an den Bezirksgrenzen durchgeführt wurden, in das vorgelagerte Modell implementiert werden könnten.

Weiters ist anzumerken, dass alle Ergebnisse, die in dieser Arbeit präsentiert werden, auf der ersten Variante des Coronavirus basieren. Somit können die Resultate bei anderen Varianten, wie beispielsweise der ansteckenderen Omikron Variante, anders aussehen. Zu berücksichtigen ist aber, dass das Modell auch auf andere Varianten angewendet werden kann. Um dies umzusetzen, ist es lediglich notwendig, die Parameter des Modells in einem Zeitraum, in dem die zu untersuchende Variante vorherrschend war, nach der beschriebenen Methode zu schätzen. Danach läuft das Modell genauso wie vorher durch und es ist möglich Vergleiche zwischen den einzelnen Varianten zu ziehen.

Abschließend hat sich durch diese Pandemie in allen Bereichen des Lebens einiges verändert. Davon blieben auch die Methoden der Simulationen nicht verschont. Das Verlangen nach Erklärbarkeit und die mediale Präsenz führten dazu, dass klassische Modelle, welche seit Jahrzehnten unangefochten waren, plötzlich Konkurrenz von wesentlich einfacheren und moderneren Methoden bekommen haben. Dabei ist immer zu beachten, dass sowohl das eine als auch das andere seine Vor- und Nachteile hat und die Einschränkung auf eine Methodik die Fähigkeit zur Problemlösung erheblich begrenzt.

Literaturverzeichnis

- Bicher, Martin, Claire Rippinger, Dominik Brunmeir, Christoph Urach & Niki Popper (2020): *Agent-Based COVID-19 Simulation Model*, [online] https://www.dwh.at/projects/covid-19/Covid19_Model_v20200929.pdf [abgerufen am 14.12.2021].
- Bicher, Martin, Christoph Urach & Niki Popper (2018): “GEPOC ABM: A GENERIC AGENT-BASED POPULATION MODEL FOR AUSTRIA”, in: *2018 Winter Simulation Conference (WSC)*, S. 2656–2667, DOI: 10.1109/WSC.2018.8632170.
- Bonabeau, Eric (2002): “Agent-based modeling: Methods and techniques for simulating human systems”, in: *Proceedings of the National Academy of Sciences*, S. 7280–7287, DOI: 10.1073/pnas.082080899.
- De Marchi, Scott & Scott E. Page (2014): “Agent-Based Models”, in: *Annual Review of Political Science*, S. 1–20, DOI: 10.1146/annurev-polisci-080812-191558.
- Dittler, Ullrich & Christian Kreidl (2021): “Eine kurze Chronologie der Covid-19-Pandemie im Frühjahr 2020”, in: *Wie Corona die Hochschullehre verändert: Erfahrungen und Gedanken aus der Krise zum zukünftigen Einsatz von eLearning*, Springer Fachmedien Wiesbaden, S. 1–13, DOI: 10.1007/978-3-658-32609-8_1.
- Hemmer, Christoph J, Hilte F Geerdes-Fenge & Emil C Reisinger (2020): “COVID-19: epidemiology and clinical facts”, in: *Der Radiologe*, S. 893–898, DOI: 10.1007/s00117-020-00741-y.
- Johns Hopkins University, (JHU) (2022): *COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)*, [online] <https://coronavirus.jhu.edu/map.html> [abgerufen am 25.04.2022].
- Macal, C. M. (2016): “Everything you need to know about agent-based modelling and simulation”, in: *Journal of Simulation*, S. 144–156, DOI: 10.1057/jos.2016.7.
- Macal, C. M. & M. J. North (2010): “Tutorial on agent-based modelling and simulation”, in: *Journal of Simulation*, S. 151–162.
- Nishiura, Hiroshi, Natalie M. Linton & Andrei R. Akhmetzhanov (2020): “Serial interval of novel coronavirus (COVID-19) infections”, in: *International Journal of Infectious Diseases*, S. 284–286, DOI: <https://doi.org/10.1016/j.ijid.2020.02.060>.
- Park, Minah, Alex R. Cook, Jue Tao Lim, Yinxiaohe Sun & Borame L. Dickens (2020): “A Systematic Review of COVID-19 Epidemiology Based on Current Evidence”, in: *Journal of Clinical Medicine*, DOI: 10.3390/jcm9040967.

- Rabadan, Raul (2020): *Das Coronavirus verstehen*, 1. Aufl., Berlin Heidelberg, Deutschland: Springer, ISBN: 9783662624296.
- Shamil, Md. Salman, Farhanaz Farheen, Nabil Ibte haz, Irtesam Mahmud Khan & M. Sohel Rahman (2021): “An Agent-Based Modeling of COVID-19: Validation, Analysis, and Recommendations”, in: *Cognitive computation*, S. 1–12, ISSN: 1866-9956.
- Wallentin, Gudrun, Dana Kaziyeva & Eva Reibersdorfer-Adelsberger (2020): “Covid-19 intervention scenarios for a long-term disease management”, in: *International journal of health policy and management*, S. 508–516, ISSN: 2322-5939.
- Wilensky, U. & W. Rand (2015): *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*, MIT Press, [online] <https://books.google.at/books?id=LQrhBwAAQBAJ>.

Abbildungsverzeichnis

1	Zusammenfassung der Definitionen für <i>Agent Based Models</i>	10
2	Aufbau der Zustände in einem <i>Agent Based Model</i>	21
3	Aufbau eines Tagesverlaufs im Modell „COVID-19 ABM“	28
4	Kontaktnetzwerk im Modell für ganz Österreich	31
5	Gesundheitsstatus Weg im Modell für Österreich	33
6	Geografische Daten von Bezirksgrenzen und Straßennetz	39
7	Zustandsverlauf vom Modell für den Bezirk Melk	41
8	Ausschnitt der Pendelwahrscheinlichkeitsmatrix	44
9	Datensatz der <i>Agents</i> nach der Agentsinitialisierung am Beispiel des Bezirks Melk	46
10	Ausschnitt aus entstehender Kontakttable für einen Tag	47
11	Test für die Monte Carlo Simulation	50
12	Ergebnisse des Modells ohne Lockdown	51
13	Ergebnisse des Modells mit Lockdown	52
14	Ergebnisse des Modells mit aufgeweichtem Lockdown	53
15	Ergebnisse des Modells ohne Freizeitkontaktreduktion	54
16	Ergebnisse des Modells nur mit Freizeitkontaktreduktion	55
17	Verlauf der Realdaten in der ersten Phase der Pandemie	56
18	Exponentieller Verlauf in den ersten 14 Tagen	57
19	Realdatenvergleich ohne Lockdown	59
20	Realdatenvergleich mit Lockdown	60
21	Dunkelziffer des Realdatenvergleichs	61
22	Realdatenvergleich mit Kontaktreduktion	62

A Tabellen

Herkunft:	Name:	Zugriff:	URL:
Statistik Austria	Bevölkerung 2019 nach Altersgruppen, Geschlecht und Bezirken bzw. NUTS 3-Regionen	15.12.2021	https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/volkszaehlungen_registerzaehlungen_abgestimmte_erwerbsstatistik/bevoelkerung_nach_demographischen_merkmalen/index.html
Statistik Austria	Bevölkerung 2019 nach Alter in Einzeljahren, Geschlecht und Bundesland	15.12.2021	
Statistik Austria	Privathaushalte nach Haushaltgröße, Bundesländern und Alter der Haushaltsreferenzperson - Jahresdurchschnitt 2021	29.12.2021	https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/haushalte_familien_lebensformen/haushalte/index.html
Statistik Austria	Schulen 2020/21 nach Gemeinden und Schultypen	4.1.2022	https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bildung/schulen/schulbesuch/index.html
Statistik Austria	Schülerinnen und Schüler 2020/21 nach Gemeinden und Schultypen	4.1.2022	
Statistik Austria	Kindertageshime, Gruppen und Kinder nach Gemeinden 2020	4.1.2022	https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bildung/kindertagesheime_kinderbetreuung/index.html
Statistik Austria	Gemeindeergebnisse der Abgestimmten Erwerbsstatistik und Arbeitsstättenzählung 2019 (Gebietsstand 2019)	5.1.2022	https://www.statistik.at/web_de/statistiken/wirtschaft/unternehmen_arbeitsstaetten/arbeitsstaetten_ab_az_2011/index.html

Tabelle 1: Alle verwendeten Datensätze mit Zugriffsdatum, URL und Herkunft.

Herkunft:	Name:	Zugriff:	URL:
Open Data Österreich	CovidFaele_Timeline_GKZ	4.2.2022	https://www.data.gv.at/katalog/dataset/covid-19-zeitliche-darstellung-von-daten-zu-covid19-fallen-je-bezirk/resource/9eb08d45-ff99-40f1-90cd-7b3659b0bc8d?view_id=f79ec160-714c-4a3b-984f-bbc59c70ac89
OpenStreet Map	Straßennetz und Grenzen	mittels API	https://www.openstreetmap.org/#map=7/47.714/13.349

Tabelle 1: Alle verwendeten Datensätze mit Zugriffsdatum, URL und Herkunft.

Variable	Verteilung	Herkunft
Inkubationszeit	Skalierte β -Verteilung mit Minimum zwei Tagen, Maximum 14 Tagen und Erwartungswert 5,1 Tagen.	Bicher, Rippinger et al. (2020)
Latenzzeit	Immer Inkubationszeit minus zwei Tage	
Reaktionszeit	$X \sim Weibull(4,29; 1,65)$	
Erholungszeit erkannt	$X \sim Tri(11; 14; 16)$	
Erholungszeit unerkannt	$X \sim Tri(1; 5; 7)$	
Testwahrscheinlichkeit	(3,0; 9,0; 22,0; 20,0; 24,0; 28,0; 20,0; 21,0; 33,0; 58,0) Prozent pro Zehnjahres Gruppen	

Tabelle 2: Zeitenvariablen mit deren Verteilungen und Herkunft

Variable	Verteilung	Herkunft
Wahrscheinlichkeit einer Infektion	$p_{household} = 0,25; p_{work} = p_{school} = p_{kindergarden} = p_{leisure} = 0,05$	Bicher, Rippinger et al. (ebd.)
Freizeitkontakte pro Tag	$X \sim \Gamma(6,11; 1,0)$	
Arbeitsplatzkontakte pro Tag	$X \sim \Gamma(5,28; 1,0)$	
Schulkontakte pro Tag (wird auch für Kindergartenkontakte verwendet)	$X \sim \Gamma(4,64; 1,0)$	

Tabelle 3: Weitere Variablen mit Verteilung und Herkunft.

Variable	Verteilung	Herkunft
Verteilung der Pendler	$[p_{Gemeinde}; p_{Bezirk}; p_{ausserhalb}]$ pro Bundesland: Bgld: [0,26; 0,21; 0,53] Ktn: [0,46; 0,16; 0,38] NÖ: [0,29; 0,2; 0,51] OÖ: [0,35; 0,24; 0,41] Sbg: [0,46; 0,21; 0,33] Stmk: [0,42; 0,23; 0,35] T: [0,41; 0,3; 0,29] Vbg: [0,37; 0,35; 0,28] W: [0,01; 0,04; 0,95]	Bicher, Rippinger et al. (2020)

Tabelle 3: Weitere Variablen mit Verteilung und Herkunft.