

Verbesserung der R-Wert-Berechnung

Peter Möller und Harry Drewes

25.2.2021

Inhalt

Einführung

Verbesserungen der R-Wert-Berechnung

1. Verschiedene Modelle im Vergleich
2. Der R-Wert einfach berechnet
3. Fehlalarm-Analyse
4. Fehler in Abhängigkeit vom R-Wert
5. Zusammenfassung

Anhang

1. Waren die Maßnahmen gegen die erste Welle wirksam?
2. Fehlerberechnung
3. Modellbeschreibung
4. Das arithmetische und das geometrische Mittel
5. Fehlalarmanalyse
6. Prognose der Neuinfizierten und Todesfälle

Zusammenfassung

Literatur und Quelle der Daten

Einführung

Mit Unterstützung der HAW Hamburg haben wir ein Modell entwickelt und weiter optimiert, um den R-Wert noch präziser zu berechnen. Je genauer man den R-Wert berechnen kann, desto besser können wir untersuchen, was die Maßnahmen gebracht haben. Für die erste Welle haben wir gezeigt, dass die Maßnahmen im März die exponentielle Ausbreitung des Virus gestoppt haben (siehe Anhang 1). Die gleiche Analyse haben wir für den zweiten Lockdown gemacht. Die R-Wert-Berechnung des Robert Koch-Instituts (RKI) konnten wir verbessern. Ersetzt man das arithmetische Mittel im RKI-Modell durch das geometrische Mittel, wird das Infektionsgeschehen optimal beschrieben.

Die Reproduktionszahl R einfach erklärt

Die Reproduktionszahl R gibt im Durchschnitt an, wie viele andere Menschen ein Infizierter in einem bestimmten Zeitraum, z.B. innerhalb von 4 Tagen, ansteckt. Ist die Zahl größer als 1, dann nimmt die Anzahl der Neuinfizierten zu. Bei $R=2$ verdoppelt sich die Anzahl der Neuinfizierten pro Tag dann alle 4 Tage (durchgezogene Linie in Abbildung 1). Die Punkte geben die erste und die zweite Verdoppelung von 1.000 auf 2.000 und danach auf 4.000 nach jeweils 4 Tagen an. Ist $R=1$, dann bleibt die Anzahl der Neuinfizierten gleich (gepunktete Linie). Ist R kleiner als 1, dann nimmt die Anzahl der Neuinfizierten ab (gestrichene Linie). Bei $R=0,5$ halbiert sich die Anzahl der Neuinfizierten pro Tag alle 4 Tage.

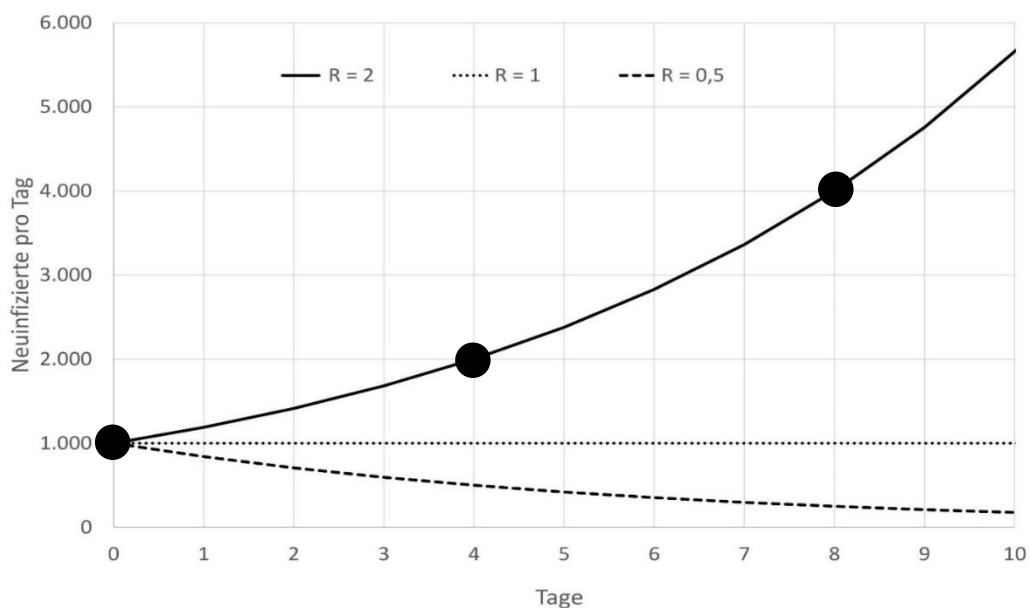


Abbildung 1. Neuinfizierte pro Tag für $R=2$, $R=1$ und $R=0,5$.

Verbesserungen der R-Wert-Berechnung

1. Verschiedene Modelle im Vergleich

Um die verschiedenen R-Wert-Berechnungen miteinander zu vergleichen, verwenden wir eine Testfunktion, welche die erste Welle widerspiegelt (siehe Anhang 2). Unsere Testfunktion bzw. R-Wert-Funktion gibt den zeitlichen Verlauf der „wahren“ R-Werte an. Aus der R-Wert-Funktion werden die Neuinfektionszahlen pro Tag berechnet. Diese Neuinfektionszahlen werden dann mit einer realistischen Wochentagschwankung (systematische Abweichungen) und mit Rauschen (zufällige Abweichungen) überlagert.

Anschließend werden die R-Werte mit den zu vergleichenden Methoden bestimmt und die mittleren Abweichungen zum „wahren“ R-Wert berechnet.

Das Ergebnis war für uns überraschend. Das RKI gibt zwei verschiedene R-Werte an, nämlich einen „sensitiven“ und einen „geglätteten“ Wert. Unser Modell nennen wir Drewes-Möller-Modell (DM-Modell). Das DM-Modell wird ausführlich im Anhang 3 beschrieben. In der Tabelle 1.1 ist der jeweilige Fehler der R-Wert-Berechnung angegeben. In der Abbildung 1.1 haben wir zunächst „RKI-sensitiv“ mit dem „DM-Modell“ verglichen. Man sieht „RKI-sensitiv“ schwankt sehr stark und der Fehler ist mit 22 Prozent sehr hoch. „RKI-geglättet“ ist schon wesentlich besser (siehe Abbildung 1.2), aber mit 9 Prozent Fehler immer noch schlechter als das DM-Modell mit 6 Prozent. Ersetzt man das arithmetische Mittel im RKI-Modell durch das geometrische Mittel (RKI-geom, siehe Abbildung 1.3) wird das Infektionsgeschehen optimal beschrieben und der Fehler liegt nur noch bei 6 Prozent. Die mathematischen Details sind im Anhang 4 beschrieben. Diese Analysen haben wir in Zusammenarbeit mit Dr. G. Fläschner von der ETH Zürich durchgeführt.

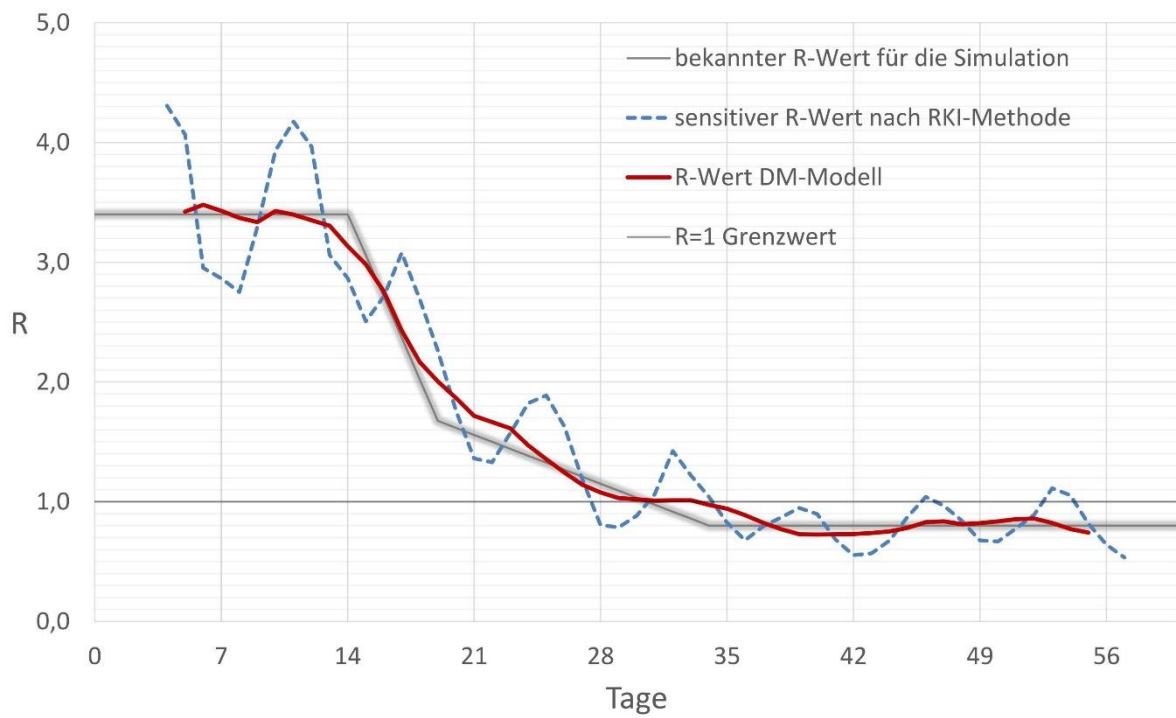


Abbildung 1.1. Vergleich von zwei R-Wert-Berechnungen:
RKI-sensitiv und DM-Modell

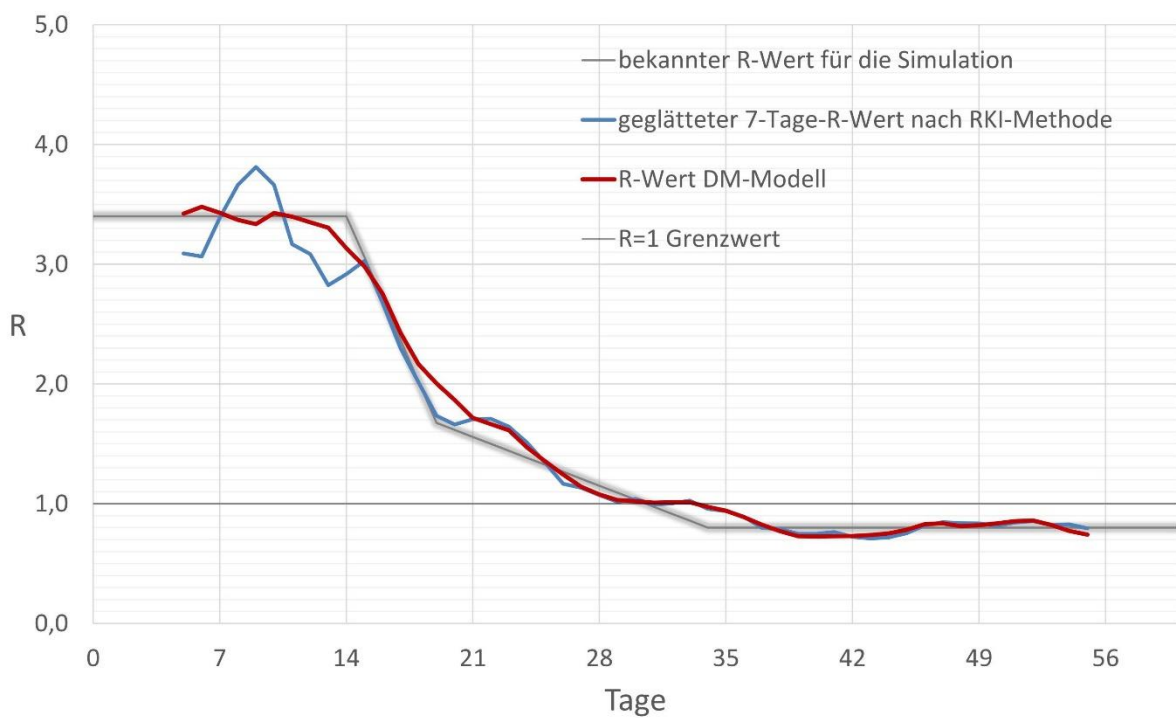


Abbildung 1.2. Vergleich von zwei R-Wert-Berechnungen:
RKI-geglättet und DM-Modell

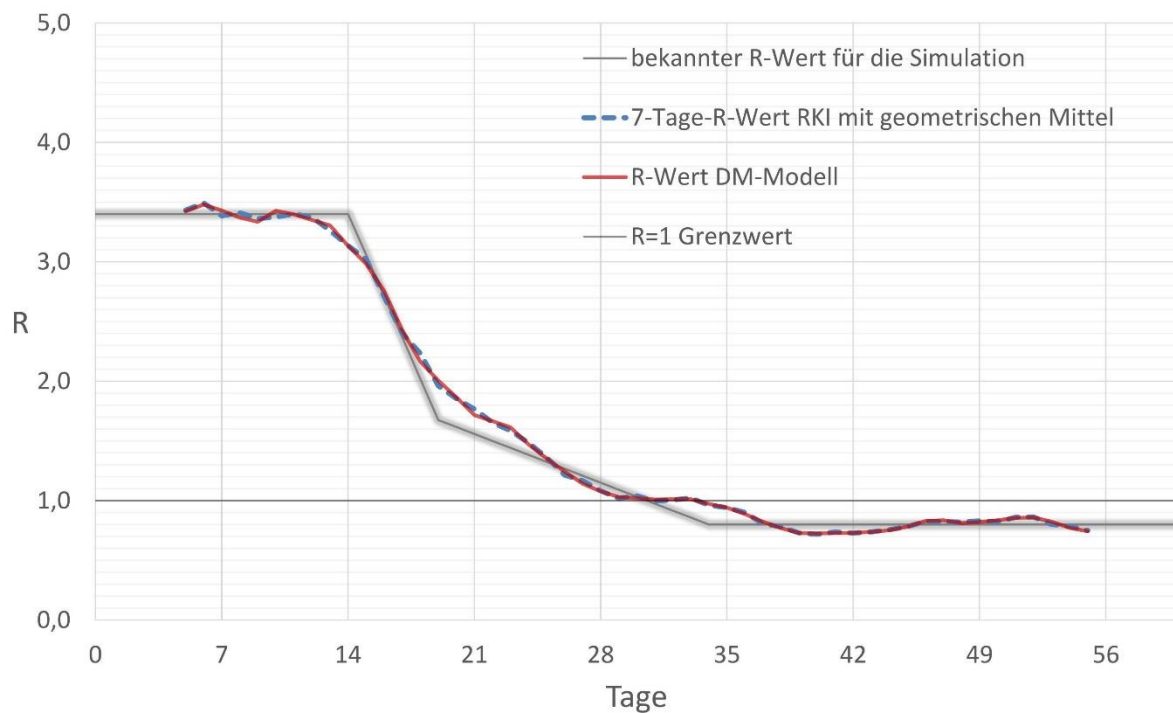


Abbildung 1.3. Vergleich von zwei R-Wert-Berechnungen:
RKI-geom und DM-Modell

Modell	Fehler in Prozent
RKI-sensitiv	22
RKI-geglättet	9
DM-Modell	6
RKI-geom	6

Tabelle 1.1. Der Fehler in Prozent für vier verschiedene R-Wert-Berechnungen.

Je genauer man den R-Wert berechnen kann, desto besser können wir untersuchen, was die Maßnahmen gebracht haben. Für die erste Welle haben wir gezeigt, dass die Maßnahmen im März die exponentielle Ausbreitung des Virus gestoppt haben. Diese Methode werden wir auch für die Untersuchung des zweiten Lockdowns ab November 2020 und den darauffolgenden Verschärfungen einsetzen.

2. Der R-Wert einfach berechnet

Wir haben uns ein neues Frühwarnsystem ausgedacht, dass jeder, ohne groß zu rechnen, anwenden kann. Ist die Anzahl der Neuinfizierten pro Tag (NI) heute größer als der Wert vor einer Woche, dann ist R größer als 1. Mit einem Taschenrechner kann man sogar den R-Wert abschätzen:

$$NI\text{-Wert von heute} / NI\text{-Wert vor einer Woche} \approx R\text{-Faktor}$$

Die genaue Formel lautet:

$$\text{Schätzwert für } R = \left(\frac{NI \text{ von heute}}{NI \text{ vor einer Woche}} \right)^{\frac{4}{7}}$$

Die 4 bedeutet das serielle Intervall* von 4 Tagen und die 7 die Zahl der Tage zwischen den verglichenen Werten. Diese Werte sind in Abbildung 2.1 für die erste Welle als Punkte markiert. Überraschenderweise zeigen diese einfach berechneten Schätzwerte das gleiche zeitliche Verhalten wie die Reproduktionszahl R .

*Das serielle Intervall definiert das durchschnittliche Intervall vom Beginn der Erkrankung eines ansteckenden Falles bis zum Erkrankungsbeginn, eines von diesem angesteckten Falles. Das RKI gibt für das serielle Intervall 4 Tage an.

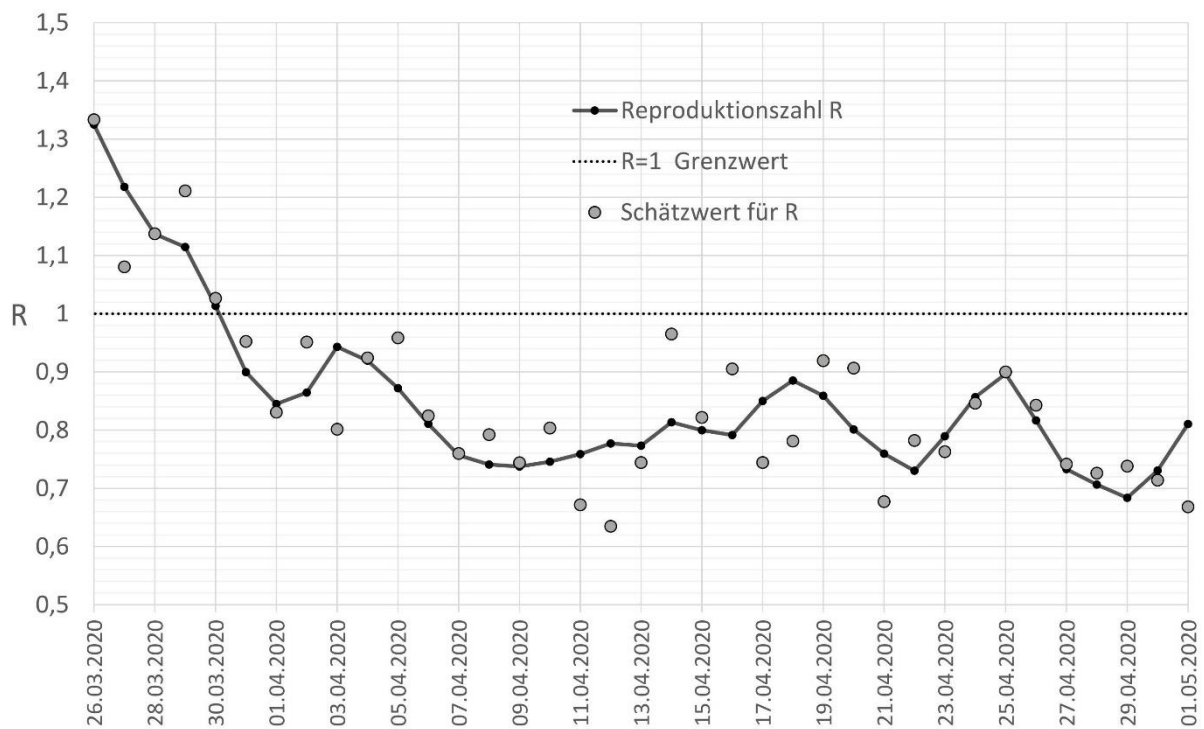


Abbildung 2.1. Der Schätzwert für R im Vergleich zum R -Wert mit dem DM-Modell berechnet.

In der Tabelle 2.1 sind drei verschiedene Methoden den R -Wert zu berechnen miteinander verglichen. Überraschenderweise ist DM-Schätz mit einem Fehler von 10 Prozent deutlich besser als der sensitive Wert vom RKI mit 22 Prozent (siehe auch Abbildung 2.2) und fast so gut wie RKI-geglättet mit 9 Prozent (siehe auch Abbildung 2.3). Wie man in der Abbildung 2.2 deutlich sieht, glättet DM-schätz die Wochentagschwankungen viel besser als RKI-sensitiv.

Modell	Fehler in Prozent
RKI-sensitiv	22
DM-Schätz	10
RKI-geglättet	9

Tabelle 2.1. Der Fehler in Prozent für drei verschiedene R -Wert-Berechnungen.

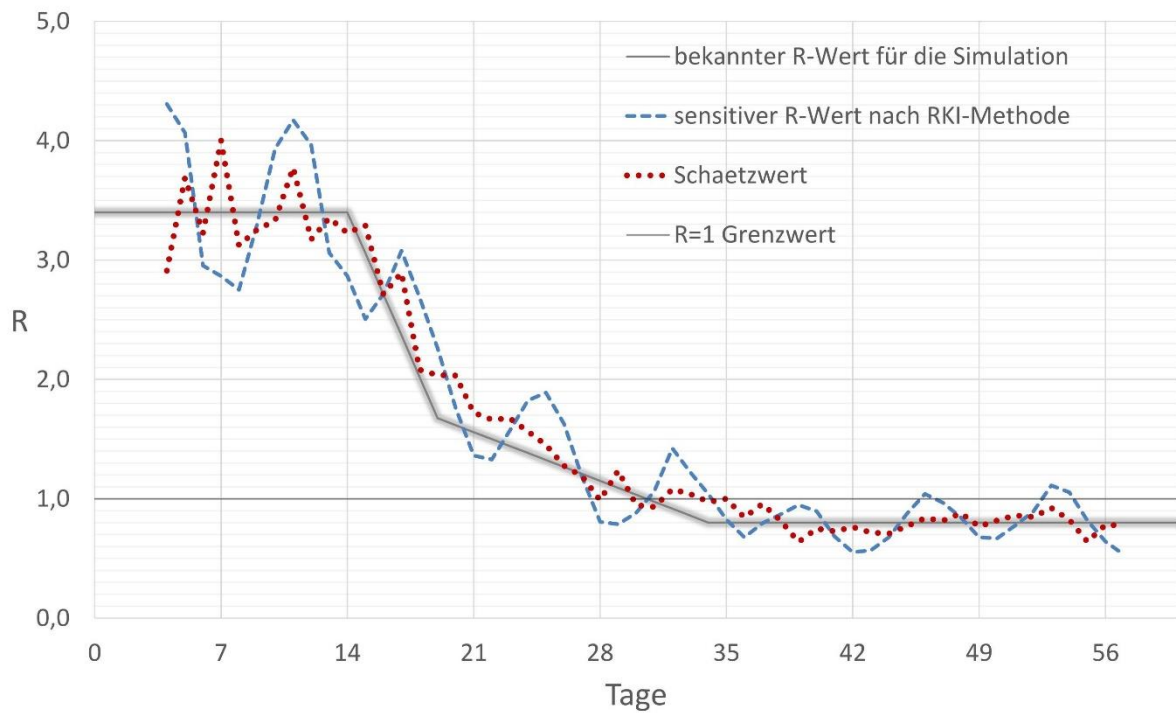


Abbildung 2.2. Vergleich von RKI-sensitiv und DM-Schätz.

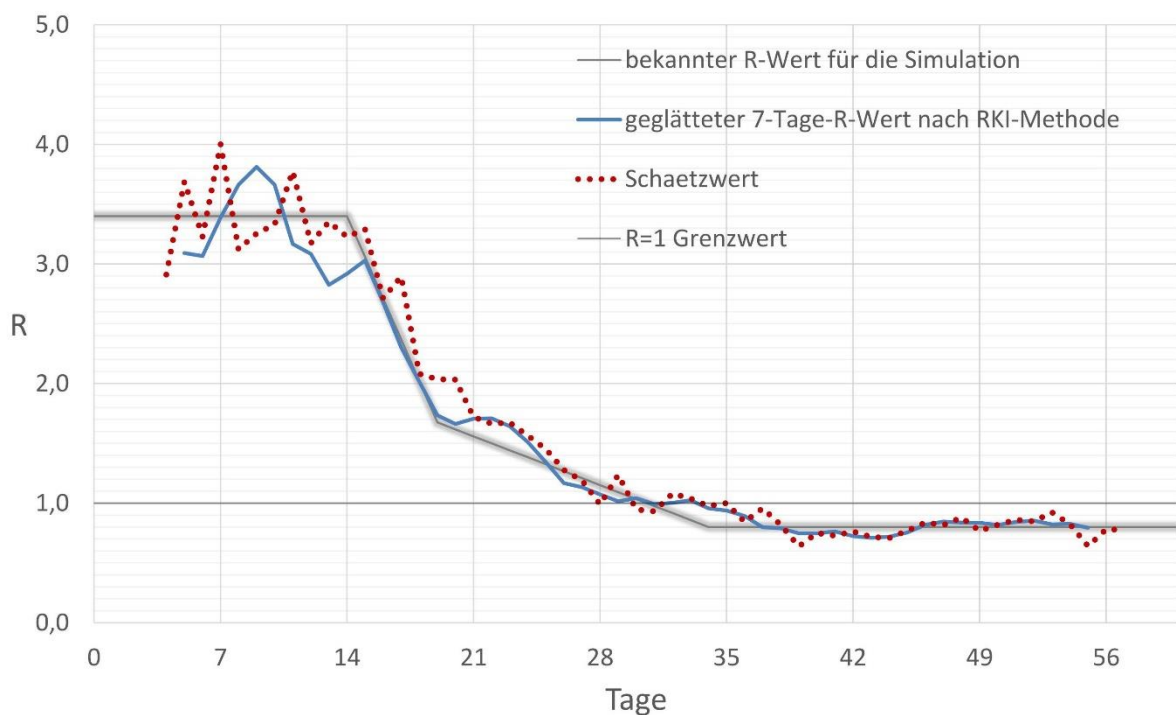


Abbildung 2.3. Vergleich von RKI-geglättet und DM-Schätz.

Mit DM-Schätz kann sich auch der Laie nur mit Hilfe eines Taschenrechners sehr schnell ein eigenes Bild vom Infektionsgeschehen machen. Natürlich kann dieser Wert nicht die viel bessere Berechnung mit dem DM-Modell oder mit RKI-geom ersetzen.

3. Fehlalarm-Analyse

Das Ziel der Maßnahmen im November 2020 war es, den R-Wert deutlich unter eins zu bringen. Haben wir ein niedriges Neuinfizierten-Niveau erreicht, können wir wieder schrittweise lockern, aber nur so weit, dass der R-Wert im Mittel bei eins liegt. Steigt der R-Wert über einen Grenzwert, sollte ein Alarm ausgelöst werden. Wird ein Alarm ausgelöst, kann die Politik schnell und mit gezielten Maßnahmen reagieren. Fehlalarme müssen vermieden werden, denn sie würden die Menschen nur verunsichern. Ein Alarm sollte nach unserer Meinung dann ausgelöst werden, wenn der R-Wert über 1,2 liegt oder wenn der Inzidenzwert über 50 ist.

Aus Abbildung 3.1 kann man die Anzahl der Fehlalarme für verschiedene R-Wert-Berechnungsmethoden abzählen. In der Tabelle 3.1 sind die Ergebnisse zusammengefasst. Der sehr einfach zu berechnende R-Wert DM-Schätz löst überraschenderweise im Zeitraum von 2.000 Tagen mit 14 – das entspricht 0,7 Fehlalarme in 100 Tagen – deutlich weniger Fehlalarme aus, als RKI-sensitiv mit 263 – das entspricht 13 Fehlalarmen in 100 Tagen. Alle anderen Modelle lösen keinen einzigen Fehlalarm in 2.000 Tagen aus.

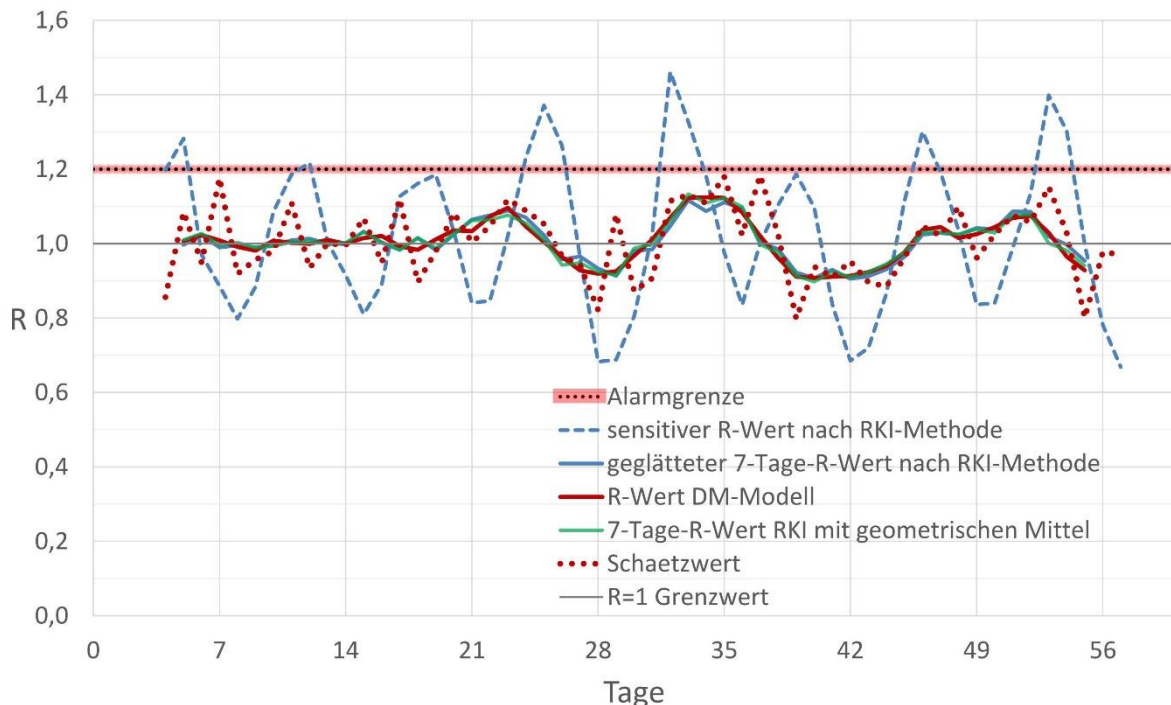


Abbildung 3.1. Fehlalarme für 5 verschiedene R-Wert-Berechnungen.

Modell	Fehlalarme
RKI-sensitiv	263
DM-Schätz	14
RKI-geglättet	0
DM-Modell	0
RKI-geom	0

Tabelle 3.1. Anzahl der Fehlalarme für 5 verschiedene R-Wert-Berechnungen im Zeitraum von 2.000 Tagen.

4. Fehler in Abhängigkeit vom R-Wert

Die Fehler für RKI-geom und dem DM-Modell sind konstant bei 4%. Sie hängen also nicht vom R-Wert ab (siehe Abbildung 4.1 und Tabelle 4.1). RKI-geglättet zeigt dagegen eine deutliche Abhängigkeit vom R-Wert. Bei $R=1$ ist der Fehler am niedrigsten und liegt auch bei 4%. Bei $R=3$ ist dagegen der Fehler mit 12% dreimal so groß.

Der Unterschied zwischen dem arithmetischen und dem geometrischen Mittel ist der Grund dafür (siehe Anhang 4). Da R-geglättet mit dem arithmetischen Mittel, R-geom aber mit dem geometrischen Mittel rechnet, ist die Abweichung zwischen den beiden R-Werten nur dann vernachlässigbar klein, wenn sich das Infektionsgeschehen nicht stark ändert (Beispiel: $R=1$). Für starke Änderungen im Infektionsgeschehen (Beispiel: $R=3$) muss mit R-geom oder mit dem DM-Modell gerechnet werden.

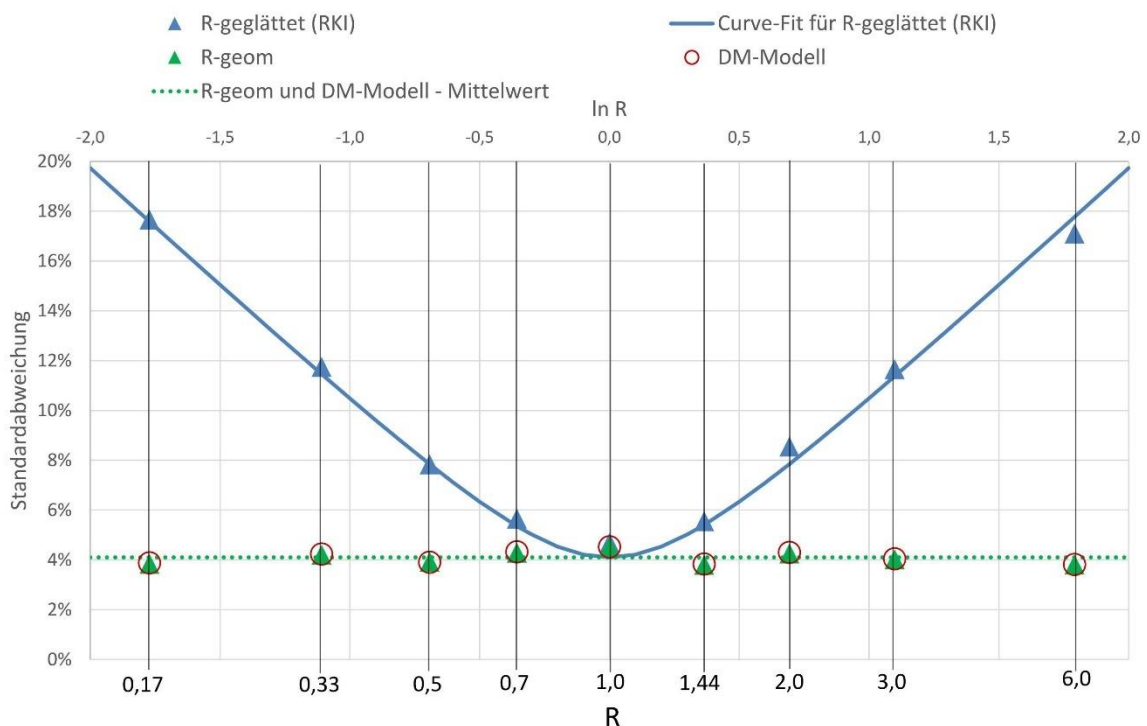
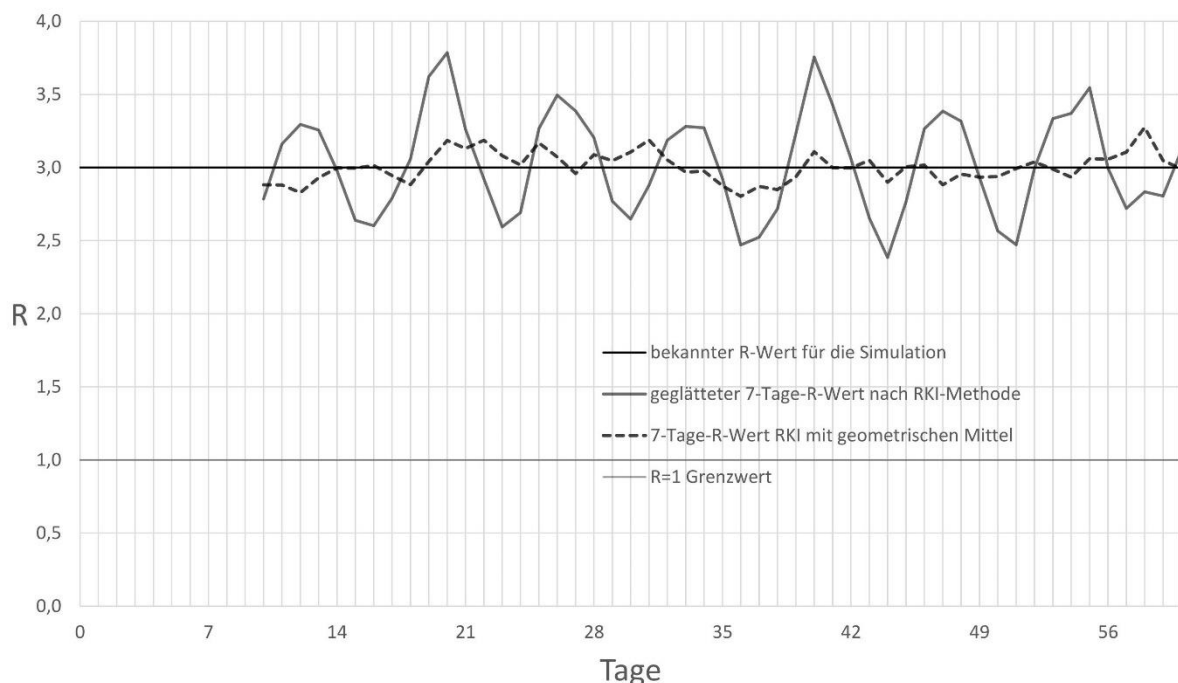


Abbildung 4.1. Fehler in Abhängigkeit vom R-Wert

Modell	Fehler für R=1	Fehler für R=3
RKI-geglättet	4 %	12 %
DM-Modell	4 %	4 %
RKI-geom	4 %	4 %

Tabelle 4.1. Fehler für R=1 und R=3 für 3 verschiedene R-Wert-Berechnungen.

In Abbildung 4.2. ist RKI-geglättet für R=3 verglichen mit RKI-geom. Auch hier ist das Ergebnis eindeutig. RKI-geom kann das Infektionsgeschehen besser beschreiben als RKI-geglättet. RKI-geglättet zeigt sehr starke Wochentagschwankungen, RKI-geom dagegen nicht. Der Unterschied zwischen dem arithmetischen und dem geometrischen Mittel ist der Grund dafür (siehe Anhang 4). Wir empfehlen daher bei der R-Wert Berechnung das arithmetische Mittel durch das geometrische Mittel zu ersetzen.



*Abbildung 4.2. Vergleich von zwei R-Wert-Berechnungen:
RKI-geglättet und RKI-geom*

Starke Schwankungen verwirren die Menschen und lassen viele an der Wissenschaft zweifeln. Wissenschaftler sollten alles dafür tun, mit ihren Aussagen und Berechnungen die Menschen zu überzeugen und nicht zu verwirren. Wir

haben es noch nicht geschafft. Wir alle entscheiden mit unserem Verhalten über den weiteren Verlauf der Pandemie.

5. Zusammenfassung

Für RKI-geom und dem DM-Modell liegt der Fehler zwischen 4% (R-Wert der Testfunktion konstant) und 6% (Knickstellen im R-Wert-Verlauf der Testfunktion, wie in Abbildung 1.1 gezeigt). Beide Berechnungsmethoden beschreiben das Infektionsgeschehen sehr gut und lösen keine Fehlalarme aus (siehe Tabelle 5.1)

Modell	Fehler	Fehlalarme
RKI-sensitiv	22 %	sehr viele Fehlalarme
DM-Schätz	10 %	wenig Fehlalarme
RKI-geglättet	9 %	keine Fehlalarme
DM-Modell	6 %	keine Fehlalarme
RKI-geom	6 %	keine Fehlalarme

Abbildung 5.1. Der Fehler und die Anzahl der Fehlalarme für 5 verschiedene R-Wert-Berechnungen.

Anhang

1. Waren die Maßnahmen gegen die erste Welle wirksam?

Am 9.03., 16.03. und 23.03.2020 traten in Deutschland Maßnahmen gegen die Verbreitung des Virus in Kraft. Waren diese Maßnahmen die Ursache für die Verbesserung des R-Wertes?

Abbildung A1.1 zeigt die gute Übereinstimmung der *R-Werte* (Punkte) mit unserem DM-Modell (durchgezogenen Linie). Die Abbildung verdeutlicht, dass die Maßnahmen tatsächlich in einem zeitlichen Zusammenhang stehen. Nach der ersten Maßnahme, der Absage von Großveranstaltungen, hat sich der R-Wert von 3,2 auf 1,35 verbessert. Nach der zweiten Maßnahme, der Bund-Länder-Vereinbarung, auf 1,13 und schließlich auf 0,80. Maßnahmen im März konnten also tatsächlich die exponentielle Ausbreitung des Virus stoppen und eine Katastrophe verhindern.

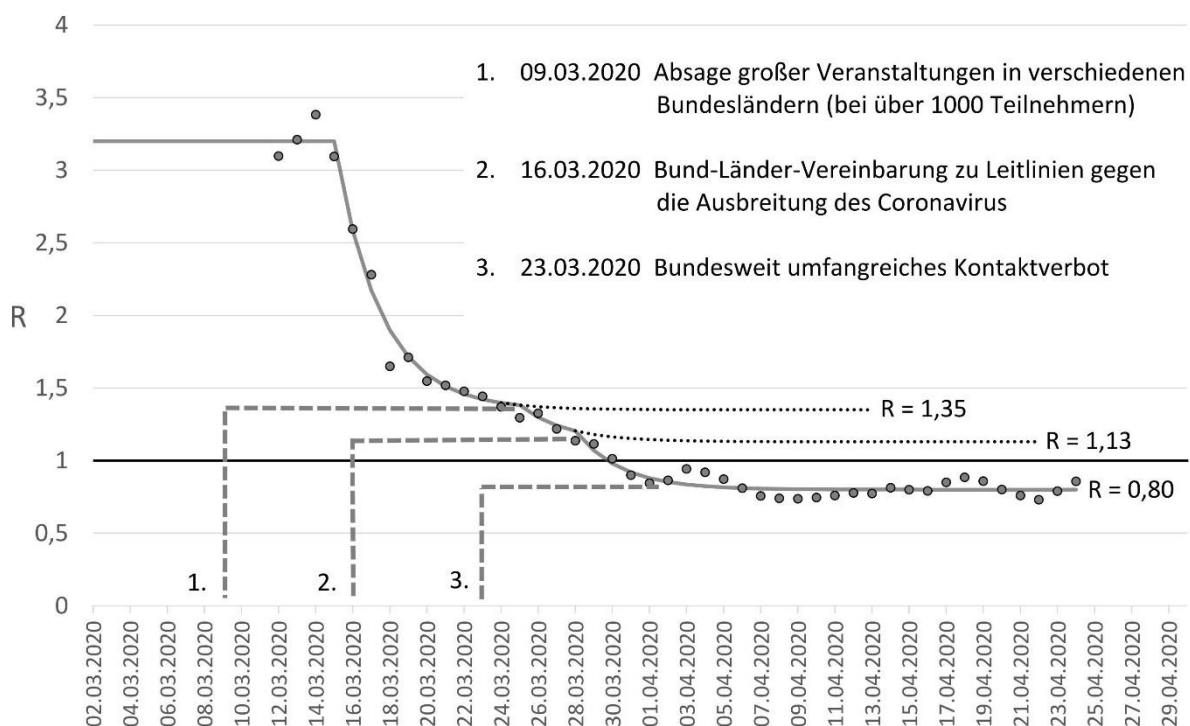


Abbildung A1.1. Analyse der Maßnahmen für Deutschland.

Reproduktionszahl R mit dem DM-Modell berechnet mit Daten aus [D1].

Simulation von lokalen Ausbrüchen

Das von uns entwickelte Modell lässt sich auch sehr gut auf lokale Ausbrüche anwenden (siehe Abbildung A1.2). Die Übereinstimmung zwischen der berechneten Kurve (gepunktet) und den gemittelten Daten (durchgezogene Kurve) ist beeindruckend. Details zur Analyse von lokalen Ausbrüchen ist in [1] zu finden.

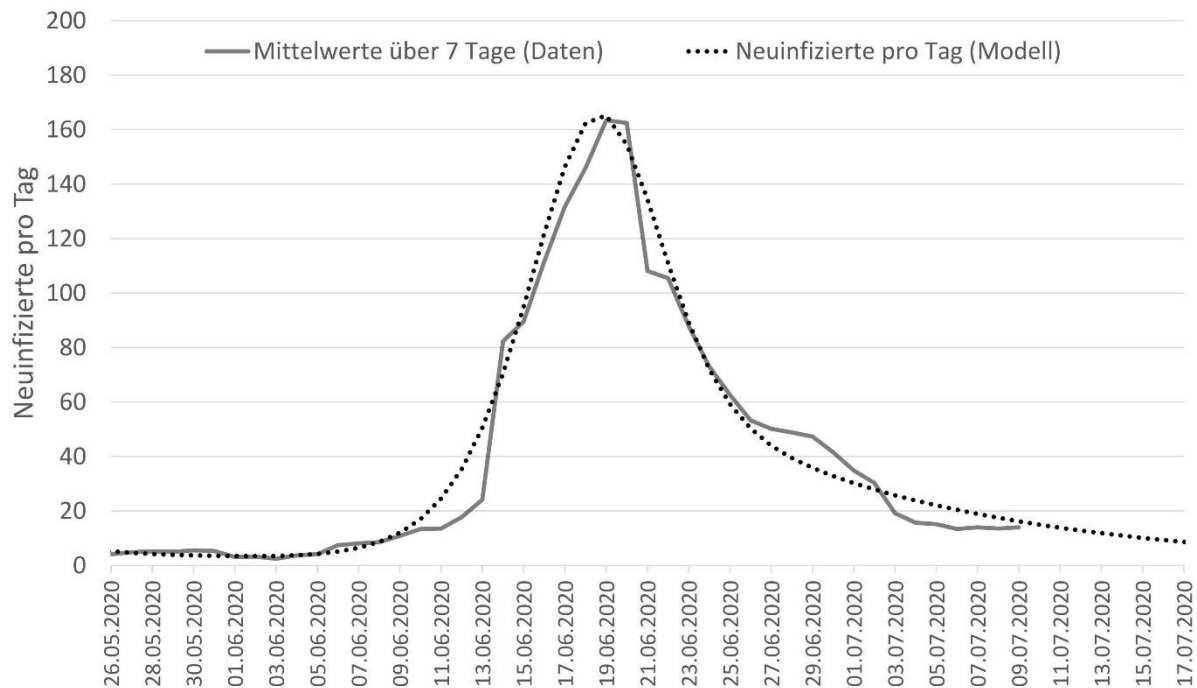


Abbildung A1.2. Neuinfizierte pro Tag im Kreis Gütersloh gemittelt über 7 Tage im Vergleich mit dem DM-Modell. Quelle der Daten [D2].

2. Fehlerberechnung

R-Wert-Berechnungen anhand realer Daten sind immer Schätzungen, die mit statistischen Unsicherheiten behaftet sind. Da der „wahre“ R-Wert normalerweise nie bekannt ist, fällt es schwer, verschiedene Modelle der R-Wert-Berechnung untereinander zu vergleichen.

Um die verschiedenen Modelle dennoch miteinander vergleichen zu können, verwenden wir Testfunktionen (siehe Abbildung A2.1). Eine Testfunktion bzw. R-Wert-Funktion gibt den zeitlichen Verlauf der „wahren“ R-Werte an. Aus der R-Wert-Funktion werden die Neuinfektionszahlen pro Tag berechnet. Diese Neuinfektionszahlen werden dann mit einer realistischen Wochentagschwankung (systematische Abweichungen) und mit Rauschen (zufällige Abweichungen)

überlagert (siehe Abbildung A2.2). Für die simulierten Neuinfektionszahlen kennen wir dann den „wahren“ R-Wert.

So können wir für die verschiedenen untersuchten Berechnungsmodelle jeweils die relativen Abweichungen

$$\frac{R_{Modell} - R_{wahr}}{R_{wahr}}$$

angeben. Unserer Maßzahl für den Vergleich ist die Wurzel der mittleren Fehlerquadratsumme (RMSE). Um den Fehler genügend genau zu bestimmen, haben wir jeweils 1000 Abweichungen herangezogen. Die Abbildungen 1.1 bis 1.3 und die Tabelle 1.1 zeigen die Ergebnisse.

Die Abbildungen A2.1. und A2.2 beschreiben den Verlauf der ersten Welle sehr gut.

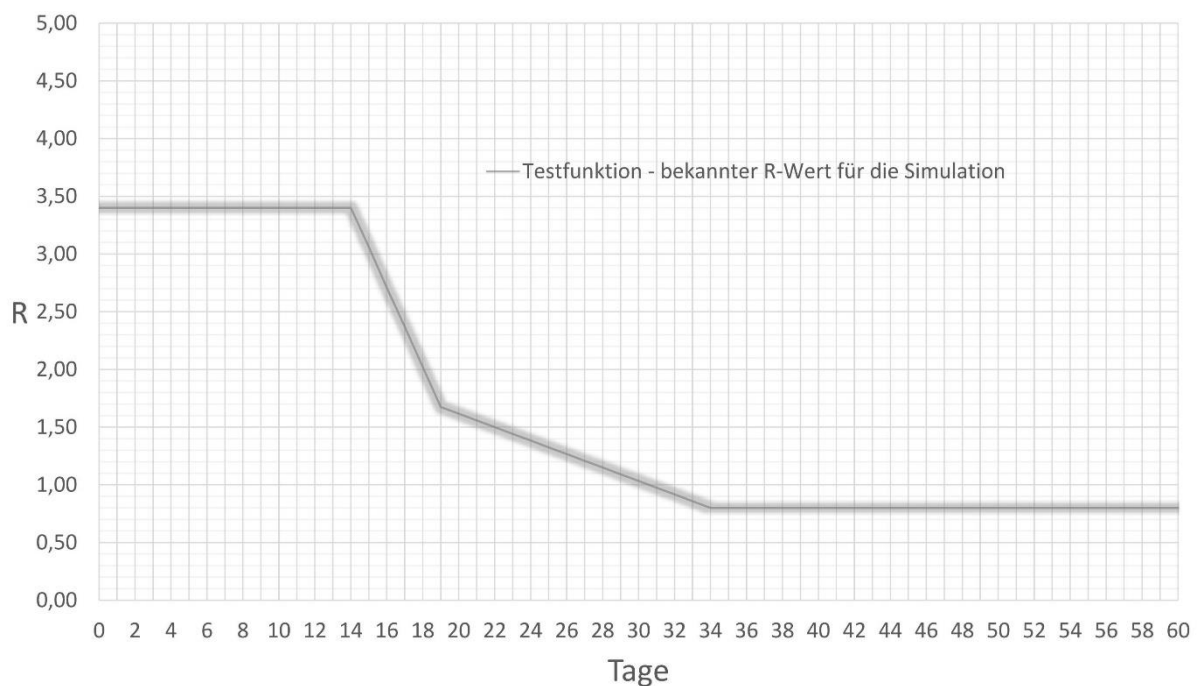


Abbildung A2.1. Testfunktion für die Abbildungen 1.1.1 bis 1.1.3. für die erste Welle.

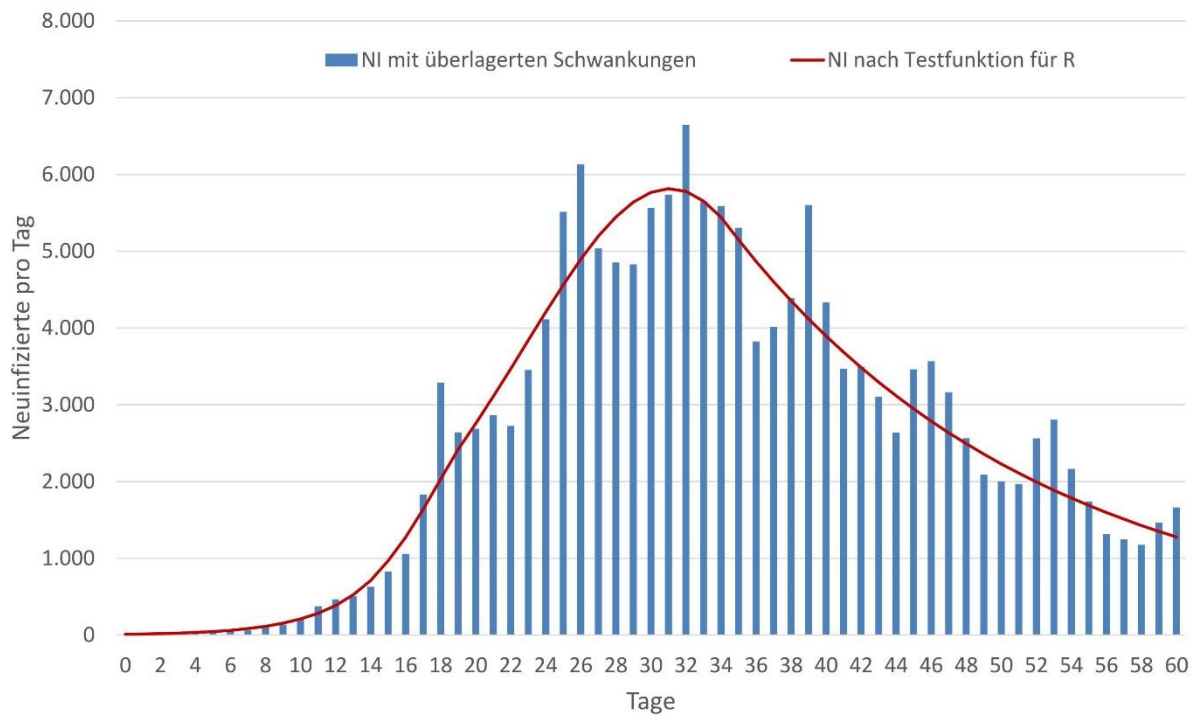


Abbildung A2.2. Anzahl der Neuinfizierten pro Tag.

Für Spezialisten:

Um den Fehler allgemeingültig zu rechnen für verschiedene Epidemien, müsste man sich entweder auf den Faktor a beziehen, um den die Neuinfektionen NI innerhalb eines Tages steigen oder fallen, bzw. auf die Steigung $k = \ln a$. Dann wäre man unabhängig vom seriellen Intervall t_{si} . Unsere Berechnung basiert auf $t_{si} = 4$.

Der zufällige Fehler lässt sich näherungsweise durch folgende Formel berechnen:

$$z_R = 0,41 * \sigma$$

Die zufälligen relativen Abweichungen der Neuinfektionszahlen NI sind näherungsweise normalverteilt. σ ist die Standardabweichung der Normalverteilung. Für unsere Modellrechnungen haben wir mit $\sigma = 0,10$ gerechnet (entsprechend 10 % vom jeweiligen Sollwert des Modells).

Der Fehler für die Wochentagschwankungen, nur für RKI-geglättet, beträgt näherungsweise:

$$w_R = 0,38 * A * |\ln R|$$

A ist die relative Amplitude der Wochentagschwankungen. Für unsere Modellrechnungen haben wir mit $A = 0,25$ gerechnet (entsprechend 25 % vom jeweiligen Sollwert des Modells). R ist hier der „wahre R-Wert“ der Testfunktion.

Der Gesamtfehler für RKI-geglättet ergibt sich dann zu:

$$wz_R = \sqrt{z_R^2 + w_R^2}$$

Das DM-Modell und RKI-geom zeigen keine Abhängigkeit von den Wochentagschwankungen. Für diese Modelle ist $w_R = 0$. Die Näherungsformeln geben den kleinstmöglichen Fehler an. Bei Knickstellen im Verlauf von R kann die R-Wert-Berechnung nicht unmittelbar folgen. Dies führt zu größeren Abweichungen (Gilt für alle Methoden der R-Wert-Berechnung).

3. Modellbeschreibung

Wie sind wir vorgegangen?

Zunächst haben wir zwei verschiedene Ansätze verfolgt. Ein „konventioneller“ Ansatz, bei der Simulationen auf der Basis der Lösung von Differentialgleichungen gerechnet werden, und ein neuer Ansatz aus der Regelung dynamischer Systeme. Uns war schnell klar, dass der Ansatz nicht nur einfacher ist, sondern auch ein viel größeres Potential hat.

Wir verwenden Infiziertenzahlen, die für viele Länder der Welt zur Verfügung stehen. Damit sind Vergleiche zwischen Ländern möglich. Täglich aktualisierte Daten stehen bei „Humanitarian Data Exchange“ [D1] zum Download bereit. Genutzt werden die täglich herausgegebenen kumulierten Neuinfizierten N . Die Anzahl der Neuinfizierten pro Tag $N/$ ergibt sich dann als Differenz zwischen aufeinanderfolgenden Werten. Das mathematische Modell besteht auf den Komponenten Datenverarbeitung, Datenanalyse und Prognose.

Datenverarbeitung

Die Anzahl der Neuinfizierten pro Tag $N/$ zeigt starke Wochentagschwankungen, verursacht durch den sogenannten Meldeverzug vor und nach den Wochenenden. Diese Schwankungen filtern wir durch gleitende Mittelwertbildung über 7 Tage heraus.

Um die zeitabhängige Reproduktionszahl R zu bestimmen, wird eine exponentielle Funktion an die Datenpunkte N angepasst. Dazu werden die Zahlen N zunächst logarithmiert. Die exponentielle Funktion lässt sich dann durch eine Geradengleichung darstellen. Die Steigung k der Geraden wird über die Minimierung der quadratischen Abweichungen von 11 aufeinanderfolgenden Datenpunkten $\ln(N)$ von dieser Gerade bestimmt (einfache lineare Regression ELR). Mit Hilfe eines kleinen Tricks können wir den Einfluss der Wochentagschwankung noch weiter verkleinern. Statt die Abweichungen von 11 Datenpunkten in einem Schritt zu minimieren, machen wir das in zwei Schritten mit jeweils 10 Datenpunkten, die um einen Tag zeitversetzt sind. Das ergibt zwei Steigungen k_1 und k_2 . Die Steigung k ist dann der Mittelwert $k = \frac{1}{2} * (k_1 + k_2)$. Diese Berechnung wird für jeden Tag neu durchgeführt. Dieses Verfahren ist eine sehr einfache, aber trotzdem sehr effektive Methode, um die starken Schwankungen zu unterdrücken. So können Berechnungen, wie sie in der Abbildung A1.1 durchgeführt werden, gelingen.

Aus der Steigung k und dem seriellen Intervall t_{si} wird R berechnet:

$$R = e^{k \cdot t_{si}}$$

Als serielles Intervall t_{si} bezeichnet man das durchschnittliche Zeitintervall vom Beginn der Erkrankung eines Infizierten Menschen bis zum Erkrankungsbeginn eines von diesem angesteckten. Das Robert-Koch-Institut gibt für das serielle Intervall 4 Tage an [2]. Diesen Wert übernehmen wir in unsere Rechnungen.

Datenanalyse

Mit Hilfe der Datenanalyse wird z.B. die Wirksamkeit von Maßnahmen ausgewertet. Abbildung A1.1 zeigt die Vorgehensweise. Es wird eine Folge von Reproduktionszahlen bestimmt, die den Verlauf der aus den Daten ermittelten Werte (Punkte) in guter Näherung beschreiben. Im vorliegenden Fall ist das die Folge: 3,20 – 1,35 – 1,13 – 0,80 als Werte für R .

4. Das arithmetische und das geometrische Mittel

Der geglättete R-Wert vom RKI zeigt gute Ergebnisse, wenn R ungefähr 1 ist, aber hat Probleme, wenn die Anzahl der Neuinfizierten exponentiell ansteigen, wie beispielsweise zum Beginn der ersten und zweiten Welle. Dieses Problem konnten wir lösen, indem wir das arithmetische Mittel durch das geometrische Mittel ersetzt haben. Um den Unterschied zwischen dem arithmetischen und dem geometrischen Mittel zu verstehen, betrachten wir ein Zahlenbeispiel.

Das arithmetische Mittel von 1 und 9 ist 5. Die Rechnung dazu lautet:

$$\frac{1 + 9}{2} = 5$$

Das geometrische Mittel von 1 und 9 ist 3. Die Rechnung dazu lautet:

$$\sqrt{1 * 9} = 3$$

Das geometrische Mittel ist immer kleiner oder gleich dem arithmetischen Mittel. Es wird bei prozentualen Veränderungen oder bei der Berechnung von Wachstumsraten eingesetzt. Irrtümlicherweise wird hier oft das arithmetische Mittel verwendet.

Warum man hier das geometrische Mittel verwenden muss, machen wir uns an einem zweiten Zahlenbeispiel klar. Alle 4 Tage verdoppelt sich die Anzahl der Neuinfizierten von 1 auf 2 und danach auf 4 (siehe Abbildung A4.1). Der Mittelwert liegt bei 2. Der geometrische Mittelwert liegt auch bei 2. Das arithmetische Mittel liegt dagegen bei 2,33. Bei exponentiellen Veränderungen muss man also das geometrische Mittel verwenden.

$$\frac{1 + 2 + 4}{3} = 2,33 > 2 = \sqrt[3]{1 * 2 * 4}$$

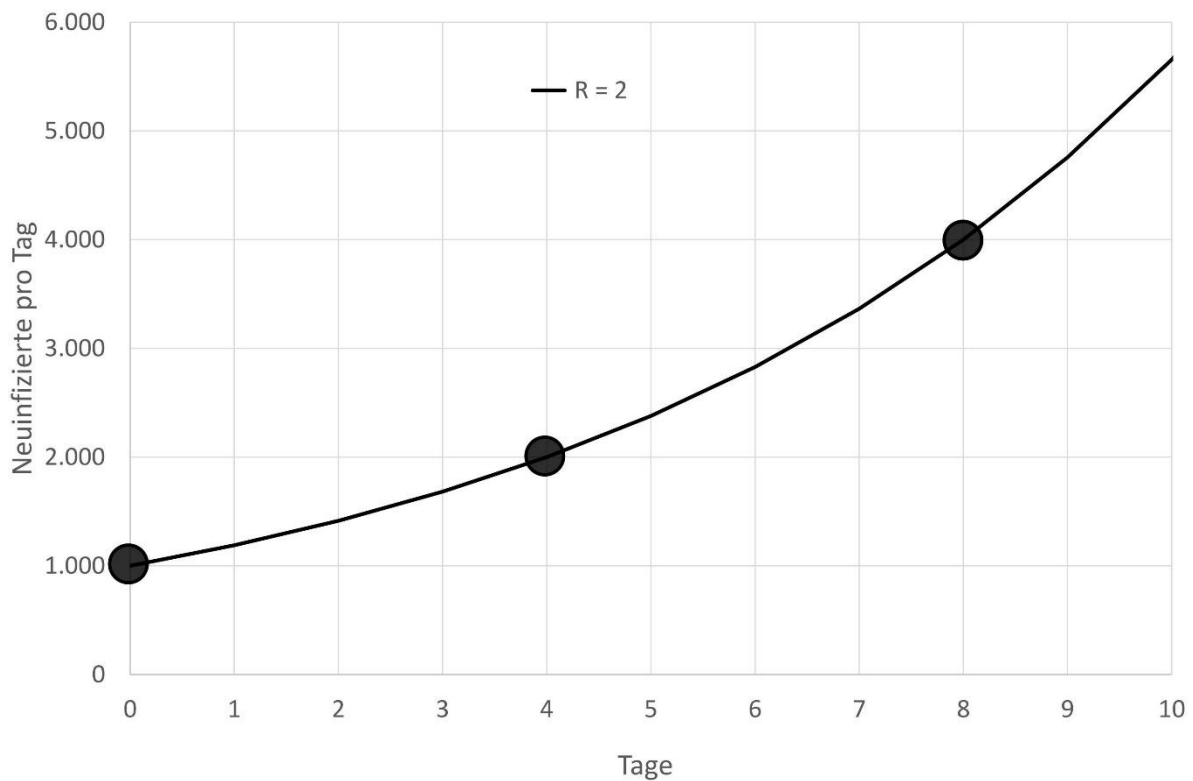


Abbildung A4.1. Neuinfizierte pro Tag für $R=2$

Warum hängt der Fehler vom R-Wert ab?

Im Kapitel 4 haben wir gesehen, dass die Abweichung von R-geglättet zu R-geom vom R-Wert abhängt und für $R=1$ ist die Abweichung Null. Der Grund dafür ist ganz einfach. Für zwei gleiche Zahlen ist die Abweichung zwischen den arithmetischen und geometrischen Mittel auch Null. Das machen wir uns an einem Zahlenbeispiel klar.

$$\frac{2 + 2}{2} = 2 = 2 = \sqrt[2]{2 * 2}$$

Das arithmetische und das geometrische Mittel sind gleich. Sind dagegen die Zahlen unterschiedlich, in unseren Beispiel 1 und 9, dann ist das arithmetische Mittel größer als das geometrische Mittel.

$$\frac{1 + 9}{2} = 5 > 3 = \sqrt[2]{1 * 9}$$

Da R-geglättet mit dem arithmetischen Mittel, R-geom aber mit dem geometrischen Mittel rechnet, ist die Abweichung zwischen den beiden R-Werten nur dann vernachlässigbar klein, wenn sich das Infektionsgeschehen nicht stark

ändert. Für starke Änderungen im Infektionsgeschehen muss mit R-geom oder mit dem DM-Modell gerechnet werden.

Zusammenfassung

Je genauer man den R-Wert berechnen kann, desto besser können wir untersuchen, was die Maßnahmen gebracht haben. Für die erste Welle haben wir gezeigt, dass die Maßnahmen die exponentielle Ausbreitung des Virus gestoppt haben. Die gleiche Analyse haben wir für die zweite Welle gemacht.

Außerdem haben wir die R-Wert-Berechnung des Robert Koch-Instituts (RKI) verbessert. Ersetzt man das arithmetische Mittel im RKI-Modell durch das geometrische Mittel wird das Infektionsgeschehen optimal beschrieben. Für RKI-geom und dem DM-Modell liegt der Fehler bei durchschnittlich 5 Prozent. Beide Berechnungsmethoden beschreiben das Infektionsgeschehen sehr gut und lösen sehr wenig Fehlalarme aus.

Literatur

- [1] P. Möller, H. Drewes, Corona – Zahlen richtig verstehen, Dr. Köster, 2020
- [2] an der Heiden M, Harmouda O: Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting (2020)

Quellen der Daten

- [D1] Für alle anderen Diagramme: **Humanitarian Data Exchange**
<https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases>
- [D2] Für Gütersloh: **RKI**
<https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4>. Verwendet wurden die „Fälle kumuliert“.