

Vorhersagen sind schwer, vor allem die Zukunft betreffend: Kurzzeitprognosen in der Pandemie

Johannes Bracher, Daniel Wolffram, Tilmann Gneiting und Melanie Schienle

Epidemiologische Modellierungen stehen seit Beginn der COVID-19-Pandemie im Fokus der Öffentlichkeit und haben politische Entscheidungen maßgeblich beeinflusst. Neben retrospektiven Einordnungen und längerfristigen Szenario-Prognosen können sie auch zur Vorhersage kurzfristiger Trends dienen. In mehreren internationalen Kollaborationsprojekten werden derartige Vorhersagen systematisch zusammengetragen, evaluiert und in Ensemble-Vorhersagen kombiniert. Das Rückgrat dieser Analysen bilden mathematisch-statistisch fundierte Maße der Vorhersagequalität, sogenannte *proper scoring rules*. Systematische Evaluationsstudien zeigen: Die Kombination verschiedener Vorhersagen in Ensemble-Prognosen kann die Vorhersagegüte verbessern. Vor allem Fallzahlen sind jedoch nur für sehr kurze Horizonte zuverlässig prognostizierbar, und eine angemessene Quantifizierung der Vorhersageunsicherheit ist von zentraler Bedeutung.

Epidemiologische Modelle: Beschreiben, verstehen, prognostizieren

Mathematische und statistische Modelle sind ein Werkzeug, um theoretische Annahmen mit Beobachtungen zu kombinieren und so Aussagen über bisher Unbekanntes oder Unbeobachtetes zu treffen. Es überrascht deshalb nicht, dass sich in der von großer Unsicherheit geprägten COVID-19-Pandemie viel Aufmerksamkeit auf derartige Modelle richtet. Die Palette der Fragestellungen seit Januar 2020 ist dabei breit: Wie groß war der Ausbruch der bisher unbekanntesten Krankheit in Wuhan bereits, als erste Fallmeldungen publik wurden [10]? Wie lange ist die Inkubationszeit [12], also die Zeit, bis im Anschluss an eine Ansteckung Symptome auftreten? Wie hoch ist die Basisreproduktionszahl, d. h. die mittlere Zahl neuer Infektionen, die ein Infizierter in Abwesenheit von Gegenmaßnahmen oder Vorimmunitäten verursacht [18]? Und eine der am meisten in der Öffentlichkeit diskutierten Fragen: Wie geht es weiter?

Hinter dieser Frage verbergen sich eigentlich gleich zwei Aufgabenstellungen, die klar zu unterscheiden sind: Mittels sogenannter *Szenario-Projektionen* sollen mögliche Verläufe über längere Zeiträume unter bestimmten Rahmenbedingungen abgeschätzt werden. Etwa: Was ist für den Winter zu erwarten, wenn eine bestimmte Impfquote erreicht wird, ein bestimmtes Niveau an nicht-pharmazeutischen Interventionen bestehen bleibt und keine neuen Varianten auftreten, die den Impfschutz umgehen? Diese Projektionen können nur sehr bedingt mit den später beobachteten Daten abgeglichen werden, da in der Regel keines der betrachteten Szenarien exakt die Gegebenheiten, die in der Realität eintraten, widerspiegelt. Anders verhält es sich mit

Kurzzeitvorhersagen: Diese treffen eine empirisch verifizierbare Aussage über den weiteren Verlauf der Pandemie, ohne Bedingung auf hypothetische Annahmen. Damit sind sie prädestiniert für die Validierung und Evaluierung von Modellen, welche im öffentlichen Diskurs oft gefordert wird [6]. Aufgrund der Schnelllebigkeit des Geschehens sind derartige Vorhersagen jedoch nur für sehr kurze Zeithorizonte von wenigen Tagen oder Wochen möglich und sinnvoll [16]. Die stets verbleibende beträchtliche Unsicherheit über den weiteren Verlauf sollte dabei in Form einer *probabilistischen* Vorhersage mitkommuniziert werden.

Das Forecast Hub-Prinzip: Vergleichbarkeit und Kombination von Vorhersagen in Echtzeit

Zu Beginn der Pandemie fokussierte sich die öffentliche und politische Wahrnehmung oft auf Vorhersagen einzelner besonders sichtbarer Modelle wie die des *Institute for Health Metrics and Evaluation* (IHME [9]; siehe [11] für eine kritische Einordnung). Zwar wurde schnell deutlich, dass verschiedene Modelle zu sehr unterschiedlichen Ergebnissen führten, doch ein direkter Vergleich war oft nicht möglich oder zumindest sehr mühsam.

Um die verschiedenen Modelle in vergleichbarer Form zusammenzuführen und zu evaluieren, baute ab April 2020 ein Team um den Biostatistiker Prof. Nicholas Reich (University of Massachusetts Amherst) den *US COVID-19 Forecast Hub* auf [15]: In einem öffentlichen Repository (github.com/reichlab/covid19-forecast-hub/) werden wöchentlich in standardisiertem Format Vorhersagen von Fällen, Hospitalisierungen und Todesfällen zusammengetra-

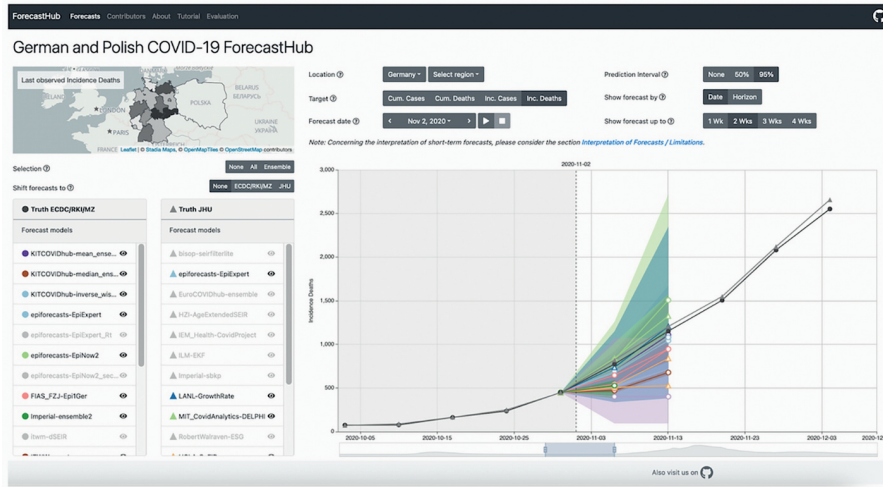


Abbildung 1. Online-Dashboard zur Visualisierung von Kurzzeitvorhersagen. Der Screenshot zeigt Vorhersagen für Todesfälle in Deutschland vom 2. November 2020.

gen und mittels eines interaktiven Dashboards visualisiert (covid19forecasthub.org). Die Vorhersagen bestehen aus 23 Vorhersagequantilen und decken die nächsten vier Wochen ab (im Wissen, dass die Unsicherheit vor allem für die Fallzahlen rapide anwächst). Die Zahl der in der Plattform vertretenen Modelle wuchs dabei stetig von anfänglich nur einer Handvoll auf mittlerweile nahe 100. Deren Vorhersagen werden wöchentlich in eine Ensemble-Prognose gemittelt, welche auch von den US Centers for Disease Prevention and Control (CDC) an die breite Öffentlichkeit kommuniziert wird [4].

Dank des konsequenten *Open-Science*-Ansatzes des Projektes (basierend auf anpassbarer quelloffener Software) fand dieses Format rasch auch hier Anwendung: Seit Mai 2020 betreibt unsere Gruppe am Karlsruher Institut für Technologie (KIT) und Heidelberger Institut für Theoretische Studien (HITS) einen deutsch-polnischen COVID-19 Forecast Hub (kitmetricslab.github.io/forecasthub/). Etwa zwanzig unabhängige Modellierergruppen aus Deutschland, Polen und weiteren Ländern haben seitdem Vorhersagen in die Plattform eingespielt. Abbildung 1 zeigt das zugehörige Visualisierungstool mit Vorhersagen für Todesfälle im November 2020. Im April 2021 startete das *European Center for Disease Prevention and Control* eine europaweite Plattform im selben Format (covid19forecasthub.eu/), in die das deutsch-polnische Projekt mittlerweile eingegliedert ist.

Beispielsweise in der Meteorologie ist die Kombination von Vorhersagemodellen seit Jahrzehnten Standard [7], und auch in der Epidemiologie gab es ähnliche Ansätze bereits vor der Pandemie [17]. Die Erfahrung zeigt, dass derartige Ensemble-Vorhersagen zuverlässiger sind und oft selbst die besten Einzelmodelle ausstechen. Wie später im Text noch ausführlicher diskutiert, hat sich dies in vielen Belangen auch für Vorhersagen des pandemischen Geschehens bestätigt [3, 5]. Doch um derartige Aussagen treffen zu können, stellt sich zunächst die Frage, wie die Qualität einer probabilistischen Vorhersage überhaupt gemessen werden kann.

Evaluierung probabilistischer Vorhersagen: Proper scoring rules

Im Folgenden bezeichnen wir mit Y die vorherzusagende Zufallsvariable mit Ergebnisraum Ω , und mit y die zugehörige Realisierung. Unsere probabilistische Vorhersage für Y ist gegeben durch eine Verteilung F aus einer konvexen Klasse \mathcal{F} von Wahrscheinlichkeitsmaßen. Als Maß für den Grad der Übereinstimmung zwischen der Vorhersage F und der Beobachtung y verwenden wir eine *scoring rule* [8]

$$S : \mathcal{F} \times \Omega \rightarrow \overline{\mathbb{R}},$$

wobei $\overline{\mathbb{R}} = [-\infty, \infty]$ die erweiterte reelle Gerade bezeichnet. Wir verwenden negativ gerichtete Scores, sodass kleinere Werte für bessere Vorhersagegüte stehen. Von besonderem Interesse sind *proper scoring rules*. Eine scoring rule heißt *proper* bezüglich der Klasse \mathcal{F} , falls

$$\mathbb{E}_F[S(F, Y)] \leq \mathbb{E}_F[S(G, Y)] \quad \text{für alle } F, G \in \mathcal{F}. \quad (1)$$

Derartige Scores schaffen also für eine Modelliererin, die glaubt, dass Y der Verteilung F folgt, einen Anreiz, auch diese Verteilung F als Vorhersage auszugeben. Wäre dem nicht so, würde der Score „unehrliche Vorhersagen“ ermutigen.

Ein Score mit vielen weiteren wünschenswerten theoretischen Eigenschaften ist der *logarithmische Score*

$$\log S(F, y) = -\log f(y),$$

wobei f die prädiktive Dichte- oder Zähldichtefunktion von F bezeichnet. Eine häufig verwendete Alternative ist der *continuous ranked probability score* (CRPS),

$$\text{CRPS}(F, y) = \int_{-\infty}^{\infty} (F(x) - \mathbf{1}\{x \geq y\})^2 dx,$$

wobei F hier die prädiktive Verteilungsfunktion und $\mathbf{1}$ die Indikatorfunktion bezeichnet. Sowohl der logarithmische Score als auch (unter endlichen ersten Momenten) der CRPS

sind sogar strikt proper, d. h. in (1) gilt Gleichheit genau dann, wenn $F = G$.

Der CRPS kann (wiederum unter der Annahme endlicher erster Momente) auch dargestellt werden als

$$\text{CRPS}(F, y) = \mathbb{E}_F |Y - y| - \frac{1}{2} \mathbb{E}_F |Y - Y'|,$$

wobei Y und Y' zwei unabhängige Ziehungen aus F sind. Diese Darstellung verdeutlicht, dass der CRPS eine direkte Verallgemeinerung des absoluten Fehlers ist: Legt F alle Wahrscheinlichkeitsmasse auf einen Punkt \hat{y} , so verschwindet der zweite Term und es verbleibt schlicht $|\hat{y} - y|$.

In unserem Anwendungsfall liegen keine vollständigen prädiktiven Verteilungsfunktionen vor, sondern lediglich 23 Quantile. Ein gängiger Score zur Evaluation des τ -Quantils q_τ der Vorhersageverteilung F ist der stückweise lineare Quantilsscore

$$\text{QS}_\tau(F, y) = 2 \times \{\mathbf{1}(y \leq q_\tau) - \tau\} \times (q_\tau - y),$$

der oft auch als *pinball loss* bezeichnet wird. Eine intuitivere Darstellung ergibt sich, wenn wir die Quantilvorhersagen zu den Niveaus $\frac{\alpha}{2} \in (0, \frac{1}{2})$ und $1 - \frac{\alpha}{2}$ gemeinsam betrachten, die ein zentrales Vorhersageintervall $[l, u]$ mit nominaler Überdeckung $1 - \alpha$ bilden. Dies führt zum Intervallscore,

$$\begin{aligned} \text{IS}_\alpha(F, y) &= (u - l) + \frac{2}{\alpha} (l - y) \mathbf{1}(y < l) + \frac{2}{\alpha} (y - u) \mathbf{1}(y > u) \\ &= \frac{1}{\alpha} (\text{QS}_{\alpha/2}(F, y) + \text{QS}_{1-(\alpha/2)}(F, y)), \end{aligned}$$

dessen drei Komponenten als Maß für die Streuung der Vorhersage und Strafterme für zu hoch oder zu niedrig angesetzte Vorhersagen interpretierbar sind.

Es kann gezeigt werden, dass

$$\text{CRPS}(F, y) = \int_0^1 \text{QS}_\tau(F, y) d\tau = \int_0^1 \frac{\text{IS}_\alpha(F, y)}{2\alpha} d\alpha,$$

sodass der CRPS in der Praxis approximiert werden kann als

$$\text{CRPS}(F, y) \approx \frac{1}{K} \sum_{k=1}^K \text{QS}_{\tau_k}(F, y),$$

wenn $\tau_1, \dots, \tau_K \in (0, 1)$ äquidistant über das Einheitsintervall verteilt sind. Dieser Score findet in allen Forecast Hub-Projekten Anwendung. Da er in [1] über den Intervallscore motiviert wird, wird er auch als *weighted interval score* (WIS) bezeichnet.

Evaluierung von Fall- und Todesfallvorhersagen während der zweiten Welle (Oktober bis Dezember 2020)

Um die Vorhersagen verschiedener Modelle systematisch mittels der beschriebenen Maße zu evaluieren, nahmen wir im Oktober 2020 die Präregistrierung einer prospektiven Studie vor [2]. Diese der klinischen Forschung entlehnte Vorgehensweise stellt durch detaillierte Festlegung des Evaluationsrahmens Transparenz und Unverfälschtheit der Analyse sicher. In [3] evaluierten wir Vorhersagen im Zeitraum

Oktober–Dezember 2020, der mit der zweiten epidemischen Welle in Deutschland und Polen zusammenfiel. Abbildung 2 zeigt Ein-Wochen-Vorhersagen für Fall- und Todesfallzahlen ausgewählter Einzelmodelle und des Medianensembles. Das Medianensemble ist eine pragmatische Aggregierungsform, bei der das α -Quantil jeweils durch den Median der α -Quantile der Mitgliedsmodelle gegeben ist.

Zu Beginn des Plateaus in den Fallzahlen Anfang November ist ein klarer *overshoot* der Vorhersagen zu beobachten, also ein Fortschreiben des vorherigen wachsenden Trends, der sich nicht in den Beobachtungen findet. Dies ist ein gängiges Muster: Umkehr- und Wendepunkte in den Fallzahlen waren, wie von vielen Beteiligten erwartet, sehr schwer vorherzusagen, und wurden nur in wenigen Ausnahmen von einzelnen Modellen korrekt antizipiert. Die Vorhersage von Todeszahlen stellte sich als deutlich weniger schwierig heraus, vor allem da vorangegangene Fall- und Hospitalisierungszahlen eine gute Einordnung erlauben.

Dies spiegelt sich auch in einer formalen Evaluierung mit dem *weighted interval score* wieder. Abbildung 3 zeigt den mittleren WIS, den ausgewählte Modelle für Fälle und Todesfälle bei verschiedenen Vorhersagehorizonten erzielten (für die in Abbildung 2 dargestellte Zeitspanne von Oktober bis Dezember 2020). Das untere Ende der grauen Fläche steht dabei für die Vorhersagegüte eines einfachen und im Voraus festgelegten Baseline-Modells. Dieses nutzt zur Vorhersage für alle Horizonte den letzten beobachteten Wert als Erwartungswert einer negativen Binomialverteilung, deren Streuungsparameter aus den letzten fünf Beobachtungen geschätzt wird [3]. Wie klar zu erkennen ist, gab dieses einfache Modell die erfolgreichsten Vier-Wochen-Vorhersagen für Fälle ab. Dies liegt im Wesentlichen daran, dass ein Großteil des untersuchten Zeitraums einem Plateau und damit exakt dem vom Baseline-Modell angenommenen Muster entspricht (und in der Tat schlugen fast alle eingereichten Modelle die Baseline in späteren Phasen). Nichtsdestotrotz unterstreichen diese Beobachtungen, dass Fallprognosen auf sehr kurze Vorhersagehorizonte beschränkt sein sollten. Bei Todesfällen hingegen übertrafen die meisten Modelle die Vorhersagegüte der Baseline, und die Ensemble-Vorhersage lieferte mit die besten Prognosen.

Ausblick

Die Vorhersage der Ausbreitung von Infektionskrankheiten ist ein junges Feld, und die Zuverlässigkeit der Prognosen reicht nicht an die etwa der modernen Wettervorhersage heran. Die Gründe hierfür liegen in der schwierigeren Datenlage, vor allem aber in der Tatsache, dass es sich bei einer Epidemie um ein soziales Phänomen handelt [13]. Menschliches Verhalten hat sich in den verschiedensten Bereichen als schwer zu prognostizieren erwiesen – und während der Pandemie kommt erschwerend hinzu, dass Vorhersagen zu *self-defeating prophecies* werden können. Gerade weil ein starker Anstieg vorhergesagt wurde, tritt dieser nicht ein, wenn ein Teil der Bevölkerung im Lichte derartiger Prognosen das Verhalten ändert. Zwar kann sicherlich keines

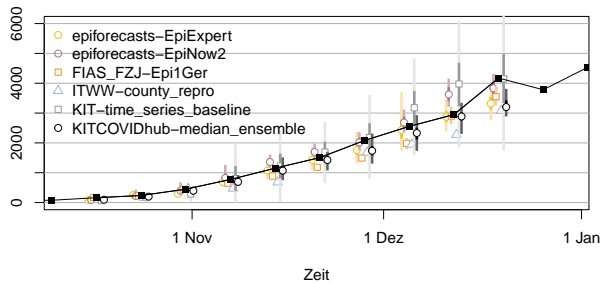
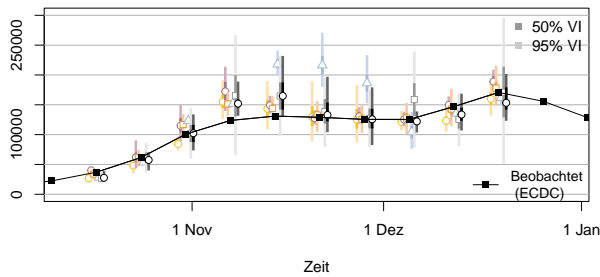


Abbildung 2. Ein-Wochen-Vorhersagen für wöchentliche Fall- (oben) und Todesfallzahlen (unten) in Deutschland im vierten Quartal 2020. Die Grafik zeigt prädiktive Mediane und zentrale 50 % sowie 95 % Vorhersageintervalle (VI) ausgewählter Einzelmodelle und des Medianensembles. Die betrachteten Modelle sind im unteren Bild angegeben und in [3] genauer beschrieben.

der hier betrachteten Modelle von sich behaupten, alleine einen derartigen Effekt zu haben – doch bewegen sich alle im Spannungsfeld von Erwartungen und Verhaltensanpassungen.

Neben diesen epidemiologischen Herausforderungen gibt es auch eine ganze Reihe offener mathematisch-statistischer Fragen im Zusammenhang mit Echtzeitvorhersagen in der Pandemie. Eine in diesem Beitrag nur am Rande angesprochene Frage ist, wie verschiedene Vorhersagen am besten in ein Ensemble kombiniert werden. Das hier erwähnte Medianensemble hat den Vorteil einer gewissen Robustheit – doch lässt es außer Acht, dass bestimmte Modelle in der Vergangenheit bessere Vorhersagen getroffen haben als andere. In unserer Arbeit [3] haben wir lediglich ein heuristisches Verfahren basierend auf inversen vergangenen WIS-Werten getestet, um solche Modelle stärker zu gewichten. Dieses brachte keine systematischen Verbesserungen gegenüber dem Medianensemble. Untersuchungen basierend auf dem größeren Vorhersagearchiv des US Forecast Hubs [14] zeigen, dass es auch mit fortgeschritteneren Quantilsregressionsverfahren nicht leicht ist, nennenswerte Verbesserungen zu erzielen. Wie Ensembles gewinnbringend datengesteuert angepasst werden können, ist Gegenstand unserer weiteren Forschung und der mehrerer anderer Gruppen. Auch im Hinblick auf die Evaluierung der Vorhersagen verbleiben zahlreiche methodische Probleme. So sind viele Verfahren zur Messung der Kalibrierung von Quantilvorhersagen auf stetige Zielgrößen ausgelegt – und bei diskretwertigen Zielgrößen wie in den Forecast Hubs können Artefakte entste-

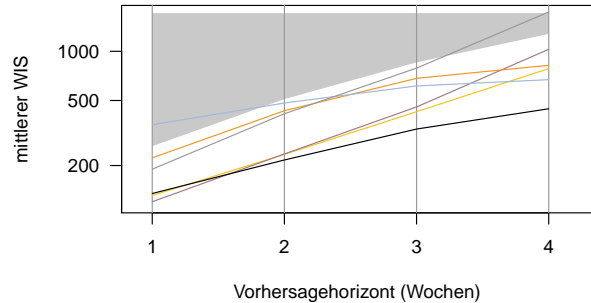
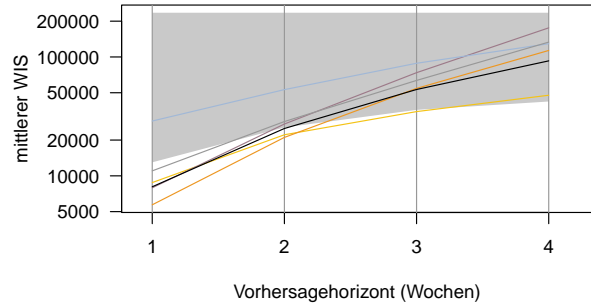


Abbildung 3. Mittlerer *weighted interval score* (WIS) verschiedener Modelle für Fall- (oben) und Todesfallzahlen (unten) in Deutschland im vierten Quartal 2020, nach Vorhersagehorizont in Wochen. Das untere Ende der grauen Fläche steht für die Performance eines einfachen Baseline-Modells. Dargestellt sind nur Modelle aus Abbildung 2, für vollständige Ergebnisse siehe [3].

hen, die selbst eigentlich perfekt kalibrierte Vorhersagen als fehlgeleitet erscheinen lassen.

Danksagung. Wir danken der Helmholtz-Gemeinschaft (Projekt SIMCARD) und der Klaus Tschira Stiftung für die Unterstützung unserer Arbeit.

Literatur

- [1] J. Bracher, E. L. Ray, T. Gneiting, and N. G. Reich. Evaluating epidemic forecasts in an interval format. *PLOS Computational Biology*, 17(2):e1008618, 2021.
- [2] J. Bracher, the German and Polish COVID-19 Forecast Hub Team, and Participants. Study protocol: Comparison and combination of real-time COVID19 forecasts in Germany and Poland. Deposited 8 October 2020, Registry of the Open Science Foundation, <https://osf.io/k8d39>, 2020.
- [3] J. Bracher, D. Wolfram, J. Deuschel, K. Görden, J.L. Ketterer, A. Ullrich, S. Abbott, M.V. Barbarossa, D. Bertsimas, S. Bhatia, M. Bodych, N.I. Bosse, J.P. Burgard, J. Fuhrmann, S. Funk, K. Gogolewski, Q. Gu, S. Heyder, T. Hotz, Y. Kheifetz, H. Kirsten, T. Krueger, E. Krymova, M.L. Li, J.H. Meinke, K. Niedzielewski, T. Ożański, F. Rakowski, M. Scholz, S. Soni, A. Srivastava, J. Zieliński, D. Zou, T. Gneiting, and M. Schienle. A pre-registered short-term forecasting study of COVID-19 in Germany and Poland during the second wave. *Nature Communications*, 12:5173, 2021.
- [4] Centers for Disease Control and Prevention. CDC ensemble report, 2021. www.cdc.gov/coronavirus/2019-ncov/science/forecasting/forecasting-us.html.
- [5] E. Y. Cramer, E. L. Ray, V. K. Lopez, J. Bracher, A. Brennan, A. J. Castro Rivadeneira, A. Gerding, T. Gneiting, K. H. House, Y. Huang, et al. Evaluation of individual and ensemble proba-

- bilistic forecasts of COVID-19 mortality in the US. medRxiv, 2021.
- [6] U. Dirnagl. Politikberatung, bis der Elefant mit dem Rüssel wackelt! *Labor Journal*, 2021. Published 10 May 2021, www.laborjournal.de/rubric/narr/narr/n_21_05.php.
- [7] T. Gneiting and A. E. Raftery. Weather forecasting with ensemble methods. *Science*, 310(5746):248–249, 2005.
- [8] T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.
- [9] IHME COVID-19 Forecasting Team. Modeling COVID-19 scenarios for the United States. *Nature Medicine*, 27:94–105, 2021.
- [10] N. Imai, I. Dorigatti, A. Cori, C. Donnelly, S. Riley, and N. M. Ferguson. Report 2: Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China, 2020. Imperial College London (22-01-2020), [Doi:10.25561/77150](https://doi.org/10.25561/77150).
- [11] N. P. Jewell, J. A. Lewnard, and B. L. Jewell. Caution warranted: using the Institute for Health Metrics and Evaluation model for predicting the course of the COVID-19 pandemic. *Annals of Internal Medicine*, 173(3):226–228, 2020.
- [12] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of Internal Medicine*, 172(9):577–582, 2020.
- [13] K. R. Moran, G. Fairchild, N. Generous, K. Hickmann, D. Osthus, R. Priedhorsky, J. Hyman, and S. Y. Del Valle. Epidemic forecasting is messier than weather forecasting: The role of human behavior and internet data streams in epidemic forecast. *Journal of Infectious Diseases*, 214(suppl_4):S404–S408, 2016.
- [14] E. L. Ray, L. C. Brooks, J. Bien, J. Bracher, A. Gerding, A. Rumack, M. Biggerstaff, M. A. Johansson, R. Tibshirani, and N. G. Reich. Challenges in training ensembles to forecast COVID-19 cases and deaths in the United States. Blog post, International Institute of Forecasters, forecasters.org/blog/2021/04/09/challenges-in-training-ensembles-to-forecast-covid-19-cases-and-deaths-in-the-united-states/, 2021.
- [15] E. L. Ray, N. Wattanachit, J. Niemi, A. Hannan Kanji, K. House, E. Y. Cramer, J. Bracher, A. Zheng, T. K. Yamana, X. Xiong, et al. Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the US. *MedRxiv*, 2020.
- [16] N. G. Reich and C. Rivers. Scientists want to predict COVID-19’s long-term trajectory. Here’s why they can’t. *Washington Post*, 2020. Published 15 September 2020, www.washingtonpost.com/outlook/2020/09/15/scientists-want-predict-covid-19s-long-term-trajectory-heres-why-they-cant/.
- [17] N. G. Reich, L. C. Brooks, S. J. Fox, S. Kandula, C. J. McGowan, E. Moore, D. Osthus, E. L. Ray, A. Tushar, T. K. Yamana, et al. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proceedings of the National Academy of Sciences*, 116(8):3146–3154, 2019.
- [18] J. Riou and C. L. Althaus. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance*, 25(4):2000058, 2020.

Dr. Johannes Bracher

Lehrstuhl für Statistik und Ökonometrie und KIT-Zentrum MathSEE,
 Karlsruher Institut für Technologie (KIT), Postfach 6980, 76049 Karlsruhe
 Heidelberger Institut für Theoretische Studien, Schloss-Wolfsbrunnengasse 35, 69118 Heidelberg
johannes.bracher@kit.edu

Daniel Wolffram

Lehrstuhl für Statistik und Ökonometrie und KIT-Zentrum MathSEE,
 Karlsruher Institut für Technologie (KIT), Postfach 6980, 76049 Karlsruhe
 Heidelberger Institut für Theoretische Studien, Schloss-Wolfsbrunnengasse 35, 69118 Heidelberg
 HIDS4Health – Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg
daniel.wolffram@kit.edu

Prof. Dr. Tilmann Gneiting

Heidelberger Institut für Theoretische Studien,
 Schloss-Wolfsbrunnengasse 35, 69120 Heidelberg
 Institut für Stochastik und KIT-Zentrum MathSEE, Karlsruher Institut für Technologie (KIT),
 Postfach 6980, 76131 Karlsruhe
tilmann.gneiting@h-its.org

Prof. Dr. Melanie Schienle

Lehrstuhl für Statistik und Ökonometrie und KIT-Zentrum MathSEE,
 Karlsruher Institut für Technologie (KIT), Postfach 6980, 76049 Karlsruhe
 Heidelberger Institut für Theoretische Studien, Schloss-Wolfsbrunnengasse 35, 69118 Heidelberg
melanie.schienle@kit.edu

Kurz nachdem die COVID-19-Pandemie Anfang 2020 Europa erreicht hatte, verteidigte Johannes Bracher seine Doktorarbeit in Biostatistik und Epidemiologie an der Universität Zürich, um im Anschluss eine Postdoc-Stelle in den Arbeitsgruppen von Melanie Schienle und Tilmann Gneiting am Karlsruher Institut für Technologie (KIT) und Heidelberger Institut für Theoretische Studien (HITS) anzutreten. Seither arbeiten wir gemeinsam mit Daniel Wolffram, der im Dezember 2020 am KIT eine Masterarbeit in Mathematik zu diesem Thema abschloss, und Kolleg(inn)en weltweit an Kurzzeitprognosen für COVID-19. Wir sind dankbar für die konstruktive Zusammenarbeit mit den Modellierungsteams, ohne deren grundlegende Arbeit und unermüdlichen Einsatz epidemiologische Vorhersagen in Echtzeit undenkbar wären.