STATE SPACE MODELS FOR REGIONAL EPIDEMIOLGICAL INDICATORS

STEFAN HEYDER

October 2023 – Draft v 0.1

ABSTRACT

Short summary of the contents in English…

ZUSAMMENFASSUNG

Kurze Zusammenfassung des Inhaltes in deutscher Sprache…

PUBLICATIONS AND CONTRIBUTIONS

This thesis consists of mostly unpublished work. During my time as a PhD student I have, however, been fortunate to collarborate with many scientists on problems in mathematical epidemiology with a focus on COVID-19, which resulted in several publications. In this section I want to clarify what my contributions to these publications were and which contributions of the present thesis are new.

[1] J. Bracher et al. "A Pre-Registered Short-Term Forecasting Study of COVID-19 in Germany and Poland during the Second Wave." In: *Nature Communications* 12.1 (1 Aug. 27, 2021), p. 5173. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25207-0. URL: https://www.nature.com/articles/s41467-021-25207-0 (visited on 09/30/2021).

[2] Johannes Bracher et al. "National and Subnational Short-Term Forecasting of COVID-19 in Germany and Poland during Early 2021." In: *Communications Medicine* 2.1 (1 Oct. 31, 2022), pp. 1–17. ISSN: 2730-664X. DOI: 10.1038/s43856-022-00191-8. URL: https://www.nature.com/articles/s43856-022-00191-8 (visited on 11/16/2022).

[3] Jan Pablo Burgard, Stefan Heyder, Thomas Hotz, and Tyll Krueger. "Regional Estimates of Reproduction Numbers with Application to COVID-19." Aug. 31, 2021. arXiv: 2108.13842 [stat]. URL: http://arxiv.org/abs/2108.13842 (visited on 09/30/2021).

[4] Sara M. Grundel, Stefan Heyder, Thomas Hotz, Tobias K. S. Ritschel, Philipp Sauerteig, and Karl Worthmann. "How to Coordinate Vaccination and Social Distancing to Mitigate SARS-CoV-2 Outbreaks." In: *SIAM Journal on Applied Dynamical Systems* 20.2 (Jan. 1, 2021), pp. 1135–1157. DOI: 10.1137/20M1387687. URL: https://epubs.siam.org/doi/abs/10.1137/20M1387687 (visited on 01/21/2022).

[5] Sara Grundel, Stefan Heyder, Thomas Hotz, Tobias K. S. Ritschel, Philipp Sauerteig, and Karl Worthmann. "How Much Testing and Social Distancing Is Required to Control COVID-19? Some Insight Based on an Age-Differentiated Compartmental Model." In: *SIAM Journal on Control and Optimization* 60.2 (Apr. 2022), S145–S169. ISSN: 0363-0129, 1095-7138. DOI: 10.1137/20M1377783. URL: https://epubs.siam.org/doi/10.1137/20M1377783 (visited on 11/16/2022).

[6] Thomas Hotz, Matthias Glock, Stefan Heyder, Sebastian Semper, Anne Böhle, and Alexander Krämer. "Monitoring the Spread of COVID-19 by Estimating Reproduction Numbers over Time." Apr. 18, 2020. arXiv: 2004.08557 [q-bio, stat]. URL: http://arxiv.org/abs/2004.08557 (visited on 07/20/2020).

[7] K. Sherratt et al. *Predictive Performance of Multi-Model Ensemble Forecasts of COVID-19 across European Nations*. June 16, 2022. DOI: 10.1101/2022.06.16.22276024. URL: http://medrxiv.org/lookup/doi/10.1101/2022.06.16.22276024 (visited on 11/28/2022). preprint.

Du musst bereit sein Dinge zu tun.

— A meme on the internet, 2022.

ACKNOWLEDGMENTS

Put your acknowledgments here.

CONTENTS

1 INTRODUCTION

2 EPIDEMIOLOGICAL CONSIDERATIONS

- COVID-19 induced unprecedented interest in epidemiological modelling from all disciplines, but also mathematics
- this chapter highlights challenges that epi modelling brings and what desirable outcomes would be from an applied perspective
- mathematical epidemiology concerns itself with modelling epidemiolocial systems, from small (local outbreaks) to large (epi/-pandemics)
- conclusions from analysis only as good as the model and the data are
- dependening on goal and circumstances different methods are applicable
- by its nature, data are observational so causal claims difficult
- in this thesis I focus on models for larger-scale epidemics, techniques would be flexible enough to deal with smaller scale as well, as long as latent states are gaussian

2.1 OBJECTIVES OF EPIDEMIOLOGICAL MODELLING

MONITORING

- monitoring is real-time scenario, interested in current developments, i.e. recent past and near future. complicated by potentially slow reporting, data revisions
- informs decision makers on whether measures should be taken
- ForecastHub(s) provide platform that creates ensemble forecast to obtain better predictions [2, 3, 11, 12]

RETROSPECTIVE ANALYSIS

- evaluation of measures taken, want interpretation as causal as possible
- informs decision makes on which measures were effective and how much
- difficult due to usual reasons: poor data quality, observational data, causual structure difficult, early/late adoption makes timing of measurements difficult
- cite some papers that did this [4, 7, 10]

SCENARIO MODELLING

- concerns itself with modelling the impact that variants, seasonality etc. have in specific scenarios
- find out whether there is already paper of ECDC to cite

2.2 AVAILABLE DATA AND ITS QUALITY

- surprising amount of data available, but quality questionable,
- in Germany have data on reported cases and deaths by gender, age group, county, with reporting date of case and for some cases even date of symptom onset
- reporting of cases is regulated by Infektionsschutzgesetz
- parallel dataset for reports of hospitalisations
- have description section from Nowcasting draft here
- descriptive statistics of German COVID-19 data set
- even larger datasets that compile this for europe + EFTA (?) by ECDC or by world (JHU)
- quality of reported case data is potentially too low
  - reporting delays
  - weeakday effects
  - testing regime changing (2G/3G)
  - ...
- data on commuting

2.3 MEASURES OF EPIDEMIC SPREAD

This section consists of the ideas published in [9], but has been rewritten to fit better into this thesis.

- not only epidemic spread but also speed of proliferation is of interest, enables forecasts
- measuring speed difficult: data problems ... (look at AK book article)
-

2.3.1 *Growth Factor*

2.3.2 *Reproduction number*

2.3.3 *Other indicators*

2.3.4 *Usefulness of indicators*

2.4 DESSIDERATA FOR EPIDEMIOLOGICAL MODELS

- we want models to be able to include as much data as possible, while still being numerically tractable

REGIONAL DEPENDENCIES AND EFFECTS

- German case data are reported on Landkreis level, performing analysis of each individual is not sensible
- inhabitants travel between regions, and measures were taken on on regional level as well
- effects are not really spatial: euclidean distance is not so much of an issue but how closely connected regions are (give some examples)
- also want to account for other regional effects such as different socio-economic settings ...

TEMPORAL CORRELATION

INTERPRETABILITY

3 STATE SPACE MODELS

State space models are a versatile class of statistical models which allow to model non-stationary time series data and come along with straight-forward interpretation. The main idea of these models is to introduce unobserved **latent states** whose joint distribution is given by a Markov process and model the observed time series conditional on theses states. By exploiting this structure, inference in state space models becomes computationally efficient, i.e. the complexity of algorithms is linear with respect to the number n of time points considered.

An additional advantage, that will become more explicit in Section 3.1, is that state space models allow to interpret the modeled dynamics of latent states which makes

Definition 3.1 (State Space Model)**.** A **state space model** is a discrete time stochastic process $(X_t, Y_t)_{t=0,\dots,n-1}$ taking values in a measurable space $\mathcal{X} \times \mathcal{Y}$ such that

1. The marginal distribution of the **states** $(X_0, \dots, X_{n-1})$ is a discrete time Markov process, i.e. for $t = 1, \dots, n-1$

$$\mathbf{P}\left(X_t \in B | X_0, \dots, X_{t-1}\right) = \mathbf{P}\left(X_t \in B | X_{t-1}\right) \tag{3.1}$$

for all measurable $B \subseteq \mathcal{X}$.

2. Conditional on the state X_t and observation Y_{t-1}, Y_t is independent of X_s and Y_{s-1}, $s < t$, i.e.

$$\mathbf{P}\left(Y_t \in B | X_0, \dots, X_t, Y_0, \dots, Y_{t-1}\right) = \mathbf{P}\left(Y_t \in B | X_t, Y_{t-1}\right)$$

for all measurable $B \subseteq \mathcal{Y}$.

For notational convenience we will write $X_{s:t} = (X_s, \dots, X_t)$ for the vector that contains all states from s to t, dropping the index if we consider the whole set of observations, so $X = X_{0:n-1}$ Similarly we set $Y_{s:t} = (Y_s, \dots, Y_t)$ and $Y = Y_{0:n-1}$.

picutre of dependency structure

Remark. Contrary to the standard definition of a state space model, our Definition 3.1 allows Y_t to depend on Y_{t-1}. This is not a limitation of the standard definition: given a state space model of the form in Definition 3.1 we can transform it to the standard form by choosing states $(X_t, Y_t) \in \mathcal{X} \times \mathcal{Y}$ and observations $Y_t \in \mathcal{Y}$ such that the state space model becomes a stochastic process on $(\mathcal{X} \times \mathcal{Y}) \times \mathcal{Y}$.

Additionally most computations and inferences in this thesis will condition on the observations Y and as such Y may be treated as fixed. The only exception to this is in simulation studies where we sample from the joint distribution of (X, Y).

As the models considered in Chapter 4 will make extensive use of state space models with this dependency structure we opt to use this non-standard definition here.

In most models we consider in this thesis we use $\mathcal{X} = \mathbf{R}^m$, $\mathcal{Y} = \mathbf{R}^p$ or $\mathcal{Y} = \mathbf{Z}^p$ so that $\mathcal{X}$ is m dimensional and $\mathcal{Y}$ is p dimensional and equip these spaces with the usual Borel σ-Algebras.

check whether there are models that violate this

Most models that I consider in this thesis will admit densities for the state transitions w.r.t. a common dominating measure $\mu_{\mathcal{X}}$ and similar for the observations w.r.t. a (potentially different) domination measure $\mu_{\mathcal{Y}}$.

Notation (Densities, conditional densities). I will use the standard abuse of notation for densities that makes the type of density „obvious" from the arguments used. This means that $p(x)$ is the density for all states X, $p(x_t|x_{t-1})$ the conditional density of $X_t|X_{t-1}$ and similarly for observations: $p(y|x)$ is the density of all observations Y conditional on all states X.

Note that this notation also implicitly includes the time t and allows for changes in, e. g., the state transition over time.

When densities stem from a parametric model parametrized by $\theta \in \Theta \subseteq \mathbf{R}^k$ and the dependence of the model on θ is of interest, i.e. because we try to estimate θ, we indicate this by adding a subscript to the densities. If the dependence is not of interest, e.g. because θ is fixed, I will usually omit θ for better readability.

In this notation, the joint density of a parametric state space model factorizes as

$$\begin{aligned} p_\theta(x,y) &= p_\theta(x_0,\dots,x_{n-1},y_0,\dots,y_{n-1}) \\ &= p_\theta(x_0)\prod_{t=1}^{n-1} p_\theta(x_t|x_{t-1})\prod_{t=0}^{n-1} p_\theta(y_t|x_t,y_{t-1}), \end{aligned}$$

where $p_\theta(y_0|x_0,y_{-1}) = p_\theta(y_0,x_0)$.

As inferences we make in this thesis depend on the state space model only through the likelihood we identify almost sure versions of (X,Y) with itself, i.e. all equations involving X or Y are understood almost surely.

Given data $(y_t)_{t=0,\dots,n-1}$ that may be modeled with a state space model the practitioner is confronted with several tasks, which provide the structure of this chapter:

1. Choosing a suitable, usually parametric, class of state space models that include the effects of interest.

2. Fitting such a parametric model to the data at hand by either frequentist or Bayesian techniques.

3. Infer about the latent states X from the observations Y by determining, either analytically or through simulation, the smoothing distribution $X|Y$.

The first step, item 1, requires that the practitioner specifies a joint probability distribution for the states and observations (Section 3.1). Due to the assumed dependency structure this boils down to specifying transition kernels for the states and observations. The setting Definition 3.1 is too abstract to perform inference in so further assumptions on the types of distributions for the latent states and observations are needed. In this chapter we will discuss linear gaussian state space models (Section 3.2), where both the posterior distribution and the likelihood are analytically available. For the epidemiological application we have in mind these are however insufficient due to the non-linear behaviour of incidences and the low count per region (Section 2.4). Such observations are better modeled with distributions on the natural numbers, i.e. with a Poisson or negative binomial distribution, leading to the class of logconcave Gaussian state space models (Section 3.3).

Regarding the second step, item 2, a frequentist practitioner will want to perform maximum likelihood inference on θ. While asymptotic confidence intervals for θ can be derived both theoretically and practically [6, Chapter 7], they are, in the context of this thesis, usually of little interest. We choose to view this fitting as an Empirical Bayes procedure and our main practial interest lies in analyzing the posterior distribution $X|Y$.

To obtain the maximum likelihood estimates $\hat\theta$ one needs access to the likelihood

$$p(y) = \int_{\mathcal{X}^n} p(x,y) \mathrm{d}\, x, \tag{3.2}$$

which is usually not analytically available. Direct numerical evaluation of Equation (3.2) is hopeless due to the high dimensionality of the state space $\mathcal{X}^n$. Instead we will resort to simulation based inference by importance sampling (see Section 3.4), an alternative would be particle filters [5].

The performance of these simulations depends crucially on constructing distributions that are close to the posterior $p(x|y)$ but are easy to sample from. To this end, we construct suitable Gaussian state space models (Section 3.5) in which sampling from the posterior is analytically possible. This will be a good strategy if the target posterior $p(x|y)$ can be well approximated by a Gaussian distribution — otherwise, we may want to account for multiple modes by considering mixtures of Gaussian state space models or account for heavy tails with t-distributed errors (Section 3.6).

3.1 MODELLING EPIDEMIOLOGICAL DESSIDERATA WITH STATE SPACE MODELS

3.2 LINEAR GAUSSIAN STATE SPACE MODELS

- joint model is gaussian
- filtering distribution obtained by Kalman filter
- smoothing distribution obtained by Kalman smoother
- variants: sqrt filter / precision filter
- gaussian likelihood analytically available, MLE can be found by numerical methods (gradient descent or EM, depending on problem)
- computation is efficient: linear in time dimension n
- Y_{t+1} may also depend on Y_t as we will target the conditional distribution anyways

Gaussian linear state space model (GLSSM) are the working horses of most methods used in this thesis because they are analytically tractable and computationally efficient. Indeed for fixed dimension of states m and observations p the runtime of algorithms that we consider in this thesis is $\mathcal{O}(n)$.

Definition 3.2 (Gaussian linear state space model (GLSSM))**.** A GLSSM is a state space model where states obey the transition equation

$$X_{t+1} = A_t X_t + u_t \varepsilon_{t+1} \qquad t = 0, \dots, n-1, \tag{3.3}$$

and observations obey the observation equation

$$Y_t = B_t X_t + v_t + \eta_t \qquad t = 0, \dots, n. \tag{3.4}$$

Here $A_t \in \mathbf{R}^{m \times m}$ and $B_t \in \mathbf{R}^{p \times m}$ are matrices the specify the systems dynamics. The **innovations** ε_{t+1} and **measurement noise** η_t are independent from one another and from the starting value $X_0 \mathcal{N}(\mathbf{E}X_0, \Sigma_0)$.

Furthermore, $\varepsilon_{t+1} \sim \mathcal{N}(0, \Sigma_t)$ and $\eta_t \sim \mathcal{N}(0, \Omega)$ are centered Gaussian random variables and $u_{t+1}, t = 1, \dots, n-1$, $v_t, t = 0, \dots, n$ are deterministic biases.

The defining feature of a GLSSM is that its joint distribution is Gaussian.

Lemma 3.1. *A state space model can be written as a GLSSM if and only if its joint distribution is Gaussian.*

technicailty: distringuish between a.s. version

Proof. □

As the joint distribution of (X, Y) is Gaussian, so are conditional distributions of states given observations. Two such distributions are of interest: the **filtering distribution** is the conditional distribution of X_t given all observations until time t, that is $Y_{0:t}$. When $t < n$ this is distinct from the **smoothing distribution**, i.e. the distribution of X_t given all observations Y.

Note that the filtering distributions does not specify a valid joint distribution for the states, but the smoothing does.

Both distributions may be obtained efficiently using the celebrated Kalman filter and smoother algorithms .

cite correclty

Algorithm 1: Kalman filter

Input: observations $y = (y_0, \dots, y_n)$, GLSSM
Output: filtered expectations $\hat{X}_{t|t}$, covariance matrices $\Xi_{t|t}$, likelihood $p(y)$
Initialization;
$\hat{y}_{0|-1} = B_0 \hat{x}_{0|-1} + v_t$;
$\Psi_{0|-1} = B_0 \Sigma_0 B_0^T + \Omega_0$;
Prediction;
Filter;

Notice that the Kalman filter calculates the likelihood $p(y)$ while filtering — this is possible because of the dependency structure of the state space model — this makes inference via maximum likelihood possible in GLSSMs.

The Kalman smoother computes the marginal distributions $X_t|Y$ for $t = 0, \dots, n-1$ and, owing to the Markov structure of the states, these are enough to specify the joint distribution $X|Y$, allowing to simulate from it.

Algorithm 2: Kalman smoother

Input: observations $y = (y_0, \dots, y_n)$, GLSSM
Output: filtered expectations $\hat{X}_{t|t}$ and covariance matrices $\Xi_{t|t}$
Initialization;
$\hat{y}_{0|-1} = B_0 \hat{x}_{0|-1} + v_t$;
$\Psi_{0|-1} = B_0 \Sigma_0 B_0^T + \Omega_0$;
Prediction;
Filter;

Algorithm 3: Forwards filter, backwards smoother [8, Proposition 1]

Input: TODO
Output: TODO
Something;

The modeling capacity of GLSSM is, however, limited: most interesting phenomena follow neither linear dynamics nor are well modeled by a Gaussian distribution. Nevertheless, linearization of non-linear dynamics suggests that GLSSMs may have some use as approximations to these more complicated phenomena, provided they are sufficiently close to Gaussian models, e.g. unimodal and without heavy tails. We start to move away from linear Gaussian models by allowing observations that are non-Gaussian.

3.3 LOGCONCAVE GAUSSIAN STATE SPACE MODELS

- replace gaussian observations with log concave observations
- motivation for logconcave distributions: posterior has unique mode, because up to constants $\log p(x|y) = -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) + \log p(y|x)$ so $\log p(x|y)$ is concave
- not restricted to same type of distribution per time step (though in ISSSM it will be)
- Laplace approximation sensible for these types of models: single mode
- special case: exponential family distributions

3.4 IMPORTANCE SAMPLING

Suppose we have a function $h : \mathcal{X} \to \mathbf{R}$ whose integral

$$\zeta = \int_{\mathcal{X}} h(x) \mathrm{d}\, x$$

we want to compute. Furthermore suppose that we can write

$$\int_{\mathcal{X}} h(x) \mathrm{d}\, x = \int_{\mathcal{X}} f(x) \mathrm{d}\, \mathbf{P}(x)$$

for a probability measure $\mathbf{P}$ and function $f : \mathcal{X} \to \mathbf{R}$. Let $\mathbf{G}$ be another measure on $\mathcal{X}$ such that $f\mathbf{P}$ is absolutely continuous with respect to

$\mathbf{G}$, $f\mathbf{P} \ll \mathbf{G}$ and let $v = \frac{\mathrm{d} f\mathbf{P}}{\mathrm{d}\,\mathbf{G}}$ be the corresponding Radon-Nikodym derivative. Then

$$\zeta = \int_{\mathcal{X}} h(x) \mathrm{d}\, x = \int_{\mathcal{X}} f(x) \mathrm{d}\, \mathbf{P}(x) = \int_{\mathcal{X}} v(x) \mathrm{d}\, \mathbf{G}(x)$$

which suggests to estimate ζ by Monte-Carlo integration:

$$\hat{\zeta} = \frac{1}{N} \sum_{i=1}^{N} v(X_i)$$

for $X_i \overset{\text{i.i.id}}{\sim} \mathbf{G}$, $i = 1, \ldots, N$.

If one is not interested in a particular h but rather in an approximation of $\mathbf{P}$ and $\mathbf{P}$ is absolutely continuous with respect to $\mathbf{G}$, then one may view

$$\hat{\mathbf{P}}_N = \frac{1}{N} \sum_{i=1}^{N} v(X_i) \delta_{X_i}$$

as a particle approximation of $\mathbf{P}$. In this setting [1] shows that the random measure $\hat{\mathbf{P}}_N$ converges to $\mathbf{P}$ at rate $\mathcal{O}\left(\frac{1}{N}\right)$ in an appropriate metric.

check

To perform importance sampling one must be able to evaluate the weights v. In a bayesian setting this is usually infeasible: if $\mathbf{P}$ is a posterior then the integration constant of its density is intractable. In this case one can usually evaluate the weights up to a constant, i.e. $w(x) \propto_x \frac{\mathrm{d}\,\mathbf{P}}{\mathrm{d}\,\mathbf{G}}(x)$ is available. The missing constant is then $\int w(x) \mathrm{d}\, \mathbf{G}$ which is itself amenable to importance sampling.

This leads to the self-normalized importance sampling weights $W_i = \frac{w(X_i)}{\sum_{i=1}^N w(X_i)}$ and Monte Carlo estimates $\hat{\zeta} = \sum_{i=1}^N W_i f(X_i)$ and particle approximation $\hat{\mathbf{P}}_N = \sum_{i=1}^N W_i \delta_{X_i}$.

In both cases one can show that once second moments of w with respect to $\mathbf{G}$ exist the Monte-Carlo estimates are consistent and asymptotically normal at the usual rates, see [5, Chapter 8].

Importance sampling is useful in situations where simulation from $\mathbf{P}$ is not feasible or when Monte Carlo integration with respect to $\mathbf{P}$ is unattractive due to high variance estimates.

- importance sampling as a variance reduction technique
- importance sampling as a technique to make intractable distributions tractable
- importance sampling vs. other methods:
  - vs. ABC
  - vs. MCMC
  - vs. INLA (isn't this MCMC?)
- measuring how good IS performs: ESS and other measures
- related results regarding performance of IS (Chatterje, Agapiou)

3.4.1 *Laplace approximation*

- approximate at mode, problematic if posterior is not unimodal (but then gaussian approximation probably not worth it)

3.4.2 *Cross entropy method*

-

3.4.3 *Efficient importance sampling*

3.5 GAUSSIAN IMPORTANCE SAMPLING FOR STATE SPACE MODELS

As the likelihood of a general state space model is neither analytically nor numerically tractable one has to resort to Monte-Carlo techniques. Recall that the likelihood is a high-dimensional integral of the form

$$\ell(\theta) = p_\theta(y) = \int p_\theta(y,x) \mathrm{d}\, x = \int p_\theta(y|x)p_\theta(x)\mathrm{d}\, x = \mathbf{E}p_\theta(y|X).$$

By the standard law of large numbers we can approximate $\ell(\theta)$ by

$$\hat{\ell}(\theta) = \frac{1}{N}\sum_{i=1}^{N} p_\theta(y|X^i)$$

for $N \in \mathbf{N}$ samples $X^i \overset{\text{i.i.id}}{\sim} p(x)$. However, the variance of $\hat{\ell}(\theta)$ is likely to be very high if samples X^i are drawn from the prior distribution $p(x)$ as they are not informed by the observations y. As $p_\theta(x|y) \propto p_\theta(x,y)$ a more promising approach would be to use samples $X^i \sim p_\theta(x|y)$, but this distribution is usually not available.

While bayesian computational approaches such as MCMC are able to generate (approximate) samples from this posterior distribution, importance sampling tries to find a distribution close to the target and re-weighs samples to ensure unbiased estimates of $\ell(\theta)$.

cite

3.6 ACCOUTING FOR MULTIMODALITY AND HEAVY TAILS

3.7 MAXIMUM LIKELIHOOD ESTIMATION

4 ANALYSIS OF SELECTED MODELS

4.1 SPATIAL REPRODUCTION NUMBER MODEL

1. essentially the Regional model presented in ECMI

4.2 REGIONAL GROWTH FACTOR MODEL

4.3 NOWCASTING HOSPITALIZATIONS

5 DISCUSSION

6

CONCLUSION (?)

A

IMPLEMENTATION IN PYTHON

B

PROOFS (?)

[1] S. Agapiou, O. Papaspiliopoulos, D. Sanz-Alonso, and A. M. Stuart. *Importance Sampling: Intrinsic Dimension and Computational Cost*. Jan. 14, 2017. DOI: 10.48550/arXiv.1511.06196. arXiv: 1511.06196 [stat]. URL: http://arxiv.org/abs/1511.06196 (visited on 04/03/2023). preprint.

[2] J. Bracher et al. "A Pre-Registered Short-Term Forecasting Study of COVID-19 in Germany and Poland during the Second Wave." In: *Nature Communications* 12.1 (1 Aug. 27, 2021), p. 5173. ISSN: 2041-1723. DOI: 10.1038/s41467-021-25207-0. URL: https://www.nature.com/articles/s41467-021-25207-0 (visited on 09/30/2021).

[3] Johannes Bracher et al. "National and Subnational Short-Term Forecasting of COVID-19 in Germany and Poland during Early 2021." In: *Communications Medicine* 2.1 (1 Oct. 31, 2022), pp. 1–17. ISSN: 2730-664X. DOI: 10.1038/s43856-022-00191-8. URL: https://www.nature.com/articles/s43856-022-00191-8 (visited on 11/16/2022).

[4] Jan M. Brauner et al. "Inferring the Effectiveness of Government Interventions against COVID-19." In: *Science* 371.6531 (Feb. 19, 2021), eabd9338. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.abd9338. URL: https://www.science.org/doi/10.1126/science.abd9338 (visited on 07/06/2023).

[5] Nicolas Chopin and Omiros Papaspiliopoulos. *An Introduction to Sequential Monte Carlo*. Springer Series in Statistics. Cham, Switzerland: Springer, 2020. 378 pp. ISBN: 978-3-030-47844-5.

[6] J. Durbin and S. J. Koopman. *Time Series Analysis by State Space Methods*. 2nd ed. Oxford Statistical Science Series 38. Oxford: Oxford University Press, 2012. 346 pp. ISBN: 978-0-19-964117-8.

[7] Seth Flaxman et al. "Estimating the Effects of Non-Pharmaceutical Interventions on COVID-19 in Europe." In: *Nature* 584.7820 (Aug. 2020), pp. 257–261. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2405-7. pmid: 32512579. URL: https://www.nature.com/articles/s41586-020-2405-7 (visited on 08/28/2020).

[8] Sylvia Frühwirth-Schnatter. "Data Augmentation and Dynamic Linear Models." In: *Journal of Time Series Analysis* 15.2 (1994), pp. 183–202. ISSN: 1467-9892. DOI: 10.1111/j.1467-9892.1994.tb00184.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9892.1994.tb00184.x (visited on 06/08/2022).

[9] Stefan Heyder and Thomas Hotz. "Measures of COVID-19 Spread." In: *Covid-19 pandisziplinär und international: Gesundheitswissenschaftliche, gesellschaftspolitische und philosophische Hintergründe*. Ed. by Alexander Kraemer and Michael Medzech. Medizin, Kultur, Gesellschaft. Wiesbaden: Springer Fachmedien, 2023, pp. 51–66. ISBN: 978-3-658-40525-0. DOI: 10.1007/978-3-

658-40525-0_3. URL: https://doi.org/10.1007/978-3-658-40525-0_3 (visited on 10/21/2023).

[10] Yeganeh Khazaei, Helmut Küchenhoff, Sabine Hoffmann, Diella Syliqi, and Raphael Rehms. "Using a Bayesian Hierarchical Approach to Study the Association between Non-Pharmaceutical Interventions and the Spread of Covid-19 in Germany." In: *Sci Rep* 13.1 (Nov. 2, 2023), p. 18900. ISSN: 2045-2322. DOI: 10.1038/s41598-023-45950-2. URL: https://www.nature.com/articles/s41598-023-45950-2 (visited on 11/10/2023).

[11] Evan L. Ray et al. "Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S." In: *medRxiv* (Aug. 22, 2020), p. 2020.08.19.20177493. DOI: 10.1101/2020.08.19.20177493. URL: https://www.medrxiv.org/content/10.1101/2020.08.19.20177493v1 (visited on 09/02/2020).

[12] K. Sherratt et al. *Predictive Performance of Multi-Model Ensemble Forecasts of COVID-19 across European Nations*. June 16, 2022. DOI: 10.1101/2022.06.16.22276024. URL: http://medrxiv.org/lookup/doi/10.1101/2022.06.16.22276024 (visited on 11/28/2022). preprint.

DECLARATION

Put your declaration here.

Ilmenau, October 2023

Stefan Heyder