

TEMP

TIDSSKRIFT FOR HISTORIE

EPIDEMI HISTORIE



HVORDAN MATEMATISK MODELLERING KAN SUPPLERE HISTORIEN

Mæslinger på Fyn i 1874

Rasmus Kristoffer Pedersen

I den historiske epidemiologi benyttes et utal af kvantitative metoder til at undersøge historisk datamateriale fra grundlæggende statistik til generaliserede lineære modeller og Machine Learning. Et velkendt eksempel er beregning af overdødelighed som følge af en epidemi; altså estimerer for, hvor mange dødsfald der fandt sted under en given epidemi i forhold til, hvad man ellers ville have forventet.¹ Denne form for beregning går langt tilbage og ses ligeledes også i moderne historisk litteratur.²

I denne artikel beskrives en bestemt kategori af kvantitative metoder, *mekanis-mebaseret matematisk modellering*, med det formål at inspirere fremtidige samarbejder mellem historieforskningen og matematisk modellering.

Matematisk modellering har tidligere bidraget til den historiske epidemiologi. Som eksempler herpå er blandt andet beskrivelser af, hvordan mæslinger spredtes i og imellem engelske byer i midten af 1900-tallet, menneskelige ektoparasitters betydning for Den Sorte Død i Europa, og hvordan den københavnske koleraepidemi i 1853 spredtes i byen.³ Disse eksempler fokuserer på den epidemiologiske indsigt, som matematisk modellering kan generere fra et historisk kildegrundlag, mens det historiefaglige primært assisterer i analysen og bidrager med kontekst.

Samarbejde mellem matematisk modellering og historie, hvori fokus er på historiefaglige spørgsmål, er også mulige, men det ses sjældnere. Dette er, i denne forfatters optik, en skam, da der er visse overlap mellem, hvordan de to fag anskuer erkendelse: Ligesom beretninger fra historiske kilder ikke kan læses som en umiddelbar beskrivelse af fortiden, men skal fortolkes i en historisk kontekst, så er det for en matematisk modellør også klart, at en model ikke *er* virkeligheden, men blot en abstraktion af denne og derfor skal fortolkes meget nøje, hvis den skal benyttes til at lave slutninger om virkelige forhold.

Denne artikel giver et eksempel på, hvordan et muligt samarbejde om histo-

1 Andreasen m.fl.: "Epidemiologic Characterization of the 1918 Influenza", 270-78.

2 Se eksempelvis Bremer og Fenger: "Om Influenza-Epidemierne", 125-38. For en gennemgang af den medicinske statistiks indtog i Danmark refereres til Skydsgaard: "It's probably in the air", 215-36. Moderne eksempler på overdødelighedsberegning nævnes eksempelvis Hinde: "A review of methods", 82-92 samt Connor og Hinde: "Mortality in town and countryside", 54-67.

3 Bjørnstad m.fl.: "Dynamics of Measles Epidemics", 185-202. Dean m.fl.: "Human Ectoparasites", 1304-9, Phelps m.fl.: "Cholera Epidemics of the Past", 641-49. Yderlig diskussion om pest, herunder også af tværfaglige bidrag, gives af Walløe: "Medieval and Modern Plague", 59-73.

riefaglige spørgsmål kunne se ud set fra den matematiske modellerings perspektiv. For at demonstrere, hvordan modellering kan bidrage, tages der her udgangspunkt i en serie af mæslingeepidemier i 1874, som er dokumenteret i medicinalberetningerne for Fyn, Langeland og Ærø i henholdsvis 1873, 1874 og 1875. Detaljerne om kildegruppen beskrives nærmere senere i artiklen. Først defineres mere nøjagtigt, hvad der her menes med *matematisk modellering*.

MATEMATISK MODELLERING

I naturvidenskaben dækker *modellering* over en bred vifte af relaterede felter og metoder. Grundlæggende handler modellering om, at viden om ét område ekstrapoleres til et andet. Typisk tages udgangspunkt i et modelsystem, som man kan manipulere og ændre på, i modsætning til det system, det er en model for. Som eksempel herpå kunne en model for en cancertumors vækst i et menneske være en lignende tumors vækst i en mus. Tilsvarende er en matematisk model typisk noget, man kan regne på, hvori forhåbningen er, at de matematiske årsagssammenhænge, der gælder for modellen, også gælder for virkeligheden. Viser en model af sygdomsspredning for eksempel, at forlængede sygdomsforløb fører til flere smittede individer under en epidemi, så kan det indikere, at det sandsynligvis også er tilfældet for virkelig sygdomsspredning.

Modeller kan have meget forskellige formuleringer og blive til gennem forskellige overvejelser og antagelser. Her diskuteres en bestemt type, nemlig *mekanisme-baserede* matematiske modeller. Sådanne modeller tager f.eks. udgangspunkt i bestemte biologiske eller sociale mekanismer, som man mener gælder i virkeligheden, så modellens bestanddele hver især repræsenterer noget virkeligt. Et eksempel kunne være et tal, der repræsenterer antallet af cancerceller i en tumor. I en model for smittespredning kunne et tal f.eks. repræsentere antallet af nære kontakter, man typisk har dagligt.

Begrænsningen til mekanismer, der kan beskrives matematisk, adskiller denne type modellering fra statistiske metoder som generaliserede lineære modeller eller difference-in-difference modeller. Hvor statistiske modeller typisk søger at kvantificere potentielle årsagssammenhænge mellem forskellige målte størrelser, så tager mekanisme-baserede modeller udgangspunkt i konkrete sammenhænge på forhånd ofte med det mål, at modellen i sidste ende stemmer overens med observationer fra virkeligheden. Man undersøger således ikke en bred vifte af potentielle sammenhænge, men fokuserer i stedet på spørgsmål om konkrete mekanismer ved at benytte modellen som et *matematisk mikroskop*.⁴ Et eksempel kunne være spørgsmål såsom effekten af bestemte samfundsrestriktioner på begrænsning af sygdomsspredning. Baseret på forhåndsantagelser om, hvordan restriktioner ændrer bl.a. antallet af nære kontakter, håber modelløren at lære noget om de restriktioner, der om ikke andet gælder for en idealiseret version af verden, der følger de antagelser, som ligger til grund for modellen.

SIR-MODELLEN

Epidemiske sygdomme har længe været et vigtigt omdrejningspunkt for matematisk

4 Ottesen: "The Mathematical Microscope", 97-118.

modellering som beskrevet i betydningsfulde udgivelser af blandt andre William Kermack og A.G. McKendrick samt Roy M. Anderson og Robert M. May.⁵ En matematisk model, der optræder i disse udgivelser samt i meget efterfølgende arbejde, er den såkaldte SIR-model. Modellens navn refererer til de tre grupperinger, som en befolkning opdeles i: De modtagelige (*Susceptible*, S), de infektiøse eller smitsomme (*Infectious*, I) og de, der er kommet sig oven på infektionen og nu er immune (*Recovered*, R). Modellen er illustreret i figur 1.

Modellen indeholder to grundlæggende mekanismer: At infektiøse individer kan smitte modtagelige, og at infektiøse kommer sig efter en given periode og opnår livslang immunitet. Dette er tilstrækkeligt til at frembringe en udvikling i antallet af smittetilfælde, der kvalitativt minder om det, der ses fra epidemisk data.⁶

Modellens mekanismer baseres på sygdomsbiologi (f.eks. infektionsvarighed) og samfundet, hvori sygdommen spreder sig (f.eks. sandsynligheden for at infektiøse og modtagelige individer mødes). Modelundersøgelser kan derfor fokuseres forskelligt afhængigt af, hvad der allerede er kendt. Er en sygdom velbeskrevet biologisk, kan modellen bruges til at undersøge samfundet, hvori sygdommen optræder.

Det er også muligt at udvide modellen og tilføje yderligere mekanismer. Dette kunne være tab af immunitet over tid eller et ekstra stadie, der håndterer en *latensperiode* mellem eksponering til sygdommen, før man selv er smitsom. Denne sidstnævnte udvidelse, kaldet SEIR-modellen, er også illustreret i figur 1 og er særligt relevant for mæslinger, hvor latensperioden er af en betydelig længde.

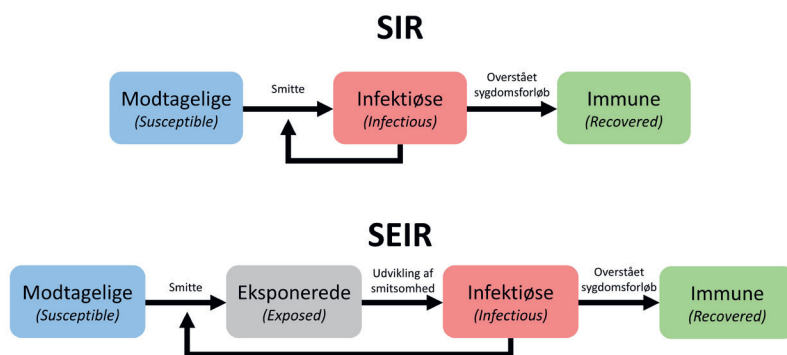
Det er et aktivt forskningsområde i sig selv at undersøge, hvordan udvidelser eller andre ændringer af modellen har betydning for den modellerede epidemis forløb. Denne type modellering kan benyttes både hypotesetestende og hypotesegenererende: Hypotesetestende ved at inkludere kendte mekanismer fra virkeligheden for at se, om resultatet stemmer overens med observationer, og hypotesegenererende ved at foreslå alternative mekanismer, som beskrives matematisk, men som ikke nødvendigvis er veletablerede (eller i det hele taget observerbare) fra det virkelige system, der modelleres. Denne forskel på formålet med modellering har betydning for, i hvilken grad det videnskabelige bidrag er kvalitativt eller kvantitativt. For samarbejde på tværs af matematisk modellering og historieforskning kan det derfor være af yderste vigtighed at være bevidst, og enige, om modelleringens bidrag til forskningsspørgsmål. Et eksempel på et kvalitativt bidrag er idéen om superspredning af SARS-CoV-2 (coronavirus), som blandt andre blev foreslået i en model af Nielsen m.fl.⁷ Et andet eksempel er udvidelse og undersøgelse af en model med målet at beskrive noget fundamentalt om smittespredning. Som koncept findes det i Andreas Eilersens bidrag i dette temanummer.⁸

5 McKendrick og Kermack: "A contribution", 700-721; Anderson og May: *Infectious Diseases of Humans*.

6 En populærvidenskabelig artikel om SIR-modellens grundlag kan findes online på videnskab.dk. Pedersen: "Tilbage til begyndelsen" (03.04.2024).

7 Nielsen m.fl.: "COVID-19 Superspreading", 118301.

8 Eilersen: "Matematisk modellering".



Figur 1. Øverste del viser SIR-modellen og de tre stadier i modellen: modtagelige, infektøse og immune. Pilene fra Modtagelige til Infektøse og fra Infektøse til Immune illustrerer, hvordan individer tænkes at bevæge sig fra stadie til stadie, mens pilen fra Infektøse til den pil, der er markeret Smitte, er for at vise, at antallet af infektøse individer har indflydelse på, hvor meget ny smitte der er i den modellerede befolkning. I den nederste halvdel vises SEIR-modellen, hvor der indgår et ekstra stadie i sygdomsforløbet i form af en latensperiode, hvori man er eksponeret for sygdommen, men endnu ikke er smitsom og derfor ikke kan smitte andre, før man rykker videre til det infektøse stadie efter nogen tid.

DET INVERSE PROBLEM

En anden måde at arbejde med matematiske modeller på er ved at sammenholde modellens resultater direkte med observationer fra virkeligheden for kvantitativt at sige noget om de mekanismer, der indgår i modellen. Hvor man typisk benytter en model til at lave beregninger som f.eks. forudsigelser af vejret inden for en bestemt tidshorisont, så beskæftiger arbejdet her sig med det, man kalder *det inverse problem*, da målet ikke er at beregne noget observerbart ud fra kendte parametre (talværdier), men omvendt at estimere værdien af disse parametre ud fra kendte observationer.⁹ I en vejrmodel kunne dette være at vurdere, hvor hurtigt et vejrphænomen bevæger sig ud fra temperaturmålinger over tid.

Det konkrete inverse problem i denne artikel er følgende: Ud fra observationer af månedlige smittetilfælde i de fynske lægekredse i 1874, hvilke værdier skal SEIR-modellens parametre have for at genskabe epidemiens forløb?

I SEIR-modellen indgår tre parametre, som jeg her refererer til med de græske bogstaver β (beta), σ (sigma) og γ (gamma), der hver især repræsenterer forskellige aspekter af en mæslingeepidemi. Løst beskrevet styrer β , hvor meget infektøse smitter, σ styrer latensperioden, altså hvor længe der går, fra man smittes, til man kan smitte andre, og γ styrer smitsomhedsperioden, altså hvor længe der dernæst går, før man stopper med at smitte. En beregning med $\beta = 2$ vil således være en model af en situation, hvor infektøse smitter mere end i en beregning med $\beta = 1$.

Foruden den praktiske udfordring med at bestemme gode værdier for parametrene, så er et vigtigt aspekt af det inverse problem her også fortolkningen af, hvad

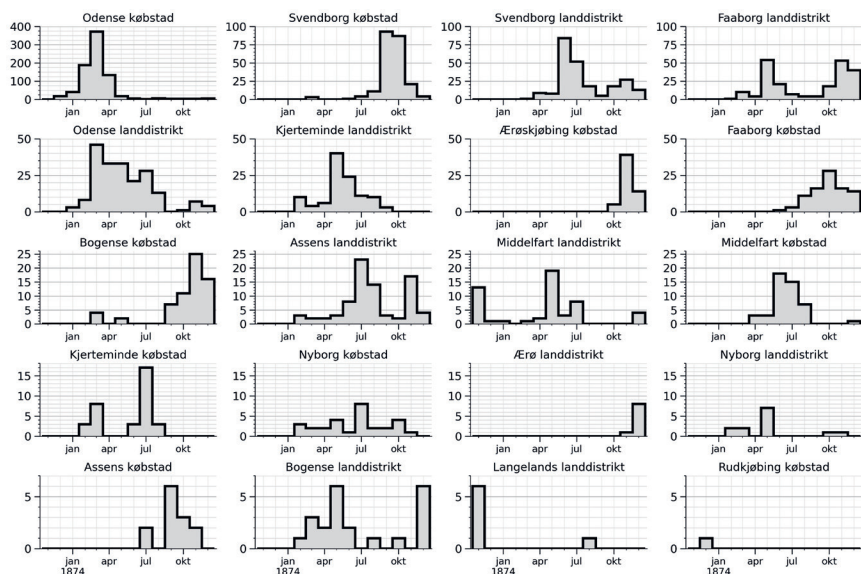
9 Aster m.fl.: *Parameter Estimation*.

disse værdier fortæller os om samfundet, hvori mæslinger spredte sig. Som det ses senere i denne artikel, så kan undersøgelsen af det inverse problem eksempelvis antyde, at mæslinger spredtes langt hurtigere i Kerteminde købstad end i Faaborg købstad. Hvordan denne information kan indgå i historiske slutninger om samfundsmæssige forskelle på de to købstæder, er et komplekst videnskabsteoretisk spørgsmål, hvor tværfaglig dialog er nødvendig. Dertil kommer naturligvis også historiefaglige overvejelser over tilbivelsen af de registrerede mæslingetilfælde og medicinalberetningerne som kilde.

MEDICINALBERETNINGER FOR FYN 1874

Medicinalberetningerne udgør et vigtigt kildemateriale til forskning i smitsomme sygdomme i Danmark.¹⁰ Denne kildegruppe består af skriftlige årsberetninger, som læger og kirurger blev pålagt at skrive fra 1803, til Det Kgl. Sundhedskollegium. Fra 1862 indgår skematiske optegnelser af sygdomstilfælde af bestemte sygdomme. For Fyn indeholder disse skemaer i 1870'erne månedlige antal for nye tilfælde af blandt andet mæslinger for såvel købstæder som landdistrikter. Her ses det, at der de fleste steder i årene 1871-1873 kun rapporteredes få enkeltstående tilfælde af mæslinger. To undtagelser er 34 tilfælde i Nyborg købstad i udgangen af 1872 og begyndelsen af 1873 samt en større epidemi med 220 tilfælde i Langelands lægedistrikt i foråret 1873.

I december 1873 observeredes 17 mæslingetilfælde i Odense købstad, som i løbet af 1874 voksede til en betydelige epidemi med samlet 779 tilfælde, og som spredte sig til resten af Fyn. Særligt i købstæderne ses tydeligt epidemier, der løb hen over flere måneder. I figur 2 vises data for de købstæder og landdistrikter, som undersøges.



Figur 2. Månedlige tilfælde af mæslinger rapporteret i medicinalberetningerne for 1874 samt for november og december 1873. Bemærk, at y-aksen ikke er den samme i de forskellige paneler. Udover i øverste række er y-aksen dog den samme langs hver række af paneler.

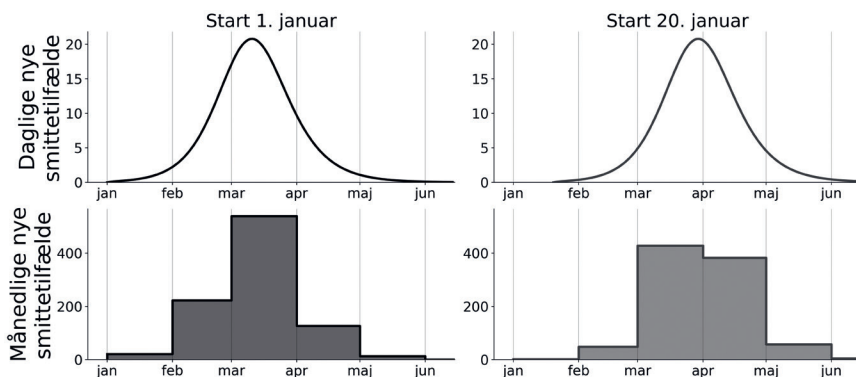
10 Bonderup m.fl.: *af yderste Vigtighed*.

Nærværende undersøgelse antager, at der i købstæderne kun rapporteredes cirka 30 % af det faktiske antal tilfælde, mens det i landdistrikterne må være endnu lavere, omkring 10 %.¹¹ Jeg skalerer derfor de rapporterede tilfælde ved at dividere med 0,3.

SAMMENHOLDNING AF MODELLENS OUTPUT OG DATA

Valget af, hvilken matematisk model man benytter til at undersøge et givent problem, er sjældent trivielt, men de metodiske overvejelser bag dette og konsekvenser deraf går ud over formålet med denne artikel. Jeg benytter her den nævnte SEIR-model.¹² En simulering giver antallet af modtagelige, eksponerede, infektiøse og immune på valgte tidspunkter.¹³

Antallet af daglige nye smittede i simuleringen holdes op imod observeret data. Her opstår to udfordringer, der er illustrative for de metodiske overvejelser, man som matematisk modellør støder på i denne slags arbejde: Den ene er, at data er på månedsbasis, og det andet, at modellen giver antallet af smittede på en given dag siden en given startdato i beregningen. Når en startdato er valgt, kan antallet af daglige smittede beregnes og summeres op på månedsbasis. I figur 3 illustreres, hvordan en forskel i startdato ændrer den månedlige sum af nye smittetilfælde.



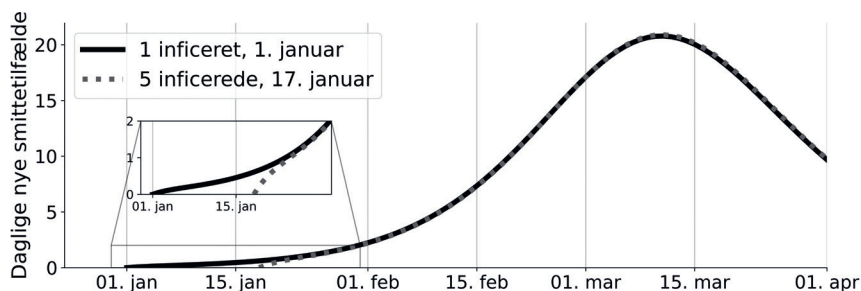
Figur 3. I øverste række vises de daglige nye antal smittetilfælde for en typisk beregning af SEIR-modellen (parameterværdier: $\beta = 0,4$, $\sigma = \gamma = 1/7$, $N = 1000$ med $S_0 = 999$, $I_0 = 1$, $E_0 = 0$), startende 1. januar og 20. januar for hhv. venstre og højre kolonne. De nederste paneler viser, hvor mange månedlige tilfælde de to beregninger giver, når der summeres på månedsbasis.

11 Disse overvejelser om rapporteringsraten er baseret på antagelsen om, at alle, der overlever barndommen, må antages at have været smittet med mæslinger. Sammenlignes antallet af voksne personer med de registrerede smittetilfælde, ses det, at kun en brøkdel af smittetilfælde er registreret (Gunning m.fl.: "Conserved Patterns", 1-9.) I et studie, der endnu ikke er udgivet, vises tilsvarende tal at gælde for Danmark (Andreasen m.fl.: "Temporal and Geographical Changes")

12 SEIR-modellen kan matematisk formuleres på flere forskellige måder. I dette arbejde benyttes konkret en formulering som et system af forbundne ordinære differentialligninger.

13 I dette tilfælde er der tale om numerisk integration af differentialligningssystemet foretaget på computer. Ordet *simulering* benyttes i daglig tale, om end der strengt taget er tale om en deterministisk *beregning*, der giver et resultat, som er garanteret til at ligge inden for en vis fejlmargen af en løsning til ligningssystemet.

Det kan have stor betydning for den månedlige sum af tilfælde at rykke startdatoen én enkelt dag, hvorfor det er hensigtsmæssigt i stedet at ændre på antallet af smittede på startdatoen.¹⁴ Dette har primært betydning for de første dage af epidemien, se figur 4, og har minimal effekt på de månedlige smittetilfælde.



Figur 4. Eksempel på daglige nye smittetilfælde af to SEIR-model beregninger. Den sorte linje starter med ét smittetilfælde den 1. januar, mens den grå stiplede linje i stedet starter med fem smittetilfælde den 17. januar. Efter få dage ses det, at den grå kurve, der starter den 17. januar, ligger næsten nøjagtigt oven på den sorte kurve med epidemier, der følger hinanden i resten af forløbet. Modelparametre som i figur 3. Årsagen bag afvigelsen i starten, og hvordan den kan reduceres, udelades fra diskussionen her.

PARAMETRE OG EN LØSNING AF DET INVERSE PROBLEM

For at løse det inverse problem, der er blevet introduceret tidligere, ledes efter en beregning af modellen, som *passer godt* til den observerede data. Formelt skal der bruges et mål for, hvad der menes med at passe godt til data. Som et mål for dette benyttes typisk kvadratsummen af 'fejlen', hvilket her betyder forskellen mellem data og den månedlige opsummerede antal nye smittetilfælde i en modelberegning. På den måde kan man kvantitativt beskrive, hvorvidt én bestemt modelberegning stemmer bedre overens med data end en anden beregning.

Som beskrevet ovenfor indgår der i SEIR-modellen tre parametre, β , σ og γ . Hertil kommer tre yderligere parametre, nemlig antallet af eksponerede og infektiøse på dag nul, E_0 og I_0 , samt størrelsen af den modellerede befolkning, N . Samlet set giver dette seks forskellige parametre, der har betydning.¹⁵

Før det undersøges, hvilke værdier af disse parametre som giver det bedste resultat (i forhold til det valgte mål), er der dog en sidste vigtig overvejelse. Data består i de fleste tilfælde af højst fem-seks månedlige tal. Det er således meget lidt data

¹⁴ Dette skyldes også mere praktiske årsager, som grundlæggende handler om, at beregningens *startbetingelser* skal vælges nøje alligevel.

¹⁵ Afhængigt af, hvordan modellen formuleres, bør antallet af immune på dag nul også tælles med. Her antages antallet af immune at være nul. Dette har ikke betydning for beregningen, men dog betydning for fortolkningen af resultatet. Dag nul i beregningerne sættes til den første dag i den første måned, der observeres smittetilfælde i data. Læsere, der allerede er bekendt med SEIR-modellen, bør bemærke, at definitionen af parameteren her adskiller sig fra den almindelige notation, idet der i beregningen kun ses på den modtagelige del af befolkningen og ikke den samlede befolkning.

til at vurdere værdierne af seks parametre. Dertil kommer, at der for nogle modeller kan være parametre, hvor man ved at sænke værdien af én parameter og øge værdien af en anden kan opnå samme resultat. Detaljeret matematisk analyse af modellen og *identificerbarheden* af parametrene kan beskrive denne udfordring formelt.¹⁶ Om end det ikke er nok i sig selv, er det hensigtsmæssigt at begrænse, hvor mange parametre man ændrer på i hver simulering. Det antages derfor, for eksemplets skyld, at antallet af eksponerede personer ved simuleringens begyndelse er nul, samt at parametrene σ og γ er $1/7$, svarende til en gennemsnitlig sygdomsforløb på 14 dage, i overensstemmelse med estimater for mæslinger fra litteraturen.¹⁷

Der findes mange metoder til at bestemme, hvilke værdier af de resterende tre parametre der passer godt til data. *Parameterestimering* er i sig selv et forskningsområde af en betydelig størrelse.¹⁸

Forskellige *optimeringsmetoder* giver bud på værdier for parametre, der passer så godt på data, somme tider med yderlig information om eksempelvis, hvor sikker metoden er i disse gæt. Her benyttes Pythonbiblioteket *lmfit*.¹⁹

RESULTATER

De tre resterende modelparametre β , I_0 og N estimeres separat for hver af de fynske købstæder ud fra de månedlige smittetilfælde (korrigeret for underrapportering som beskrevet tidligere). Assens og Nyborg købstæder oplevede i 1874 kun meget få sporadiske tilfælde og udelades derfor fra analysen. I landdistrikterne oplevedes ligeledes epidemier, særligt i Kerteminde, Faaborg, Svendborg og Odense landdistrikter. Epidemien i Odense landdistrikt følger ikke en typisk epidemisk kurve og udelades her. I Faaborg og Svendborg landdistrikter ses det umiddelbart fra data, at der er tale om to separate epidemier. Under en antagelse om, at disse epidemier sker i forskellige dele af landdistrikterne, estimeres parametre for epidemierne separat. Grundet få smittetilfælde udelades de resterende landdistrikter fra analysen.

I figur 5 ses modelberegninger med de estimerede parametre, sammenlignet med månedlige data. De estimerede parametre kan ses i tabel 1 sammen med, hvilken dato antallet af nye smittetilfælde topper i beregningen.

Tabel 1 (næste side). Oversigt over de estimerede parametre. Parameteren β er relateret til, hvor hurtigt mæslinger spredtes fra person til person. I_0 er en startbetingelse for modelberegningen, hvilket i denne sammenhæng ikke har nogen betydningsfuld fortolkning. Estimaterne for N beskriver antallet af personer i den modtagelige befolkning. Sidste kolonne viser, hvilken dato epidemien toppede i modelberegningen.

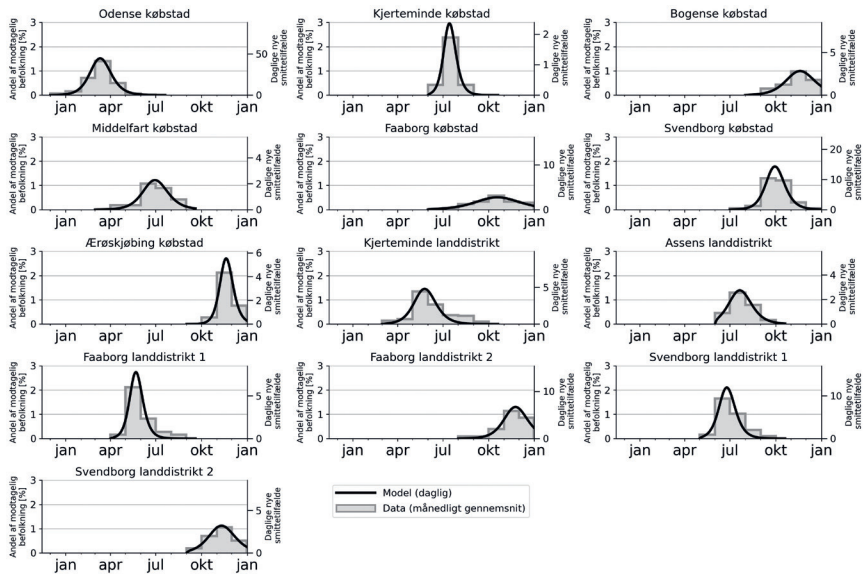
16 Eisenberg m.fl.: "Identifiability and Estimation", 84-102; Dankwa m.fl.: "Structural Identifiability", 100643.

17 Anderson: "Directly Transmitted Viral Infections", 1-37.

18 Aster m.fl.: *Parameter Estimation*.

19 Newville m.fl.: "lmfit" (03.04.2024).

Købstad/landdistrikt	β	I_0	N	Maksimumdato
Odense	0,33	1,01	2941	1874-03-11
Kerteminde	0,54	0,23	79	1874-07-14
Bogense	0,27	0,31	283	1874-11-19
Middelfart	0,30	0,05	187	1874-06-29
Faaborg	0,22	1,05	541	1874-10-18
Svendborg	0,36	0,26	803	1874-09-29
Ærøskøbing	0,49	0,01	204	1874-11-20
Assens landdistrikt	0,31	3,40	197	1874-07-21
Kerteminde landdistrikt	0,32	0,41	330	1874-05-25
Faaborg landdistrikt Første epidemi	0,50	0,46	284	1874-05-22
Faaborg landdistrikt Anden epidemi	0,31	0,14	516	1874-11-24
Svendborg landdistrikt Første epidemi	0,40	2,10	567	1874-06-25
Svendborg landdistrikt Anden epidemi	0,28	2,06	285	1874-11-10



Figur 5. Resultater for de syv købstæders epidemier og de fem landdistriktsepidemier. De sorte streger viser modelberegningen med en daglig tidsopløsning. I baggrunden vises den månedlige data, skaleret til et gennemsnit, der svarer til dagligt niveau. Skalaen på venstre side af panelerne viser, hvor stor en andel der dagligt smittes ud af størrelsen på den modtagelige befolkning som estimeret i beregningen. På akse på højre side vises det tilsvarende i antal personer på daglig tidsopløsning.

ANALYSE OG FORTOLKNING AF RESULTATER

Parameteren I_0 har primært betydning for epidemiens timing, og den konkrete værdi, parameteren er estimeret til at have, er derfor afhængig af, hvilken dato simuleringen startes. Selvom en separat analyse viser, at en reduktion af I_0 til nogen grad kan kompenseres ved at øge β , så er der ikke nogen betydelig indsigt at fortolke ud fra de estimerede værdier af I_0 .

Estimatet for parameteren N er et mål for, hvor mange der ved epidemiens begyndelse tænkes at være modtagelige over for mæslinger, og kan fortolkes som antallet af børn født (eller tilflyttet) siden sidste epidemi. Om end ældre børn og voksne forventes at indgå i epidemien, fremgår det af medicinalberetningerne, at de fleste rapporterede tilfælde var blandt børn under 15. For eksempel ses for hele 1874 i Svendborg kun 11 tilfælde blandt voksne over 15 år, ud af samlet 224 tilfælde. Dette stemmer overens med estimater fra litteraturen af, at 90-98 % af mæslingetilfælde er blandt børn under 10 år.²⁰

Estimaterne for N ses at skalere med købstædernes størrelse, med Odense størst og Svendborg og Faaborg købstæder derefter. Sammenlignet med folketællingen i 1880 udgør estimaterne for den modtagelige befolkning mellem 8 og 16 % af købstædernes samlede befolkning. I Faaborg landdistrikt ses først en epidemi i en mindre befolkningsgruppe med 284 modtagelige, hvorefter den anden epidemi ramte en gruppe med 516 modtagelige. For Svendborg landdistrikt ses modsat, at epidemien først nåede en større befolkning for dernæst at bevæge sig til en mindre befolkning efter epidemien i Svendborg købstad.

Parameteren β er en kompleks størrelse, der dækker over flere aspekter af en epidemis forløb, der ikke kan adskilles fra hinanden, herunder kontaktmønstre på tværs af aldersgrupper, sygdommens smitsomhed samt social adfærd, der kan ændre på, hvor meget sygdommen smitter. Derudover har befolkningens størrelse i sig selv også betydning for β i den formulering af modellen, der her benyttes. Jeg fortolker derfor β som et mål for, hvor hurtigt epidemien forløber om end betydningen af β er mere kompleks.²¹ I nederste panel af figur 5 ses det, at epidemierne i Kerteminde og Ærøskøbing købstæder samt den første epidemi i Faaborg landdistrikt forløb hurtigt. Dette ses også i de estimerede værdier af β på henholdsvis 0,54, 0,49 og 0,50. Omvendt ses de mere langstrakte epidemier i f.eks. Bogense og Faaborg at være udtrykt i lavere værdier, hvor β er henholdsvis 0,27 og 0,22.

Modellen gør det også muligt at bestemme en dato, hvor antallet af daglige nye smittetilfælde var på sit højeste. Selvom smittespredning fra område til område kan ses fra den oprindelige data, så kan dette forsimple beskrivelsen af epidemiens timing til en enkel dato. Det ses, at den første epidemi i Faaborg landdistrikt toppede fem måneder før epidemien i Faaborg købstad, mens den anden epidemi i landdistriktet skete meget kort tid efter, kun lidt over en måned. En lignende situation ses i Svendborg,

20 Anderson: "Directly Transmitted Viral Infections", 1-37.

21 Parameteren β er relateret til det effektive kontakttal R_e , som er et udtryk for, hvor mange en infektøs i gennemsnit smitter i løbet af deres smitteforløb. Grundet valget af parameterværdi for γ gælder det at $R_e = 7 \cdot \beta$. Bemærk, at det effektive kontakttal R_e er forskelligt fra det basale kontakttal R_0 , der for mæslinger forventes at være omkring 13.

hvor landdistriktets første epidemi var tre måneder før epidemien i købstaden, og den anden epidemi var blot halvanden måned efter. I Kerteminde landdistrikt og købstad ses det også, at epidemiene er forskudt med omkring halvanden måned, startende i landdistriktet.

DISKUSSION

Epidemiologiske aspekter af en serie af mæslingeepidemier på Fyn i 1874 er her blevet analyseret med perspektiver fra matematisk modellering. Resultatet af analysen giver to forskelligartede bidrag. Det ene i form af rekonstruktionen af epidemiernes forløb som vist i figur 5, hvor månedlige data fra kildematerialet nu kan ses repræsenteret som daglige nye smittetilfælde. Sådanne rekonstruerede forløb er abstraktioner, der afviger fra det faktiske kildemateriale, men maler et billede af, hvordan epidemiene sandsynligvis blev observeret fra dag til dag.

Det andet bidrag af analysen er den kvantitative værdi af de estimerede parametre og fortolkningen af disse. Forud for vores beregninger antog vi, at parametrene relateret til et sygdomsforløbs varighed, σ og γ , var kendte på forhånd. En konsekvens heraf er, at forskelle i estimerede parameterværdier er udtryk for samfundsmæssige forskelle. Med forbehold for resultaternes validitet giver parametrene således mål for, hvordan mæslinger spredtes forskelligt i de analyserede købstæder.²² Hvor der for Odense købstad observeredes en langsomt spredende epidemi i en stor befolkning af modtagelige, var der i Kerteminde købstad tale om en epidemi, som meget hurtigt spredte sig blandt købstadens langt mindre befolkning. Denne observation reflekterer muligvis sociale forskelle i de to købstæder. Det kan dog også være et udtryk for sygdommens spredning inden for lægedistrikterne og den restriktion, som månedlig tidsopløsning af data sætter for analysen: Den langsomme epidemi i Odense kan i virkeligheden være resultatet af flere mindre, men hurtigt spredende epidemier i forskellige dele af byen, der i de månedlige opsummeringer ligner én epidemi. De matematiske resultater indikerer altså således, at de månedlige opgørelser fra medicinalberetningerne ikke i sig selv er fyldestgørende for at kortlægge mæslingeepidemien i Odense købstad.

Inden parameterestimering skaleredes data for at tage højde for smittetilfælde, der ikke er blevet indberettet. Dette har betydning for størrelsen af parameteren N . Antages der f.eks. i stedet, at kun 20 % af tilfælde rapporteres, bliver den estimerede værdi af N 50 % højere.²³ Hvor stor en andel af smittetilfælde man forventer bliver rapporteret, har med andre ord stor betydning for estimatet af den modtagelige befolkningsgruppe i en købstad. Lokale forskelle i rapportering kan således ligge til grund for en betydelig del af de observerede forskelle i antallet af smittetilfælde i lægekredsene, foruden den forskel som befolkningsstørrelserne i sig selv udgør. Epidemiens forløb og hastighed er dog uændret af denne skalering.

22 Udfordringer med identificerbarheden af modellens parametre og korrelation mellem de estimerede parametre vidner om en begrænset validitet af resultaterne, hvilket også har indflydelse på fortolkningen af disse.

23 Da 0,3 divideret med 0,2 er 1,5, vurderes det faktiske antal tilfælde 50 % højere, når der skaleres med 20 % sammenlignet med 30 %.

I analysen var det muligt at bestemme en given dato, hvor antallet af smitte-tilfælde i hver lokal epidemi toppede. Dette kan gøre det nemmere at give kvalificerede bud på, hvordan smitten bevægede sig rundt på Fyn end ud fra den månedlige data alene. Epidemien i Odense købstad efterfulgtes af en langstrakt periode med smitte i det omkringliggende Odense landdistrikt, som kan forklares med flere mindre epidemier i landsbyer i landdistriktet. Smitten rykkede dernæst til Faaborg landdistrikt, hvor den ses som den første epidemi i landdistriktet. Smitten bevægede sig dernæst mod vest, til Assens landdistrikt. Den efterfølgende epidemi i Faaborg købstad og senere igen i det omkringliggende Faaborg landdistrikt kan meget vel være forbundet med smitten i Assens landdistrikt, om end det ikke er muligt at vurdere det ud fra dette arbejde, især når de er baseret på blot et enkelt kildegrundlag. Sådanne spørgsmål kunne undersøges med mere avancerede matematiske modeller, hvori der også tages højde for mekanismer bag smittespredning fra område til område.

Den præsenterede analyse illustrerer nogle af de metodiske overvejelser og resultater, der opstår i forbindelse med matematisk modellering af konkrete epidemier. I matematisk modellering, som i ethvert andet videnskabeligt område, er det vigtigt at forholde sig kritisk til de opnåede resultater. Et estimat for antallet af børn, der er modtagelige over for mæslinger i en given købstad, vil aldrig være mere end netop det: et *estimat*. Utsigtede eller endda ubevidste antagelser i modellens formulering kan have stor betydning for, hvilken talværdi man får.²⁴

Metoderne fra matematisk modellering skal ses som et forsøg på at gå et spadestik dybere, end hvad der er muligt ud fra en (historisk) analyse baseret på det historiske kildegrundlag alene. Særligt når datamaterialet er småt, kan matematiske værktøjer udvide rækkevidden af fortolkninger af kilder, selvom der, en anelse selvmodsigende, er flere udfordringer med validiteten af de matematiske resultater, når antallet af observationer er lavt.

Hypotesegenererende modelarbejde kan bane vejen for nye idéer, f.eks. om, hvordan sygdomme spredtes i historien, og hypotesetestende arbejde kan undersøge, om en given fortolkning passer med observeret data. Men begge typer arbejde baseres på en forforståelse af den historiske kontekst og for epidemier også om sygdommens spredningsbiologi og patologi. Hvis denne forforståelse mangler, bliver modelberegningens beskrivelser af samfund og sygdom i bedste fald unøjagtige og i værste fald direkte forkerte. Dertil kommer også, at selvom en matematisk undersøgelse kan være korrekt, nyskabende og baseret på både grundig matematisk og historisk viden, så kan de opnåede resultater være historisk uinteressante.

I dette arbejde var det overordnede spørgsmål, hvilke værdier af SEIR-modelens parametre, der skaber et epidemiforløb, som ligner det, man ser i data for mæslinger i 1874. Selvom fortolkninger af disse parametre medtages i en historisk analyse, så var de undersøgte spørgsmål ikke baseret på historiefaglige overvejelser. Hvilke spørgsmål der har størst værdi at undersøge i en given historisk analyse, er svært at vide på

24 Dertil kommer også usikkerhed i de estimerede værdier. Til de beskrevne estimater hører en matematisk usikkerhed, der kan beskrives formelt, hvilket her har været udeladt af pædagogiske årsager.

forhånd og kræver et nært samarbejde og god dialog på tværs af historie, matematik og epidemiologi.

Derudover sætter tværfagligt arbejder store krav til nøjagtige beskrivelser af såvel baggrund for data, konteksten hvori den er indsamlet, samt fortolkning af resultaterne af kvantitative undersøgelser. Med en stigende kompleksitet af forskellige kvantitative metoder inden for både matematisk og statistisk modellering er godt samarbejde på tværs af faggrænser nødvendigt for korrekt valg og brug af metode. I denne artikel er et område inden for matematisk modellering beskrevet detaljeret med et håb om at understøtte fremtidig forskning inden for historisk epidemiologi og for at inspirere til samarbejde og erfaringsudveksling mellem matematik og historie generelt. Med den betydning for samfundet, som epidemier og pandemier har haft og til stadighed har, er det nødvendigt at gøre brug af ethvert redskab tilgængeligt for at uddybe vores indsigt i sygdommene selv og samfundene, hvori de spredtes.

PUBLICERET MATERIALE

- Anderson, Roy M.: "Directly transmitted viral and bacterial infections of man". I Roy M. Anderson (red.): *The Population Dynamics of Infectious Diseases: Theory and Applications*, Springer US, 1982, 1–37.
- Anderson, Roy M. og Robert M. May: *Infectious Diseases of Humans*, Oxford: Oxford University Press, 1991.
- Andreasen, Viggo, Cécile Viboud og Lone Simonsen: "Epidemiologic Characterization of the 1918 Influenza Pandemic Summer Wave in Copenhagen: Implications for Pandemic Control Strategies", *The Journal of Infectious Diseases*, 197 (2), 2008, 270–78 doi: 10.1086/524065.
- Andreasen, Viggo, Maarten van Wijhe, Mathias Mølbak Ingholt, Mads Linnet Perner og Lone Simonsen: "Temporal and geographical changes of reporting rates and case fatality for Measles in Denmark 1870 - 1960" (Poster, Epidemics7 - International Conference on Infectious Disease Dynamics, 2019) <https://rucforsk.ruc.dk/ws/files/75872668/2019PosterEpidemics.pdf> (03.04.2024).
- Aster, Richard C., Brian Borchers og Clifford H. Thurber: *Parameter Estimation and Inverse Problems*, Elsevier, 2019, doi: 10.1016/b978-0-12-804651-7.00003-1.
- Bjornstad, Ottar N., Barbel F. Finkenstadt og Bryan T. Grenfell: "Dynamics of Measles Epidemics: Estimating Scaling of Transmission Rates Using a Time Series SIR Model", *Ecological Monographs*, 72.2, 2002, 169 doi: 10.2307/3100023.
- Bonderup, Gerda, Jørgen Mikkelsen og Lisbeth Skjernov: *"af yderste Vigtighed for det hele Borgersamfunds Tryghed": medicinalberetninger og deres anvendelsesmuligheder i historisk forskning*, København: Selskabet til Udgivelse af Kilder til Dansk Historie, 2005.
- Bremer, A. F. og C. E. Fenger: "Om Influenza-Epidemierne i Danmark i Aarene 1825 til 1844 med Bemærkninger om disse Epidemiers Indflydelse på Dødelighedsforholdene". I *De permanente Comiteers Arbejder i Aarene 1846 og 1847*, 1848, 213-232.

- Connor, Dilece og Andrew Hinde: "Mortality in town and countryside in early modern England", *Local Population Studies*, 89, 2012, 54-67 doi: 10.35488/lps89.2012.54.
- Dankwa, Emmanuelle A., Andrew F. Brouwer og Christl A. Donnelly: "Structural identifiability of compartmental models for infectious disease transmission is influenced by data type", *Epidemics*, 41 (September), 2022, 100643 doi: 10.1016/j.epidem.2022.100643.
- Dean, Katharine R., Fabienne Krauer, Lars Walløe, Ole Christian Lingjærde, Barbara Bramanti, Nils Chr. Stenseth og Boris V. Schmid: "Human ectoparasites and the spread of plague in Europe during the Second Pandemic", *Proceedings of the National Academy of Sciences of the United States of America*, 115 (6), 2018, 1304-9 doi: 10.1073/pnas.1715640115
- Eilersen, Andreas: "Matematisk modellering", *Temp – tidsskrift for historie*, 28, 2024, 110-126.
- Eisenberg, Marisa C., Suzanne L. Robertson og Joseph H. Tien: "Identifiability and estimation of multiple transmission pathways in cholera and waterborne disease", *Journal of Theoretical Biology*, 324, 2013, 84–102 doi: 10.1016/j.jtbi.2012.12.021.
- Gunning, Christian E., Erik Erhardt og Helen J. Wearing: "Conserved patterns of incomplete reporting in pre-vaccine era childhood diseases", *Proceedings of the Royal Society B: Biological Sciences*, 281 (1794), 2014 doi: 10.1098/rspb.2014.0886.
- Hinde, Andrew: "A review of methods for identifying mortality 'crises' using parish record data", *Local Population Studies*, 39 (84), 2010, 82–92.
- McKendrick, William Ogilvy og A.G. Kermack: "A contribution to the mathematical theory of epidemics", *Proceedings of the Royal Society of London*, 115 (772), 1927, 700-721 doi: 10.1098/rspa.1927.0118.
- Newville, Matt, Renee Otten, Andrew Nelson og Till Stensitzki: "lmfit/lmfit-py" (Zenodo, 2023) <https://zenodo.org/records/7887568> (03.04.2024) doi: 10.5281/zenodo.7887568.
- Nielsen, Bjarke Frost, Lone Simonsen og Kim Sneppen: "COVID-19 Superspreading Suggests Mitigation by Social Network Modulation", *Physical Review Letters*, 126 (11), 2021, 118301 doi: 10.1103/PhysRevLett.126.118301.
- Ottesen, Johnny T.: "The Mathematical Microscope – Making the Inaccessible Accessible". I Bernhelm Booß-Bavnbek, Beate Klösgen, Jesper Larsen, Flemming Pociot og Erik Renström (red.): *BetaSys : Systems biology of regulated exocytosis in pancreatic β -cells*, New York, NY: Springer New York, 2011, s. 97-118 doi: 10.1007/978-1-4419-6956-9_6.
- Pedersen, Rasmus Kristoffer: "Tilbage til begyndelsen: Lav dine egne corona-kurver", Videnskab.dk, 2021 <https://videnskab.dk/krop-sundhed/tilbage-til-begyndelsen-lav-dine-egne-corona-kurver/> (03.04.2024).
- Phelps, Matthew, Mads Linnet Perner, Virginia E. Pitzer, Viggo Andreasen, Peter K.M. Jensen og Lone Simonsen: "Cholera Epidemics of the Past Offer New Insights

Into an Old Enemy”, *The Journal of Infectious Diseases*, 217 (4), 2018, 641-49
doi: 10.1093/infdis/jix602.

Skydsgaard, Morten A.: “It’s probably in the air: Medical meteorology in Denmark, 1810-1875”, *Medical History*, 54 (2), 2010, 215-36, doi: 10.1017/S0025727300006724.

Walløe, Lars: “Medieval and Modern Bubonic Plague: Some Clinical Continuities”, *Medical History*, 52 (S27), 2008, 59-73, doi: 10.1017/S0025727300072094.

RASMUS KRISTOFFER PEDERSEN
POSTDOC, PH.D.
PANDEMIX CENTER, INSTITUT FOR NATURVIDENSKAB OG MILJØ
ROSKILDE UNIVERSITET
RAKRPE@RUC.DK

ABSTRACT (EN)

How Mathematical Modelling can Supplement History: Measles on Funen in 1874

In 1874, Funen and the islands nearby were struck by a series of measles epidemics – first in the market town of Odense during the first months of the year and later in other market towns and rural districts. From 1803, Danish doctors and surgeons were required to write an annual medical report for the Royal College of Health (Det Kongelige Sundhedskollegium), in which they were to give details of any epidemics that had occurred in their district. From 1862, the reports also had to include a table with monthly reported cases of different diseases, including measles. Using these reports, it is possible to track the measles epidemic from month to month and district to district.

This article describes how specific methods from mathematical modelling can be used to investigate historical data, such as that from the measles epidemic on Funen. The classic “SIR”-model and the extended “SEIR”-model are mathematical models, which describe the progress of epidemics of infectious diseases such as measles. By relating model calculations to data, it is possible to describe aspects of the epidemic which cannot be determined from the data alone. For example, it makes it possible to estimate the number of children susceptible to measles before the outbreaks of 1874.

Although this article is concerned with the outbreaks of measles on Funen in 1874, methods like those presented here can be used in other contexts. The work described here is used as an example to present the concept of mathematical modelling with the aim of inspiring future interdisciplinary work between mathematicians and historians.