

Automated and Early Detection of Disease Outbreaks

AEDDO

Master Thesis



Automated and Early Detection of Disease Outbreaks
AEDDO

Master Thesis
August, 2023

By
Kasper Schou Telkamp

Copyright: Reproduction of this publication in whole or in part must include the customary bibliographic citation, including author attribution, report title, etc.

Cover photo: Vibeke Hempler, 2012

Published by: DTU, Department of Applied Mathematics and Computer Science,
Richard Petersens Plads, Building 324, 2800 Kgs. Lyngby Denmark
www.compute.dtu.dk

ISSN: [0000-0000] (electronic version)

ISBN: [000-00-0000-000-0] (electronic version)

ISSN: [0000-0000] (printed version)

ISBN: [000-00-0000-000-0] (printed version)

Approval

This master thesis has been prepared over six months at the Section for Dynamical Systems under the Department of Applied Mathematics and Computer Science at the Technical University of Denmark, and the Department of Epidemiology Research / Model group at Statens Serum Institut in partial fulfillment for the degree Master of Science in Engineering, MSc Eng., Quantitative Biology and Disease Modelling.

Kasper Schou Telkamp - s170397

.....
Signature

.....
Date

Abstract

In recent years there has been a surge in interest and development of statistical methods for automated and early detection of infectious disease outbreaks. This growth can largely be attributed to the widespread availability of computerized databases, such as The Danish Microbiology Database (MiBa). The accessibility of these databases has provided a wealth of data that can be analyzed, facilitating the use of statistical outbreak detection methods.

However, at the time of writing this thesis, there are no fully automated detection procedures in place at Statens Serum Institut (SSI), and the identification of outbreaks relies on the work of individual epidemiologists. Consequently, there is a momentous demand for an automated procedure for the early identification of disease outbreaks.

This thesis presents the development of a novel method for automated and early detection of disease outbreaks, with a focus on comparing its performance to state-of-the-art methods from the literature and the methods currently employed at SSI. The main objective of this research is to evaluate the effectiveness of the proposed method and its potential for improving outbreak detection.

The methodology employed in the novel method centers around an innovative utilization of generalized linear mixed effects models and hierarchical generalized linear models. These models offer a robust framework for capturing the underlying dynamics of disease spread and allow for the incorporation of various covariates and random effects.

Through case studies involving foodborne illnesses, including *Listeriosis*, *Shigellosis*, Shiga toxin (verotoxin)-producing *Escherichia coli*, and *Salmonellosis*, as well as a simulation study, the key findings of this research highlight the promising potential of the novel method. The developed approach demonstrates good performance in detecting disease outbreaks compared to established methods, including the Farrington method and the subsequently improved Noufaily method. Moreover, it effectively controls the number of “false alarms”, thereby enhancing the reliability of the outbreak detection method.

Keywords— Automated outbreak detection, Prospective outbreak detection, Disease surveillance, Generalized linear mixed effects models, Hierarchical generalized linear models, Foodborne illnesses

Acknowledgements

I would like to express my heartfelt gratitude to my supervisors, Lasse Engbo Christiansen and Jan Kloppenborg Møller, for their invaluable guidance, support, and expertise throughout the entire process of conducting my master thesis. Their extensive knowledge, insightful feedback, and dedication have been instrumental in shaping the direction and quality of this research.

Lasse Engbo Christiansen's exceptional guidance and expertise in mathematical modeling have been crucial in providing me with a solid foundation and understanding of the subject matter. His valuable insights, constructive criticism, and attention to detail have immensely contributed to the development of this thesis. Furthermore, Lasse facilitated my project and provided me with the best possible opportunity to succeed with my thesis by offering me a work desk at Statens Serum Institut.

I am deeply grateful to Jan Kloppenborg Møller for his continuous support and mentorship. His vast experience with hierarchical models and their application has been a huge asset to this project. His guidance in refining research methodologies, analyzing data, and interpreting results has been invaluable.

I would like to extend my sincere appreciation to both supervisors for their patience, encouragement, and willingness to dedicate their time and expertise to help me overcome challenges and achieve the objectives of this research. Their unwavering support and belief in my abilities have been a constant source of motivation throughout this journey.

I would also like to acknowledge the Section for Dynamical Systems under the Department of Applied Mathematics and Computer Science at the Technical University of Denmark, and the Department of Epidemiology Research / Model group at Statens Serum Institut for providing a conducive environment for research and for the resources that have facilitated the completion of this thesis.

Finally, I would like to express my gratitude to my family, friends, and loved ones for their unwavering support, compassion, and encouragement throughout this endeavor. I am especially grateful to my wife, Cecilia Schou Telkamp, who, aside from providing me with endless support, listening to all my complaints, and hosting several reading sessions, has also raised our two oldest daughters and given me the beautiful gift of a newborn boy during this period of conducting my thesis.

Notation

The mathematical notation in this master thesis is adapted from Madsen and Thyregod 2011. All vectors are column vectors. Vectors and matrices are emphasized using a bold font. Lowercase letters are used for vectors and uppercase letters are used for matrices. Transposing is denoted with the upper index T . Random variables are always written using uppercase letters. Thus, it is not possible to distinguish between a multivariate random variable and a matrix. However, variables and random variables are assigned to letters from the last part of the alphabet (X, Y, Z, U, V, \dots), while constants are assigned to letters from the first part of the alphabet (A, B, C, D, \dots). From the context it should be possible to distinguish between a matrix and a random vector.

It is assumed that the reader has a basic knowledge in the areas of statistics.

Contents

Approval	ii
Abstract	iii
Acknowledgements	iv
Notation	v
Acronyms	ix
1 Introduction	1
2 Setting the scene	3
3 Surveillance data in Denmark	5
3.1 Data collection and data quality	5
3.2 Introducing the case studies	5
4 Methods	15
4.1 State-of-the-art outbreak detection algorithm	15
4.2 Novel outbreak detection algorithm	17
4.3 General mixed effects models	18
4.4 Hierarchical generalized linear models	20
4.5 Parameter estimation	25
4.6 Scoring rule	27
5 Case studies	31
5.1 Outbreak detection in Shiga toxin (verotoxin)-producing <i>Escherichia coli</i>	31
5.2 Outbreak detection in other case studies	40
5.3 Challenges in statistical outbreak detection	47
5.4 Performance of statistical outbreak detection algorithms	51
5.5 Potential extension of the novel algorithm	55
6 Simulation study	59
6.1 The simulated baseline data	59
6.2 The simulated outbreaks	60
6.3 Evaluation measures	62
6.4 Results of the simulations	62
7 Discussion	67
7.1 Alarm fatigue and the implications for outbreak detection	67
7.2 For and against for the two modeling frameworks in the novel method	67
7.3 Harnessing the potential of MiBa	68
8 Conclusion	69
Bibliography	71
A Some probability functions	77
A.1 The Poisson distribution model	77
A.2 The Gamma distribution model	77

A.3 The Negative Binomial distribution	78
A.4 The Gaussian distribution model	78
B State-of-the-art detection algorithm	79
B.1 Controls	79
C Clusters reported to The Food- and waterborne Outbreak Database	81
D Alarms by state-of-the-art methods in the <i>Listeriosis</i> case study	89
E Figures and tables related to the case studies	92
E.1 <i>Listeriosis</i>	92
E.2 <i>Shigellosis</i>	93
E.3 Shiga toxin (verotoxin)-producing <i>Escherichia coli</i>	95
E.4 <i>Salmonellosis</i>	98

Acronyms

DTU Technical University of Denmark.

ACF autocorrelation function.

CRAN Comprehensive R Archive Network.

DVFA Danish Veterinary and Food Administration.

SSI Statens Serum Institut.

GLM Generalized Linear Model.

WGS Whole Genome Sequencing.

LIST *Listeriosis*.

SHIL *Shigellosis*.

STEC Shiga toxin (verotoxin)-producing *Escherichia coli*.

SALM *Salmonellosis*.

HUS *Hemolytic uremic syndrome*.

PCR Polymerase Chain Reaction.

FPR False Positive Rate.

POD Probability Of Detection.

MiBa The Danish Microbiology Database.

COVID-19 Coronavirus disease 19.

HAIBA The Healthcare-Associated Infections Database.

FUD The Food- and waterborne Outbreak Database.

1 Introduction

The fight against infectious diseases not only requires proper treatment of patients and implementation of preventive measures but also demands early detection of emerging disease outbreaks. Timely identification and intervention can mean the difference between containing an outbreak or facing a devastating epidemic.

Statistical outbreak detection starts with the detection of an aberrant number of cases of a particular disease in a given time and space. When an increase in the number of cases is detected, an alarm, viz. the signal, is raised by the detection method. Next, an epidemiologist assesses the public health relevance of the aberration, to determine if further investigation is warranted.

In Denmark, the surveillance of infectious diseases is mainly carried out by Statens Serum Institut (SSI), which plays a pivotal role in national and international disease preparedness. The high quality of the Danish surveillance registers, facilitated by a robust regulatory community framework, presents a significant opportunity for data analysis and enables timely detection of aberrant case counts. More to, the establishment of MiBa by SSI in 2010 (Voldstedlund et al. 2014) has further strengthened the ability to monitor diseases promptly.

In this master thesis, emphasis is placed on diseases caused by foodborne illnesses, as they are often associated with specific food sources and provide an opportunity for intervention and preventive measures. Currently, several approaches are employed at SSI to detect outbreaks of foodborne illnesses. These include:

1. **Reports from doctors:** Physicians may report cases of foodborne illnesses they encounter in their practice to the relevant authorities.
2. **Citizen reports:** Individuals may directly contact food or health authorities to report suspected cases of foodborne illnesses.
3. **Cluster identification through laboratory surveillance:** Clusters of cases can be identified through routine laboratory testing and surveillance of samples from patients with suspected foodborne illnesses.
4. **Identification of identical "fingerprints":** When bacteria or viruses are type-tested, the presence of identical fingerprints among multiple cases can strongly indicate a common source of infection.

These approaches contribute to the early detection and investigation of infectious disease outbreaks, including foodborne diseases. However, at the time of writing this thesis, there are only partial automated procedures in place at SSI, and the identification of outbreaks relies on the work of individual epidemiologists. Consequently, there is a momentous demand for fully automated procedures for the early identification of disease outbreak signals. Such automated methods could effectively guide the work of epidemiologists at SSI and other health institutes, enhancing the overall surveillance of infectious diseases at national and international levels.

This master thesis proposes a novel method for automated and early detection of disease outbreaks. The method utilizes hierarchical models in an innovative way, offering an alternative to state-of-the-art outbreak detection algorithms. The models are specifically formulated to model the case count observations y and assess the corresponding

random effects u in a subset of diseases within the mandatory notification system. Subsequently, the random effects are directly utilized in the detection algorithm to characterize and identify outbreak signals.

The main objective of this research is to assess the effectiveness and practicality of automated outbreak detection methods. To achieve this, a comparative analysis is conducted, evaluating the ability of the novel method to detect outbreak signals in case studies. Specifically, this thesis considers case studies of *Listeriosis* (LIST), *Shigellosis* (SHIL), Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC), and *Salmonellosis* (SALM). The performance of the proposed method is benchmarked against state-of-the-art methods identified in the literature as well as the methods currently employed at SSI.

It is important to highlight that while the case studies in this thesis are disease-specific, the statistical algorithms utilized are not limited to these particular diseases. They can be applied to detect abnormal case counts in any type of disease, enabling early detection and timely response. The versatility of these algorithms allows for their application in various disease surveillance systems, contributing to the enhancement of public health efforts.

In addition to the case studies, a simulation study adapted from Noufaily et al. (2013) is conducted to evaluate the performance of the novel outbreak detection algorithm compared to the state-of-the-art algorithms. The objective is to assess the ability of the automated procedures to timely detect disease outbreaks. By leveraging the potential of hierarchical models, this research seeks to improve the timeliness and accuracy of disease outbreak detection.

The novel outbreak detection algorithm proposed in this master thesis is open-source and can be accessed at <https://github.com/telkamp7/AEDDO>

2 Setting the scene

Outbreak investigations have a long history, dating back to John Snow's iconic removal of the handle of London's Broad Street pump during the cholera epidemic in 1865 (Tulchin-sky 2018). Indeed, while John Snow's work was groundbreaking for his time, modern disease outbreak investigations require more advanced and sophisticated techniques.

Today, epidemiologists and public health professionals utilize a range of tools and methodologies to effectively tackle disease outbreaks. Here, laboratory-based approaches play a crucial role in outbreak investigations and may involve technique such as molecular epidemiology (Honardoost, Rajabpour, and Vakil 2018; Struelens and Brisson 2013) and, more recently, Whole Genome Sequencing (WGS) (Koeser et al. 2012; Baldry 2010).

However, in recent years, there has been a growing interest in statistical methods for automated and early detection of disease outbreaks. These methodologies encompass various statistical techniques, including regression analysis, time series methodology, methods inspired by statistical process control, approaches incorporating spatial information, and multivariate outbreak detection. A comprehensive review of these methods can be found in studies by Buckeridge (2007) and Unkel et al. (2012).

To establish a golden standard, this master thesis will focus on the method initially introduced by Farrington et al. (1996) and the subsequent improvements proposed by Noufaily et al. (2013). These methods offer advanced statistical tools for detecting and monitoring disease outbreaks and are currently *the* methods of choice at European public health institutes (Hulth et al. 2010). They can be accessed through the R package called **surveillance** developed by Salmon, Schumacher, and Höhle (2016).

It is a well known fact, that one limitation of these detection algorithms is an occasional lack of specificity, leading to false alarms that can overwhelm the epidemiologist with verification tasks (Bédubourg and Strat 2017). Therefore, in this master thesis, these established methods will be compared to a novel outbreak detection algorithm based on hierarchical models. The thesis introduces this new algorithm as an innovative approach to outbreak detection and aims to assess its performance in comparison to already existing methods.

While hierarchical models have earned a reputation within ecology (Bolker et al. 2009; Zuur et al. 2009), urban energy modeling (Palmer Real et al. 2022; Real et al. 2021), and other fields, their application in the automatic detection of disease outbreaks is relatively unproven. However, there is a promising paper by Heisterkamp, Dekkers, and Heijne (2006) that applied a hierarchical time series model to detect infectious disease outbreaks in empirical data from *Rubella* and *Salmonella*. The authors concluded that the method is a powerful and versatile way of analyzing time series of routinely recorded laboratory data.

Common to the aforementioned methods is that no accumulation of evidence takes place, and they detect aberrations only when the count of the currently monitored time point is above the threshold. However, not all disease outbreaks are characterized by a sudden increase in the number of reported cases. Specifically, some continuous-source outbreaks may exhibit a small but consistent increase in the number of cases, making them difficult for statistical outbreak detection methods to identify. To address this challenge, a method known as the Negative Binomial cumulative sum was introduced by Höhle and Paul (2008) and is also implemented in the R package **surveillance**. This method allows

for the detection of sustained shifts by accumulating evidence over multiple time points. However, for the purposes of this thesis, the focus will be on the prospective detection of disease outbreaks.

3 Surveillance data in Denmark

This chapter delves into the data collection methods and quality assurance procedures within the Danish surveillance system. Moreover, it introduces the case studies selected for this master thesis, which include *Listeriosis* (LIST), *Shigellosis* (SHIL), Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC), and *Salmonellosis* (SALM). The epidemiological background of the diseases, along with notable outbreaks investigated and documented by Statens Serum Institut (SSI) in the period from 2008 to 2022, is presented.

3.1 Data collection and data quality

In Denmark, the surveillance of infectious diseases is conducted by SSI. This surveillance system plays a pivotal role in national and international disease preparedness. It encompasses more than just the collection and registration of disease data; it also involves the prompt and ongoing dissemination of knowledge to the relevant authorities responsible for treatment, prevention, and control. This comprehensive approach ensures efficient communication and facilitates appropriate measures to address infectious diseases.

The quality of the Danish surveillance registers is maintained at a high standard, thanks to The National Board of Health Statutory Order on Physicians' Notification of Infectious Diseases (<https://www.retsinformation.dk/eli/ita/2000/277>). This order specifies that several diseases¹ are individually notifiable by physicians and general practitioners. Notifications consist of essential patient information and are submitted in paper form to both the Ministry of Health and to SSI. This rigorous notification process ensures accurate and comprehensive data collection for disease surveillance purposes in Denmark. In addition to the individually notifiable diseases, SSI has implemented a laboratory notification system for numerous microorganisms. Clinical-microbiological laboratories are obligated to report the identification of specific microorganisms, along with relevant patient information. These data are then stored in The Danish Microbiology Database (MiBa) (Voldstedlund et al. 2014).

MiBa is a nationwide and automatically updated database specifically designed to collect and store microbiological test results. In order to utilize the data from MiBa, the information in the test results needs to have a standardized structure with common codes and terminology. MiBa employs national standards to harmonize the data, which initially may be structured in diverse formats. The standards currently used are XRPT05, which is widely employed for the exchange of microbiological test results in the healthcare system, and a specific standard called XRPT06. These standards are regularly revised and exist in various versions². The national surveillance system focuses on diseases of a severe nature, those that are highly contagious, and the majority of vaccine-preventable diseases.

3.2 Introducing the case studies

For the scope of this master thesis, only a specific subset of diseases from the mandatory notification system is considered. This subset consists of LIST, SHIL, STEC, and SALM.

¹For a full list of diseases see https://www.ssi.dk/sygdomme-beredskab-og-forskning/anmeldelse-af-sygdomme/lovpligtige-meldesystemer/individ_anmeldelses_sygdomme

²Information on national standards and codes within the healthcare domain can be found on MedCom's website (<https://medcom.dk/>).

These diseases have been chosen for analysis and investigation based on various characteristics, such as seasonality, incidence, and severity. Additionally, these diseases are associated with documented outbreaks investigated and documented by SSI in the past decade, which adds to their relevance for the study. However, not all outbreaks are equally documented. Thus, for a full overview over the disease clusters reported in The Food- and waterborne Outbreak Database (FUD), refer to Table C.1. For visualization purposes, this figure only includes data from 2012 to 2022.

The count observations are observed in the period from January 2008 to December 2022. An epidemic curve graph for selection of the data for each of the diseases considered in this thesis is shown in Figure 3.1.

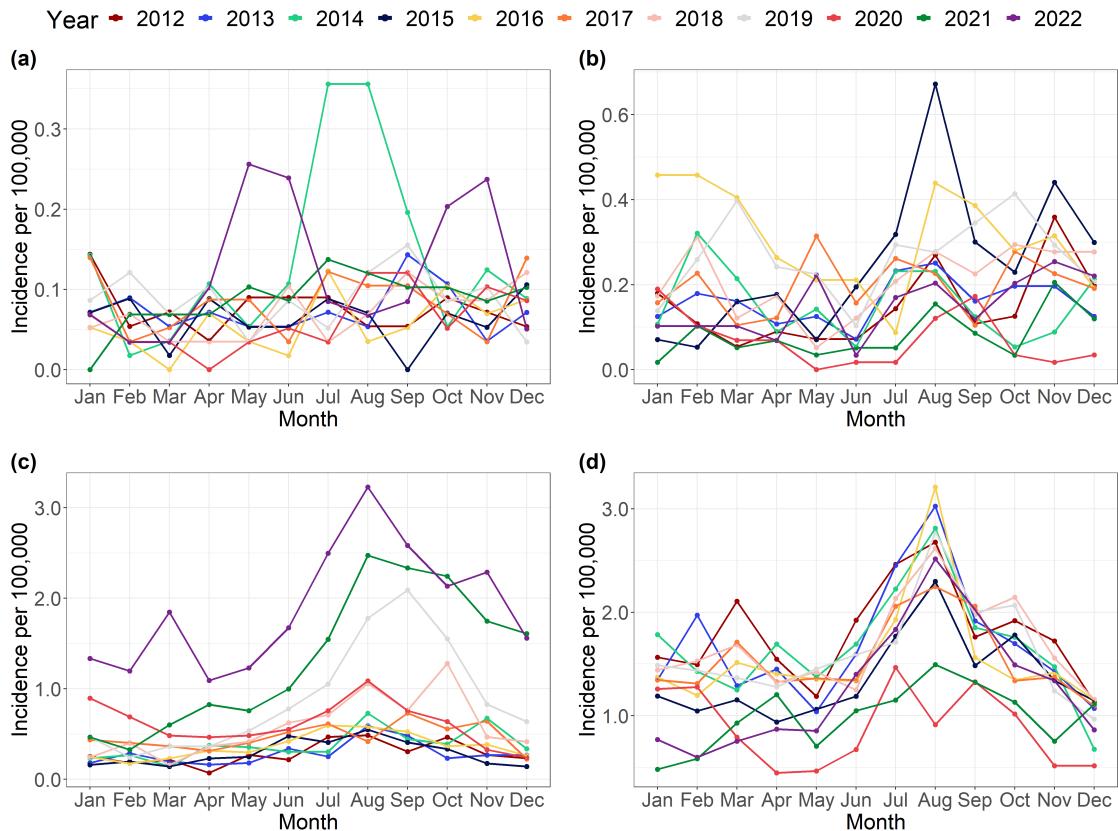


Figure 3.1: Epidemic curve showing the incidence per 100,000 in Denmark, 2012-2022, for the subset of diseases considered in this master thesis. **(a)** LIST, **(b)** SHIL, **(c)** STEC, and **(d)** SALM.

Evidently, the subset of diseases in Figure 3.1 exhibits varying incidences and different levels of seasonal patterns on an annual basis. It is interesting to note that the incidences generally peak in August, which can be attributed to several factors, including:

- **Increased travel activity:** Especially when individuals travel to countries with unsafe drinking water, poor sanitation, and insufficient hygiene practices.
- **Large social gatherings:** Events such as weddings, large festivals, or other gatherings where a significant number of people consume potentially contaminated food or drinking water.

Moreover, one could hypothesize that the warmer climate during the summer could po-

tentially directly influence the proliferation of bacteria, which in turn may have an impact on the transmission and spread of these diseases.

In general, there is a noticeable decrease in the number of observed cases starting from March 2020 and continuing until January 2021. This decline can be attributed to the strict lockdown measures implemented in Denmark in response to the Coronavirus disease 19 (COVID-19) pandemic. These measures, which involved restrictions on movement and social interactions, likely played a significant role in reducing the transmission of infectious diseases, including the ones being investigated.

From Figure 3.1a, it is evident that the incidence of LIST remains relatively low but stable over time. However, there are six specific months that stand out due to higher incidence rates. These months include July and August 2014, May and June 2022, and October and November 2022. These periods show a notable increase in the number of reported cases compared to the rest of the time series.

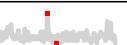
In Figure 3.1b, the incidence of the SHIL exhibits a consistent pattern over the years, with sporadic peaks occurring in the time series. Notable examples include a peak in August 2015, where the incidence reaches its highest level in the entire series. Additionally, January of 2016 stands out compared to other years. From the series alone, it is not evident whether there is a significant seasonality in the data. Another interesting observation is the substantial increase in incidence during August and September 2020 compared to the period leading up to these months.

Furthermore, in Figure 3.1c, a notable increase in the amplitude of the seasonal variation in STEC can be observed starting from 2018, with incidences doubling compared to the preceding years. At a first glance, this increase in the incidences might be recognized as a serious, reoccurring outbreak of the disease, but a more reasonable explanation can be found. Up to 2018, most departments of clinical microbiology used culture-based methods as a diagnostic test for bacterial pathogens and the process of changing the test method to Polymerase Chain Reaction (PCR) methods was ongoing (Svendsen et al. 2023). In general, PCR resulted in higher incidences compared to other test methods, which is no surprise as higher sensitivity is well documented for PCR (Buss et al. 2015; Knabl, Grutsch, and Orth-Höller 2016).

Overall, the highest incidences among the diseases considered in this thesis are observed for SALM in Figure 3.1d. SALM exhibits a notable pattern of high incidence throughout the observed period. However, there is a significant drop in incidences observed in 2020, which is consistent with the impact of the COVID-19 pandemic on disease surveillance and reporting. Additionally, 2021 shows generally lower incidences compared to previous years.

Some summary statistics for each of the diseases considered in this master thesis are gathered in Table 3.1.

Table 3.1: Summary statistics of the monthly count observations for the subset of diseases considered in this master thesis. Time series: normalized observations (0-1), first time points minimum and maximum count (red)

Case definition	Min	Max	Mean	Median	Std. Deviation	Time series
LIST	0	20	4.909	5	3.145	
SHIL	0	38	10.941	11	6.253	
STEC	2	190	33.177	21	33.108	
SALM	17	588	107.677	86	78.708	

In Table 3.1, it is readily seen that the diseases exhibit different statistical properties, including variations in mean number of monthly cases, standard deviation, and range. These variations indicate that each disease has its own unique epidemiological characteristics.

From an epidemiological perspective, all the diseases within the selected subset pose a significant risk of infection and can vary in terms of severity for affected individuals. Therefore, early identification of disease outbreaks is of utmost importance in order to promptly implement necessary interventions. Timely detection allows for swift and targeted actions to control the spread of these diseases and mitigate their impact on public health.

3.2.1 Epidemiological background and notable outbreaks of *Listeriosis*

Listeriosis (LIST) is a foodborne illness that is caused by consuming food contaminated with *Listeria monocytogenes*. The bacteria is ubiquitous in the environment, found in moist environments, soil, water, decaying vegetation, and animals. Furthermore, it can survive and even grow under refrigeration and other food preservation measures.

This disease primarily affects pregnant women, unborn or newborn babies, the elderly, and individuals with weakened immune systems. The disease is associated with high mortality (Goulet et al. 2012) and manifests in three ways: sepsis, meningitis, and mother-to-child transmission. Pregnancy-associated listeria can have severe consequences for the fetus or newborn, including miscarriage, stillbirth, neonatal sepsis, and meningitis (Awofisayo et al. 2015). LIST is uncommon among individuals in other demographic groups.

In 2014, Whole Genome Sequencing (WGS) was employed by SSI as the routine testing of *Listeria* isolates in Denmark (Wingstrand et al. 2015). This method involves mapping the entire DNA of the bacteria and enables SSI to identify cases where patients are infected with identical *Listeria* bacteria. However, it is important to note that for this thesis, the DNA typing data is unavailable for use.

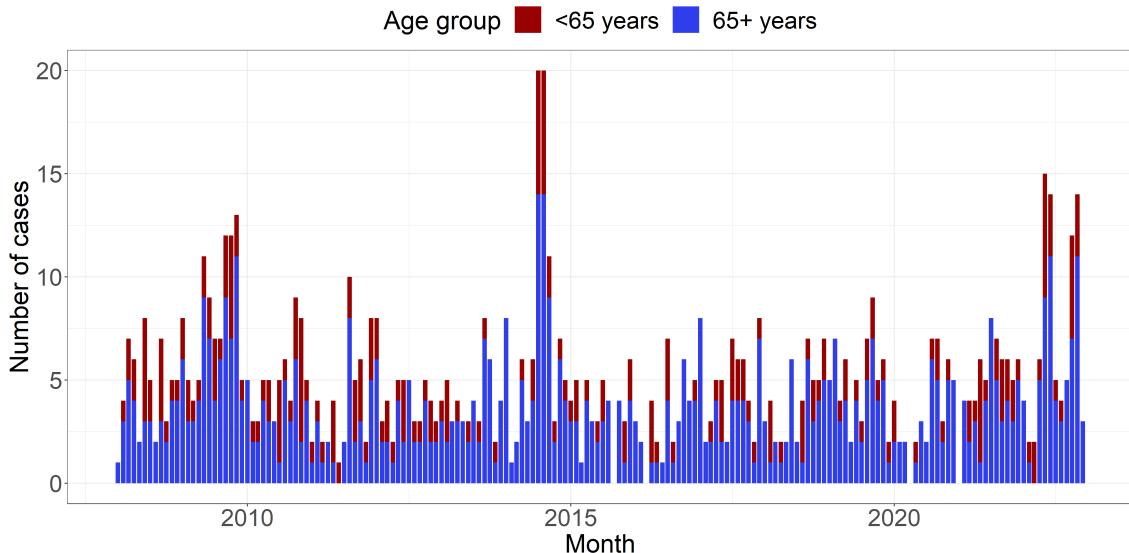


Figure 3.2: A stacked bar graph illustrating the number of monthly LIST cases observed in the period from 2008 to 2022 for the age groups below 65 years and 65 years and above.

The first notified outbreak in the 20th century occurred in 2009 (Smith et al. 2011). This outbreak affected 8 people, of which two died. The patient samples were isolated between the 6th and 11th of May and it was believed that the cause was infected beef meat from a meals-on-wheels delivery.

A notable outbreak investigated by SSI occurred between September 2013 and October 2014 (Kvistholm Jensen et al. 2016). This LIST outbreak involved a total of 41 cases, resulting in 17 deaths. Deli meat products from a specific company were identified as the source of the outbreak. The high mortality rate in this outbreak may be attributed to the consumption of these products in nursing homes and hospitals, where patients are more vulnerable. Following the discovery of *Listeria* at a production facility, the Danish Veterinary and Food Administration (DVFA) recalled all products from the company. Following the outbreak, a series of control activities and research project were initiated to improve the control and management of *Listeria* in food production (Takeuchi-Storm et al. 2023).

In another LIST outbreak investigated by SSI, the source was traced back to cold-smoked and cured salmon products (Schjorring et al. 2017). A total of 5 related cases were identified, with 4 of them occurring in August 2017, and the fifth case in May 2017.

In some cases, despite extensive investigations, the source of contamination in an outbreak cannot always be identified. Such was the case in an unresolved outbreak that took place between the 13th of May and the 6th of June, 2022. During this period, a total of nine cases were infected with the same type of *Listeria*, with the majority of affected patients located in the Capital Region of Denmark. Despite thorough efforts, the specific source of contamination remained unknown.

Early identification of outbreaks caused by *Listeria monocytogenes* is crucial to implement timely interventions and mitigate the impact of the disease. Otherwise, these outbreaks can persist over an extended period. SSI has successfully resolved several long-spanned outbreaks in the last decade. For example, one investigation revealed that a single outbreak was actually two simultaneous outbreaks both caused by the consumption of smoked fish. Each outbreak consisted of ten cases and spanned from May 2013

to July 2015 (Gillesberg Lassen et al. 2016).

Other documented (Helwigh and Müller 2018; Helwigh, Petersen, and Torpdahl 2020; Lassen et al. 2023) long-spanned outbreaks investigated by SSI include:

- A cold-smoked fish outbreak with 9 cases spanning from December 2016 to February 2019.
- A prolonged outbreak with 6 cases from 2016 to 2019, traced back to a local green-grocer.
- An outbreak with 8 cases from October 2021 to June 2022, caused by a deli meat product.
- Two unresolved outbreaks with 9 cases and 12 cases from the end of 2018 to November 2021 and October 2020 to May 2022, respectively.

The use of WGS in these outbreaks provided the ability to link cases that occurred over a period of years and revealed that they were, in fact, continuous-source outbreaks.

In a recent outbreak investigated by SSI, DVFA, and the National Food Institute at the Technical University of Denmark (DTU), fish patties were identified as the source of contamination. This outbreak occurred from August 2022 to December 2022 and affected a total of 11 cases (Lassen et al. 2023). In Figure 3.3 a timeline plot indicating the start and duration of outbreaks investigated and documented by SSI is visualized.

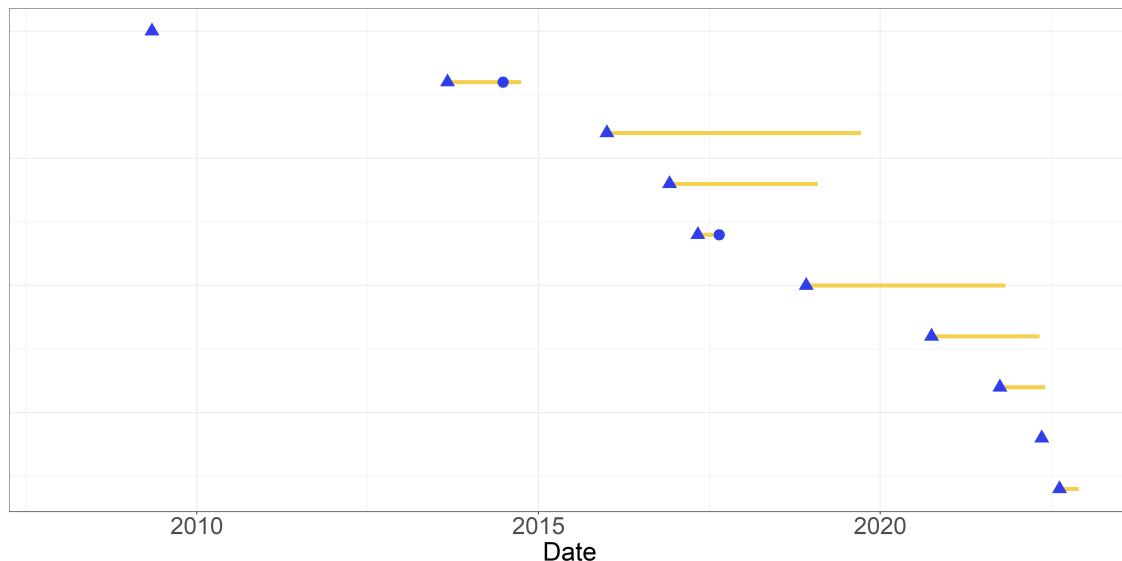


Figure 3.3: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of LIST by SSI. Time of detection (circle) is indicated when the information is available.

Evidently, the figure shows that many of the LIST outbreaks are long-spanned and overlap each other.

3.2.2 Epidemiological background and notable outbreaks of *Shigellosis*

Shigellosis (SHIL) is a diarrheal illness that is caused by a group of bacteria called *Shigella*. There are four types of *Shigella* bacteria, namely: *Shigella dysenteriae*, *Shigella boydii*, *Shigella flexneri*, and *Shigella sonnei*. The latter is the most common species in Denmark.

The bacteria are highly contagious and can be transmitted through direct person-to-person contact, consumption of contaminated food, or ingestion of water contaminated with human feces. SHIL infections are most commonly observed in children under the age of 5, individuals traveling to regions with poor sanitation and unsafe water and food practices, as well as gay, bisexual, and other men who have sex with men.

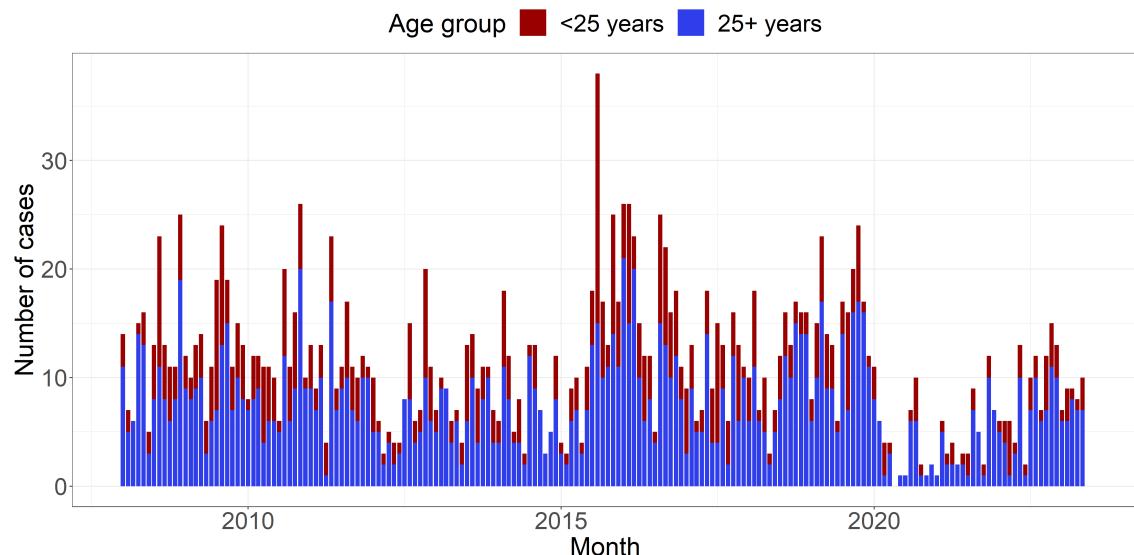


Figure 3.4: A stacked bar graph illustrating the number of monthly SHIL cases observed in the period from 2008 to 2022 for the age groups below 25 years and 25 years and above.

In Denmark, another significant cause of SHIL outbreaks is the importation of contaminated vegetables. This was evident in several incidents, including a 2007 outbreak where 215 individuals fell ill after consuming imported contaminated baby corn (Lewis et al. 2009), a smaller outbreak in 2009 linked to sugar snap peas from Kenya (Muller et al. 2009), and a 2020 outbreak associated with fresh mint as the source of infection (Helwig, Petersen, and Torpdahl 2020).

The 2020 outbreak is indeed a significant focus of this study, as it serves as a benchmark for evaluating the effectiveness of outbreak detection algorithms. It took place from the 22nd of August to the 9th of September and was investigated by SSI in collaboration with the DVFA and the National Food Institute at DTU. The outbreak affected 44 patients, mainly concentrated in the Capital Region of Denmark. The investigation identified at least five independent events where individuals subsequently developed SHIL.

3.2.3 Epidemiological background and notable outbreaks of Shiga toxin (verotoxin)-producing *Escherichia coli*

Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC) primarily spreads through contaminated food. Less common sources of infection include contaminated drinking and bathing water, as well as direct or indirect contact with infected animals. Cattle and other ruminants are primary reservoirs for STEC serotypes that are frequently associated with human disease (Menge 2020). Therefore, in Denmark, the source of infection is often products derived from beef, non-heat-treated dairy products, or other foods such as ready-to-eat vegetables, leafy greens, vegetable sprouts, and berries contaminated with feces from cows. *Hemolytic uremic syndrome* (HUS) is a severe complication that, in some cases, particularly in children, can develop following an infection with STEC.

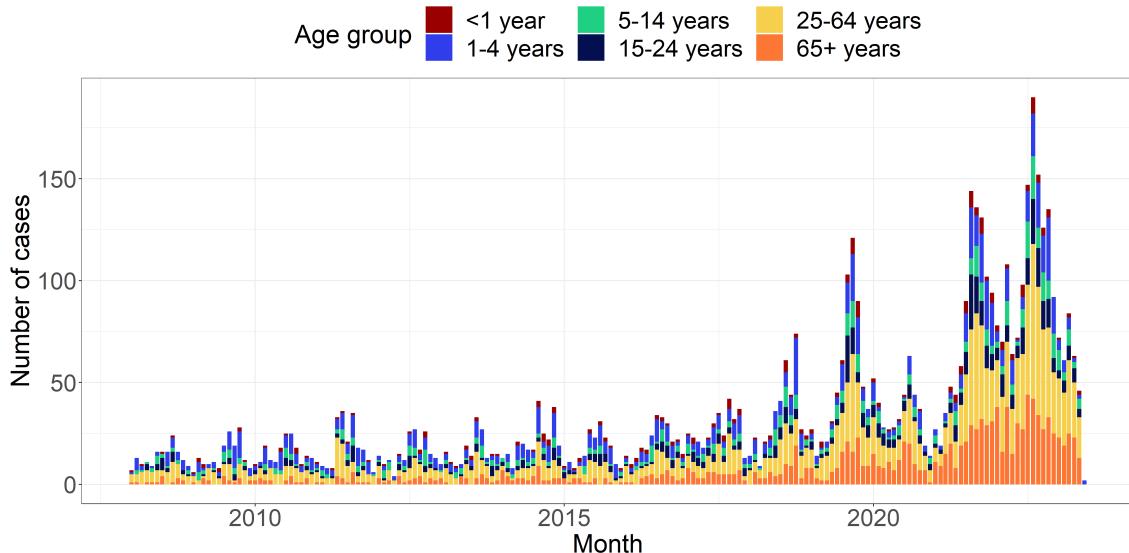


Figure 3.5: A stacked bar graph illustrating the number of monthly STEC cases observed in the period from 2008 to 2022 for the six age groups.

In general, stool samples are commonly used for diagnostic purposes in cases of STEC infections. Until 2018, most clinical microbiology departments relied on culture-based methods to detect and identify STEC bacteria in stool samples. However, in recent years, PCR methods have been increasingly adopted as a replacement for culture-based methods in the diagnosis of STEC infections (Svendsen et al. 2023). PCR methods offer advantages such as increased sensitivity and faster turnaround time, contributing to their growing popularity in clinical laboratories.

It is important to note that not all patients are routinely tested for STEC, and therefore, physicians need to specifically request STEC testing when submitting stool samples.

One of the earliest documented STEC outbreaks occurred in 2007, involving 18 laboratory-confirmed cases over a six-week period. The outbreak primarily affected children in day-care settings, and most patients experienced mild symptoms without bloody diarrhea. Investigations indicated a specific brand of organic beef sausage as the likely source of infection.

In September to October 2012, a STEC outbreak with a high risk of HUS was observed. Thirteen cases were diagnosed, with eight individuals developing HUS. Epidemiological investigations suggested that ground beef was the vehicle of the outbreak (Soborg et al. 2013; Helwigh, Sørensen, and Ethelberg 2013).

More recent outbreaks include a 38-case outbreak from September to November 2018, with a suspected association with beef sausage as the source of infection (Helwigh, Petersen, and Müller 2019). Additionally, there were two unresolved outbreaks with 11 and 14 cases occurring from May to July 2019 and from December 2021 to January 2022, respectively. The latter outbreak included three cases of HUS.

3.2.4 Epidemiological background and notable outbreaks of *Salmonellosis*

Salmonellosis (SALM) is a bacterial disease that primarily affects the intestinal tracts of humans. The *Salmonella* bacteria are commonly found in the intestines of animals and

humans and are excreted in feces. Human infection typically occurs through the consumption of contaminated food or water. *Salmonella* infections are often associated with the consumption of raw or under cooked meat, poultry, eggs or egg products, as well as unpasteurized milk.

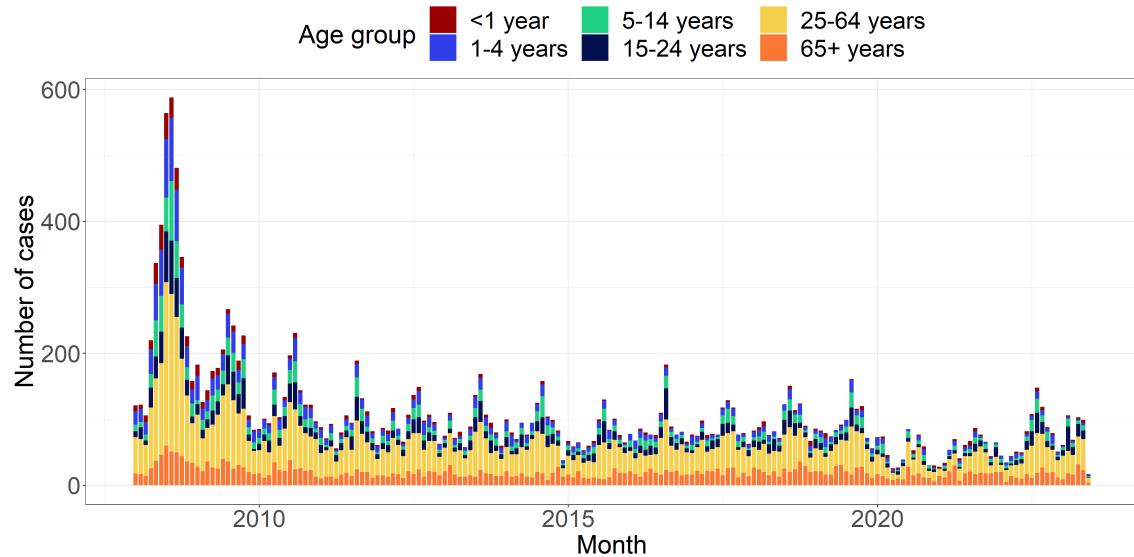


Figure 3.6: A stacked bar graph illustrating the number of monthly SALM cases observed in the period from 2008 to 2022 for the six age groups.

It is worth noting that an increasing proportion of infections in Denmark are now observed in connection with international travel, particularly since *Salmonella* has been eliminated from commercial chicken flocks in Denmark, making Danish eggs and poultry meat free from the bacteria. However, imported meat products can still pose a risk of contamination.

Numerous clusters of SALM are reported to FUD every year, and the disease remains a major public health issue. This was also evident in one of the longest lasting outbreaks with *Salmonella*. A total of 172 cases were reported between March and September 2010 (Kuhn et al. 2013; Helwigh, Krogh, and Müller 2011). Multiple other outbreaks are investigated and well-documented by SSI. However, it is not feasible to describe all of them in detail. For a full overview over the clusters of SALM reported to FUD, refer to Table C.1. Also, Figure 3.7 shows a timeline plot for some of the outbreaks of SALM documented by SSI and possible collaborators.

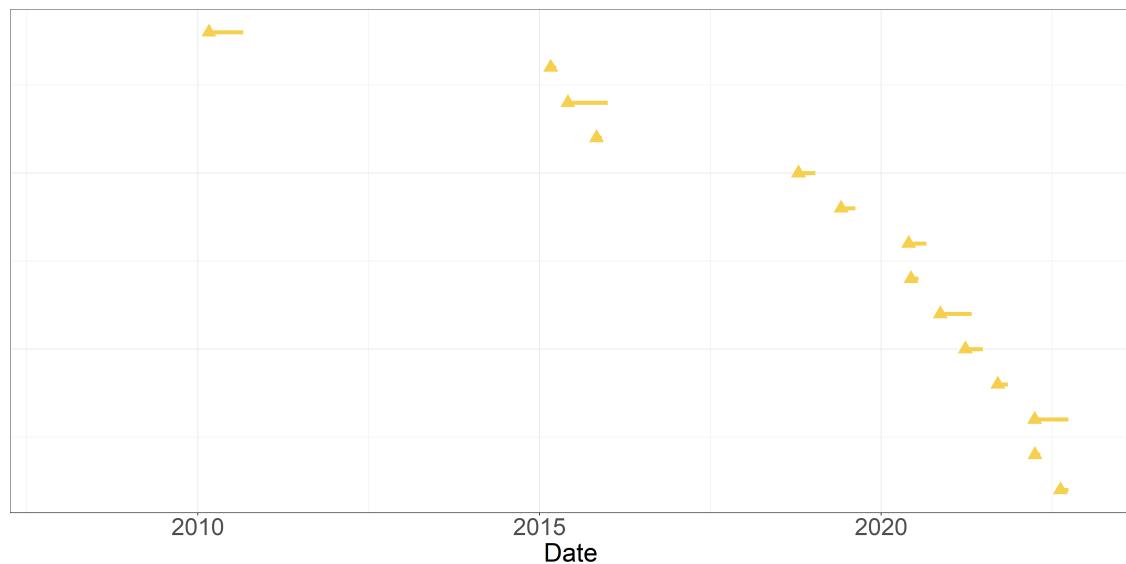


Figure 3.7: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of SALM by SSI and possible collaborators.

4 Methods

In this chapter, the current state-of-the-art methods for disease outbreak detection will be outlined (Section 4.1). Furthermore, the novel outbreak detection algorithm will be introduced, along with the theory related to generalized mixed effects models (Section 4.3) and hierarchical generalized linear models (Section 4.4). The presentation of theory related to these modeling frameworks are mostly inspired by Madsen and Thyregod (2011).

In this thesis, these modeling frameworks are utilized to analyze the count observations denoted as y , but more importantly they play a crucial role in assessing the unobserved random variables or random effects represented by u , which are directly employed in the detection algorithm for characterizing outbreaks.

Due to the complexity of generalized mixed effects models, obtaining closed-form solutions is generally not feasible. Therefore, an overview of the Laplace approximation technique will be provided in Section 4.3, which allows for approximating the likelihood function in these models. Additionally, the implementation of these models in the R programming language will be presented in Section 4.5. Finally, an outline of the *logarithmic score*, which is used to objectively evaluate the performance of the models in the novel method, is presented in Section 4.6.

4.1 State-of-the-art outbreak detection algorithm

In this section, the original method introduced by Farrington et al. (1996) and the subsequent improvements proposed by Noufaily et al. (2013) will be outlined. These methods are recognized as the current state-of-the-art for disease outbreak detection and will be used as benchmarks to evaluate the performance of the novel outbreak detection algorithm proposed in this thesis. Both of the aforementioned methods have been implemented in the R package called **surveillance** by Salmon, Schumacher, and Höhle (2016), which can be accessed from the Comprehensive R Archive Network (CRAN) at <https://cran.r-project.org/web/packages/surveillance/index.html>. The presentation of the two methods are strongly inspired by Salmon, Schumacher, and Höhle (2016).

Both methods follow the same steps in the algorithm. The first step involves fitting an over-dispersed Poisson Generalized Linear Model (GLM) with a log link to the reference data. In this model, the baseline count y_t corresponding to the baseline time point t is assumed to have an expected value λ_t and a variance $\phi\lambda_t$, where $\phi \geq 1$ is ensured to account for over-dispersion. The systematic component of the model includes only a linear time trend in the frequency of reports. Therefore, the systematic component can be expressed as

$$\log(\lambda_t) = \alpha + \beta t \quad (4.1)$$

The original method incorporated seasonal effects by considering counts from comparable periods in past years for threshold calculation. This approach is similar to the one used by Stroup et al. (1989). The baseline weeks, which are used as reference, are determined by two integers: b the number of years back, and w the window half-width. For a given current week x of year h , only data from weeks $x - w$ to $x + w$ of years $h - b$ to $h - 1$ are considered, resulting in a total of $n = b(2w + 1)$ baseline weeks. The default values are $b = 5$ and $w = 3$, resulting in a total of $n = 35$ baseline values.

However, Noufaily et al. (2013) demonstrated that the algorithm performs better when utilizing more historical data without disregarding seasonality. To achieve this, the author introduced a 10-level factor with a 7-week reference period and nine additional 5-week periods in each year. As a result, the systematic component of the model is modified as follows

$$\log(\lambda_t) = \alpha + \beta t + \delta_{j(t)} \quad (4.2)$$

In this equation, $j(t)$ represents the seasonal factor level corresponding to time point t . The reference week t_0 is always associated with the reference seasonal level, denoted by $j(t_0) = 0$ and $\delta_0 = 0$.

The idea of incorporating more data while preserving seasonality has been further expanded in the implementation of the method in the **surveillance** R package. The package allows the user to choose an arbitrary number of periods in each year. Consequently, the systematic component is adjusted as follows

$$\log(\lambda_t) = \alpha + \beta t + \delta_{c(t)} \quad (4.3)$$

In this equation, $c(t)$ represents the coefficients of a zero-order spline with `noPeriods + 1` knots. It can be conveniently represented as a `noPeriods`-level factor that captures seasonality. The function $c(t)$ indicates which season or period of the year t belongs to.

Furthermore, Noufaily et al. (2013) demonstrated that it is beneficial to exclude the last 26 weeks before t_0 from the baseline calculation. This exclusion helps prevent a reduction in sensitivity when an outbreak has recently started before t_0 .

In the second step, the algorithm predicts the expected number of counts λ_{t_0} for the current time point t_0 using the fitted GLM. Both methods differ in their assumptions for calculating the upper bound U_{t_0} .

The original method assumes that a transformation of the prediction error, denoted as $g(y_{t_0} - \hat{\lambda}_{t_0})$, follows a normal distribution. For example, when using the identity transformation $g(x) = x$, the assumption becomes

$$y_{t_0} - \hat{\lambda}_{t_0} \sim N(0, V(y_{t_0} - \hat{\lambda}_{t_0})) \quad (4.4)$$

The upper bound of the prediction interval is then calculated based on this distribution. The variance of the prediction error is given by

$$V(y_{t_0} - \hat{\lambda}_{t_0}) = V(y_{t_0}) + V(\hat{\lambda}_{t_0}) = \phi \lambda_{t_0} + V(\hat{\lambda}_{t_0}) \quad (4.5)$$

Here, $V(\hat{y}_{t_0})$ represents the variance of an observation, and $V(\hat{\lambda}_{t_0})$ represents the variance of the estimate. The threshold, defined as the upper bound of a one-sided $(1 - \alpha) \cdot 100\%$ prediction interval, is calculated as

$$U_{t_0} = \hat{\lambda}_{t_0} + z_{1-\alpha} \hat{V} \left(\sqrt{y_{t_0} - \hat{\lambda}_{t_0}} \right) \quad (4.6)$$

However, this method's weakness lies in the assumption of normality itself. Therefore, an alternative assumption was presented in Noufaily et al. (2013). This approach assumes that y_{t_0} follows a Negative Binomial distribution, denoted as $\text{NB}(\lambda_{t_0}, \nu)$, where λ_{t_0} represents the mean and $\nu = \frac{\lambda_{t_0}}{\phi-1}$ represents the over-dispersion parameter. In this parameterization, the expected value of y_t remains λ_t , and the variance of y_t is $\phi\lambda_t$, with $\phi > 1$. If $\phi \leq 1$, a Poisson distribution is assumed for the observed count. The threshold is defined as a quantile of the Negative Binomial distribution using the plug-in estimates $\hat{\lambda}_{t_0}$ and $\hat{\phi}$.

In the final step, the observed count y_{t_0} is compared to the upper bound U_{t_0} , and an alarm is raised if $y_{t_0} > U_{t_0}$. The fitting of the GLM in both methods involves three important steps.

First, the algorithm optionally performs a power transformation to correct for skewness and stabilize the variance of the data.

Next, the significance of the time trend is checked. The time trend is included in the model only if it is statistically significant at a chosen significance level, there are more than three years of reference data, and there is no over-extrapolation due to the time trend.

Finally, past outbreaks are reweighted based on their Anscombe residuals. If the Anscombe residual of a count exceeds a certain weight threshold, it is re-weighted in a second fitting of the GLM. In the original method by Farrington et al. (1996), a re-weighting threshold of 1 was used. However, Noufaily et al. (2013) suggests using a value of 2.56 for the weight threshold to make the re-weighting procedure less drastic, as it also reduces the variance of the observations. For a more detailed description of the re-weighting scheme employed in the Farrington and Noufaily method, refer to Section 3.6 in Farrington et al. (1996) and Section 2.1 in Noufaily et al. (2013), respectively.

4.2 Novel outbreak detection algorithm

In this section, the novel algorithm for the prospective detection of disease outbreaks proposed in this thesis is outlined. The algorithm utilizes a generalized mixed effects model or a hierarchical generalized linear model as a modeling framework to model the count observations y and assess the unobserved random effects u . These random effects are used directly in the detection algorithm to characterize an outbreak. The theoretical foundations of these models will be further discussed in Sections 4.3 and 4.4.

The first step involves fitting either a hierarchical Poisson Normal or Poisson Gamma model with a log link to the reference data. Here, it is possible for the user to include an arbitrary number of covariates by supplying a model formula. In order to account for structural changes in the time series, e.g. an improved and more sensitive diagnostic method or a new screening strategy at hospitals, a rolling window with width k is used to estimate the time-varying model parameters. Also, it is assumed that the count is proportional to the population size n . Hence, in terms of the canonical link the model for the fixed effects is

$$\log(\lambda_{it}) = \mathbf{x}_{it}\boldsymbol{\beta} + \log(n_{it}), \quad i = 1, \dots, m, \quad t = 1, \dots, T \quad (4.7)$$

Here \mathbf{x}_{it} and $\boldsymbol{\beta}$ are p -dimensional vectors of covariates and fixed effects parameters respectively, where p denotes the number of covariates or fixed effects parameters, m denotes the number of groups, and T denotes the length of the period.

In the second step of the algorithm, as a new observation becomes available, the algorithm infers the one-step ahead random effect u_{it_1} for each group using the obtained

model estimates $\hat{\theta}_{t_0}$. Here, t_0 represents the current time point, and t_1 represents the one-step ahead time point. The threshold U_{t_0} for detecting outbreaks is defined as a quantile of the distribution of the random effects in the second stage model. This threshold can be calculated based on either a Gaussian distribution using the plug-in estimate $\hat{\sigma}_{t_0}$ or a Gamma distribution using the plug-in estimate $\hat{\phi}_{t_0}$. The choice of distribution depends on the specific modeling framework and assumptions used in the analysis.

In the final step, the inferred random effect \hat{u}_{it_1} is compared to the upper bound U_{t_0} , and an alarm is raised if $\hat{u}_{it_1} > U_{t_0}$. If an outbreak is detected, the related observation y_{it_1} is omitted from the parameter estimation in the future. Thus, resulting in a smaller sample size for the rolling window until that specific observation is discarded.

The above algorithm is implemented in an R function, which can be accessed from the associated GitHub repository at `src/models/aeddo.R`.

4.3 General mixed effects models

In this section selected theory related to generalized mixed effects models is presented. Moreover, an outline of the Laplace approximation is included.

The general mixed effects model can be represented by its likelihood function

$$L_M(\boldsymbol{\theta}; \mathbf{y}) = \int_{\mathbb{R}^q} L(\boldsymbol{\theta}; \mathbf{u}, \mathbf{y}) d\mathbf{u} \quad (4.8)$$

where \mathbf{y} is the observed random variable, $\boldsymbol{\theta}$ is the model parameters to be estimated and \mathbf{U} is the q unobserved random variables. The likelihood function L is the joint likelihood of both the observed and the unobserved random variables. The likelihood function for estimating $\boldsymbol{\theta}$ is the marginal likelihood L_M obtained by integrating out the unobserved random variables. In general it is difficult to solve the integral in (4.8) if the number of unobserved random variables is more than a few and hence numerical methods must be used. Thus, an outline of the Laplace approximation is included in this section.

4.3.1 Hierarchical models

It is useful to formulate the model as a hierarchical model containing a *first stage model*

$$f_{Y|u}(\mathbf{y}; \mathbf{u}, \boldsymbol{\beta}) \quad (4.9)$$

which is a model for the observed random variables given the unobserved random variables, and a *second stage model*

$$f_U(\mathbf{u}; \boldsymbol{\Psi}) \quad (4.10)$$

which is a model for the unobserved random variables. Here $\boldsymbol{\beta}$ represent the fixed effects parameters and $\boldsymbol{\Psi}$ is a model parameter. The total set of parameters is $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\Psi})$. Hence the joint likelihood is given as

$$L(\boldsymbol{\beta}, \boldsymbol{\Psi}; \mathbf{u}, \mathbf{y}) = f_{Y|u}(\mathbf{y}; \mathbf{u}, \boldsymbol{\beta}) f_U(\mathbf{u}; \boldsymbol{\Psi}) \quad (4.11)$$

To obtain the likelihood for the model parameters $(\boldsymbol{\beta}, \boldsymbol{\Psi})$ the unobserved random variables are integrated out. The likelihood function for estimating $(\boldsymbol{\beta}, \boldsymbol{\Psi})$ is as in (4.8) the marginal likelihood

$$L_M(\boldsymbol{\beta}, \boldsymbol{\Psi}; \mathbf{y}) = \int_{\mathbb{R}^q} L(\boldsymbol{\beta}, \boldsymbol{\Psi}; \mathbf{u}, \mathbf{y}) d\mathbf{u} \quad (4.12)$$

where q is the number of unobserved random variables, and $\boldsymbol{\beta}$ and $\boldsymbol{\Psi}$ are the parameters to be estimated.

4.3.2 Laplace Approximation

The Laplace approximation will be outlined in the following. A thorough description of the Laplace approximation in nonlinear mixed effects models is found in Wolfinger and Lin (1997).

For a given set of model parameters $\boldsymbol{\theta}$ the joint log-likelihood $\ell(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y}) = \log(L(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y}))$ is approximated using a second order Taylor approximation around the optimum $\tilde{\mathbf{u}} = \hat{\mathbf{u}}_{\boldsymbol{\theta}}$ of the log-likelihood function w.r.t. the unobserved random variables \mathbf{u} , i.e.,

$$\ell(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y}) \approx \ell(\boldsymbol{\theta}, \tilde{\mathbf{u}}, \mathbf{y}) - \frac{1}{2}(\mathbf{u} - \tilde{\mathbf{u}})^T \mathbf{H}(\tilde{\mathbf{u}})(\mathbf{u} - \tilde{\mathbf{u}}) \quad (4.13)$$

where the first-order term of the Taylor expansion disappears since the expansion is done around the optimum $\tilde{\mathbf{u}}$ and $\mathbf{H}(\tilde{\mathbf{u}}) = -\ell''_{uu}(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y})|_{\mathbf{u}=\tilde{\mathbf{u}}}$ is the negative Hessian of the joint log-likelihood evaluated at $\tilde{\mathbf{u}}$.

It is readily seen that the joint log-likelihood for the hierarchical model specified in (4.9) and (4.10) is

$$\ell(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y}) = \ell(\boldsymbol{\beta}, \boldsymbol{\Psi}, \mathbf{u}, \mathbf{y}) = \log f_{Y|u}(\mathbf{y}; \mathbf{u}, \boldsymbol{\beta}) + \log f_U(\mathbf{u}; \boldsymbol{\Psi}) \quad (4.14)$$

which implies that the Laplace approximation becomes

$$\ell_{M,LA}(\boldsymbol{\theta}, \mathbf{y}) = \log f_{Y|u}(\mathbf{y}; \tilde{\mathbf{u}}, \boldsymbol{\beta}) + \log f_U(\tilde{\mathbf{u}}, \boldsymbol{\Psi}) - \frac{1}{2} \log \left| \frac{\mathbf{H}(\tilde{\mathbf{u}})}{2\pi} \right| \quad (4.15)$$

4.3.3 Formulation of the generalized mixed effects model

The generalized mixed effects model utilized in this thesis to model the count observations \mathbf{y} and assess the random effects \mathbf{u} is presented in Definition 4.1, along with the joint likelihood function for the first and second stage models.

Definition 4.1 (Hierarchical Poisson Normal model). In order to simplify the notation, the probability density functions are presented for a specific observation. Hence, the subscripts indicating the group and time are omitted. The conditional distribution of the count observations is assumed to be a Poisson distribution with intensities λ

$$f_{Y|u}(y; u, \boldsymbol{\beta}) = \frac{\lambda \exp(u)^y}{y!} \exp(-\lambda \exp(u)) \quad (4.16)$$

Also, it is assumed that the count is proportional to the population size n . Hence, in terms of the canonical link for the Poisson distribution the model for the fixed effects is

$$\log(\lambda_{it}) = \mathbf{x}_{it}\boldsymbol{\beta} + \log(n_{it}), \quad i = 1, \dots, m, \quad t = 1, \dots, T \quad (4.17)$$

The probability density function for the distribution of the random effects is assumed to follow a zero mean Gaussian distribution, $\mathbf{u} \sim \mathbf{N}(\mathbf{0}, I\sigma^2)$, i.e.

$$f_U(u; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \quad (4.18)$$

where σ is a model parameter.

Henceforth, the total set of parameters are $\theta = (\beta, \sigma)$ and the model can be formulated as a two-level hierarchical model

$$Y|\mathbf{u} \sim \text{Pois}(\lambda \exp(\mathbf{u})) \quad (4.19a)$$

$$\mathbf{u} \sim \mathbf{N}(\mathbf{0}, I\sigma^2) \quad (4.19b)$$

The joint likelihood for the count observations y and the random effects \mathbf{u} becomes

$$L(\beta, \sigma; u_{it}, y_{it}) = \prod_{t=1}^T \prod_{i=1}^m \frac{(\lambda_{it} \exp(u_{it}))^{y_{it}}}{y_{it}!} \exp(-\lambda_{it} \exp(u_{it})) \prod_{t=1}^T \prod_{i=1}^m \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{u_{it}^2}{2\sigma^2}\right) \quad (4.20)$$

4.4 Hierarchical generalized linear models

In this section selected theory related to hierarchical generalized linear models is presented. The model class was initially formulated by Lee and Nelder (1996) as a natural generalization of the GLM to also incorporate random effects. A starting point in hierarchical modelling is an assumption that the distribution of random effects may be modeled by an exponential dispersion family. This family of models were first introduced by Fisher and Russell (1922), and has proven to play an important role in mathematical statistics because of their simple inferential properties. The exponential dispersion family considers a family of distributions, which can be written on the form

$$f_Y(y; \theta) = c(y, \phi) \exp(\phi\{\theta y - \kappa(\theta)\}) \quad (4.21)$$

Here the parameter $\phi > 0$ is called the *precision parameter*, which in some cases represents a shape parameter as for the Gamma distribution. In other cases the precision parameter represents an over-dispersion that is not related to the mean. These distributions combine with the so-called *standard conjugate distributions* in a simple way, and lead to marginal distributions that may be expressed in a closed form suited for likelihood calculations. For an introduction to the concept of *standard conjugate distributions* and the definition of a hierarchical generalized linear model, refer to Section 6.3 and Section 6.5 of Madsen and Thyregod (2011), respectively.

In general, when the conditional distribution of $Y|u$ is a Poisson distribution, and the conjugate distribution is assumed to be a Gamma distribution with mean value 1, it follows that the distribution of Y is a Negative Binomial distribution.

To further motivate this choice of distribution for the second stage model, an illustrative example focusing on *Listeriosis* (LIST) is presented.

Example 4.1 (Variation between observed cases of LIST). Usually, it is reasonable to assume a Poisson distribution for count data, where the expected value and variance are both equal to λ . However, it should be noted that this assumption may not always hold true when modeling count data.

Table 4.1: The distribution of observations with 1, 2, ⋯ , 10+ cases of LIST.

Number of LIST cases	Number of times observed	Poisson expected	Negative binomial expected
0	4	6.26	14.44
1	5	15.52	21.41
2	28	25.66	24.75
3	25	31.83	24.58
4	26	31.58	22.01
5	34	26.11	18.27
6	19	18.51	14.31
7	14	11.48	10.72
8	10	6.33	7.73
9	3	3.14	5.41
10+	12	2.34	10.46

Take, for example, the monthly cases of LIST analyzed in this thesis. Table 4.1 presents the distribution of these cases from 2008 to 2022, along with the expected numbers from both the Poisson and Negative Binomial distributions. The distribution of cases over time is visualized in Figure 4.1.

As observed in both the table and figure, the empirical distribution of LIST has notably heavier tails compared to the Poisson distribution. Additionally, the mean of the Poisson distribution for the monthly LIST cases is $\hat{\lambda} = \bar{y} = 4.96$, while the empirical variance is $s^2 = 10$, which is considerably larger than the mean. These findings indicate overdispersion, suggesting that a Negative Binomial model would be more appropriate for modeling the data.

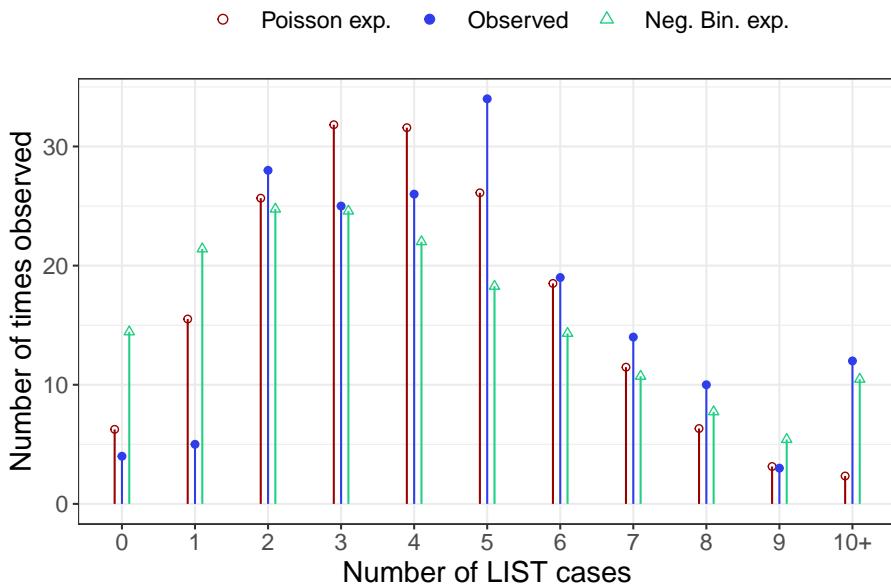


Figure 4.1: Empirical distribution of monthly LIST cases together with the expected numbers from both the Poisson and Negative Binomial distributions

4.4.1 Formulation of the hierarchical model

The hierarchical model used in this thesis to model the count observations y and assess the random effects u is presented in Definition 4.2, along with the derivation of the marginal distribution of Y and the joint likelihood function.

Definition 4.2 (Compound Poisson Gamma model). In the compound Poisson Gamma model the conditional distribution of the count observations are assumed to be a Poisson distribution with intensities λ

$$f_{Y|u}(y; u, \beta) = \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \quad (4.22)$$

The probability density function for the random effects u are assumed to follow a reparametrized Gamma distribution with mean 1, $u \sim G(1/\phi, \phi)$ that is

$$f_u(u; \phi) = \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi}\right)^{1/\phi-1} \exp(-u/\phi) \quad (4.23)$$

Subsequently, the model can be formulated as a two-level hierarchical model

$$Y|u \sim \text{Pois}(\lambda u) \quad (4.24a)$$

$$u \sim G(1/\phi, \phi) \quad (4.24b)$$

Given (4.22) and (4.23), the probability function for the marginal distribution of Y is determined from

$$\begin{aligned}
g_Y(y; \beta, \phi) &= \int_{u=0}^{\infty} f_{Y|u}(y; u, \beta) f_u(u; \phi) du \\
&= \int_{u=0}^{\infty} \frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi}\right)^{1/\phi-1} \exp(-u/\phi) du \\
&= \frac{\lambda^y}{y! \Gamma(1/\phi) \phi^{1/\phi}} \int_{u=0}^{\infty} u^{y+1/\phi-1} \exp(-u(\lambda\phi + 1)/\phi) du
\end{aligned} \tag{4.25}$$

In (4.25) it is noted that the integrand is the kernel in the probability density function for a Gamma distribution, $G(y + 1/\phi, \phi/(\lambda\phi + 1))$. As the integral of the density shall equal one, it is found by adjusting the norming constant that

$$\int_{u=0}^{\infty} u^{y+1/\phi-1} \exp\left(-u/(\phi/(\lambda\phi + 1))\right) du = \frac{\phi^{y+1/\phi} \Gamma(y + 1/\phi)}{(\lambda\phi + 1)^{y+1/\phi}} \tag{4.26}$$

Therefore, it is shown that the marginal distribution of Y is a Negative Binomial distribution, $Y \sim NB(1/\phi, 1/(\lambda\phi + 1))$. The probability function for Y is

$$\begin{aligned}
P[Y = y] &= g_Y(y; \beta, \phi) \\
&= \frac{\lambda^y}{y! \Gamma(1/\phi) \phi^{1/\phi}} \frac{\phi^{y+1/\phi} \Gamma(y + 1/\phi)}{(\lambda\phi + 1)^{y+1/\phi}} \\
&= \frac{\Gamma(y + 1/\phi)}{\Gamma(1/\phi) y!} \frac{1}{(\lambda\phi + 1)^{1/\phi}} \left(\frac{\lambda\phi}{\lambda\phi + 1}\right)^y \\
&= \binom{y + 1/\phi - 1}{y} \frac{1}{(\lambda\phi + 1)^{1/\phi}} \left(\frac{\lambda\phi}{\lambda\phi + 1}\right)^y, \quad \text{for } y = 0, 1, 2, \dots
\end{aligned} \tag{4.27}$$

where the following convention is used

$$\binom{z}{y} = \frac{\Gamma(z+1)}{\Gamma(z+1-y)y!} \tag{4.28}$$

for z real and y integer values. Consequently, the mean and variance of Y are given by

$$E[Y] = \lambda \quad V[Y] = \lambda(\lambda\phi + 1) \tag{4.29}$$

The joint likelihood function for estimating (β, ϕ) is

$$L(\beta, \phi; y_{it}) = \prod_{t=1}^T \prod_{i=1}^m \binom{y_{it} + 1/\phi - 1}{y_{it}} \frac{1}{(\lambda_{it}\phi + 1)^{1/\phi}} \left(\frac{\lambda_{it}\phi}{\lambda_{it}\phi + 1}\right)^{y_{it}} \tag{4.30}$$

4.4.2 Inference on individual groups

Consider the compound Poisson Gamma model in (4.24), and assume that a value $Y = y$ has been observed.

The conditional distribution of u for given $Y = y$ is found using Bayes Theorem. In order to simplify the notation, the subscript indicating the group and time are omitted.

$$\begin{aligned}
g_u(u|Y=y) &= \frac{f_{y,u}(y,u)}{g_Y(y; \lambda, \phi)} \\
&= \frac{f_{y|u}(y; u) g_u(u)}{g_Y(y; \lambda, \phi)} \\
&= \frac{1}{g_Y(y; \lambda, \phi)} \left(\frac{(\lambda u)^y}{y!} \exp(-\lambda u) \frac{1}{\phi \Gamma(1/\phi)} \left(\frac{u}{\phi} \right)^{1/\phi-1} \exp(-u/\phi) \right) \\
&\propto u^{y+1/\phi-1} \exp(-u(\lambda\phi + 1)/\phi)
\end{aligned} \tag{4.31}$$

Here, the *kernel* of the probability density function is identified

$$u^{y+1/\phi-1} \exp(-u(\lambda\phi + 1)/\phi) \tag{4.32}$$

as the kernel of a Gamma distribution, $G(y + 1/\phi, \phi/(\lambda\phi + 1))$, i.e. the conditional distribution of u for given $Y = y$ can be written as

$$u|Y=y \sim G(y + 1/\phi, \phi/(\lambda\phi + 1)) \tag{4.33}$$

The mean of the conditional distribution is given by:

$$\mathbb{E}[u|Y=y] = \frac{y\phi + 1}{\lambda\phi + 1} \tag{4.34}$$

And the variance of the conditional distribution is:

$$\text{V}[u|Y=y] = \frac{(\phi^2 + \phi)}{(\lambda\phi + 1)^2} \tag{4.35}$$

These formulas provide the mean and variance of the conditional distribution of u given the observed value $Y = y$.

4.4.3 Why is the Gamma distribution chosen to represent the variation between months?

The choice of the Gamma distribution for modeling the random effects has been motivated by several reasons. Firstly, the support of the Gamma distribution, which ranges from 0 to infinity, aligns with the mean-value space, denoted as \mathcal{M} , for the Poisson distribution. This ensures that the random effects are constrained within a meaningful range for the underlying Poisson process.

Secondly, the two-parameter family of Gamma distributions offers considerable flexibility, encompassing a wide range of shapes and distributions that can span from exponential-like distributions to fairly symmetrical distributions on the positive real line. This flexibility allows the model to capture various patterns and characteristics observed in the data.

Additionally, the choice of the Gamma distribution has benefits in terms of the derivation of the marginal distribution of the response variable Y . The kernel $u^{\alpha-1} \exp(-u/\beta)$ of the Gamma distribution used for modeling the random effects exhibits a similar structure to the kernel $u^y \exp(-u)$ of the likelihood function corresponding to the sampling distribution of

Y . This similarity facilitates the analytical computation of the integral involved in deriving the marginal distribution, as it can be expressed in terms of known functions.

Overall, the Gamma distribution is selected due to its alignment with the mean-value space of the Poisson distribution, its flexibility in capturing diverse distributions, and its analytical convenience in computing the marginal distribution of the response variable.

4.5 Parameter estimation

In this section, the parameter estimation and implementation of the models used in the novel outbreak detection algorithm are presented. These model are implemented in R using the open-source R package **TMB** (Template Model Builder) developed by Kristensen et al. (2016). This package facilitates efficient maximum likelihood estimation and uncertainty calculations for the parameter set $\theta = (\beta, \Psi)$ and random effects u . The presentation of the parameter estimation conducted in **TMB** is strongly inspired by Chapter 2 in Kristensen et al. (2016) and Section 5.10 in Madsen and Thyregod (2011).

TMB maximizes a user-provided objective function in the form of a C++ template, to estimate the maximum likelihood for the parameter set $\theta = (\beta, \Psi)$. In the following code chunk, the C++ template for the hierarchical Poisson normal model specified in (4.19), is shown:

```
#include <TMB.hpp>
template<class Type>
Type objective_function<Type>::operator() ()
{
    // R input data
    DATA_VECTOR(y);                                // Count data
    DATA_VECTOR(x);                                // Population size
    DATA_MATRIX(X);                                // Design matrix
    PARAMETER_VECTOR(u);                           // Random effects
    // Parameters
    PARAMETER_VECTOR(beta);                         // Fixed effects parameters
    PARAMETER(log_sigma_u);                        // Model parameter
    vector<Type> lambda = exp(X*beta-log(x)+u); // Construct 'lambda'
    Type sigma_u = exp(log_sigma_u);                // And the model parameters
    Type mean_ran = Type(0);
    // Objective function
    Type f = 0;                                     // Declare the objective
    f -= sum(dnorm(u,mean_ran,sigma_u,true));      // Calculate the objective
    f -= sum(dpois(y,lambda,true));                 // Calculate the objective
    return f;
}
```

Refer to the associated GitHub repository at `src/models/PoissonGamma.cpp` to access the C++ template for the hierarchical Poisson Gamma model. The objective function maximizes the marginal log-likelihood function, which integrates out the random effects u

$$\ell_M(\theta; y) = \int_{\mathbb{R}^q} \ell(\theta; u, y) du \quad (4.36)$$

where $\ell(\theta, u, y)$ is the joint log-likelihood function of the data given the parameters and random effects. The maximizer \hat{u}_θ of the joint log-likelihood $\ell(\theta; u, y)$ with respect to the

random effects \mathbf{u} is defined as:

$$\hat{\mathbf{u}}_{\boldsymbol{\theta}} = \arg \max_{\mathbf{u}} \ell(\boldsymbol{\theta}; \mathbf{u}, \mathbf{y}) \quad (4.37)$$

Using $H(\hat{\mathbf{u}}_{\boldsymbol{\theta}})$ to denote the negative Hessian of the joint log-likelihood evaluated at $\hat{\mathbf{u}}_{\boldsymbol{\theta}}$; i.e.,

$$H(\hat{\mathbf{u}}_{\boldsymbol{\theta}}) = -\ell''_{uu}(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y})|_{\mathbf{u}=\hat{\mathbf{u}}_{\boldsymbol{\theta}}} \quad (4.38)$$

The Laplace approximation for the marginal log-likelihood $\ell_M(\boldsymbol{\theta})$ is

$$\ell_{M,LA}(\boldsymbol{\theta}, \mathbf{y}) = \ell(\boldsymbol{\theta}, \mathbf{u}, \mathbf{y}) - \frac{1}{2} \log \left| \frac{H(\hat{\mathbf{u}}_{\boldsymbol{\theta}})}{2\pi} \right| \quad (4.39)$$

Our estimate of $\boldsymbol{\theta}$ minimizes the negative log of the Laplace approximation, i.e.,

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} -\ell_{M,LA}(\boldsymbol{\theta}, \mathbf{y}) \quad (4.40)$$

The minimization of the Laplace approximation for the marginal likelihood is then performed using conventional R optimization routines (e.g., BFGS) to optimize the objective and obtain our estimate $\hat{\boldsymbol{\theta}}$. Uncertainty of the estimate $\hat{\boldsymbol{\theta}}$, or any differentiable function of the estimate $\phi(\hat{\boldsymbol{\theta}})$, is obtained by the δ -method:

$$\text{V}(\phi(\hat{\boldsymbol{\theta}})) = -\phi'_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) \left(\Delta^2 \ell_{M,LA}(\hat{\boldsymbol{\theta}}, \mathbf{y}) \right)^{-1} \phi'_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}})^T \quad (4.41)$$

Additionally, **TMB** utilizes Automatic Differentiation (AD) techniques (Griewank and Walther 2008) to evaluate first, second, and potentially third-order derivatives. This approach enhances the computational efficiency and accuracy of the parameter estimation process in the implemented models. Therefore, even though the random effects are analytically integrated out in the hierarchical Poisson Gamma model, and the Laplace approximation is not needed, implementing the joint log-likelihood function in **TMB** can still result in more efficient computations. For a comprehensive introduction to the concept of AD, it is recommended to read Section 2.1 and Section 2.2 of Fournier et al. (2012).

It is important to note that the implementation of the hierarchical Poisson Normal model requires the use of the Laplace approximation to approximate the likelihood function. To investigate the impact on run time, the optimization of both the hierarchical Poisson Normal and Poisson Gamma models is compared using a simulated data set with $n = 1000$. The simulated series $Y = \{y_1, \dots, y_n\}$ follows a Poisson distribution with a rate of $\lambda = 5$, i.e., $Y \sim \text{Pois}(5)$. After generating the series, the following model for the fixed effects is optimized in both modeling frameworks:

$$\log(\lambda) = \beta_{\text{Intercept}} \quad (4.42)$$

Finally, the optimization routines for each of the modeling frameworks is run 100 times and the run time across these iterations are collected using the R package **microbenchmark**, which can be accessed from the CRAN at [https://CRAN.R-project.org/package=](https://CRAN.R-project.org/package=microbenchmark) microbenchmark. The resulting run times are visualized as a violin plot in Figure 4.2.

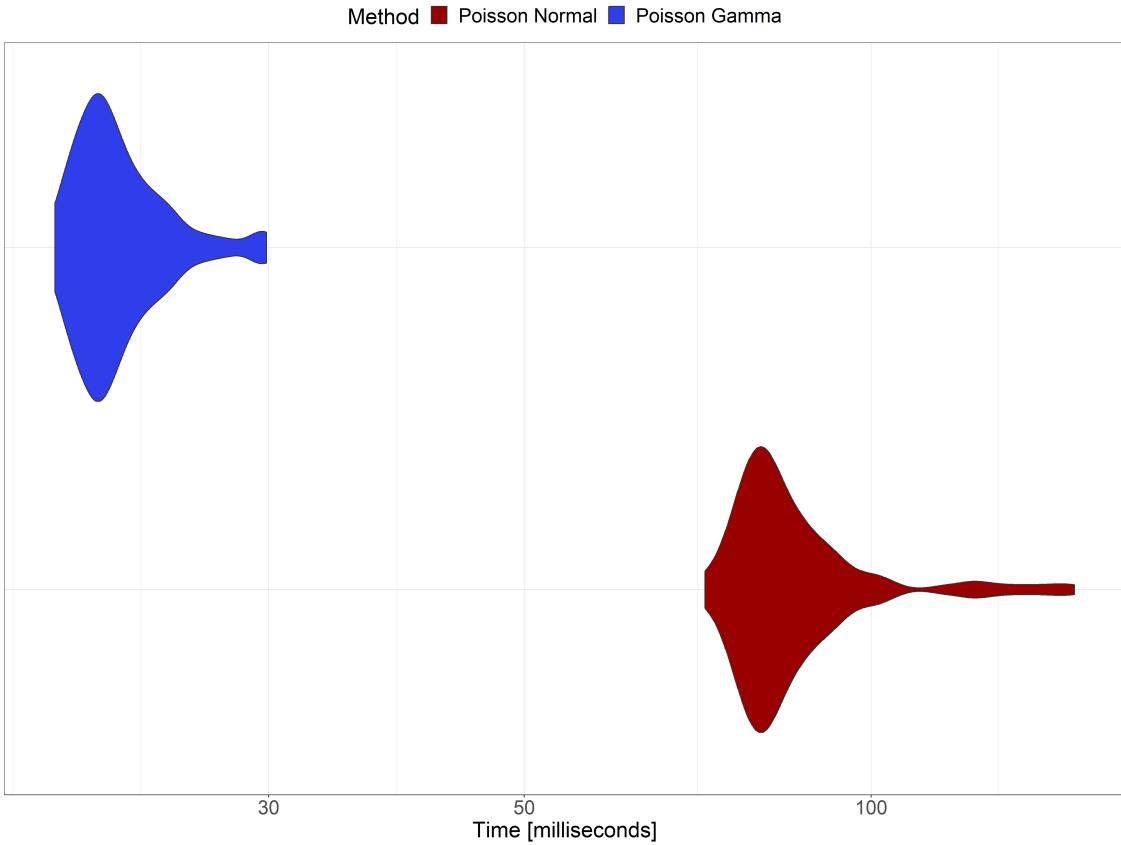


Figure 4.2: A violin plot (Hintze and Nelson 1998) of the 100 run times for the optimization routines in each modeling framework.

Figure 4.2 shows a violin plot comparing the running times for the optimization routines in each modeling framework. The plot demonstrates that the run time in the hierarchical Poisson Gamma model is remarkably faster, up to a factor of $4\times$, compared to the hierarchical Poisson Normal model. This indicates that the Poisson Gamma model may offer computational advantages in terms of run time efficiency.

4.6 Scoring rule

In this section, the scoring rule used to evaluate the overall score of the models is outlined. The approach is inspired by Bjerregård, Møller, and Madsen (2021). For a given time series $y_t = y_1, \dots, y_N$, each forecast and its corresponding realized observation pair (G_t, y_t) is evaluated. The overall score of the model is then reported as the average score:

$$\bar{S}(G, y) = \frac{1}{N} \sum_{t=1}^N S(G_t, y_t) \quad (4.43)$$

One commonly used scoring rule is the *logarithmic score* derived from likelihood theory, which is defined as $S(G, y) = -\log(G(y))$ (Good 1992; Gneiting and Raftery 2007). In this thesis, this scoring rule is based on the probability mass function and is equivalent to the log-likelihood of the forecast model. The calculation of the logarithmic score is shown in Example 4.2, which is adapted from Bjerregård, Møller, and Madsen (2021).

Example 4.2 (Calculation of the logarithmic score). The Gamma distribution is used to represent the probabilistic forecast in this example. The Gamma distribution is parametrized by two parameters, shape α and rate β , and its probability density function (PDF) is given by:

$$f(y) = \frac{1}{\Gamma(\alpha)\beta} \left(\frac{y}{\beta}\right)^{\alpha-1} \exp\left(-\frac{y}{\beta}\right) \quad (4.44)$$

In this example, the parameters of the true model, denoted as f , is chosen to be $(\alpha, \beta) = (3, 3)$. 10 observations are simulated, denoted as y_1, y_2, \dots, y_{10} , which are shown in Table 4.2.

Table 4.2: 10 simulated observations following a G(3,3)-distribution.

i	1	2	3	4	5	6	7	8	9	10
y_i	1.278	2.233	0.657	0.831	1.281	0.287	1.363	1.612	0.790	0.161

The true model f is compared to a competing model, denoted as g , which is a Gamma distribution with parameters $(\alpha, \beta) = (3, 8)$. The true model f , the competing model g , and the observations y are illustrated in Figure 4.3.

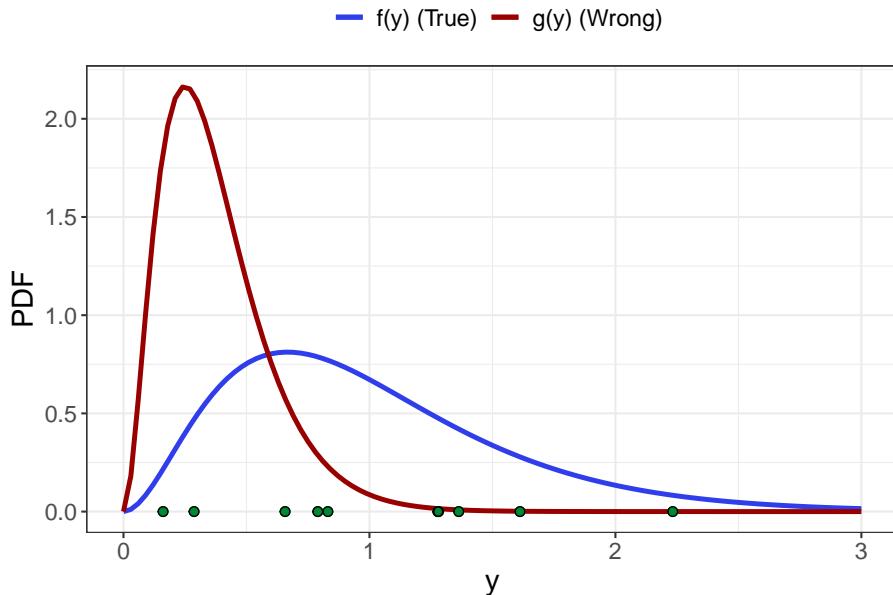


Figure 4.3: Observations (green dots) were simulated from a G(3,3)-distribution. The true model f (blue) is shown, along with the PDF for the competing model g (red), which follows a G(3,8)-distribution.

The logarithmic score of the true model f for the first observation, $y_1 = 1.278$, is calculated as follows

$$\begin{aligned}
-\log(f(y_1)) &= -\log(f(1.278)) \\
&= -\log\left[\frac{1}{\Gamma(3)4}\left(\frac{1.278}{4}\right)^{3-1} \exp(-1.278/4)\right] \\
&= -1.156
\end{aligned} \tag{4.45}$$

Similarly, the logarithmic scores for the other observations can be calculated using the same formula. The individual logarithmic scores for all 10 observations are presented in Table 4.3.

$$\begin{aligned}
\bar{S}(f, \mathbf{y}) &= 1.1 \\
\bar{S}(g, \mathbf{y}) &= 3.22
\end{aligned} \tag{4.46}$$

Table 4.3: Logarithmic scores of the two different Gamma models w.r.t. the 10 individual observations.

i	1	2	3	4	5	6	7	8	9	10
$S(f, y_i)$	1.16	3.86	0.00	0.23	1.16	0.18	1.37	2.03	0.17	0.83
$S(g, y_i)$	4.19	10.71	0.55	1.47	4.21	-0.75	4.74	6.40	1.25	-0.60

Among the 10 observations, 8 of them are more likely to occur under the true model f compared to the competing model g . The final key quantity, the average logarithmic score, $\bar{S}(G, \mathbf{y})$, can be calculated using Equation (4.43).

The *logarithmic score* has desirable properties as it captures all possible information about the observed data in relation to the model. However, it has a potential drawback in that it heavily penalizes unlikely observations. Consequently, even small changes in the tails of a density forecast can lead to significant changes in the *logarithmic score*, even when the overall shape of the density remains unchanged.

5 Case studies

This chapter presents the findings obtained from applying both the state-of-the-art and novel outbreak detection algorithms to the subset of diseases examined in this master thesis. To demonstrate the practical application of these algorithms, a case study focusing on Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC) will be presented. The results of the remaining case studies will be summarized more concisely, highlighting the differences observed. For a complete collection of figures and tables related to the case studies, please refer to Appendix E.

It is widely recognized that effective monitoring of a surveillance time series necessitates accurate modeling of the time series prior to assessing aberrations. Therefore, the performance of these algorithms in identifying outbreaks within the selected diseases will be thoroughly discussed and analyzed. Both the state-of-the-art and the novel algorithms take into account trends and seasonality, which will be addressed in the analysis. Additionally, a comparison of the two modeling frameworks used in the novel algorithm will be presented, providing valuable insights into the strengths and limitations of generalized mixed effects models and hierarchical generalized linear models.

Moreover, this chapter outlines some of the challenges encountered during the process of conducting this thesis, specifically related to statistical outbreak detection. These challenges are showcased using real case studies of diseases in the mandatory notification system, including *Listeriosis* (LIST), *Shigellosis* (SHIL) and *Salmonellosis* (SALM). Several key topics will be discussed, including the role of overdispersion in statistical outbreak detection, the influence of context and observational bias, and the approach to handling diseases with numerous alarms. Furthermore, a comparative analysis will be conducted to evaluate the effectiveness of different methods in detecting outbreaks.

For a detailed presentation of the data used to generate these results, please refer to Chapter 3.

5.1 Outbreak detection in Shiga toxin (verotoxin)-producing *Escherichia coli*

The initial analysis focuses on Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC). The data set comprises monthly counts of Danish STEC cases, denoted as y_{it} , where $i = 1, \dots, 6$ represents the six age groups and $t = 1, \dots, T$ represents the time period of $T = 180$ months starting in 2008. The first step involves applying the state-of-the-art outbreak detection algorithms to identify potential outbreaks in the time series.

Hereafter, the novel algorithm is utilized, and different models for the fixed effects are proposed. Both the hierarchical Poisson Normal and Poisson Gamma models are considered as modeling frameworks. To determine the most appropriate models for further analysis, the performance of the different fixed effects models is compared using the average logarithmic score, denoted as $S(G, y)$, within each of the modeling frameworks. The average logarithmic score provides a measure of how well the models fit the observed data and allows for objective comparison between different models.

5.1.1 Applying the state-of-the-art outbreak detection algorithm to STEC

To investigate outbreak detection, the method by Farrington et al. (1996) and the subsequently improved method by Noufaily et al. (2013) are initially explored. These methods

can be employed using the `farringtonFlexible` function, which is available from the R package called **surveillance**. The methods are controlled using specific control arguments. Here, a period of $b = 3$ years and a window half-width of $w = 2$ is chosen for reference data. Consequently, the number of baseline months used to calculate the thresholds are based on a total of $n = 15$ months. Also, the threshold calculation is based on a significance level of $\alpha = 0.05$. To access the other specific control arguments used in this thesis, refer to Appendix B.1.

For the analysis of STEC, the reference values are based on data collected from January 2008 to February 2011. Subsequently, surveillance is performed using the data spanning from March 2011 to December 2022. The resulting time series and alarms are visualized in Figure 5.1.

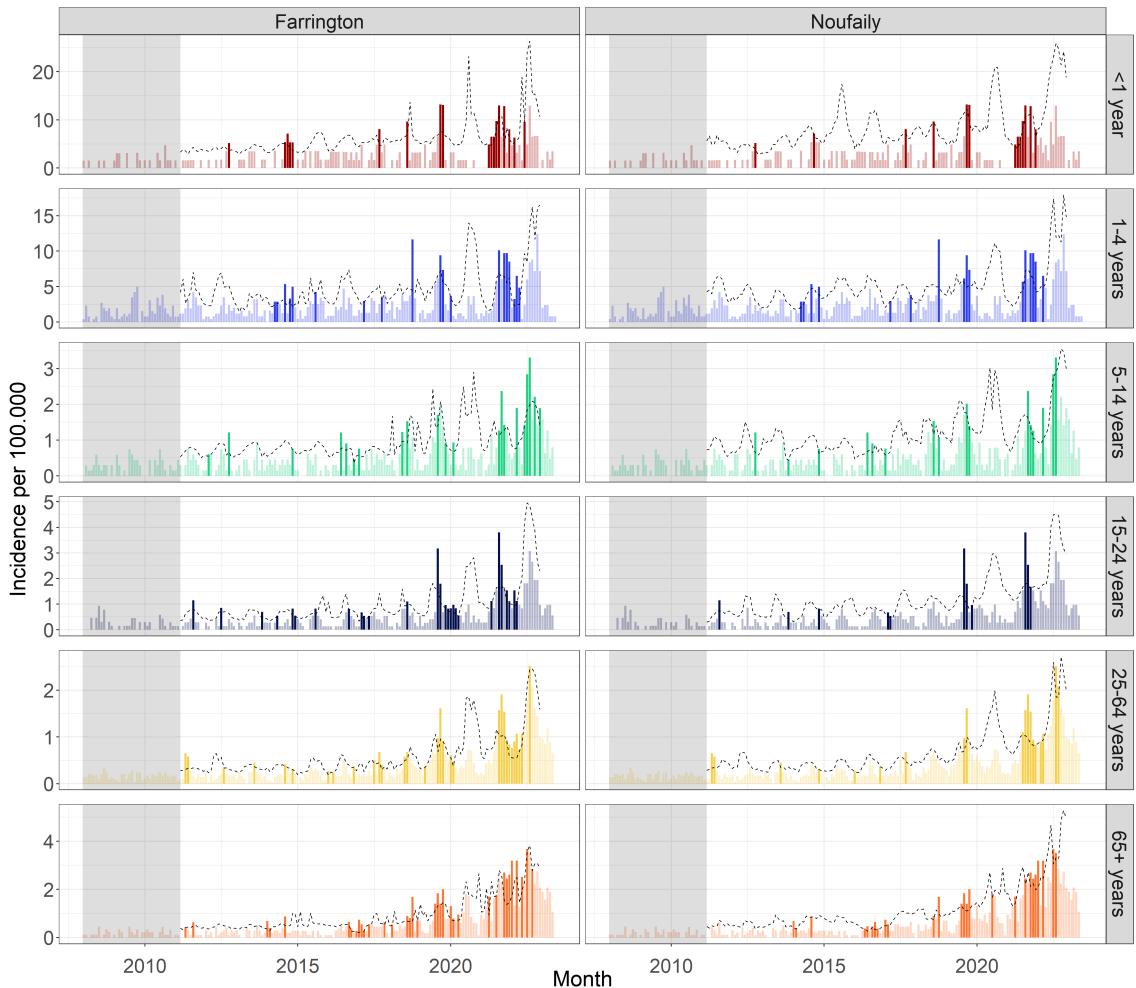


Figure 5.1: Monthly STEC incidence per 100,000 in six age groups in Denmark, 2008-2022. Monitored by Farrington (left) and Noufaily (right) methods. Reference data (grey area) for the estimation of model parameters are in the period from January 2008 to March 2011. Threshold (dashed line) is computed for observation time points outside reference data. Alarms are triggered (dark color) if observations exceed threshold.

Multiple alarms are displayed in Figure 5.1 for both the Farrington and Noufaily methods. The Farrington method generates a total of 151 alarms, while the Noufaily method produces a notably lower count of 101 alarms. It is not surprising that the Farrington method

generates more alarms, as it is known for its high sensitivity. This means that the method is able to detect very small aberrations in the time series, but it also leads to a greater number of “false positive” alarms.

The substantial number of alarms generated by both methods raises valid concerns regarding their practical utility. The sheer volume of alarms can overwhelm epidemiologists and pose challenges in effectively prioritizing and investigating potential outbreaks. It is essential to consider the overall disease picture and avoid overly burdening epidemiologists with verification tasks. Striking a balance between sensitivity and specificity is crucial to ensure that the generated alarms are meaningful and actionable, leading to efficient outbreak detection and response.

As a general rule, alarms that are triggered simultaneously in multiple age groups or during concurrent time points are of particular importance. Such patterns indicate potential outbreaks that warrant further attention and investigation by epidemiologists. By analyzing and interpreting the alarms in context, epidemiologists can gain a more thorough understanding of the outbreak situation and make informed decisions about resource allocation and intervention strategies.

When comparing the results from the Farrington and Noufaily methods in Figure 5.1, it is interesting to note that the thresholds are inflated towards the end of the series. Specifically, the Noufaily method exhibits a higher threshold compared to the Farrington method. This behavior is expected since the Noufaily method employs a less drastic approach in weighting prior outbreaks, allowing more information to be transmitted from previous outbreaks. As a result, the Noufaily method allows for a greater variation in the series and may generate fewer alarms. This characteristic highlights the different sensitivities of the two methods.

5.1.2 Applying the novel outbreak detection algorithm to STEC

The subsequent investigation focuses on the application of the novel outbreak detection algorithm to monitor STEC for potential outbreak signals. In this analysis, a rolling window with a width of $k = 36$ months is selected for the parameter estimation. Thus, the reference values are established using monthly data collected from January 2008 to December 2010. Similar to when applying the state-of-the-art algorithms, the time series is recursively monitored for outbreaks by moving the rolling window one month at a time using data from January 2011 to December 2022.

To identify outbreaks, an observation is classified as such if the one-step ahead random effect u_{it_1} for the specific age group exceeds the upper bound U_{t_0} . In this case study, the upper bound is determined as the 90% quantile of the distribution of random effects obtained from the second stage model. This threshold helps identify significant deviations in the outbreak intensity.

Introducing the fixed effects models

To ensure effective monitoring, a series of models for the fixed effects are proposed. These models aim to capture different aspects of the disease dynamics and improve the accuracy of outbreak detection.

In the initial model, the intensity λ_{it} is assumed to be solely dependent on the age group and the population size n_{it} . The model is formulated as:

$$\log(\lambda_{it}) = \beta(\text{ageGroup}_i) + \log(n_{it}) \quad (5.1)$$

Here, $\beta(\text{ageGroup}_i)$ represents the fixed effect specific to the age group i , capturing the

age group's influence on the outbreak intensity. The term $\log(n_{it})$ acts as an offset, accounting for the population size at time t for age group i . This model will be referred to as the *constant* model.

In the first extension of the constant model, the trend over time across all age groups is incorporated. The extended model is termed the *trend* model and is formulated as:

$$\log(\lambda_{it}) = \beta(\text{ageGroup}_i) + \beta_{\text{trend}} t + \log(n_{it}) \quad (5.2)$$

In this model, β_{trend} quantifies the rate of change in the outbreak intensity over time. By including this parameter, the model accounts for the overall trend in the outbreak intensity across all age groups, allowing for a more comprehensive analysis of the outbreak dynamics.

In addition to the models described earlier, another model is proposed, which incorporates a seasonality component represented by Fourier terms into the analysis. This model builds upon the constant model and assumes an annual seasonality pattern. This model is termed the *seasonality* model. The intensity λ_{it} is modeled as follows:

$$\log(\lambda_{it}) = \beta(\text{ageGroup}_i) + \sin\left(\frac{2\pi \cdot \tau_t}{12}\right)\beta_{\text{sin}} + \cos\left(\frac{2\pi \cdot \tau_t}{12}\right)\beta_{\text{cos}} + \log(n_{it}) \quad (5.3)$$

In this model, τ_t represents the time period within a year, ranging from 1 to 12 (corresponding to the months of January to December). The parameters β_{sin} and β_{cos} capture the effect of the seasonal pattern on the outbreak intensity. By including sine and cosine functions of τ_t , the model accounts for the periodic fluctuations in the outbreak intensity observed throughout the year. This allows for the detection and analysis of seasonality patterns in the outbreak data, providing insights into the seasonal variations in disease occurrence.

In addition to the models described earlier, a final model is proposed that combines both trend and seasonality components. This model builds upon the previous models and includes the effects of both the overall trend over time and the seasonal patterns. Thus, the model is labeled as the *combined* model. The intensity λ_{it} is modeled as follows:

$$\log(\lambda_{it}) = \beta(\text{ageGroup}_i) + \beta_{\text{trend}} t + \sin\left(\frac{2\pi \cdot \tau_t}{12}\right)\beta_{\text{sin}} + \cos\left(\frac{2\pi \cdot \tau_t}{12}\right)\beta_{\text{cos}} + \log(n_{it}) \quad (5.4)$$

Performance evaluation and parameter estimates of hierarchical models

The proposed fixed effects models are subsequently implemented and estimated in the two different modeling frameworks: the hierarchical Poisson Normal and Poisson Gamma models. The performance of these models is evaluated using the average logarithmic score, $\bar{S}(G, y)$. An excerpt of these results, namely for the models that minimizes $\bar{S}(G, y)$ in both modeling frameworks, are summarized in Table 5.1. For the full table of results, refer to Table E.3.

Table 5.1: An excerpt of the STEC modeling results, namely for the models that minimizes the average logarithmic score, $\bar{S}(G, y)$. The parameters are estimated in a rolling window with width $k = 36$ and the estimates at time t_0 , i.e. the last time point, are presented in this table for both modeling frameworks. Confidence intervals for the parameter estimates are calculated using 95% profile likelihood confidence intervals.

	Parameter	Estimate	95% CI
Poisson Normal			
$\bar{S}(G, y) = 14.18$	$\hat{\beta}_{trend}$	0.04	[0.04, 0.05]
	$\hat{\beta}_{<1year}$	12.56	[12.3, 12.81]
	$\hat{\beta}_{1-4years}$	15.23	[15.03, 15.41]
	$\hat{\beta}_{5-14years}$	15.68	[15.46, 15.88]
	$\hat{\beta}_{15-24years}$	16.19	[15.99, 16.38]
	$\hat{\beta}_{25-64years}$	18.74	[18.56, 18.9]
	$\hat{\beta}_{65+years}$	17.60	[17.43, 17.77]
	$\hat{\beta}_{\sin}$	-0.34	[-0.44, -0.25]
	$\hat{\beta}_{\cos}$	-0.06	[-0.15, 0.04]
	$\log(\hat{\sigma})$	-1.12	[-1.33, -0.92]
Poisson Gamma			
$\bar{S}(G, y) = 14.15$	$\hat{\beta}_{trend}$	0.04	[0.04, 0.05]
	$\hat{\beta}_{<1year}$	12.61	[12.34, 12.85]
	$\hat{\beta}_{1-4years}$	15.28	[15.09, 15.47]
	$\hat{\beta}_{5-14years}$	15.72	[15.51, 15.92]
	$\hat{\beta}_{15-24years}$	16.23	[16.03, 16.42]
	$\hat{\beta}_{25-64years}$	18.77	[18.6, 18.94]
	$\hat{\beta}_{65+years}$	17.65	[17.47, 17.82]
	$\hat{\beta}_{\sin}$	-0.33	[-0.43, -0.24]
	$\hat{\beta}_{\cos}$	-0.04	[-0.14, 0.05]
	$\log(\hat{\phi})$	-2.21	[-2.63, -1.82]

It is shown, that the combined model for the fixed effects minimizes the logarithmic score for both modeling frameworks, indicating a better fit to the STEC data. As a result, these models have been selected for further investigation regarding their ability to detect outbreaks.

Moreover, it is worth noting that the parameter estimates and confidence intervals for the fixed effects are consistent across the two modeling frameworks, despite the differences in assumptions about the distribution of random effects. This suggests that the choice of modeling framework does not have a substantial impact on the estimation of the fixed effects parameters in this analysis. However, a neat feature of the hierarchical Poisson Gamma model is, that the maximum likelihood estimates for the fixed effects parameters can be directly interpreted as the contribution from the specific parameter to the disease intensity. This is possible due to the properties of the expected value described in Equation (4.29). However, this direct interpretation does not hold for the parameter estimates obtained in the hierarchical Poisson Normal model. Nevertheless, it can be observed that both sets of parameter estimates are similar throughout the entire time period, indicating that this discrepancy is not significant.

It is noteworthy that all parameter estimates, except one, are statistically significant at t_0 . The only parameter estimate that is not significant is $\hat{\beta}_{cos}$, as its 95% confidence interval overlaps with zero in both modeling frameworks.

A slight, yet consistent, increase in the contribution from the age group to the overall intensity of STEC cases is observed in Figure 5.2 over the analyzed period, with only one instance of a decrease in intensity immediately after the 2020 COVID-19 lockdown.

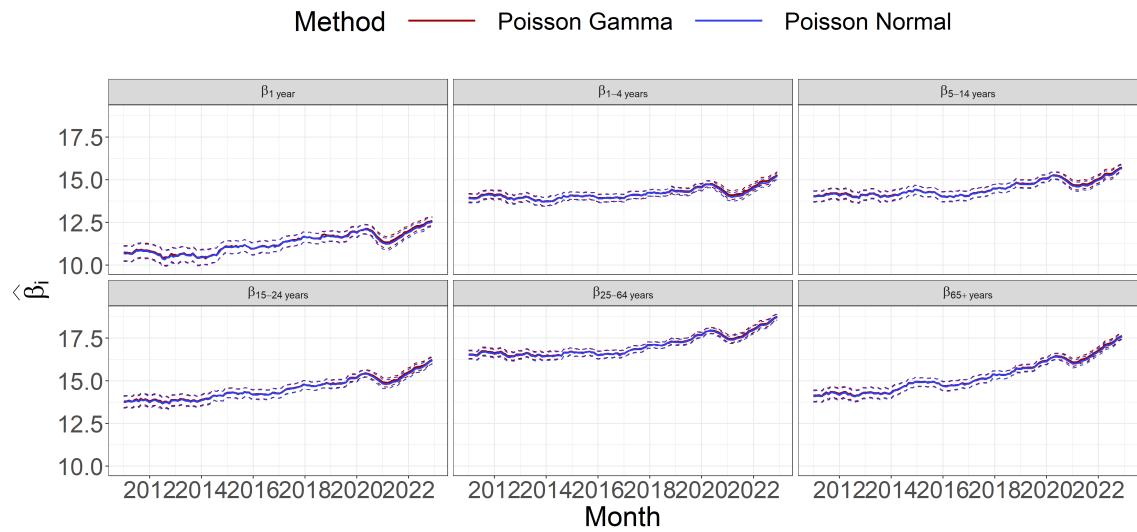


Figure 5.2: Age group contribution in STEC in the time period between January 2011 and December 2022. Recursive estimates (solid line) and 95% profile likelihood confidence intervals (dashed line) of $\hat{\beta}(ageGroup_i)$ for the combined model. Estimates are obtained using a rolling window with a width of $k = 36$ months and discarding observations that are characterized as alarms.

For the most, the maximum likelihood estimate for $\hat{\beta}_{trend}$ indicates a positive trend during the observed period (See Figure 5.3). However, it is important to note that the trend is not consistently statistically significant. The first significant positive trend is observed in the mid-2014, lasting until 2016. Another significant positive trend is observed in 2018, coinciding with the transition from a culture-based diagnostic method to a PCR method in most departments of clinical microbiology (Svendsen et al. 2023).

Notably, a significant drop in the trend is observed in the mid-2020, immediately after the COVID-19 lockdown, lasting until the mid-2021. Since then, the trend has been steadily increasing and is currently at an all-time high for the observed period.

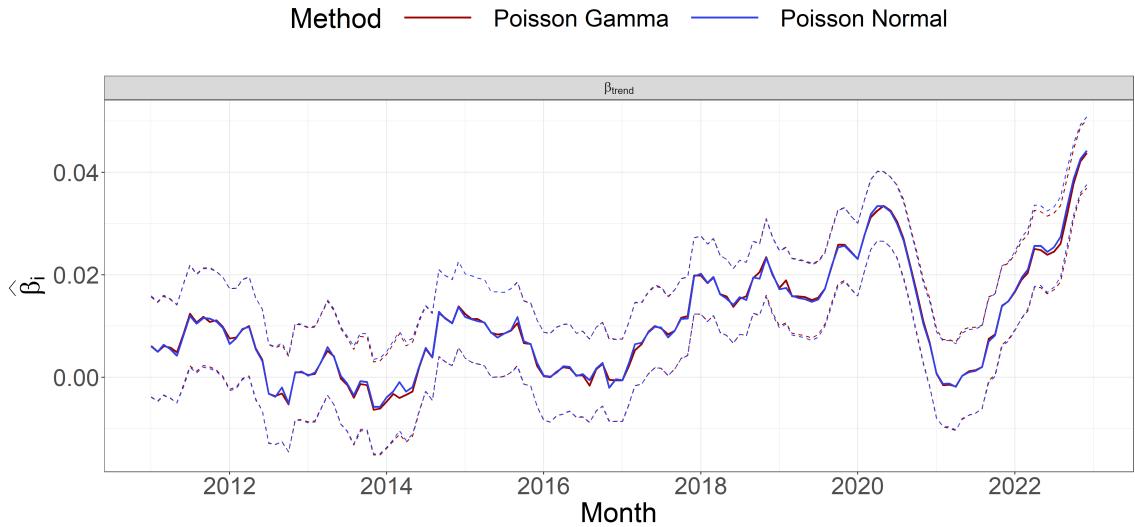


Figure 5.3: Disease trend in STEC in the time period between January 2011 and December 2022. Recursive estimates (solid line) and 95% profile likelihood confidence intervals (dashed line) of $\hat{\beta}_{trend}$ for the combined model. Estimates are obtained using a rolling window with a width of $k = 36$ months and discarding observations that are characterized as alarms.

Figure 5.4 clearly demonstrates the presence of significant and pronounced seasonality in the observed period for STEC. Moreover, the estimated seasonality parameters exhibit variability over time.

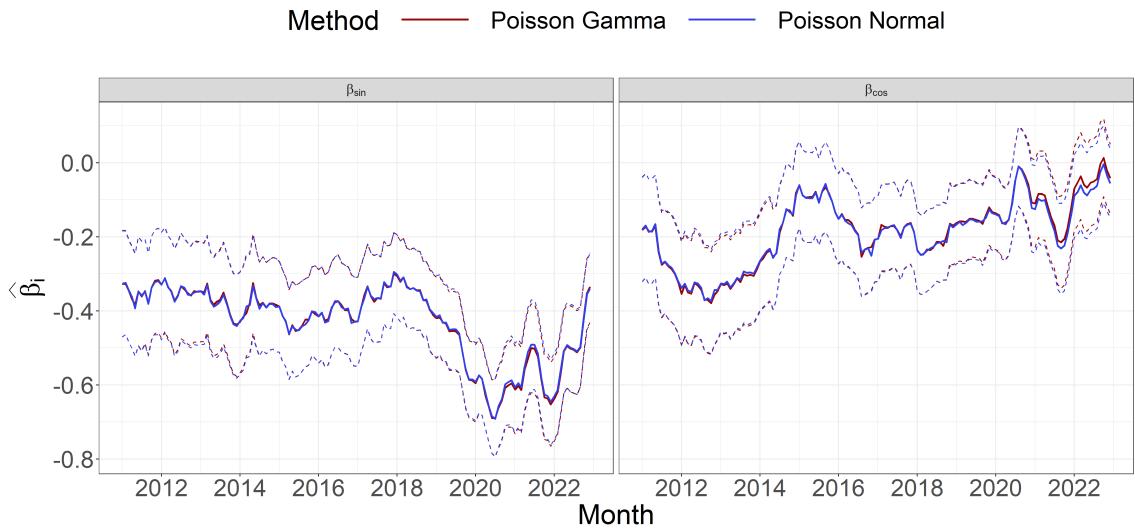


Figure 5.4: Disease seasonality in STEC in the time period between January 2011 and December 2022. Recursive estimates (solid line) and 95% profile likelihood confidence intervals (dashed line) of $\hat{\beta}_{\cos}$ and $\hat{\beta}_{\sin}$ for the combined model. Estimates are obtained using a rolling window with a width of $k = 36$ months and discarding observations that are characterized as alarms.

To further understand the variation in the seasonal component, the amplitude ω and phase

ρ are calculated using the following formulas:

$$\omega = \sqrt{(\beta_{\sin}^2 + \beta_{\cos}^2)} \quad \rho = -\arctan\left(\frac{\beta_{\cos}}{\beta_{\sin}}\right) \quad (5.5)$$

Due to the challenges of propagating uncertainty estimates, bootstrap confidence intervals are calculated based on 10,000 replicates for each parameter, in each time point, and for each modeling framework. The maximum likelihood estimates for the amplitude and phase is visualized along with its corresponding bootstrap confidence intervals in Figure 5.5.

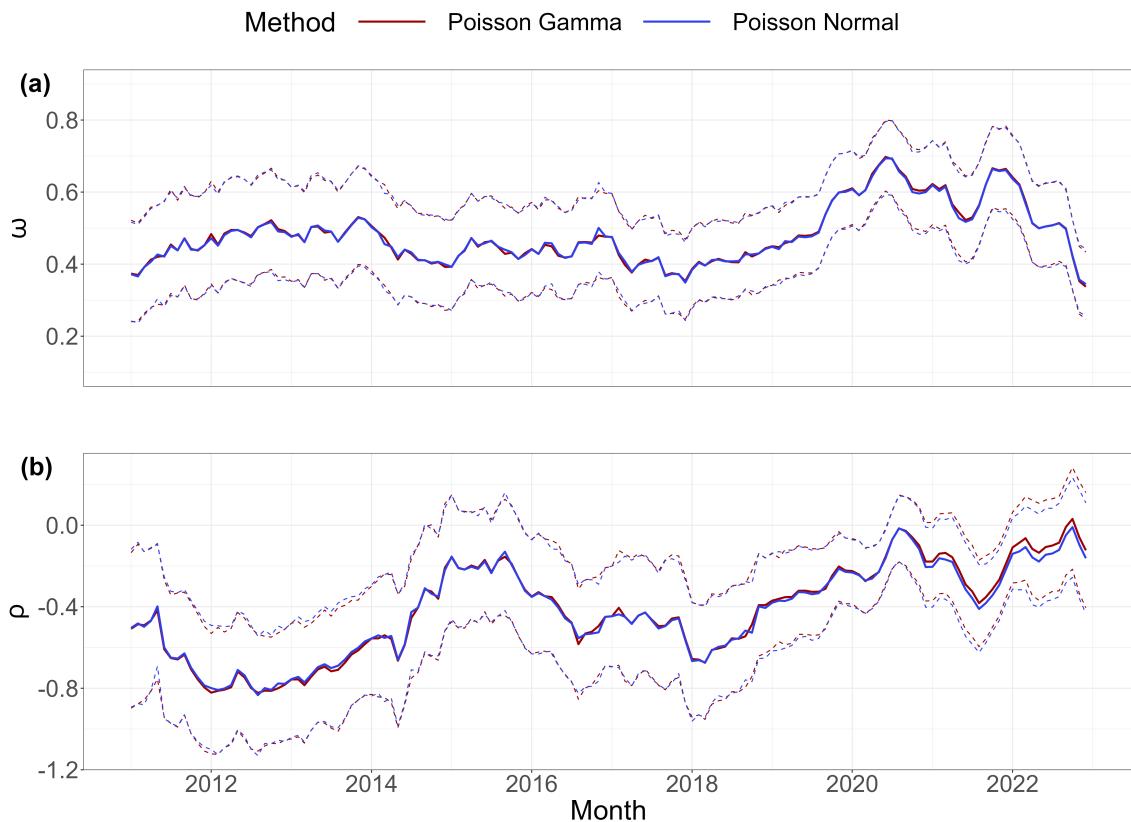


Figure 5.5: Illustration of the amplitude **(a)** and phase **(b)** along with their corresponding 95% bootstrap confidence intervals for each of the modeling frameworks.

It is worth noting that the amplitude of the seasonal component exhibits temporal variation. Particularly, an increase in amplitude is observed after 2018, coinciding with the introduction of a more sensitive PCR diagnostic method. This change in diagnostic practices may have influenced the detection and reporting of STEC cases, resulting in an apparent intensification of the seasonal pattern. However, it is more likely attributed to a change in observational bias rather than a true increase in disease intensity. The varying amplitude highlights the importance of considering temporal changes and external factors when modeling and interpreting disease outbreak data.

The estimated variance $\hat{\sigma}$ for the hierarchical Poisson Normal model and the estimated dispersion $\hat{\phi}$ for the hierarchical Poisson Gamma model are visualized in Figure 5.6. These parameters provide insights into the variability of the random effects and play a

crucial role in determining the threshold for detecting outbreaks. The values of $\hat{\sigma}$ and $\hat{\phi}$ help assess the magnitude of deviations from the expected disease intensity and aid in distinguishing significant aberrations from normal fluctuations in the data

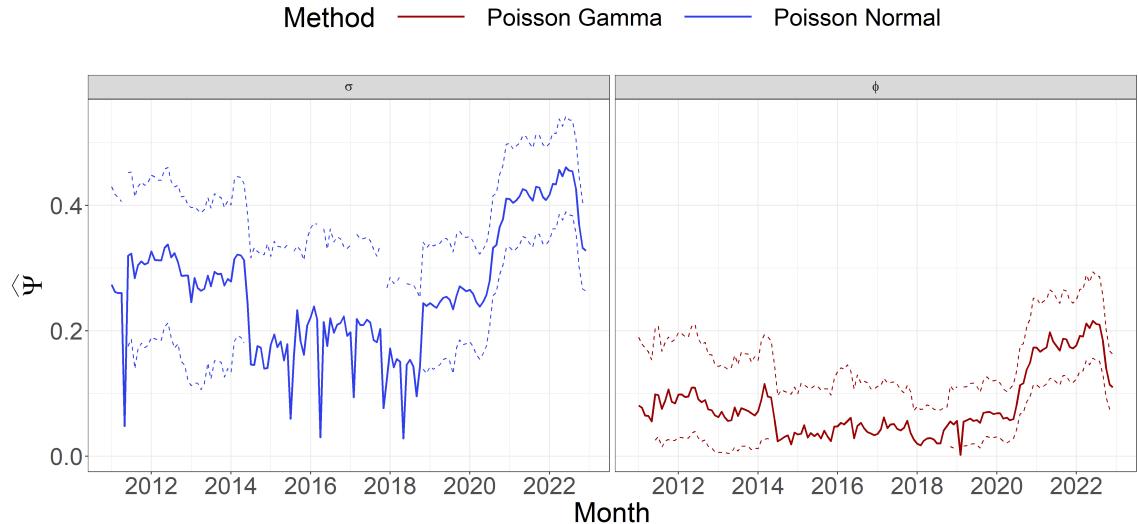


Figure 5.6: Variability and dispersion in STEC in the time period between January 2011 and December 2022. Recursive estimates (solid line) and 95% profile likelihood confidence intervals (dashed line) of $\hat{\phi}$ and $\hat{\sigma}$. Estimates are obtained using a rolling window with a width of $k = 36$ months and discarding observations that are characterized as alarms.

It should be noted that the inability of the second stage model parameters (σ or ϕ) to converge during certain time periods does not affect the performance of the novel outbreak detection algorithm. These periods typically coincide with time periods with relatively few reported cases, i.e. periods with no outbreaks. Generally, these periods occur when the sample data show no overdispersion. In such cases, the time series can be adequately modeled by a Poisson distribution, and therefore the second stage model is not informed. Specifically, this becomes relevant in the case study of LIST, where the example of missing overdispersion is very explicit. The role of overdispersion in the disease data will be discussed in Section 5.3.

Visualization of estimated one-step ahead random effects and upper bounds

Figure 5.7 illustrates the estimated one-step ahead random effects \hat{u}_{it_1} for each age group, along with their corresponding upper bounds U_{t_0} , in both modeling frameworks.

The random effects u_{it_1} represent the deviations from the expected outbreak intensity in the subsequent time period for each age group. These random effects provide information on whether there is an unusual or unexpected increase or decrease in the outbreak intensity compared to the expected values.

The upper bounds U_{t_0} are calculated based on the 90% quantile of the distribution of the random effects obtained from the second stage model. They serve as threshold values to identify potential outbreaks. If the one-step ahead random effects u_{it_1} exceed the upper bounds U_{t_0} , it suggests a significant deviation from normal variation and indicates a potential outbreak signal.

By visualizing the one-step ahead random effects u_{it_1} and the upper bounds U_{t_0} together,

epidemiologists can easily identify periods or age groups where the random effects surpass the upper bounds, highlighting potential outbreaks that require further investigation and monitoring.

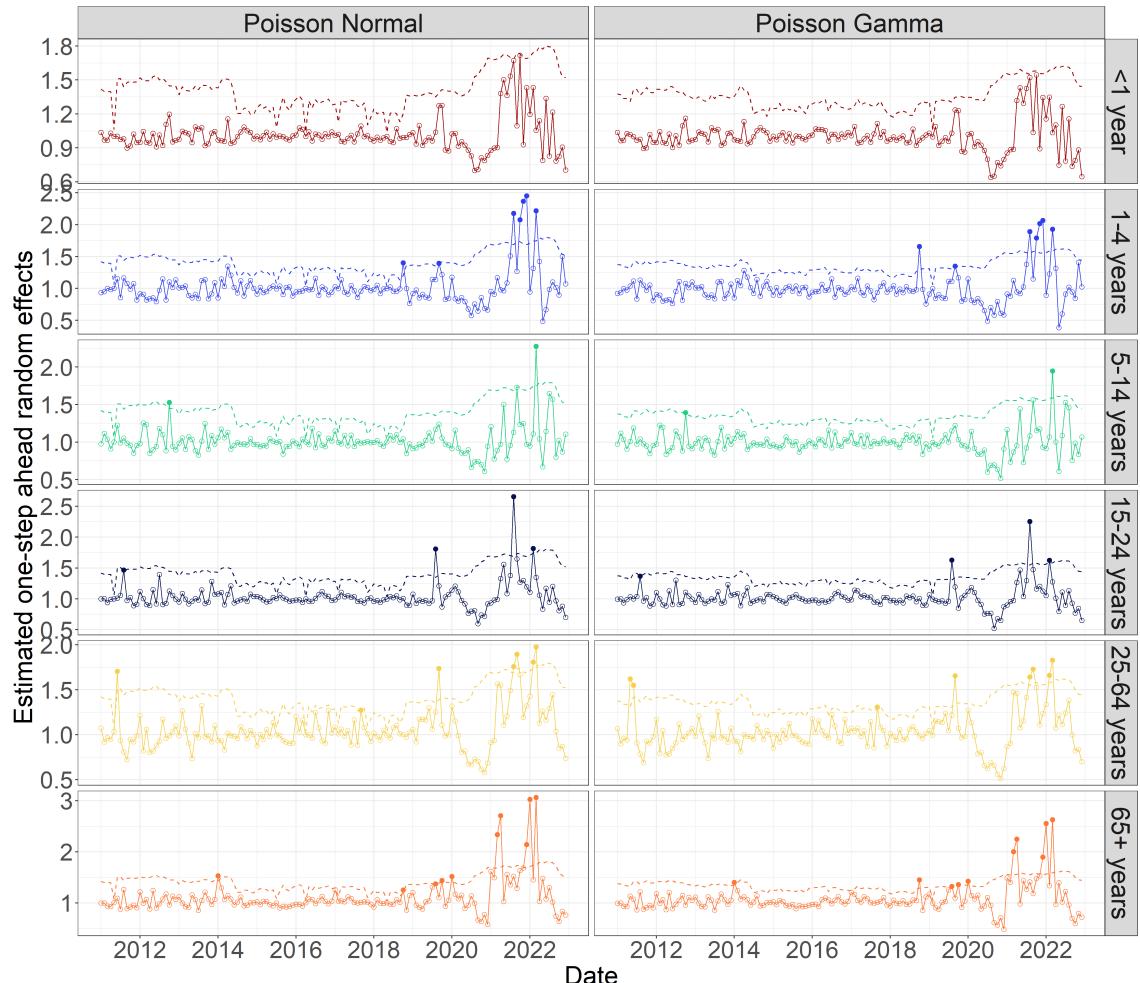


Figure 5.7: Estimated one-step ahead random effects \hat{u}_{it_1} (circles) for STECs in Denmark, 2011-2022, are shown for the combined model. An alarm is raised (solid circle) if \hat{u}_{it_1} exceeds the threshold U_{t_0} (dashed line). In the hierarchical Poisson Normal model (left), the random effects are exponentiated to $\exp(\hat{u}_{it_1})$ to transform them into the same domain as the hierarchical Poisson Gamma model (right).

From Figure 5.7, it is evident that the estimated one-step ahead random effects u_{it_1} surpass the threshold U_{t_0} on multiple occasions. Specifically, a total of 30 alarms are generated using the hierarchical Poisson Normal framework, while 31 alarms are generated in the hierarchical Poisson Gamma framework. Notably, during the period from March 2021 to March 2022, a significant number of alarms are triggered in multiple age groups in both modeling frameworks.

5.2 Outbreak detection in other case studies

In this section, the results of applying the state-of-the-art and novel outbreak detection algorithms to the other case studies, including *Listeriosis* (LIST), *Shigellosis* (SHIL), and *Salmonellosis* (SALM), are presented.

As for the STEC case study, the reference data used in the state-of-the-art outbreak detection algorithms is based on monthly data from January 2008 to February 2011. Subsequent surveillance is conducted using data from March 2011 to December 2022. In the novel outbreak detection algorithm, a rolling window of width $k = 36$ is used. Thus, the reference data is established using monthly data collected from January 2008 to December 2010. The time series is then monitored for outbreaks using data from January 2011 to December 2022.

It is also important to mention that different age groupings are utilized in each of the case studies.

5.2.1 Analysis of outbreak detection in *Listeriosis*

Intuitively, it is not common to employ statistical methods for monitoring and detecting outbreaks of *Listeriosis* (LIST). The current state-of-the-art approach for detecting outbreaks of this disease relies heavily on laboratory-based methods, specifically Whole Genome Sequencing (WGS). WGS enables epidemiologists to link cases that have occurred over an extended period, even months or years apart, and identify them as part of a continuous-source outbreak. Statistical methods may face challenges in detecting such outbreaks due to their complex and prolonged nature. However, it is crucial to include this case study in the thesis to highlight both the strengths and limitations of the novel and state-of-the-art methods.

The data used in this case study consist of monthly Danish LIST case counts from the laboratory notification system. These counts are denoted as y_{it} . Here, the subscript i distinguishes between two age groups, with $i = 1$ representing the age group below 65 years and $i = 2$ representing the age group of 65 years and above. As in the STEC case study, the subscript t denotes the time.

Applying the state-of-the-art outbreak detection algorithms to *Listeriosis*

The Farrington and Noufaily methods are both investigated on their ability to detect outbreaks of LIST. The resulting series is visualized in Figure 5.8 together with the alarms triggered by each of the methods.

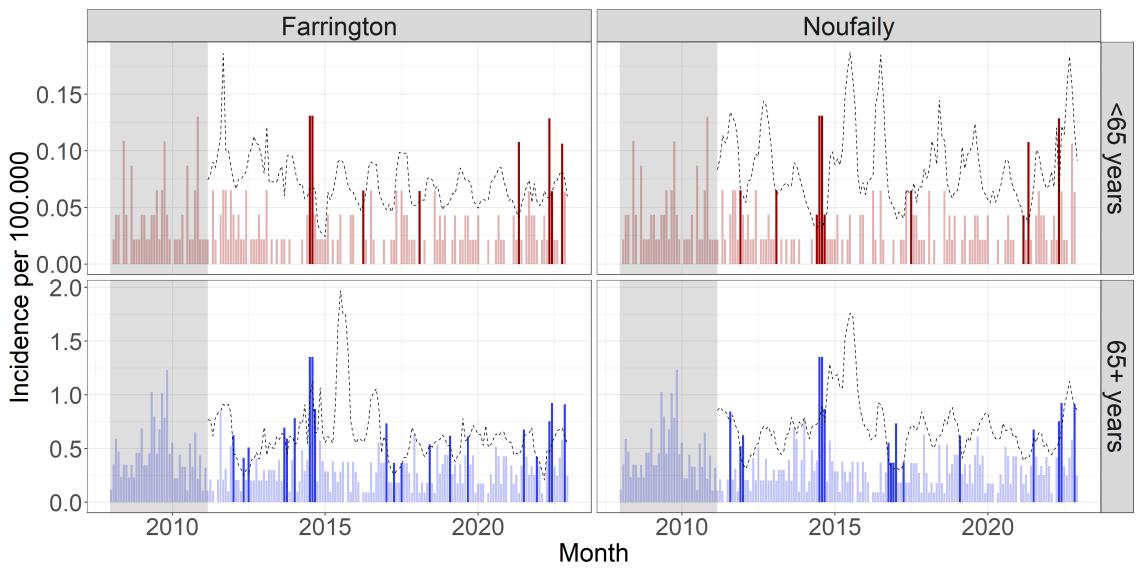


Figure 5.8: Monthly LIST incidence per 100,000 in Denmark, 2008-2022. Monitored by Farrington (left) and Noufaily (right) methods. Reference data (grey area) for the estimation of model parameters are in the period from January 2008 to March 2011. Threshold (dashed line) is computed for observations time points outside reference data. Alarms are triggered (dark color) if observations exceed threshold.

In Figure 5.8, multiple alarms can be observed for both the Farrington and Noufaily methods. The Farrington method triggers a total of 28 alarms, while the Noufaily method produces slightly fewer alarms, specifically 26 alarms.

Interestingly, both methods demonstrate the ability to correctly raise alarms coinciding with ongoing outbreak investigations by SSI. One notable example is the outbreak that involved 41 cases and resulted in 17 deaths (Wingstrand et al. 2015). The Farrington method successfully flags this event in September 2013, with alarms occurring sporadically in the subsequent period. The Noufaily method also identifies this outbreak, although it first does so in June 2014.

It is worth noting that the other observations flagged by the methods may be related to some of the numerous other long-spanned outbreaks that occurred concurrently in the period from 2016 to 2021. Also, several alarms are raised sporadically in 2022 coinciding with the outbreaks investigated by SSI caused by cold-cut and fish patties, respectively (Lassen et al. 2023). For a full list of the alarms raised by the methods see Table D.1.

Applying the novel outbreak detection algorithm to *Listeriosis*

For the LIST data, it is found that only the constant age group model for the fixed effects is able to converge. However, even the constant fixed effects models struggle to accurately capture the underlying process of the disease. Figure 5.9 provides a visual representation of this, showing the estimated one-step ahead random effects \hat{u}_{it_1} and the calculated thresholds U_{t_0} . It can be observed that the estimated random effects and the calculated thresholds collapse for prolonged periods, indicating that there is insufficient information in the data to inform the second stage model. This is particularly evident in the hierarchical Poisson Normal model, where the estimate of the variance parameter $\hat{\sigma}$ approaches zero. The role of overdispersion in the disease data will be further discussed in Section 5.3.

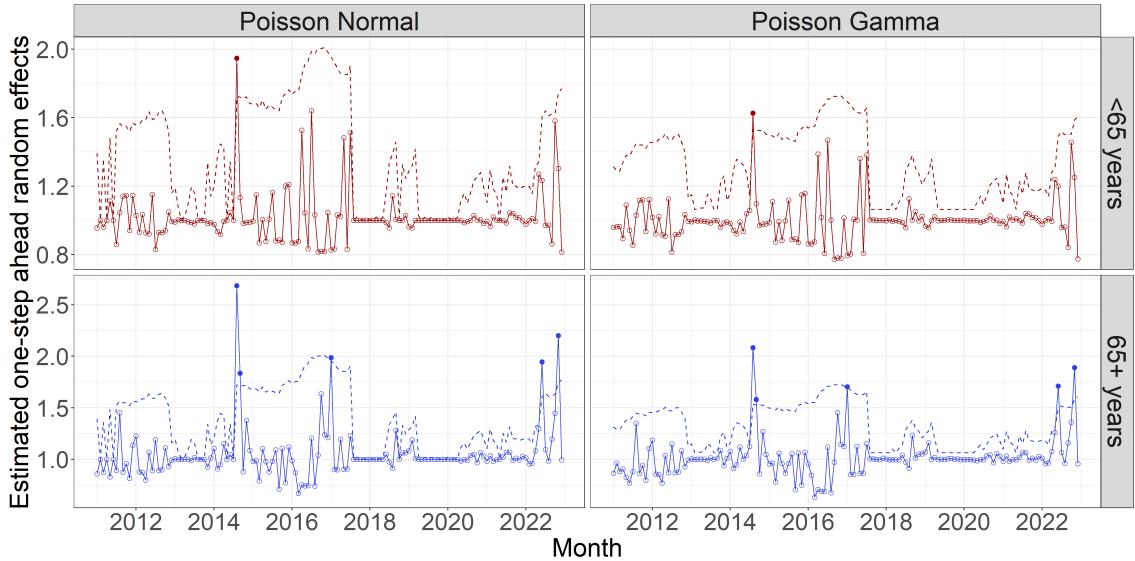


Figure 5.9: Estimated one-step ahead random effects \hat{u}_{it_1} (circles) for LIST in Denmark, 2011-2022, are shown for the constant model described. An alarm is raised (solid circle) if \hat{u}_{it_1} exceeds the threshold U_{t_0} (dashed line). In the hierarchical Poisson Normal model (left), the random effects are exponentiated to $\exp(\hat{u}_{it_1})$ to transform them into the same domain as the hierarchical Poisson Gamma model (right).

Nevertheless, it is worth noting that both modeling frameworks identify the same observations as outbreaks, all of which coincide with one or more concurrent outbreaks investigated by SSI. The first two observations characterized as outbreaks occur in August and September 2014 coinciding with the outbreak investigated by SSI that affected a total of 41 individuals. Notably, the investigation of this outbreak was initiated on the 26th of July by SSI (Wingstrand et al. 2015). The third outbreak is in January 2017 coinciding with multiple investigations, and the last two outbreaks are in June and November 2022 coinciding with cold-cut and fish patties, respectively (Lassen et al. 2023).

5.2.2 Analysis of outbreak detection in *Shigellosis*

The *Shigellosis* (SHIL) data set consists of monthly Danish SHIL case counts categorized into two age groups. These counts are represented as y_{it} , where the subscript $i = 1$ corresponds to the age group below 25 years, and $i = 2$ corresponds to the age group of 25 years and above.

Applying the state-of-the-art outbreak detection algorithm to *Shigellosis*

In Figure 5.10, a significant number of alarms are observed sporadically in the time series for both of the state-of-the-art methods. For the Farrington method, the number of alarms is 54, while for the Noufaily method, the number of alarms is 36.

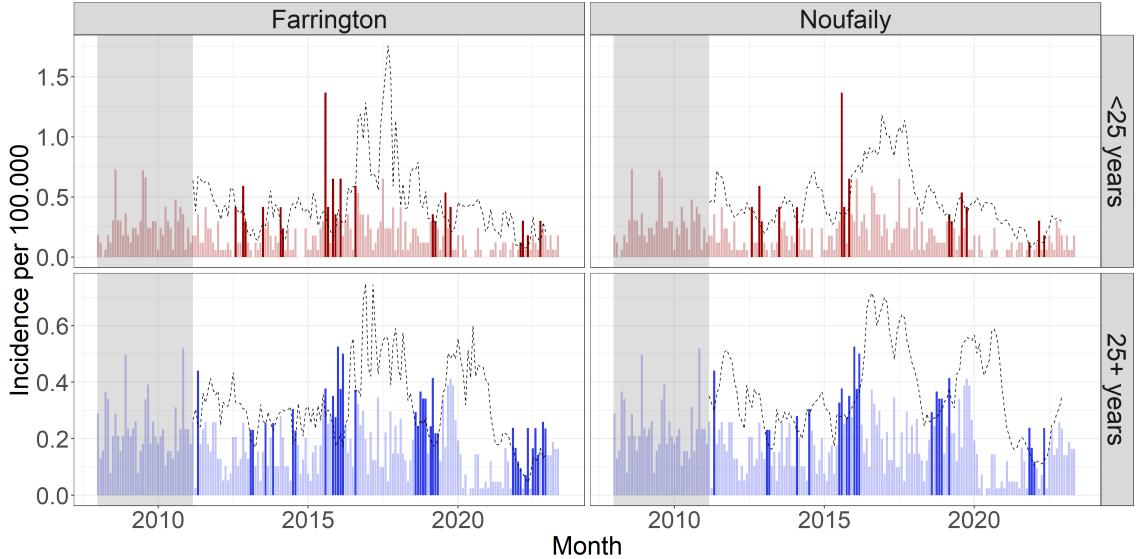


Figure 5.10: Monthly SHIL incidence per 100,000 in Denmark, 2008-2022. Monitored by Farrington (left) and Noufaily (right) methods. Reference data (grey area) for the estimation of model parameters are in the period from January 2008 to March 2011. Threshold (dashed line) is computed for observations time points outside reference data. Alarms are triggered (full opacity) if observations exceed threshold.

It is noteworthy that some of the alarms occur in clusters over specific time periods and are raised simultaneously in both age groups. These clustering patterns are particularly evident during periods with multiple alarms, such as from July 2015 to April 2016, from July 2018 to October 2019, and from November 2021 to December 2022. Surprisingly, none of these periods coincide with outbreaks investigated by SSI, and all methods fail to identify the outbreak investigated in August and September 2020.

However, it is important to recognize the significant challenge faced by statistical algorithms in detecting the outbreak during the summer of 2020. This difficulty arises due to the concurrent COVID-19-related lockdown, which results in different disease dynamics and a relatively low number of SHIL cases compared to similar periods in other years. This highlights the importance of considering the context in which the statistical outbreak detection algorithm is applied. Further discussion on this topic can be found in Section 5.3.

Applying the novel outbreak detection algorithm to *Shigellosis*

For SHIL it was found that the trend model for the fixed effects minimizes the average logarithmic score, $\bar{S}(G, y)$, in both modeling frameworks. Thus, the estimated one-step ahead random effects \hat{u}_{it_1} and thresholds U_{t_0} visualized in Figure 5.11 are calculated using this model.

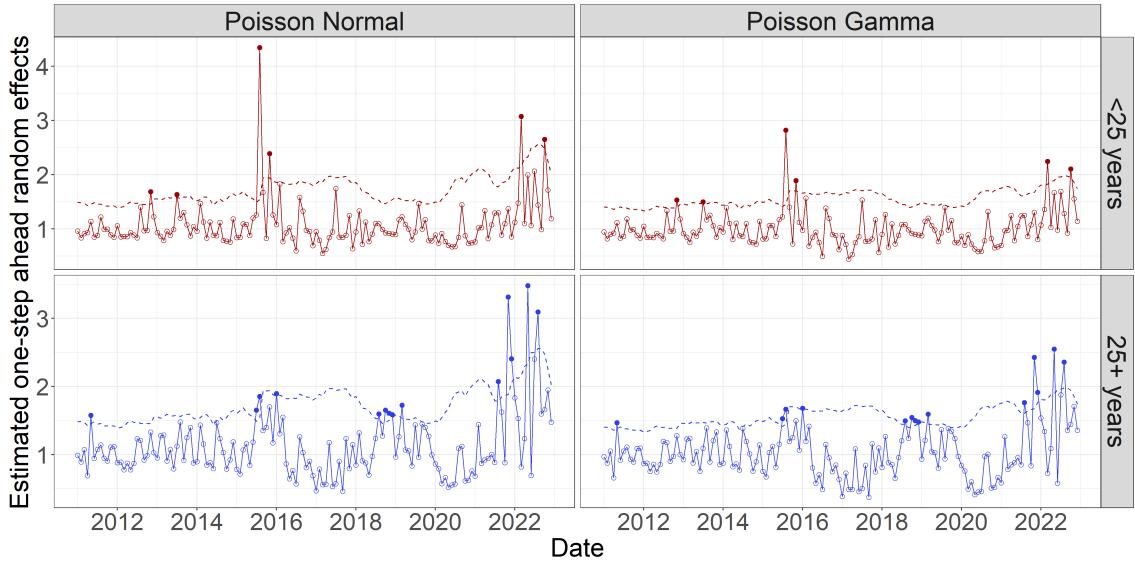


Figure 5.11: Estimated one-step ahead random effects \hat{u}_{it_1} (circles) for SHIL in Denmark, 2011-2022, are shown for the trend model. An alarm is raised (solid circle) if \hat{u}_{it_1} exceeds the threshold U_{t_0} (dashed line). In the hierarchical Poisson Normal model (left), the random effects are exponentiated to $\exp(\hat{u}_{it_1})$ to transform them into the same domain as the hierarchical Poisson Gamma model (right).

It is worth noting that both modeling frameworks agree on the raising an alarm in the same 20 observations, indicating consistency in their detection. Additionally, it is observed that the novel outbreak detection algorithms identify fewer alarms compared to the state-of-the-art algorithms. However, it is important to highlight that both algorithms detect outbreaks during certain overlapping time periods, demonstrating their ability to capture similar patterns of disease occurrence.

However, it is also noted that the novel algorithm failed to identify the outbreak investigated by SSI and collaborators in the summer of 2020. As previously described when applying the state-of-the-art methods, this outbreak is particularly challenging to detect for statistical outbreak detection algorithms.

5.2.3 Analysis of outbreak detection in *Salmonellosis*

Historically, *Salmonellosis* (SALM) is associated with frequent disease outbreaks. This is no different during the time period investigated here. The data set consist of monthly counts of Danish SALM cases, denoted as y_{it} . Here, the subscript i differentiates between the six age groups $i = 1, \dots, 6$. The subscript t represents time.

Applying the state-of-the-art outbreak detection algorithm to *Salmonellosis*

From Figure 5.12, it is evident that both of the state-of-the-art methods generate an overwhelming number of alarms across all six age groups. The Farrington method produces 131 alarms, while the Noufaily method produces 89 alarms. Although these methods exhibit high sensitivity, the excessive number of alarms undermines their usefulness and practicality in a real-world setting.

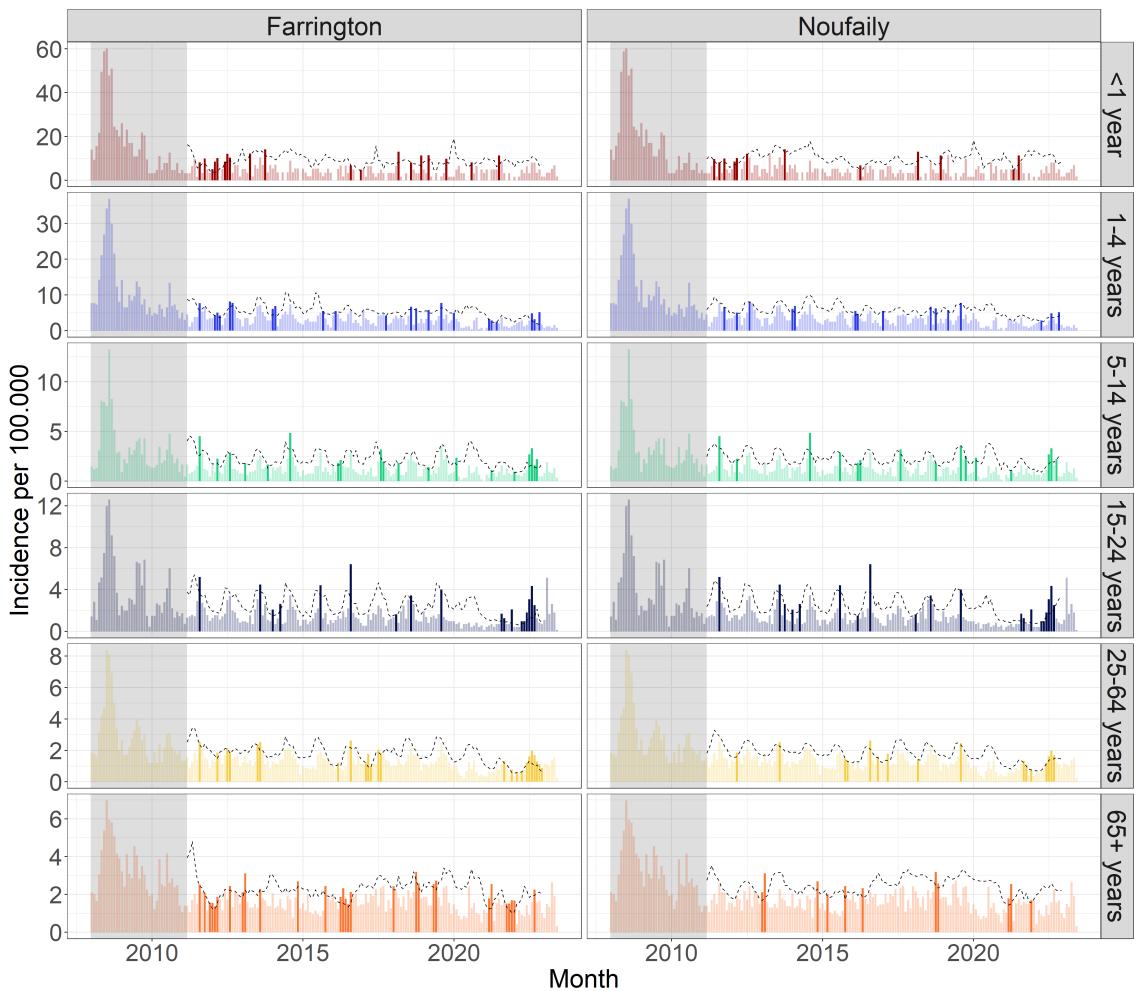


Figure 5.12: Monthly SALM incidence per 100,000 in Denmark, 2008–2022. Monitored by Farrington (left) and Noufaily (right) method. Reference data (grey area) for the estimation of model parameters are in the period from January 2008 to March 2011. Threshold (dashed line) is computed for observations time points outside reference data. Alarms are triggered (full opacity) if observations exceed threshold.

Applying the novel outbreak detection algorithm to *Salmonellosis*

In Figure 5.13 the estimated one-step ahead random effects \hat{u}_{it_1} are visualized together with the calculated threshold U_{t_0} .

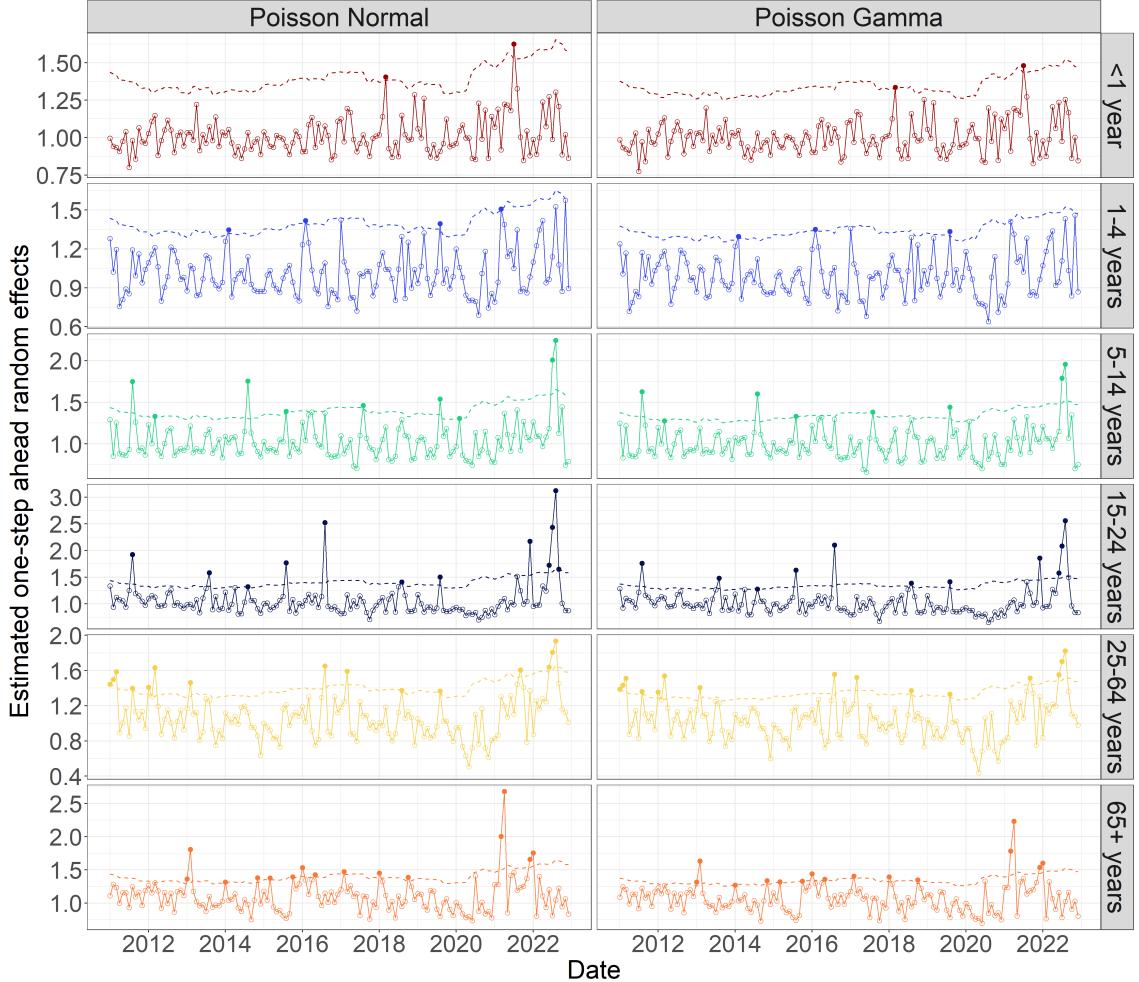


Figure 5.13: Estimated one-step ahead random effects \hat{u}_{it_1} (circles) for SALM in Denmark, 2011-2022, are shown for the combined model. An alarm is raised (solid circle) if \hat{u}_{it_1} exceeds the threshold U_{t_0} (dashed line). In the hierarchical Poisson Normal model (left), the random effects are exponentiated to $\exp(\hat{u}_{it_1})$ to transform them into the same domain as the hierarchical Poisson Gamma model (right).

In total, the hierarchical Poisson Normal framework generates 57 alarms, while the hierarchical Poisson Gamma framework generates 54 alarms. This demonstrates that the novel outbreak detection algorithm has lower sensitivity compared to the state-of-the-art algorithms.

However, the large number of alarms raises valid concerns about how to handle them effectively. In the novel algorithm, observations that trigger an alarm are identified and excluded from future parameter estimation to prevent inflating the parameter estimates. This approach aims to mitigate the impact of false alarms and maintain the accuracy of the modeling process. However, as the alarms accumulate, the effective sample window for parameter estimation diminishes, reducing the amount of available data for estimating reliable parameters. This challenge is expanded upon in Section 5.3.

5.3 Challenges in statistical outbreak detection

In this section, the challenges encountered in this master thesis regarding statistical outbreak detection are outlined. These challenges encompass specific issues related to the

use of hierarchical modeling as well as more generalized challenges that are relevant to statistical modeling in general.

The discussion will focus on the role of overdispersion in statistical outbreak detection, the impact of context and bias, and the handling of a large number of alarms.

5.3.1 Investigating the role of overdispersion in statistical outbreak detection for *Listeriosis*

Convergence issues were observed in the hierarchical models during the case studies, particularly in the LIST case study. Only the constant fixed effects models were able to converge, indicating challenges with the other models. The low incidence rate in the LIST data and the absence of overdispersion are likely factors contributing to these convergence difficulties.

To further investigate the lack of convergence, the empirical mean and variance of the LIST data series in the two age groups are visualized in Figure 5.14 using a rolling window approach. The rolling window width was set to match the parameter estimation window, which was 36 months. This visualization provides a closer examination of the patterns in the data and helps to identify potential factors affecting convergence.

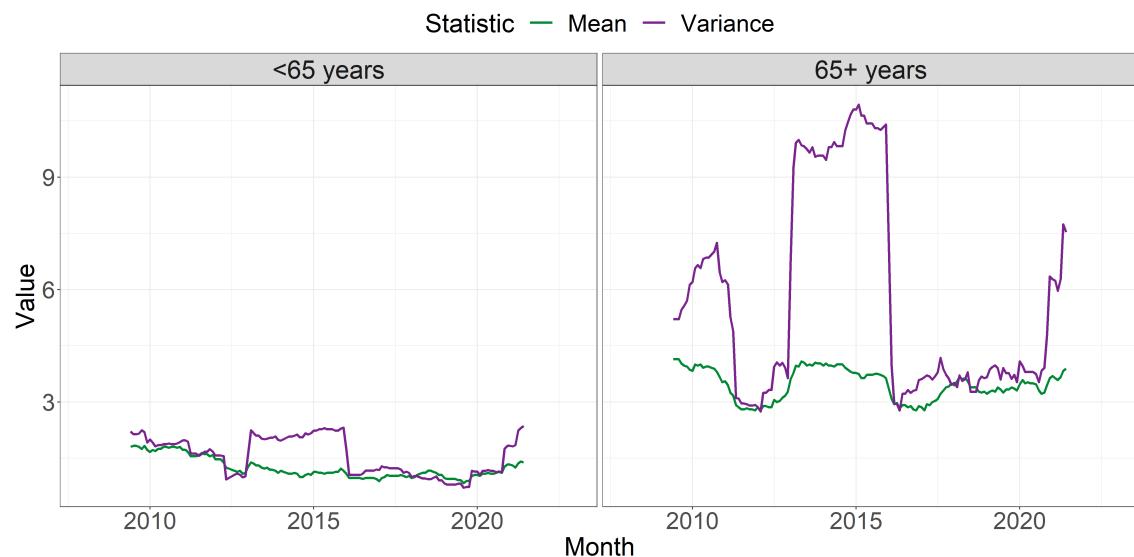


Figure 5.14: Empirical mean and variance of the LIST data in the age group below 65 years and 65 years and above

The figure clearly shows prolonged periods in the sample window where there is no overdispersion, as indicated by the similarity between the empirical mean and variance. The duration of the window with overdispersion is directly influenced by the chosen window size, and even a small number of observations can introduce notable overdispersion.

It is worth noting that the periods where the estimated second stage model parameters ($\hat{\sigma}$ or $\hat{\phi}$) fail to converge align with periods where statistical evidence for an ongoing outbreak is not feasible. These periods are characterized by a low number of cases and, more importantly for the hierarchical models, the absence of overdispersion in the sample window.

As the number of cases are very low, it is also unlikely that an outbreak is occurring. Regardless, it is not possible for statistical models to detect outbreak signals in these

periods.

In the case of the novel outbreak detection algorithm, particular attention is given to periods with overdispersion. This is because the modeling frameworks assume that the distribution of random “noise” can be modeled by a Gaussian distribution or a Gamma distribution. In the absence of overdispersion, count data can be adequately modeled by a Poisson distribution alone, making an approach based on GLMs more suitable for these periods.

5.3.2 The impact of context and observational bias

Deploying an outbreak detection algorithm blindly, without considering the context and potential changes in observational bias, can lead to unreliable results. It is essential to closely monitor the algorithms and critically assess the model inputs and the specific context in which it operates. This is evident in the case study of SHIL, where an outbreak investigated by SSI in August and September 2020 was challenging to detect using statistical methods due to the concurrent COVID-19-related lockdown and the resulting decrease in reported cases.

Considering the specific circumstances and potential changes in observational bias is crucial when interpreting outbreak detection results. It is important to determine whether the number of cases during an outbreak is considered aberrant compared to similar periods in other years. However, in situations where there is a sudden decrease in reported cases, such as during a lockdown, detecting outbreaks using statistical methods becomes more challenging.

An important potential extension of automated statistical outbreak detection algorithms is the incorporation of data drift monitoring. This involves detecting changes in the underlying data distribution and adapting the model accordingly to maintain its predictive performance. Specifically, detecting sudden decreases in case counts can be considered a form of data drift.

Indeed, sudden decreases in case counts can be treated similarly to sudden increases in case counts, and detecting such data drift can be incorporated into hierarchical models using the unobserved one-step ahead random effects u_{it_1} . One approach is to define a lower threshold L_{t_0} by selecting an appropriate quantile of the random effects distribution obtained from the second stage model, such as the 5% or 10% quantile.

As a proof of concept, Figure 5.15 presents the visualization of the estimated one-step ahead random effects for the SHIL data, accompanied by a lower bound defined as the 10% quantile in the random effects distribution.

The extension to the novel method demonstrates potential, with a total of 5 and 6 data drift alarms raised in the hierarchical Poisson Normal and Poisson Gamma model, respectively. These data drift alarms are visualized in Figure 5.15. When a data drift alarm occurs, it should be of concern to the epidemiologist, prompting further investigation into the reasons behind the abnormally low numbers. It is crucial to critically assess the input data and the fixed effects model employed recursively to ensure optimal performance of the automated procedure.

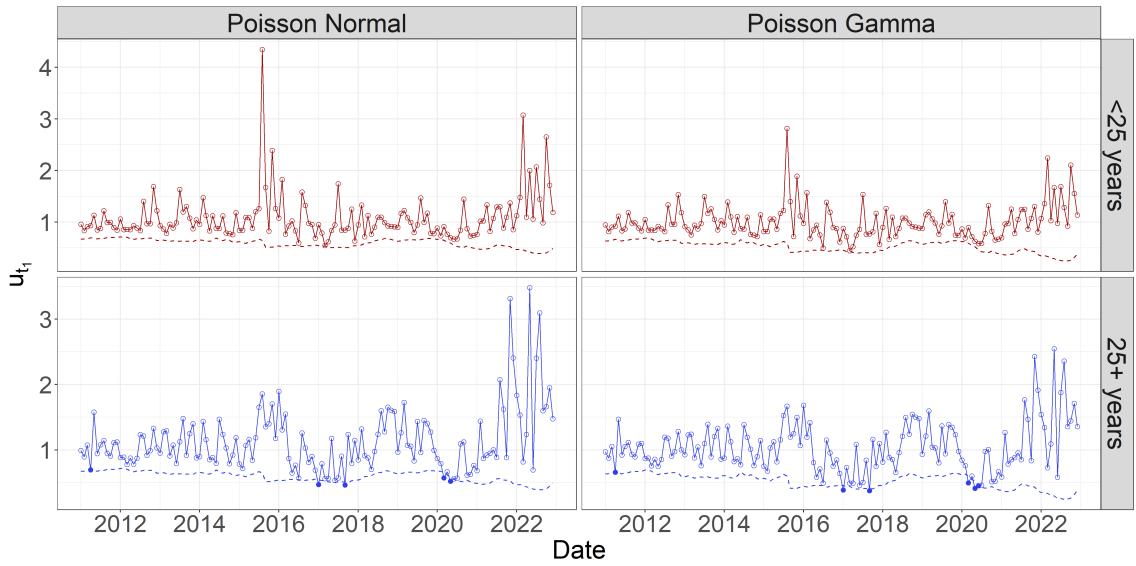


Figure 5.15: Estimated one-step ahead random effects \hat{u}_{it_1} (circles) for SHIL in Denmark, 2011-2022, are shown for the trend fixed effects model described in (5.2). An alarm is raised (solid circle) if \hat{u}_{it_1} transcend the lower threshold L_{t_0} (dashed line) based on the 10% quantile in the random effects distribution. In the hierarchical Poisson Normal model (left), the random effects are exponentiated to $\exp(\hat{u}_{it_1})$ to transform them into the same domain as the hierarchical Poisson Gamma model (right).

The inclusion of data drift monitoring can improve the algorithms' ability to adapt to changing outbreak patterns and contextual factors, leading to more reliable and accurate detection outcomes.

5.3.3 How do we handle many alarms?

Dealing with diseases that generate frequent alarms presents a challenge in statistical outbreak detection. The novel outbreak detection algorithm addresses this challenge by excluding flagged observations from future parameter estimation. This exclusion reduces the sample window for parameter estimation and results in a loss of information and reduced variation in the data. The purpose of this exclusion is to prevent inflated parameter estimates.

On the other hand, the state-of-the-art detection algorithm employs a re-weighting scheme using Anscombe residuals. It has been observed that the Farrington method's correction for past outbreaks is too drastic, and a higher re-weighting threshold of 2.58 rather than 1 is recommended in the Noufaily method to improve specificity (Noufaily et al. 2013).

Effectively handling previous outbreaks in statistical outbreak detection requires domain-specific knowledge from epidemiologists to determine the most appropriate approach. While the novel method completely excludes past outbreaks, this may not be the optimal practice as it reduces the sample window. A clear example of this can be seen in the SALM case study. In Figure 5.16, the effective number of observations used for parameter estimation at a given point in time is visualized.

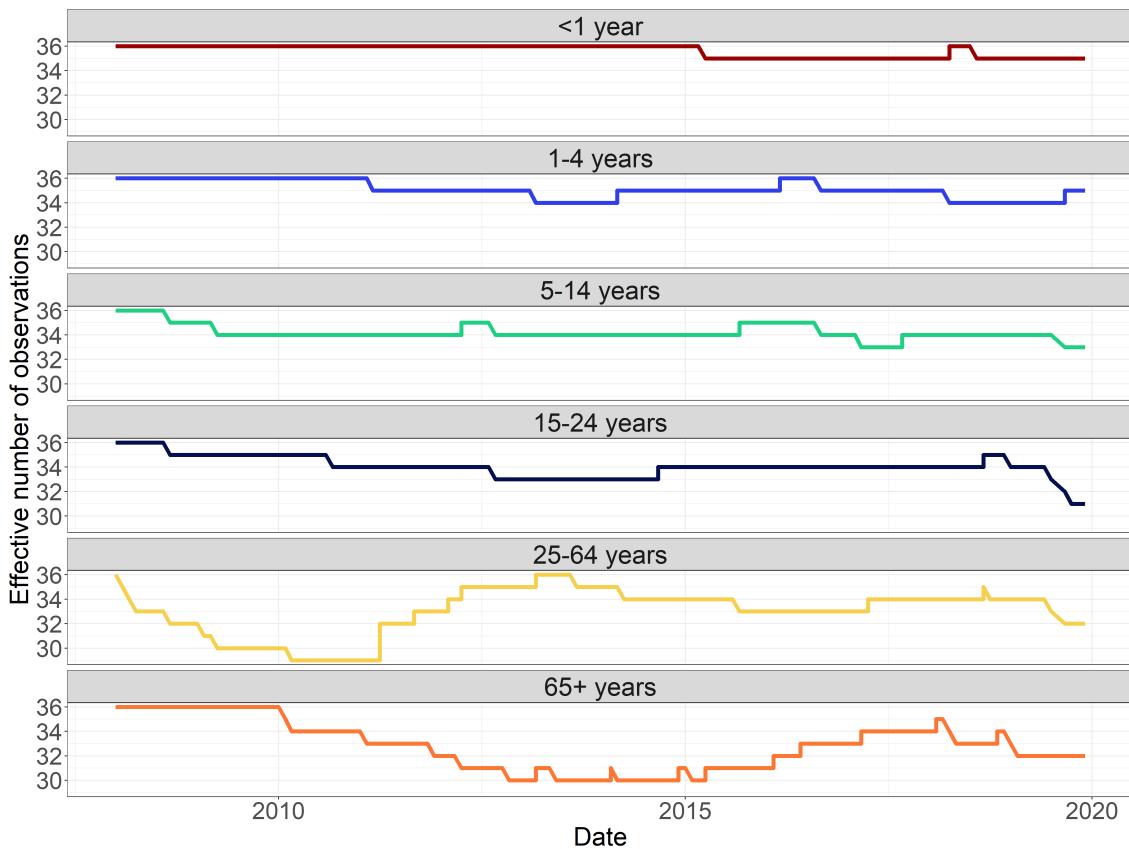


Figure 5.16: Effective number of observations in the SALM case study, that enter the parameter estimation at a given point in time.

If the method is employed in an automated setting, it is essential to report the size of the sampling window to the epidemiologists so that they can interpret the results appropriately. In extreme cases, one could imagine that the number of observations used for parameter estimation could approach zero, leaving the model with insufficient data to adequately describe the process.

5.4 Performance of statistical outbreak detection algorithms

In this section, the performance of both the state-of-the-art and novel outbreak detection algorithms is evaluated to assess their ability to identify outbreaks in the case studies. A comparison is made between the alarms raised by each method using a timeline plot. Additionally, the first identified cases in the outbreaks investigated by SSI are included in the plot to provide context and evaluate the accuracy and timeliness of the outbreak detection algorithms.

The overall caveat in evaluating methods for automated an early detection of disease outbreaks, is the problem of defining an outbreak. In the case studies examined in this thesis, an outbreak is defined as something officially investigated and documented by SSI. However, this definition is neither standardized, nor reproducible, and depends largely on what the individual epidemiologists have perceived as an outbreak. Thus, in the absence of a gold standard definition for outbreak, these documented investigations have been used.

Figure 5.17 demonstrates that the state-of-the-art outbreak detection algorithms, espe-

cially the Farrington method, raise a higher number of alarms compared to the novel algorithms in the LIST case study.

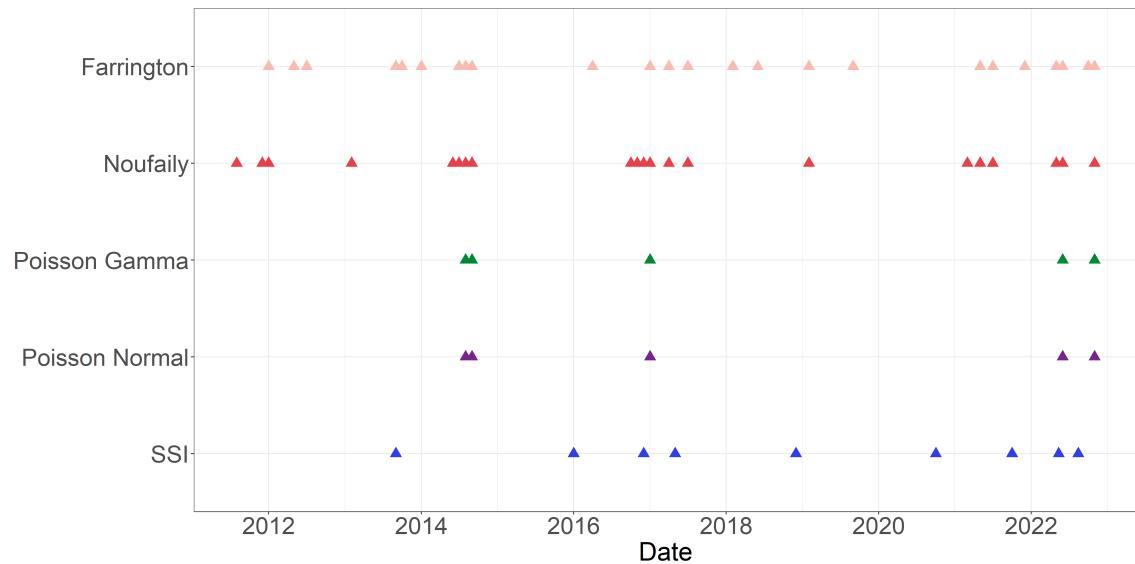


Figure 5.17: Alarm plot displaying alarms for LIST time series using four different algorithms, along with outbreaks investigated by SSI.

In general, SSI initiates an outbreak investigation when there are either 2 cases with identical bacteria over a 12-week period or 3 cases at any given time. Therefore, it is crucial to have a method with high sensitivity for automated LIST outbreak detection. As discussed earlier, the Farrington method demonstrates lower sensitivity compared to the other algorithms, which may impact its suitability for detecting LIST outbreaks.

Furthermore, it was discovered that the hierarchical models faced inherent convergence issues due to the lack of overdispersion in the data. This resulted in insufficient information to adequately inform the second stage model and impacted the performance of the models in detecting outbreaks. Consequently, the novel method had prolonged periods where the estimated random effects and thresholds collapse.

However, it is worth noting that the novel outbreak detection algorithms successfully identified five outbreaks that aligned with ongoing investigations conducted by SSI.

In the SHIL case study, the incidence of cases is slightly higher, resulting in a relatively higher number of alarms generated by both the state-of-the-art and novel outbreak detection methods. This can be observed in Figure 5.18. It is important to note that only one well-documented outbreak of SHIL was investigated by SSI in the observed period. As a result, the number of alarms generated by the methods may be overestimated in terms of the priority given to investigating outbreaks in this specific disease.

Additionally, it is worth noting that none of the algorithms were able to identify the outbreak in the summer of 2020, as it occurred during the COVID-19-related lockdown period when there were very few reported cases. This highlights the challenge of detecting outbreaks in periods of low incidence or during exceptional circumstances such as a lockdown.

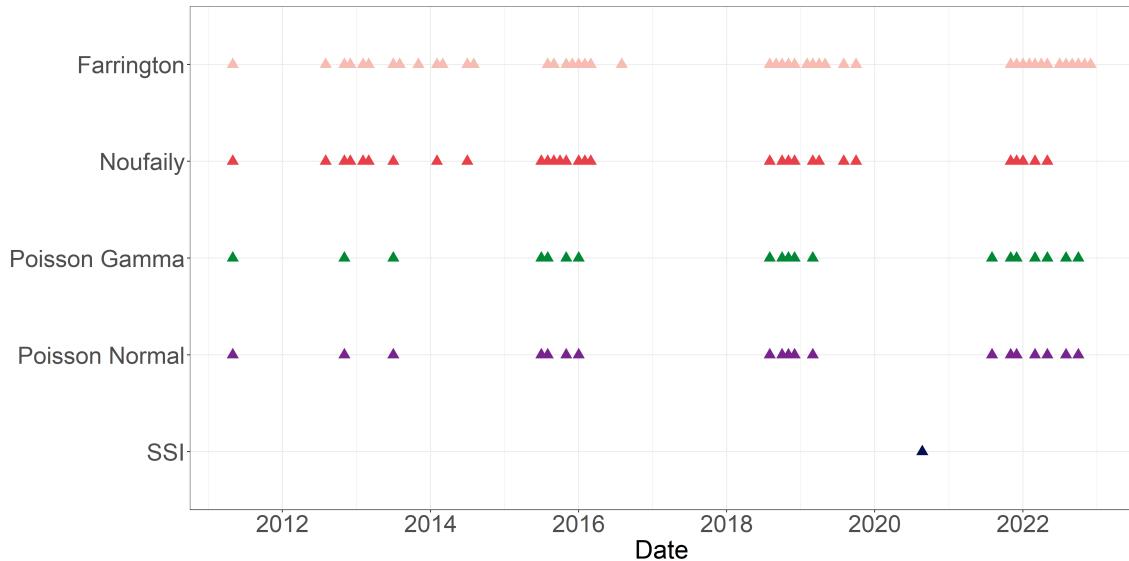


Figure 5.18: Alarm plot displaying alarms for SHIL time series using four different algorithms, along with outbreaks investigated by SSI.

Furthermore, it is interesting to mention that a total of 10 clusters with SHIL have been reported to the FUD between 2008 and 2022. Unfortunately, due to the limitations of this thesis, it was not possible to retrieve the necessary information to provide a detailed description of these outbreaks. Therefore, with the available information, it is not possible to reject the possibility that some of the alarms raised by the algorithms coincided with one or more of these clusters.

When examining the STEC data, the substantial difference in sensitivity between the state-of-the-art and novel outbreak detection methods becomes evident. Figure 5.19 clearly demonstrates that both state-of-the-art methods generate a significant number of alarms. This excessive number of alarms raises concerns about the practicality and feasibility of these methods in real-world applications.

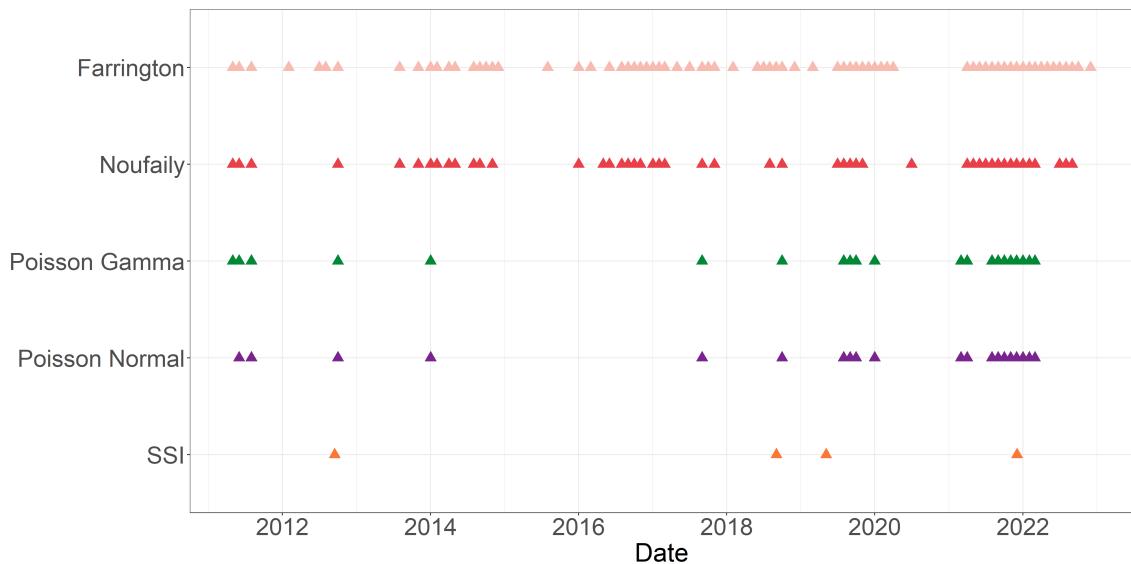


Figure 5.19: Alarm plot displaying alarms for the STEC time series using four different algorithms, along with outbreaks investigated by SSI.

It is worth noting that both the state-of-the-art and novel outbreak detection algorithms are capable of correctly raising alarms at several, if not all, of the investigated outbreaks. This highlights the effectiveness of these algorithms in identifying potential outbreaks and their ability to contribute to timely public health interventions. However, it is important to consider other factors such as the number of false positives and the overall usability of the methods in real-world settings.

In the case study of SALM, a large number of alarms are generated by both the state-of-the-art and novel outbreak detection algorithms. These alarms are depicted in Figure 5.20. The vast number of alarms can be attributed to the inherent nature of the disease, as SALM is known for frequently causing outbreaks. Therefore, the high number of alarms generated in this case is not unexpected.

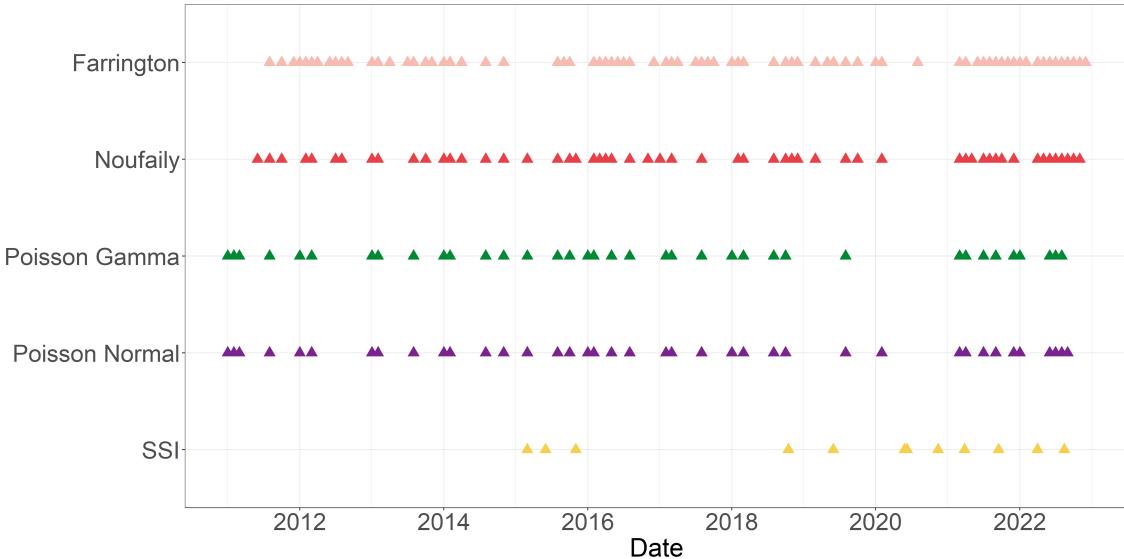


Figure 5.20: Alarm plot displaying alarms for SALM time series using four different algorithms, along with outbreaks investigated by SSI. Note that there are several more clusters with SALM, than what is included in this plot. This is because of missing or poor documentation of these clusters. However, for an overview over the clusters reported to FUD, refer to Table C.1.

For future work, it is indeed beneficial to consider monitoring individual serotypes of *Salmonella* bacteria instead of monitoring them collectively. This approach would enable more specific and targeted outbreak detection. By focusing on specific serotypes, the algorithms can potentially reduce the number of false alarms and enhance the overall effectiveness of outbreak signal detection for SALM infections.

Furthermore, monitoring individual serotypes provides an opportunity to prioritize outbreaks of *Salmonella* serotypes that are known to cause particularly severe diseases. This allows for a proactive response in identifying and containing outbreaks that pose a higher risk to public health. By tailoring the detection algorithms to specific serotypes, public health authorities can effectively prioritize their efforts and allocate resources accordingly, leading to more efficient and targeted outbreak management strategies.

5.5 Potential extension of the novel algorithm

A potential extension of the novel outbreak detection method proposed in this master thesis involves incorporating correlation, either in time or between groups. To highlight this, an example considering the SALM case study is presented. The autocorrelation function (ACF) plot in Figure 5.21 provides insights into the temporal dependencies and correlations within the data.

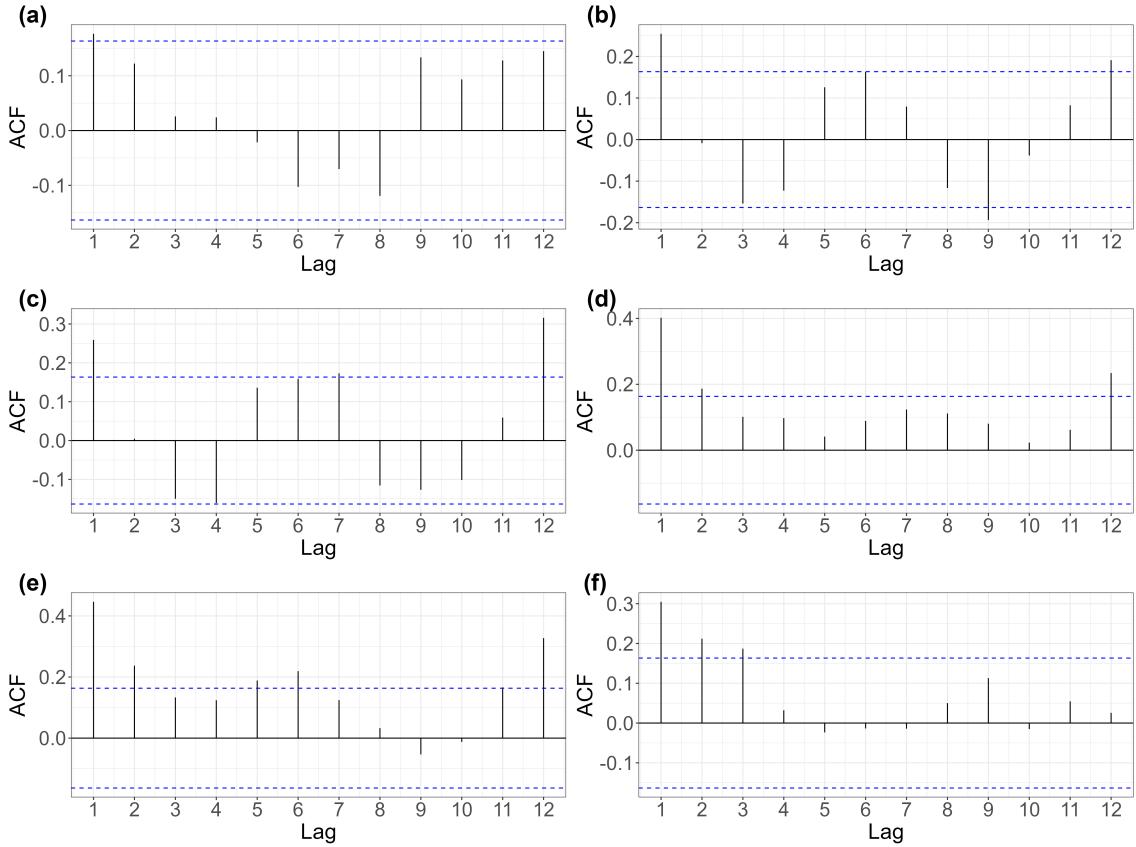


Figure 5.21: ACF for the estimated one-step ahead random effects \hat{u}_{it_1} for the six age groups in the SALM case study. The plots are organized in increasing age group, i.e. (a) <1 year, (b) 1-4 years, (c) 5-14 years, (d) 15-24 years, (e) 25-64 years, and (f) 65+ years.

The ACF plot reveals significant autocorrelation at lag 1 across all age groups, indicating a strong temporal dependency. Some age groups, such as 15-24 years, 25-64 years, and 65+ years, exhibit significant autocorrelation at lag 2. Additionally, a notable seasonal autocorrelation is observed at lag 12 in most age groups. These findings suggest the presence of underlying patterns and correlations that can be further explored and leveraged in the development of the outbreak detection method.

The correlation plot in Figure 5.22 provides insights into the intergroup correlation, highlighting potential shared patterns or dependencies across different age groups.

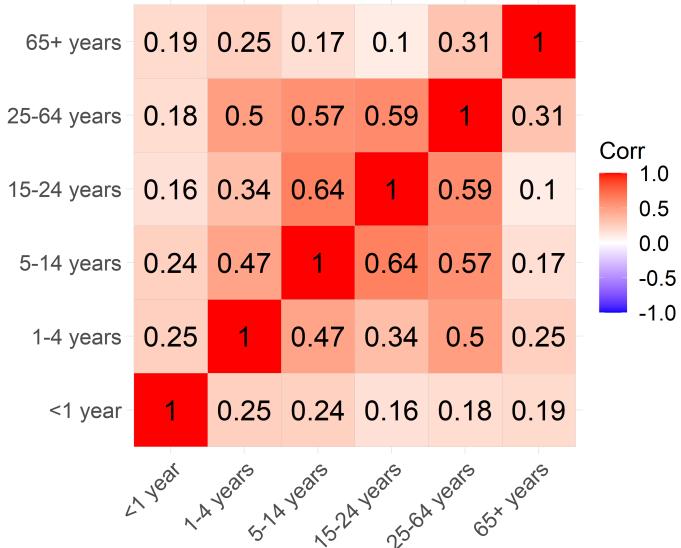


Figure 5.22: Correlation of the estimated one-step ahead random effects \hat{u}_{it_1} for the six age groups in the SALM case study.

The plot clearly shows a strong intergroup correlation between the age groups, particularly between 5-14 years, 15-24 years, and 25-64 years. This suggests that these age groups may exhibit similar disease patterns or dynamics.

Analyzing the intergroup correlation can provide valuable information for understanding the dynamics and interactions between different age groups. It can help identify commonalities and differences in disease patterns, enabling targeted intervention strategies and efficient resource allocation.

Further exploration and analysis of the intergroup correlation can enhance our understanding of disease dynamics and improve the accuracy and effectiveness of the outbreak detection method, leading to better public health outcomes.

6 Simulation study

In this chapter, a simulation study is conducted to evaluate the performance of the novel outbreak detection algorithm compared to state-of-the-art algorithms. The simulations cover various scenarios, adapted from the study by Noufaily et al. (2013).

The chapter begins by describing the method used to simulate the baseline data. These data are generated using a Negative Binomial model with a time-dependent mean $\mu(t)$. Next, the assumptions regarding the simulated outbreaks are outlined, including the outbreak size and distribution in time.

The evaluation measures used to assess the performance of the outbreak detection algorithms are then presented. These measures are designed to capture relevant quantities in the context of outbreak detection.

Finally, the chapter presents the results of applying both the state-of-the-art and the novel outbreak detection algorithms to the simulated data. The performance of both algorithms is evaluated based on the simulation results.

6.1 The simulated baseline data

The simulated baseline data is generated using a Negative Binomial model with a mean parameter μ and a variance parameter $\phi\mu$. The dispersion parameter ϕ is assumed to be greater than or equal to 1. The mean $\mu(t)$ is defined by a linear predictor that includes a trend component and a seasonality component represented by Fourier terms.

The equation for $\mu(t)$ is given as:

$$\mu(t) = \exp\left(\theta + \beta_t + \sum_{j=1}^m \left(\gamma_1 \cos\left(\frac{2\pi jt}{52}\right) + \gamma_2 \sin\left(\frac{2\pi jt}{52}\right)\right)\right) \quad (6.1)$$

In this equation, m represents the number of Fourier terms used to model seasonality. When $m = 0$, it indicates the absence of seasonality, while $m = 1$ corresponds to annual seasonality.

To cover a wide range of data sets encountered in practical applications, 28 different parameter combinations are generated. These combinations vary in terms of trends (represented by different values of β), seasonalities (represented by different values of γ_1 and γ_2), baseline frequencies of reports (represented by different values of θ), and dispersion (represented by different values of ϕ). The specific parameter values for the 28 scenarios are provided in Table 6.1.

Table 6.1: Parameters and criteria utilized to generate the 28 scenarios.

Scenario	θ	ϕ	β	γ_1	γ_2	m	Trend
1	0.10	1.5	0.0000	0.00	0.00	0	0
2	0.10	1.5	0.0000	0.60	0.60	1	0
3	0.10	1.5	0.0025	0.00	0.00	0	1
4	0.10	1.5	0.0025	0.60	0.60	1	1
5	-2.00	2.0	0.0000	0.00	0.00	0	0

Table 6.1: (*continued*)

Scenario	θ	ϕ	β	γ_1	γ_2	m	Trend
6	-2.00	2.0	0.0000	0.10	0.30	1	0
7	-2.00	2.0	0.0050	0.00	0.00	0	1
8	-2.00	2.0	0.0050	0.10	0.30	1	1
9	1.50	1.0	0.0000	0.00	0.00	0	0
10	1.50	1.0	0.0000	0.20	-0.40	1	0
11	1.50	1.0	0.0030	0.00	0.00	0	1
12	1.50	1.0	0.0030	0.20	-0.40	1	1
13	0.50	5.0	0.0000	0.00	0.00	0	0
14	0.50	5.0	0.0000	0.50	0.50	1	0
15	0.50	5.0	0.0020	0.00	0.00	0	1
16	0.50	5.0	0.0020	0.50	0.50	1	1
17	2.50	3.0	0.0000	0.00	0.00	0	0
18	2.50	3.0	0.0000	1.00	0.10	1	0
19	2.50	3.0	0.0010	0.00	0.00	0	1
20	2.50	3.0	0.0010	1.00	0.10	1	1
21	3.75	1.1	0.0000	0.00	0.00	0	0
22	3.75	1.1	0.0000	0.10	-0.10	1	0
23	3.75	1.1	0.0010	0.00	0.00	0	1
24	3.75	1.1	0.0010	0.10	-0.10	1	1
25	5.00	1.2	0.0000	0.00	0.00	0	0
26	5.00	1.2	0.0000	0.05	0.01	1	0
27	5.00	1.2	0.0001	0.00	0.00	0	1
28	5.00	1.2	0.0001	0.05	0.01	1	1

To simulate the baseline data without outbreaks, 100 replicates are generated for each of the 28 parameter scenarios. Each replicate consist of a time series of size $T = 624$ weeks.

The 624 weeks are divided into three periods: weeks 1-313 are used for training, weeks 313-575 are considered as baseline weeks, and weeks 576-624 are designated as the test weeks for evaluation. In Figure 6.2, one randomly selected realization for scenario 5, 7, 12, and 28 are visualized.

The simulation results are based on the test weeks of all the replicates, totaling $100 \times 49 = 4900$ weeks, for each of the 28 data scenarios and each method investigated.

6.2 The simulated outbreaks

The outbreaks starting in week t_i are simulated using the following procedure. First, a constant value k is chosen at random. The size of the outbreak, denoted as v , is then generated randomly from a Poisson distribution with a mean equal to k times the standard deviation of the baseline count in that scenario.

Next, the outbreak is distributed randomly in time according to a discretized log-normal distribution with a mean of 0 and a standard deviation of 0.5, represented as $Z \sim [\text{LN}(0, 0.5^2)]$. This is achieved by drawing v random numbers, which correspond to the outbreak size,

from the specified log-normal distribution and then rounding down these numbers to the nearest integer.

The probability mass function for the discretized log-normal distribution is visualized in Figure 6.1.

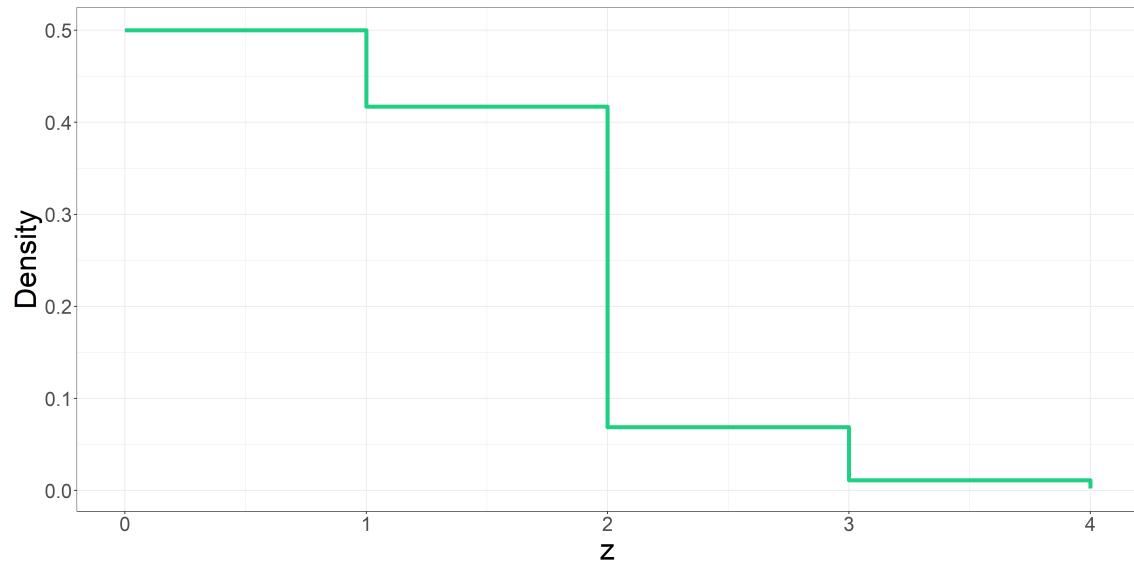


Figure 6.1: Stairstep plot of the probability mass function for the discretized log-normal distribution with a mean of 0 and a standard deviation of 0.5, i.e. $Z \sim \lfloor \text{LN}(0, 0.5^2) \rfloor$.

Typically, outbreak durations of 2-4 weeks are observed when values of k are in the range of 2-10. To simulate the outbreaks, the outbreak cases are added to the baseline count in week $t_i + z_i$, where t_i represents the start time of the outbreak and z_i represents the number of weeks after the start of the outbreak. The start and end times of the outbreaks are recorded for evaluating the performance of the methods.

To simulate outbreaks, the following procedure is followed:

- **Outbreaks in baseline weeks:** For each data series, four outbreaks are generated. The start time of each outbreak is randomly selected from the baseline weeks (weeks 313-575). The value of k is sampled randomly with replacement from the set $\{2, 3, 5, 10\}$. It should be noted that different outbreaks are generated for each of the 2800 runs.
- **Outbreaks in current weeks:** For each data series, one outbreak is generated. The start time of the outbreak is randomly chosen from the last 49 weeks (weeks 576-624). The value of k is sampled randomly in the range of 1 to 10. Similar to the previous case, different outbreaks are generated for each of the 2800 runs.

One randomly chosen realization for scenario 8, 12, 13, and 20 are visualized in Figure 6.2.

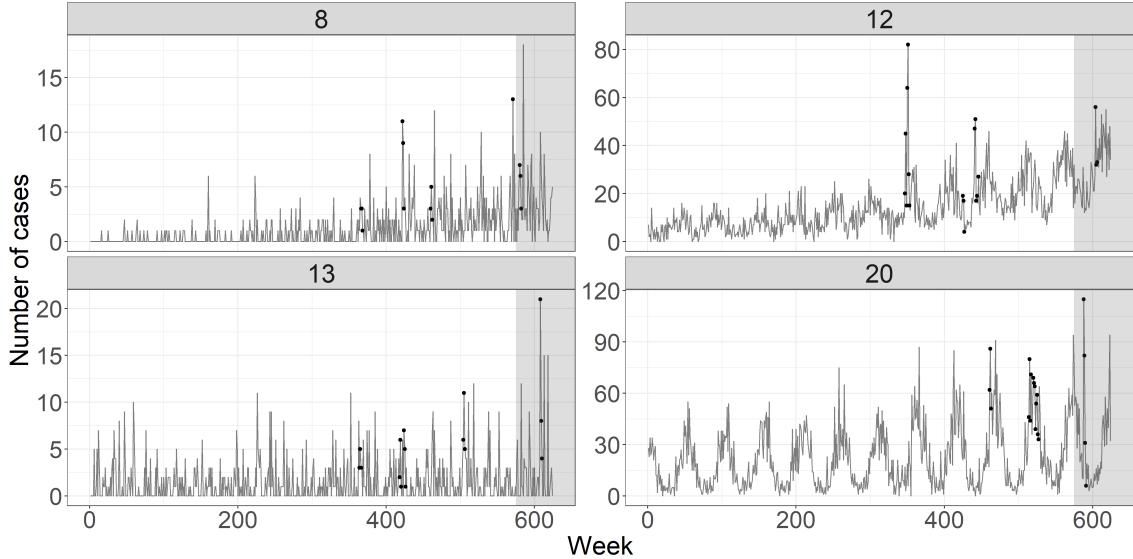


Figure 6.2: Plots of one randomly chosen realization for scenario 8, 12, 13, and 20 (see Table 6.1). During outbreaks (circles), outbreak cases are added to the baseline data. Four outbreaks are added during the baseline weeks and 1 outbreak is added during the test weeks. The results are based on the data obtained in the test weeks (grey area).

Evidently, the scenarios vary in their epidemiological characteristics, such as seasonality, trend, and incidence.

6.3 Evaluation measures

To evaluate the performance of the outbreak detection system in the absence and presence of outbreaks, several measures are employed to assess its effectiveness. These measures are specifically designed to capture relevant quantities in the given context.

In the absence of outbreaks in the data, one of the primary measures used is the False Positive Rate (FPR). This is calculated for each of the 28 scenarios, before the addition of the simulated outbreaks to the baseline data. The FPR is determined by calculating the proportion of the 49 weeks and 100 replicates in which the observed value exceeds the threshold in the absence of any simulated outbreaks.

Another measure is the Probability Of Detection (POD) of an outbreak. Likewise, this is calculated for each of the 28 scenarios, but this time it is in the presence of the simulated outbreaks. The algorithm is applied iteratively for the 49 current weeks, and an outbreak is considered detected if the observed value exceeds the threshold at least once within the start and end times of the outbreak. The POD of an outbreak is then determined as the proportion of outbreaks detected out of the 100 runs.

It is important to note that the FPR is a rate per week, while the POD is a rate per realization. These evaluation measures are chosen because they provide insights into the performance of the detection system on individual time series.

6.4 Results of the simulations

In this section, the results of the simulation study are presented. Both the state-of-the-art and novel outbreak detection algorithms are applied to the same simulated data, allowing for the evaluation of their performance in a setting where the true parameters and process are known.

The data used consists of weekly counts of a simulated disease, denoted as y_t , where $t = 1, \dots, T$ represents the time period of $T = 624$ weeks. In order to evaluate the state-of-the-art and novel outbreak detection algorithms, it is decided to include 5 years of reference data, corresponding to a total of 260 weeks.

The first step involves applying the state-of-the-art outbreak detection algorithms, with specific control arguments determined for the Farrington and Noufaily methods. These control arguments for the algorithms are specified exactly as in Noufaily et al. (2013), in order to reproduce the same results. Notably, the thresholds calculated in the state-of-the-art algorithms are based on the 99.5% quantiles. Also, in the Noufaily method, it is decided to exclude the last 26 weeks before the current week to avoid adaption of the model to emerging outbreaks, which would reduce its sensitivity. In the Farrington method, only 3 weeks are excluded before the current week. To access the full configuration of the state-of-the-art methods, refer to Appendix B.1.

Hereafter, the novel outbreak detection algorithm is applied. For this purpose, the model for the fixed effects in each of the modeling frameworks is defined as

$$\log(\lambda_t) = \beta_{intercept} + \beta_{trend}t + \sin\left(\frac{2\pi \cdot \tau_t}{52}\right)\beta_{\sin} + \cos\left(\frac{2\pi \cdot \tau_t}{52}\right)\beta_{\cos} \quad (6.2)$$

Unlike the state-of-the-art algorithms, the thresholds in the novel algorithm are based on the 95% quantile of the distribution of the random effects in the second stage model. This threshold is calculated based on either a Gaussian distribution using the plug-in estimate $\hat{\sigma}$ or a Gamma distribution using the plug-in estimate $\hat{\phi}$ depending on the modeling framework.

False Positive Rates

In general, the method introduced by Farrington et al. (1996) tends to have relatively higher FPRs compared to the other methods. This observation is consistent with the results presented in Table 6.2. However, this outcome is not surprising since the Farrington method is known to be overly sensitive, making it more prone to producing false alarms. On the other hand, the improved method by Noufaily et al. (2013) outperforms the other methods by minimizing the FPR.

Additionally, it is noted that the novel algorithms, using the two different modeling frameworks, perform somewhere in between the two state-of-the-art algorithms in terms of minimizing the FPR.

Table 6.2: Summary statistics of the FPRs obtained in the 28 scenarios using the four different methods.

Method	median(FPR)	mean(FPR)	sd(FPR)	min(FPR)	max(FPR)
Farrington	0.0217	0.0306	0.0310	0	0.2174
Noufaily	0.0000	0.0091	0.0153	0	0.1250
Poisson Normal	0.0208	0.0165	0.0203	0	0.1277
Poisson Gamma	0.0208	0.0165	0.0203	0	0.1277

Upon examining Figure 6.3, it becomes even more apparent that the Noufaily method consistently outperforms the other methods. Furthermore, it is interesting to note that scenario 8 and 15 consistently pose challenges for all methods, while scenarios 13-20 prove to be particularly problematic for the novel method.

A closer look reveals that scenarios 13-20 have the highest overdispersion parameters, indicating increased variability in the data. On the other hand, scenario 8 incorporates both a steep trend and a seasonality component, which can complicate the detection process for all methods.

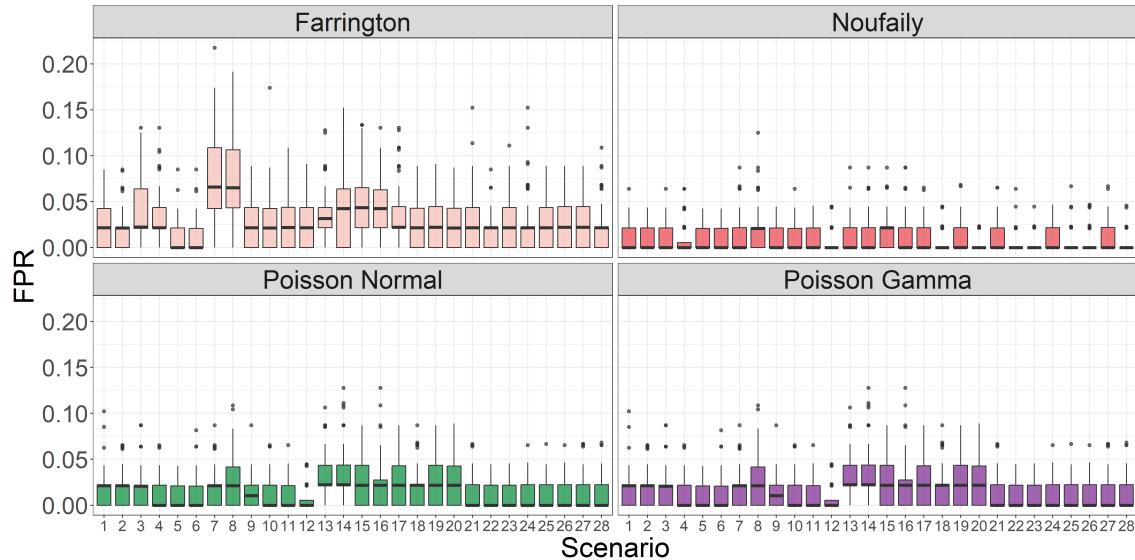


Figure 6.3: FPRs obtained in each of the 28 scenarios for each of the methods applied.

Probability an outbreak is detected

As expected, the POD of an outbreak increases with the size of k . Intuitively, when the outbreak size v is larger, it becomes more likely to be detected by the outbreak detection algorithms. In Table 6.3, it is evident that the Farrington method performs very well in terms of POD, closely followed by the novel method using either modeling framework. Moreover, it can be seen that the Noufaily method is outperformed by the other methods, w.r.t. POD.

The high performance of the Farrington method can be attributed to its sensitivity in detecting outbreaks. Similarly, the novel method utilizing both modeling frameworks demonstrates its effectiveness in detecting outbreaks of varying sizes.

Table 6.3: Summary statistics of the POD of an outbreak of size k times the standard deviations of the baseline data for each of the methods applied.

Method	k	median(POD)	mean(POD)	sd(POD)	min(POD)	max(POD)
Farrington	2	0.2614	0.2850	0.1792	0.0000	0.6250
	4	0.4615	0.4926	0.1729	0.0909	0.7778
	6	0.8452	0.7997	0.1742	0.3750	1.0000
	8	0.9706	0.9140	0.1468	0.3750	1.0000
	10	1.0000	0.9324	0.0957	0.7143	1.0000
Noufaily	2	0.1111	0.1275	0.1081	0.0000	0.4000
	4	0.3000	0.3276	0.1821	0.0909	0.7500
	6	0.7208	0.6819	0.2493	0.1250	1.0000
	8	0.8452	0.8224	0.1931	0.2500	1.0000
	10	1.0000	0.9131	0.1086	0.7000	1.0000

Table 6.3: (continued)

Method	k	median(POD)	mean(POD)	sd(POD)	min(POD)	max(POD)
Poisson Normal	2	0.2435	0.2670	0.1742	0.0000	0.7143
	4	0.4410	0.4596	0.1899	0.1000	0.8182
	6	0.8258	0.7583	0.2222	0.2353	1.0000
	8	1.0000	0.9034	0.1621	0.3750	1.0000
	10	1.0000	0.9448	0.1077	0.5714	1.0000
Poisson Gamma	2	0.2435	0.2670	0.1742	0.0000	0.7143
	4	0.4410	0.4596	0.1899	0.1000	0.8182
	6	0.8258	0.7583	0.2222	0.2353	1.0000
	8	1.0000	0.9034	0.1621	0.3750	1.0000
	10	1.0000	0.9448	0.1077	0.5714	1.0000

In Figure 6.4, the variability in PODs of outbreaks can be observed across the 28 scenarios. The level of variability in POD is generally low when the outbreak size factor k is set to 1, indicating that only a few outbreaks are detected in these scenarios. Similarly, when k is set to 10, indicating that almost all outbreaks are detected, the variability in POD is also low.

On the other hand, the variability in POD across the scenarios is highest when k is set to 5, and around 75% of the outbreaks are detected.

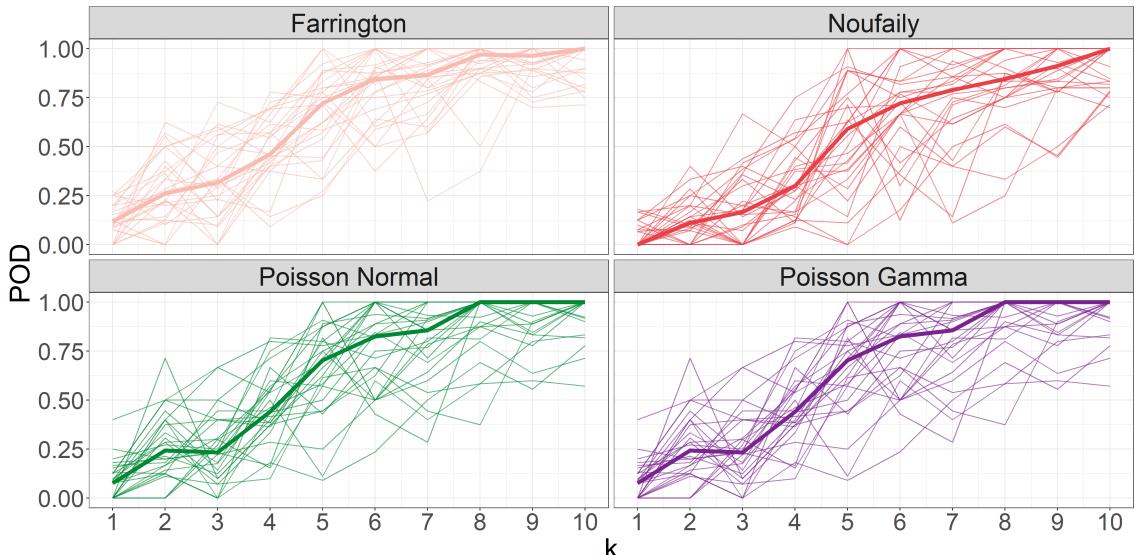


Figure 6.4: POD of an outbreak of a random size v drawn from a Poisson distribution with mean equal to k times the standard deviations of the baseline data. The x-axis shows increasing values of k . The POD for each scenario is plotted along with the median curves (bold) across all 28 scenarios.

It is important to bear in mind that an outbreak of size v is randomly distributed in time according to a discretized log-normal distribution with a mean of 0 and a standard deviation of 0.5, denoted as $Z \sim \lfloor \text{LN}(0, 0.5^2) \rfloor$. The probability mass function of Z is shown in Figure 6.1. From the figure, it can be observed that 50% of the outbreak cases are added to the

same week as the outbreak starts, 42% are added to the following week, and only 7% are added two weeks after the start. Therefore, the simulated outbreak cases are not observed in a single week only but rather in several concurrent weeks.

Consequently, an outbreak of size v generated from a Poisson distribution with a mean equal to k times the standard deviation of the baseline series is perceived to be relatively smaller than initially perceived in the simulation setup. For example, an outbreak of size $k = 4$ times the standard deviation may only be perceived as an outbreak signal of size 2 times the standard deviation in an individual week.

7 Discussion

In this chapter, several topics that have not been previously discussed will be addressed. This includes addressing the implications of overly sensitive outbreak detection methods, providing a summary of the considerations when deciding on a modeling framework in the novel method, and exploring the potential of MiBa-based surveillance.

7.1 Alarm fatigue and the implications for outbreak detection

The field of automated disease outbreak detection has witnessed significant advancements in recent years, with the development of sophisticated statistical methods and the utilization of computerized databases. These advancements have tremendous potential in supplementing epidemiologists and help them effectively prioritize their efforts and allocate resources accordingly, leading to more efficient and targeted outbreak management strategies. However, an emerging concern in the field is the phenomenon known as “alarm fatigue” and its implications for statistical outbreak detection.

Alarm fatigue refers to the desensitization and decreased responsiveness that can occur when epidemiologists are exposed to frequent or excessive alarms from a statistical outbreak detection method. Specifically, the method introduced by Farrington et al. (1996) are known for its occasional lack of specificity, leading to false alarms that can overwhelm the epidemiologists with verification tasks. Consequently, its usefulness and practicality in real-world settings are undermined. In an effort to address the number of false alarms, the subsequently improved method by Noufaily et al. (2013) was introduced. While it was found that this method greatly reduced the number of alarms while maintaining a good overall performance in detecting outbreaks, it has been unable to establish itself as a reliable method.

An outbreak detection algorithm that can effectively control the number of false alarms, while still maintaining the power to detect genuine outbreaks, is in increasingly high demand. In this thesis it is shown, that the proposed novel method have immense potential as a candidate tool for outbreak detection. It considerably reduces the number of alarms generated compared to the Farrington and Noufaily methods, while still maintaining the ability to generate alarms coinciding in time with outbreak investigations by SSI and during the simulated outbreaks. Hopefully, this result highlights the potential of the novel method and the utilization of hierarchical models as a framework for outbreak detection in general.

7.2 For and against for the two modeling frameworks in the novel method

The hierarchical Poisson Gamma model offers the advantage of directly interpreting parameter estimates as their actual contribution to disease intensity, thanks to the first moment of the resulting Negative Binomial distribution. In contrast, the hierarchical Poisson Normal model lacks a known expected value for the distribution of $Y|u$, and necessitates numerical methods like the Laplace approximation for parameter estimation and inference.

However, despite the differences in assumptions about the distribution of random effects, this master thesis demonstrates that parameter estimates and confidence intervals for the fixed effects model remain consistent across both modeling frameworks. Consequently,

the choice between the two frameworks may depend more on data characteristics and computational considerations rather than the interpretation of parameter estimates.

A notable advantage of the hierarchical Poisson Normal model is its ease of extension to incorporate correlation in time or between groups by including a covariance matrix in the second stage model. This extension is relatively straightforward within this framework. In contrast, implementing correlation in the hierarchical Poisson Gamma model is more complex and requires further investigation. However, including correlation has the potential to enhance the outbreak detection method's performance and capture additional patterns in the data, as observed in Figures 5.21 and 5.22 for the STEC case study.

Another significant difference between the two modeling frameworks is the run time efficiency. The hierarchical Poisson Gamma model exhibits faster computation, up to a factor of $4\times$ faster, compared to the hierarchical Poisson Normal model. However, it is crucial to note that for the task of outbreak detection, real-time decision-making may not heavily depend on efficient run time. The outbreak detection procedure is typically performed on a weekly basis, minimizing the practical importance of the computational advantages offered by the hierarchical Poisson Gamma model.

7.3 Harnessing the potential of MiBa

In the future, the utilization of MiBa-based surveillance has immense potential for disease surveillance. It has already demonstrated its value in various surveillance systems, such as The Healthcare-Associated Infections Database (HAIBA) for monitoring hospital-acquired infections (Condell et al. 2016; Gubbels et al. 2017) and the COVID-19 surveillance system (Schønning et al. 2021).

HAIBA, launched in 2015, was the first fully automated surveillance system built on MiBa data. It provides monitoring capabilities for hospital-acquired infections, enabling health-care professionals to track and manage these infections more effectively. Similarly, the COVID-19 surveillance system, developed during 2020 and 2021, utilizes MiBa data to monitor and respond to the COVID-19 pandemic.

In addition to these systems, MiBa-based surveillance includes monitoring respiratory infections (such as influenza, pertussis, Mycoplasma pneumonia, and respiratory syncytial virus) and sexually transmitted diseases like chlamydia. While these surveillance systems currently have partial automation in data processing, there are plans to fully automate them in the near future.

Expanding on the field of automated disease outbreak detection is crucial to fully harness the potential of MiBa. By developing advanced algorithms and methodologies, it becomes possible to automatically analyze MiBa data and detect disease outbreak signals in a timely manner. This can lead to early identification of outbreaks, allowing for prompt interventions and preventive measures.

This master thesis serves as a proof of concept and can serve as a catalyst for future research and development in automated disease outbreak detection. The focus should be on customizing methods to effectively utilize the valuable MiBa data, as this has the potential to greatly improve the efficiency of detecting and responding to infectious disease outbreaks. By fully leveraging the capabilities of MiBa-based surveillance and continuously improving automated detection methods, the overall disease surveillance efforts can be strengthened, leading to better protection of public health.

8 Conclusion

This master thesis aimed to address the potential of automating the early detection of disease outbreaks. The primary objective was to develop a novel method for outbreak detection and compare its performance with established methods from the literature, such as the method proposed by Farrington et al. (1996) and the improved method by Noufaily et al. (2013). By utilizing innovative approaches based on hierarchical models, important advancements were achieved in the field of outbreak detection.

The key finding of this study underscores the immense potential of the developed method in automating and detecting disease outbreaks. Through the analysis of various case studies, the novel approach successfully identified anomalous counts coinciding with outbreak investigations conducted by SSI. Also, it considerably reduced the number of alarms generated compared to the Farrington and Noufaily methods. Furthermore, in the simulation study, the developed method exhibited a lower FPR compared to the Farrington method, but a little higher FPR than the Noufaily method, while maintaining a high POD for varying sizes of simulated outbreaks.

These findings highlight the effectiveness and robustness of the novel method in accurately detecting disease outbreaks, offering a promising avenue for improving public health surveillance and response systems. By integrating this method into existing public health frameworks, such as MiBa, the preparedness and response capabilities can be enhanced, ultimately leading to improved outcomes in disease prevention and control.

Bibliography

- Madsen, Henrik and Poul Thyregod (2011). *Introduction to general and generalized linear models*. English. Texts in statistical science. CRC Press. ISBN: 9781420091557.
- Voldstedlund, M et al. (2014). "The Danish Microbiology Database (MiBa) 2010 to 2013". In: *Eurosurveillance* 19.1, 20667. DOI: <https://doi.org/10.2807/1560-7917.ES2014.19.1.20667>. URL: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2014.19.1.20667>.
- Noufaily, Angela et al. (Mar. 2013). "An Improved Algorithm for Outbreak Detection in Multiple Surveillance Systems". en. In: *Online Journal of Public Health Informatics* 32.7, pp. 1206–1222.
- Tulchinsky, Theodore H. (2018). "Chapter 5 - John Snow, Cholera, the Broad Street Pump; Waterborne Diseases Then and Now". In: *Case Studies in Public Health*. Ed. by Theodore H. Tulchinsky. Academic Press, pp. 77–99. ISBN: 978-0-12-804571-8. DOI: <https://doi.org/10.1016/B978-0-12-804571-8.00017-2>. URL: <https://www.sciencedirect.com/science/article/pii/B9780128045718000172>.
- Honardoost, Maryam, Azam Rajabpour, and Ladan Vakil (2018). "Molecular epidemiology; New but impressive". In: *Medical Journal of the Islamic Republic of Iran* 32.1. Cited by: 4; All Open Access, Gold Open Access, Green Open Access. DOI: 10.14196/mjiri.32.53. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065203514&doi=10.14196%2fmjiri.32.53&partnerID=40&md5=31ba202c2698dce55f73f5bfb6efb1a7>.
- Struelens, M J and S Brisse (2013). "From molecular to genomic epidemiology: transforming surveillance and control of infectious diseases". In: *Eurosurveillance* 18.4, 20386. DOI: <https://doi.org/10.2807/ese.18.04.20386-en>. URL: <https://www.eurosurveillance.org/content/10.2807/ese.18.04.20386-en>.
- Koeser, Claudio U. et al. (June 2012). "Rapid Whole-Genome Sequencing for Investigation of a Neonatal MRSA Outbreak". In: *NEW ENGLAND JOURNAL OF MEDICINE* 366.24, pp. 2267–2275. ISSN: 0028-4793.
- Baldry, Susannah (2010). "Attack of the clones". In: *Nature Reviews Microbiology* 8.6. Cited by: 9; All Open Access, Bronze Open Access, p. 390. DOI: 10.1038/nrmicro2369. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-77952857458&doi=10.1038%2fnrmicro2369&partnerID=40&md5=228603d5ee11cdcdf04bea3f5de7c9a8>.
- Buckeridge, David L. (2007). "Outbreak detection through automated surveillance: A review of the determinants of detection". In: *Journal of Biomedical Informatics* 40.4. Public Health Informatics, pp. 370–379. ISSN: 1532-0464. DOI: <https://doi.org/10.1016/j.jbi.2006.09.003>. URL: <https://www.sciencedirect.com/science/article/pii/S1532046406000980>.
- Unkel, Steffen et al. (2012). "Statistical methods for the prospective detection of infectious disease outbreaks: a review". In: *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 175.1, pp. 49–82. ISSN: 09641998, 1467985X. URL: <http://www.jstor.org/stable/41409708> (visited on 02/15/2023).
- Farrington, C. P. et al. (1996). "A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease". In: *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159.3, pp. 547–563. ISSN: 09641998, 1467985X. URL: <http://www.jstor.org/stable/2983331> (visited on 01/27/2023).
- Hulth, A. et al. (2010). "Practical usage of computer-supported outbreak detection in five european countries". In: *Eurosurveillance* 15.36. Cited by: 31, pp. 1–6. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-77957607700&partnerID=40&md5=ff916d06bcf38218b2388d0874eed9f8>.

- Salmon, Maëlle, Dirk Schumacher, and Michael Höhle (2016). "Monitoring Count Time Series in R: Aberration Detection in Public Health Surveillance". In: *Journal of Statistical Software* 70.10, pp. 1–35. DOI: 10.18637/jss.v070.i10. URL: <https://www.jstatsoft.org/index.php/jss/article/view/v070i10>.
- Bédubourg, Gabriel and Yann Le Strat (2017). "Evaluation and comparison of statistical methods for early temporal detection of outbreaks: A simulation-based study". In: *PLoS ONE* 12.7. Cited by: 27; All Open Access, Gold Open Access, Green Open Access. DOI: 10.1371/journal.pone.0181227. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85024904807&doi=10.1371%2fjournal.pone.0181227&partnerID=40&md5=0a18530ff75e9efd8643ca4a743097f7>.
- Bolker, Benjamin M. et al. (2009). "Generalized linear mixed models: a practical guide for ecology and evolution". In: *Trends in Ecology & Evolution* 24.3, pp. 127–135. ISSN: 0169-5347. DOI: <https://doi.org/10.1016/j.tree.2008.10.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0169534709000196>.
- Zuur, Alain F. et al. (2009). *Mixed Effects Models and Extensions in Ecology with R*. English. Statistics for Biology and Health. Springer New York, NY. ISBN: 9780387874586.
- Palmer Real, Jaume et al. (2022). "A data-driven framework for characterising building archetypes: A mixed effects modelling approach". English. In: *Energy* 254. ISSN: 0360-5442. DOI: 10.1016/j.energy.2022.124278.
- Real, J. Palmer et al. (2021). "Simulating heat load profiles in buildings using mixed effects models". English. In: vol. 2069. 1. IOP Publishing. DOI: 10.1088/1742-6596/2069/1/012138.
- Heisterkamp, Simon H., Arnold L. M. Dekkers, and Janneke C. M. Heijne (2006). "Automated detection of infectious disease outbreaks: hierarchical time series models". In: *Statistics in Medicine* 25.24, pp. 4179–4196. DOI: <https://doi.org/10.1002/sim.2674>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.2674>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.2674>.
- Höhle, Michael and Michaela Paul (2008). "Count data regression charts for the monitoring of surveillance time series". In: *Computational Statistics and Data Analysis* 52.9. Cited by: 56, pp. 4357–4368. DOI: 10.1016/j.csda.2008.02.015. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-42749087852&doi=10.1016%5C%2fj.csda.2008.02.015&partnerID=40&md5=8d18fdad2a28d99ecbda5441a8fbeae3>.
- Svendsen, Anna Tølbøll et al. (2023). "The incidence of laboratory-confirmed cases of enteric pathogens in Denmark 2018: a national observational study". In: *Infectious Diseases* 55.5. PMID: 36868794, pp. 340–350. DOI: 10.1080/23744235.2023.2183253. eprint: <https://doi.org/10.1080/23744235.2023.2183253>. URL: <https://doi.org/10.1080/23744235.2023.2183253>.
- Buss, Sarah N. et al. (2015). "Multicenter evaluation of the BioFire FilmArray gastrointestinal panel for etiologic diagnosis of infectious gastroenteritis". In: *Journal of Clinical Microbiology* 53.3. Cited by: 327; All Open Access, Bronze Open Access, Green Open Access, pp. 915–925. DOI: 10.1128/JCM.02674-14. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84923376418&doi=10.1128%2fJCM.02674-14&partnerID=40&md5=a31098bbc8437c1849360924f660dacd>.
- Knabl, L., I. Grutsch, and D. Orth-Höller (2016). "Comparison of the BD MAX® Enteric Bacterial Panel assay with conventional diagnostic procedures in diarrheal stool samples". In: *European Journal of Clinical Microbiology and Infectious Diseases* 35.1. Cited by: 22, pp. 131–136. DOI: 10.1007/s10096-015-2517-4. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84954361539&doi=10.1007%2fs10096-015-2517-4&partnerID=40&md5=2800622c88a129b3ed2c8dbf533d2a39>.

- Goulet, Véronique et al. (2012). "Incidence of listeriosis and related mortality among groups at risk of acquiring listeriosis". In: *Clinical infectious diseases* 54.5, pp. 652–660.
- Awofisayo, A. et al. (Jan. 2015). "Pregnancy-associated listeriosis in England and Wales". In: *EPIDEMIOLOGY AND INFECTION* 143.2, pp. 249–256. ISSN: 0950-2688. DOI: 10.1017/S0950268814000594.
- Wingstrand, Anne et al., eds. (2015). *Annual Report on Zoonoses in Denmark 2014*. English. Annual Report on Zoonoses in Denmark. DTU Food.
- Smith, B. et al. (2011). "Outbreak of listeriosis caused by infected beef meat from a meals-on-wheels delivery in Denmark 2009". In: *Clinical Microbiology and Infection* 17.1, pp. 50–52. ISSN: 1198-743X. DOI: <https://doi.org/10.1111/j.1469-0691.2010.03200.x>. URL: <https://www.sciencedirect.com/science/article/pii/S1198743X14609125>.
- Kvistholm Jensen, Anne et al. (Mar. 2016). "Whole-genome Sequencing Used to Investigate a Nationwide Outbreak of Listeriosis Caused by Ready-to-eat Delicatessen Meat, Denmark, 2014". In: *Clinical Infectious Diseases* 63.1, pp. 64–70. ISSN: 1058-4838. DOI: 10.1093/cid/ciw192. eprint: <https://academic.oup.com/cid/article-pdf/63/1/64/6711273/ciw192.pdf>. URL: <https://doi.org/10.1093/cid/ciw192>.
- Takeuchi-Storm, Nao et al. (2023). "Presence and Persistence of Listeria monocytogenes in the Danish Ready-to-Eat Food Production Environment". In: *Hygiene* 3.1, pp. 18–32. ISSN: 2673-947X. DOI: 10.3390/hygiene3010004. URL: <https://www.mdpi.com/2673-947X/3/1/4>.
- Schjorring, Susanne et al. (Dec. 2017). "Cross-border outbreak of listeriosis caused by cold-smoked salmon, revealed by integrated surveillance and whole genome sequencing (WGS), Denmark and France, 2015 to 2017". In: *EUROSURVEILLANCE* 22.50, pp. 8–12. ISSN: 1025-496X. DOI: 10.2807/1560-7917.ES.2017.22.50.17-00762.
- Gillesberg Lassen, S. et al. (2016). "Two listeria outbreaks caused by smoked fish consumption—using whole-genome sequencing for outbreak investigations". In: *Clinical Microbiology and Infection* 22.7, pp. 620–624. ISSN: 1198-743X. DOI: <https://doi.org/10.1016/j.cmi.2016.04.017>. URL: <https://www.sciencedirect.com/science/article/pii/S1198743X16301148>.
- Helwigh, Birgitte and Luise Müller, eds. (2018). *Annual Report on Zoonoses in Denmark 2017*. English. Annual Report on Zoonoses in Denmark. Technical University of Denmark.
- Helwigh, Birgitte, Channie Kahl Petersen, and Mia Torpdahl, eds. (2020). *Annual Report on Zoonoses in Denmark 2019*. English. Annual Report on Zoonoses in Denmark. Technical University of Denmark.
- Lassen, Brian et al., eds. (2023). *Annual Report on Zoonoses in Denmark 2022*. English. Annual Report on Zoonoses in Denmark. DTU Food.
- Lewis, H. C. et al. (2009). "Outbreaks of Shigella sonnei infections in Denmark and Australia linked to consumption of imported raw baby corn". In: *Epidemiology and Infection* 137.3, pp. 326–334. DOI: 10.1017/S0950268808001829.
- Muller, L. et al. (June 2009). "Imported fresh sugar peas as suspected source of an outbreak of Shigella sonnei in Denmark, April – May 2009". In: *Eurosurveillance* 14.24. ISSN: 1560-7917.
- Menge, Christian (2020). "The Role of Escherichia coli Shiga Toxins in STEC Colonization of Cattle". In: *Toxins* 12.9. ISSN: 2072-6651. DOI: 10.3390/toxins12090607. URL: <https://www.mdpi.com/2072-6651/12/9/607>.
- Soborg, B et al. (2013). "A verocytotoxin-producing E. coli outbreak with a surprisingly high risk of haemolytic uraemic syndrome, Denmark, September-October 2012". In: *Eurosurveillance* 18.2, 20350. DOI: <https://doi.org/10.2807/ese.18.02.20350-en>. URL: <https://www.eurosurveillance.org/content/10.2807/ese.18.02.20350-en>.

- Helwigh, Birgitte, Anna Irene Vedel Sørensen, and Steen Ethelberg (2013). *Annual Report on Zoonoses in Denmark 2012*. English. Annual Report on Zoonoses in Denmark. The National Food Institute, Technical University of Denmark.
- Helwigh, Birgitte, Channie Kahl Petersen, and Luise Müller, eds. (2019). *Annual Report on Zoonoses in Denmark 2018*. English. 2nd ed. Annual Report on Zoonoses in Denmark. Technical University of Denmark.
- Kuhn, K. G. et al. (2013). "A long-lasting outbreak of *Salmonella Typhimurium* U323 associated with several pork products, Denmark, 2010". English. In: *Epidemiology and Infection* 141.2, pp. 260–268. ISSN: 0950-2688. doi: 10.1017/S0950268812000702.
- Helwigh, Birgitte, Anne Louise Krogh, and Luise Müller (2011). *Annual Report on Zoonoses in Denmark 2010*. English. ISSN: 0909-3837. National Food Institute, Technical University of Denmark.
- Stroup, D. F. et al. (Mar. 1989). "Detection of aberrations in the occurrence of notifiable diseases surveillance data". In: *Statistics in medicine* 8.3, pp. 323–329. ISSN: 0277-6715. doi: 10.1002/sim.4780080312.
- Wolfinger, Russell D. and Xihong Lin (1997). "Two Taylor-series approximation methods for nonlinear mixed models". In: *Computational Statistics and Data Analysis* 25.4. Cited by: 87, pp. 465–490. doi: 10.1016/S0167-9473(97)00012-1. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0031222254&doi=10.1016%2fS0167-9473%2897%2900012-1&partnerID=40&md5=dbf7ca1e47d836029af4044ff9880463>.
- Lee, Y. and J. A. Nelder (1996). "Hierarchical Generalized Linear Models". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 58.4, pp. 619–678. ISSN: 00359246. URL: <http://www.jstor.org/stable/2346105> (visited on 05/16/2023).
- Fisher, R. A. and Edward John Russell (1922). "On the mathematical foundations of theoretical statistics". In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 222.594-604, pp. 309–368. doi: 10.1098/rsta.1922.0009. eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.1922.0009>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.1922.0009>.
- Kristensen, Kasper et al. (2016). "TMB: Automatic Differentiation and Laplace Approximation". In: *Journal of Statistical Software* 70.5, pp. 1–21. doi: 10.18637/jss.v070.i05.
- Griewank, Andreas and Andrea Walther (2008). *Evaluating Derivatives*. Second. Society for Industrial and Applied Mathematics. doi: 10.1137/1.9780898717761. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9780898717761>. URL: <https://epubs.siam.org/doi/abs/10.1137/1.9780898717761>.
- Fournier, David A. et al. (2012). "AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models". In: *Optimization Methods and Software* 27.2, pp. 233–249. doi: 10.1080/10556788.2011.597854. eprint: <https://doi.org/10.1080/10556788.2011.597854>. URL: <https://doi.org/10.1080/10556788.2011.597854>.
- Hintze, Jerry L. and Ray D. Nelson (1998). "Violin Plots: A Box Plot-Density Trace Synnergism". In: *The American Statistician* 52.2, pp. 181–184. doi: 10.1080/00031305.1998.10480559. eprint: <https://www.tandfonline.com/doi/pdf/10.1080/00031305.1998.10480559>. URL: <https://www.tandfonline.com/doi/abs/10.1080/00031305.1998.10480559>.
- Bjerregård, Mathias Blicher, Jan Kloppenborg Møller, and Henrik Madsen (2021). "An introduction to multivariate probabilistic forecast evaluation". In: *Energy and AI* 4, p. 100058. ISSN: 2666-5468. DOI: <https://doi.org/10.1016/j.egyai.2021.100058>. URL: <https://www.sciencedirect.com/science/article/pii/S2666546821000124>.
- Good, I. J. (1992). "Rational Decisions". In: *Breakthroughs in Statistics: Foundations and Basic Theory*. Ed. by Samuel Kotz and Norman L. Johnson. New York, NY: Springer

- New York, pp. 365–377. ISBN: 978-1-4612-0919-5. doi: 10.1007/978-1-4612-0919-5_24.
URL: https://doi.org/10.1007/978-1-4612-0919-5_24.
- Gneiting, Tilmann and Adrian E Raftery (2007). “Strictly Proper Scoring Rules, Prediction, and Estimation”. In: *Journal of the American Statistical Association* 102.477, pp. 359–378. doi: 10.1198/016214506000001437. eprint: <https://doi.org/10.1198/016214506000001437>.
URL: <https://doi.org/10.1198/016214506000001437>.
- Condell, O. et al. (2016). “Automated surveillance system for hospital-acquired urinary tract infections in Denmark”. In: *Journal of Hospital Infection* 93.3. Cited by: 7, pp. 290–296. doi: 10.1016/j.jhin.2016.04.001. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84965047492&doi=10.1016%2fj.jhin.2016.04.001&partnerID=40&md5=02773b9c98c0bb1aa96633de924d50fb>.
- Gubbels, Sophie et al. (2017). “National automated surveillance of hospital-acquired bacteremia in Denmark using a computer algorithm”. In: *Infection Control and Hospital Epidemiology* 38.5. Cited by: 16, pp. 559–566. doi: 10.1017/ice.2017.1. URL: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85017657147&doi=10.1017%2fice.2017.1&partnerID=40&md5=3a1da21170a0d88a4c14f9b932776d7d>.
- Schønning, Kristian et al. (2021). “Electronic reporting of diagnostic laboratory test results from all healthcare sectors is a cornerstone of national preparedness and control of COVID-19 in Denmark”. In: *APMIS* 129.7, pp. 438–451. doi: <https://doi.org/10.1111/apm.13140>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/apm.13140>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/apm.13140>.

A Some probability functions

This chapter serves as a reference, specifying notation, properties, and moments related to the distributions used in this master thesis. The presentation of these probability functions are greatly inspired by Madsen and Thyregod (2011). The densities and first two moment are briefly summarized in Table A.1.

Name	Support	Density	$E[Y]$	$V[Y]$
Poisson Pois(λ)	$0, 1, 2, \dots$ $\lambda \in \mathbb{R}_+$	$\frac{\lambda^y}{y!} \exp(-\lambda)$	λ	λ
Gamma G(α, β)	\mathbb{R}_+ $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+$	$\frac{1}{\Gamma(\alpha)\beta} \left(\frac{y}{\beta}\right)^{\alpha-1} \exp(-y/\beta)$	$\alpha\beta$	$\alpha\beta^2$
Neg. Bin. NB(r, p)	$0, 1, 2, \dots$ $r \in \mathbb{R}_+, p \in]0, 1]$	$\binom{r+y-1}{y} p^r (1-p)^y$	$\frac{r(1+p)}{p}$	$\frac{r(1-p)}{p^2}$
Normal N(μ, σ^2)	\mathbb{R} $\mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}_+$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$	μ	σ^2

Table A.1: Density, support, mean value, and variance for a number of distributions used in this master thesis.

A.1 The Poisson distribution model

A discrete random variable Y , that can assume the values $0, 1, 2, \dots$ is said to follow a Poisson distribution with the parameter λ , if the density function of Y is of the form

$$f(y) = \frac{\lambda^y}{y!} \exp(-\lambda) \quad (\text{A.1})$$

for $y = 0, 1, 2, \dots$, where $0 < \lambda$. The notation $Y \sim \text{Pois}(\lambda)$ is short for the above.

The mean and variance are

$$E[Y] = \lambda \quad V[Y] = \lambda \quad (\text{A.2})$$

A.2 The Gamma distribution model

A continuous random variable Y , which can take all real non-negative values, is said to follow the Gamma distribution with parameters α and β if the distribution of Y can be described by a density function of the form

$$f(y) = \frac{1}{\Gamma(\alpha)\beta} \left(\frac{y}{\beta}\right)^{\alpha-1} \exp(-y/\beta) \quad (\text{A.3})$$

for $y \in \mathbb{R}_+$, with $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}_+$. Short notation for the above is $Y \sim G(\alpha, \beta)$. The parameter α is called the *shape parameter*, and β is a *scale parameter* for the distribution.

The mean and variance are

$$E[Y] = \alpha\beta \quad V[Y] = \alpha\beta^2 \quad (\text{A.4})$$

A.3 The Negative Binomial distribution

A discrete random variable Z with support $0, 1, 2, \dots$, is said to be distributed according to the Negative Binomial distribution model if the probability mass function for Z is

$$f(z) = \binom{r+z-1}{z} p^r (1-p)^z \quad (\text{A.5})$$

for $z = 0, 1, 2, \dots$. The Negative Binomial distribution describes the “waiting time”, i.e., the number of “successes” before the r ’th failure in a sequence of independent Bernoulli trials where the probability of success in each trial is p . Short notation for this is $Z \sim \text{NB}(r, p)$.

The mean and variance are

$$\mathbb{E}[Y] = \frac{r(1+p)}{p} \quad \mathbb{V}[Y] = \frac{r(1-p)}{p^2} \quad (\text{A.6})$$

A.4 The Gaussian distribution model

A continuous random variable Y , which can take all real values, is said to follow the Gaussian distribution with parameters μ and σ^2 if the distribution of Y can be described by a density function of the form

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) \quad (\text{A.7})$$

for $y \in \mathbb{R}$, with $\mu \in \mathbb{R}$ and $\sigma^2 \in \mathbb{R}_+$. Short notation for the above is $Y \sim \mathcal{N}(\mu, \sigma^2)$. The parameter μ represents the mean, and σ^2 the variance of the distribution.

Hence, the mean and variance are

$$\mathbb{E}[Y] = \mu \quad \mathbb{V}[Y] = \sigma^2 \quad (\text{A.8})$$

B State-of-the-art detection algorithm

B.1 Controls

In the function `farringtonFlexible`, which is accessible from the R package `surveillance`, users can select either the original Farrington method or the improved method by Noufaily by specifying the appropriate `control` arguments. The choice of algorithm variant is determined by the contents of the `control` slot. In the example provided, `con.farrington` indicates the use of the original method, while `con.noufaily` represents the options for the improved method.

B.1.1 Case studies

Here, the specific `control` arguments used in the case studies is provided.

```
con.farrington <- list(  
  range = NULL, b = 3, w = 2,  
  reweight = TRUE, weightsThreshold = 1,  
  verbose = TRUE, glmWarnings = TRUE,  
  alpha = 0.05, trend = TRUE, pThresholdTrend = 0.05,  
  limit54 = c(0,4), powertrans = "2/3",  
  fitFun = "algo.farrington.fitGLM.flexible",  
  populationOffset = TRUE,  
  noPeriods = 1, pastWeeksNotIncluded = NULL,  
  thershodMethod = "delta"  
)  
  
con.noufaily <- list(  
  range = NULL, b = 3, w = 2,  
  reweight = TRUE, weightsThreshold = 2.58,  
  verbose = TRUE, glmWarnings = TRUE,  
  alpha = 0.05, trend = TRUE, pThresholdTrend = 1,  
  limit54 = c(0,4), powertrans = "2/3",  
  fitFun = "algo.farrington.fitGLM.flexible",  
  populationOffset = TRUE,  
  noPeriods = 10, pastWeeksNotIncluded = NULL,  
  thershodMethod = "nbPlugin"  
)
```

B.1.2 Simulation study

Here, the specific `control` arguments used in the simulation studies is provided.

```
con.farrington <- list(  
  range = NULL, b = 5, w = 3,  
  reweight = TRUE, weightsThreshold = 1,  
  verbose = TRUE, glmWarnings = TRUE,  
  alpha = 0.005, trend = TRUE, pThresholdTrend = 0.05,  
  limit54 = c(0,4), powertrans = "2/3",  
  fitFun = "algo.farrington.fitGLM.flexible",  
  populationOffset = TRUE,  
  noPeriods = 1, pastWeeksNotIncluded = 3,
```

```
    thersholdMethod = "delta"
)

con.noufailey <- list(
  range = NULL, b = 5, w = 3,
  reweight = TRUE, weightsThreshold = 2.58,
  verbose = TRUE, glmWarnings = TRUE,
  alpha = 0.005, trend = TRUE, pThresholdTrend = 0.05,
  limit54 = c(0,4), powertrans = "2/3",
  fitFun = "algo.farrington.fitGLM.flexible",
  populationOffset = TRUE,
  noPeriods = 10, pastWeeksNotIncluded = 26,
  thersholdMethod = "nbPlugin"
)
```

C Clusters reported to The Food- and waterborne Outbreak Database

In this chapter, all the clusters of *Listeriosis* (LIST), *Shigellosis* (SHIL), Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC), and *Salmonellosis* (SALM) reported to The Food- and waterborne Outbreak Database (FUD) is presented. These clusters are summarized in Table C.1.

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022.

Pathogen	No. of patients	Setting	Suspected source	FUD no.
Salmonella	43	Private party	Unknown	809
Shigella flexneri	17	Private party	Unknown	795
S. Agona	4	Canteen	Composite meal	810
S. Agona	NA	General outbreak	Unknown	859
S. Enteritidis	40	Private party	Unknown	873
S. Enteritidis	NA	Tourists from Greece	Unknown	861
S. Derby	NA	General outbreak	Pork	875
S. Derby	22	General outbreak	Unknown	874
S. Kottbus	5	General outbreak	Unknown	784
S. Typhimurium U312b	NA	Slaughterhouse	Pork	863
S. Typhimurium U288	39	Food producer	Pork	855
S. Typhimurium DT 135b	109	General outbreak	Unknown	854
S. Typhimurium DT 3b	61	General outbreak	Unknown	853
S. Typhimurium DT 120	53	Shop	Pork (ham)	852
S. Stanley	5	General outbreak	Fresh vegetables	860
S. Chester	9	General outbreak	Unknown	796
Listeria monocytogenes	3	Other	Other meat	878
Listeria monocytogenes	8	Restaurant/catering	Beef	887
S. Enteritidis	4	Other	Unknown	871
S. Enteritidis PT 13a	NA	General outbreak	Eggs	917
S. Enteritidis PT 13a	86	Swimming meet	Eggs	928
S. Enteritidis PT 6a	NA	Abroad	Unknown	906
S. Enteritidis PT 11	NA	Regionel outbreak	Unknown	892
S. Enteritidis PT 8	150	General outbreak	Eggs	891
S. Typhimurium	53	Regionel outbreak	Unknown	945
S. Typhimurium DT 17	8	General outbreak	Unknown	902
S. Typhimurium DT 135a	90	General outbreak	Unknown	854
S. Typhimurium U292a	228	General outbreak	Unknown	788
S. Typhimurium DT 3a	30	General outbreak	Unknown	883
Shigella sonnei	4	Private party	Molluscs/Shellfish	894
Shigella sonnei	NA	General outbreak	Fresh vegetables	888
VTEC O157	3	Private home	Unknown	997
Listeria monocytogenes	9	National	Fish	1035
S. Enteritidis	NA	Tourists in Egypt	Unknown	977

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022. (continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
S. Enteritidis	NA	Tourists in Spain	Unknown	1038
S. Typhimurium DT104	8	Regional	Unknown	967
S. Typhimurium U292	19	National	Unknown	1010
S. Typhimurium DT10	7	Restaurant, Bulgaria	Unknown	1027
S. Typhimurium DT41	9	Tourists in Egypt	Unknown	1044
S. Typhimurium U323	172	National	Pork/pork products	979
S. Typhimurium DT120/DT7	20	National	Pork/deer product	996
S. 4,5,12:i:- U311	9	National	Unknown	1045
S. 4,5,12:i:- DT120	13	National	Unknown	995
S. Infantis	87	Hotel	Composite meal	1039
S. Virchow	3	Private party	Chicken	994
S. Umbilo	4	National	Unknown	1000
VTEC	103	Canteen	Composite meal	1098
Salmonella spp.	4	Restaurant	Buff et meals	1074
S. Typhimurium DT104	NA	Shop	Pork products (imp/dk)	1116
S. Typhimurium DT120	22	National	Pork products (imp)	1067
S. Aberdeen	20	Private party	Buff et meals	1101
S. Strathcona	NA	National	Tomatoes (imp)	1112
Shigella flexneri	38	Private party	Buff et meals	1088
Shigella flexneri	32	Private party	Buff et meals	1087
S. Bareilly	11	Restaurant	Unknown	1215
S. Mikawasima	3	National	Unknown	1206
S. 4,5,12:i:-, MLVA0006	64	Canteens	Beef	1191
S. 4,5,12:i:-, MLVA0006	24	National	Beef	1192
S. 1,4,5,12:i:-, MLVA0126	9	Private party	Composite meal	1185
S. 4,5,12:i:-, MLVA0201	6	Restaurant	Buffet meal	1199
S. Saintpaul	3	Shop	Duck (imp)	1193
S. Typhimurium DT120, MLVA0007	15	Shop	Pork	1245
S. Typhimurium DT120, MLVA0995	7	Regional	Unknown	1186
S. Typhimurium DT120, MLVA0006	6	Regional	Unknown	1200
S. Poona	10	National	Unknown	1174
VTEC O157	4	Regional	Unknown	1159
VTEC O157	14	National	Beef	1210
S. Enteritidis	12	National	Unknown	1327
S. Mikawasima	11	National	Unknown	1314
S. Typhimurium DT120, MLVA0007b	22	Shop	Pork	1245
S. Typhimurium U312, MLVA0550	43	National	Pork	1254

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022.
(continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
S. Typhimurium, MLVA0008	7	Regional	Unknown	1305
S. Typhimurium, MLVA0642	6	Restaurant	Composite meal	1348
S. Typhimurium, MLVA0817	7	National	Unknown	1258
L. monocytogenes, MLST224d,e	41	National	Cold cuts	1373
L. monocytogenes, MLST391	8	National	Unknown	1376
L. monocytogenes, MLST399	6	Hospital	Composite meal	1384
L. monocytogenes, MLST6	6	National	Fish	1385
S. Agona	8	National	Unknown	1317
S. Enteritidis, MLVA0206c,g	5	National	Unknown	1327
S. Enteritidis, MLVA0017	4	National	Chicken (imp)	1391
S. Enteritidis, MLVA0019	18	National	Eggs	1379
S. Infantis	8	Restaurant	Chicken (imp)	1380
S. Infantis	5	Regional	Unknown	1370
Salmonella 1,4,5,12:i:-, MLVA0201	25	National	Pork	1372
S. Typhimurium, MLVA1788	12	National	Unknown	1410
Salmonella 1,4,12:i:-, MLVA0007	22	Shop	Pork	1368
Salmonella 1,4,5,12:i:-, MLVA1277	19	National	Minced beef	1374
Salmonella 1,4,5,12:i:-, MLVA0008	38	Sports event	Unknown	1378
Salmonella 1,4,5,12:i:-, MLVA0334	5	Private party	Pork	1446
Shigella sonnei	5	Canteen	Sugar snaps (imp)	1408
VTEC O103:H2, eae and ehxA, vtx1a	5	National	Unknown	1377
VTEC O157:H-, eae, vtx1a, vtx2a	4	Institution	Unknown	1392
VTEC O157:H7, eae, vtx1a, vtx2a	7	Restaurant	Beef	1409
L. monocytogenes	2	Unknown	Unknown	1452
MLST224a				
Salmonella 4,5,12:i:-, MLVA0479b	6	National	Unknown	1487
Salmonella Newport	6	Restaurant	Composite meal	1453
Salmonella Oranienburgc	14	National	Unknown	1469
VTEC O157:H7, eae, vtx1a, vtx2c	3	Restaurant	Kebab	1463

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022. (continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
L. monocytogenes	2	National	Fish	1376
MLST391				
L. monocytogenes	2	National	Composite meal	1384
MLST399				
L. monocytogenes MLST6	4	National	Fish	1385
L. monocytogenes ST6	3	Regional	Unknown	1574
L. monocytogenes ST4	7	National	Cold cuts of meat	1525
Salmonella Enteritidis, MLVA0051d	6	Restaurant	Unknown	1531
Salmonella O:4,12;H:i:-, MLVA1776	5	National	Unknown	1548
Salmonella O:4,5,12;:-:, MLVA0126	16	Regional	Pork meat	1521
Salmonella O:4,5,12;H:i:-, MLVA0338b	12	International	Dried spicy snack sausage	1504
Salmonella O:4,5,12;H:i:-, MLVA0478b	9	National	Unknown	1515
Salmonella O:4,5,12;H:i:-, MLVA0617b,e	5	National	Composite meal	1558
Salmonella Reading	4	National	Unknown	1527
Salmonella Szentes	3	National	Unknown	1514
Salmonella Typhimurium, MLVA0642	6	National	Unknown	1526
Salmonella Worthington	3	Regional	Unknown	1518
VTEC O103:H2, ST17	6	National	Unknown	1511
VTEC O121:H19, vtx2a, eae, ehxA, ST655	5	Farm	Unknown	1520
VTEC O157:H7, vtx1a, vtx2a, eae, ehxA, ST11	2	Regional	Unknown	1528
VTEC O157:H7, vtx1a, vtx2c, ST11	6	Regional	Beef/Kebab	1530
L. monocytogenes, ST1b	1	National	Unknown	1592
L. monocytogenes, ST1247	6	National	Unknown	1559
L. monocytogenes, ST55c	2	National	Unknown	1632
L. monocytogenes, ST6c	2	International	Sweet corn (imp)	1618
L. monocytogenes, ST8d	6	National	Smoked salmon (imp)	1597
Salmonella Agona, ST13	10	National	Unknown	1589
Salmonella Bovismorbificans, ST1499b	7	National	Unknown	1593
Salmonella Enteritidis, ST11	7	National	Unknown	1581
Salmonella Enteritidis, ST11	6	National	Unknown	1582
Salmonella Enteritidis, ST11	7	Private party	Composite meal	1585

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022.
(continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
Salmonella O:4,5,12,H:i:-, ST 34	10	Catering	Composite meal	1601
Salmonella O:4,5,12,H:i:-, ST 34c	8	National	Unknown	1645
Salmonella O:4,5,12;H:i:-, ST34	5	National	Unknown	1596
Salmonella O:4,12;H:i:-	13	National	Pork meat	1577
Salmonella O:4,5,12;H:i:-, MLVA0617b	16	National	Composite meal	1558
Salmonella O:4,5,12;H:i:-, MLVA0617, ST34	4	Regional	Unknown	1624
Salmonella Tennessee, ST3288c	8	Regional	Unknown	1600
Salmonella Typhimurium	6	Regional	Unknown	1580
Salmonella Typhimurium, MLVA1919	5	National	Unknown	1572
Salmonella Typhimurium, ST19	13	National	Dried pork sausage (imp)	1603
Salmonella Typhimurium, ST2212	3	International	Dried pork sausage (imp)	1615
Salmonella Worthington, ST592	4	National	Unknown	1599
STEC O103:H2, ST17	4	National	Unknown	1621
STEC O131:H5	15	Restaurant	Buffet meal	1584
STEC O157:H7, ST11	3	National	Unknown	1620
STEC O157:H7, ST11	2	Regional	Unknown	1622
Listeria monocytogenes, ST20a	4	Regional	Unknown	1691
Listeria monocytogenes, ST8	5	National	Unknown	1652
Salmonella O:4,[5],12:i:-, ST34	11	Private party	Pork meat	1681
Salmonella Enteritidis, ST11	10	National	Unknown	1699
Salmonella Kottbus, ST212	12	National	Unknown	1690
Salmonella Mikawasima, ST1815	9	National	Unknown	1689
Salmonella Newport, ST45	5	National	Unknown	1765
Salmonella O:4,[5],12:i:-, ST34	17	National	Pork meat	1710
Salmonella O:4,[5],12:i:-, ST5296	43	National	Raw pork sausage, pork meat	1713
Salmonella Typhimurium, ST19	4	National	Pork meat	1666
Salmonella Typhimurium, ST19	6	National	Unknown	1762

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022. (continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
Salmonella Typhimurium, ST36	8	National	Leafy greens, rocket	1675
STEC O111:H8, ST16	7	National	Unknown	1702
STEC O26:H11, ST21	39	National	Cured dried beef sausage	1707
Listeria monocytogenes, ST1e	3	Retail, delicatessen	Salads	1592
Salmonella Coeln, ST1995	26	National	Unknown	1790
Salmonella Derby, ST682	11	National	Pork meat	1787
Salmonella Enteritidis, ST11	8	National	Unknown	1845
Salmonella London, ST155	4	National	Unknown	1820
Salmonella Mikawasima, ST1815	3	International	Vegetables, lettuces	1828
Salmonella Muenchen, ST82	4	International	Unknown	1801
Salmonella 4,[5],12:i:-, ST34#79	14	National	Pork meat	1772
Salmonella 4,[5],12:i:-, ST34#107	5	National	Pork meat	1771
Salmonella 4,[5],12:i:-, ST34#34	57	International	Minced beef / beef meat	1728
STEC O157:H7, ST11	13	National	Unknown	1791
Listeria monocytogenes, ST7#7	4	National	Unknown	1914
Listeria monocytogenes, ST394#1	2	International	Hot-smoked trout	1910
Listeria monocytogenes, ST451#2	2	Regional	Hot-smoked fish products	1890
Salmonella Coeln, ST1955#5	6	Regional	Unknown	1909
Salmonella Dublin, ST10#22	7	National	Unknown	1908
Salmonella Kasenyi, ST4546#1	12	National	Unknown	1888
Salmonella Kottbus, ST1669#1	36	Restaurant	Unknown	1879
Salmonella Strathcona, ST2559#1	25	International	Unknown	1883
Salmonella Typhimurium, ST36#6	7	National	Unknown	1898
Salmonella Typhimurium, ST19#60	4	Private party	Unknown	1913
Salmonella 4,[5],12:i:-, ST34#25	6	National	Unknown	1901
Salmonella 4,[5],12:i:-, ST34#123	9	National	Unknown	1900
Salmonella 4,[5],12:i:-, ST34#127	9	National	Unknown	1899

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022. (continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
Shigella sonnei	44	National	Fresh mint (imp)	1893
STEC O55:H7 ST335#1	8	National	Unknown	1911
Listeria monocytogenes, ST398#1	3	National	Unknown	2005
Listeria monocytogenes, ST7#4	9	National	Unknown	1970
Listeria monocytogenes, ST7#6	5	National	Unknown	1966
Listeria monocytogenes, ST5#1	10	National	Unknown	1941
Listeria monocytogenes, ST11#1	11	National	Unknown	1939
Salmonella	8	Take-away	Chicken meat	2008
Bovismorbificans, ST1499#4				
Salmonella Braenderup, ST22#2	41	International	Sugar melons	1954
Salmonella Chester, ST1954#6	8	International	Unknown	1989
Salmonella Enteritidis, ST11#117	3	International	Eggs	2019
Salmonella Enteritidis, ST11#142	26	National	Eggs	2009
Salmonella Litchfield, ST214#1	7	International	Unknown	1991
Salmonella Mikawasima, ST1815#13	4	National	Unknown	1981
Salmonella Montevideo, ST2327#1	11	National	Unknown	1964
Salmonella 4,[5],12:i,-, ST34#146	11	National	Unknown	2039
Salmonella Typhimurium, ST36#9	9	National	Unknown	1992
Salmonella Typhimurium, ST36#7	52	International	Dietary supplement, psyllium	1917
Salmonella multiple serotypes	6	International	Tahini	1988
Shigella sonnei, ST152#4	9	Canteen	Buffet meals	2031
Shigella sonnei, ST152#3	3	National	Unknown	1948
Shigella sonnei, ST152#2	12	Canteen	Ready-to-eat salads	1940
STEC O12:H6, ST583#10	9	National	Unknown	2033
STEC O157:H7, ST11#19	12	National	Unknown	2032
STEC O103:H2, ST12#22	5	National	Unknown	1973
STEC O103:H2, ST17#23	7	National	Unknown	1972
STEC O26:H11, ST21#14	4	National	Unknown	1959
E. coli multiple types (EAEC, ETEC, STEC)	11	Take-away	Mixed food	2079

Table C.1: Food- and waterborne disease clusters of LIST, SHIL, STEC, and SALM reported in The Food- and waterborne Outbreak Database (FUD) (n=229), 2008-2022. (continued)

Pathogen	No. of patients	Setting	Suspected source	FUD no.
Listeria monocytogenes, ST8	9	National	Cold-cut meat	2074
Listeria monocytogenes, ST8	2	National	Unknown	2098
Listeria monocytogenes, ST37	9	Regional	Unknown	2080
Listeria monocytogenes, ST37	5	National	Unknown	2006
Listeria monocytogenes, ST7	10	National	Fish patties	2127
Listeria monocytogenes, ST1607	4	National	Unknown	1969
Salmonella Ball ST3502	3	International	Unknown	2118
Salmonella Enteritidis ST11	5	National	Unknown	2161
Salmonella Enteritidis ST11	24	National	Unknown	2084
Salmonella Jukestown ST5005	6	International	Unknown	2078
Salmonella Mikawasima ST185	9	National	Unknown	2087
Salmonella Strathcona ST2559	6	National	Unknown	2117
Salmonella 4,[5],12:i:-, ST34	11	National	Unknown	2116
Salmonella 4,[5],12:i:-, ST34	6	National	Minced beef	2086
Salmonella 4,[5],12:i:-, ST34	4	International	Chocolate	2067
Salmonella Typhimurium ST36	15	National	Unknown	2100
Salmonella Typhimurium ST36	5	Regional	Unknown	2088
STEC O145:H28 ST32 (stx2a)	11	National	Unknown	2083
STEC O26:H11 ST21 (stx2a) and Campylobacter jejuni ST19	5	Farm	Raw milk from cow	2076

D Alarms by state-of-the-art methods in the *Listeriosis* case study

In this chapter, the alarms generated by the Farrington et al. (1996) method and the subsequently improved Noufaily et al. (2013) method is summarized in Table D.1.

Table D.1: Longtable

Date	Age group	y_t	Farrington		Noufaily	
			Threshold	Alarm	Threshold	Alarm
2011-08-01	<65 years	2	5.194	FALSE	6.173	FALSE
	65+ years	8	8.021	FALSE	7.089	TRUE
2011-12-01	<65 years	3	3.940	FALSE	2.818	TRUE
	65+ years	5	7.927	FALSE	3.944	TRUE
2012-01-01	<65 years	2	3.407	FALSE	2.765	FALSE
	65+ years	6	4.345	TRUE	3.366	TRUE
2012-05-01	<65 years	1	3.594	FALSE	4.022	FALSE
	65+ years	4	2.444	TRUE	4.149	FALSE
2012-07-01	<65 years	0	4.424	FALSE	4.351	FALSE
	65+ years	5	2.715	TRUE	5.386	FALSE
2013-02-01	<65 years	3	5.565	FALSE	2.801	TRUE
	65+ years	2	4.349	FALSE	5.168	FALSE
2013-09-01	<65 years	1	2.749	FALSE	4.097	FALSE
	65+ years	7	6.321	TRUE	7.232	FALSE
2013-10-01	<65 years	0	4.384	FALSE	3.976	FALSE
	65+ years	6	4.551	TRUE	6.524	FALSE
2014-01-01	<65 years	0	3.920	FALSE	2.789	FALSE
	65+ years	8	5.530	TRUE	8.012	FALSE
2014-06-01	<65 years	2	2.751	FALSE	1.457	TRUE
	65+ years	4	5.615	FALSE	7.483	FALSE
2014-07-01	<65 years	6	3.093	TRUE	1.711	TRUE
	65+ years	14	10.223	TRUE	8.139	TRUE
2014-08-01	<65 years	6	3.093	TRUE	1.429	TRUE
	65+ years	14	11.690	TRUE	9.242	TRUE
2014-09-01	<65 years	2	3.061	FALSE	1.902	TRUE
	65+ years	9	6.705	TRUE	7.673	TRUE

Table D.1: (*continued*)

Date	Age group	y_t	Threshold	Alarm	Threshold	Alarm
2016-04-01	<65 years	3	2.474	TRUE	3.652	FALSE
	65+ years	1	5.990	FALSE	6.471	FALSE
2016-10-01	<65 years	0	3.012	FALSE	2.840	FALSE
	65+ years	6	9.634	FALSE	4.166	TRUE
2016-11-01	<65 years	0	2.609	FALSE	2.524	FALSE
	65+ years	4	4.978	FALSE	3.561	TRUE
2016-12-01	<65 years	1	2.420	FALSE	2.368	FALSE
	65+ years	4	6.653	FALSE	3.359	TRUE
2017-01-01	<65 years	0	2.228	FALSE	1.857	FALSE
	65+ years	8	6.005	TRUE	3.566	TRUE
2017-04-01	<65 years	1	2.440	FALSE	1.863	FALSE
	65+ years	4	2.789	TRUE	3.997	TRUE
2017-07-01	<65 years	3	4.541	FALSE	2.904	TRUE
	65+ years	4	3.736	TRUE	7.894	FALSE
2018-02-01	<65 years	3	2.618	TRUE	3.872	FALSE
	65+ years	1	6.040	FALSE	6.948	FALSE
2018-06-01	<65 years	0	3.569	FALSE	6.717	FALSE
	65+ years	6	5.935	TRUE	6.853	FALSE
2019-02-01	<65 years	0	2.622	FALSE	3.543	FALSE
	65+ years	7	5.962	TRUE	6.526	TRUE
2019-09-01	<65 years	2	3.446	FALSE	4.647	FALSE
	65+ years	7	6.686	TRUE	7.582	FALSE
2021-03-01	<65 years	2	2.432	FALSE	1.870	TRUE
	65+ years	2	5.097	FALSE	4.526	FALSE
2021-05-01	<65 years	5	1.905	TRUE	1.692	TRUE
	65+ years	1	5.209	FALSE	4.848	FALSE
2021-07-01	<65 years	0	2.802	FALSE	3.105	FALSE
	65+ years	8	7.258	TRUE	7.149	TRUE
2021-12-01	<65 years	1	2.595	FALSE	3.115	FALSE
	65+ years	5	4.979	TRUE	6.479	FALSE
2022-05-01	<65 years	6	2.714	TRUE	2.997	TRUE
	65+ years	9	5.562	TRUE	6.692	TRUE
2022-06-01	<65 years	3	2.957	TRUE	5.696	FALSE
	65+ years	11	6.700	TRUE	8.252	TRUE

Table D.1: (*continued*)

Date	Age group	y_t	Threshold	Alarm	Threshold	Alarm
2022-10-01	<65 years	5	3.864	TRUE	7.405	FALSE
	65+ years	7	8.298	FALSE	11.943	FALSE
2022-11-01	<65 years	3	3.481	FALSE	5.569	FALSE
	65+ years	11	7.213	TRUE	10.783	TRUE

E Figures and tables related to the case studies

In this chapter, a full overview over the figures and tables related to the case studies of *Listeriosis* (LIST), *Shigellosis* (SHIL), Shiga toxin (verotoxin)-producing *Escherichia coli* (STEC), and *Salmonellosis* (SALM) are included.

E.1 Listeriosis

Table E.1: LIST modeling results. The average logarithmic score, $\bar{S}(G, y)$, is computed for the constant fixed effects model. The parameters are estimated in a rolling window with width $k = 36$ and the estimates at time t_0 , i.e. the last time point, are presented in this table for both modeling frameworks. Confidence intervals for the parameter estimates are calculated using 95% profile likelihood confidence intervals.

	Parameter	Estimate	95% CI
Poisson Normal			
$\bar{S}(G, y) = 3.75$	$\hat{\beta}_{<65\text{years}}$	15.5977	[15.24, 15.91]
	$\hat{\beta}_{65+\text{years}}$	15.1128	[14.83, 15.35]
	$\log(\hat{\sigma})$	-0.8077	[-1.77, -0.35]
Poisson Gamma			
	$\hat{\beta}_{<65\text{years}}$	15.7017	[15.38, 16.01]
	$\hat{\beta}_{65+\text{years}}$	15.2058	[14.96, 15.45]
	$\log(\hat{\phi})$	-1.5751	[-3.57, -0.67]

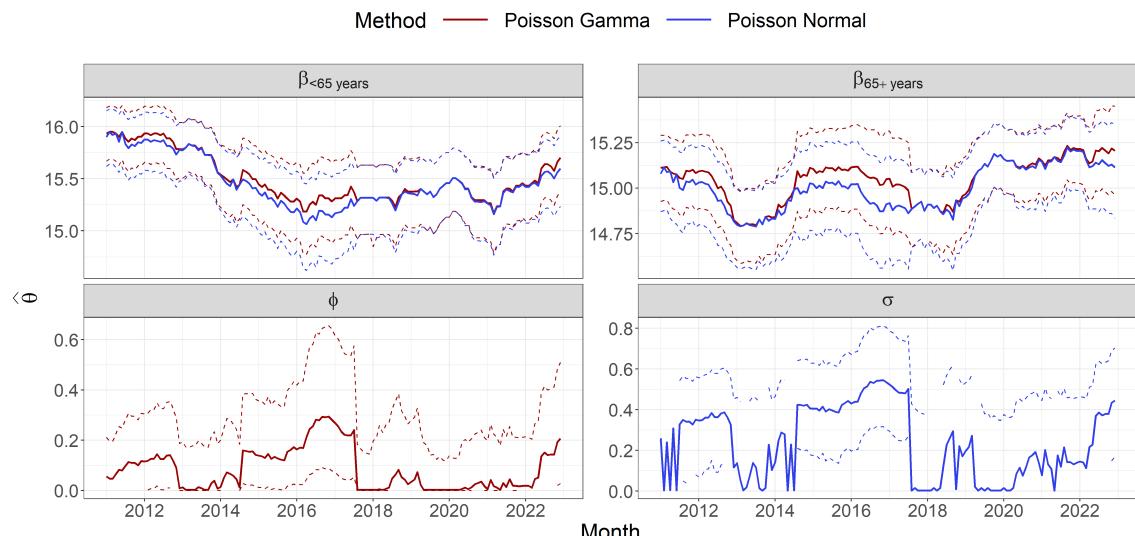


Figure E.1: Considering LIST in the time period between January 2011 and December 2022. Estimate (solid line) and 95% profile likelihood confidence interval (dashed line) of $\hat{\theta}$ for the constant fixed effects model estimated using the rolling window with a width of $k = 36$ months.

E.2 *Shigellosis*

Table E.2: SHIL modeling results. The average logarithmic score, $\bar{S}(G, y)$, is computed for all the fixed effects models. The parameters are estimated in a rolling window with width $k = 36$ and the estimates at time t_0 , i.e. the last time point, are presented in this table for both modeling frameworks. Confidence intervals for the parameter estimates are calculated using 95% profile likelihood confidence intervals.

	Parameter	Estimate	95% CI
Poisson Normal			
Constant			
$\bar{S}(G, y) = 5.3$	$\hat{\beta}_{<25\text{years}}$	14.5005	[14.11, 14.84]
	$\hat{\beta}_{25+\text{years}}$	16.5011	[16.2, 16.76]
	$\log(\hat{\sigma})$	-0.5421	[-0.98, -0.19]
Trend			
$\bar{S}(G, y) = 5.23$	$\hat{\beta}_{trend}$	0.0270	[0.01, 0.05]
	$\hat{\beta}_{<25\text{years}}$	15.0142	[14.55, 15.44]
	$\hat{\beta}_{25+\text{years}}$	17.0063	[16.59, 17.4]
	$\log(\hat{\sigma})$	-0.5975	[-1.06, -0.24]
Seasonality			
$\bar{S}(G, y) = 5.28$	$\hat{\beta}_{<25\text{years}}$	14.5056	[14.14, 14.83]
	$\hat{\beta}_{25+\text{years}}$	16.4979	[16.22, 16.73]
	$\hat{\beta}_{\sin}$	-0.3420	[-0.62, -0.07]
	$\hat{\beta}_{\cos}$	0.3249	[0.05, 0.6]
	$\log(\hat{\sigma})$	-0.7804	[-1.45, -0.35]
Combined			
$\bar{S}(G, y) = 5.25$	$\hat{\beta}_{trend}$	0.0260	[0.01, 0.05]
	$\hat{\beta}_{<25\text{years}}$	15.0083	[14.55, 15.43]
	$\hat{\beta}_{25+\text{years}}$	16.9888	[16.58, 17.37]
	$\hat{\beta}_{\sin}$	-0.1290	[-0.41, 0.15]
	$\hat{\beta}_{\cos}$	0.2826	[0.01, 0.56]
	$\log(\hat{\sigma})$	-0.7095	[-1.27, -0.32]
Poisson Gamma			
Constant			
$\bar{S}(G, y) = 5.3$	$\hat{\beta}_{<25\text{years}}$	14.6592	[14.31, 15]
	$\hat{\beta}_{25+\text{years}}$	16.6720	[16.41, 16.94]
	$\log(\hat{\phi})$	-1.0376	[-1.87, -0.39]
Trend			
$\bar{S}(G, y) = 5.24$	$\hat{\beta}_{trend}$	0.0239	[0.01, 0.04]
	$\hat{\beta}_{<25\text{years}}$	15.1054	[14.67, 15.53]
	$\hat{\beta}_{25+\text{years}}$	17.1086	[16.73, 17.5]
	$\log(\hat{\phi})$	-1.1757	[-2.08, -0.5]
Seasonality			
$\bar{S}(G, y) = 5.27$	$\hat{\beta}_{<25\text{years}}$	14.6073	[14.27, 14.93]

Table E.2: (continued)

Parameter	Estimate	95% CI
$\hat{\beta}_{25+years}$	16.5995	[16.36, 16.84]
$\hat{\beta}_{\sin}$	-0.3498	[-0.62, -0.08]
$\hat{\beta}_{\cos}$	0.3317	[0.05, 0.61]
$\log(\hat{\phi})$	-1.4570	[-2.68, -0.67]
Combined		
$\bar{S}(G, y) = 5.25$	$\hat{\beta}_{trend}$	0.0240 [0.01, 0.04]
	$\hat{\beta}_{<25years}$	15.0931 [14.67, 15.51]
	$\hat{\beta}_{25+years}$	17.0766 [16.7, 17.47]
	$\hat{\beta}_{\sin}$	-0.1271 [-0.4, 0.15]
	$\hat{\beta}_{\cos}$	0.2854 [0.01, 0.56]
	$\log(\hat{\phi})$	-1.3831 [-2.49, -0.64]

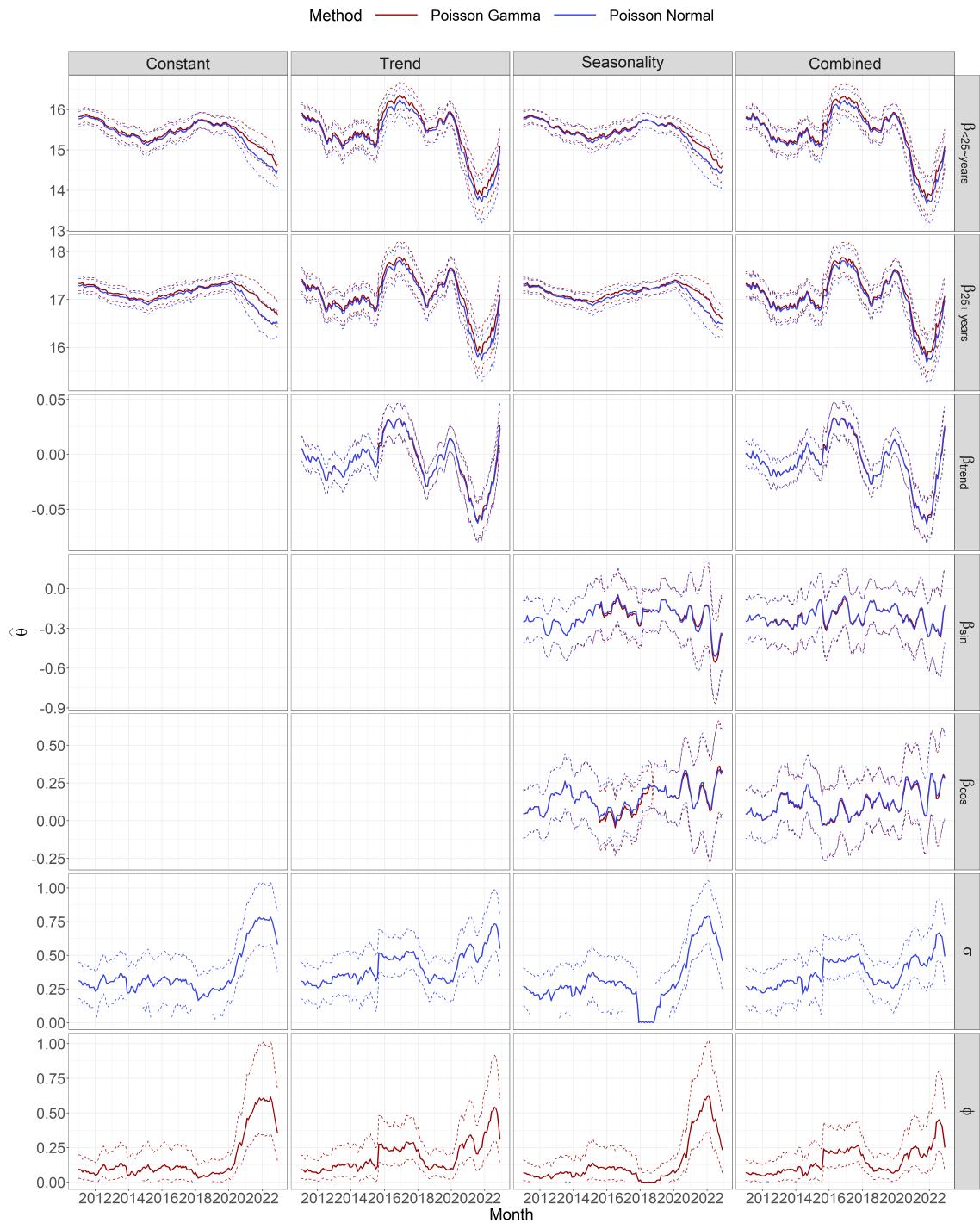


Figure E.2: Considering SHIL in the time period between January 2011 and December 2022. Estimate (solid line) and 95% profile likelihood confidence interval (dashed line) of $\hat{\theta}$ for the considered fixed effects models estimated using the rolling window with a width of $k = 36$ months.

E.3 Shiga toxin (verotoxin)-producing *Escherichia coli*

Table E.3: STEC modeling results. The average logarithmic score, $\bar{S}(G, y)$, is computed for all the fixed effects models. The parameters are estimated in a rolling window with width $k = 36$ and the estimates at time t_0 , i.e. the last time point, are presented in this table for both modeling frameworks. Confidence intervals for the parameter estimates are calculated using 95% profile likelihood confidence intervals.

	Parameter	Estimate	95% CI
Poisson Normal			
Constant			
$\bar{S}(G, y) = 17.89$	$\hat{\beta}_{<1year}$	11.7028	[11.41, 11.97]
	$\hat{\beta}_{1-4years}$	13.8781	[13.61, 14.13]
	$\hat{\beta}_{5-14years}$	14.6666	[14.42, 14.9]
	$\hat{\beta}_{15-24years}$	15.0082	[14.75, 15.26]
	$\hat{\beta}_{25-64years}$	17.3903	[17.15, 17.63]
	$\hat{\beta}_{65+years}$	16.2941	[16.03, 16.54]
	$\log(\hat{\sigma})$	-0.8608	[-1.12, -0.63]
Trend			
$\bar{S}(G, y) = 14.87$	$\hat{\beta}_{trend}$	0.0508	[0.04, 0.06]
	$\hat{\beta}_{<1year}$	12.6621	[12.38, 12.93]
	$\hat{\beta}_{1-4years}$	15.3254	[15.11, 15.54]
	$\hat{\beta}_{5-14years}$	15.7844	[15.55, 16.01]
	$\hat{\beta}_{15-24years}$	16.3031	[16.08, 16.52]
	$\hat{\beta}_{25-64years}$	18.8483	[18.65, 19.05]
	$\hat{\beta}_{65+years}$	17.7083	[17.51, 17.91]
	$\log(\hat{\sigma})$	-0.8741	[-1.05, -0.71]
Seasonality			
$\bar{S}(G, y) = 14.58$	$\hat{\beta}_{<1year}$	11.7751	[11.48, 12.06]
	$\hat{\beta}_{1-4years}$	14.4147	[14.18, 14.64]
	$\hat{\beta}_{5-14years}$	14.8871	[14.64, 15.13]
	$\hat{\beta}_{15-24years}$	15.4167	[15.19, 15.64]
	$\hat{\beta}_{25-64years}$	17.9597	[17.76, 18.16]
	$\hat{\beta}_{65+years}$	16.8271	[16.62, 17.03]
	$\hat{\beta}_{\sin}$	-0.5125	[-0.64, -0.38]
	$\hat{\beta}_{\cos}$	-0.0983	[-0.23, 0.03]
	$\log(\hat{\sigma})$	-0.5651	[-0.71, -0.42]
Combined			
$\bar{S}(G, y) = 14.18$	$\hat{\beta}_{trend}$	0.0442	[0.04, 0.05]
	$\hat{\beta}_{<1year}$	12.5572	[12.3, 12.81]
	$\hat{\beta}_{1-4years}$	15.2251	[15.03, 15.41]
	$\hat{\beta}_{5-14years}$	15.6763	[15.46, 15.88]
	$\hat{\beta}_{15-24years}$	16.1862	[15.99, 16.38]
	$\hat{\beta}_{25-64years}$	18.7354	[18.56, 18.9]
	$\hat{\beta}_{65+years}$	17.5977	[17.43, 17.77]
	$\hat{\beta}_{\sin}$	-0.3402	[-0.44, -0.25]
	$\hat{\beta}_{\cos}$	-0.0556	[-0.15, 0.04]

Table E.3: (continued)

	Parameter	Estimate	95% CI
	$\log(\hat{\sigma})$	-1.1158	[-1.33, -0.92]
Poisson Gamma			
Constant			
$\bar{S}(G, y) = 19.79$	$\hat{\beta}_{<1year}$	11.7954	[11.53, 12.04]
	$\hat{\beta}_{1-4years}$	13.8443	[13.61, 14.08]
	$\hat{\beta}_{5-14years}$	14.8027	[14.6, 15]
	$\hat{\beta}_{15-24years}$	15.1239	[14.91, 15.33]
	$\hat{\beta}_{25-64years}$	17.2872	[17.07, 17.51]
	$\hat{\beta}_{65+years}$	16.1486	[15.91, 16.39]
	$\log(\hat{\phi})$	-2.4552	[-3.89, -1.7]
Trend			
$\bar{S}(G, y) = 14.94$	$\hat{\beta}_{trend}$	0.0499	[0.04, 0.06]
	$\hat{\beta}_{<1year}$	12.7353	[12.46, 13]
	$\hat{\beta}_{1-4years}$	15.4253	[15.22, 15.64]
	$\hat{\beta}_{5-14years}$	15.8612	[15.64, 16.09]
	$\hat{\beta}_{15-24years}$	16.3741	[16.16, 16.59]
	$\hat{\beta}_{25-64years}$	18.9102	[18.72, 19.11]
	$\hat{\beta}_{65+years}$	17.7576	[17.56, 17.96]
	$\log(\hat{\phi})$	-1.7281	[-2.07, -1.41]
Seasonality			
$\bar{S}(G, y) = 14.63$	$\hat{\beta}_{<1year}$	11.9352	[11.66, 12.21]
	$\hat{\beta}_{1-4years}$	14.6130	[14.4, 14.83]
	$\hat{\beta}_{5-14years}$	15.0472	[14.82, 15.28]
	$\hat{\beta}_{15-24years}$	15.5434	[15.33, 15.77]
	$\hat{\beta}_{25-64years}$	18.0941	[17.9, 18.3]
	$\hat{\beta}_{65+years}$	16.9806	[16.79, 17.19]
	$\hat{\beta}_{\sin}$	-0.5067	[-0.63, -0.38]
	$\hat{\beta}_{\cos}$	-0.0750	[-0.21, 0.05]
	$\log(\hat{\phi})$	-1.1541	[-1.42, -0.89]
Combined			
$\bar{S}(G, y) = 14.15$	$\hat{\beta}_{trend}$	0.0438	[0.04, 0.05]
	$\hat{\beta}_{<1year}$	12.6075	[12.34, 12.85]
	$\hat{\beta}_{1-4years}$	15.2839	[15.09, 15.47]
	$\hat{\beta}_{5-14years}$	15.7232	[15.51, 15.92]
	$\hat{\beta}_{15-24years}$	16.2339	[16.03, 16.42]
	$\hat{\beta}_{25-64years}$	18.7678	[18.6, 18.94]
	$\hat{\beta}_{65+years}$	17.6451	[17.47, 17.82]
	$\hat{\beta}_{\sin}$	-0.3350	[-0.43, -0.24]
	$\hat{\beta}_{\cos}$	-0.0415	[-0.14, 0.05]
	$\log(\hat{\phi})$	-2.2059	[-2.63, -1.82]

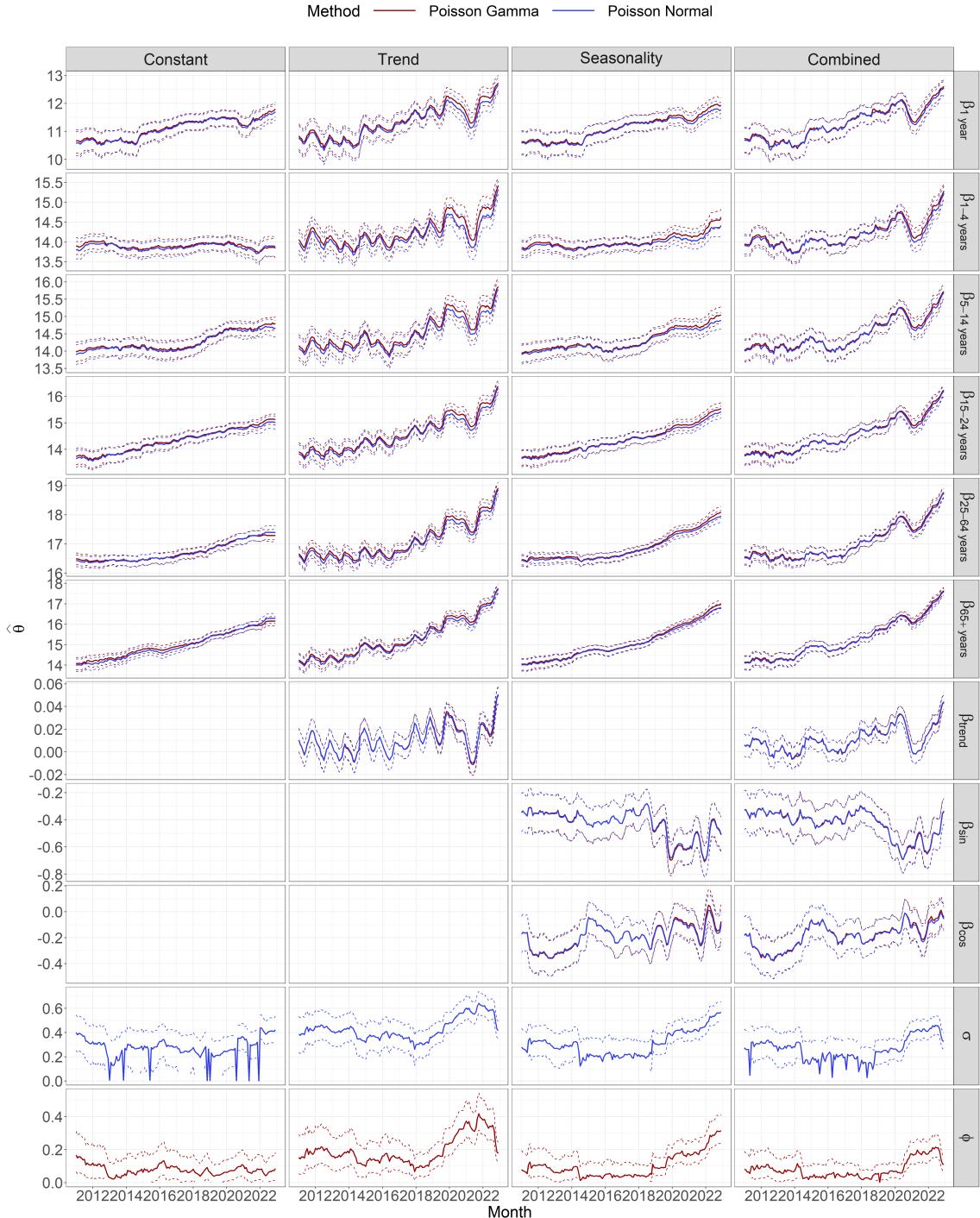


Figure E.3: Considering STEC in the time period between January 2011 and December 2022. Estimate (solid line) and 95% profile likelihood confidence interval (dashed line) of $\hat{\theta}$ for the considered fixed effects models estimated using the rolling window with a width of $k = 36$ months.

E.4 Salmonellosis

Table E.4: SALM modeling results. The average logarithmic score, $\bar{S}(G, y)$, is computed for all the fixed effects models. The parameters are estimated in a rolling window with width $k = 36$ and the estimates at time t_0 , i.e. the last time point, are presented in this table for both modeling frameworks. Confidence intervals for the parameter estimates are calculated using 95% profile likelihood confidence intervals.

	Parameter	Estimate	95% CI
Poisson Normal			
Constant			
$\bar{S}(G, y) = 21.43$	$\hat{\beta}_{<1year}$	11.7785	[11.52, 12.02]
	$\hat{\beta}_{1-4years}$	13.9431	[13.75, 14.13]
	$\hat{\beta}_{5-14years}$	14.9380	[14.74, 15.12]
	$\hat{\beta}_{15-24years}$	15.0651	[14.87, 15.26]
	$\hat{\beta}_{25-64years}$	17.9479	[17.81, 18.08]
	$\hat{\beta}_{65+years}$	16.6484	[16.51, 16.79]
	$\log(\hat{\sigma})$	-1.1211	[-1.35, -0.91]
Trend			
$\bar{S}(G, y) = 18.48$	$\hat{\beta}_{trend}$	0.0185	[0.01, 0.03]
	$\hat{\beta}_{<1year}$	12.1153	[11.83, 12.39]
	$\hat{\beta}_{1-4years}$	14.3075	[14.07, 14.54]
	$\hat{\beta}_{5-14years}$	15.3753	[15.14, 15.6]
	$\hat{\beta}_{15-24years}$	15.6231	[15.39, 15.84]
	$\hat{\beta}_{25-64years}$	18.3590	[18.16, 18.56]
	$\hat{\beta}_{65+years}$	16.9722	[16.77, 17.18]
	$\log(\hat{\sigma})$	-0.8790	[-1.05, -0.71]
Seasonality			
$\bar{S}(G, y) = 18.4$	$\hat{\beta}_{<1year}$	11.8322	[11.59, 12.06]
	$\hat{\beta}_{1-4years}$	13.9902	[13.81, 14.16]
	$\hat{\beta}_{5-14years}$	14.9578	[14.77, 15.13]
	$\hat{\beta}_{15-24years}$	15.0840	[14.89, 15.27]
	$\hat{\beta}_{25-64years}$	17.9905	[17.87, 18.11]
	$\hat{\beta}_{65+years}$	16.6379	[16.51, 16.76]
	$\hat{\beta}_{\sin}$	-0.2556	[-0.35, -0.16]
	$\hat{\beta}_{\cos}$	-0.0713	[-0.16, 0.02]
	$\log(\hat{\sigma})$	-1.3075	[-1.61, -1.06]
Combined			
$\bar{S}(G, y) = 17.89$	$\hat{\beta}_{trend}$	0.0128	[0.01, 0.02]
	$\hat{\beta}_{<1year}$	12.0119	[11.74, 12.28]
	$\hat{\beta}_{1-4years}$	14.2187	[13.99, 14.44]
	$\hat{\beta}_{5-14years}$	15.2753	[15.06, 15.49]
	$\hat{\beta}_{15-24years}$	15.5252	[15.31, 15.74]
	$\hat{\beta}_{25-64years}$	18.2585	[18.08, 18.44]
	$\hat{\beta}_{65+years}$	16.8735	[16.68, 17.06]
	$\hat{\beta}_{\sin}$	-0.2763	[-0.38, -0.17]
	$\hat{\beta}_{\cos}$	-0.1079	[-0.21, -0.01]

Table E.4: (continued)

	Parameter	Estimate	95% CI
	$\log(\hat{\sigma})$	-1.0284	[-1.23, -0.85]
Poisson Gamma			
Constant			
$\bar{S}(G, y) = 20.53$	$\hat{\beta}_{<1year}$	11.8293	[11.57, 12.08]
	$\hat{\beta}_{1-4years}$	14.0385	[13.84, 14.23]
	$\hat{\beta}_{5-14years}$	14.9949	[14.79, 15.19]
	$\hat{\beta}_{15-24years}$	15.2123	[15.02, 15.41]
	$\hat{\beta}_{25-64years}$	18.1409	[18, 18.29]
	$\hat{\beta}_{65+years}$	16.7187	[16.57, 16.87]
	$\log(\hat{\phi})$	0.0000	[NA, 0.03]
Trend			
$\bar{S}(G, y) = 18.6$	$\hat{\beta}_{trend}$	0.0186	[0.01, 0.03]
	$\hat{\beta}_{<1year}$	12.2004	[11.92, 12.47]
	$\hat{\beta}_{1-4years}$	14.3957	[14.17, 14.63]
	$\hat{\beta}_{5-14years}$	15.4791	[15.26, 15.7]
	$\hat{\beta}_{15-24years}$	15.7329	[15.52, 15.95]
	$\hat{\beta}_{25-64years}$	18.4455	[18.25, 18.64]
	$\hat{\beta}_{65+years}$	17.0377	[16.84, 17.24]
	$\log(\hat{\phi})$	-1.7579	[-2.1, -1.44]
Seasonality			
$\bar{S}(G, y) = 18.41$	$\hat{\beta}_{<1year}$	11.8688	[11.63, 12.09]
	$\hat{\beta}_{1-4years}$	14.0303	[13.85, 14.2]
	$\hat{\beta}_{5-14years}$	14.9963	[14.82, 15.17]
	$\hat{\beta}_{15-24years}$	15.1163	[14.93, 15.3]
	$\hat{\beta}_{25-64years}$	18.0284	[17.91, 18.15]
	$\hat{\beta}_{65+years}$	16.6702	[16.54, 16.8]
	$\hat{\beta}_{\sin}$	-0.2525	[-0.34, -0.16]
	$\hat{\beta}_{\cos}$	-0.0677	[-0.16, 0.02]
	$\log(\hat{\phi})$	-2.6121	[-3.2, -2.13]
Combined			
$\bar{S}(G, y) = 17.95$	$\hat{\beta}_{trend}$	0.0124	[0.01, 0.02]
	$\hat{\beta}_{<1year}$	12.0685	[11.8, 12.33]
	$\hat{\beta}_{1-4years}$	14.2816	[14.06, 14.5]
	$\hat{\beta}_{5-14years}$	15.3429	[15.13, 15.55]
	$\hat{\beta}_{15-24years}$	15.6014	[15.4, 15.8]
	$\hat{\beta}_{25-64years}$	18.3038	[18.13, 18.49]
	$\hat{\beta}_{65+years}$	16.9192	[16.74, 17.11]
	$\hat{\beta}_{\sin}$	-0.2706	[-0.37, -0.17]
	$\hat{\beta}_{\cos}$	-0.1049	[-0.2, -0.01]
	$\log(\hat{\phi})$	-2.0716	[-2.48, -1.7]

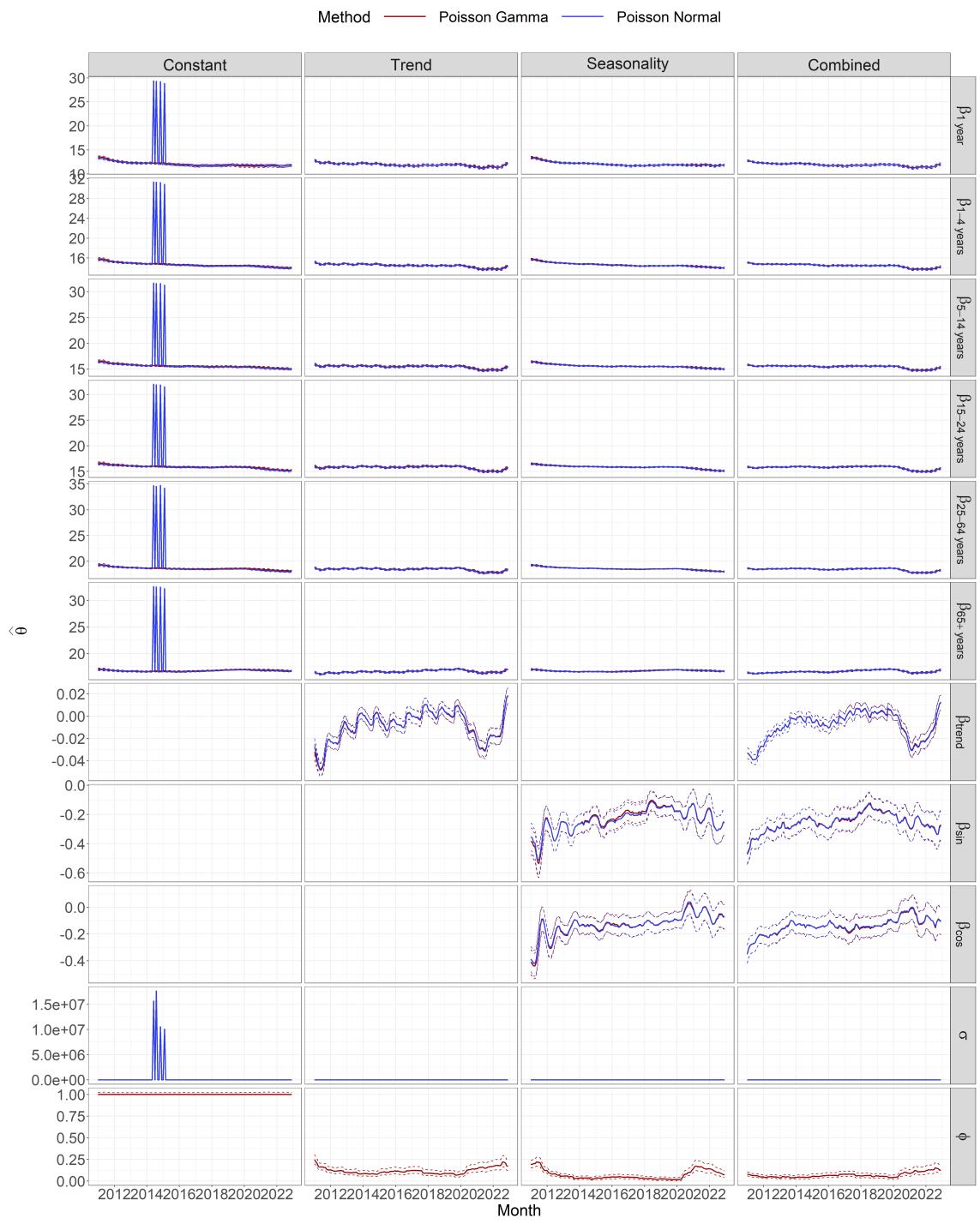


Figure E.4: Considering SALM in the time period between January 2011 and December 2022. Estimate (solid line) and 95% profile likelihood confidence interval (dashed line) of $\hat{\theta}$ for the considered fixed effects models estimated using the rolling window with a width of $k = 36$ months.

Technical
University of
Denmark

Richard Petersens Plads, Building 324
2800 Kgs. Lyngby
Tlf. 4525 1700

www.compute.dtu.dk