**DTU Compute**
Department of Applied Mathematics and Computer Science

# 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 thesis has been prepared over six months at the Section for Dynamical Systems, Department of Applied Mathematics and Computer Science, at the Technical University of Denmark, DTU, in collaboration with Epidemiologisk Forskning / Modelgruppen at Statens Serum Institut, SSI, in partial fulfilment for the degree Master of Science in Engineering, MSc Eng., Quantitative Biology and Disease Modelling.

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

Kasper Schou Telkamp - s170397

.....

*Signature*

.....

*Date*

## Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

## Acknowledgements

**Lasse Engbo Christiansen**, Senior Researcher, Statens Serum Institut
Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

**Jan Kloppenborg Møller**, Associate Professor, Technical University of Denmark
Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

# Contents

# 1 Introduction

This chapter is an overview of the methods that we propose to solve an **important problem**.

| instruction | expectation | rating |
|---|---|---|
| good | high | 25 |
| good | high | 0 |
| good | high | -16 |
| good | high | 5 |
| good | high | 11 |
| good | high | -6 |

# 2 Literature

Here is a review of existing methods.

Farrington et al. (1996) was improved in Noufaily et al. (2013)

and Salmon, Schumacher, and Höhle (2016)

Yang, Santillana, and Kou (2015) and Ning, Yang, and Kou (2019) uses Google Search terms to predict influenza-like ilness (ILI).

Vega, Jose Eugenio Lozano, et al. (2013) introduces the *Moving Epidemic Method* (MEM), which is used to compare the intensity level of ILI's in Vega, José E. Lozano, et al. (2015).

An analysis of ILI spread in Sweden showed that, rates in some large countries could vary considerably from one region to another Skog et al. (2014).

Different methods for monitoring influenza surveillance using only recent data is presented in Cowling et al. (2006)

Hutwagner et al. (1997) applied cumulative sums (CUSUM), to detect Salmonella outbreaks in US.

Costagliola et al. (1991) implement a simple regression model and calculates a 95% CI for a non-epidemic curve and use this threshold to alert when an epidemic begins.

Stern and Lightfoot (1999) discusses an automated algorithm to detect disease outbreaks with salmonella in Australia

# 3 Dataset

Test

| Date | ageGroup | landsdel | caseDef | cases | n |
|---|---|---|---|---|---|
| 2008-01-01 | <1 year | København by | AIDS | 0 | 10120 |
| 2009-01-01 | <1 year | København by | AIDS | 0 | 10288 |
| 2010-01-01 | <1 year | København by | AIDS | 0 | 10654 |
| 2011-01-01 | <1 year | København by | AIDS | 0 | 11199 |
| ... | ... | ... | ... | ... | ... |
| 2019-12-01 | 85+ years | Nordjylland | Shiga- og veratoxin producerende E. coli. | 0 | 14153 |
| 2020-12-01 | 85+ years | Nordjylland | Shiga- og veratoxin producerende E. coli. | 0 | 14613 |
| 2021-12-01 | 85+ years | Nordjylland | Shiga- og veratoxin producerende E. coli. | 0 | 14976 |
| 2022-12-01 | 85+ years | Nordjylland | Shiga- og veratoxin producerende E. coli. | 0 | 15203 |

# 4 Methods

### 4.0.1 Hierachical Poisson Normal model

**Fomulation**

$$Y_t^a | u_t^a \sim \mathsf{Pois}\big(w_t^a \lambda_a \exp(u_t^a)\big) \tag{4.1a}$$

$$u_t^a \sim \mathsf{N}(0, \sigma^2) \tag{4.1b}$$

# Bibliography


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).

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.

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.

Yang, Shihao, Mauricio Santillana, and S. C. Kou (2015). “Accurate estimation of influenza epidemics using Google search data via ARGO”. In: *Proceedings of the National Academy of Sciences* 112.47, pp. 14473–14478. DOI: 10.1073/pnas.1515373112. eprint: https://www.pnas.org/doi/pdf/10.1073/pnas.1515373112. URL: https://www.pnas.org/doi/abs/10.1073/pnas.1515373112.

Ning, Shaoyang, Shihao Yang, and S. C. Kou (Mar. 2019). “Accurate regional influenza epidemics tracking using Internet search data”. In: *Scientific Reports* 9.1. DOI: 10.1038/s41598-019-41559-6. URL: https://doi.org/10.1038/s41598-019-41559-6.

Vega, Tomás, Jose Eugenio Lozano, et al. (2013). “Influenza surveillance in Europe: establishing epidemic thresholds by the Moving Epidemic Method”. In: *Influenza and Other Respiratory Viruses* 7.4, pp. 546–558. DOI: https://doi.org/10.1111/j.1750-2659.2012.00422.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1750-2659.2012.00422.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1750-2659.2012.00422.x.

Vega, Tomás, José E. Lozano, et al. (2015). “Influenza surveillance in Europe: comparing intensity levels calculated using the moving epidemic method”. In: *Influenza and Other Respiratory Viruses* 9.5, pp. 234–246. DOI: https://doi.org/10.1111/irv.12330. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/irv.12330. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/irv.12330.

Skog, Lars et al. (July 2014). “Spatiotemporal characteristics of pandemic influenza”. In: *BMC Infectious Diseases* 14.1. DOI: 10.1186/1471-2334-14-378. URL: https://doi.org/10.1186/1471-2334-14-378.

Cowling, Benjamin J et al. (Aug. 2006). “Methods for monitoring influenza surveillance data”. In: *International Journal of Epidemiology* 35.5, pp. 1314–1321. ISSN: 0300-5771. DOI: 10.1093/ije/dyl162. eprint: https://academic.oup.com/ije/article-pdf/35/5/1314/14112661/dyl162.pdf. URL: https://doi.org/10.1093/ije/dyl162.

Hutwagner, L.C. et al. (1997). “Using Laboratory-Based Surveillance Data for Prevention: An Algorithm for Detecting Salmonella Outbreaks”. In: *Emerging Infectious Diseases* 3.3. Cited by: 131; All Open Access, Gold Open Access, Green Open Access, pp. 395–400. DOI: 10.3201/eid0303.970322. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0031181478&doi=10.3201%2feid0303.970322&partnerID=40&md5=0b5a2fea57368381526824f978f96ee3.

Costagliola, D et al. (Jan. 1991). “A routine tool for detection and assessment of epidemics of influenza-like syndromes in France.” In: *American Journal of Public Health* 81.1, pp. 97–99. DOI: 10.2105/ajph.81.1.97. URL: https://doi.org/10.2105/ajph.81.1.97.

Stern, L. and D. Lightfoot (1999). “Automated outbreak detection: A quantitative retrospective analysis”. In: *Epidemiology and Infection* 122.1. Cited by: 55; All Open Access, Bronze Open Access, Green Open Access, pp. 103–110. DOI: 10.1017/S0950268898001939. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-0033065916&doi=10.1017%2fS0950268898001939&partnerID=40&md5=cc79bbef077747d0cfe003ef5ee3b1c5.

# A Title

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Technical
University of
Denmark

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

www.compute.dtu.dk