

# How well can we predict weekly death counts?

Jonas Schöley

[jschoeley@health.sdu.dk](mailto:jschoeley@health.sdu.dk)

# The forecasting perspective on excess deaths

**The credibility of any COVID-19 pandemic excess deaths model**

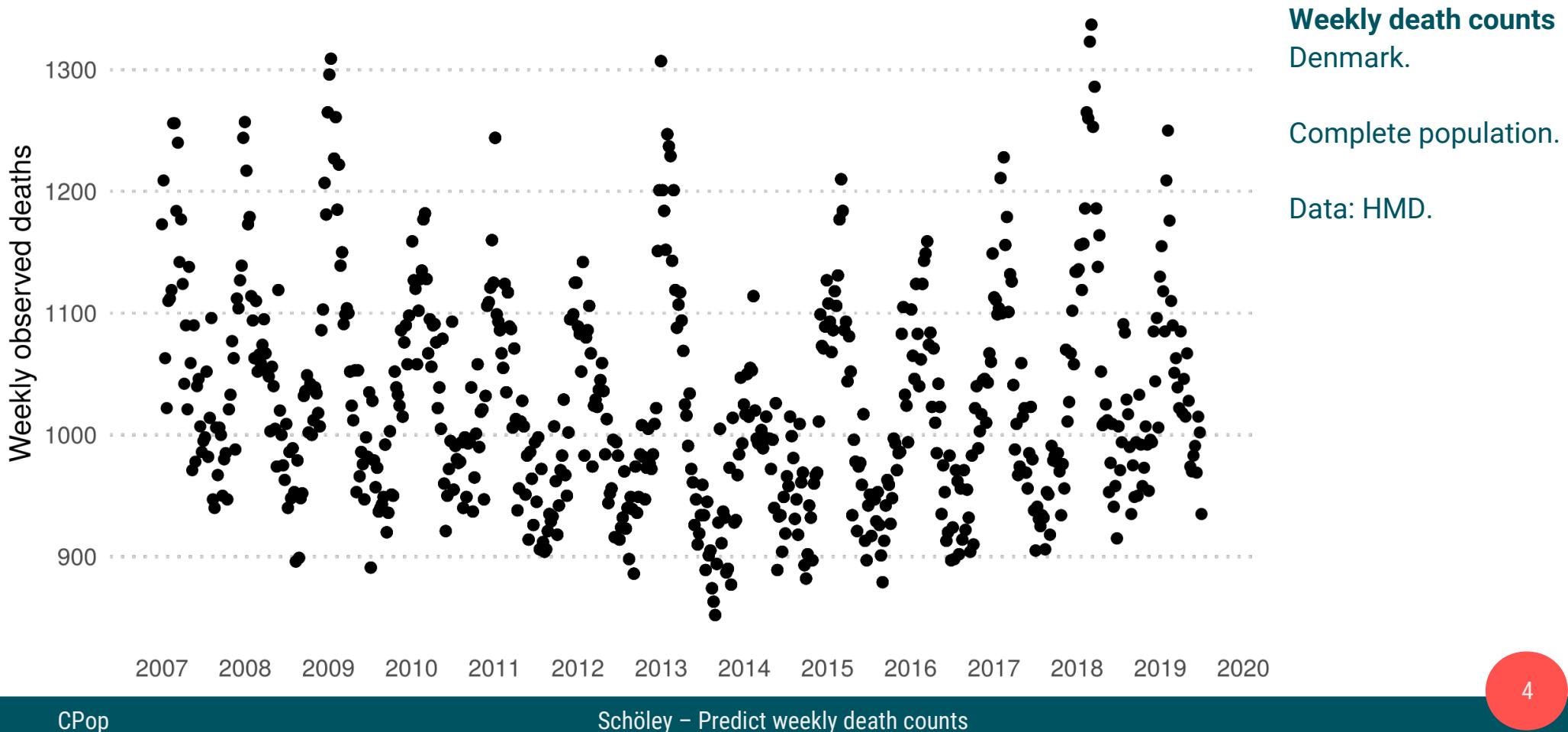
**derives from its ability to correctly forecast death counts in non-COVID situations**

# The challenge

Given 3 years of Danish weekly death counts by age and sex

How well can we predict weekly deaths over the next year?

# Training-test split of time-series

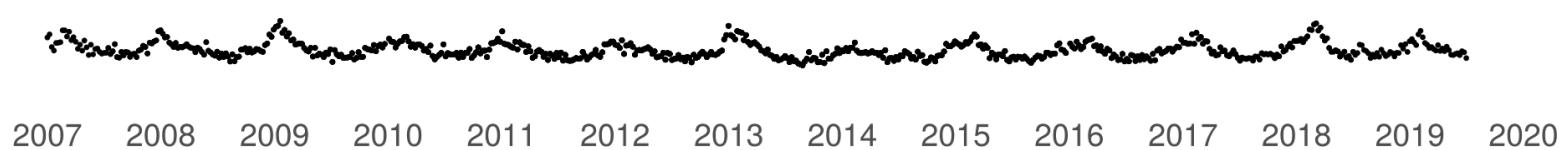


# Training-test split of time-series

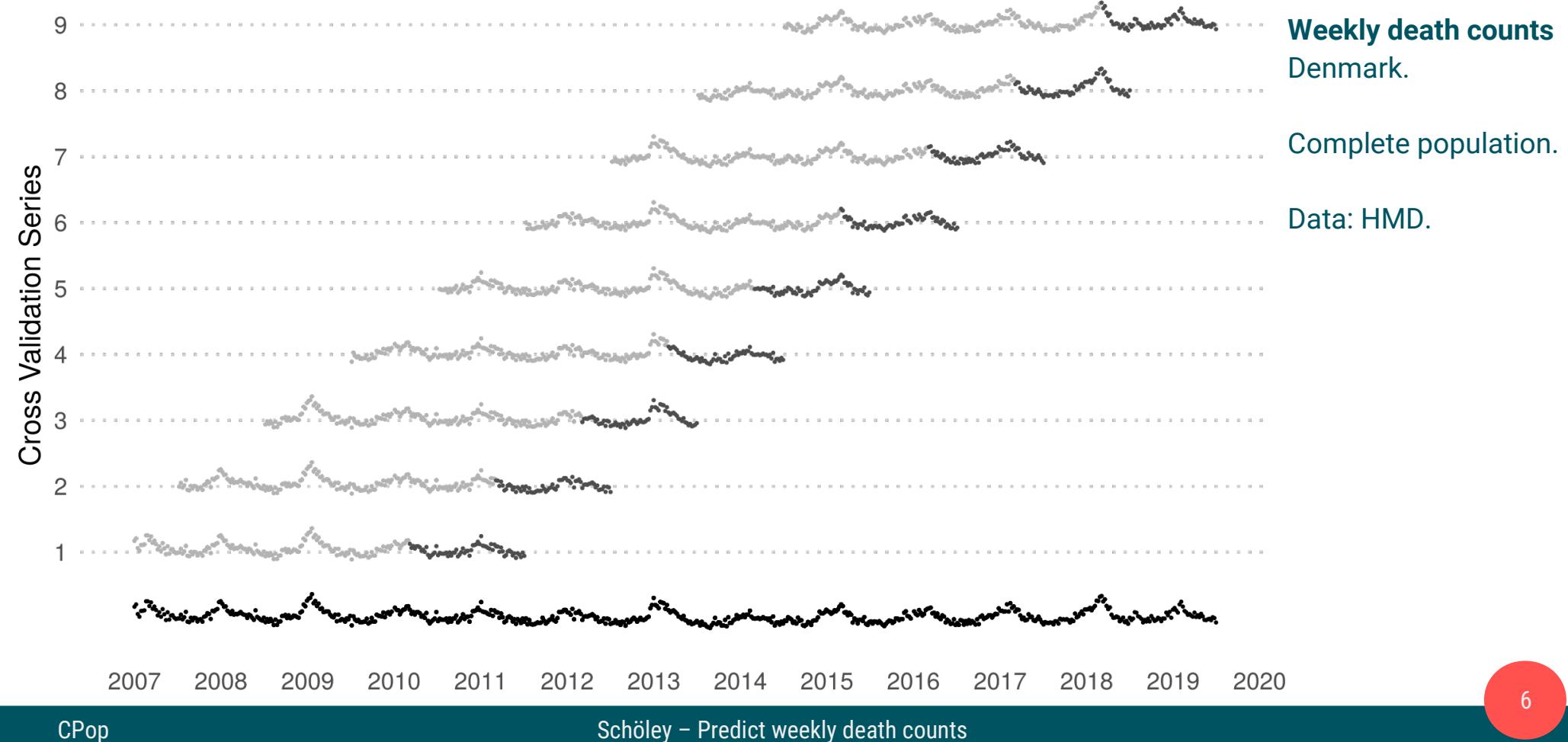
Weekly death counts  
Denmark.

Complete population.

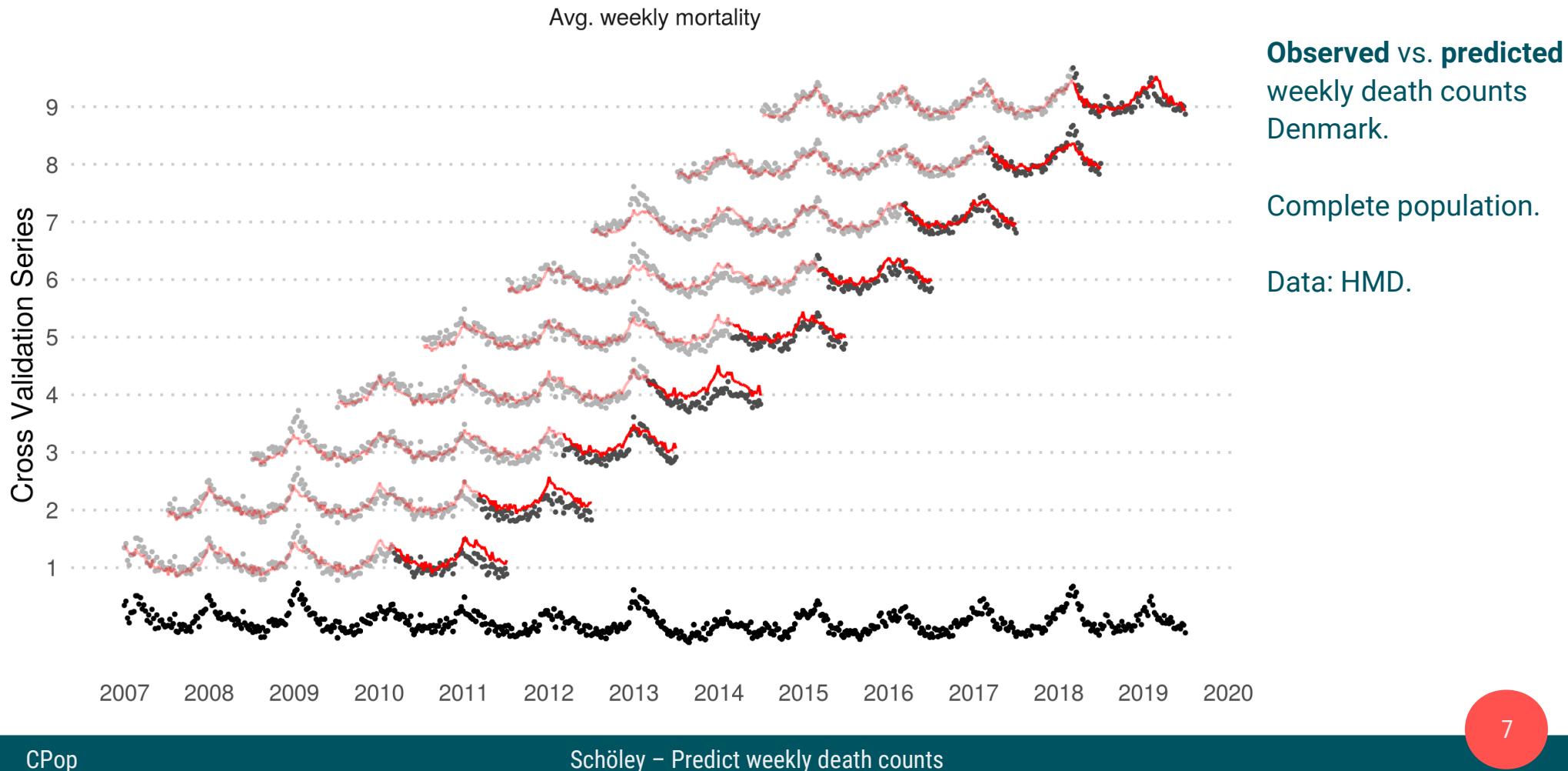
Data: HMD.



# Training-test split of time-series



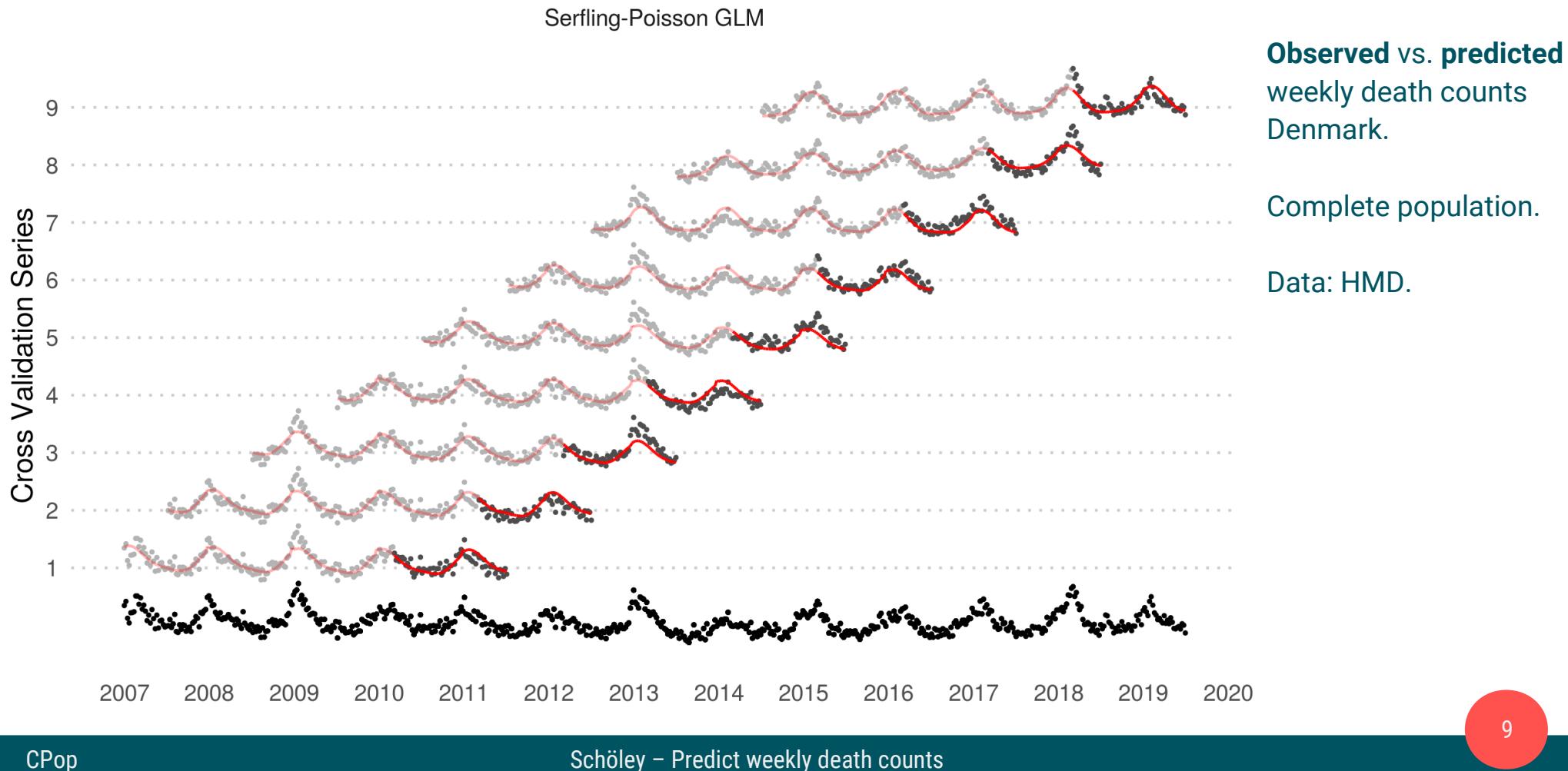
# The models



# The models

```
'Serfling-Poisson GLM', 'glm', list(
  formula = formula(
    observed_deaths ~
      # log linear long term trend
      weeks_since_origin*sex*age_group +
      # seasonality
      # full year period
      sin(2*pi*weeks_into_season_year/(365.25/7))*sex*age_group +
      cos(2*pi*weeks_into_season_year/(365.25/7))*sex*age_group +
      # half year period
      sin(2*pi*weeks_into_season_year/(365.25/2/7))*sex*age_group +
      cos(2*pi*weeks_into_season_year/(365.25/2/7))*sex*age_group +
      # adjustment for new years eve
      new_year*sex*age_group +
      # exposures
      offset(log(exposure)))
  ),
  family = quasipoisson(link = 'log')
)
```

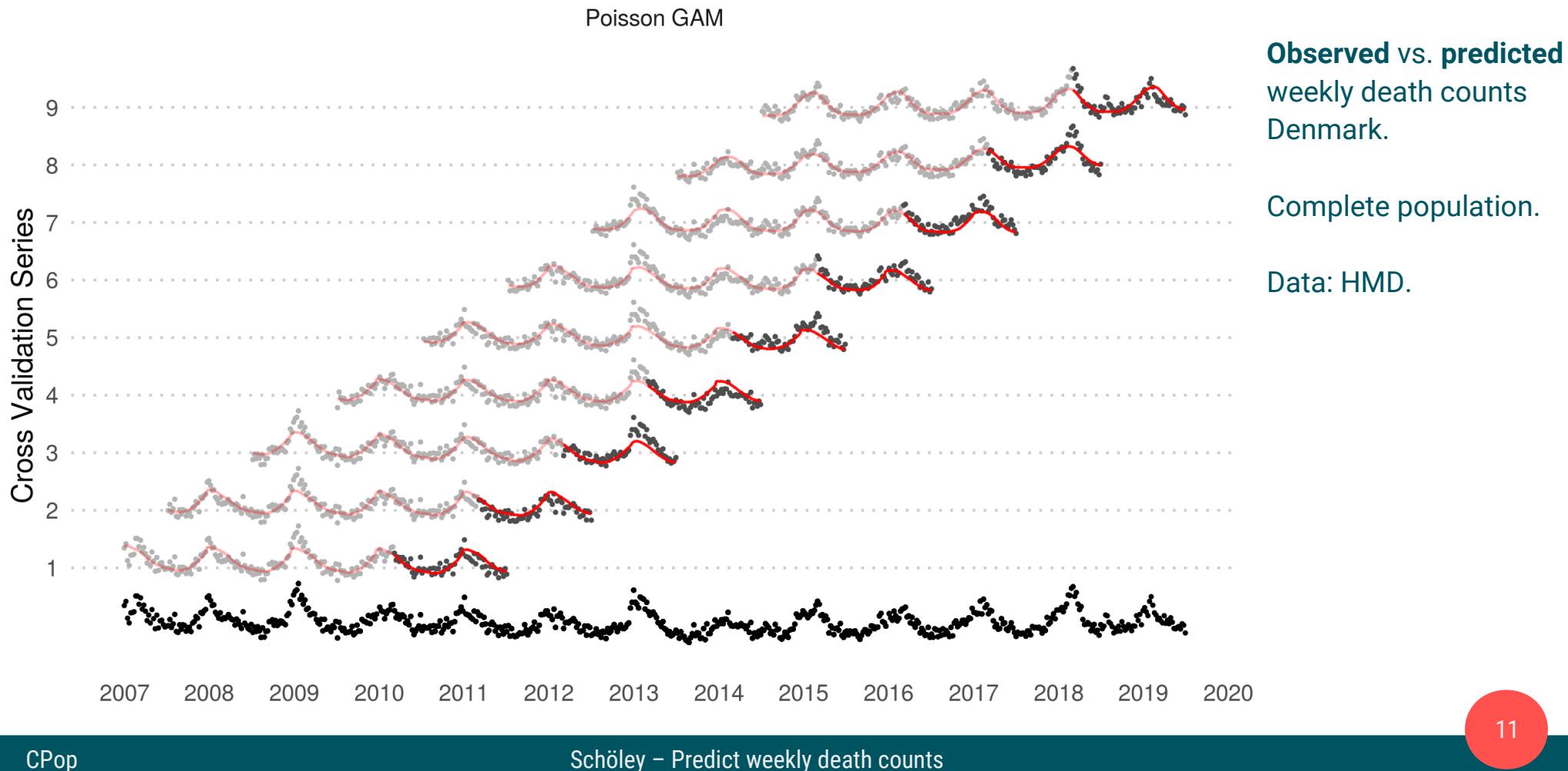
# The models



# The models

```
'Poisson GAM', 'gam', list(  
  formula = formula(  
    observed_deaths ~  
      1 + sex + age_group +  
      # log linear long term trend  
      weeks_since_origin*sex_age_interaction +  
      # penalized cyclic spline for seasonality  
      s(weeks_into_season_year, bs = 'cp', k = 52, by = sex_age_interaction) +  
      # adjustment for new years eve  
      new_year*sex_age_interaction +  
      # exposures  
      offset(log(exposure)))  
,  
  family = quasipoisson(link = 'log')  
)
```

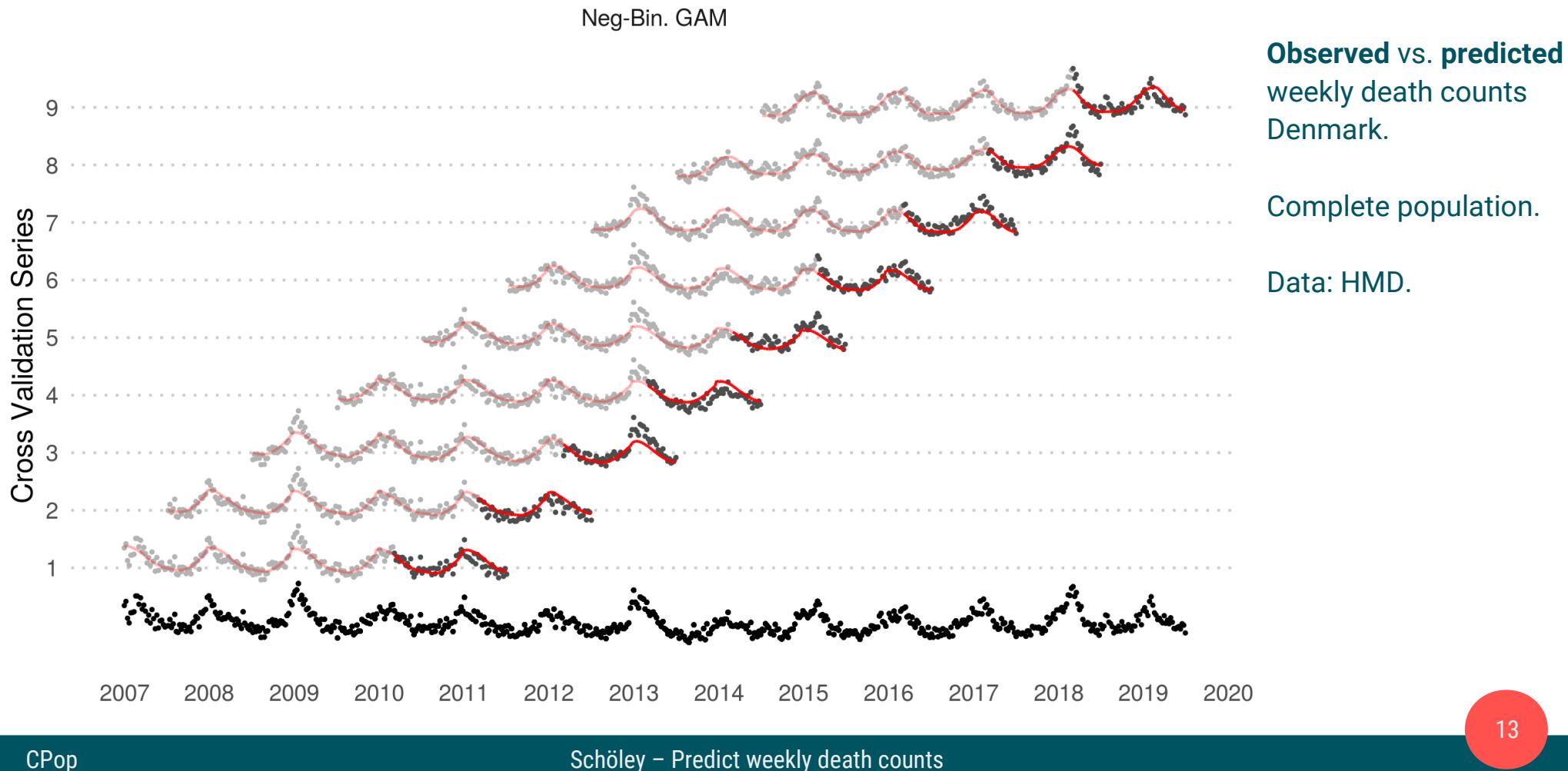
# The models



# The models

```
'Neg-Bin. GAM', 'gam', list(
  formula = formula(
    observed_deaths ~
      1 + sex + age_group +
      # log linear long term trend
      weeks_since_origin*sex_age_interaction +
      # penalized cyclic spline for seasonality
      s(weeks_into_season_year, bs = 'cp', k = 52, by = sex_age_interaction) +
      # adjustment for new years eve
      new_year*sex_age_interaction +
      # exposures
      offset(log(exposure)))
  ),
  family = nb(link = 'log')
)
```

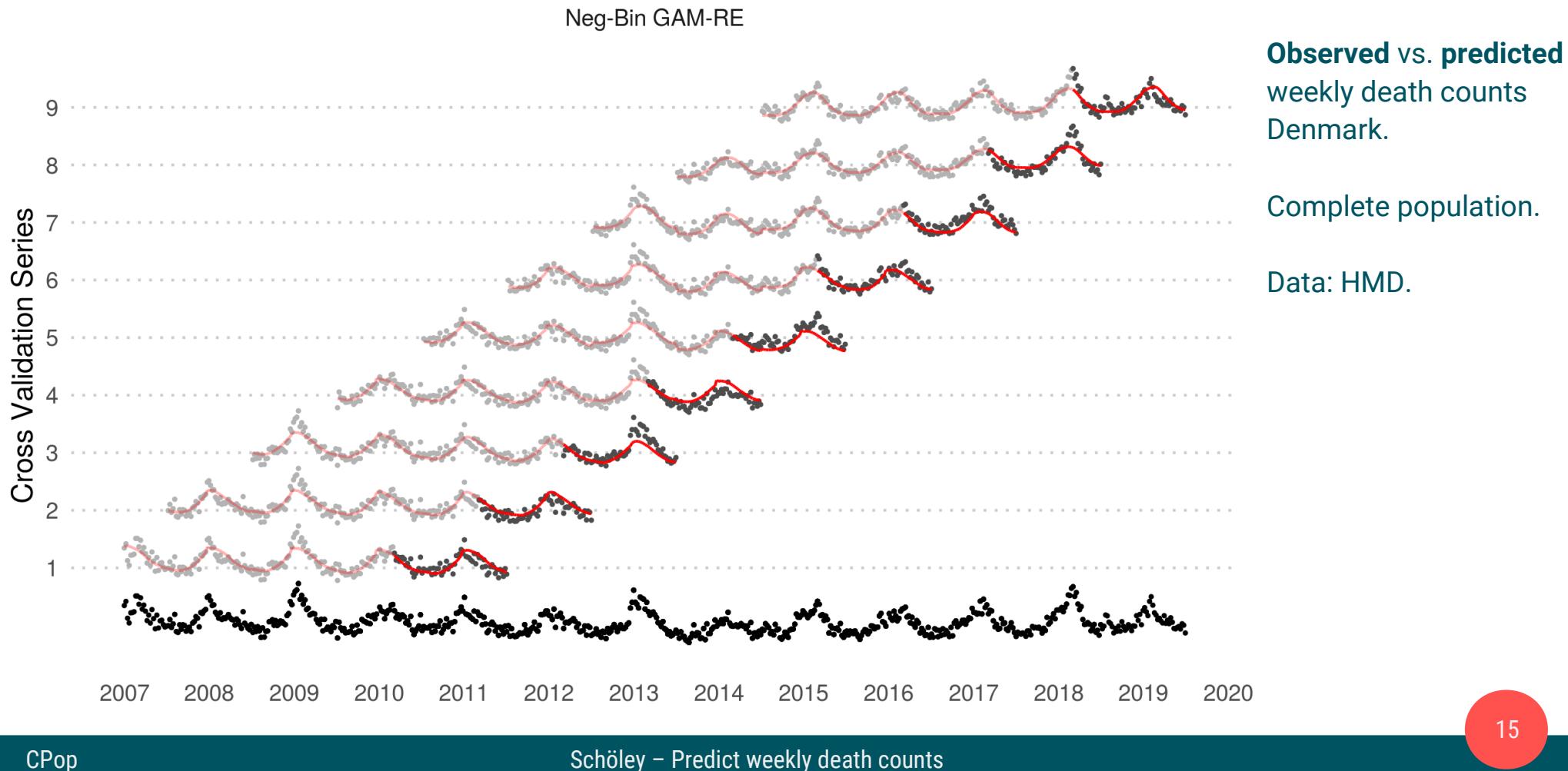
# The models



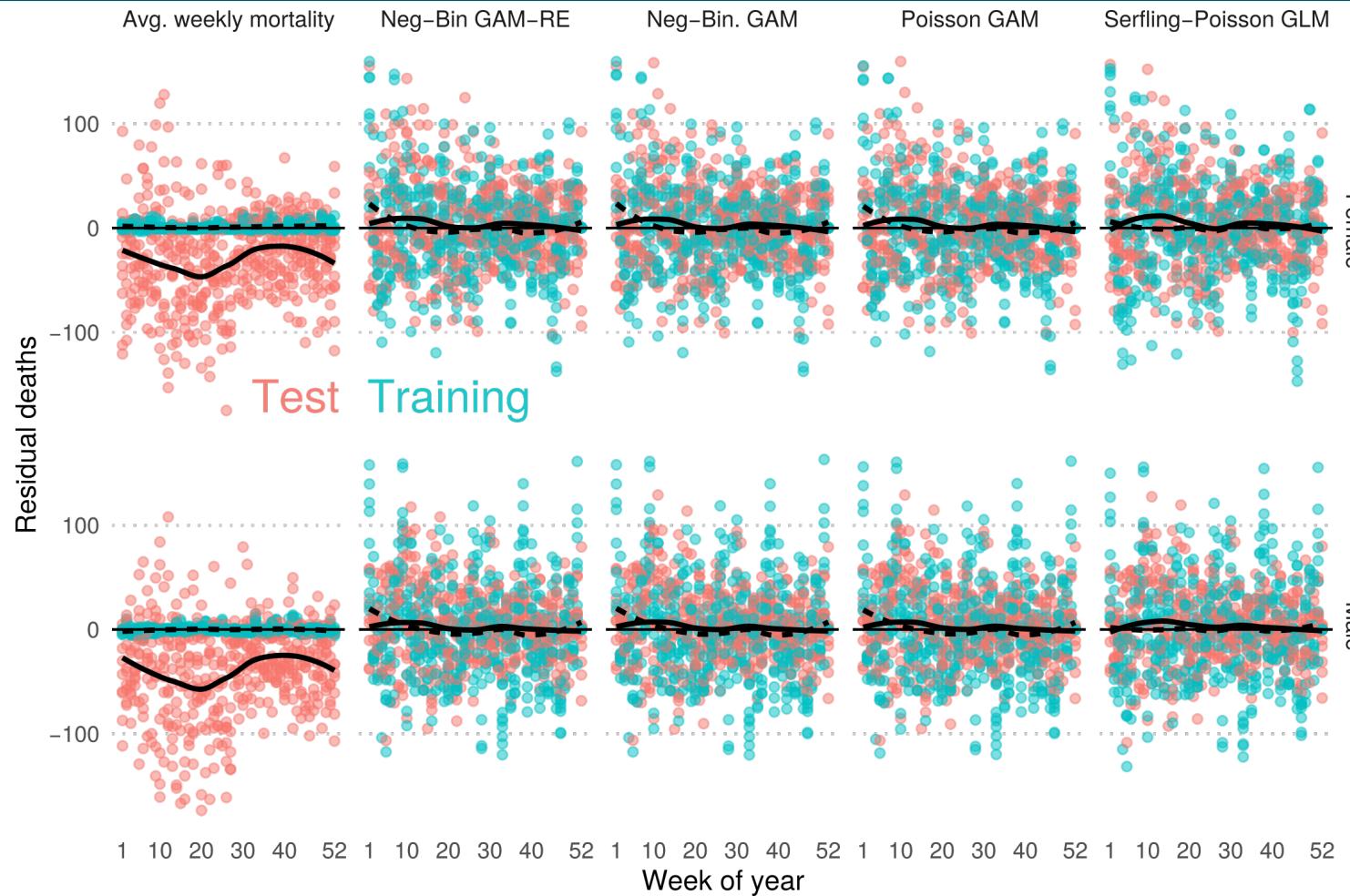
# The models

```
'Neg-Bin GAM-RE', 'gam', list(  
  formula = formula(  
    observed_deaths ~  
      1 + sex + age_group +  
      # log linear long term trend  
      weeks_since_origin*sex_age_interaction +  
      # penalized cyclic spline for seasonality  
      s(weeks_into_season_year, bs = 'cp', k = 52, by = sex_age_interaction) +  
      # adjustment for new years eve  
      new_year*sex_age_interaction +  
      # season year random effect adjustment  
      s(season_year_sex_age_interaction, bs = 're') +  
      # exposures  
      offset(log(exposure)))  
,  
  family = nb(link = 'log')  
)
```

# The models

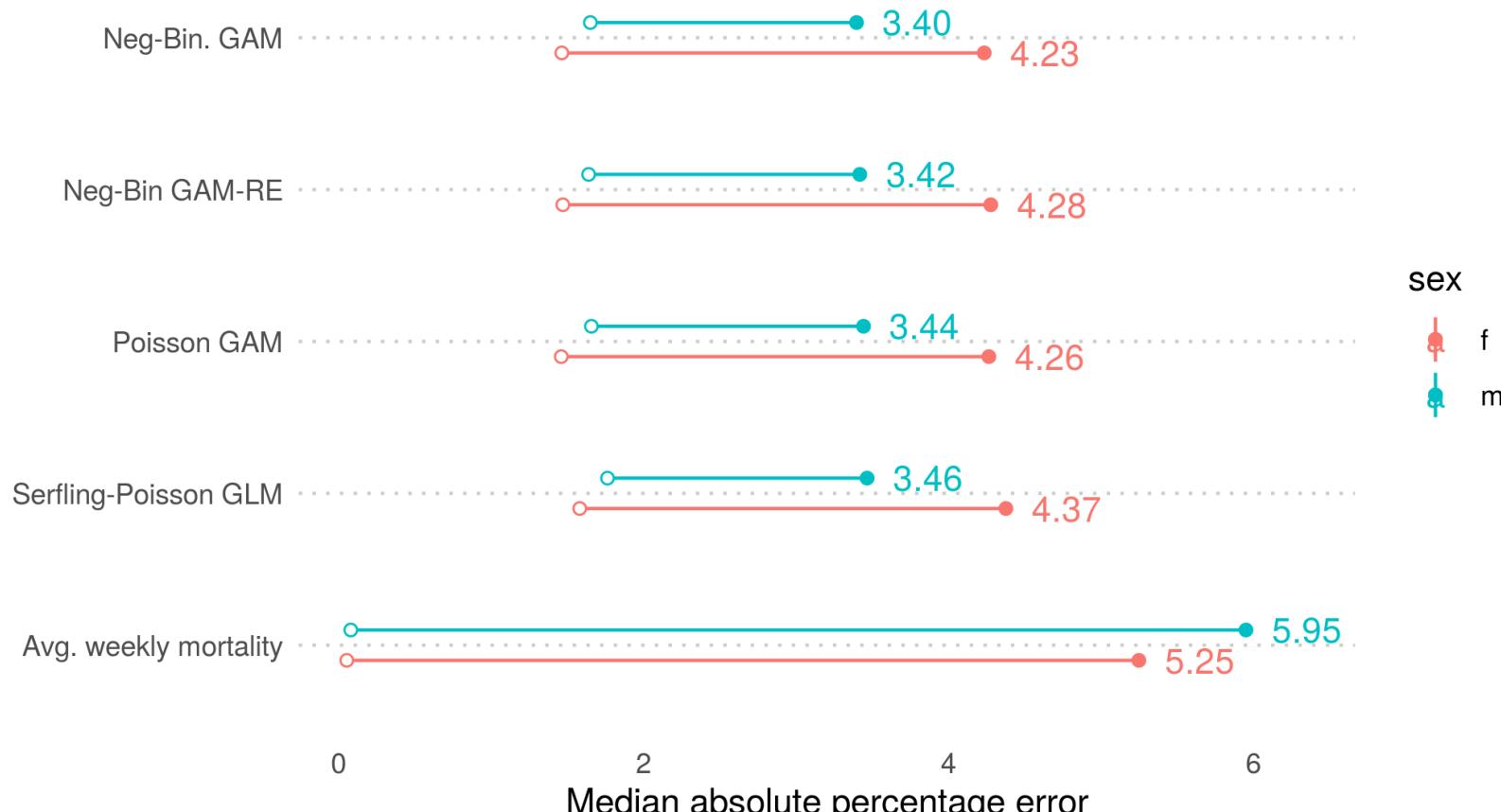


# Absolute weekly prediction error over test sets



Observed minus  
predicted weekly death  
counts Denmark.

# Median absolute prediction error

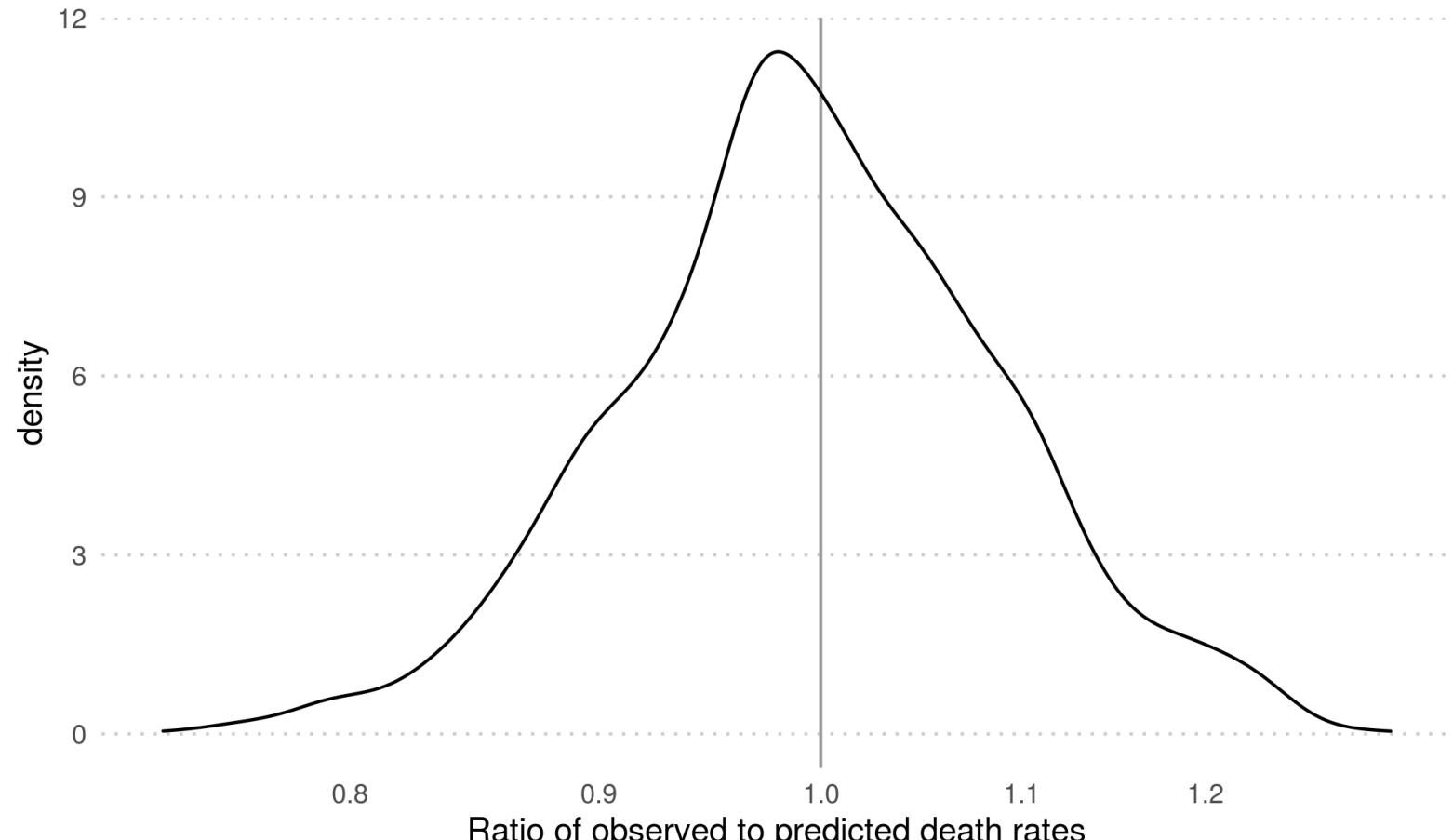


Median absolute prediction error of death counts on test set.

Complete Danish population by sex.

sex  
f  
m

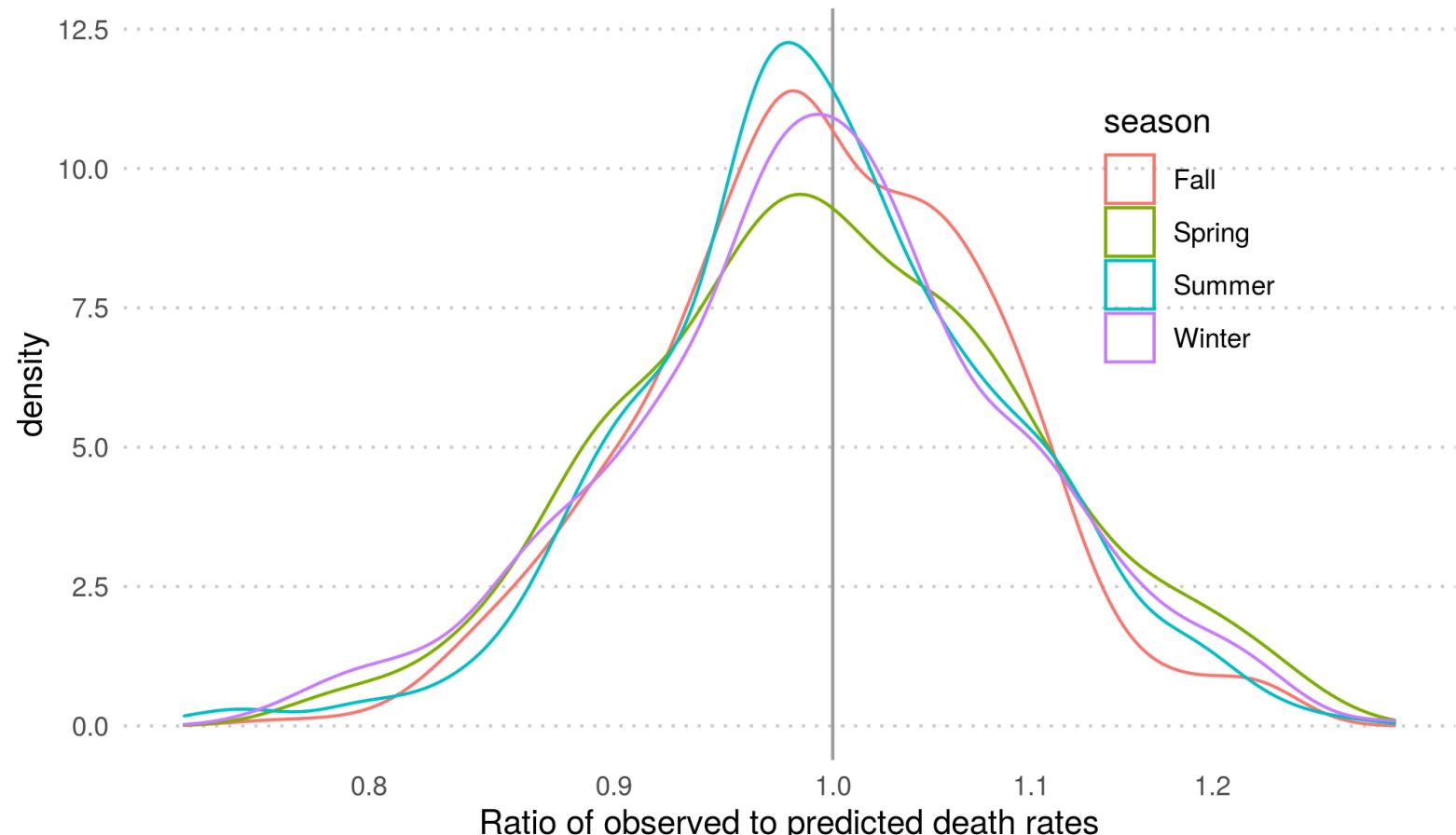
# Out-of-sample residual distribution



Relative prediction error of death rates on test set.

Danish males ages 75–85.

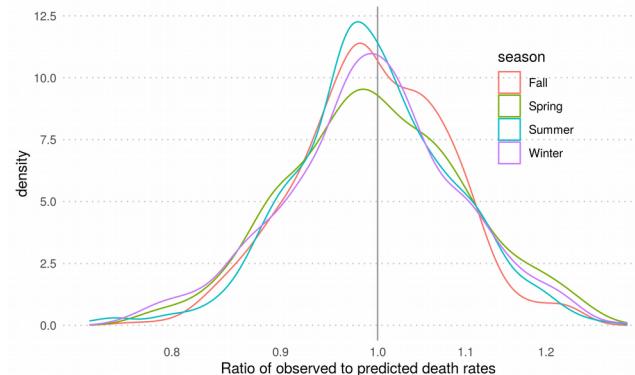
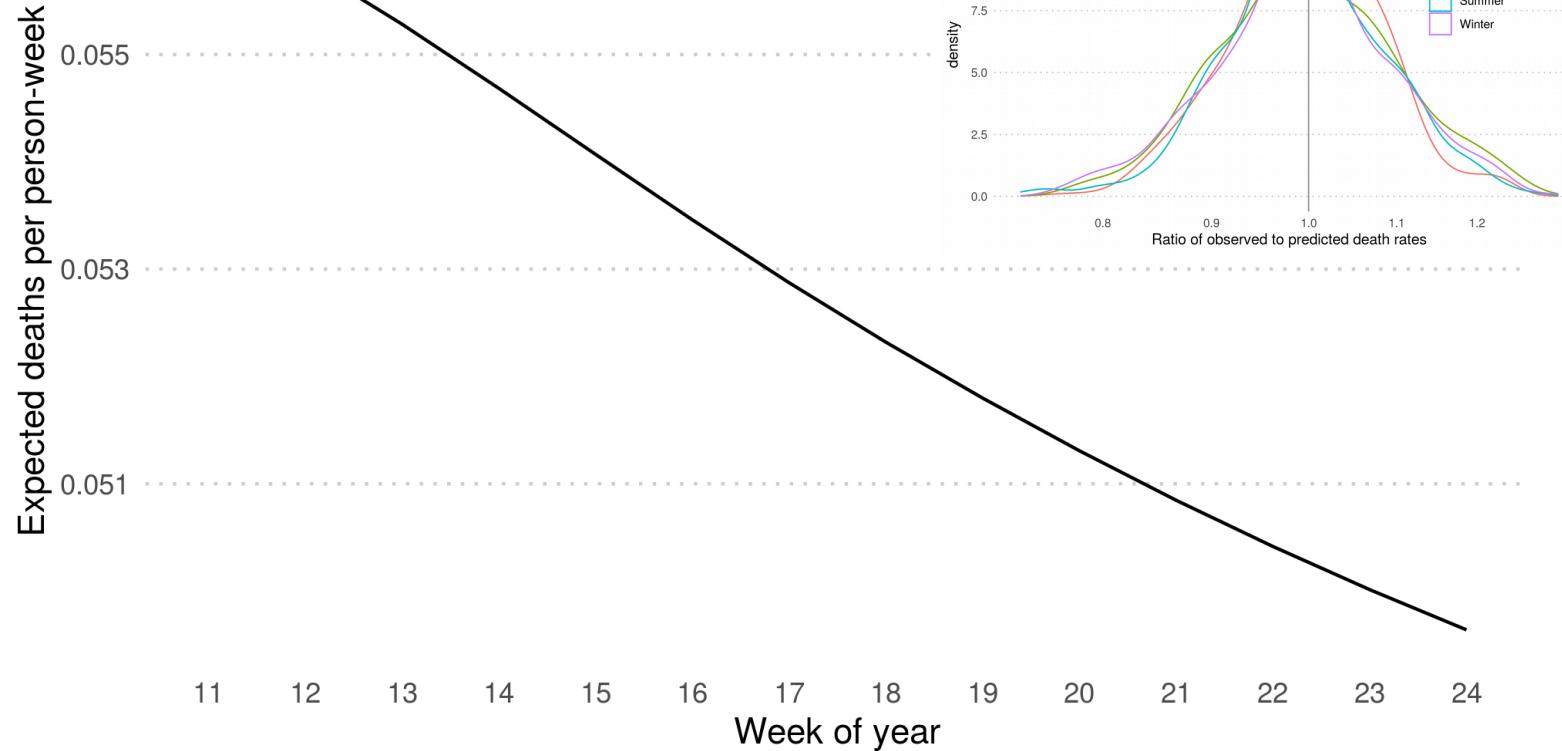
# Out-of-sample residual distribution



Relative prediction error of death rates on test set by season.

Danish males ages 75–85.

# Residual bootstrapping

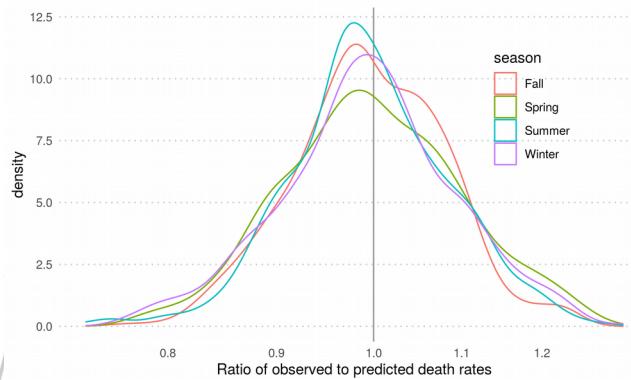
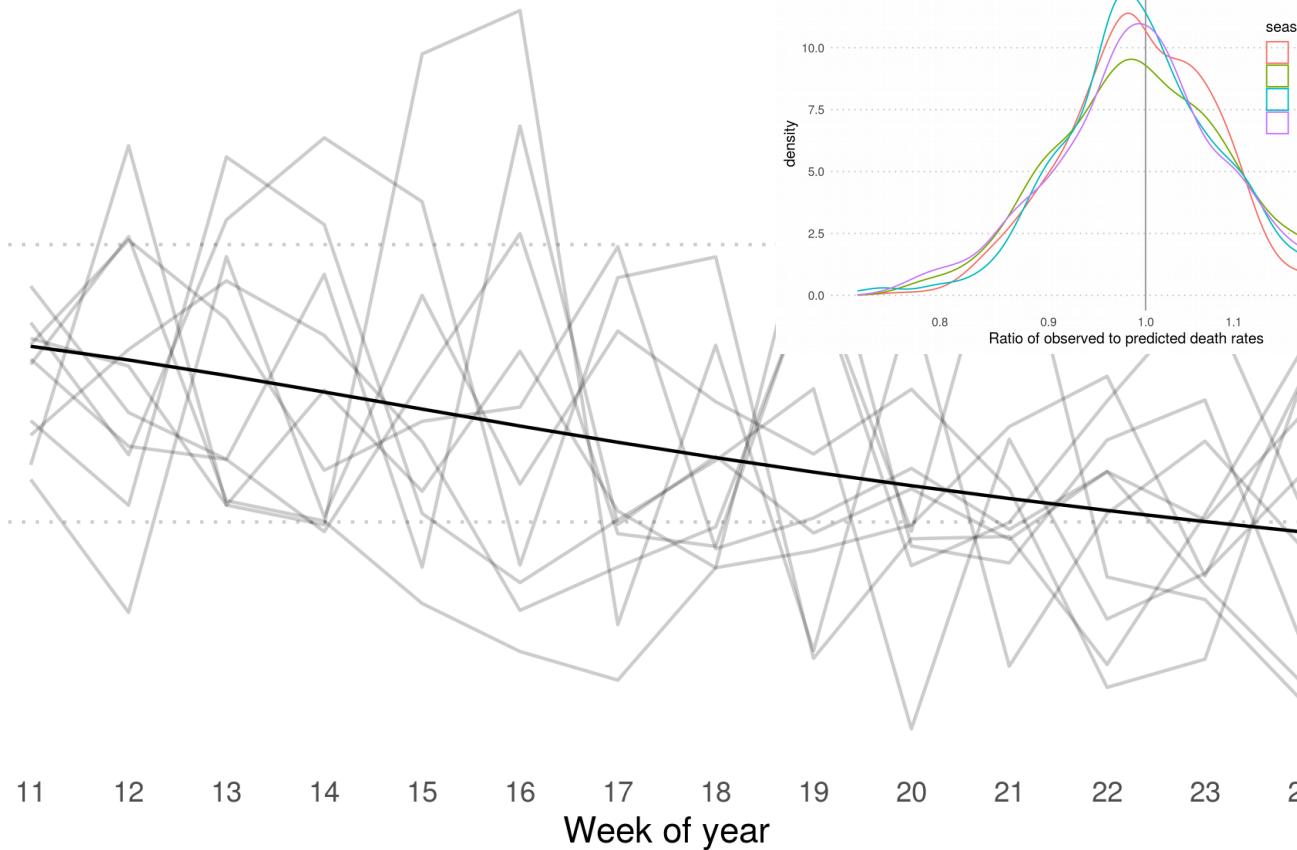


Expected death rates by week of year.

Danish males ages 75–85, year 2009.

# Residual bootstrapping

Expected deaths per person-week

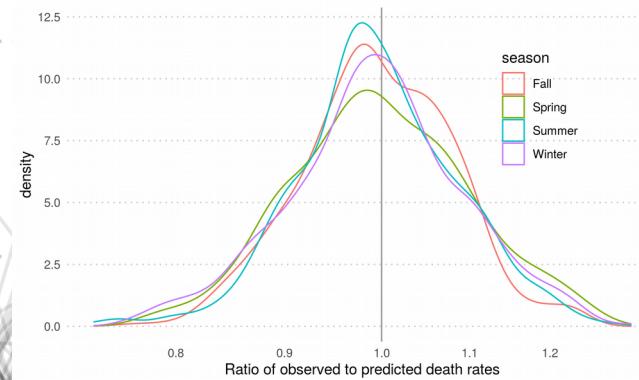
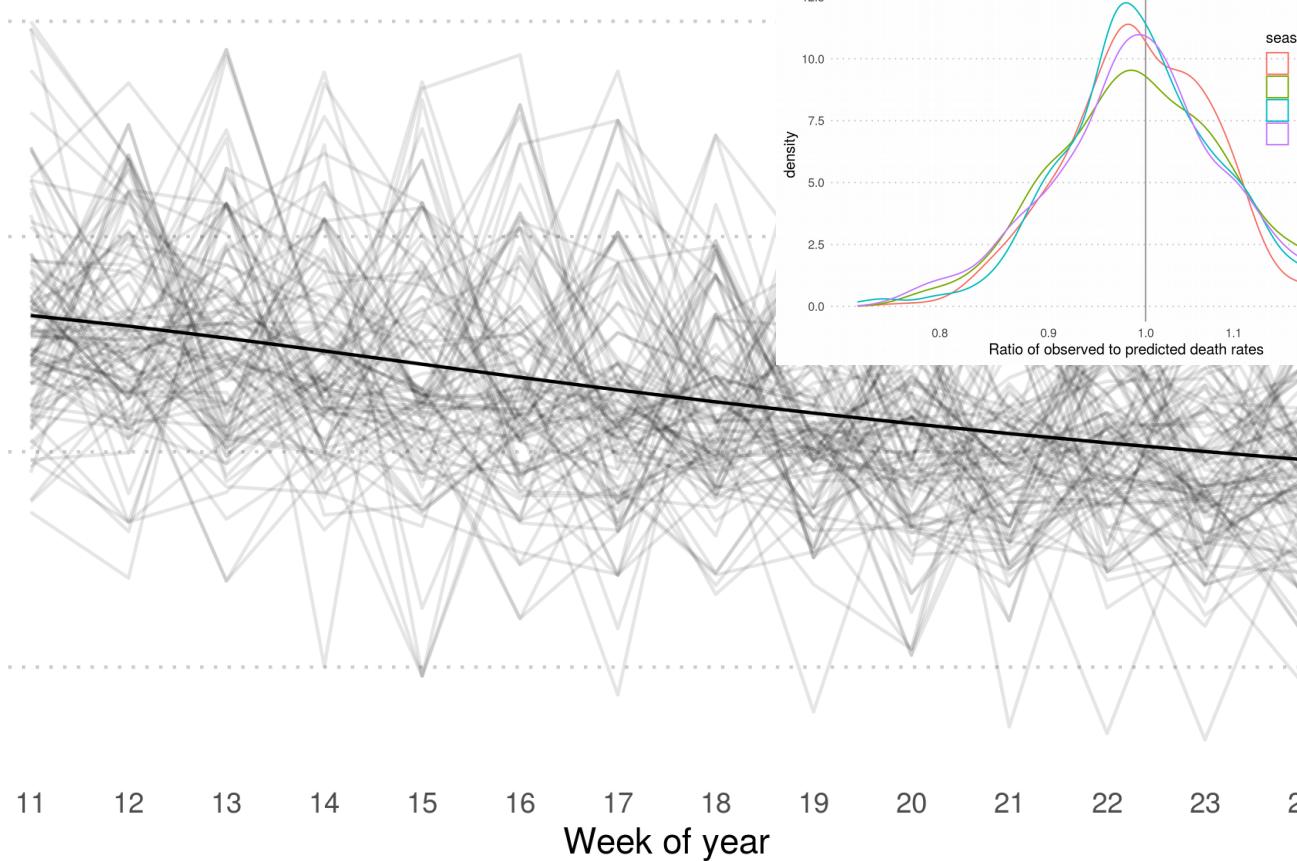


10 bootstrapped death rates by week of year.

Danish males ages 75–85, year 2009.

# Residual bootstrapping

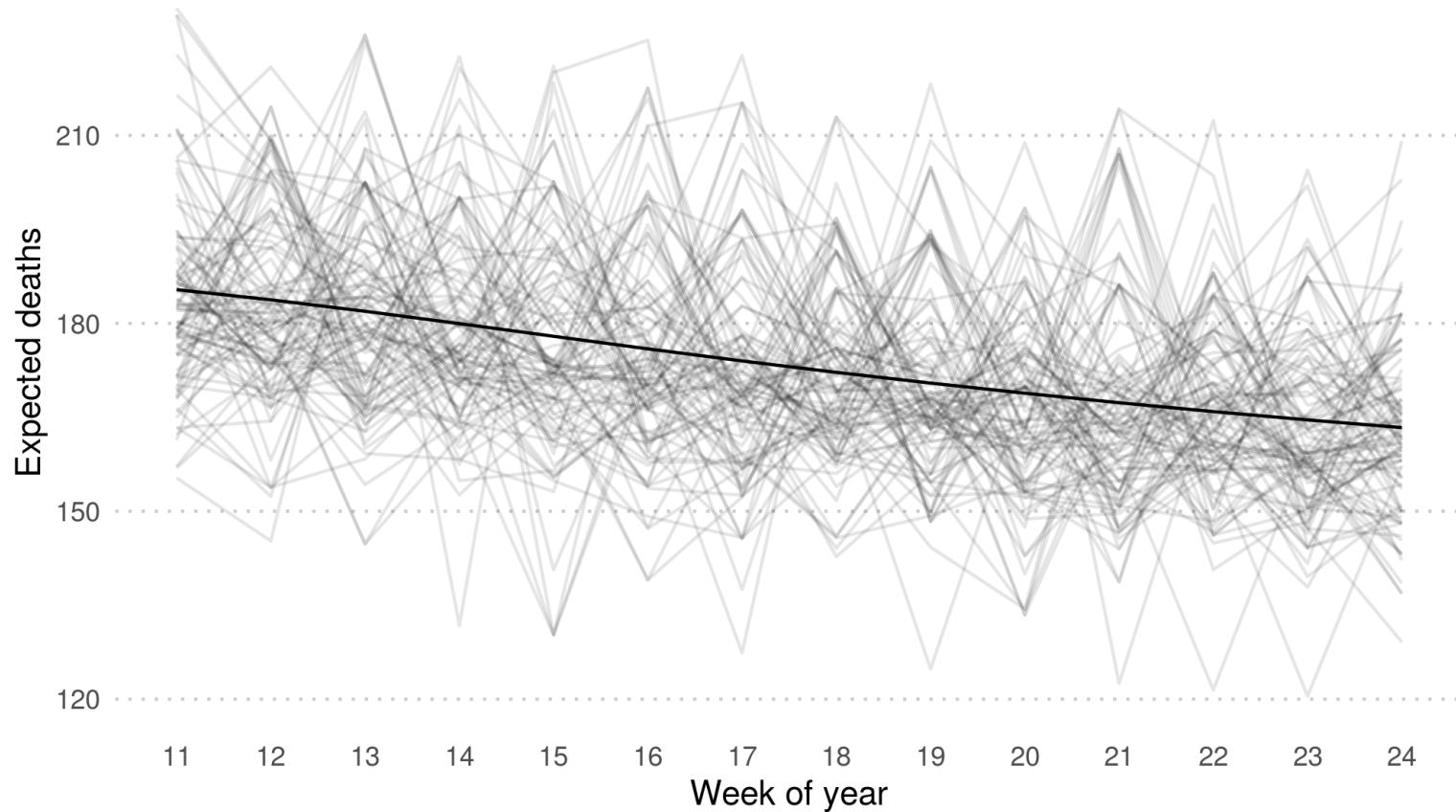
Expected deaths per person-week



100 bootstrapped death rates by week of year.

Danish males ages 75–85, year 2009.

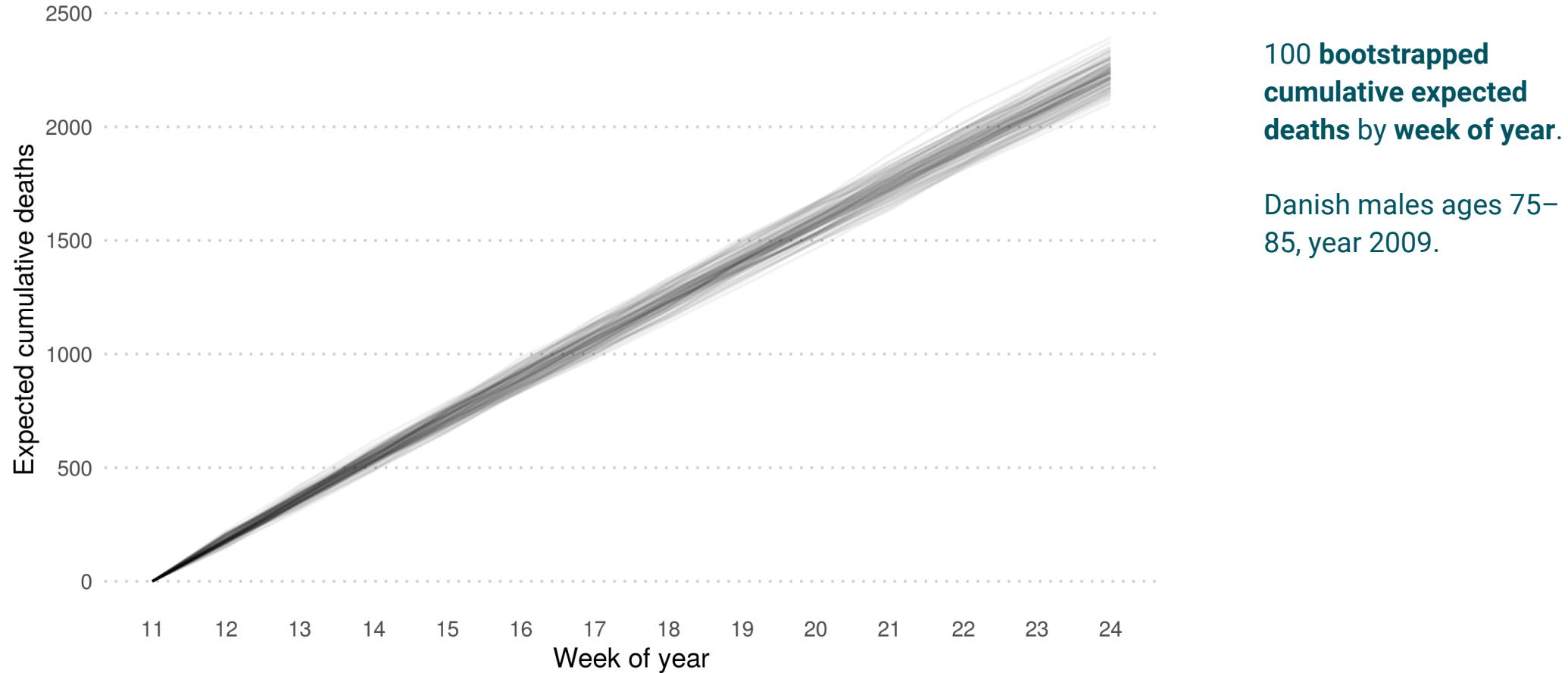
# Residual bootstrapping



100 bootstrapped death counts by week of year.

Danish males ages 75–85, year 2009.

# Residual bootstrapping



# Residual bootstrapping



**Distribution of  
cumulative expected  
death counts at week  
24**

Danish males ages 75–  
85, year 2009.

# **Reproducible analysis**

[github.com/jschoeley](https://github.com/jschoeley)

Jonas Schöley

[jschoeley@health.sdu.dk](mailto:jschoeley@health.sdu.dk)

Twitter: [@jschoeley](https://twitter.com/jschoeley)