

Hidden Markov models for prospective surveillance

Al Ozonoff*, Paola Sebastiani

Boston University School of Public Health
Department of Biostatistics

aozonoff@bu.edu

4/27/06

Influenza surveillance

- Models for influenza
- Data
- National P+I mortality

Hidden Markov Models

Model fitting

Forecasting

Future work

Influenza surveillance

Models for influenza

Serfling's method for influenza.

- Traditional approach: model respiratory illness as sinusoid (Serfling's method).
- Problem: sinusoid fits data poorly during epidemic periods (i.e. winter-time increase in flu activity).
- Implication for prospective surveillance: decreased performance (i.e. lower power for detection of outbreaks) during winter months.

Data

Influenza surveillance

- Models for influenza
- **Data**
- National P+I mortality

Hidden Markov Models

Model fitting

Forecasting

Future work

122 Cities influenza surveillance system:

- CDC operated program running continuously since 1962.
- Weekly counts attributed to pneumonia and influenza (P&I). Reporting lag of 2-3 weeks.
- Approx 25% coverage of U.S. pop'n. Used by CDC for determining epidemic influenza (Serfling).
- Age-specific counts available. 122 cities divided into 9 administrative regions, roughly 14 cities per region.
- Limitations: difficult to accurately attribute deaths to influenza; mortality known to lag morbidity (e.g. ILI activity); dynamics may differ from morbidity (depending on circulating viral strains).

National P+I mortality

Influenza surveillance

- Models for influenza
- Data
- National P+I mortality

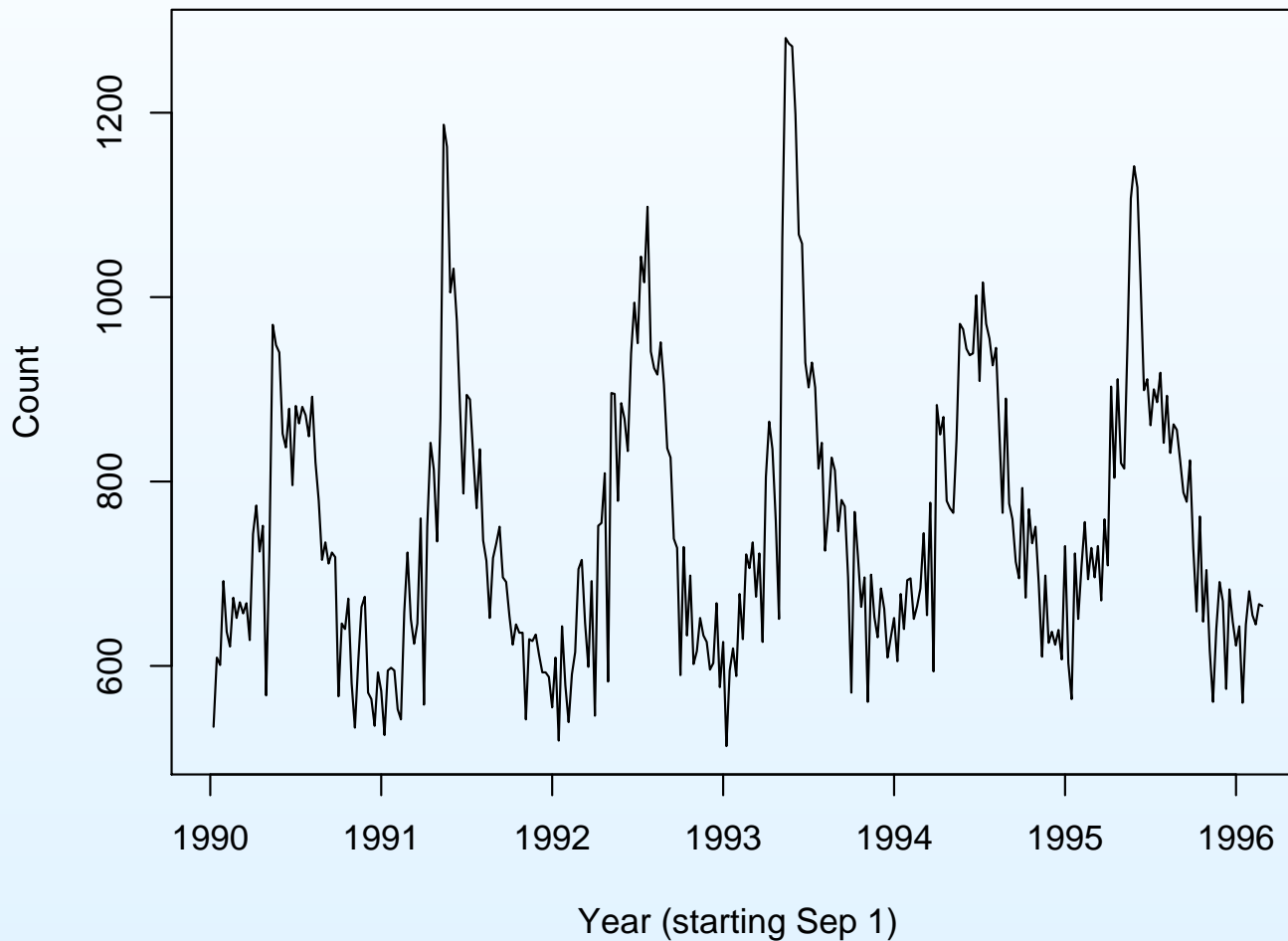
Hidden Markov Models

Model fitting

Forecasting

Future work

Weekly P&I mortality 1990–1996



National P+I mortality

Influenza surveillance

- Models for influenza
- Data
- National P+I mortality

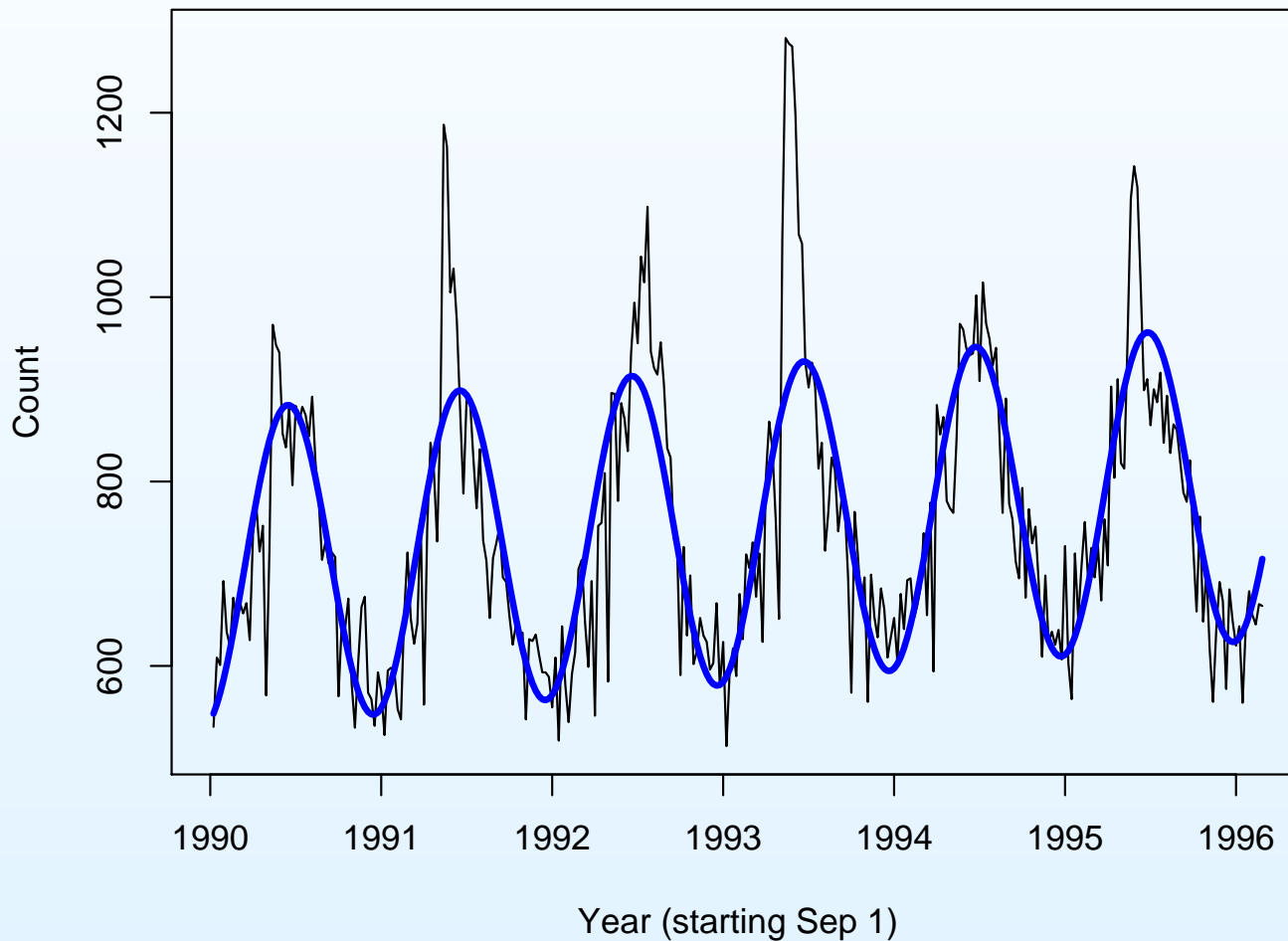
Hidden Markov Models

Model fitting

Forecasting

Future work

Weekly P&I mortality 1990–1996



National P+I mortality

Influenza surveillance

- Models for influenza
- Data
- National P+I mortality

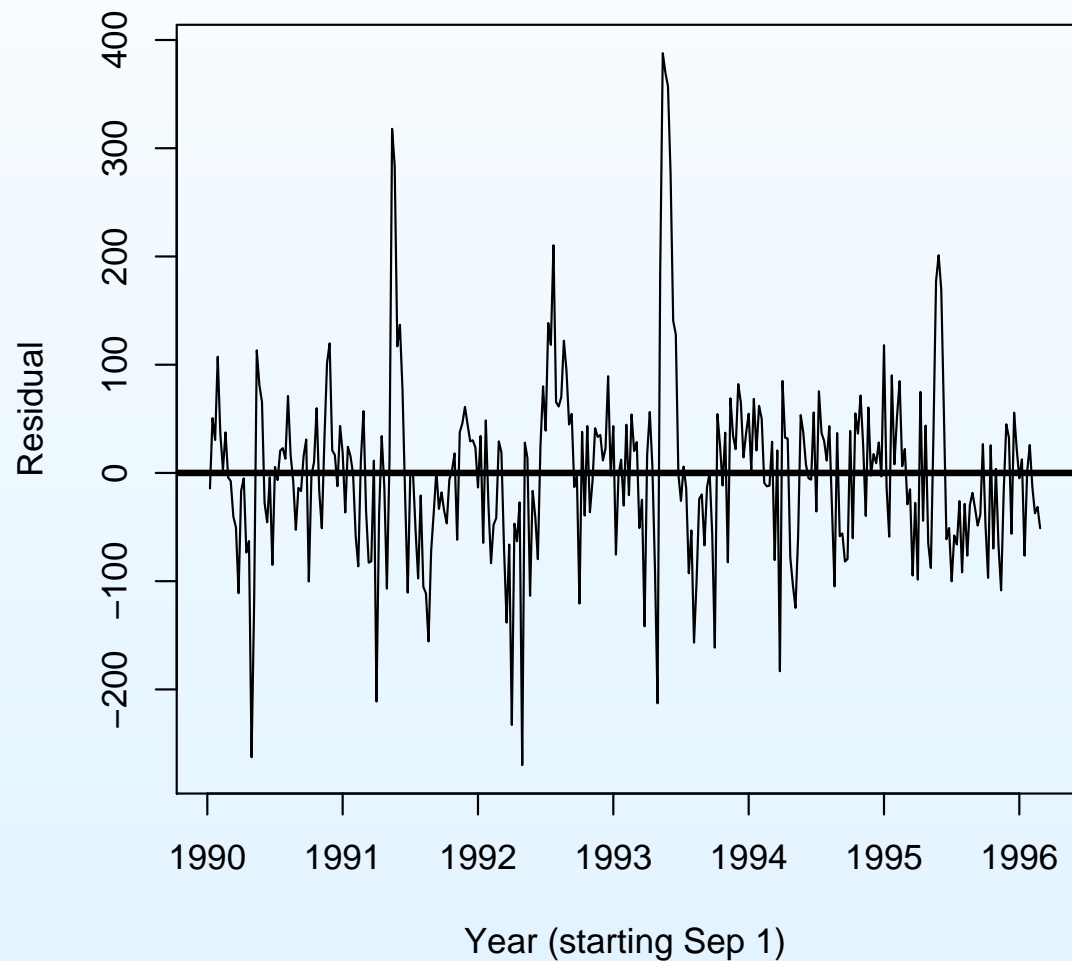
Hidden Markov Models

Model fitting

Forecasting

Future work

Residuals from sinusoidal model



Influenza surveillance

Hidden Markov Models

- Classical approach
- Other approaches
- HMMs
- Evaluation

Model fitting

Forecasting

Future work

Hidden Markov Models

Classical approach

Influenza surveillance

Hidden Markov Models

● **Classical approach**

● Other approaches

● HMMs

● Evaluation

Model fitting

Forecasting

Future work

- Serfling's model: underlying seasonal baseline is roughly sinusoidal. May be driven by temp; annual patterns (e.g. school year); dynamics of disease.

$$Y_t = \alpha_0 + \alpha_1 t + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) + \epsilon_t$$

- Large deviations above this baseline indicate epidemic state. Integrating residuals allows calculation of “excess mortality” i.e. mortality attributed to influenza above what would be expected, accounting for seasonal variation.
- Performs well for what it is asked to do. Not good at one-step-ahead predictions, since model fit is poor during epidemic state.

Influenza surveillance

Hidden Markov Models

● Classical approach

● **Other approaches**

● HMMs

● Evaluation

Model fitting

Forecasting

Future work

Other approaches

- Periodic regression with auto-regressive component (PARMA). Used in syndromic surveillance settings. Better model fit thanks to AR component.
- “Method of analogues”: non-parametric forecasting. Outperforms many methods in one-step-ahead prediction (Viboud 2003). Non-parametric \Rightarrow ignores and obscures knowledge about mechanism of disease.
- Nuño and Pagano developing mixed models approach using Gaussian with phase shift as random effect. Also incorporate bimodal Gaussian for occasional dual-wave behavior.

Hidden Markov Models (HMMs)

Influenza surveillance

Hidden Markov Models

- Classical approach
- Other approaches
- **HMMs**
- Evaluation

Model fitting

Forecasting

Future work

Our approach: HMMs.

- ‘Hidden’ (latent, unobserved) discrete random variable, representing some aspect of disease process.
- Observed variables are modeled, conditional upon the hidden state. Know which state \Rightarrow know distribution of observed random variable.
- Markov property: conditional probability of state change (transition probability) depends only on the value of latent state at previous time point. Thus specify the Markov model for k states with a $k \times k$ matrix of transition probabilities, and the distributions of the observed data conditional on the hidden state.
- Parameter estimation using Bayesian inference Using Gibbs Sampling (BUGS). Freeware available, e.g. WinBUGS, BRUGS.

WinBUGS screen shot

Influenza surveillance

Hidden Markov Models

- Classical approach
- Other approaches
- **HMMs**
- Evaluation

Model fitting

Forecasting

Future work

The screenshot displays the WinBUGS14 software interface. The main window shows a model file named 'hmm-ar-2-state.txt' with the following code:

```
##### Model
model;
{
  epsilon[1] <- 1
  #mu[1] <- 534
  b[1] <- 534
  for(t in 2:N)
  {
    ind[t] ~ dbern( p.epsilon[ epsilon[t-1] ] )
    epsilon[t] <- ind[t] + 1
    b[t] <- alpha + beta.0*t + beta.1*sin(t*2*pi/52.3) + (beta.2)*cos(t*2*pi/52.3)
    mu[t] <- b[t] + alpha.e*(ind[t]) + gamma.e*(ind[t])*(x[t-1] - b[t-1])
    sigma.eps[t] <- ind[t]*sigma[1] + (1-ind[t])*sigma[2]
    x[t] ~ dnorm(mu[t], sigma.eps[t])
  }
  alpha ~ dnorm(a.coef,prec.a)
  beta.0 ~ dnorm(p.coef,prec.coef)
  beta.1 ~ dnorm(p.coef,prec.coef)
  beta.2 ~ dnorm(p.coef,prec.coef)
  alpha.e ~ dpois(p.muind)
  gamma.e ~ dnorm(p.coef,prec.coef)
  p.epsilon[1] ~ dbeta(alpha.1,alpha.2)
  p.epsilon[2] ~ dbeta(alpha.1,alpha.2)
  sigma[1] ~ dgamma(alpha.1,alpha.2)
  sigma[2] ~ dgamma(alpha.1,alpha.2)
}

##### Data
list(
  N=320, a.coef=700, prec.a=10, p.coef=0, pi=3.141593,
  prec.coef=0.001, p.muind=250, alpha.1=1, alpha.2=1,
  x =
  c(534,609,601,692,637,621,674,652,669,657,668,628,743,774,724,7
  52,568,726,970,949,940,852,837,879,796,882,863,881,872,849,892,
  821,779,715,734,711,723,718,567,646,640,673,581,533,604,664,675
  ,571,564,535,593,573,525,595,598,595,553,542,658,723,651,624,64
  6,760,558,751,842,813,735,869,1187,1163,1005,1031,974,878,787,8
  94,889,827,771,835,736,714,652,717,733,751,696,691,654,623,645,
  636,636,542,629,627,634,612,593,593,588,555,609,519,643,579,539
  ,591,615,705,715,647,599,692,546,752,755,809,583,896,895,779,88
  5,868,833,939,994,950,1044,1016,1098,941,923,916,951,905,836,82
```

The interface includes three tool windows:

- Specification Tool:** Contains buttons for 'check model', 'load data', 'compile', 'load inits', and 'gen inits'. It also has a 'num of chains' field set to 1 and a 'for chain' dropdown set to 1.
- Update Tool:** Contains fields for 'updates' (4000), 'refresh' (100), 'iteration' (5000), and 'thin' (1). It also has checkboxes for 'over relax' and 'adapting'.
- Sample Monitor Tool:** Shows the current node 'mu' with 'chains' set to 1 and 'to' set to 1. It displays a list of percentiles: 2.5, 5, 10, 25, median, 75, 90, 95, and 97.5. Buttons for 'clear', 'set', 'trace', 'history', 'density', 'stats', 'coda', 'quantiles', 'bgr diag', and 'auto cor' are visible.

Hidden Markov Models (HMMs)

Influenza surveillance

Hidden Markov Models

- Classical approach
- Other approaches
- **HMMs**
- Evaluation

Model fitting

Forecasting

Future work

Model fitting in WinBUGS:

- Sequence of hidden states is treated as a free parameter and fit simultaneously with other model coefficients.
- Computational demanding for long time series; parameter space has order k^n .
- Convergence via Gibbs sampling may be an issue, esp for misspecified models.
- Latent variable provides information about mechanism of disease. Epidemic and non-epidemic behavior can be modeled separately.

Hidden Markov Models (HMMs)

Influenza surveillance

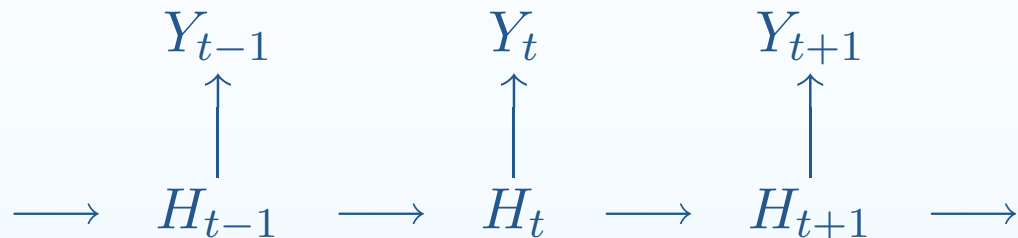
Hidden Markov Models

- Classical approach
- Other approaches
- **HMMs**
- Evaluation

Model fitting

Forecasting

Future work



Y_t are observed data i.e. weekly P&I counts.

H_t are the hidden states (for us, 2-state model).

Arrows indicate conditional dependencies.

$$Y_t \sim \alpha_0 + \alpha_1 t + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) \mid H_t = 0$$

$$Y_t \sim \left(\alpha_0 + \alpha_e\right) + \alpha_1 t + \beta_1 \sin\left(\frac{2\pi t}{52}\right) + \beta_2 \cos\left(\frac{2\pi t}{52}\right) \mid H_t = 1$$

Evaluation

Influenza surveillance

Hidden Markov Models

- Classical approach
- Other approaches
- HMMs
- **Evaluation**

Model fitting

Forecasting

Future work

Evaluation scheme for model performance:

- Systematically investigate various HMMs and evaluate (with other approaches) using RMSE on one-step-ahead predictions.
- Use fixed period (e.g. 1992-1996) to fit all models, and subsequent year (1997) for predictions. Repeat on other time periods so evaluation is not dependent on time period chosen. “Virtual prospective surveillance” (Siegrist).
- Compare HMMs; Serfling’s method; PARMA; other methods?

Influenza surveillance

Hidden Markov Models

Model fitting

- Models
- Model fits
- Residuals

Forecasting

Future work

Model fitting

Influenza surveillance

Hidden Markov Models

Model fitting

● **Models**

● Model fits

● Residuals

Forecasting

Future work

Models

1. Traditional cyclic model (Serfling). OLS regression with terms for intercept, linear trend, two periodic terms for sinusoid with phase shift.
2. Periodic auto-regression (PARMA) with fixed order (2,0) fits cyclic model plus additional ARMA terms.
3. Naive 2-state HMM. Non-epidemic state follows Serfling. Epidemic state modeled with simple mean shift.
4. 2-state AR-HMM. Non-epidemic state, data follow PARMA. Epidemic state auto-regresses deviation from cyclic baseline.

Serfling

[Influenza surveillance](#)

[Hidden Markov Models](#)

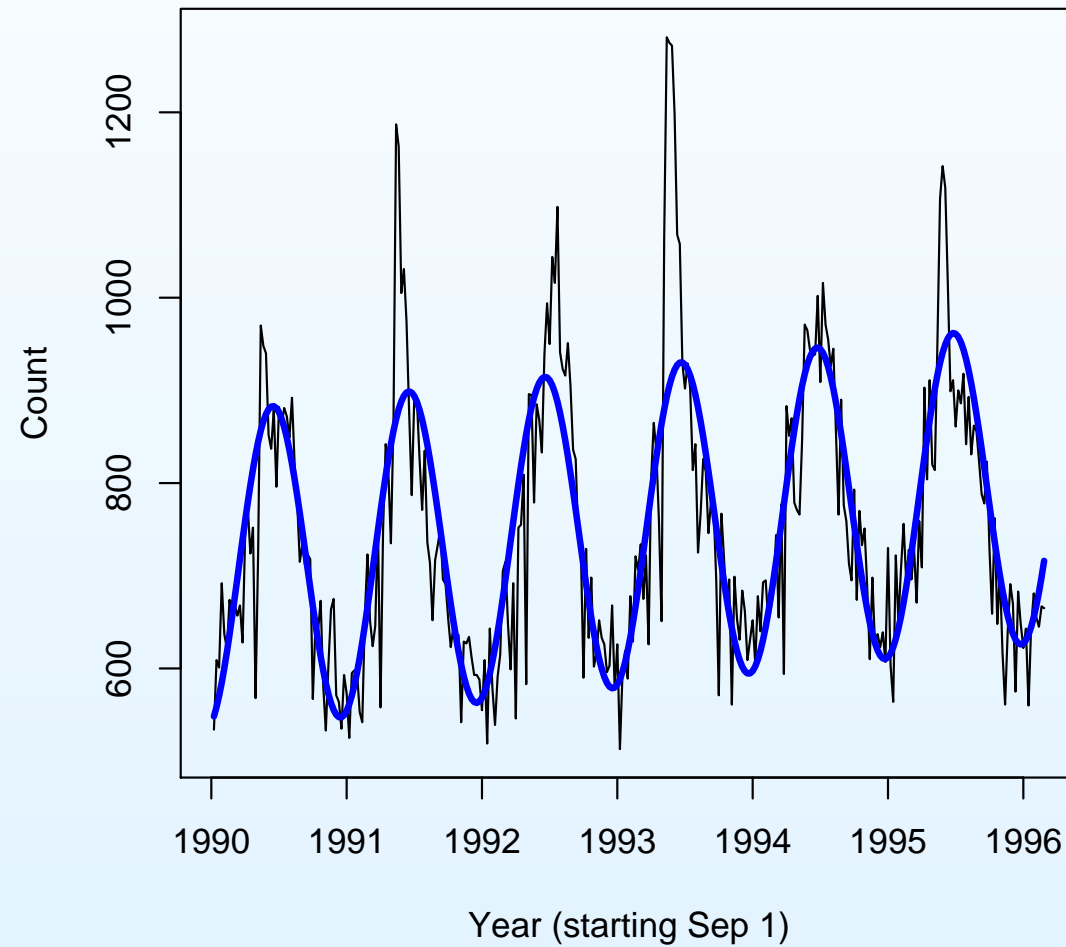
[Model fitting](#)

- Models
- **Model fits**
- Residuals

[Forecasting](#)

[Future work](#)

Serfling's model



Simple HMM

Influenza surveillance

Hidden Markov Models

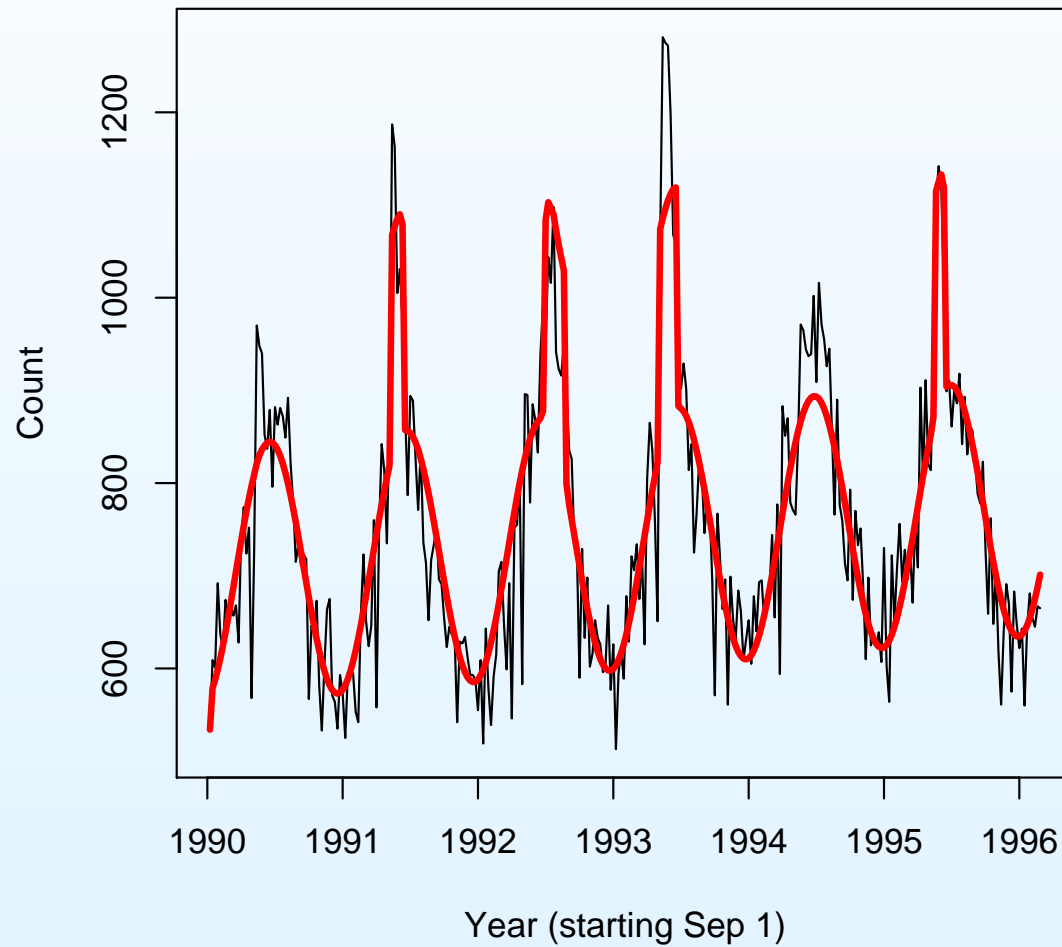
Model fitting

- Models
- **Model fits**
- Residuals

Forecasting

Future work

Simple HMM



PARMA

Influenza surveillance

Hidden Markov Models

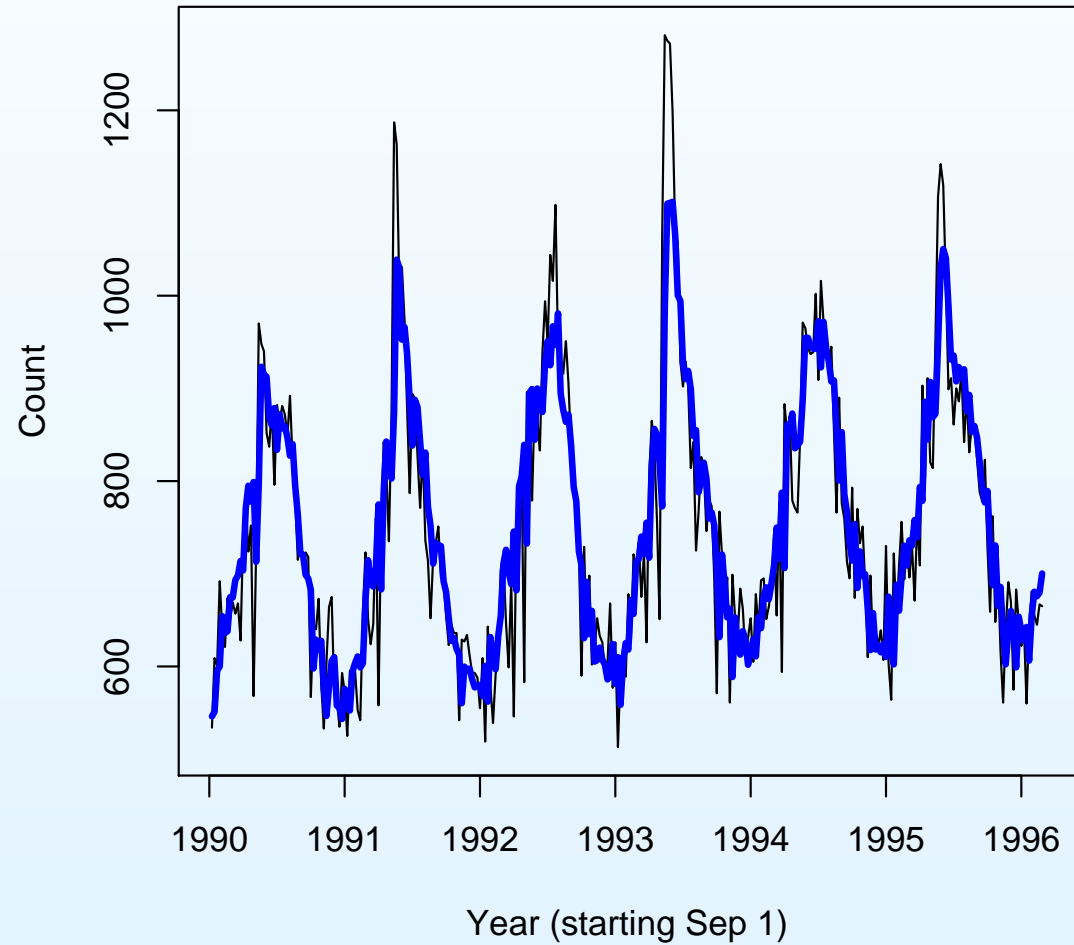
Model fitting

- Models
- Model fits
- Residuals

Forecasting

Future work

PARMA model



AR-HMM

Influenza surveillance

Hidden Markov Models

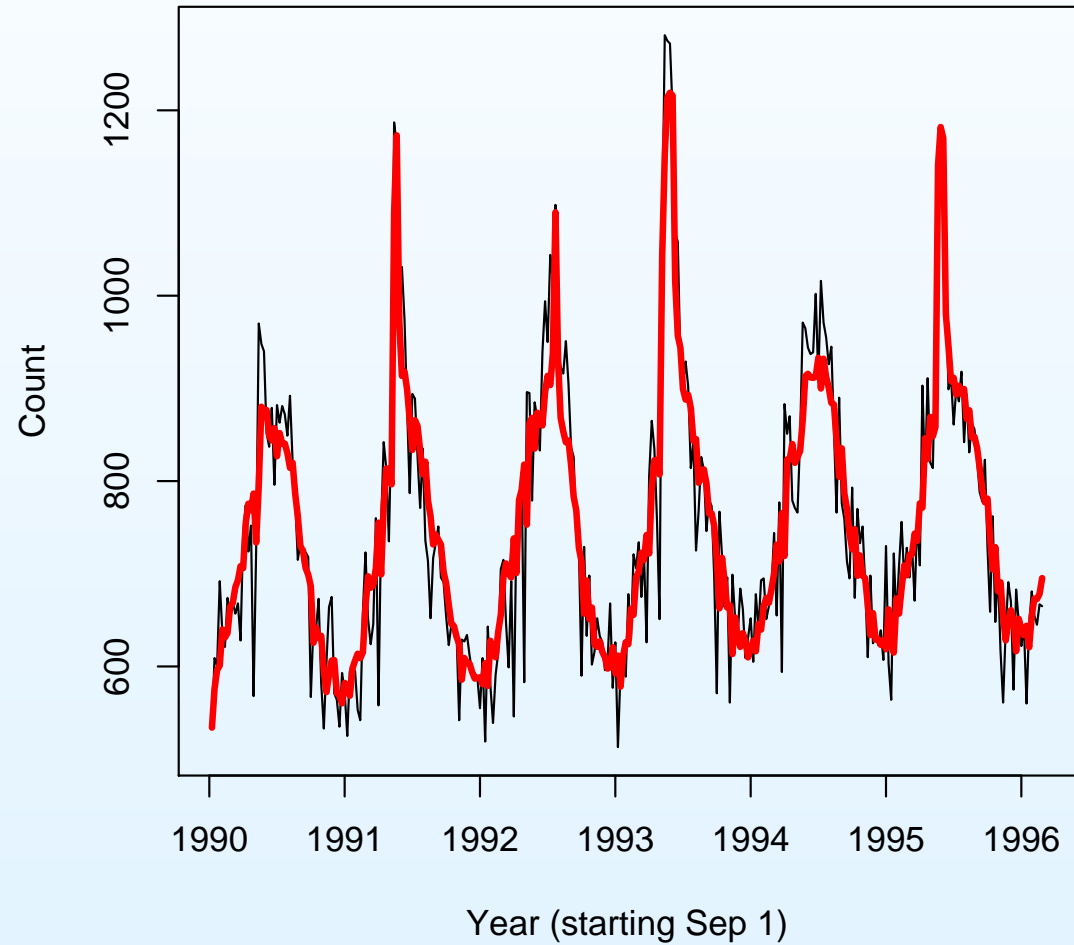
Model fitting

- Models
- **Model fits**
- Residuals

Forecasting

Future work

AR-HMM



Residuals

Influenza surveillance

Hidden Markov Models

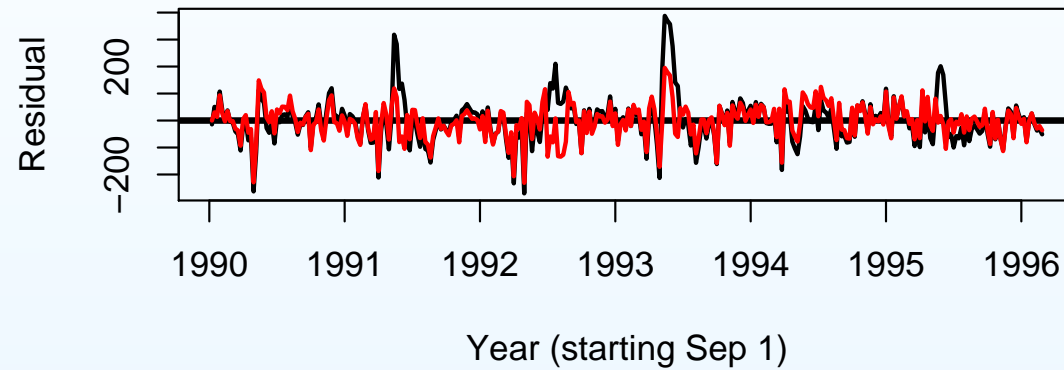
Model fitting

- Models
- Model fits
- Residuals

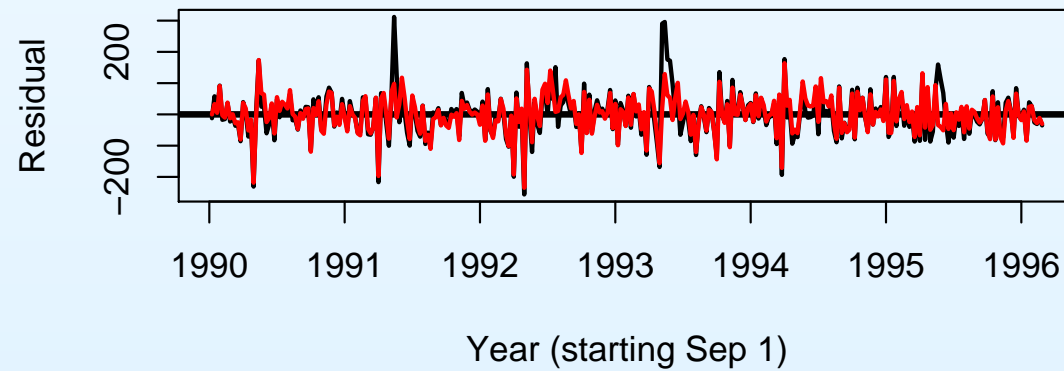
Forecasting

Future work

Residuals – Serfling/HMM



Residuals – PARMA/AR-HMM



Residuals

Influenza surveillance

Hidden Markov Models

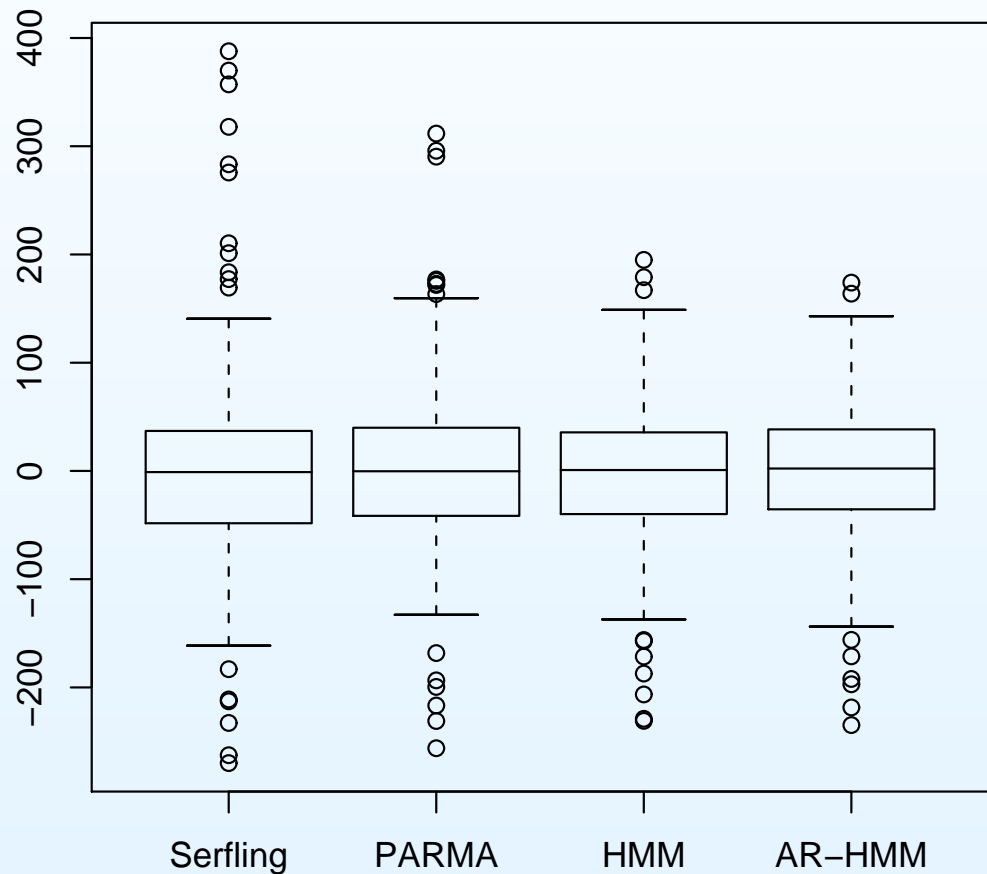
Model fitting

- Models
- Model fits
- Residuals

Forecasting

Future work

Model residuals



Residuals

Influenza surveillance

Hidden Markov Models

Model fitting

- Models
- Model fits
- **Residuals**

Forecasting

Future work

Both HMMs provide a roughly 25% reduction in RMSE from Serfling, roughly 10% reduction for PARMA.

Model	RMSE
Serfling	83.3
PARMA	72.0
Simple HMM	63.7
AR-HMM	60.4

ACF of residuals

[Influenza surveillance](#)

[Hidden Markov Models](#)

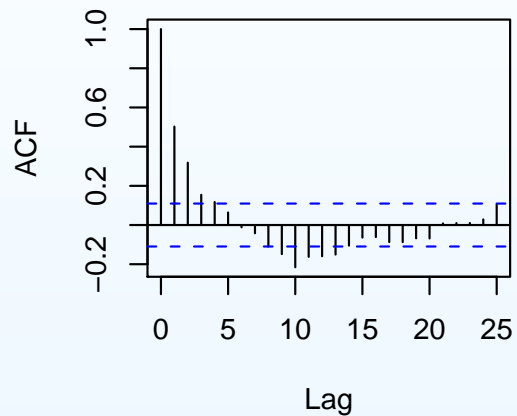
[Model fitting](#)

- Models
- Model fits
- **Residuals**

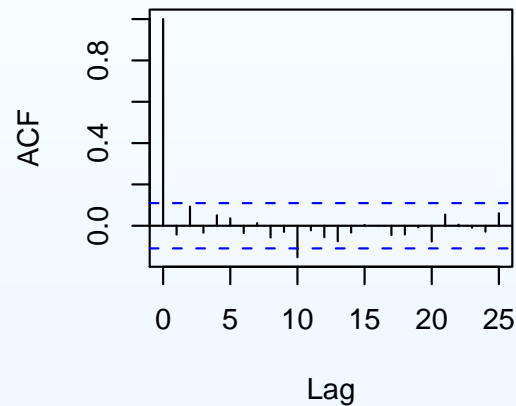
[Forecasting](#)

[Future work](#)

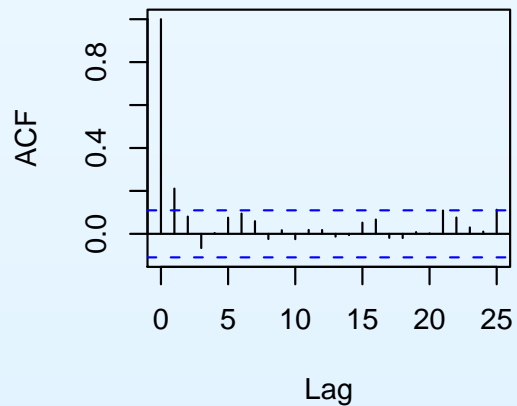
Serfling



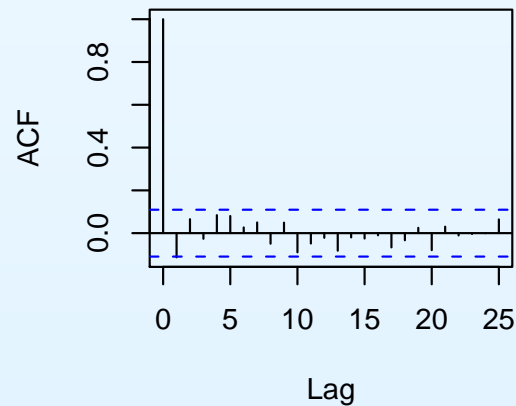
PARMA



HMM



AR-HMM



Discussion

Preliminary conclusions from model-fitting:

- HMMs offer superior model fit during epidemic periods. AR components do not offer much improvement during non-epidemic period.
- Both models with AR component eliminate auto-correlation of residuals. Important for use of control chart detection methods (e.g. Shewhart, CUSUM).
- Interpretability of latent variable (for two-state models) provides immediate benefit beyond better fit.
- Many-state models ($k > 2$) prove difficult to fit for convergence reasons.

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

- Forecasting
- Data
- Results
- Discussion

Future work

Forecasting

Forecasting

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

● **Forecasting**

● Data

● Results

● Discussion

Future work

Framework for evaluating forecasting performance:

- Three time periods chosen (essentially) at random. Five year periods, with four years for model fit/parameter estimation, and one year to evaluate forecasts (RMSE of one-step-ahead residuals).
- Evaluated PARMA and AR-HMM models using AR(2) components. Included Serfling and LVCF (lag 1 predictors) for comparison purposes.
- For HMMs, *ad hoc* procedure to choose state for prediction. Ratio of previous non-epidemic : epidemic residuals greater than 1.5 \Rightarrow use epidemic state prediction.

Data

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

● Forecasting

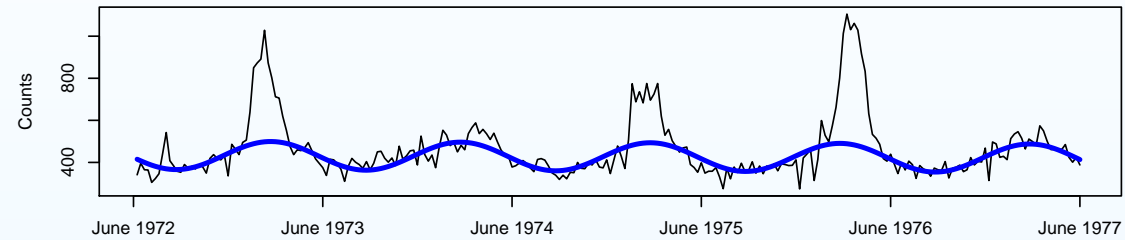
● **Data**

● Results

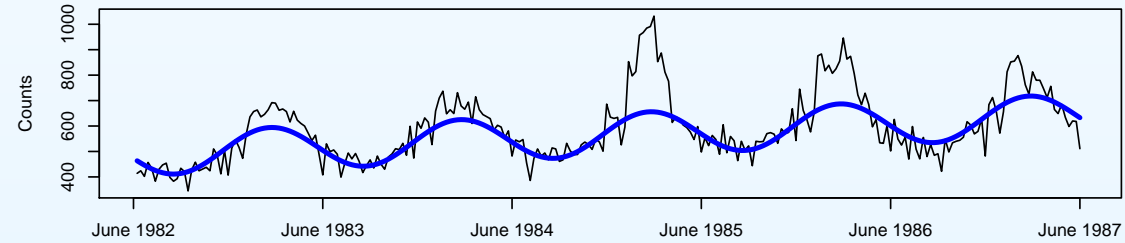
● Discussion

Future work

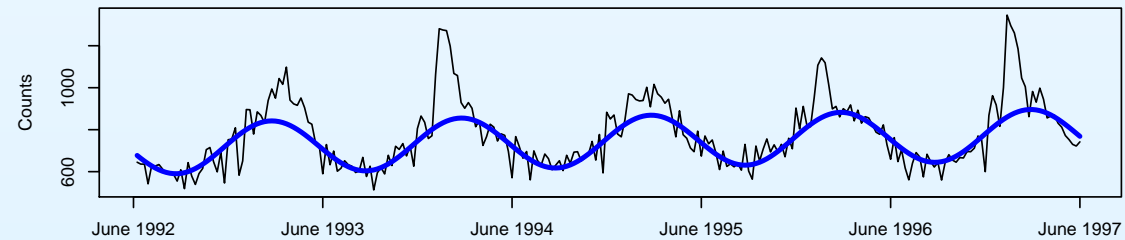
1972-1977



1982-1987



1992-1997



Results

Results from forecasting methods depend on characteristics of epidemic season. HMMs generally outperform other methods.

Model	1972-77	1982-87	1992-97
LVCf	53.7	70.6	92.9
Serfling	36.1	68.1	129.9
PARMA	64.3	66.8	90.4
AR-HMM	45.3	62.3	94.1

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

- Forecasting
- Data
- **Results**
- Discussion

Future work

Discussion

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

- Forecasting
- Data
- Results
- Discussion

Future work

Preliminary conclusions from model-fitting:

- HMMs perform adequately for forecasting, with modest gains over other methods.
- Significant advantage provided by using all available data, rather than excluding portions of the time series as is customary with Serfling and PARMA models. AR component achieves better fit, thanks to more precise parameter estimation.
- Difficulty for forecasting is deciding on hidden state to use for prediction. More work needed to resolve this issue.

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

Future work

- Integration
- Diffusion of influenza

Future work

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

Future work

● **Integration**

● Diffusion of influenza

Integration

Bayesian methodology for integration of multiple time series:

- Developed for gene expression data.
- Bottom-up heuristic search to aggregate time series data; likelihood criterion using model specification to identify “clusters” of time series.
- Hypothesis: cluster assignments will vary over time; possibly dependent on circulating strain, point of origin; less evidence of diffusion in recent years.

Diffusion of influenza

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

Future work

● Integration

● Diffusion of influenza

Evidence for diffusion dynamics?

- Standardization of multiple time series to allow for direct comparison across geographic regions.
- Comparison of standardized counts across distances to quantify diffusion over course of surveillance period. Use L^2 norm, cross-correlation, for dissimilarity measure between series.
- Eventual goal: development of true spatio-temporal model for influenza activity.

Influenza surveillance

Hidden Markov Models

Model fitting

Forecasting

Future work

- Integration
- Diffusion of influenza

Acknowledgements

- Many thanks to David Banks, Doug Kelly, SAMSI directors and staff, and participants in the AD working group.
- The author gratefully acknowledges the contributions of his co-author Dr. Sebastiani, research assistant Suporn Sukpraprut, and doctoral candidate Ling Wang.
- Thanks also to collaborators at the Harvard School of Public Health: Marcello Pagano, Laura Forsberg, Bethany Hedt, Caroline Jeffery, Justin Manjourides, and Miriam Nuño.
- Assistance in acquiring influenza data provided by Willie J Anderson and Felicia J Dawkins (CDC/CoCHIS/NCPHI/DISSS).
- Research partially supported by a pilot grant originating from the Blood Center of Wisconsin, via NIAID grant U19 AI62627-02.