

Nowcasting influenza epidemics using a sentinel network based influenza surveillance system

Baltazar Nunes¹ Isabel Natário² M. Lucília Carvalho³

¹Departamento de Epidemiologia, Instituto Nacional de Saúde Dr. Ricardo Jorge

²Departamento de Matemática, Faculdade de Ciências e Tecnologia (UNL); CEAUL

³Departamento de Estatística e Investigação Operacional, Faculdade de Ciências (UL);
CEAUL

September 8, 2012

European Congress of Epidemiology, Porto, 2012

Introduction

Timeliness of public health surveillance systems

- 1 **Timeliness** of a public health surveillance system is crucial for its capacity of a **timely intervention**;
- 2 In Europe the epidemiological **surveillance of influenza is supported by general practitioners (GP) sentinel networks**, that weekly report epidemiological bulletins with influenza-like illness rates and the number of ILI cases tested positive for influenza;
- 3 These bulletins are issued **between Wednesday (country level) and Friday (European level)**, reporting the previous week observed influenza estimates, **representing a 2 to 4 days reporting delay**;

Introduction

Nowcasting from incomplete information

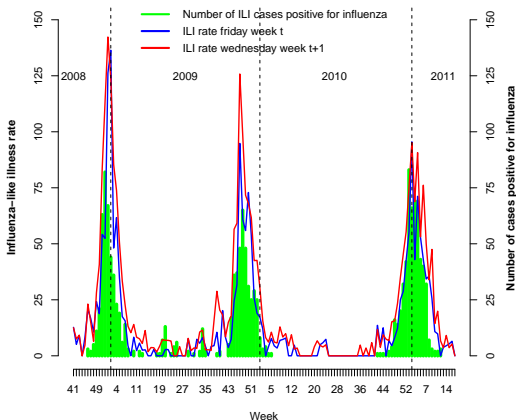
- 1 Some **surveillance systems** use web interfaces or computer routines that can provide daily data streams **accessing the current situation**;
- 2 The process of **predicting the present week situation** using the available incomplete information available, has received the term **nowcasting** and has high public health interest;
- 3 **Examples of nowcasting applied to surveillance are:** influenza A(H1N1) hospitalizations during 2009 pandemic in the Netherlands (Donker et al EJE 2011), the outbreak of haemolytic-uremic syndrome in Germany (Heiden et al ESCAIDE 2011) and excess mortality estimation during June 2011 England and Wales heat wave (Green HK et al JECH 2012).

Development of a **statistical model** that, on a weekly bases, uses **all data collected until Friday of week t** by the surveillance system to nowcast two measures of interest:

- 1 *the week t official ILI incidence rate (only known on Wednesday of week $t + 1$);*
- 2 *the week t influenza activity state: epidemic or non-epidemic.*

Influenza-like illness incidence rates of week t , calculated by Friday of week t - $y_{t(t)}$ - and by **Wednesday of week $t + 1$ (official rate - $y_{t(t+1)}$).**

source: GP sentinel network ("Médicos-Sentinel") and Sentinel Emergency Services network ("Serviços de Urgência Sentinela") from INSA.



A two states (non-epidemic $S_t = 0$, epidemic $S_t = 1$) non-homogeneous hidden Markov model (HMM) was proposed to estimate the official rate $y_{t(t+1)}$:

$$y_{t(t+1)} = \begin{cases} \mu + \beta_1 \cos\left(\frac{2\pi t}{52}\right) + \beta_2 \sin\left(\frac{2\pi t}{52}\right) + \theta_{0,1}y_{t(t)0} + e_{t,0} & S_t = 0 \\ \mu + \beta_1 \cos\left(\frac{2\pi t}{52}\right) + \beta_2 \sin\left(\frac{2\pi t}{52}\right) + \theta_{1,1}y_{t(t)1} + \theta_{1,2}y_{t(t)1}^2 + e_{t,1} & S_t = 1 \end{cases}$$

where $e_{t,i} \sim N(0, \tau_i)$, with precision $\tau_0 > \tau_1$, $t = 1, \dots, T$ and $i \in \{0, 1\}$, and

$$y_{t(t)0} = \begin{cases} y_{t(t)} & v_{t-1(t)} \leq 20 \\ 0 & \text{otherwise} \end{cases},$$

$$y_{t(t)1} = \begin{cases} y_{t(t)} & v_{t-1(t)} \geq 1 \\ 0 & \text{otherwise} \end{cases}.$$

are two new covariates based on the **observation of that same week t ILI rate on Friday of week t ($y_{t(t)}$)** and on the **number of ILI cases tested positive for influenza in the week before $t - 1$ ($v_{t-1(t)}$)**.

Three models were proposed with different influenza state-transition probabilities

matrices $\Gamma^t = (\gamma_{j,i}^t)$, where $\gamma_{j,i}^t = P[S_t = i | S_{t-1} = j]$:

- **Model 1:** $\gamma_{j,i}^t = \frac{\exp(\alpha_{j,i,0} + \alpha_{j,i,1}Y_t(t))}{(1 + \exp(\alpha_{j,i,0} + \alpha_{j,i,1}Y_t(t)))}$
- **Model 2:** $\gamma_{j,i}^t = \frac{\exp(\alpha_{j,i,0} + \alpha_{j,i,1}Y_t(t) + \alpha_{j,i,2}V_{t-1}(t))}{(1 + \exp(\alpha_{j,i,0} + \alpha_{j,i,1}Y_t(t) + \alpha_{j,i,2}V_{t-1}(t)))}$
- **Model 0:** $\gamma_{j,i}^t = \gamma_{j,i}$, considered time invariant for comparison purposes.

for any $j, i \in \{0, 1\}$ and $j \neq i$.

Methods

MCMC algorithm for parameter estimation

Hidden influenza activity states and model parameters estimates were obtained numerically using a **bayesian approach via Markov Chain Monte Carlo**. More specifically:

- 1 The influenza activity state sequence \mathbf{s}^T is generated by the *forward filtering -backward sampling* algorithm (Chib 1996), which allows sampling each week state from the states joint distribution;
- 2 $\tau_i, \mu, \beta, \theta_0, \theta_1$ and each i -th row of the Γ^k matrix in the homogeneous model are generated via Gibbs sampler from their full conditional distribution;
- 3 For the **non-homogeneous models** the transition probabilities matrix parameters $\alpha_{0,1}$ and $\alpha_{1,0}$ are sampled using a **random-walk Metropolis-Hastings algorithm**.

Methods

Nowcasting weekly influenza activity states and ILI rates

Consider that **we want to nowcast week T , on Friday**, using previous information until week $T - 1$ and the early calculation of the ILI rate of week T .

The estimated **probability that week T belongs to the epidemic** influenza activity state is:

$$\hat{P}[S_T = 1] = \sum_{k=1}^K \frac{\hat{P}[S_T = 1 | \Psi^{(k)}]}{K}$$

where $\hat{P}[S_T = 1 | \Psi^{(k)}] = \hat{P}[S_{T-1} = 1 | \Psi^{(k)}] \gamma_{1,1}^{T(k)} + \hat{P}[S_{T-1} = 0 | \Psi^{(k)}] \gamma_{0,1}^{T(k)}$.

The **nowcast of the official ILI rate for week T** :

$$\hat{y}_{T(T+1)} = \sum_{k=1}^K \frac{y_{T(T+1)|s_T^{(k)}=1}^{(k)} \hat{P}[S_T = 1 | \Psi^{(k)}] + y_{T(T+1)|s_T^{(k)}=0}^{(k)} (1 - \hat{P}[S_T = 1 | \Psi^{(k)}])}{K}.$$

Methods

Computation and models comparison

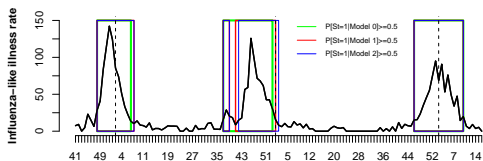
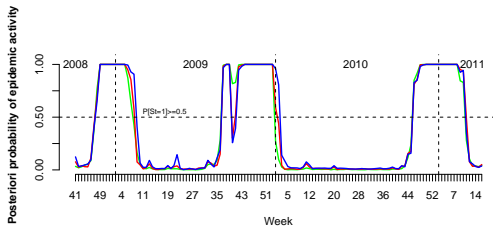
- 1 The MCMC algorithm was run for 200,000 iterations with a burn-in of 50,000 and a thinning of 100, for the non-homogeneous models, and for 100,000 iterations with a burn-in of 25,000 and a thinning of 50, for the homogeneous one;
- 2 The MCMC output convergence was evaluated by the observation of the trace and auto-correlation functions of the parameters runs, and by the application of the statistic of Gelman-Rubin and by the Raftery and Lewis method;
- 3 The Bayes factor was used in order to compare the models fit to data. Marginal likelihoods were computed numerically using the Chib method, for the homogeneous model, and Chib and Jeliazkov for the non homogeneous models;
- 4 All results were obtained using specific programs implemented in the R computing language.

Results

Classification of each week in epidemic or non-epidemic

Panel 1: Posterior mean of the probabilities of epidemic influenza activity (Model 0: green; Model 1: red; Model 2: blue).

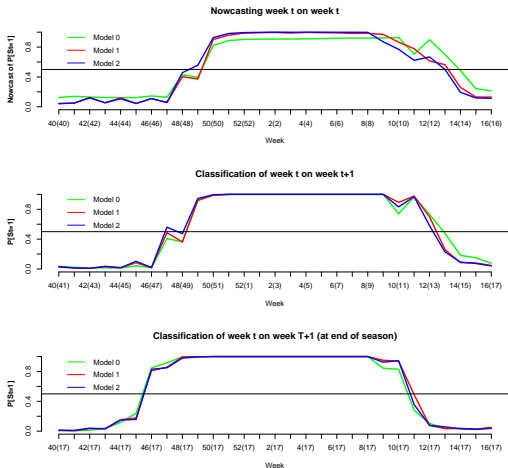
Panel 2 : Influenza-like illness rates, reported by Wednesday (solid line); periods of epidemic activity according to model fitted and probability threshold of influenza epidemic activity (colored boxes)



Results

Real-time nowcast of 2010-2011 season results

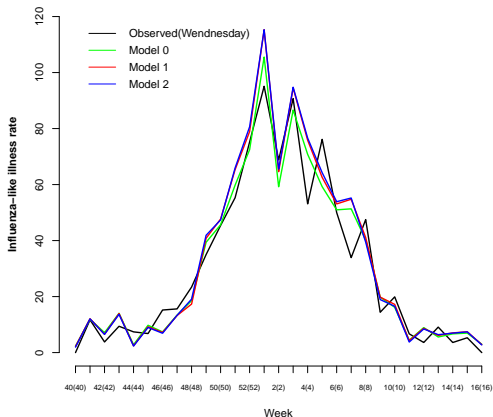
Weekly mean posteriori probabilities of epidemic influenza activity (season 2010-11) calculated, Panel 1: in the current week (nowcast); Panel 2: in the following week; Panel 3: at end of the season. (.) week of the calculus.



Results

Real-time nowcast of 2010-2011 season results

ILI rate nowcast for season 2010-11. (.) week of the calculus.



Main achievements and discussion

- The non-homogenous HMM presents the best fit and is able to identify four epidemic waves in the studied period, one in 2008-09 and 2010-11 and two in the pandemic season;
- The models proposed were able to nowcast the ILI incidence rate and the influenza activity state by Friday of the same week - reducing reporting delay in 5 days;
- The non-homogeneous Model 2 was able to nowcast the epidemic state two weeks before the homogeneous one;
- At the end of the epidemic the probability of being in the epidemic state, obtained by the non-homogeneous model, decreased more rapidly;

Main achievements and discussion

- **Non-homogeneous models are more demanding** in terms of computational time - does not influence the timeliness of the nowcast;
- Results were obtained with the available three seasons 2008-2009 to 2010-2011 - **application to a higher number of seasons is necessary**;

Conclusions and further developments

- This work showed the advantage of using non-homogeneous HMM to nowcast the ILI incidence rate and the influenza activity state in the context of a public health surveillance system;
- This advantage is achieved because the non-homogeneous HMM enables the inclusion of covariates to model the influenza activity state-transition probabilities;
- It was also demonstrated that the incomplete information collected by the Portuguese GP sentinel system, by Friday of the current week, enables the early estimation of the ILI rate and detection of the epidemic start and end;
- Further applications of these models to a higher number of seasons and also to other public health surveillance systems are suggested in order to consolidate the observed added value of this approach.

Thank you for your attention!