ADDRESSING THE PROBLEM OF LACK OF REPRESENTATIVENESS ON SYNDROMIC SURVEILLANCE SCHEMES

Isabel Natário
CEAUL; Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
email: icn@fct.unl.pt

M. Lucília Carvalho
CEAUL; Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, Portugal

M. Gabriela Gomes
Instituto Gulbenkian de Ciência, Oeiras, Portugal

Abstract:

A major concern with some contagious diseases has recently led to an enormous effort to monitor population health status by several different means.

However, some of the implemented monitoring schemes are only able to follow a non-random sample of individuals from the population, by relying on volunteer participation as is the case of GRIPNET, for example, for monitoring symptoms of influenza like illness (ILI) in Portugal.

This work presents a modeling approach to overcome this poor data characteristic, allowing its use for the estimation of the true population disease picture. We use a state space model, where we run two processes in parallel - a process describing the non observable states of the population concerning the presence/absence of disease, and an observational process resulting from the monitoring.

We then use resampling importance sampling estimation techniques, in a Bayesian framework, which enables us to estimate the population states and, thus, the corresponding disease incidence curves.

Keywords: Syndromic surveillance, state space models, importance sampling.

2000 Mathematics Subject Classification: 62P10, 60G35, 93E11

1 Introduction

A sudden outbreak of some contagious diseases as influenza or SARS (severe acute respiratory syndrome), for example, can be very society disturbing, being essencial its timely and effective detection in order to contain it. The awareness of this has led to an huge public and private investment on the monitoring of population health state, frequently based on symptoms rather than on confirmed diagnosis - syndromic surveillance. These systems are set to sound an alarm, that must be further investigated, but they have already proven to be successful in spotting clusters of cases of disease[1].

Some examples of well established syndromic surveillance systems are BioSTORM[2], Biological Spatio-Temporal Outbreak Reasoning Module, AEGIS[3], Automated Epidemiological Geotemporal Integrated Surveillance, ESSENCE II[4], Electronic Surveillance System for the Early Notification of Community-Based Epidemics, HealthMap[5], Rede médicos sentinela do Instituto Dr. Ricardo Jorge, Gripenet[6].

Several of the implemented monitoring schemes have often to trade off between true and false positives[7] and may suffer from imperfect knowledge data with many causes. One aspect that we are concerned with has to do with those schemes that are only able to follow a non-random sample of the individuals from the population, as a consequence of relying on volunteer participation. This is the case of GRIPNET for monitoring symptoms of influenza like illness (ILI) in Portugal.

Gripenet is a syndromic surveillance scheme set up in Portugal in 2005 for detection of ILI diseases. Participation is volunteer and internet based and people joint to answer weekly questionnaires about symptoms related to ILI diseases, such as presence/absence of sudden fever, nasal congestion, etc.. For someone to represent an ILI case (ILI case definition) he/she should satisfy all of: at least one respiratory symptom (running nose or coughing or sore throat or chest pain), muscle pain or severe headache, temperature greater than 38º Celsius and sudden rise of the fever.

For addressing the problem of the poor and non-random data chara-cte-ris-tic associated to some of the surveillances going on we present an idea that allows the use of these data for estimating the true population disease status, based on a state space approach borrowed from wild animal population dynamics modeling [8], [9], [10]. Monitoring data constitutes an observational process that runs in parallel with the non-observable states of the diseased and non-diseased population, allowing their estimation and, consequently, the estimation of the unknown disease incidence. Estimation is then carried out by resampling importance

sampling, in a Bayesian setting.

Section 2 describes the approach just mentioned and gives details on the estimation, Section 3 describes the results obtained by applying this methodology to Gripenet 2006 data and finally Section 4 concludes.

2 A State Space Models Approach

2.1 State-space models

A state space model can be used when the development over time of a system is determined by an unobserved time series, **the state process**, with which a parallel time-series of observations is associated, **the observation process**, being the relation between the two specified by the state space model itself [11].

Let $\mathbf{n}_t, t = 0, 1, \ldots, T$ denote the state process and $\mathbf{y}_t, t = 1, \ldots, T$ the observation process, completely observable and a function of the state process, either observed with or without error. The population dynamics is described by the development in time of $\mathbf{n}_t$, that is usually modeled to accommodate random variation and that can consist of several subprocesses.

Represent $g(\cdot)$ and $f(\cdot)$ the probability density (or mass) functions (pdf) of the state and the observation process, respectively. The state space model is then described by the initial state, the state process (for which we assume first order Markovian property) and the observation process pdfs:

$$
\begin{aligned}
& g_0(\mathbf{n}_0; \boldsymbol{\theta}) \\
& g_t(\mathbf{n}_t | \mathbf{n}_{t-1}, \ldots, \mathbf{n}_0; \boldsymbol{\theta}) = g_t(\mathbf{n}_t | \mathbf{n}_{t-1}; \boldsymbol{\theta}) \\
& f_t(\mathbf{y}_t | \mathbf{n}_t; \boldsymbol{\theta}),
\end{aligned}
$$

where $\boldsymbol{\theta}$ is a vector of parameters and where we assume that $\mathbf{y}_t$ given $\mathbf{n}_t$ is independent of all other states and observations.

It is possible and frequently desirable to modularize the state process pdf into separate but linked sub-processes pdfs, that succeed and describe the state evolution in time in a way that the input to one pdf is the output of the previous one. Although we gain flexibility like this, the price to pay is an increase in the complexity of the state pdf (that have to be integrated over the sub-processes appropriately) and in the likelihood.

2.2 A state space model for Gripenet

Gripenet was set up in Portugal for doing surveillance on ILI diseases, relying on volunteer participation of persons through internet weekly questionnaires about their symptoms on these - typically about 2000-3000 persons join per year. Naturally that the sample we get is not random and, in order to be able to use of the information it holds and based on it, we propose a state space model for modeling the true Portuguese ILI incidence. The data we use is the 2006 Gripenet answers database[18], during the Influenza epidemic season - from November 2005 to May 2006.

Let the state of the population in each successive week of epidemic season 2005/2006, $\{\mathbf{n}_t,\ t = 0, 1, \ldots, T\}$, be an unobserved vector of the population with and without ILI symptoms in each week. Let further divide these on those under surveillance of Gripenet scheme and those that were not:

$$\mathbf{n}'_t = \left(n_{i,t}(\overline{GN})\ \ n_{ni,t}(\overline{GN})\ \ n_{i,t}(GN)\ \ n_{ni,t}(GN)\right),$$

denoting, for week t, $n_{i,t}(\overline{GN})$ and $n_{ni,t}(\overline{GN})$ the ILI symptomatic and the ILI non-symptomatic population, respectively, under Gripenet surveillance and $n_{i,t}(GN)$ and $n_{ni,t}(GN)$ the same for those that were not surveilled by Gripenet. Note that this might further be divided according to other demographic characteristics such as gender, age, place of residence, known for all Gripenet respondents.

2.2.1 Population state process

Following an approach for modeling wild animal population dynamics[8], [9], [10], we choose the population state process to be a stochastic one, based on the following deterministic general process, $\mathbf{n}_t = P\,\mathbf{n}_{t-1}$, where P is a development projection matrix ([12], p.33), as we can consider the population to be divided into two stages: displaying ILI symptoms or not.

We assume that the births and deaths are negligible during that period, so that the size of the population remains constant equal to $N = n_{i,t}(GN) + n_{i,t}(\overline{GN}) + n_{ni,t}(GN) + n_{ni,t}(\overline{GN})$.

The complexity of the population dynamics can be better captured and modeled by further sub-dividing the population state process into sub-processes that consecutively succeed in time, each of which only depend on the subprocess immediately before. Each subprocess corresponds to a matrix so that the general projection matrix is given by the product of these matrices, with obvious advantages.

Here we assume that the sub-processes happen during consecutive time periods in the same order each week, corresponding to matrix C of population class transition due to become ILI symptomatic or by recovering from an ILI setting, and to matrix GN of being or not under Gripenet surveillance scheme, so that $P = GN\,C$.

The corresponding stochastic formulation is done in terms of conditional expected values of the state process that we assume to be a first order Markov process:

$$E[\mathbf{n}_t|\mathbf{n}_{t-1}] = P\,\mathbf{n}_{t-1},$$

describing P the average effect of a set of stochastic processes, such that:

$$E[\mathbf{n}_t|\mathbf{n}_{t-1}] = GN\,C\,\mathbf{n}_{t-1}.$$

Note that, the second subprocess, from now on denominated just Gripenet, only divides the population into those that are followed by Gripenet scheme and those who are not.

The Markov hypothesis assures that the process is completely defined if we know the state process distribution in a certain time t conditionally on the process in the previous time point, $\mathbf{n}_t \sim H_t[\mathbf{n}_{t-1}]$, which can be further decomposed according to the sub-processes as:

$$\mathbf{u}_t^C \sim \mathbf{H}_t^C[\mathbf{n}_{t-1}] \qquad\qquad \mathbf{n}_t = \mathbf{u}_t^{GN} \sim \mathbf{H}_t^{GN}[\mathbf{u}_t^C]$$

where, for each time t, $\mathbf{u}_t^C$ represents a realization of the state vector after the subprocess of class transition and $\mathbf{u}_t^{GN}$ represents the state vector after the subprocess Gripenet.

For the subprocess of class transition, let C_i denotes the probability of a healthy person become ILI symptomatic (equal to everybody), with binomial distribution for the number of persons that become symptomatic,

$$\mathbf{u}_t^C \sim \mathbf{H}_t^C(\mathbf{n}_{t-1}) : \begin{pmatrix} u_{i,t}^C & = & X[\mathbf{n}_{ni,t-1}] \\ u_{ni,t}^C & = & N - u_{i,t}^C \end{pmatrix},$$

where those who became ILI are

$$X[\mathbf{n}_{ni,t-1}] \sim \text{Binomial}\left(n_{ni,t-1}(GN) + n_{ni,t-1}(\overline{GN})), C_i\right),$$

and $u_{i,t}^C$ represents the total new number of ILI symptomatic persons in the population at time t (what is important for incidence); $u_{ni,t}^C$ represents the non symptomatics and the non-new syntomatics at time t being

given by the total population (constant) minus the new ILI symptomatic persons at time t.

The second subprocess divides the population into those followed by Gripenet and those who are not, allowing different participation probabilities for symptomatic and for the asymptomatic. Being p_i and p_{ni} the probabilities of the ILI and non-ILI population, respectively, entering Gripenet at time t,

$$\mathbf{n}_t = \mathbf{u}_t^{GN} \sim \mathbf{H}_t^{GN}(\mathbf{u}_t^C) : \begin{pmatrix} n_{i,t}(\overline{GN}) & = & u_{i,t}^C - n_{i,t}(GN) \\ n_{ni,t}(\overline{GN}) & = & u_{ni,t}^C - n_{ni,t}(GN) \\ n_{i,t}(GN) & = & Y[u_{i,t}^C] \\ n_{ni,t}(GN) & = & Z[u_{ni,t}^C] \end{pmatrix}.$$

where the number of ILI persons and the number of non-ILI persons in Gripenet at time t are given respectively by:

$$Y[u_{i,t}^C] \sim \text{Binomial}\left(u_{i,t}^C, p_i\right),$$

$$Z[u_{ni,t}^C] \sim \text{Binomial}\left(u_{ni,t}^C, p_{ni}\right).$$

To summarize, the state process pdfs are then given by:

$$g_0(\mathbf{n}_0; \boldsymbol{\theta})$$
$$g_t(\mathbf{n}_t|\mathbf{n}_{t-1}; \boldsymbol{\theta}) = \int_{\mathbf{u}_t^C} g^C(\mathbf{u}_t^C|\mathbf{n}_{t-1}; \boldsymbol{\theta})\; g^{GN}(\mathbf{n}_t|\mathbf{u}_t^C; \boldsymbol{\theta})\; \mathrm{d}\mathbf{u}_t^C,$$

with parameters (C_i, p_i, p_{ni})

2.2.2 Observational process

The observational data on the population can be a deterministic or a stochastic function of the unknown states and a complete realization of this process is here denoted by $\{\mathbf{y}_t, t = 0, 1, \ldots, T\}$.

In Gripenet the observations are the ILI and non-ILI persons being followed there at week t, $\mathbf{y}_t = (y_{i,t}, y_{ni,t})$, which we assume to constitute an independent measure with errors of the population states $n_{i,t}(GN)$ and $n_{ni,t}(GN)$, respectively. As such we have the following error model defining the observation process pdf, $f_t(\mathbf{y}_t|\mathbf{n}_t; \boldsymbol{\theta})$, assuming a constant coefficient of variation:

$$
\begin{aligned}
&y_{i,t}|\mathbf{n}_t \sim N(n_{i,t}(GN), \psi_1^2 n_{i,t}(GN)^2) \\
&y_{ni,t}|\mathbf{n}_t \sim N(n_{ni,t}(GN), \psi_2^2 n_{ni,t}(GN)^2) \\
&\mathbf{y}_t|\mathbf{n}_t \sim N(n_{i,t}(GN), \psi_1^2 n_{i,t}(GN)^2) \times N(n_{ni,t}(GN), \psi_2^2 n_{ni,t}(GN)^2),
\end{aligned}
$$

with parameters (ψ_1, ψ_2).

2.3 Estimation

The natural inference setting here is the Bayesian one. Within that we have further to specify a prior distribution on $\boldsymbol{\theta} = (C_i, p_i, p_{ni}, \psi_1, \psi_2)$, $g_0(\boldsymbol{\theta})$. Thus, a complete specification of the probability distribution for states, observations and parameter vector, including the intermediate states is (with $\mathbf{y}^T = (y_1, \ldots, y_T)$):

$$
\begin{aligned}
&P(\mathbf{n}_0, \mathbf{u}_1^C, \ldots, \mathbf{u}_T^C, \mathbf{u}_1^{GN} = \mathbf{n}_1, \ldots, \mathbf{u}_T^{GN} = \mathbf{n}_T, \mathbf{y}^T, \boldsymbol{\theta}) = \\
&g_0(\boldsymbol{\theta}) \times g_0(\mathbf{n}_0|\boldsymbol{\theta}) \times \prod_{t=1}^{T} \left\{ f_t(\mathbf{y}_t|\mathbf{n}_t, \boldsymbol{\theta}) \times g^C(\mathbf{u}_t^C|\mathbf{n}_{t-1}; \boldsymbol{\theta}) \times g^{GN}(\mathbf{n}_t|\mathbf{u}_t^C; \boldsymbol{\theta}) \right\}
\end{aligned}
$$

All the inferences we might be interested in, not only about the parameters given the data but also about the population states in each time point t given all the observations until the previous time point (one-step ahead prediction, [13]) and also including that time point (filtering [13]) as well as the expectation of the states given all the observations and parameters (smoothing) result from integrations over these joint distributions, which are often not trivial.

Consequently, Monte Carlo simulation based approaches are an unavoidable alternative, such as sequential importance sampling methods (that we have chosen) or Markov Chain Monte Carlo Methods, which yields estimates of the likelihood and simulates from the posterior distributions.

Importance sampling is a technique that is used when direct sampling from a target pdf $p(x)$ is not feasible, but we can generate samples from an alternative and easier trial pdf $q(x)$, and then weight them properly to use them as samples from the target distribution. When importance sampling is done using $q(x)$, the resulting sample can be resampled according to conveniently chosen weights to become a sample from $p(x)$, and this is sequential importance sampling, SIS. Sequential importance sampling with resampling, SISR, is a technique where, for space state models,

the generation of the unknown states is carried out combining the two approaches above:

- $q(\mathbf{n}_t) = g_t(\mathbf{n}_t|\mathbf{n}_{t-1})$;
- The weights for the SISR are proportional to the observation pdf, $f(\mathbf{y}_t|\mathbf{n}_t, \theta)$. Resampling results in selecting the states that are the "better" choices according to observations.

For more details see for example [14].

To do estimation we have used the Sequential Importance Sampling with Resampling algorithm proposed by Liu and West in 2001 [15], described next:

1. Simulate a initial parameter and state vectors from prior: particles $\left(\mathbf{n}_0^{[j]}, \theta^{[j]}\right)$, $j = 1, \ldots, J$ (J large).
2. Project each particle forward to the first time period using the state process distribution:
$$\mathbf{n}_1^{[j]} \sim H_1[\mathbf{n}_0^{[j]}, \theta^{[j]}], \quad j = 1, \ldots, J.$$
3. Estimate $g(\mathbf{n}_1|\mathbf{y}^1)$ (filtered state distribution), by using the observation process distribution to calculate a likelihood weight:
$$w_1^{[j]} = \frac{f_1\big(\mathbf{y}_1|\mathbf{n}_1^{[j]}, \theta^{[j]}\big)}{\sum_{i=1}^{J} f_1\big(\mathbf{y}_1|\mathbf{n}_1^{[i]}, \theta^{[i]}\big)}, \quad j = 1, \ldots, J,$$
and then take a weighted resample from the particles.
4. The process is then repeated for subsequent time periods:
   (a) Using the state process distribution to project forward to the next time period;
   (b) Correcting the resulting predicted state distribution using the weighted resample, with weights calculated according to the observation process:
$$w_m^{[j]} = \frac{f_m\big(\mathbf{y}_m|\mathbf{n}_m^{[j]}, \theta^{[j]}\big)}{\sum_{i=1}^{J} f_m\big(\mathbf{y}_m|\mathbf{n}_m^{[i]}, \theta^{[i]}\big)}, \quad j = 1, \ldots, J,$$

This algorithm yields estimates of $\mathbf{n}_t|\mathbf{y}^t$, ($\mathbf{y}^t = (y_1, \ldots, y_t)$) and parameter densities at each time point t. At the last time point T we get an estimate of the posterior density of θ.

To overcome a problem of "particle deplection" (particles with relative large sizes tend to be chosen many times and dominate) we have implemented Kernel smoothing of parameter vectors at each time step, adding a small perturbation to parameter values, increasing the diversity of parameters values in vicinity of parameter space and auxiliary particle filter, an initial "auxiliary" resample is taken from the population at time t, with weights calculated according to the expected likelihood of the states at time $t + 1$, given the data as time $t + 1$. This resampled set of particles is then projected forward from time t to time $t + 1$, and "corrected" using likelihood weights just as with filter, except that the likelihood weights must take account of the auxiliary resampling stage.

2.4 Details for Gripenet model

For the Gripenet model we are not only interested in estimating the number of new ILI cases each week on the population (incidence), from the population state space, given the observations, but we are also interested in estimating the model parameters given the data, specially the probabilities associated with being followed by Gripenet and becoming ILI.

The prior parameters we have considered here are detailed in table 1. We have based their choice on prior knowledge about the quantities they refer to, namely two reports produced by the Instituto Nacional de Saúde Dr. Ricardo Jorge [16] and [17], and some available data about Gripenet participation numbers [18].

Parameter	Prior Distribution	Mean	Standard Deviation
C_i	Beta(0.862,11.631)	0.069	0.069
p_i	Beta(0.9996,4997)	0.002	0.002
p_{ni}	Beta(0.9996,4997)	0.002	0.002
ψ_1	Gamma(1,0.1)	0.1	0.1
ψ_2	Gamma(1,0.1)	0.1	0.1

Table 1: Parameter prior distributions considered.

The states were initialized like (being N_i (respectively, N_{ni}) the initial number of ILI (respectively, non-ILI) persons in the population):

$$
\begin{aligned}
\mathbf{n}_0 &= \begin{cases} n_{i,0}(\overline{GN}) = N_i - y_{i,0} \\ n_{ni,0}(\overline{GN}) = N_{ni} - y_{ni,0} \\ n_{i,0}(GN) = y_{i,0} \\ n_{ni,0}(GN) = y_{ni,0} \end{cases} = \\
&= \begin{cases} n_{i,0}(\overline{GN}) = y_{i,0} * \left(\frac{1}{p_i} - 1\right) & \leftarrow \text{From } N_i p_i = y_{i,0} \\ n_{ni,0}(\overline{GN}) = y_{ni,0} * \left(\frac{1}{p_{ni}} - 1\right) & \leftarrow \text{From } N_i p_{ni} = y_{ni,0} \\ n_{i,0}(GN) = y_{i,0} \\ n_{ni,0}(GN) = y_{ni,0} \end{cases}
\end{aligned}
$$

Force $n_{i,0}(\overline{GN}) + n_{ni,0}(\overline{GN}) + n_{i,0}(GN) + n_{ni,0}(GN) \approx 10,000,000$. We have considered an initial number of iterations of $J = 500000$.

3 Results on Gripenet

Figure 3 depicts the estimated population states and also the observations. From the panels related to the states in Gripenet, we can see that the estimates are quite nicely mimicking the observation patterns, with some exceptions, more towards the end of the estimating period. As such, and taking this as an indicator of a nice states estimates, we are able, by adding the estimates related to ILI diseased persons, to get an estimate of the ILI incidence curve. The not so good estimation we got at the end of the time period considered is probably related to the particle deplection problem mention before, that might have not be completely solved here.

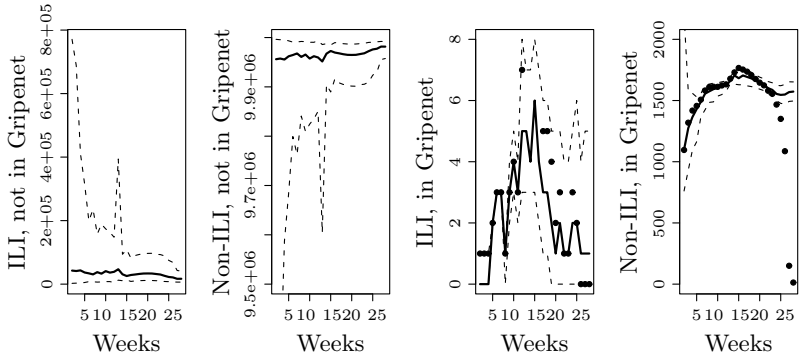

Figure 1: State estimates - quartiles - and observations (dots) - from left to right, ILI and non-ILI not under Gripenet surveillance and ILI and non-ILI under Gripenet surveillance.

4 Discussion

In this paper we have proposed a method to overcome the lack of representativeness of data that so often appears in syndromic surveillance schemes. These data, although it can not be analyzed with the usual statistical techniques to infer for the rest of the population, is valuable and our proposal allows the desirable extrapolation.

For the particular case of ILI diseases and Gripenet example, the estimated number of persons in each population class enables us to estimate ILI incidence curve and further to compare it to other curves obtained from different surveillance schemes as, for example in this case, with the sentinel doctors network.

Of course that this methodology has its weaknesses, namely the computer intensive effort that is associated with its estimation, the necessity of being able to get informative priors to improve estimation, the fact that the estimation method used here starts failing if the total times points T begins to be too large [10]. Further more, there are some assumptions that we have to make to simplify the modeling procedures, namely related to the time points of class transitions, ordering of the sub-processes, time of occurrence of the observations in relation to the state space flow.

Nonetheless we believe on its value as it overcomes in an elegant way the problem of how to analyze this data, how can this data be used in an useful way, as it holds a great potencial for the description of the ILI disease panorama.

On going work related to this application comprehends the inclusion of demographic characteristics in the population states definition - Gripenet respondents fill in a previous questionnaire for these, namely age and gender, which complicates greatly the estimation problem, proportional to the parameter space and the state space grow in complexity.

5 Acknowledgements

IN was supported by project "Reinfection thresholds and the management of recurrent infections", a Marie Curie Excellence Grant, during part of this research.

References

[1] DL Cooper, GE Smith, M Regan, S Large, PP Groenewegen (2008). Tracking the spatial diffusion on influenza and norovirus using tele-

health data: a spatiotemporal analysis of syndromic data. *BMC Medicine*. **6**:16 doi:10.1186/1741-7015-6-16.

[2] MJ O'Connor, D Buckeridge, MK Choy, M Crubezy, Z Pincus, MA Musen (2003). *BioSTORM: A System for Automated Surveillance of Diverse Data Sources*. AMIA Annual Symposium Proceedings.

[3] BY Reis, C Kirby, LE Hadden, K Olson, AJ McMurry, JB Daniel, KD Mandl (2007). AEGIS: A robust and scalable real-time public health surveillance system. *Journal of the American Medical Informatics Association*. **14**: 581-588

[4] J Lombardo, H Burkom, E Elbert, S Magruder, SH Lewis, W Loschen, J Sari, C Sniegoski, R Wojcik, J Pavlin (2003). A Systems Overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). *J Urban Health*, **80**(2 Suppl 1):i32-i42.

[5] JS Brownstein, CC Freifeld, BY Reis, KD Mandl (2008). Surveillance Sans Frontières: Internet-Based Emerging Infectious Disease Intelligence and the HealthMap Project. *PLoS Med* **5**(7): e151.

[6] SP van Noort, M Muehlen, H Rebelo de Andrade, C Koppeschaar, JM Lima Lourenço, MG Gomes, MG (2007). Gripenet: an internet-based system to monitor influenza-like illness uniformly across Europe. *Euro Surveill.*, **12**(7):pii=722.

[7] MA Stoto, M Schonlau, LT Mariano (2004). Syndromic surveillance: it is worth the effort? *Chance*, **17**:19-24.

[8] ST Buckland, KB Newman, L Thomas, NB Koesters (2004). State-space models for the dynamics of wild animal populations. *Ecological Modeling*, **171**: 157-175.

[9] L Thomas, ST Buckland, KB Newman, J Harwood (2005). A unified framework for modelling wild population dynamics. *Australian New Zealand Journal Statistics*, **47**:19-34.

[10] KB Newman, ST Buckland, ST Lindley, L Thomas, C Fernández (2006). Hidden process models for animal population dynamics. *Ecological Applications*, **16**: 74-86.

[11] J Durbin, SJ Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

[12] Caswell, H. (2001). *Matrix Population Models - 2nd Edition*. Sinauer Associates, Inc. Publishers.

[13] M West, J Harrison (1997). *Bayesian forecasting and dynamic models* - 2nd edition. Springer.

[14] A Doucet AM Johansen (2009). *A Tutorial on Particle Filtering and Smoothing: Fifteen years Later*. In Handbook of Nonlinear Filtering, eds D Crisan, B Rozovsky. Oxford University Press.

[15] J Liu, M West (2001). *Combining parameter and state estimation in simulation-based filtering*. In sequential Monte Carlo Methods in Practice, eds A Doucet, N Freitas, N Gordon. New-York: Springer-Verlag.

[16] Departamento de Epidemiologia do INSA (2007). *Gripe 2007 - um estudo sobre comportamentos face à "gripe" - relatório*. Instituto Nacional de Saúde Dr. Ricardo Jorge.

[17] Departamento de Epidemiologia do INSA (2009). *Médicos Sentinela, o que se fez em 2007 - relatório de actividades 21*. Instituto Nacional de Saúde Dr. Ricardo Jorge.

[18] `http://www.gripenet.pt/`