FluDetWeb: an interactive web-based system for the early detection of the onset of influenza epidemics.

David V. Conesa Guillén

Grup d’Estadística Espacial i Temporal en Epidemiologia i Medi Ambient

Universitat de València

joint work with A. López-Quílez, M.A. Martínez-Beneito and F. Verdejo
1 Introduction
**Introduction**

- **Surveillance systems** have become increasingly important due to new challenging public health issues.
- Early detection of the onset of a disease outbreak is one of the most challenging objectives of epidemiological surveillance systems.
- Detecting the onset of an outbreak as soon as possible would imply timely interventions which could suppose a great impact on the number of lives saved.
- Among other diseases, influenza is of special interest: Influenza epidemics occur virtually every year and result in substantial disease, death and expense.
Introduction

- **Surveillance systems** have become increasingly important due to new challenging public health issues.
- **Early detection of the onset of a disease outbreak** is one of the most challenging objectives of epidemiological surveillance systems.
- Detecting the onset of an outbreak as soon as possible would imply timely interventions which could suppose a great impact on the number of lives saved.
- Among other diseases, **influenza** is of special interest: Influenza epidemics occur virtually every year and result in substantial disease, death and expense.
Introduction

- **Surveillance systems** have become increasingly important due to new challenging public health issues.
- Early detection of the onset of a disease outbreak is one of the most challenging objectives of epidemiological surveillance systems.
- Detecting the onset of an outbreak as soon as possible would imply timely interventions which could suppose a great impact on the number of lives saved.
- Among other diseases, influenza is of special interest: Influenza epidemics occur virtually every year and result in substantial disease, death and expense.
Introduction

- **Surveillance systems** have become increasingly important due to new challenging public health issues.

- **Early detection of the onset of a disease outbreak** is one of the most challenging objectives of epidemiological surveillance systems.

- Detecting the onset of an outbreak as soon as possible would imply timely interventions which could suppose a great impact on the number of lives saved.

- Among other diseases, **influenza** is of special interest: Influenza epidemics occur virtually every year and result in substantial disease, death and expense.
Several surveillance methods to track influenza virus activity:

- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency room visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Several surveillance methods to track influenza virus activity:

- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Several surveillance methods to track influenza virus activity:
- real-time internet survey, over the counter sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...
... “each method only captures a portion of infections within the community with different timeliness and specificity”.

David Conesa, GEet●E-mail (UV)
Several surveillance methods to track influenza virus activity:
- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Several surveillance methods to track influenza virus activity:

- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.

Several surveillance methods to track influenza virus activity:
- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...
... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Several surveillance methods to track influenza virus activity:
- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Several surveillance methods to track influenza virus activity:
- real-time internet survey, over the count sales,
- prescription pharmaceutical sales, absenteeism,
- syndromic/sentinel surveillance, laboratory test orders,
- hospital admission surveillance,
- emergency rooms visits, pneumonia and influenza mortality, etc.

There is no such thing as the best method ...

... “each method only captures a portion of infections within the community with different timeliness and specificity”.
Different methods result in different **types of influenza surveillance data**: mortality rates, hospital admissions and emergency room visits, school or work place absenteeism rates, number of emergency phone calls, nurse call line data, etc.

- Sentinel networks are surveillance systems, formed with volunteer practitioners, that provide knowledge about any disease in a short space of time.
Different methods result in different types of influenza surveillance data: mortality rates, hospital admissions and emergency room visits, school or work place absenteeism rates, number of emergency phone calls, nurse call line data, etc.

Sentinel networks are surveillance systems, formed with volunteer practitioners, that provide knowledge about any disease in a short space of time.
Different methods result in different types of influenza surveillance data: mortality rates, hospital admissions and emergency room visits, school or work place absenteeism rates, number of emergency phone calls, nurse call line data, etc.

Sentinel networks are surveillance systems, formed with volunteer practitioners, that provide knowledge about any disease in a short space of time.

Weekly influenza incidence rates in Valencia
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:

- describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
- providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:
- describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
- providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:
- describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
- providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:
- describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
- providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:

1. describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
2. providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:

1. describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and
2. providing an enhanced web implementation of that prospective influenza surveillance methodology.
A large variety of statistical algorithms for automated monitoring of influenza surveillance have been proposed.

Some of them use hidden Markov models to segment the time series of influenza into epidemic and non-epidemic phases.

Bayesian studies are not new in the surveillance literature, but in recent years there has been increasing interest in them.

One of the surveillance contexts to which Bayesian methodology is perfectly suited is in quantifying the probability of being in an epidemic phase at any given moment.

Our main interest is twofold:

1. describing a Bayesian Markov switching model to determine the epidemic and non-epidemic periods from influenza surveillance data, and

2. providing an enhanced web implementation of that prospective influenza surveillance methodology.
2 Modeling time series of influenza data
Our model is based on a segmentation of the series of differences into epidemic and non epidemic phases using a hidden Markov model.

Taking into account that:
- in the epidemic phase changes are greater and inter-related and
- the non-epidemic dynamic is characterized by small random changes around zero,

we model the conditional distribution of the differences either:
- as an autoregressive process of order 1 if the system is in the epidemic phase
- or as a Gaussian white noise process if it is not.
Formulation of the model

• Our model is based on a **segmentation of the series of differences** into epidemic and non-epidemic phases using a **hidden Markov model**.

• **Taking into account that:**
  
  • in the epidemic phase **changes are greater and inter-related** and  
  • the non-epidemic dynamic is characterized by small random changes around zero,  

• we model the conditional distribution of the differences either:
  
  • as an autoregressive process of order 1 if the system is in the epidemic phase  
  • or as a Gaussian white noise process if it is not.
**Formulation of the model**

- Our model is based on a **segmentation of the series of differences** into epidemic and non-epidemic phases using a **hidden Markov model**.

- **Taking into account that:**
  - in the epidemic phase **changes are greater and inter-related** and
  - the non-epidemic dynamic is characterized by small random changes around zero,

- we model the conditional distribution of the differences either:
  - as an autoregressive process of order 1 if the system is in the epidemic phase
  - or as a Gaussian white noise process if it is not.
Formulation of the model

- Our model is based on a segmentation of the series of differences into epidemic and non epidemic phases using a hidden Markov model.
- Taking into account that:
  - in the epidemic phase changes are greater and inter-related and
  - the non-epidemic dynamic is characterized by small random changes around zero,
- we model the conditional distribution of the differences either:
  - as an autoregressive process of order 1 if the system is in the epidemic phase
  - or as a Gaussian white noise process if it is not.
Formulation of the model

- Our model is based on a segmentation of the series of differences into epidemic and non epidemic phases using a hidden Markov model.

- Taking into account that:
  - in the epidemic phase changes are greater and inter-related and
  - the non-epidemic dynamic is characterized by small random changes around zero,

- we model the conditional distribution of the differences either:
  - as an autoregressive process of order 1 if the system is in the epidemic phase
  - or as a Gaussian white noise process if it is not.
Formulation of the model

- Our model is based on a segmentation of the series of differences into epidemic and non-epidemic phases using a hidden Markov model.

- Taking into account that:
  - in the epidemic phase changes are greater and inter-related and
  - the non-epidemic dynamic is characterized by small random changes around zero,

- we model the conditional distribution of the differences either:
  - as an autoregressive process of order 1 if the system is in the epidemic phase
  - or as a Gaussian white noise process if it is not.
Our model is based on a segmentation of the series of differences into epidemic and non epidemic phases using a hidden Markov model.

Taking into account that:
- in the epidemic phase changes are greater and inter-related and
- the non-epidemic dynamic is characterized by small random changes around zero,

we model the conditional distribution of the differences either:
- as an autoregressive process of order 1 if the system is in the epidemic phase
- or as a Gaussian white noise process if it is not.
Formulation of the model (2)

Then if $Y_{i,j}$ denotes the difference between two consecutive weeks:

$$
Y_{i,j} | \text{Non-epidemic} \sim \mathcal{N}(0, \sigma^2_{0,j})
$$

$$
Y_{i,j} | \text{Epidemic} \sim \mathcal{N}(\rho Y_{i-1,j}, \sigma^2_{1,j})
$$

The key in the hidden Markov model is that the unobserved variable $Z_{i,j}$ that indicates the system state (epidemic and non-epidemic) follows a two-state Markov chain of order 1 with transition probabilities:

$$
P_{k,l} = P(Z_{i+1,j} = l | Z_{i,j} = k).
$$
Formulation of the model (2)

Then if $Y_{i,j}$ denotes the difference between two consecutive weeks:

$$Y_{i,j} | \text{Non-epidemic} \sim \text{N}(0, \sigma_{0,j}^2)$$
$$Y_{i,j} | \text{Epidemic} \sim \text{N}(\rho Y_{i-1,j}, \sigma_{1,j}^2)$$

The key in the hidden Markov model is that the unobserved variable $Z_{i,j}$ that indicates the system state (epidemic and non-epidemic) follows a two-state Markov chain of order 1 with transition probabilities:

$$P_{k,l} = P(Z_{i+1,j} = l | Z_{i,j} = k).$$
Once the model is determined, we must estimate its parameters.

Our choice: Bayesian paradigm. Requires specification of the prior distributions of each parameter involved in the model:

\[
\begin{align*}
\rho & \sim \text{Unif}(-1, 1) \\
\theta_{\text{low}} &= \lambda[1] \\
\theta_{\text{mid1}} &= \lambda[2] \\
P_{1,1} & \sim \text{Beta}(0.5, 0.5) \\
\theta_{\text{mid2}} &= \lambda[3] \\
P_{0,0} & \sim \text{Beta}(0.5, 0.5) \\
\theta_{\text{sup}} &= \lambda[4] \\
\sigma_{0,j} & \sim \text{Unif}(\theta_{\text{low}}, \theta_{\text{mid1}}) \\
\sigma_{1,j} & \sim \text{Unif}(\theta_{\text{mid2}}, \theta_{\text{sup}})
\end{align*}
\]

where \(\{\lambda[1], \lambda[2], \lambda[3], \lambda[4]\}\) corresponds to the ordered sequence of the variables \(\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}\) which follow the following prior distribution:

\[
\lambda_j \sim \text{Unif}(a, b) \quad j = 1, \ldots, 4
\]

These expressions do not yield analytical estimates. MCMC methods are necessary: our choice WinBUGS.
Once the model is determined, we must estimate its parameters.

Our choice: Bayesian paradigm. Requires specification of the prior distributions of each parameter involved in the model:

\[ \rho \sim \text{Unif}(-1, 1) \]
\[ P_{1,1} \sim \text{Beta}(0.5, 0.5) \]
\[ P_{0,0} \sim \text{Beta}(0.5, 0.5) \]
\[ \sigma_{0,j} \sim \text{Unif}(\theta_{\text{low}}, \theta_{\text{mid}1}) \]
\[ \sigma_{1,j} \sim \text{Unif}(\theta_{\text{mid}2}, \theta_{\text{sup}}) \]

\[ \theta_{\text{low}} = \lambda[1] \]
\[ \theta_{\text{mid}1} = \lambda[2] \]
\[ \theta_{\text{mid}2} = \lambda[3] \]
\[ \theta_{\text{sup}} = \lambda[4] \]

where \( \{\lambda[1], \lambda[2], \lambda[3], \lambda[4]\} \) corresponds to the ordered sequence of the variables \( \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} \) which follow the following prior distribution:

\[ \lambda_j \sim \text{Unif}(a, b) \quad j = 1, \ldots, 4 \]

These expressions do not yield analytical estimates. MCMC methods are necessary: our choice WinBUGS.
Once the model is determined, we must estimate its parameters.

Our choice: Bayesian paradigm. Requires specification of the prior distributions of each parameter involved in the model:

\[
\begin{align*}
\rho & \sim \text{Unif}(-1, 1) \\
P_{1,1} & \sim \text{Beta}(0.5, 0.5) \\
P_{0,0} & \sim \text{Beta}(0.5, 0.5) \\
\sigma_{0,j} & \sim \text{Unif}(\theta_{\text{low}}, \theta_{\text{mid1}}) \\
\sigma_{1,j} & \sim \text{Unif}(\theta_{\text{mid2}}, \theta_{\text{sup}})
\end{align*}
\]

where \(\{\lambda[1], \lambda[2], \lambda[3], \lambda[4]\}\) corresponds to the ordered sequence of the variables \(\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}\) which follow the following prior distribution:

\[
\lambda_j \sim \text{Unif}(a, b) \quad j = 1, \ldots, 4
\]

These expressions do not yield analytical estimates. MCMC methods are necessary: our choice WinBUGS.
3 Fludetweb
Fludetweb: implementation of the method

- We have implemented previous methodology using a **client-server architecture with a web-based client application design**.

- **Available on-line from:**
  http://www.geeitema.org/meviepi/fludetweb/
Fludetweb: implementation of the method

- We have implemented previous methodologically using a client-server architecture with a web-based client application design.

- Available on-line from:
  http://www.ggeeitema.org/meviepi/fludetweb/.
We use a master-slave intranet architecture in order to take advantage of the available computers (with usual statistical software R and WinBUGS installed) in our department.

In particular, there are various computers acting as slaves and connected via intranet with the server, which acts as the master.
We use a **master-slave intranet architecture** in order to take advantage of the available computers (with usual statistical software $R$ and WinBUGS installed) in our department.

In particular, there are **various computers acting as slaves and connected via intranet with the server, which acts as the master.**
The system has been implemented as a **three-tier architecture** by separating its functions into three separate layers.

1. **Presentation layer**;
2. **Business logic tier**; and
3. **Data layer**.
The system has been implemented as a **three-tier architecture** by separating its functions into three separate layers:

1. **Presentation layer**;
2. **Business logic tier**; and
3. **Data layer**.
The system has been implemented as a three-tier architecture by separating its functions into three separate layers.

1. **Presentation layer**;
2. **Business logic tier**; and
3. **Data layer**.
The system has been implemented as a **three-tier architecture** by separating its functions into three separate layers:

1. **Presentation layer**;
2. **Business logic tier**; and
3. **Data layer**.
Users can **introduce and edit their own data** consisting of a series of influenza weekly incidence rates.
After completion of the process, the system returns the probability of being in an epidemic phase and, when the probability is greater than 0.5, it also returns the probability of an increase in the incidence rate during the following week.

It also provides two graphs: one of a comparison of the weekly rates of the last two seasons, and another with the weekly rates of all the seasons indicating those weeks with a posterior probability of being in an epidemic phase greater than 0.5.
After completion of the process, the system returns the probability of being in an epidemic phase and, when the probability is greater than 0.5, it also returns the probability of an increase in the incidence rate during the following week.

It also provides two graphs: one of a comparison of the weekly rates of the last two seasons, and another with the weekly rates of all the seasons indicating those weeks with a posterior probability of being in an epidemic phase greater than 0.5.
4 Conclusions
Conclusions

- The modeling presented can be used for monitoring surveillance data:
  - good results obtained when using data from Sentinel Networks,
  - generates sensitive, specific and timely alerts,
  - can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

- Possible extensions:
  - To include a spatial component,
  - to incorporate information from sources, including public health.

- Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
Conclusions

- The modeling presented can be used for monitoring surveillance data:
  - good results obtained when using data from Sentinel Networks,
  - generates sensitive, specific and timely alerts,
  - can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

- Possible extensions:
  - To include a spatial component,
  - to incorporate information from sources, including public health,

- Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
Conclusions

- The modeling presented can be used for monitoring surveillance data:
  - good results obtained when using data from Sentinel Networks,
  - generates sensitive, specific and timely alerts,
  - can be adapted in order to work with other data and for any other surveillance system in which the activity lasts the whole year.

- Possible extensions:
  - To include a spatial component,
  - To incorporate information from sources, including public health.

- Ease of use and its on-line availability can make FludetWeb a valuable tool for public health practitioners.
The modeling presented can be used for monitoring surveillance data:

- good results obtained when using data from Sentinel Networks,
- generates sensitive, specific and timely alerts,
- can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

Possible extensions:

- To include a spatial component,
- to incorporate information from sources, including public health.

Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
Conclusions

The modeling presented can be used for monitoring surveillance data:
- good results obtained when using data from Sentinel Networks,
- generates sensitive, specific and timely alerts,
- can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

Possible extensions:
- To include a spatial component,
- to incorporate information from sources, including public-health.
- Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
Conclusions

- The modeling presented can be used for monitoring surveillance data:
  - good results obtained when using data from Sentinel Networks,
  - generates sensitive, specific and timely alerts,
  - can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

- Possible extensions:
  - To include a spatial component,
  - to incorporate information from sources, including public-health.

- Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
Conclusions

- The modeling presented can be used for monitoring surveillance data:
  - good results obtained when using data from Sentinel Networks,
  - generates sensitive, specific and timely alerts,
  - can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

- Possible extensions:
  - To include a spatial component,
  - to incorporate information from sources, including public-health.

- Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
The modeling presented can be used for monitoring surveillance data:
- good results obtained when using data from Sentinel Networks,
- generates sensitive, specific and timely alerts,
- can be adapted in order to work with other data and for any other surveillance system in which the activity last the whole year.

Possible extensions:
- To include a spatial component,
- to incorporate information from sources, including public-health.

Ease of use and its on-line availability can make Fludetweb a valuable tool for public health practitioners.
References


