# Using Participatory Web-based Surveillance Data to Improve Seasonal Influenza Forecasting in Italy

Daniela Perrotta
ISI Foundation
Turin, Italy
daniela.perrotta@isi.it

Michele Tizzoni
ISI Foundation
Turin, Italy
michele.tizzoni@isi.it

Daniela Paolotti
ISI Foundation
Turin, Italy
daniela.paolotti@isi.it

## ABSTRACT

Traditional surveillance of seasonal influenza is generally affected by reporting lags of at least one week and by continuous revisions of the numbers initially released. As a consequence, influenza forecasts are often limited by the time required to collect new and accurate data. On the other hand, the availability of novel data streams for disease detection can help in overcoming these issues by capturing an additional surveillance signal that can be used to complement data collected by public health agencies. In this study, we investigate how combining both traditional and participatory Web-based surveillance data can provide accurate predictions for seasonal influenza in real-time fashion. To this aim, we use two data sources available in Italy from two different monitoring systems: traditional surveillance data based on sentinel doctors reports and digital surveillance data deriving from a participatory system that monitors the influenza activity through Internet-based surveys. We integrate such digital component in a linear autoregressive exogenous (ARX) model based on traditional surveillance data and evaluate its predictive ability over the course of four influenza seasons in Italy, from 2012-2013 to 2015-2016, for each of the four weekly time horizons. Our results show that by using data extracted from a Web-based participatory surveillance system, which are usually available one week in advance with respect to traditional surveillance, it is possible to obtain accurate weekly predictions of influenza activity at national level up to four weeks in advance. Compared to a model that is only based on data from sentinel doctors, our approach significantly improves real-time forecasts of influenza activity, by increasing the Pearson's correlation up to 30% and by reducing the Mean Absolute Error up to 43% for the four weekly time horizons.

## Keywords

Influenza; Web-based Surveillance; Participatory Surveillance; Modeling; Forecasts


*WWW 2017,* April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052670

## 1. INTRODUCTION AND MOTIVATION

Seasonal influenza epidemics occur annually during winter in temperate regions, resulting in around 3-5 million cases of severe illness and 250,000-500,000 deaths worldwide each year [4]. Accurate influenza incidence predictions are needed to estimate in advance, rapidly and reliably, the number of influenza cases and to properly prepare for and respond to the unfolding of the influenza epidemics. In developed and developing countries, national syndromic (i.e. based on observed symptoms) surveillance systems monitor levels of influenza-like illness (ILI) cases among the general population by gathering information from physicians, known as sentinel doctors, who record the number of people seeking medical attention and presenting ILI symptoms. These traditional surveillance systems for seasonal influenza are generally affected by reporting lags of at least one week and by continuous revision of the numbers initially released [14]. Thus, influenza forecasts are often limited by the time required to collect new, accurate data.

The pervasive use of digital communication technologies for public health [37] and the increasing community engagement in public health have fostered the birth of surveillance systems based on the possibility for single individuals to monitor and report their own health status through Web-based platforms [46]. These so-called participatory surveillance systems aim at capturing seasonal influenza activity directly from the general population through Internet-based surveys. Participatory surveillance systems for seasonal influenza are currently running in 13 countries worldwide and collect, aggregate and communicate data in real time during the course of every influenza season. Specifically, the systems that are currently online are: Influenzanet, a network of Web platforms running in eleven European countries [1], FluNearYou in the United States [16, 18, 42] and FluTracking in Australia [12, 13, 20, 21, 31]. Participatory surveillance systems have proven to be accurate and reliable for ILI surveillance, as the detected timing and relative intensities of influenza epidemics are consistent with those reported by sentinel doctors [18, 20, 30, 35, 44]. Furthermore, it has been shown that participatory surveillance systems can also provide relevant information to estimate age-specific influenza attack rates [17, 32, 35] and influenza vaccine effectiveness [13, 22], to assess health care usage [43], and to estimate risk factors for ILI [6].

Previous studies have developed a range of model-inference systems to generate real-time influenza forecasts [40], some of them exploiting the availability of large scale Web data from search queries and social media, such as Google [24, 39], Yahoo [36], Twitter [9, 15, 19, 26, 34, 33, 41, 47], Wikipedia [23, 25, 29] and also by including participatory surveillance data in their models [38]. However, the large majority of published studies have focused on forecasting the influenza activity in the United States only. This is mainly due to the availability of advanced Natural Language Processing algorithms for the English language and the easy access to ILI data provided by the Centers for Disease Control and Prevention (CDC), also in the context of the "Forecasting influenza season challenge" [8]. To what extent similar forecasting approaches can be extended to non-native English speaking countries remains largely unexplored, with a few exceptions [47].

In this study, we combine traditional surveillance data and Web-based participatory surveillance data to improve forecasts of seasonal influenza activity in an European country, Italy, for which such a forecasting approach has never been tested. To this aim, we use two influenza activity monitoring systems. The first one is Influnet [5], the national surveillance system for influenza syndromes in Italy, which is coordinated by the Italian National Institute of Health (ISS). The second one is Influweb [2], a Web-based participatory surveillance system, part of the Influenzanet network, that monitors ILI activity in Italy since 2008.

Our approach is based on linear autoregressive models that integrate ILI prevalence for the current week as reported from the Influweb platform to generate predictions in a real-time fashion up to four weeks ahead of the release of Influnet report. We retrospectively evaluate the predictive ability of our forecasting models over the course of four influenza seasons in Italy, from 2012-2013 to 2015-2016, for each of the four weekly time horizons.

# 2. DATASET

In this section, we describe the data collection for the two surveillance systems, highlighting their main characteristics and explaining how they have been used in our analysis.

## 2.1 Influnet

Influenza activity in Italy is officially monitored by the Italian National Institute of Health, "Istituto Superiore di Sanità", through a system called Influnet. ILI incidence data reported by Influnet represent the ground truth for this study. The Influnet system collects data from a network of sentinel General Practitioners (GPs) and compiles a weekly report in which the national and regional incidence rates by age group are published during the winter season, generally from week 42 to week 17 of the calendar year. The system covers about 2% of the Italian population. Influnet data are published with at least one-week lag and typically new reports provide a first estimate of the weekly ILI incidence which is then updated in the following weeks as more data from sentinel GPs are recorded.

In this study, we distinguished between final *revised* reports, i.e. ILI data that are no longer being revised and are available only at the end of the influenza season, and weekly *unrevised* reports, i.e. ILI data that are actually available with one-week lag and subject to weekly revision until the end of the influenza season. We collected the Influnet revised and unrevised reports for five influenza seasons, from 2011-2012 to 2015-2016, from week 47 to week 17. Reports for weeks 2011-52 and 2013-51 were not available. Weekly reports released on week $WW$ of the year $YYYY$ are available at the following URLs: `http://www.iss.it/binary/iflu/cont/Influnet_YYYY_WW.pdf`. Final revised reports are available at the following URLs: `http://www.iss.it/flue/index.php?lang=1&anno=2016&tipo=13`. All reports were accessed and downloaded on September 1, 2016.

## 2.2 Influweb

Influweb collects weekly symptoms reports in Italy with the aid of self-selected volunteers from the general population [44]. Generally, the data collection is active from November to May during each influenza season. Participants, upon registration, are invited to provide some general background information (e.g. postal code of residence, their birth month and year, their vaccination status, their household composition etc.; further details about the background questions are available upon request). Once they are enrolled in the system they are invited by means of a weekly e-mail reminder to report whether or not they experienced respiratory, gastrointestinal and systemic symptoms (see the full list in [11]).

ILI cases among participants are determined by applying the ECDC case definition [3] to the set of symptoms reported by volunteers. Accordingly, ILI is defined as the acute onset (within a few hours) of symptoms, including at least one among fever or feverishness, malaise, headache and myalgia and at least one among cough, sore throat and shortness of breath. The day of symptoms onset determines the day of ILI onset. Weekly ILI incidence is determined by dividing the number of ILI onsets by the number of active participants per week. When retrospectively analyzing these data, inclusion criteria to determine active participants may vary and depend on the specific aim of the study [7, 11, 44, 45]. Here, we consider as active those participants who filled the background survey and at least two symptoms surveys, to avoid sporadic participation [11, 45]. ILI cases detected in the first symptoms survey were not taken into account because of a potential correlation between symptom presence and willingness to join the web-platform [6, 10, 45]. Participants are considered active for two weeks around the week of reporting.

The fact that volunteers are self-selected is a well known issue that usually affects participatory systems and causes the sample to be non-representative of the general population. In the case of the Influenzanet platforms, self-selection biases have been discussed and quantified in a previous work [11] in which all the relevant aspects for epidemiological analyses, including geography, mobility, demographic, socio-economic and health indicators, have been extensively explored and compared with national statistics. More recently, the representativeness of the users sample has been thoroughly investigated by restricting the analysis only to the Italian Web-platform [35]. In particular, the representativeness of the sample has been examined in terms of age, gender and geographic distribution with respect to the general Italian population and to the Influnet population sample during three influenza seasons, from 2012-2013 to 2014-2015. Results have shown that, despite the existing participation biases, the epidemiological signal extracted from Influweb correlates well with the ILI signal detected by the Influnet sentinel network, capturing both the timing and the relative intensity of seasonal flu activity in Italy. In addition, differently from Influnet, Influweb data are available in real-time and the system can detect ILI cases from the general population and not only from medically attended patients.

Here, we have used Influweb data from five influenza seasons, from 2011-2012 to 2015-2016. The data collection period generally starts at the beginning of the influenza season in November, through a first e-mail reminder to the participants already enrolled in the system and a number of press releases to attract new volunteers among the general population. Thus, the system needs some weeks to reach a stable cohort of participants and early data are usually noisy and must be discarded to avoid biases due to a variable sample of reporting participants. To address this issue, for each influenza season we have selected a starting week according to a threshold on the number of active participants of at least 25% of the total number of participants in the previous season. Accordingly, the influenza seasons under study start on weeks 2011-47, 2012-48, 2013-47, 2014-47 and 2015-47. We have also tested different thresholds and 25% has been found as the optimal value to remove the initial noise and obtain the largest number of data points at the same time, for all the seasons under study.

In Table 1, we report the number of registered users, the average number of weekly active users and the total number of symptoms surveys for all influenza seasons considered in this work from the starting week until the last week of active GP surveillance.

Data collected from the Influweb platform are gathered according to the Italian regulations on privacy which states that only aggregated and anonymized data can be published and shared. Raw data are available upon request from third parties wishing to conduct scientific research and upon discussion with other members of the Influenzanet Scientific Committee. Weekly incidence data aggregated at country level for the current influenza season (2016-2017) are publicly available at the following URLs (in Italian): `http://www.epicentro.iss.it/problemi/influenza/flunews.asp`. Weekly incidence data used in this study are available at the following URLs: `https://www.influweb.it/it/dati/`.

**Table 1: Participation to Influweb during the five influenza seasons under study.**

| Season | Number of registered users | Average number of weekly active users | Total number of symptoms surveys |
|---|---|---|---|
| 2011-2012 | 2,270 | 1,085 | 14,681 |
| 2012-2013 | 3,057 | 1,165 | 15,789 |
| 2013-2014 | 3,513 | 1,244 | 18,471 |
| 2014-2015 | 3,753 | 1,290 | 20,224 |
| 2015-2016 | 4,053 | 1,224 | 20,823 |

## 3. METHODS

In this section we describe three models we have used to produce seasonal influenza forecasts: one model is based on data from GP surveillance reports only, while the other two combine data from both traditional influenza surveillance and Web-based participatory surveillance. We also present evaluation metrics we used to quantify the forecasting accuracy of the three models over the four influenza seasons under study.

Following the same methodology adopted in previous works [27, 34, 38], we considered as our reference a baseline model that considers only ILI data from the traditional GP surveillance Influnet. The baseline model consists of a linear autoregressive model with three weekly lagged components (AR3) as independent variables and takes the form:

$$y_{w+k} = \alpha_1 \widetilde{y}_{w-1} + \alpha_2 \widetilde{y}_{w-2} + \alpha_3 \widetilde{y}_{w-3} \tag{1}$$

where $y_w$ denotes the ILI incidence value at week $w$ and $y$ indicates the value reported by the *revised* reports, while $\widetilde{y}$ indicates the *unrevised* report that is most recent at the time of the forecast.

According to the value of $k$, a distinction can be made between *nowcasting* ($k = 0$), i.e. inferring the present ILI incidence value that Influnet will report in the following week, and *forecasting* ($k > 0$), i.e. predicting the ILI incidence value in $k$ weeks. ILI predictions generated with the baseline model are then contrasted with those produced by the forecasting models that integrate data from participatory surveillance, to assess their added value.

As mentioned above, one advantage of the participatory system Influweb is that data are available in real-time, not just for the previous week as for Influnet. Thus, we used a forecasting model that integrates the real-time ILI signal detected by Influweb into the epidemiological signal of Influnet. First, we produced ILI predictions by including the previous three weeks of ILI incidence reported by Influnet and the Influweb signal for the current week in a linear autoregressive exogenous (ARX) model, as described in [34]. This model, hereafter called $ARX_{1w}$, takes the following form:

$$y_{w+k} = \sum_{i=1}^{3} \alpha_i \widetilde{y}_{w-i} + \gamma_0 \widetilde{z}_w \tag{2}$$

Then, we further extended the ARX model by adding other three weekly lagged terms of the Influweb ILI incidence, (hereafter called $ARX_{4w}$), as follows:

$$y_{w+k} = \sum_{i=1}^{3} \alpha_i \widetilde{y}_{w-i} + \sum_{j=0}^{3} \gamma_j \widetilde{z}_{w-j} \tag{3}$$

In the previous models, the regression coefficients $\alpha_i$ and $\gamma_i$ are estimated separately for different values of $k$ by a least-squares regression.

We considered the 2011-2012 influenza season as a training set for all the models and the subsequent four seasons, from 2012-2013 to 2015-2016, for validation purposes. The goal was to use all available information, at a given point in time, to produce accurate predictions of the ILI incidence one, two, three and four weeks ahead of the release of Influnet reports, effectively predicting ILI three weeks into the future. To this aim, the training set was dynamically increased to include all available information at a given week. In other words, for our first prediction on week 2012-48, the training set included 23 weeks of historical data, for the second prediction on week 2012-49, the training set consisted of 24 weeks, including the values at week 2012-48, and so on for all subsequent weeks.

For each forecasting model we report four different evaluation metrics: Pearson correlation, mean absolute error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE). The definitions of all evaluation metrics are given below, conforming to the following notation: $x_i$ indicates the predicted value at time $i$, $y_i$ corresponds to the ground truth value at time $i$, and, finally, $\overline{x}$ and $\overline{y}$ denote the average of the values $\{x_i\}$ and $\{y_i\}$, respectively.

- The Pearson's Correlation is a measure of the linear dependence between two variables and is defined as:

$$r = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}} \tag{4}$$

- The Mean Absolute Error is a measure of the average of the absolute errors between predicted and true values and is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\, y_i - x_i \,| \tag{5}$$

- The Root Mean Squared Error (RMSE) is a measure of the difference between predicted and true values and is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2} \tag{6}$$

- The Mean Absolute Percent Error (MAPE) is a measure of prediction accuracy of a forecasting method and is defined as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - x_i|}{y_i} \times 100 \tag{7}$$

These similarity metrics were calculated for the time period from the beginning of the 2012-2013 influenza season, week 2012-48, to the end of the 2015-2016 influenza season, week 2016-16.

Moreover, for each forecasting model we evaluate both the timing and magnitude of the peaks in the influenza seasons with respect to the ground truth. In particular, for each value of $k$, we report the number of weeks lag between the predictions and the ground truth and the percent error (PE) around the peak, that is a measure of the discrepancy between predicted and true values and is defined as:

$$PE = \frac{|y_i - x_i|}{y_i} \times 100 \tag{8}$$

The peak analysis is performed separately for each of the four influenza seasons, but for simplicity's sake, we report only the minimum and maximum number of weeks lag and the average together with the minimum and maximum percent error around the peak intensity.

# 4. RESULTS

Figure 1 shows the ILI predictions, while Figure 2 shows the weekly errors associated to each model. Table 2 displays all the values of the evaluation metrics and the peak analysis

**Figure 1: Comparison between the ground truth (black) and the forecasting models: baseline model (blue), $ARX_{1w}$ (orange) and $ARX_{4w}$ (purple), for the four time horizons: a) k=0; b) k=1; c) k=2; d) k=3.**

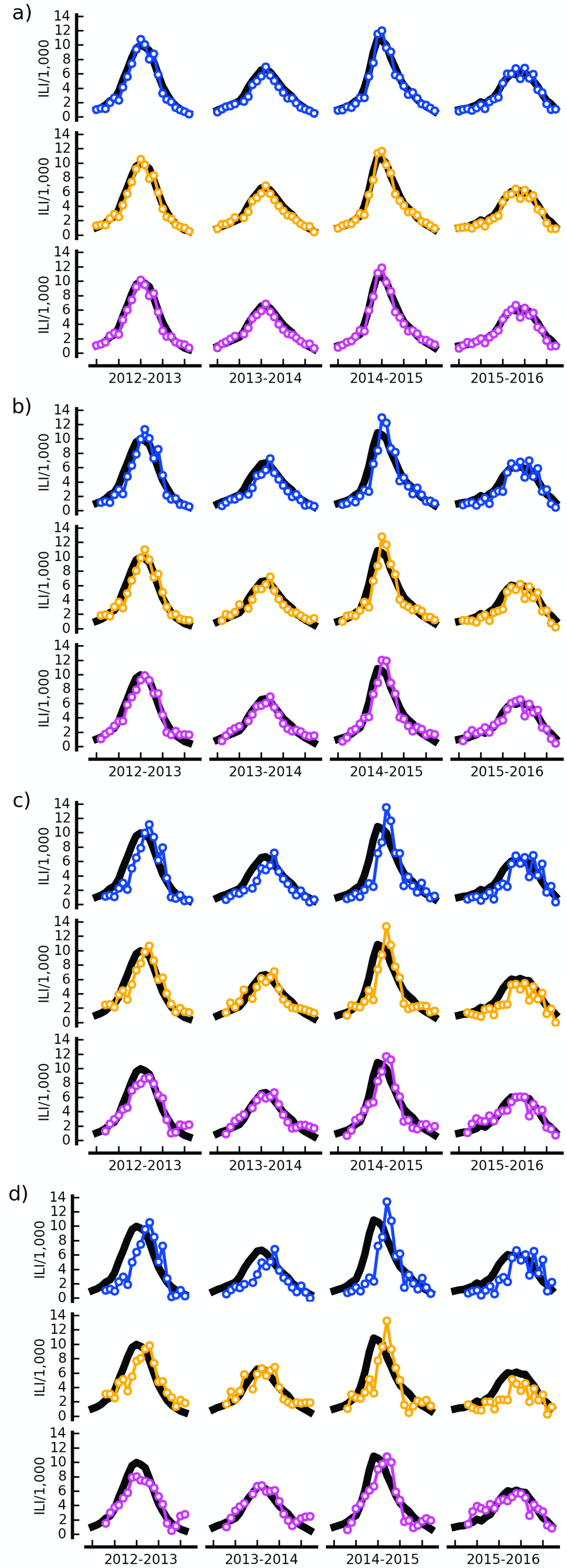


**Table 2: Similarity metrics and peak analysis of the forecasting models with respect to the ground truth. The best performing model per metric is bold faced.**

| | | Similarity Metrics | | | | Peak Analysis | |
|---|---|---|---|---|---|---|---|
| | Model | CORR | MAE (ILI/1,000) | RMSE (ILI/1,000) | MAPE (%) | weeks lag [Min, Max] | PE (%) Mean [Min, Max] |
| $k = 0$ | baseline | 0.981 | 0.45 | 0.5955 | **12.93** | [0, 1] | 8.6 [4.8, 10.9] |
| | $ARX_{1w}$ | 0.983 | 0.42 | 0.5419 | 12.95 | [0, 1] | 5.5 [3.5, 7.7] |
| | $ARX_{4w}$ | **0.984** | **0.39** | **0.5052** | 13.13 | [0, 1] | 5.6 [1.9, 9.0] |
| $k = 1$ | baseline | 0.934 | 0.83 | 1.1174 | **22.39** | [1, 2] | 13.7 [8.9, 19.3] |
| | $ARX_{1w}$ | 0.942 | 0.73 | 0.9635 | 24.31 | [1, 2] | 9.3 [0.8, 18.1] |
| | $ARX_{4w}$ | **0.957** | **0.63** | **0.7952** | 25.63 | [0, 1] | 5.5 [0.7, 10.6] |
| $k = 2$ | baseline | 0.835 | 1.29 | 1.7497 | 33.79 | [2, 3] | 13.8 [7.7, 24.3] |
| | $ARX_{1w}$ | 0.864 | 1.05 | 1.4488 | **32.72** | [2, 3] | 11.7 [6.5, 23.0] |
| | $ARX_{4w}$ | **0.918** | **0.84** | **1.0885** | 34.62 | [0, 2] | 5.2 [0.3, 12.0] |
| $k = 3$ | baseline | 0.677 | 1.80 | 2.4287 | 45.23 | [0, 3] | 9.7 [2.3, 23.2] |
| | $ARX_{1w}$ | 0.751 | 1.42 | 1.9404 | 43.23 | [0, 3] | 10.7 [1.7, 22.1] |
| | $ARX_{4w}$ | **0.879** | **1.02** | **1.3058** | **42.93** | [0, 3] | 7.2 [0.6, 19.7] |

obtained by comparing the baseline model and the forecasting models $ARX_{1w}$ and $ARX_{4w}$ for the four weekly time horizons, corresponding to $k = 0, 1, 2, 3$, against the ground truth. The Pearson's correlation increases with respect to the baseline model of about 0.2% and 0.3%, respectively for the $ARX_{1w}$ and the $ARX_{4w}$ models, for $k = 0$; 0.8% and 2.5% for $k = 1$; 3.4% and 9.9% for $k = 2$; 10.9% and 29.9% for $k = 3$. The MAE decreases with respect to the baseline model of about 6.7% and 13.3%, respectively for the $ARX_{1w}$ and the $ARX_{4w}$ models, for $k = 0$; 12.0% and 24.1% for $k = 1$; 18.6% and 34.9% for $k = 2$; 21.1% and 43.3% for $k = 3$. The root mean squared error (RMSE) decreases with respect to the baseline model of about 9.0% and 15.2%, respectively for the $ARX_{1w}$ and the $ARX_{4w}$ models, for $k = 0$; 13.8% and 28.8% for $k = 1$; 17.2% and 37.8% for $k = 2$; 20.1% and 46.2% for $k = 3$. The mean absolute percentage error (MAPE) increases with respect to the baseline model of about 0.2% and 1.5%, respectively for the $ARX_{1w}$ and the $ARX_{4w}$ models, for $k = 0$; 8.6% and 14.5% for $k = 1$; for $k = 2$, the MAPE decreases of about 3.2% for the $ARX_{1w}$ model, while it increases of about 2.5% for the $ARX_{4w}$ model; finally, for $k = 3$, the MAPE decreases of about 4.4% and 5.1%, respectively for the $ARX_{1w}$ and the $ARX_{4w}$ models.

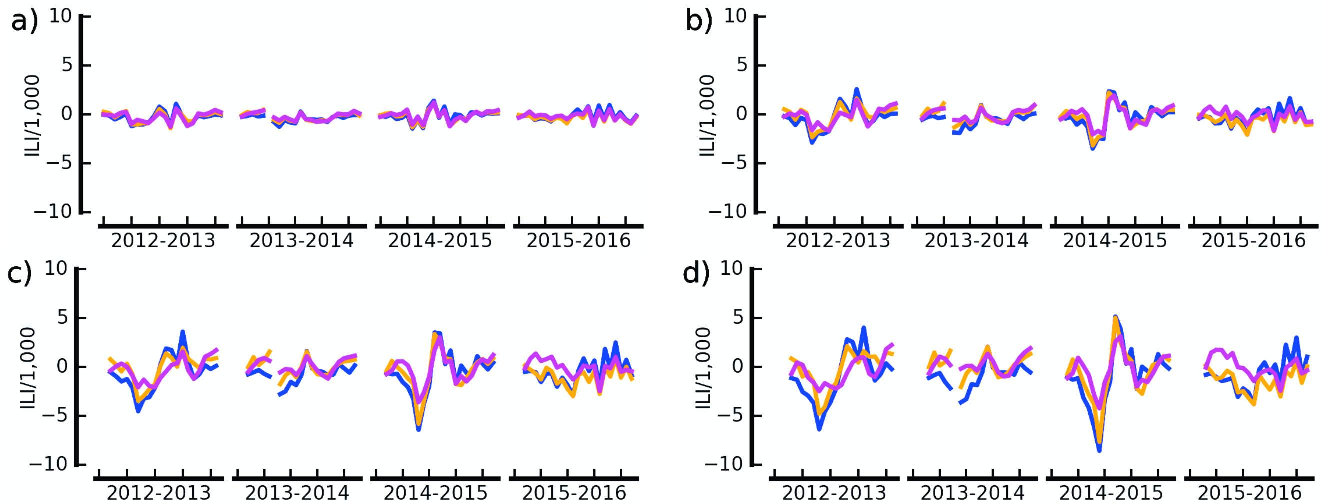
As observed in Figure 1, the predictions curves produced by the baseline model undergo a slight shift forward. This trend is then corrected by incorporating a source of data that integrates the current value of the ILI incidence and contributes to readjust the predictions and partially remove the shift.

In the nowcasting task ($k = 0$) the three forecasting models show a comparable performance in predicting the peak with 0 or 1-week lag. The $ARX_{1w}$ model predicts the peak with the same range of weeks lag with respect to the baseline model, but with a lower percent error around the peak, for all the four weekly time horizons. The $ARX_{4w}$ model predicts the peak with a greater accuracy with respect to the baseline model and the $ARX_{1w}$ model for $k = 1, 2$ and with a lower percent error around the peak, for all the four weekly time horizons. As expected, the ability of the forecasting models to capture the peaks in the influenza seasons decays as the time horizon increases, as shown in Figure 2.

# 5. DISCUSSION

In this study, we have investigated how real-time forecasts of seasonal influenza activity in Italy can be improved by integrating surveillance data from the GP based monitoring system, called Influnet, with data from a Web-based participatory platform, called Influweb. Indeed, ILI incidence estimates produced by traditional surveillance systems such as Influnet undergo weekly revisions during the influenza season as more data from sentinel doctors are collected and official reports are usually released with at least one-week lag. On the other hand, participatory surveillance data such as those produced by Influweb are available as soon as participants report their health status. To produce accurate seasonal influenza forecasts, we have leveraged on this real-time component and built two forecasting models incorporating the weekly lagged terms of the ILI incidence of both surveillance systems in a linear autoregressive exogenous (ARX) model. One is the $ARX_{1w}$ model, which combines incidence data from GP surveillance with a single term of the Influweb incidence for the current week. The second is the $ARX_{4w}$ model which integrates incidence data from traditional surveillance with four weekly lagged terms of the Influweb incidence. We have found that both models outperform the predictions based on GP surveillance data only. The $ARX_{4w}$ model achieves the best performance by predicting the unfolding of the influenza epidemic four weeks in advance and showing the best accuracy in terms of lowest MAE and RMSE and highest correlation with the ground truth across all time horizons. These results highlight the added value provided by the integration of a digital real-time participatory component into seasonal influenza forecasting models. For sensitivity analysis, we also tested the performance of a linear autoregressive model with three weekly lagged components (AR3) based only on Influweb

**Figure 2: Errors associated with each forecasting models, baseline model (blue), $ARX_{1w}$ (orange) and $ARX_{4w}$ (purple), are displayed for the four time horizons: a) k=0; b) k=1; c) k=2; d) k=3.**




ILI incidence. This model resulted to be less accurate than the baseline model (results not presented) showing that, although ILI retrospective estimates tracked by Influweb correlate well with the ground truth data, a model relying only upon Web-based surveillance data is not able to capture the main indicators of the epidemic season.

Our study relies on the assumption that ILI incidence reported by GP surveillance represents the best available measure of influenza activity in Italy. Such assumption may not be completely accurate, as GP surveillance can sometimes over- or underestimate the true burden of the epidemic, depending on what ILI fraction corresponds to real influenza cases and on the reporting rates of sentinel medical doctors [28]. However, sentinel incidence data is considered a highly reliable indicator of influenza activity and adjusting sentinel data for under-ascertainment would require additional information, such as consultancy rates by age groups or reporting rates by GPs, which is usually not available.

The advantages of an integrated framework have been explored also in previous works [29, 34, 38] in which it has been shown that, at least in the United States, the most effective approaches aimed at improving influenza forecasts combine information from multiple flu predictors extracted from various Web data sources such as Twitter, Google Trends, Wikipedia, including data from participatory surveillance. In countries where the penetration of social media is low or Natural Language Processing algorithms do not reach the same accuracy as for the English language and such abundance of different data sources may not be harnessed with equal efficacy, participatory surveillance systems remain not only an important tool to measure the levels of flu activity but also a data source that, combined with the traditional GP surveillance, can be used to provide real-time accurate forecasts.

In the next future, we will extend the present work to include all the countries in which the Influenzanet project has been deployed so far. Another promising extension of our study lies in the integration of other Web sources, such as Twitter or Wikipedia, into our framework. While the majority of the existing studies are focused on the United States and the English language, it would be of interest to understand how much real-time flu forecasts can be improved in countries like Italy, where the penetration of Twitter is more limited, and how this compares to the results obtained by assimilating participatory surveillance data into predictive models.

## 6. ACKNOWLEDGMENTS

We want to thank all the Italian participants who took part to the Influweb data collection. This work has been supported by the EU Cimplex Grant agreement n. 641191 under the H2020 Framework program.

## 7. REFERENCES

[1] https://www.influenzanet.eu.
[2] https://www.influweb.it.
[3] Influenza case definition. available at: http://ecdc.europa.eu/en/healthtopics/influenza/surveillance/pages/influenza_case_definitions.aspx.
[4] Influenza (seasonal), fact sheet number 211. available at: http://www.who.int/mediacentre/factsheets/fs211/en/index.html.
[5] Influnet website. http://www.iss.it/flue/index.php?lang=1.
[6] A. J. Adler, K. T. Eames, S. Funk, and W. J. Edmunds. Incidence and risk factors for influenza-like-illness in the uk: online surveillance using flusurvey. *BMC infectious diseases*, 14(1):1, 2014.
[7] P. Bajardi, A. Vespignani, S. Funk, K. T. Eames, W. J. Edmunds, C. Turbelin, M. Debin, V. Colizza, R. Smallenburg, C. E. Koppeschaar, et al. Determinants of follow-up participation in the internet-based european influenza surveillance platform influenzanet. *Journal of medical Internet research*, 16(3):e78, 2014.
[8] M. Biggerstaff, D. Alper, M. Dredze, S. Fox, I. C.-H. Fung, K. S. Hickmann, B. Lewis, R. Rosenfeld, J. Shaman, M.-H. Tsou, et al. Results from the

centers for disease control and prevention's predict the 2013–2014 influenza season challenge. *BMC Infectious Diseases*, 16(1):357, 2016.

[9] D. A. Broniatowski, M. J. Paul, and M. Dredze. National and local influenza surveillance through twitter: an analysis of the 2012-2013 influenza epidemic. *PloS one*, 8(12):e83672, 2013.

[10] E. Brooks-Pollock, N. Tilston, W. J. Edmunds, and K. T. Eames. Using an online survey of healthcare-seeking behaviour to estimate the magnitude and severity of the 2009 h1n1v influenza epidemic in england. *BMC infectious diseases*, 11(1):1, 2011.

[11] P. Cantarelli, M. Debin, C. Turbelin, C. Poletto, T. Blanchon, A. Falchi, T. Hanslik, I. Bonmarin, D. Levy-Bruhl, A. Micheletti, et al. The representativeness of a european multi-center network for influenza-like-illness participatory surveillance. *BMC public health*, 14(1):1, 2014.

[12] S. J. Carlson, C. B. Dalton, D. N. Durrheim, J. Fejsa, et al. Online flutracking survey of influenza-like illness during pandemic (h1n1) 2009, australia. *Emerg Infect Dis*, 16(12):1960–1962, 2010.

[13] S. J. Carlson, D. N. Durrheim, and C. B. Dalton. Flutracking provides a measure of field influenza vaccine effectiveness, australia, 2007–2009. *Vaccine*, 28(42):6809–6810, 2010.

[14] P. Chakraborty, P. Khadivi, B. Lewis, A. Mahendiran, J. Chen, P. Butler, E. O. Nsoesie, S. R. Mekaru, J. S. Brownstein, M. V. Marathe, et al. Forecasting a moving target: Ensemble models for ili case count predictions. In *SDM*, pages 262–270. SIAM, 2014.

[15] L. Chen, K. T. Hossain, P. Butler, N. Ramakrishnan, and B. A. Prakash. Flu gone viral: Syndromic surveillance of flu on twitter using temporal topic models. In *2014 IEEE International Conference on Data Mining*, pages 755–760. IEEE, 2014.

[16] R. Chunara, S. Aman, M. Smolinski, and J. S. Brownstein. Flu near you: An online self-reported influenza surveillance system in the usa. *Online Journal of Public Health Informatics*, 5(1), 2012.

[17] R. Chunara, E. Goldstein, O. Patterson-Lomba, and J. S. Brownstein. Estimating influenza attack rates in the united states using a participatory cohort. *Scientific reports*, 5:9540, 2015.

[18] A. W. Crawley. Flu near you: Comparing crowdsourced reports of influenza-like illness to the cdc outpatient influenza-like illness surveillance network, october 2012 to march 2014. In *2014 CSTE Annual Conference*. Cste, 2014.

[19] A. Culotta. Towards detecting influenza epidemics by analyzing twitter messages. In *Proceedings of the first workshop on social media analytics*, pages 115–122. ACM, 2010.

[20] C. B. Dalton, S. J. Carlson, M. T. Butler, E. Elvidge, and D. N. Durrheim. Building influenza surveillance pyramids in near real time, australia. *Emerging infectious diseases*, 19(11):1863, 2013.

[21] C. B. Dalton, S. J. Carlson, L. McCallum, M. T. Butler, J. Fejsa, E. Elvidge, and D. N. Durrheim. Flutracking weekly online community survey of influenza-like illness: 2013 and 2014. *Communicable diseases intelligence quarterly report*, 39(3):E361, 2015.

[22] W. J. Edmunds and S. Funk. Using the internet to estimate influenza vaccine effectiveness. *Expert review of vaccines*, 11(9):1027–1029, 2012.

[23] N. Generous, G. Fairchild, A. Deshpande, S. Y. Del Valle, and R. Priedhorsky. Global disease monitoring and forecasting with wikipedia. *PLoS Comput Biol*, 10(11):e1003892, 2014.

[24] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2009.

[25] K. S. Hickmann, G. Fairchild, R. Priedhorsky, N. Generous, J. M. Hyman, A. Deshpande, and S. Y. Del Valle. Forecasting the 2013–2014 influenza season using wikipedia. *PLoS Comput Biol*, 11(5):e1004239, 2015.

[26] V. Lampos and N. Cristianini. Nowcasting events from the social web with statistical learning. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(4):72, 2012.

[27] D. Lazer, R. Kennedy, G. King, and A. Vespignani. The parable of google flu: traps in big data analysis. *Science*, 343(6176):1203–1205, 2014.

[28] S. A. McDonald, A. M. Presanis, D. D. Angelis, W. van der Hoek, M. Hooiveld, G. Donker, and M. E. Kretzschmar. An evidence synthesis approach to estimating the incidence of seasonal influenza in the netherlands. *Influenza and Other Respiratory Viruses*, 8(1):33–41, nov 2013.

[29] D. J. McIver and J. S. Brownstein. Wikipedia usage estimates prevalence of influenza-like illness in the united states in near real-time. *PLoS Comput Biol*, 10(4):e1003581, 2014.

[30] D. Paolotti, A. Carnahan, V. Colizza, K. Eames, J. Edmunds, G. Gomes, C. Koppeschaar, M. Rehn, R. Smallenburg, C. Turbelin, et al. Web-based participatory surveillance of infectious diseases: the influenzanet participatory surveillance experience. *Clinical Microbiology and Infection*, 20(1):17–21, 2014.

[31] A. Parrella, C. B. Dalton, R. Pearce, J. C. Litt, and N. Stocks. Aspren surveillance system for influenza-like illness: A comparison with flutracking and the national notifiable diseases surveillance system. *Australian family physician*, 38(11):932, 2009.

[32] O. Patterson-Lomba, S. Van Noort, B. J. Cowling, J. Wallinga, M. G. M. Gomes, M. Lipsitch, and E. Goldstein. Utilizing syndromic surveillance data for estimating levels of influenza circulation. *American journal of epidemiology*, 179(11):1394–1401, 2014.

[33] M. J. Paul and M. Dredze. You are what you tweet: Analyzing twitter for public health. *ICWSM*, 20:265–272, 2011.

[34] M. J. Paul, M. Dredze, and D. Broniatowski. Twitter improves influenza forecasting. *PLOS Currents Outbreaks*, 2014.

[35] D. Perrotta, A. Bella, C. Rizzo, and D. Paolotti. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in italy. *PloS one*, 12(1):e0169801, 2017.

[36] P. M. Polgreen, Y. Chen, D. M. Pennock, F. D. Nelson, and R. A. Weinstein. Using internet searches for influenza surveillance. *Clinical infectious diseases*, 47(11):1443–1448, 2008.

[37] M. Salathe, L. Bengtsson, T. J. Bodnar, D. D. Brewer, J. S. Brownstein, C. Buckee, E. M. Campbell, C. Cattuto, S. Khandelwal, P. L. Mabry, et al. Digital epidemiology. *PLoS Comput Biol*, 8(7):e1002616, 2012.

[38] M. Santillana, A. T. Nguyen, M. Dredze, M. J. Paul, E. O. Nsoesie, and J. S. Brownstein. Combining search, social media, and traditional data sources to improve influenza surveillance. *PLoS Comput Biol*, 11(10):e1004513, 2015.

[39] S. V. Scarpino, N. B. Dimitrov, and L. A. Meyers. Optimizing provider recruitment for influenza surveillance networks. *PLoS Comput Biol*, 8(4):e1002472, 2012.

[40] J. Shaman, A. Karspeck, W. Yang, J. Tamerius, and M. Lipsitch. Real-time influenza forecasts during the 2012–2013 season. *Nature communications*, 4, 2013.

[41] A. Signorini, A. M. Segre, and P. M. Polgreen. The use of twitter to track levels of disease activity and public concern in the us during the influenza a h1n1 pandemic. *PloS one*, 6(5):e19467, 2011.

[42] M. S. Smolinski, A. W. Crawley, K. Baltrusaitis, R. Chunara, J. M. Olsen, O. Wójcik, M. Santillana, A. Nguyen, and J. S. Brownstein. Flu near you: crowdsourced symptom reporting spanning 2 influenza seasons. *American journal of public health*, 105(10):2124–2130, 2015.

[43] N. L. Tilston, K. T. Eames, D. Paolotti, T. Ealden, and W. J. Edmunds. Internet-based surveillance of influenza-like-illness in the uk during the 2009 h1n1 influenza pandemic. *BMC Public Health*, 10(1):650, 2010.

[44] S. P. van Noort, C. T. Codeço, C. E. Koppeschaar, M. Van Ranst, D. Paolotti, and M. G. M. Gomes. Ten-year performance of influenzanet: Ili time series, risks, vaccine effects, and care-seeking behaviour. *Epidemics*, 13:28–36, 2015.

[45] Y. Vandendijck, C. Faes, and N. Hens. Eight years of the great influenza survey to monitor influenza-like illness in flanders. *PLoS One*, 8(5):e64156, 2013.

[46] O. P. Wójcik, J. S. Brownstein, R. Chunara, and M. A. Johansson. Public health for the people: participatory infectious disease surveillance in the digital age. *Emerging themes in epidemiology*, 11(1):1, 2014.

[47] Q. Zhang, C. Gioannini, D. Paolotti, N. Perra, D. Perrotta, M. Quaggiotto, M. Tizzoni, and A. Vespignani. Social data mining and seasonal influenza forecasts: The fluoutlook platform. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 237–240. Springer, 2015.