https://doi.org/10.20965/jaciii.2021.p0138

Paper:

Forecasting Influenza Based on Autoregressive Moving Average and Holt-Winters Exponential Smoothing Models

Guohun Zhu*1,*2, Liping Li*3,*4, Yuebin Zheng*1, Xiaowei Zhang*3,*4, and Hui Zou*1

*1School of EEE, Guilin University of Electronic Technology
No.1 Jinji Road, Guilin, Guangxi 541004, China
E-mail: g.zhu@uq.edu.au
*2School of ITEE, The University of Queensland
78-626 General Purpose South Building, St Lucia, Queensland 4067, Australia
*3School of Public Health, Shantou University
No.22 Xinling Road, Shantou, Guangdong 515041, China
*4Shantou University Medical College
No.22 Xinling Road, Shantou, Guangdong 515041, China

[Received October 23, 2020; accepted November 30, 2020]

Influenza outbreaks can be effectively prevented if further outbreaks are predicted as early as possible. This article proposes an autoregressive integrated moving average (ARIMA) model and a Holt-Winters exponential smoothing (HWES) model to analyze tweet data for predicting influenza outbreaks and to visualize the number of flu-infection-related tweets with heat maps. First, textual influenza data for Australia from June 2015 to June 2017 are collected through the Twitter Application Programming Interface (API). Next, the ARIMA and HWES models are applied to predict the difference between the flu tweets and confirmations from the Centers for Disease Control and Prevention. Finally, a visualized heat map based on influenza topics validates the modeling analysis in two different time zones. The results show that the average relative error of the ARIMA (HWES) model is 7.25% (11.29%) for the one-week flu forecast.

Keywords: influenza, ARIMA model, exponential smoothing model, Twitter data, heat map

1. Introduction

Influenza, also known as the flu, is caused by influenza viruses and can easily spread to people via mosquito bites. In Australia, there have been 100,494 laboratory-confirmed influenza advisories in just six months in 2019 [1]. Influenza may lead to serious complications and can easily lead to lung infection. In general, influenza can cause pneumonia, which can be life-threatening. Up to half a million people worldwide die annually from the flu [2]. At present, vaccines are effective in preventing and managing seasonal influenza [3–5]. However, as globalization accelerates, human mobility also rises rapidly. Migrations cause infectious influenza viruses to spread quickly from one area to another in different ways [6]. As a result, it is difficult to forecast and prevent influenza if scientists predict influenza outbreak based only on weather conditions or other factors.

Currently, methodologies for influenza prediction are widely classified into three categories: the use of 1) search engines for influenza surveillance, 2) social media for influenza surveillance, or 3) data from the Centers for Disease Control and Prevention (CDCs) or sentinel hospitals to predict influenza conditions. Polgreen et al. [7] used the search volume of Yahoo's search words related to influenza to study its correlation with the flu. Google developed a flu outbreak prediction software, Google Flu, based on its own engine [8]. Although its application is relatively extensive, its accuracy has not met expectations, and real-time monitoring of influenza cases has been overestimated [9–11]. Twitter is another important media resource for the detection of influenza [12–14]. However, these studies only collected tweet data for less than 18 months, indicating that the algorithm has not yet been validated by seasonal flu outbreaks. China has established a national influenza network laboratory for purposes of influenza prediction and prevention [15]. Influenza network data mainly comes from physical hospitals, epidemic prevention stations, and other institutions, and only a small part of the data comes from search engine data. He and Tao [16] used the sentinel to monitor hospitals' monitoring data and conducted an in-depth study on the incidence of influenza cases in Wuhan through the autoregressive integrated moving average (ARIMA) model. Although these efforts are carried out in full swing worldwide, the accuracy of the predictions still do not meet expectations. With the rapid development of science and technology, today's society has gradually entered the era of big data [17]. In this era, because most people usually have only mild cold or flu symptoms and do not choose to go to hospitals immediately; they often choose to post relevant news on social media such as Twitter and Weibo. This makes data from search engines and social media more time-sensitive and realistic than those of the CDC

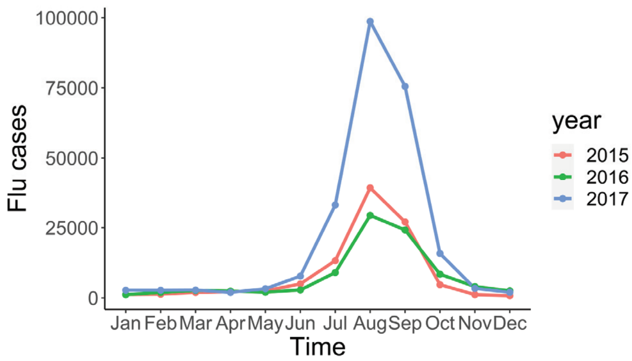


Fig. 1. Australian Ministry of Health 2016–2018 influenza incidence statistics.

and sentinel hospitals, and this is precisely why these data can help the CDC to take effective measures to prevent influenza early.

This study investigates the ARIMA and Holt-Winters exponential smoothing (HWES) models to predict influenza trends in Australia from 2015 to 2017. Both models are suitable for time series analysis in a variety of fields, and the latter includes correlation algorithms for seasonal fluctuation sequences. The Twitter Application Programming Interface (API) is used to capture tweets in real time. Then, Natural Language Processing (NLP) technology is used to filter keywords from the contents of tweets. Finally, based on the extracted keywords and geolocation information, the ARIMA and HWES models are established. The experimental results compare the accuracy of both models and show that ARIMA performs better in predicting influenza trends.

2. Methodology

2.1. Influenza Data Collection and Preprocessing

Located between 10°S–42°S and 113°E–153°E, Australia is the only country in the world that spans the entire continent and crosses both tropical and temperate climate zones. Since Australia is located in the southern hemisphere and China is located in the northern hemisphere, its season is completely opposite to that of the southern hemisphere. Specifically, in a year, autumn is from March to May, winter is from June to August, spring is from September to November, and summer is from December to February. This study obtained 2016–2018 Australian influenza incidence statistics from the Australian Ministry of Health's official website (http://www.health.gov.au/). **Fig. 1**, which reflects influenza incidence in Australia, illustrates clear seasonal patterns, with the highest incidence between late August and early September.

2.2. ARIMA Models

The ARIMA model is one of the most commonly used models in time series analysis [18]. The model mainly predicts the future value of the time series based on the historical and current values in the time series, and it can be influenced by the transformation of other related variables [19]. The ARIMA model must be applied to a stationary time series. For unstable time series data, the d-order differential needs to be converted into a stable time series [18, 20]. The sequence differential conversion formula is:

$$w = \Delta^{\alpha} x = (1-L)^{\alpha} x, \qquad (1)$$

where w is an α-order single-order sequence, and it is a stationary sequence. The ARIMA model prediction formula is as follows:

$$\begin{aligned} y_t &= \mu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} \\ &\quad + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-p}, \end{aligned} \qquad (2)$$

where p is the lag number of the time series data itself, also called the auto-regressive (AR) term; q is the lag number of the prediction error, also called the moving average (MA) item.

2.3. Holt-Winters Exponential Smoothing Model

The HWES model is a kind of exponential smoothing prediction method. It is suitable for time series with increasing or decreasing trend, seasonality, and description by addition model [21].

The HWES method relies on three parameters to estimate the level, slope, and seasonality of the current time point [22].

$$F_{t+m} = (S_t + b_t m) I_{t-L+m}, \qquad (3)$$

where L is the length of the season, and I is the seasonal correction index. The steady, trending, and seasonal equations are:

$$S_t = \alpha \cdot \frac{x_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}), \quad 0 < \alpha < 1, \qquad (4)$$

$$b_t = \beta (S_t - S_{t-1}) + (1-\beta) b_{t-1}, \quad 0 < \beta < 1, \qquad (5)$$

$$I_t = \gamma \cdot \frac{x_t}{S_t} + (1-\gamma) I_{t-L}, \quad 0 < \gamma < 1. \qquad (6)$$

In Eqs. (4)–(6), α is used to estimate the current time level; β is used to estimate the slope of the trend portion of the current time; and γ is used to estimate the seasonal portion of the current time. The values of the three parameters, α, β, and γ are the same. In the interval between 0 and 1, the closer the parameter value is to 0, the smaller the weight of the recent observation value for the future prediction value. The closer the parameter value is to 1, the more recent is the observation, and the observation would have greater weight for future predictions.

2.4. Twitter Data Collections

As a social tool, Twitter is used extensively on Australian social media. In this study, Twitter user data of six administrative states in Australia are crawled by Twitter API [23]. The crawled data have been saved into a mongo database during the specific design experiment. The data include tweet text, time of creation, and the geolocation

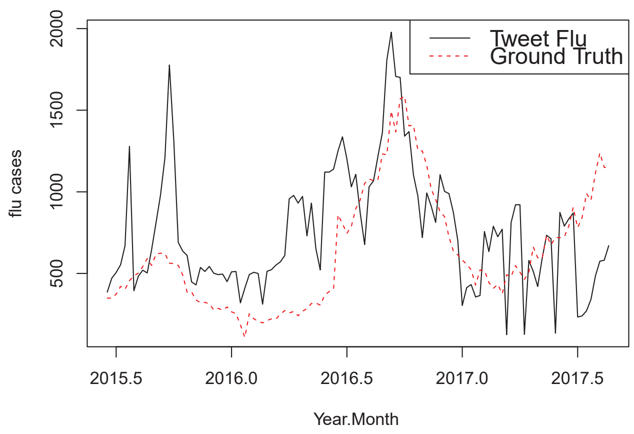


Fig. 2. Influenza-related tweets and ground truth from May 2015 to July 2017.

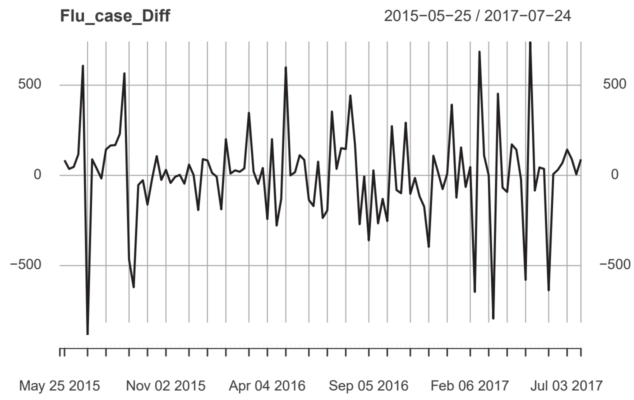


Fig. 3. Sequence first-order difference map of influenza cases in Australia from May 2015 to July 2017.

fields. Then, in the database, keyword extraction of the tweet of the demand time and the demand location is carried out, and statistics on the number of tweets about influenza (discussed in the region) are studied. The crawl tool is implemented by Java. The Twitter data collection source code in Java language can be downloaded from https://github.com/Guohun/CrawlTweets2Mongo.

The number of tweets discussed during the flu episode is measured in days, and there are no missing values. Before the data are used, they are converted into weekly data so that one can observe its time series. After the conversion, the time series in weeks is delayed by one week, as shown in **Fig. 2**. It can be seen that the time series graph in weekly units is a non-stationary sequence with an obvious trend. Moreover, there are obvious seasonal features – the middle of August to early October in 2016 is a period with high incidence of influenza, and this period is positively correlated with the statistical trend of influenza cases from the health department of Australia. This means that social media can capture the flu trends about one or two weeks earlier than ground truth from health agencies.

2.5. Visualization of Flu Distributions

There are many influenza trend predictions that use social media [12, 13, 24, 25]. However, few of them study the influenza epidemic using Google map visualization methods. This section introduces a heat map method to visualize differences in influenza distribution in different periods.

Heat map is a diagram that shows the distribution of human movements and social events in a map using colors as highlights. Unlike the heat map generally used for transport research, this study applies the heat map that shows influenza distributions based on Australian tweets and compares the trends with hospital-confirmed flu reports.

3. Experiments

3.1. Data Preprocessing

The HWES and ARIMA models have different requirements for the smoothness of the processed sequence; specifically, the ARIMA (HWES) model is suitable for a stationary (any) time series. Following the ADF unit root test principle, the ADF unit root test was performed on the sequence of influenza cases in Australia from the first week of 2015 to the 30th week of 2017. The p-value of the test result is 0.0524, a significance level greater than 5%, which indicates that the sequence is a non-stationary sequence. To facilitate the ARIMA model processing, this study used the first-order difference method to process the sequence of influenza cases from the 22nd week of 2015 to the 30th week of 2017. Following the principle of the ADF unit root test, the processed sequence was subjected to the ADF unit root test. The p-value of the student test result is 0.01, less than the 5% significance level. This indicates that the processed sequence is a stationary sequence.

3.2. Determining Model Parameters

For the ARIMA model, it is necessary to determine its parameters according to the time series of the autocorrelation function (ACF), partial autocorrelation function (PACF), and Bayesian information criterion (BIC). The autocorrelation function and the partial autocorrelation function of the first-order difference sequence of the difference between tweets influenza data and ground truth from the 22nd week of 2015 to the 30th week of 2017 are shown in **Fig. 3**. **Fig. 4** shows that the autocorrelations at lags 1, 2, and 5 exceed the significance bounds, and partial correlations at lags 1 and 5 exceed the significance bounds. As a result, the ARIMA optimal model is determined as *ARIMA(1,1,1)*.

To predict influenza using the number of tweets, we combined the number of tweets with influenza incidents from May 2015 to July 2017. First, the trend, seasonality, and random fluctuations were partially decomposed

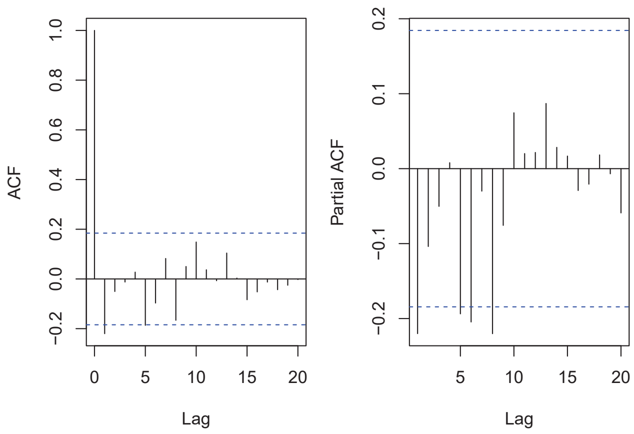


Fig. 4. First-order difference sequence autocorrelation graph and partial autocorrelation graph of tweet influenza and ground truth cases.

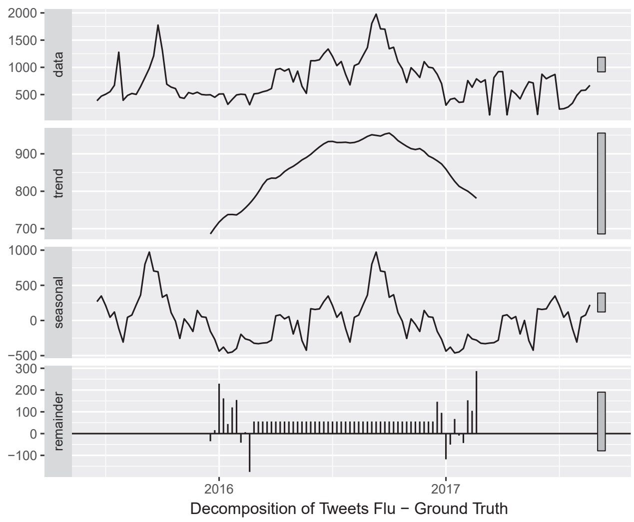


Fig. 5. Decomposition of the difference between tweet influenza and ground truth cases from 2015 to 2017.

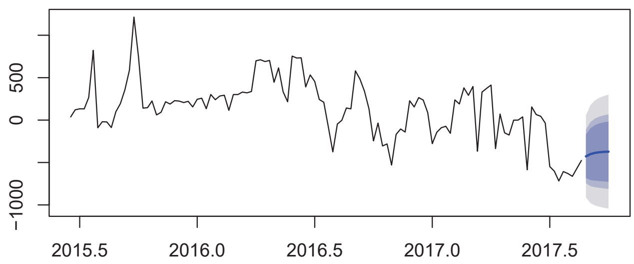
from the combined data to investigate where the differential data are roughly seasonal.

Figure 5 shows (from top to bottom) the differences between tweet flu and ground truth numbers, trends, and seasonal and stochastic fluctuations. A trend is seen in the period from August to September. It is possible to use this ARIMA model and the HWES model to predict influenza in August of 2017 because the seasonal trends (as shown in **Fig. 5**) could be predicted with an adjacent time. Thus, if the differential between flu tweets and ground truth cases can be correctly predicted, the real number of influenza in Australia can be calculated from the predicted differential.

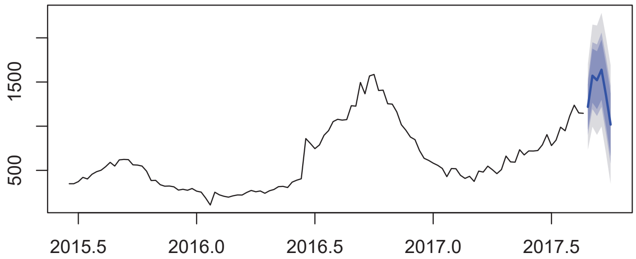
4. Results

4.1. ARIMA Model Prediction Results

The ARIMA model was used to predict the sequence values of influenza cases in Australia from the 31st to the 36th week of 2017. According to Section 3.2, the ACF autocorrelation test is performed on the *ARIMA(1,1,1)* model. The ARIMA model used the difference between ground truth and twitter flu data as input, as shown in **Fig. 6(a)**. The blue shaded curve in **Fig. 6(b)** represents the predicted data for the number of influenza-related tweets from the 31st week of 2017 to the 36th week of 2017. The black curve represents the historical number of influenza-related cases from the 22nd week of 2015 to the 30th week of 2017. The six-week predicted data fits are from 70% to 95% confidence intervals.



(a) Difference between flu tweets and ground truth for predicting flu cases in the 31st to 36th weeks of 2017



(b) Flu prediction based on (a)

Fig. 6. ARIMA prediction results.

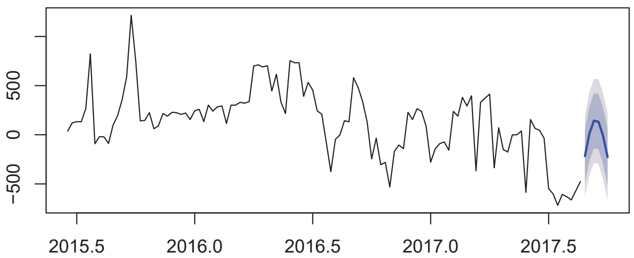
4.2. HoltWinters Exponential Smoothing Model Prediction Results

As shown in **Fig. 2**, the ground truth of flu cases is not really seasonal. This section applies the HWES model to predict the flu cases. Similar to the ARIMA prediction, the difference in number of Flu tweets and ground truth are applied into the HWES model. **Fig. 7** shows the flu trend prediction in 2017 from the 31st to the 36th week.

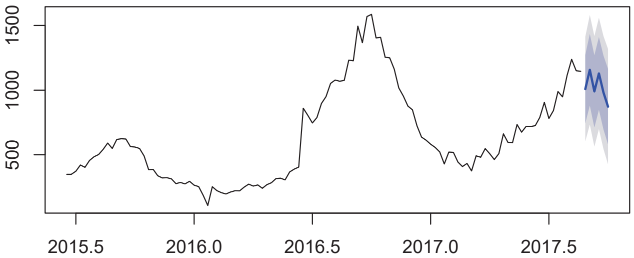
Based on the actual values of influenza cases in Australia from the 31st to the 36th week of 2017, the sequence values predicted by the ARIMA and the HWES models are compared below. As shown in **Table 1**, the root mean square error (RMSE) of the ARIMA model prediction is 435.45 and the average error is 22.71%, which is smaller than those of the HWES model, for which the RMSE is 638.33 and the average error is 31.90%. However, it can be observed that in both models, the error in the forecast data for the first week is lower.

4.3. Visualization of the Flu Distributions

This section examines the heat map from the flu-related tweet distribution in Australia at two different periods.



(a) Difference between flu tweets and ground truth for predicting flu cases in the 31st to 36th weeks of 2017



(b) Flu prediction based on (a)

Fig. 7. HWES prediction results.

Fig. 8. The heat map of Australian tweets on flu topics from 21:24:41 on August 8 to 21, 2016.

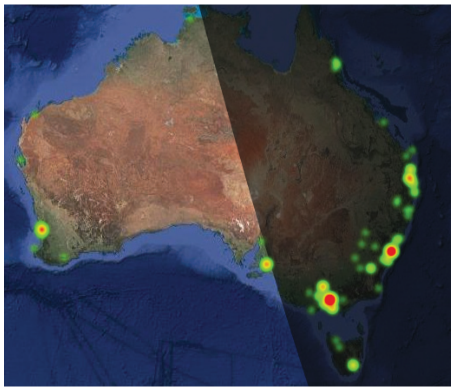


Fig. 9. The heat map of Australian tweets on flu topics at 21:24:41 from October 10 to 23, 2016.

Table 1. Comparison of the six-week prediction accuracy of the ARIMA vs. the HWES model.

Year/ Week	Real flu numbers	ARIMA		HWES	
		Predicted values	Error	Predicted values	Error
2017/31	1136	1218	7.25%	1008	11.29%
2017/32	1088	1571	44.48%	1157	6.38%
2017/33	1653	1520	8.01%	990	40.10%
2017/34	1764	1639	7.04%	1129	35.99%
2017/35	1796	1345	25.08%	982	45.32%
2017/36	1831	1018	44.39%	872	52.35%

Table 2. Comparison of the influenza laboratory surveillance information and tweets flu detection in Australia.

Year	Weeks	Processed number of specimens	Total number of influenza-positive viruses	Number of tweets on flu topics
2015	23–24	856	134	1725
2015	32–33	1546	378	2061
2015	41–42	878	153	1530
2016	23–24	1556	112	2394
2016	32–33	3685	1002	1764
2016	41–42	2417	237	1136

The first one represents summary tweets of the heat map regarding flu topics on August 2016, as shown in **Fig. 8**. The second one corresponds to the same duration using the same flu topics on October 2016, as drawn in **Fig. 9**. The recording time for both is two weeks, using real-time stream tweets API.

Comparing **Figs. 8** and **9**, there are two small hot spots regarding flu tweets in central Australia. Furthermore, the green and red points in August between Sydney and Melbourne (on the right side of the map) are higher than those in October. In other words, the flu trend in tweets shows that the flu outbreak among Australians in August was as mild as that in October 2016.

To check heat map performance, influenza cases reported by the Australian CDC over these four weeks are presented in **Table 2**. The CDC data can be obtained from https://apps.who.int/flumart/Default?ReportNo=12.

Table 2 shows that according to CDC data, the influenza cases during the 32nd to the 33rd week in August 2015 and 2016 were more severe than those in October 2015 and 2016, which agrees with the influenza reports detected in tweets. This also aligns with the ground truth reports in **Fig. 2**.

5. Discussions

Much of the existing research has shown that social media analysis can predict an influenza outbreak [12, 13, 24, 25] by comparing reported CDC-confirmed data. The results in **Tables 1** and **2** also confirm that influenza patterns in tweets may predict trends using the ARIMA model. Furthermore, predictions based on the ARIMA model can reach 22.71% accuracy, which is higher than in Achrekar et al. [24], at 31.8% RMSE for predicting the response to influenza.

Existing methods have been applied to forecast influenza outbreaks using Google Flu Trends [8], tweets [12, 13, 24], clinical pediatric outpatients' data [16], and even collaborative international resources [15]. Google Flu Trends are unreliable in cities, regions, and countries, especially when flu season and media coverage are different [25]. Predictions based on clinical case reports [15, 16] would be more accurate than the ones based on social media. However, responses based on clinical data are delayed because of the difficulty of collecting clinical data in a timely manner. Our research is based on Twitter, which contains a Geo tag on tweets. As such, the proposed methodology could be reliable in cities and regions, compared with Google Flu Trends. In addition, in comparison to two existing influenza predictions on Twitter, the results of our study are based on two-year tweets; this makes it more reliable than studies that used Twitter in other regions [12, 13, 24]. Zuccon et al. [12] studied Australian flu with three-month tweets, which is too short for the seasonal peak to be detected, unless tweets were manually selected. Therefore, our studies are robust and reliable, compared with similar studies of short duration.

Table 2 shows that the weekly influenza-related topics in the heat map tweets are roughly similar to those in the CDC samples. It confirms that influenza-related topics from tweets can effectively and quickly identify influenza outbreaks and trends using natural language processing technology. However, these flu tweets were expressions of Twitter users' feelings, in which most of the users relayed their suspicion of exposure to the flu virus. Therefore, there is a big difference between the tweet statistics and the positive confirmation by the CDC about influenza infection. **Table 2** raises an open-ended issue as well. How many patients diagnosed by the CDC can be obtained through tweets?

6. Conclusions

This study is the first to display Australian influenza distributions in two different time zones using more than two years of tweets. It used differentials in influenza tweets and validated influenza clinical data to predict influenza outbreak. An ARIMA model and a HWES model were applied to analyze differential data.

Training data were collected between week 22 of 2015 and week 30 of 2017. Data from week 36 to week 40 in 2017 (from the end of August to the beginning of September) were used for validation. The trend analyzed was found to be positively correlated with the seasonal trend based on flu statistics from the health sector in Australia. The results show that social media data could rapidly detect influenza outbreaks, as compared with clinical reports. Regarding the accuracy of the one-week influenza prediction in **Table 1**, the mean relative error of the ARIMA and HWES models are 7.25% and 11.29%, respectively. The predicted results in **Tables 1** and **2** and the two visualized heat maps in **Figs. 8** and **9** show that the flu outbreak in August based on tweets was significantly higher than those in October. The major advantage of our study is that our data on tweets covered two years, while most of existing studies covered less than 18 months of tweet data [12, 13, 24]. This implies that our model is robust and reliable because of verification of two influenza seasons.

Acknowledgements

This work was supported by the Guangxi Cloud Computing and Big Data Collaborative Innovation Center (No.YD16E18).

References:

[1] I. Coalition, Influenza Activity Surveillance 2019, https://www.immunisationcoalition.org.au/news-media/2019-influenza-statistics/ [accessed June 24, 2019]

[2] H. Brody, "Influenza," Nature, Vol.573, No.7774, p. S49, 2019.

[3] L. A. Grohskopf, L. Z. Sokolow, S. J. Olsen, J. S. Bresee, K. R. Broder, and R. A. Karron, "Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices – United States, 2015–16 influenza season," Morbidity and Mortality Weekly Report (MMWR), Vol.64, No.30, pp. 818, 2015.

[4] S. Dhakal and S. L. Klein, "Host Factors Impact Vaccine Efficacy: Implications for Seasonal and Universal Influenza Vaccine Programs," J. of Virology, Vol.93, No.21, pp. e00797-19, 2019.

[5] S. V. Vemula, E. E. Sayedahmed, S. Sambhara, and S. K. Mittal, "Vaccine approaches conferring cross-protection against influenza viruses," Expert Review of Vaccines, Vol.16, No.11, pp. 1141-1154, 2017.

[6] V. N. Petrova and C. A. Russell, "The evolution of seasonal influenza viruses," Nature reviews Microbiology Vol.16, No.1, pp. 47-60, 2018.

[7] P. M. Polgreen, Y. Chen, D. M. Pennock, F. D. Nelson, and R. A. Weinstein, "Using internet searches for influenza surveillance," Clinical Infectious Diseases, Vol.47, No.11, pp. 1443-1448, 2008.

[8] D. Lazer, R. Kennedy, G. King, and A. Vespignani, "The parable of Google Flu: traps in big data analysis," Science, Vol.343, No.6176, pp. 1203-1205, 2014.

[9] D. Butler, "When Google got flu wrong," Nature, Vol.494, No.7436, pp. 155-156, 2013.

[10] B. M. Althouse et al., "Enhancing disease surveillance with novel data streams: challenges and opportunities," EPJ Data Science, Vol.4, No.1, p. 17, 2015.

[11] H. Woo, Y. Cho, E. Shim, J.-K. Lee, C.-G. Lee, and S. H. Kim, "Estimating influenza outbreaks using both search engine query data and social media data in South Korea," J. of Medical Internet Research, Vol.18, No.7, p. e177, 2016.

[12] G. Zuccon et al., "Automatic detection of tweets reporting cases of influenza like illnesses in Australia," Health Information Science and Systems, Vol.3, Article No.S4, 2015.

[13] Y. Wang, K. Xu, Y. Kang, H. Wang, F. Wang, and A. Avram, "Regional influenza prediction with sampling twitter data and PDE model," Int. J. of Environmental Research and Public Health, Vol.17, No.3, pp. 678, 2020.

[14] M. J. Paul, M. Dredze, and D. Broniatowski, "Twitter improves influenza forecasting," PLoS Currents, Vol.1, No.6, 2014.

[15] Y. Shu et al., "A ten-year China-US laboratory collaboration: improving response to influenza threats in China and the world, 2004–2014," BMC Public Health, Vol.19, No.3, pp. 520, 2019.

[16] Z. He and H. Tao, "Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study," Int. J. of Infectious Diseases, Vol.74, pp. 61-70, 2018.

[17] J. S. Beckmann and D. Lew, "Reconciling evidence-based medicine and precision medicine in the era of big data: challenges and opportunities," Genome Medicine, Vol.8, No.1, p. 134, 2016.

[18] P. J. Brockwell and R. A. Davis, "Introduction to time series and forecasting," Springer, 2002.

[19] L. Liu, R. S. Luan, F. Yin, X. P. Zhu, and Q. Lü, "Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model," Epidemiology and Infection, Vol.144, No.1, pp. 144-151, 2016.

[20] J. D. Hamilton, "Time Series Analysis," Princeton University Press, 1994.

[21] D. C. Montgomery, L. A. Johnson, and J. S. Gardiner, "Forecasting and Time Series Analysis," 2nd Edition, McGraw-Hill, 1990.

[22] C. Chris and M. Yar, "Holt-Winters Forecasting: Some Practical Issues," The Statistician, Vol.37, No.2, pp. 129-140, 1988.

[23] Twitter, Twitter libraries, https://developer.twitter.com/en/docs/developer-utilities/twitter-libraries.html [accessed May 31, 2019]

[24] H. Achrekar, A. Gandhe, R. Lazarus, S. H. Yu, and B. Liu. "Predicting Flu Trends Using Twitter Data," Proc. of the IEEE Conf. on Computer Communications Workshops, pp. 702-707, 2011.

[25] D. R. Olson et al., "Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales," Plos Computational Biology, Vol.9, No.10, p. e1003256, 2013.

Name:
Guohun Zhu

Affiliation:
School of EEE, Guilin University of Electronic Technology
School of ITEE, University of Queensland

Address:
No.1 Jinji Road, Guilin, Guangxi Province, China
78-626 General Purpose South Building, St Lucia, Queensland 4067, Australia

Brief Biographical History:
2005- Tokyo Institute of Technology
2006- Guilin University of Electronic Technology
2011- University of Southern Queensland
2015- University of Queensland

Main Works:
• Data analysis, machine learning, graph theory, and biomedical engineering

Membership in Academic Societies:
• The Institute of Electrical and Electronics Engineers (IEEE), Member
• Australian Neuroscience Society (ANS)
• The Australian Association of Computational Neuroscientists and Neuromorphic Engineers (NeuroEng)

Name:
Liping Li

Affiliation:
School of Public Health, Shantou University
Shantou University Medical College

Address:
No.22 Xinling Road, Shantou, Guangdong 515000, China

Brief Biographical History:
1992- Shantou University Medical College
2003 Received Ph.D. in Occupational and Environmental Health from School of Public Health, Beijing University
2020- School of Public Health, Shantou University

Main Works:
• Epidemiology and occupational medicine

Membership in Academic Societies:
• The Injury Prevention and Control Section, Chinese Preventive Medicine Association (CPMA), Director

Name:
Yuebin Zheng

Affiliation:
Guilin University of Electronic Technology

Address:
No.1 Jinji Road, Guilin, Guangxi Province, China

Brief Biographical History:
2016-2019 Guilin University of Electronic Technology

Main Works:
• Computer Science and Technology

Name:
Xiaowei Zhang

Affiliation:
School of Public Health, Shantou University
Shantou University Medical College

Address:
No.22 Xinling Road, Shantou, Guangdong 515000, China

Brief Biographical History:
2017 Graduated from Guangdong Medical University with a Major in Statistics
2018- Data Statistical Analyst, Medical Technology Company

Main Works:
• Data analysis and data mining

Name:
Hui Zou

Affiliation:
Guilin University of Electronic Science and technology

Address:
No.1 Jinji Road, Guilin, Guangxi 541004, China

Brief Biographical History:
2016-2020 Guilin University of Electronic Science and Technology

Main Works:
• Computer science, information security