WILEY

DOI: 10.1111/zph.13114

# ORIGINAL ARTICLE

Forecasting seasonal influenza activity in Canada—Comparing seasonal Auto-Regressive integrated moving average and artificial neural network approaches for public health preparedness

Armin Orang<sup>1</sup> | Olaf Berke<sup>1,2</sup> | Zvonimir Poljak<sup>1,2</sup> | Amy L. Greer<sup>1,2</sup> | Erin E. Rees<sup>3</sup> | Victoria Ng<sup>2,4</sup>

<sup>1</sup>Department of Population Medicine, University of Guelph, Guelph, Ontario, Canada

<sup>2</sup>Centre for Public Health and Zoonoses, Ontario Veterinary College, University of Guelph, Guelph, Ontario, Canada

<sup>3</sup>Public Health Risk Sciences Division, National Microbiology Laboratory Branch, Public Health Agency of Canada, Saint-Hyacinthe, Québec, Canada

<sup>4</sup>Public Health Risk Sciences Division, National Microbiology Laboratory Branch, Public Health Agency of Canada, Guelph, Ontario, Canada

## Correspondence

Armin Orang, Department of Population Medicine, University of Guelph, 50 Stone Road E., Guelph, ON, Canada N1G 2W1. Email: aorang@uoguelph.ca

#### **Funding information**

Canadian Safety and Security Program (CSSP); Defence Research and Development Canada (DRDC); Public Health Agency of Canada, Grant/Award Number: CSSP-2018-CP-2334

## Abstract

**Introduction:** Public health preparedness is based on timely and accurate information. Time series forecasting using disease surveillance data is an important aspect of preparedness. This study compared two approaches of time series forecasting: seasonal auto-regressive integrated moving average (SARIMA) modelling and the artificial neural network (ANN) algorithm. The goal was to model weekly seasonal influenza activity in Canada using SARIMA and compares its predictive accuracy, based on root mean square prediction error (RMSE) and mean absolute prediction error (MAE), to that of an ANN.

**Methods:** An initial SARIMA model was fit using automated model selection by minimizing the Akaike information criterion (AIC). Further inspection of the autocorrelation function and partial autocorrelation function led to 'manual' model improvements. ANNs were trained iteratively, using an automated process to minimize the RMSE and MAE.

**Results:** A total of 378, 462 cases of influenza was reported in Canada from the 2010–2011 influenza season to the end of the 2019–2020 influenza season, with an average yearly incidence risk of 20.02 per 100,000 population. Automated SARIMA modelling was the better method in terms of forecasting accuracy (per RMSE and MAE). However, the ANN correctly predicted the peak week of disease incidence while the other models did not.

**Conclusion:** Both the ANN and SARIMA models have shown to be capable tools in forecasting seasonal influenza activity in Canada. It was shown that applying both in tandem is beneficial, SARIMA better forecasted overall incidence while ANN correctly predicted the peak week.

## KEYWORDS

ANN, disease surveillance, SARIMA, seasonal influenza, time series

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

 $\ensuremath{\mathbb{C}}$  2024 The Authors. Zoonoses and Public Health published by Wiley-VCH GmbH.

# <sup>2</sup> WILEY

# 1 | INTRODUCTION

Seasonal influenza is caused by the influenza virus that targets the respiratory system with potentially fatal consequences. Approximately 3500 deaths are recorded annually in Canada due to influenza (Government of Canada, 2019). There are four types of influenza viruses: Influenza A, Influenza B, Influenza C and Influenza D with only three affecting humans: A, B and C (World Health Organization, 2018). The elderly, pregnant women, children under 5 years and individuals with chronic conditions are more susceptible to adverse outcomes due to influenza-related complications (Government of Canada, 2019). This study presents a time series analysis of seasonal influenza incidence. Type C cases are infrequent and typically result in mild illness thus the focus was on type A and type B (World Health Organization, 2018).

Seasonal influenza follows an annual pattern, which peaks in the fall and winter months and begins to decline around mid-spring (Lagacé-Wiens et al., 2010). Despite the seasonal pattern, the magnitude of the seasonal peak is difficult to estimate. The seasonal characteristic of the influenza virus is still poorly understood; however, it has been theorized that one of the primary driving factors is the antigenic variation that occurs within the virus season to season (Lagacé-Wiens et al., 2010; Yaari et al., 2013). Variation refers to the seasonal mutation of genes within the virus, thus making a person susceptible to reinfection with the influenza virus (Lagacé-Wiens et al., 2010). This mutation combined with the possibility of co-circulation of different subtypes may largely determine the expected magnitude and duration of the influenza season as well as the epidemic potential (Lagacé-Wiens et al., 2010).

Time series forecasting can be applied to aid in disease surveillance and serve as part of an early warning system for efficient allocation of public health resources and timely initiation of public health control measures. Time series techniques have been effectively used for analysis and forecasting of seasonal influenza (Choi & Ahn, 2020; Cong et al., 2019; Kraay et al., 2017; Song et al., 2016). For the present study, the focus was on SARIMA (seasonal ARIMA) models and machine learning algorithms using the ANN (artificial neural network) algorithm.

Traditional time series analysis utilizes the SARIMA modelling approach proposed by Box and Jenkins (Box & Jenkins, 1970). SARIMA modelling has been extensively used and reviewed (Hyndman & Athanasopoulos, 2021; Zeger et al., 2006). Cong et al. demonstrated the robustness of SARIMA in infectious disease modelling by accurately predicting seasonal influenza in Mainland China (Cong et al., 2019). Xu et al. (Xu et al., 2017) and Tian et al. (Tian et al., 2019) reported similar findings for forecasting the incidence of mumps in Zibo City, China, and hand, foot and mouth disease in China.

Machine learning algorithms including ANN have been growing in popularity for time series forecasting. Unlike traditional time series forecasting (which is model-based, i.e. a probability distribution), ANNs are data-driven algorithms and learn from data (i.e. extract patterns from data they are trained on) (Zhang

#### Impacts

- Time series analysis indicates a year-to-year growth in seasonal peaks for seasonal influenza activity in Canada from the 2010–2011 influenza season to the 2019–2020 influenza season.
- The study showed an automated SARIMA model had the better forecasting accuracy than a 'manual' built SARIMA model as well as the ANN.
- Although SARIMA had better forecast accuracy in measures of RMSE and MAE, public health may benefit from applying the ANN as well since it correctly predicted the week at which the seasonal influenza incidence peaked.

et al., 1998). Therefore, ANNs require sufficient data (i.e. long time series) to train the algorithm for a given situation. Although, high data volume should not be a replacement for high data veracity or it may introduce the problem of 'Big Data Hubris' (Fuller et al., 2017). ANNs are thus advantageous in situations where underlying relationships in the data are complex and largely unknown. The literature presents several instances where the proficiency of the ANN was compared to the robustness of the SARIMA model in forecasting. Zhang and Qi showed that, assuming data is sufficiently pre-processed, neural networks will outperform the SARIMA model in terms of predictive accuracy (Zhang & Qi, 2005). Furthermore, Zhang et al. produced a comparison study of typhoid fever in China, where the traditional SARIMA model was compared to three separate ANN models (Zhang et al., 2013). Based on a simulation study employing the root mean squared prediction error (RMSE) and the mean absolute prediction error (MAE) as accuracy measures, the three ANNs outperformed the SARIMA modelling approach. Contrary to this, Berke et al. forecast the monthly cryptosporidiosis incidence for Ontario and found no evidence for better performance of the ANN over the SARIMA approach as measured by the MAE and RMSE (Berke et al., 2020).

The goal of this study was to present and compare time series forecasts based on the SARIMA model and ANN algorithm, using weekly seasonal flu incidence risk per 100,000 population for Canada from the start of the 2010–2011 flu season to the end of the 2018–2019 season. The study objectives were to:

- Conduct a descriptive time series analysis of seasonal influenza activity in Canada.
- Conduct a traditional time series analysis using the SARIMA model to study seasonal influenza activity in Canada and forecast future incidence.
- Train an ANN algorithm to forecast seasonal influenza incidence in Canada.
- Compare the forecast from the SARIMA and ANN approaches via their RMSE and MAE to learn about their relative predictive accuracy.

## 2 | MATERIALS AND METHODS

The weekly incidence data are publicly available and were acquired through the global web-based tool *FluNet* from the World Health Organization database (World Health Organization, n.d.). Forecasting was done for the 2019–2020 influenza season. According to the FluWatch calendar from the Public Health Agency of Canada, the influenza season begins on the 35th week (Government of Canada, 2020). Thus, for consistency, this was set as the beginning of every influenza season in the data set. The data set was then split into training data (35th week in 2010 to the 34th week in 2019 [August 25th]) and validation data weekly incidence from the 35th week in 2019 (August 26) to the 34th week in 2020 (August 23) for analysis. Data were adjusted for population growth by factoring in population-at-risk (Statistics Canada, n.d.). All data analysis was done in R 3.6.2 (Team RC, 2019) and RStudio (RStudio Team, 2019). Ethics approval was not required as this study only used publicly available surveillance data.

Data pre-processing was performed to filter and prepare the data for analysis. Two weeks of missing data were identified, week 33 from 2011, and week 41 from year 2012. Using weekly reports from FluWatch (Government of Canada, 2013) the missing data were imputed. For practical reasons, equal observation frequency between years is important. In 2015, there were 53 weeks as opposed to the usual frequency of 52 weeks; therefore, the weekly counts for week 52 and week 53 from 2015 were aggregated and recorded as week 52.

The objectives of time series analysis are generally three-fold: describe (i.e. identify components), filter (i.e. remove noise) and forecast (i.e. predict future observations). Additional objectives include outbreak detection and to estimate intervention effects, but the focus of this study was the former three. The following methods were utilized to achieve those goals. For exploratory and descriptive purposes, the seasonal and trend decomposition using LOESS (STL) method was employed (Cleveland et al., 1990). STL filters the data through the application of the LOESS smoother to reveal the trend, seasonal and remainder components. Trend refers to the long-term changes in the incidence, while the seasonal component refers to cyclical variation. The remaining variation, which is unexplained by trend or seasonality, is the remainder component (Cleveland et al., 1990).

All SARIMA models and ANN algorithms were fit with an automated Box–Cox transformation to normalize the skewed distribution. The forecasts were automatically back-transformed as described by Hyndman (Hyndman & Athanasopoulos, 2021). Back-transformation ensures that results can be interpreted on the original observation scale, which is the incidence risk scale in this study.

# 2.1 | Seasonal autoregressive integrated moving average (SARIMA)

Full notation for the SARIMA model is SARIMA(p, d, q)(P, D, Q)<sub>s</sub>, where p and q are the order of lagged observations and white noise

error terms represented by the auto-regressive (AR) and moving average (MA) elements, respectively, while *d* represents the order for difference filter to remove a trend component. Similarly, *P*, *D* and *Q* represent the orders at the seasonal length *s*, that is, *D* is the order of a difference filter to remove a seasonal component.

For this study, two SARIMA models were built. First, an automated SARIMA model was fit by maximum likelihood estimation (MLE) through an automated model selection process, combining unit root tests and minimizations of the Akaike information criterion (AIC). Differencing order at the trend and seasonal length was determined through the STL decomposition plot. When inspection of the autocorrelation function (ACF) and partial autocorrelation function (PACF) indicated the presence of residual autocorrelation, manual methods were employed to find a better fitting model, that is, the second, manual SARIMA model was fit. The manual model revised the automated model through an iterative process, where subsequent models were fitted with slight adjustments made to parameters. Adjustments were determined from inspecting their ACF and PACF graphs. Residual autocorrelation was further assessed through the Ljung-Box Q test (Ljung & Box, 1978). Lag (k) was set at 20 as suggested by Petris, Petrone and Campagnoli (Petris et al., 2009). Significant residual autocorrelation (p-value <0.05) signalled a lack of fit. As the objective of this study was forecasting accuracy, forecasts from both manual and automated SARIMA models were explored to see if and how much the efforts of manual model optimization improved accuracy. SARIMA residuals are assumed Gaussian white noise, and this was assessed through quantile-quantile plots. If assumption of normality was not met, residuals were bootstrapped for more accurate forecasting intervals (Hyndman & Athanasopoulos, 2021). The selected model as fitted to the 2010-2019 training data was then used to forecast the weekly incidence for the 2019-2020 validation data. The 80% and 95% prediction intervals were based on bootstrapped residuals and visualized together with the point forecasts.

## 2.2 | Artificial neural network

The ANN is a machine learning algorithm, which extends the linear regression model (18) to a non-linear combination of weighted input data until these are summarized into output, that is, a forecast value. In time series analysis, the lagged values of a time series are its predictors for future values, like that of an autoregressive model. Weights are determined through a learning algorithm, which minimize a cost function (e.g., RMSE or MAE) (Hyndman & Athanasopoulos, 2021).

This study implemented a multilayer feed-forward neural network denoted as neural network autoregressive (NNAR). In a feedforward neural network, data is passed only forward, where each layer receives inputs from the previous layer and is then merged through a weighted linear combination. The result is further modified by a linear or non-linear function (such as identity or sigmoid transformations) before being sent to the output layer (Hyndman & Athanasopoulos, 2021). Through this process, an ANN can capture even non-linear behaviour in a time series. Figure 1 shows a simple framework of this procedure.

The NNAR is denoted similar to that of SARIMA models and as follows (Berke et al., 2020): NNAR(p, P, k)<sub>s</sub> where p indicates the number of lagged inputs, P indicates the number of seasonal lags to be used as inputs and k denotes the number of neurons in the hidden layer. Inside the hidden layer, inputs are linearly combined and then passed through an activation function (which is either a linear or non-linear function) to the next neuron or the output, that is, the final forecast. S indicates the seasonal length. The order p was found as the optimal number of lags for a linear autoregressive model according to the AIC. P was held constant at 1 to control for linear seasonality. The initial algorithm assessed p = 7. Successive algorithms were built by incrementally increasing k. The algorithm architecture was chosen through MAE and RMSE. Ensemble forecasts were generated from 100 random starts to provide variations in the training weights for each model. The NNAR algorithm was applied to the 2010-2019 training data and then tested on the 2019-2020 validation data. A prediction interval was based on 1000 bootstrapped sample paths (Hyndman & Athanasopoulos, 2021).

Forecasts from the SARIMA model and the NNAR were compared through the Diebold–Mariano test (Diebold & Mariano, 1995). The Diebold–Mariano test assesses equal predictive accuracy between forecasts by comparing their loss differential (Diebold, 2015). A significant result (*p*-value  $\leq 0.05$ ) indicates that the forecasts are different from one another thus variations in the forecasts are not due to chance alone.

## 3 | RESULTS

There were a total of 378,658 cases of influenza (type A and type B) with an average yearly seasonal incidence of 20.019 cases per 100,000 and 0.385 incidence cases per 100,000 cases per week from 2010 to 2020. The lowest incidence was recorded in the 2011-2012 influenza season with a total of 12,690 reported cases (37.136 incident cases per 100,000 population). The maximum incidence occurred in the 2017-2018 influenza season with a total of 64,250 reported cases (174.568 incident cases per 100,000 population). Week 1 of year 2015 had the highest number of cases in

a week with a total of 5313 reported cases (14.881 incident cases per 100,000 population). The lowest number of incident cases in a week was zero reported cases and was reported eight times. The time series of weekly reported influenza incidences in Canada from the years 2010–2011 to the end of the 2019–2020 influenza season is shown in Figure 2.

The time series indicated a strong seasonal component and slight positive trend in seasonal peaks. The STL decomposition plot (Figure 3) supported these results: the trend component accounted for only a small proportion of the total variation in the data and seasonal peaks developed at regular intervals during the winter months. Thus, a differencing order of 1 was applied at the seasonal length and differencing at the trend level was not needed to achieve stationarity.

The model selected through the automated process was SARIMA(1, 0, 2)(2, 1, 0)<sub>52</sub>. A total of 13 models were additionally fit to the training dataset through the iterative SARIMA model-building process. SARIMA(4, 0, 6)(2, 1, 1)<sub>52</sub> was selected as the best-fitting model for forecasting. Models had a Box–Cox transformation parameter  $\lambda$ =0.12 with SARIMA(4, 0, 6)(2, 1, 1)<sub>52</sub> having a smaller  $\Delta$ AIC (difference in AIC) of 65.71. Quantile–quantile plots indicate a violation of the Gaussian assumption of the residual process as overdispersion of the residuals is apparent in both models. Forecasting residuals were thus bootstrapped for better prediction intervals.

The ACF and PACF from the manual model indicated the presence of some residual autocorrelation, contrary to the Ljung-Box Q test *p*-value = 0.74. For the automated model Ljung-Box test signalled autocorrelation (*p*-value <0.001). The ACF and PACF showed remaining correlation at both the observational (weekly) and seasonal level for both models. Figure 4 displays time series and forecast from SARIMA(4, 0, 6)(2, 1, 1)<sub>52</sub>. Figure 5 displays time series and forecast from the automated SARIMA(1, 0, 2)(2, 1, 0)<sub>52</sub>. The figures present 80% and 95% prediction intervals.

A total of 15 ANNs were trained and an ANN of order NNAR(7, 1, 1)<sub>52</sub> was determined to be the best according to its RMSE and MAE. NNAR(7, 1, 1)<sub>52</sub> indicates an autoregressive order of seven, a seasonal autoregressive order of one and one node within the hidden layer. The last seven observations plus the first seasonal observation, that is,  $y_{t-1}$ ,  $y_{t-2}$ ,  $y_{t-3}$ ,  $y_{t-52}$  were linearly combined to form one node of a single hidden layer. Similar to Figures 4–6 displays the time



**FIGURE 1** Multilayer feed-forward neural network. The inputs (i.e., autoregressive parameters) are passed forward through the network and are then combined through a weighted linear combination. The result is then modified by the activation function (generally the identity or sigmoid function) to the next layer, where again a weighted combination of inputs is transformed by an activation function and passed on to the next layer until the output layer is reached presenting the final prediction.





FIGURE 2 Time series plot of weekly influenza incidence in Canada from the 2010-2011 to the end of the 2019-2020 flu season. The dotted vertical line indicates break between the training data and validation dataset.



FIGURE 3 Weekly influenza incidence in Canada from 2010-2011 to 2018-2019 flu season (top panel) split into its three additive components (trend, seasonality and remainder) through STL decomposition.

series and forecast of NNAR (1, 7) with 80% and 95% prediction intervals.

Table 1 shows summary statistics with the RMSE and MAE from the 2019 to 2020 seasonal forecasts for all stated models. Observed weekly incidences and rounded forecasts for ANN and SARIMA models are also shown in Figure 7. The Diebold-Mariano test was significant (p-value <0.001). Table 2 shows the total observed and forecasted influenza incidence for the 2019-2020 influenza season.



**FIGURE 4** Time series and weekly forecasts including forecasting intervals of the 2019–2020 influenza season in Canada from SARIMA(4, 0, 6)(2, 1, 1)<sub>52</sub>. The blue line indicates the forecasted values with darker and lighter shades of blue to indicate 80% and 95% prediction intervals, respectively.



**FIGURE 5** Time series and weekly forecasts including forecasting intervals of the 2019–2020 influenza season in Canada from SARIMA(1, 0, 2)(2, 1, 0)<sub>52</sub>. The blue line indicates the forecasted values with darker and lighter shades of blue to indicate 80% and 95% prediction intervals, respectively.

Additionally, it presents the predicted number of cases at peak with the week in which this occurred for each.

The automated SARIMA(1, 0, 2)(2, 1, 0)<sub>52</sub> outperformed the manual SARIMA model (SARIMA(4, 0, 6)(2, 1, 1)<sub>52</sub>) with the MAE being smaller by 0.337 incident cases per 100,000 population and the RMSE smaller by 0.661 incident cases per 100,000 population.

The machine learning algorithm NNAR  $(7,1,1)_{52}$  was similarly, outperformed by the automated SARIMA model in predicting



**FIGURE 6** Time series and weekly forecasts including forecast interval of the 2019–2020 influenza season in Canada from an NNAR(7, 1, 1)<sub>52</sub>. The blue line indicates the forecasted values with darker and lighter shades of blue to indicate 80% and 95% prediction intervals, respectively.

TABLE 1 AIC and p-values of the Ljung–Box test for SARIMA models and predictive accuracy measurements of SARIMA models and the ANN.

Model	AIC	Ljung-Box test	RMSE (per 100,000 population)	MAE (per 100,000 population)
SARIMA(1, 0, 2)(2, 1, 0) <sub>52</sub> (automated)	419.81	<0.001	1.700	1.094
SARIMA(4, 0, 6)(2, 1, 1) <sub>52</sub> (manual)	354.11	0.74	2.361	1.431
NNAR(7, 1, 1) <sub>52</sub>	N/A	N/A	2.570	1.674



the weekly incidence of influenza per 100,000 population during the 2019–2020 flu season. The automated SARIMA model had a smaller MAE by 0.654 incident cases per 100,000 population and a RMSE smaller by 0.864 incident cases per 100,000 population. The Diebold–Mariano test indicates the forecasts are different from one another thus variation in forecasts is not due to chance alone.

**TABLE 2** Observed and forecasted incidence risk of the 2019–2020 seasonal flu year as total number of cases, peak influenza cases and week in which peak occurred.

Model	Total incidence risk (per 100,000)	Peak incidence risk (per 100,000)	Week at which peak occurred
Observed	141.792	12.601	23
SARIMA(1, 0, 2)(2, 1, 0) <sub>52</sub> (automated)	139.187	10.908	20
SARIMA(4, 0, 6)(2, 1, 1) <sub>52</sub> (manual)	105.343	7.390	20
NNAR(7, 1, 1) <sub>52</sub>	109.982	8.316	23

# 4 | DISCUSSION

-WILEY

As expected, the seasonal influenza incidence in Canada had a dominant seasonal pattern with peaks in fall and winter months during 2010–2019 influenza seasons. There was no substantial long-term trend. As stated previously, the seasonal behaviour may be attributed to annual changes in genomes of the virus making the host population susceptible to reinfection. Seasonal variation may also be attributed to the host population's behaviour of increased crowding in the fall and winter months. This leads to greater person-toperson contact giving the virus a greater opportunity to spread and persist in the population. Furthermore, increased indoor heating in the colder temperature creates a continuous recirculated body of air with low humidity. This creates an ideal environment for viral pathogens to exist (Andrew, 2016; Fitch et al., 1991; Lofgren et al., 2007; Lowen & Steel, 2014; McCullers et al., 1999).

Evidently, manual model optimization did not improve forecasts of the SARIMA model. The automated SARIMA model was most accurate in forecasting the total number of incident cases per 100,000 population as well as the number of cases per 100,000 at peak. The ANN was accurate in predicting the week in which incidence risk peaked. Examining Figure 7, the manual SARIMA model appeared reasonably accurate in predicting the earlier part of the flu season. Differences in predictive accuracy were more pronounced during the seasonal peak. All models largely overestimated the true incidence risk in the later part of the 2020 influenza season. This was likely due to the COVID-19 control measures that were introduced around that time and abruptly ended the influenza season.

Automated SARIMA modelling demonstrates the ability to capture and forecast seasonal influenza data accurately and, in this case study outperformed the manual SARIMA and the ANN in terms of both RMSE and MAE. ANNs are still powerful to capture and forecast a time series; however, as ANNs are data-driven, they require longer time series to improve predictive performance. In terms of computational cost, automated SARIMA modelling was also more efficient than both ANNs and manual SARIMA modelling. However, the significant Ljung-Box test of the automated model signalled the presence of residual autocorrelation. As evident by measures of RMSE and MAE, the point forecasts are still reliable and automated SARIMA model is the better method for forecasting purposes. However, residual autocorrelation indicates that there is still information left in the residuals that can be used to explain the spread of disease. The manual model improves on this and therefore, for explanatory purposes manually building a model is preferred.

# 4.1 | Limitations

Analysis and forecasts are limited by the quality of the data. The dataset likely underestimates the true incidence of influenza as it is restricted to only the reported cases. Therefore, forecasts may underpredict the true incidence risk of disease; however, this does not reflect capabilities of the model as models are only capable of predicting reported cases.

As observed, the PIs of the models are wide. Given the variation in the amplitude of the seasonal peaks, this result may be expected as the prediction intervals reflect the variation in the data. Furthermore, the prediction intervals are limited to the size of the training data, that is, the information given by the fitted model. As the training data only consisted of nine influenza seasons, the information used for forecasting is relatively small. Provided a longer time series, point forecasts are likely to improve and result in narrower PIs.

## 4.2 | Implications and future research directions

Real-time infectious disease forecasting has historically been a great challenge for the public health sector. As demonstrated in this study, forecasting methods such as the SARIMA models provide a useful tool to further refine allocation of health resources and make more informed decisions to combat seasonal influenza activity in Canada. Reporting the point forecasts and PIs will be beneficial for public health as the former provides an estimate of the likely scenario while the upper limit of the PI estimates the worse scenario.

Future research may take into consideration the flu seasons in Oceania as they precede the flu seasons in North America by a few months. Thus, flu season data from Oceania is expected to improve forecasts of the flu seasons in Canada in terms of seasonal peaks and trends, as well as improve forecast accuracy.

The forecasting approaches considered here target the influenza incidence risk for all of Canada; however, to identify geographic disease clusters more precise forecasts are desirable. This aligns with the concept of precision public health (Dowll et al., 2016). Hierarchical forecasting models provide individual forecasts for sub-populations

# 5 | CONCLUSION

To conclude, seasonal influenza incidence in Canada exhibits a strong seasonal component with no long-term trend. Machine learning algorithms can provide accurate forecasts and are growing in popularity. In this case study, however, the automated SARIMA(1, 0, 2)(2, 1, 0)<sub>52</sub> outperformed the ANN based on RMSE and MAE. Machine learning algorithms generally benefit from having longer time series and thus should still be employed in those scenarios. However, machine learning algorithms, often described as a 'black box', provide little insight on explanations of the disease process. Conversely, SARIMA models provide detailed explanations on how past observations affect future cases (Shmueli, 2010). Therefore, SARIMA models and ANNs can complement each other when the time series is long to provide better explanations of disease patterns.

## FUNDING INFORMATION

This study was funded by the Canadian Safety and Security Program (CSSP) from Defence Research and Development Canada (DRDC) awarded to the Public Health Agency of Canada CSSP-2018-CP-2334.

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data used for this study are publicly available through the global web-based tool FluNet from the World Health Organization database.

## ORCID

Armin Orang b https://orcid.org/0000-0002-4069-9771

## REFERENCES

- Andrew, A. C. (2016). Overcrowding and mortality during the influenza pandemic of 1918: Evidence from US Army camp a. A. Humphreys, Virginia. American Journal of Public Health, 106(4), 642–644. https:// doi.org/10.2105/AJPH.2015.303018
- Berke, O., Trotz-Williams, L., & de Montigny, S. (2020). Good times bad times: Automated forecasting of seasonal cryptosporidiosis in Ontario using machine learning. *Canada Communicable Disease Report*, 46, 192–197. https://doi.org/10.14745/ccdr.v46i06a07
- Box, G., & Jenkins, G. (1970). Time series analysis; Forecasting and Control. Holden-Day.
- Choi, S. B., & Ahn, I. (2020). Forecasting seasonal influenza-like illness in South Korea after 2 and 30weeks using Google trends and influenza data from Argentina. *PLoS One*, 15, e0233855. https://doi.org/ 10.1371/journal.pone.0233855

- Cleveland, B. R., Cleveland, S. W., McRae, E. J., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on Loess. *Journal of Official Statistics*, 6(1), 3–73.
- Cong, J., Ren, M., Xie, S., & Wang, P. (2019). Predicting seasonal influenza based on SARIMA model, in mainland China from 2005 to 2018. International Journal of Environmental Research and Public Health, 16(23), 1–8. https://doi.org/10.3390/ijerph16234760
- Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. Journal of Business & Economic Statistics, 33(1), 1–9. https:// doi.org/10.1080/07350015.2014.983236
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253–263. https:// doi.org/10.1080/07350015.1995.10524599
- Dowll, S., Blazes, D., & Desmond-Hellman, S. (2016). Four steps to precision public health. *Nature*, 540, 189–191.
- Fitch, W., Leiter, J., Li, X., & Palese, P. (1991). Positive Darwinian evolution in human influenza a viruses. National Academy of Sciences, 88(10), 4270–4274.
- Fuller, D., Buote, R., & Stanley, K. (2017). A glossary for big data in population and public health: Discussion and commentary on terminology and research methods. *Journal of Epidemiology and Community Health*, 71(11), 1113-1117. https://doi.org/10.1136/ jech-2017-209608
- Government of Canada. (2013). FluWatch. http://publications.gc.ca/site/ eng/9.507424/publication.html
- Government of Canada. (2019). Flu(influenza): For health professionals. https://www.canada.ca/en/public-health/services/diseases/fluinfluenza/health-professionals.html
- Government of Canada. (2020). FluWatch weeks calander. https://www. canada.ca/en/public-health/services/diseases/flu-influenza/influ enza-surveillance/fluwatch-weeks-calendar.html
- Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice (3rd ed.). OTexts.
- Kraay, A., Gicquelais, R. E., Roller, R., Lemoi, K., Bochniewicz, E. M., & Fox, S. J. (2017). Forecasting seasonal influenza in the United States using nonlinear time series analysis. 1–15.
- Lagacé-Wiens, P. R. S., Rubinstein, E., & Gumel, A. (2010). Influenza epidemiology-past, present, and future. *Critical Care Medicine*, 38(Suppl 4), 1–9. https://doi.org/10.1097/CCM.0b013e3181cbaf34
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. https://doi.org/10.1093/ biomet/65.2.297
- Lofgren, E., Fefferman, N. H., Naumov, Y. N., Gorski, J., & Naumova, E. N. (2007). Influenza seasonality: Underlying causes and modeling theories. *Journal of Virology*, 81(11), 5429–5436. https://doi.org/10. 1128/jvi.01680-06
- Lowen, A. C., & Steel, J. (2014). Roles of humidity and temperature in shaping influenza seasonality. *Journal of Virology*, 88(14), 7692– 7695. https://doi.org/10.1128/jvi.03544-13
- McCullers, J. A., Wang, G. C., He, S., & Webster, R. G. (1999). Reassortment and insertion-deletion are strategies for the evolution of influenza B viruses in nature. *Journal of Virology*, *73*(9), 7343–7348. https:// doi.org/10.1128/jvi.73.9.7343-7348.1999
- Petris, G., Petrone, S., & Campagnoli, P. (2009). Dynamic linear models with R. Springer Science & Business Media.
- RStudio Team. (2019). RStudio: Integrated development for R.
- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289-310. https://doi.org/10.1214/10-STS330
- Song, X., Xiao, J., Deng, J., Kang, Q., Zhang, Y., & Xu, J. (2016). Time series analysis of influenza incidence in Chinese provinces from 2004 to 2011. Medicine (United States), 95(26), 1–7. https://doi.org/10.1097/ MD.000000000003929

#### 10 WILEY

- Statistics Canada Population estimates on July 1st, by age and sex. https:// doi.org/10.25318/1710000501-eng
- Team RC. (2019). R a language and environment for statistical computing.
- Tian, C. W., Wang, H., & Luo, X. M. (2019). Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. Epidemiology and Infection. 147. 2008–2010. https:// doi.org/10.1017/S095026881800362X
- World Health Organization. (2018). Influenza (Seasonal). https://www. who.int/news-room/fact-sheets/detail/influenza-(seasonal)
- World Health Organization FluNet. https://www.who.int/influenza/ gisrs\_laboratory/flunet/en/
- Xu, Q., Li, R., Liu, Y., Luo, C., Xu, A., Xue, F., Xu, Q., & Li, X. (2017). Forecasting the incidence of mumps in Zibo City based on a SARIMA model. International Journal of Environmental Research and Public Health, 14(8), 1-11. https://doi.org/10.3390/ijerph14080925
- Yaari, R., Katriel, G., Huppert, A., Axelsen, J. B., & Stone, L. (2013). Modelling seasonal influenza: The role of weather and punctuated antigenic drift. Journal of the Royal Society Interface, 10(84), 20130298. https://doi.org/10.1098/rsif.2013.0298
- Zeger, S. L., Irizarry, R., & Peng, R. D. (2006). On time series analysis of public health and biomedical data. Annual Review of Public Health, 27, 57–79. https://doi.org/10.1146/annurev.publhealth.26.021304. 144517
- Zhang, G., Eddy Patuwo, B., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of

Forecasting, 14(1), 35-62. https://doi.org/10.1016/S0169-2070(97) 00044-7

- Zhang, G. P., & Qi, M. (2005). Neural network forecasting for seasonal and trend time series. European Journal of Operational Research, 160(2), 501-514. https://doi.org/10.1016/j.ejor.2003.08.037
- Zhang, X., Liu, Y., Yang, M., Zhang, T., Young, A. A., & Li, X. (2013). Comparative study of four time series methods in forecasting typhoid fever incidence in China. PLoS One, 8(5), e63116. https://doi. org/10.1371/journal.pone.0063116

How to cite this article: Orang, A., Berke, O., Poljak, Z., Greer, A. L., Rees, E. E., & Ng, V. (2024). Forecasting seasonal influenza activity in Canada-Comparing seasonal Auto-Regressive integrated moving average and artificial neural network approaches for public health preparedness. Zoonoses and Public Health, 00, 1–10. https://doi.

org/10.1111/zph.13114