Forecasting Measles in the European Union Using the Adaptive Neuro-Fuzzy Inference System

Erkut İnan İşeri 1, Kaan Uyar 2, Ümit İlhan 2

1Department of Electrical and Electronic Engineering, Near East University Faculty of Engineering, Nicosia, Cyprus
2Department of Computer Engineering, Near East University Faculty of Engineering, Nicosia, Cyprus

ORCID IDs of the authors: E.I. 0000-0003-1470-9742; K.U. 0000-0002-5608-9898; Ü.I. 0000-0002-4914-8749.

BACKGROUND/AIMS
Measles is one of the diseases that cause child mortality. The measles forecasting is essential in planning the fight against the disease and reducing the risk of the vaccine stocks expiration. Governments and health institutions estimate the measles vaccine requirements using certain equations, which are generally based on the size of the target population and the past consumption records. There are several studies that have examined the measles forecasting and conducted a vaccine requirement assessment.

MATERIAL and METHODS
This study uses a forecasting model that employs an adaptive neuro-fuzzy inference system (ANFIS) based on clustering. In this study, the measles data were derived using the World Health Organization (WHO) Measles and Rubella Surveillance Data, which cover the period from January 2011 to March 2018 and include 28 European Union member countries. Out of total 87 monthly measles cases, 80% were used for training, and 20% were chosen for testing.

RESULTS
In addition to the mean square error, the root mean square error, normalized root mean square error, mean absolute error, and mean absolute percentage error were calculated.

CONCLUSION
The model created for this purpose has demonstrated that the predictions made for the data collected between January 2011 and March 2018 were successful.

Keywords: Measles, forecasting, European Union

INTRODUCTION
Measles is still one of the leading causes of child mortality. The measles forecasting is essential in planning the fight against the disease and reducing the risk of the vaccine stocks expiration.

Governments and health institutions estimate the measles vaccine requirements using certain equations, which are generally based on the size of the target population and the past consumption records. There are several studies that have examined the measles forecasting and conducted a vaccine requirement assessment using tools, such as the Statistical Analysis System (1), the World Health Organization (WHO) Measles Programmatic Risk Assessment Tool (2-7), Statistical Package for the Social Sciences (8), nonlinear forecasting and chaos (9), and Markov Chain Monte Carlo methods (10).

Measles cases in the European Union (EU)/European Economic Area principally occur in unvaccinated populations, affecting both adults and children (11). The European Centre for Disease Prevention and Control annual report shows that as many as 80% of teenagers and young adults who contracted measles in 2017 had not been vaccinated (11, 12).

MATERIAL and METHODS
According to the WHO Measles and Rubella Surveillance Data showing the distribution of measles cases by country and by month between January 2011 and March 2018 (Table 1), there were a total of 1,353,222 measles cases worldwide.
The WHO indicates that many cases are not reported. For the same period, a total of 86,246 measles cases were reported by the EU member states (Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom).

This study uses a forecasting model in MATLAB (MathWorks Inc., USA) that employs an adaptive neuro-fuzzy inference system (ANFIS) based on clustering that was applied to the EU measles cases. Fuzzy clustering is a data-clustering technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. ANFIS uses a hybrid learning algorithm to tune the parameters of a Sugeno-type fuzzy inference system. The algorithm uses a combination of the least-squares and back-propagation gradient descent methods to model a training dataset. ANFIS also validates the models using a checking dataset to test for overfitting of the training data.

Out of the total 87 monthly EU measles cases (Table 3), 67 (80%) were used for training, and 17 (20%) of them were chosen for testing. In this study, the mean square error (MSE), root mean square error (RMSE), normalized root mean square error (NRMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) were calculated using the following equations, which are the indexes of forecasting accuracy:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (Y_i - P_i)^2, \tag{1}
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (Y_i - P_i)^2}, \tag{2}
\]

\[
NRMSE = \frac{RMSE}{\text{Mean of actual data}},
\]

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |Y_i - P_i|,
\]

\[
MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{Y_i - P_i}{Y_i} \right|.
\]
where $Y_i$ is an actual value, $P_i$ is the forecasted value of the $i$-th data obtained, and $N$ is the number of data. In our forecasting model, we assumed the relationship as

$$ y(k) = F(y(k-3), y(k-2), y(k-1)) $$

where $y(k)$ are the measles cases by $k$-th month, $y(k-1)$ is the measles cases by $(k-1)$-th month, $y(k-2)$ is the measles cases by $(k-2)$-th month, and $y(k-3)$ is the measles cases by $(k-3)$-th month.

Since this study was based on merely a statistical dataset freely available online (13), we did not use any human materials. So, an ethic committee approval and informed consent were not necessary. This study was performed in accordance with the principles of the Declaration of Helsinki.

RESULTS

The study was aimed at predicting the number of measles cases expected in the following month given the data for previous three months. The data cover all measles cases in the EU area between January 2011 and March 2018 by month. The data collected for March 2018 are not complete, but we included them because they were listed.

The client data were subjected to the subtractive clustering procedure using MATLAB. The algorithm was repeated for the cluster radii 0.1 through 0.9, while keeping the accept ratio, reject ratio, and squash factor constant at 0.5, 0.15, and 1.25, respectively. The training and testing of the FIS with the radius 0.2 at 600 epochs resulted in the lowest training error of 282.261 and testing error of 503.821. The ANFIS graphical outputs show the training points in Figure 1 and testing points in Figure 2.

The input membership functions obtained from the subtractive clustering using MATLAB are of the Gaussian type. Each input space shown in Figure 3 generalizes the input data submitted to the ANFIS training.

The model is constructed with five clusters, hence five rules as shown in Figure 4. The model structure has three inputs, each made of five Gaussian membership functions. The five rules determined as a result of the training of the network. The first order Sugeno-type reasoning was conducted on the forth layer where the firing strengths were obtained.

The crisp output was received on the 6th layer after the aggregation process. Figure 5 shows the crisp input data and the crisp output.
output of the number of expected measles cases for the following month.

The surface plot is a three-dimensional graphics tool in MATLAB that shows the shape of the FIS with three-dimensional plots at a time. A particular instance of the three consecutive months is shown in Figure 6. A general idea can be obtained from the surface plot by observing the shape of the relationship between the given parameters.

**DISCUSSION**

Measles cases are considered to be epidemiologically linked, confirmed by laboratory findings and clinical cases as reported to the WHO. Some countries report measles cases at irregular intervals, providing multiple months of data in a single month period, and missing future months are reported as no cases; hence, they are expected to be updated as data becomes available (11).

In this study, a forecasting model was created using ANFIS to predict the number of future measles cases in the EU area. Such a work was particularly required to control the vaccine stocks and organize an effective distribution through the member countries. The model created for this purpose demonstrated that the predictions made for the data collected between January 2011 and March 2018 were successful.

**Ethics Committee Approval:** The authors declared that the research was conducted according to the principles of the World Medical Association Declaration of Helsinki “Ethical Principles for Medical Research Involving Human Subjects” (amended in October 2013).

**Informed Consent:** N/A.

**Peer-review:** Externally peer-reviewed.


**Conflict of Interest:** The authors have no conflicts of interest to declare.

**Financial Disclosure:** The authors declared that this study has received no financial support.

**REFERENCES**

12. Wise J. European countries are urged to carry out catch-up campaigns as measles outbreaks continue. BMJ 2018; 361: k1771. [CrossRef]