

RESEARCH ARTICLE

Evaluation of models for multi-step forecasting of hand, foot and mouth disease using multi-input multi-output: A case study of Chengdu, China

Xiaoran Geng¹, Yue Ma¹, Wennian Cai¹, Yuanyi Zha², Tao Zhang¹, Huadong Zhang³, Changhong Yang⁴, Fei Yin^{1*}, Tiejun Shui^{5*}

1 West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China, **2** Kunming Medical University, Kunming, China, **3** Chongqing Center for Disease Control and Prevention, Chongqing, China, **4** Sichuan Center for Disease Control and Prevention, Chengdu, China, **5** Yunnan Center for Disease Control and Prevention, Kunming, China

✉ These authors contributed equally to this work.

* yinfeiscuedu@163.com (FY); shuitiesjunyjkedu@163.com (TS)



OPEN ACCESS

Citation: Geng X, Ma Y, Cai W, Zha Y, Zhang T, Zhang H, et al. (2023) Evaluation of models for multi-step forecasting of hand, foot and mouth disease using multi-input multi-output: A case study of Chengdu, China. *PLoS Negl Trop Dis* 17(9): e0011587. <https://doi.org/10.1371/journal.pntd.0011587>

Editor: Marilia Sá Carvalho, Fundacao Oswaldo Cruz, BRAZIL

Received: January 18, 2023

Accepted: August 11, 2023

Published: September 8, 2023

Copyright: © 2023 Geng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The HFMD data underlying the results in the study cannot be shared publicly because of the limitation of data availability in the data management rule of Sichuan Center for Disease Control and Prevention. The datasets generated and/or analyzed during the current study are available from Sichuan Center for Disease Control and Prevention (<https://www.sccdc.cn/>, email: zxbgs@sccdc.cn). The meteorological data can be obtained from the China Meteorological Data Sharing Service System

Abstract

Background

Hand, foot and mouth disease (HFMD) is a public health concern that threatens the health of children. Accurately forecasting of HFMD cases multiple days ahead and early detection of peaks in the number of cases followed by timely response are essential for HFMD prevention and control. However, many studies mainly predict future one-day incidence, which reduces the flexibility of prevention and control.

Methods

We collected the daily number of HFMD cases among children aged 0–14 years in Chengdu from 2011 to 2017, as well as meteorological and air pollutant data for the same period. The LSTM, Seq2Seq, Seq2Seq-Luong and Seq2Seq-Shih models were used to perform multi-step prediction of HFMD through multi-input multi-output. We evaluated the models in terms of overall prediction performance, the time delay and intensity of detection peaks.

Results

From 2011 to 2017, HFMD in Chengdu showed seasonal trends that were consistent with temperature, air pressure, rainfall, relative humidity, and PM₁₀. The Seq2Seq-Shih model achieved the best performance, with RMSE, sMAPE and PCC values of 13.943–22.192, 17.880–27.937, and 0.887–0.705 for the 2-day to 15-day predictions, respectively. Meanwhile, the Seq2Seq-Shih model is able to detect peaks in the next 15 days with a smaller time delay.

(<https://data.cma.cn/data/cdcindex/cid/0b9164954813c573.html>). The air pollution data is available from Department of Ecology and Environment of Sichuan Province. (<http://www.scaep.cn/sthj/c104334/scemc.shtml>).

Funding: This work was supported by the National Natural Science Foundation of China (81872713 to FY, and 82373689 to YM); Sichuan Science and Technology Program (2021YFS0181 to YM, and 2022YFS0641 to TZ); Chongqing Science and Technology Program (cstc2020jscx-cylhX0003 to TZ, URL: <http://kjj.cq.gov.cn/>). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Conclusions

The deep learning Seq2Seq-Shih model achieves the best performance in overall and peak prediction, and is applicable to HFMD multi-step prediction based on environmental factors.

Author summary

Hand, foot and mouth disease (HFMD) remains a serious public health concern in China. It is important to predict trends and understand peaks in the number of cases in advance for its prevention and control. The aim of this study was to consider the influence of meteorology and air pollution on the transmission of HFMD and establish multi-step prediction models for HFMD. In this study, we compared the performance of the Shih attention mechanism with the Luong attention mechanism, Seq2Seq model and LSTM model for future multi-day HFMD prediction based on multi-input multi-output. It was found that the Seq2Seq-Shih model performed best in predicting the trend for the next 2 days-15 days with RMSE, sMAPE and PCC values of 13.943~22.192, 17.880~27.937, and 0.887~0.705, respectively. Meanwhile, it was able to predict the peak within 15 days earlier by 17 days. This is the first study to use Seq2Seq models and perform daily multi-step prediction of HFMD. This study demonstrates the benefit of the Shih attention mechanism in multivariate time series multi-step prediction of infectious diseases.

1. Introduction

Hand, foot and mouth disease (HFMD) is a common infectious disease caused by enteroviruses that usually affects children under 5 years [1]. HFMD was first reported in New Zealand [2], and most outbreaks in recent decades have occurred in the Asia-Pacific region [3]. As a country in this region, China listed HFMD as a notifiable infectious disease in 2008 [4]. HFMD was ranked in the top three of all notifiable infectious diseases in China from 2008 to 2019, with an average annual number of approximately 1.87 million cases [5]. According to the study of economic costs associated with patients diagnosed with HFMD in China in 2012–2013, the average quality-adjusted life years (QALYs) lost for mild and severe patients were 6.9 and 13.7 per 1000 cases, respectively. In addition, the average total costs in United States dollars (USD) for mild and severe patients were \$1072 and \$3051, respectively [6]. In China, HFMD is still a serious public health issue. Therefore, it is necessary to forecast HFMD trends to detect peaks timely, providing more time for public health departments to make decisions and reduce the risk of HFMD outbreaks.

Some environmental factors related to the spread of HFMD need to be considered when forecasting its trends. It has been shown that temperature, humidity, air pressure, rainfall and sunshine duration are associated with HFMD [7–9]. In addition, air pollution factors such as PM₁₀, SO₂ and NO₂ are also associated with HFMD [10–12]. Furthermore, these environmental factors not only have mostly nonlinear associations with HFMD but also have delayed and cumulative effects on the number of HFMD cases.

In recent years, there have been many studies on the prediction of HFMD [13–23]. For example, Tian et al. developed a seasonal autoregressive integrated moving average (SARIMA) model using monthly data [15]. Xie et al. used the Prophet model to predict daily HFMD cases in Hubei Province, China [19]. Yoshida et al. predicted the weekly number of HFMD cases in Japan [20]. In brief, most of these studies are single time-point predictions, usually predicting

the incidence of HFMD on a future day (or week/month). However, for public health departments, the predictions of a single point in time are limited in terms of how far in advance they can detect a peak in the number of HFMD cases. This requires multi-day (week) forecasting to help public health departments to understand the future trends of HFMD cases in advance, to identify peaks early, and to have sufficient time to make decisions. However, there are few studies on the multi-day (week) forecasting of HFMD.

Multi-day (week) forecasting is achieved by multi-step forecasting, i.e., using data from past time points to forecast more than one future time point. There are three main types of multi-step forecasting approaches: direct multi-step [24], iterative multi-step [25], and multi-input multi-output forecasting [26]. Both direct multi-step and iterative multi-step forecasting is achieved by single-step forecasting, which will cause the accumulation of errors and make the forecasting task more difficult. However, multi-input multi-output forecasting is achieved by outputting a prediction vector, which not only avoids error accumulation, but also preserves the random dependences among the numbers of HFMD cases. In addition, considering that meteorological and air pollution factors have delayed and cumulative effects on HFMD and that their effects are nonlinear, we use recursive neural networks to learn these features and achieve multi-day (week) forecasting with multi-input multi-output.

The long-short term memory network (LSTM) model [27] has achieved good performance in predicting infectious diseases such as COVID-19 [28], influenza [29], and dengue fever [30]. It can learn complex nonlinearity, memorize historical information using internal storage units [31], and achieve multi-output through ‘multi-objective regression’.

The sequence-to-sequence(Seq2Seq) model [32] also enables multi-step prediction using LSTM units via an encoder-decoder structure, which converts one sequence (historical cases of HFMD) into another sequence (future cases of HFMD). However, when the length of the input sequence increases, it is more difficult for the encoder to fully encode and store all the information into the context vector, and then the predicted values generated by the decoder are less accurate. The introduction of the attention mechanism into the Seq2Seq model may improve the situation. The attention mechanism learns the most relevant information from all previous data to predict the target sequence [33]. That is, the attention mechanism is able to identify the optimal lag in which the input HFMD, meteorological and air pollution variables have the greatest impact on predicting the number of HFMD cases.

The capital of Sichuan Province, Chengdu is situated in southwest China, west of the Sichuan Basin, and has a humid subtropical monsoon climate. The annual incidence of HFMD in Chengdu showed an increasing trend during 2009–2018, with an average annual incidence of 250.2 per 100,000 person-years [34]. In addition, the major serotypes of HFMD in Chengdu have been changing over the years [35]. Thus, it is extremely important to pay long-term attention to HFMD in Chengdu and to forecast the trend of HFMD for its prevention and control.

In this study, we compared the performance of the Seq2Seq-Shih, Seq2Seq-Luong, LSTM and Seq2Seq models in terms of overall prediction, peak prediction accuracy and prediction time delay by considering the effects of meteorological and air pollution on HFMD in Chengdu city. Seq2Seq-Luong and Seq2Seq-Shih represent the Seq2Seq models using the Luong attention mechanism [36] and the Shih attention mechanism [37], respectively. To our knowledge, this is the first study to use the Seq2Seq model and to make multi-step predictions of HFMD on a daily scale, which can provide public health departments with an advanced understanding of the future trends of HFMD.

2. Materials and methods

2.1. Ethics statement

Our study was approved by the Institutional Review Board of the West China School of Public Health, Sichuan University. All HFMD surveillance data in this study were obtained from the Chinese Disease Prevention and Control Information System. The study was conducted at the population level. Therefore, this study did not involve confidential information and did not require informed consent.

2.2. Data collection and processing

The number of daily HFMD cases in Chengdu from January 1, 2011 to December 31, 2017 was obtained from the Sichuan Center for Disease Control and Prevention. Only cases in patients under 15 years were included in this study since children account for the vast majority of HFMD cases (more than 99% of all cases) [34]. The HFMD surveillance data is not publicly available but it is available on request from the Sichuan Provincial Center for Disease Control and Prevention.

Daily meteorological data were obtained from the site monitoring data of the China Meteorological Data Sharing Service System, including daily average temperature ($^{\circ}\text{C}$), sunshine duration (h), air pressure (hPa), rainfall (mm), wind speed (m/s) and relative humidity (%). Daily air pollution data were collected from urban monitoring data of the China National Environmental Monitoring Center, including PM_{10} ($\mu\text{g}/\text{m}^3$), SO_2 ($\mu\text{g}/\text{m}^3$) and NO_2 ($\mu\text{g}/\text{m}^3$).

First, missing values were addressed. During the study period, there were no missing values for HFMD and meteorological data, 0.16% missing values for SO_2 and 0.12% missing values for both NO_2 and PM_{10} . We used linear interpolation to fill in these missing values. Next, features were selected. According to the results of Spearman correlation analysis (S1 Fig), air pressure was highly correlated with average temperature, which was excluded from this study because of the stronger correlation between average temperature and HFMD. Next, the data were normalized. To enable the model to handle all features in a balanced way, we normalized each feature to the range [0, 1] using the Min-Max scaling method. Then, the data were divided into 3 parts in chronological order: the training set (80%) was used to build the LSTM, Seq2Seq, Seq2Seq-Luong and Seq2Seq-Shih models, the validation set (10%) was used to adjust the hyperparameters, and the test set (10%) was used to evaluate the overall and peak performance of the models. The overall modeling flow chart is shown in the S2 Fig.

2.3. LSTM model

LSTM is a recurrent neural network that is widely used to process temporal data. It stores and controls information flow by constructing internal states and using three gates (input gate, forget gate, and output gate) [38]. The LSTM uses the network activation values of the previous time step as the input of current time step to influence the output of the current time step [39]. This LSTM structure was able to learn the delayed and cumulative effects of influencing factors on HFMD in this study. In addition, LSTM stores network activation values in internal states to learn serial long-term correlations of the number of HFMD cases. Moreover, neural networks are capable of approximating arbitrary nonlinear functions. These properties enable LSTM to model complex multivariate time series [40].

Fig 1 describes the structure of the LSTM model. At each time step, the forget gate (f_t), the input gate (i_t), the gate of the memory cell (\tilde{C}_t), the memory cell (C_t), the output gate (o_t) and the output (h_t) are calculated by the nonlinear function (σ or \tanh). Here, h_{t-1} , x_t , W and b represent the hidden state of the previous time step, the input of the current time step, the weight matrix

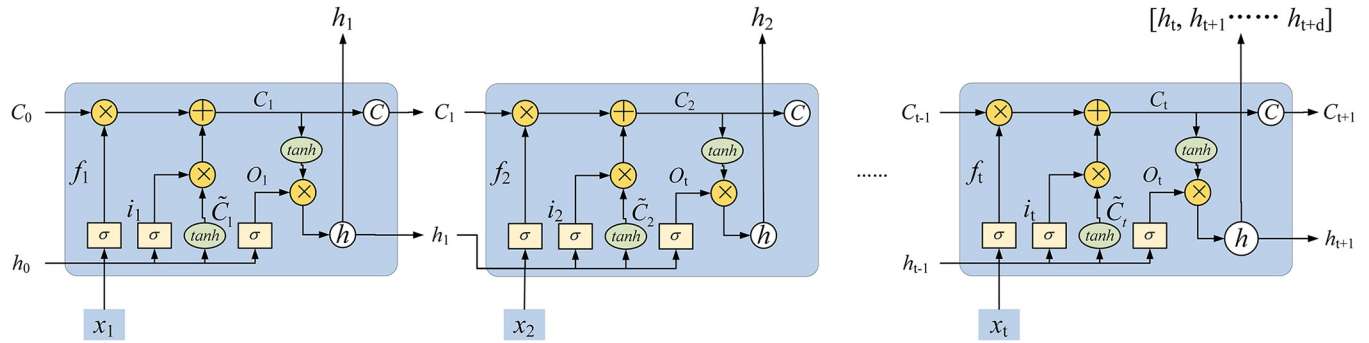


Fig 1. The structure of the LSTM model.

<https://doi.org/10.1371/journal.pntd.0011587.g001>

and the bias, respectively. The last time step outputs row vectors to achieve multi-step prediction.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \tag{3}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{4}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{5}$$

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

2.4. Seq2Seq model

The Seq2Seq model uses an encoder-decoder structure to convert variable-length input sequences into variable-length output sequences [41], which is originally applied to machine translation and has been increasingly applied to time series prediction in recent years [42]. In our study, the encoder-decoder structure used recurrent neural networks. The encoder learns the information of the input sequence and generates the context vector, while the decoder predicts the output sequence based on the context vector. The LSTM neural unit is the most widely used to learn the long-term dependence of time series. Therefore, the processing unit of both the encoder and decoder in this study used the LSTM unit.

Fig 2 illustrates the structure of the Seq2Seq model. Here, we assumed that $\mathbf{X}(x_1, x_2, x_3, \dots, x_t)$ represents multivariate time series including the number of HFMD cases and meteorological and air pollution factors at time step t , and y_t represents the number of HFMD cases at time step t . The encoder generates hidden states $h_{en,t}$ for each time step of the input sequence \mathbf{X} . The last time-step hidden state is the context vector C . The decoder takes the context vector, the previous hidden state and the previous output as input, calculates the current hidden state output $h_{de,t+d}$, and then obtains the current predicted value \hat{y}_{t+d} with a nonlinear function. f_1 and f_2 in Eqs (8) and (9) refer to the activation function.

$$[h_{en,1}, h_{en,2}, \dots, h_{en,t}] = \text{LSTM}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_t) \tag{7}$$

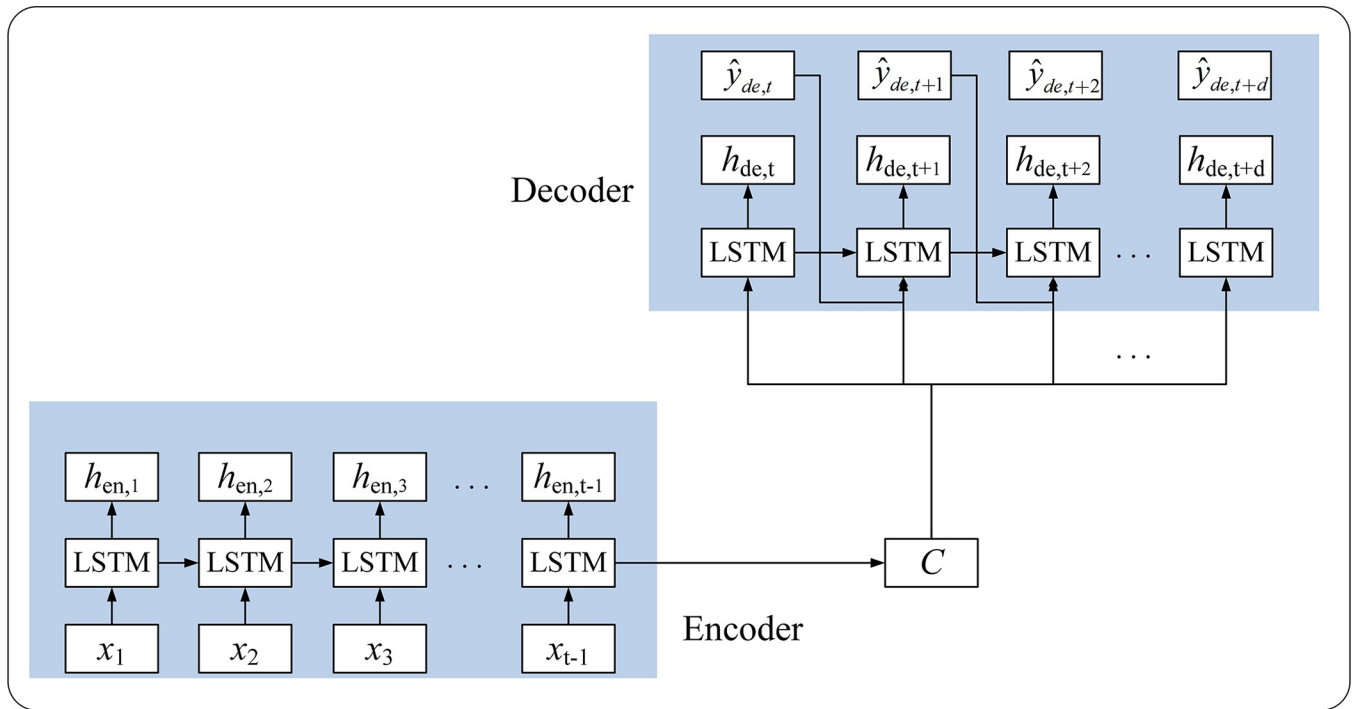


Fig 2. The structure of the Seq2Seq model.

<https://doi.org/10.1371/journal.pntd.0011587.g002>

$$h_{de,t+d} = f_1(\hat{y}_{t+d-1}, C, h_{de,t+d-1}) \tag{8}$$

$$\hat{y}_{t+d} = f_2(h_{de,t+d}) \tag{9}$$

2.5. Seq2Seq model based on the Luong attention mechanism

The Seq2Seq model makes predictions from the perspective of contextual information in the learning data, and its encoder compresses the contextual information of the input sequence into a fixed-length vector. However, the attention mechanism makes predictions in terms of the relevance of data information, learns the correlations between input and output sequences and focuses on the data features useful for prediction to generate a dynamic contextual vector. The Luong mechanism is proposed for considering all input words in a natural language processing task and the relative importance of each word [36]. Considering the similarity between natural language processing and time series prediction, studies using this method in the fields of energy load [43], wind power [44] and building energy [45] are increasing. The structure of the Seq2Seq model using the Luong attention mechanism is shown in Fig 3.

The context vector v_t in the Luong attention mechanism is calculated by a weighted sum of the hidden states of the encoder and the corresponding attention scores. In this study, we used the ‘dot product’ of the Luong attention mechanism to calculate the attention score. Eq (11) calculates the correlation between the encoder hidden state $h_{en,i}$ and decoder hidden state $h_{de,j}$ which is used to express the importance of $h_{en,i}$ for predicting $h_{de,j}$. The attention score a_{ij} is aligned using the *softmax* function to highlight the weight of important information.

$$a_{ij} = \exp(\text{score}(h_{de,j}, h_{en,i})) / \sum_{i=1}^{t-1} \exp(\text{score}(h_{de,j}, h_{en,i})) \tag{10}$$

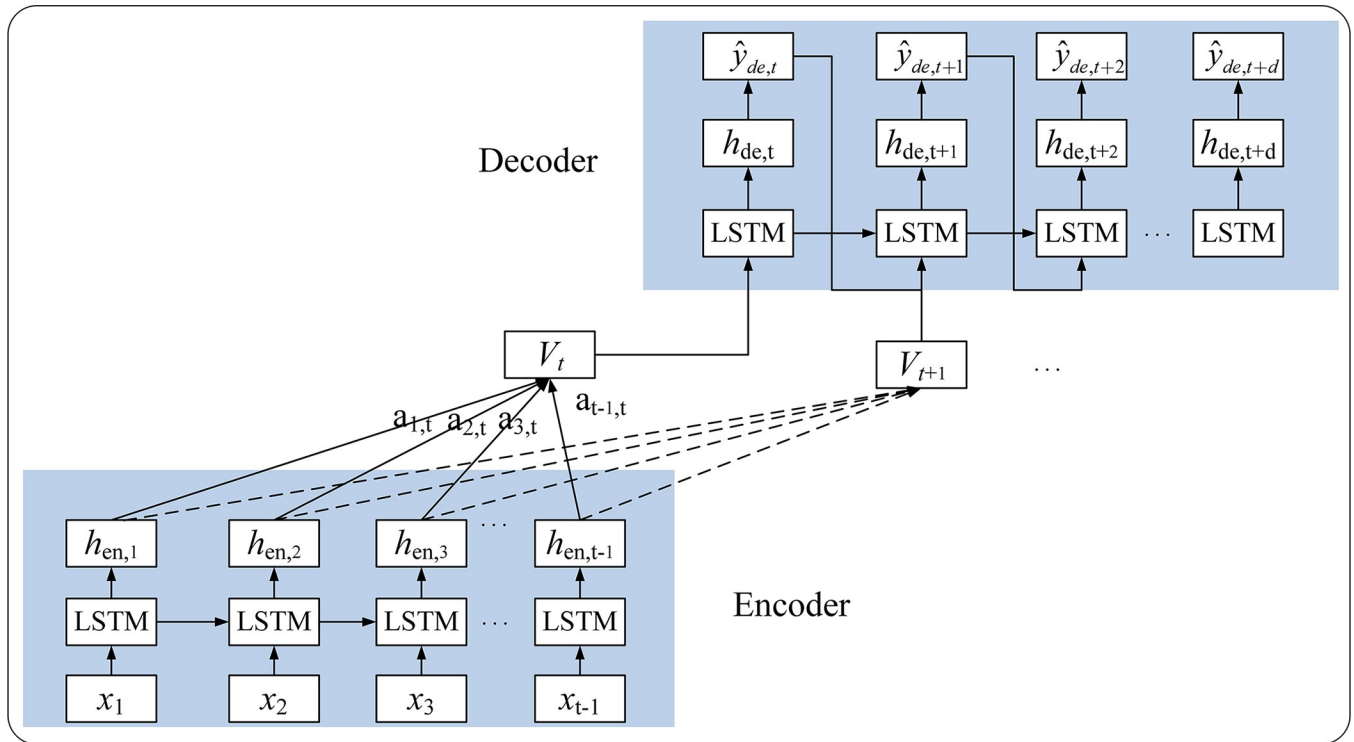


Fig 3. The structure of the Seq2Seq model using the Luong attention mechanism.

<https://doi.org/10.1371/journal.pntd.0011587.g003>

$$score(h_{de,j}, h_{en,i}) = h_{de,j}^T \cdot h_{en,i} \tag{11}$$

$$v_t = t - 1 \sum_{i=1}^{t-1} a_{i,j} h_{en,i} \tag{12}$$

2.6. Seq2Seq model based on the Shih attention mechanism

The Shih attention mechanism focuses on multivariate time series forecasting [31], which selects the relevant variables in each time step on the basis of capturing the time dependence of the time series, while the Luong attention mechanism selects all features of the relevant time step. The encoder of the Seq2Seq model based on these two attention mechanisms is the same, but the decoding process calculates the weights in a different way. The structure of the Seq2Seq model using the Shih attention mechanism is illustrated in Fig 4.

Given the encoder hidden state matrix $\mathbf{H} = [h_1, h_2, \dots, h_{t-1}]$, $\mathbf{H} \in \mathbb{R}^{m \times (t-1)}$, i.e., the encoder extracts the m -dimensional information of the data, and the parameters of the LSTM units are shared for each time step; the row vectors of the \mathbf{H} matrix express the same dimensional information. The Shih attention mechanism uses convolutional neural networks to capture the m -dimensional information of the hidden states and produces a matrix \mathbf{H}^C . The context vector v_t is obtained by weighting the sum of the attention score and the row vector H_i^C . This method also uses the ‘dot product’ to calculate the attention scores α_i of each row of \mathbf{H}^C with the decoder hidden state h_t . W_a represents the connection weights of H_i^C and h_t . Notably, the Luong attention mechanism uses the *softmax* function to scale up the attention of local important information, while the Shih attention mechanism uses the *sigmoid* function to obtain

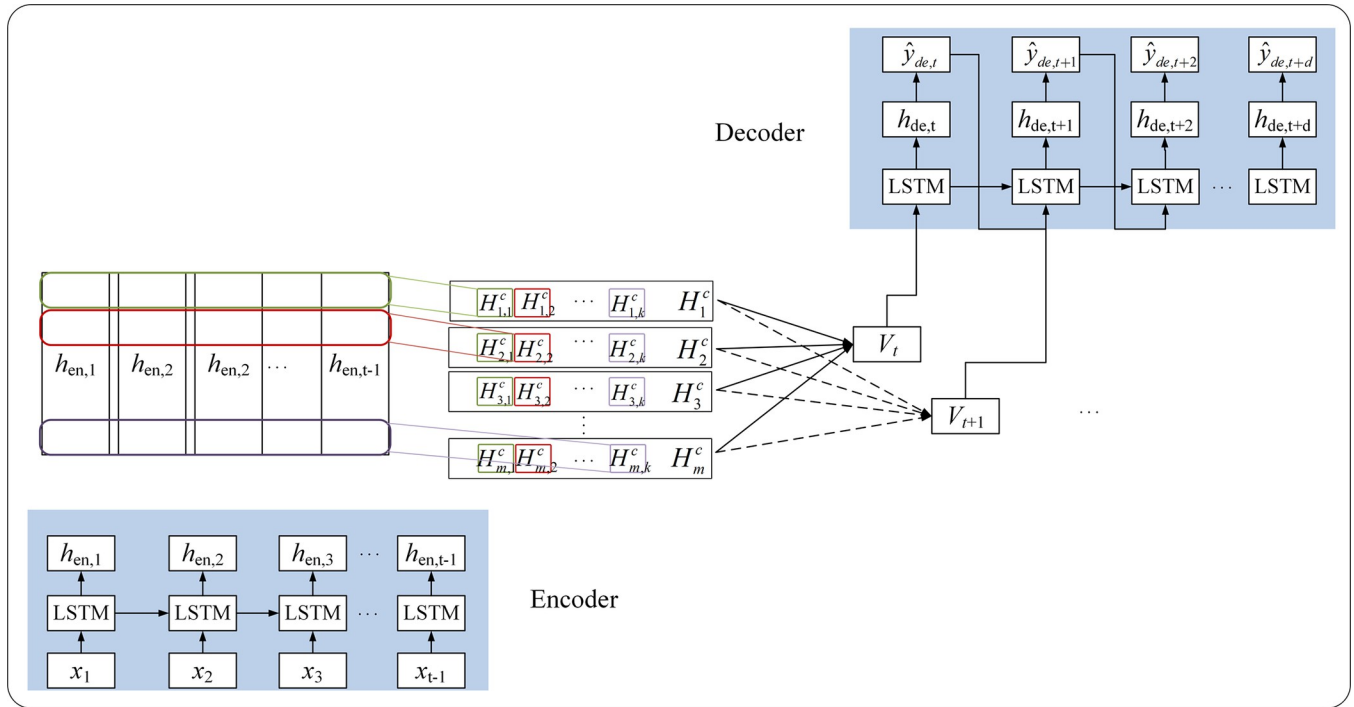


Fig 4. The structure of the Seq2Seq model using the Shih attention mechanism (different colored rectangles indicate 1-D CNN filters).

<https://doi.org/10.1371/journal.pntd.0011587.g004>

more helpful variables.

$$\alpha_i = \text{sigmoid}((H_i^C)^T W_a h_t) \tag{13}$$

$$v_t = m \sum_{i=1}^m \alpha_i H_i^C \tag{14}$$

2.7. Parameter selection and model training

We list some parameters used to train the model. The time step is an important parameter of the neural network. We used the autocorrelation coefficient to calculate the correlation coefficient of the number of HFMD case series at different moments. As seen from Fig 5, the autocorrelation coefficients were tailing and decreasing, and the partial autocorrelation coefficients (PACF) were larger at lags of 1–7 days and then showed a trend with a period of 7 days. However, the trend of the PACF at lags of 1–7 days was different from the PACF during the period, so the time step was set to 14 in this study. The number of hidden layer neurons was searched from the set {16, 32, 64}, the number of encoder layers was obtained from {1, 2}, and the number of decoder layers was set to 1. Moreover, to avoid model overfitting, we also used a dropout parameter to allow neurons to be lost with a 0.2 probability. In the training of models, the loss function was the mean square error (MSE). The number of epochs was determined by the error convergence curve. In addition, considering the randomness of internal parameter initialization during neural network training, we conducted 10 replicate experiments and took the average of the predicted values from the replicate experiments as the final prediction. All models were implemented in Python 3.7 and deployed using Keras 2.6.0 and TensorFlow-gpu 2.6.0.

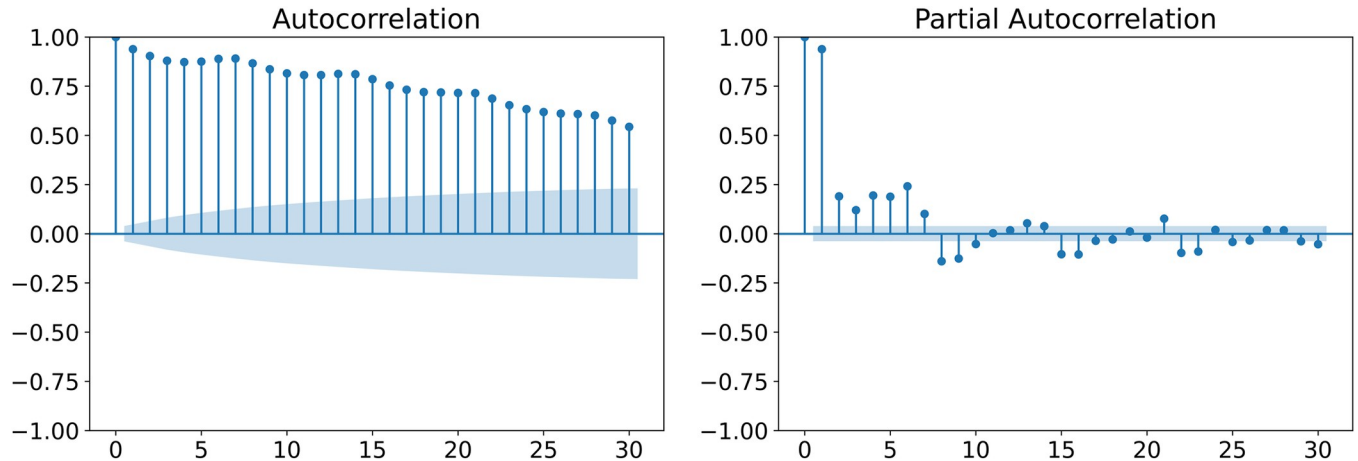


Fig 5. The autocorrelation and partial autocorrelation of HFMD series in Chengdu.

<https://doi.org/10.1371/journal.pntd.0011587.g005>

2.8. Model evaluation

2.8.1. Overall performance evaluation. To compare the performance of models, the root mean square error (RMSE), symmetric mean absolute percentage error (sMAPE) and Pearson correlation coefficient (PCC) were calculated.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \tag{15}$$

$$sMAPE = \frac{2 \times 100\%}{n} \sum_{i=1}^n \frac{|y_i^* - y_i|}{|y_i| + |y_i^*|} \tag{16}$$

$$PCC = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)(y_i^* - \bar{y}_i^*)}{\sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^n (y_i^* - \bar{y}_i^*)^2}} \tag{17}$$

where n represents the length of the observation sequence, y_i^* is the predicted value of the i th observation, y_i is the actual value of the i th observation, \bar{y}_i is the mean of the observation sequence, and \bar{y}_i^* is the mean of the predicted sequence.

2.8.2. Peak performance evaluation. We also evaluated the ability of each model to predict the time point and intensity of HFMD peaks. The HFMD peak was defined as the day with the highest number of HFMD cases during the HFMD season. We calculated the difference between the observed peak time t_{peak} and the predicted peak time \hat{t}_{peak} . Moreover, the relative error between the predicted peak value y_{max}^* and the observed peak value y_{max} was also calculated.

$$delay = |t_{peak} - \hat{t}_{peak}| \tag{18}$$

$$magnituderror = \frac{|y_{max} - y_{max}^*|}{y_{max}} \tag{19}$$

The test set was used to evaluate the performance of the model, so two peak periods were selected from the test set for evaluation in this study. The first peak period was from May 26,

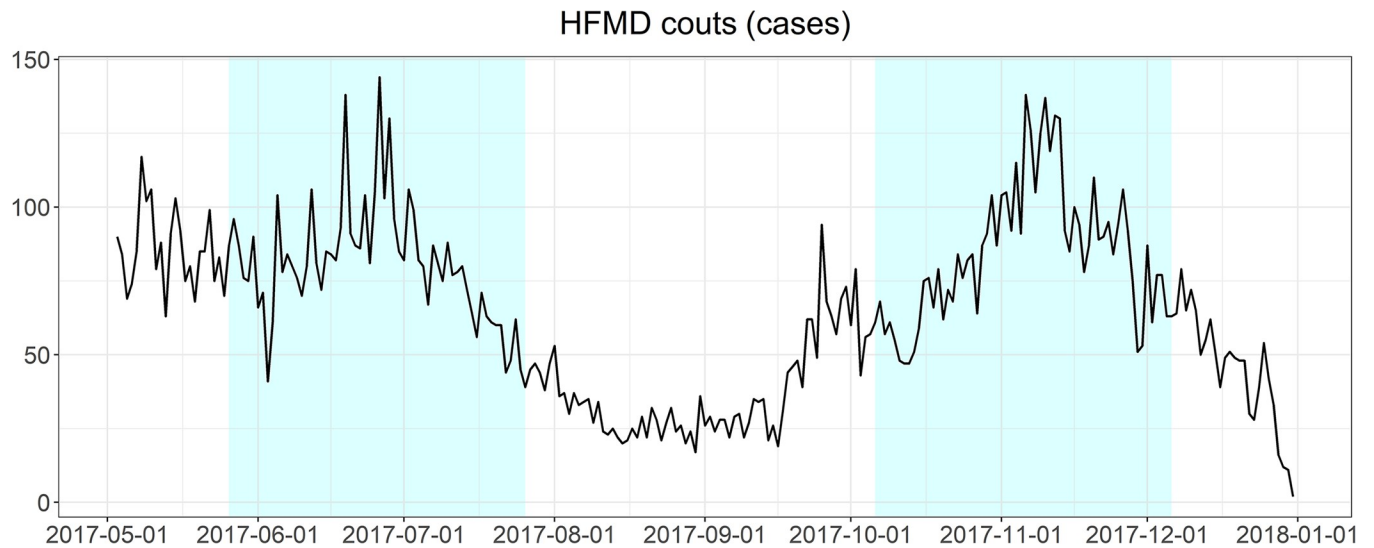


Fig 6. The trends in the number of HFMD cases during the two peak periods in the test set (shaded).

<https://doi.org/10.1371/journal.pntd.0011587.g006>

2017 to July 26, 2017 and the second peak period was from October 6, 2017 to December 6, 2017 (Fig 6).

3. Results

3.1. Descriptive analysis

In Chengdu, there were 184,610 HFMD cases in children under 15 years. Approximately 72 HFMD cases were reported daily on average, with a maximum of 303 cases. The average values of wind speed, sunshine duration, temperature, relative humidity, rainfall, PM₁₀, SO₂ and NO₂ were 1.22 m/s, 2.74 hours, 16.52°C, 79.32%, 2.65 mm, 112.14 μg/m³, 21.96 μg/m³ and 53.36 μg/m³, respectively. Fig 7 shows that there are two annual peaks in the number of HFMD cases, one from April to July and another from October to December. In most years, the first peak is slightly higher than the second, except for 2013 and 2016. Moreover, we observed a clear and consistent seasonal pattern in the number of HFMD cases, temperature, rainfall, relative humidity and PM₁₀. The detailed data description of the HFMD, meteorological and air pollution variables are shown in Table 1.

3.2. Comparison of models in overall prediction

Table 2 summarizes the means and standard deviations of the metrics for the 10 replicate experiments when the model takes the optimal parameters at forecasting steps $h = 2, 3, 6, 9, 12$ and 15, with the bolded numbers indicating the best mean values of the metrics at the forecasting horizons. In addition, we showed the experimental results with bar charts to visually compare the differences of the models (Fig 8). Fig 9 shows the prediction trends of the models at different forecasting horizons and the observed trends.

As shown in Table 2 and Fig 8, the metrics gradually decreased as the forecasting horizon increased, except for the RMSE metric of the Seq2Seq-Luong model at forecasting horizon T+3. In terms of RMSE, the error of the Seq2Seq-Shih model increased from 13.943 at forecasting horizon T+2 to 22.192 at forecasting horizon T+15, with the smallest value for each forecasting horizon and the slowest increase. In terms of sMAPE, the Seq2Seq-Shih model from T+2 to T+15 also obtained the smallest error for each forecasting horizon. In terms of overall

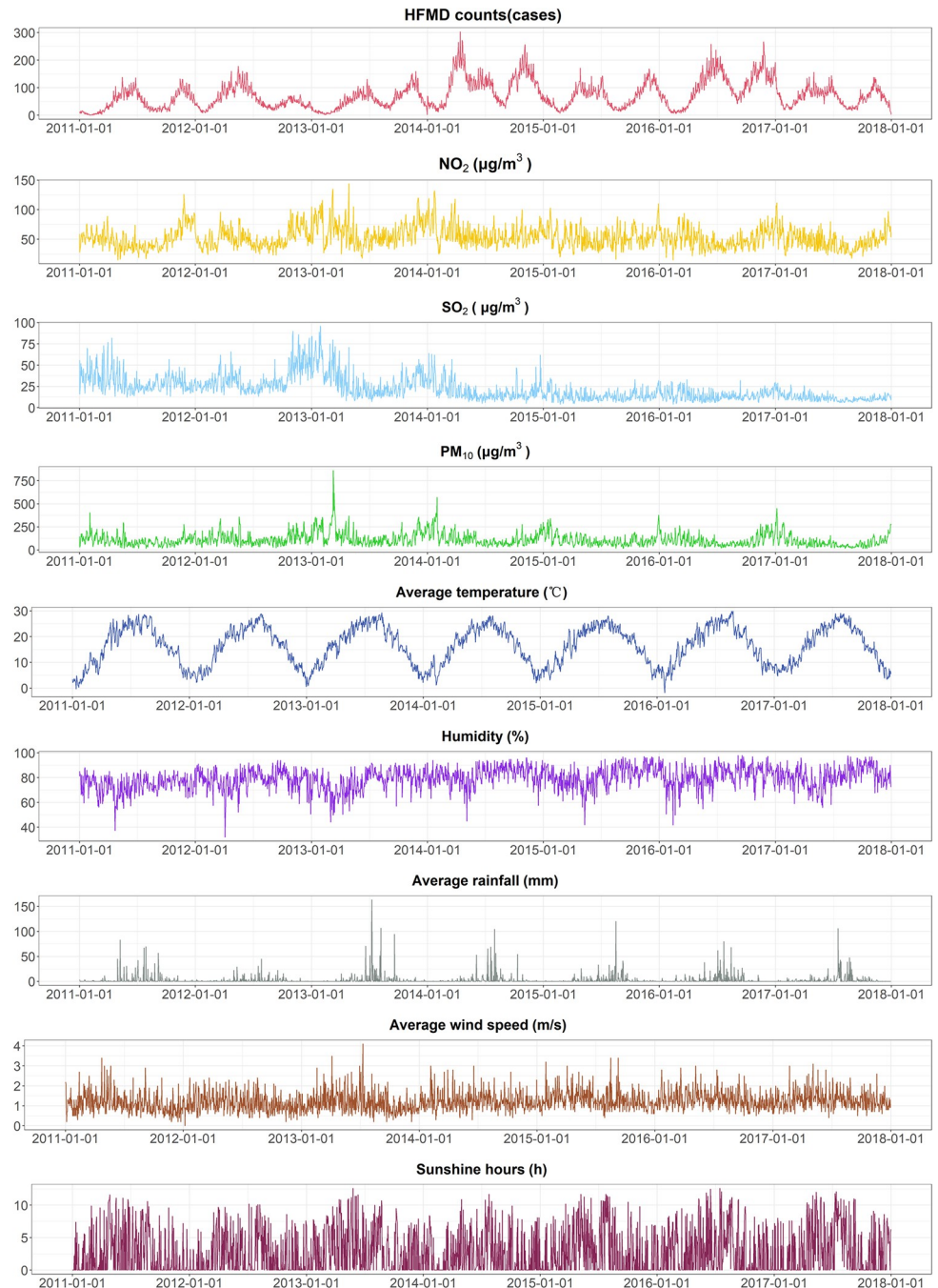


Fig 7. Time series of daily HFMD, meteorological and air pollution variables in Chengdu from 2011 to 2017.

<https://doi.org/10.1371/journal.pntd.0011587.g007>

consistency between predicted and actual values, the Seq2Seq-Shih model showed better performance for each forecasting horizon, obtaining the largest correlation of 0.887 at forecasting horizon T+2.

Fig 9 indicates that the four models were able to predict the overall trend of HFMD. As the forecasting step increased, the delay in predicting the peak increased and the intensity of the

Table 1. Descriptive analysis of daily HFMD cases and meteorological and air pollution variables in Chengdu from 2011 to 2017.

Variables	Mean±SD	Min	P ₂₅	Median	P ₇₅	Max
HFMD (cases)	72.2±48.46	0	34	64	100	303
wind speed (m/s)	1.22±0.48	0	0.90	1.10	1.50	4.10
sunshine duration (h)	2.74±3.32	0	0	1	5.10	12.60
temperature (°C)	16.52±7.4	-1.80	9.70	17.40	23	29.90
relative humidity (%)	79.32±8.64	32	74.10	80	85.30	98
rainfall (mm)	2.65±9.41	0	0	0	1.20	163.40
PM ₁₀ (µg/m ³)	112.14±71.43	15	63	95	141	862
SO ₂ (µg/m ³)	21.96±13.79	4	12	18	28	96
NO ₂ (µg/m ³)	53.36±17.98	15	40	50	63	144

Note: SD represents standard deviation, P_x represents the xth percentile.

<https://doi.org/10.1371/journal.pntd.0011587.t001>

peak was also underestimated. Overall, the prediction curves of the Seq2Seq-Shih model were generally the closest to the observed curves.

3.3. Comparison of models in peak prediction

Table 3 shows the average time delay (difference between the predicted and actual peak day) predicted by the model for the two peaks in the test set. At forecasting horizon T+3, the predicted peaks were delayed by 5 days compared to the actual peaks. The Seq2Seq-Shih model

Table 2. Comparison of evaluation metrics of models at different forecasting horizons.

Metrics	Model	Forecasting horizons					
		2	3	6	9	12	15
RMSE	LSTM	16.358 (0.301)	17.361 (0.263)	20.113 (4.915)	22.288 (3.769)	23.635 (1.773)	25.512 (1.053)
	Seq2Seq	15.552 (0.418)	16.290 (0.536)	17.587 (0.700)	20.816 (1.670)	23.971 (1.109)	26.118 (0.837)
	Seq2Seq-Luong	17.182 (1.577)	19.765 (4.282)	19.255 (1.151)	21.539 (1.385)	23.184 (1.922)	25.413 (1.048)
	Seq2Seq-Shih	13.943 (1.418)	14.337 (0.900)	15.150 (1.253)	18.635 (0.505)	20.239 (0.539)	22.192 (1.124)
sMAPE	LSTM	21.522 (0.485)	22.701 (0.451)	28.279 (6.250)	31.162 (5.005)	32.861 (2.856)	35.229 (1.645)
	Seq2Seq	19.719 (1.232)	20.498 (1.233)	22.306 (1.372)	26.372 (2.915)	30.505 (1.847)	32.655 (1.434)
	Seq2Seq-Luong	27.134 (5.813)	26.852 (5.394)	27.783 (5.657)	31.094 (6.489)	30.239 (5.748)	34.585 (2.970)
	Seq2Seq-Shih	17.880 (2.291)	19.525 (1.997)	19.839 (2.661)	24.445 (1.622)	26.192 (1.322)	27.937 (1.346)
PCC	LSTM	0.871 (0.001)	0.859 (0.002)	0.808 (0.010)	0.737 (0.007)	0.661 (0.018)	0.576 (0.014)
	Seq2Seq	0.858 (0.002)	0.844 (0.004)	0.818 (0.008)	0.760 (0.012)	0.712 (0.013)	0.664 (0.018)
	Seq2Seq-Luong	0.855 (0.005)	0.845 (0.010)	0.795 (0.011)	0.740 (0.017)	0.688 (0.029)	0.632 (0.021)
	Seq2Seq-Shih	0.887 (0.001)	0.884 (0.002)	0.868 (0.003)	0.786 (0.007)	0.753 (0.014)	0.705 (0.017)

Note: (x) represents the standard deviation of the metrics for 10 replicate experiments.

<https://doi.org/10.1371/journal.pntd.0011587.t002>

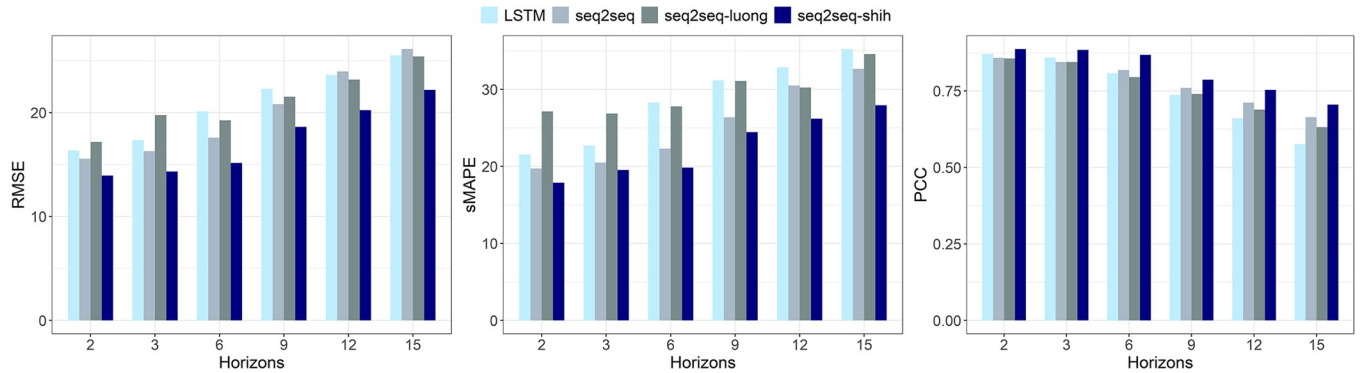


Fig 8. The RMSE, sMAPE and PCC values of models at different forecasting horizons.

<https://doi.org/10.1371/journal.pntd.0011587.g008>

had the smallest time delay for all forecasting horizons, with a 2-day delay in predicting the peak for the next two days and a 16.5-day delay in predicting the peak within the next 15 days. The Seq2Seq-Luong model had the largest average time delay for all forecasting horizons.

Table 4 illustrates the average relative errors of the models in predicting the intensities of the two peaks in the test set. The maximum average relative errors of the LSTM, Seq2Seq, Seq2Seq-Luong and Seq2Seq-Shih models were 0.258, 0.203, 0.271, and 0.207, respectively. The LSTM model had the smallest relative errors at forecasting horizons T+2 and T+3, and the Seq2Seq-Shih model had the second smallest. At forecasting horizon T+6, the peak predicted by the Seq2Seq-Shih model was the closest to the actual peak. The relative error of the peak predicted by the Seq2Seq model was the smallest at forecasting horizons T+9, T+12 and

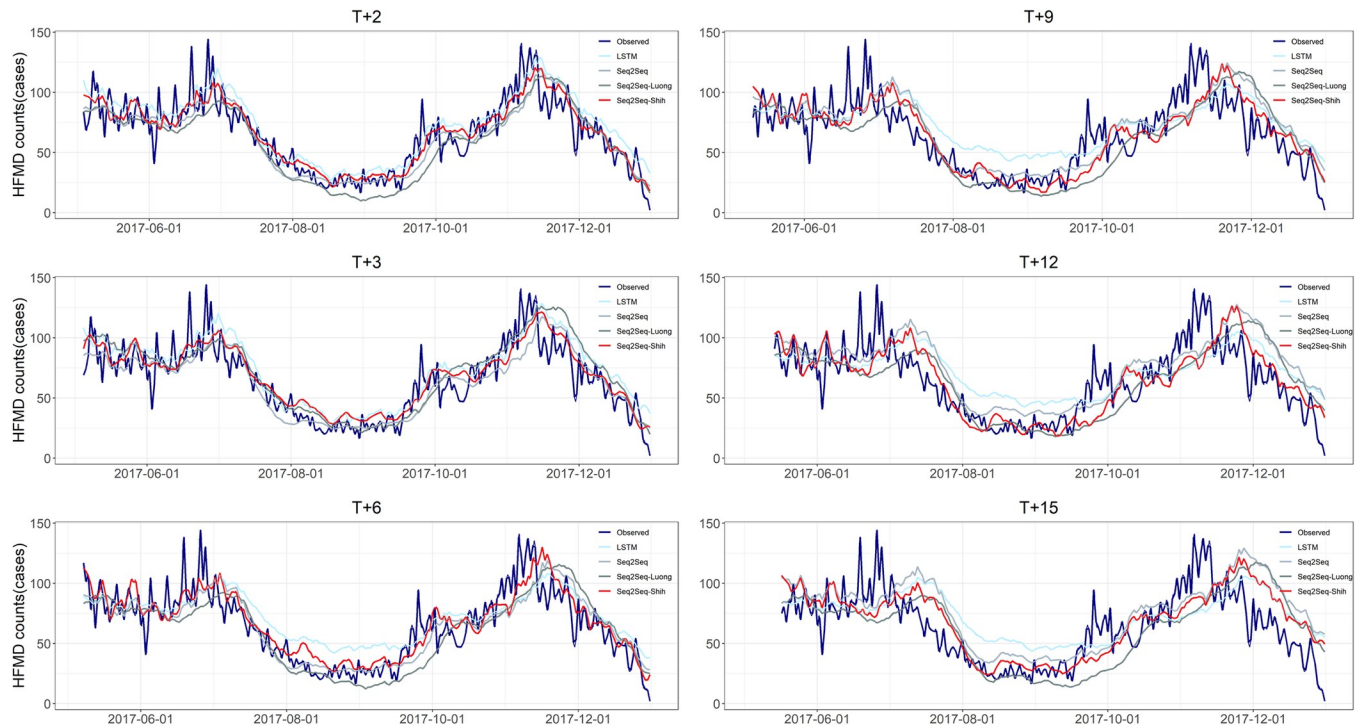


Fig 9. Predictive performance of LSTM, Seq2Seq, Seq2Seq-Luong and Seq2Seq-Shih models for the test set at forecasting steps 2, 3, 6, 9, 12 and 15.

<https://doi.org/10.1371/journal.pntd.0011587.g009>

Table 3. Comparison of peak prediction time delay at different forecasting horizons (unit: days).

Model	Forecasting horizons (<i>h</i>)					
	2	3	6	9	12	15
LSTM	4	5	8	11	14	17
Seq2Seq	4	5	8	11	14	17
Seq2Seq-Luong	5	5	11.5	14.5	17.5	26
Seq2Seq-Shih	2	5	7	10	13	16.5

<https://doi.org/10.1371/journal.pntd.0011587.t003>

T+15, followed by the Seq2Seq-Shih model. Although the LSTM and Seq2Seq models performed better in peak intensity prediction at some forecasting horizons, they had larger time delays. In addition, they were not as good as the Seq2Seq-Shih model in overall prediction (Fig 9). Therefore, the Seq2Seq-Shih model performed better.

4. Discussion

In this study, we evaluated the performance of four deep learning models in forecasting future multi-day trends and detecting peaks in the number of HFMD cases in Chengdu. The Seq2Seq-Shih model performed best in overall prediction, and was able to detect a possible peak incidence within half a month, 17 days in advance. This study provides suggestions for multi-day ahead prediction of HFMD to help local health departments respond to upcoming outbreaks in a timely and rapid manner.

The number of infectious disease cases has a strong autocorrelation, with the number of infections at the current moment correlated with the number of recent cases [46]. Therefore, in this study, the time step parameter was determined from the autocorrelation coefficient. The time step parameter was set to 14 for all deep learning models, which is consistent with previous studies on the association between meteorological and air pollution variables and HFMD in Chengdu. There were time lags between environmental factors and the risk of HFMD, with a temperature lag of 0–10 days, a relative humidity lag of 0 day, a wind speed lag of 0–6 days, a PM₁₀ lag of 0–14 days, a SO₂ lag of 0–14 days, and a NO₂ lag of 0–7 days [11,12,47]. This allowed models to capture the impact of environmental factors on HFMD.

The Seq2Seq model has been used in the time series prediction of thermal load [48], temperature [49,50] and electric vehicle charging demand [51]. In the field of infectious diseases, only malaria [52] and COVID-19 [53] have been studied using this model. Previous studies have found that the Seq2Seq model performs better than the LSTM model [50,52,54]. We also found that the Seq2Seq model outperformed the LSTM model in terms of sMAPE and PCC as well as peak relative error when the forecasting step was greater than 6 days in this study. This may be due to that the Seq2Seq model takes into account the correlation between the number of HFMD cases when making multi-step predictions, making it superior to the LSTM model when making long-term predictions. Surprisingly, the Seq2Seq-Luong model performed poorly in this study, although it showed stronger predictive power than the standalone in

Table 4. Comparison of peak prediction error at different forecasting horizons (metrics: magnitude error).

Model	Forecasting horizons (<i>h</i>)					
	2	3	6	9	12	15
LSTM	0.116	0.114	0.238	0.246	0.258	0.252
Seq2Seq	0.193	0.203	0.196	0.156	0.134	0.135
Seq2Seq-Luong	0.264	0.196	0.258	0.251	0.271	0.258
Seq2Seq-Shih	0.187	0.190	0.151	0.175	0.164	0.207

<https://doi.org/10.1371/journal.pntd.0011587.t004>

other applications [31,45,54]. The advantage of the Luong attention mechanism is that it can learn the importance of the data with different lags in the input sequence for the predicted values; however, this advantage does not seem to be reflected in this study, and regarding this result, we may need more research to determine whether the Seq2Seq-Luong model is truly inferior to the Seq2Seq model in HFMD prediction. In addition, the Seq2Seq-Shih model outperformed the Seq2Seq-Luong and Seq2Seq models, which is consistent with previous findings [31]. This indicates that for the prediction of HFMD, considering the contributions of input variables with different lags to the prediction improves the prediction performance.

Previous studies have used various statistical errors to measure the closeness of the predicted and observed sequences, such as root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). These errors provide a general comparison of two sequences. The peak of an epidemic is one of the most important focuses, and knowledge of its magnitude and timing is vital from the perspective of health service providers [55]. The earlier an outbreak is detected, the sooner public health departments can trigger control measures [56]. Therefore, it is also important to assess the performance of peak predictions when making epidemic predictions. For example, Xu et al. used week differences to measure the performance in predicting influenza peaks [57]. Ertem et al. performed model evaluation using peak week error and peak magnitude error metrics [58]. From the perspective of applications, accurate overall forecasting provides epidemiological trends and guides early planning and resource allocation [59], and accurate forecasting of peak times and intensities informs public health departments of changes in the demand for local resources, as well as combines thresholds for outbreak warning [60]. In this study, we evaluated peak predictions using the time delay and magnitude relative error in addition to prediction error and correlation coefficient metrics. The results showed that the Seq2Seq-Shih model had the smallest difference between the predicted peak day and the actual peak day for all forecasting horizons. There was no model consistently had the smallest relative error in peak prediction. However, the Seq2Seq-Shih model remained in the top two performance in peak magnitude prediction. This suggests that the Seq2Seq-Shih model is able to predict the upcoming peak accurately with a smaller time delay.

Previous studies with multi-step prediction of HFMD were conducted at a weekly or monthly scales [14,18]. In terms of model training, integrating daily data into weekly or monthly data results in a smaller sample size of data, which can easily cause the model to be overfitted. Moreover, if meteorological and air pollution data are aggregated into weekly or monthly averages on a daily basis, the number of HFMD cases is inaccurately predicted. In contrast, data with finer temporal resolution reflect trends more accurately and therefore may improve the performance of the model [61]. In addition, daily-scale forecasts of the number of HFMD cases can provide more timely information to facilitate adequate preparation of local medical departments for a possible upcoming HFMD outbreak. Furthermore, our study is the first to predict HFMD trends using the Seq2Seq model, and the results consistently showed that the Seq2Seq-Shih model performed better in overall and peak prediction.

Our study also has some limitations. First, social factors including population flow and population density also influence the spread of HFMD. Further studies are needed to collect and include relevant data in the model. Second, the results obtained in this study were based on Chengdu city. Further studies in other regions are needed to validate the advantages of the Seq2Seq-Shih model for HFMD prediction.

5. Conclusions

This study evaluated the performance of four deep learning models in predicting multi-day trends of HFMD based on Chengdu city. The Seq2Seq-Shih model showed high accuracy in

predicting future multi-day the number of HFMD cases, along with accurate prediction of peak time and intensity. This study can help public health departments monitor HFMD, understand future HFMD trends and deploy prevention and control measures in advance of an upcoming peak.

Supporting information

S1 Table. Comparison of models used in this study with ARIMA, CNN, XGBoost and random forest models.

(DOCX)

S2 Table. The hyperparameters adjustment of models in this study at all forecasting horizons (the best parameters are marked in gray for the background).

(DOCX)

S3 Table. The results of the local sensitivity analysis of models in this study at forecasting horizon T+2.

(DOCX)

S1 Fig. The Spearman correlation coefficients of HFMD with meteorological and air pollution variables in Chengdu, 2011–2017.

(TIF)

S2 Fig. The overall modeling flow chart of this study.

(TIF)

Author Contributions

Conceptualization: Xiaoran Geng.

Data curation: Wennian Cai, Yuanyi Zha, Huadong Zhang.

Formal analysis: Xiaoran Geng.

Funding acquisition: Yue Ma, Tao Zhang, Fei Yin.

Methodology: Xiaoran Geng.

Project administration: Fei Yin.

Software: Xiaoran Geng, Yue Ma.

Supervision: Yue Ma, Tao Zhang, Fei Yin, Tiejun Shui.

Visualization: Wennian Cai, Changhong Yang.

Writing – original draft: Xiaoran Geng.

Writing – review & editing: Yue Ma, Tao Zhang, Fei Yin.

References

1. Xing W, Liao Q, Viboud C, Zhang J, Sun J, Wu JT, et al. Hand, foot, and mouth disease in China, 2008–12: an epidemiological study. *Lancet Infect Dis.* 2014; 14:308–18. [https://doi.org/10.1016/S1473-3099\(13\)70342-6](https://doi.org/10.1016/S1473-3099(13)70342-6) PMID: 24485991
2. ROBINSON CR, DOANE FW, RHODES AJ. Report of an outbreak of febrile illness with pharyngeal lesions and exanthem: Toronto, summer 1957; isolation of group A Coxsackie virus. *Can Med Assoc J.* 1958; 79:615–21. PMID: 13585281

3. Aswathyraj S, Arunkumar G, Alidjinou EK, Hober D. Hand, foot and mouth disease (HFMD): emerging epidemiology and the need for a vaccine strategy. *Med Microbiol Immunol (Berl)*. 2016; 205:397–407. <https://doi.org/10.1007/s00430-016-0465-y> PMID: 27406374
4. Zhu Z, Zhu S, Guo X, Wang J, Wang D, Yan D, et al. Retrospective seroepidemiology indicated that human enterovirus 71 and coxsackievirus A16 circulated widely in central and southern China before large-scale outbreaks from 2008. *Virology*. 2010; 7:300. <https://doi.org/10.1186/1743-422X-7-300> PMID: 21050463
5. National Health Commission of the People's Republic of China. <http://www.nhc.gov.cn/wjw/index.shtml>. Accessed 18 Jul 2022.
6. Zheng Y, Jit M, Wu JT, Yang J, Leung K, Liao Q, et al. Economic costs and health-related quality of life for hand, foot and mouth disease (HFMD) patients in China. *PLOS ONE*. 2017; 12:e0184266. <https://doi.org/10.1371/journal.pone.0184266> PMID: 28934232
7. Cheng Q, Bai L, Zhang Y, Zhang H, Wang S, Xie M, et al. Ambient temperature, humidity and hand, foot, and mouth disease: A systematic review and meta-analysis. *Sci Total Environ*. 2018; 625:828–36. <https://doi.org/10.1016/j.scitotenv.2018.01.006> PMID: 29306826
8. S Y, L W, Y D, H L, Y L, Q L, et al. Short-Term Effects of Meteorological Factors and Air Pollutants on Hand, Foot and Mouth Disease among Children in Shenzhen, China, 2009–2017. *Int J Environ Res Public Health*. 2019; 16. <https://doi.org/10.3390/ijerph16193639> PMID: 31569796
9. Qi H, Li Y, Zhang J, Chen Y, Guo Y, Xiao S, et al. Quantifying the risk of hand, foot, and mouth disease (HFMD) attributable to meteorological factors in East China: A time series modelling study. *Sci Total Environ*. 2020; 728:138548. <https://doi.org/10.1016/j.scitotenv.2020.138548> PMID: 32361359
10. Wei Q, Wu J, Zhang Y, Cheng Q, Bai L, Duan J, et al. Short-term exposure to sulfur dioxide and the risk of childhood hand, foot, and mouth disease during different seasons in Hefei, China. *Sci Total Environ*. 2019; 658:116–21. <https://doi.org/10.1016/j.scitotenv.2018.11.481> PMID: 30577010
11. Yin F, Ma Y, Zhao X, Lv Q, Liu Y, Li X, et al. Analysis of the effect of PM10 on hand, foot and mouth disease in a basin terrain city. *Sci Rep*. 2019; 9:3233. <https://doi.org/10.1038/s41598-018-35814-5> PMID: 30824722
12. Peng H, Chen Z, Cai L, Liao J, Zheng K, Li S, et al. Relationship between meteorological factors, air pollutants and hand, foot and mouth disease from 2014 to 2020. *BMC Public Health*. 2022; 22:998. <https://doi.org/10.1186/s12889-022-13365-9> PMID: 35581574
13. Liu L, Luan RS, Yin F, Zhu XP, Lü Q. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. *Epidemiol Infect*. 2016; 144:144–51. <https://doi.org/10.1017/S0950268815001144> PMID: 26027606
14. Liu S, Chen J, Wang J, Wu Z, Wu W, Xu Z, et al. Predicting the outbreak of hand, foot, and mouth disease in Nanjing, China: a time-series model based on weather variability. *Int J Biometeorol*. 2018; 62:565–74. <https://doi.org/10.1007/s00484-017-1465-3> PMID: 29086082
15. Tian CW, Wang H, Luo XM. Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol Infect*. 2019; 147:e82. <https://doi.org/10.1017/S095026881800362X> PMID: 30868999
16. Liu W, Bao C, Zhou Y, Ji H, Wu Y, Shi Y, et al. Forecasting incidence of hand, foot and mouth disease using BP neural networks in Jiangsu province, China. *BMC Infect Dis*. 2019; 19:828. <https://doi.org/10.1186/s12879-019-4457-6> PMID: 31590636
17. Meng D, Xu J, Zhao J. Analysis and prediction of hand, foot and mouth disease incidence in China using Random Forest and XGBoost. *PloS One*. 2021; 16:e0261629. <https://doi.org/10.1371/journal.pone.0261629> PMID: 34936688
18. Wang Y, Cao Z, Zeng D, Wang X, Wang Q. Using deep learning to predict the hand-foot-and-mouth disease of enterovirus A71 subtype in Beijing from 2011 to 2018. *Sci Rep*. 2020; 10:12201. <https://doi.org/10.1038/s41598-020-68840-3> PMID: 32699245
19. Xie C, Wen H, Yang W, Cai J, Zhang P, Wu R, et al. Trend analysis and forecast of daily reported incidence of hand, foot and mouth disease in Hubei, China by Prophet model. *Sci Rep*. 2021; 11:1445. <https://doi.org/10.1038/s41598-021-81100-2> PMID: 33446859
20. Yoshida K, Fujimoto T, Muramatsu M, Shimizu H. Prediction of hand, foot, and mouth disease epidemics in Japan using a long short-term memory approach. *PLoS ONE*. 2022; 17:e0271820. <https://doi.org/10.1371/journal.pone.0271820> PMID: 35900968
21. Chen Y, Chu CW, Chen MIC, Cook AR. The utility of LASSO-based models for real time forecasts of endemic infectious diseases: A cross country comparison. *J Biomed Inform*. 2018; 81:16–30. <https://doi.org/10.1016/j.jbi.2018.02.014> PMID: 29496631
22. Jayaraj VJ, Hoe VCW. Forecasting HFMD Cases Using Weather Variables and Google Search Queries in Sabah, Malaysia. *Int J Environ Res Public Health*. 2022; 19:16880.

23. Verma S, Razzaque MA, Sangtongdee U, Arpnikanondt C, Tassaneetrithep B, Arthan D, et al. Hand, Foot, and Mouth Disease in Thailand: A Comprehensive Modelling of Epidemic Dynamics. *Comput Math Methods Med.* 2021; 2021:6697522. <https://doi.org/10.1155/2021/6697522> PMID: 33747118
24. Cox DR. Prediction by Exponentially Weighted Moving Averages and Related Methods. *J R Stat Soc Ser B Methodol.* 1961; 23:414–22.
25. Chevillon G. DIRECT MULTI-STEP ESTIMATION AND FORECASTING. *J Econ Surv.* 2007; 21:746–85.
26. Bontempi G. Long term time series prediction with multi-input multi-output local learning. *Proc 2nd Eur Symp Time Ser Predict TSP ESTSP08.* 2008.
27. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput.* 1997; 9:1735–80. <https://doi.org/10.1162/neco.1997.9.8.1735> PMID: 9377276
28. Alassafi MO, Jarrah M, Alotaibi R. Time series predicting of COVID-19 based on deep learning. *Neuro-computing.* 2022; 468:335–44. <https://doi.org/10.1016/j.neucom.2021.10.035> PMID: 34690432
29. Kara A. Multi-step influenza outbreak forecasting using deep LSTM network and genetic algorithm. *Expert Syst Appl.* 2021; 180:115153.
30. Nguyen V-H, Tuyet-Hanh TT, Mulhall J, Minh HV, Duong TQ, Chien NV, et al. Deep learning models for forecasting dengue fever based on climate data in Vietnam. *PLoS Negl Trop Dis.* 2022; 16:e0010509. <https://doi.org/10.1371/journal.pntd.0010509> PMID: 35696432
31. Yin C, Dai Q. A deep multivariate time series multistep forecasting network. *Appl Intell.* 2022; 52:8956–74.
32. Sutskever I, Vinyals O, Le Q. Sequence to Sequence Learning with Neural Networks. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, editors. *Advances in Neural Information Processing Systems 27* (nips 2014). La Jolla: Neural Information Processing Systems (nips); 2014.
33. Bian J, Wang L, Scherer R, Wozniak M, Zhang P, Wei W. Abnormal Detection of Electricity Consumption of User Based on Particle Swarm Optimization and Long Short Term Memory With the Attention Mechanism. *IEEE Access.* 2021; 9:47252–65.
34. Han Y, Chen Z, Zheng K, Li X, Kong J, Duan X, et al. Epidemiology of Hand, Foot, and Mouth Disease Before and After the Introduction of Enterovirus 71 Vaccines in Chengdu, China, 2009–2018. *Pediatr Infect Dis J.* 2020; 39:969. <https://doi.org/10.1097/INF.0000000000002745> PMID: 32433221
35. Han Y, Zheng K, Chen Z, Li X, Kong J, Duan X, et al. Epidemiological characteristics of hand, foot, and mouth disease before the introduction of enterovirus 71 vaccines in Chengdu, China. *Int J Infect Dis.* 2020; 101:347.
36. Luong T, Pham H, Manning CD. Effective Approaches to Attention-based Neural Machine Translation. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing.* Lisbon, Portugal: Association for Computational Linguistics; 2015. p. 1412–21.
37. Shih S-Y, Sun F-K, Lee H. Temporal pattern attention for multivariate time series forecasting. *Mach Learn.* 2019; 108:1421–41.
38. Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. In: 1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470). 1999. p. 850–5 vol. 2.
39. Sak H, Senior A, Beaufays F. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.* Singapore, Singapore; 2014. p. 338–42.
40. Malhotra P, Vig L, Shroff G, Agarwal P. Long Short Term Memory networks for anomaly detection in time series. In: *23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2015—Proceedings.* Bruges, Belgium; 2015. p. 89–94.
41. Luong M-T, Le QV, Sutskever I, Vinyals O, Kaiser L. Multi-task sequence to sequence learning. In: *4th International Conference on Learning Representations, ICLR 2016—Conference Track Proceedings.* San Juan, Puerto rico; 2016.
42. Zhou Y, Li Y, Wang D, Liu Y. A multi-step ahead global solar radiation prediction method using an attention-based transformer model with an interpretable mechanism. *Int J Hydrog Energy.* 2023; 48:15317–30.
43. Zhang G, Bai X, Wang Y. Short-time multi-energy load forecasting method based on CNN-Seq2Seq model with attention mechanism. *Mach Learn Appl.* 2021; 5:100064.
44. Zhang Y, Li Y, Zhang G. Short-term wind power forecasting approach based on Seq2Seq model using NWP data. *Energy.* 2020; 213:118371.
45. Li G, Li F, Ahmad T, Liu J, Li T, Fang X, et al. Performance evaluation of sequence-to-sequence-Attention model for short-term multi-step ahead building energy predictions. *Energy.* 2022;259:124915.

46. Imai C, Armstrong B, Chalabi Z, Mangtani P, Hashizume M. Time series regression model for infectious disease and weather. *Environ Res.* 2015; 142:319–27. <https://doi.org/10.1016/j.envres.2015.06.040> PMID: 26188633
47. Yin F, Zhang T, Liu L, Lv Q, Li X. The Association between Ambient Temperature and Childhood Hand, Foot, and Mouth Disease in Chengdu, China: A Distributed Lag Non-linear Analysis. *Sci Rep.* 2016; 6:27305. <https://doi.org/10.1038/srep27305> PMID: 27248051
48. Lu Y, Tian Z, Zhou R, Liu W. Multi-step-ahead prediction of thermal load in regional energy system using deep learning method. *Energy Build.* 2021; 233:110658.
49. Tabrizi SE, Xiao K, Van Griensven Thé J, Saad M, Farghaly H, Yang SX, et al. Hourly road pavement surface temperature forecasting using deep learning models. *J Hydrol.* 2021; 603:126877.
50. Fang Z, Crimier N, Scanu L, Midelet A, Alyafi A, Delinchant B. Multi-zone indoor temperature prediction with LSTM-based sequence to sequence model. *Energy Build.* 2021; 245:111053.
51. Yi Z, Liu XC, Wei R, Chen X, Dai J. Electric vehicle charging demand forecasting using deep learning model. *J Intell Transp Syst.* 2022; 26:690–703.
52. Kamana E, Zhao J, Bai D. Predicting the impact of climate change on the re-emergence of malaria cases in China using LSTMSeq2Seq deep learning model: a modelling and prediction analysis study. *BMJ Open.* 2022; 12:e053922. <https://doi.org/10.1136/bmjopen-2021-053922> PMID: 35361642
53. Kim Y, Park C-R, Ahn J-P, Jang B. COVID-19 outbreak prediction using Seq2Seq + Attention and Word2Vec keyword time series data. *PLOS ONE.* 2023; 18:e0284298. <https://doi.org/10.1371/journal.pone.0284298> PMID: 37099535
54. Du S, Li T, Yang Y, Horng S-J. Multivariate time series forecasting via attention-based encoder–decoder framework. *Neurocomputing.* 2020; 388:269–79.
55. Tabataba FS, Chakraborty P, Ramakrishnan N, Venkatramanan S, Chen J, Lewis B, et al. A framework for evaluating epidemic forecasts. *BMC Infect Dis.* 2017; 17:345. <https://doi.org/10.1186/s12879-017-2365-1> PMID: 28506278
56. Herrera JL, Srinivasan R, Brownstein JS, Galvani AP, Meyers LA. Disease Surveillance on Complex Social Networks. *PLoS Comput Biol.* 2016; 12:e1004928. <https://doi.org/10.1371/journal.pcbi.1004928> PMID: 27415615
57. Xu Q, Gel YR, Ramirez Ramirez LL, Nezafati K, Zhang Q, Tsui K-L. Forecasting influenza in Hong Kong with Google search queries and statistical model fusion. *PLOS ONE.* 2017; 12:e0176690. <https://doi.org/10.1371/journal.pone.0176690> PMID: 28464015
58. Ertem Z, Raymond D, Meyers LA. Optimal multi-source forecasting of seasonal influenza. *PLOS Comput Biol.* 2018; 14:e1006236. <https://doi.org/10.1371/journal.pcbi.1006236> PMID: 30180212
59. Pley C, Evans M, Lowe R, Montgomery H, Yacoub S. Digital and technological innovation in vector-borne disease surveillance to predict, detect, and control climate-driven outbreaks. *Lancet Planet Health.* 2021; 5:e739–45. [https://doi.org/10.1016/S2542-5196\(21\)00141-8](https://doi.org/10.1016/S2542-5196(21)00141-8) PMID: 34627478
60. Nsoesie EO, Brownstein JS, Ramakrishnan N, Marathe MV. A systematic review of studies on forecasting the dynamics of influenza outbreaks. *Influenza Other Respir Viruses.* 2014; 8:309–16. <https://doi.org/10.1111/irv.12226> PMID: 24373466
61. Zimmer C, Leuba SI, Yaesoubi R, Cohen T. Use of daily Internet search query data improves real-time projections of influenza epidemics. *J R Soc Interface.* 2018; 15:20180220. <https://doi.org/10.1098/rsif.2018.0220> PMID: 30305417