Original Article

Prediction of Trends and Bioclimatic Factors Influencing the Monthly Incidence of Zoonotic Cutaneous Leishmaniasis Using Arima and Sarima Time Series Models in Maraveh Tappeh County, Golestan Province, Iran

Rezvan Rajabzadeh[1], Seyed Hamid Hosseini[2], *Aioub Sofizadeh[3]

[1]Assistant Professor of Epidemiology, Non-Communicable Diseases Research Center, Sabzevar University of Medical Sciences, Sabzevar, Iran
[2]Assistant Professor, Non-Communicable Diseases Research Center, Sabzevar University of Medical Sciences, Sabzevar, Iran
[3]Infectious Diseases Research Center, Golestan University of Medical Sciences, Gorgan, Iran

***Corresponding Author:** Dr Aioub Sofizadeh, E-mail: a_sofizadeh@yahoo.com

(Received 10 Sep 2025; accepted 16 Dec 2025)

Abstract

Background: Zoonotic cutaneous leishmaniasis (ZCL) is a significant vector-borne disease in northeastern Iran, strongly affected by climatic conditions. Maraveh Tappeh County in Golestan Province is an endemic area with considerable annual case numbers. This study aimed to predict monthly ZCL trends and identify key bioclimatic factors influencing disease occurrence using ARIMA and SARIMA time series models.

Methods: This analytical cross-sectional study used monthly confirmed ZCL case data from 2003 to 2018, obtained from the Maraveh Tappeh County Health Center. Climatic variables, including temperature indices, relative humidity indices, total monthly precipitation and number of rainy days, were collected from the local meteorological office. Stationarity was assessed using the Augmented Dickey–Fuller test and autocorrelation patterns were evaluated through ACF and PACF plots. ARIMA (AutoRegressive Integrated Moving Average) and SARIMA (Seasonal AutoRegressive Integrated Moving Average) models were developed, with the optimal model selected based on AIC and BIC criteria. Cross-correlation analysis examined associations between climatic variables and ZCL incidence at lags of 0–5 months.

Results: A total of 1,301 ZCL cases were reported over the 16 years, with marked monthly and seasonal variability. Incidence peaked in November and reached its lowest level in June. The ARIMA (2,0,2)–SARIMA $(0,0,1)_{12}$ model demonstrated the best predictive performance. Significant positive correlations were observed between ZCL incidence and relative humidity, precipitation and number of rainy days at short lags (0–2 months), while inverse associations appeared at longer lags (5 months) ($p<0.05$).

Conclusion: Relative humidity and precipitation are key drivers of ZCL dynamics in Maraveh Tappeh. Incorporating SARIMA models into surveillance systems may improve outbreak prediction and support timely prevention and control strategies.

Keywords: Cutaneous leishmaniasis; *Leishmania major*; Forecasting; Time series analysis; Iran

Introduction

Leishmaniasis, recognized as one of the most significant vector-borne parasitic diseases, is caused by various species of *Leishmania* and transmitted in the Old World through the bite of infected female sand flies of the genus *Phlebotomus*. In Iran, the disease presents in two major clinical and epidemiological forms: cutaneous leishmaniasis (CL) and visceral leishmaniasis (VL). Cutaneous leishmaniasis is further divided into two principal types: Zoonotic cutaneous leishmaniasis (ZCL), primarily caused by *Leishmania major*, with *Phlebotomus papatasi* as the main vector and wild rodents as the reservoir hosts; and anthroponotic cutaneous leishmaniasis (ACL), typically caused by *L. tropica*, with *Ph. sergenti* as the main vec-

http://jad.tums.ac.ir
Published Online: December 30, 2025

tor and humans and dogs serving as the primary and secondary reservoirs, respectively. Zoonotic cutaneous leishmaniasis occurs in different regions of Iran, including central, northeastern, western, south-western and south-eastern areas, where the distribution patterns of vectors and reservoirs vary according to ecological, climatic and seasonal conditions (1).

Golestan Province is considered one of the most important endemic foci of ZCL in Iran, with multiple studies confirming the presence of a complete transmission cycle (2–7). In this province, *L. major* is the predominant causative agent, transmitted mainly by *Ph. papatasi*, which maintains close ecological associations with rodent reservoir hosts, notably *Rhombomys opimus* and *Meriones libycus* (2). According to systematic reviews and analytical data from Golestan Province during 2010–2017, the annual number of reported CL cases ranged from 360 to 1,766, with a mean annual incidence rate of 31.7 per 100,000 population. In the northeastern counties, Gonbad Kavus and Maraveh Tappeh are classified as highly endemic, with incidence rates of 153 and 117 per 1,000 population, respectively, highlighting strong focal clustering and the need for targeted control measures (11, 12). Field observations indicate that demographic changes, expansion of human settlements into reservoir habitats and increased human vector contact are significant contributors to the persistence of active ZCL foci in this region (4–7).

Entomological and zoological surveys indicate significantly higher densities of vector sand flies and wild rodent reservoirs in the northeastern endemic zones of the province compared to other areas (5, 6, 13, 14). This elevated distribution is largely attributed to favorable geographic and climatic conditions, including optimal temperature, relative humidity, vegetation cover and soil properties that support the breeding and survival of *Ph. papatasi* and the reproduction of *R. opimus* and *M. libycus* (13, 14). Therefore, the relationship between these climatic and geographical factors and the distribution of *Ph. papatasi* has been investigated in various studies, for example: In a study conducted in Golestan Province (13), slope, altitude, annual mean temperature and the normalized difference vegetation index (NDVI) were introduced as the most significant factors in the distribution of *Ph. papatasi*. In another Iranian study, the mean temperature of the wettest quarter, slope, precipitation seasonality and precipitation of the wettest quarter were identified as key factors associated with the species' distribution (14).

These climatic and geographical factors also affect the distribution of *R. opimus* as a main reservoir host of ZCL, as in a study conducted in Golestan Province, mean temperature (°C) of driest quarter, maximum temperature (°C) of warmest quarter and altitude from the sea level (m) had more effect on the distribution of *R. opimus* (15). Therefore, it can be said that temperature, precipitation and relative humidity are known as the most significant factors associated with the distribution, growth and development of *Ph. papatasi* and *R. opimus* and also the distribution of ZCL (16).

A time series is a set of time-ordered observations of a process where the intervals between observations remain constant (for example, weeks, months, years and minor deviations in the intervals are acceptable) (17). Time series data are often distinguished from other types of longitudinal data by the number and source of observations and due to its unique structure, a time series exhibits features that are either absent or less prominent in cross-sectional and longitudinal data types (18). One of the most important uses of time series is in predicting diseases such as leishmaniasis, which is very important in the care of this disease. Forecasting of leishmaniasis is done with various tools and software with the goals of identifying active foci and areas at risk of the disease, predicting disease epidemics and responding quickly to them, the scope of the spread of vectors and reservoirs and identifying climatic and geographical factors affecting their spread, of which time series is one of the most important of these tools (19–23).

Epidemiological investigations in Kalaleh, Maraveh Tappeh and Gonbad Kavus Counties have revealed similar transmission patterns, with seasonal peaks in autumn and early winter (3–5) and time series are a very useful tool for predicting seasonal diseases such as cutaneous leishmaniasis, due to their ability to model recurring patterns and accurately identify temporal relationships with environmental factors. This allows for quantitative prediction of future cases and proactive resource management (24, 25).

Given that multiple anthropogenic, biological and environmental factors contribute to the occurrence, endemicity and distribution of leishmaniasis, the disease is inherently sensitive to environmental and climatic variability (26). Evidence indicates that environmental features directly influence the dynamics and distribution of leishmaniasis by affecting the three core elements of its transmission cycle: parasite (*Leishmania* spp.), vectors (sand flies) and reservoirs (wild rodents and other animal hosts), potentially altering the epidemiology of the disease at local, national and global levels (27, 28). Considering the established role of climatic parameters in the occurrence and persistence of CL and recognizing Maraveh Tappeh County as one of the principal foci of the disease in Golestan Province, the present study was conducted in 2019 to forecast CL incidence trends and identify the most influential climatic factors associated with transmission in this area. A time series modeling approach was applied to analyze epidemiological and climatic datasets, enabling scientific monitoring and predictive assessment of the disease's future trajectory. The findings aim to serve as an evidence-based decision-support tool for health authorities and local policymakers in the planning and implementation of targeted control and prevention strategies.

Materials and Methods

Study area

Maraveh Tappeh County, located in the northeast of Golestan Province, borders Turkmenistan to the north, North Khorasan Province to the east and Gonbad Kavus and Kalaleh counties to the south and west, respectively (Fig. 1). Owing to its unique geographical setting ranging from a semi-arid to arid climate, with diverse topography including plains, tropical mountains and forested highlands and the presence of vectors (*Ph. papatasi*) and primary reservoirs (*R. opimus*, *M. libycus*), this county is recognized as one of the most important endemic foci of ZCL in Iran. The study site was selected based on prior epidemiological evidence and a surveillance reporting coverage rate exceeding 95%, ensuring comprehensive data capture from all health service units in the county.

Study design and population

This time-series analytical study was based on a 16-year time-series dataset, enabling the detection of temporal trends, seasonal patterns and short-term forecasts. According to the national guideline, a confirmed case was defined as a patient with a clinically compatible skin lesion and laboratory confirmation through microscopic examination (direct smear with Giemsa staining). The entire covered population (~65,000 individuals) served as the reference population.

Data collection

Monthly epidemiological data on confirmed cutaneous leishmaniasis (CL) cases in Maraveh Tappeh County from 2003 to 2018 were obtained from the County Health Center and verified through the Health Information System (HIS) and standard surveillance forms. Data were analyzed for the 16-year period from January 2003 to December 2018. Model calibration was based on this observed dataset and forecasts were then generated for the subsequent 12-month period (January–December 2019).

Data were entered into structured checklists and imported into Stata (version 14.0; StataCorp LP, College Station, TX, USA) as the dependent variable. All records were aggregated on a monthly basis to reflect the epidemiological pattern of cutaneous leishmaniasis, which typically exhibits distinct seasonal peaks during autumn and early winter and lower incidence from late spring to summer. Monthly climatic variables, including minimum, maximum and mean air temperature (°C); minimum, maximum and mean relative humidity (%RH); total precipitation (mm); and number of rainy days, were obtained from the local meteorological station and entered into the software as independent variables.

A total of 1301 confirmed CL cases were reported during the observed period (2003–2018) with >95% verified surveillance coverage, as documented in the national leishmaniasis registry (Golestan Provincial Center for Communicable Diseases Control). The 159 cases mentioned in the results correspond to the model-predicted total for 2019, representing the 12-month forecast horizon beyond the observed dataset.

Data preprocessing

Data preprocessing was performed in Stata v14.0. Visual inspection using boxplots and statistical screening revealed no extreme outliers in the climatic or epidemiologic time series; therefore, no removal or Winsorization was required. The dataset contained no missing observations. Accordingly, all statistical analyses were performed on the complete dataset without the need for imputation or sensitivity analyses. Non-normally distributed variables were normalized using natural logarithmic transformation. Time-series stationarity was assessed using the Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests; non-stationary series were differenced (regular or seasonal) to achieve stationarity. The selection of climatic lags (0–5 months) was based on the duration of the vector's life cycle (~6–8 weeks) and the incubation period of *L. major* (~2–4 months).

Time series modeling

Time series, defined as sequences of regularly collected observations, allow the identification of trends, seasonal structures and forecasting potential. For predictive modeling, we employed Autoregressive Integrated Moving Average (ARIMA) and its seasonal extension SARIMA (P, D, Q, S), which are among the most robust linear time series models for stationary data. It is important to note that these models were fitted as univariate time series models, using monthly CL incidence as the sole dependent variable. To account for annual seasonality, the seasonal length S was set to 12 months (29, 30). An ARIMA model is characterized by three parameters: p, the order of the autoregressive (AR) component; d, the degree of differencing required to attain stationarity; and q, the order of the moving average (MA) component. In seasonal SARIMA models, the parameters P, D, Q and S are additionally included (31, 32). The optimal parameter combination was determined using autocorrelation (ACF) and partial autocorrelation (PACF) plots, together with information criteria (AIC and BIC).

The ARIMA and SARIMA models applied in this study are linear time series approaches that relate present observations to past values and stochastic errors. Although nonlinear or hybrid alternatives such as ARIMAX, SARIMAX, or machine learning architectures (e.g., GRU, LSTM) can capture more complex dependencies, these methods were not adopted here. The relatively limited dataset length (2003–2018), the study's interpretive epidemiological purpose and the need for transparent parameter interpretation favored the use of parsimonious linear models. ARIMA/SARIMA formulations have been widely employed in infectious disease time series analyses be-cause they offer an effective balance between interpretability and predictive performance, enabling both short-term forecasting and mechanistic epidemiological explanation.

Model evaluation and selection

Candidate models were compared using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with preference given to the model yielding the lowest values. Statistical significance of model coefficients was determined at a 95% confidence level ($p<0.05$). Model residuals were evaluated using the Ljung-Box and Portmanteau tests to verify independence and adequacy.

Since the primary aim of this analysis was to characterize temporal associations between climatic variables and cutaneous leishmaniasis (CL) incidence rather than to develop an operational forecasting model, the entire dataset (January 2003–December 2018) was used for model calibration. Model validity and predictive adequacy were confirmed through residual diagnostics (Ljung–Box and Portmanteau tests) and information criteria (AIC and BIC) rather than through an out-of-sample evaluation period.

After model calibration using data from January 2003 to December 2018, the final ARIMA/SARIMA models were employed to generate short-term forecasts extending 12 months ahead (January–December 2019). This horizon, corresponding to a full annual seasonal cycle, enabled validation of the model's ability to reproduce monthly trends and seasonal peaks beyond the calibration period.

Since monthly CL incidence data represent count outcomes with right-skewed distribution, values were log-transformed [$\log(y+1)$] before ARIMA/SARIMA model fitting to achieve variance stabilization and normality of residuals. Consequently, the estimated parameters and residual standard deviations (e.g., 0.445) presented in Table 2 are reported on the transformed scale. Forecasts were subsequently back-transformed to the original count scale for interpretation and graphical presentation.

Analysis of climatic factors

The association between climatic factors and monthly CL incidence was analyzed using cross-correlation functions (CCF) at lags of 0–5 months. Lag selection was guided by both biological plausibility and statistical evaluation, corresponding to the approximate developmental cycle of *Ph. papatasi* (~6–8 weeks) and the incubation period of *L. major* (~2–4 months). All analyses were performed using Stata v14.0 (StataCorp LP, College Station, TX, USA). Preprocessing and data management were carried out with the generate, replace and summarize commands. Normality of variables was checked using histogram and kdensity and missing values ≤5% were imputed via a centered moving average procedure. Stationarity of each time series was assessed using Augmented Dickey–Fuller (AD-Fuller) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests; where required, differencing was applied using gen Dvar = D.var.

Before conducting cross-correlation analyses, multicollinearity among the climatic predictors was assessed. Variance Inflation Factors (VIF) were calculated using the VIF command, with all values falling below 5, indicating no concerning multicollinearity. This was further supported by examining Pearson correlation matrices, which revealed strong correlations between variables such as minimum and maximum temperatures. Where high correlations were detected, separate analyses were conducted for each climatic variable group to ensure valid inference. Cross-correlation analyses were then conducted using the ccf and corrgram commands and only statistically significant lagged associations ($p<0.05$) were reported. Results were interpreted in terms of short-term predictive effects or lagged impacts.

It is critical to note that all ARIMA and SARIMA models were fitted as univariate time series models, using monthly CL incidence as the sole dependent variable. Climatic variables were not entered as covariates within the ARIMA/SARIMA framework; rather, their tem-

http://jad.tums.ac.ir
Published Online: December 30, 2025

poral associations with CL incidence were assessed separately through CCF analyses to identify time-lagged relationships. Seasonal adjustment and model fitting for the univariate CL series were performed with the ARIMA command and model diagnostics were obtained from estat ic, ac, pac and Portmanteau (Ljung–Box) tests. To prevent overfitting, the final lag configuration for the univariate ARIMA/ SARIMA model was chosen by minimizing the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), ensuring model parsimony, reproducibility and biological consistency in the climatic–epidemiological time series relationships. This approach ensured that the forecasting model remained parsimonious and based solely on the historical incidence data.

Results

From 2003 to 2018, a total of 1301 cases of CL were identified within the covered population of Maraveh Tappeh County. The highest annual incidence was reported in 2018 with 159 cases, whereas the lowest was observed in 2006 with only 10 cases. Monthly distribution analysis indicated that November had the highest incidence, with 397 cases, while June had the lowest, with 10 cases (Fig. 2). This seasonal peak aligns with the behavior and life cycle of the vector *Ph. papatasi*, whose population density peaks in late summer. Considering the incubation period of *L. major* (<4 months), the November peak is biologically plausible.

The autocorrelation plot (Fig. 3) showed a regular sinusoidal pattern with significant peaks at 12- and 24-month lags, indicating stability of the transmission cycle and the annual recurrence of the disease. Such marked seasonality supports the application of seasonal time series models such as SARIMA, which are appropriate for datasets with consistent periodic fluctuations.

Following the fitting of multiple models and comparison using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), root mean square error (RMSE), and residual diagnostics (Ljung Box and Portmanteau tests), the ARIMA (2,0,2) SARIMA $(0,0,1)_{12}$ specification was identified as the optimal predictive model. Beyond its superior statistical performance, this model is theoretically consistent with the annual transmission pattern of CL: the seasonal component (S=12) captures year-to-year fluctuations, while the non-seasonal component models short-term variations and stochastic noise.

As illustrated in Fig. 4, the chosen model successfully predicted overall trends and seasonal peaks for the following year without a substantial rise in predictive error. Comparison of the SARIMA-based forecast with observed ZCL cases in 2019 demonstrated a good agreement in overall trend and seasonal variation, indicating acceptable predictive performance of the model.

At a significance level of 0.05, statistically significant positive associations were identified between monthly CL incidence and specific humidity and precipitation indices at defined lag intervals:

- Mean relative humidity at lag 0 months (β=0.17) and lag 2 months (β=0.14)
- Minimum relative humidity at lag 2 months (β=0.16)
- Maximum relative humidity at lag 0 months (β=0.16)
- Total precipitation at lag 2 months (β=0.21)
- Number of rainy days at lag 0 months (β=0.21) and lag 2 months (β=0.17)

From ecological and epidemiological perspectives, these findings indicate that increased humidity and rainfall improve survival, feeding and breeding conditions for *Ph. papatasi* and stabilize reservoir habitats for *R. opimus* and *M. libycus*. The two-month lag corresponds to the vector's developmental cycle (~6–8 weeks) and the incubation period of *L. major* (up to four months). By fostering vegetation growth and maintaining soil moisture, precipitation and humidity, burrow systems and other microhabitats for both vectors and reservoirs. The two-month delay likely represents the period needed for environmental changes to impact vector abundance and human transmission. In contrast, at longer lags such as five months, inverse associations were occasionally observed, potentially reflecting secondary climatic impacts on vector population dynamics or altered reservoir behavior in dry seasons (Table 1).

The observed patterns align with vector-borne disease dynamics theory, whereby climatic variables exert lag-specific effects on transmission parameters. Under such conditions, ARIMA/SARIMA models combining autoregressive (AR), moving average (MA) and seasonal (SAR/SMA) terms can mathematically capture these dependencies. Given that the series was rendered stationary following the Augmented Dickey Fuller (ADF) test, the use of a model without differencing (d=0) was both appropriate and efficient.

As detailed in Table 2, multiple ARIMA and SARIMA models incorporating various climatic variables were evaluated. Comparisons were made using AIC, residual standard deviation, and the Portmanteau test for residual independence. The models with the lowest AIC values, approximately 245.23 for "number of rainy days" and 805.15 for "total monthly precipitation," and reduced residual dispersion achieved the highest predictive accuracy. The final ARIMA $(2,0,2)(0,0,1)_{12}$ model was deemed adequate, as the Portmanteau test confirmed that the residuals were independent ($p>0.05$). Although some alternative models fitted with temperature (mean, minimum, maximum) or humidity (minimum, maximum, mean) variables occasionally displayed higher implied R^2 values or simpler structures, they ranked lower than the final model in both fit indices and epidemiological consistency with the seasonal disease pattern.

Finally, the ARIMA (2,0,2) SARIMA $(0,0,1)_{12}$ model was selected for its superior statistical fit, relative simplicity and ability to represent the annual CL transmission cycle alongside relevant climatic influences. This selection ensures that forecasts are both statistically robust and biologically consistent with the life cycle of *Ph. papatasi*.

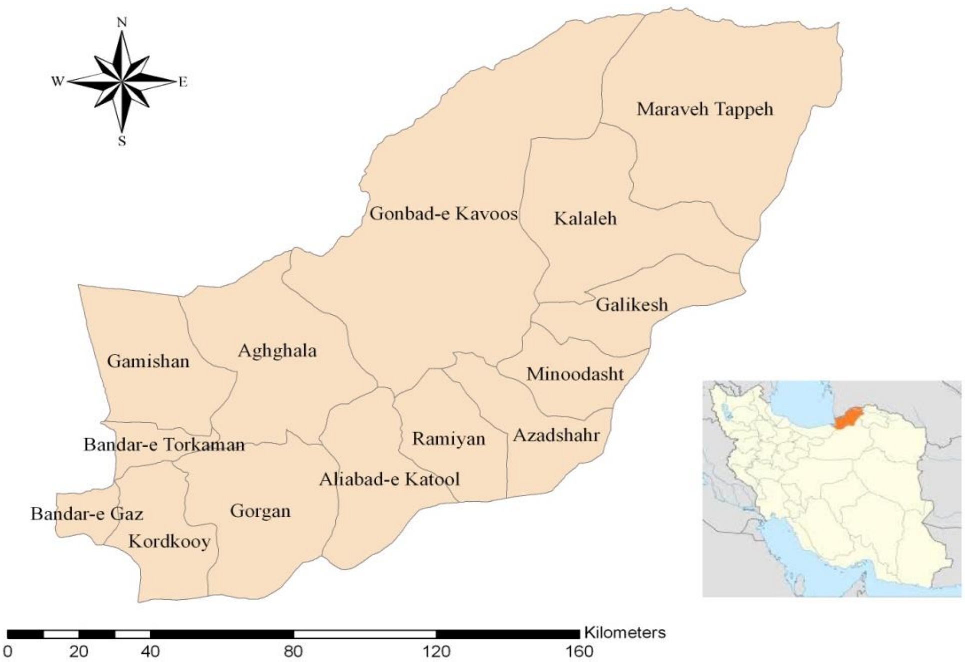

Fig. 1. Geographic location of Maraveh Tappeh County in Golestan Province, Iran

Table 1. Cross-correlation coefficients between climatic variables and monthly cutaneous leishmaniasis cases at lags 0–5 months in Maraveh Tappeh County, Golestan Province, Iran, 2003–2018

Time-lag (months)	Minimum temperature	Maximum temperature	Mean monthly temperature	Minimum relative humidity	Maximum relative humidity	Mean monthly relative humidity	Total monthly precipitation	Number of rainy days
5	0.12	0.08	0.10	0.05	0.06	0.01	-0.15*	-0.17*
4	-0.18*	-0.14*	-0.15*	0.08	0.06	0.07	0.02	0.07
3	-0.09	-0.06	-0.07	0.06	0.04	0.06	-0.02	-0.02
2	-0.15*	-0.16*	-0.16*	0.16*	0.09	0.14*	0.21*	0.17*
1	0.03	0.03	0.03	0.06	0.09	0.11	-0.04	0.01
0	-0.03	-0.04	-0.04	0.12	0.16*	0.17*	0.08	0.21*

*Significant at p< 0.05

Table 2. ARIMA and SARIMA models fitted to monthly cutaneous leishmaniasis data in Maraveh Tappeh County, Golestan Province, Iran, 2003–2018

Model ARIMA(p,d,q)SARIMA(P,D,Q,S)	** AR1	** AR2	** AR3	** MA1	** MA2	** SAR	** SMA	SD of residuals	AIC **	P-value (Portmanteau) residual
ARIMA(2,0,2) SARIMA(1,0,0,12)	1.71*	-0.97*	-	-1.05*	0.41*	-0.10	-	2.39	889.61	0.4836
ARIMA(2,0,2) SARIMA(1,0,0,12)	1.71*	-0.98*	-	-1.14*	0.48*	-0.09	-	1.97	973.206	0.2083
ARIMA(2,0,2) SARIMA(1,0,1,12)	1.72*	-0.98*	-	-1.09*	0.44*	0.55	0.54	2.62	927.76	0.3284
ARIMA(3,0,1) SARIMA(1,0,0,12)	1.18*	-0.19	-0.30*	-0.82*	-	0.16*		7.14	1311.31	0.1427
ARIMA(2,0,2) SARIMA(1,0,1,12)	1.70*	-0.97*	-	-1.56*	0.85*	0.97*	-0.90*	5.95	1269.77	0.4390
ARIMA(1,0,2) SARIMA(1,0,1,12)	-0.85*		-	1.04*	0.25*	0.10*	-0.92*	6.23	1262.37	0.2193
ARIMA(2,0,2) SARIMA(0,0,1,12)	1.73*	-1.0*	-	-	-1.72*	-	1.0	1.92	805.1556	0.8016
ARIMA(2,0,2) SARIMA(0,0,1,12)	1.71*	-1.0*	-	-1.73*	1.0	0.02	0.43*	0.445	245.2338	0.9621

*Significant at p< 0.05; **AIC (Akaike information criteria), AR (autoregressive), MA (moving average), SMA (seasonal moving average), SAR (seasonal autoregressive)

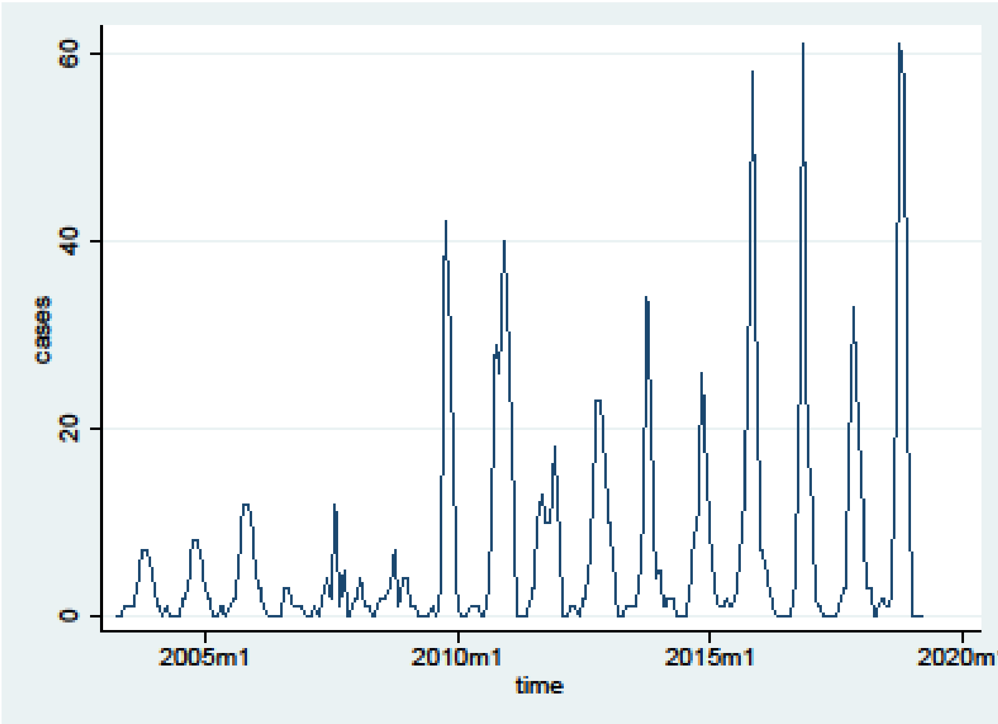

Fig. 2. Temporal trend of monthly cutaneous leishmaniasis cases in Maraveh Tappeh County, Golestan Province, Iran, 2003–2018

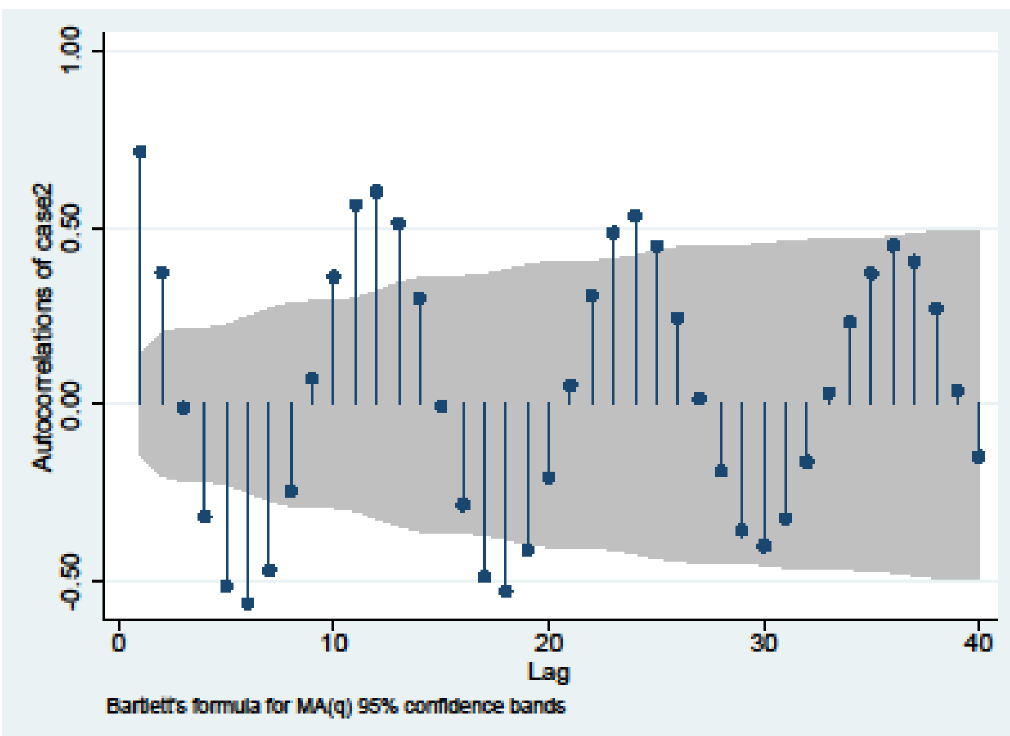

Fig. 3. Autocorrelation function (ACF) of monthly cutaneous leishmaniasis incidence used for time-series model identification in Maraveh Tappeh County, Golestan Province, Iran, 2003–2018

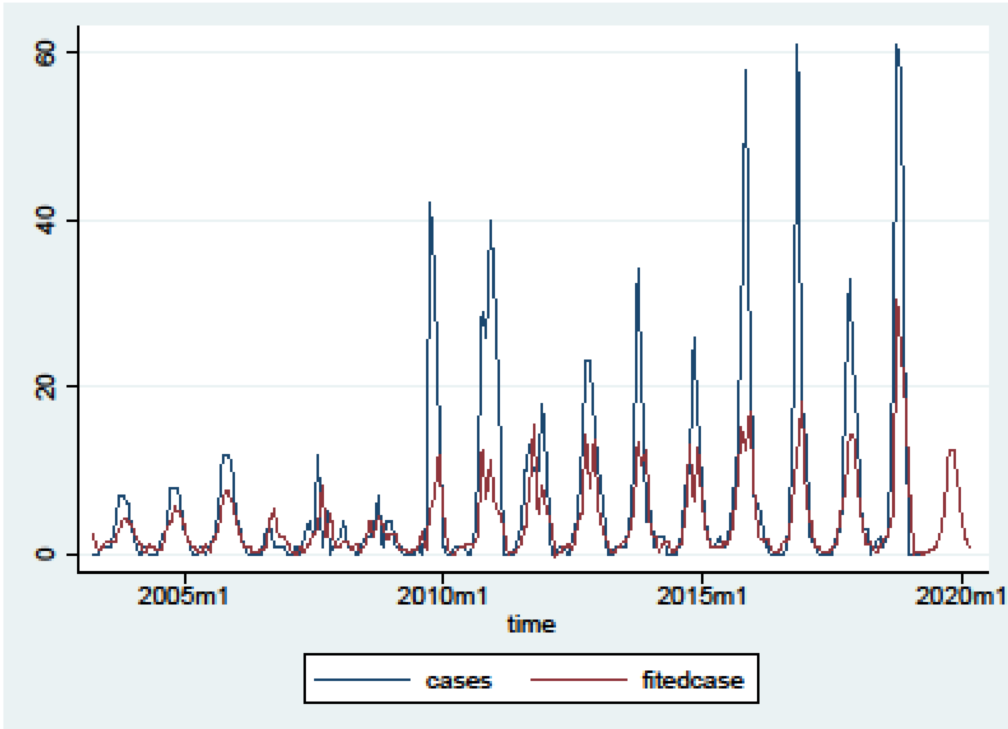

Fig. 4. Observed and predicted monthly cutaneous leishmaniasis cases for 2019 in Maraveh Tappeh County, Golestan Province, Iran

Discussion

This study, encompassing data from 2003 to 2018, not only provides a quantitative profile of CL cases in Maraveh Tappeh County but also delineates the causal–ecological mechanisms underpinning observed temporal patterns. Over the study period, 1301 CL cases were documented in the county's catchment population. Strikingly, 82% of cases occurred within a narrow four-month window (September–December), supporting a pronounced seasonal cycle driven by the synchrony between the life cycle of the vector (*Ph. papatasi*), activity patterns of rodent reservoirs and human exposure behaviors. Comparable seasonal dynamics have been reported in other investigations conducted in Golestan Province (2–6).

From a vector epidemiology perspective, the peak of *Ph. papatasi* activity in Golestan Province occurs in August and September (2, 5, 7). This seasonal peak coincides with a subsequent lag in human CL incidence consistent with the incubation period of *L. major* (<4 months) (33), indicating sustained transmission and accumulation of cases in the post vector-peak phase. This temporal alignment is critical for defining optimal intervention windows (e.g., residual spraying, rodent control) and projecting near-term incidence trends.

In the present analysis, the ARIMA (2,0,2)–SARIMA $(0,0,1)_{12}$ model exhibited the best fit and predictive performance. Comparable modeling efforts in southern Fars (34) identified SARIMA $(4,1,4)(0,1,0)_{12}$ as optimal; in western Iran (35), SARIMA$(1,0,1)(1,0,0)_{12}$ was selected; and in Shahroud (36), SARIMA (3,1,1) $(0,1,2)_4$ (AIC=324.3, BIC=317.7, RMSE=0.167) emerged as the best performer. Given the seasonality of CL in these regions, with incidence rising from around October and declining during winter, such forecasts are both epidemiologically consistent and expected.

Transmission of CL is shaped by multiple determinants, including vector abundance, reservoir host population dynamics and climatic geographic parameters. We examined the effects of monthly mean temperature, relative humidity, rainfall and rainy day frequency on incidence.

Climatic correlation analyses revealed that temperature, rainfall and relative humidity jointly depict a complex environmental biological interaction network. Among these, temperature is a key determinant, as sand fly development is tightly constrained within 16–35 °C; within this range, higher temperatures shorten development time (7, 33, 37). However, the significant inverse correlations observed at 2- and 4-month lags, contrary to the simplistic assumption that higher temperatures invariably enhance vector proliferation, suggest threshold effects. Extreme heat may increase larval mortality, change nighttime activity patterns, or suppress blood-feeding behavior. Temperatures above the developmental range of sand flies, by creating aestivation conditions, cause the development of sand flies to be inactive and lead to a decrease in the incidence of cutaneous leishmaniasis. This inverse relationship may also be related to the secondary role of reservoir behavior in hot phases of the year and their reduced contact with the vector (7, 37). Parallel inverse relationships have been reported in Sri Lanka and Khuzestan (27, 38), whereas studies in Isfahan (39) and Costa Rica (40) have demonstrated positive temperature incidence associations. Such discrepancies between cooler temperate zones like Isfahan (39) and humid tropical settings like Costa Rica (40) emphasize the geographic specificity of climate disease linkages. In Khuzestan (27), annual mean temperature correlated positively with incidence and in Fars (34), positive associations were detected at a 3-month lag. Findings from Sabzevar (41) mirrored the present results, showing negative correlations with monthly mean temperature.

Rainfall and rainy day count also emerged as significant predictors. The bimodal effect, positive associations at shorter lags and negative associations at a 5-month lag, suggests a dual role: acutely, rainfall promotes suitable breeding habitat via increased moisture and vegetation cover; over longer periods, flooding and habitat degradation may suppress vector populations. Post-rain vegetation growth can boost reservoir rodent abundance, with timing contingent on breeding cycles that may lag rainfall (33). Here, total rainfall correlated positively at a 2-month lag and rainy day frequency correlated positively at 0 and 2-month lags; both factors were inversely correlated at a 5-month lag. Comparable positive rainfall CL associations have been reported in Brazil (42), Sabzevar (41), and Fars (34) at 3- and 9-month lags. Conversely, studies in Fars (34) and Kerman (43) have documented negative associations, and an Isfahan investigation (39) observed no significant relationship. In Morocco (44), minimum temperature and rainfall were both positively linked to wet-type and dry-type CL; similar patterns have been noted in Africa, Europe, and Iran (45–47).

Relative humidity demonstrated a stable positive correlation at 0, 2 and 4 month lags, reflecting its role in maintaining vector physiological viability. This relationship aligns with reports from Sabzevar (41), Isfahan (39) and Khuzestan (27). Humidity is a recognized determinant in both ACL and visceral leishmaniasis (45, 46), with effects documented even in hyper-arid endemic zones, indicating that even marginal changes in ambient moisture can tip habitat suitability in favor of vectors. From a public health perspective, these findings support embedding climate-driven models within early warning surveillance systems. Identification of optimal predictor lags (2–4 months) suggests real-time climatic inputs can reliably forecast near-term incidence, optimizing resource allocation, intervention timing and deployment of mobile health units. Also, based on two to four month lags, climatic data can serve as early indicators predicting increases in cutaneous leishmaniasis incidence approximately 2–4 months in advance." However, projected climate-change scenarios, including warming trends, altered precipitation regimes and humidity variability, pose potential disruptions to these associations, necessitating adaptive forecasting frameworks. Future work integrating outputs from regional climate models with SARIMA or hybrid approaches such as ARIMAX could enhance policy-relevant predictive capacity. This study not only identifies the seasonal and climatic drivers of CL incidence in Maraveh Tappeh but also underscores the value of integrating climatic, land cover, and behavioral factors into multi-scale predictive models. Such models could enhance early warning capabilities and improve response strategies for leishmaniasis control.

This study has certain limitations that should be considered when interpreting the results. Due to its ecological time-series design, causal inferences between climatic factors and cutaneous leishmaniasis incidence cannot be definitively established and the findings primarily reflect temporal associations. Additionally, the use of routine surveillance data may be associated with some degree of underreporting, particularly for mild or non-referred cases. Finally, the predictive models were developed based on historical patterns and therefore, unexpected climatic fluctuations or large-scale interventions could influence long-term forecast accuracy.

Despite these considerations, the application of a long-term dataset and a biologically coherent SARIMA model provides robust and policy-relevant insights for disease monitoring and early warning in endemic areas. The absence of a formal sensitivity analysis to assess the robustness of the results to alternative lag specifications or model formulations should be considered a limitation of this study.

Conclusion

A 16-year analysis from Maraveh Tappeh County shows that cutaneous leishmaniasis follows a clear seasonal pattern strongly influenced by temperature, precipitation and relative humidity. Climatic effects were nonlinear and lag-dependent, with short-term conditions favoring vector activity and longer-term or extreme conditions limiting vector survival. The temporal consistency between *Ph. papatasi* activity and subsequent human cases, along with identified 2–4-month lags, supports the use of climate-based early-warning systems. ARIMA and SARIMA models effectively captured and forecasted disease trends, highlighting their value for optimizing the timing of preventive interventions. These results underscore the importance of location-specific strategies that integrate climatic, land-use and vegetation data, particularly in the context of climate-driven shifts in CL transmission.

Acknowledgements

The authors gratefully acknowledge the Vice Chancellor for Research and Technology of Golestan University of Medical Sciences for supporting this project. We also extend our appreciation to the Maraveh Tappeh County Health Center and the Maraveh Tappeh County Meteorological Department for their valuable assistance in providing epidemiological and climatic data.

Ethical consideration

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. The research protocol was reviewed and approved by the Research Ethics Committee of Golestan University of Medical Sciences, Iran (Approval Code: IR.GOUMS. REC.1397.250), under the supervision of the Vice Chancellor for Research and Technology. Data were analyzed at an aggregate level and no personal identifiers were collected or used in this study.

http://jad.tums.ac.ir
Published Online: December 30, 2025

Conflict of interest statement

The authors declare there is no conflict of interest.

References

1- Yaghoobi-Ershadi MR (2012) Phlebotomine sand flies (Diptera: Psychodidae) in Iran and their role on *Leishmania* transmission. J Arthropod-Borne Dis. 6: 1–7.
2- Rassi Y, Sofizadeh A, Abai MR, Oshaghi MA, Rafizadeh S, Mohebail M, Mohtarami F, Salahi R (2008) Molecular detection of *Leishmania major* in the vectors and reservoir hosts of Cutaneous Leishmaniasis in Kalaleh District, Golestan Province, Iran. Iran J Arthropod-Borne Dis. 2(2): 21–27.
3- Cherabin M, Sofizadeh A, Palideh AR, Gharavi AH, Gharavi M (2012) Epidemiological characteristics of cautaneous leishmaniasis in Maraveh Tappeh district, Golestan Province during 2006- 2010. J Zabol Uni Med Sci. 4(1): 19-27 (Persian).
4- Sofizadeh A, Faraji Far AA, Cherabin M, Badiei F, Cherabin M, Sarli J, YapangGharavi, Mehravaran A (2013) Cutaneous leishmaniasis in GonbadKavoos, North of Iran (2009–11), an epidemiological study. J Gorgan Uni Med Sci. 14 (4): 100–106 (Persian).
5- Sofizadeh A, Ghorbani M, Gorganli Davaji A, Gharemeshk Gharavi A (2015) Epidemiological status of cutaneous leishmaniasis and ecological characteristics of sandflies in Maraveh-Tappeh County, Golestan Province, Iran (2011–2012). Qom Univ Med Sci J. 9(6): 53–65 (Persian).
6- Sofizadeh A, Vatandoost H, Rassi Y, Hanafi-Bojd AA, Rafizadeh S (2016) Spatial analyses of relation between rodent's active burrows and incidence of zoonotic cutaneous leishmaniasis in Golestan Province, northeastern of Iran. J Arthropod-Borne Dis. 10(4): 569–576.
7- Agh-Ataby MD, Sofizadeh A, Ozbaki GM, Malaki-Ravasan N, Ghanbari MR, Mozafari O (2016) Ecoepidemiological characteristics of a hypoendemic focus of zoonotic cutaneous leishmaniasis in north Iran (southeast of Caspian Sea). J Vector Borne Dis. 53: 248–256.
8- Roshanghalb M, Parvizi P (2012) Isolation and determination of *Leishmania major* and *Leishmania turanikain* in *Phlebotomus papatasi* main vector of Zoonotic Cutaneous leishmaniasis in Turkmen Sahra, Golestan Province. J Mazandaran Uni Med Sci. 21(Supplement 1): 74–83.
9- Akhoundi M, Mohebali M, Asadi M, Mahmodi MR, Amraei K, Mirzaei A (2013) Molecular characterization of *Leishmania* spp. In reservoir hosts in endemic foci of zoonotic cutaneous leishmaniasis in Iran. Folia Parasitol. 60(3): 218–224.
10- Bordbar A, Parvizi P (2014) High density of *Leishmania major* and rarity of other mammals' Leishmania in zoonotic cutaneous leishmaniasis foci, Iran. Trop Med Int Health. 19(3): 355–363.
11- Mozafari O, Sofizadeh A, Shoraka HR, Namroodi J, Kalteh EA (2020) Eco-epidemiology of cutaneous leishmaniasis in Golestan Province, northeastern Iran: a systematic review. Jorjani Biomed J. 8(1): 60–78.
12- Jorjani O, Mirkarimi K, Charkazi A, Dadban Shahamat Y, Mehrbakhsh Z, Bagheri A (2019) The epidemiology of cutaneous leishmaniasis in Golestan Province, Iran: A cross-sectional study of 8-years. Parasite Epidemiol Control. 5: e00099.
13- Sofizadeh A, Rassi Y, Vatandoost H, Hanafi-Bojd AA, Mollalo A, Rafizadeh S, Akhavan AA (2017) Predicting the distribution of *Phlebotomus papatasi* (Diptera: Psychodidae), the primary vector of zoonotic cutaneous leishmaniasis, in Golestan Province of Iran using ecological niche modeling: Comparison of MaxEnt and GARP Models. J Med Entomol. 54 (2): 312–320.

14- Hanafi-Bojd AA, Yaghoobi-Ershadi MR, Haghdoost AA, Akhavan AA, Rassi Y, Karimi A, Charrahy Z (2015) Modeling the distribution of cutaneous leishmaniasis vectors (Psychodidae: Phlebotominae) in Iran: a potential transmission in disease prone areas. J Med Entomol. 52(4): 557–565.
15- Sofizadeh A, Hanafi-Bojd AA, Shoraka HR (2018) Modeling spatial distribution of *Rhombomys opimus* as the main reservoir host of zoonotic cutaneous leishmaniasis in northeastern Iran. J Vector Borne Dis. 55(4): 297–304.
16- Hosseini SH, Allah-Kalteh E, Sofizadeh A (2021) the effect of geographical and climatic factors on the distribution of *Phlebotomus papatasi* (Diptera: Psychodidae) in Golestan Province, an endemic focus of zoonotic cutaneous leishmaniasis in Iran, 2014. J Arthropod-Borne Dis. 15(2): 225–235.
17- Cowpertwait PS, Metcalfe A (2009) Introductory Time Series with R. Springer-Verlag, New York.
18- Jebb AT, Tay L, Wang W and Huang Q (2015) Time series analysis for psychological research: examining and forecasting change. Front Psychol. 6: 727.
19- Sharafi M, Ghaem H, Tabatabaee HR, Faramarzi H (2017) Forecasting the number of zoonotic cutaneous leishmaniasis cases in south of Fars Province, Iran using seasonal ARIMA time series method. Asian Pac J Trop Med. 10(1): 79–86.
20- Lewnard JA, Jirmanus L, Júnior NN, Machado PR, Glesby MJ, Ko AI, Carvalho EM, Schriefer A, Weinberger DM (2014) Forecasting temporal dynamics of cutaneous leishmaniasis in Northeast Brazil. PLoS Negl Trop Dis. 8(10): e3283.
21- Bozorg-Omid F, Kafash A, Jafari R, Akhavan AA, Rahimi M, Foroushani RF, Youssefi F, Shirzadi MA, Ostadtaghizadeh A, Hanafi-Bojd AA (2023) Predicting current and future high-risk areas for vectors and reservoirs of cutaneous leishmaniasis in Iran. Sci Rep. 13: 11546.
22- Ahangar Cani M, FARNAGHI M (2019) Providing a model for cutaneous leishmaniasis risk mapping using GIS and neural network algorithm. Geographical Data. 28(109): 7–24 (Persian).
23- Parto F, Valadan Zoij MJ, Yousefi F, Alatab S (2025) Forecasting zoonotic cutaneous leishmaniasis risk through integrated simulated remote-sensing data and epidemiological records. Iran J Remote Sens GIS. (In press).
24- Kumari S, Muthulakshmi P (2024) SARIMA Model: An efficient machine learning technique for weather forecasting. Procedia Comput Sci. 235: 656–670.
25- Peng PY, Duan HY, Xu L, Sun JQ, Ma LJ (2025) A seasonal autoregressive integrated moving average (SARIMA) forecasting model to predict the epidemic trends of scrub typhus in China. PLOS ONE. 20(6): e0325905.
26- Correa Antonialli S, Torres TG, Paranhos Filho AC, Tolezano JE (2007) Spatial analysis of American visceral leishmaniasis in Mato Grosso do Sul State, Central Brazil. J Infect. 54: 509–514.
27- Azimi F, Shirian S, Jangjoo S, Ai A, Abbasi T (2017) Impact of climate variability on the occurrence of cutaneous leishmaniasis in Khuzestan Province, Southwestern Iran. Geospat Health. 12(1): 15–22.
28- Rodriguez-Morales A (2005) Ecoepidemiology and satellite epidemiology: new tools for managing public health problems. Rev Peru Med Exp Salud Publica. 22: 54–63 (Spanish).
29- Milić M, Milojković J, Jeremić M (2022) Optimal neural network model for short-term prediction of confirmed cases in the COVID-19 pandemic. Mathematics. 10(20): 3804.
30- Kontopoulou VI, Panagopoulos AD, Kakkos I, Matsopoulos GK (2023) A re-

view of ARIMA vs. Machine learning approaches for time series forecasting in data driven networks. Future Internet. 15: 255.

31- ArunKumar KE, Kalaga DV, Sai Kumar CM, Kawaji M, Timothy MB (2022) Comparative analysis of gated recurrent units (GRU), long short-term memory (LSTM) cells, autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA) for forecasting COVID-19 trends. Alex Eng J. 16(10): 7585–7603.

32- Hakem A, El Khiat A, Ezzahidi A, Bouhout S, Ait Ali D, El Houate B, Boutaayamou I (2025) Incidence and prediction of cutaneous leishmaniasis cases and its related factors in an endemic area of Southeast Morocco: Time series analysis. Acta Trop. 264: 107579.

33- Rassi Y, Hanafi-Bojd AA (2006) Sandflies, vectors of Leishmaniasis. Noavarane- Elm Publication, Tehran (Persian).

34- Sharafi M, Ghaem H, Tabatabaee HR, Faramarzi H (2017) Forecasting the number of zoonotic cutaneous leishmaniasis cases in south of Fars Province, Iran using seasonal ARIMA time series method. Asian Pac J Trop Med. 10(1): 79–86

35- Nikonahad A, Khorshidi A, Ghaffari HR, Aval HE, Miri M, Amarloei A, Nourmoradi H, Mohammadi A (2017) A time series analysis of environmental and metrological factors impact on cutaneous leishmaniasis incidence in an endemic area of Dehloran, Iran. Environ Sci Pollut Res. 24: 14117–14123.

36- Majidnia M, Ahmadabadi Z, Zolfaghari P, Khosravi A (2023) Time series analysis of cutaneous leishmaniasis incidence in Shahroud based on ARIMA model. BMC Public Health. 23: 1190.

37- Oshaghi MA, Maleki Ravasan N, Javadian E, Rassi Y, Sadraei J, Enayati AA, Vatandoost H, Zare Z, Emami SN (2009) Application of predictive degree day model for field development of sandfly vectors of visceral leishmaniasis in northwest of Iran. J Vector Borne Dis. 46(4): 247–255.

38- Wijerathna T, Gunathilaka N (2023) Time series analysis of leishmaniasis incidence in Sri Lanka: evidence for humidity-associated fluctuations. Int J Biometeorol. 67: 275–284.

39- Arvin AA, Mozafari GA, NoruzBagheri J (2013) Investigating the relationship between leishmaniasis outbreak and climatic parameters in the Eastern Region of Isfahan. J Nat Haz Environ. 4(2): 43–60 (Persian).

40- Chaves LF, Pascual M (2006) Climate cycles and forecasts of cutaneous leishmaniasis, a nonstationary vector-borne disease. PLoS Med. 3(8): e295.

41- Akbari E, Mayvaneh E, Entezari A, Nazari M (2014) Survey of the role of bioclimatic factors in the outbreak of Cutaneous Leishmaniasis. Iran J Epidemiol. 10(3): 65–74 (Persian).

42- Viana GMC, Nascimento MDSB, Rabelo EMF, Neto JAD, Junior JRB, Galvao CS, Santos ACD, Junior OMS, Oliveria RAS, Guimaraes RS (2011) Relationship between rainfall and temperature: observations on the cases of visceral leishmaniasis in São Luis Island, State of Maranhão, Brazil. Rev Soc Bras Med Trop. 44(6): 722–724.

43- Afshar PJ, Bahrampour A, Shahesmaeili A (2022) Determination of the trend of incidence of cutaneous leishmaniasis in Kerman Province 2014-2020 and forecasting until 2023. A time series study. PLoS Negl Trop Dis 16(4): e0010250.

44- El Idrissi Saik I, Talimi H, Badri B, Bouhout S, Fissoune R, Riyad M, Lemrani M (2025) Cutaneous leishmaniasis in Casablanca-Settat region (Morocco): spatio-temporal analysis of disease dynamic and machine learning based case prediction. medRxiv. 01.28.25321261.

45- Bounoua L, Kahime K, Houti L, Blakey T, Ebi KL, Zhang P, Imhoff ML, Thoma KJ, Dudek C, Sahabi SA, Messouli M, Makhlouf B, El Laamrani A, Boumezzough A (2013) Linking climate to incidence of zoonotic cutaneous leishmaniasis (*L. major*) in pre-Saharan North Africa. Int J Environ Res Public Health. 10: 3172–3191.

46- Medenica S, Miladinović-Tasić N, Stojanović NM, Lakićević N, Rakočević B (2023) Climate variables related to the incidence of human leishmaniosis in Montenegro in southeastern Europe during seven decades (1945-2014). Int J Environ Res Public Health. 20(3): 1656.

47- Navi Z, Salahi-Moghaddam A, Habibi-Nokhandan M, Mohebali M, Hajjaran H, Fuentes MV (2024) A geomedical survey: is there an association between climatic conditions and *Leishmania* species distribution in Iran during the years 1999-2021? Acta Parasitol. 69(1): 769–775.