Document heading   doi:

Modeling seasonal leptospirosis transmission and its association with rainfall and temperature in Thailand using time−series and ARIMAX analyses

Sudarat Chadsuthi[1], Charin Modchang[1,2], Yongwimon Lenbury[3,4], Sopon Iamsirithaworn[5], Wannapong Triampo[1,2,6*]

[1]Biophysics Group, Department of Physics, Faculty of Science Mahidol University, Rama VI, Ratchathewi District, Bangkok Thailand
[2]ThEP Center CHE, 328 Si Ayutthaya Road, Bangkok, Thailand
[3]Centre of Excellence in Mathematics, CHE, 328 Si Ayutthaya Road, Bangkok, Thailand
[4]Department of Mathematics, Faculty of Science Mahidol University, Rama VI, Ratchathewi District, Bangkok Thailand
[5]Bureau of Epidemiology, Department of Disease Control, Ministry of Public Health, Tivanond Road, Nonthaburi, Thailand
[6]Institute for Innovative Learning Mahidol University, 999, Phuttamonthon 4 Road, Salaya, Nakorn Pathom Thailand

ARTICLE INFO

ABSTRACT

Objective: To study the number of leptospirosis cases in relations to the seasonal pattern, and its association with climate factors. **Methods:** Time series analysis was used to study the time variations in the number of leptospirosis cases. The Autoregressive Integrated Moving Average (ARIMA) model was used in data curve fitting and predicting the next leptospirosis cases. **Results:** We found that the amount of rainfall was correlated to leptospirosis cases in both regions of interest, namely the northern and northeastern region of Thailand, while the temperature played a role in the northeastern region only. The use of multivariate ARIMA (ARIMAX) model showed that factoring in rainfall (with an 8 months lag) yields the best model for the northern region while the model, which factors in rainfall (with a 10 months lag) and temperature (with an 8 months lag) was the best for the northeastern region. **Conclusions:** The models are able to show the trend in leptospirosis cases and closely fit the recorded data in both regions. The models can also be used to predict the next seasonal peak quite accurately.

1. Introduction

Leptospirosis is a worldwide zoonosis, which causes more than 500 000 cases per year[1]. Leptospirosis is caused by spirochete pathogenic bacteria genus *Leptospira interrogans*[2]. These organisms live in many wild and domestic animal species acting as potential reservoir for human infection. The infection in humans is usually caused by direct contacts with products of infected animals, mainly urine, and/or indirect contacts with contaminated environment, especially water contaminated with *Leptospira* spp. Ingestion of the bacteria and human−to−human transmission are rare. The disease mostly occurs in tropical and subtropical zones such as Thailand[3] and India[4]. By indirect contacts, *Leptospira* may get transmitted through breaks or cuts of skin, or mucous membranes via contaminated water or moist soil[5]. Leptospira organisms can survive for a long period in natural aqueous environments[6]. Agricultural workers are most infected during cultivated rice activity in marshy land[7] or capturing fish or animals. Infection also occurs to those who swim or wade, applying fertilizer or plowing in wet field in contaminated water[7]. Flooding can determine the size of outbreak since it may carry urine of infected animals to

distant places[8].

Seasonal pattern of leptospirosis has been observed in Thailand and it corresponds to the rainy season. The peak of incident cases correlates with the peaks in rainfall and temperature. The association between rainfall and leptospirosis cases in Thailand was previously described by Triampo *et al*[9]. By solving deterministic SIR model, they found that the amount of rainfall could cause the seasonal increase in incident rates in the Phrae, a northern province, and the Nakhon Ratchasima province, in the northeastern part of Thailand. They can fit the temporal evolution of incident cases in a year of outbreak. However, they did not consider the incidence of leptospirosis cases during a lengthy period or analyze the prediction. To describe and predict seasonal patterns in leptospirosis infection, climate variations, such as the amount of rainfall and/or temperature in that region, needs to be incorporated. Desvars *et al*. found that the amount of rainfall 2 months earlier and the average temperature that month were correlated with the time series of infected cases[10]. Other researchers have also shown that there was a link between leptospirosis and rainfall[11].

Some statistical tools have proved to be useful in measuring the correlation between climate and leptospirosis cases leading to improved models for prediction. Auto Regressive Integrated Moving Average (ARIMA) modeling may be the most famous tool to analyze the time series data. Through modeling with time series, steps–ahead predictions are shown to be more accurate than those obtained by other statistical methods[12]. This model has been widely used in epidemiology to monitor and predict infectious diseases[13] and found to be successful in monitoring and predicting the spread of infectious diseases such as Influenza in Hong Kong (China) and Arizona, USA[14], Dengue in Rio de Janeiro, Brazil[13], and malaria in Bhutan[15]. The prediction of the number of cases for the upcoming year will generate useful information for planning of public health service and medical care.

In this work, we aim to investigate and describe the seasonal pattern of leptospirosis cases in the northern and northeastern region of Thailand during 2003–2009. To determine the seasonal pattern of leptospirosis, we evaluated the association between leptospirosis cases and the amount of rainfall and temperature. We used ARIMA(X) models to fit the number of reported cases of leptospirosis. Additionally, the next seasonal trend was predicted based on the previous cases number and climate data.

2. Materials and methods

2.1 Data sources

A case of leptospirosis was first reported in Thailand in 1942 by Yunibandhu[16]. Most cases were reported from the northern and northeastern regions, which are the tropical and agricultural areas. In this study, monthly cases of leptospirosis were retrieved from the National Disease Surveillance Report (506) of the Bureau of Epidemiology, the Ministry of Public Health[17]. From 2003 to 2009, the incidence rates per 100 000 population of reported leptospirosis were between 9 and 18 in northeastern region and between 4 and 8 in northern region. Most cases were suspected leptospirosis cases, based on clinical diagnoses made by attending physicians. A small portion of case–patients were lab–confirmed cases of leptospirosis, mostly tested at the beginning of epidemic season.

The time series of monthly rainfall and average temperature data were taken from the Research Data Archive (RDA), which is maintained by the Computational and Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR). The NCAR is sponsored by the National Science Foundation (NSF). The original data are available from the RDA (http://dss.ucar.edu) in dataset number ds512.0. We extracted the climate data from 32 stations, 19 stations are in the northern region and 13 stations are in northeastern region. We find the average climate data from those stations to represent the regional climate.

2.2. Analysis

The 84 months data on both the northern and northeastern regions are divided into 2 sets; first 72 months data is used for model fitting propose and the last 12 months data was used for prediction. In this study, we use the software R, version 2.13.0 (The R Foundation for Statistical Computing, http://www.R–project.org) for time series data analysis and graphic displays. The time series data is analyzed and fitted by using Autoregressive Integrated Moving Average, ARIMA(p, d, q)x(P, D, Q)s model[18,19], where: p and P are the auto regressive (AR) order and seasonal order of AR, respectively; d and D are the differencing order and seasonal differencing order, respectively; q and Q are the moving average (MA) order and seasonal moving average order, respectively s is the seasonal period.

The ARIMA models are the most general class of models for fitting and forecasting time series data. ARIMA consists of 3 parts a) Auto regression, b) Moving average, and c) differencing order. Applying ARIMA, we first need to check the time series data by autocorrelation function (ACF) plot. If the time series are stationary, the ACF plot will show the fluctuated pattern. But if time series are non–stationary, we have to remove the non–stationary term, by using log transform and/or differencing order, then re–

test the stationary character of the time series by ACF plot and Dickey–Fuller test (*P*–value < 0.01). We then find the p and q orders by using cut–off time lag of ACF and partial autocorrelation function (PACF) plots. The coefficients of ARIMA model are estimated by using mean square method, which assists in the goodness of fit process. For fitting non–seasonal univariate ARIMA model, we try numerous different values of p and q orders. The residuals of model results are checked for independence of the noise term through ACF and PACF plots. If the residuals have nearly white noise, the goodness of fit is examined through calculated Akaike's Information Criterion (AIC) and root mean square error (RMSE). The best–fitted model is used to predict the next leptospirosis season. For seasonal–ARIMA model, we first remove the trend and seasonal component, after which we carry out the same procedure as described above.

To investigate the climate factor, we incorporate climate variables as input series in the ARIMA model, called ARIMAX model. We calculate the cross–correlations function (CCF) between the pre–whitened climate series with leptospirosis cases series to find significant time lags at *P*–value < 0.05. Similar to the univariate ARIMA model, we estimate the coefficients of ARIMAX associated with the lagged climate variable. We test the different time lags of climate input one at a time before combining them together. The multivariate model is re–fitted and used to predict future leptospirosis season.

3. Results

The numbers of monthly cases from January 2003 to

December 2008 have been analyzed by using ARIMA with climate factors. There are 3 696 reported cases in the northern region and 15 334 cases in the northeastern region. The results for the northern region will be presented first, followed by that for the northeastern region.

3.1. Northern region

To analyze the data with ARIMA model, we first consider the log transformation of the time series data as shown in Figure 1 to stabilize the variance and subsequent differencing, either non–seasonal or seasonal, of the first differencing order. As can be seen from the ACF plot, the time series appears to be stationary (*P* < 0.01 by the Dickey–Fuller test). We use the ACF and PACF plots to identify the order of ARIMA model. The plots of ACF and PACF for the first differencing order of log of leptospirosis cases are shown in Figure 1. The ACF plot shows strong positive autocorrelation at time lags 1, 12, 24 and 36, which suggests a seasonal component. For the first difference log series, the ACF cut off is lag 2 and PACF cut off is lag 1, while for the seasonal order difference series, the ACF decreases slowly and PACF cuts off at lag 1 (data not shown).

We fitted the data with several univariate ARIMA (p, d, q) and ARIMA (p, d, q)x(P, D, Q)s with different orders and excluded the models in which the residual is not likely to be white noise. For non–seasonal model, we found that ARIMA models show autocorrelation of residuals by ACF and PACF plots. We excluded the non–seasonal ARIMA for fitting leptospirosis cases series in the northern region. But for seasonal ARIMA, we found that the residuals of several ARIMA orders do not show evidence of autocorrelation.
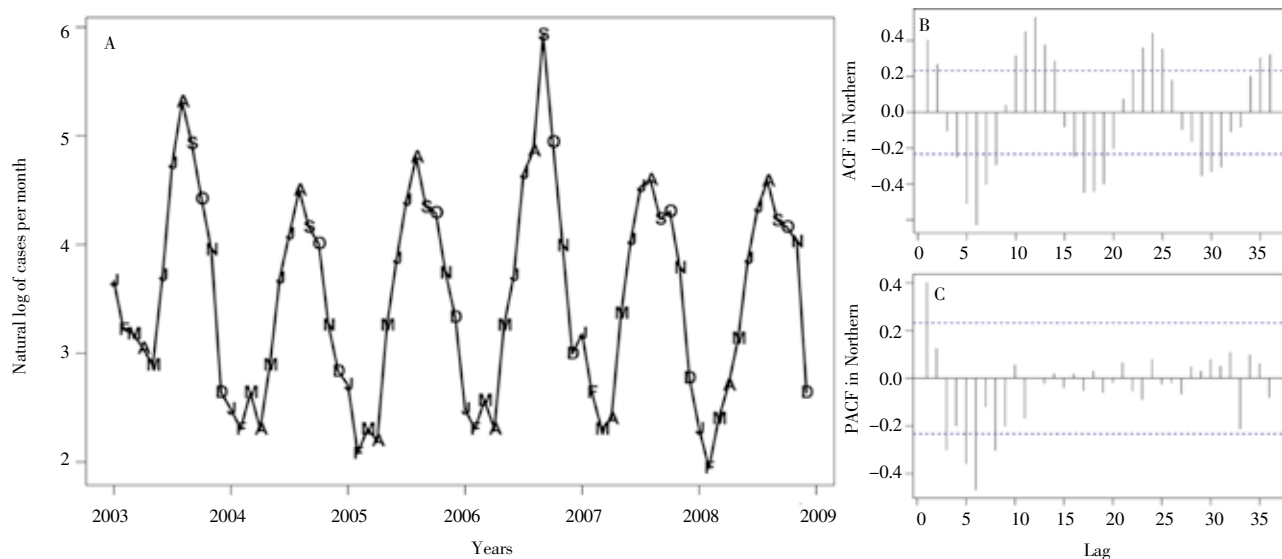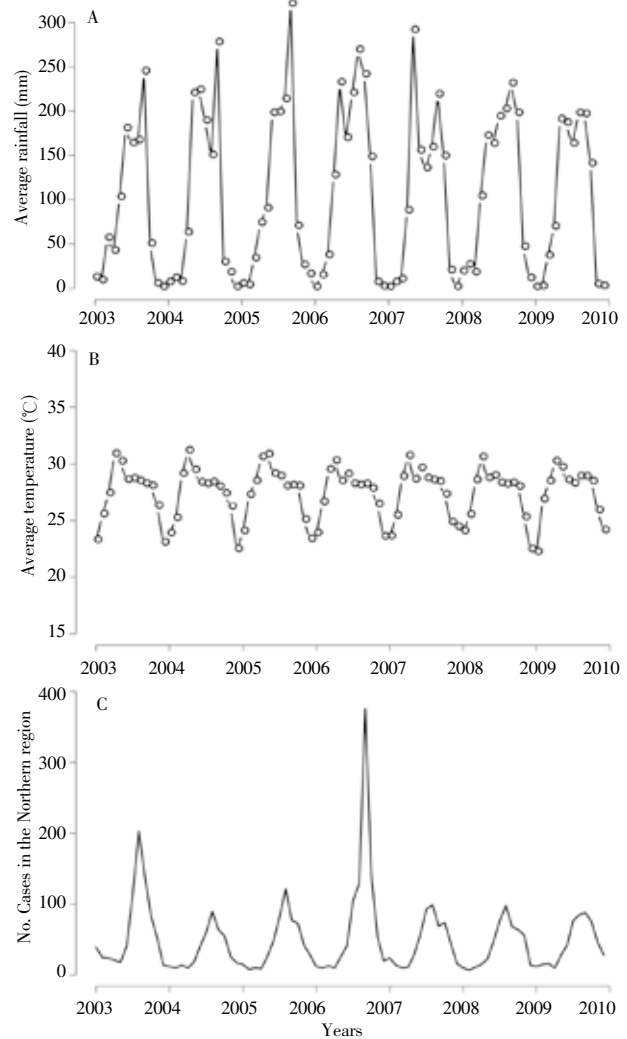

Figure 1. The natural logarithm of leptospirosis cases in northern Thailand (left panel).
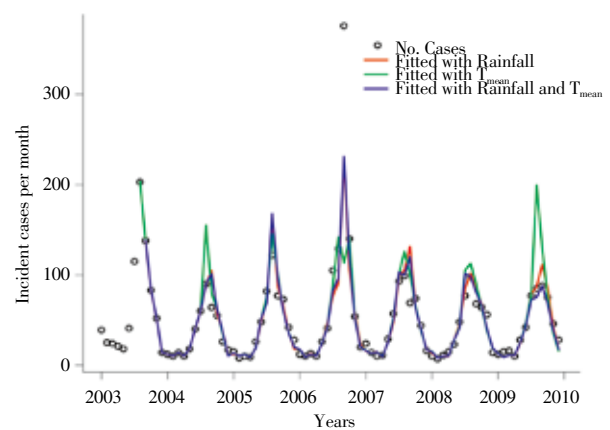The ACF and PACF of the first differencing order of log of cases in the northern region are shown in the right panel; x–axis represents the time lags in month(s). Dotted line indicates 95% confidence interval.

The performance of seasonal–ARIMA models are shown in Table 1. We found that ARIMA(0,0,1)(1,1,1)$_{12}$ model shows the best performance for prediction, but it also shows high AIC and RMSE of fitting. For the lowest AIC, ARIMA(1,0,0)(0,1,1)$_{12}$ is the model, while the ARIMA(1,0,1)(0,1,1)$_{12}$ model has the smallest RMSE, but the relative difference of RMSE for fitting of those models is only 1%. The relative difference of RMSE for prediction between the best–predicted model and the best–AIC model is 7.8%. When we consider the significant p–value of estimated coefficients, we find that all coefficients in the ARIMA(1,0,0)(0,1,1)$_{12}$ model are significant. So, we used the ARIMA(1,0,0)(0,1,1)$_{12}$ model , which has the lowest AIC and good fitting, for future analysis.

We hypothesize that climate factors are correlated with the number of leptospirosis cases. Figure 2 shows the monthly average of the amount of rainfall and temperature (T$_{mean}$) compared with the time series of infected cases during 2003–2009. To incorporate the climate factor as input series in the model, we find the cross–correlations between the pre–whitened climate series and leptospirosis series (Table 2). We find significant correlations of the rainfall factor at lags 8 and 9, while there are no significant lags for the T$_{mean}$ factor. We then test several orders of seasonal ARIMAX models, estimated with each of climate factors for several time lags even though that time lag does not appear significant, as shown in Table 3. We find that the only rainfall factor at time lags 7 and 8 appear significant in the ARIMAX model, when we test one by one time lag or climate factor. The multivariate model, which has the best fit and smallest AIC, is ARIMAX(1,0,1)(0,1,1)$_{12}$ associated with rainfall and T$_{mean}$ as covariates, while ARIMAX(0,0,1)(1,1,1)$_{12}$ associated with rainfall provides the best prediction. Concerning the *P*–value of estimated coefficients, we found 3 good models, which are ARIMAX(0,0,1)(1,1,1)$_{12}$ associate with Rainfall (Lag 7), ARIMAX(1,0,1)(0,1,1)$_{12}$ associated with Rainfall (Lag 8), and ARIMAX(1,0,0)(0,1,1)$_{12}$ associated with rainfall (Lag 8) and T$_{mean}$ (Lag7). The relative difference of fitting RMSE between models associated with rainfall (Lag 8) and associated with both rainfall (Lag 8) and T$_{mean}$ (Lag 7) is only 2.7%, while the relative difference of AIC is 2.3%. We also fitted the combination of rainfall at time lags 7 and 8, but the *P*–value of coefficients at time lag 7 is very high. Figure 3 shows the fitted model associated to each climate variables and combined variables with the leptospirosis cases series. Overall results suggest that the ARIMAX(1,0,1)(0,1,1)$_{12}$ model associated with rainfall (Lag 8) is the most appropriate. Although the result of the ARIMAX model with rainfall and temperature is similar, fewer variables should be better for monitoring and prediction.

Figure 2. The recorded climate time series: average rainfall (measured in mm) and average temperature (measured in ℃), compared with the time series of leptospirosis cases in the northern region.

Figure 3. Fitted and predicted values of ARIMAX(1,0,1)(0,1,1)$_{12}$ associated with rainfall (red line), T$_{mean}$ (green line), and both rainfall and T$_{mean}$ (blue line) in the northern region.

3.2. Northeastern region

We carry out the same procedure here as for the northern region. Figure 4 shows the log of time series, and ACF and
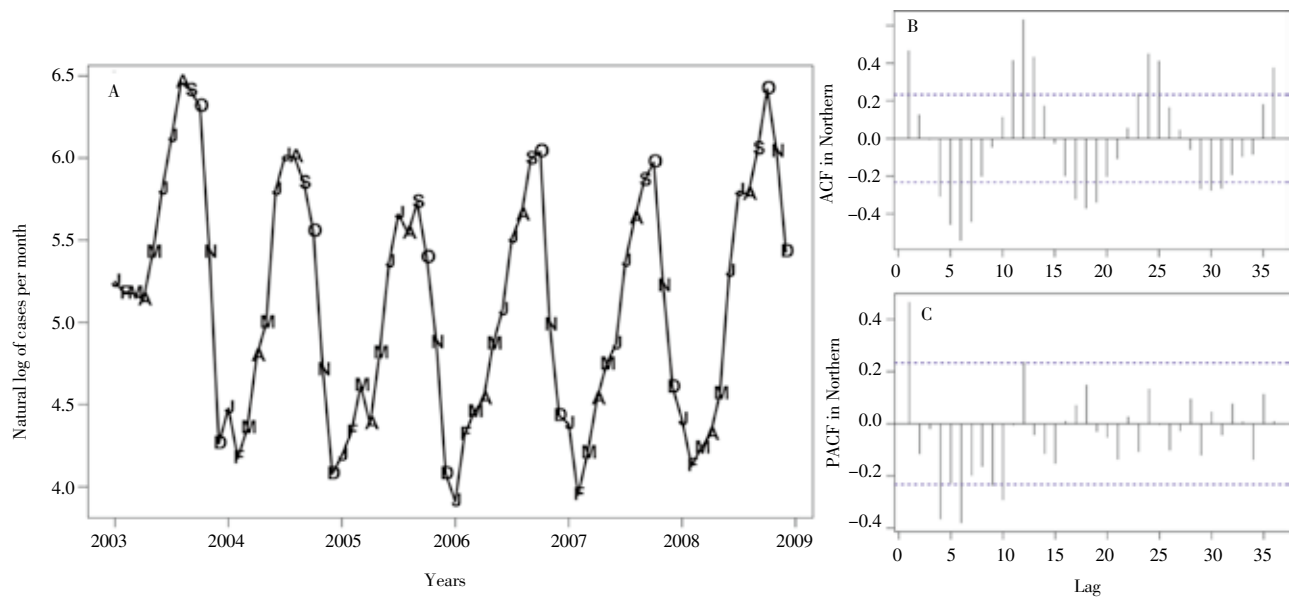
Figure 4. The natural logarithm of leptospirosis cases in northeastern Thailand (left panel).
The ACF and PACF of the first differencing order of log of cases in the northeastern region are shown in the right panel; x–axis represents the time lags in month(s). Dotted line indicates 95% confidence interval.

PACF plots for the first difference log series. Both ACF and PACF are cut off at lag 1, while in the seasonal order difference series the ACF decreases slowly and PACF is cut off at lag 1 (data not shown). With the ACF plot, the data also shows the seasonal pattern as in the northern region. Both non–seasonal and seasonal univariate ARIMA models are found. The residuals of non–seasonal model appear to be autocorrelated. The performance of the univariate seasonal model is shown in Table 4. The best model appears to be ARIMA$(1,0,0)(0,1,1)_{12}$, which has the smallest RMSE and AIC. We further find the cross–correlations of the pre–whitened climate factors series and data series. Figure 5 shows the monthly average of the amount of rainfall and temperature (T_{mean}) compared with the time series of infected cases during 2003–2009. The cross–correlations are shown in Table 5. For the monthly average rainfall factor, we find that the significant time lag is 10, while the significant time lags of monthly average temperature are 0, 1, and 8. We have fitted each climate factors at one time lag as input series. The performance of the models is shown in Table 6. It is found that only rainfall factor gives significant coefficients as in the northern region, while the average temperature factor yields P–value of 0.0535. We also combined both climate factors. Figure 6 shows the comparison of climate variable models. We found that the model with only rainfall or temperature could not fit the data series in certain years. This figure suggests that ARIMA$(1,0,0)(0,1,1)_{12}$ associated with average rainfall (Lag 10) and average temperature (Lag 8) is the best model. This model has also the smallest RMSE and AIC of fitting, with p–value of coefficients less than 0.05. So we use this multivariate model to predict the outbreak in the year 2009. The results will be useful to monitor and predict leptospirosis incidence in the forthcoming year.
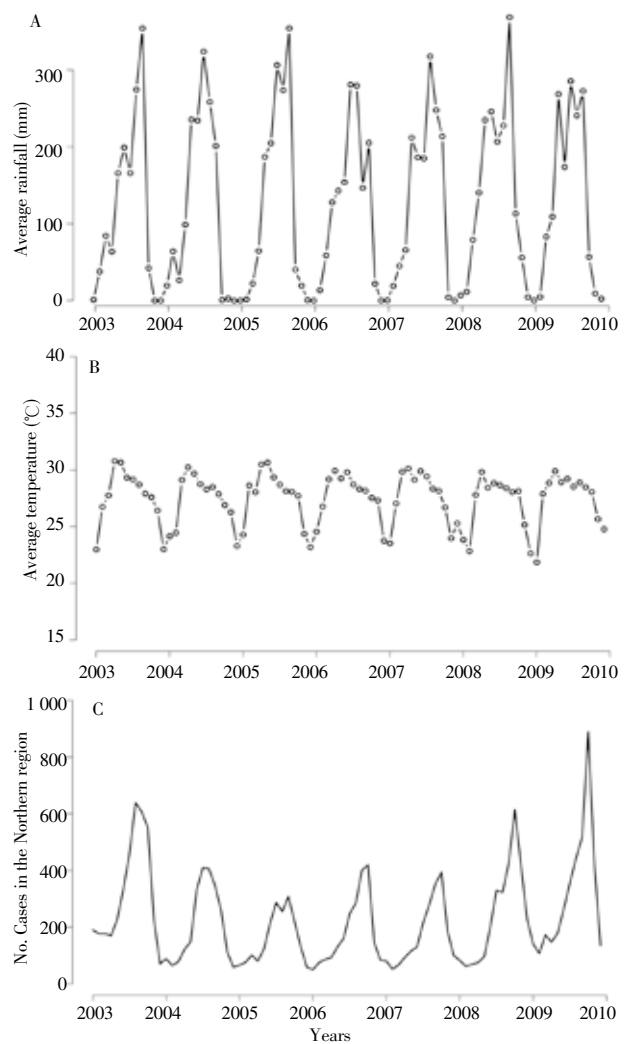

Figure 5. The recorded climate time series, average rainfall (measured in mm) and average temperature (measured in ℃), compared with the time series of leptospirosis cases in the northeastern region.

Table 1

Summary of ARIMA model fitting parameters in northern Thailand.

Model	Fit		Pred.	AR		MA		SAR		SMA	
	RMSE	AIC	RMSE	Coef.	p–value	Coef.	p–value	Coef.	P–value	Coef.	P–value
ARIMA(0,0,1)(0,1,1)$_{12}$	0.2970	61.24	0.2153	–	–	0.4343	<.0001			−1.0000	0.0028
ARIMA(0,0,1)(1,1,1)$_{12}$	0.2958	63.22	0.2140	–	–	0.4339	<.0001	−0.0231	0.8911	−0.9999	0.0048
ARIMA(1,0,0)(0,1,1)$_{12}$	0.2830	55.59	0.2307	0.5686	<.0001	–	–	–	–	−0.9995	0.0060
ARIMA(1,0,1)(0,1,1)$_{12}$	0.2827	57.50	0.2328	0.6291	0.0079	−0.0836	0.7802	–	–	0.3869	0.0098

ARIMA = Autoregressive Integrated Moving Average; Fit = Fitting results; RMSE = Root Mean Square Error; AIC = Akaike's Information Criterion; Pred. = Prediction of ARIMA model; Coef. = Coefficient of AR (autoregressive), MA (moving average), SAR (seasonal– autoregressive), and SMA (seasonal–moving average) part in ARIMA model.

Table 2

Northern: Cross–correlations between the pre–whitened climate series and cases series.

Variable	Lag										
	0	1	2	3	4	5	6	7	8	9	10
Rainfall	0.1480	−0.1200	−0.0090	0.0650	−0.0190	0.0650	−0.1090	0.2340	−0.4820[***]	0.2880[*]	−0.0140
P–value	0.2135	0.3183	0.9399	0.5975	0.8787	0.6035	0.3831	0.0602	0.0001	0.0219	0.9122
T$_{mean}$	0.0060	−0.1700	0.1260	0.1290	−0.1960	−0.094	0.0970	0.1670	−0.2230	0.1690	−0.0550
P–value	0.9570	0.1556	0.2998	0.2922	0.1089	0.4508	0.4378	0.1826	0.0771	0.1856	0.6700

Table 3

Summary of ARIMAX model fitting parameters in northern Thailand.

Model	Fit		Pred. RMSE	Climate variables		
	RMSE	AIC		Vars	Coef.	P–value
ARIMAX(0,0,1)(1,1,1)$_{12}$ with Rainfall	0.2539	45.95	0.2405	Rainfall (Lag 7)	0.0253	0.0365
ARIMAX(1,0,1)(0,1,1)$_{12}$ with Rainfall	0.2294	33.56	0.2433	Rainfall (Lag 8)	−0.0401	0.0005
ARIMAX(0,0,1)(1,1,1)$_{12}$ with Rainfall	0.2582	44.69	0.2415	Rainfall (Lag 9)	0.0211	0.1181
ARIMAX(1,0,1)(0,1,1)$_{12}$ with T$_{mean}$	0.2536	41.66	0.2429	T$_{mean}$(Lag 7)	2.6454	0.0996
ARIMAX(1,0,0)(0,1,1)$_{12}$ with T$_{mean}$	0.2530	41.66	0.2614	T$_{mean}$ (Lag 8)	−2.0686	0.2057
ARIMAX(1,0,1)(0,1,1)$_{12}$ with Rainfall	0.2288	35.24	0.2451	Rainfall (Lag 7)	−0.0077	0.5684
				Rainfall (Lag 8)	−0.0439	0.0008
ARIMAX(1,0,1)(0,1,1)12 with Rainfall and T$_{mean}$	0.2233	32.82	0.2519	Rainfall (Lag 8)	−0.0370	0.0004
				T$_{mean}$ (Lag 7)	2.4636	0.0839

ARIMAX = Autoregressive Integrated Moving Average with input series; Fit = Fitting results; RMSE = Root Mean Square Error; AIC = Akaike's Information Criterion; Pred. = Prediction of ARIMA model; Coef. = Coefficient of climate variables; Lag = time lag of climate variables.

Table 4

Summary of ARIMA mode fitting parameters in northeastern Thailand.

Model	Fit		Pred.	AR	MA	SAR		SMA	
	RMSE	AIC	RMSE	Coef.	P–value	Coef.	P–value	Coef.	P–value
ARIMA(1,0,0)(1,1,0)$_{12}$	0.2365	14.21	0.2556	0.8028	<0.0001	−0.2836	0.0457	–	–
ARIMA(1,0,0)(0,1,1)$_{12}$	0.2164	9.63	0.2370	0.8855	<0.0001	–	–	−0.6640	0.0113
ARIMA(1,0,1)(0,1,0)$_{12}$	0.2458	17.72	0.2935	0.7869	<0.0001	−0.0673	0.7669	–	–
ARIMA(1,0,0)(0,1,0)$_{12}$	0.2461	15.80	0.3134	0.7504	<0.0001	–	–	–	-

ARIMA= Autoregressive Integrated Moving Average; Fit = Fitting results; RMSE=Root Mean Square Error; AIC = Akaike's Information Criterion; Pred. = Prediction of ARIMA model; Coef. = Coefficient of AR (autoregressive), MA (moving average), and SMA (seasonal–moving average);

Table 5

Northeastern region: Cross–correlations between the pre–whitened climate series and cases series.

Variable	Lag										
	0	1	2	3	4	5	6	7	8	9	10
Rainfall	0.1600	−0.1790	0.0160	−0.0990	0.0730	0.0570	0.0220	−0.0820	0.1040	−0.2100	0.3330[**]
P–value	0.1784	0.1362	0.8960	0.4169	0.5528	0.6460	0.8611	0.5161	0.4141	0.0987	0.0081
T$_{mean}$	−0.2400[*]	0.2800[*]	0.0160	−0.0070	−0.1440	−0.0290	0.1250	0.1060	−0.2650[*]	0.1770	−0.0420
P–value	0.0427	0.0181	0.8970	0.9570	0.2421	0.8171	0.3192	0.3993	0.0345	0.1662	0.7468

Table 6

Summary of ARIMAX model fitting parameters in northeastern Thailand.

Model	Fit		Pred	Climate variables		
	RMSE	AIC	RMSE	Vars	Coef.	*P*–value
ARIMAX $(1,0,0)(0,1,1)_{12}$ with Rainfall	0.2265	12.50	0.2956	Rainfall (lag9)	−0.0088	0.2280
ARIMAX $(1,0,0)(0,1,1)_{12}$ with Rainfall	0.2182	8.53	0.2845	Rainfall (lag10)	0.0161	0.0196
ARIMAX $(1,0,0)(0,1,1)_{12}$ with T_{mean}	0.2194	11.11	0.2038	T_{mean} (lag0)	−0.4794	0.4693
ARIMAX $(1,0,0)(0,1,1)_{12}$ with T_{mean}	0.2177	10.34	0.1836	T_{mean} (lag1)	0.9066	0.1325
ARIMAX $(1,0,0)(0,1,1)_{12}$ with T_{mean}	0.2196	9.35	0.3204	T_{mean} (lag8)	−1.2018	0.0535
ARIMAX $(1,0,0)(0,1,1)_{12}$ with Rainfall, T_{mean}	0.2088	5.69	0.3129	Rainfall (lag10)	0.0151	0.0242
				T_{mean} (lag8)	−1.3618	0.0256

ARIMAX = Autoregressive Integrated Moving Average with input series; Fit = Fitting results; RMSE=Root Mean Square Error; AIC = Akaike's Information Criterion; Pred. = Prediction of ARIMA model; Coef. = Coefficient of climate variables; Lag = time lag of climate variables.
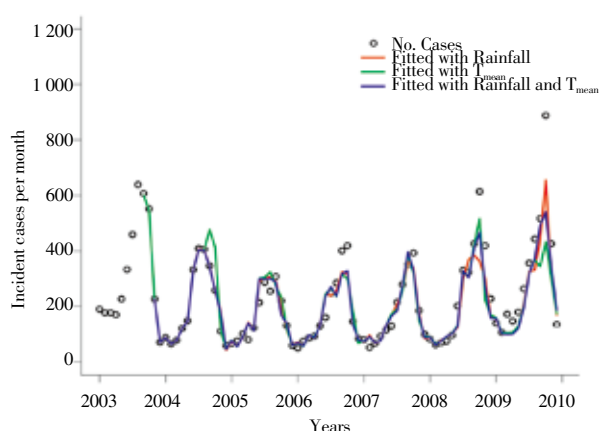
Figure 6. Fitted and predicted values of ARIMAX$(1,0,0)(0,1,1)_{12}$ associate with rainfall (red line), with T_{mean} (green line), and with rainfall and T_{mean} (blue line) in the northeastern region.

4. Discussion

To predict the number of leptospirosis cases, we first investigated the non–seasonal and seasonal ARIMA models without the environmental factors. It is found that only seasonal ARIMA model has the ability to predict the future leptospirosis cases for both the northern and northeastern region. The best univariate model is ARIMA(1,0,0)(0,1,1)12, where current number of leptospirosis cases depend on the number of cases in the previous month as well as the moving average in the previous season. In the multivariate ARIMA model, the result shows the appropriate model for the northern region is the model that only includes the rainfall factor while the model which includes rainfall and temperature factors is the best model for the northeastern region.

The monthly number of reported cases of leptospirosis all year round showed a seasonality pattern. The peak in the seasonal curve in the northern region occurred in August during the rainy season, while the peak in the seasonal curve in the northeastern region occurs between August and October. The significant time lag of average rainfall is 8 months in the northern region, and the correlation is negative.

During the rainy season, the soil will accumulate moisture, leading to small or large water pools. These facilitate leptospira organisms' growth in water–soaked soils. The bacteria can survive for 1 to 2 months[10,20]. In the natural water source, *leptospira* spp. can survive even with exposure to UV–A for 6 hours[21]. These may mean greater chance for people to be exposed to leptospira organisms during their agricultural activities. In the northern and northeastern region, the rainy season period is about 6–8 months every year. This may allow longer exposure to leptospira organisms. Flooding is related to the outbreaks in both human and animals, since it leads to more water contaminated with the urine of infected animals. In 2002, Ward showed that there was a correlation between rainfall and canine leptospirosis[22]. In many tropical countries, the heavy rain and/or flood increased the size of outbreaks, via indirect transmission from contaminated water after flooding [23–24].

In this work, we have found that the correlation between pre–whitened temperature and leptospirosis cases is only significant in the northeastern region. Although the temperature patterns and the 7 years average temperatures were almost the same, about 28 ℃, in both regions, the average numbers of reported leptospirosis cases in the northern region were fewer than that observed in the northeastern region by a factor of 4. The small number of cases may increase the error and cause the low variation when we added more input series. The temperature of about 28–30 ℃ is optimum for lepospires growth[25–28], corresponding to the rainy season. Leptospirosis occurs in countries with warm–climate rather than those in the temperate zone, where the average temperature is about 23.5 ℃. Leptospires are able to survive longer in a warm and humid environment. In higher temperature, human and animals may participate in water–based activities such as swimming, bathing or drinking. These activities increase contact probability of human and leptospires[8].

Our model is able to show the trend of reported leptospirosis curve and predict the future occurrence relatively accurate. However, it still cannot accommodate the extreme values[10]. Leptospirosis are also determined by other factors, which influence the survival of leptospira such as the water's concentration of oxygen[29], or water pH [29,30]. Variations in daily human activities might increase the probability of infection as well. In this work, we also did not include the data on cyclones or climate depressions in our model. However, this model may prove to be most useful in our attempt at monitoring and predicting leptospirosis cases in order to be well prepared to provide adequate and timely

health care services.

References

[1] Fraga TR, Barbosa AS, Isaac L. Leptospirosis: Aspects of innate immunity, immunopathogenesis and immune evasion from the complement system. *Scand J Immunol* 2011; **73**(5): 408–419.

[2] DebMandal M, Mandal S, Pal NK. Serologic evidence of human leptospirosis in and around Kolkata, India: A clinico– epidemiological study. *Asian Pac J Trop Med* 2011; **4**(12): 1001– 1006.

[3] Tangkanakul W, Smits H, Jatanasen S, Ashford D. Leptospirosis: an emerging health problem in Thailand. *Southeast Asian J Trop Med Public Health* 2005; **36**(2): 281–188.

[4] Trivedi TH, Kamath SA. Leptospirosis: Tropical to subtropical India. *JAPI* 2010; **58**:351.

[5] Radl C, Müller M, Revilla–Fernandez S, Karner–Zuser S, de Martin A, Schauer U, et al. Outbreak of leptospirosis among triathlon participants in Langau, Austria, 2010. *Wien Klin Wochenscha* 2011; **123**: 751–755.

[6] Ridzlan F, Bahaman A, Khairani–Bejo S, Mutalib A. Detection of pathogenic *Leptospira* from selected environment in Kelantan and Terengganu, Malaysia. *Trop Biomed* 2010; **27**(3): 632–638.

[7] Victoriano A, Smythe L, Gloriani–Barzaga N, Cavinta L, Kasai T, Limpakarnjanarat K, et al. Leptospirosis in the Asia Pacific region. *BMC Infect Dis* 2009; **9**(1):147.

[8] Lau CL, Smythe LD, Craig SB, Weinstein P. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? *Trans R Soc Trop Med Hyg* 2010; **104**(10): 631–638.

[9] Triampo W, Baowan D, Tang I, Nuttavut N, Wong–Ekkabut J, Doungchawee G. A simple deterministic model for the spread of leptospirosis in Thailand. *Int J Biol Med Sci* 2007; **2**(1): 22–26.

[10] Desvars A, JÈgo S, Chiroleu F, Bourhy P, Cardinale E, Michault A. Seasonality of human leptospirosis in reunion island (Indian Ocean) and its association with meteorological data. *PLoS One* 2011; **6**(5): e20377.

[11] Robertson C, Nelson T, Stephen C. Spatial epidemiology of suspected clinical leptospirosis in Sri Lanka. *Epidemiol Infect* 2011; **1**(1): 1–13.

[12] Permanasari AE, Rambli DRA, Dominic PDD. Performance of univariate forecasting on seasonal diseases: the case of tuberculosis. In: Arabnia HRR, Tran QN. (eds.). Software Tools and algorithms for biological systems. New York: Springer; 2011, p.171–179.

[13] Luz PM, Mendes BVM, CodeÁo CT, Struchiner CJ, Galvani AP. Time series analysis of dengue incidence in Rio de Janeiro, Brazil. *Am J Trop Med Hyg* 2008; **79**(6): 933–939.

[14] Soebiyanto RP, Adimi F, Kiang RK. Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLoS One* 2010; **5**(3): e9450.

[15] Wangdi K, Singhasivanon P, Silawan T, Lawpoolsri S, White N, Kaewkungwal J. Development of temporal modelling for forecasting and prediction of malaria infections using time–series and ARIMAX analyses: A case study in endemic districts of Bhutan. *Malaria J* 2010; **9**(1): 251.

[16] Yunibandha J. First report of Weilís disease in Thailand. *J Med Assoc Thailand* 1943; **26**: 83–136.

[17] Bureau of Epidemiology DoDC, Ministry of Public Health, Thailand. National notifiable disease surveillance (Report 506). 2003–2009. [online] Avaliable from: http://www.boe.moph.go.th/ boedb/surdata/disease.php?ds=43. [Accessed September, 14, 2011].

[18] Cryer JD, Chan K–S. *Time series analysis: with applications in R*. 2 ed. New York: Springer Science Business Media; 2008.

[19] Chatfield C. *Time–series forecasting*. 1 ed. Florida: Chapman &C Hall/CRC; 2000.

[20] Biggs HM, Bui DM, Galloway RL, Stoddard RA, Shadomy SV, Morrissey AB, et al. Leptospirosis among hospitalized febrile patients in northern Tanzania. *Am J Trop Med Hyg* 2011; **85**(2): 275–281.

[21] Chadsuthi S, Wong–Ekkabut J, Triampo W, Doungchawee G, Triampo D. Comparison of the effects of UV–A radiation on Leptospira interrogan serovar Bataviae, Canicola and Pomona. *Afr J Biotechnol* 2010; **9**(21): 3196–3206.

[22] Ward MP. Seasonality of canine leptospirosis in the United States and Canada and its association with rainfall. *Prev Vet Med* 2002; **56**(3): 203–213.

[23] Sethi S, Sharma N, Kakkar N. Increasing trends of leptospirosis in Northern India: A clinico–epidemiological study. *PLoS Negl Trop Dis* 2010; **4**(1): e579.

[24] Mohan ARM, Cumberbatch A, Adesiyun AA, Chadee DD. Epidemiology of human leptospirosis in Trinidad and Tobago, 1996 – 2007: A retrospective study. *Acta Trop* 2009; **112**(3): 260– 265.

[25] Adler B, De La Peña Moctezuma A. *Leptospira* and leptospirosis. *Vet Microbiol* 2010; **140**(3–4): 287–296.

[26] Sakhaee E, Abdollah pour GR. Detection of leptospiral antibodies by microscopic agglutination test in north–east of Iran. *Asian Pac J Trop Biomed* 2011; **1**(3): 227–229.

[27] DebMandal M, Mandal S, N Kumar Pal. Is jaundice a prognosis of leptospirosis? *Asian Pac J Trop Dis* 2011; **1**(4): 279–281.

[28] Khan SA, Dutta P, Borah J, Chowdhury P, Topno R, Baishya M, et al. Leptospirosis presenting as acute encephalitis syndrome (AES) in Assam, India. *Asian Pac J Trop Dis* 2011; **2**(2): 151–153.

[29] Xue F, Dong H, Wu J. Transcriptional responses of Leptospira interrogans to host innate immunity: significant changes in metabolism, oxygen tolerance, and outer membrane. *PLoS Negl Trop Dis* 2010; **4**(10): e857.

[30] Parker J, Walker M. Survival of a pathogenic Leptospira serovar in response to combined in vitro pH and temperature stresses. *Vet Microbiol* 2011; **152**(1 – 2): 146–150.