Infection and Drug Resistance 



ORIGINAL RESEARCH

# Comparison of ARIMA and Bayesian Structural Time Series Models for Predicting the Trend of Syphilis Epidemic in Jiangsu Province

Fengquan Zhang[1,*], Yanyan Li[2,*], Xinxiao Li[2], Bingjie Zhang[2], Chenlu Xue[2], Yongbin Wang[2]

[1]Center for Experimental Teaching, School of Public Health, Xinxiang Medical University, Xinxiang, Henan Province, 453003, People's Republic of China; [2]Department of Epidemiology and Health Statistics, School of Public Health, Xinxiang Medical University, Xinxiang, Henan Province, 453003, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yongbin Wang, Department of Epidemiology and Health Statistics, School of Public Health, Xinxiang Medical University, Xinxiang, Henan Province, 453003, People's Republic of China, Email wybwho@163.com



**Purpose:** This study sets out to explore the forecasting value in syphilis incidence of the Bayesian structural time series (BSTS) model in Jiangsu Province.

**Methods:** The seasonal autoregressive integrated moving average (ARIMA) and BSTS models were constructed using the series from January 2017 to December 2021, and the prediction accuracy of both models was tested using the series from January 2022 to November 2022.

**Results:** From January 2017 to November 2022, the total number of syphilis cases in Jiangsu Province was 170629, with an average monthly notification cases of 2403. The optimal model was ARIMA (1,0,0) (0,1,1) 12 (AIC = 663.12, AICc = 664.05, and BIC = 670.60). The model coefficients were further tested: AR1 = 0.48 (t = 3.46, P < 0.001), and SMA1 =−0.48 (t =−2.32, P = 0.01). The mean absolute deviation, mean absolute percentage error, root mean square error, and root mean square percentage error from the BSTS model were smaller than those from the ARIMA model. The total number of syphilis cases predicted by the BSTS model from December 2022 to December 2023 in Jiangsu Province was 29902 (95% CI: 16553 ~ 42,401), with a monthly average of 2300 (95% CI: 1273 ~ 3262) cases.

**Conclusion:** Syphilis is a seasonal disease in Jiangsu Province, and its incidence is still at a high level. The BSTS model is superior to the ARIMA model in dynamically predicting the incidence trend of syphilis in Jiangsu Province and has better application value.

**Keywords:** syphilis, trends, forecast, BSTS model, ARIMA model

## Introduction

Syphilis is an infectious disease caused by Treponema pallidum. The disease is mainly transmitted through direct contact with sex, blood, and mother to child. It can also be transmitted through contact with daily items or medical devices contaminated by patients. It is highly infectious and is generally susceptible to infection among the population. In recent years, the incidence rate of syphilis has increased.[1] Syphilis is prevalent worldwide. According to WHO estimates, there are approximately 12 million new cases of syphilis worldwide each year, mainly in South Asia, Southeast Asia, and Africa. In the past years, syphilis has grown rapidly in China and has become the sexually transmitted disease with the largest number of reported cases. Among the reported syphilis cases, latent syphilis accounts for the majority, and primary and secondary syphilis are also relatively common. In patients with primary syphilis, within 3 weeks after infection, there is a single ulcer or multiple lesions in the genitals or other related body parts, accompanied by local Lymphadenopathy; these symptoms are usually painless and self-resolved. The primary lesion subsides after 6–8 weeks, followed by secondary manifestations, including fever, headache, and papules on the side, shoulders, arms, chest, or back, typically involving the palms and soles of the feet. As the signs and symptoms subside, the patient enters a latent



Received: 5 February 2024
Accepted: 9 December 2024
Published: 20 December 2024

period, which may last for many years. Patients with a latent period of 1–2 years are still considered infectious, as the risk of secondary syphilitic recurrence is 25%.[2,3] The number of reported cases of congenital syphilis is also increasing. In Western Europe, the United States, and China, Syphilis has increased significantly in key populations such as gay men.[4,5] Syphilis is still prevalent in low- and middle-income countries.[6] In 2016, WHO launched the initiative to eliminate congenital syphilis. Syphilis is a highly contagious sexually transmitted disease that is classified as a Class B disease for prevention and management. It can cause serious harm to human physical, psychological, and reproductive health and is a public health and social issue of key concern. Syphilis is the second-leading cause of adverse pregnancy outcomes, including fetal and neonatal deaths,[7] resulting in a neglected issue and poses considerable challenges to many health systems.

The severity and complexity of syphilis make it a significant public health concern. Its infectivity is extremely strong, and its clinical manifestations after infection are complex, which can cause damage and pathological changes to almost all tissues and organs of the body. In addition to direct disease, it will also increase the risk of HIV infection. The incidence of syphilis has increased among HIV-infected patients both in Europe and in the USA, and especially in the homosexual/bisexual transmission group.[8] If not treated early, this disease may persist for several years in a series of clinical stages and lead to irreversible neurological or cardiovascular complications. In the late stage, it can cause dysfunction, tissue damage, and even death. Preliminary data from 2021 shows that there are over 171000 syphilis cases in all stages, an increase of 68% compared to 2017. In only one year from 2020 to 2021, syphilis cases will increase by 28%, which indicates that syphilis is still a major public health problem even during the COVID-19.[9]

In this study, Bayesian Structural Time Series (BSTS) model was used to predict the incidence trend of syphilis in Jiangsu Province, and its prediction effect was compared with the seasonal autoregressive integrated moving average (ARIMA) model to explore the application value, with a view to providing a reference basis for the formulation of syphilis prevention and control policies and the optimal allocation of health resources.

# Materials and Methods

## Material

Obtain the syphilis incidence data from the Jiangsu Provincial Health and Family Planning Commission from January 2017 to November 2022. The series from January 2017 to December 2021 was used as the training set, it proved the predictive potential of BSTS method and its applicability and sufficiency in estimating the epidemiological trend of syphilis incidence rate, and that from January 2022 to November 2022 as the test set to verify the prediction accuracy of the model. This study protocol was approved by the study institutional review board of the Xinxiang Medical University (No: XYLL-2019072). We collect data anonymously. This data is second-hand and publicly available, so it does not need ethics.

## ARIMA Model

The ARIMA model consists of three parts: "AR" represents an autoregressive model, "I" represents a difference, and "MA" represents a moving average model. For periodic and seasonal time series, it can be expressed as ARIMA (p, d, q) (P, D, Q)S (p is the number of non-seasonal AR orders, d is the number of non-seasonal differences, q is the number of non-seasonal MA orders, P is the number of seasonal AR orders, D is the number of seasonal differences, and Q is the number of seasonal MA orders).[10,11] The ARIMA model modeling process includes:[12–14] ① Testing for stationarity. The ARIMA model requires the data to be stationary. Various tests, such as the Augmented Dickey–Fuller (ADF) test, can determine if a series is stationary. If the series is not stationary, it is essential to use differential and/or data transformation methods to stabilize it. ② Identifying ARIMA Parameters: The next step in constructing an ARIMA model is to identify the appropriate orders of the AR and MA terms, as well as their seasonal counterparts. This can be done using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. Once the orders of the terms are identified, the parameters of the ARIMA model can be estimated using the "coeftest" function in R. The estimated model can then be summarized to review the coefficients of the AR and MA components, alongside associated statistics (eg, *p*-values).

A further step may include utilizing criteria such as Akaike Information Criterion (AIC), corrected AIC (AICc) or Bayesian Information Criterion (BIC) for model selection, favoring models with lower values. ③ Model diagnostics are crucial to confirm the adequacy of the fitted model. The residuals (difference between the predicted values and the actual observations) should resemble white noise; they must be normally distributed with a mean of zero. Tools such as the Ljung–Box test, ACF of residuals, and Q-Q plots can help assess these conditions. Any indications of autocorrelation in the residuals may require model adjustments or selection of a different set of parameters. ④ Model prediction: With a validated ARIMA model, we can now make forecasts. The model can predict future values based on historical data, capturing the underlying patterns and seasonality. It is vital to evaluate forecast accuracy periodically as new data become available, and to adjust the model as necessary to accommodate any changes in the data characteristics over time.

## BSTS Model

Bayesian structural time series (BSTS) model generally uses the spike-slab prior regression component and Kalman filters for the time series component. The BSTS model is a random state space model that can study trends, seasonality, and regression components separately.[15] The BSTS model combines prior information with a likelihood function to generate a final Bayesian model. Use Bayesian model averaging and Kalman filtering to generate more accurate predictions.[16] Due to the use of Bayesian methods combined with Markov Chain Monte Carlo (MCMC) to estimate the distribution of parameters, predictive values after intervention are estimated based on parameter distribution. Local linear trends quickly adopt local changes, making them more suitable for short-term prediction.[17] When predicting BSTS, it does not rely on specific assumptions in the data. It can energize the posterior uncertainties of different components in the sequence, impose a prior probability on the model, and smooth out the uncertainties in the sequence.[18] Bayesian methods can be effectively used to predict the pattern of this epidemic.[19] When the ARIMA model predicts, it is assumed that the time series has a linear trend of change, and when external explicit variables are included, it is easy to lead to overfitting. 15 Short-term prediction performance is relatively high, but when the data generation mechanism is highly non-linear, it may lead to prediction failure. BSTS models screen variables spike and slab regression to avoid overfitting. It can effectively compensate for the shortcomings of ARIMA, which is also the reason why its prediction accuracy is higher than ARIMA.

## Statistical Analysis

Using the Hodrick–Prescott (HP) and seasonal index methods to analyze long-term trends, cycles, and seasonal index (SI) of the data. The seasonal index calculation of syphilis is based on the average value of the same month over the years as the numerator, and the monthly average value of each year as the denominator. The study was analyzed using R software. The ARIMA model is calculated using the auto.ARIMA function available in the prediction package, while the BSTS package is used to obtain the prediction of the BSTS model. The prediction accuracy of the proposed model is compared with that of the ARIMA model using the mean absolute deviation (MAD), mean absolute percentage error (MAPE), Root-mean-square deviation (RMSE), and root mean square percentage error (RMSPE). The above measures can effectively determine the prediction accuracy of the model. These prediction accuracy standards indicate that an improvement in prediction accuracy was observed under the BSTS model. Bayesian methods can be effectively used to predict the patterns of this epidemic.

# Results

From January 2017 to November 2022, the total number of syphilis cases in Jiangsu Province was 170629, the annual incidence rate is 2.58/100000 people. The highest incidence rate of syphilis is 3.12/100000 people in 2021, and the lowest incidence rate of syphilis is 2.58/100000 people in 2017. With an average monthly notification cases of 2403. The cycle and trend pattern of syphilis incidence sequence decomposed by HP technology (Figure 1) indicated that the overall incidence of syphilis in Jiangsu Province was on the rise, with a certain cyclical pattern. The seasonal index obtained by the seasonal index method (Table 1) from January to December for syphilis was 0.86, 0.73, 0.95, 1.0, 1.08, 1.08, 1.14, 1.08, 1.06, 0.99, 0.96, and 0.98, indicating that there were seasonal fluctuations in the incidence of syphilis in Jiangsu

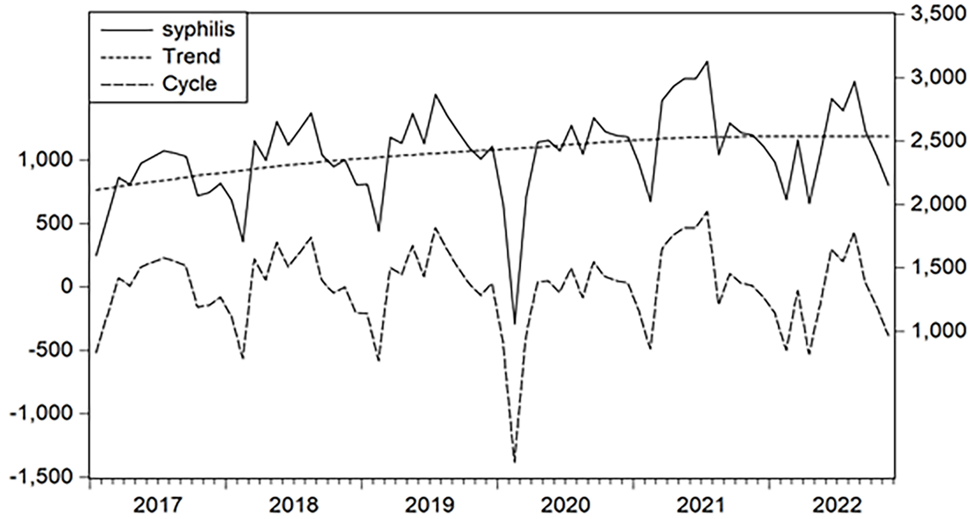


**Figure 1** Trend and periodic sequence of syphilis incidence and HP in Jiangsu province from January 2017 to November 2022.

Province, with the lowest incidence from January to February each year, followed by October to December; The highest incidence occurred in July, followed by April June and August September.

## ARIMA and BSTS Model Parameter Selection

### ARIMA Model

We fitted the incidence series of syphilis in Jiangsu Province from January 2017 to November 2022 based on the modeling steps of the ARIMA model. The ACF and PACF plots showed that syphilis had a significant correlation, but there was no significant correlation in the time series after difference (Figure 2a). We thus identified five candidate models, by simulating the ARIMA (1,0,0) $(0,1,1)_{12}$ structure, the corresponding AIC = 663.12, AICc = 664.05, and BIC = 670.60 were minimized in all candidate models. Likewise, we ran the "auto.arima" function in R, which also identified the ARIMA (1,0,0) $(0,1,1)_{12}$ as the best possible model. Therefore, the structure is selected as the optimal model. The model coefficients were further tested: AR1 = 0.48 ($t$=3.46, $P$<0.001), SMA1=−0.48 ($t$=−2.32, $P$=0.01), and the coefficients passed statistical tests. The ACF and PACF plots of the model residuals suggested that the different lag correlation coefficients basically fell within 95% *CI* (Figure 2b); the Ljung-Box Q test results indicated that the model residual sequence belongs to a white noise series ($\chi^2$=22.74, $P$=0.83). Therefore, the ARIMA (1,0,0) $(0,1,1)_{12}$ structure has passed all diagnoses, indicating that it is capable of fully fitting the incidence trend of syphilis data.

### BSTS Model

Trend components include the local-level model (assuming that the trend follows a random walk pattern, suitable for short-term prediction), the local linear trend model (assuming that both the mean and slope of the trend follow a random walk pattern, suitable for short-term prediction), and the semi-local linear trend model (assuming that the horizontal component changes in a random walk pattern, but the slope changes in an autoregressive process centered on a non-zero value, suitable for long-term prediction) and shared local-level model (assuming that the trend follows a multivariate

**Table 1** Seasonal Index of Syphilis Obtained by Seasonal Index Method from January to December

| Month | Seasonal Index | Month | Seasonal Index | Month | Seasonal Index |
|---|---|---|---|---|---|
| 1 | 0.86 | 5 | 1.08 | 9 | 1.06 |
| 2 | 0.73 | 6 | 1.08 | 10 | 0.99 |
| 3 | 0.95 | 7 | 1.14 | 11 | 0.96 |
| 4 | 1.00 | 8 | 1.08 | 12 | 0.98 |

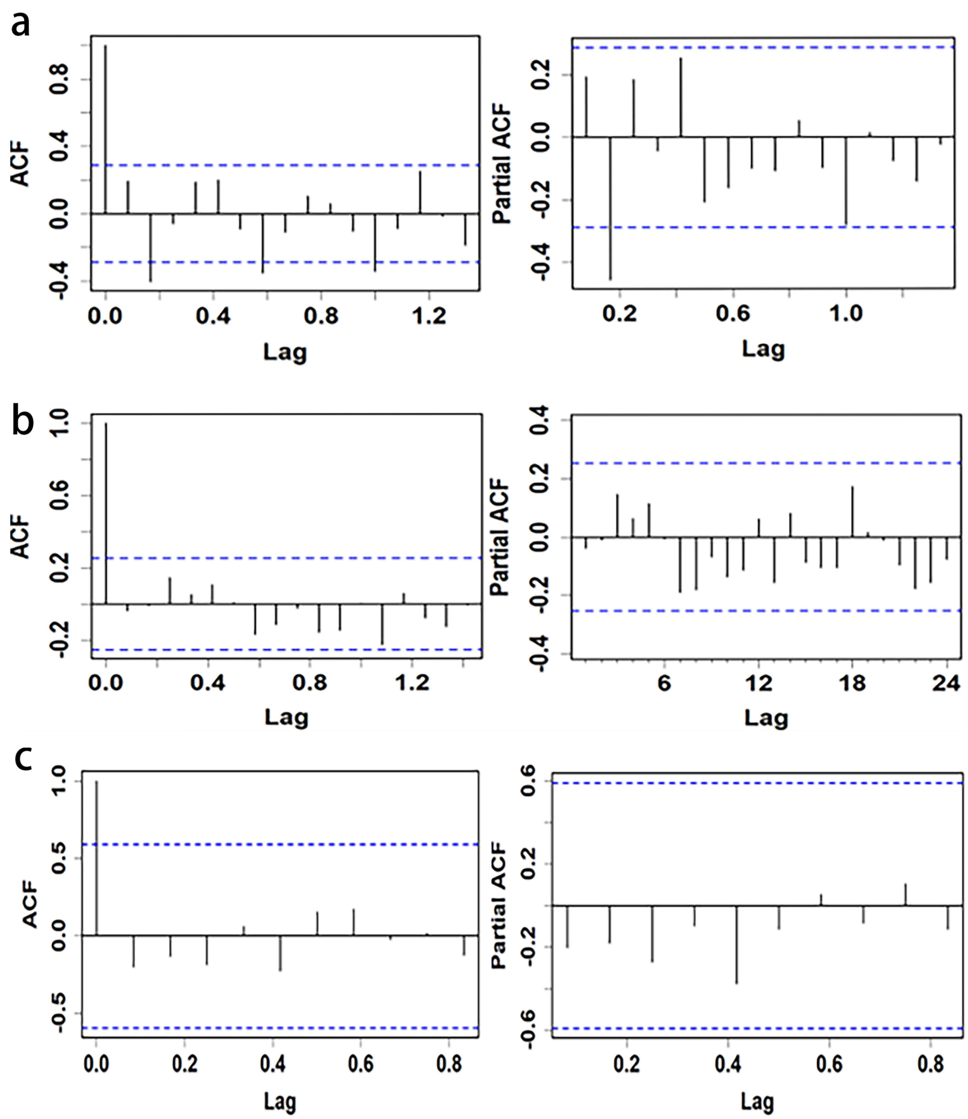


**Figure 2** ACF and PACF Plots after differential analysis of syphilis incidence sequence(**a**). ACF and PACF Plots of ARIMA (1,0,0) $(0,1,1)_{12}$ model residuals(**b**). ACF and PACF Plots of BSTS model residuals(**c**).

random walk pattern, suitable for short-term prediction).[20] The best BSTS model was created by modeling the morbidity series of syphilis using the 1000 times MCMC algorithm and using semi-local linear and seasonal state components. The model diagnosis results showed that the correlation coefficients in the residual ACF and PACF plots fell within 95% *CI* (Figure 2c); The Ljung-Box *Q* test results showed a $\chi^2$=0.863 (*P*=0.65), indicating that the model residual series belongs to a white noise series. The diagnosis results suggested that using BSTS model to simulate syphilis data is sufficient and appropriate.

## Prediction Results and Accuracy Evaluation Between BSTS and ARIMA Models

Table 2 lists the prediction results of the optimal ARIMA (1,0,0) $(0,1,1)_{12}$ and BSTS methods on the incidence of syphilis from January 2022 to November 2022. In order to test the prediction robustness of the model, the ARIMA and BSTS models were further reconstructed using the data from January 2017 to November 2022. The performance comparison results using the optimal model ARIMA (1,0,0) $(0,1,1)_{12}$ and BSTS models for the prediction from January 2022 to November 2022 showed that the prediction error indicators of MAD, MAPE, RMSE, and RMSPE under BSTS model were all smaller than ARIMA (Table 3). The above results indicated that the prediction accuracy of BSTS was higher

**Table 2** Predicted Values of ARIMA (1,0,0) $(0,1,1)_{12}$ and BSTS for Syphilis Incidence from January 2022 to November 2022

| Time | Observed Values | ARIMA | | BSTS | |
|---|---|---|---|---|---|
| | | Forecast | 95% CI | Forecast | 95% CI |
| 2022–01 | 2334 | 2286 | 1851~2722 | 2293 | 1813~2752 |
| 2022–02 | 2039 | 1880 | 1397~2362 | 1805 | 1170~2461 |
| 2022–03 | 2512 | 2719 | 2227~3212 | 2568 | 1824~3245 |
| 2022–04 | 2009 | 2878 | 2383~3373 | 2618 | 1850~3333 |
| 2022–05 | 2397 | 2972 | 2477~3468 | 2727 | 1872~3573 |
| 2022–06 | 2836 | 2915 | 2419~3410 | 2664 | 1763~3546 |
| 2022–07 | 2741 | 3094 | 2598~3589 | 2891 | 1919~3896 |
| 2022–08 | 2971 | 2639 | 2143~3134 | 2614 | 1599~3684 |
| 2022–09 | 2578 | 2802 | 2307~3298 | 2697 | 1618~3757 |
| 2022–10 | 2381 | 2702 | 2207~3198 | 2593 | 1497~3685 |
| 2022–11 | 2155 | 2579 | 2183~3174 | 2579 | 1504~3710 |

**Table 3** Comparison of Prediction Accuracy Between ARIMA (1,0,0) $(0,1,1)_{12}$ and BSTS Models

| Index | ARIMA Model | BSTS Model |
|---|---|---|
| MAD | 335.55 | 245.82 |
| MAPE | 14.40 | 10.57 |
| RMSE | 407.33 | 295.08 |
| RMSPE | 0.18 | 0.13 |

than that of the ARIMA, and the prediction results are robust. Therefore, using the data from January 2017 to November 2022, the BSTS model was reconstructed, and the number of syphilis cases from December 2022 to December 2023 was predicted (Table 4 and Figure 3). It was showed that the total number of new cases in the next 13 months was 29902 (95% *CI*: 16553 ~ 42,401), with a monthly average of 2300 (95% *CI*: 1273 ~ 3262) cases. The number of syphilis cases was still high and at a high level.

## Discussion

The emergence and resurgence of syphilis as a public health concern in various regions around the world have generated significant attention from health professionals, researchers, and policymakers alike. In Jiangsu Province, China, the increasing incidence of syphilis poses challenges for health authorities tasked with developing effective prevention and control strategies. With the advent of sophisticated statistical methods, such as the ARIMA model and the BSTS model, the ability to accurately forecast disease trends has become more attainable and practical. Employing statistical models to

**Table 4** Prediction Value of BSTS for Syphilis Incidence from November 2022 to December 2023

| Time | Observed Values | Forecast | 95% CI | Time | Observed Values | Forecast | 95% CI |
|---|---|---|---|---|---|---|---|
| 2022–12 | 1340 | 2241 | 1693~2699 | 2023–07 | 3526 | 2602 | 1462~3689 |
| 2023–01 | 1308 | 2057 | 1508~2620 | 2023–08 | 3657 | 2471 | 1275~3563 |
| 2023–02 | 2173 | 1664 | 962~2387 | 2023–09 | 3290 | 2420 | 1139~3605 |
| 2023–03 | 2589 | 2348 | 1524~3157 | 2023–10 | 3548 | 2299 | 953~3461 |
| 2023–04 | 2686 | 2287 | 1350~3149 | 2023–11 | 3267 | 2259 | 928~3519 |
| 2023–05 | 3036 | 2475 | 1502~3358 | 2023–12 | 2758 | 2285 | 886~3607 |
| 2023–06 | 3038 | 2494 | 1371~3587 | | | | |

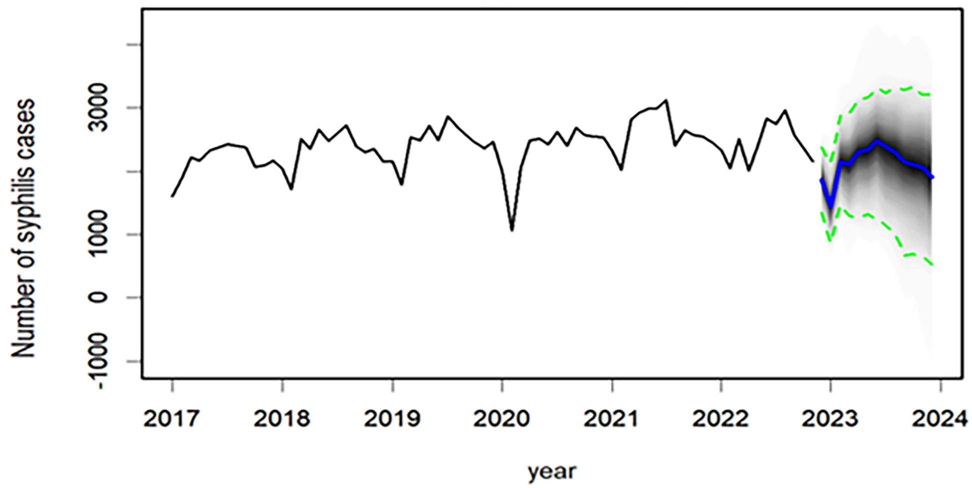


**Figure 3** Prediction of syphilis incidence by BSTS model from November 2022 to December 2023.

predict the incidence rates facilitates timely decision-making and effective allocation of resources to combat the spread of this infection. In the research, various approaches to predicting the occurrence of epidemics were introduced. ARIMA model was used to model and predict the incidence rate of syphilis in Jiangsu Province and performed well. The ARIMA model is a widely utilized method for time series forecasting, particularly in situations where the data exhibit temporal dependencies. This model operates on the principle that past values of a variable can inform future values. The advantage of ARIMA lies in its capacity to handle both autoregressive and moving average components while integrating differencing to ensure stationarity. However, it is ineffective for predicting the occurrence of infectious diseases that may be influenced by a number of variables because it can only extract linear relationships from time series data. ARIMA algorithm to forecast the incidence rate in the future. Despite the fact that these models are useful, their forecasting ability has some drawbacks. When using covariates, the overfitting will arise because the ARIMA model's prediction is dependent on the preceding value of the series and the prior prediction error. In order to overcome these limitations, Bayesian modeling methods have received attention in recent years. In contrast to the traditional ARIMA approach, the BSTS model introduces a more flexible framework for time series analysis. The BSTS model incorporates Bayesian inference, allowing for the integration of prior knowledge and uncertainty into the modeling process. It builds upon the components of trend, seasonality, and regressors to provide a comprehensive view of the underlying processes that shape the data. One of the significant advantages of the BSTS model is its ability to capture unobserved factors that may influence syphilis incidence.[17] For instance, the model can incorporate seasonal effects related to social behaviors or public health interventions, which might not be adequately accounted for in traditional ARIMA models. By leveraging the BSTS model, researchers can develop predictions that encompass a broader range of influencing variables, ultimately offering a more nuanced forecast of syphilis trends.

It is reported that BSTS model has better performance than ARIMA model in predicting the incidence of tuberculosis in China.[18] Simultaneously, it is better than ARIMA model in predicting the trend of COVID-19 in Pakistan.[16] Navid Feroze used the BSTS model to predict the COVID-19 pattern and the causal impact of the blockade of the affected countries in India, Brazil, the United States, Russia, and the United Kingdom. The results showed that the BSTS model produced more accurate predictions than the ARIMA model.[15] In forecasting the incidence rate of syphilis in Jiangsu, the BSTS model performed better than the ARIMA model. As a result, it has been demonstrated that the BSTS model is a promising substitute for the ARIMA model for forecasting disease trends. It has been established that the BSTS model is successful at forecasting the incidence rate of infectious diseases in the future. The time series prediction algorithm offers some distinctive benefits. First of all, they can make precise predictions by completely utilizing the time information of the original dataset. Second, because the modeling procedure is straightforward, the model can be applied generally. The time series model investigated in this study can be used to forecast the incidence of syphilis, allowing

Jiangsu Province's relevant departments to plan and implement quick-response actions that will maximize the prevention and control effect of STDs and resource mobilization.

The implications of accurately forecasting syphilis incidence trends using the ARIMA and BSTS models are multifaceted. First, precise forecasts enable public health authorities to anticipate surges in infection rates, thereby facilitating preemptive measures. This capability is particularly critical in resource allocation and planning preventive campaigns targeted at high-risk populations. Moreover, insights obtained from these statistical models can inform the design and implementation of health education programs. Understanding when and where syphilis cases are likely to rise can help health educators tailor their messages and outreach strategies effectively. For example, if forecasts indicate a spike in cases during specific months, targeted awareness campaigns can be scheduled to address sexual health and safe practices. Additionally, the integration of statistical forecasting can strengthen surveillance systems. By continually updating the models with new data, health authorities can refine their predictions and respond dynamically to changing trends. This responsiveness is vital in an era where the epidemiology of sexually transmitted infections is influenced by factors such as mobility, behavior changes, and public policy. The use of ARIMA and BSTS models to forecast syphilis incidence is not merely an academic exercise; it carries profound implications for health policy and system-level interventions. Policymakers can utilize these predictions to justify the allocation of resources toward testing, treatment, and preventive measures such as condom distribution and education about safe sexual practices. Furthermore, the findings from these models can effectively inform legislative discussions on public health funding. By demonstrating the potential for rising infection rates, health authorities can advocate for increased budget allocations to bolster public health initiatives aimed at reducing syphilis incidence. Moreover, these models can play a pivotal role in fostering collaboration among various stakeholders, including governmental agencies, non-governmental organizations (NGOs), and community-based organizations. A shared understanding of the trend forecasts can unify efforts toward common public health goals, leading to more comprehensive and effective campaigns. The research results showed that the incidence of syphilis is still on the rise in recent years, and there are still many obstacles that need to be overcome before the syphilis epidemic is fully controlled. During the epidemic, most of the medical resources were used to fight against COVID-19, and the policy restricted people's activities and incubation period, as well as people with mild symptoms who were afraid of medical infection and other factors, people did not go to hospital in time. The easing of policies has led to a sharp increase in the number of syphilis. Jiangsu Province exhibits an increase in syphilis. This is in line with findings from US studies.[21] Syphilis control funding is limited, and relevant agencies use medical resources and funds to prevent and contain the COVID-19 pandemic. Syphilis may become more widespread during the COVID-19 outbreak as a result of this. Plans for syphilis prevention could stall. The findings of this research indicated that there is a general increase in the incidence of syphilis in Jiangsu province, with significant seasonality and periodicity. Normal high incidence months are July and August. It might be connected to changes in hormone levels caused by the seasons. The findings of this research are in agreement with that of study conducted in China's Guangdong Province.[22] The seasonal index method is used to obtain the seasonal index of syphilis, which shows that the incidence of syphilis in Jiangsu Province has seasonal characteristics. The annual incidence trough of syphilis is in the Spring Festival per year, reaching the lowest level in January and February of each year, but then rapidly rising to a relatively high level. This is consistent with the research results of Zhixin Zhu et al on syphilis in China.[23,24] The reason may be that the patient did not see a doctor or delayed seeing a doctor during the Spring Festival. We will improve syphilis prevention and control during this time, create a complete syphilis prevention and control system, and promote early syphilis detection, early diagnosis, early treatment, and early control. Effective public health interventions should be implemented by the relevant agencies, such as monitoring syphilis cases, improving case report accuracy, and lowering the number of unreported, under-reported, and false reports. Boost funding, broaden syphilis screening among the populace, and raise the discovery threshold. We will enhance the management of floating populations and change unhealthy societal norms. The elderly, expectant mothers, adolescents, and high-risk groups should receive more information and health instruction. Additionally, it is important to improve supervision, share information on the prevention and treatment of STDs like syphilis, conduct screening for high-risk groups, increase screening for important groups, and stop the spread of syphilis. Jiangsu Province should implement comprehensive management strategies to lower the incidence rate of syphilis.

## Conclusion

The incidence trend of syphilis in Jiangsu Province is generally on the rise, with obvious seasonal and cyclical patterns; compared to ARIMA model, BSTS model has better application value in predicting the incidence trend of syphilis, and can be used to provide reference and technical support for the formulation of syphilis prevention and control strategies and measures.

## Data Sharing Statement

All the data supporting the findings of the work are contained within the study.


## Acknowledgments

We appreciate Jiangsu Health Commission for providing the monthly number of syphilis cases.


## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.


## Funding

There is no funding to report.


## Disclosure

The authors report no conflicts of interest in this work.

## References


1. Arando Lasagabaster M, Otero Guerra L. Syphilis. *Enferm Infecc Microbiol Clin*. 2019;37(6):398–404. doi:10.1016/j.eimc.2018.12.009
2. Rolfs RT, Joesoef MR, Hendershot EF, et al. A randomized trial of enhanced therapy for early syphilis in patients with and without human immunodeficiency virus infection. The syphilis and HIV study group. *N Engl J Med*. 1997;337(5):307–314. doi:10.1056/NEJM199707313370504
3. Peeling RW, Mabey D, Kamb ML, Chen XS, Radolf JD, Benzaken AS. Syphilis. *Nat Rev Dis Primers*. 2017;3(1):17073. doi:10.1038/nrdp.2017.73
4. Abara WE, Hess KL, Neblett Fanfair R, Bernstein KT, Paz-Bailey G. Syphilis trends among men who have sex with men in the United States and Western Europe: a systematic review of trend studies published between 2004 and 2015. *PLoS One*. 2016;11(7):e0159309. doi:10.1371/journal.pone.0159309
5. Nicoll A, Hamers FF. Are trends in HIV, gonorrhoea, and syphilis worsening in Western Europe? *BMJ*. 2002;324(7349):1324–1327. doi:10.1136/bmj.324.7349.1324
6. Green T, Talbot MD, Morton RS. The control of syphilis, a contemporary problem: a historical perspective. *Sex Transm Infect*. 2001;77(3):214–217. doi:10.1136/sti.77.3.214
7. Kojima N, Klausner JD. An update on the global epidemiology of syphilis. *Curr Epidemiol Rep*. 2018;5(1):24–38. doi:10.1007/s40471-018-0138-z
8. Pialoux G, Vimont S, Moulignier A, Buteux M, Abraham B, Bonnard P. Effect of HIV infection on the course of syphilis. *AIDS Rev*. 2008;10(2):85–92.
9. Ramchandani MS, Cannon CA, Marra CM. Syphilis: a modern resurgence. *Infect Dis Clin North Am*. 2023;37(2):195–222. doi:10.1016/j.idc.2023.02.006
10. Wang YW, Shen ZZ, Jiang Y. Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLoS One*. 2018;13(9):e0201987. doi:10.1371/journal.pone.0201987
11. Jian Y, Zhu D, Zhou D, et al. ARIMA model for predicting chronic kidney disease and estimating its economic burden in China. *BMC Public Health*. 2022;22(1):2456. doi:10.1186/s12889-022-14959-z
12. Swaraj A, Verma K, Kaur A, Singh G, Kumar A, Melo de Sales L. Implementation of stacking based ARIMA model for prediction of Covid-19 cases in India. *J Biomed Inform*. 2021;121:103887. doi:10.1016/j.jbi.2021.103887
13. Meng P, Huang J, Kong D. Prediction of incidence trend of influenza-like illness in Wuhan based on ARIMA model. *Comput Math Methods Med*. 2022;2022:6322350. doi:10.1155/2022/6322350
14. Zeng Q, Li D, Huang G, et al. Time series analysis of temporal trends in the pertussis incidence in Mainland China from 2005 to 2016. *Sci Rep*. 2016;6(1):32367. doi:10.1038/srep32367
15. Feroze N. Forecasting the patterns of COVID-19 and causal impacts of lockdown in top five affected countries using Bayesian structural time series models. *Chaos, Solitons Fractals*. 2020;140:110196. doi:10.1016/j.chaos.2020.110196
16. Feroze N, Abbas K, Noor F, Ali A. Analysis and forecasts for trends of COVID-19 in Pakistan using Bayesian models. *PeerJ*. 2021;9:e11537. doi:10.7717/peerj.11537

17. Vavilala H, Yaladanda N, Krishna Kondeti P, et al. Weather integrated malaria prediction system using Bayesian structural time series model for northeast states of India. *Environ Sci Pollut Res Int*. 2022;29(45):68232–68246. doi:10.1007/s11356-022-20642-y
18. Ding W, Li Y, Bai Y, Li Y, Wang L, Wang Y. Estimating the effects of the COVID-19 outbreak on the reductions in tuberculosis cases and the epidemiological trends in China: a causal impact analysis. *Infect Drug Resist*. 2021;14:4641–4655. doi:10.2147/IDR.S337473
19. Liu J, Spakowicz DJ, Ash GI, et al. Bayesian structural time series for biomedical sensor data: a flexible modeling framework for evaluating interventions. *PLoS Comput Biol*. 2021;17(8):e1009303. doi:10.1371/journal.pcbi.1009303
20. Scott S, Varian H. Predicting the present with bayesian structural time series. *Int J Math Model Numerical Optimisation*. 2014;5(1/2):4–23. doi:10.1504/ijmmno.2014.059942
21. Nazir A, Masood W, Ahmad S, et al. Rise of syphilis surge amidst COVID-19 pandemic in the USA: a neglected concern. *Ann Med Surg Lond*. 2022;80:104239. doi:10.1016/j.amsu.2022.104239
22. Tan NX, Tan GX, Yang LG, et al. Temporal trends in syphilis and gonorrhea incidences in Guangdong Province, China. *J Infect Dis*. 2014;209 (3):426–430. doi:10.1093/infdis/jit496
23. Zhu Z, Zhu X, Zhan Y, Gu L, Chen L, Li X. Development and comparison of predictive models for sexually transmitted diseases-AIDS, gonorrhea, and syphilis in China, 2011-2021. *Front Public Health*. 2022;10:966813. doi:10.3389/fpubh.2022.966813
24. Zhang X, Zhang T, Pei J, Liu Y, Li X, Medrano-Gracia P. Time series modelling of syphilis incidence in China from 2005 to 2012. *PLoS One*. 2016;11(2):e0149401. doi:10.1371/journal.pone.0149401