ARTIGO ORIGINAL

Estimates of gestational syphilis in Brazil using mathematical and artificial intelligence models

Estimativas sobre sífilis gestacional no Brasil com modelos matemáticos e de inteligência artificial

Maria Clara da Silva Maia[1], Antonio Manoel Ferreira Raymundo[1], Daiane Conceição de Araújo[1], Kelly de Almeida Schlager[1], Ketlin Angelin[1], Lídhia Cainnã de Souza Araújo[1], Daniela Teixeira Borges[2], Renata dos Santos Rabello[2]

1. Discente do curso de medicina, Universidade Federal da Fronteira Sul (UFFS), campus Passo Fundo, Chapecó, SC, Brasil. 2. Docente do curso de medicina, Universidade Federal da Fronteira Sul (UFFS), campus Passo Fundo, Chapecó, SC, Brasil.

Resumo

Objetivo: estimar novos casos de sífilis em gestantes no Brasil entre 07/2024 e 06/2027 utilizando os modelos Autorregressivo Integrado de Médias Móveis Sazonal (SARIMA), Memória de Curto-Longo Prazo (LSTM) e Suavização Exponencial Complexa (CES). **Métodos**: foram utilizados dados do Sistema de Informação de Agravos de Notificação referentes às notificações mensais entre janeiro de 2014 e junho de 2024. A validação ocorreu pelas métricas Erro Percentual Absoluto Médio (MAPE), Raiz do Erro Quadrático Médio (RMSE) e Erro Absoluto Médio (MAE). A implementação foi em Python. **Resultados**: a série não foi estacionária pelo teste de Dickey-Fuller aumentado, teve tendência crescente pelo teste de Mann-Kendall e autocorrelação dos dados pelo teste de Ljung-Box. O SARIMA (RMSE = 987; MAPE = 11%; MAE = 726) e o CES (RMSE = 1225; MAPE = 14%; MAE = 875) tiveram bom desempeno nas métricas de erro, mostrando-se adequados para a série em questão. Porém, o LSTM (RMSE = 811; MAPE = 9%; MAE = 610) foi superior a ambos. Até maio de 2027, os modelos divergiram em suas previsões. O SARIMA enxergou queda de casos, e o CES, alta. O LSTM estimou uma diminuição seguida de aumento de notificações. **Conclusões**: todos os modelos testados mostraram-se confiáveis para prever os casos de sífilis em gestantes para os próximos 3 anos no Brasil, mas o LSTM demonstrou maior precisão. Portanto, recomenda-se seu uso na tomada de decisão dos gestores públicos.

Palavras chave: infecções bacterianas; infecções sexualmente transmissíveis; algoritmos de predição; planejamento em saúde.

Abstract

Objective: tto estimate new cases of syphilis among pregnant women in Brazil between July 2024 and June 2027 using Seasonal Autoregressive Integrated Moving Average (SARIMA), Long Short-Term Memory (LSTM), and Complex Exponential Smoothing (CES) models. **Methods**: data from the Notifiable Diseases Information System on monthly notifications between January 2014 and June 2024 were used. Validation was performed using Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) metrics. Implementation was carried out in Python. **Results:** the time series was found to be non-stationary (Augmented Dickey-Fuller test), exhibited an upward trend (Mann-Kendall test), and showed autocorrelation (Ljung-Box test). SARIMA (RMSE = 987; MAPE = 11%; MAE = 726) and CES (RMSE = 1225; MAPE = 14%; MAE = 875) performed well regarding error metrics, proving suitable for the series in question. However, LSTM (RMSE = 811; MAPE = 9%; MAE = 610) outperformed both. The models diverge in their forecasts through May 2027: SARIMA predicts a decline in cases, while CES predicts an increase; LSTM estimates a decrease followed by a rise in notifications. **Conclusions**: all tested models proved reliable for forecasting syphilis cases among pregnant women in Brazil over the next three years, but LSTM demonstrated greater accuracy. Therefore, its use is recommended to support decision-making by public administrators.

Keywords: bacterial infections; sexually transmitted diseases; prediction algorithms; health planning.

INTRODUÇÃO

A sífilis congênita permanece como um importante problema de saúde pública global, refletindo falhas nos sistemas de prevenção, diagnóstico e tratamento oportuno durante o pré-natal[1]. A persistência de altas taxas da doença em países em desenvolvimento sugere que os esforços de eliminação ainda enfrentam barreiras estruturais, sociais e culturais que dificultam a adesão às estratégias de controle[2].

Essa doença é considerada evento sentinela da qualidade da de gestantes infectadas[3]. Nesse contexto, torna-se importante o desenvolvimento de pesquisas que olhem para a sífilis na gestação. Afinal, sua prevenção depende não apenas do diagnóstico precoce e do tratamento adequado, mas também do fortalecimento das redes de vigilância epidemiológica e da ampliação da cobertura de pré-natal[1].

Nesse contexto, a previsão de séries temporais é chave para a prevenção precoce de epidemias e melhor direcionamento das campanhas em saúde pública[4]. Essa ferramenta tem como objetivo determinar a ocorrência de epidemias futuras e, assim, viabilizar campanhas de prevenção mais assertivas ao possibilitar decisões mais embasadas dos gestores públicos[5]. Afinal, com a capacidade de identificar surtos de forma precoce, promove-se o uso otimizado de recursos e viabiliza-se

Correspondence: Maria Clara da Silva Maia. E-mail: maria.maia@estudante.uffs.edu.br
Conflict of interesse: There is no conflict of interest on the part of any of the authors.
Received: 2026 Ago 10; Revised: 2026 Ago 28; 2026 ; Accepted: 2026 Sep 1

a implementação de medidas preventivas antes que a situação alcance níveis críticos[6].

Assim, o presente estudo objetiva estimar a notificação de novos casos de sífilis na gestação no Brasil para os próximos três anos a partir de previsão de séries temporais, utilizando os modelos autorregressivo integrado de médias móveis sazonal (SARIMA), de memória de curto-longo prazo (LSTM) e suavização exponencial complexa (CES).

MÉTODOS

Tipo de estudo

Trata-se de um estudo observacional, transversal, de caráter descritivo e com abordagem quantitativa, que foi realizado na Universidade Federal da Fronteira Sul – Campus Passo Fundo, no período de março a setembro de 2025.

Fonte de dados

Foram utilizados dados secundários disponíveis no Sistema de Informação de Agravos de Notificação (SINAN), referentes às notificações de novos casos mensais de sífilis na gestação entre janeiro de 2014 e junho de 2024 no Brasil, de acordo com o ano de notificação e município de residência. Todos os registros foram incluídos na análise. Não houve cálculo e definição de tamanho da amostra, pois, por incluir todos os registros, a análise caracteriza-se como um censo.

Análise dos dados

Os dados foram extraídos automaticamente em formato de planilha, dispensando a digitação manual. A planilha foi utilizada como base de dados para os modelos SARIMA, CES e LSTM, implementados em linguagem Python, por meio das bibliotecas StatsForecast e Keras. O código foi escrito e executado em nuvem através da ferramenta Google Colabs, que é gratuita.

As bibliotecas Statsmodels e pyMannKendall foram utilizadas para decompor a série temporal e aplicar os testes de Mann-Kendall, Dickey-Fuller Aumentado e Ljung-Box. Esse último foi aplicado nos resíduos e na própria série temporal como forma de ratificar a presença de autocorrelação nos dados e independência dos ruídos da previsão. As métricas de erro foram calculadas por meio da biblioteca SkLearn.

Essas bibliotecas surgiram em um contexto de desenvolvimento de versões para Python de ferramentas estatísticas já presentes na linguagem R, objetivando tornar o Python a principal linguagem para análise estatística[7]. Portanto, foi considerada válida a sua experimentação na presente pesquisa.

Modelos preditivos

O estudo da série temporal inicia-se com sua decomposição em três componentes: tendência, sazonalidade e irregularidade. A análise da tendência pode ser realizada pelo teste de Mann-Kendall e, da estacionaridade, pelo teste de Dickey-Fuller Aumentado. A irregularidade não pode ser analisada[4].

Entre os modelos matemáticos mais utilizados para prever epidemias, destaca-se o ARIMA (modelo autorregressivo integrado de médias móveis). No entanto, ele não capta a influência da sazonalidade. Por isso, emprega-se, também, o modelo SARIMA (ARIMA sazonal)[8,9]. O SARIMA é definido por parâmetros que determinam os componentes sazonais e não sazonais da série (p, d, q) (P, D, Q)s, e sua configuração é auxiliada pela função de autocorrelação (ACF) e autocorrelação parcial (PACF)[8,10].

As Redes Neurais Recorrentes (RNN) constituem um modelo de rede neural voltado à análise de séries temporais. Cada neurônio recebe os dados provenientes da camada anterior e o resultado de sua própria ativação em etapas anteriores, o que possibilita à rede explorar a autocorrelação dos dados em seus processos decisórios[11]. A arquitetura de Memória de Longo e Curto Prazo (LSTM) é um tipo de RNN que supera dificuldades de versões anteriores e introduz mecanismos adicionais de controle por meio de três portas: esquecimento, entrada e saída[12].

A suavização exponencial simples utiliza uma média ponderada de dados históricos para prever valores futuros. Requer apenas um parâmetro de nível (α) entre 0 e 1 para controlar a rapidez com que a influência das observações passadas diminui. É considerado adequado apenas para dados de séries temporais curtas, sem tendência clara ou padrão sazonal13. Quando se objetiva prever séries com tendência e/ou sazonalidade, recomenda-se utilizar a suavização exponencial complexa. Nesse contexto, além do nível, existem parâmetros de tendência (β) e sazonalidade (γ)[14].

O teste de Ljung-Box foi aplicado ao final da execução dos modelos para detecção de autocorrelação nos ruídos. A validação ocorreu por métricas como erro percentual absoluto médio (MAPE), raiz do erro quadrático médio (RMSE) e erro absoluto médio (MAE), sendo que menores valores indicam previsões mais precisas.

Aspectos éticos

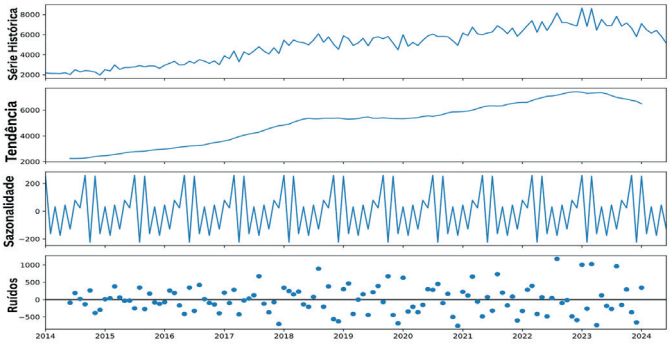
Como os dados foram coletados de domínio público, foi dispensada pela Comissão Nacional de Ética em Pesquisa (CONEP) a submissão ao Comitê de Ética em Pesquisa (CEP/CONEP) com Seres Humanos da UFFS, conforme a Resolução nº 510, de 07 de abril de 2016, do Conselho Nacional de Saúde (CNS). Por serem dados secundários e agregados por municípios, não existe risco de identificação pessoal dos sujeitos envolvidos.

RESULTADOS

A decomposição da série temporal de casos de sífilis na gestação no Brasil (figura 1) indica tendência de aumento progressivo entre 2014 e 2018, chegando a quase 6.000 casos mensais

no pico do período. Após esse crescimento, permaneceu em estabilidade, mantendo-se em torno de 5.500 casos por mês. A partir de 2021, a série voltou a crescer, aproximando-se de 8.000 casos reportados. No início de 2024, houve leve queda das notificações. O componente sazonal indicou padrão anual recorrente com aumento dos casos no final do ano. Os resíduos oscilaram em até 1.000 casos, refletindo variações não explicadas pela tendência ou sazonalidade.

Figura 1. Decomposição da série temporal de casos de sífilis na gestação no Brasil entre janeiro de 2014 e junho de 2024 em tendência, sazonalidade e resíduos respectivamente



Fonte: própria, 2025

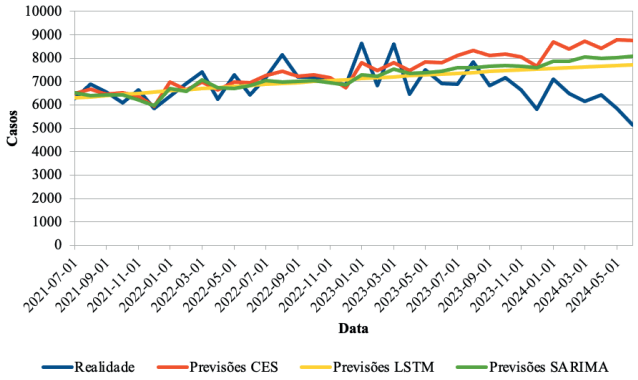
O teste de Dickey-Fuller aumentado indicou que a série não era estacionária (p = 0,30); o teste de Mann-Kendall confirmou uma tendência crescente. O teste de Ljung-Box (p = 1,27 × 10-110) apontou a autocorrelação dos dados, reforçando a possibilidade de aplicar métodos de previsão. Com base na função AutoARIMA da biblioteca StatsForecast, foi selecionado o modelo SARIMA (0,1,1)(1,0,0)[12] e α = 1,38; β = 1,01; γ = 0,01 para o CES como a ajuste ótimo entre as combinações testadas. Paralelamente, foi ajustado um modelo LSTM utilizando 400 neurônios, função de ativação ReLU, otimizador Adam e 200 epochs de treinamento.

A figura 2 apresenta a comparação entre as previsões dos modelos SARIMA, LSTM e CES entre as predições de ambos com os valores reais observados no período de julho de 2021 a junho de 2024. Nota-se que os três modelos acompanharam consistentemente a curva real de casos notificados nos dois primeiros anos. Todavia, a partir do terceiro ano, a confiabilidade cai de forma importante. Nessa reta final da validação, o LSTM foi o que mais se aproximou do que veio a ser a situação da época.

O desempenho dos modelos, avaliado por RMSE, MAE e MAPE, comprovou esse fato. Afinal, o LSTM obteve RMSE = 811 casos, MAPE = 9% e MAE = 610 casos. Enquanto o SARIMA ficou com RMSE = 987 casos, MAPE = 11% e MAE = 726 casos e, o CES, RMSE = 1225 casos, MAPE = 14% e MAE = 875 casos.

A análise dos resíduos pelo teste Q de Ljung-Box indicou presença de autocorrelação nos ruídos para os dois modelos matemáticos, com p = 1,02x10-7 para o SARIMA, p = 5,00x10-12 para o CES. Isso sugere que nem todos os padrões de dependência temporal foram completamente capturados. Mesmo com o LSTM alcançando maior precisão da estimativa, os erros continuaram sem independência (p = 8,58x10-6).

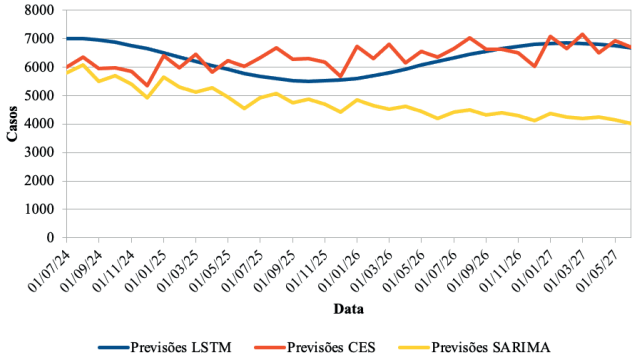
Figura 2. Comparação entre valores reais e previsões dos modelos SARIMA, LSTM e a média aritmética de suas previsões para os casos de sífilis na gestação no Brasil para um horizonte preditivo de 3 anos.



Fonte: própria, 2025

A figura 3 e a tabela 1 ilustram as projeções das notificações mensais de sífilis congênita no Brasil entre julho de 2024 e junho de 2027. Observa-se que cada modelo faz uma projeção diferente, gerando dúvidas de análise. O SARIMA prevê que a queda de casos presente no final da série continuará. O CES afirma que a tendência de alta anterior persistirá, tornando aquela queda passageira. Quanto ao LSTM, modelo que se mostrou o mais confiável pelas métricas de erro, foi estimado que haverá, sim, uma redução nas notificações; porém, apenas até 2026, quando ficará próximo dos 5.000 casos. Depois, ocorrerá nova elevação até alcançar o patamar ao redor de 7.000 casos, ponto que o CES também anteviu para junho de 2027.

Figura 3. Projeções de casos de sífilis na gestação no Brasil de julho de 2024 a junho de 2027 pelos modelos SARIMA, LSTM e CES



Fonte: própria, 2025

Tabela 1. Projeções de casos mensais de sífilis na gestação no Brasil de julho de 2024 a junho de 2027 pelos modelos SARIMA, LSTM e CES

Data	SARIMA	LTSM	CES
Jul/2024	5791	7002	6007
Ago/2024	6074	7004	6360
Set/2024	5514	6952	5965
Out/2024	5701	6872	5971
Nov/2024	5408	6767	5844
Dez/2024	4923	6644	5349
Jan/2025	5659	6506	6394
Fev/2025	5307	6358	5976
Mar/2025	5125	6207	6464
Abr/2025	5275	6058	5824
Mai/2025	4952	5917	6233
Jun/2025	4544	5789	6017
Jul/2025	4915	5681	6323
Ago/2025	5077	5596	6684
Set/2025	4758	5538	6288
Out/2025	4864	5512	6296
Nov/2025	4697	5516	6171
Dez/2025	4421	5551	5678
Jan/2026	4840	5614	6724
Fev/2026	4640	5703	6307
Mar/2026	4536	5813	6797
Abr/2026	4621	5937	6158
Mai/2026	4437	6071	6568
Jun/2026	4204	6206	6354
Jul/2026	4416	6337	6662
Ago/2026	4509	6459	7024
Set/2026	4326	6568	6630
Out/2026	4387	6662	6639
Nov/2026	4292	6739	6515
Dez/2026	4134	6796	6024
Jan/2027	4373	6833	7072
Fev/2027	4259	6845	6656
Mar/2027	4200	6835	7147
Abr/2027	4249	6802	6510
Mai/2027	4143	6750	6922
Jun/2027	4011	6680	6709

DISCUSSÃO

Um estudo com dados globais, regionais e nacionais para quatro defeitos congênitos principais (cardiopatias congênitas, defeitos do tubo neural, anomalias congênitas do sistema digestivo e síndrome de Down) utilizou o modelo Bayesiano de idade-período-coorte para prever as tendências futuras de incidência dessas condições. As previsões sugeriram que a carga de defeitos congênitos continuará a seguir a tendência de declínio já presente atualmente[15]. Isso demonstra que os modelos preditivos são opção para monitorar a epidemiologia de condições que afetam recém-nascidos, sejam elas infecciosas ou não.

Foram analisadas as curvas das notificações de sífilis primária, secundária e congênita na Coreia do Sul. A autocorrelação dos dados foi negativa para sífilis congênita e positiva para sífilis adquirida apenas em homens[10]. Esse resultado vai de encontro ao nosso, que observou a autocorrelação nos dados para sífilis na gestação. Isso indica que o comportamento da curva de notificações das infecções é muito particular em cada região.

O modelo SARIMA foi utilizado para prever casos de sífilis adquirida em homens no México. A série temporal foi não estacionária com tendência crescente de 2010 a 2015 e mudança de tendência a partir de 2016. A independência na distribuição dos ruídos pelo Teste de Ljung-Box ($p = 0,76$) foi utilizada para confirmar a viabilidade do modelo preditivo[16]. Nossos resultados também apresentaram uma série não estacionária com tendência crescente. Contudo, considerando que não se conseguiu obter independência na distribuição dos ruídos das previsões para sífilis na gestação, pode-se pensar que a transmissão é mais bem captada na população em geral em vez de em recortes.

Utilizando o modelo LSTM para prever os casos de sífilis adquirida no Brasil para o dia seguinte, obteve-se uma previsão média de 317,63 casos e um RMSE médio de 86,97 casos, representando 27,7% da quantidade média do número de casos. Quando os dados diários foram acumulados em semanas, o erro do modelo aumentou[17]. Esse estudo conversa com o nosso ao reiterar o baixo erro do LSTM em prever casos de sífilis. As métricas de precisão dele são maiores do que as encontradas por nós; porém, a diferença de extensão entre os horizontes preditivos explica esse fenômeno.

Os casos de sífilis congênita na província de Guangdong, na China, foram previstos com o modelo ARIMA. A série se mostrou não estacionária, e a precisão do modelo foi avaliada pela distribuição aleatória dos ruídos pelo Teste de Ljung-Box ($p > 0,05$). O número de casos previstos para 2021 foi 48, indicando uma epidemia de baixo nível[18]. Nosso estudo também concluiu que a série era não estacionária, porém não obtivemos a aleatoriedade do erro nem com o SARIMA, nem com os demais métodos. Desse modo, o ARIMA e seus derivados parecem captar melhor as variações de Guangdong do que as brasileiras.

Os modelos LightGBM e XGBoost foram utilizados para prever a probabilidade de uma criança nascer com sífilis. Para tal, foram utilizados dados do pré-natal, vida sexual materna e diagnóstico de sífilis na gestação. O resultado da validação do melhor ajuste

não foi considerado satisfatório, e a análise foi interrompida[19]. Isso indica que as técnicas de predição ainda não possuem limitações, como a dificuldade de levar em consideração todas as variáveis que influenciam o desfecho. Nosso estudo também teve problemas de adequação, como evidenciado pela não independência dos ruídos. Entretanto, as baixas métricas de erro sustentam a viabilidade da aplicação dos modelos testados.

Foram utilizados os modelos ARIMA, ERNN, ARIMA-ERNN e LSTM para prever as notificações de AIDS, gonorreia e sífilis na China. Para a previsão de 1 ano, os MAPEs para sífilis foram 9,80%, 9,55%, 8,67% e 5,79%, respectivamente. Para a previsão de 5 anos, foram 21,88%, 24,00%, 20,18% e 11,20%, respectivamente4. A superioridade do LSTM foi reafirmada pelo nosso estudo. Assim, mesmo com a confiabilidade dos modelos sendo uma variável dinâmica, reforça-se a indicação de preferir o uso do LSTM no caso da sífilis em gestantes.

Um estudo descreveu a tendência temporal de novos casos de cardiopatia congênita em recém-nascidos em Jinhua de 2019 a 2020 e buscou um modelo apropriado para prever essa tendência. O ARIMA apresentou vantagens relativas em relação ao método aditivo de Holt-Winters ES, embora sua precisão não tenha sido ideal[20]. Comparando com o desempenho que obtivemos com o SARIMA e o CES, que seriam modelos equivalentes aos testado em Jinhua, repetiu-se esse padrão da superioridade do ARIMA e seus derivados. Percebe-se, portanto, certa constância em termos de desempenhos dos modelos.

Foram utilizados os modelos de suavização por médias móveis simples e da equação das tendências temporais para prever os casos futuros de sífilis congênita até 2025 no estado do Rio Grande do Norte. As previsões apontam relativa estabilidade dos casos[21]. Nenhum dos modelos previu estabilidade nas notificações para os próximos três anos. Seja um aumento ou uma queda, todos previram mudança de tendência. Portanto, reforça-se que a disseminação da sífilis entre gestantes tem características diferentes pelos estados do país.

A suavização exponencial complexa (modelo Holt-Winters) foi utilizada para prever os casos de sífilis na gestação em Porto Alegre. A prevalência no parto foi subestimada pelo modelo em 2021 e 2022 em cerca de três pontos percentuais[22]. Esse resultado diverge do encontrado em nosso estudo, tendo em vista que a suavização exponencial superestimou os casos no último dos três anos utilizados como validação. Essa diferença pode se dar devido a características específicas da sífilis em Porto Alegre ou por serem recortes temporais distintos.

Uma previsão utilizando o modelo matemático de suavização exponencial indicou aumento das notificações de sífilis congênita no Brasil, podendo chegar a 29.073 casos notificados em 2025[23]. Nosso estudo previu 227.120 casos de sífilis na gestação para todo o Brasil em 2025 com o modelo de menor erro (LSTM). Esse resultado indica que a transmissão vertical ocorreria em 12,8% das vezes.

Na Região Sudeste do Brasil no período, a série temporal de casos de sífilis gestacional revelou tendência crescente significativa ($R2 = 0,98$; $p < 0,001$), saindo de 12.892 casos em 2014 para 39.876 em 2023. O estado de São Paulo concentrou a maior aceleração absoluta[24]. Esse achado vai de encontro à tendência crescente também observada na série temporal do Brasil inteiro e às previsões de aumento dos casos nos próximos três anos feitas pelos modelos LSTM e CES.

CONCLUSÃO

Os modelos preditivos são usados e validados em todo o mundo na previsão de casos de doenças infecciosas. Nesse contexto, os resultados deste estudo corroboram a aplicabilidade dessa técnica, desde que observem o horizonte preditivo de forma crítica. Para a sífilis na gestação, as previsões dos modelos entraram em desacordo, o que reforça as limitação das estimativas futuras. Entretanto, pode-se tomar como norte as previsões do LSTM, que obtive MAPE < 10%, o que é considerado alta precisão. Segundo ele, as notificações diminuirão até ficar próximo aos 5.000 casos no início de 2026, quando voltarão a subir até próximo aos 7.000 casos em junho de 2027.

AGRADECIMENTOS

O primeiro autor recebeu uma bolsa de Iniciação Científica por essa pesquisa da Universidade Federal da Fronteira Sul (UFFS).

REFERÊNCIAS

1. Cooper JM, Sánchez PJ. Congenital syphilis. Semin Perinatol. 2018 Apr; 42(3): 176-84. doi: 10.1053/j.semperi.2018.02.005.

2. Finelli L, Berman SM, Koumans EH, Levine WC. Congenital syphilis. Bull World Health Organ. 1998; 76(Suppl 2): 126-8.

3. Wendel GD Jr, Sheffield JS, Hollier LM, Hill JB, Ramsey PS, Sánchez PJ. Treatment of syphilis in pregnancy and prevention of congenital syphilis. Clin Infect Dis. 2002 Oct; 35(Suppl 2): S200-9.

4. 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 Aug; 10: 966813. doi: 10.3389/fpubh.2022.966813.

5. Biggerstaff M, Johansson M, Alper D, Brooks LC, Chakraborty P, Farrow DC, et al. Results from the second year of a collaborative effort to forecast influenza seasons in the United States. Epidemics. 2018 Sep; 24: 24-33. doi: 10.1016/j.epidem.2018.02.003.

6. Lutz CS, Huynh MP, Schroeder M, Anyatonwu S, Dahlgren FS, Danyluk G, et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. BMC Public Health. 2019 Dec; 19(1):1659. doi: 10.1186/s12889-019-7966-8.

7. McKinney W, Seabold S, Perktold J. Time series analysis in Python with statsmodels. In: Proceedings of the 10th Python in Science Conference (SciPy 2011); 2011 Jun 28-Jul 3; Austin, Texas. Austin: SciPy; 2011. p. 57-61. doi: 10.25080/Majora-ebaa42b7-012.

8. Wang ZD, Li J, Chen Y, Zhang X, Liu Y, Huang H. Analysis and forecasting of syphilis trends in mainland China based on hybrid time series models. Epidemiol Infection. 2024 May; 152: e93. doi: 10.1017/S0950268824000694.

9. Weng RX, Fu HL, Zhang CL, Ye JB, Hong FC, Chen YM. Time series analysis and forecasting of Chlamydia trachomatis incidence using surveillance data from 2008 to 2019 in Shenzhen, China. Epidemiol Infect. 2020 Mar; 148: e76. doi: 10.1017/S0950268820000680.

10. Song Y, Choi H. Seasonality of syphilis in males through the 2011 to 2019 mandatory surveillance period: A cross-sectional study in South Korea. Medicine (Baltimore). 2023 Dec;102(50): e36723. doi: 10.1097/MD.0000000000036723.

11. Orssatto AJ, Rizzi CB, Rizzi RL, Brun AL. Aplicação de séries temporais na estimação do número de casos de dengue em Cascavel-PR. Rev Bras Comput Aplic. 2022;14(3): 37-50.

12. Campanhol AMC. Explicabilidade de modelo Long Short-Term Memory (LSTM) para previsão de dengue em Porto Alegre [trabalho de conclusão de curso]. Santa Maria (RS): Universidade Federal de Santa Maria; 2024.

13. Silva WS, et al. Seleção de um método de previsão de demanda oriundo da aquicultura na Bacia do Rio doce em Minas Gerais (MG) e Espírito Santo (ES). Revista Produção Online, 2025; 25(3): 5336. doi: 10.14488/1676-1901.v25i3.5336.

14. Alves CC, Hoepers E, Corazza EJ, Santos GJ, Cristofolini, Cruz AC. Aplicação de métodos estatísticos com suavização exponencial dupla e tripla para previsão de demanda na gestão de estoques. Rev Prod Online. 2019; 19(3): 1001-26. doi: 10.14488/1676-1901.v19i3.3539.

15. Bai Z, Han J, An J, Wang H, Du X, Yang Z, et al. The global, regional, and national patterns of change in the burden of congenital birth defects, 1990–2021: an analysis of the global burden of disease study 2021 and forecast to 2040. eClinicalMedicine. 2024 Oct; 77:102873. doi: 10.1016/j.eclinm.2024.102873.

16. Ibáñez-Cervantes G, León-Garcia G, De-León CV, Castro-Escarpulli G, Bandala C, Sosa-Hernández O, et al. Epidemiological behavior and current forecast of syphilis in Mexico: increase in male population. Public Health. 2020 Aug; 185: 386-93. doi: 10.1016/j.puhe.2020.05.057.

17. Silva GVS. Deep learning com modelos baseados em LSTM para previsão de casos de sífilis adquirida no Brasil [trabalho de conclusão de curso]. Natal (RN): Universidade Federal do Rio Grande do Norte; 2022.

18. Tang X, Chen W, Tang SQ, Zhao PZ, Ling L, Wang C. The evaluation of preventive and control measures on congenital syphilis in Guangdong Province, China: a time series modeling study. Infection. 2022 Oct; 50(5): 1179-90. doi: 10.1007/s15010-022-01791-1.

19. Aleixo RP. Predição de surtos de dengue e diagnóstico de sífilis congênita utilizando aprendizado de máquina [dissertação]. São Paulo (SP): Universidade de São Paulo; 2022.

20. Xu W, Shao Z, Lou H, Qi J, Zhu J, Li D, et al. Prediction of congenital heart disease for newborns: comparative analysis of Holt-Winters exponential smoothing and autoregressive integrated moving average models. BMC Med Res Methodol. 2022 Oct;22(1):257.

21. Raimundo DML, Silva JP, Oliveira RS, Pereira LF, Souza MA. Sífilis congênita: análise de tendência temporal e projeção de casos. Acta Paul Enferm. 2025; 38: eAPE0003054. doi: 10.37689/acta-ape/2025AO003054.

22. Echegaray F, Silveira MF, Costa JC, Basso LA, Nielsen-Saines K. Repercussions of the COVID-19 pandemic on maternal and congenital syphilis in South Brazil: a time series analysis 2010-2022. BMC Infect Dis. 2025 Apr; 25(1):528. doi: 10.1186/s12879-025-10901-x.

23. Carneiro DO, Jantsch LB, Oliveira LC, Pimenta RM. Utilização do modelo de suavização exponencial para avaliação temporal da situação epidemiológica da sífilis congênita no Brasil. Dialog Ciênc. 2023; 3(1): 19-34. doi: 10.7447/1678-0493.2023v3n1p19-34.

24. Borges OA, Lusvarghi EF, Pires VS, Silva AM, Libório LG, Rocha S. Perfil epidemiológico e tendência temporal da sífilis gestacional e congênita na Região Sudeste do Brasil: uma análise de 2014 a 2023. J Health Biol Sci. 2026; 14(1): e6367. doi: 10.12662/2317-3076jhbs.v14i1.6367.pe6367.2026.

Como citar este artigo/How to cite this article: Maia MCS, Raymundo AMF, Araújo DC, Schlager KA, Angelin KAraújo LCS, et al. Estimativas sobre sífilis gestacional no Brasil com modelos matemáticos e de inteligência artificial. J Health Biol Sci. 2026; 14(1): e6617.