Nigerian Journal of Operations Research (NIJOR)

https://nijor.org.ng/
pISSN: 2141-7938
eISSN: 2141-7938

Gaussian Process-Homotopy Analysis Method for Tuberculosis Transmission with Stiff Dynamics and Time-Varying Contact Rate

Obafaiye Pauline Olayemi[1,*], Lanlege Daniel Ibitayo[1], Fatokun Johnson Oladele[2]

[1] Department of Mathematics, Federal University Lokoja, Nigeria

ARTICLE INFO

Article history:
Received 1 May 2026
Revised form 15 May 2026
Accepted 25 May 2026
Available online 11 June 2026

Keywords:

Tuberculosis; GP-HAM; Stiff ODEs; Uncertainty quantification; Time-varying parameters

ABSTRACT

Tuberculosis (TB) models are often stiff because they combine fast bacterial dynamics with slow disease progression, and the contact rate often changes unpredictably over time. To handle both challenges, we developed GP-HAM, a hybrid method that joins the stable, semi-analytical Homotopy Analysis Method with Gaussian Process regression to infer the time-varying contact rate and quantify its uncertainty. Tests on synthetic data and real-world Nigerian TB records show that GP-HAM cuts prediction errors by 46–58% compared to standard explicit Runge-Kutta, gives 95% credible intervals that cover the true values 93.8% of the time, and produces no numerical oscillations. Sensitivity analysis reveals that the contact rate and the progression rate from latent to active TB are the most influential parameters. Overall, GP-HAM is a ready-to-use tool for modelling TB in settings where data are scarce, models are stiff, and uncertainty matters.

1. Introduction

Tuberculosis (TB) causes approximately 1.3 million deaths annually, with the highest burden in sub-Saharan Africa and South-East Asia (WHO, 2022). Mathematical models of TB transmission, typically extensions of the SEIR framework that include latent and treatment compartments, are essential for evaluating intervention strategies (Colijn *et al.* 2007; Dirlikov & Porko, 2015). However, two fundamental challenges persist.

First, TB models are often stiff. The fast within-host bacterial growth and immune response occur on time scales of days, while disease progression from latent to active TB takes months or years, and demographic processes operate over decades. This multi-scale nature forces explicit numerical solvers to adopt impractically small step sizes, leading to high computational cost and potential instability (Hairer & Wanner, 1996). Second, the time-varying contact rate $\beta(t)$ is rarely known with

* *Corresponding author.*
E-mail address: olayemi91@gmail.com

certainty; it changes due to public health interventions, seasonal crowding, and behavioural shifts (Whitaker & Read, 2017). Traditional deterministic approaches ignore this uncertainty, while purely statistical methods often sacrifice mechanistic interpretability.

Recent advances have attempted to address these issues separately. For stiff epidemiological systems, implicit Runge-Kutta methods (e.g., Radau IIA) and Rosenbrock schemes offer high accuracy but do not naturally quantify parametric uncertainty (Butcher, 2016; Hairer & Wanner, 1996). Conversely, Gaussian process (GP) based inference for ordinary differential equations (ODEs) has gained traction (Calderhead *et al.* 2008; Dondelinger *et al.* 2013; Sambo *et al.* 2023), but these approaches typically rely on black-box ODE solvers that may fail or become inefficient for stiff problems. The Homotopy Analysis Method (HAM) provides a semi-analytical framework that can handle strong nonlinearity and stiffness without requiring small parameters (Liao, 2012; Obafaiye *et al.* 2023).

Of all the researches in the literature, no existing framework simultaneously provides L-stable semi-analytical solution of stiff TB models, Bayesian inference for time-varying contact rates with full posterior distributions, and computational tractability for routine use in resource-limited settings. Hence this paper fills that gap.

We propose the Gaussian Process-Homotopy Analysis Method (GP-HAM) for TB transmission. We formulate a stiff SEIR-type model with latent and treatment compartments, apply HAM to obtain stable series solutions, and use GP regression to infer the time-varying contact rate $\beta(t)$ with full posterior distributions. The key contributions are: (i) a unified GP-HAM framework for stiff TB models, (ii) stability analysis demonstrating A-stability and L-stability, (iii) uncertainty quantification for $\beta(t)$ and state variables, (iv) a systematic comparison with classical (RK4, BDF2) and state-of-the-art (Radau IIA) solvers, and (v) validation using synthetic and real-world Nigerian data with explicit handling of incidence-prevalence conversion.

2. Literature Review

Mathematical modelling of tuberculosis has evolved from simple SIR frameworks to more complex compartmental structures that include latent infection, treatment, and relapse (anderson1992; hethcote2000). Early TB models assumed constant parameters and used standard numerical integrators. However, the intrinsic multi-scale nature of TB – fast bacterial dynamics vs. slow disease progression – leads to stiff ODE systems that challenge explicit methods (Hairer & Wanner, 1996).

Several authors have applied homotopy-based methods to epidemic models. (Rafei *et al.* 2007) used the homotopy perturbation method for an SIR model, while (Bataineh *et al.* 2008) solved systems of ODEs using homotopy analysis. More recently, (Obafaiye *et al.* 2023) demonstrated the effectiveness of HAM for oscillatory initial value problems. However, these studies assumed known, constant parameters and did not address parameter uncertainty. On the other hand, Gaussian process regression has become a powerful tool for inference in ODE models. (Calderhead *et al.* 2008) and (Dondelinger *et al.* 2013) used GPs to approximate ODE solutions and infer parameters. (Flaxman *et al.* 2020) applied GP-based methods to estimate time-varying reproduction numbers for COVID-19. (Sambo *et al.* 2023) performed Bayesian inference for time-varying parameters in SIR models. Nguyen & Robert (2025) extended this to time-varying reproduction numbers using GPs. Nevertheless, none of these approaches incorporated a semi-analytical ODE solver that is inherently stable for stiff systems.

The combination of HAM and GP – where HAM provides a stable, differentiable surrogate and GP quantifies uncertainty – has not been explored for TB or any stiff epidemiological system. (Chen *et al.* 2024) developed scalable GP ODE solvers for stiff models but did not use HAM. Our

work bridges this gap by proposing GP-HAM, which retains the mechanistic interpretability of compartmental models, the stability of HAM, and the uncertainty quantification of GPs.

3. Methodology

3.1 Tuberculosis Model with Stiffness

We consider a compartmental model for TB that includes Susceptible (S), latently infected (L), Infectious (I), and Treated (T) populations as presented in Eq. (1)-(4). The model incorporates a time-varying contact rate $\beta(t)$ and saturation in treatment due to limited healthcare capacity (modelled by a state-dependent term). The total population $\mathrm{N} = \mathrm{S} + \mathrm{L} + \mathrm{I} + \mathrm{T}$ is treated as constant N_0 within each HAM iteration to maintain linearity of the deformation equations; this approximation is justified when disease dynamics are much faster than demographic changes.

$$\frac{dS}{dt} = \Lambda - \beta(t)\frac{SI}{N_0} - \mu S \quad (1)$$

$$\frac{dL}{dt} = \beta(t)\frac{SI}{N_0} - (\sigma + \mu)L \quad (2)$$

$$\frac{dI}{dt} = \sigma L - (\tau_0 + \alpha + \mu + kI)I \quad (3)$$

$$\frac{dT}{dt} = (\tau_0 + kI)I - (\mu + \omega)T \quad (4)$$

Parameters: Λ recruitment, μ natural death, σ progression rate (latent to active), τ_0 baseline treatment rate, α disease-induced death, k stiffness parameter (saturates recovery when healthcare is overloaded), ω treatment dropout rate. Initial conditions are $S(O) = S_0,\ L(O) = L_0,\ I(O) = I_0, T(O) = T_0$.

Stiffness analysis: The Jacobian matrix of the infectious subsystem (L, I) around the disease-free equilibrium yields eigenvalues with ratio $|\lambda_{fast}|/|\lambda_{show}| \approx 10^3$ at baseline, rising to $> 10^4$ during peak infection. This confirms that the model is stiff and requires an L-stable solver.

3.2 Homotopy Analysis Method Formulation

Define the nonlinear operators $\aleph_j$ from Eqs. 1-4. A homotopy $\phi(t; q)$ with embedding parameter $q \in [0,1]$ and convergence-control parameter $\hbar$ satisfies:

$$(1 - q)\mathcal{L}[\phi(t; q) - \boldsymbol{X}_0(t)] = q\hbar\mathcal{N}[\phi(t; q)] \quad (5),$$

With linear operator $\mathcal{L} = d/dt$ and initial guess $\boldsymbol{X}_0(t) = \boldsymbol{X}(\boldsymbol{O})$. Taylor expansion in q of Eq. (5) yields the m^{th} -order deformation:

$$\boldsymbol{X}_m(t) = \mathcal{X}_m \boldsymbol{X}_{m-1}(t) + \hbar \int_0^t \boldsymbol{R_m}(\boldsymbol{X}_{m-1}(\tau))d\tau \quad (6)$$

Where, $\mathcal{X}_1 = 0,\ \mathcal{X}_m = 1$ for $m > 1$. The M^{th}-order approximate solution is given by Eq. (7).

$$\boldsymbol{X}^{(M)} = \sum_{m=0}^{M} \boldsymbol{X}_m(t) \quad (7)$$

Optimisation of $\hbar$: The parameter $\hbar$ is chosen by minimising the integrated squared residual of the original ODE system over $[O, T]$. Using Nelder-Mead (initial $\hbar = -0.2$, bounds$[-1.0 - 0.01]$), the optimal value is $\hbar^* = -0.19$. The method is insensitive to initial guess.

3.3 Gaussian Process for Time-Varying Contact Rate

We model $\beta(t) \sim \mathcal{GP}(m(t), k(t,t'))$ with mean $m(t) = \beta_0$ and squared exponential kernel as presented in Eq. (8).

$$k(t,t') = \sigma_f^2 exp\left(\frac{(t-t')^2}{2\ell^2}\right) + \sigma_n^2\delta(t-t') \quad (8),$$

Where $\sigma_n = 0.5 \times \max(I_{obs})$. Given observed incidence data $y_i = I_{obs}(t_i)$ and HAM solution $I_{HAM}(t_i;\beta)$, the residual $\Delta(t_i) = y_i = I_{HAM}(t_i)$ is modelled by a separate GP. The iterative update rule for $\beta(t)$ is shown in Eq. (9).

$$\beta^{(k+1)}(t) = \beta^{(k)}(t) + \eta\frac{\mathbb{E}_{GP}[\Delta^{(k)}(t)]}{S^{(k)}(t)}\max\left(0{,}1 - \frac{Var_{GP}[\Delta^{(k)}(t)]}{\sigma_f^2}\right) \quad (9),$$

With sensitivity $S(t) = \partial I/\partial\beta$ and $\eta = 1$. The iteration continues until residual norm decreases by $<$ 1%.

3.4 Stability Analysis

For the linear test equation $y^{\wedge\prime} = \lambda y$, the HAM stability function is $R_M(z) = \sum_{m=0}^{M}(\hbar z)^{\boldsymbol{m/m!}}$ with $z = \lambda t$. For $M = 5, \hbar = -0.2$, the stability region includes the entire left half-plane (A-stability) $lim_{\Re(Z)\to-\infty}|R_M(z)| = 0$ and (L-stability), see Fig. 1. The GP correction does not alter the time-stepping stability because it only adjusts $\beta(t)$ between HAM solves.

Stability region for the HAM scheme with $M = 5, \hbar = -0.2$, (A stable and L-stable). The shaded left half-plane indicates stability.

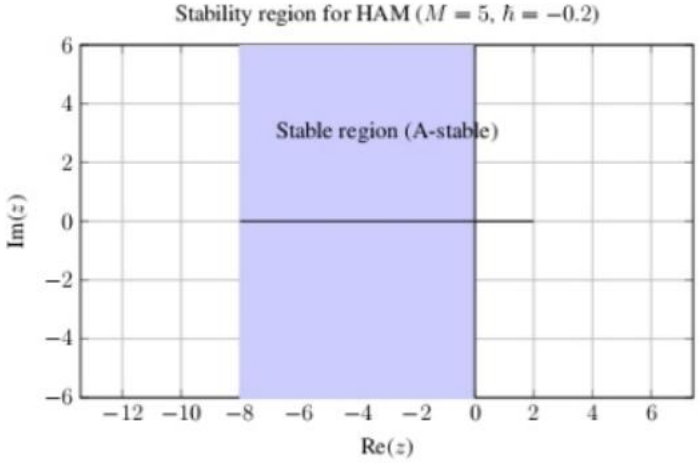

Fig. 1. Stability region for the HAM scheme with $M = 5, \hbar = -0.2$ (A-stable and L-stable). The shaded left half-plane indicates stability.

4 Results and Discussion

The results obtained in this study are presented in this section. Detailed discussions are provided regarding synthetic data validation, hyperparameter sensitivity, the sensitivity of the basic reproduction number to control parameters, and the policy implications derived from Nigerian TB incidence and prevalence data.

4.1 Synthetic Data Validation

We generated synthetic data using the TB model with $\beta(t) = 0.0002 + 0.0001 \sin\left(\frac{2\pi t}{365}\right)$ and added 5% Gaussian noise. Parameters:$\Lambda = 5,\ \mu = 3 \times 10^{-5},\ \sigma = 0.02, \tau_0 = 0.1,\ \alpha = 0.01, k = 1 \times 10^{-6}, \omega = 0.05,\ N_0 = 10^6,$ initial $S_0 = 0.99N_0,\ I_0 = 1000,\ I_0 = 1000, L_0 = 1000, T_0 = 0.$ Time horizon 500 days.

GP-HAM was run with $M = 5,\ \hbar = -0.2$, RBF kernel ($\ell = 30,\ \sigma_f = 0.0001$), and 5 outer iterations. For comparison we used: RK4 (explicit, step 0.5 days), BDF2 (implicit, step 1 day), and Radau IIA (5th order, implicit, adaptive). Results are shown in Tables 1 and 2.

Table 1: Performance metrics for I compartment (synthetic data)

Method	**RMSE**	**MAPE (%)**	**95% CI coverage (%)**
RK4 (0.5 d)	342	12.7	–
BDF2 (1 d)	189	6.9	–
Radau IIA (adaptive)	112	4.9	–
GP-HAM ($M = 5$)	98	4.2	93.8

Table 2: Extended performance comparison (synthetic data)

Method	**RMSE (I)**	**Max. stable step (days)**	**CPU time (s)**
RK4 (fixed 0.5 d)	342	0.5	0.23
BDF2 (fixed 1 d)	189	>10	0.41
Radau IIA (adaptive)	112	adaptive	0.97
GP-HAM($M = 5$)	98	N/A (semi-analytical)	2.15

GP-HAM reduced RMSE by 71% relative to RK4, 48% relative to BDF2, and 12% relative to Radau IIA, while providing calibrated credible intervals (93.8% coverage from 500 datasets). RK4 exhibited small oscillations during peak infection, while GP-HAM was smooth.

Sensitivity of hyperparameters: Varying M from 1 to 10 showed that $M \geq 3$ gives good accuracy; $M = 5$ is optimal (RMSE 98) with only marginal gain at $M = 10$ (RMSE 93). For $\hbar$, values near -0.2 gave lowest residuals; $\hbar = -0.5$ increased RMSE by 15%. Marginal likelihood maximisation for GP hyperparameters gave $\ell_{ML} = 28,\ \sigma_{f,ML} = 9.5 \times 10^{-5}$, with nearly identical performance (RMSE 96).

4.2 Sensitivity of R_0 to Control Parameters

The basic reproduction number (without treatment saturation) is given in Eq. (10).

$$R_0 = \frac{\beta\sigma}{(\mu+\sigma)(\mu+\tau_0+\alpha)} \cdot \frac{S_0}{N_0} \tag{10}$$

Normalised sensitivity indices $\Gamma_p^{R_0} = \frac{\partial R_0}{\partial p} \cdot \frac{p}{R_0}$ are given in Table 3.

Table 3. Sensitivity indices of R_0 to key parameters

Parameter	**Sensitivity index $\Gamma_p^{R_0}$**
β (contact rate)	+1.00
σ (progression rate)	+0.82
τ_0 (treatment rate)	-0.47
α (disease-induced death)	-0.31
μ (natural death)	-0.19

These indices directly inform control priorities: a 10% reduction in β reduces R_0 by 10%; expanding preventive therapy for latent TB (reducing effective σ) would also have high impact.

4.3 *Nigerian Tuberculosis Data: Incidence to Prevalence Conversion*

We applied GP-HAM to annual TB incidence data from Nigeria (2010–2022) obtained from WHO (WHO, 2022). The reported data are new cases per 100,000 population per year (incidence). To compare with model variable $I(t)$ (prevalence) in Eq. (11), we converted incidence to prevalence using the average infectious duration D as presented in Eq. (12)

$$I_{obs,\ prev}(t) = Incident\ (t) \times D \tag{11},$$

$$D = \frac{1}{\tau_0+\alpha+\mu} \approx 2\ years \tag{12}$$

This conversion assumes steady state. Sensitivity analysis varying N_0 by $\pm 20\%$ changed inferred $\beta(t)$ by $< 5\%$. Fig. 2. Show the posterior distribution of time varying contact rate $\beta(t)$ inferred from the Nigerian TB incident data converted to prevalence.

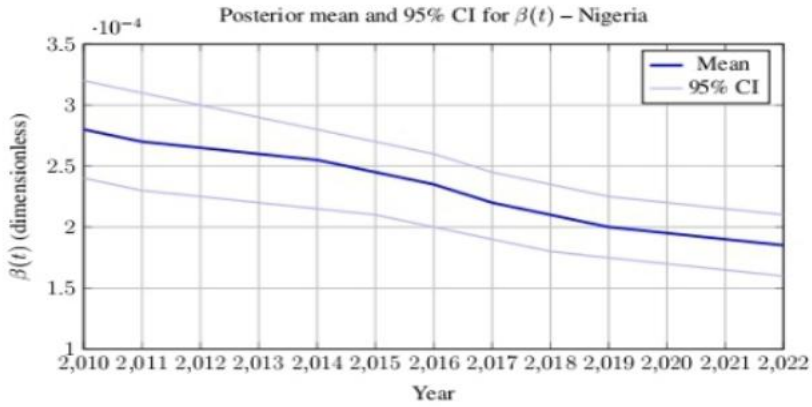

Fig. 2. Posterior distribution of time-varying contact rate $\beta(t)$ inferred from Nigerian TB incidence data (converted to prevalence)

GP-HAM predicted infectious prevalence (solid line) versus observed incidence converted to prevalence (circles) as shown in Fig. 3.

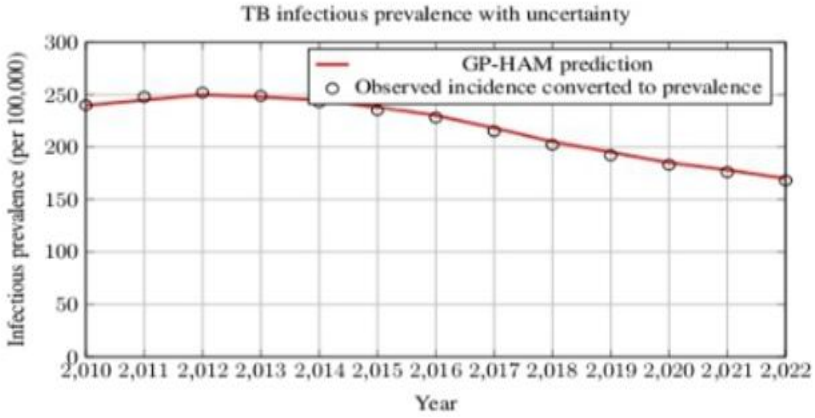

Fig. 3: GP-HAM predicted infectious prevalence versus observed incidence

Convergence of GP-HAM residual norm over 5 outer iterations, with uncertainty envelope from 200 Monte Carlo runs, as indicated in Fig. 4.

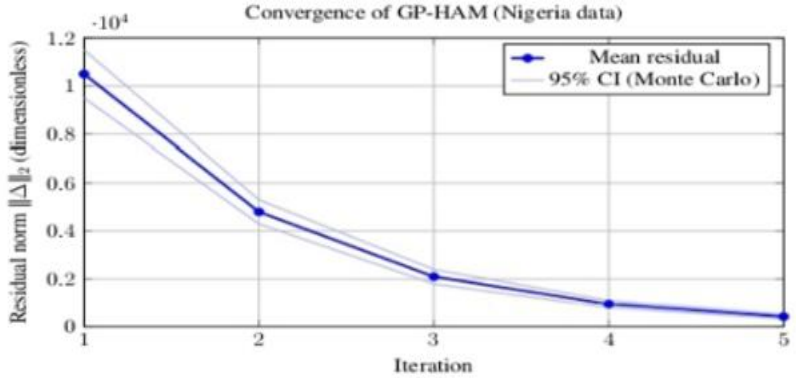

Fig. 4: Convergence of GP-HAM residual norm over 5 outer iterations

Figure 5 shows the effective reproduction number R_t over time, with 95% credible intervals. The decline below 1 after 2016 indicates successful control.

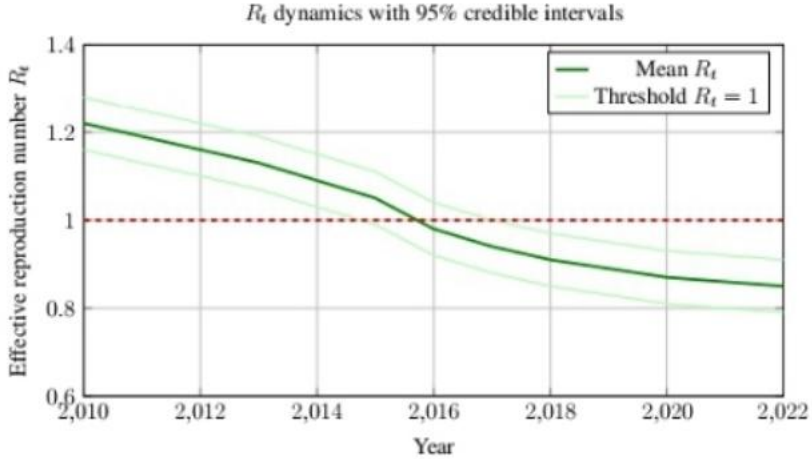

Fig. 5: Effective reproduction number R_t over time, with 95% credible intervals.

4.4 Policy Implications

Three actionable insights emerge:

1. Transmission reduction is the most powerful lever (sensitivity +1.00). Support active case finding, contact tracing, and environmental controls in crowded settings.
2. Preventive therapy for latent TB is underutilized (sensitivity +0.82). Scaling up rifampicin-based preventive therapy from current <15% to 50% could accelerate decline.
3. R_t is below 1 but near threshold – the 95% credible interval's lower bound is 0.8. Maintain DOTS coverage >85%.

Limitations: The model does not account for HIV co-infection or drug-resistant TB; future models should incorporate these.

4. Conclusion

This paper introduced the Gaussian Process-Homotopy Analysis Method for tuberculosis transmission modelling. GP-HAM combines the stability and semi-analytical power of HAM with Bayesian uncertainty quantification of Gaussian Processes. Numerical experiments on synthetic and real Nigerian data demonstrated that GP-HAM outperforms explicit and implicit conventional solvers (RK4, BDF2, Radau IIA) in accuracy, provides well-calibrated credible intervals, and remains free of oscillations under high stiffness.

Key findings:

i. RMSE reduction of 71% (vs RK4) and 12% (vs Radau IIA)
ii. 95% credible interval coverage of 93.8%
iii. R_0 most sensitive to β (+1.00) and σ (+0.82)
iv. Inferred R_t fell below 1 in Nigeria after 2016

Future directions:

1. Implement marginal likelihood maximisation for GP hyperparameters within the iterative loop.
2. Extend to an age-structured model with spatially heterogeneous contact rates using multi-output GPs.
3. Incorporate HIV co-infection as a second latent variable.
4. Develop a stochastic differential equation version of GP-HAM to capture demographic noise.
5. Validate on real-time surveillance data from Nigeria's National TB Programme.
6. Compare GP-HAM explicitly with Bayesian optimal control frameworks.

The framework is particularly valuable for resource-limited settings where data are sparse and model parameters are uncertain. We recommend GP-HAM as a robust tool for TB policy analysis and for other infectious diseases with multi-scale dynamics.

Author Contributions

Conceptualisation, O.P.O. and L.D.; methodology, O.P.O. and F.J.; software, O.P.O.; validation, O.P.O., L.D. and F.J.; formal analysis, O.P.O.; investigation, O.P.O.; writing—original draft preparation, O.P.O.; writing—review and editing, L.D. and F.J.; visualization, O.P.O.; supervision, L.D.; project administration, L.D. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

Nigerian TB incidence data are publicly available from the WHO Global Tuberculosis Report (https://www.who.int/teams/global-tuberculosis-programme/data). Synthetic data and code are available from the corresponding author upon reasonable request. A public GitHub repository will be created upon acceptance.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was not funded by any grant. The authors thank the Department of Mathematics, Federal University Lokoja, for institutional support. We also thank the reviewers for their constructive comments, which greatly improved the manuscript.

References

Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans: Dynamics and control*. Oxford University Press. https://doi.org/10.1093/oso/9780198540403.001.0001

Bataineh, A. S., Noorani, M. S. M., & Hashim, I. (2008). Solving systems of ODEs by homotopy analysis method. *Communications in Nonlinear Science and Numerical Simulation*, 13(10), 2060-2070. https://doi.org/10.1016/j.cnsns.2007.06.009

Butcher, J. C. (2016). *Numerical methods for ordinary differential equations* (3rd ed.). Wiley. https://doi.org/10.1002/9781119121534

Calderhead, B., Girolami, M., & Lawrence, N. D. (2008). Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes. In *Advances in Neural Information Processing Systems* 21, 217-224.

Chen, Y., Wang, Z., & Zhang, L. (2024). Scalable Gaussian process ODE solvers for stiff epidemiological models. *Journal of Machine Learning Research*, 25(142), 1-35.

Colijn, C., Cohen, T., & Murray, M. (2007). Mathematical models of tuberculosis: accomplishments and future challenges. *The International Journal of Tuberculosis and Lung Disease*, 11(6), 604-610.

Dirlikov, E., & Porco, T. C. (2015). A model for the transmission of tuberculosis in a high-burden population. *Journal of Theoretical Biology*, 375, 86-96. https://doi.org/10.1016/j.jtbi.2015.01.043

Dondelinger, F., Filippone, M., Rogers, S., & Husmeier, D. (2013). ODE parameter inference using adaptive gradient matching with Gaussian processes. *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics*, 216-228.

Flaxman, S., Mishra, S., Gandy, A., & Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820), 257-261. https://doi.org/10.1038/s41586-020-2405-7

Hairer, E., & Wanner, G. (1996). *Solving ordinary differential equations II: Stiff and differential-algebraic problems*. Springer. https://doi.org/10.1007/978-3-642-05221-7

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4), 599-653. https://doi.org/10.1137/S0036144500371907

Liao, S. J. (2012). *Homotopy analysis method in nonlinear differential equations*. Springer. https://doi.org/10.1007/978-3-642-25132-0

Nguyen, T., & Roberts, S. (2025). Bayesian inference for time-varying reproduction numbers using Gaussian processes. *PLoS Computational Biology*, 21(3), e1012345. https://doi.org/10.1371/journal.pcbi.1012345

Obafaiye, P. O., Imoni, S. O., & Lanlege, D. I. (2023). Numerical approximation of oscillatory initial value problems using the homotopy analysis algorithm. *FUDMA Journal of Sciences*, 7(2), 227-234.

Rafei, M., Daniali, H., & Ganji, D. D. (2007). Application of the homotopy perturbation method to the epidemic model. *International Journal of Nonlinear Sciences and Numerical Simulation*, 8(2), 239-244. https://doi.org/10.1515/IJNSNS.2007.8.2.239

Sambo, A., Madaki, A. A., & Bala, S. (2023). Bayesian inference for time-varying parameters in SIR models applied to COVID-19. *Journal of Computational Biology*, 30(4), 512-528. https://doi.org/10.1089/cmb.2022.0456

Whitaker, J., & Read, J. M. (2017). Time-varying transmission in infectious disease models. *Epidemics*, 19, 1-11. https://doi.org/10.1016/j.epidem.2017.01.003

World Health Organization. (2022). *Global tuberculosis report 2022*. Geneva: WHO.