A Bayesian Semi-Mechanistic Framework for Modelling Antimicrobial Resistance

Oraya Srimokla[1*], Sean Cavany[1], James Petrie[2], Mark Pritchard[1], Anna Nebykova[1], Gisela Aguilar[1], Christiane Dolecek[1], and Ben Cooper[1]

[1]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine
University of Oxford

[2]Independent Researcher

[*]Corresponding author: oraya.srimokla@ndm.ox.ac.uk

Abstract

Antimicrobial resistance is recognised as a global health threat but trends in the prevalence of resistance are often obscured by heterogeneity between data sources. *Acinetobacter baumannii* is a globally significant cause of antibiotic-resistant healthcare-associated infections. Data on *A. baumannii* resistance in Asia are available from 18 data sources, each using different methods to identify infections and resistance. Mechanistic mathematical models attempt to explain trends in resistance with biologically meaningful parameters but do not account for data heterogeneity. Purely statistical models can account for heterogeneity but have limited utility in predicting responses to interventions. Here we use a semi-mechanistic model assuming that trends in prevalence follow logistic growth dynamics, described by two parameters representing the coefficient for resistance acquisition and the loss of resistance. Using a Bayesian framework to account for uncertainty and missing data, we assume that these two parameters could vary by country, time and data source, and that neighbouring countries could exhibit spatial autocorrelation. We compared six models with different assumptions for these parameters, along with a stacked ensemble of these models. For four out of seven pathogen-antibiotic combinations considered, the stacked model performed best. For the other pathogen-antibiotic combinations, models that accounted for heterogeneity between data sources always performed better than models that did not. This framework improves existing approaches by offering a

more biologically plausible structure that can quantify the effects of covariates on resistance dynamics and effectively handle heterogeneous data sources. This methodology provides a potentially valuable foundation for better understanding AMR resistance dynamics and evaluating public health interventions.

Author Summary

Antimicrobial resistance occurs when disease-causing microorganisms become unaffected by the drugs designed to destroy them. This is a growing public health threat since these medicines become ineffective and can make infections difficult to treat. Understanding resistance trends across countries and over time is challenging because available surveillance data is sparse and comes from many different sources that can vary in terms of how the data is collected, the types of patients being treated, and the type of facility the data is being collected. We developed a way to model resistance trends by accounting for differences between data sources, something existing models do not consider, while maintaining the biological nature of how resistance growth is modelled. We applied our method, consisting of 6 different models, using data on a pathogen in Asia from 2000 to 2022. Our results show that accounting for variability between data sources improves our ability to predict resistance trends. Creating models that can more accurately estimate resistance patterns are essential for public health planning and evaluating health interventions. Our framework provides a potentially valuable way to better understand resistance trends and can be used as a foundation to model and assess different public health interventions.

1 Introduction

Antimicrobial resistance (AMR) threatens global health by compromising infection treatment and straining health systems worldwide [1, 2, 3, 4]. AMR occurs when microorganisms such as bacteria, viruses, parasites, and fungi evolve to gain resistance against antimicrobial medications, such as antibiotics [3]. When resistance develops, existing medicines become ineffective, making infections more difficult to treat. In response to the growing threat of AMR, international collaborations have been established to coordinate global action. The 2015 Global Action Plan (GAP) on AMR [5] led to the institution of AMR surveillance systems such as the WHO Global Antimicrobial Resistance and Use Surveillance System (GLASS) [6], the publication of national targets, and the creation of national action plans [7]. The 2024 UN General Assembly political declaration on AMR further strengthened these commitments [8]. Despite these efforts, significant challenges remain in understanding AMR trends to support surveillance, direct policy-setting, and strengthen national action plan adherence [9].

Mathematical models have become essential tools for understanding AMR dynamics and can be broadly classified as phenomenological or mechanistic. Phenomenological models typically describe relationships in data without making assumptions about the underlying biological or transmission mechanisms [10]. For example, Naghavi et al. used stacked ensemble modelling and spatiotemporal Gaussian process regression to estimate the proportion of resistant isolates for various pathogen-drug combinations from 1990 to 2021 [4]. In contrast, mechanistic models aim to explicitly describe mechanisms with model parameters having biological or behavioral interpretation [10]. For example, Kachalov et al. developed a compartmental model to simulate the spread of carbapenem-resistant *Klebsiella pneumoniae* (CRK) and extended-spectrum beta-lactamase-producing (ESBL) *Klebsiella pneumoniae* in community and hospital settings [11], while similar modelling approaches have been used to evaluate the effect of antimicrobial stewardship interventions [12, 13, 14]. While mechanistic models offer more biological interpretability and enable intervention evaluation, they often require more detailed data for model parameterisation and can result in greater model complexity [15, 16]. On the other hand, phenomenological models are simpler to fit and computationally efficient, but offer less model flexibility to explore underlying mechanisms of resistance [16]. Both modelling approaches face substantial data challenges when applied to AMR. Resistance data typically originate from various data sources, including clinical settings, national surveillance programs, and site-specific studies, each with distinct sampling protocols, reporting practices, and data processing standards [14, 17, 18, 3, 19]. Additional challenges include sparse data coverage, particularly in low and middle-income countries, bias in available data towards healthcare-seeking populations, and low blood culture utilisation rates [14, 20, 21, 22, 23]. While many AMR models have applied a frequentist approach, Bayesian modelling frameworks are increasingly applied due to their ability to implement regularisation by incorporating prior knowledge to adjust the estimates based on the defined priors and as a result, reduce variance and improve generalisation performance [24]. This additionally allows this framework to supplement incomplete data, constrain extreme prevalence estimates driven by data sparsity at the boundaries of the observation range, and generate Bayesian credible intervals that reflect the uncertainty from both the data and prior information [25, 26, 27, 28, 29, 30]. For example, Browne et al. applied a Bayesian spatiotemporal modelling framework to estimate the prevalence of resistance in *Salmonella enterica* serovars Typhi and Paratyphi A from 1990 to 2019 for 75 countries, using multiple sources of data and including locations with no data [25].

Semi-mechanistic models offer a hybrid approach to bridge the gap between intractable mechanistic models and less informative phenomenological models. These models use a biologically-plausible underlying structure to capture resistance dynamics, such as a logistic growth model, while the functional relationship between the model parameters and covariates are estimated phenomenologically. For example, by fitting dif-

ferent models to 887 pathogen-drug combinations, Emons et al. found that temporal trajectories of resistance support a logistic growth model with an initial increase in resistance followed by stabilisation at an intermediate level [31]. Combining a semi-mechanistic structure with a Bayesian modelling framework allows the model to maintain biological interpretability while accounting for incomplete data and constraining extreme estimates. This approach additionally provides a foundation to efficiently explore different model components that can account for country and data source-level variation, spatiotemporal effects, and covariates while quantifying uncertainty through Bayesian credible intervals [25, 32, 33]. To more realistically capture resistance patterns and intervention effects, time-varying parameters can be incorporated to allow model dynamics and resistance growth rates to adapt over time rather than remain static. As individual models may not be able to independently capture complex resistance dynamics, Bhatt et al. explored the use of an ensemble approach using stacked generalisation to combine models within a Gaussian process framework and nonlinearly maps covariates directly to disease prevalence [34, 35]. While methods presented by Bhatt et al. are effective for systems at equilibrium, this approach can bypass intrinsic disease dynamics, which may not be suitable for non-equilibrium systems such as AMR where resistance is commonly dynamic [31].

In this study, we develop a Bayesian semi-mechanistic modelling framework with time-varying parameters to estimate the prevalence of resistance for *Acinetobacter baumannii*, a critical nosocomial pathogen on the WHO priority pathogen list [36]. Building on previous studies [31, 37, 34], we extend a logistic growth model to describe resistance trends in terms of time-varying and biologically interpretable disease parameters, and nonlinearly map covariates to our model parameters. Our framework compares six model variants that account for different combinations of country and data source-level variation, spatial correlation effects, and covariates for seven pathogen-antibiotic combinations across Asia from 2000 to 2022. To our knowledge, this is the first application of a Bayesian semi-mechanistic modelling framework that simultaneously integrates time-varying parameters, country and source-level variation, spatial correlation effects, and covariate information for AMR prevalence estimation. We additionally apply a model stacking approach to generate ensemble predictions from the best combination of models [34]. This framework provides a methodological framework for estimating the prevalence of AMR that integrates biological interpretability, time-varying dynamics, and spatial correlation while accounting for data constraints.

2 Methods

2.1 Data

Resistance data for *Acinetobacter baumannii* was obtained from the Global Research on Antimicrobial Resistance 2024 global burden study [4] and contains around 1.3 million isolates for 113 countries globally from 1991 to 2022. This dataset consists of resistance data from 31 sources including surveillance programs, individual testing facilities, and systematic reviews across seven antibiotic categories. The antibiotic combinations included are aminoglycosides, antipseudomonal penicillins, beta lactam/beta-lactamase inhibitor combinations, carbapenems, fluoroquinolones, fourth-generation cephalosporins, and third-generation cephalosporins. For this analysis, we focus on countries in Asia from 2000 to 2022, which consists of around 300000 isolates across 18 data sources, 32 countries, and 7 pathogen-antibiotic combinations (see Section A in S1 Text for a summary of the data).

Covariate data were selected from on previous studies identifying possible associations with resistance in *Acinetobacter baumannii* [38, 39, 40, 41, 42]. These included time-dependent and country-level gridded mean temperature data from the Copernicus Climate Change Service [43], along with estimates of total human antibiotic consumption (defined daily doses per 1000 inhabitants per day (DDD per 1000)) and the fraction of out-of-pocket healthcare expenditures from the 2024 Global Burden of Resistance study [4, 44].

2.2 Model Structure

We model the prevalence of resistance (the proportion of resistant isolates), defined as the proportion of total isolates that are resistant to a specific antibiotic, using parameters from a two-compartment ordinary differential equation (ODE) model to describe the dynamics between drug-sensitive and drug-resistant strains (see Section B in S1 Text). This framework assumes that resistance dynamics are decoupled from the dynamics of sensitive strains. Under this assumption, the proportion of resistant isolates, P_t, follows logistic growth dynamics, similarly characterised by Emons et al. [31].

The resistance prevalence at time t, can be defined as:

$$P_{c,t} = \frac{1 - \frac{\gamma_c}{\beta_c}}{1 + \left(\frac{1-\frac{\gamma_c}{\beta_c}}{P_{0,c}} - 1\right) e^{-(\beta_c - \gamma_c)t}}, \beta_c \neq \gamma_c \tag{1}$$

$P_{0,c}$ is the initial resistance prevalence of country c, β_c is a coefficient describing the rate of resistance for country c, and γ_c is the rate at which resistant strains become sensitive for country c. We parameterise this

model in terms of β and γ, rather than the standard carrying capacity and growth rate as this parameterisation demonstrated more consistent model convergence. In this model, the growth rate can be defined as $\beta - \gamma$ and the carrying capacity can be defined as $(1 - \gamma/\beta)$ (see Section B in S1 Text).

We extend this approach to allow these parameters to vary over time in order to capture temporal changes in resistance dynamics. This is achieved by constructing piecewise continuous growth curves where parameters can change between consecutive time points (see Algorithm 1). At each new year t_i, the model updates the initial resistance prevalence to match the final prevalence from the previous year t_{i-1}, creating a continuous trajectory from the model (see Algorithm 1). The pseudocode incorporating the time-varying structure is shown in Algorithm 1.

For models without covariates, each country has country-specific but temporally constant values for β_c, γ_c, $P_{0,c}$ and produces a singular growth curve per country. For models incorporating covariates, parameters can vary by country and year (see Table 1). In general, we expect both β and γ could vary over time. We simplify parameter estimation by assuming γ_c is constant over time. We represent $\beta_{c,t}$ as a function of time- and country-specific covariates:

$$\beta_{c,t} = exp(log_{\beta_c} + \sum_{k=1}^{K} \theta_k \cdot X_{k,c,t}) \tag{2}$$

where log_{β_c} is the log-transformed β_c for the baseline coefficient value of describing the gain of resistance, $X_{k,c,t}$ are the K normalised covariates for country c at year t, and θ_k are the estimated covariate coefficients (see Algorithm 1 and Section C in S1 Text). This structure ensures that β remains positive while allowing time-varying dynamics through the addition of covariates.

We incorporate random effects at multiple levels to account for variability in resistance dynamics. Six model variants are explored with different combinations of country and data source-level variation, spatial correlation effects, and covariates (see Table 2).

Country-level random effects allow β, γ, and P_0 to vary by country c with hierarchical priors (see Table 1 and Section D in S1 Text). Data source-level random effects account for variation in the data sources that can result from differences in data collection protocols and sampling methods. These source-level effects are implemented through a multiplicative scaling factor:

$$P_{c,s,t} = \left(\frac{1 - \frac{\gamma_c}{\beta_c}}{1 + \left(\frac{1-\frac{\gamma_c}{\beta_c}}{P_{0,c}} - 1\right) e^{-(\beta_c - \gamma_c)t}}\right) * \eta_s \tag{3}$$

where η_s is the scaling factor for source s, the first term is the baseline level of resistance with country-level random effects at country c and time t (see Table 1 and Section E in S1 Text). Spatial effects are

Algorithm 1 Resistance Probability Calculation with Logistic Growth Model

Initialize vectors
Set initial conditions for parameters
for $i = 1$ to N :number Of Observations with example for one country **do**
    Set initial resistance value at first observation
    After the first observation, set the parameter values
    $\beta_{\text{current}} \leftarrow exp(log\beta_c[\text{country}[i]] + X[i,] * \theta)$
    $\gamma_{\text{current}} \leftarrow \gamma_c[\text{country}[i]]$
    if repeated year values are observed **then**
        prob$[i] \leftarrow$ lastI
        Handle repeated years data logging
    else
        Calculate probability using previous value:
        $\text{prob}[i] \leftarrow \dfrac{\frac{\beta_{\text{current}} - \gamma_{\text{current}}}{\beta_{\text{current}}}}{1 + \left(\frac{\beta_{\text{current}} - \gamma_{\text{current}}}{\beta_{\text{current}}} - 1\right) \cdot e^{-(\beta_{\text{current}} - \gamma_{\text{current}}) \cdot (t[i] - \text{lastYear})}}$
    end if
    Update tracking variables:
    lastI $\leftarrow$ prob$[i]$
    lastCountry $\leftarrow$ country$[i]$
    lastYear $\leftarrow t[i]$

    Apply likelihood with bounds:
    $p_{\text{bounded}} \leftarrow \max(\min(\text{prob}[i], 0.99999), 0.00001)$
    $y[i] \sim \text{Binomial}(\text{cases}[i], p_{\text{bounded}})$
end for

incorporated through Gaussian process regression to capture spatial autocorrelation between neighbouring countries [45, 46]. To establish relationships between geographically close sites, we use a kernel function to define the covariance between countries according to the scaled distance (see Section F in S1 Text):

$$K_{i,j} = \zeta_\epsilon^2 * exp(-\rho^2 * D_{i,j}^2) \tag{4}$$

where $D_{i,j}$ is the scaled Euclidean distance from geographical midpoints between countries i and j, ζ_ϵ^2 is the maximum covariance between sites for model parameter ϵ, and ρ^2 is the rate covariance declines with distance. This covariance function is used to estimate spatially-correlated random effects for each country and is used to define $beta_c$, $gamma_c$, $P_{0,c}$ [45, 46]. For P_0, the spatial term is added to the population-level mean value in logit space:

$$P_{0,c} = \text{logit}^{-1}(P_{0_\text{logit}} + k_{P_0,c}) \tag{5}$$

For β and γ, the spatial effects are added similarly added to their population-level mean value in log space to ensure positivity:

where $k_{\theta,c}$ represents the spatial random effect for parameter θ for country c (see Table 1).

$$\beta_c = exp(log_\beta + k_{\beta,c}) \tag{6}$$

$$\gamma_c = exp(log_\gamma + k_{\gamma,c}) \tag{7}$$

Observations of resistant isolates are assumed to follow a binomial distribution:

$$y_{c,s,t} \sim \text{Binomial}(\text{N}_{c,s,t}, P_{c,t}) \tag{8}$$

where N_t is the total number of isolates tested for a given country-source combination at time t, and $P_{c,t}$ is the proportion of resistant isolates corresponding to the value of the logistic growth function at time t and country c. $P_{c,t}$ becomes $P_{c,s,t}$ when source-level effects are included.

Ensemble modelling: To combine predictions across model variants, we apply Bayesian model stacking using the loo (v2.8.0) package in R [47]. Model stacking generates ensemble predictions by finding the combination of models that maximises the leave-one-out (LOO) predictive density and the expected log pointwise predictive density (ELPD) of the ensemble predictions [34, 35, 48]. LOO predictive densities are obtained using the leave-one-out cross-validation (LOO-CV), a method for estimating pointwise out-of-sample prediction accuracy from a fitted Bayesian model [47]. The best linear combination of model weights for each pathogen-antibiotic combination was determined using the loo_model_weights() function and used to determine the weights for each of models when combining predictive distributions [47, 35].

Countries with Missing Data: The Gaussian process framework enables spatial imputation by using the estimated maximum covariance between sites ζ^2, and the rate covariance declines with distance ρ^2 to estimate the covariance function for countries with unobserved resistance data. New values are predicted following a multivariate normal predictive distribution:

$$p(\mathbf{f}_*|\mathbf{f}) \sim \mathcal{N}(\mathbf{K}_*^T\mathbf{K}^{-1}\mathbf{f}, \mathbf{K}_{**} - \mathbf{K}_*^T\mathbf{K}^{-1}\mathbf{K}_*) \tag{9}$$

where f is the observed spatial effects for countries with resistance data, f_* is the spatial effects for countries with no resistance data, K is the covariance matrix of old countries with resistance data, K_* is the covariance matrix with a combination of countries with and without resistance data, and K_{**} is the covariance matrix of countries with no resistance data [49, 50].

Table 1. Definitions of model parameters for six model variants. Parameters are defined based on logistic growth dynamics where $P_{c,t}$ represents the proportion of resistant isolates at time t and country c. Here Model C: Model with country-level random effects, C+X: Country-level random effects model and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with source-level effects, SP+S+X: Spatial model with source-level effects and covariates.

Model	**Rate of Resistance Coefficient (β)**	**Loss of Resistance Rate (γ)**	**Initial Prevalence (P_0)**	**Source Effect (η_s)**
Model C	β_c (Country-specific)	γ_c (Country-specific)	$P_{0,c}$ (Country-specific)	None ($\eta_s = 1$)
Model C + X	$\beta_{c,t} = \exp(\log \beta_c + \sum \theta_k \cdot X_{k,c,t})$ (country-specific+time-varying covariates)	γ_c	$P_{0,c}$	None ($\eta_s = 1$)
Model C + S	β_c	γ_c	$P_{0,c}$	$P_{c,s,t} = P_{c,t} \times \eta_s$
Model SP	$\beta_c = \exp(\log \beta + k_{\beta,c})$ (spatial effect k_β)	$\gamma_c = \exp(\log \gamma + k_{\gamma,c})$ (spatial effect k_γ)	$P_{0,c} = \text{logit}^{-1}(P_{0_logit} + k_{P_0,c})$ (spatial effect k_{P_0})	None ($\eta_s = 1$)
Model SP + S	$\beta_c = \exp(\log \beta + k_{\beta,c})$	$\gamma_c = \exp(\log \gamma + k_{\gamma,c})$	$P_{0,c} = \text{logit}^{-1}(P_{0_logit} + k_{P_0,c})$	$P_{c,s,t} = P_{c,t} \times \eta_s$
Model SP + S + X	$\beta_{c,t} = \exp(\log \beta + k_{\beta,c} + \sum \theta_k \cdot X_{k,c,t})$ (spatial effect k_β + time-varying covariates)	$\gamma_c = \exp(\log \gamma + k_{\gamma,c})$	$P_{0,c} = \text{logit}^{-1}(P_{0_logit} + k_{P_0,c})$	$P_{c,s,t} = P_{c,t} \times \eta_s$

2.3 Model Fitting and Evaluation

The modelling framework is implemented using RStan version 2.32.6, which implements Hamiltonian Monte Carlo (HMC) sampling to estimate the model parameters including β_c, γ_c, $P_{0,c}$, population-level mean values ($P_{0,logit}$, log_β, and log_γ), associated random effects terms (e.g. source s scaling factor η_s, maximum covariance between sites ζ_ϵ^2 for model parameter ϵ, and the rate covariance declines with distance ρ^2), and the estimated covariate coefficients θ_k for the Asia region from 2000 to 2022 for each of the 7 pathogen-antibiotic combinations.

Prior distributions for the models were selected to be weakly informative or informative when domain knowledge suggested biologically plausible parameter ranges (see Section C to F in S1 Text). We conducted prior and posterior predictive checks. Prior predictive checks generate simulated data according to the prior distributions to assess if the priors are appropriate and produce reasonable model outputs [51]. After fitting

Table 2. Model variants and the model components that make up each model. Here Model C: Model with country-level random effects, C+X: Country-level random effects model and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with source-level effects, SP+S+X: Spatial model with source-level effects and covariates.

	Model C	**Model C + X**	**Model C + S**	**Model SP**	**Model SP + S**	**Model SP + S + X**
Country-level Random Effects	✓	✓	✓			
Covariates		✓				✓
Data source-level Random Effects			✓		✓	✓
Spatial-level Random Effects				✓	✓	✓

the model to empirical data, posterior predictive checks were done by simulating new replicated data sets based on the fitted model and comparing them to the observed data [51].

For each of the models, we assessed convergence using standard MCMC diagnostics including traceplots and Rhat statistics (with Rhat < 1.01). We additionally examined the number of divergent transitions, maximum tree depth, and pairs plots to identify issues with convergence or model parameterisation.

Prior to fitting the model to empirical data, we fit the model to simulated data with known parameter values to verify parameter recoverability. Once the models were fitted, predictions were generated by extracting and sampling from the parameters' posterior distribution. Parameter estimates were then used to generate growth curves from 2000 to 2022 of each country for all six model variants.

To demonstrate the framework's ability to generate predictions for countries with missing data, we use the Gaussian process structure in our models with spatial effects (SP, SP+S, SP+S+X) and refit the models without five randomly selected whole-country held-out data. We then generated predictions for these held-out countries using the Gaussian process feature and compared these trajectories to the actual observed resistance data for that country.

The performance of six individual model variants for each pathogen-antibiotic combination are compared using the Leave-One-Out Cross-Validation Information Criterion (LOOIC) obtained from applying the LOO-CV method via the loo (v2.8.0) package in R. The LOOIC value is computed from the ELPD, with lower values indicating better expected out-of-sample predictive performance.

To evaluate predictive accuracy on held-out data, we fitted models to a training dataset and evaluated the predictions on a test set created by reserving the most recent four years of data for each country. Predictive accuracy for this assessment was quantified using the mean absolute error (MAE) between observed and predicted values. This additional evaluation assessed the models' ability to forecast future resistance trends.

3 Results

To ensure the reproducibility of the model results, the parameter posterior densities are extracted from fitting simulated data with the country-level random effects model with a β_c, γ_c, and $P_{0,c}$ parameterisation (see Section G in S1 Text). The resulting posterior densities show that the models are able to recover the true values of the parameters (see Section G in S1 Text). Additional, posterior predictive checks further confirmed the model was able to capture the mean and standard deviation of the observed data.

Following model fitting, model predictions are generated for all countries in Asia from 2000 to 2022. Individual model fits for five countries with available resistance data are shown in Fig 1. From this figure, we observed all six model variants captured similar temporal trajectories for each country. The models were also able to capture logistic growth trends as well as decay and stabilisation trends, as shown in Japan from 2000 to 2022. For models with covariates, we observe year-to-year fluctuations for model estimates as a result of the time-varying covariates. We also observe more narrow credible intervals for countries with more data and larger credible intervals for countries with fewer resistance isolate data. These trends can similarly be observed within a single country. For example, in the early 2000s where little resistance data was available in Viet Nam, larger credible intervals were observed and narrowed as more data became available.

From our model comparison, the model incorporating both country-level and source-level random effects (C+S) produced the lowest LOOIC value across 5 of 7 antibiotic categories, showing the best predictive performance (see Table 3). For the remaining two antibiotic categories (aminoglycosides and antipseudomonal penicillins), the model combining spatial and source-level effects with covariates (SP+S+X) were shown to have the best performance. From these results, the best performing models (C+S and SP+S+X) both incorporate source-level random effects and are able to account for data-source variation.

Model C with only country-level random effects consistently had the second highest LOOIC, while the base spatial model (SP) had the highest LOOIC values for 6 of 7 antibiotic categories. These results suggest that country-level effects and spatial effects alone are not able to capture heterogeneity in the resistance data. In contrast, spatial models incorporating source-level effects and/or covariates had relatively good predictive performance, ranking in the top three across all antibiotic combinations. The results additionally demonstrate that including covariate information with the spatial and source-level effects model (SP+S) improves model performance across all 7 combinations.

From our whole-country held-out data model predictions, Fig 2 shows the imputed estimates for Singapore and Viet Nam, two of the five countries that were held-out from fitting, and their neighbouring countries (see Section H in S1 Text). While the model estimates for these held-out countries have large credible intervals,

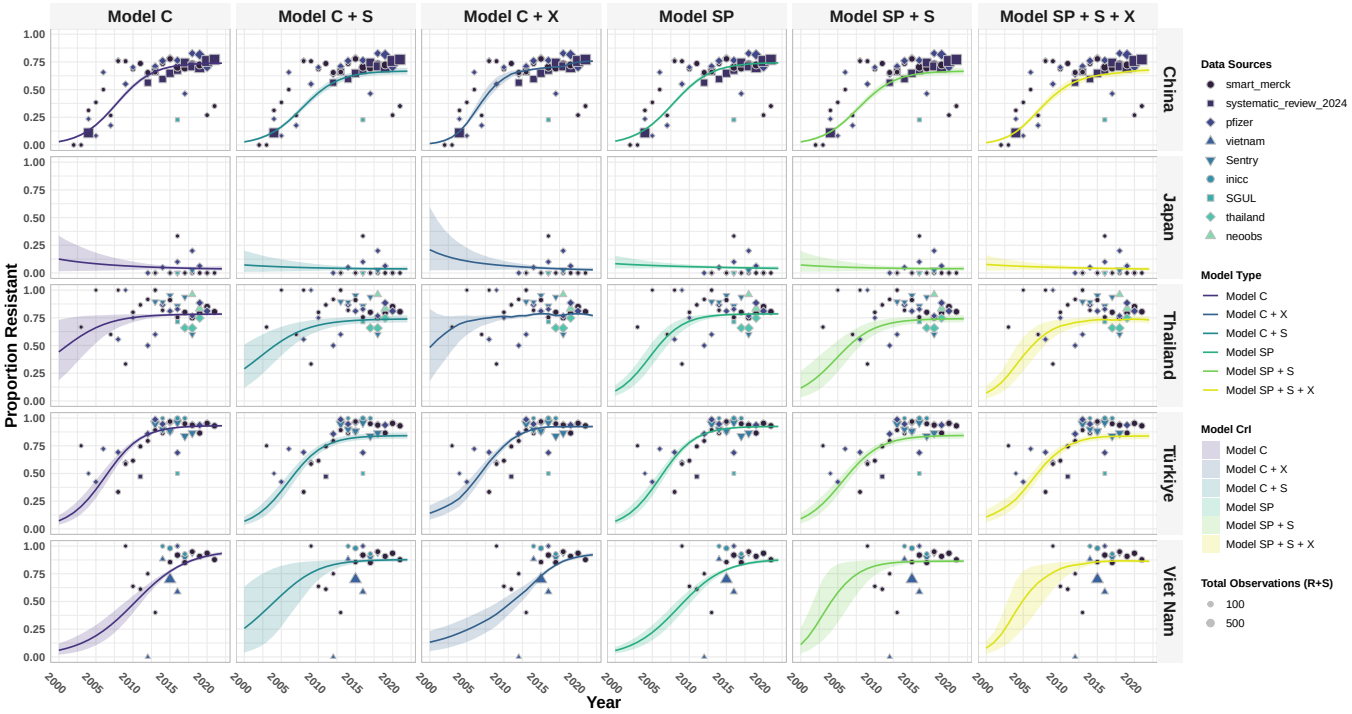


Fig. 1. Individual model fits using β γ, P_0 parameterisation of the logistic growth function for resistance of *Acinetobacter baumannii* and carbapenems model training for Asia. Shows model estimates for 5 countries each model variant as the column and each country as the row. Model C: Model with country-level random effects, C+X: Country-level random effects and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with Gaussian process regression and source-level effects, SP+S+X: Spatial model with Gaussian process regression, source-level effects, and covariates. Solid lines represent the mean posterior prediction and the shaded regions show the 95% credible intervals with the different colours corresponding to each model type. The points show the observed resistance data used for training the models.

Antibiotic Class/Subclass	**Model Comparison using LOOIC differences from best model**					
	Model C	**Model C+X**	**Model C+S**	**Model SP**	**Model SP+S**	**Model SP+S+X**
Aminoglycosides	431.3	430.4	38.3	466.9	4.9	**0**
Antipseudomonal penicillins	200.8	136.6	49.2	221.0	71.8	**0**
Beta lactam/beta-lactamase inhibitor combinations	196.0	210.0	**0**	230.0	55.4	47.3
Carbapenems	378.2	329.9	**0**	432.2	100.9	95.8

Fluoroquinolones	663.6	567.7	**0**	722.9	27.7	13.3
Fourth-generation cephalosporins	131.9	132.7	**0**	130.8	31.7	30.7
Third-generation cephalosporins	312.2	221.9	**0**	334.3	56.6	9.5

Table 3. Comparison of fitted models for each antibiotic class using the leave-one-out cross-validation information criterion (LOOIC). Relative differences in LOOIC calculated from the best model (shown in bold with 0). Lower values indicate better predictive performance. Model C: Model with country-level random effects, C+X: Country-level random effects and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with Gaussian process regression and source-level effects, SP+S+X: Spatial model with Gaussian process regression, source-level effects, and covariates.

the mean estimates are generally able to capture trends observed in the data and their neighbouring countries.

Model stacking is further explored to understand the best model combination to improve predictive accuracy. The model stacking weights are shown in Table 4.The model incorporating both country-level and source-level random effects (C+S) was assigned the highest weight for 5 of 7 antibiotic categories, indicating optimal prediction accuracy if the sampling is heavily weighted using this model. For the remaining two antibiotics categories, aminoglycosides and antipseudomonal penicillins, the highest weights were assigned to the spatial model with source-level effects (SP+S) and the country-level effects model with covariates, respectively. Model C with only country-level random effects, and/or the model with only spatial effects generally received lower weightings across different antibiotic categories. Finally, these weights are rounded to the nearest percent and used to determine the number of samples allocated to specific models to generate stacked model predictions. Examples of stacked model predictions for 9 countries are shown Fig 3.

Antibiotic Class/Subclass	**Model Stacking Weights**					
	Model C	**Model C+X**	**Model C+S**	**Model SP**	**Model SP+S**	**Model SP+S+X**
Aminoglycosides	0.16	0.12	0.04	0.10	**0.32**	0.26
Antipseudomonal penicillins	0.03	**0.34**	0.12	0.02	0.26	0.23

Beta lactam/beta-lactamase inhibitor combinations	0.08	$1.06*10^{-6}$	**0.63**	0.13	0.06	0.10
Carbapenems	$2.50*10^{-7}$	0.29	**0.48**	0.16	0.05	0.02
Fluoroquinolones	0.02	0.18	**0.44**	0.02	0.28	0.06
Fourth-generation cephalosporins	$4.26*10^{-6}$	0.06	**0.54**	0.23	$8.03*10^{-6}$	0.17
Third-generation cephalosporins	0.02	0.26	**0.37**	0.02	$8.27*10^{-7}$	0.33

Table 4. Comparison of fitted models weights for each antibiotic category using model stacking. The best model, having the highest weights, are shown in bold. Higher weights indicate better predictive performance. Model C: Model with country-level random effects, C+X: Country-level random effects and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with Gaussian process regression and source-level effects, SP+S+X: Spatial model with Gaussian process regression, source-level effects, and covariates

For models incorporating source-level random effects, the MAE was calculated for each country-source-time combination and for all other models, it was calculated for each country-time combination. The MAE values for all pathogen-antibiotic combinations are shown in Table 5. The stacked model had the lowest MAE value for 4 of 7 antibiotic categories, followed by the spatial model with source-level random effects and covariates (SP+S+X). Similar to the LOOIC model comparison, models only incorporating country-level random effects (C) or only spatial effects (SP), generally tend to have a slightly lower predictive accuracy. The spatial model (SP) was found to have the largest MAE values for 4 of 7 antibiotic categories. As this model does not account for any source-level effects or include any covariates, the lower performance could suggest that the estimates are more influenced by their neighbouring countries, leading to a higher error when those neighbours have different resistance dynamics.

Antibiotic Class/Subclass	**Mean Absolute Error**						
	Model C	**Model C+X**	**Model C+S**	**Model SP**	**Model SP+S**	**Model SP+S+X**	**Stacked Model**
Aminoglycosides	0.1684	0.1672	0.1525	0.1646	0.1513	**0.1511**	0.1532

Antipseudomonal penicillins	0.1284	0.1277	0.1222	0.1294	0.1257	0.1236	**0.1210**
Beta lactam/beta-lactamase inhibitor combinations	0.1861	0.1854	0.1872	0.1848	0.1890	**0.1847**	0.1951
Carbapenems	0.1406	0.1403	0.1362	0.1423	0.1392	0.1374	**0.1361**
Fluoroquinolones	0.1187	0.1188	0.1215	0.1250	0.1215	0.1203	**0.1177**
Fourth-generation cephalosporins	0.1288	0.1302	0.1231	0.1273	0.1219	0.1221	**0.1216**
Third-generation cephalosporins	0.1333	0.1340	**0.1259**	0.1429	0.1367	0.1370	0.1267

Table 5. Comparison of mean absolute error (MAE) calculated for four years of held-out data across countries, years, and sources (for source-specific model). The MAE was calculated for a held-out test set containing the most recent four years of data for each country. The best model for each antibiotic category is shown in bold. Lower MAE values indicate more accurate model predictions. Model C: Model with country-level random effects, C+X: Country-level random effects and covariates, C+S: Country-level random effects and source-level effects, SP: Spatial model with Gaussian process regression, SP+S: Spatial model with Gaussian process regression and source-level effects, SP+S+X: Spatial model with Gaussian process regression, source-level effects, and covariates.

The convergence of all six model variants were visually checked for all 7 pathogen-antibiotic combinations using trace plots and pairs plots (see Section I in S1 Text). These plots were generated for each pathogen-antibiotic combination and each individual model fit to ensure stability of model runs. All plots exhibit convergence of all four chains in the trace plot and do not show model divergences which would have been indicated through red points in the pairs plot.

4 Discussion

We developed a Bayesian semi-mechanistic modelling framework accounting for time-varying parameters, country and source-level variation, covariates, and spatial correlation effects to estimate the prevalence of AMR. Our approach compares 6 model variants applied to seven *Acinetobacter baumannii* -antibiotic combinations across Asia from 2000 to 2022 and explores the applications of model stacking to generate ensemble predictions. Through our model comparisons, models incorporating source-level random effects consistently outperformed other models, as shown by the lowest LOOIC values. We account for spatial correlation effects

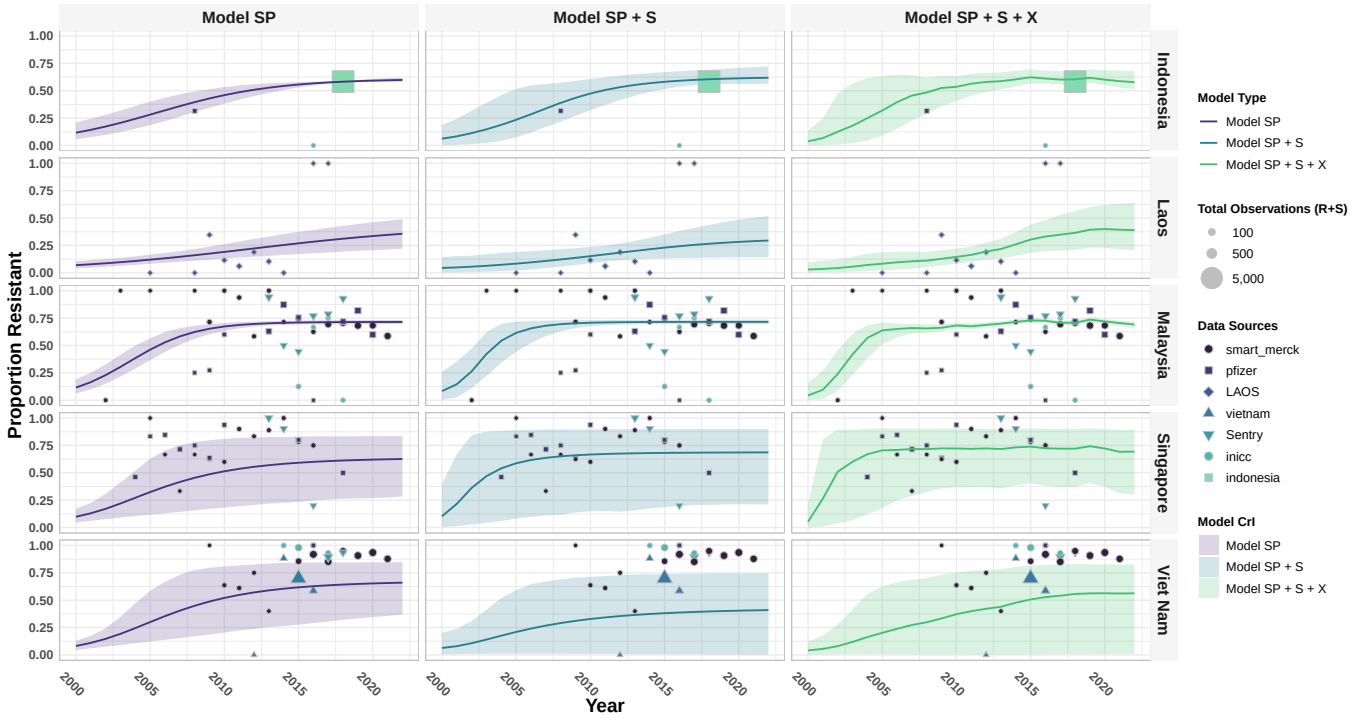
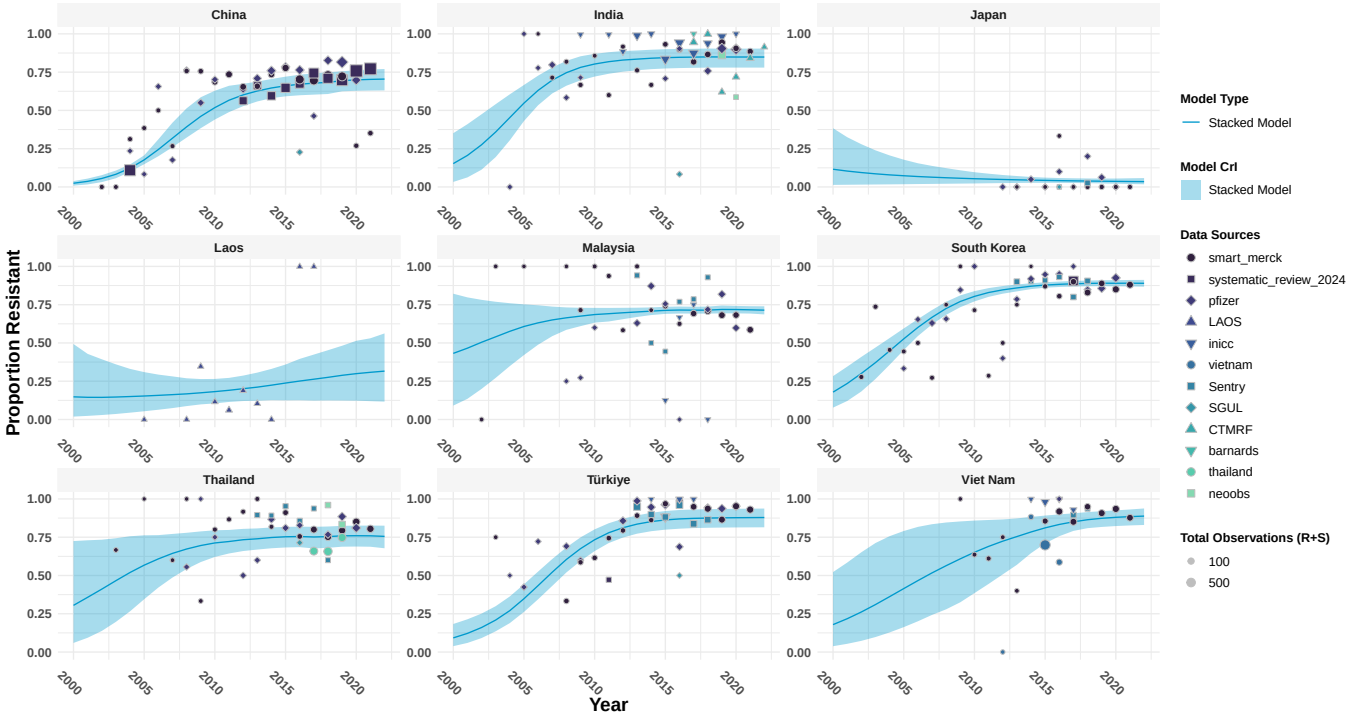


Fig. 2. Individual model fits using $\beta\,\gamma$, P_0 parameterisation of the logistic growth function for resistance of *Acinetobacter baumannii* and carbapenems model training for Asia with five whole-country held-out data. Shows model estimates for 5 countries, with two countries, Singapore and Viet Nam, held-out from fitting, and the remaining three countries are their neighbouring countries included in the model fitting. Each spatial model variant is shown as the column and each country as the row. SP is the spatial model with Gaussian process regression, SP+S is the spatial model with source-level effects, and SP+S+X is the spatial model with source-level effects and covariates. Solid lines represent the mean posterior prediction and the shaded regions show the 95% credible intervals with the different colours corresponding to each model type. The points show the observed resistance data used for training the models.

in our models and observe an improvement in model performance when including covariates in the spatial and source-level effects model (SP+S). This framework is able to capture resistance dynamics of AMR across different countries, and enables predictions for countries with no data through spatial correlation.

We extend existing methods by taking a semi-mechanistic approach in order to capture a biologically-plausible structure of resistance dynamics within a less complex fitting framework. This structure models the prevalence of resistance following logistic growth dynamics derived from a two-compartment ODE model. By representing the spread of resistance in terms of parameters describing the coefficient of acquiring resistance and the rate of losing resistance, we were able to include covariates that represented time-varying pressures, spatial correlation, and heterogeneity between data sources. We further incorporated a time-varying structure to capture temporal changes in resistance dynamics as a result of non-linearly mapping time-dependent covariates to our model parameters. This allows models to more realistically capture year-to-year fluctuations of resistance as a result of population-level shifts in factors like antibiotic consumption



(a)

Fig. 3. Stacked model predictions using β γ, P_0 parameterisation of the logistic growth function for resistance of *Acinetobacter baumannii* and carbapenems for Asia from 2000 to 2022. Shows stacked model estimates for 9 countries. Solid lines represent the mean posterior prediction and the shaded regions show the 95% credible intervals with the different colours corresponding to each model type. The points show the observed resistance data used for training the models.

and healthcare policies. Our Bayesian semi-mechanistic framework additionally accounts for country and data source-level variation, spatial correlation effects, and covariates to estimate the prevalence of AMR.

Our model comparison revealed that incorporating a model structure capable of accounting for data source-level variability in our data improves AMR model prediction accuracy. We find that models incorporating source-level random effects either with the country-level effects model or the spatial model had the best performance across all antibiotic categories, shown by the lowest LOOIC. AMR data originate from various sources which have differences in patient care settings where longer stay settings typically see more resistance, patient characteristics, local resistance prevalence (within country variation), and culture practices (sampling protocols and laboratory methods). Accounting for source-level variability can capture these differences and result in better model predictive performance. Our final model stacking weights similarly reflect this as the highest weight is assigned to models with data source-level effects for 6 of 7 antibiotic categories.

Incorporating covariates generally improved model performance across both types of models, the model with only country-level effects (C) and the spatial effects model (SP). For the model with only country-level effects (C), adding covariates (C+X) reduced the LOOIC values for 5 of 7 antibiotic categories. This is similarly observed when comparing the model weights used for stacking. We observed a similar pattern when adding covariates to our spatial models. When comparing the spatial model (SP), the spatial model with data source-level effects (SP+S), and the spatial model with data source-level effects and covariates (SP+S+X), we see a decrease in LOOIC values across all antibiotic categories when covariates are added to the spatial model with data-source level effects. While covariates alone were not able to improve the predictive performance as much as source-level effects, adding covariates to model C and SP can improve model performance and can additionally improve performance when combined with source-level effects.

When comparing the base spatial model (SP) to the country-level effects model (C), the addition of spatial correlation effects did not consistently reduce LOOIC values across all antibiotic categories. Similarly, combining spatial effects with source-level effects (SP+S) did not always improve predictive performance relative to the country and data source-level effects (C+S). While this is the case, incorporating a spatial correlation structure provides a useful feature that enables the estimation of resistance trends for countries without observed resistance data. Through the Gaussian process structure of the spatial model, we can estimate trends for data-sparse countries from their neighbouring countries. Our current implementation uses geographic distance to define the spatial correlation structure and assumes neighbouring countries share similar resistance patterns. However, alternative spatial definitions can be explored within this framework. For example, human movement patterns through travel and migration can be more mechanistically relevant

when considering a spatial structure in terms of resistance dynamics.

This study has several limitations. Our model parameterisation is derived from a two-compartment system that simplifies the dynamics of drug-resistant and drug-sensitive strains. Under this assumption, we derive our logistic growth model in terms of a two-compartment system. While our model parameterises the logistic growth model in terms of β, γ, and P_0, this simplifies more complex real-world dynamics of drug-resistant and drug-sensitive strains. Additionally, this parameterisation requires that β and γ are not equal, since this would result in the logistic growth equation being undefined. Our current spatial correlation structure uses geographic distance to define the spatial correlation structure between neighbouring countries. The spatial correlation structure with human mobility can be explored to provide a more mechanistically relevant structure.

In this work, we developed a Bayesian semi-mechanistic modelling framework designed to address existing challenges in AMR modelling. This framework accounts for data source heterogeneity, covariate data, and spatial correlation structures to model AMR dynamics, while maintaining biological interpretability. Using a model stacking approach to generate ensemble predictions, this methodology can highlight the predictive strengths of the different model structures and enables resistance estimation in regions without surveillance data. This methodology offers a potentially valuable tool to better understand the dynamics of global resistance trends and provides a foundation for evaluating public health interventions.

Acknowledgements

I would like to thank the GRAM team at IHME and Oxford and all our GRAM data collaborators for making this work possible. I would also like to thank the UK Department of Health and Social Care's Fleming Fund, and the Wellcome Trust for funding this project.

References

1. Global Antimicrobial Resistance and Use Surveillance System (GLASS) report: antibiotic use data for 2022. Available from: https://www.who.int/publications/i/item/9789240108127 [Accessed on: 2025 May 6]

2. Prestinaci F, Pezzotti P, and Pantosti A. Antimicrobial resistance: a global multifaceted phenomenon. Pathogens and Global Health 2015 Oct; 109:309–18. DOI: 10.1179/2047773215Y.0000000030. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4768623/

3. Ahmed SK, Hussein S, Qurbani K, Ibrahim RH, Fareeq A, Mahmood KA, and Mohamed MG. Antimicrobial resistance: Impacts, challenges, and future prospects. Journal of Medicine, Surgery, and Public Health 2024 Apr; 2:100081. DOI: 10.1016/j.glmedi.2024.100081. Available from: https://www.sciencedirect.com/science/article/pii/S2949916X24000343

4. Naghavi M et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. The Lancet 2024 Sep; 404. Publisher: Elsevier:1199–226. DOI: 10.1016/S0140-6736(24)01867-1. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(24)01867-1/fulltext

5. Global action plan on antimicrobial resistance. Available from: https://www.who.int/publications/i/item/9789241509763 [Accessed on: 2025 May 6]

6. Global Antimicrobial Resistance and Use Surveillance System (GLASS). Available from: https://www.who.int/initiatives/glass [Accessed on: 2025 May 6]

7. Target. Global Database for Tracking Antimicrobial Resistance (AMR) Country Self- Assessment Survey (TrACSS). Available from: http://amrcountryprogress.org/ [Accessed on: 2025 May 6]

8. UN General Assembly High-Level Meeting on antimicrobial resistance 2024. Available from: https://www.who.int/news-room/events/detail/2024/09/26/default-calendar/un-general-assembly-high-level-meeting-on-antimicrobial-resistance-2024 [Accessed on: 2025 Nov 19]

9. Limmathurotsakul D, Dunachie S, Fukuda K, Feasey NA, Okeke IN, Holmes AH, Moore CE, Dolecek C, Doorn HRv, Shetty N, Lopez AD, and Peacock SJ. Improving the estimation of the global burden of antimicrobial resistant infections. The Lancet Infectious Diseases 2019 Nov; 19. Pub-

lisher: Elsevier:e392–e398. DOI: 10.1016/S1473-3099(19)30276-2. Available from: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30276-2/fulltext

10. Porgo TV, Norris SL, Salanti G, Johnson LF, Simpson JA, Low N, Egger M, and Althaus CL. The use of mathematical modeling studies for evidence synthesis and guideline development: A glossary. Research Synthesis Methods 2019 Mar; 10:125–33. DOI: 10.1002/jrsm.1333. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6491984/

11. Kachalov VN, Nguyen H, Balakrishna S, Salazar-Vizcaya L, Sommerstein R, Kuster SP, Hauser A, Wiesch PAz, Klein E, and Kouyos RD. Identifying the drivers of multidrug-resistant Klebsiella pneumoniae at a European level. PLOS Computational Biology 2021 Jan; 17. Publisher: Public Library of Science:e1008446. DOI: 10.1371/journal.pcbi.1008446. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008446

12. Durazzi F, Pezzani MD, Arieti F, Simonetti O, Canziani LM, Carrara E, Barbato L, Onorati F, Remondini D, and Tacconelli E. Modelling antimicrobial resistance transmission to guide personalized antimicrobial stewardship interventions and infection control policies in healthcare setting: a pilot study. Scientific Reports 2023 Sep; 13. Publisher: Nature Publishing Group:15803. DOI: 10.1038/s41598-023-42511-5. Available from: https://www.nature.com/articles/s41598-023-42511-5

13. Changruenngam S, Modchang C, and Bicout DJ. Modelling of the transmission dynamics of carbapenem-resistant Klebsiella pneumoniae in hospitals and design of control strategies. Scientific Reports 2022 Mar; 12. Publisher: Nature Publishing Group:3805. DOI: 10.1038/s41598-022-07728-w. Available from: https://www.nature.com/articles/s41598-022-07728-w

14. Niewiadomska AM, Jayabalasingham B, Seidman JC, Willem L, Grenfell B, Spiro D, and Viboud C. Population-level mathematical modeling of antimicrobial resistance: a systematic review. BMC Medicine 2019 Apr; 17:81. DOI: 10.1186/s12916-019-1314-9. Available from: https://doi.org/10.1186/s12916-019-1314-9

15. Miyama T, Jung Sm, Hayashi K, Anzai A, Kinoshita R, Kobayashi T, Linton NM, Suzuki A, Yang Y, Yuan B, Kayano T, Akhmetzhanov AR, Nishiura H, Miyama T, Jung Sm, Hayashi K, Anzai A, Kinoshita R, Kobayashi T, Linton NM, Suzuki A, Yang Y, Yuan B, Kayano T, Akhmetzhanov AR, and Nishiura H. Phenomenological and mechanistic models for predicting early transmission data of COVID-19. Mathematical Biosciences and Engineering 2022; 19. Cc_license_type: cc_by Primary_atype: Mathematical Biosciences and Engineering Subject_term: Research article Subject_term_id: Research article:2043–

55. DOI: 10.3934/mbe.2022096. Available from: http://www.aimspress.com/article/doi/10.3934/mbe.2022096

16. Zhou Z, Shyti I, Kim J, and You L. Predicting population dynamics of antimicrobial resistance using mechanistic modeling and machine learning. Advanced Drug Delivery Reviews 2025 Oct; 225:115661. DOI: 10.1016/j.addr.2025.115661. Available from: https://www.sciencedirect.com/science/article/pii/S0169409X25001462

17. Nateghi Haredasht F, Amrollahi F, Maddali MV, Marshall N, Ma SP, Cooper LN, Johnson AO, Wei Z, Medford RJ, Kanjilal S, Banaei N, Deresinski S, Goldstein MK, Asch SM, Chang A, and Chen JH. Antibiotic Resistance Microbiology Dataset (ARMD): A Resource for Antimicrobial Resistance from EHRs. Scientific Data 2025 Jul; 12. Publisher: Nature Publishing Group:1299. DOI: 10.1038/s41597-025-05649-7. Available from: https://www.nature.com/articles/s41597-025-05649-7

18. Nazir A, Nazir A, Zuhair V, Aman S, Sadiq SUR, Hasan AH, Tariq M, Rehman LU, Mustapha MJ, and Bulimbe DB. The Global Challenge of Antimicrobial Resistance: Mechanisms, Case Studies, and Mitigation Approaches. Health Science Reports 2025 Jul; 8:e71077. DOI: 10.1002/hsr2.71077. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC12284435/

19. Holt KE, Carey ME, Chandler C, Cross JH, Dyson ZA, Furnham N, Glover RE, Virgo M, and Knight GM. Tools and challenges in the use of routine clinical data for antimicrobial resistance surveillance. npj Antimicrobials and Resistance 2025 May; 3. Publisher: Nature Publishing Group:37. DOI: 10.1038/s44259-025-00105-3. Available from: https://www.nature.com/articles/s44259-025-00105-3

20. Frost I, Kapoor G, Craig J, Liu D, and Laxminarayan R. Status, challenges and gaps in antimicrobial resistance surveillance around the world. Journal of Global Antimicrobial Resistance 2021 Jun; 25:222–6. DOI: 10.1016/j.jgar.2021.03.016. Available from: https://www.sciencedirect.com/science/article/pii/S2213716521000825

21. Ashley EA, Shetty N, Patel J, Doorn R van, Limmathurotsakul D, Feasey NA, Okeke IN, and Peacock SJ. Harnessing alternative sources of antimicrobial resistance data to support surveillance in low-resource settings. Journal of Antimicrobial Chemotherapy 2019 Mar; 74:541–6. DOI: 10.1093/jac/dky487. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6406030/

22. Dunachie SJ, Day NP, and Dolecek C. The challenges of estimating the human global burden of disease of antimicrobial resistant bacteria. Current Opinion in Microbiology 2020 Oct; 57:95–101. DOI:

10.1016/j.mib.2020.09.013. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7763986/

23. Lim C, Hantrakun V, Teerawattanasook N, Srisamang P, Teparrukkul P, Sumpradit N, Turner P, Day NP, Cooper BS, Peacock SJ, and Limmathurotsakul D. Impact of low blood culture usage on rates of antimicrobial resistance. The Journal of Infection 2021 Mar; 82:355–62. DOI: 10.1016/j.jinf.2020.10.040. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC7994019/

24. Kalina J. A Bayesian Framework for Regularized Estimation in Multivariate Models Integrating Approximate Computing Concepts. 2025 Sep. DOI: 10.48550/arXiv.2509.10045. Available from: http://arxiv.org/abs/2509.10045 [Accessed on: 2025 Nov 19]

25. Browne AJ, Chipeta MG, Fell FJ, Haines-Woodhouse G, Kashef Hamadani BH, Kumaran EAP, Robles Aguilar G, McManigal B, Andrews JR, Ashley EA, Audi A, Baker S, Banda HC, Basnyat B, Bigogo G, Ngoun C, Chansamouth V, Chunga A, Clemens JD, Davong V, Dougan G, Dunachie SJ, Feasey NA, Garrett DO, Gordon MA, Hasan R, Haselbeck AH, Henry NJ, Heyderman RS, Holm M, Jeon HJ, Karkey A, Khanam F, Luby SP, Malik FR, Marks F, Mayxay M, Meiring JE, Moore CE, Munywoki PK, Musicha P, Newton PN, Pak G, Phommasone K, Pokharel S, Pollard AJ, Qadri F, Qamar FN, Rattanavong S, Reiner B, Roberts T, Saha S, Saha S, Shakoor S, Shakya M, Simpson AJ, Stanaway J, Turner C, Turner P, Verani JR, Vongsouvath M, Day NPJ, Naghavi M, Hay SI, Sartorius B, and Dolecek C. Estimating the subnational prevalence of antimicrobial resistant Salmonella enterica serovars Typhi and Paratyphi A infections in 75 endemic countries, 1990–2019: a modelling study. The Lancet Global Health 2024 Mar; 12:e406–e418. DOI: 10.1016/S2214-109X(23)00585-5. Available from: https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(23)00585-5/fulltext

26. Bamber R, Sullivan B, Gorman L, Lee WWY, Avison MB, Dowsey AW, and Williams PB. A Bayesian Model Based on Local Phenotypic Resistance Data to Inform Empiric Antibiotic Escalation Decisions. Infectious Diseases and Therapy 2024 Sep; 13:1963–81. DOI: 10.1007/s40121-024-01011-3. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11343932/

27. Howard A, Green PL, Velluva A, Gerada A, Hughes DM, Brookfield C, Hope W, and Buchan I. Bayesian estimation of the prevalence of antimicrobial resistance: a mathematical modelling study. Journal of Antimicrobial Chemotherapy 2024 Sep; 79:2317–26. DOI: 10.1093/jac/dkae230. Available from: https://doi.org/10.1093/jac/dkae230

28. Cherny SS, Chowers M, and Obolski U. Bayesian network modeling of patterns of antibiotic cross-resistance by bacterial sample source. Communications Medicine 2023 May; 3. Publisher: Nature

Publishing Group:61. DOI: 10.1038/s43856-023-00289-7. Available from: https://www.nature.com/articles/s43856-023-00289-7

29. Zhang M, Wang C, and O'Connor A. A Bayesian approach to modeling antimicrobial multidrug resistance. PLoS ONE 2021 Dec; 16:e0261528. DOI: 10.1371/journal.pone.0261528. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8716034/

30. Rosato C, Green PL, Harris J, Maskell S, Hope W, Gerada A, and Howard A. Bayesian Calibration to Address the Challenge of Antimicrobial Resistance: A Review. IEEE access : practical innovations, open solutions 2024; 12:100772–91. DOI: 10.1109/ACCESS.2024.3427410. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7616450/

31. Emons M, Blanquart F, and Lehtinen S. The evolution of antibiotic resistance in Europe, 1998–2019. PLOS Pathogens 2025 Apr; 21. Publisher: Public Library of Science:e1012945. DOI: 10.1371/journal.ppat.1012945. Available from: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1012945

32. Bertagnolio S, Suthar AB, Tosas O, and Weezenbeek KV. Antimicrobial resistance: Strengthening surveillance for public health action. PLOS Medicine 2023 Jul; 20. Publisher: Public Library of Science:e1004265. DOI: 10.1371/journal.pmed.1004265. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1004265

33. Rempel OR and Laupland KB. Surveillance for antimicrobial resistant organisms: potential sources and magnitude of bias. Epidemiology & Infection 2009 Dec; 137:1665–73. DOI: 10.1017/S0950268809990100. Available from: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/surveillance-for-antimicrobial-resistant-organisms-potential-sources-and-magnitude-of-bias/A1C115B1099BECEA5231EE2E220C2EAD

34. Bhatt S, Cameron E, Flaxman SR, Weiss DJ, Smith DL, and Gething PW. Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. Journal of The Royal Society Interface 2017 Sep. Publisher: The Royal Society. DOI: 10.1098/rsif.2017.0520. Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2017.0520

35. Yao Y, Vehtari A, Simpson D, and Gelman A. Using Stacking to Average Bayesian Predictive Distributions (with Discussion). Bayesian Analysis 2018 Sep; 13. Publisher: International Society for Bayesian Analysis:917–1007. DOI: 10.1214/17-BA1091. Available from: https://projecteuclid.org/journals/bayesian-analysis/volume-13/issue-3/Using-Stacking-to-Average-Bayesian-Predictive-Distributions-with-Discussion/10.1214/17-BA1091.full

36. WHO bacterial priority pathogens list, 2024: Bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. Available from: https://www.who.int/publications/i/item/9789240093461 [Accessed on: 2025 Nov 19]

37. Fallach N, Dickstein Y, Silberschein E, Turnidge J, Temkin E, Almagor J, Carmeli Y, and Consortium obotDA. Utilising sigmoid models to predict the spread of antimicrobial resistance at the country level. Eurosurveillance 2020 Jun; 25. Publisher: European Centre for Disease Prevention and Control:1900387. DOI: 10.2807/1560-7917.ES.2020.25.23.1900387. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.23.1900387

38. Howard A, O'Donoghue M, Feeney A, and Sleator RD. Acinetobacter baumannii. Virulence 2012 May; 3:243–50. DOI: 10.4161/viru.19700. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3442836/

39. Kim YA, Kim JJ, Won DJ, and Lee K. Seasonal and Temperature-Associated Increase in Community-Onset Acinetobacter baumannii Complex Colonization or Infection. Annals of Laboratory Medicine 2018 May; 38:266–70. DOI: 10.3343/alm.2018.38.3.266. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5820073/

40. Filho SPF, Pereira MS, Nobre Rodrigues JL, Guimarães RB, Ribeiro da Cunha A, Corrente JE, Campos Pignatari AC, and Fortaleza CMCB. Seasonality and weather dependance of Acinetobacter baumannii complex bloodstream infections in different climates in Brazil. PLoS ONE 2021 Aug; 16:e0255593. DOI: 10.1371/journal.pone.0255593. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8376013/

41. Li W, Liu C, Ho HC, Shi L, Zeng Y, Yang X, Huang Q, Pei Y, Huang C, and Yang L. Association between antibiotic resistance and increasing ambient temperature in China: an ecological study with nationwide panel data. The Lancet Regional Health – Western Pacific 2023 Jan; 30. Publisher: Elsevier. DOI: 10.1016/j.lanwpc.2022.100628. Available from: https://www.thelancet.com/journals/lanwpc/article/PIIS2666-6065(22)00243-7/fulltext

42. Kritsotakis EI and Groves-Kozhageldiyeva A. A systematic review of the global seasonality of infections caused by *Acinetobacter* species in hospitalized patients. Clinical Microbiology and Infection 2020 May; 26:553–62. DOI: 10.1016/j.cmi.2019.09.020. Available from: https://www.sciencedirect.com/science/article/pii/S1198743X19305051

43. Temperature and precipitation gridded data for global and regional domains derived from in-situ and satellite observations. Available from: https://cds.climate.copernicus.eu/datasets/

insitu-gridded-observations-global-and-regional?tab=overview [Accessed on: 2025 Apr 6]

44. Browne AJ, Chipeta MG, Haines-Woodhouse G, Kumaran EPA, Hamadani BHK, Zaraa S, Henry NJ, Deshpande A, Reiner RC, Day NPJ, Lopez AD, Dunachie S, Moore CE, Stergachis A, Hay SI, and Dolecek C. Global antibiotic consumption and usage in humans, 2000–18: a spatial modelling study. The Lancet. Planetary Health 2021 Nov; 5:e893–e904. DOI: 10.1016/S2542-5196(21)00280-1. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC8654683/

45. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and STAN. 2nd ed. New York: Chapman and Hall/CRC, 2020 Mar. DOI: 10.1201/9780429029608

46. Stewart PS. An occupancy model with a Gaussian process for spatial autocorrelation in Stan. 2022 May. Available from: https://peter-stewart.github.io/blog/gaussian-process-occupancy-tutorial/ [Accessed on: 2025 Jul 23]

47. Vehtari A, Gabry J, Magnusson M, Yao Y, Bürkner PC, Paananen T, and Gelman A. loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models. 2015 Jun. DOI: 10.32614/CRAN.package.loo. Available from: https://CRAN.R-project.org/package=loo [Accessed on: 2025 Nov 19]

48. Model averaging/weighting via stacking or pseudo-BMA weighting — loo_model_weights. Available from: https://mc-stan.org/loo/reference/loo_model_weights.html [Accessed on: 2025 Nov 19]

49. Gaussian Processes. Available from: https://mc-stan.org/docs/stan-users-guide/gaussian-processes.html [Accessed on: 2025 Nov 19]

50. Wang J. An Intuitive Tutorial to Gaussian Process Regression. Computing in Science & Engineering 2023 Jul; 25. arXiv:2009.10862 [stat]:4–11. DOI: 10.1109/MCSE.2023.3342149. Available from: http://arxiv.org/abs/2009.10862

51. Posterior and Prior Predictive Checks. Available from: https://mc-stan.org/docs/stan-users-guide/posterior-predictive-checks.html [Accessed on: 2025 Nov 19]

52. Shigemoto K. Various Logistic Curves in SIS and SIR Models. European Journal of Mathematics and Statistics 2023 Jan; 4. arXiv:2211.04186 [physics]:1–6. DOI: 10.24018/ejmath.2023.4.1.185. Available from: http://arxiv.org/abs/2211.04186

53. Bacaër N. Verhulst and the logistic equation (1838). *A Short History of Mathematical Population Dynamics*. Ed. by Bacaër N. London: Springer, 2011 :35–9. DOI: 10.1007/978-0-85729-115-8_6. Available from: https://doi.org/10.1007/978-0-85729-115-8_6 [Accessed on: 2025 Apr 6]

54. Besag J. Spatial Interaction and the Statistical Analysis of Lattice Systems. Journal of the Royal Statistical Society: Series B (Methodological) 1974; 36. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1974.tb00999.x:192–225. DOI: 10.1111/j.2517-6161.1974.tb00999.x. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1974.tb00999.x

55. Besag J, York J, and Mollié A. Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 1991 Mar; 43:1–20. DOI: 10.1007/BF00116466. Available from: https://doi.org/10.1007/BF00116466

56. Riebler A, Sørbye SH, Simpson D, and Rue H. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Statistical Methods in Medical Research 2016 Aug; 25. Publisher: SAGE Publications Ltd STM:1145–65. DOI: 10.1177/0962280216660421. Available from: https://doi.org/10.1177/0962280216660421

57. Morris M, Wheeler-Martin K, Simpson D, Mooney SJ, Gelman A, and DiMaggio C. Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. Spatial and Spatio-temporal Epidemiology 2019 Nov; 31:100301. DOI: 10.1016/j.sste.2019.100301. Available from: https://www.sciencedirect.com/science/article/pii/S1877584518301175

58. Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data. Available from: https://mc-stan.org/learn-stan/case-studies/icar_stan.html [Accessed on: 2025 May 5]

Supporting Information

S1 Text. Supporting Model Outputs

Combined file of supporting text and figures.

A Data Summary

Figures to summarise data.

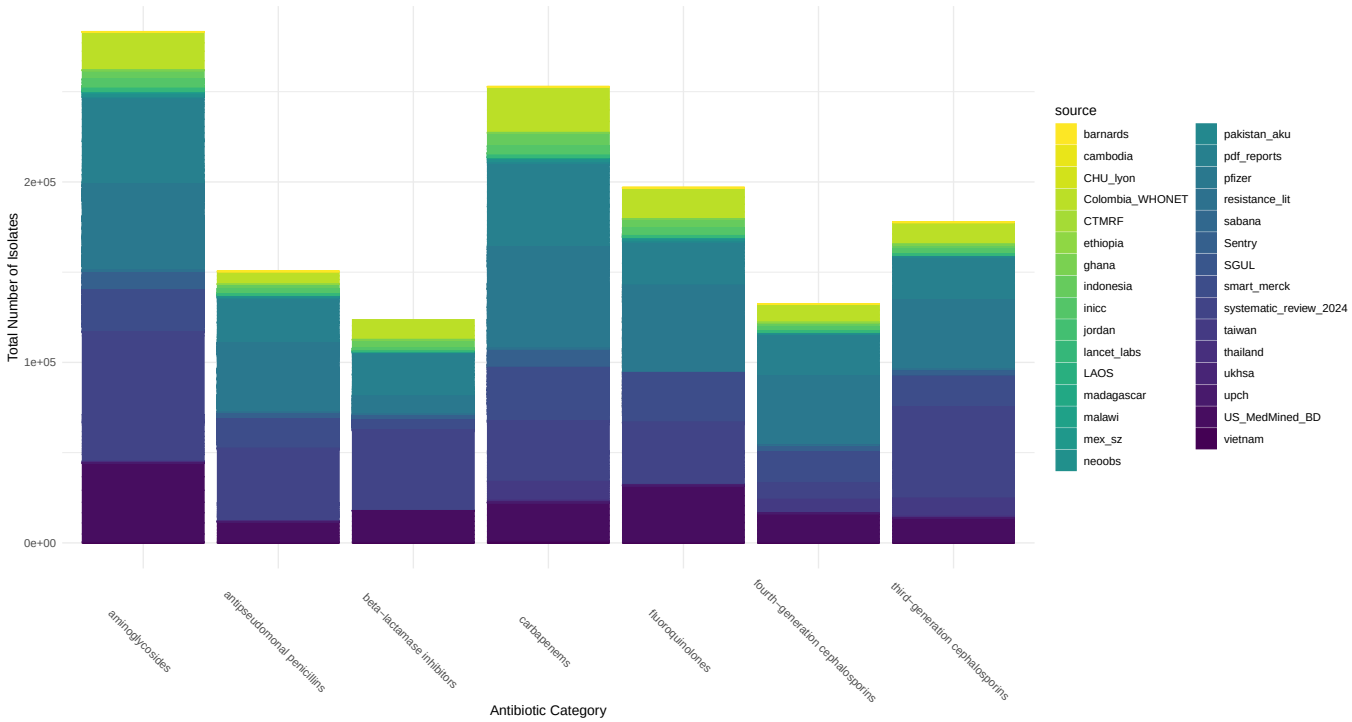


Fig. 4. Total number of isolates for global *Acinetobacter baumannii* data. The total number of isolates for 113 countries from 1991 to 2022 are shown across 7 antibiotic categories.

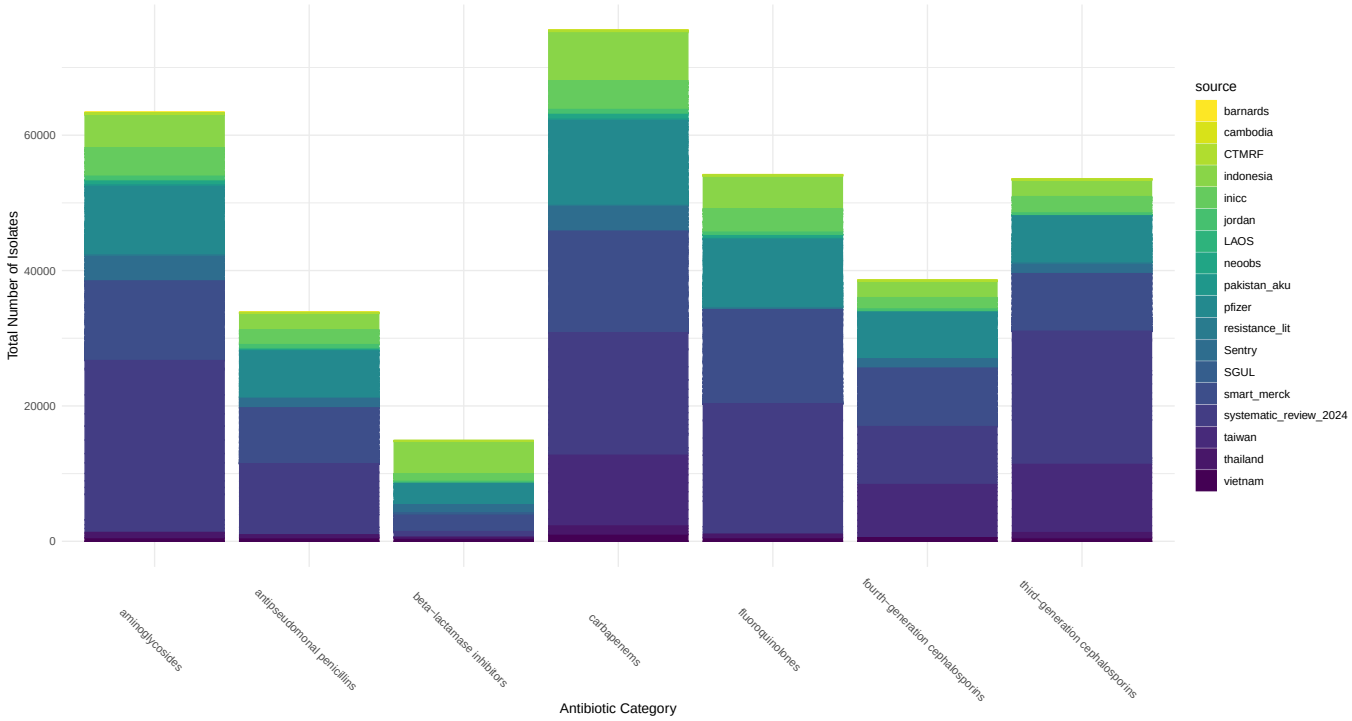


Fig. 5. Total number of isolates for *Acinetobacter baumannii* data across Asia. The total number of isolates for 32 countries in Asia from 2000 to 2022 are shown across 7 antibiotic categories.

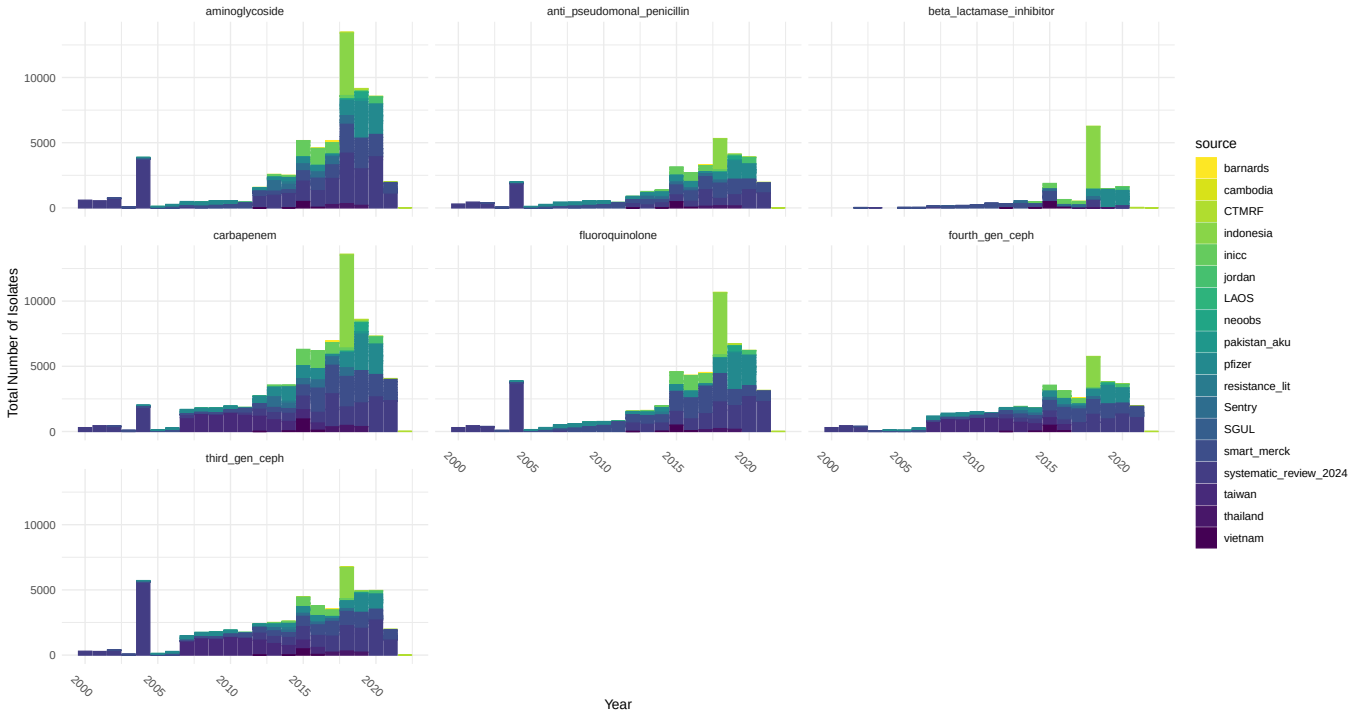


Fig. 6. Total number of isolates for *Acinetobacter baumannii* data across Asia per year. The total number of isolates for 32 countries in Asia from 2000 to 2022 are shown across 7 antibiotic categories.

B Model Derivation

A two-compartment ODE model can be used to represent the general dynamics of a drug-resistant and drug-sensitive strains. This framework assumes that the resistance dynamics do not substantially affect the dynamics of sensitive strains.

The differential equations describing the dynamics of this two-compartment model are shown in Equation 10 and Equation 11. Here, the S compartment describes the drug-sensitive strain and I describes the drug-resistant strain. In these equations, β is a coefficient that can describe how sensitive strains gain resistance and γ is the rate at which resistant strains become sensitive.

$$\frac{dS}{dt} = \frac{-\beta}{N} SI + \gamma I \tag{10}$$

$$\frac{dI}{dt} = \frac{\beta}{N} SI - \gamma I \tag{11}$$

We assume a constancy of the population N to obtain a proportional value for drug-resistant isolates similarly shown in previous studies and can rewrite Equation 11 in terms of I, β, γ under this assumption.

$$\frac{dS}{dt} + \frac{dI}{dt} = 0 \tag{12}$$

$$\frac{dI}{dt} = \frac{\beta}{N}(N - I)I - \gamma I \tag{13}$$

$$= \beta(1 - \frac{I}{N})I - \gamma I \tag{14}$$

$$= \beta I - \frac{I^2}{N}\beta - \gamma I \tag{15}$$

$$= (\beta - \gamma)I - \frac{I^2}{N}\beta \tag{16}$$

An analytical solution to the two-compartment model is the Verhulst equation, also known as the logistic growth function (see Equation 19). We use the logistic growth function as an analytical solution to this two-compartment model [52]. This equation can be expressed in terms of the initial proportion of resistance P_0, a maximum limiting value for the population K, and the growth rate r (see in Equation 18) [53].

$$\frac{dP}{dt} = rP(1 - \frac{P}{K}) \tag{17}$$

$$= rP - \frac{r}{K}P^2 \tag{18}$$

$$P(t) = \frac{K}{1 + (\frac{K - P_0}{P_0})e^{-rt}} \tag{19}$$

To derive Equation 19 in terms of β and γ, we structure Equation 16 in terms of Equation 18 to find the growth rate r and carrying capacity K in terms of β and γ.

$$r = \beta - \gamma \tag{20}$$

$$K = \frac{\beta - \gamma}{\beta}N \tag{21}$$

By setting $N = 1$, we can estimate the proportion of resistant strains to overall strains (resistant and susceptible). This allows us to express Equation 19 in terms of β and γ, as shown in Equation 22.

$$P(t) = \frac{1 - \frac{\gamma}{\beta}}{1 + (\frac{1-\frac{\gamma}{\beta}}{P_0} - 1)e^{-(\beta-\gamma)t}} \tag{22}$$

C Covariate Models

For the models with covariates, the β prior is set to explore the logarithmic space while the priors for P_0 and γ remain the same as the country-level effects model.

In Equation 23, θ_k is the coefficient of the covariate k and is set to have a weakly informative normal prior distribution. The covariate values are then multiplied by their estimated coefficient value, summed over the total number of covariates, and then added to their estimated country-level β_c term in the log space. The final value is then exponentially transformed.

$$\theta_k \sim \text{Normal}(0, 1) \tag{23}$$

$$log_\beta \sim \text{Normal}(0, 1) \tag{24}$$

$$\sigma_{log_\beta} \sim \text{Exponential}(0.5) \tag{25}$$

$$log_{\beta_c} \sim \text{Normal}(log_\beta, \sigma_{log_\beta}) \tag{26}$$

$$\beta_{c,t} = exp(log_{\beta_c} + \sum_{k=1}^{K} \theta_k \cdot X_{k,c,t}) \tag{27}$$

D Country-level Priors

The prior distributions are shown. Here, we have two options we explored for the initial resistance P_0 prior. Option 1: We select a weakly informative normal prior distribution for P_{0_logit} since we do not have specific information on this parameter on a logit scale and having a hierarchical structure. Option 2: We select a beta(2,2) prior distribution for the initial resistance value for each country with no hierarchical structure. The first option was preferred for fitting models starting at the initial time point for each country. The second option was preferred for fitting models all from the same year and starting at previous years where data is sparse since it is computationally more efficient. For this analysis, we use the second option for the initial resistance prior as we are generating predictions from 2000 onwards and data is sparse. For β and γ we

select an exponential prior distribution with a value of three since the value for these parameters have a higher probability density closer to 0 and should be positive. $\sigma_{P_{0_\text{logit}}}$, σ_β, and σ_γ all have an exponential prior distribution with a value of 0.5 as we expect this value to be positive and expect to have a slightly higher probability density closer to 0 since we do not expect major differences between countries for these parameters.

Prior distributions for the hierarchical model:

$$\text{Option 1 for } P_0 \text{ (logit):} \tag{28}$$

$$P_{0_\text{logit}} \sim \text{Normal}(0, 1) \tag{29}$$

$$\sigma_{P_{0_\text{logit}}} \sim \text{Exponential}(0.5) \tag{30}$$

$$P_{0_\text{logit},c} \sim \text{Normal}(P_{0_\text{logit}}, \sigma_{P_{0_\text{logit}}}) \tag{31}$$

$$\tag{32}$$

$$\text{Option 2 for } P_0\text{:} \tag{33}$$

$$ProbP_{0,c} \sim \text{beta}(2, 2) \tag{34}$$

$$\tag{35}$$

$$\beta \sim \text{Exponential}(3) \tag{36}$$

$$\sigma_\beta \sim \text{Exponential}(0.5) \tag{37}$$

$$\beta_c \sim \text{Normal}(\beta, \sigma_\beta) \tag{38}$$

$$\tag{39}$$

$$\gamma \sim \text{Exponential}(3) \tag{40}$$

$$\sigma_\gamma \sim \text{Exponential}(0.5) \tag{41}$$

$$\gamma_c \sim \text{Normal}(\gamma, \sigma_\gamma) \tag{42}$$

After defining these priors, we can define the parameter values used as inputs into Equation 1

$$P_{0,c} = \text{logit}^{-1}(P_{0_\text{logit},c}) \text{ or } ProbP_{0,c} \tag{43}$$

$$\beta_c = \beta_c \tag{44}$$

$$\gamma_c = \gamma_c \tag{45}$$

E Data source-level Models

The prior distributions for the source-level parameters are shown. The models incorporating source-level random effects extend the country-level model and use the same priors as shown in the previous section for β_c, γ_c, and $P_{0,c}$. Here, the source-dependent multiplicative factor η_s is set to be bounded between 0 and 1 to ensure positive constraints and with a beta distribution density concentrated near 1. The structure of this model shows that the underlying curve for each country should be similar and the variation between sources should be relatively small when scaling the underlying curve.

$$\eta_s \sim \text{beta}(5, 1) \tag{46}$$

$$\tag{47}$$

F Spatial Effects

To incorporate spatial effects into the model, an initial approach was taken using the Besag-York-Mollié (BYM) model and extract the intrinsic conditional autoregressive component to smooth estimates based on neighbouring countries and to account for spatial interactions based on pairs of countries [54, 55, 56]. We incorporate these spatial models following the implementation of BYM model in Stan [57, 58]. We add in parameter ϕ to account to encode the neighbour relationships in terms of the pairwise difference formula as shown in Equation 48. The pairwise difference formula is calculated between i and j if they are neighbours and under the condition that i cannot equal j. A soft sum-to-zero constraint is also applied to ϕ to soft-centers the mean close to zero and is shown in Equation 49 [57]. A precision term, τ_ϕ, is added to estimate the standard deviation of ϕ, as shown in Equation 51. The spatial component, $\sigma_\phi * \phi_{c,i}$, is added related to I_0 and β as we expect neighbouring regions to have similar initial resistance rates and influence on transmission. When testing models with spatial effects, we only include source and spatial effects as we now relate regions to their neighbours.

$$p(\phi) \propto \exp\left(-\frac{1}{2}\sum_{i\sim j}(\phi_i - \phi_j)^2\right) \tag{48}$$

$$\sum(\phi) \sim \mathcal{N}(0, 0.001 \cdot C) \tag{49}$$

$$\tau_\phi \sim \text{Gamma}(1, 0.1) \tag{50}$$

$$\sigma_\phi = \text{inv}(\sqrt{\tau_\phi}) \tag{51}$$

$$I_{0,s_spatial,i} = \text{logit}^{-1}(I_{0_\text{logit},s,i} + \sigma_\phi * \phi_{c,i}) \tag{52}$$

$$\beta_{s_spatial,i} = \beta_{s,i} + \sigma_\phi * \phi_{c,i} + \sum_{j=1}^{n} \theta_j \cdot X_{j,i,c} \tag{53}$$

$$\gamma_{s_spatial,i} = \gamma_{s,i} \tag{54}$$

After testing spatial effects in terms of the BYM model, an alternate method was explored in terms of Gaussian process regression to include spatial autocorrelation to reduce model run times and improve model convergence [45, 46].

To create a relationship between sites that are closer together, a kernel function can be used to define the covariance between countries based on distance. The kernel functions is shown as Equation 55 where D is the distance between countries, ζ^2 is the maximum covariance between sites, and ρ^2 is the rate that covariance declines. This relationship is then used to estimate the a spatial intercept term for each country following methods from previous examples [45, 46].

$$K_{i,j} = \zeta^2 * exp(-\rho^2 * D_{i,j}^2) \tag{55}$$

We incorporate the spatial effects intercept term,k, for P_0, β, and γ. For P_0, the spatial term is added to the population-level term and transformed to the inverse logit space as shown in Equation 56. For both $beta$ and $gamma$, we add the spatial term in the log space and then exponentiated the sum of the spatial term and the population level parameter. The final β and γ terms used in the final probability equation is shown in Equations 57 and 58.

$$P_{0,c} = \text{logit}^{-1}(P_{0_\text{logit}} + k_{P_0,c}) \tag{56}$$

$$\beta_c = exp(log_\beta + k_{\beta,c}) \tag{57}$$

$$\gamma_c = exp(log_\gamma + k_{\gamma,c}) \tag{58}$$

G Posteriors Densities from Simulated Data Model Fits

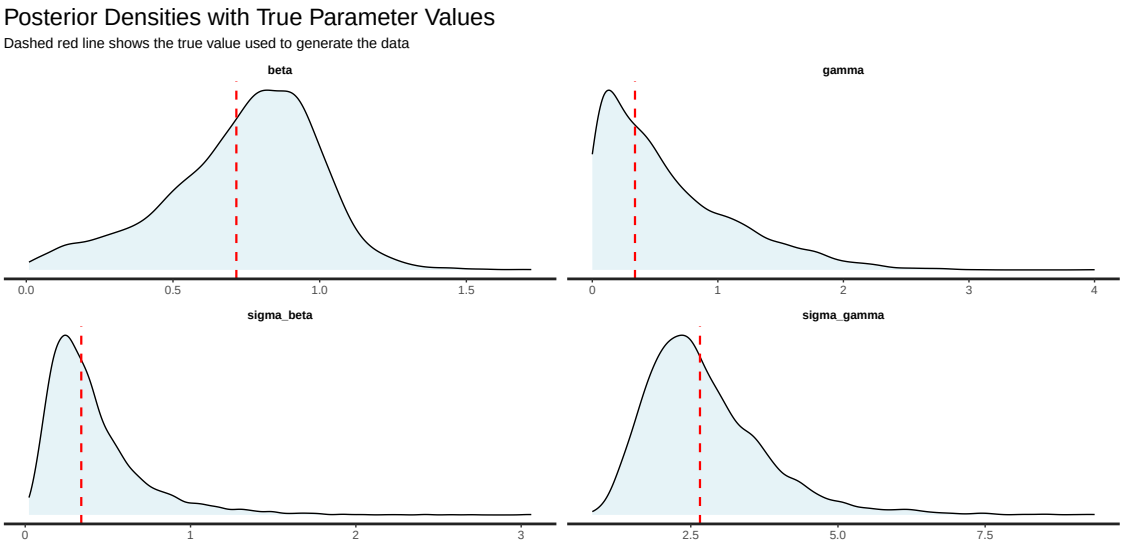
The models are fitted using simulated data with known model parameters to confirm parameter estimates are recoverable. The parameter posterior densities resulting from fitting simulated data with the country-level random effects model with a β, γ, I_0 (initial resistance prevalence P_0 in our model) parameterisation are shown. The resulting posterior densities are shown with the corresponding true parameter value used to generate data.

H Model predictions for 5 whole-country held-out data

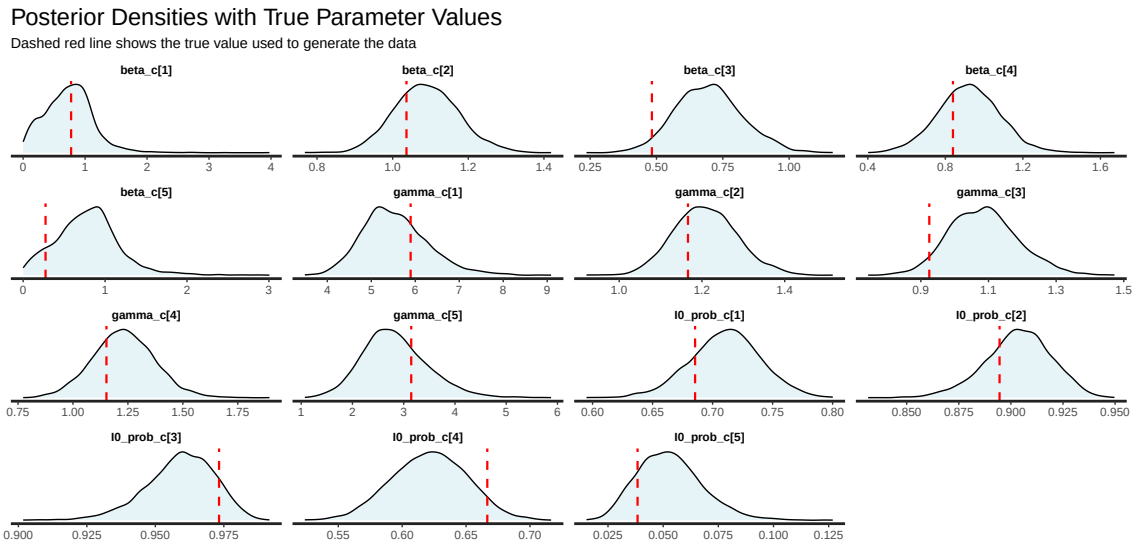
Gaussian process spatial imputation using models with spatial-effects (SP, SP+S, SP+S+X) to generate estimates. We refit the models without five randomly selected whole-country held-out data to show the generated model estimates compared to the actual observed resistance data for these five countries.

I Model Trace and Pair Plots

Model convergence was checked through trace and pair plots generated after fitting models to *Acinetobacter baumannii* resistance data for each pathogen-antibiotic combination. Example pair and trace plots for carbapenems are shown.



(a)



(b)

Fig. 7. Posterior densities from data simulation using β γ parameterisation of the logistic growth function a) shows the population level β and γ and b) shows the country-specific posterior densities recovered using the true parameter values shown by the dashed red line.

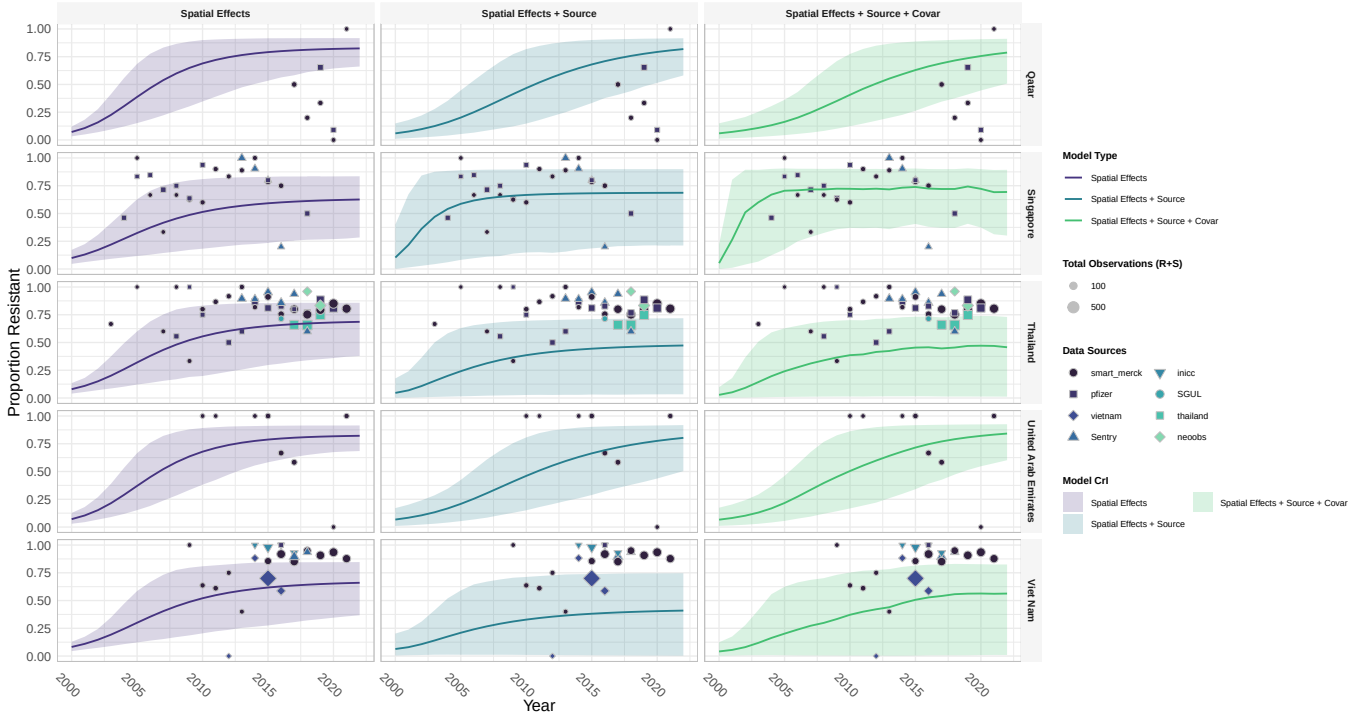
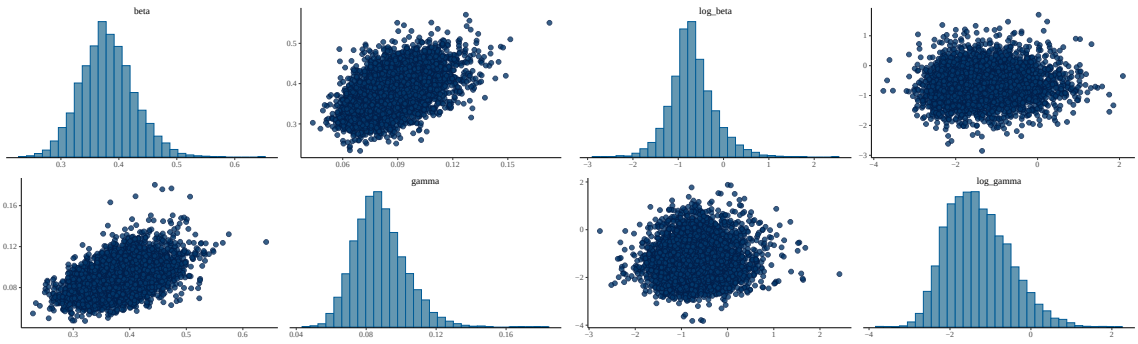
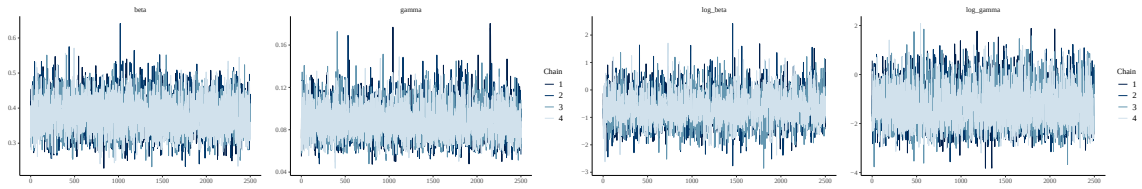


Fig. 8. Estimates generated for whole-country held-out data using Gaussian process structure in the spatial models for Asia from 2000 to 2022 Shows model estimates for 5 countries with each model variant as the column and each country as the row. Spatial effects (SP) is the spatial model with Gaussian process regression, Spatial Effects+Source (SP+S) is the spatial model with source-level effects, and Spatial Effects+Source+Covar (SP+S+X) is the spatial model with source-level effects and covariates. Solid lines represent the mean posterior prediction and the shaded regions show the 95% credible intervals with the different colours corresponding to each model type. The points show the observed resistance data used for training the models.



(a) **(b)**

Fig. 9. Pairs plot of β and γ the using β γ, P_0 parameterisation of the logistic growth function for resistance of *Acinetobacter baumannii* and carbapenems model trained for Asia from 2000 to 2022. a) Shows pairs plots of country-level random effects b) pairs plot for base spatial-effects model



(a) **(b)**

Fig. 10. Trace plots of β and γ the using β γ, P_0 parameterisation of the logistic growth function for resistance of *Acinetobacter baumannii* and carbapenems model trained for Asia from 2000 to 2022. a) Shows trace plots of the country-level random effects model b) trace plot for the spatial-effects model