mathematics

MDPI

Article

Short-Term Predictions of the Trajectory of Mpox in East Asian Countries, 2022–2023: A Comparative Study of Forecasting Approaches

Aleksandr Shishkin [1], Amanda Bleichrodt [1], Ruiyan Luo [1], Pavel Skums [2], Gerardo Chowell [1,*] and Alexander Kirpich [1,*]

[1] Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, GA 30303, USA; rluo@gsu.edu (R.L.)
[2] School of Computing, University of Connecticut, Storrs, CT 06269, USA
* Correspondence: gchowell@gsu.edu (G.C.); akirpich@gsu.edu (A.K.)

Abstract: The 2022–2023 mpox outbreak exhibited an uneven global distribution. While countries such as the UK, Brazil, and the USA were most heavily affected in 2022, many Asian countries, specifically China, Japan, South Korea, and Thailand, experienced the outbreak later, in 2023, with significantly fewer reported cases relative to their populations. This variation in timing and scale distinguishes the outbreaks in these Asian countries from those in the first wave. This study evaluates the predictability of mpox outbreaks with smaller case counts in Asian countries using popular epidemic forecasting methods, including the ARIMA, Prophet, GLM, GAM, *n*-Sub-epidemic, and Sub-epidemic Wave frameworks. Despite the fact that the ARIMA and GAM models performed well for certain countries and prediction windows, their results were generally inconsistent and highly dependent on the country, i.e., the dataset, as well as the prediction interval length. In contrast, *n*-Sub-epidemic Ensembles demonstrated more reliable and robust performance across different datasets and predictions, indicating the effectiveness of this model on small datasets and its utility in the early stages of future pandemics.

Keywords: forecasting; ARIMA; Prophet; GAM; mpox; monkeypox; Asia; China; South Korea; Japan; Thailand; model comparison; epidemiology; incidence

MSC: 92-04; 92-08; 92-10; 92Bxx; 92D25; 92D30

Citation: Shishkin, A.; Bleichrodt, A.; Luo, R.; Skums, P.; Chowell, G.; Kirpich, A. Short-Term Predictions of the Trajectory of Mpox in East Asian Countries, 2022–2023: A Comparative Study of Forecasting Approaches. *Mathematics* **2024**, *12*, 3669. https://doi.org/10.3390/math12233669

Academic Editor: Takashi Suzuki

Received: 25 September 2024
Revised: 14 November 2024
Accepted: 20 November 2024
Published: 23 November 2024

1. Introduction

Epidemics and outbreaks of infectious diseases pose significant challenges to global public health systems, necessitating robust surveillance, early detection, and effective response strategies. In May 2022, a multi-country outbreak of mpox began, spreading a viral disease, previously endemic primarily in Central and West Africa [1], worldwide. In response, the WHO declared the mpox outbreak a public health emergency of international concern (PHEIC) on 23 July 2022, citing the rapid global spread and evolving transmission dynamics [1]. The declaration mobilized a coordinated international response, including heightened surveillance, public awareness campaigns, and the re-purposing of smallpox vaccines to mitigate transmission.

Mpox (formerly known as monkeypox) is caused by the mpox virus (MPXV), a member of the Orthopoxvirus genus, which also includes the variola virus responsible for smallpox [2]. While mpox is generally less severe than smallpox, these two diseases share similarities in clinical presentation, including fever, rash, and lymphadenopathy. The 2022–2023 outbreak of mpox marked a notable departure from multiple historical patterns, as mpox, previously confined to African regions, spread extensively across diverse geographic regions, including Europe, North and South America, and Asia. Moreover,

the diversity of symptoms observed in the 2022–2023 outbreak, ranging from classic presentations to atypical single lesions, presented diagnostic challenges and underscored the importance of accurate surveillance and clinical characterization.

The initial cluster of cases was identified in the United Kingdom, with the first case reported in London on 6 May 2022 [3]. Subsequent investigations traced the source of infection to a patient with recent travel history from Nigeria, where mpox is endemic. Outbreaks were reported worldwide, with the most affected countries being Brazil, Canada, France, Germany, Spain, the United Kingdom, and the United States. Notably, the outbreak primarily, but not exclusively, affected men who have new or multiple sexual contacts [4]. This highlights the importance of outbreak studies to prevent mortality and morbidity in vulnerable groups.

Despite the WHO efforts, the outbreak persisted, albeit at reduced levels, prompting the WHO to declare an end to the PHEIC in May 2023 [1]. Pockets of transmission, however, continued, with a total of 92,226 confirmed cases and 185 deaths reported across 117 countries as of 4 May 2023 [5]. Throughout 2022, Asia experienced the fewest cases across all regions, with only approximately 20 cases across all countries in the region [6]. Nevertheless, the transmission dynamics of the outbreak have also quickly evolved, with confirmed cases emerging later in countries across Asia, including South Korea, China, Japan, and Thailand among the primarily impacted countries in 2023. The outbreak appeared at different times in each country. The first case was in South Korea on 22 June 2022 [7]. The next index case was in Thailand on 21 July 2022 [8]. Several days later, on 25 July 2022, Japan registered the first mpox case [9]. Finally, China registered its index case on 16 September 2022 [10]. This new geographic expansion into Southeast Asia indicates the ongoing nature of the outbreak and highlights the need for comprehensive global surveillance and response mechanisms, even in previously non-endemic regions.

Interventions varied among the countries. China implemented limited measures, and concerns were raised about the sufficiency of vaccinations and monitoring efforts [11]. South Korean authorities claimed to establish quarantine measures for travelers from affected countries and expand its testing capabilities [7]. Similarly, Thai authorities introduced border surveillance for mpox and started a public awareness campaign. Officials reported securing a significant amount of vaccines to contain the outbreak [12]. Japanese authorities launched consultation and public awareness interventions. The Tokyo Metropolitan Government also established an mpox testing system at the Tokyo Metropolitan Institute of Public Health [13].

Forecasting is important for public health resource allocation, and public health organizations are in need of reliable forecasting tools [14]. This study specifically focuses on the second stage of the mpox outbreak in Southeast Asian countries, with a particular emphasis on the predictability of its dynamics using standard epidemiological forecasting tools. The countries most impacted by the mpox outbreak have attracted the most research interest, and thus, epidemic onset has been studied and forecasted most thoroughly for the European and North American regions [15–18]. However, the nature of the epidemics and the associated data are quite different between the first and subsequent waves, with the first wave having significantly higher case counts. Consequently, those models tuned for high-count data may perform differently on low-count data. The extent of this difference remains an open question. This study addresses this question by evaluating the short-term forecasting performance of ensemble Sub-epidemic frameworks and other time-series models. Understanding the capabilities and limitations of these forecasting approaches should guide epidemiology researchers and practitioners in selecting optimal computational tools for different scenarios of potential future outbreaks. This knowledge will inform future outbreak response strategies and enhance preparedness for emerging infectious threats.

2. Methods

2.1. Data Sources

The data utilized in this study were sourced from publicly available statistics provided by the World Health Organization (WHO), collected and published during 2022–2023 [5]. To avoid numerical instabilities in the estimates, we used reported weekly confirmed case counts for the analysis of the 2023 mpox outbreak in South Korea, China, Japan, and Thailand. The WHO definition of confirmed cases used in this dataset is as follows: "The person with MPXV infection laboratory confirmed by detection of unique sequences of viral DNA by real-time polymerase chain reaction (PCR) c and/or sequencing" [19]. The reported time series for weekly reports contained only data points with case counts above zero. Weeks with no reported cases were assigned zero counts to fill those gaps. The same studied period was selected for all four countries from 2 February 2023, until 28 December 2023, so each country's time series would have the same length and starting point. If the reporting of cases for some counties started after 2 February 2023 or ended earlier than 28 December 2023, those non-reported weekly counts were coded as zeros. The summaries of the resulting raw weekly counts in South Korea, China, Japan, and Thailand are presented in Figure 1. For the analysis in the ensemble framework, processed cumulative data were utilized, while other considered models utilized weekly case count data.

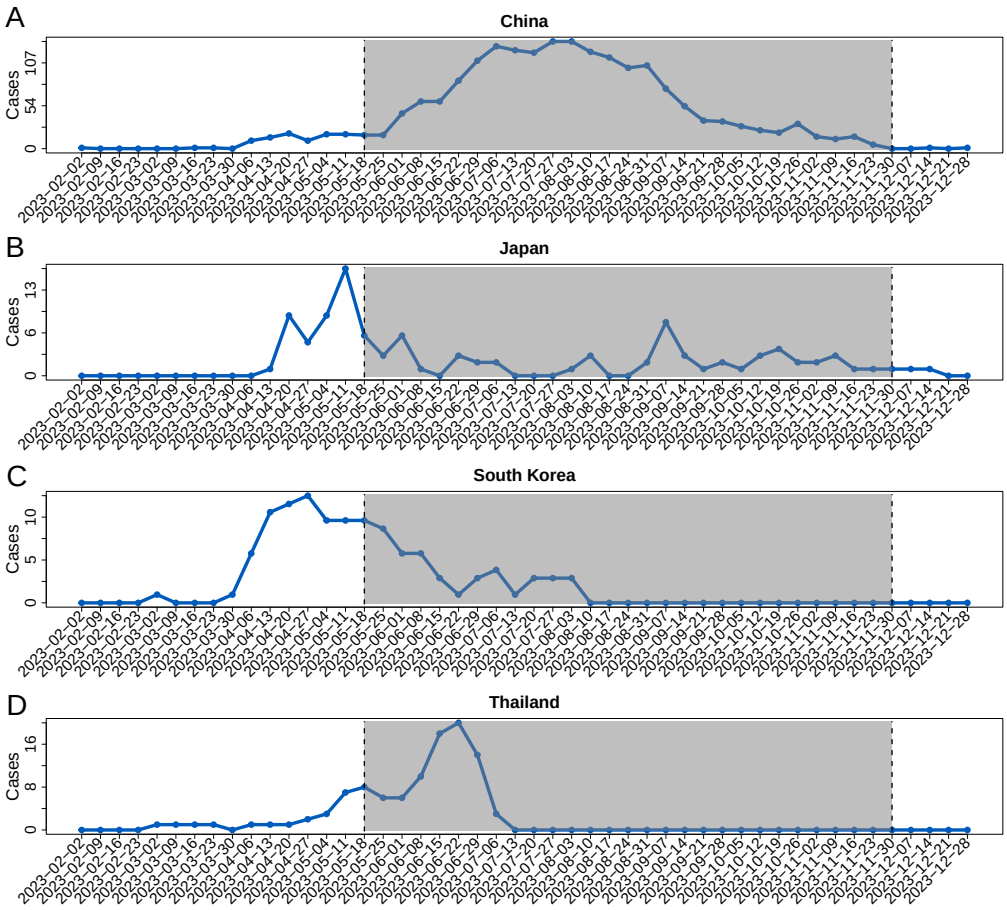

Figure 1. Mpox weekly cases (weekly incidence) for China (**A**), Japan (**B**), South Korea (**C**), and Thailand (**D**) from 2 February 2023 to 28 December 2023. Weekly incidence predictions were issued for the dates that are in the highlighted area from 18 May 2023 to 30 November 2023.

2.2. Analyzed Models

The presented analysis encompasses a diverse set of forecasting methodologies to predict weekly incidence, totaling eleven distinct models [15]. These models include variants of

the n-Sub-epidemic framework and the Sub-epidemic Wave (spatial wave) framework [20]. Additionally, for references and performance and prediction comparisons, a generalized additive model (GAM) [21], a generalized linear model (GLM), an auto-regressive integrated moving average model (ARIMA) [22], and Facebook's Prophet model [23] can be referenced. Detailed formulations for all these models, except the GLM, are presented in subsequent sections to elucidate the underlying principles and techniques utilized for forecasting.

For the model fitting, `StatModPredict` `R` Shiny app [24] was used for the ARIMA, GAM, and Prophet models, which provided a graphical user interface to run implementations of those models in `R` language for statistical computing [24]. For the n-Sub-epidemic and spatial-wave frameworks, the publicly available `Matlab` implementations by the model authors were adopted from the corresponding GitHub repositories [25,26].

The incidence predictions were made for every week starting from 20 April 2023 to 28 December 2023 (end of the available data). Each prediction utilized 12 previous time points (weeks) of incidence data for the model calibration. This resulted in forecasts beginning at the 13th time point (week) from the beginning of the time series, i.e., from 20 April 2023. Predictions of different lengths were also considered for the performance evaluation of the model. In particular, one, two, three, and four incidence points (weeks) ahead were predicted. For the models implemented in the `StatModPredict` suite [24], the default parameters for each model set by the suite were unitized, which are outlined in the Appendix A. For the n-Sub-epidemic and Sub-epidemic Wave frameworks, the default set of parameters was used as well, as they were set in the `options.m` file, except those described in the Appendix A.

2.2.1. The n-Sub-Epidemic Modeling Framework

The n-Sub-epidemic modeling framework involves partitioning the epidemic into overlapping Sub-epidemic trajectories indexed by $i = 1, 2, \ldots, n$. These Sub-epidemics could be used to delineate various geographical regions, demographic groups, or temporal stages within the epidemic progression [27]. A tutorial guide on the use of the n-Sub-epidemic modeling framework for infectious disease forecasting has been published [28].

Mathematically, the framework is formulated by modeling each Sub-epidemic incidence trajectory independently. For instance, each model accounts for factors of the i-th Sub-epidemic such as size, growth rate, and transmission dynamics.

The formulation of the n-Sub-epidemic framework is formulated as a system of differential equations. The equations have the same structure and differ by the set of parameters and variables that characterize each i-th Sub-epidemic ($i = 1, 2, \ldots, n$) and have the form:

$$\frac{dC_i(t)}{dt} = r_i A_{i-1}(t) C_i(t)^{p_i} \left(1 - \frac{C_i(t)}{K_{0i}}\right). \tag{1}$$

In Formula (1) for the i-th epidemic wave, the change in the number of weekly cumulative cases (i.e., weekly cumulative incidence) $C_i(t)$ is defined as $\frac{dC_i(t)}{dt}$ at time t and is defined by a set of parameters (r_i, p_i, K_{0i}) that is specific to the wave i. The set of parameters consists of r_i—the growth rate; p_i—the growth scaling coefficient; and K_{0i}—the final size of the Sub-epidemic. The indicator variable A_{i-1} is auxiliary and is represented by

$$A_i(t) = \begin{cases} 1, C_{i-1}(t) > C_{thr} \\ 0, \text{otherwise} \end{cases} \quad \text{for } i = 2, ..., n, \tag{2}$$

where $A_1(1)$ is set to 1 for the first Sub-epidemic Wave $i = 1$. For n waves, the total of $3n + 1$ parameters has to be estimated. The initial number of mpox cases is parameterized in the first wave as $C_1(0) = I_0$, where I_0 is the initial number of cases at the starting point. The cumulative curve at time t is modeled as

$$C_{tot}(t) = \sum_{i=1}^{n} C_i(t). \tag{3}$$

Overall, the n-Sub-epidemic modeling framework provides a structured and systematic approach to dissecting disease outbreaks, enabling researchers to delve into various aspects of the epidemic and derive insights critical for effective public health responses [27]. The framework is a generalization of the logistic growth model, which corresponds to the special case $n = 1$.

2.2.2. Sub-Epidemic Wave Modeling Framework

The spatial wave framework is a modeling approach designed to capture the complex trajectories of disease outbreaks, particularly those characterized by linked, overlapping, and synchronous Sub-epidemics known as "waves" that represent weekly cumulative case counts (i.e., cumulative incidence) over time. This framework, which is an extension of the n-Sub-epidemic modeling paradigm, has already been proven effective in depicting diverse epidemic patterns, including prolonged plateaus and multi-peak trajectories observed in various outbreaks [20]. The corresponding tutorial for using the spatial wave Sub-epidemic model has been published [29].

The framework could also be formulated as a system of coupled differential equations. Each Sub-epidemic Wave, denoted by i, is represented by the equation

$$\frac{dC_i(t)}{dt} = rA_{i-1}(t)C_i(t)^p\left(1 - \frac{C_i(t)}{K_i}\right). \tag{4}$$

Here, $\frac{dC_i(t)}{dt}$ is the change in the number of cumulative reported incidence $C_i(t)$ for Sub-epidemic i. The values $C_i(t)$ represent the cumulative number of cases for Sub-epidemic $i = 1, 2, \ldots, n$ at the time t, while K_i indicates the size of the Sub-epidemic i. The parameter $A_i(t)$ is the same as in the n-Sub-epidemic model (2).

The initial Sub-epidemic size is denoted as K_0 while the subsequent Sub-epidemics' sizes are denoted K_i and are described by the following formula:

$$K_i = K_0 e^{-q(i-1)}. \tag{5}$$

The growth rate per unit of time is denoted by r and the "scaling of growth" parameter by p. Unlike the n-Sub-epidemic model, the parameters r and p remain constant across Sub-epidemics, with K_i dependent on the decline rate q.

To characterize a Sub-epidemic Wave composed of multiple Sub-epidemics $n > 1$, a total of five parameters (r, p, C_{thr}, q, K_0) are needed for an arbitrary n where $i = 1, 2, \ldots, n$. However, when only one Sub-epidemic is present ($n = 1$), four parameters are sufficient. Typically, the maximum number of considered Sub-epidemics is up to five ($n \leq 5$), which form the combined epidemic wave trajectory.

The curve of total weekly cases (weekly incidence) could be given by the following formula:

$$dC_{total}(t) = \sum_{i=1}^{n} C_i(t). \tag{6}$$

2.2.3. Autoregressive Integrated Moving Average (ARIMA) Model

The ARIMA model is a widely used method for time series analysis and forecasting future values based on past observations and moving averages. For this study, a time series is a number of weekly case counts (i.e., incidence) for each week of the studied period. The ARIMA model consists of three main components: autoregression (AR), differencing or integration (I), and moving average (MA) [22].

1. Autoregression (AR): This component models the relationship between an observation and a number of lagged observations (i.e., observations at previous time steps). The $AR(p)$ component, where p is the number of lagged observations, can be represented as

$$AR(p) = \phi_1 Y(t-1) + \phi_2 Y(t-2) + ... + \phi_p Y(t-p), \tag{7}$$

where $Y(t-i)$ is the value of the time series at time unit $t-i$ for $i = 1, 2, \ldots, p$ previous time intervals, and $\phi_1, \phi_2, ..., \phi_p$ are the parameters of the model.

2. Differencing (I): This component aims to make the time series stationary before model fitting by differencing the consecutive observations. The order d indicates the number of times differencing must be performed between the observations to obtain a new transformed series $y(t)$. For example, for $d = 0$ the transformed variable is $y(t) = Y(t)$, and for $d = 1$, the transformed variable is $y(t) = Y(t) - Y(t-1)$.

3. Moving Average (MA): This component represents the sum of the parameters (averages) that characterize the time series trend with the corresponding errors. The $MA(q)$ component can be represented as

$$MA(q) = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \tag{8}$$

where c is the global parameter of the time series, $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model, and ε_{t-j} are the error terms at times $t-j$ for $j = 1, 2, \ldots q$.

The complete ARIMA model formulation has the form

$$y(t) = c + \sum_{i=1}^{p} \phi_i \, y(t-i) + \sum_{j=1}^{q} \theta_j \, \epsilon_{t-j} + \epsilon_t, \tag{9}$$

where $y(t)$ is the transformed series after the d-th differencing, c is the global parameter, φ_i is autoregressive coefficients (parameters) for $i = 1, 2, \ldots, p$, θ_j are moving average coefficients (parameters) for $j = 1, 2, \ldots, q$, ϵ_{t-j} refers to $j = 0, 1, 2, \ldots, q$, which are random errors at times t_j.

Alternatively, the classic ARIMA model could be rewritten in a more concise form using backward shift operator notations to account for lags [30]. In particular, for a backward shift operator B,

$$\begin{aligned} &BY(t) = Y(t-1) \text{ and similar} \\ &B^p Y(t) = B\left(B^{p-1} Y(t)\right) \text{ for } p > 1, \end{aligned} \tag{10}$$

which is useful for short representation of differencing. For example, for $d = 1$,

$$y(t) = Y(t) - Y(t-1) = Y(t) - BY(t) = (1-B)Y(t), \tag{11}$$

while the same is true for a d-th order difference, i.e., $(1-B)^d Y(t)$. Using the backward shift operator formulas,

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p, \\ \theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q, \end{aligned} \tag{12}$$

we have a more concise representation of the classic ARIMA model with the d-th order differences:

$$\phi(B)(1-B)^d Y(t) = c + \theta(B)\varepsilon_t. \tag{13}$$

2.2.4. Prophet Model

The Prophet model is a forecasting tool developed by Facebook's Core Data Science team [23]. It is designed to handle time series data with strong seasonal patterns and multiple seasonality. Prophet is particularly useful for forecasting in business settings where datasets may contain irregularities such as missing data points, outliers, or changes in trend, but it has also been widely used in public health [15,31–35]. In this analysis, the Prophet model was used to predict weekly mpox case counts (i.e., weekly incidence). For the study, the Prophet model implementation from the `Prophet` package in `R` was used.

Prophet utilizes an additive model that decomposes a time series into several components: trend, seasonality, holidays, and error [23]. It employs a piecewise linear or logistic growth curve to model the trend and Fourier series to introduce seasonality. The model also incorporates special events or holidays that may influence the time series. It can also include change points, which are dates on which the trend is changing. They could be either specified manually or estimated automatically. The general form of the model is

$$Y(t) = g(t) + s(t) + h(t) + \epsilon(t), \tag{14}$$

where $Y(t)$ is the mpox cases at time point t, $g(t)$ is the trend component, $s(t)$ is the periodical seasonal component, $h(t)$ is the irregular component (like holidays, for example), and $\epsilon(t)$ is modeling error at time t. The default parametrization was used for the periodic component, while the irregular events and holidays were not modeled as described in Appendix A.

2.2.5. Generalized Additive Model

The GAM is a flexible regression model that extends the linear regression framework by allowing for non-linear relationships between the predictors and the response variable [21,36]. In this study, the response variable is weekly mpox case counts at time t, while the predictor is time t. GAMs are particularly useful for capturing complex relationships in data without making strong assumptions about the functional form. `R` implementation of GAM from the package `mgcv` was used in this study.

In a GAM, the response variable is modeled as a sum of smooth functions of the predictors, along with potentially linear terms and other parametric components. The `R` implementation of this model from the package `mgcv` was used. With time t as the only predictor, we fit the GAM model:

$$Y(t) = \beta_0 + g(t) + \epsilon(t),$$

where $g(t)$ is represented with the following basis expansion: $g(t) = \sum_{j=1}^{k} \beta_j f_j(t)$. Here, $\{f_1(t), f_2(t), \ldots, f_k(t)\}$ represent basis functions and $\{\beta_0, \beta_1, \beta_2, \ldots, \beta_k\}$ are coefficients to be estimated within the model. GAMs are widely used in various fields, including environmental science, epidemiology, and finance, where relationships between variables may be non-linear and complex. They offer advantages such as interpretability, flexibility, and the ability to incorporate domain knowledge through the specification of smooth terms [37].

2.3. Forecasting Performance Comparison

To compare the performance of the selected models, four performance metrics were considered, which included mean squared error (MSE), mean absolute error (MAE), coverage 95% prediction interval (95% PI), and weighted interval score (WIS). MSE and MAE are widely used in regression analysis to evaluate how well regression fits the data. However, they are also applicable to time series analysis [38]. WIS is a metric used to evaluate the accuracy of probabilistic forecasts, particularly in the context of forecasting intervals or quantiles. It is particularly useful for assessing the performance of prediction intervals or full predictive distributions rather than point forecasts. The WIS combines information about the central tendency and the dispersion of the forecast distribution, providing a comprehensive measure of forecast quality [39]. The coverage of a 95% prediction interval (PI) measures the proportion of observed values that fall within the 95% PIs [30] and provides the alternative metric to evaluate the performance. Each performance metric was calculated by comparing weekly case counts predicted by the corresponding models with the reported case counts, while the exact formulas for metric calculation are outlined below in the text.

MAE quantifies the average absolute difference between predicted and actual values computed over predicted time points:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

where y_i is the observed value, $\hat{y}_i$ is the model-predicted value, and n is the number of predicted observations.

MSE quantifies the average squared difference between predicted and actual values, computed over predicted time points, while giving greater weight to larger errors:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where y_i is the observed value, $\hat{y}_i$ is the model-predicted value, and n is the number of predicted observations.

The coverage of the 95% prediction interval assesses the proportion of the observed values that fall within the prediction interval, which is computed to capture 95% of future observations. It is calculated as follows:

$$\text{Coverage 95\% PI} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{I}(L_i \leq y_i \leq U_i)$$

where L_i and U_i are the lower and upper bounds of the 95% prediction interval for the i-th observation, and $\mathbb{I}(\cdot)$ is an indicator function that returns 1 if the observed value y_i falls within the interval and 0 otherwise.

WIS evaluates the accuracy of probabilistic predictions by considering both the width and coverage of prediction intervals. For a 95% prediction interval, the WIS is calculated as

$$\text{WIS} = \frac{U_i - L_i}{2} + \frac{2}{\alpha}(L_i - y_i)\mathbb{I}(y_i < L_i) + \frac{2}{\alpha}(y_i - U_i)\mathbb{I}(y_i > U_i)$$

where α represents the nominal level of the interval (e.g., 0.05 for a 95% interval), L_i and U_i are the lower and upper bounds of the interval, and $\mathbb{I}(\cdot)$ is an indicator function [40].

To perform thorough comparisons, the forecasting metrics were computed for each point in the prediction period. As a result, a total of 36 predictions from 20 April 2023 to 28 December 2023, for four prediction lengths (horizons) and countries, were evaluated. The considered horizons were one, two, three, and four units (i.e., weekly intervals). The considered performance metrics were averaged for each prediction length and country, so there was a single number for each metric, model, country, and prediction horizon combination.

3. Results

3.1. Time Series Observational Analysis

The reported weekly incidence for the four studied countries are illustrated in Figure 1. In particular, each country experienced a single major wave, albeit at different times and at quite different scales. In China, the single wave peaked in August 2023, in Japan and South Korea in May 2023, and in Thailand in June 2023. Shapes of incidence curves are also different. In Japan and South Korea, a major wave was followed by several minor spikes, and China experienced only one wave with a longer period. Thailand had a minor wave before the major wave. China experienced significantly more cases during its major wave than the other three countries.

3.2. Forecasting Performance Comparisons

The averages of the considered metrics for each prediction length and country are presented in Figures 2–5. In total, there are 36 predictions from 20 April 2023 to 28 December 2023 for each prediction length (horizon) and country.

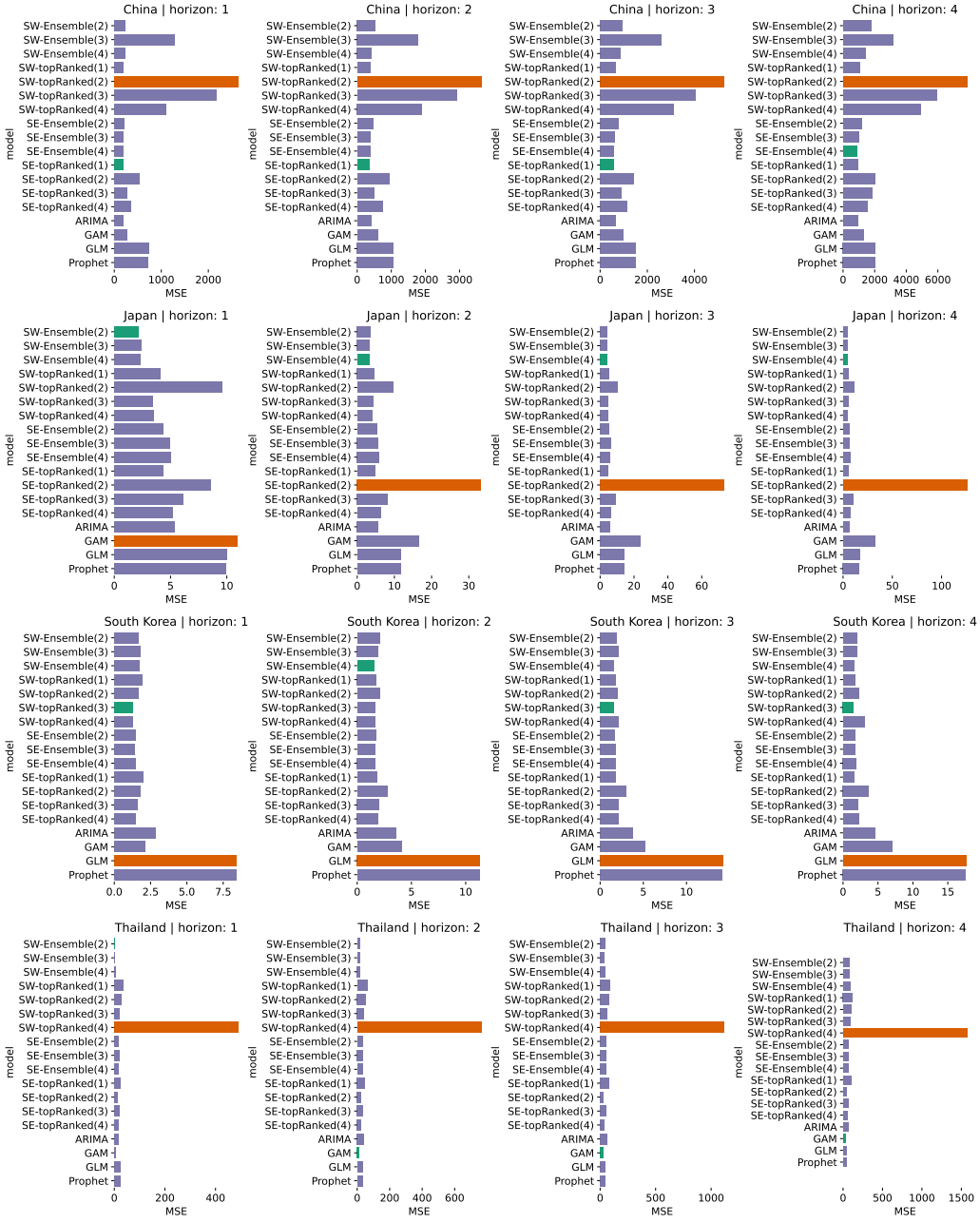

Figure 2. MSE of predictions across China, Japan, South Korea, and Thailand and four prediction horizons (1, 2, 3, and 4 weeks forward). Teal bars are the best metric for the particular combination of location and horizon, and dark orange bars are the models with the worst metrics. SW stands for Sub-epidemic Wave and SE for n-Sub-epidemic frameworks. Ranked models are singular models used for the ensembles.

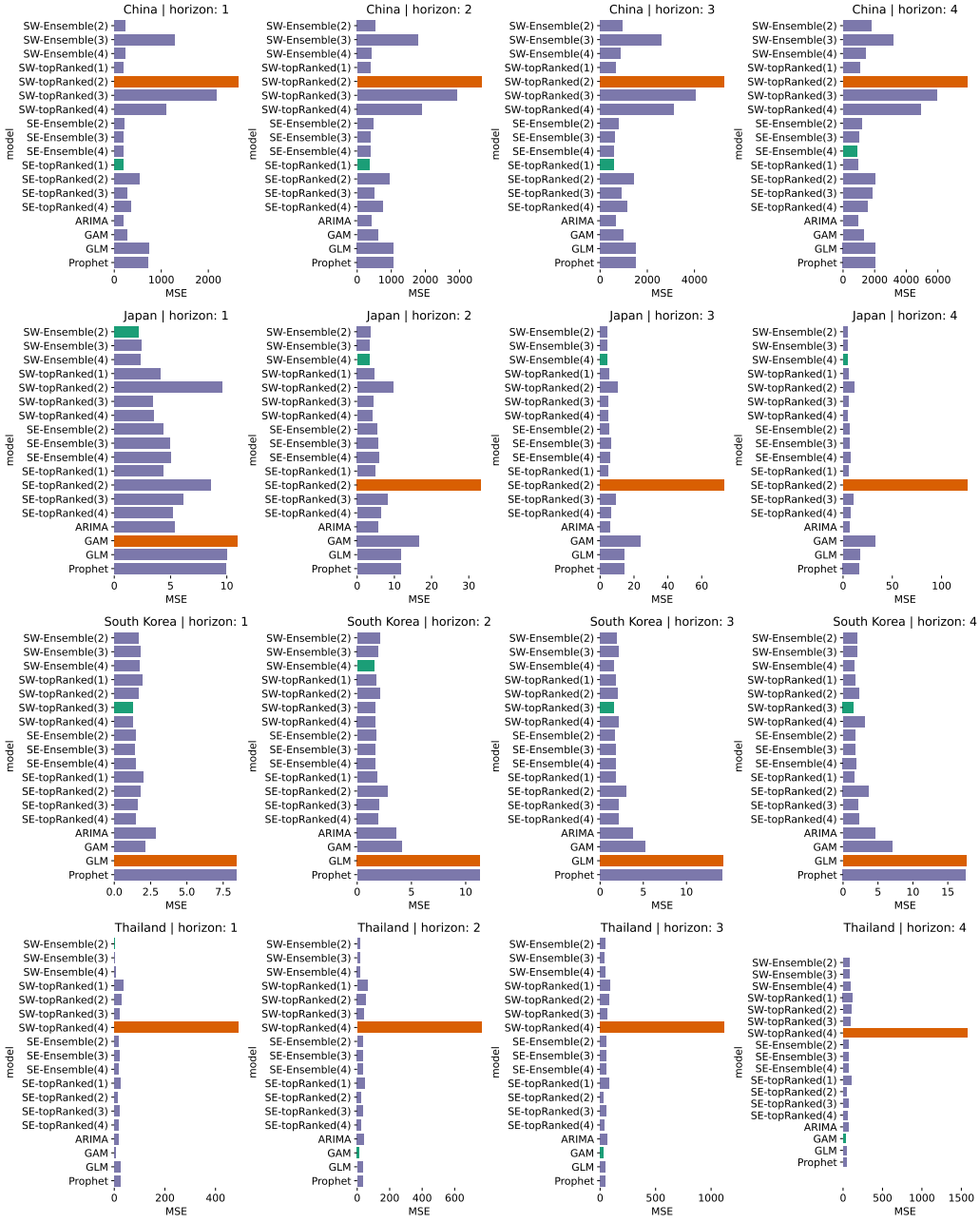

Figure 3. MAE of predictions across China, Japan, South Korea, and Thailand and four prediction horizons (1, 2, 3, and 4 weeks forward). Teal bars are the best metric for the particular combination of location and horizon, and dark orange bars are the models with the worst metrics. SW stands for Sub-epidemic Wave and SE for n-Sub-epidemic frameworks. Ranked models are singular models used for the ensembles.

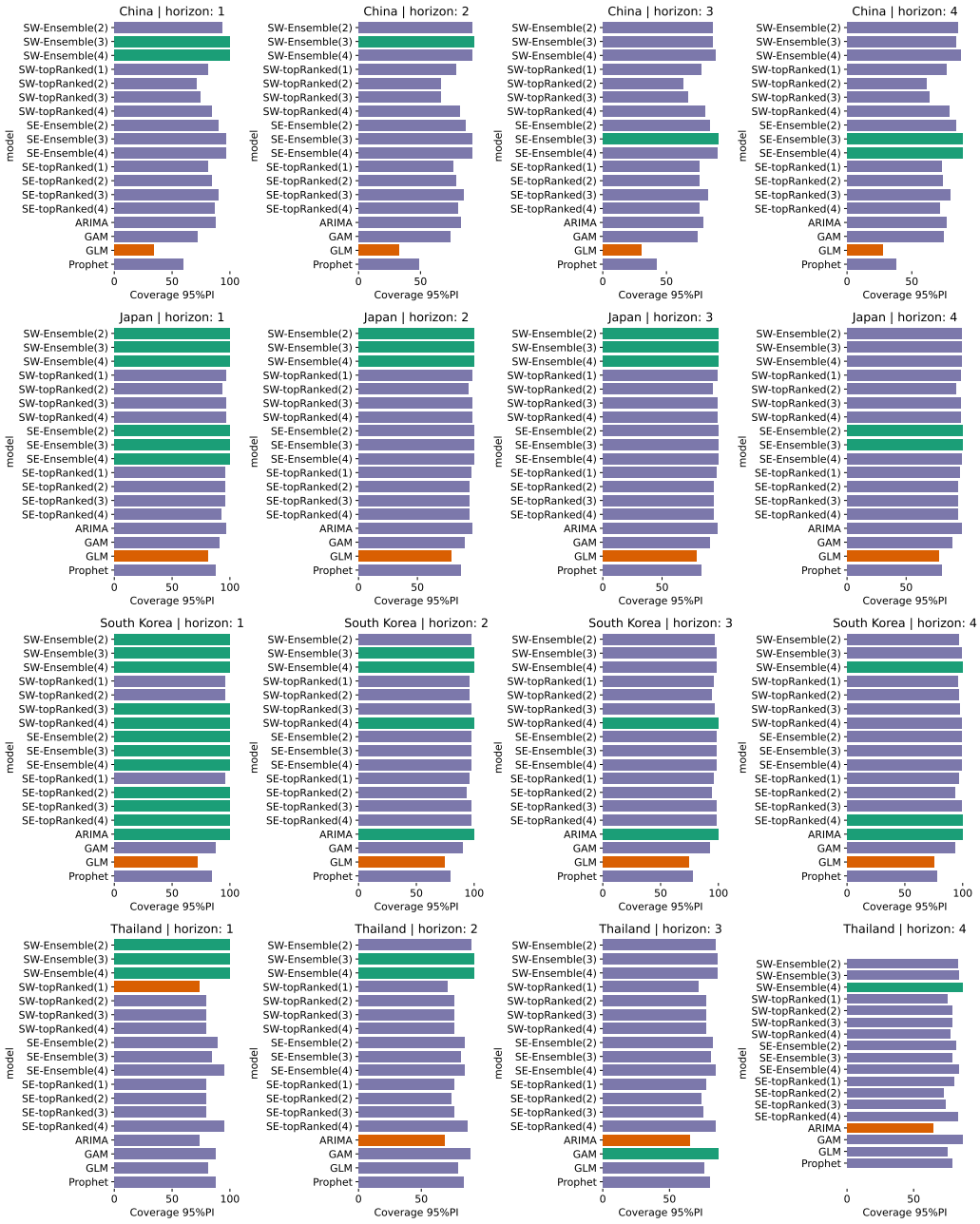

Figure 4. Coverage 95% Prediction Interval metrics across China, Japan, South Korea, and Thailand and four prediction horizons (1, 2, 3, and 4 weeks forward). Teal bars are the best metric for the particular combination of location and horizon, and dark orange bars are the models with the worst metric. SW stands for Sub-epidemic Wave and SE for n-Sub-epidemic frameworks. Ranked models are singular models used for the ensembles.

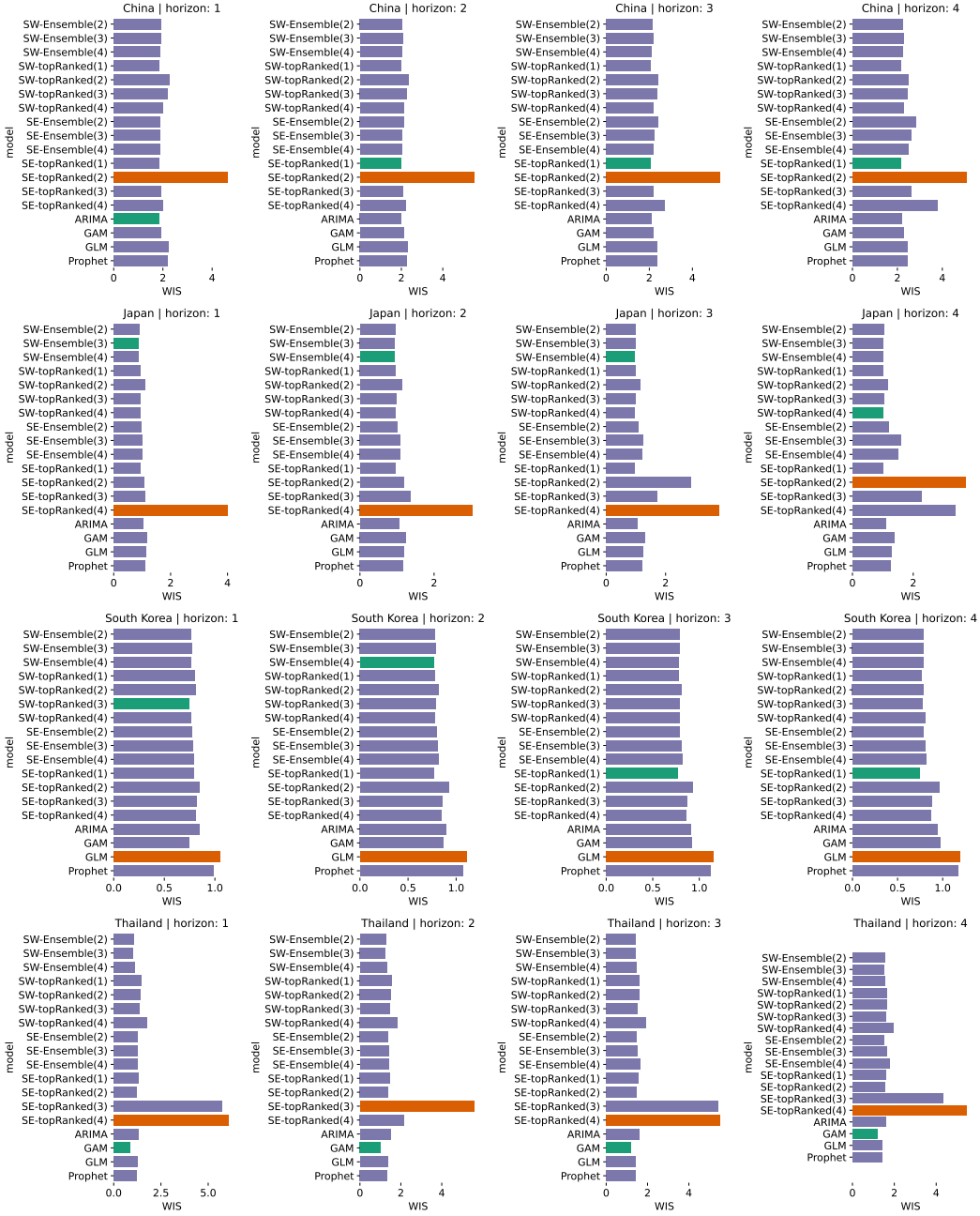

Figure 5. WIS metrics for predictions across China, Japan, South Korea, and Thailand and four prediction horizons (1, 2, 3, and 4 weeks forward). The metric is on a logarithmic scale with a base of 10. Furthermore, all values were multiplied by 10, so no value would not go below 0. Teal bars are the best metric for the particular combination of location and horizon, and dark orange bars are the models with the worst metrics. SW stands for Sub-epidemic Wave and SE for n-Sub-epidemic frameworks. Ranked models are singular models used for the ensembles.

3.2.1. China

The *n*-Sub-epidemic first-ranked models generally performed the best according to the MSE metrics (especially for the horizons 1, 2, and 3), with the *n*-Sub-epidemic (SE) ensemble (horizon 4) of four models and the first-ranked Sub-epidemic Wave (SW) model following closely. MAE was best for the SW first-ranked model on horizon 1, while on other horizons, SE first-ranked models were best. Other ranked Sub-epidemic Wave models showed the worst metrics across all models for all four horizons. For the metric of 95% PI coverage in China, ensembles of both SE (horizons 3, 4) and SW (horizons 1, 2) showed the

best results, while the GLM performed the worst across all horizons, scoring only 27.6% on horizon 4. The WIS metric was the best for ARIMA on horizon 1 and the first-ranked SE model for the other horizons, with the second-ranked SE model achieving the maximum value on all horizons.

3.2.2. Japan

The MSE was lowest for the SW ensembles of two (for horizon 1) and four models (horizon 2, 3, and 4), while it was highest for GAM on horizon 1 and the second-ranked SE model on horizons 2, 3, and 4, with a maximum of 126 at horizon 4. The MAE was best for the SW ensembles of two for horizon 1, four models for horizons 2 and 3, and the fourth-ranked SW model for horizon 4. GAM had the worst MAE. The 95% PI coverage was generally good across all models, reaching 100%, indicating they captured real data well in their predictions, although GAM and Prophet were the least successful. In terms of WIS, the SW ensembles of two and four models, as well as the fourth-ranked SW model, performed the best, while the second and fourth SE models had the highest WIS.

3.2.3. South Korea

The MSE was lowest for the SW ensemble of four models for horizon 2 and the third-ranked SW model for horizons 1, 3, and 4. The greatest MSE was observed with GLM on all horizons, with a maximum value of 17.7 at horizon 4, closely followed by Prophet. The third-ranked SW model (horizons 2, 3, and 4) and GAM (horizon 1) had the best MAE, while GLM and Prophet had the worst across all horizons. The 95% PI coverage was good across all models, with GLM and Prophet having the lowest metrics. The SW ensemble of four models for horizons 2 and 3, the third-ranked SW model for horizon 1, and the fourth-ranked SE model for horizon 4 had the lowest WIS, whereas GLM had the highest values across all horizons.

3.2.4. Thailand

GAM (horizons 2, 3, and 4) and the SW ensemble of two models (horizon 1) had the best MSE here, with a minimum of 3.5 on horizon 1. The fourth-ranked SW model had the highest MSE by far across all horizons, reaching 1573 on horizon 4. Similarly, GAM had the best MAE, with a minimum of 0.9 at horizon 1, and the fourth-ranked SW model had the most error for all horizons. The 95% PI coverage was best for the SW ensembles of two, three, and four models (horizons 1, 2, and 4), as well as for GAM (horizon 3). The first-ranked SW model (horizon 1) and ARIMA (horizons 2, 3, and 4) had the lowest coverage among all models. The WIS metric was lowest for GAM on all horizons, while the third (horizon 2) and fourth-ranked (horizons 1, 3, and 4) SE models performed poorly.

By comparing the results for the studied countries, conclusions can be drawn about model performance for different settings. In particular, for countries with low incidence counts and earlier onset, such as Japan, South Korea, and Thailand, ARIMA and GAM showed the worst performance. In contrast, for countries with larger counts and later onset, such as China, the performance was the best. On the other hand, Prophet and GLM consistently showed poor performance, while the SW and SE frameworks demonstrated consistently good performance across all countries.

4. Discussion

In this study, we evaluated and compared the performances of seasonal ARIMA, GLM, GAM, Prophet, Sub-epidemic Wave, and *n*-Sub-epidemics models for the selected East Asian countries of China, South Korea, Japan, and Thailand for the mpox outbreak for the period of time from 2 February 2023 to 28 December 2023. Specifically, the "second wave" of mpox during 2023 was more localized in Southeast Asia and had much lower counts, which allowed us to evaluate the models' performance for much smaller case counts in a more localized setting. The ensembles of the *n*-Sub-epidemic and Sub-epidemic Wave models demonstrated good performance consistently across different sets of data and

prediction horizons and across all measured metrics. They performed slightly better than other models on the Chinese time series, possibly due to greater case counts than in the other datasets and a single big wave with a few minor waves. Other models' performances were heavily dependent on the particular time series. GAM outperformed other models on Thailand data. This may be due to the structure of the GAM model and specifics of the Thailand outbreak since there were fewer case counts than in China but greater than in South Korea and Japan, and an increasing magnitude of waves during the studied period.

On the other hand, ARIMA generally outperformed GAM on the other datasets and scenarios. Those datasets can be characterized by low total case counts and big but short initial waves followed by several minor waves. While the GLM and Prophet models showed poor performance overall, the GLM model most likely had low performance due to its simplicity, while the Prophet models seemed to depend heavily on the selection of its parameters (especially those that are responsible for the seasonality), which may not always be an obvious task. Changes in calibration parameters for the Prophet model can influence its performance significantly, thus requiring a lot of tuning to perform better. Those results were obtained from low-case data and for countries that were ready for the outbreak and introduced measures to stop mpox from spreading. Since the n-Sub-epidemic and Sub-epidemic Waves frameworks showed consistently good performance on both high and low case counts, we can conclude that they are suitable for future mpox outbreak prediction, no matter their size.

By predicting the spread and intensity of the disease, public health authorities can implement targeted interventions, allocate resources efficiently, and initiate timely vaccination campaigns. Accurate forecasting enables the identification of high-risk areas and populations, thereby facilitating the deployment of preventive measures and reducing transmission rates. This proactive approach is essential in mitigating the impact of mpox and safeguarding the health of at-risk communities. That is why reliable forecasting is so crucial.

The study has the following limitations. Firstly, the analyzed period is less than a year due to the nature of the pandemic in the selected countries, so the annual seasonality of the epidemic could not possibly be established or modeled. Also, the limited number of available data points could have influenced the stability of the performance metrics. The utilization of longer time series from later periods could change model ranking, while such data were not available for the presented study. Another limitation is that all models included in the study did not in any way adjust or incorporate behavioral change and interventions into the model. That means that the inclusion of external factors outside just case counts alone could change model performance. Some models like ARIMAX can include external variables [41]. Prophet can include irregularity in the model as "holidays" [42]. However, there is no single standardized and established approach to incorporate all those changes into the modeling framework, including information about outside factors that may influence the pandemic, as well as information about the exact numbers of such factors and possible interactions between them.In conclusion, the consistently good performance of Sub-epidemic ensemble frameworks was observed for mpox incidence forecasting, and these results are not limited to mpox outbreaks. The *n*-Sub-epidemic ensemble framework was also successfully used to predict SARS-COV-2, SARS, Ebola, and Plague cases [20,27–29]. It should be noted, however, that due to the complexity of the model, this framework is computationally intensive and thus requires significant time to compute predictions. Consequently, features such as parallel computing can be utilized to improve performance.

Author Contributions: A.S.: Data curation, Formal analysis, Investigation, Code writing, Software, Writing—original draft, Writing—review and editing. A.B.: Software, Code writing, Writing—original draft. R.L.: Conceptualization, Writing—original draft. P.S.: Conceptualization, Writing—original draft. G.C.: Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing. A.K.: Conceptualization, Data curation,

Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data were obtained from WHO public sources and are publicly available using the provided link [43]. The corresponding data sources have also been provided and duplicated on GitHub [43,44]. The entire analysis source code in `R`, `MATLAB` and `Python` has also been made publicly available on GitHub [44]. The versions used for the presented analysis were v. 4.2.2 for `R`, v. R2022b for `MATLAB`, and v. 3.10.0 for `Python`.

Conflicts of Interest: The authors declare no conflicts interests.

Abbreviations

The following abbreviations are used in this manuscript:

ARIMA	auto-regressive integrated moving average
ARIMAX	autoregressive integrated moving average with explanatory variable
DNA	deoxyribonucleic acid
GAM	generalized additive model
GLM	generalized linear model
MA	moving average
MAE	mean absolute error
MSE	mean squared error
MPXV	mpox virus
PCR	polymerase chain reaction
PI	prediction interval
WHO	World Health Organization
WIS	weighted interval score

Appendix A. Model Parameters

Appendix A.1. n-Sub-Epidemic Framework

The model is fitted using non-linear least squares and assuming a normal distribution for the error structure. The uncertainty has been quantified using parametric bootstrap methods [27] using the default number of bootstraps equal to 300. The model was set to accommodate a maximum of three Sub-epidemics.

Appendix A.2. Sub-Epidemic Wave Framework

For this model, the non-linear least squares estimation method was utilized, while anormal distribution was used for the error structure. The uncertainty was quantified using parametric bootstrap methods [20] using the default number of bootstraps equal to 300. The maximum number of Sub-epidemics in the model was set to two.

Appendix A.3. ARIMA

For the presented ARIMA model implementations, parameters d were set to 2. The considered values for p were $\{0, 1, \ldots, 10\}$ and the considered values for q were $\{0, 1, \ldots, 5\}$.

Appendix A.4. Prophet

The following model parameters were used for the model. The growth trend was set as linear. Seasonality was "auto". No irregular components (i.e., holidays) were specified. All other parameters were left for the default ones used in the package `Prophet` in `R`.

Appendix A.5. GAM

For this model, normal error distribution was used. For the smoothing term, we used P-splines (value `ps` for the `bs` argument in the `mgcv::s` function) and the default dimension of the basis.

References

1. World Health Organization: WHO Declares End of Mpox Emergency, Calls for Sustained Efforts for Long-Term Management of the Disease. 2024. Available online: https://www.paho.org/en/news/11-5-2023-who-declares-end-mpox-emergency-calls-sustained-efforts-long-term-management-disease (accessed on 4 May 2024).
2. Karagoz, A.; Tombuloglu, H.; Alsaeed, M.; Tombuloglu, G.; AlRubaish, A.A.; Mahmoud, A.; Smajlović, S.; Ćordić, S.; Rabaan, A.A.; Alsuhaimi, E. Monkeypox (mpox) virus: Classification, origin, transmission, genome organization, antiviral drugs, and molecular diagnosis. *J. Infect. Public Health* **2023**, *16*, 531–541. [CrossRef] [PubMed]
3. Aden, D.; Zaheer, S.; Kumar, R.; Ranga, S. Monkeypox (Mpox) outbreak during COVID-19 pandemic—Past and the future. *J. Med. Virol.* **2023**, *95*, e28701. [CrossRef] [PubMed]
4. Multi-Country Monkeypox Outbreak: Situation Update. Available online: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON396 (accessed on 4 May 2024).
5. 2022-24 Mpox (Monkeypox) Outbreak: Global Trends. 2024. Available online: https://worldhealthorg.shinyapps.io/mpx_global/ (accessed on 4 May 2024).
6. Endo, A.; Jung, S.M.; Miura, F. Mpox emergence in Japan: Ongoing risk of establishment in Asia. *Lancet* **2023**, *401*, 1923–1924. [CrossRef]
7. S. Korea Reports First Two Monkeypox Cases—Reuters World Asia-Pacific. 22 June 2022. Available online: https://www.reuters.com/world/asia-pacific/skorea-reports-first-two-monkeypox-cases-yonhap-2022-06-21/ (accessed on 21 November 2024).
8. Thailand Confirms First Monkeypox Infection—Reuters World Asia-Pacific. 21 June 2022. Available online: https://www.reuters.com/world/asia-pacific/thailand-confirms-first-monkeypox-infection-2022-07-21/ (accessed on 21 November 2024).
9. Japan Confirms First Monkeypox Case as It Steps up Preparation for Outbreak. Japan Times. 25 July 2022. Available online: https://www.japantimes.co.jp/news/2022/07/25/national/japan-monkeypox-outbreak-preparation/ (accessed on 21 November 2024).
10. China's Chongqing City Reports One Imported Monkeypox Case—Reuters World China. 16 September 2022. Available online: https://www.reuters.com/world/china/chinas-chongqing-city-reports-one-imported-monkeypox-case-2022-09-16/ (accessed on 21 November 2024).
11. Yang, Z. China Is Suddenly Dealing with Another Public Health Crisis: Mpox. Available online: https://www.technologyreview.com/2023/08/01/1077047/china-public-health-crisis-mpox-outbreak/ (accessed on 4 May 2024).
12. Thailand Declares National Monkeypox Alert After Emergency Meeting. 2022. Available online: https://www.nationthailand.com/in-focus/40018102 (accessed on 4 May 2024).
13. The Response of the Tokyo Metropolitan Government to Mpox. Available online: https://www.hokeniryo.metro.tokyo.lg.jp/kansen/icdc/mpox.files/Mpox.pdf (accessed on 4 May 2024).
14. Kaftan, D.; Kim, H.Y.; Ko, C.; Howard, J.S.; Dalal, P.; Yamamoto, N.; Braithwaite, R.S.; Bershteyn, A. Performance analysis of mathematical methods used to forecast the 2022 New York City Mpox outbreak. *J. Med. Virol.* **2024**, *96*, e29791. [CrossRef]
15. Bleichrodt, A.; Luo, R.; Kirpich, A.; Chowell, G. Evaluating the forecasting performance of ensemble Sub-epidemic frameworks and other time series models for the 2022–2023 mpox epidemic. *R. Soc. Open Sci.* **2024**, *11*, 240248. [CrossRef]
16. Yasmin, F.; Hassan, M.M.; Zaman, S.; Aung, S.T.; Karim, A.; Azam, S. A forecasting prognosis of the monkeypox outbreak based on a comprehensive statistical and regression analysis. *Computation* **2022**, *10*, 177. [CrossRef]
17. Iftikhar, H.; Khan, M.; Khan, M.S.; Khan, M. Short-term forecasting of monkeypox cases using a novel filtering and combining technique. *Diagnostics* **2023**, *13*, 1923. [CrossRef]
18. Manohar, B.; Das, R. Artificial neural networks for the prediction of monkeypox outbreak. *Trop. Med. Infect. Dis.* **2022**, *7*, 424. [CrossRef]
19. Surveillance, Case Investigation and Contact Tracing for Monkeypox: Interim Guidance. Available online: https://www.who.int/publications/i/item/WHO-MPX-Surveillance-2024.1 (accessed on 4 May 2024).
20. Chowell, G.; Tariq, A.; Hyman, J.M. A novel Sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Med.* **2019**, *17*, 164. [CrossRef]
21. Liu, H. *Generalized Additive Model*; Department of Mathematics and Statistics University of Minnesota Duluth: Duluth, MN, USA, 2008.
22. Shumway, R.H.; Stoffer, D.S. ARIMA models. In *Time Series Analysis and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 75–163.
23. Taylor, S.J.; Letham, B. Forecasting at scale. *Am. Stat.* **2018**, *72*, 37–45. [CrossRef]
24. Bleichrodt, A.; Phan, A.; Luo, R.; Kirpich, A.; Chowell-Puente, G. StatModPredict: A User-Friendly R-Shiny Interface for Fitting and Forecasting with Statistical Models. *Stat. Model.* **2024**, *preprint*. [CrossRef]
25. Chowell-Puente, G. SubEpiPredict: A MATLAB Toolbox for Fitting and Forecasting Epidemic Trajectories Using the Ensemble n-Subepidemic Framework. 2024. Available online: https://github.com/gchowell/ensemble_n-subepidemic_framework (accessed on 4 May 2024).
26. Chowell-Puente, G. SpatialWavePredict: A MATLAB Toolbox for Fitting and Forecasting Epidemic Trajectories Using the Spatial Wave Sub-Epidemic Framework. 2024. Available online: https://github.com/gchowell/SpatialWavePredict-Toolbox (accessed on 4 May 2024).

27. Chowell, G.; Dahal, S.; Tariq, A.; Roosa, K.; Hyman, J.M.; Luo, R. An ensemble n-Sub-epidemic modeling framework for short-term forecasting epidemic trajectories: Application to the COVID-19 pandemic in the USA. *PLoS Comput. Biol.* **2022**, *18*, e1010602. [CrossRef] [PubMed]
28. Chowell, G.; Dahal, S.; Bleichrodt, A.; Tariq, A.; Hyman, J.M.; Luo, R. SubEpiPredict: A tutorial-based primer and toolbox for fitting and forecasting growth trajectories using the ensemble n-Sub-epidemic modeling framework. *Infect. Dis. Model.* **2024**, *9*, 411–436. [CrossRef] [PubMed]
29. Chowell, G.; Tariq, A.; Dahal, S.; Bleichrodt, A.; Luo, R.; Hyman, J.M. SpatialWavePredict: A tutorial-based primer and toolbox for forecasting growth trajectories using the ensemble spatial wave Sub-epidemic modeling framework. *BMC Med. Res. Methodol.* **2024**, *24*, 131. [CrossRef] [PubMed]
30. Hyndman, R.J.; Athanasopoulos, G. Forecasting: Principles and Practice. 2018. Available online: https://otexts.com/fpp2 (accessed on 4 May 2024).
31. Satrio, C.B.A.; Darmawan, W.; Nadia, B.U.; Hanafiah, N. Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and Prophet. *Procedia Comput. Sci.* **2021**, *179*, 524–532. [CrossRef]
32. Navratil, M.; Kolkova, A. Decomposition and forecasting time series in the business economy using Prophet forecasting model. *Cent. Eur. Bus. Rev.* **2019**, *8*, 26. [CrossRef]
33. Samal, K.K.R.; Babu, K.S.; Das, S.K.; Acharaya, A. Time series based air pollution forecasting using SARIMA and Prophet model. In Proceedings of the 2019 International Conference on Information Technology and Computer Communications, Singapore, 16–18 August 2019; pp. 80–85.
34. Kirpich, A.; Shishkin, A.; Weppelmann, T.A.; Tchernov, A.P.; Skums, P.; Gankin, Y. Excess mortality in Belarus during the COVID-19 pandemic as the case study of a country with limited non-pharmaceutical interventions and limited reporting. *Sci. Rep.* **2022**, *12*, 5475. [CrossRef]
35. Shishkin, A.; Lhewa, P.; Yang, C.; Gankin, Y.; Chowell, G.; Norris, M.; Skums, P.; Kirpich, A. Excess mortality in Ukraine during the course of COVID-19 pandemic in 2020–2021. *Sci. Rep.* **2023**, *13*, 6917. [CrossRef]
36. James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; Taylor, J. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.
37. Baayen, R.H.; Linke, M. An introduction to the generalized additive model. In *A Practical Handbook of Corpus Linguistics*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 563–591.
38. Botchkarev, A. Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology. *arXiv* **2018**, arXiv:1809.03006.
39. Bracher, J.; Ray, E.L.; Gneiting, T.; Reich, N.G. Evaluating epidemic forecasts in an interval format. *PLoS Comput. Biol.* **2021**, *17*, e1008618. [CrossRef]
40. Package 'Scoringutils' Documentation. Available online: https://cran.r-project.org/web/packages/scoringutils/scoringutils.pdf (accessed on 4 August 2024).
41. Peter, Ď.; Silvia, P. ARIMA vs. ARIMAX—Which approach is better to analyze and forecast macroeconomic time series. In Proceedings of the 30th International Conference Mathematical Methods in Economics, Karviná, Czech Republic, 11–13 September 2012; Volume 2, pp. 136–140.
42. Aziza, V.N.; Moh'd, F.H.; Maghfiroh, F.A.; Notodiputro, K.A.; Angraini, Y. Performance comparison of sarima intervention and Prophet models for forecasting the number of airline passenger at Soekarno-Hatta international airport. *BAREKENG J. Ilmu Mat. Dan Terap.* **2023**, *17*, 2107–2120. [CrossRef]
43. World Health Organization WHO—Detailed Case Data. 2024. Available online: https://worldhealthorg.shinyapps.io/mpx_global/#42_Trends_in_cases (accessed on 4 August 2024).
44. Shishkin, A. Mpox_Forecast: Short-Term Predictions of the Trajectory of Mpox in East Asian Countries, 2022–2023: A Comparative Study of Forecasting Approaches. 2024. Available online: https://github.com/keder/mpox_forecast (accessed on 4 August 2024).