

SARS-CoV-2 transmission dynamics in Mozambique and Zimbabwe during the First Three Years of the Pandemic

Roselyn F Kaondera-Shava, Marta Galanti, Matteo Perini, Jiyeon Suh, Shannon M Farley, Sergio Chicumbe, Ilesh Jani, Annette Cassy, Ivalda Macicame, Naisa Manafe, Wafaa El-Sadr, and Jeffrey Shaman

Abstract

The 2019 emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and its rapid spread created a public health emergency of international concern. However, the impact of the pandemic in Sub-Saharan Africa, as documented in cases, hospitalizations and deaths, appears far lower than in the Americas, Europe, and Asia. Characterization of the transmission dynamics is critical for understanding how SARS-CoV-2 spread and the true scale of the pandemic. Here, to better understand SARS-CoV-2 transmission dynamics in two southern African countries, Mozambique and Zimbabwe, we developed a dynamic model-Bayesian inference system to estimate key epidemiologic parameters, namely the transmission and ascertainment rates. Total infection burdens (reported and unreported) during the first three years of the pandemic were reconstructed using a model-inference approach. Transmission rates rose with each successive wave, which aligns with observations in other continents. Ascertainment rates were found to be low matching prior assumption and consistent with other African countries. Overall, the estimated disease burden was higher than the documented cases, indicating need for improved reporting and surveillance. These findings aid understanding of COVID-19 disease and respiratory virus transmission dynamics in two African countries little investigated to date, and can help guide future public health planning and control strategies.

Keywords: SARS-CoV-2, COVID-19, VOC, dynamic modeling, Bayesian inference, transmission rate, ascertainment rate, disease burden

1 Introduction

Following its emergence in China in late December 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (COVID-19) spread globally and produced an unprecedented public health crisis [44, 49]. The World Health Organization (WHO) declared the COVID-19 outbreak a pandemic on March 11th, 2020 [8]. In less than three years, in various parts of the world, including two southern African countries, i.e. Mozambique and Zimbabwe, SARS-CoV-2 triggered multiple waves of new infections, often driven by new variants of concern (VOC) [30].

The COVID-19 pandemic put enormous pressure on public health security; to mitigate COVID-19 spread, most African countries imposed travel restrictions or banned international travel and implemented curfews, lockdowns and other social distancing interventions beginning in March and April 2020 [31]. The initial cases of COVID-19 in Mozambique and Zimbabwe were reported on March 22nd and March 20th 2020 [20, 35], respectively, and the pandemic subsequently spread to all provinces in both countries. By December 2022, three waves of infection, i.e., Alpha, Delta, and Omicron (BA.1), had occurred.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Characterization of transmission dynamics is critical for understanding how SARS-CoV-2 spread and the full extent of its impact on local populations. Statistical and modeling studies have provided important insights into disease transmission dynamics by estimating critical epidemiological parameters (such as the reproductive number, ascertainment rates, and total infections) in various studies. For example, the time-varying reproductive number characterizes virus transmissibility at a particular point in time; this quantity has been estimated using stochastic models coupled with Bayesian inference for France and Ireland [6], Switzerland [32], and Scandinavian countries [19]. Oduro *et al.* [28] used a time varying reproduction estimation approach for sub-Saharan Africa and estimated the time-varying reproductive number while also accounting for depletion of the susceptible population. In addition, a model-inference approach was used to estimate the background population characteristics such as population susceptibility for South Africa [46], the United States at county resolution [34], and India [45].

Many SARS-CoV-2 infections are unreported; hence, confirmed case counts do not fully reflect epidemic dynamics. Understanding the level of ascertainment and the impact of undocumented cases on transmission is crucial to plan and implement control strategies. Models coupled with Bayesian inference methods have been used to investigate the role of undocumented infections in China [22], as well as for various regions (countries and US states) [7], 54 African countries [18], and 3 North African countries: Algeria, Egypt, and Morocco [11]. Russel *et al.* [36] fitted a Bayesian Gaussian process model to estimate under-ascertainment in 210 countries and territories.

In spite of this vast body of literature, characterization of SARS-CoV-2 transmission dynamics (or any respiratory virus) in the Southern African region remains mostly lacking, except for South Africa [46, 40, 48, 25]. An autoregressive integrated moving average (ARIMA) model was utilized to forecast the trend of the disease in four African countries which reported the most cases: South-Africa, Egypt, Nigeria and Ghana [23] and in Southern Africa [38]. Shoko *et al.* [39] used a support vector regression for short-term forecasting of the disease in Zimbabwe. However, the reported COVID-19 impact (cases, hospitalizations, and deaths) in Africa has likely underestimated the actual extent of infection and thus the transmission dynamics. As a consequence, investigation of key epidemiological characteristics over time and by VOC wave is needed.

Here we utilize a model-inference framework to study COVID-19 disease dynamics in Mozambique and Zimbabwe during the first three years of the pandemic, for the three VOCs Alpha, Delta and Omicron (BA.1). The dynamic model was coupled with a Bayesian inference method and province-specific case data. We accounted for undocumented cases to estimate key epidemiologic parameters, namely the transmission rate and the ascertainment rate at the end of each variant outbreak. The findings help advance understanding of the dynamics of SARS-CoV-2 and give information on the transmission dynamics of respiratory viruses in two African nations that have received little attention to date.

2 Methodology

2.1 Epidemic transmission model

To describe the prevalent characteristics of COVID-19 in each province, i , for Mozambique and Zimbabwe, we used a Susceptible-Exposed-Infectious-Recovered (SEIR) compartmental model, similar to ones used in recent studies on COVID-19 [42, 41, 22, 16]. The infected compartment was split into two compartments; infected reported (I_i^r), individuals who have tested positive, and infected unreported (I_i^u), individuals who are infected with the virus but have not been tested and thus not reported. The susceptible population, S_i , is not yet

infected, the exposed population, E_i , has been infected with the virus but do not yet transmit the virus. There is no reinfection of the recovered population, R_i , and N_i is the total population.

The model (see Figure 1) was coupled with data assimilation and case data in order to estimate the transmission rate, β_i and the ascertainment rate, α_i , at the end of each variant wave outbreak. The other parameters, i.e., average latency period, Z , average duration of infection, D and reduction of infection rate for unreported infected individuals, μ were kept constant (see Table S5). We also maintained a constant human population for the study period with no births or deaths.

2.1.1 Model assumptions

The model is based on the following assumptions:

- (i) there is no human movement between the provinces
- (ii) individuals who are infected (reported and unreported) can transmit the virus to those susceptible,
- (iii) there is no re-infection during each variant wave period.

2.1.2 Disease dynamics

The system of ordinary differential equations was adapted from Li *et al.* [22]. Figure 1 presents a schematic of model transmission dynamics, and a description of the model variables and parameters is shown in Table S3.

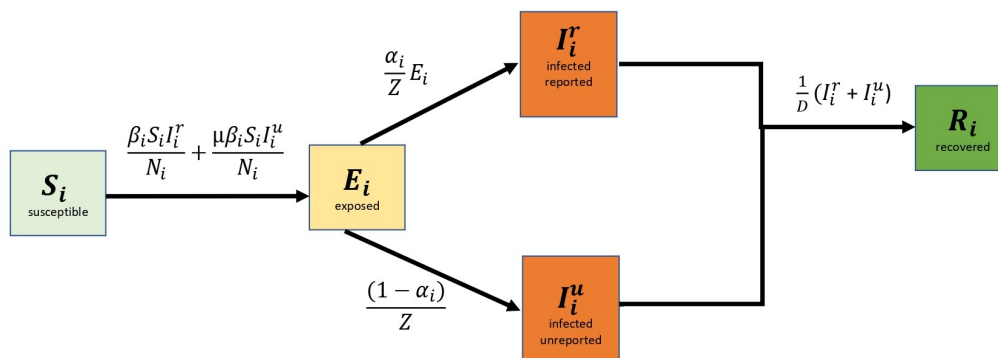


Figure 1: Schematic presentation of simulated disease dynamics. The solid arrows demonstrate the movement from one disease state to another.

Model disease dynamics are described by a system of ordinary differential equations, Equation (S1)–Equation (S4) which are subject to the non-negative initial conditions and satisfy the conservation of the population, i.e., $N_i = S_i + E_i + I_i^r + I_i^u + R_i$ for each province i . The reproductive number at a particular time, $Rt_i = \alpha_i \beta_i D + (1 - \alpha_i) \mu \beta_i D$ was obtained through the next generation approach (see Section S1.3 and Section S1.4 for detailed explanation).

In our analysis, the SEIR model was integrated stochastically forward in time. A Latin hypercube sampling (LHS) technique was used to select a random set of initial variables and parameter combinations from prior ranges, which are assumed to be uniformly distributed. The initial priors for the variables and parameters are given in Table S5. A 4th-order Runge Kutta (RK4) scheme was used for stochastic integration; in particular, to introduce stochasticity to the right hand side of the system, a random sample from a Poisson distribution, was used for

each step of the RK4 scheme to determine each term as shown in Equation (S6)–Equation (S9).

For both countries all provinces were run simultaneously as a metapopulation model without representation of human movement between provinces, due to the absence of robust mobility data. Attempted simulations using a gravity model to estimate inter-provincial movement did not yield credible results (described in Section S1.9). As a consequence, we assumed disease transmission dynamics to be principally determined by local within-province processes.

2.1.3 Piecewise Simulation and Inference

The metapopulation model form was coupled with the ensemble adjustment Kalman filter (EAKF), a Bayesian sequential ensemble filtering method, and province-specific COVID-19 case data. We assume no re-infection during each wave; piecewise simulation and parameter estimation were performed separately for each variant wave: Alpha, Delta and Omicron (BA.1). The data periods for this piecewise inference are shown in Table S4. Thus, for each province i , we produced parameter estimates for the transmission rate, β_i and ascertainment rate, α_i , for each variant wave period. The procedure for parameter estimation using the EAKF is summarized in Algorithm S1.

The model is reinitialized for each wave and does not explicitly represent re-infections. In Mozambique, for the Delta and Omicron waves, the final S_i estimated range of the prior wave (Alpha and Delta, respectively) was used as the susceptible initial prior, e.g. the *final* S_i range of Alpha wave = *initial prior* S_i range of Delta wave (see Table S5). In Zimbabwe, because data were only available for the Omicron (BA.1) wave, *initial prior* S_i was set to that of Mozambique for the same outbreak, as both countries are in the same region and experienced similar epidemic patterns.

The depletion of susceptibles, S_i at the end of an outbreak, reflected an approximation of the total new infections (both reported and unreported), given ongoing filter adjustment. Consequently, to estimate the cumulative disease burden from the beginning of the pandemic across subsequent waves, computation was carried out as follows:

$$\begin{aligned} \text{proportion of total new infections} &\approx \frac{\text{cumulative new}(I_i^r + I_i^u)}{N_i} \\ &\approx \text{cumulative loss in } S_i \\ &\approx 1 - \frac{S_i}{N_i} \text{ (at end of outbreak)} \end{aligned} \quad (1)$$

2.2 The SEIR-EAKF framework

The EAKF assimilation algorithm has been used in previous studies of infectious disease dynamics to assimilate observations and conduct inference [10, 33, 37]. When coupled with the dynamic model, Equation (S1)–Equation (S4), the combined SEIR-EAKF system iteratively optimizes the distribution of model state variables and parameters whenever new observations become available [3].

In general, sequential ensemble filtering is the problem of estimating the probability of the system state at a given time \mathbf{x}_t conditional on observations O_t taken up until and inclusive of time t . After model initialization, the system integrates an ensemble of simulations forward in time to compute the prior distribution for the model state variables, parameters, and the model-simulated case data. At the time of observation, the system is halted and the EAKF is used to update the state variables and parameters based on those model-generated prior

estimates and case data. The EAKF algorithm is applied with prescribed observational error variance (OEV), see Section 2.2.3. In this fashion, the ensemble simulations of the observed state variables (incidence) are updated to align with observations (see Anderson [3] for algorithmic details). The updates (posteriors) are determined by computing the Kalman gain using the latest observation and the distribution of current model states (the prior). The unobserved state variables and model parameters are then adjusted by the EAKF using cross ensemble co-variability. The process is then repeated, with the posterior integrated to the next observation. Through this iterative optimization process, the ensemble of model simulations is better aligned to simulate current outbreak dynamics and estimate key parameters.

Here, a 300-member ensemble of simulations using the SEIR model Equation (S1)–Equation (S4) was coupled with the EAKF and case data. The state vector at time t is: $\mathbf{x}_t = (\mathbf{s}_t, \theta_t)$, where $\mathbf{s}_t = [S_i, E_i, I_i^r, I_i^u]$ is the vector of the local state variables at time t and $\theta_t = [\mu, Z, D, \beta_i, \alpha_i]$ is the vector of model parameters. Using observation O_t^i at time t , the posterior distribution of the system state is derived by applying Bayes' rule (posterior \propto prior \times likelihood) to incorporate the new information. The EAKF deterministically adjusts the prior distribution to a posterior using Bayes rule while assuming a Gaussian distribution for both the prior and likelihood, which allows estimation of the first two moments (mean and covariance), leaving the higher-order moments unchanged. Bayes' rule provides a target for updating the system state given an observation:

$$p(\mathbf{x}_t | O_{1:t}^i) \propto p(\mathbf{x}_t | O_{1:t-1}^i) p(O_{1:t}^i | \mathbf{x}_t).$$

We used the daily number of new reported cases in province i on a given day t , O_t^i , as observations. Unobserved variables, such as the susceptible population S , and model parameters, were adjusted in accordance with covariant relationships with the observed variable i.e. reported new infections, $\left(\alpha_i \frac{E_i}{Z}\right)$ that arise naturally as a result of the dynamics of the system (see Anderson [3] for algorithmic details).

2.2.1 State variables and parameter updates

State variables and the inferred model parameters (β_i and α_i) were updated locally for each province. Estimating the parameters locally was motivated by the discrepancies found between population size and total incidence: Nampula, the most densely populated province in Mozambique, reported the fewest COVID-19 cases, whereas Maputo Cidade, the least populated province, reported the highest number of cases across all variant waves. This discrepancy may be due to the limited healthcare accessibility in Nampula and lower population density compared to Maputo Cidade, which has the highest level of healthcare access and greatest population density among Mozambique provinces [12]. Similarly, for Zimbabwe, Harare province is the most densely populated province with better healthcare facilities compared to other provinces. Such differences underpinned an expectation of different contact and ascertainment rates and motivated deriving local β_i and α_i estimates for both countries.

The EAKF prediction–update cycle is performed sequentially, and an update is triggered by the arrival of new data (daily in this study) [3]. For the observable state variables, the i^{th} ensemble member is updated by Equation (S13), while, for the unobserved variables/parameters x^i , the i^{th} ensemble member is updated by Equation (S14).

2.2.2 Initial priors

The initial 300-member ensemble of state variables and parameters were randomly drawn from a uniform distribution using Latin hypercube sampling and the initial range of values presented in Table S5; other parameters, D , Z , and μ were fixed. During filtering, ensemble members

could move outside initial ranges; in such instances, the individual ensemble member would be resampled from the prior range.

The initial prior range of $\alpha_i = [0.004 - 0.1]$ (which is low, corresponding to a maximum of one reported case for every ten infected) was motivated by the study of Han *et al.* [18], which estimated an overall daily report rate (ascertainment rate) ranging from $\alpha = [0.0002 - 0.3]$, among all African countries. Furthermore, the study by Evans *et al.* [17] highlighted that the lower-than-expected case burdens in Madagascar were explained solely by detection rates of 0.1–1% (or $\alpha = [0.001 - 0.01]$). On the other hand, due to the heterogeneity of reported cases across provinces in Mozambique, the initial prior lower bound for α_i was varied by province for both countries (see Table S5). For province i , we applied the formula: lower bound $\alpha_i = (\text{total reported cases})/0.5N_i$, assuming 50% of the population was infected at the end of the three variant waves (Alpha, Delta, Omicron (BA.1)). An exception was Maputo Cidade which had a different initial prior range $\alpha_i = [0.14, 0.2]$, because, despite being the less populous province it reported the highest number of cases. This may be attributed to Maputo Cidade being the capital city with greater access to healthcare and testing facilities compared to the rest of the country. Generally, the remaining Mozambique provinces had lower bounds ranging $\alpha_i = [0.004 - 0.03]$, whereas for Zimbabwe, lower bounds were $\alpha_i = [0.01 - 0.04]$. The S_i initial priors varied for each wave because the model was run piecewise for the periods as described above and in Table S5. Due to missing Zimbabwe COVID-19 provincial case data for earlier waves, the initial S prior for the Omicron (BA.1) variant wave was based on that of Mozambique.

2.2.3 Observational error variance

Using the EAKF requires specification of error for both the simulated model output and observations. The model error may be easily computed as the variance of the 300-member ensemble. For each O_t^i , we specify the OEV at time t as:

$$OEV_t^i = \max\left(4, \frac{(O_t^i)^2}{4}\right). \quad (2)$$

An OEV of this form has been used successfully for inference and forecasting for infectious diseases such as influenza [47, 33] and COVID-19 [22].

2.2.4 Filter divergence

As successive observations are assimilated, there is a tendency for the variance between the ensemble members to decrease due to repeated filter adjustment. This may potentially lead to filter divergence, in which the ensemble error variance is so minimal that the fitting process essentially ignores the observations [3]. To prevent filter divergence, the prior ensemble was inflated by a multiplicative factor $\lambda = 0.015$, before each daily assimilation and calculation of the posterior (see Equation (S15)). The inflation was applied to all state variables and estimated parameters, i.e. β_i (transmission rate) and α_i (ascertainment rate).

2.3 Synthetic Testing

To validate EAKF inference with the SEIR model, we generated a synthetic, model-simulated COVID-19 outbreak (defined as the 'truth'), defined by model Equation (S1)–Equation (S4) and we tested the ability of the model/filter framework to identify the true model state variables and parameters. Specifically, we generated nine synthetic datasets using different scenarios combinations of transmission rates, $\beta_i = (0.8, 1.3, 1.8)$ and ascertainment rates, $\alpha_i = (0.01, 0.04, 0.09)$, where the chosen parameter values were the same across all provinces for each outbreak, as shown in Table S6. Synthetic observations of new daily reported cases

(reported new infections, $\left(\alpha_i \frac{E_i}{Z}\right)$) were then generated by adding normally distributed random observational error calculated as described in Equation (2). The resulting time series of synthetic error-laden observational records were smoothed using a 7-day moving average, and then used for assimilation in the combined SEIR-EAKF framework to determine whether the system can reliably identify known parameter values.

2.4 Parameter estimation using real data

In order to estimate the transmission rate, β_i and ascertainment rate, α_i , for each variant wave (Alpha, Delta and Omicron (BA.1)), we fitted the model to province specific case data (i.e., number of new reported cases) as reported by Mozambique Ministry of Health-SIGILIA and DISA [29] and Zimbabwe COVID-19 hub [1] for the period between March 2020 to December 2022. Due to missing Zimbabwe COVID-19 provincial data we only estimated parameters for the Omicron (BA.1) variant wave in Zimbabwe.

The SEIR-EAKF framework was applied in isolation to the eleven Mozambique provinces and, separately, the ten Zimbabwe provinces, as illustrated in Figure S1. The Mozambique COVID-19 data was transformed from weekly cases to daily cases using a linear interpolation method, and piecewise parameter estimation was carried out for the model's local parameters, β_i and α_i , for each variant wave as described in Table S4 and Section S1.7.

3 Results

3.1 Data description

COVID-19 data

The reported daily COVID-19 cases used for this study span three variant waves (Alpha, Delta, and Omicron (BA.1)) during the first three years of the pandemic, i.e. from March 30th 2020 to December 26th 2022 for Mozambique and from November 18th 2021 to January 26th 2022 for Zimbabwe. The available Zimbabwe provincial data began in the middle of the Delta variant wave, therefore, we only utilized the data covering the Omicron (BA.1) wave. At the provincial scale, COVID-19 daily cases were obtained from the Zimbabwe COVID-19 hub [1] and weekly cases for Mozambique were obtained from the Mozambique Ministry of Health-SIGILIA and DISA [29]. Heat maps of daily and weekly counts of COVID-19 cases, for each variant period, in each province of both countries are presented in Figure 2 while plots of the disaggregated time series of COVID-19 infections for the provinces are reported in Figure S1.

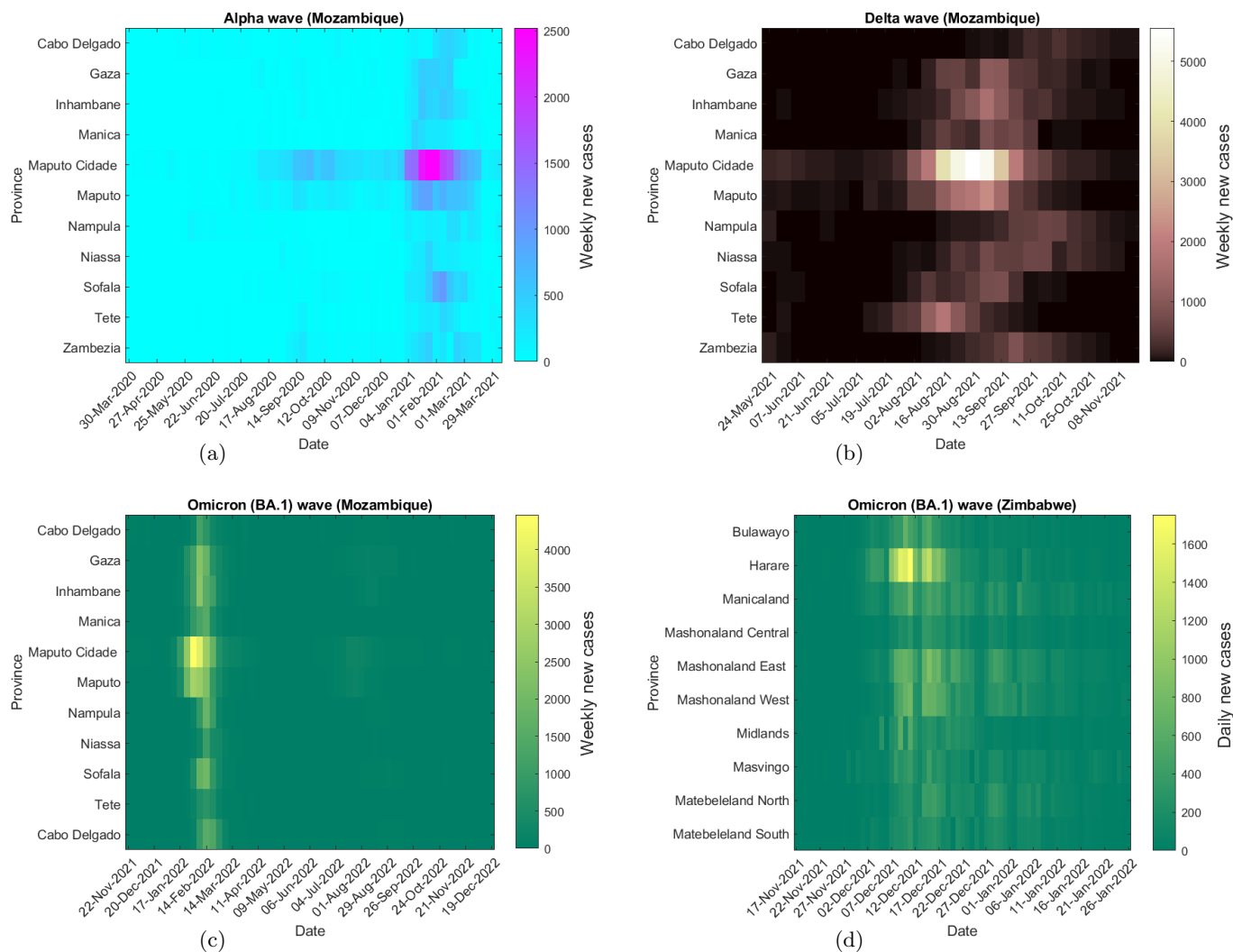


Figure 2: Heat maps of daily and weekly provincial level COVID-19 infections in Mozambique and Zimbabwe between March 2020 and December 2022, based on data released by the Mozambique Ministry of Health-SIGILIA and DISA [29] and Zimbabwe COVID-19 hub [1].

Population data

Provincial population data were taken from the Zimbabwe National Statistics Agency [2] and Instituto Nacional de Estatística Moçambique [9] for Zimbabwe and Mozambique, respectively.

3.2 Validation of the model-inference system using synthetic outbreaks

The synthetic observations of COVID-19 cases, along with their defined OEV, were used to test whether the SEIR-EAKF system could estimate unobserved state variables and parameters (β_i and α_i) accurately. Creffig:paramsSynth depicts the parameter estimates for the simulation period; the posterior parameters estimates (β_i and α_i) converged near their target values and the reproductive number at the end of the variant wave, Rt_i was correctly estimated. The ensemble posterior distributions at the end of the outbreak were well constrained for all the generated outbreaks as shown in Figure S3. The results indicate that an ensemble of size 300 was sufficient to capture the target 'true' parameters.

3.3 Mozambique parameter estimates

The estimates for β_i and α_i at the end of each variant outbreak, along with their associated 95% credible intervals (CrIs) are presented in Table 1. With the exception of Maputo Cidade, for all provinces, the mean estimates across all three waves were in the range $\alpha_i = [0.08 - 0.09]$. On the other hand, for Maputo Cidade, with its unique α_i prior range of $[0.14, 0.2]$ (see Section S1.7), the estimate was around $\alpha_i = 0.17$ across all variant waves. The local β_i mean estimates and by extension Rt_i increased with each successive wave: the mean estimate ranges of β_i were $[0.6 - 0.7]$ ($Rt_i = [1.4 - 1.8]$), for the Alpha wave, slightly higher, $\beta_i = [0.6 - 1.4]$ ($Rt_i = [1.4 - 4]$), for the Delta wave, and highest for the Omicron (BA.1) wave, $\beta_i = [1.1 - 1.7]$ ($Rt_i = [3 - 5]$).

Table 1: EAKF estimated parameters and credible intervals at the end of each Mozambique variant period.

Province	Parameter	Prior	EAKF estimate (Alpha)		EAKF estimate (Delta)		EAKF estimate (Omicron (BA.1))	
			mean	95% CrI	mean	95% CrI	mean	95% CrI
transmission rates								
Cabo Delgado	β_1	[0.5, 2.0]	0.6038	(0.5032, 0.8037)	1.1571	(0.7873, 1.6191)	1.1528	(1.0684, 1.4876)
Gaza	β_2		0.5643	(0.5036, 0.7021)	1.4886	(0.6507, 1.9467)	1.5630	(0.6812, 1.9322)
Inhambane	β_3		0.5830	(0.5040, 0.7898)	1.2970	(0.6593, 1.9414)	1.6998	(0.9363, 1.9593)
Manica	β_4		0.5526	(0.5021, 0.6768)	0.6362	(0.5073, 1.3342)	1.4411	(0.6655, 1.9614)
Maputo Cidade	β_5		0.6681	(0.5157, 0.9132)	1.1154	(0.5938, 1.9074)	1.5664	(0.7519, 1.9561)
Maputo	β_6		0.6136	(0.5135, 0.8306)	1.3520	(0.6281, 1.9099)	1.5299	(0.8293, 1.9683)
Nampula	β_7		0.5645	(0.5035, 0.6679)	0.9902	(0.6117, 1.7501)	1.1174	(0.5324, 1.9316)
Niassa	β_8		0.6059	(0.5026, 0.7684)	0.7763	(0.5156, 1.5475)	1.5890	(0.7134, 1.9788)
Sofala	β_9		0.6060	(0.5046, 0.8032)	0.5680	(0.5035, 0.9613)	1.3762	(0.7603, 1.9652)
Tete	β_{10}		0.6082	(0.5036, 0.8303)	0.6315	(0.5107, 1.1309)	1.1819	(0.5908, 1.9180)
Zambezia	β_{11}		0.5796	(0.5055, 0.7656)	0.8914	(0.5229, 1.7148)	1.3361	(0.6657, 1.9511)
ascertainment rates								
Cabo Delgado	α_1	[0.007, 0.1]	0.0923	(0.0744, 0.0993)	0.0833	(0.0158, 0.0990)	0.0798	(0.0301, 0.0983)
Gaza	α_2	[0.02, 0.1]	0.0883	(0.0606, 0.0998)	0.0622	(0.0264, 0.0978)	0.0718	(0.0316, 0.0991)
Inhambane	α_3	[0.03, 0.1]	0.0904	(0.0554, 0.0998)	0.0627	(0.0330, 0.0965)	0.0881	(0.0570, 0.0991)
Manica	α_4	[0.01, 0.1]	0.0922	(0.0739, 0.0996)	0.0881	(0.0426, 0.0987)	0.0836	(0.0203, 0.0994)
Maputo Cidade	α_5	[0.14, 0.2]	0.1739	(0.1419, 0.1988)	0.1760	(0.1477, 0.1990)	0.1729	(0.1411, 0.1977)
Maputo	α_6	[0.03, 0.1]	0.0836	(0.0558, 0.0985)	0.0640	(0.0334, 0.0986)	0.0677	(0.0379, 0.0970)
Nampula	α_7	[0.004, 0.1]	0.0911	(0.0658, 0.0995)	0.0318	(0.0174, 0.0754)	0.0867	(0.0163, 0.0987)
Niassa	α_8	[0.01, 0.1]	0.0930	(0.0684, 0.0998)	0.0812	(0.0371, 0.0992)	0.0888	(0.0403, 0.0979)
Sofala	α_9	[0.01, 0.1]	0.0872	(0.0519, 0.0997)	0.0849	(0.0455, 0.0995)	0.0827	(0.0176, 0.0998)
Tete	α_{10}	[0.009, 0.1]	0.0902	(0.0684, 0.0995)	0.0862	(0.0468, 0.0996)	0.0787	(0.0468, 0.0996)
Zambezia	α_{11}	[0.005, 0.1]	0.0902	(0.0543, 0.0996)	0.0884	(0.0320, 0.0990)	0.0879	(0.0513, 0.0996)
reproductive number								
Cabo Delgado	Rt_1		1.5367	(1.2760, 2.0486)	2.9193	(1.9882, 4.0874)	2.9056	(1.3940, 4.8801)
Gaza	Rt_2		1.4317	(1.2738, 1.7767)	3.6994	(1.6392, 4.8517)	3.9199	(1.7308, 4.8177)
Inhambane	Rt_3		1.4748	(1.2827, 2.0140)	3.2424	(1.6336, 4.8583)	4.3049	(2.3873, 4.9700)
Manica	Rt_4		1.4100	(1.2765, 1.7275)	1.6184	(1.2840, 3.3278)	3.6355	(1.6712, 4.9141)
Maputo Cidade	Rt_5		1.7843	(1.3755, 2.4536)	2.9981	(1.5922, 5.1157)	4.1683	(1.9782, 5.2456)
Maputo	Rt_6		1.5550	(1.3017, 2.0922)	3.3779	(1.5619, 4.8312)	3.8328	(2.1190, 4.9489)
Nampula	Rt_7		1.4404	(1.2798, 1.7002)	2.4349	(1.5105, 4.2809)	2.7798	(1.3296, 4.8345)
Niassa	Rt_8		1.5362	(1.2773, 1.9420)	1.9631	(1.3020, 3.8910)	4.0304	(1.8178, 4.9980)
Sofala	Rt_9		1.5218	(1.2829, 2.0447)	1.4373	(1.2688, 2.4461)	3.4886	(1.9383, 4.9815)
Tete	Rt_{10}		1.5395	(1.2799, 2.1200)	1.5965	(1.2932, 2.8883)	2.9824	(1.5107, 4.8073)
Zambezia	Rt_{11}		1.4710	(1.2825, 1.9245)	2.2548	(1.3290, 4.3444)	3.4098	(1.6867, 4.9622)

Figure S4–Figure S6 illustrate the time evolution of the parameter posteriors (means and 95% CrIs) for the Alpha, Delta and Omicron (BA.1) variant waves, respectively, plotted against their prior range. The initial drift in the first month, is likely a reflection of the time needed for the EAKF to learn the system [3]. The α_i estimates at the end of the outbreaks were closer to the upper bound of 0.2 for Maputo Cidade and 0.1 for the rest of the provinces. For the Alpha and Delta waves β_i converges to a solution during January 2021 and August 2021, respectively, as cases rise, on the other hand, for the Omicron (BA.1) wave, β_i convergence to a solution is less clear, with a broader posterior and higher mean estimate.

The posterior fits of new reported cases to province-specific actual data [29] were generally good for all provinces in Mozambique for all variant waves, as demonstrated in Figure S7–Figure S9. The model-inference framework was able to reproduce even complicated outbreak structures characterized by multiple peaks such as during the Alpha wave in Nampula,

Tete and Zambezia.

In order to check the consistency of the parameter estimates for each variant wave, presented in Table 1, SEIR-EAKF system simulations were repeated 20 times. The mean parameter estimates (β_i and α_i), and the corresponding mean Rt_i for the runs are illustrated in Figure 3 and show consistency for most estimates. Figure S12 (a)-(c) shows the posterior mean estimates of susceptibility % at the end of each variant wave. An estimated 75 – 95% of the population remained susceptible at the end of the Alpha wave (in March 2021), furthermore, 15 – 50% and 5 – 30% remained susceptible after the Delta (in October 2021) and Omicron (BA.1) (in March 2022) waves, respectively.

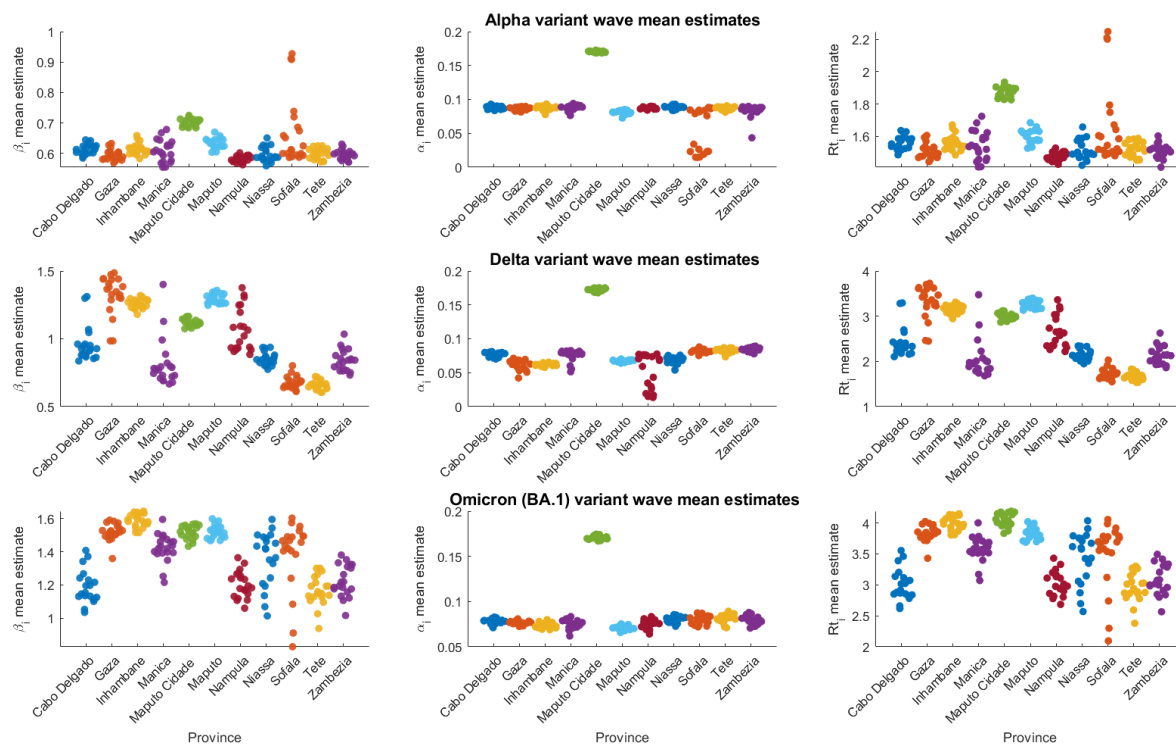


Figure 3: Distribution of mean parameter estimates (α_i and β_i , and the corresponding Rt_i) for the 11 Mozambique provinces at the end of variant waves (Alpha, Delta and Omicron (BA.1)) for 20 runs.

To quantify the burden of COVID-19 infections at the end of each variant period, we estimated the total cumulative infections, i.e., both reported and unreported infections by applying Equation (1). Figure 4(a)–Figure 4(c) display the spatial distribution of the estimated total new infections (reported and unreported) across all Mozambique provinces. We estimated an increase in the proportion of the infected population across the three VOC (Alpha, Delta and Omicron (BA.1)): the Alpha wave estimates were below 20% for the majority of the country; in contrast, Maputo Cidade and Maputo, showed the highest burdens of 72% and 57%, respectively. The cumulative infection estimates increased from 6 – 72% (Alpha wave), to 46 – 88% (Delta wave), to 74 – 95% (Omicron BA.1 wave).

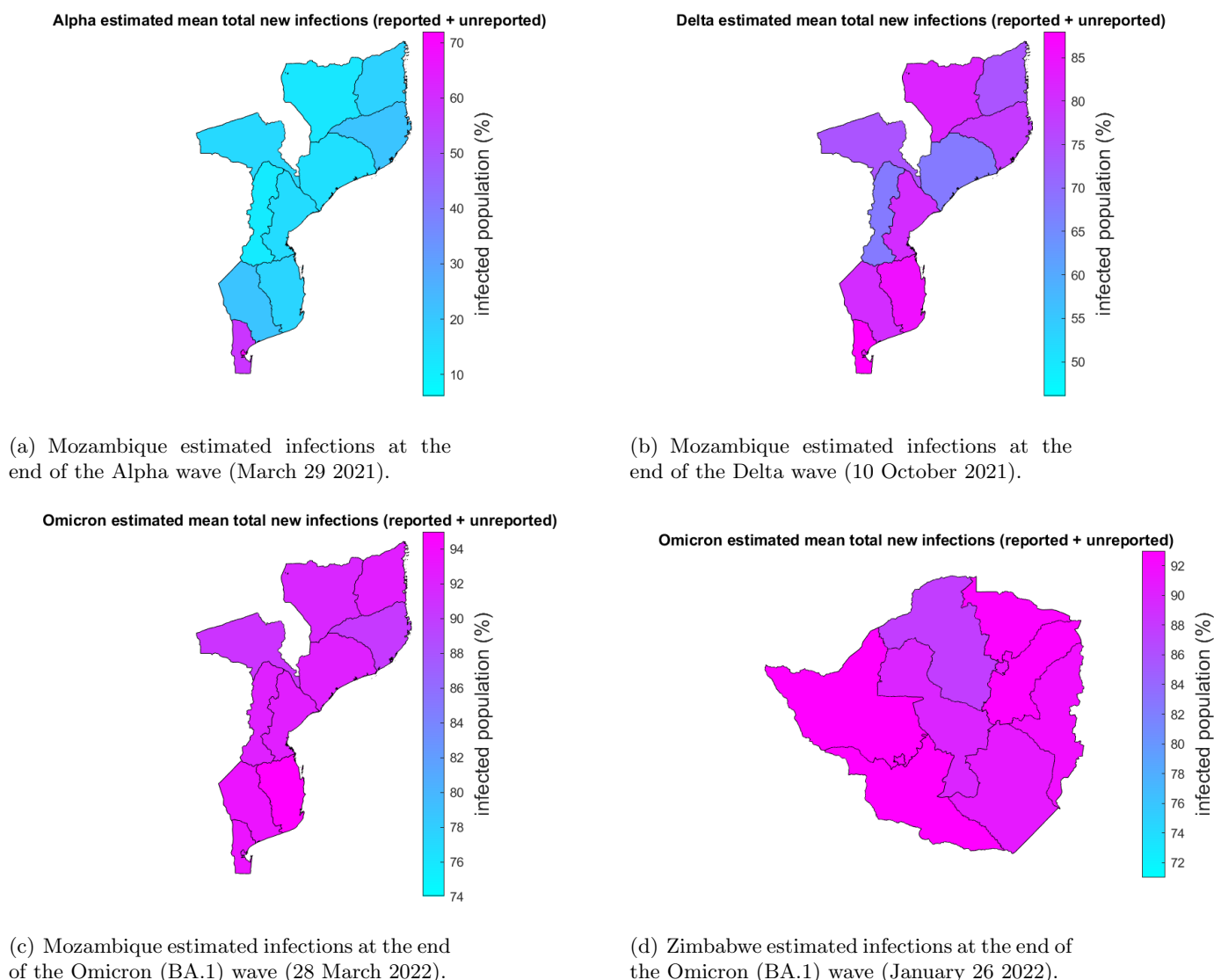


Figure 4: Spatial distribution of estimated SARS-CoV-2 infections (both reported and unreported) per VOC for Mozambique and Zimbabwe provinces between March 2020 and December 2022 [29, 1]. Colors toward red indicate more infections.

3.4 Zimbabwe parameter estimates

Table 2 presents the parameters estimates (β_i and α_i) and associated 95% CrIs at the end of the Omicron (BA.1) variant outbreak. The estimates for the ascertainment rate were in the range $\alpha_i = [0.08 - 0.09]$, which is in line with Mozambique estimates (shown in Table 1). On the other hand, the mean estimates for the transmission rate ranged from $\beta_i = [1.1 - 1.7]$ for all provinces with a reproductive number of $Rt_i = [3 - 5]$. The time evolution of α_i and β_i parameter posteriors (means and 95% CrI) over the course of the Omicron (BA.1) wave, plotted against their prior range is demonstrated in Figure S10. The α_i estimates drifted higher towards the upper bound (0.1) during December 2021, a level maintained until the end of the outbreak. The convergence of β_i to a solution is less clear, with a broader posterior and higher mean estimate.

Figure S11 demonstrates a good model fit to provincial data [1] for the outbreak with the posterior estimates capturing the outbreak peaks for all ten provinces. The consistency of the parameter estimates Table 2 was evaluated by running the model 20 times, and the distribution

of the mean parameter estimates (β_i and α_i) at the end of the outbreak, and the corresponding mean Rt_i for the runs are shown in Figure 5. Figure S12 (d) shows the posterior mean estimates of susceptibility % at the end of the Omicron (BA.1) wave (in January 2022), where 15 – 40% of the population remained susceptible at the end of that outbreak.

Table 2: EAKF estimated parameters and credible intervals for the Zimbabwe Omicron (BA.1) variant period

Province	Parameter	Prior	EAKF estimate (Omicron (BA.1))	
			mean	95% CrI
transmission rates				
Bulawayo	β_1	[0.5, 2.0]	1.6529	(1.0081, 1.9813)
Harare	β_2		1.6407	(1.0550, 1.9762)
Manicaland	β_3		1.1327	(0.5971, 1.9040)
Mashonaland Central	β_4		1.2506	(0.7189, 1.9148)
Mashonaland East	β_5		1.5438	(0.8128, 1.9457)
Mashonaland West	β_6		1.4040	(0.8200, 1.9323)
Midlands	β_7		1.3669	(0.6988, 1.9702)
Masvingo	β_8		1.5457	(0.8012, 1.9652)
Matebeleland North	β_9		1.6567	(1.0512, 1.9716)
Matebeleland South	β_{10}		1.6446	(0.9344, 1.9848)
ascertainment rates				
Bulawayo	α_1	[0.04, 0.1]	0.0911	(0.0668, 0.0997)
Harare	α_2	[0.03, 0.1]	0.0924	(0.0744, 0.0997)
Manicaland	α_3	[0.02, 0.1]	0.0864	(0.0387, 0.0990)
Mashonaland Central	α_4	[0.01, 0.1]	0.0929	(0.0589, 0.0997)
Mashonaland East	α_5	[0.03, 0.1]	0.0871	(0.0562, 0.0984)
Mashonaland West	α_6	[0.02, 0.1]	0.0930	(0.0694, 0.0997)
Midlands	α_7	[0.01, 0.1]	0.0917	(0.0534, 0.0993)
Masvingo	α_8	[0.02, 0.1]	0.0876	(0.0474, 0.0989)
Matebeleland North	α_9	[0.04, 0.1]	0.0857	(0.0499, 0.0992)
Matebeleland South	α_{10}	[0.03, 0.1]	0.0846	(0.0489, 0.0991)
reproductive number				
Bulawayo	Rt_1		5.4010	(3.3035, 6.4817)
Harare	Rt_2		5.3709	(3.4402, 6.4669)
Manicaland	Rt_3		3.6866	(1.9242, 6.2350)
Mashonaland Central	Rt_4		4.0797	(2.3522, 6.2600)
Mashonaland East	Rt_5		5.0435	(2.6502, 6.3627)
Mashonaland West	Rt_6		4.6015	(2.6877, 6.3179)
Midlands	Rt_7		4.4749	(2.2919, 6.4159)
Masvingo	Rt_8		5.0560	(2.0240, 3.0989)
Matebeleland North	Rt_9		5.4064	(3.4283, 6.4578)
Matebeleland South	Rt_{10}		5.3699	(3.0497, 6.4785)

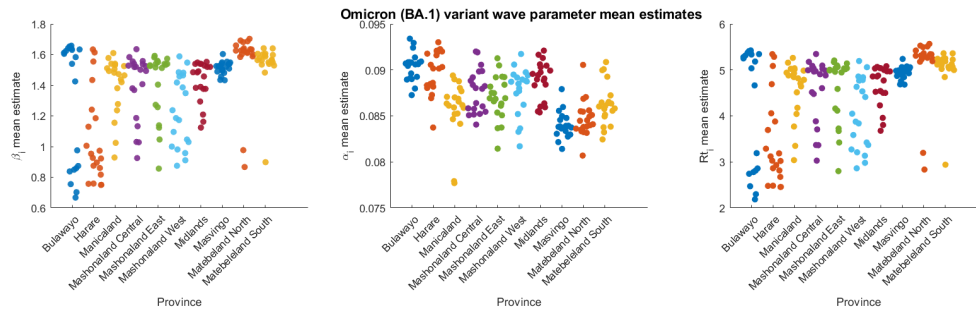


Figure 5: Distribution of mean parameter estimates (α_i , β_i , and corresponding Rt_i) for the 10 Zimbabwe provinces at the end of Omicron (BA.1) wave for 20 runs.

The spatial distribution of the estimated mean total new COVID-19 infections (both reported and unreported) at the end of the Omicron (BA.1) wave is summarized in Figure 4(d). The findings indicate that 71 – 93% of the population had been infected.

4 Discussion

Although the emergence of COVID-19 and its rapid spread created a public health emergency of international concern, the impact of the pandemic in Sub-Saharan Africa, as documented in cases, hospitalizations and deaths, appears far lower than in the Americas, Europe, and Asia [43]. We utilized a model-inference framework to shed light on disease dynamics in two southern African countries, i.e., Mozambique and Zimbabwe, during the first three years of the pandemic. While accounting for undocumented cases, we estimated key epidemiologic parameters namely the transmission rate, and the ascertainment rate at the end of each outbreak. Our study provides several insights into the disease burden of infections (reported and unreported) for three VOCs, i.e. Alpha, Delta and Omicron (BA.1). While we focus on the Mozambique and Zimbabwe case, our framework can be applied to other African countries given the low documented impact of the disease in this region.

The ascertainment rate, α , estimates were very low for Mozambique and Zimbabwe provinces, matching our prior assumption. Our estimates in both countries is in line with the studies by Han *et al.* [18] and Evans *et al.*[17]. This indicates that the majority of infections were never documented as cases; by extension, many severe cases requiring hospitalization or resulting in death may have been missed due to limited access to healthcare and testing facilities. The estimated low ascertainment rate highlights the need for enhanced reporting and surveillance mechanisms, with special emphasis on Sub-Saharan Africa. Mozambique experienced natural disasters and conflict which hampered access to healthcare facilities: Prior to the pandemic, the country was devastated by two consecutive deadly and destructive cyclones (Idai and Kenneth), which caused massive destruction to infrastructure; consequently, accessibility in some places became more difficult. Furthermore, conflict in northern Mozambique produced additional challenges during the pandemic by displacing many people, hence making it difficult to control the disease [26].

The transmission rate estimates, and by extension reproductive number, rose with each successive wave. This matches observations in the Americas, Africa, Europe and Asia that the Omicron variant, in particular, was more transmissible than the preceding Delta variant [21, 24], which in turn had greater transmissibility than the preceding Alpha variant [15, 27]. Mozambique and Zimbabwe deployed public health and social measures to mitigate the spread of COVID-19. These included travel restrictions, lockdown, social distancing measures, compulsory mask wearing, contact tracing and testing, school closures and use of personal protective equipment among health workers [4, 5, 13, 14].

Our model was fitted to provincial reported cases. We find that representing and accounting for unreported cases across provinces is crucial for estimating the true disease burden, i.e. new infections (reported and unreported), at the end of each variant wave. The low numbers of confirmed cases in Africa were generally a poor indicator of true incidence of infection, which may be attributed to limited testing capacity and health system access. Overall, the estimated disease burden in Mozambique and Zimbabwe was much higher than the documented reported cases for each VOC (see [29, 1], Table S1, and Table S2). The estimates highlight that by the end of the Omicron (BA.1) wave, almost the entire population had been infected, which is in agreement with the estimates of 93% cumulative total infection after the same wave in South Africa [46]. Our results indicate that in African countries, efforts must be made to improve the timely reporting and surveillance of public health threats.

There are limitations to our model: Even though human mobility plays a crucial role in the transmission of SARS-CoV-2, the unavailability of inter-provincial human movement data for both Mozambique and Zimbabwe was a drawback. We explored utilizing a gravity

model to estimate human movement between the provinces. This approach, however, did not yield credible results; hence, we implemented the metapopulation model without movement. Additionally, there was limited provincial case data for Zimbabwe, which covered only the Omicron (BA.1) variant. Furthermore, the model does not explicitly account for re-infections, but the filter adjusts the susceptibles, S_i , thereby implicitly accounting for increasing susceptibility due to immune escape.

In conclusion, we developed an inference-based transmission model to aid in understanding the evolving dynamics of SARS-CoV-2 in Mozambique and Zimbabwe. By taking unreported cases into account, we estimated key epidemiologic characteristics, i.e. the transmission rate and ascertainment rate. This approach and findings are relevant for countries with less comprehensive surveillance systems. The findings of this study on the disease burden can help guide future public health planning. In particular, they shed light on respiratory virus transmission dynamics in two African countries little investigated to date.

References

- [1] African-Surveyors-Connect. Zimbabwe-covid-19-data, 2022. URL https://github.com/African-Surveyors-Connect/Zimbabwe-COVID-19-Data/blob/master/time_series_data/daily_provincial_records.csv.
- [2] Zimbabwe National Statistics Agency. Provinces, 2022. URL <https://www.citypopulation.de/en/zimbabwe/cities/>.
- [3] Jeffrey L Anderson. An ensemble adjustment Kalman filter for data assimilation. *Monthly weather review*, 129(12):2884–2903, 2001. doi: [https://doi.org/10.1175/1520-0493\(2001\)129%3C2884:AEAKFF%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129%3C2884:AEAKFF%3E2.0.CO;2).
- [4] Imen Ayouni, Jihen Maatoug, Wafa Dhouib, Nawel Zammit, Sihem Ben Fredj, Rim Ghammam, and Hassen Ghannem. Effective public health measures to mitigate the spread of COVID-19: a systematic review. *BMC public health*, 21(1):1015, 2021. doi: <https://doi.org/10.1186/s12889-021-11111-1>.
- [5] Rosário Betho, Marcia Chelengo, Sam Jones, Michael Keller, Ibraimo Hassane Mussagy, Dirk van Seventer, and Finn Tarp. The macroeconomic impact of COVID-19 in Mozambique. *WIDER Working Papers 2021.(93): 38 pp. many ref.*, 2021. doi: <https://doi.org/10.35188/UNU-WIDER/2021/033-7>.
- [6] Bernard Cazelles, Clara Champagne, Benjamin Nguyen-Van-Yen, Catherine Comiskey, Elisabeta Vergu, and Benjamin Roche. A mechanistic and data-driven reconstruction of the time-varying reproduction number: Application to the COVID-19 epidemic. *PLoS computational biology*, 17(7):e1009211, 2021. doi: <https://doi.org/10.1371/journal.pcbi.1009211>.
- [7] Carson C Chow, Joshua C Chang, Richard C Gerkin, and Shashaank Vattikuti. Global prediction of unreported SARS-CoV2 infection from observed COVID-19 cases. *MedRxiv*, 2020. doi: <https://doi.org/10.1101/2020.04.29.20083485>.
- [8] Domenico Cucinotta and Maurizio Vanelli. WHO declares COVID-19 a pandemic. *Acta bio medica: Atenei parmensis*, 91(1):157, 2020. doi: <https://doi.org/10.23750/2Fabm.v91i1.9397>.
- [9] Instituto Nacional de Estatística Moçambique. Provinces, 2022. URL <https://www.citypopulation.de/en/mozambique/cities/>.

- [10] Nicholas B DeFelice, Eliza Little, Scott R Campbell, and Jeffrey Shaman. Ensemble forecast of human west nile virus cases and mosquito infection rates. *Nature Communications*, 8(1): 14592, 2017. doi: <https://doi.org/10.1038/ncomms14592>.
- [11] Salih Djilali, Lahbib Benahmadi, Abdessamad Tridane, and Khadija Niri. Modeling the impact of unreported cases of the COVID-19 in the North African countries. *Biology*, 9(11):373, 2020. doi: <https://doi.org/10.3390/biology9110373>.
- [12] António dos Anjos Luis and Pedro Cabral. Geographic accessibility to primary healthcare centers in mozambique. *International journal for equity in health*, 15:1–13, 2016. doi: <https://doi.org/10.1186/s12939-016-0455-0>.
- [13] Tafadzwa Dzinamarira, Solomon Mukwenha, Rouzeh Eghtessadi, Diego F Cuadros, Gibson Mhlanga, and Godfrey Musuka. Coronavirus disease 2019 (COVID-19) response in Zimbabwe: a call for urgent scale-up of testing to meet national capacity. *Clinical Infectious Diseases*, 72(10):e667–e674, 2021. doi: <https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1301/5899551>.
- [14] Mathias Dzobo, Itai Chitungo, and Tafadzwa Dzinamarira. Covid-19: a perspective for lifting lockdown in Zimbabwe. *The Pan African Medical Journal*, 35(Suppl 2), 2020. doi: [10.11604/pamj.2020.35.2.23059](https://doi.org/10.11604/pamj.2020.35.2.23059).
- [15] Rebecca Earnest, Rockib Uddin, Nicholas Matluk, Nicholas Renzette, Sarah E Turbett, Katherine J Siddle, Christine Loreth, Gordon Adams, Christopher H Tomkins-Tinch, Mary E Petrone, et al. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA. *Cell Reports Medicine*, 3(4), 2022. doi: <https://doi.org/10.1016/j.xcrm.2022.100583>.
- [16] Ralf Engbert, Maximilian M Rabe, Reinhold Kliegl, and Sebastian Reich. Sequential data assimilation of the stochastic seir epidemic model for regional covid-19 dynamics. *Bulletin of mathematical biology*, 83(1):1, 2021. doi: <https://doi.org/10.1007/s11538-020-00834-8>.
- [17] Michelle V Evans, Andres Garchitorea, Rado JL Rakotonanahary, John M Drake, Benjamin Andriamihaja, Elinambinina Rajaonarifara, Calistus N Ngonghala, Benjamin Roche, Matthew H Bonds, and Julio Rakotonirina. Reconciling model predictions with low reported cases of covid-19 in sub-saharan africa: insights from madagascar. *Global Health Action*, 13(1):1816044, 2020. doi: <https://doi.org/10.1080/16549716.2020.1816044>.
- [18] Qing Han, Nicola Bragazzi, Ali Asgary, James Orbinski, Jianhong Wu, and Jude Dzevela Kong. Estimation of epidemiological parameters and ascertainment rate from early transmission of COVID-19 across Africa. *Royal Society Open Science*, 10(9):230316, 2023. doi: <https://doi.org/10.1098/rsos.230316>.
- [19] Agus Hasan, Hadi Susanto, Venansius Tjahjono, Rudy Kusdiantara, Endah Putri, Nuning Nuraini, and Panji Hadisoemarto. A new estimation method for COVID-19 time-varying reproduction number using active cases. *Scientific reports*, 12(1):6675, 2022. doi: <https://doi.org/10.1038/s41598-022-10723-w>.
- [20] The Herald. Zimbabwe confirms first COVID-19 case, 2020. URL <https://www.herald.co.zw/zimbabwe-confirms-first-covid-19-case/>.
- [21] Christoph Jung, Dorota Kmiec, Lennart Koepke, Fabian Zech, Timo Jacob, Konstantin MJ Sparrer, and Frank Kirchhoff. Omicron: what makes the latest SARS-CoV-2 variant of concern so concerning? *Journal of virology*, 96(6):e02077–21, 2022. doi: <https://doi.org/10.1128/jvi.02077-21>.

- [22] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2). *Science*, 368(6490):489–493, 2020. doi: <https://doi.org/10.1126/science.abb3221>.
- [23] Adewale F Lukman, Rauf I Rauf, Oluwakemi Abiodun, Olajumoke Oludoun, Kayode Ayinde, and Roseline O Ogundokun. COVID-19 prevalence estimation: Four most affected African countries. *Infectious Disease Modelling*, 5:827–838, 2020. doi: <https://doi.org/10.1016/j.idm.2020.10.002>.
- [24] Smriti Mallapaty. COVID-19: How Omicron overtook Delta in three charts. *Nature*, 10, 2022.
- [25] Rendani Mbuyha and Tshilidzi Marwala. Bayesian inference of COVID-19 spreading rates in South Africa. *PloS one*, 15(8):e0237126, 2020. doi: <https://doi.org/10.1371/journal.pone.0237126>.
- [26] Vánio A Mugabe, Eduardo S Gudo, Osvaldo F Inlamea, Uriel Kitron, and Guilherme S Ribeiro. Natural disasters, population displacement and health emergencies: multiple public health threats in Mozambique. *BMJ global health*, 6(9):e006778, 2021. doi: <https://doi.org/10.1136/bmjgh-2021-006778>.
- [27] United Nations. Delta variant drives Africa COVID threat to ‘whole new level’: WHO warns; ‘dominant’ in Europe by August, 2021. URL <https://news.un.org/en/story/2021/07/1095152#>.
- [28] Michael Safo Oduro, Seth Arhin-Donkor, Louis Asiedu, Damazo T Kadengye, and Samuel Iddi. SARS-CoV-2 incidence monitoring and statistical estimation of the basic and time-varying reproduction number at the early onset of the pandemic in 45 sub-Saharan African countries. *BMC Public Health*, 24(1):612, 2024. doi: <https://doi.org/10.1186/s12889-024-18184-8>.
- [29] Ministry of Health-SIGILIA and DISA. Covid-19 cases. Mozambique Ministry of Health.
- [30] World Health Organization. Tracking SARS-CoV-2 variants, 2024. URL <https://www.who.int/activities/tracking-SARS-CoV-2-variants/>.
- [31] World Health Organization et al. A systematic approach to monitoring and analysing public health and social measures (PHSM) in the context of the COVID-19 pandemic: underlying methodology and application of the PHSM database and PHSM Severity Index. Technical report, World Health Organization. Regional Office for Europe, 2020.
- [32] Damiano Pasetto, Joseph C Lemaitre, Enrico Bertuzzo, Marino Gatto, and Andrea Rinaldo. Range of reproduction number estimates for COVID-19 spread. *Biochemical and biophysical research communications*, 538:253–258, 2021. doi: <https://doi.org/10.1016/j.bbrc.2020.12.003>.
- [33] Sen Pei, Sasikiran Kandula, Wan Yang, and Jeffrey Shaman. Forecasting the spatial transmission of influenza in the United States. *Proceedings of the National Academy of Sciences*, 115(11):2752–2757, 2018. doi: <https://doi.org/10.1073/pnas.1708856115>.
- [34] Sen Pei, Teresa K Yamana, Sasikiran Kandula, Marta Galanti, and Jeffrey Shaman. Burden and characteristics of COVID-19 in the United States during 2020. *Nature*, 598(7880):338–341, 2021. doi: <https://doi.org/10.1038/s41586-021-03914-4>.
- [35] National Post. Mozambique confirms first coronavirus case, 2020. URL <https://nationalpost.com/pmnm/health-pmn/mozambique-confirms-first-coronavirus-case>.

- [36] Timothy W Russell, Nick Golding, Joel Hellewell, Sam Abbott, Lawrence Wright, Carl AB Pearson, Kevin van Zandvoort, Christopher I Jarvis, Hamish Gibbs, Yang Liu, et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. *BMC medicine*, 18(1):332, 2020. doi: <https://doi.org/10.1186/s12916-020-01790-9>.
- [37] J Shaman, A Karspeck, W Yang, J Tamerius, and M Lipsitch. Real-time influenza forecasts during the 2012-2013 season. *Nat Commun.*, 4:2837, 2013. doi: <https://doi.org/10.1038/ncomms3837>.
- [38] Claris Shoko and Peter Njuho. ARIMA model in predicting of COVID-19 epidemic for the southern Africa region. *African journal of infectious diseases*, 17(1):1–9, 2023. doi: <https://doi.org/10.21010/Ajid>.
- [39] Claris Shoko and Caston Sigauke. Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data. *American Journal of Infection Control*, 51(10):1095–1107, 2023. doi: <https://doi.org/10.1016/j.ajic.2023.03.010>.
- [40] Sheetal Prakash Silal, Juliet RC Pulliam, Gesine Meyer-Rath, Lise Jamieson, Brooke E Nichols, Jared Norman, Rachel Hounsell, Saadiyah Mayet, Frank Kagoro, and Harry Moultrie. The National COVID-19 Epi Model (NCEM): Estimating cases, admissions and deaths for the first wave of COVID-19 in South Africa. *PLOS Global Public Health*, 3(4): e0001070, 2023. doi: <https://doi.org/10.1371/journal.pgph.0001070>.
- [41] Deshun Sun, Li Duan, Jianyi Xiong, and Daping Wang. Modeling and forecasting the spread tendency of the COVID-19 in China. *Advances in Difference Equations*, 2020:1–16, 2020. doi: <https://doi.org/10.1186/s13662-020-02940-2>.
- [42] Kangkang Wan, Jing Chen, Changming Lu, Lanlan Dong, Zhicheng Wu, and Lianglu Zhang. When will the battle against novel coronavirus end in Wuhan: A SEIR modeling analysis. *Journal of global health*, 10(1), 2020. doi: <https://doi.org/10.7189%2Fjogh.10.011002>.
- [43] Worldometers.info. Reported cases and deaths by country or territory, 2024. URL <https://www.worldometers.info/coronavirus/>.
- [44] Aiping Wu, Yousong Peng, Baoying Huang, Xiao Ding, Xianyue Wang, Peihua Niu, Jing Meng, Zhaozhong Zhu, Zheng Zhang, Jiangyuan Wang, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell host & microbe*, 27(3):325–328, 2020. doi: [10.1016/j.chom.2020.02.001](https://doi.org/10.1016/j.chom.2020.02.001).
- [45] Wan Yang and Jeffrey Shaman. COVID-19 pandemic dynamics in India, the SARS-CoV-2 Delta variant and implications for vaccination. *Journal of the Royal Society Interface*, 19(191):20210900, 2022. doi: <https://doi.org/10.1098/rsif.2021.0900>.
- [46] Wan Yang and Jeffrey L Shaman. COVID-19 pandemic dynamics in South Africa and epidemiological characteristics of three variants of concern (Beta, Delta, and Omicron). *Elife*, 11, 2022. doi: <https://doi.org/10.7554/eLife.78933>.
- [47] Wan Yang, Alicia Karspeck, and Jeffrey Shaman. Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLoS computational biology*, 10(4):e1003583, 2014. doi: <https://doi.org/10.1371/journal.pcbi.1003583>.
- [48] Zebin Zhao, Xin Li, Feng Liu, Gaofeng Zhu, Chunfeng Ma, and Liangxu Wang. Prediction of the COVID-19 spread in African countries and implications for prevention and control: A case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. *Science of the Total Environment*, 729:138959, 2020. doi: <https://doi.org/10.1016/j.scitotenv.2020.138959>.

- [49] Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798):270–273, 2020. doi: 10.1038/s41586-020-2012-7.