Mathematical Biosciences and Engineering

https://www.aimspress.com/journal/mbe

MBE, 22(10): 2686–2719.
DOI: 10.3934/mbe.2025099
Received: 13 March 2025
Revised: 28 June 2025
Accepted: 14 July 2025
Published: 26 August 2025

Research article

Autopsy of SARS-CoV-2 spread dynamics in Ecuador using data assimilation techniques: A tale of two provinces

Paula Castro[1,2] **and Juan Carlos De los Reyes**[1,2,*]

[1] Departamento de Matemática, Escuela Politécnica Nacional, Quito, Ecuador

[2] MODEMAT Research Center in Mathematical Modeling and Optimization, Quito, Ecuador

* **Correspondence:** Email: juan.delosreyes@epn.edu.ec.

Abstract: In this article, we considered a Bayesian approach to estimating the evolution of the COVID-19 pandemic in Ecuador, providing the first rigorous analysis of its progression in the country. Specifically, we applied variational data assimilation to estimate the parameters and initial conditions of a compartmental SARS-CoV-2 propagation model while accounting for structural data uncertainty through error covariance matrices. These optimized parameters correspond to maximum-a-posteriori (MAP) estimates, which, in a second stage, allow us to infer the posterior distribution of the parameters. We considered two different data sources: the official count of positive COVID-19 tests from the Ecuadorian Public Health Ministry (MSP) and an estimate of COVID-19-related deaths derived from excess mortality data recorded by the Ecuadorian Civil Registry (RC). We regard RC data as the closest approximation to the actual number of COVID-19 cases. The results highlight that, although there are differences between the estimates obtained using MSP data–generated in real time during the pandemic–and those based on positive cases inferred from excess mortality, the trends in the computed effective reproduction numbers coincide. This suggests that the methodology presented in this paper, and applied in real time during the pandemic, was able to accurately capture the evolution of the pandemic in Ecuador. Additionally, we conducted a comparative analysis of Ecuador's two most populous provinces, Pichincha and Guayas, which experienced the pandemic very differently, particularly in its initial stages. This study aimed to improve our understanding of the virus's spread in these provinces and provide insights into how epidemiological dynamics can vary within the same country.

Keywords: data assimilation; epidemiological modeling; COVID-19; parameter estimation; Bayesian inference

1. Introduction

A new type of coronavirus (SARS-CoV-2) appeared in November 2019, and due to its extensive spread, the World Health Organization (WHO) declared a pandemic on March 11, 2020 [1]. Ecuador reported its first confirmed case of SARS-CoV-2 infection on February 29th, 2020, and by April 2020, it became one of the worst-hit countries in the region [2]. Guayaquil, the capital city of Guayas, was the most affected city in Ecuador [3]. In response to the escalating crisis, the Research Center in Mathematical Modeling and Optimization (MODEMAT) started studying and tracking the transmission of the SARS-CoV-2 virus in Ecuador using mathematical models, see [4].

In Ecuador, the primary source of information was the official count of COVID-19-positive cases reported by the Public Health Ministry (MSP). Unfortunately, the lack of widespread testing and non-coordinated strategies between official entities in collecting and reporting cases increased the data uncertainty, which was exceptionally high compared to other countries in the region. Relying on just one source of information could in fact lead to biased results (see, e.g., [5]). To evaluate this, we consider the mortality excess registered during the COVID-19 pandemic as another source of reliable information. In this respect, even though not all excess deaths are directly related to the pandemic, there exists a positive correlation between excess mortality from all causes and COVID-19 deaths; see, e.g., [6]. To obtain a good estimation of excess mortality, we modeled the time series of deaths before the pandemic using a seasonal autoregressive integrated moving average (SARIMA) model. Using this, we forecasted for the following months, which was then compared with the registered deaths to derive an estimation of COVID-19-related deaths from the difference.

Both data sources have their strengths and weaknesses. The lack of widespread tests makes the Public Health Ministry (MSP) data less reliable. Conversely, mortality data from the Ecuadorian Civil Registry (RC) is generally considered less uncertain. However, it requires additional time to be consolidated, with a lag of up to four weeks between the occurrence of a death and its inclusion in the official records. For our study, we consider excess mortality and the number of COVID-19-positive cases inferred from it as the closest estimate of the true number of COVID-19 cases in the country.

To address the uncertainty inherent in COVID-19 data, we employ variational data assimilation techniques in conjunction with a compartmental susceptible-exposed-infectious-removed (SEIR) model. This model distinguishes between reported (symptomatic and documented) and unreported (asymptomatic or mild and undocumented) infections, thereby enhancing the capacity to capture the dynamics of SARS-CoV-2 transmission (see, for example, [7–9] for further insights). Variational data assimilation is used to estimate both the unknown parameters and the initial conditions of the SEIR model, accounting for observational uncertainty via error covariance matrices.

Notably, unlike other parameter-estimation approaches applied to COVID-19 modeling (e.g., [10, 11]), our method performs joint estimation of parameters and initial conditions—an approach known to yield more robust results in other applied fields. A related methodology for COVID-19 modeling, although based on Bayesian filtering rather than variational data assimilation, is found in [12], where the authors infer model parameters and initial conditions for a SEIR-type model, requiring, however, a reliable and consistent stream of epidemiological data.

After obtaining an initial estimate of the parameters by solving the variational data assimilation problem, we perform a Laplace approximation to infer an approximate posterior of the model parameters. The observational data required are sourced from two different datasets: the Ecuadorian Public

Health Ministry (MSP) and the Civil Registry (RC). It is important to note that the observational data, considered closer to the true values, are those retrieved from this last source. Therefore, the predictions of the *reported infected* population obtained with this data will be regarded as the closest representation of the actual pandemic state at any given time.

In this study, we compare the results obtained using data from the Public Health Ministry (MSP) with those derived from the Civil Registry (RC) to assess the accuracy of our model's predictions, which were used during the pandemic to inform health policy decisions. We note that the estimates made during the pandemic were based on real-time MSP data, whereas the second set of estimates could only be performed retrospectively, once the Civil Registry data had been processed and published—often weeks or even months later.

Our results reveal how the COVID-19 pandemic unfolded in Ecuador, based on a rigorous methodological framework. To the best of the authors' knowledge, there is only one other study, in Ecuador, regarding the evolution of the pandemic in the first three months since the first case was unveiled, which exclusively utilizes official data on reported cases from MSP (with a high degree of uncertainty) and a simple probabilistic model [13]. Since the evolution of the pandemic in Ecuador exhibits significant geographic heterogeneity, in this paper, we focused on a comparative analysis of the pandemic's progression in the provinces of Guayas and Pichincha. We chose these provinces for three primary reasons: First, due to their population, as they are the most populous provinces in the country. Second, because of the different socioeconomic conditions that each of the two presents (living conditions, hospital infrastructure, overcrowding, etc.). Third, because of their distinct geographical features, with Guayas situated on the coast and Pichincha in the Andes at an altitude of around 2500 meters above sea level.

This document is structured as follows: In Section 2, we analyze the data and discuss the necessary preprocessing steps, including the calculation of expected mortality to determine excess mortality. Section 3 delves into the variational data assimilation process for estimating initial conditions and parameters for the compartmental SARS-CoV-2 model, outlining the specific problem, its mathematical treatment, and numerical solution. Section 4 showcases the results achieved through the proposed methodology, illustrating the COVID-19 case evolution in the provinces of Guayas and Pichincha taking into account the different data sources. Finally, in Section 5, we offer a comprehensive discussion, draw conclusions, and address limitations stemming from our work.

2. Data preparation

We analyze data from March 2020, when the pandemic hit the country, to August 2021, when the vaccination process started for most of the Ecuadorian population. We used two data sources: the official count of COVID-19-positive cases and the mortality data. We describe them below.

- **COVID-19 positive cases:** This data comes from the official statistics of confirmed COVID-19 cases reported by the Ecuadorian Public Health Ministry (MSP). The considered individuals are those who, due to their symptomatology, have undergone a PCR test with a positive result.
- **COVID-19 deaths:** The mortality data is retrieved from the Ecuadorian Civil Registry (RC) [14]. Although this data is assumed to be more reliable, it takes between three and four weeks to be consolidated by the authorized governmental entity; therefore, it is not possible to analyze it in real-time. Below, we describe the processes to first estimate the mortality excess and, second, the

observed documented infected population from the Civil Registry data.

2.1. *Mortality excess estimation via the SARIMA forecast model*

To calculate the COVID-19 deaths, we compute the excess of mortality, which refers to the difference between the total number of deaths in a crisis event, such as the COVID-19 pandemic, compared to those expected under normal conditions. Expected mortality can be computed using various forecasting methods, including time series or regression-based methods; see [15] and the references therein. In this work, we employ a seasonal autoregressive integrated moving average (SARIMA) model to forecast expected deaths from February 4, 2020 to September 2, 2021. We use the historical data of daily deaths from January 2, 2017 to February 3, 2020. The final 182 days of the historical period, corresponding to 16% of the data, were used to verify the accuracy of the prediction model.

Selecting appropriate values for the parameters of the $SARIMA(p,d,q)(P,D,Q,s)$ forecast model requires careful analysis of the time series data, including checks for stationarity and examination of autocorrelation structures. Based on this analysis, which involved applying differencing, evaluating ACF/PACF plots, and comparing models using the Akaike Information Criterion (AIC), the final model selected for forecasting expected deaths was $SARIMA(1,1,2)(1,1,0,52)$. The detailed derivation, justification for parameter selection, diagnostic validation of this model, and an explanation of the parameter notation are provided in Appendix A.

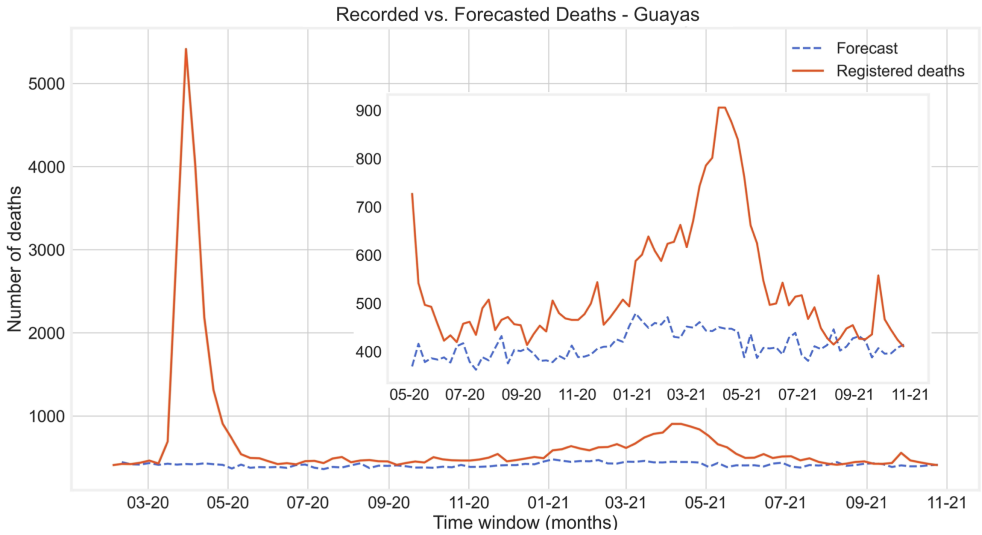


Figure 1. Comparison between recorded and forecasted deaths in Guayas from February 2020 to October 2021. A zoomed image of the results from May 2020 to October 2021 is shown inside.

Subsequently, the validated model was used to generate forecasts of expected deaths. By subtracting these forecasted values from the registered deaths recorded by the Civil Registry, we computed the excess mortality. As an example, we depict the registered (solid line) vs. the forecasted deaths (dotted line) for Guayas in Figure 1. The mortality excess from February 3rd, 2020 to July 23rd, 2021, computed as the difference between the forecasted and the registered deaths on those days, was 15,174 (for Pichincha) and 23,672 (for Guayas). These values significantly surpass the COVID-19 deaths reported

by the Ecuadorian Health Ministry for these provinces during the same period, which indicated 6750 for Pichincha and 5995 for Guayas [16]. It is noteworthy that during the early months of the pandemic, numerous reports highlighted the underreporting of COVID-19 deaths in Ecuador, prompting the government to suspend disaggregated information at the provincial level from July 26th onwards.

The significant gap observed between forecasted and recorded deaths may indicate that the real impact of COVID-19 in Ecuador was much worse than what was reported by the Ecuadorian official means. Underreporting could have been caused by inadequate testing and the misattribution of COVID-19 deaths to other causes. Moreover, besides the differences between mortality data obtained with both data sources, there is a notorious difference between the mortality excess registered in Guayas and Pichincha, Ecuador's two most populated provinces, with approximately 4,387,434 and 3,228,233 inhabitants, respectively [17].

2.2. Observed documented infected population from mortality excess

This study uses two different sources for observational data on COVID-19 infections: the official positive case counts from the Ecuadorian Public Health Ministry (MSP) and an estimate derived from excess mortality data recorded by the Ecuadorian Civil Registry (RC). The reason for incorporating RC data stems from the significant uncertainties associated with the official MSP case counts, particularly due to testing limitations.

To obtain a potentially more robust, albeit retrospective, measure of the infection burden less susceptible to testing biases, we use excess mortality data. This subsection outlines the methodology used to transform the estimated excess mortality (calculated as described in Appendix A) into an estimate of the documented infected population, which serves as the observational input (I^{obs}) for our data assimilation framework when using the Civil Registry data stream, allowing for a comparative analysis against results obtained using MSP data. The core of this transformation relies on the case fatality rate (CFR).

According to [18], the case fatality rate (CFR) represents the proportion of individuals diagnosed with a disease who eventually die from it. Once a pandemic has ended, the CFR can be calculated by dividing the total number of known deaths by the total number of confirmed cases; see, e.g., [18, 19]. However, estimating this ratio during an ongoing pandemic requires a delay correction to account for the time between case confirmation and death. As noted by [19], a real-time CFR can be calculated as

$$CFR = \frac{m(X)}{I(X - T)},$$

where $m(X)$ represents the number of deaths at day X and $I(X - T)$ denotes the number of cumulative cases (or infected population) at day $X - T$. T represents the time from case confirmation to death. In our specific case, we consider 15 days for this period as an average of those reported on the MIDAS Online COVID-19 portal [20]. We remark that assuming a fixed average duration ($T = 15$ days) is indeed an approximation. In reality, the time from diagnosis to death is highly variable and depends on numerous factors that our model does not capture at the individual level.

Given that our study was conducted in hindcast, we already have the CFR values. Therefore, we can estimate the number of deaths, m, on day X as the number of cumulative cases on day $(X-T)$, multiplied by the case fatality rate, i.e., $m = CFR \times I$. Furthermore, since the infected population is divided into reported and unreported cases, with approximately one-third of infections being asymptomatic, we can

obtain estimates for the reported and unreported case fatality rates. This allows us to refine the formula as follows:

$$m = 66.86\%\ I\ CFR_r + 33.14\%\ I\ CFR_u, \tag{2.1}$$

where the subindices r and u denote if the rates correspond to the reported and unreported cases, respectively. In Eq (2.1), the approximated one-third proportion between symptomatic and asymptomatic cases is taken as 0.3314, which corresponds to the average value reported in the literature [21–23]. We remark that the fraction of reported and unreported cases is highly context-dependent. It depends on test availability and socio-economic factors, among others.

The case fatality rate (CFR) of the reported infected population varies by region and time period. For instance, according to the US National Epidemiological Surveillance System, the CFR for reported COVID-19 cases from February 29, 2020 to December 6, 2021, was 2.17% for Pichincha and 6.84% for Guayas; see [24]. Additionally, according to Yang and coauthors [25], asymptomatic patients were younger than the symptomatic ones. Consequently, a plausible estimate for the CFR of the undocumented infected population can be derived from the rate assigned to the younger population (0–19 years old), which in Ecuador is 0.24% [24].

It is important to note that excess mortality does not distinguish between deaths caused directly by COVID-19 and those resulting from other causes; see, e.g., [26]. Nonetheless, the analysis remains valid, as there is evidence of a positive correlation between COVID-19-related deaths and all-cause excess mortality [6].

3. Variational data assimilation problem

Variational data assimilation aims to determine the most accurate estimate of a dynamical system's state by minimizing a cost functional that quantifies the disparity between the model predictions and observed data. Its formulation is rooted in a Bayesian framework and may account for nonlinear dynamics, as well as integrate observations within a specified time interval.

In brief, if we have knowledge of the *prior* probability density function of the model variables x, denoted as $p(x)$, and the conditional probability density function of the observations y given x, represented as $p(y|x)$, then we can derive the *posterior* distribution of the unknown model variables x given the observations y, expressed as $p(x|y)$, by applying Bayes' formula:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}. \tag{3.1}$$

If the observations and the background information are normally distributed, with error covariance matrices $\mathbf{C}$ and $\mathbf{B}$, respectively, then the conditional probability density function $p(y|x)$ takes the form:

$$p(y|x) = \frac{1}{(2\pi)^{n/2}|\mathbf{C}|^{1/2}} \exp\left\{-\frac{1}{2}(y - \mathbf{H}(x))^T \mathbf{C}^{-1}(y - \mathbf{H}(x))\right\}, \tag{3.2}$$

where n is the number of available observations, and $\mathbf{H}$ is the observation operator. Likewise, the *prior* probability density function $p(x)$ is given by

$$p(x) = \frac{1}{(2\pi)^{m/2}|\mathbf{B}|^{1/2}} \exp\left(-\frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b)\right), \tag{3.3}$$

where m and x^b are the number of model variables to estimate and the background estimate of them, respectively.

In the Bayesian formulation, the objective is to find the best estimator, i.e., the one that maximizes the *posterior* probability density function given by (3.1). Since $p(y)$ is independent from x, this maximization problem reads as follows:

$$\max_x \quad p(y|x)p(x). \tag{3.4}$$

The parameter maximizing the posterior distribution, i.e., $x_{MAP} = \arg\max_x p(y|x)p(x)$ is called the maximum-a-posteriori (MAP) estimate.

On the other hand, taking into account that the logarithmic function is increasing and continuous, the maximization problem (3.4) is equivalent to the following one:

$$\max_x \quad \ln(p(y|x)) + \ln(p(x)). \tag{3.5}$$

Substituting the probability density functions (3.2) and (3.3) in the problem above, we obtain the equivalent minimization problem:

$$\min_x \quad \frac{1}{2}(y - \mathbf{H}(x))^T \mathbf{C}^{-1}(y - \mathbf{H}(x)) + \frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b), \tag{3.6}$$

which is a robust balance between the background information and the observed data. Problem (3.6) considers a single observation acquired at a given instant. In cases where observations are obtained at different moments in a time window, the Bayesian formulation described above has to be adapted to account for the system's evolution and the variation of the uncertainty measured by the covariance matrices.

Furthermore, when the estimation problem extends beyond determining the initial state of a dynamical system to also include the estimation of its governing parameters, the minimization problem exemplified by (3.6) must be modified. This involves incorporating background information for both the parameters and the initial state through different error covariance matrices: $\mathbf{B}$, the background error covariance for the model parameters, and $\mathbf{Q}_0$, the background error covariance for the system state.

3.1. Problem statement

In our case, the model variables to be estimated, considering the information given by the observations, are the *parameters and initial conditions* of the SARS-CoV-2 virus spread model proposed in [7, 27–29]. The mathematical formulation used to solve the parameter estimation problem jointly with an estimation of the initial conditions of the virus propagation model reads as follows:

$$\min_{\mathbf{u}_0, \rho \in \mathcal{U}_{ad}} \quad \sum_{i,j} (I^r(t_i) - I_i^{obs}) \mathbf{C}_{ij}^{-1} (I^r(t_j) - I_j^{obs})) + (\rho - \rho^b)^T \mathbf{B}^{-1} (\rho - \rho^b) \tag{3.7a}$$

$$+(\mathbf{u}_0 - \mathbf{u}_0^b)^T \mathbf{Q}_0^{-1} (\mathbf{u}_0 - \mathbf{u}_0^b) = J(\bar{\rho}, \bar{\mathbf{u}}_0)$$

$$\text{s.t.} \quad \frac{dS}{dt} = -\frac{\beta S I^r}{N} - \frac{\mu \beta S I^u}{N}, \qquad S(t_0) = S_0, \tag{3.7b}$$

$$\frac{dE}{dt} = \frac{\beta S I^r}{N} + \frac{\mu \beta S I^u}{N} - \frac{E}{Z}, \qquad E(t_0) = E_0, \tag{3.7c}$$

$$\frac{dI^r}{dt} = \alpha\frac{E}{Z} - \frac{I^r}{D}, \qquad I^r(t_0) = I_0^r, \tag{3.7d}$$

$$\frac{dI^u}{dt} = (1-\alpha)\frac{E}{Z} - \frac{I^u}{D}, \qquad I^u(t_0) = I_0^u, \tag{3.7e}$$

$$\frac{dR}{dt} = \frac{I^r + I^u}{D}, \qquad R(t_0) = R_0. \tag{3.7f}$$

The contagion and spread of SARS-CoV-2 are modeled by the SEIR-type system given by Eqs (3.7b)–(3.7f), where S, E, I^r, I^u, R, and N represent the susceptible, exposed, documented infected, undocumented infected, removed, and total number of inhabitants of the studied province. To capture mathematically the particularities of the virus, such as the presence of asymptomatic infections, the infected population is divided into two separate compartments: documented and undocumented. The documented infected population comprises individuals who have undergone PCR testing due to their symptoms, yielding positive results. Conversely, the undocumented cases represent positive infections with mild or no symptoms, potentially exposing a larger portion of the population to the virus as they may go unnoticed. Additionally, the SEIR model strongly depends on the following coefficients:

- β: transmission rate due to documented or symptomatically infected patients.
- μ: multiplicative factor that reduces the transmission rate of undocumented patients.
- α: fraction of the documented infected population developing severe symptoms.
- Z and D: average latency period and the duration of the infection, respectively.

In the cost functional (3.7a), $\rho = (\beta, \mu, \alpha, Z, D)^T$ represents a parameter's vector, and $\mathcal{U}_{ad}$ is the admissible set formed by the initial prior range of the model parameters as follows: $0.8 \leq \beta \leq 1.5$, $0.2 \leq \mu \leq 1$, $0.02 \leq \alpha \leq 1$, Z and D between 2 and 5 days. The bounds of the intervals as well as the background information of the parameters, ρ^b, are determined by the values given in the extensive COVID-19 literature (see, e.g., [7] and the reference therein). In particular, for the ranges of the model parameters, we use the work of Li et al. [7], where the intervals were computed considering a 95% confidence level, using Latin hypercube sampling (LHS) with $n = 300$ simulations. $\mathbf{u}_0$ is the vector of initial conditions, while $\mathbf{u}_0^b$ represents its background information. For the current assimilation window k, the background initial condition $\mathbf{u}_0^b$ will be set to $\mathbf{u}_{k-1}$, the final state vector of the SEIR model obtained from the analysis of the immediately preceding window $(k-1)$. The measured or observed data, I^{obs}, is given by the official statistics on positive COVID-19 cases reported by the Ecuadorian Health Ministry or the number of cases inferred from the mortality excess obtained from the Ecuadorian Civil Registry. The error covariance matrices appearing in (3.7a), $\mathbf{C}$, $\mathbf{B}$, and $\mathbf{Q}_0$, are built as described in 3.3. The solutions to problem (3.7) are optimal parameters $\bar{\rho} = (\beta, \mu, \alpha, Z, D)$ and an optimal initial condition $\bar{\mathbf{u}}_0 = (E_0, I_0^r, I_0^u, R_0)$, that can be used to estimate the SARS-CoV-2 propagation model outputs, $\mathbf{u}(t) = (S(t), E(t), I^r(t), I^u(t), R(t))$, for a given period of time $[t_0, t_n]$. We recall that the obtained optimal parameters correspond to MAP estimates, which are single-point values. Obtaining their full probability distribution requires additional steps, such as employing a Laplace approximation method. We will explore this process further in Section 4.1.

It is important to mention that the SEIR system (3.7b)–(3.7f) studies the virus's behavior and its spreading in an isolated province or region. In the Ecuadorian case, studying the isolated model makes sense since one of the measures implemented by the government during the pandemic outbreak was the prohibition of traveling outside provinces.

3.2. Solution via a Lagrangian approach

The minimization problem (3.7), where the objective functional $J(\rho, \mathbf{u}_0)$ is minimized subject to the SEIR differential equation (3.7b)–(3.7f), is addressed using techniques from optimal control [30, 31]. Specifically, we employ the Lagrangian approach to derive the first-order necessary optimality conditions, which comprise:

- The state equations, which are the original SEIR model ODEs (3.7b)–(3.7f).
- The adjoint equations, a backward system describing the dynamics of the Lagrange multipliers (adjoint variables) $\lambda(t)$.
- The gradient of the cost functional J with respect to the control parameters ρ and initial conditions $\mathbf{u}_0$, expressed in terms of the state and adjoint variables.

The detailed derivation of the specific adjoint system and the gradient expressions for our problem can be found in Appendix B.1.

Since the variable corresponding to the model parameters ρ belongs to the admissible set $\mathcal{U}_{ad}$, the optimality condition of the problem is given by the following variational inequality:

$$\nabla J(\bar{\rho}, \bar{\mathbf{u}}_0)^T [(\rho, \mathbf{u}_0) - (\bar{\rho}, \bar{\mathbf{u}}_0)] \geq 0, \qquad \forall \rho \in \mathcal{U}_{ad}, \forall \mathbf{u}_0 \in \mathbb{R}^4_+, \tag{3.8}$$

where the partial derivatives forming the gradient $\dfrac{\partial J}{\partial \rho}$ and $\dfrac{\partial J}{\partial \mathbf{u}_0}$ are given by (B.4) and (B.5), respectively. Note that since the box constraints for ρ and $\mathbf{u}_0$ are convex sets: $\mathcal{U}_{ad} = \{\rho \in \mathbb{R}^5 : \rho_\ell \leq \rho \leq \rho_u\}$ where for each $i \in \{1, \ldots, 5\}$, ρ_i^ℓ and ρ_i^u represent the lower and upper bounds for each parameter, and, for each $j \in \{1, \ldots, 4\}$, the lower bound of $\mathbf{u}_0^j$ is taken as 0, then, the variational inequality (3.8) may be replaced by a pointwise projection [31, Section 5.2], as follows:

$$\begin{aligned} &\bar{\rho}_i - P\left(\bar{\rho}_i - \gamma \nabla_{\rho_i} J\right) = 0, && \text{for } i \in \{1, \ldots, 5\}, \\ &\bar{\mathbf{u}}_0^j - P\left(\bar{\mathbf{u}}_0^j - \gamma \nabla_{\mathbf{u}_0^j} J\right) = 0, && \text{for } j \in \{1, \ldots, 4\}, \end{aligned}$$

for every $\gamma > 0$. $P : \mathbb{R}^n \to \Omega$ represents the projection operator defined by:

$$P(x)_i = \begin{cases} a_i, & \text{if } x_i \leq a_i, \\ x_i, & \text{if } a_i < x_i < b_i, \\ b_i, & \text{if } b_i \leq x_i, \end{cases}$$

where $\Omega = \{x \in \mathbb{R}^n : a_i \leq x_i \leq b_i\}$.

Furthermore, the robustness and local optimality of the obtained solution can be assessed by verifying the second-order sufficient conditions, as detailed in Appendix B.2. In particular, the solution to the variational data assimilation problem (3.7) is a local minimizer if the second derivative of the Lagrangian functional is positive definite at that point. This condition is satisfied, for instance, when $|\mathbf{C}_{ii}^{-1}(I^r - I^{obs})|$ is sufficiently small.

3.3. Building error covariance matrices

One of the critical elements in the variational formulation is the construction of the error covariance matrices appearing in the cost functional (3.7a). $\mathbf{C}$ represents the error covariance matrix of the observations, and $\mathbf{B}$ and $\mathbf{Q}_0$ are the error covariance matrices of the background information of the model

parameters and the initial condition, respectively. Next, we explain how the error covariance matrices are built for our specific problem and data.

Matrix **C** - Health Ministry data

The matrix **C** is computed by taking into account the effectiveness of the PCR tests in detecting the SARS-CoV-2 virus and the number of PCR tests applied to the population. Specifically, **C** is a diagonal matrix where each diagonal element σ_{ii}^2 takes the form:

$$\sigma_{ii}^2 = \frac{\kappa}{\eta_i}\sigma_{PCR}^2, \tag{3.9}$$

where σ_{PCR} is the standard deviation inherent to PCR tests. It refers to the baseline level of uncertainty or variability associated with using PCR tests as a diagnostic tool for counting cases. The parameter κ represents a benchmark for weekly testing capacity, set at 4000 tests per million inhabitants, reflecting levels achieved in countries with effective and widespread testing campaigns during the pandemic, such as Germany, Vietnam, and South Korea (see, e.g., [32,33]). The variable η_i is the number of PCR tests per million inhabitants performed in each province during a week of study.

Matrix **C** - Civil Registry data

Since the observed infected population is estimated based on excess mortality, the observation error is directly related to the error in computing this excess. Given that excess mortality is defined as the difference between reported and expected deaths, we assume that the primary source of error in excess mortality arises from inaccuracies in forecasting expected deaths.

The residuals of the forecasted expected mortality, obtained using a seasonal ARIMA model, form a stationary time series with constant variance; see, e.g., [44, Section 1.4]. In our case, the standard deviations of these residuals for Pichincha and Guayas are $\sigma_P = 27.29$ and $\sigma_G = 45.25$, respectively.

Matrix **B**

The background knowledge of the parameters, like their confidence intervals and mean values, are taken from the extensive COVID-19 literature. We use the work of Li et al. [7], where the authors set the initial prior range of the model parameters as follows: $0.8 \le \beta \le 1.5$, $0.2 \le \mu \le 1$, $0.02 \le \alpha \le 1$, a mean latency period Z between 2 and 5 days, and an average duration of the infection, D, in infected patients from 2 to 5 days, to find the elements of the diagonal matrix **B**. In fact, in the source literature, the intervals were computed at $1-\tilde{\alpha} = 95\%$ confidence level, using LHS with $n = 300$ simulations. This information lets us retrieve a measure of uncertainty of the background knowledge of the parameters. In fact, according to [34], the lower and upper bounds of the intervals are computed as follows:

$$\bar{\varsigma} \pm t_{1-\frac{\tilde{\alpha}}{2}} s_\varsigma, \tag{3.10}$$

where ς represents any of the model parameters $\{\beta, \mu, \alpha, Z, D\}$ and $\bar{\varsigma}$ its mean value. Additionally, for the given values of n and $\tilde{\alpha}$, $t_{1-\frac{\tilde{\alpha}}{2}} = 1.968$. Given the lower ($\varsigma_{\text{low}}$) and upper bounds ($\varsigma_{\text{up}}$) of each model parameter, we can compute its standard deviation, s_ς, from Eq (3.10). In fact, $s_\varsigma = \frac{\varsigma_{\text{up}} - \bar{\varsigma}}{t_{1-\frac{\tilde{\alpha}}{2}}}$. The square values of these elements will form the diagonal matrix **B**.

Matrix $\mathbf{Q}_0$

The matrix $\mathbf{Q}_0$ measures the problem's sensitivity concerning the initial conditions. To build this matrix for the k-period, we form a set of perturbed initial conditions $\mathbf{u}_{0_k}^{\ell} = (E_0^{\ell}, Ir_0^{\ell}, Iu_0^{\ell}, R_0^{\ell})$, by considering the solution vector at the end of the previous period, i.e., $\mathbf{u}_{k-1}$, and random Gaussian noise as follows:

$$\mathbf{u}_{0_k}^{\ell} = \mathbf{u}_{k-1} + \varepsilon_{\ell}, \qquad \varepsilon_{\ell} \sim N(0, \sigma_{\ell}^2),$$

where $\sigma_{\ell} = 0.2\mathbf{u}_{k-1}$ and $\ell = 1, \ldots, M$. Solving the SARS-CoV-2 propagation model with these initial conditions results in an ensemble of M state vectors $\mathbf{u}_k^{\ell} = (E^{\ell}, Ir^{\ell}, Iu^{\ell}, R^{\ell})$. The main idea is to emulate the uncertainty of the system by the variability of a collection of solution vectors. In this sense, matrix $\mathbf{Q}_0$ is estimated using the information of these ensemble members as follows:

$$\mathbf{Q}_0 = \frac{1}{M-1} \sum_{\ell=1}^{M} (\mathbf{u}_k^{\ell} - \bar{\mathbf{u}})(\mathbf{u}_k^{\ell} - \bar{\mathbf{u}})^T, \tag{3.11}$$

where $\bar{\mathbf{u}} = \dfrac{1}{M} \displaystyle\sum_{\ell=1}^{M} \mathbf{u}_k^{\ell}$.

3.4. Numerical algorithms

In this section, we describe the algorithms we use to solve the data assimilation problems that appear in the paper. We have chosen a *discretize-then-optimize* approach as it offers advantages in finding the numerical solution to the minimization problem by employing quasi-Newton methods. In particular, an updating formula, without solving large linear systems, is available in this context, which significantly improves the computational efficiency.

3.4.1. Projected quasi-Newton methods

Projected methods are used to solve minimization problems with box constraints, $\min_{w \in U_{ad}} f(w)$, with $U_{ad} = \{w : a_i \leq w_i \leq b_i\}$. Generally, these methods use the descent directions of the unconstrained problem and project the next iteration on the admissible set, that is,

$$w_{k+1} = P_{U_{ad}}(w_k + \alpha_k d_k), \tag{3.12}$$

where P represents the projection operator, d_k is a descent direction, and α_k is a projected line search parameter; see, e.g., [31, Section 5.3] .

For projected quasi-Newton methods, it is necessary to use the information of the reduced Hessian or its approximations (denoted as H_k) based on the estimation of $\epsilon-$active sets:

$$A^{\epsilon}(w) = \{i : a_i \leq w_i \leq a_i + \epsilon \text{ or } b_i \geq w_i \geq b_i - \epsilon\},$$

with $0 \leq \epsilon \leq \min_i\{b_i - a_i\}/2$. Its complement forms the $\epsilon-$inactive set, $I^{\epsilon}(w)$. The reduced Hessian, or an approximation, is defined as follows:

$$\tilde{R}(w_k, \epsilon_k, H_k) \quad = R_{A^{\epsilon_k}(w_k)} + R_{I^{\epsilon_k}(w_k)} H_k R_{I^{\epsilon_k}(w_k)} \quad = \begin{cases} \delta_{ij}, \text{if } i \in A^{\epsilon_k}(w_k) \text{ or } j \in A^{\epsilon_k}(w_k) \\ (H_k)_{ij}, \text{ otherwise,} \end{cases}$$

where $(R_S(w))_i = w_i$ if $i \in S$, and $(R_S(w))_i = 0$ if this is not the case, for any index set S. The model iterations will be given as follows:

$$w_{k+1} = P_{U_{ad}}(w_k - \alpha_k \tilde{R}(w_k, \epsilon_k, H_k)^{-1} \nabla f(w_k)).$$

BFGS projected method

If the approximation of the Hessian matrix is given by the BFGS method (Broyden-Fletcher-Goldfarb-Shanno), its reduced matrix is given by:

$$H_{k+1} = R_{I_{k+1}} H_k R_{I_{k+1}} - R_{I_{k+1}} \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k} R_{I_{k+1}} + \frac{y_k^{\#}(y_k^{\#})^T}{s_k^T y_k^{\#}},$$

where $I_{k+1} = I^{\epsilon_{k+1}}(w_{k+1})$, $s_k = w_{k+1} - w_k$, and $y_k^{\#} = R_{I_{k+1}}(\nabla f(w_{k+1}) - \nabla f(w_k))$. In some problems, it is helpful to get updates for the generalized inverse of H_k, i.e., B_k. According to Kelley [35] in Section 5.5.3, the update of this generalized inverse, B_k, has the form:

$$B_{k+1} = \left(I - \frac{s_k^{\#}(y_k^{\#})^T}{(y_k^{\#})^T s_k^{\#}}\right) R_{I_{k+1}} B_k R_{I_{k+1}} \left(I - \frac{y_k^{\#}(s_k^{\#})^T}{(y_k^{\#})^T s_k^{\#}}\right) + \frac{s_k^{\#}(s_k^{\#})^T}{(y_k^{\#})^T s_k^{\#}}, \tag{3.13}$$

where $s_k^{\#} = R_{I_{k+1}}(w_{k+1} - w_k)$. In particular, using the update of the generalized inverse, it yields

$$d_k = -B_k \nabla f(w_k) - R_{A^{\epsilon_k}(w_k)} \nabla f(w_k).$$

In our implementation, we used the reduced matrix update given by (3.13). To see the form of this matrix for our specific problem, let us set the following values: in the iteration k, $w_k = (\rho_k, \mathbf{u}_0^k)$. Let $U_{ad} = \{w = (\rho, \mathbf{u}_0) : \rho \in \mathcal{U}_{ad}, \text{ and } 0 \le \mathbf{u}_0 \le U\}$ represent the admisible set, where $U > 0$ is a high upper bound for the initial conditions, and $\mathcal{U}_{ad} = \{\rho : \rho_\ell \le \rho \le \rho_u\}$ where ρ_ℓ and ρ_u denote the lower and upper bound for the model parameters, respectively. Note that $w = (\rho, \mathbf{u}_0) \in \mathbb{R}^9$, and then, for a given value of ϵ, the ϵ-active set has the form

$$\begin{gathered} A^{\epsilon}(w) = \{i : (\rho_\ell)_i \le w_i \le (\rho_\ell)_i + \epsilon \text{ or } (\rho_u)_i - \epsilon \le w_i \le (\rho_u)_i \text{ if } i \in \{1, \dots, 5\}, \text{ or} \\ 0 \le w_i \le \epsilon \text{ or } U - \epsilon \le w_i \le U \text{ if } i \in \{6, \dots, 9\}\}. \end{gathered}$$

We will denote $I_{k+1} = I^{\epsilon}(w_{k+1}) = (A^{\epsilon}(w_{k+1}))'$, $s_k^{\#} = R_{I_{k+1}}((\rho_{k+1}, \mathbf{u}_0^{k+1}) - (\rho_k, \mathbf{u}_0^k))$, and

$$y_k^{\#} = R_{I_{k+1}} \left(\left[\frac{\partial J(\rho_{k+1}, \mathbf{u}_0^{k+1})}{\partial \rho}, \frac{\partial J(\rho_{k+1}, \mathbf{u}_0^{k+1})}{\partial \mathbf{u}_0} \right]^T - \left[\frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \rho}, \frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \mathbf{u}_0} \right]^T \right),$$

where $\dfrac{\partial J}{\partial \rho}$ and $\dfrac{\partial J}{\partial \mathbf{u}_0}$ are given by (B.4) and (B.5), respectively. Replacing these terms in (3.13), we obtain an update of the generalized inverse B_k, and consequently, a descent direction

$$d_k = -B_k \left[\frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \rho}, \frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \mathbf{u}_0} \right]^T - R_{A^{\epsilon_k}(w_k)} \left[\frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \rho}, \frac{\partial J(\rho_k, \mathbf{u}_0^k)}{\partial \mathbf{u}_0} \right]^T. \tag{3.14}$$

4. Results

In this section, we analyze and compare the evolution of the COVID-19 pandemic in two provinces of Ecuador: Pichincha and Guayas. One of our goals is to compare the estimated models built during the pandemic—based on the official COVID-19 case counts reported by the Public Health Ministry—with those derived from Civil Registry data, which only became available after our initial estimates were made. Since the latter dataset and its corresponding estimated models are considered the most reliable representations of the pandemic's state, we use them as references to assess the accuracy of our initial estimates. Additionally, we examine the evolution of the *documented infected population* and the *effective reproduction number*, two key indicators of the pandemic's progression, based on the two datasets.

4.1. Estimation of COVID-19 model parameters

To obtain the model parameters and their approximate posterior distributions, we proceed in two steps. First, we determine the optimal parameters (ρ) and initial conditions ($\mathbf{u}_0$) by solving the variational data assimilation problem (3.7). We proceed iteratively using the projected BFGS method described in Section 3.4.1, until a stopping criterion is met. Since the computations are performed weekly, the initial point for the algorithm is set to the parameters obtained in the previous week, ρ_{k-1}, while for the initial conditions, $\mathbf{u}_0$, the observations at the first day of the studied week are considered. The BFGS direction is determined by (3.14), and the parameters and initial conditions are updated following (3.12). The stopping criterion is defined by $e(w_k) = \|w_k - P_{U_{ad}}(w_k - \nabla f(w_k))\| \leq \texttt{tol}$, where $w_k = (\rho_k, \mathbf{u}_0^k)$ and $\texttt{tol} > 0$ is a fixed tolerance value. Specifically, we set $\texttt{tol} = 1 \times 10^{-4}$; see, e.g., [31, Chapter 5.3].

If the iterations of the projected BFGS start near a non-degenerate local minimum and a good initial approximation to the reduced Hessian is available, it is expected that the iterations of the method will converge q-superlinearly; see [35, Theorem 5.5.4]. This is verified also in our experiment, see Table 1, where we summarize the method performance for a subsample of weeks. We recall that the algorithm converges superlinearly if the convergence rate, p, is between 1 and 2.

Table 1. Performance of the projected BFGS method with RC data for Guayas.

Studied week	iter	$f(w)$	$e(w)$	Conv. rate (p)
March 13–19, 2020	8	3.003	1.71×10^{-11}	1.37
June 19–25, 2020	10	4.908	3.01×10^{-10}	1.10
Nov 20–26, 2020	8	4.212	2.12×10^{-10}	1.97
April 9–15, 2021	9	3.258	9.14×10^{-9}	1.74
May 28–June 3, 2021	11	14.59	1.99×10^{-8}	2.14

Since we are using observational data from two different sources—MSP and RC—the iterative process described above yields two sets of optimal solutions: $(\mathbf{u}_0^{MSP}, \rho^{MSP})$ and $(\mathbf{u}_0^{RC}, \rho^{RC})$. As mentioned previously, the obtained optimal parameters correspond to the MAP estimates, representing the mode of the posterior distribution. However, MAP estimates alone do not provide information about the uncertainty surrounding these optimal values or the full shape of their posterior distribution. Conse-

quently, to obtain a more complete understanding of the parameter uncertainty, we need to approximate the posterior distribution of the model parameters, $w = (\rho, \mathbf{u}_0)$, given the observed data, I^{obs}, i.e., $p(w|I^{obs})$.

According to Bayes' formula (Eq 3.1), the posterior distribution is proportional to the product of the likelihood $p(I^{obs}|w)$ and the prior $p(w)$. We have already identified the MAP estimate, $w_{MAP} = (\rho_{MAP}, \mathbf{u}_{0,MAP})$, which maximizes this posterior (or equivalently, its logarithm). The data assimilation approach used to find w_{MAP} (Section 3) was formulated by assuming Gaussian distributions for both the likelihood (errors in observations) and the prior knowledge (background information for w). Given these assumptions and the computed w_{MAP}, we employ the Laplace approximation to estimate the full posterior distribution.

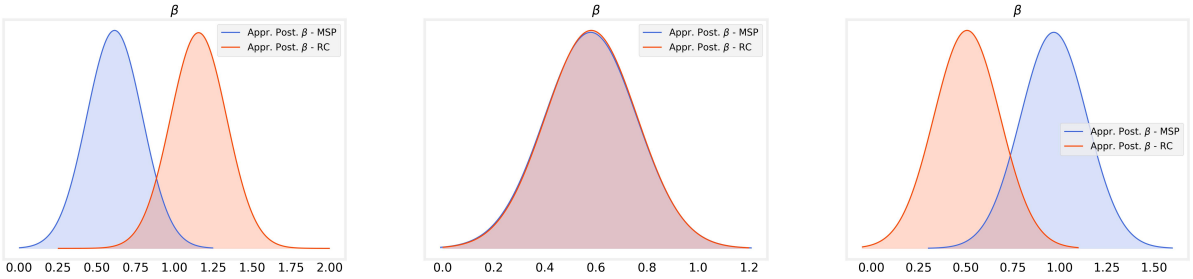
Laplace approximation exploits the properties of the approximate posterior around the mode. The core idea is that, in the vicinity of a well-defined mode (such as our MAP estimate), the posterior distribution can be effectively approximated by a multivariate Gaussian distribution, i.e.,

$$p(w|I^{obs}) \approx \mathcal{N}(w|\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{4.1}$$

This approximating Gaussian is centered at the MAP estimate, so its mean is $\boldsymbol{\mu} = w_{MAP}$. The covariance matrix $\boldsymbol{\Sigma}$ is determined by the curvature of the log-posterior at the mode, specifically $\boldsymbol{\Sigma} = \mathcal{H}^{-1}$, where $\mathcal{H}$ is the Hessian matrix of the negative log-posterior evaluated at w_{MAP} (i.e., $\mathcal{H} = -\nabla^2 \ln p(w|I^{obs})|_{w=w_{MAP}}$); see, e.g., [45, Section 13.3]. This method provides an efficient alternative to sampling-based approaches like Markov Chain Monte Carlo (MCMC) for characterizing posterior uncertainty, particularly when a well-defined mode exists.

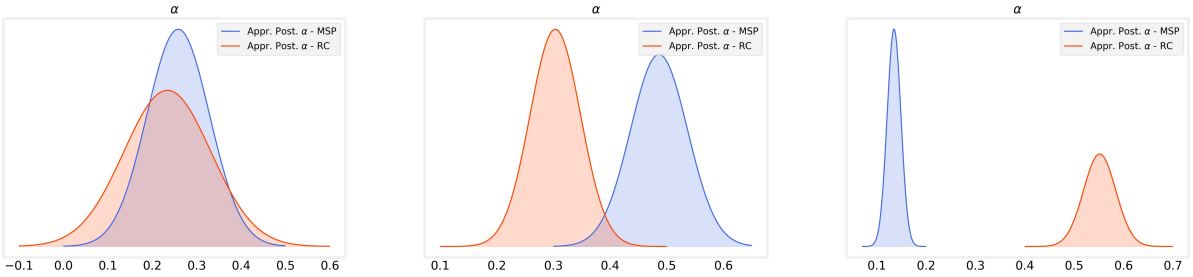
In our framework, this matrix corresponds to the Hessian of the cost functional $J(\rho, \mathbf{u}_0)$, given in (3.7a), with respect to $w = (\rho, \mathbf{u}_0)$, evaluated at $w_{MAP} = (\mathbf{u}_0^{MAP}, \rho_{MAP})$, i.e., $\nabla^2 J(\rho, \mathbf{u}_0)|_{w=w_{MAP}}$. Observe that we have computed $\nabla^2 \mathcal{L}_{\mathbf{u},\rho,\mathbf{u}_0}[(h, h_\rho, h_0), (h, h_\rho, h_0)]$ in Section B.2, which when evaluated at the optimum $w_{MAP} = (\rho_{MAP}, \mathbf{u}_0^{MAP})$, coincides with $\nabla^2 J_{\rho,\mathbf{u}_0}[(h_\rho, h_0), (h_\rho, h_0)]$; see, e.g., [31, Section 4.2]. The resulting Gaussian distribution provides an estimate of the full posterior, allowing for the quantification of uncertainty in w and the generation of samples for propagating this uncertainty to model predictions.

To assess the similarity between the posterior distributions of the parameters obtained using observations from the Civil Registry and the Public Health Ministry, we compare both distributions for selected parameters and dates. Figure 2 presents both approximate posterior distributions of the parameter β, representing the transmission rate due to symptomatically infected patients, on different dates for the province of Pichincha, while Figure 3 illustrates the same comparison for the parameter α, which indicates the fraction of the documented infected population developing severe symptoms, for the province of Guayas. As shown in the figures, for both selected parameters, there are instances where the two approximate posterior distributions closely align, as seen in Figure 2(b). However, there are also weeks when the approximate distributions are close to each other but do not overlap, as observed in Figure 3(c).



(a) April 16, 2020 **(b)** April 30, 2020 **(c)** September 17, 2020

Figure 2. Comparison of the approximate posterior distribution of β obtained using MSP data (shown in blue) and RC data (shown in red) for Pichincha.



(a) April 23, 2020 **(b)** May 21, 2020 **(c)** January 14, 2021

Figure 3. Comparison of the approximate posterior distribution of α obtained using MSP data (shown in blue) and RC data (shown in red) for Guayas.

To gain insight into the behavior of these estimated posterior distributions over the entire period, we display and compare their means. Figure 4 presents the comparison and evolution of the posterior distributions' means for the parameter β in the Guayas province, and μ in the province of Pichincha. We recall that μ is a multiplicative factor that reduces the transmission rate of undocumented patients.

To observe the percentage of difference between the parameters' estimations obtained with both data sources, we compute the relative error between them as follows: $\rho_{\text{rel.error}} = \frac{\rho_{RC} - \rho_{MSP}}{\rho_{MSP}}$ for each $\rho \in \{\beta, \mu, \alpha, Z, D\}$. Figure 5 displays the relative errors for selected model parameters in the first year of study for Guayas (top) and Pichincha (bottom).

Uncertainty quantification is crucial for parameter estimation in epidemiological models, particularly for COVID-19, where data may be incomplete, noisy, or affected by reporting delays. By characterizing the full posterior distribution rather than relying solely on MAP estimates, we obtain a more robust understanding of the variability in model predictions. This is especially important for decision-making, as predictive intervals allow policymakers to assess potential worst-case scenarios and adapt containment strategies accordingly.

4.2. Evolution of the documented infected population

If the estimation of the COVID-19 documented population is performed periodically (e.g., weekly), the solution variables—particularly the *documented infected* population—can be updated at the same

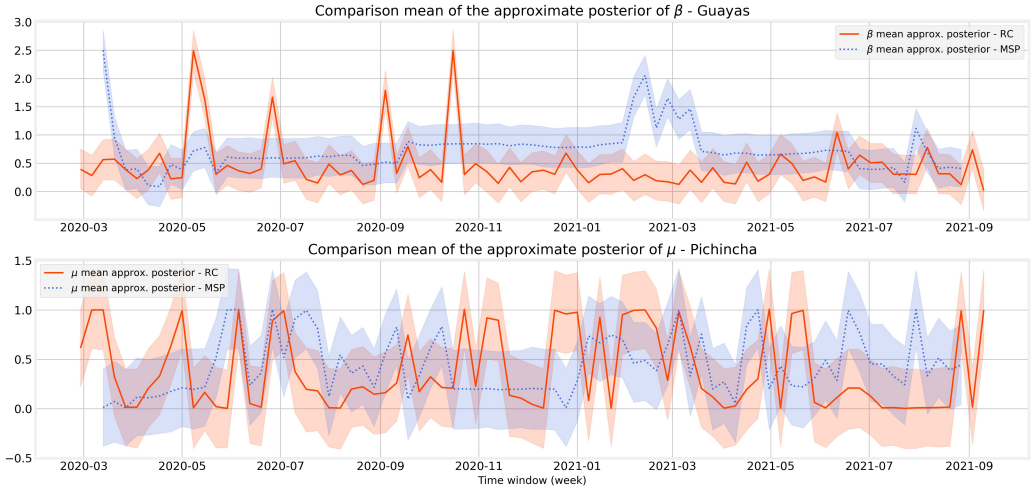


Figure 4. Comparison of the evolution of the posterior distribution mean for parameter β (top), and parameter μ (bottom) computed with MSP data (dotted line) and RC data (solid line).

frequency. We report the results for the variable I^r in Guayas and Pichincha, computed using two different data sources and the methodology described in Section 4.1:

- I^r-MSP: Obtained using the observations coming from official COVID-19 data reported by the Public Health Ministry (MSP).
- I^r-RC: Estimated using COVID-19-positive cases inferred from excess mortality data recorded by the Civil Registry (RC).

Figures 6 and 7 depict the evolution of the number of contagious individuals estimated from both data sources. It is important to highlight that the MSP estimate (I^r-MSP) was conducted in real time during the crisis using the available data at that moment. In contrast, the estimate based on RC data (I^r-RC) was performed retrospectively when a more reliable data source became available. The estimate is considered to better approximate the actual progression of the pandemic.

In Guayas' case, the difference between both series is evident, particularly in the first months of the pandemic. In fact, by analyzing the approximate dates when the pandemic hit the worst with the different data sources, we get different values for the maximum of contagious individuals and the dates when they occurred. Nevertheless, both peaks occurred between the last weeks of March and the first week of April 2020. The difference in dates can be explained by considering that estimates based on official MSP data may naturally be delayed due to the time lag between individuals taking a COVID-19 test and receiving their results. This phenomenon was particularly evident during the pandemic's first months. It is worth mentioning that the maximum number obtained with MSP data (3670 cases) is much smaller than the one obtained with the RC data (18,368 cases), highlighting the potential for underreporting in official statistics. We summarize these results in Table 2.

In Pichincha's case, it is important to remark that in mid-September 2020, the results of repressed PCR tests were released, which artificially increased the number of observed positive COVID-19 cases due to the high PCR-test positivity; see, e.g., [36, 37]. Therefore, the results obtained from MSP

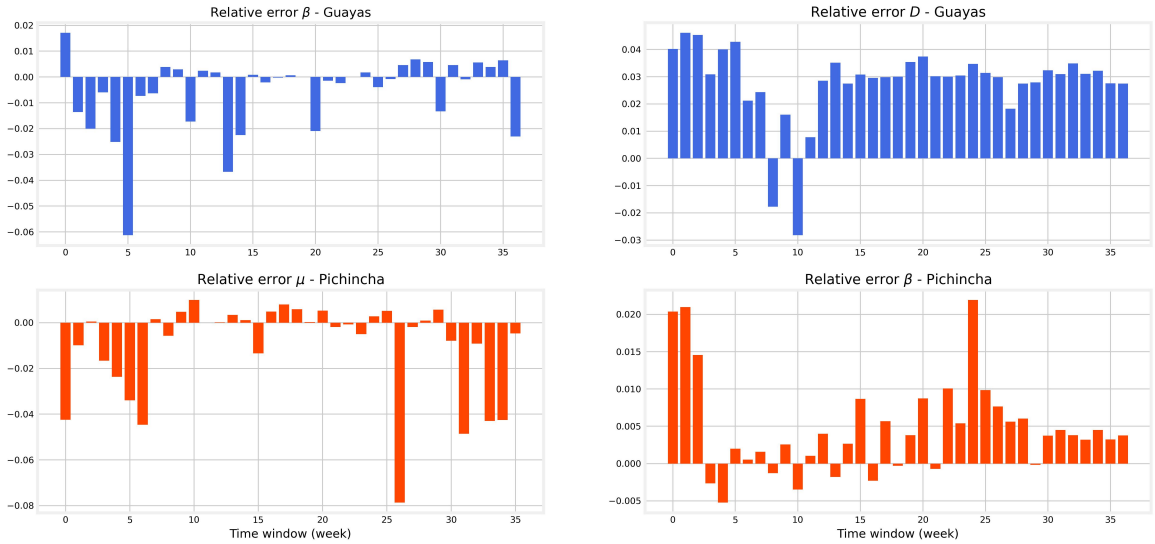


Figure 5. Evolution of the relative error for selected model parameters from March to December 2020.

data for those dates were artificially high. To deal with that, we treat these results as outliers, and to correct them, we used a linear interpolation of the generated I^r-MSP series. Likewise, the results of this province's peak infection values and dates are illustrated in Table 2. Analyzing Pichincha's series, we find that the first critical infection increment, according to both data sources, occurred in July 2020. In fact, this date was recognized by national and international media as one of the hardest of the COVID-19 pandemic in Quito (Pichincha's capital city); see, e.g., [38]. It is worth noticing that, different from Guayas' case, the estimate of the infected population obtained with RC data is lower than the one obtained with the records of the Public Health Ministry.

Table 2. Infection peaks computed with MSP and RC data for Guayas and Pichincha.

	Guayas		Pichincha	
	Max	Date	Max	Date
I^r-RC	18386	12-03-2020	1624	16-07-2020
I^r-MSP	3670	02-04-2020	7490	22-04-2021

4.3. Effective reproduction number

A critical metric for understanding epidemic dynamics and guiding public health interventions is the *effective reproduction number*, denoted by R_t. Unlike the basic reproduction number (R_0), which measures the average number of secondary infections caused by a single infectious individual in a completely susceptible population, R_t quantifies this average number of secondary infections at a specific time during the epidemic. R_t accounts for changes in population susceptibility, due to infection-acquired immunity or vaccination, and the impact of interventions like lockdowns. Therefore, tracking

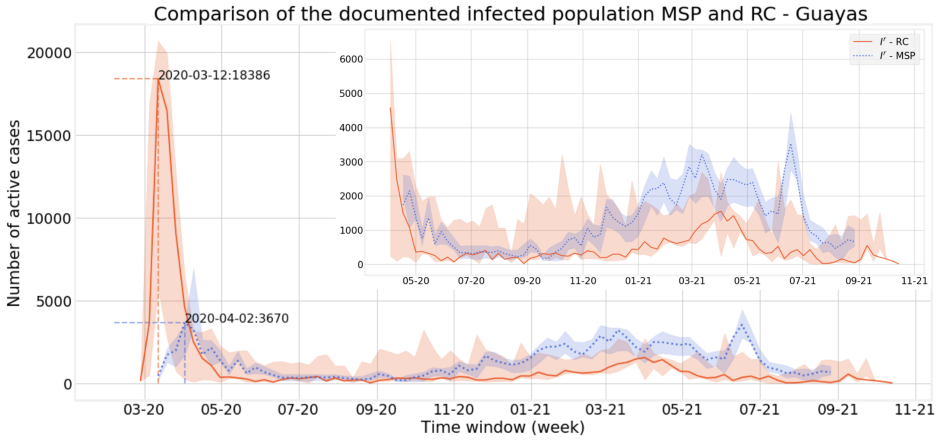


Figure 6. Comparison of the evolution of the variable *reported infected population* computed with MSP data (dotted line) and RC data (solid line) for Guayas. A zoomed image of the results from May 2020 to September 2021 is shown inside.

R_t provides a real-time overview of the epidemic's evolution and the effectiveness of control measures.

In the context of our compartmental SEIR model (Eqs (3.7b)–(3.7f)), which distinguishes between documented (I^r) and undocumented (I^u) infections, the effective reproduction number is calculated based on the estimated model parameters as:

$$R_t = \alpha\beta D + (1-\alpha)\mu\beta D \tag{4.2}$$

where β is the transmission rate due to documented cases, μ is the multiplicative factor reducing transmission from undocumented cases, α is the fraction of infections developing severe symtoms, and D is the average duration of the infection. The time-varying nature of R_t arises directly from our methodology: since the parameters (α, β, μ, D) are estimated periodically (e.g., weekly) using the Bayesian data assimilation framework described in Section 3, the resulting R_t value will reflect the conditions prevalent during that estimation window. The interpretation of R_t values is crucial. An R_t-value higher than two represents the exponential growth of the disease. Moreover, if the reproduction number is between one and two, it means a less accelerated increase in the infection. When the R_t values are lower than one, we can say that the spread of the virus is shrinking. Being steadily in this range means the infection is under control [39].

Since the SEIR model parameters are periodically updated using either MSP or RC data, we compute and compare the corresponding time series for the effective reproduction number, R_t. Figures 8 (for Guayas) and 9 (for Pichincha) illustrate the evolution of the estimated R_t comparing the estimates derived from MSP data (R_t-MSP, dashed line, from March 12, 2020) and RC data (R_t-RC, solid line, from March 5, 2020) until July 29, 2021. The shaded area represents the uncertainty of the R_t values. It was computed using deterministic bounds [40–42]. The upper bound is found by using the upper limits of the uncertainty intervals of all input parameters in the R_t formula. The lower bound is found by using the corresponding lower limits. We present the R_t series with semi-log graphs to appreciate in a better way the variations in their evolution without getting perturbed by the different scales of the series itself.

In Guayas (Figure 8), the early outbreak phase in March–April 2020 was characterized by extremely

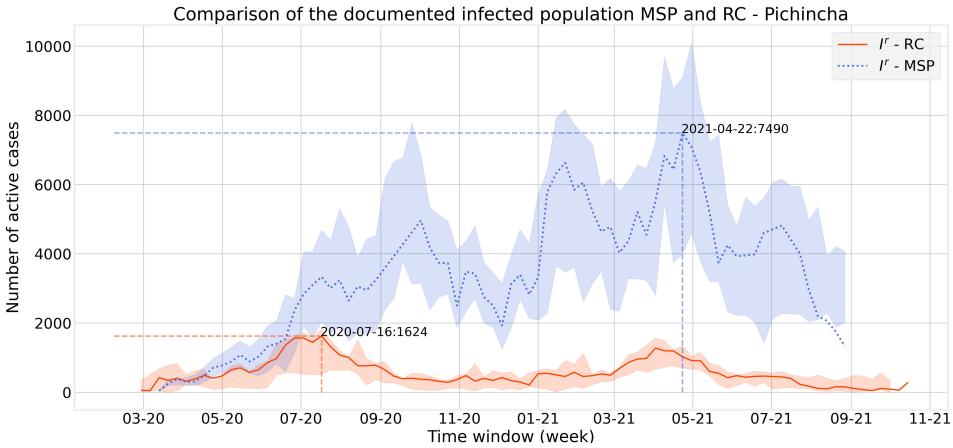


Figure 7. Comparison of the evolution of the variable *reported infected population* computed with MSP data (dotted line) and RC data (solid line) for Pichincha.

high transmission levels, both for R_t-RC and R_t-MSP estimates. This aligns with Guayas's status as one of the initial epicenters of the pandemic in the region. From mid-2020 to early 2021, transmission appeared comparatively stable, with both indicators fluctuating around 1, though R_t-RC occasionally showed values above this threshold.

In Pichincha (Figure 9), the initial R_t peak was lower but still significant. The period from mid-2020 to mid-2021 saw more frequent and sometimes prolonged phases during which R_t slightly exceeded 1. Although the general trends in R_t-RC and R_t-MSP remain broadly aligned, short-term differences are more apparent.

An almost yearly wave pattern can be observed in the R_t estimates, with peaks around March in Guayas and June in Pichincha. However, interpreting these patterns requires caution, as their meaning depends heavily on the quality of the reported data and the assumptions made in the modeling process.

5. Discussion

This work presents a comparative analysis of the evolution of the COVID-19 pandemic in two provinces of Ecuador: Pichincha and Guayas. We computed the optimal parameters involved in the SARS-CoV-2 propagation model using two different data sources. The first one relies on the official count of positive COVID-19 cases reported by the Public Health Ministry. The second, considered a closer approximation to the actual number of COVID-19 cases in Ecuador, is based on excess mortality records from the Civil Registry.

Our approach involves the optimal estimation of the model parameters while accounting for uncertainty, using variational data assimilation methods. These optimal solutions are then combined with the likelihood of the observed data given the parameters—modeled as a normal distribution. Using Bayes' theorem, we compute the posterior probability distribution of the parameters. This framework enables a Bayesian inference of the model variables, providing a more robust estimation of the pandemic's progression.

Even though the parameter estimates of the models for both datasets differ in certain weeks, the

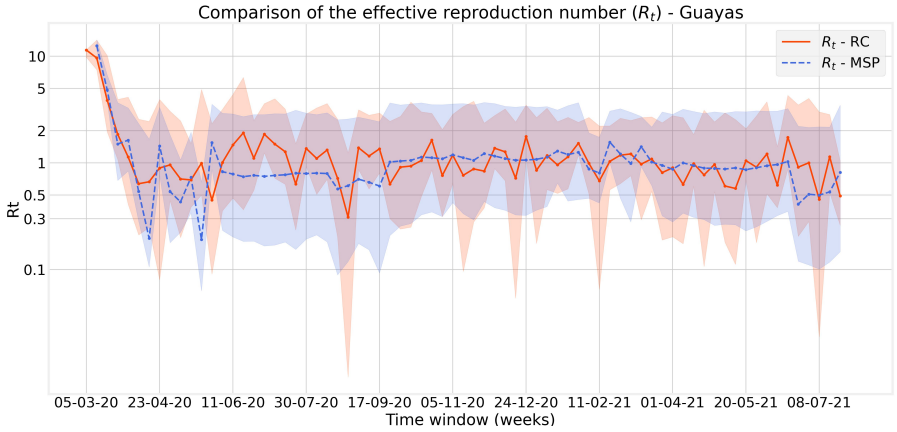


Figure 8. Evolution of the effective reproduction number in Guayas calculated with MSP and RC data from March 5, 2020 to July 29, 2021.

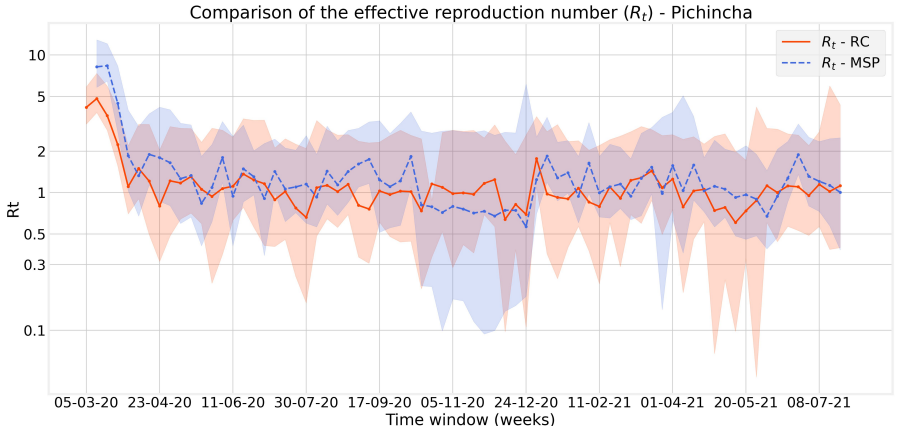


Figure 9. Evolution of the effective reproduction number in Pichincha calculated with MSP and RC data from March 5, 2020 to July 29, 2021.

evolution of the pandemic exhibits similar trends. This is particularly true for the effective reproduction number R_t, which was the primary indicator for public policy design during the pandemic. In Figures 8 and 9, it can be observed that the credible intervals of the effective reproduction numbers calculated using MSP data and RC data overlap in most of the weeks considered. In fact, in cases where they do not (e.g., September 2020), this can be attributed to the presence of outliers in the MSP data due to delayed reporting. Assuming the RC data to be closer to reality, this proximity of R_t values suggests that the data assimilation methodology presented in this paper, and used during the pandemic with MSP data, was able to accurately capture the evolution of the pandemic in Ecuador.

Regarding the comparison between the provinces of Pichincha and Guayas, a different pattern in the evolution of the pandemic is observed, regardless of whether MSP or RC data are considered. In Guayas, the pandemic had a more significant impact in the early weeks, with a sharper peak in infections compared to Pichincha. This behavior may be attributed to socioeconomic factors such as

population density and mobility, which are higher in Guayas than in Pichincha. On the other hand, in Pichincha, the pandemic had a more moderate impact, with a less pronounced peak in infections and a lower mortality rate. These differences highlight the importance of addressing the pandemic in a context-specific manner, taking into account the particular characteristics of each region.

Moreover, the estimation of the reported infected population using both datasets shows significant differences in both provinces (see Figures 6 and 7). In Guayas, the trend appears to be consistent across both datasets. In the case of Pichincha, although some peaks in the series coincide, the difference in overall trends is significant. This discrepancy may be attributed to several factors, such as a higher hospitalization rate, lower mortality at higher altitudes, or better socioeconomic conditions of the population. However, all these hypotheses require much more detailed verification.

Limitations:

One of the main limitations of our work is that we assumed no mobility between provinces in the SEIR propagation model. Studying the spatial movement of COVID-19 can be ideal for obtaining a precise picture of the pandemic state. However, in most cases, mobility data were available just for control entities, and access to that data was not granted to us. Nevertheless, the isolated SEIR model was adequate for the pandemic onset in the Ecuadorian case since mobility between provinces was prohibited.

Distinguishing the excess in mortality attributed directly to the SARS-CoV-2 infection from the one caused by non-related causes would be ideal for understanding the real impact of the pandemic. Unfortunately, there is no data available to make this distinction. Nevertheless, due to the positive correlation between all-causes mortality excess and COVID-19 deaths (see, e.g., [6]), working with the mortality excess provides a more reliable picture of the COVID-19 pandemic in Ecuador than the one obtained using the official statistics of the Public Health Ministry.

The comparison between estimates obtained using different data lets us see how the estimate can be significantly different when a reliable source of information is considered. However, the Civil Registry data takes weeks to be consolidated for the official Ecuadorian entities; therefore, we cannot develop an *online* version of the Bayesian inference approach.

Use of AI tools declaration

The authors declare that no Artificial Intelligence (AI) tools were used in the generation of the scientific content of this article. AI-based tools were used solely for minor language and style improvements, as English is not the authors' first language.

Acknowledgments

Paula Castro acknowledges partial support from the PhD Program in Applied Mathematics at Escuela Politécnica Nacional del Ecuador.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. World Health Organization, *Situation Report 51*, March 11, 2020.
2. S. Kung, M. Doppen, M. Black, I. Braithwaite, C. Kearns, M. Weatherall, et al., Underestimation of COVID-19 mortality during the pandemic, *ERJ Open Res.*, **7** (2021). https://doi.org/10.1183/23120541.00766-2020
3. H. Cevallos-Valdiviezo, A. Vergara-Montesdeoca, G. Zambrano-Zambrano, Measuring the impact of the COVID-19 outbreak in Ecuador using preliminary estimates of excess mortality, March 17–October 22, 2020, *Int. J. Infect. Dis.*, **104** (2021), 297–299. https://doi.org/10.1016/j.ijid.2020.12.045
4. P. Castro, J. C. De los Reyes, S. González-Andrade, P. Merino, Estimación de parámetros para un modelo del SARS-CoV-2 en Ecuador en presencia de incertidumbre, *Rev. Politéc.*, **47** (2021), 7–16. https://doi.org/10.33333/rp.v47n1.01
5. F. S. Lu, A. T. Nguyen, N. B. Link, M. Lipsitch, M. Santillana, Estimating the early outbreak cumulative incidence of COVID-19 in the United States: Three complementary approaches, medRxiv, 2020.
6. M. Roccetti, Excess mortality and COVID-19 deaths in Italy: A peak comparison study, *Math. Biosci. Eng.*, **20** (2023), 7042–7055. https://doi.org/10.3934/mbe.2023304
7. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2), *Science*, **368** (2020), 489–493. https://doi.org/10.1126/science.abb3221
8. D. Wu, T. Wu, Q. Liu, Z. Yang, The SARS-CoV-2 outbreak: what we know, *Int. J. Infect. Dis.*, **94** (2020), 44–48. https://doi.org/10.1016/j.ijid.2020.03.004
9. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis, *Lancet Infect. Dis.*, **20** (2020), 669–677.
10. P. Nadler, S. Wang, R. Arcucci, X. Yang, Y. Guo, An epidemiological modelling approach for COVID-19 via data assimilation, *Eur. J. Epidemiol.*, **35** (2020), 749–761. https://doi.org/10.1007/s10654-020-00676-7
11. H. Biegel, J. Lega, EpiCovDA: a mechanistic COVID-19 forecasting model with data assimilation, preprint, arXiv:2105.05471, 2021.
12. M. Daza-Torres, M. Capistrán, A. Capella, J. Christen, Bayesian sequential data assimilation for COVID-19 forecasting, *Epidemics*, **39** (2022), 100564. https://doi.org/10.1016/j.epidem.2022.100564
13. R. Fernández-Naranjo, E. Vásconez-González, K. Simbaã-Rivera, L. Gómez-Barreno, J. Izquierdo-Condoy, D. Cevallos-Robalino, et al., Statistical data driven approach of COVID-19 in Ecuador: R_0 and R_t estimation via new method, *Infect. Dis. Modell.*, **6** (2021), 232–243. https://doi.org/10.1016/j.idm.2020.12.012
14. Dirección General del Registro Civil, Identificación y Cedulación, *Cifras Defunciones*, 2021. Available from: https://www.registrocivil.gob.ec/cifras/.

15. H. Booth, L. Tickle, Mortality modelling and forecasting: A review of methods, *Ann. Actuarial Sci.*, **3** (2008), 3–43. https://doi.org/10.1017/S1748499500000440

16. Ministerio Salud Pública Ecuador, *Actualización de Casos de Coronavirus en Ecuador*, 2021. Available from: https://www.salud.gob.ec/actualizacion-de-casos-de-coronavirus-en-ecuador/.

17. Instituto Nacional de Estadísticas y Censos (INEC), *Proyección Poblacional*, 2021. Available from: https://www.ecuadorencifras.gob.ec/proyecciones-poblacionales/.

18. World Health Organization, *Estimating Mortality from COVID-19*, 2020. Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci-Brief-Mortality-2020.1.

19. J. Feng, H. Luo, Y. Wu, Q. Zhou, R. Qi, A new method for accurate calculation of case fatality rates during a pandemic: Mathematical deduction based on population-level big data, *Infect. Med.*, **2** (2023), 96–104. https://doi.org/10.1016/j.imj.2023.03.002

20. MIDAS Coordination Center, *Midas Online COVID-19 Portal*, 2020. Available from: https://github.com/midas-network/COVID-19/.

21. L. Nikolai, C. Meyer, P. Kremsner, T. Velavan, Asymptomatic SARS Coronavirus 2 infection: Invisible yet invincible, *Int. J. Infect. Dis.*, **100** (2020), 112–116. https://doi.org/10.1016/j.ijid.2020.08.076

22. P. Sah, M. Fitzpatrick, C. Zimmer, E. Abdollahi, L. Juden-Kelly, S. Moghadas, et al., Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis, *Proc. Natl. Acad. Sci.*, **118** (2021), e2109229118. https://doi.org/10.1073/pnas.2109229118

23. G. Syangtan, S. Bista, P. Dawadi, B. Rayamajhee, L. Shrestha, R. Tuladhar, et al., Asymptomatic SARS-CoV-2 carriers: a systematic review and meta-analysis, *Front. Public Health*, **8** (2021), 587374. https://doi.org/10.3389/fpubh.2020.587374

24. Ministerio de Salud Pública del Ecuador, *Informe Epidemiológico de COVID-19, Ecuador 2021, December 07, 2021*, 2021. Available from: https://www.salud.gob.ec/wp-content/uploads/2021/12/MSP_ecu_cvd19_datos_epi_20211207.pdf.

25. R. Yang, X. Gui, Y. Xiong, Comparison of clinical characteristics of patients with asymptomatic vs symptomatic coronavirus disease 2019 in Wuhan, China, *JAMA Network Open*, **3** (2020), e2010182. https://doi.org/10.1001/jamanetworkopen.2020.10182

26. H. Wang, K. Paulson, S. Pease, S. Watson, H. Comfort, P. Zheng, et al., Estimating excess mortality due to the COVID-19 pandemic: A systematic analysis of COVID-19-related mortality, 2020–2021, *Lancet*, **399** (2022), 1513–1536. https://doi.org/10.1016/S0140-6736(21)02796-3

27. M. Peirlinck, K. Linka, F. Costabal, J. Bhattacharya, E. Bendavid, J. Ioannidis, et al., Visualizing the invisible: The effect of asymptomatic transmission on the outbreak dynamics of COVID-19, *Comput. Methods Appl. Mech. Eng.*, **372** (2020), 113410. https://doi.org/10.1016/j.cma.2020.113410

28. V. Bitsouni, N. Gialelis, I. Stratis, A model for the outbreak of COVID-19: Vaccine effectiveness in a case study of italy, in *International Scientific Conference (on) Modern Methods, Problems and Applications of Operator Theory and Harmonic Analysis*, Springer, (2020), 91–107.

29. I. De Falco, A. Della Cioppa, U. Scafuri, E. Tarantino, Differential evolution to estimate the parameters of a seiar model with dynamic social distancing: the case of COVID-19 in Italy, in *Data Science for COVID-19*, Elsevier, (2021), 75–90. https://doi.org/10.1016/B978-0-12-824536-1.00005-8
30. V. M. Alekseev, V. M. Tikhomirov, S. V. Fomin, *Optimal Control*, Springer Science and Business Media, 1987.
31. J. C. De los Reyes, *Numerical PDE-Constrained Optimization*, Springer Verlag, 2015. https://doi.org/10.1007/978-3-319-13395-9
32. D. Kennedy, V. K. Vu, H. Ritchie, R. Bartlein, O. Rothschild, D. Bausch, et al., COVID-19: Identifying countries with indicators of success in responding to the outbreak, *Gates Open Res.*, **4** (2021), 62.
33. Our World in Data, *Daily COVID-19 Tests per 1000 People*, 2024. Available from: https://ourworldindata.org/grapher/daily-tests-per-thousand-people-smoothed-7-day.
34. J. Helton, F. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, *Reliab. Eng. Syst. Saf.*, **81** (2003), 23–69. https://doi.org/10.1016/S0951-8320(03)00058-9
35. C. T. Kelley, *Iterative Methods for Optimization*, SIAM, 1999. https://doi.org/10.1137/1.9781611970920
36. El Universo, *Positividad en Pruebas de Detección Para COVID-19, a Nivel Nacional*, 2022. Available from: https://www.eluniverso.com/noticias/ecuador/la-positividad-en-pruebas-para-detectar-COVID-19-ha-ido-en-aumento-desde-noviembre-en-el-pais-nota/.
37. J. Machado, *38 de Cada 100 Pruebas Tomadas en Ecuador dan Positivo por Rebrote de COVID-19*, 2022. Available from: https://www.primicias.ec/noticias/sociedad/pruebas-ecuador-COVID-positividad-omicron/.
38. S. España, *Quito Enfrenta un Pico de Coronavirus sin Camas de Cuidados Intensivos Disponibles*, 2020. Available from: https://elpais.com/sociedad/2020-07-14/quito-enfrenta-un-pico-de-coronavirus-sin-camas-de-cuidados-intensivos-disponibles.html.
39. B. Ridenhour, J. M. Kowalik, D. K. Shay, El número reproductivo básico (R_0): Consideraciones para su aplicación en la salud pública, *Am. J. Public Health*, **104** (2014), e32–e41. https://doi.org/10.2105/AJPH.2013.301704
40. R. Moore, R. Kearfott, M. Cloud, *Introduction to Interval Analysis*, SIAM, 2009. https://doi.org/10.1137/1.9780898717716
41. D. Sotiropoulos, K. Tserpes, Interval-based computation of the uncertainty in the mechanical properties and the failure analysis of unidirectional composite materials, *Math. Comput. Appl.*, **27** (2022), 38. https://doi.org/10.3390/mca27030038
42. Z. Qiu, Y. Ma, X. Wang, Comparison between non-probabilistic interval analysis method and probabilistic approach in static response problem of structures with uncertain-but-bounded parameters, *Commun. Numer. Methods Eng.*, **20** (2004), 279–290. https://doi.org/10.1002/cnm.668
43. A. Zemkoho, A basic time series forecasting course with Python, *Oper. Res. Forum*, **4** (2022), 2. https://doi.org/10.1007/s43069-022-00179-z

44. P. Brockwell, R. Davis, *Introduction to Time Series and Forecasting*, Springer, 2002. https://doi.org/10.1007/b97391

45. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, *Bayesian Data Analysis*, Chapman and Hall/CRC, 1995.

Appendix

A. Details of excess mortality estimation

As discussed in Section 2.1, one of the key data sources used in this study relies on estimating COVID-19 infections from excess mortality data provided by the Ecuadorian Civil Registry (RC). Calculating excess mortality requires a reliable forecast of the expected number of deaths that would have occurred under normal circumstances. This appendix provides the methodology, based on a seasonal autoregressive integrated moving average (SARIMA) time series model, used to generate these forecasts and validate the model's performance.

The general seasonal ARIMA model is denoted by $ARIMA(p, d, q)(P, D, Q, s)$. This notation represents the combination of non-seasonal and seasonal components:

- Non-seasonal part (p, d, q)
  - p : The order of the non-seasonal autoregressive (AR) component, indicating how many past values in the time series affect the present.
  - d : The order of non-seasonal integrated (I) differencing, representing the number of times the raw observations are differenced to achieve stationarity.
  - q : The order of the non-seasonal moving average (MA) component, indicating how many past forecast errors affect the present
- Seasonal part (P, D, Q, s)
  - P : The order of the seasonal autoregressive component.
  - D : The order of seasonal integrated differencing.
  - Q : The order of the seasonal moving average component.
  - s : The length of the seasonal cycle (e.g., $s = 7$ for daily data with weekly seasonality, $s = 52$ for weekly data with yearly seasonality).

First, we determine the necessary order of differencing to ensure stationarity. Examining the original series and its differences, Figure A1 indicates non-stationarity in the original series. Therefore, we initially set the parameter $d = 1$. It is worth mentioning another time series model for Ecuador mortality data during the COVID-19 pandemic that we are aware of [3], where the authors proposed an $ARIMA(3, 0, 1)$ model, overlooking the non-stationary nature of the series.

One way to obtain preliminary values for p and q from the data is to observe the partial autocorrelation function (PACF) and autocorrelation function (ACF) plots. An $AR(p)$ process is obtained if there is a significant spike at lag p in the PACF plot and there is no larger one beyond [43]. Likewise, we have a $MA(q)$ if such a significant spike is obtained at lag q in the ACF plot. In our case, according to the ACP and PACF plots of the first difference time series (see the second row of Figure A1), we set $p = q = 1$ as preliminary values.

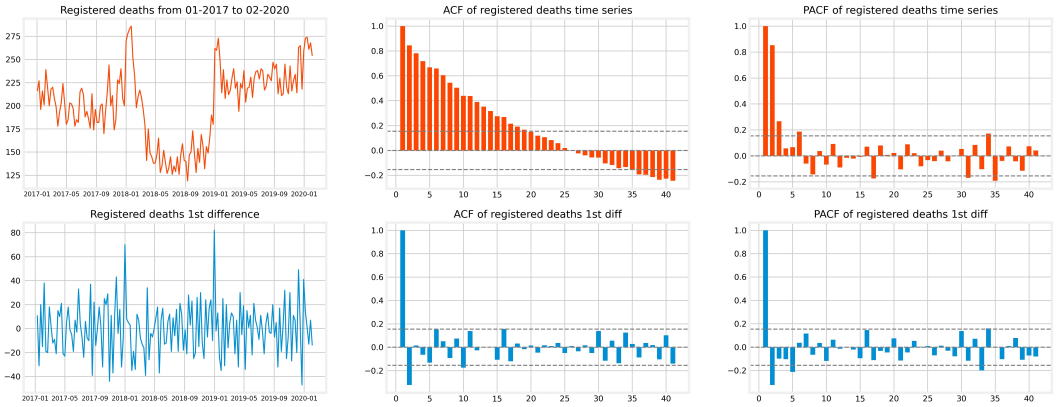


Figure A1. The first row presents the series of registered deaths in Pichincha from January 2, 2017, to February 3, 2020, and their ACF and PACF plots. The second row presents the series after performing one difference and their corresponding ACF and PACF plots.

The above procedure enables us to estimate the parameters of pure AR and MA models; therefore, rigorously verifying if another $ARIMA(p, 1, q)$ model with $p > 0$ and $q > 0$ performs better is necessary. We conduct the latter using the Akaike information criterion (AIC), i.e., we compare which parameter combination provides the lowest AIC value. By running this test, the best model is $ARIMA(1, 1, 2)$. Likewise, for the seasonal components of the model, we set $(P, D, Q, s) = (1, 1, 0, 52)$. The seasonality is set at $s = 52$ weeks. Similarly, in our case, $D = 1$. This model considers the relationship between the current observation and the observation from the same period in the previous year ($P = 1$). Additionally, we set $Q = 0$, indicating that no past seasonal errors are considered. After the model identification, $ARIMA(1, 1, 2)(1, 1, 0, 52)$, we must ensure that its residuals are uncorrelated and normally distributed. Figure A2 shows the statistics assessing the quality of the model. Observe that if the number of parameters were constant across all compared models, then minimizing AIC is equivalent to maximizing the likelihood function of the model. However, order selection typically involved comparing models where the number of parameters varied (due to different choices of p, q, P, Q). Then, this specific equivalence generally did not apply.

Visual inspection of diagnostic plots for Pichincha's residuals supports normality. The normal Q-Q plot (Figure A2, bottom left) shows residual quantiles closely aligning with theoretical normal quantiles, and the histogram (Figure A2, top right) displays an approximately symmetric, bell-shaped distribution. Formal testing corroborates this: the Jarque-Bera (JB) test for Pichincha yields a p-value of 0.49, well above the typical 0.05 significance level, leading to the non-rejection of the normality hypothesis. The observed skewness of -0.23 and kurtosis of 2.36 are consistent with this finding as they do not indicate a significant deviation. Similarly, for Guayas' time series, the Jarque-Bera test ($p = 0.41$) also leads to the non-rejection of normality, a conclusion supported by its skewness of 0.41 and kurtosis of 3.31.

B. Details of variational data assimilation and optimization

B.1. Lagrangian approach

As the constraints of problem (3.7) consist of a system of differential equations, it must be addressed as an optimal control problem. In this framework, the state vector $\mathbf{u} = (S, E, I^r, I^u, R)$ belongs to the space of absolutely continuous functions $AC([t_0, t_n], \mathbb{R}^5)$. This choice is justified by the measurability of the right-hand side of Eqs (3.7b)–(3.7f) with respect to time for a fixed $\mathbf{u}$, its differentiability with respect to $\mathbf{u}$ for any t, and the boundedness of this derivative in compact sets; see, e.g., [30, Section 2.5.1].

We compute its first-order optimality system by using the Lagrangian approach, with the Lagrangian given by:

$$\begin{aligned}\mathcal{L}(\mathbf{u}, \rho, \mathbf{u}_0, \boldsymbol{\lambda}, \boldsymbol{\theta}) = J(\mathbf{u}, \rho, \mathbf{u}_0) + \int_{t_0}^{t_n} \Bigg(& \lambda_S \left(-\frac{dS}{dt} - \frac{\beta S I^r + \mu\beta S I^u}{N} \right) + \lambda_I \left(-\frac{dI^r}{dt} + \alpha \frac{E}{Z} - \frac{I^r}{D} \right) \\ & + \lambda_E \left(-\frac{dE}{dt} + \frac{\beta S I^r + \mu\beta S I^u}{N} - \frac{E}{Z} \right) + \lambda_U \left(-\frac{dI^u}{dt} + (1-\alpha)\frac{E}{Z} - \frac{I^u}{D} \right) \\ & + \lambda_R \left(-\frac{dR}{dt} + \frac{I^r + I^u}{D} \right) \Bigg) dt + \theta_S (N - S(t_0) - E_0 - I_0^r - I_0^u - R_0) \\ & + \theta_E (E_0 - E(t_0)) + \theta_I (I_0^r - I^r(t_0)) + \theta_U (I_0^u - I^u(t_0)) + \theta_R (R_0 - R(t_0)),\end{aligned} \tag{L}$$

where $\boldsymbol{\lambda} = (\lambda_S, \lambda_E, \lambda_I, \lambda_U, \lambda_R) \in AC([t_0, t_n], \mathbb{R}^5)$ is the Lagrange multiplier associated with the problem's differential equations constraints, and $\boldsymbol{\theta} = (\theta_S, \theta_E, \theta_I, \theta_U, \theta_R) \in \mathbb{R}^5$ is the multiplier associated with the initial condition of problem (3.7).

Let U be an open set in $\mathbb{R} \times \mathbb{R}^5 \times \mathbb{R}^5$ and, for each $t \in [t_0, t_n]$, let the function $\varphi(t, \mathbf{u}(t), \rho) \colon U \longrightarrow \mathbb{R}^5$ denote the right-hand side of Eqs (3.7b)–(3.7f). Additionally, let us write the problem's initial condition as

$$\begin{bmatrix} -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{u}_0 + \begin{bmatrix} N \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} =: \mathbf{M}\mathbf{u}_0 + m.$$

Then, the mapping $e \colon \mathcal{U} \longrightarrow L^1([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5$ determined by

$$\mathcal{U} \ni (\mathbf{u}, \rho, \mathbf{u}_0) \mapsto e(\mathbf{u}, \rho, \mathbf{u}_0) = \left(\frac{d\mathbf{u}}{dt} - \varphi(\cdot, \mathbf{u}, \rho) \,, \quad \mathbf{u}(t_0) - \mathbf{M}\mathbf{u}_0 - m \right),$$

where $\mathcal{U} = \{(\mathbf{u}(\cdot), \rho, \mathbf{u}_0) : (t, \mathbf{u}(t), \rho) \in U, t_0 \le t \le t_n, \mathbf{u}_0 \in \mathbb{R}^4_+\} \subset AC([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5 \times \mathbb{R}^4$, defines a superposition operator, which is differentiable in $\mathcal{U}$; see Proposition 3 in [30, Section 2.4.1]. Moreover, $e'(\mathbf{u}, \rho, \mathbf{u}_0)$ defines the following linear and continuous mapping:

$$\begin{aligned} e'(\mathbf{u}, \rho, \mathbf{u}_0) \colon AC([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5 \times \mathbb{R}^4 &\longrightarrow L^1([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5 \\ \mathbf{h} = (h, h_\rho, h_0) \mapsto e'(\mathbf{u}, \rho, \mathbf{u}_0)\mathbf{h} &= \begin{bmatrix} \frac{d(\cdot)}{dt} - \varphi_{\mathbf{u}} & -\varphi_\rho & \mathbf{0} \\ \delta_{t_0}(\cdot) & \mathbf{0} & -\mathbf{M} \end{bmatrix} \mathbf{h} \end{aligned}$$

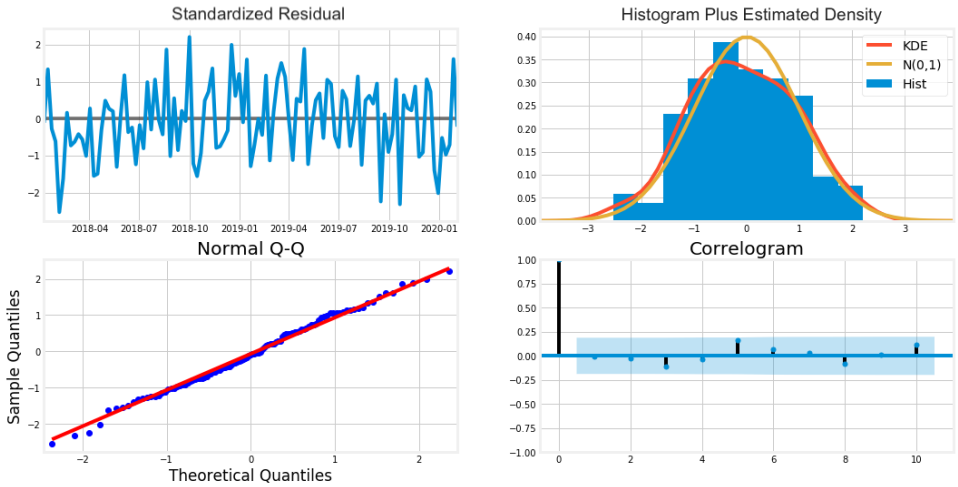


Figure A2. Pichincha's diagnostic graphics. The histogram plus estimated density of the model's residual plot (top right) serves to visually assess the normality assumption for the model residuals. The residual plot (top left) does not display any seasonality. This is confirmed by the correlogram plot (bottom right), which shows that the residuals have a low correlation with lagged versions of themselves. The normal Q–Q plot (bottom left) shows that the ordered distribution of residuals (blue dots) follows the linear trend of the samples taken from an $\mathcal{N}(0,1)$ distribution.

where $\delta_{t_0}(\cdot)$ denotes the evaluation operator, that is, $\delta_{t_0}(v) = v(t_0)$ for some v. Consequently,

$$e'(\mathbf{u},\rho,\mathbf{u}_0)\mathbf{h}(t) = \left(\frac{dh(t)}{dt} - \varphi_{\mathbf{u}}(t)h(t) - \varphi_\rho(t)h_\rho, \quad h(t_0) - \mathbf{M}h_0\right). \tag{B.1}$$

Note that, for each t, the partial derivatives of φ with respect to $\mathbf{u}$ and ρ are given by

$$\varphi_{\mathbf{u}}(t) = \begin{bmatrix} \frac{-\beta I^r - \mu\beta I^u}{N} & 0 & -\frac{\beta S}{N} & -\frac{\mu\beta S}{N} & 0 \\ \frac{\beta I^r + \mu\beta I^u}{N} & -\frac{1}{Z} & \frac{\beta S}{N} & \frac{\mu\beta S}{N} & 0 \\ 0 & \frac{\alpha}{Z} & -\frac{1}{D} & 0 & 0 \\ 0 & \frac{1-\alpha}{Z} & 0 & -\frac{1}{D} & 0 \\ 0 & 0 & \frac{1}{D} & \frac{1}{D} & 0 \end{bmatrix}, \varphi_\rho(t) = \begin{bmatrix} \frac{-S I^r - \mu S I^u}{N} & -\frac{\beta S I^u}{N} & 0 & 0 & 0 \\ \frac{S I^r + \mu S I^u}{N} & \frac{\beta S I^u}{N} & 0 & \frac{E}{Z^2} & 0 \\ 0 & 0 & \frac{E}{Z} & -\frac{\alpha E}{Z^2} & \frac{I^r}{D^2} \\ 0 & 0 & -\frac{E}{Z} & -\frac{(1-\alpha)E}{Z^2} & \frac{I^u}{D^2} \\ 0 & 0 & 0 & 0 & -\frac{I^r + I^u}{D^2} \end{bmatrix}.$$

The existence of Lagrange multipliers is guaranteed if

$$Im\left[e'(\mathbf{u},\rho,\mathbf{u}_0)\right] = \{(y(\cdot),\theta) \in L^1([t_0,t_n],\mathbb{R}^5)\times\mathbb{R}^5 : \exists (h(\cdot),h_\rho,h_0) \in AC([t_0,t_n],\mathbb{R}^5)\times\mathbb{R}^5\times\mathbb{R}^4$$
$$\text{such that } e'(\mathbf{u},\rho,\mathbf{u}_0)[h(\cdot),h_\rho,h_0] = (y(\cdot),\theta)\}$$

is a closed subspace of $L^1([t_0,t_n],\mathbb{R}^5)\times\mathbb{R}^5$; see [30, Section 3.2]. Note that for each $(y(\cdot),\theta) \in L^1([t_0,t_n],\mathbb{R}^5)\times\mathbb{R}^5$, the equation $e'(\mathbf{u},\rho,\mathbf{u}_0)[h,h_\rho,h_0](t) = (y(t),\theta)$ is equivalent to the linear differential equation, with continuous coefficients,

$$\begin{cases} \frac{dh(t)}{dt} - \varphi_{\mathbf{u}}(t)h(t) - \varphi_\rho(t)h_\rho & = y(t), \\ h(t_0) - \mathbf{M}h_0 & = \theta, \end{cases} \tag{B.2}$$

which according to [30, Section 2.5.4] possesses a solution $h(\cdot) \in AC([t_0, t_n], \mathbb{R}^5)$. Consequently, the mapping $e'(\mathbf{u}, \rho, \mathbf{u}_0)$ is surjective. Therefore, its image coincides with $L^1([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5$, from where we can conclude its closedness. On the other hand, observe that for fixed $h_\rho \in \mathbb{R}^5$ and $h_0 \in \mathbb{R}^4$, and for each $(y(\cdot), \theta) \in L^1([t_0, t_n], \mathbb{R}^5) \times \mathbb{R}^5$, the linear differential equation:

$$\begin{cases} \frac{dh(t)}{dt} - \varphi_{\mathbf{u}}(t)h(t) & = y(t) + \varphi_\rho(t)h_\rho, \\ h(t_0) & = \theta + \mathbf{M}h_0 \end{cases}$$

has a unique solution, i.e., the operator $e_{\mathbf{u}}(\mathbf{u}, \rho, \mathbf{u}_0)$ is bijective. As a result, the bijectivity of $e_{\mathbf{u}}(\mathbf{u}, \rho, \mathbf{u}_0)$ ensures the existence of Lagrange multipliers.

To obtain the adjoint equation, we compute the derivative of the Lagrangian with respect to the state variable $\mathbf{u} = (S, E, I^r, I^u, R)$ in an arbitrary direction $\boldsymbol{\nu} = (\sigma, \zeta, w, v, r) \in AC([t_0, t_n], \mathbb{R}^5)$ and set the result equal to zero (see, e.g., [30, 31]). That is, using additional integration by parts,

$$\nabla_S \mathcal{L}[\sigma] = \int_{t_0}^{t_n} \left[\frac{d\lambda_S}{dt} + \left(\frac{\beta I^r}{N} + \frac{\mu\beta I^u}{N} \right)(\lambda_E - \lambda_S) \right] \sigma \, dt - \lambda_S(t_n)\sigma(t_n) + \sigma(t_0)(\lambda_S(t_0) + \theta_S) = 0$$

$$\nabla_E \mathcal{L}[\xi] = \int_{t_0}^{t_n} \left[\frac{d\lambda_E}{dt} + \alpha\frac{\lambda_I}{Z} + (1-\alpha)\frac{\lambda_U}{Z} - \frac{\lambda_E}{Z} \right] \xi \, dt - \lambda_E(t_n)\xi(t_n) + \xi(t_0)(\lambda_E(t_0) + \theta_E) = 0$$

$$\begin{aligned} \nabla_{I^r} \mathcal{L}[w] = \int_{t_0}^{t_n} \left[\frac{d\lambda_I}{dt} + \frac{\partial J}{\partial I^r} - \frac{\beta S\,\lambda_S}{N} + \frac{\beta S\,\lambda_E}{N} - \frac{\lambda_I}{D} + \frac{\lambda_R}{D} \right] w \, dt + w(t_n)(C_{nn}^{-1}(I^r(t_n) - I_n^{obs}) \\ - \lambda_I(t_n)) + w(t_0)(\lambda_I(t_0) + \theta_I) = 0 \end{aligned}$$

$$\nabla_{I^u} \mathcal{L}[v] = \int_{t_0}^{t_n} \left[\frac{d\lambda_U}{dt} - \frac{\mu\beta S\,\lambda_S}{N} + \frac{\mu\beta S\,\lambda_E}{N} - \frac{\lambda_U}{D} + \frac{\lambda_R}{D} \right] v \, dt - v(t_n)\lambda_U(t_n) + v(t_0)(\lambda_U(t_0) + \theta_U) = 0$$

$$\nabla_R \mathcal{L}[r] = \int_{t_0}^{t_n} \frac{d\lambda_R}{dt} r \, dt - r(t_n)\lambda_R(t_n) + r(t_0)(\lambda_R(t_0) + \theta_R) = 0.$$

Since the direction $\boldsymbol{\nu}$ was taken arbitrarily, the latter yields the following adjoint system:

$$-\frac{d\lambda_S}{dt} = \left(\frac{\beta I^r}{N} + \frac{\mu\beta I^u}{N} \right)(\lambda_E - \lambda_S), \qquad \lambda_S(t_n) = 0 \tag{B.3a}$$

$$-\frac{d\lambda_E}{dt} = -\frac{1}{Z}\lambda_E + \frac{\alpha}{Z}\lambda_I + \frac{(1-\alpha)}{Z}\lambda_U, \qquad \lambda_E(t_n) = 0 \tag{B.3b}$$

$$-\frac{d\lambda_I}{dt} = \frac{\partial J}{\partial I^r} - \frac{\beta S}{N}(\lambda_S - \lambda_E) - \frac{\lambda_I}{D} + \frac{\lambda_R}{D}, \qquad \lambda_I(t_n) = C_{nn}^{-1}(I^r(t_n) - I_n^{obs}) \tag{B.3c}$$

$$-\frac{d\lambda_U}{dt} = -\frac{\mu\beta S}{N}(\lambda_S - \lambda_E) - \frac{\lambda_U}{D} + \frac{\lambda_R}{D}, \qquad \lambda_U(t_n) = 0 \tag{B.3d}$$

$$-\frac{d\lambda_R}{dt} = 0, \qquad \lambda_R(t_n) = 0. \tag{B.3e}$$

Additionally, the multiplier $\boldsymbol{\theta}$ takes the value of the adjoint at the initial time $\boldsymbol{\lambda}(t_0)$, that is:

$$\theta_S = \lambda_S(t_0), \quad \theta_E = \lambda_E(t_0), \quad \theta_I = \lambda_I(t_0), \quad \theta_U = \lambda_U(t_0), \quad \theta_R = \lambda_R(t_0).$$

Observe that from (B.3e), we obtain that $\lambda_R(t) = 0$ for each $t \in [t_0, t_n]$, and $\theta_R = 0$, which further simplifies the above adjoint system. Furthermore, the gradient of the objective function is obtained by differentiating the Lagrangian with respect to the parameters $\rho := (\beta, \mu, \alpha, Z, D)^T$ and with respect to the initial condition $\mathbf{u}_0 = (E_0, I_0^r, I_0^u, R_0)$. Note that since four out of five initial condition variables were set, the remaining one, corresponding to the susceptible population, is computed as $S_0 = N - E_0 - I_0^r - I_0^u - R_0$. Proceeding like this and computing first the derivate with respect to the parameters, we obtain:

$$\nabla_\beta J = \mathbf{B}_\beta^{-1}(\beta - \beta^b) - \int_{t_0}^{t_n} \frac{S}{N}\left(I^r + \mu I^u\right)(\lambda_S - \lambda_E)\, dt, \tag{B.4a}$$

$$\nabla_\mu J = \mathbf{B}_\mu^{-1}(\mu - \mu^b) - \int_{t_0}^{t_n} \beta \frac{S}{N} I^u\,(\lambda_S - \lambda_E)\, dt, \tag{B.4b}$$

$$\nabla_\alpha J = \mathbf{B}_\alpha^{-1}(\alpha - \alpha^b) + \int_{t_0}^{t_n} \frac{E}{Z}(\lambda_I - \lambda_u) dt, \tag{B.4c}$$

$$\nabla_Z J = \mathbf{B}_Z^{-1}(Z - Z^b) + \int_{t_0}^{t_n} \frac{E}{Z^2}(\lambda_E - \alpha\lambda_I - (1-\alpha)\lambda_U) dt, \tag{B.4d}$$

$$\nabla_D J = \mathbf{B}_D^{-1}(D - D^b) + \int_{t_0}^{t_n} \frac{1}{D^2}\left(I^r\lambda_I + I^u\lambda_U - \lambda_R(I^r + I^u)\right) dt. \tag{B.4e}$$

Likewise, computing the derivative with respect to the initial conditions, we get:

$$\nabla_{E_0} J = \left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_1 - \lambda_S(t_0) + \lambda_E(t_0), \tag{B.5a}$$

$$\nabla_{I_0^r} J = \left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_2 - \lambda_S(t_0) + \lambda_I(t_0), \tag{B.5b}$$

$$\nabla_{I_0^u} J = \left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_3 - \lambda_S(t_0) + \lambda_U(t_0), \tag{B.5c}$$

$$\nabla_{R_0} J = \left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_4 - \lambda_S(t_0) + \lambda_R(t_0), \tag{B.5d}$$

where $\left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_i$ represents the i-component of the vector $\left(\mathbf{Q}_0^{-1}(\mathbf{u}_0 - \mathbf{u}_0^b)\right)_{1\times 4}^T$.

B.2. Sufficient second-order optimality condition

To analyze second-order optimality conditions, we use the Hamiltonian associated to the problem, i.e.,

$$\begin{aligned} H\colon \mathbb{R}^5 \times \mathbb{R}^4 \times \mathbb{R}^5 \times \mathbb{R}^5 &\longrightarrow \mathbb{R} \\ (\rho, \mathbf{u}_0, \mathbf{u}(t), \boldsymbol{\lambda}(t)) &\mapsto H(\rho, \mathbf{u}_0, \mathbf{u}(t), \boldsymbol{\lambda}(t)), \end{aligned}$$

where, for a.e. $t \in (t_0, t_n)$,

$$H(\rho, \mathbf{u}_0, \mathbf{u}(t), \boldsymbol{\lambda}(t)) = J(\rho, \mathbf{u}_0) + \lambda_S(t)\left(-\frac{\beta S(t) I^r(t) + \mu\beta S(t) I^u(t)}{N}\right) + \lambda_I(t)\left(\alpha\frac{E(t)}{Z} - \frac{I^r(t)}{D}\right)$$
$$+ \lambda_E(t)\left(\frac{\beta S(t) I^r(t) + \mu\beta S(t) I^u(t)}{N} - \frac{E(t)}{Z}\right) + \lambda_U(t)\left((1-\alpha)\frac{E(t)}{Z} - \frac{I^u(t)}{D}\right) + \lambda_R(t)\left(\frac{I^r(t) + I^u(t)}{D}\right).$$

Observe that, by integrating by parts, we can rewrite the Lagrangian, using the Hamiltonian, as follows:

$$\mathcal{L}(\mathbf{u}, \rho, \mathbf{u}_0, \boldsymbol{\lambda}, \boldsymbol{\theta}) = \int_{t_0}^{t_n}\left(H(\mathbf{u}, \rho, \mathbf{u}_0, \mathbf{u}(t), \boldsymbol{\lambda}(t)) + \frac{d\boldsymbol{\lambda}(t)}{dt}\mathbf{u}(t)\right)dt$$
$$- \boldsymbol{\lambda}(t_n)\mathbf{u}(t_n) + \boldsymbol{\lambda}(t_0)\mathbf{u}(t_0) + \boldsymbol{\theta}(u_0 - \mathbf{u}(t_0)).$$

Consequently, the second derivative of the Lagrangian in a direction $[h, h_\rho, h_0]$ can be expressed by using the second derivative of the Hamiltonian, that is,

$$\nabla^2_{\mathbf{u},\rho,\mathbf{u}_0}\mathcal{L}[(h, h_\rho, h_0), (h, h_\rho, h_0)] = \int_{t_0}^{t_n}\nabla^2_{\mathbf{u}}H[h,h]dt + \int_{t_0}^{t_n}\nabla^2_{\rho}H[h_\rho,h_\rho]dt$$
$$+ \int_{t_0}^{t_n}\nabla^2_{\mathbf{u}_0}H[h_0,h_0]dt + 2\int_{t_0}^{t_n}\nabla^2_{(\mathbf{u},\rho)}H[h,h_\rho]dt, \quad \text{(B.6)}$$

since $\nabla^2_{(\mathbf{u},\mathbf{u}_0)}H = \nabla^2_{(\rho,\mathbf{u}_0)}H = \mathbf{0}_{5\times4}$. Moreover $\nabla^2_{\mathbf{u}_0}H = \mathbf{Q}_0^{-1}$, and $\nabla^2_{\mathbf{u}}H$ and $\nabla^2_{(\mathbf{u},\rho)}H$ have the form:

$$\nabla^2_{\mathbf{u}}H = \begin{bmatrix} 0 & 0 & \frac{\beta}{N}(\lambda_E(t) - \lambda_S(t)) & \frac{\mu\beta}{N}(\lambda_E(t) - \lambda_S(t)) & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{\beta}{N}(\lambda_E(t) - \lambda_S(t)) & 0 & \frac{\partial^2 J}{\partial I_r^2} & 0 & 0 \\ \frac{\mu\beta}{N}(\lambda_E(t) - \lambda_S(t)) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\nabla^2_{(\mathbf{u},\rho)}H = \begin{bmatrix} \nabla^2_{(S,\beta)}H & \nabla^2_{(S,\mu)}H & 0 & 0 & 0 \\ 0 & 0 & \nabla^2_{(E,\alpha)}H & \nabla^2_{(E,Z)}H & 0 \\ \nabla^2_{(I^r,\beta)}H & 0 & 0 & 0 & \nabla^2_{(I^r,D)}H \\ \nabla^2_{(I^u,\beta)}H & \nabla^2_{(I^u,\mu)}H & 0 & 0 & \nabla^2_{(I^u,D)}H \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

with

$$\nabla^2_{(S,\beta)}H = \left[\frac{I^r}{N}(\lambda_E(t) - \lambda_S(t)) + \frac{\mu I^u(t)}{N}(\lambda_E(t) - \lambda_S(t))\right],$$
$$\nabla^2_{(E,Z)}H = \frac{1}{Z^2}[\lambda_E(t) - \alpha\lambda_I(t) - (1-\alpha)\lambda_U(t)], \qquad \nabla^2_{(E,\alpha)}H = \frac{1}{Z}(\lambda_I(t) - \lambda_U(t)),$$
$$\nabla^2_{(I^r,\beta)}H = \frac{S}{N}(\lambda_E(t) - \lambda_S(t)), \qquad \nabla^2_{(I^r,D)}H = \frac{1}{D^2}(\lambda_I(t) - \lambda_R(t)),$$

$$\nabla^2_{(I^u,\beta)}H = \frac{\mu S}{N}(\lambda_E(t) - \lambda_S(t)), \qquad \nabla^2_{(I^u,\mu)}H = \frac{\beta S}{N}(\lambda_E(t) - \lambda_S(t)),$$
$$\nabla^2_{(I^u,D)}H = \frac{1}{D^2}(\lambda_U(t) - \lambda_R(t)), \qquad \nabla^2_{(S,\mu)}H = \frac{\beta I^u(t)}{N}(\lambda_E(t) - \lambda_S(t)).$$

Likewise, $\nabla^2_\rho H$ has the form:

$$\begin{pmatrix} \mathbf{B}_\beta^{-1} & \frac{S(t)I(t)^u}{N}(\lambda_E(t) - \lambda_S(t)) & 0 & 0 & 0 \\ \frac{S(t)I(t)^u}{N}(\lambda_E(t) - \lambda_S(t)) & \mathbf{B}_\mu^{-1} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{B}_\alpha^{-1} & \frac{E(t)}{Z^2}(\lambda_U(t) - \lambda_I(t)) & 0 \\ 0 & 0 & \frac{E(t)}{Z^2}(\lambda_U(t) - \lambda_I(t)) & \nabla^2_Z H & 0 \\ 0 & 0 & 0 & 0 & \nabla^2_D H \end{pmatrix}$$

with

$$\nabla^2_Z H = \mathbf{B}_Z^{-1} - \frac{2E(t)}{Z^3}[\lambda_E(t) - \alpha\lambda_I(t) - (1-\alpha)\lambda_U(t)],$$
$$\nabla^2_D H = \mathbf{B}_D^{-1} - \frac{2I^r(t)}{D^3}[\lambda_I(t) - \lambda_R(t)] - \frac{2I^u(t)}{D^3}[\lambda_U(t) - \lambda_R(t)].$$

Before evaluating the quantities in (B.6), and since most terms depend on the adjoint variable $\boldsymbol{\lambda}(t)$, we derive an upper bound for its norm. By writing $V(t) = \varphi_u(t)^T$ and transforming the time variable $t \to t_n - t + t_0$, the adjoint system (B.3) can be written forward in time as follows:

$$\frac{d\boldsymbol{\lambda}(t)}{dt} - V(t)\boldsymbol{\lambda}(t) = \mathbf{b}, \qquad \boldsymbol{\lambda}(t_0) = \boldsymbol{\lambda}_0, \tag{B.7}$$

where $\mathbf{b} = (0, 0, \frac{\partial J}{\partial I^r}, 0, 0)^T$, with $\frac{\partial J}{\partial I^r} = \sum_i (I^r(t_i) - I_i^{obs})C_{ii}^{-1}$. Moreover, $\boldsymbol{\lambda}_0 = (0, 0, C_{nn}^{-1}(I^r(t_n) - I_n^{obs}), 0, 0)^T$. While the explicit solution $\boldsymbol{\lambda}(t)$ could be found by using the integrating factor method, obtaining a bound is sufficient for our purpose, which is obtaining the sufficient second-order conditions. To achieve this, we start with the integral representation of the solution to (B.7):

$$\boldsymbol{\lambda}(t) = \boldsymbol{\lambda}_0 + \int_{t_0}^{t_n} V(\tau)\boldsymbol{\lambda}(\tau)d\tau + \int_{t_0}^{t_n} \mathbf{b}\, d\tau.$$

Computing the norm of the above expression and using Gronwall's inequality gives the upper bound:

$$\begin{aligned} \|\boldsymbol{\lambda}(t)\| &\leq (\|\boldsymbol{\lambda}_0\| + \|\mathbf{b}\|(t_n - t_0))\, exp\left(\int_{t_0}^{t} \|V(\tau)\|d\tau\right) \\ &\leq \kappa\left(\left|\mathbf{C}_{nn}^{-1}(I^r(t_n) - I_n^{obs})\right| + (t_n - t_0)\sum_i \left|\mathbf{C}_{ii}^{-1}(I^r(t_i) - I_i^{obs})\right|\right), \end{aligned} \tag{B.8}$$

for some $\kappa > 0$, which exists since $V(t)$ is bounded in compact sets. Observe that if the value of the observation is near the solution variable I^r in a given time window, weighted by the observation error covariance term, then the adjoint variable $\boldsymbol{\lambda}(t)$ will be sufficiently small in norm, for a.e. $t \in (t_0, t_n)$.

The closeness between observations and solutions is fundamental for obtaining second-order conditions. In fact, for sufficiently small values of $\boldsymbol{\lambda}(t) = (\lambda_S(t), \lambda_E(t), \lambda_I(t), \lambda_U(t), \lambda_R(t))$ for a.e. $t \in (t_0, t_n)$,

the matrix $\nabla^2_\rho H$ is positive definite since the terms of the diagonal matrix $\mathbf{B}^{-1}$ are strictly positive. Therefore, there exists $\delta_1 > 0$ such that $\nabla^2_\rho H[h_\rho, h_\rho] \geq \delta_1 \|h_\rho\|^2$, for every $h_\rho \in \mathbb{R}^5$. Consequently,

$$\int_{t_0}^{t_n} \nabla^2_\rho H[h_\rho, h_\rho] dt \geq \int_{t_0}^{t_n} \delta_1 \|h_\rho\|^2 dt = (t_n - t_0)\delta_1 \|h_\rho\|^2. \tag{B.9}$$

Likewise, since $\mathbf{Q}_0^{-1}$ is positive definite, there exists $\delta_2 > 0$ such that $\nabla^2_{\mathbf{u}_0} H[h_0, h_0] \geq \delta_2 \|h_0\|^2$ for every $h_0 \in \mathbb{R}^4$, and

$$\int_{t_0}^{t_n} \nabla^2_{\mathbf{u}_0} H[h_0, h_0] dt \geq (t_n - t_0)\delta_2 \|h_0\|^2. \tag{B.10}$$

Observe that (B.10) and (B.9) hold for every $(h_\rho, h_0) \in \mathbb{R}^5 \times \mathbb{R}^4$. Likewise, considering

$$h(\cdot) = (h_1(\cdot), h_2(\cdot), h_3(\cdot), h_4(\cdot), h_5(\cdot)) \in AC([t_0, t_n], \mathbb{R}^5),$$

for sufficiently small values of $\boldsymbol{\lambda}(t)$ and for a.e. $t \in (t_0, t_n)$, given that the terms of the diagonal matrix $\mathbf{C}^{-1}$ are strictly positive, it holds that $\nabla^2_{\mathbf{u}} H[h, h] = \sum_i h_3(t_i)^2 \mathbf{C}^{-1}_{ii} \geq 0$, and

$$\int_{t_0}^{t_n} \nabla^2_{\mathbf{u}} H[h, h] dt \geq 0. \tag{B.11}$$

Moreover, analyzing the matrix $\nabla^2_{(\mathbf{u},\rho)} H$, it holds that

$$\left|\nabla^2_{(\mathbf{u},\rho)} H[h, h_\rho]\right| \leq c \|\boldsymbol{\lambda}(t)\| \|h(t)\| \|h_\rho\|, \qquad \text{for some constant } c > 0.$$

Additionally, if $\mathbf{h} = (h, h_\rho, h_0)$ satisfies $e'(\mathbf{u}, \rho, \mathbf{u}_0)\mathbf{h} = (0, 0)$, i.e.,

$$\frac{dh(t)}{dt} - \varphi_{\mathbf{u}}(t)h(t) - \varphi_\rho(t)h_\rho = 0, \qquad h(t_0) - \mathbf{M}h_0 = 0, \tag{B.12}$$

we can obtain a bound of $\left|\nabla^2_{(\mathbf{u},\rho)} H[h, h_\rho]\right|$ in terms of $\|h_\rho\|$ and $\|h_0\|$. In fact, since the integral representation of (B.12) has the form

$$h(t) = \mathbf{M}h_0 + \int_{t_0}^{t_n} \varphi_u(\tau)h(\tau)d\tau + \int_{t_0}^{t_n} \varphi_\rho(\tau)h_\rho \, d\tau,$$

Gronwall's inequality yields

$$\begin{aligned} \|h(t)\| &\leq \left(\|\mathbf{M}\| \|h_0\| + \|h_\rho\| \int_{t_0}^{t} \|\varphi_\rho(\tau)\| d\tau \right) exp\left(\int_{t_0}^{t} \|\varphi_{\mathbf{u}}(\tau)\| d\tau \right) \\ &\leq \kappa_1 \left(\|\mathbf{M}\| \|h_0\| + \kappa_2 (t_n - t_0) \|h_\rho\| \right) \leq \kappa (\|h_0\| + \|h_\rho\|), \end{aligned}$$

for some $\kappa_1, \kappa_2 > 0$, which exists since $\varphi_{\mathbf{u}}(t)$ and $\varphi_\rho(t)$ are bounded in compact sets, and where $\kappa = \max\{\kappa_1 \|\mathbf{M}\|, \kappa_1 \kappa_2 (t_n - t_0)\}$. Therefore, $\left|\nabla^2_{(\mathbf{u},\rho)} H[h, h_\rho]\right| \leq c\kappa \|\boldsymbol{\lambda}(t)\| \|h_\rho\| (\|h_0\| + \|h_\rho\|)$. Observe that if $\|h_\rho\| \leq \|h_0\|$, it holds that $\left|\nabla^2_{(\mathbf{u},\rho)} H[h, h_\rho]\right| \leq 2c\kappa \|\boldsymbol{\lambda}(t)\| \|h_0\|^2$, and consequently,

$$\int_{t_0}^{t_n} \nabla^2_{(\mathbf{u},\rho)} H[h, h_\rho] dt \geq -2c\kappa (t_n - t_0) \|\boldsymbol{\lambda}(t)\| \|h_0\|^2. \tag{B.13}$$

The case $\|h_0\| \leq \|h_\rho\|$ holds analogously. On the other hand, using (B.9), (B.10), and (B.13) in (B.6), it follows that

$$\nabla^2_{\mathbf{u},\rho,\mathbf{u}_0}\mathcal{L}[(h,h_\rho,h_0),(h,h_\rho,h_0)] \geq (t_n - t_0)\left(\delta_1\|h_\rho\|^2 + (\delta_2 - 2c\kappa\|\boldsymbol{\lambda}(t)\|)\|h_0\|^2\right) \geq \delta\|(h_\rho,h_0)\|^2,$$

where $\delta := (t_n - t_0)\min\{\delta_1, \delta_2 - 2c\kappa\|\boldsymbol{\lambda}(t)\|\}$ which is positive given that, accordingly to (B.8), $\|\boldsymbol{\lambda}(t)\|$ is bounded and sufficiently small for a.e. $t \in (t_0,t_n)$, given that the observations I^{obs} and the solution variable I^r are close enough. Consequently, the sufficient second-order optimality condition holds for every $\mathbf{h} = (h,h_\rho,h_0) \in AC([t_0,t_n],\mathbb{R}^5) \times \mathbb{R}^5 \times \mathbb{R}^4$ satisfying $e'(\mathbf{u},\rho,\mathbf{u}_0)\mathbf{h} = (0,0)$ if $\|\mathbf{C}_{ii}^{-1}(I^r - I^{obs})\|$ is sufficiently small.

AIMS Press