Using mobile phone data to estimate dynamic population changes and improve the understanding of a pandemic: A case study in Andorra

Alex Berke[1*], Ronan Doorley[1], Luis Alonso[1], Vanesa Arroyo[2], Marc Pons[2], Kent Larson[1]

1 Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA
2 Andorra Recerca + Innovació, Andorra

* aberke@mit.edu

Abstract

Compartmental models are often used to understand and predict the progression of an infectious disease such as COVID-19. The most basic of these models consider the total population of a region to be closed. Many incorporate human mobility into their transmission dynamics, usually based on static and aggregated data. However, mobility can change dramatically during a global pandemic as seen with COVID-19, making static data unsuitable. Recently, large mobility datasets derived from mobile devices have been used, along with COVID-19 infections data, to better understand the relationship between mobility and COVID-19. However, studies to date have relied on data that represent only a fraction of their target populations, and the data from mobile devices have been used for measuring mobility within the study region, without considering changes to the population as people enter and leave the region.

This work presents a unique case study in Andorra, with comprehensive datasets that include telecoms data covering 100% of mobile subscribers in the country, and results from a serology testing program that more than 90% of the population voluntarily participated in. We use the telecoms data to both measure mobility within the country and to provide a real-time census of people entering, leaving and remaining in the country. We develop multiple SEIR (compartmental) models parameterized on these metrics and show how dynamic population metrics can improve the models. We find that total daily trips did not have predictive value in the SEIR models while country entrances did. As a secondary contribution of this work, we show how Andorra's serology testing program was likely impacted by people leaving the country. Overall, this case study suggests how using mobile phone data to measure dynamic population changes could improve studies that rely on more commonly used mobility metrics and the overall understanding of a pandemic.

Introduction

At the start of the COVID-19 pandemic, nonpharmaceutical interventions (NPIs) were widely deployed in an effort to stymie the rate of new infections. These interventions included stay-at-home orders and restrictions on economic activity, which were used as a means to reduce contact and hence transmission rates, effectively limiting mobility. Country border restrictions were also put in place to reduce the chance of importing the

virus through inter-country travel. At the same time, tests became more available to better track population infection rates [1]. There has been an influx of data and research used to study the efficacy of various interventions [2–4]. In particular, this work addresses the use of population movement data.

Research preceding COVID-19 has indicated a close relationship exists between human mobility and the spread of infectious disease [5]. Past studies have shown how mobility data, such as commuter trips, can be used to improve disease forecasting models [6]. These earlier works highlighted the importance of combining their modeling frameworks with mobility data to address potential future emergent respiratory viruses, while also citing a lack of real-time mobility data as a limitation. In the wake of COVID-19, such real-time mobility data became widely available to study the pandemic, largely collected through airlines or via mobile phones. For example, early research modeling the effect of travel restrictions and the international spread of COVID-19 used a combination of airline transportation data to estimate international travel, and mobile phone data from Baidu LBS [7] to estimate domestic population movement [8]. Mobility data collected from mobile phones has also been made available by Google [9], Facebook [10], Safegraph [11], transit apps [12], telecoms, and other companies [13]. Metrics based on these sources have been used to model or predict COVID-19 transmission rates [14–18] as well as to verify model results [19], with the assumption that changes in transmission rates are correlated with changes in the mobility metrics.

However, the relationship between these data sources and COVID-19 remains unclear. The published mobility metrics are often aggregated statistics representing the number of trips taken, such as measured through transit apps, or based on foot traffic to points-of-interest (POIs). Furthermore, these data sources are limited, reporting on a fraction of the population (for example, Safegraph has one of the larger U.S. datasets yet covers only about 10% of the U.S. population and acknowledges reporting bias [20]). Likewise, other studies using reported cases microdata or air travel data to analyze the risks of importing the virus via inter-country travel (e.g. [21, 22]) are also limited by data sources that only report on a fraction of the true data.

Contribution. This work presents a unique case study in Andorra, with comprehensive datasets that include telecoms data covering 100% of mobile subscribers in the country, and results from a serology testing program that more than 90% of the population voluntarily participated in. Previous work used these data sources to compare various mobility metrics and infection rates with retrospective correlation analysis [23]. This work builds upon these previous findings and develops compartmental epidemic models.

At the start of the pandemic in Andorra, border restrictions and economic lockdowns drastically reduced country entrances and internal country mobility. This study includes that period as well as when restrictions were lifted. The mobile phone data are used to estimate mobility metrics representing trips, similar to related works, as well as to conduct a real-time census and estimate metrics that represent the dynamic population changes, such as daily country entrances. These data are then used to improve the understanding of the pandemic in Andorra in multiple ways.

First, we show how Andorra's serology testing program, conducted in May 2020, was likely impacted by people leaving the country. We then show how the estimated country entrances data can improve epidemiological (SEIR) models that otherwise rely on mobility measured by trips. Related works have used meta-population SEIR models where the modeled sub-populations are dynamic, yet based on static census commuting data or based on a combination of POI visits and static commuting data (e.g. [24]). In contrast, this work uses comprehensive telecoms data to estimate a real-time census to more accurately capture the changing dynamics of the population during the period of

study.

We develop and test multiple (SEIR) models that differ in how they parameterize transmission rates based on the trips and entrances metrics developed in this work. The models are simple, where their purpose is to illustrate how different types of mobility information can be better incorporated into SEIR models.

Finally, we use the best model to simulate a hypothetical counterfactual, representing a scenario where economic and border restrictions had not been put in place, and trips and entrances metrics had not drastically reduced.

Outline. Before presenting our methods and results, we provide background information, with a timeline of events around the start of COVID-19 in Andorra, and the features of the country that contribute to a unique case study. We also provide background information about compartmental epidemic models to guide the reader in the presentation of our models.

Background

Andorra and COVID-19

The study region of this work is the small country of Andorra, which is located in the Pyrenees mountains and shares borders with only France and Spain. The country has a population of approximately 77,000 [25], yet attracts more than 8 million visitors annually, mostly for tourism associated with skiing and nature-related activities [26]. In addition, a large number of cross-border temporary workers reside in the country, mainly employed in the tourism industry. Andorra lacks an airport or train service so the primary way to enter or exit the country is by crossing the French or Spanish border by car. The country is divided into 7 municipalities, called parishes.

Partly because of the country's small size and limited border crossings, Andorra was able to implement comprehensive policies at the start of the COVID-19 pandemic, as well as implement a serology testing program which more than 90% of the population participated in. Furthermore, there is one telecoms provider for the entire country, which contributes a comprehensive view of all mobile subscribers who spend any time in Andorra, whether they are Andorran nationals or have foreign SIM cards. The telecoms data and serology data are used in this work and are described in the Data sources and preprocessing section.

Timeline of COVID-19 cases and policies

The first COVID-19 case in Andorra was reportedly imported via Italy and confirmed March 2, 2020 [27]. Reported cases then rose rapidly in March before falling again in April (see Fig 1). On March 13, government officials ordered the closure of public establishments and a quarantine was requested of the entire population. A series of COVID-19 related policies followed and neighboring country borders were restricted. In accordance with these policies, mobility within the country dropped and border crossings ceased. Other NPIs, such as masks and hand sanitizer, were also deployed. The lockdown measures in Andorra were gradually lifted in April and May, and fully lifted starting June 1. Borders also reopened in June and border crossings resumed. Table A.1 in S1 Appendix shows a timeline of COVID-19 related events.

Nationwide serology testing program

In May of 2020, Andorra conducted a nationwide serology testing program. This resulted in the first published seroprevalence study universally testing the entire

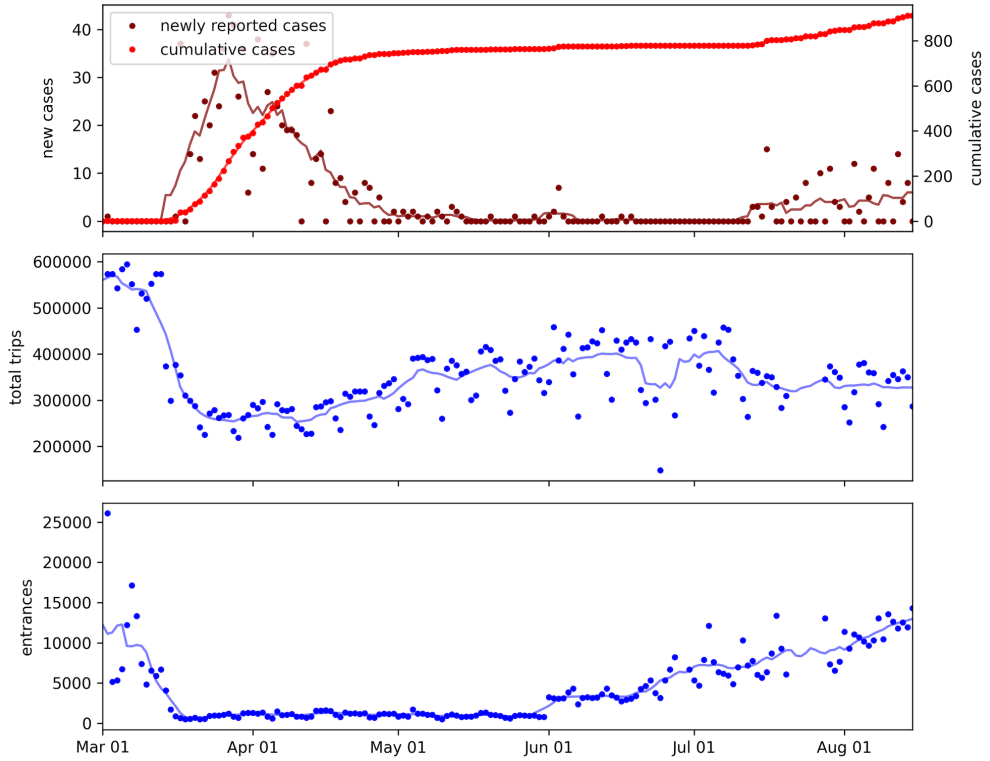

Fig 1. Daily reported cases, trips, and entrances metrics at the start of the COVID-19 pandemic in Andorra. The time series data are plotted for March to August, 2020, which covers the study period. Solid lines show values smoothed over a 7-day rolling window.

population of a country and one of the largest of its kind [28]. Anyone over the age of 2 was invited to participate in the study, including the country's temporary workers. The testing was conducted in two phases: May 4 -14, and May 18 - 28, 2020. The objectives of the second phase were (a) to track the progression of COVID-19 between the two surveys and (b) to account for indeterminate or potential false negative results from the first survey. More than 90% of the population participated voluntarily in at least one of the two surveys. However, an issue with the testing program was that many participants in the first phase did not participate in the second, limiting the data collection and impact of the two-phase study. This issue is further explored and addressed in the Results section.

SEIR models and COVID-19

SEIR models, and their variations, are compartmental models used in epidemiology. They have been widely used in forecasting COVID-19 transmission and modeling the outcomes of government policies [14, 29, 30]. The basic concept of these models is that the population is partitioned into sequential compartments, and transitions through the compartments over time. This framework was first developed by Kermack and McKendrick in 1927 [31] and has been well described more recently by Keeling et al. [32]. In short, the SEIR model takes its name from its compartments:

S = Susceptible
E = Exposed
I = Infectious
R = Removed (quarantined, recovered, or deceased)

S represents the number of Susceptible people in the population who have not yet been exposed to the virus. Individuals transition from Susceptible to Exposed after exposure to individuals in the Infectious (I) compartment. Hence the transition S to E is a function of the number of people in the Susceptible (S) and Infectious (I) compartments, as well as the transmission rate, β, and the total population size, N. The standard model considers N constant, and the following conservation holds for any time, t:

$$N = S(t) + E(t) + I(t) + R(t) \tag{1}$$

Transitions between compartments are modeled by a set of ordinary differential equations (ODEs).

$$\begin{aligned} S'(t) &= -\beta \frac{S(t)I(t)}{N(t)} \\ E'(t) &= \beta \frac{S(t)I(t)}{N(t)} - \sigma E(t) \\ I'(t) &= \sigma E(t) - \gamma I(t) \\ R'(t) &= \gamma I(t) \end{aligned} \tag{2}$$

Where
β = transmission rate of the infection
σ = latent rate
γ = removal rate

The latent rate, σ, is the average rate to become infectious after exposure (i.e. σ^{-1}=average incubation period) and the removal rate, γ, is the average rate at which individuals transition from I to R.

The modeled compartments and transitions are simplifications, yet this simple framework may be well applied to COVID-19 at the start of the pandemic, before populations were vaccinated or encountering re-infections. (Models for diseases over longer periods of time may also incorporate changes in the population via birth and death rates, while other models handle individuals becoming susceptible again [33].)

An epidemic is often characterised by the basic reproduction number, R_0. The estimation and value of the reproduction number is complex and often misrepresented, but in general it represents the expected number of secondary infections caused by a typical infected case in a population where everyone is susceptible [34, 35]. The value of R_0 can change over the course of an epidemic due to changes in human behavior and NPIs; a goal of control efforts, such as those employed during the COVID-19 pandemic, is to reduce R_0 to bring an epidemic under control [36]. Estimates for reproduction numbers have been used to understand the state of a pandemic and to measure the effectiveness of interventions [4, 30, 37–39].

$R_0 = \beta/\gamma$

R_0 is a function of both transmission rate and removal rate. The removal rate represents the rate at which infectious individuals are removed from the population and then are no longer at risk of infecting susceptible individuals. Removal might occur because they isolate, or recover and are no longer infectious, or die. The removal rate may vary due to changes in testing procedures (e.g. more proactive testing can identify

more cases and cause individuals to isolate earlier in their infectious period) or government policies (e.g. quarantine rules). Likewise, the transmission rate can change due to governmental policies and behavioral changes (e.g. staying home, wearing masks, and other NPIs).

Recent models that address COVID-19 have taken into account that transmission rates vary over time [14, 40–42]. Many models do so by incorporating mobility metrics to estimate behavioral changes and model changes in transmissibility based on these data. However, these mobility metrics are often based on sources that report on a small fraction of the population, and where the mobility metrics are aggregated statistics based on the number of trips to points of interest (POIs), which may not be the most important indicators of COVID-19 transmission. This is in contrast to the telecoms data used in this work, which covers all mobile subscribers within the country of Andorra, and is provided as a complete and unaggregated dataset, not limited to trips to POIs.

We note that any of the models referenced or presented in this work are oversimplifications of the complex dynamics of disease spread. They also suffer from unreliable case reports data, limited by the availability of tests, and reactive to changes in testing protocols [1].

Materials and methods

This section describes the SEIR models used in this work, and how they are trained and tested. It then describes data sources and preprocessing methods.

Code and data availability

All aggregated metrics and code used in this work are made available and documented in a public repository. The code includes analysis notebooks as well as the preprocessing scripts that produced the aggregated metrics. The data reporting on individuals, which was used to compute aggregate metrics, is sensitive and kept private.
`https://github.com/CityScope/CSL_Andorra_COVID_Public`

Modeling

This work develops and compares multiple SEIR models that differ in how they incorporate trips and entrances data in order to model transmission rates. The trips data measure mobility behavior within the country while the entrances data measure new country entrances (described in the Data sources and preprocessing section.)

The aim is to evaluate the relative impact of the trips and entrances data on model performance; the aim is not to build a state-of-the-art, accurate predictive model. To this end, the models are highly simplified.

Comparison models

In SEIR models, $\beta(t)$ typically represents the average number of people an infected person would expose per-unit time if everyone were susceptible. In particular, $\beta(t)$ is used to model the transition from the Susceptible to Exposed compartments. The use of $\beta(t)$ in our models is captured by the following equation from Eq (2).

$$E'(t) = \beta(t)\frac{S(t)I(t)}{N(t)} - \sigma E(t)$$

We develop multiple models that only differ in how they define $\beta(t)$.

In the following descriptions, $b0, ..., bn$ are parameters of $\beta(t)$ and are estimated during model training for each model in which they are included.

One model uses trips without entrances data (model ii). Another model uses both trips and entrances data (model iii). A model that uses neither data source is used as a baseline (model i).

Each of the models use the same framework, methods, and training and testing periods, described further below.

Model i: constant transmissibility

This is a baseline, dummy model where $\beta(t)$ is constant.

$$\beta(t) = b0$$

Model ii: transmission a function of trips data

$$\beta(t) = b0 + b1 \times trips(t)^{b2}$$

Model iii: transmission a function of trips and entrances data

In this model, the average rate at which the susceptible population is exposed can be impacted by the behavior of people within the country (e.g. mobility measured in trips) as well as the import of new cases (entrances).

$$E'(t) = \beta(t)\frac{S(t)I(t)}{N(t)} + S(t)f(entrances(t)) - \sigma E(t)$$

where

$$\beta(t) = b0 + b1 \times trips(t)^{b2}$$

$$f(entrances(t)) = \frac{I(t)}{N} \times b3 \times entrances(t)^{b4}$$

$f(entrances(t))$ represents the likelihood of new country entrants importing the virus. The term $\frac{I(t)}{N}$ reflects the assumption that the likelihood of new country entrants being infectious tracks with the timeline of infection rates in Andorra. This assumption is based on the fact that during the study period, the timeline of infections in Andorra was highly correlated with the timeline of infections in Spain and France (with Pearson correlation coefficients of 0.922 (p=0.000) and 0.932 (p=0.000), respectively), and the primary way to enter Andorra is through the Spanish or French borders. Furthermore, telecoms data showed that 86% of entrances by foreign SIMs were either Spanish or French, and when accounting for entrances by Andorran SIMs, 68% of all entrances were by Spanish or French SIMs. See section A.3 in S1 Appendix.

The above functions using entrances and trips can be combined into one equivalent expression representing transmissibility. We do this to simplify modeling and maintain a common expression for $E'(t)$.

$$E'(t) = \beta(t)\frac{S(t)I(t)}{N(t)} - \sigma E(t)$$

where

$$\beta(t) = b0 + b1 \times trips(t)^{b2} + b3 \times entrances(t)^{b4}$$

Model framework

The SEIR framework used in this work is illustrated in Fig 2 and is described by the following ODEs.

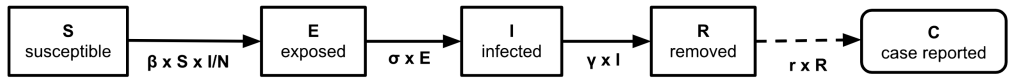

Fig 2. Schematic representing the SEIR model framework used in this work. The population is divided into compartments where individuals transition through the compartments: Susceptible, Exposed, Infected, Removed, Case reported, where the transitions are described by ODEs (Eq (3)).

$$
\begin{aligned}
S'(t) &= -\beta(t)\frac{S(t)I(t)}{N(t)} \\
E'(t) &= \beta(t)\frac{S(t)I(t)}{N(t)} - \sigma E(t) \\
I'(t) &= \sigma E(t) - \gamma I(t) \\
R'(t) &= \gamma I(t) \\
C'(t) &= rR'(t-d)
\end{aligned}
\tag{3}
$$

Where

$S(t) = N - E(t) - I(t) - R(t)$

$C(t)$ is cumulative case reports and accounts for reporting delay, d, and the reporting rate, r.

Given initial values for the compartments and the other model parameters, time series data for the compartments can be deterministically estimated by integrating over the ODEs into the future, where each compartment time series represents the compartment population on each day, t. This is done to calibrate parameters during model training as well as to generate forecasts beyond the training period.

Initial values for R and C at $t = 0$ are set based on the number of cumulative reported cases at the start of the study period. Initial values for E, I, are estimated by model training, along with γ and parameters of $\beta(t)$. The reporting rate, r, is set to $\frac{1}{11}$, estimated from the serology and case reports data (Data sources and preprocessing section). The latent rate, σ, is set to $\frac{1}{5.2}$, estimated by prior work [43]. The reporting delay, d, is set to 7, consistent with related works [19, 44] and empirical checks (see section A.7 in S1 Appendix). d is the average time from when an infectious individual is removed (isolated) to the time the case is reported, and must account for the time it takes to seek a test, for the test to be processed, and for the result to be included in reported cases data. At the start of the pandemic, tests in Andorra were sent to Spain for processing, which may have increased reporting delays. The reporting delay is incorporated into the models by shifting the trips and entrances metrics time series by d.

See table A.5 in section A.6 of S1 Appendix for a concise description of model parameters.

Training and testing

Cumulative reported cases in Andorra reached a threshold of 2 (over a 7-day average) on March 14. The serology tests, which were used to estimate the reporting rate, were conducted in May. In September, massive testing programs began and even before then, testing started to become more available. These programs and test availability increased the case identification rate, impacting both the reporting rate and the removal rate, changing the dynamics in modeling. For these reasons, the study period includes March

to August, 2020. The period of March 14 - May 31 is used for model training and the following 10 weeks are used for testing.

Training: Parameters and initial values for $E(t)$, $I(t)$ at $t = 0$ were fit with maximum likelihood estimation (MLE). Log-likelihood was computed by comparing time series values of predicted cumulative reported cases (C) to the time series of actual cumulative reported cases:

$$\textit{log-likelihood} = \sum log\ P_{k,\lambda}(k, \lambda) \tag{4}$$

Where the sum is over all days in the training data, $P_{k,\lambda}(k, \lambda)$ is the Poisson distributed probability mass function, k is actual reported cases, λ is predicted reported cases.

Parameters were optimized by minimizing the negative log-likelihood using the L-BFGS-B method [45]. See section A.5 in S1 Appendix for details.

Testing: Median absolute percentage error (MAPE) over cumulative estimates has been used in a recent framework to evaluate and compare COVID-19 models [46], where the errors incorporate an intercept shift. MAPE is similarly used to evaluate and compare the performance of models in this work. Given model training estimates S, E, I, R, C up to time t, the trained model is tested starting at time $t + 1$ as follows. The value of $C(t)$ is corrected to the true reported cases at time t and further integration over the ODEs is used to continue the simulation over the test period. The resulting C estimated over the test period is compared to actual reported cases via MAPE.

Data sources and preprocessing

Three main data sources are used in this work and are further described below: (i) serology data from the nationwide testing program conducted in May 2020, (ii) telecoms data covering all mobile subscribers in the country, (iii) official COVID-19 case and death reports. All time series metrics estimated from (ii) and (iii) are smoothed by taking the mean over a 7-day rolling window.

Serology Data

As described in the Andorra and COVID-19 section, a nationwide serology testing program was conducted in May of 2020. The program was voluntary, and conducted in 2 phases, and 91% of the population participated.

The program was conducted for a previous research study, in which the methods and results are detailed [28]. The study was approved by the Institutional Review Board of the Servei Andorra Atencio Sanitaria (register number 0720). The data were also provided to researchers in our lab as part of a research partnership. The dataset includes a unique identifier for each participant and results from the 1st and 2nd round of tests; test results were left empty when there was a lack of participation. The dataset also includes demographic information for participants, including their home parish and whether they are a temporary worker. As previously described, an issue with the serology testing program was that many of the participants from the first phase of testing did not participate in the second phase (see table A.3).

From the serology data, Bayes Theorem [47] was used to estimate the portion of the population infected up to May. With this number and the official reported cases data, we estimated a case reporting rate of $\frac{1}{11}$. This reporting rate is used in the epidemiology models described in this work.

Telecoms data and metrics

Andorra has one telecoms provider (Andorra Telecom), which provided the data for this study. Since they are the sole provider, the dataset covers 100% of mobile subscribers in the country, unlike most telecoms datasets where the market is fragmented. Each data point includes a unique ID for the subscriber, a timestamp, the coordinates of the device, and nationality for the subscriber's home network. The data have been further described in [48].

The stay-point extraction algorithm of Li et al. (2008) [49] was used to reduce the series of data points for each subscriber into a series of stay-points of 10 minutes or more within a radius of 200m or less. The stay-points represent a more concise and reliable series of places the subscriber spent time; stay-points were used to infer presence in the country, dynamic population changes, and compute the trips and entrances metrics.

There are gaps in the available telecoms data and the resulting trips and entrances metrics during the period of study (data gaps were June 28-29, and July 21-27, 2020). Missing values were imputed by taking the mean across the values from the 7 days surrounding each missing period of data.

Dynamic population inference and metrics: On each day, a subscriber was considered present in the country if they had a stay-point in the country within a 7-day window. The window accounts for unobserved subscriber devices due to a combination of inactivity, lack of reception in certain areas, or noisy data. The beginnings and endings of periods of presence were counted as entrances to and departures from the country, respectively.

Trips metrics: Daily trips for subscribers were counted as their daily number of stay points minus 1, since a new stay point is recorded when a subscriber moves beyond a 200m radius. Daily trips by subscribers were summed as a total daily trips metric.

Home inference: The home parish of each subscriber was inferred from the telecoms data, to come up with a population count for each of the 7 parishes of Andorra. This was done by first assigning each stay-point to the parish in which it was contained. Each subscriber's home parish was then determined to be the parish in which they spent the most cumulative time during night-time hours (12:00am to 6:00am). Related studies of human mobility that use cellular data have employed similar methods [50–53].

These inferred parish-level populations were compared to the published 2020 population statistics [25]. There is a Pearson correlation coefficient of 0.959 ($p < 0.001$), suggesting that the telecoms data are representative of the true population. (See table A.2 and Fig A.1 in S1 Appendix). Inferring the parish of residence is done both to check methodology as well as compare populations to serology test participation (see the Serology tests and country departures section).

COVID-19 infection data

This dataset was made available by Johns Hopkins University [54] and downloaded from OWID [55] as a time series of daily reports. Reported cases in Andorra were used for model estimation and prediction. There were cases identified in Andorra through the May serology testing program that were reported late, on June 2 [56]. This reporting error was handled by removing the excess case reports. Fig 1 plots the resulting daily new and cumulative case reports over the period of study. Reported deaths data for Andorra and its neighboring countries, Spain and France, were used in model assumptions (see section A.3 in S1 Appendix).

Results

2019 versus 2020 metrics

Before presenting our main findings, we first present the start of the pandemic in Andorra through a series of plots, and compare this period to the same period in 2019, when Andorra experienced a normal economy with tourism.

Fig 3 shows that by the start of March of 2020, there were already fewer people (mobile subscribers) in the country than in 2019. This number then substantially dropped with the start of the border restrictions and economic lockdown in mid March. There were also already fewer total daily trips being taken at the start of March, 2020, compared to 2019. This is largely due to fewer people in the country making the trips. This metric also substantially dropped at the start of the lockdown. This drop was partly due to even fewer people in the country making trips, and due to the government imposing restrictions on movement. The number of trips gradually rose again before the border restrictions were lifted in June, indicating that the population increased internal mobility. The number of daily entrances to (and departures from) Andorra also significantly dropped in mid March of 2020, as tourists and others left the country and border restrictions were imposed, limiting entry to the country. These daily metrics remained near zero throughout April and May, until border restrictions were lifted in June.

COVID-19 cases and mobility

The time series of reported COVID-19 cases is shown with the time series of the trips and entrances metrics in Fig 1. Other studies have implied that changes in case growth often lag changes in behavior and mobility metrics by 14 or more days [16, 19, 23]. However, Fig 1 shows that daily trips were able to increase throughout May of 2020 while newly reported cases remained low. Case growth did not increase again until daily entrances increased again when the border restrictions were lifted in June. This suggests that the entrances metric is more related to case growth than the trips metric in this case study. The relative predictive power of these metrics is further shown by the model results (Models results section).

Serology tests and country departures

Andorra's nationwide serology testing program conducted in May, 2020 involved two phases of testing (see the Andorra and COVID-19 section). An issue with this program was that many of the participants from the first phase of testing did not participate in the second phase, limiting the impact of the study. An important question for a country conducting such a program might be why this happened.

This drop in participation might be particularly concerning, as we found the drop in participation was more than 3 times higher among temporary workers versus the general population, and results from the testing program showed that temporary workers had higher seroprevalence (infection rates) versus the general population. See table A.4 in S1 Appendix. This might imply that a more infected demographic group was then less monitored.

By combining the serology test data with information inferred from the telecoms data, we find that test participants likely left the country after their first test.

We counted the number of mobile subscribers, by inferred home parish, who were in the country during the first and second phases of testing (May 4-14 and May 18-28, 2020). Subscribers were counted as present during a testing period if they had at least one "stay" within the period. We estimated how many subscribers left the country after

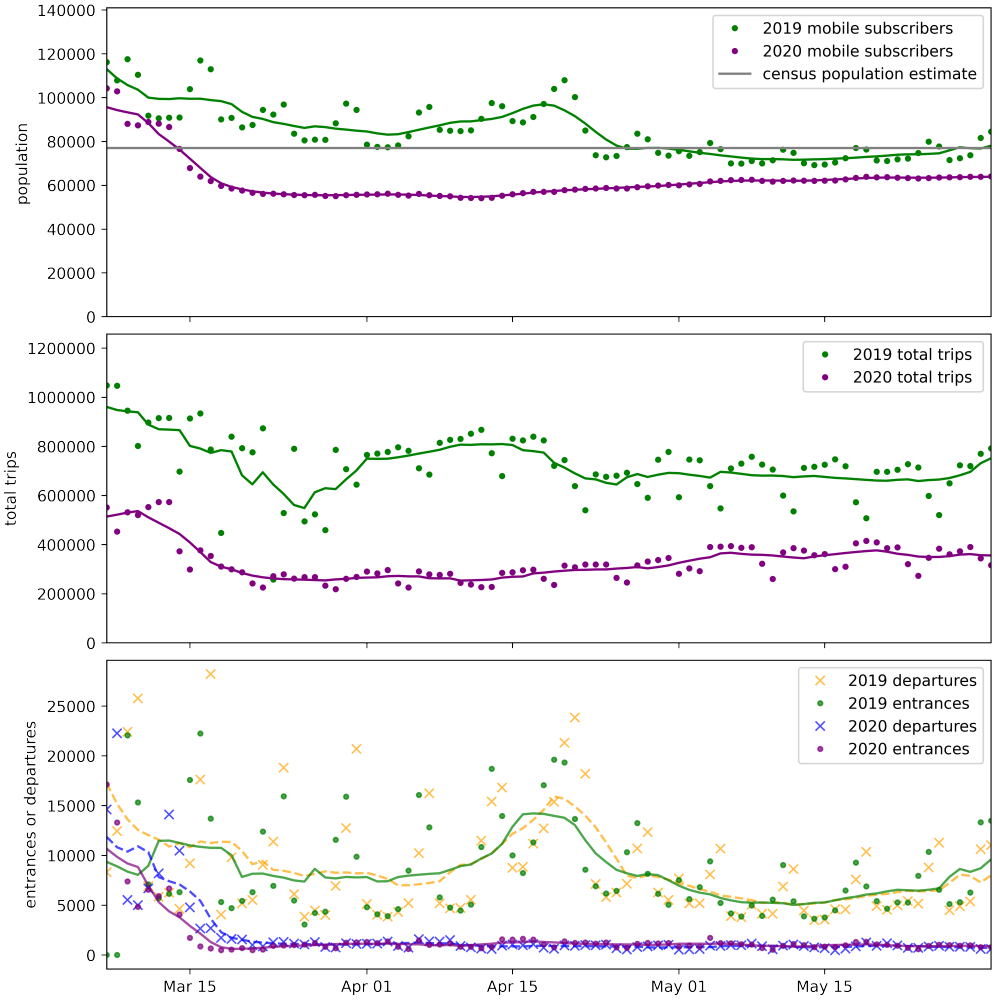

Fig 3. Estimated population, trips, country entrances and departures metrics for 2020 vs 2019. (Top) daily mobile subscribers counted as present in the country, (middle) daily total trips, and (bottom) daily country entrances and departures, for the country of Andorra during the start of the pandemic in 2020 versus the same period in 2019. All metrics are estimated from telecoms data that covers 100% of mobile subscribers in the country. Solid lines show values smoothed over a 7-day rolling window.

the first test by counting how many subscribers were present during only the first test period versus both test periods.

These numbers were compared to the parish-level serology test participant populations. Namely, the portion of serology test participants who did test 1 but not test 2 was compared to the estimated portion of mobile subscribers who left the country between test periods, and this comparison was done for each home parish. Comparing across parishes, there is a statistically significant Pearson correlation coefficient of 0.937 (p=0.0019).

To check the robustness of this result, we also restricted the May 2020 telecoms data to subscribers who had at least 7 days, or 4 nights, of data. The results are similar with Pearson correlation coefficients of 0.925 (p=0.0028), and 0.955 (p=0.0008), respectively.

To further validate that the decline in test participation was related to people leaving the country, we repeated these tests using 2019 telecoms data: we estimated the number of subscribers by home parish who were in the country during the periods May 4-14 and May 18-28 of 2019 (using 2019 telecoms data) and compared the number of subscribers who left the country between those periods to the serology test participation. In this case, there is a Pearson correlation coefficient of 0.4928 (p=0.2612). If the May 2020 subscribers had left the country for reasons not related to the pandemic, we would expect the correlation to be similar for the 2019 and 2020 data. However, the correlation for the 2019 data is much lower and not statistically significant. See table A.4 in S1 Appendix.

Models results

Simple models based on the SEIR framework, were developed to compare the impact of trips and entrances data on transmission rates and predicted infections.

The baseline, dummy model (i) assumes a constant transmission rate. For model (ii) transmission is a function of mobility measured by trips data, and for model (iii) transmission is a function of both trips and entrances data. (See the Modeling section for details.)

Models were trained over the period March 14 - May 31, 2020. Table A.5 and Fig A.5 in S1 Appendix show the parameter values for the best fit models and the corresponding time series values for the estimated R_0, the compartment populations, and the predicted reported cases, over the training period.

Models were evaluated by their prediction performance over the weeks that followed the training period. This was done using MAPE, based on the framework used by Friedman et al. to evaluate leading COVID-19 models [46]. Results for 1 - 10 forecasting weeks are shown in table 1. All models performed relatively well during the period of study. (As a point of comparison, Friedman et al. found in their global evaluation of COVID-19 models, MAPE values of 1 - 2% for 1 week forecasts and 17 - 25% for 10 week forecasts. See Fig. 3 and Fig. 5 in [46]. Note their evaluation used cumulative deaths data whereas this work uses cumulative cases data.)

The model (iii) using both trips and entrances data outperformed the other models in all but excluding the first week that followed the training period. More importantly, the model (ii) that used trips data to model transmission rates (without entrances data) had results similar to, and slightly worse than, the baseline model (i) which assumed a constant transmission rate. This is not surprising, as the data indicated trips were able to increase without impacting transmission rates (Fig 1).

This is also shown in that the best fit for model (ii) had parameters that flattened the impact of the trips data, resulting in a nearly flat reproduction number, R_0. Given that there were few new infections at the end of the training period (i.e. a smaller population in the I compartment), this resulted in relatively flat predictions for new reported cases for model (ii) over the forecasting weeks that followed the training period

(similar to model (i)). This is in contrast to the model (iii) that used both trips and entrances data, and where predictions for new reported cases closely tracked with actual predictions. See Fig 4. Overall, these estimated R_0 values are reasonable and within the range of values estimated by previous works [57].

	MAPE		
	model		
forecasting weeks	**i. constant β**	**ii. trips data**	**iii. trips & entrances data**
1	0.03	0.02	0.24
2	0.20	0.30	0.19
3	0.65	0.80	0.25
4	0.97	1.16	0.34
5	1.22	1.44	0.52
6	1.47	1.73	0.76
7	1.36	1.60	1.01
8	1.39	1.54	1.08
9	1.59	1.78	1.09
10	1.80	1.98	1.12

Table 1. MAPE results for the 3 models. The models differ in whether they incorporate trips and entrances data to model transmissibility. Model (i) is a baseline, dummy model where transmissibility is constant, model (ii) uses trips data, and model (iii) uses trips and entrances data. All models used the same framework and methods.

As a robustness check, all models were trained and tested over an additional set of training and testing periods that ended slightly earlier than those used for the main results. (The training period for the robustness check was March 14 - May 14, 2020.) The results are similar to the main results, and shown in table A.6 and Fig A.6 in section A.8 of S1 Appendix. However in this case, the model (iii) using trips and entrances data consistently outperformed the other models for all forecasting weeks.

Counterfactuals

What if Andorra had not imposed a lockdown, which caused reduced mobility? What if border restrictions had not been put in place, which caused a drop in entrances? Overall, what if the population mobility, measured in total trips and entrances, had not dropped in March?

In this section we explore such a counterfactual scenario by using the best fit model (iii) from the Models results section, which uses the trips and entrances data.

The lockdown in Andorra began on March 13, 2020, and there was a large drop in trips and entrances surrounding this date (see Fig 1). We again take a simplified approach to modeling, and create hypothetical trips and entrances data for a counterfactual scenario where mobility and border restrictions were not put in place. We do this by using the true metrics up to March 13 of 2020, and then keeping the metrics constant at the March 13 values. This is shown in Fig 5. We then estimate counterfactual case reports by using the previously fit model (i.e. we use the model parameters that were fit with the true trips and entrances time series values) and replace the model's trips and entrances data with the counterfactual data. We then run

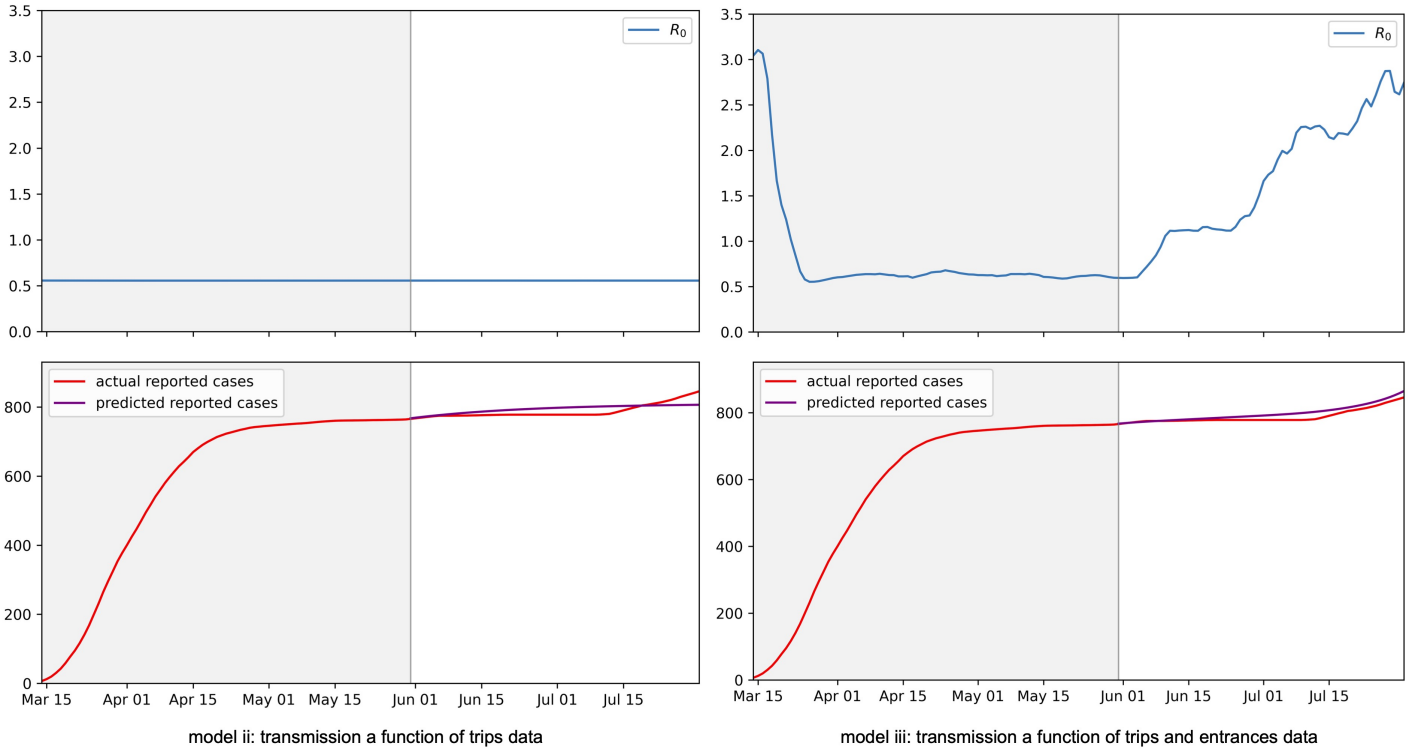

Fig 4. Fit model results. Time series values for (top) the estimated R_0 and (bottom) actual versus predicted reported cases that resulted from model training. Left: Plotted values for the model which uses just trips data. Right: Plotted values for the model which uses both trips and entrances data. Models were trained over the period March 15 - May 31 and tested over the weeks that followed. The training and testing periods are divided by gray and white backgrounds, respectively. Axes for the R_0 values are set to highlight that values were flattened for the trips data model. See Fig A.5 in S1 Appendix for plots that show the full variation in the R_0 values.

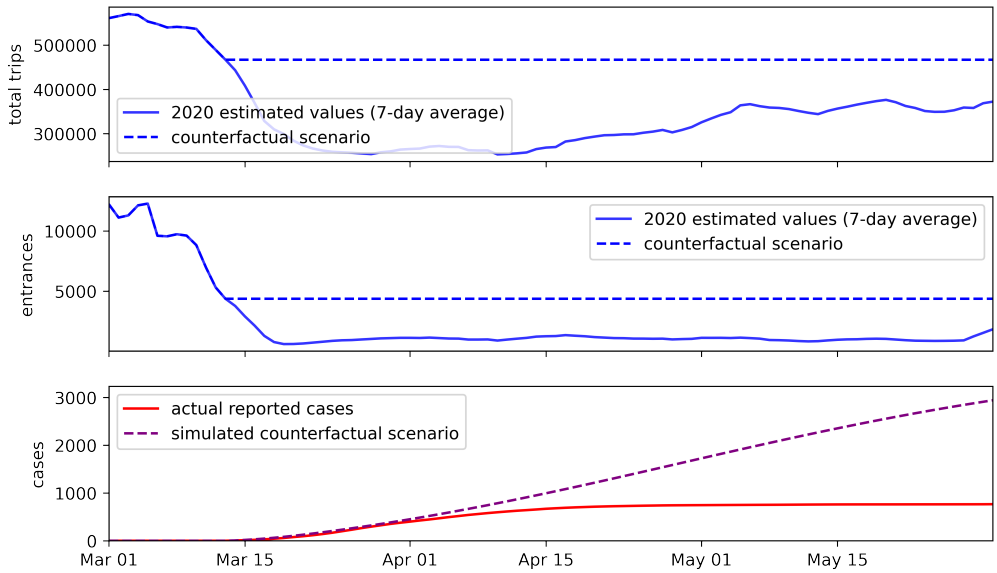

Fig 5. Counterfactual results. Top: Hypothetical total trips and entrances metrics that are used to simulate reported cases for a counterfactual scenario where mobility and border restrictions had not been put in place. Bottom: Predicted reported cases for such a counterfactual scenario, versus the actual reported cases that occurred in the true scenario.

the simulation over the same period that was used to train the original model. The result is a prediction of 2941 cumulative reported cases up to May 31, 2020 under the counterfactual model, versus the actual 766 reported cases up to May 31, under the true scenario. The difference is an additional 2175 (more than 3x as many) reported cases during this time period under the counterfactual scenario.

Discussion

When COVID-19 was introduced to Andorra at the start of March 2020, the country and its bordering neighbors responded quickly with economic and border restrictions. These interventions and other NPIs showed to be effective in Andorra, as the country brought case growth under control from March - May 2020, before the restrictions were fully lifted. The counterfactual scenario modeled in this work shows a stark alternative had the mobility changes observed during this period not occurred, with more than an estimated 3x as many cases, likely overwhelming the hospital system.

Numerous other works have also used mobility data collected from mobile phones to model the impacts of mobility restrictions on COVID-19 transmission. However, these studies have relied on data about trips, and the data represented a small sample. Other works using meta-population SEIR models where the modeled sub-populations are dynamic have been based on static census data. In contrast, this work leverages data collected from mobile phones that represent 100% of subscribers in a country.

We showed how these data could be used to build on previous works by computing daily trips metrics as well as estimating a dynamic, real-time population census. We then showed how these data can be used to improve upon the understanding of a pandemic in two main ways.

First, these data were used in order to better understand why participation in the nationwide serology testing program dropped between the first and second phases of

testing. The drop in participation may have been concerning as the second phase of testing was intended to help better detect and track the virus. This decreased ability to track the virus might have been particularly concerning because the test results showed that the temporary worker population had the highest infection rates and this population also had the largest drop in test participation. However, the analysis, which leveraged the telecoms data to estimate dynamic population changes, suggested that the decline in participation was likely due to test participants leaving the country after their first test.

Second, we showed how the dynamic population data could be used to improve epidemiological (SEIR) models that otherwise rely on mobility measured by trips. In our contribution, we developed simple SEIR models that differed in how they used the trips and entrances metrics developed through this work. These models performed well compared to the 7 global COVID-19 models evaluated by Friedman et al. (2021) [46] [14, 40, 41, 58–60], but their purpose was not to be highly accurate; the purpose of these models was to illustrate the relative importance of trips mobility data versus real-time population data, namely country entrances. In particular, for the case of Andorra, we find that the population was able to regain internal mobility measured in daily total trips with limited growth in cases, and that total trips per day did not have predictive value in the SEIR models while country entrances did.

While we show that the entrances metric had superior predictive power over the trips metric in Andorra, we do not mean to draw a direct line between country entrances and new COVID-19 cases. Changes in the entrances metric may have been highly correlated with other changes that impacted transmission rates, such as changes in COVID-19 policies and cautions.

In general, the models were limited by their simplifications. For example, there was likely an interaction effect between the trips and entrances metrics that was not captured in the models. The models also assumed that the case identification rate (and hence removal rate) and reporting rate were constant, which related works have as well (e.g. [19]). However, these rates likely changed with Andorra's increased testing. Future works can more accurately model the impacts of mobility and entrances, and the interaction between these metrics. This might also include incorporating data on the infection rates for other countries whose populations contribute to entrances. Future work can also incorporate data on testing rates to better model changes in the removal and reporting rates.

Overall, this case study suggests how using mobile phone data to measure dynamic population changes could improve studies that rely on more commonly used mobility metrics and the overall understanding of a pandemic.

Supporting information

S1 Appendix. Supplementary Appendix.

References

1. IHME. COVID-19: Estimating the historical time series of infections; 2021. Available from: http://www.healthdata.org/special-analysis/covid-19-estimating-historical-infections-time-series.
2. Haug N, Geyrhofer L, Londei A, Dervic E, Desvars-Larrive A, Loreto V, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nature human behaviour. 2020;4(12):1303–1312.
3. Adjodah D, Dinakar K, Chinazzi M, Fraiberger SP, Pentland A, Bates S, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. PLOS ONE. 2021;16(6):e0252315.
4. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2021;371(6531).
5. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proceedings of the National Academy of Sciences. 2009;106(51):21484–21489.
6. Pei S, Kandula S, Yang W, Shaman J. Forecasting the spatial transmission of influenza in the United States. Proceedings of the National Academy of Sciences. 2018;115(11):2752–2757.
7. Baidu, Inc. Baidu Qianxi platform; 2020. Available from: http://qianxi.baidu.com.
8. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020;368(6489):395–400.
9. Google LLC. Google COVID-19 Community Mobility Reports; 2021. Available from: https://www.google.com/covid19/mobility/.
10. Facebook. Facebook Data For Good: Our Work on COVID-19; 2021. Available from: https://dataforgood.fb.com/docs/covid19/.
11. Safegraph. Shelter in Place Index: The Impact of Coronavirus on Human Movement; 2021. Available from: https://www.safegraph.com/data-examples/covid19-shelter-in-place.
12. City Mapper. Citymapper Mobility Index; 2021. Available from: https://citymapper.com/cmi.
13. Warren MS, Skillman SW. Mobility changes in response to COVID-19. arXiv preprint arXiv:200314228. 2020;.
14. IHME COVID-19 forecasting team. Modeling COVID-19 scenarios for the United States. Nature medicine. 2020;.

15. Walker PG, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low-and middle-income countries. Science. 2020;369(6502):413–422.

16. Soucy JPR, Sturrock SL, Berry I, Westwood DJ, Daneman N, MacFadden DR, et al. Estimating effects of physical distancing on the COVID-19 pandemic using an urban mobility index. medRxiv. 2020;.

17. Ilin C, Annan-Phan S, Tai XH, Mehra S, Hsiang S, Blumenstock JE. Public mobility data enables covid-19 forecasting and management at local and global scales. Scientific reports. 2021;11(1):1–11.

18. Guan G, Dery Y, Yechezkel M, Ben-Gal I, Yamin D, Brandeau ML. Early Detection of COVID-19 Outbreaks Using Human Mobility Data. PloS one. 2021;16(7):e0253865.

19. Arroyo-Marioli F, Bullano F, Kucinskas S, Rondón-Moreno C. Tracking R of COVID-19: A new real-time estimation using the Kalman filter. PloS one. 2021;16(1):e0244474.

20. Ryan Fox Squire S. What about bias in the SafeGraph dataset?; 2019. Available from: `https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset`.

21. Hâncean MG, Perc M, Lerner J. Early spread of COVID-19 in Romania: imported cases from Italy and human-to-human transmission networks. Royal Society open science. 2020;7(7):200780.

22. Gilbert M, Pullano G, Pinotti F, Valdano E, Poletto C, Boëlle PY, et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. The Lancet. 2020;395(10227):871–877.

23. Doorley R, Berke A, Noyman A, Alonso L, Ribo JF, Arroyo V, et al. Mobility and COVID-19 in Andorra: Country-scale analysis of high-resolution mobility patterns and infection spread. IEEE Journal of Biomedical and Health Informatics. 2021; p. 1–1. doi:10.1109/JBHI.2021.3121165.

24. Pei S, Kandula S, Shaman J. Differential effects of intervention timing on COVID-19 spread in the United States. Science advances. 2020;6(49):eabd6370.

25. Departament-d'Estadística. Estimacions de població, gener 2020 and Estadística dels censos parroquials, gener 2020. Govern d'Andorra (Reports No A001 and A003). 2020; p. 1–11.

26. CIA. The World Factbook: ANDORRA; 2021. Available from: `https://www.cia.gov/the-world-factbook/countries/andorra/`.

27. Staff R. A 20-year old man is Andorra's first coronavirus case. Reuters;.

28. Royo-Cebrecos C, Vilanova D, López J, Arroyo V, Pons M, Francisco G, et al. Mass SARS-CoV-2 serological screening, a population-based study in the Principality of Andorra. The Lancet Regional Health-Europe. 2021;5:100119.

29. Giattino C. How epidemiological models of COVID-19 help us estimate the true number of infections; 2020. Our World in Data. Available from: `https://ourworldindata.org/covid-models`.

30. Karnakov P, Arampatzis G, Kičić I, Wermelinger F, Wälchli D, Papadimitriou C, et al. Data-driven inference of the reproduction number for COVID-19 before and after interventions for 51 European countries. Swiss medical weekly. 2020;150:w20313.

31. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london Series A, Containing papers of a mathematical and physical character. 1927;115(772):700–721.

32. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton university press; 2011.

33. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton university press; 2011.

34. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R0). Emerging infectious diseases. 2019;25(1):1.

35. Rothman KJ, Greenland S, Lash TL, et al. Modern epidemiology. vol. 3. Wolters Kluwer Health/Lippincott Williams & Wilkins Philadelphia; 2008.

36. Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford university press; 1992.

37. Camacho A, Kucharski A, Aki-Sawyerr Y, White MA, Flasche S, Baguelin M, et al. Temporal changes in Ebola transmission in Sierra Leone and implications for control requirements: a real-time modelling study. PLoS currents. 2015;7.

38. Thompson R, Stockwin J, van Gaalen RD, Polonsky J, Kamvar Z, Demarsh P, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;29:100356.

39. Chaves LF, Hurtado LA, Rojas MR, Friberg MD, Rodríguez RM, Avila-Aguero ML. COVID-19 basic reproduction number and assessment of initial suppression policies in Costa Rica. Mathematical Modelling of Natural Phenomena. 2020;15:32.

40. Li ML, Bouardi HT, Lami OS, Trikalinos TA, Trichakis NK, Bertsimas D. Forecasting COVID-19 and analyzing the effect of government interventions. MedRxiv. 2021; p. 2020–06.

41. Gu Y. COVID-19 projections using machine learning; 2021. Available from: `https://covid19-projections.com`.

42. Kounchev O, Simeonov G, Kuncheva Z. The TVBG-SEIR spline model for analysis of COVID-19 spread, and a Tool for prediction scenarios. arXiv preprint arXiv:200411338. 2020;.

43. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. New England journal of medicine. 2020;.

44. Tariq A, Lee Y, Roosa K, Blumberg S, Yan P, Ma S, et al. Real-time monitoring the transmission potential of COVID-19 in Singapore, March 2020. BMC medicine. 2020;18:1–14.

45. Byrd RH, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. SIAM Journal on scientific computing. 1995;16(5):1190–1208.

46. Friedman J, Liu P, Troeger CE, Carter A, Reiner RC, Barber RM, et al. Predictive performance of international COVID-19 mortality forecasting models. Nature communications. 2021;12(1):1–13.

47. McNeil BJ, Adelstein SJ. Determining the value of diagnostic and screening tests. Journal of Nuclear Medicine. 1976;17(6):439–448.

48. Doorley R, Alonso L, Grignard A, Maciá N, Larson K. Travel demand and traffic prediction with cell phone data: Calibration by mathematical program with equilibrium constraints. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC). IEEE; 2020. p. 1–8.

49. Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma WY. Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems; 2008. p. 1–10.

50. Kung KS, Greco K, Sobolevsky S, Ratti C. Exploring universal patterns in human home-work commuting from mobile phone data. PloS one. 2014;9(6):e96180.

51. Çolak S, Alexander LP, Alvim BG, Mehndiratta SR, González MC. Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. Transportation Research Record. 2015;2526(1):126–135.

52. Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C, et al. COVID-19 outbreak response: a first assessment of mobility changes in Italy following national lockdown. medRxiv. 2020;.

53. Phithakkitnukoon S, Smoreda Z, Olivier P. Socio-geography of human mobility: A study using longitudinal mobile phone data. PloS one. 2012;7(6):e39253.

54. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases. 2020;20(5):533–534.

55. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al.. Coronavirus Pandemic (COVID-19); 2021. Our World in Data. Available from: `https://ourworldindata.org/coronavirus`.

56. Els tests d'anticossos permeten diagnosticar 78 positius de la COVID-19, que podrien haver contagiat unes 360 persones; 2020. Govern d'Andorra. Available from: `https://www.govern.ad/comunicats/item/11629-els-tests-d-anticossos-permeten-diagnosticar-78-positius-de-la-covid-19-que-podrien-haver-contagiat-unes-360-persones`.

57. Viceconte G, Petrosillo N. COVID-19 R0: Magic number or conundrum?; 2020. Multidisciplinary Digital Publishing Institute.

58. Los Alamos national Laboratory COVID-19 Team. Los Alamos Natinoal Laboratory COVID-19 Confirmed and Forecasted Case Data; 2020. Available from: `https://covid-19.bsvgateway.org`.

59. MRC Centre for Global Infectious Disease Analysis. Imperial College COVID-19 LMIC Reports; 2020. Available from: `https://mrc-ide.github.io/global-lmic-reports`.

60. Srivastava A, Xu T, Prasanna VK. Fast and Accurate Forecasting of COVID-19 Deaths Using the SIkJa Model. arXiv preprint arXiv:200705180. 2020;.

61. COVID-19 LMIC Reports: Frequently Asked Questions; 2020. MRC Centre for Global Infectious Disease Analysis, Imperial College London. Available from: `https://mrc-ide.github.io/global-lmic-reports/FAQ.html`.

62. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17:261–272. doi:10.1038/s41592-019-0686-2.

S1 Appendix

A.1 Andorra and COVID-19

March 2	First COVID-19 case confirmed in Andorra.
March 12	Ski stations closed.
March 13	Partial confinement: school closures and recommended confinement.
March 18	Total confinement: all non-essential activities ordered closed.
April 7	Beginning of masks delivery and progressive use by population.
April 17	Allowance to 1 hour of walk in 1km radius every two days.
April 20	Phase 1 reopening: low risk economic activities resume; 1000 people return to work.
May 4	Start of 1st round of population serology screening.
May 4	Phase 2 reopening. Additional 4760 workers return to normal activity.
May 13	Increase to 2 hours of activity every day to walk or exercise.
May 18	Start of 2nd round of population serology screening.
May 18	Phase 3 reopening: additional 3300 workers returned to normal activity.
June 1	Confinement restrictions completely lifted.
June 15	French and Spanish borders opened (with restrictions on Spanish side).
June 21	Spain ended state of emergency and further lifted border controls.
July 1	Remaining Spanish border restrictions lifted.
September 1	Massive testing for teachers and school kids began.

Table A.1. Timeline of the COVID-19 related dates in Andorra.

A.2 Home parish inference

The parish-level populations inferred from the telecoms data for May 2020 were compared to published 2020 population statistics [25]. There is a Pearson correlation coefficient of 0.959 ($p<0.001$), suggesting that the telecoms data are representative of the true population.

Parish	Published population	Inferred mobile subscriber population
Andorra la Vella	22504	13555
Canillo	4371	4367
Encamp	11716	6766
Escaldes-Engordany	14626	7366
La Massana	10199	6165
Ordino	4957	2427
Sant Julià de Lòria	9374	4339

Table A.2. Published parish populations compared to parish populations inferred from telecoms data.

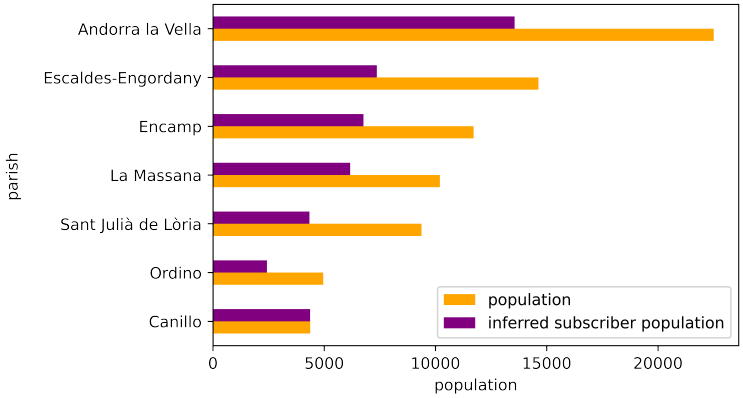

Fig A.1. Published parish populations compared to parish populations inferred from telecoms data There is a Pearson correlation coefficient of 0.959 ($p<0.001$), suggesting that the data are representative of the true population.

A.3 Comparing entrances and infection rates between Andorra, France, and Spain

The SEIR model (iii) in this work incorporates trips and entrances data to model transmission rates. It includes an assumption that the likelihood of new country entrants being infectious tracks with the timeline of infection rates in Andorra. In order to check this assumption, we compare the rates of COVID-19 between the country of Andorra and its bordering neighbors, Spain and France, during our period of study. This is done by comparing the timeline of reported deaths per 1 million residents, where data is smoothed over a 7 day rolling window. Death reports are used instead of case reports as a more stable comparison indicator in this study and others because death reports were considered to be less impacted than case reports by the dynamically changing testing procedures which varied by country [19, 61]. The Pearson correlation coefficients between Andorra and France and Andorra and Spain are 0.916 ($p < 0.001$) and 0.928 ($p < 0.001$), respectively. We note there was an error in the Spain data with negative deaths values in late May. We changed the negative values to 0 for this comparison. Without the change, the correlation was still statistically significant with Pearson correlation coefficient 0.918 ($p < 0.001$). The timeline of infections is shown in Fig A.2.

During this same period, telecoms data showed that 86% of entrances by foreign SIMs were either Spanish or French, and 68% of all entrances were by Spanish or French SIMs when accounting for entrances by Andorran SIMs. The timeline of entrances by SIM nationality is shown in Fig A.3.

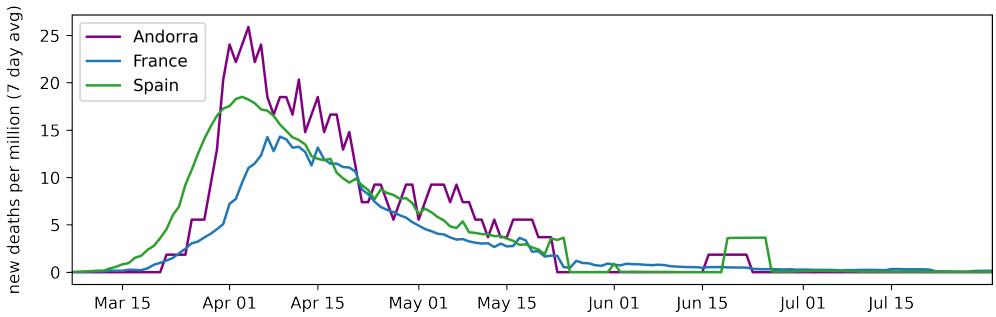

Fig A.2. Comparison of reported deaths per 1 million residents between the country of Andorra and its only bordering neighbor countries, Spain and France. The Pearson correlation coefficients between Andorra and France and Andorra and Spain are 0.916 ($p<0.001$) and 0.928 ($p<0.001$), respectively.

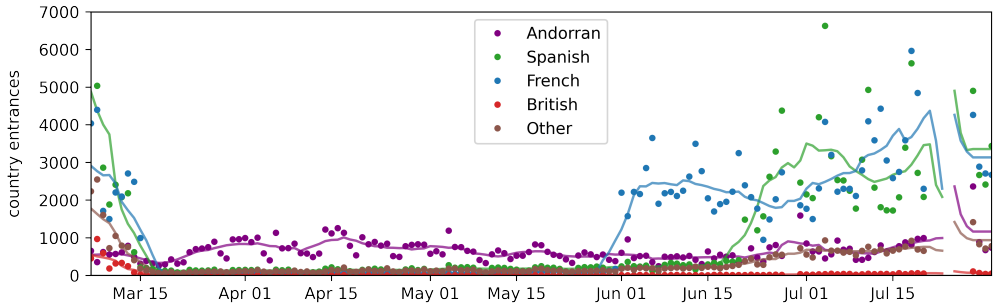

Fig A.3. The timeline of country entrances, measured via Andorra Telecom data, by mobile subscribers' SIM nationality. During the period of March through July, 2020, 86% of entrances by foreign SIMs were either Spanish or French, and 68% of all entrances were by Spanish or French SIMs when accounting for entrances by Andorran SIMs.

A.4 Serology tests and country departures

Note that the sum of survey participation varies due to missing values regarding the participants' parish of residence or temporary worker status.

	survey 1 participants (N)	survey 1 seroprevalence	survey 2 participation
all	70,626	9.7%	91.6%
temporary workers	2,714	13.3%	70.7%

Table A.3. Serological survey participation and results for temporary workers versus general population. Seroprevalence results were previously reported by Royo-Cebrecos et al. Survey 2 participation indicates how many individuals who participated in survey 1 also participated in survey 2.

Two cross sectional serological surveys were conducted in Andorra from May 4-28, 2020, using a rapid serological test (nCOV IgG/IgM) [28]. For each participant, the test data include the dates of their participation in surveys 1 and/or 2, the positive versus negative results for IgG/IgM antibodies for each round of testing, the participant's parish of residence, whether the participant is a temporary worker, and other demographic information. Table A.3 shows the number of participants and seroprevalence from survey 1 as well how many participants from survey 1 also participated in survey 2. Data for temporary workers is highlighted. Seroprevalence data is from previously reported results and was calculated based on the number of individuals who had a positive result of IgG and/or IgM [28]. The testing was voluntary; an issue with the testing was that many people who participated in the first survey did not participate in the second survey. Seroprevalence was higher among temporary workers. At the same time, temporary workers who participated in survey 1 were less likely to participate in survey 2 versus the general population. See table A.3.

parish	serology survey participants			2020 mobile subscribers present			2019 mobile subscribers present		
	survey 1 (N)	survey 1 & 2	**decline**	survey period 1	survey periods 1 & 2	**decline**	survey period 1	survey periods 1 & 2	**decline**
Andorra la Vella	4542	4262	**6.2%**	15803	15415	**2.5%**	27531	23295	**15.4%**
Canillo	4495	4017	**10.6%**	5406	5121	**5.3%**	5095	3936	**22.7%**
Encamp	10846	10273	**5.3%**	7613	7449	**2.2%**	8102	7397	**8.7%**
Escaldes-Engordany	28924	27688	**4.3%**	8336	8172	**2.0%**	9089	6961	**23.4%**
La Massana	8843	8451	**4.4%**	7389	7214	**2.4%**	9813	8365	**14.8%**
Ordino	4214	4049	**3.9%**	2821	2758	**2.2%**	3446	2964	**14.0%**
Sant Julià de Lòria	8632	8295	**3.9%**	5056	4929	**2.5%**	9201	8197	**10.9%**

Table A.4. Serological survey participation and mobile subscribers by home parish.

We counted the number of mobile subscribers, by inferred home parish, who were in the country during the first and second survey periods (May 4-14 and May 18-28, 2020). Subscribers were counted as present during a survey period if they had at least one "stay" within the period. We estimated the rate at which subscribers left the country after the first test by counting how many subscribers were present during only the first period versus both periods.

These numbers were compared to the parish-level serology test participant populations. Namely, the portion of serology survey participants who did survey 1 but not survey 2 was compared to the estimated portion of mobile subscribers who left the country between survey periods. There is a statistically significant Pearson correlation coefficient of 0.937 (p=0.0019).

To further validate that the decline in test participation was related to people leaving the country, we repeated these tests using 2019 telecoms data. In this case, there is a Pearson correlation coefficient of 0.4928 (p=0.2612), which is not statistically significant. If the May 2020 subscribers had left the country for reasons not related to the pandemic, we would expect the correlation to be similar for the 2019 and 2020 data.

A.5 Model training details

All modeling work is publicly accessible in a Python Jupyter notebook: `https://github.com/CityScope/CSL_Andorra_COVID_Public/blob/main/analysis/SEIR_models_trips_entrances.ipynb`.

Optimal parameters were estimated by minimizing the negative log-likelihood function using the L-BFGS-B method with the Python SciPy library [45, 62]. This step searches for optimal parameters by taking initial parameters which are then modified towards improved values, with specified bounds. The (minimum, maximum) bounds for γ were set to (1/10, 1/2). The bounds used for each of the parameters related to β - $b0, b1, b2, b3, b4$ - were (0, 2), with the observation that the parameters for the best fit models did not come up against these bounds. The (minimum, maximum) bounds for both $E(0)$ and $I(0)$ were set to (40, 4000), where t=0 corresponds to March 14, the first day of the training period. The values of (40, 4000) were conservatively set where the minimum was based on reported active cases, and the maximum based on reported cumulative cases. In addition, the training routine discarded models where E(0) and I(0) values differed by more than 3000.

Before describing how initial parameters were handled, first note that the optimizing function is not convex. To avoid the optimization function terminating at local minima, a grid search was used for the initial parameters. The same grid search method was used for each of the models.

In addition to the grid search routine, another step was taken to find optimal model fits: The models using the trips and entrances data were initially fit using spline approximations of these metrics, where the splines were estimated from the true metrics using knots spaced by 7 days. Fig A.4 shows the comparison of the true metrics versus their spline approximations. This step was taken to further smooth the data and ease the computational complexity of the model fitting routine. Without this improvement, the model training was slow and rarely resulted in successful outcomes, and the estimated parameters rarely varied from their initial values.

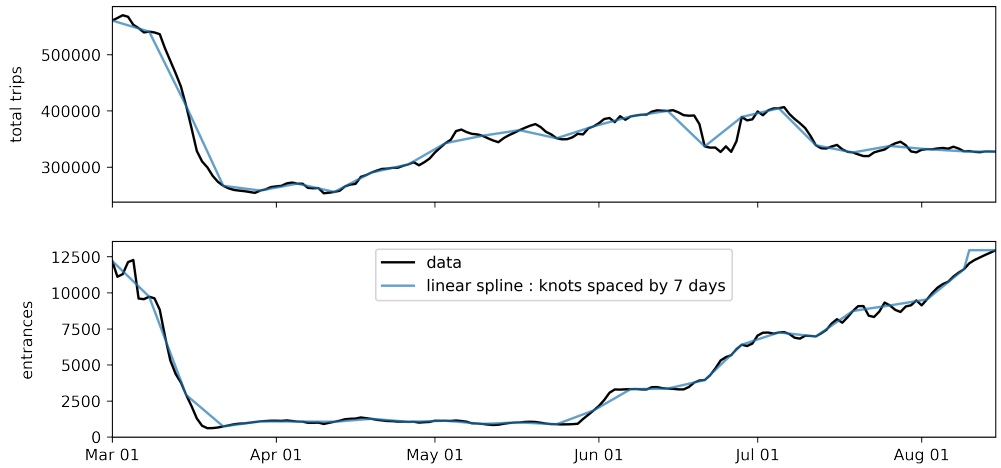

Fig A.4. Metrics and spline approximations. Daily estimated total trips and entrances metrics, and the linear spline approximations of these metrics where knots are spaced by 7 days.

After the models were fit using the spline approximations of the data, the models were finally fit again using the true data, where the parameters found via the fitting routine with the spline approximations of the data were used as the initial parameters in the L-BFGS-B method.

A.6 Fit model values and parameters

Table A.5 shows the values of the parameters for the best fit models that resulted from model training. Fig A.5 shows the corresponding time series values representing each of the best fit models. The values include R_0, the compartment populations, and the predicted reported cases.

The best fit models were determined as those with the best log-likelihood score when fit over the training data, where the training data period was March 14 - May 31, 2020. See section A.8 for values from the robustness check.

Models were trained based on the standard SEIR model where:

$$
\begin{aligned}
S'(t) &= -\beta(t)\frac{S(t)I(t)}{N(t)} \\
E'(t) &= \beta(t)\frac{S(t)I(t)}{N(t)} - \sigma E(t) \\
I'(t) &= \sigma E(t) - \gamma I(t) \\
R'(t) &= \gamma I(t) \\
C'(t) &= rR'(t-d)
\end{aligned}
$$

and

$S(t) = N - E(t) - I(t) - R(t)$

$C(t)$ is cumulative case reports and accounts for reporting delay, d, and reporting rate, r.

The 3 models in this work differed based on how they defined transmission rate and the transition from S to E. See the Modeling section for details.

Model i: baseline model with constant transmission rate

Model ii: daily transmission a function of daily trips data

Model iii: daily transmission a function of daily trips and entrances data

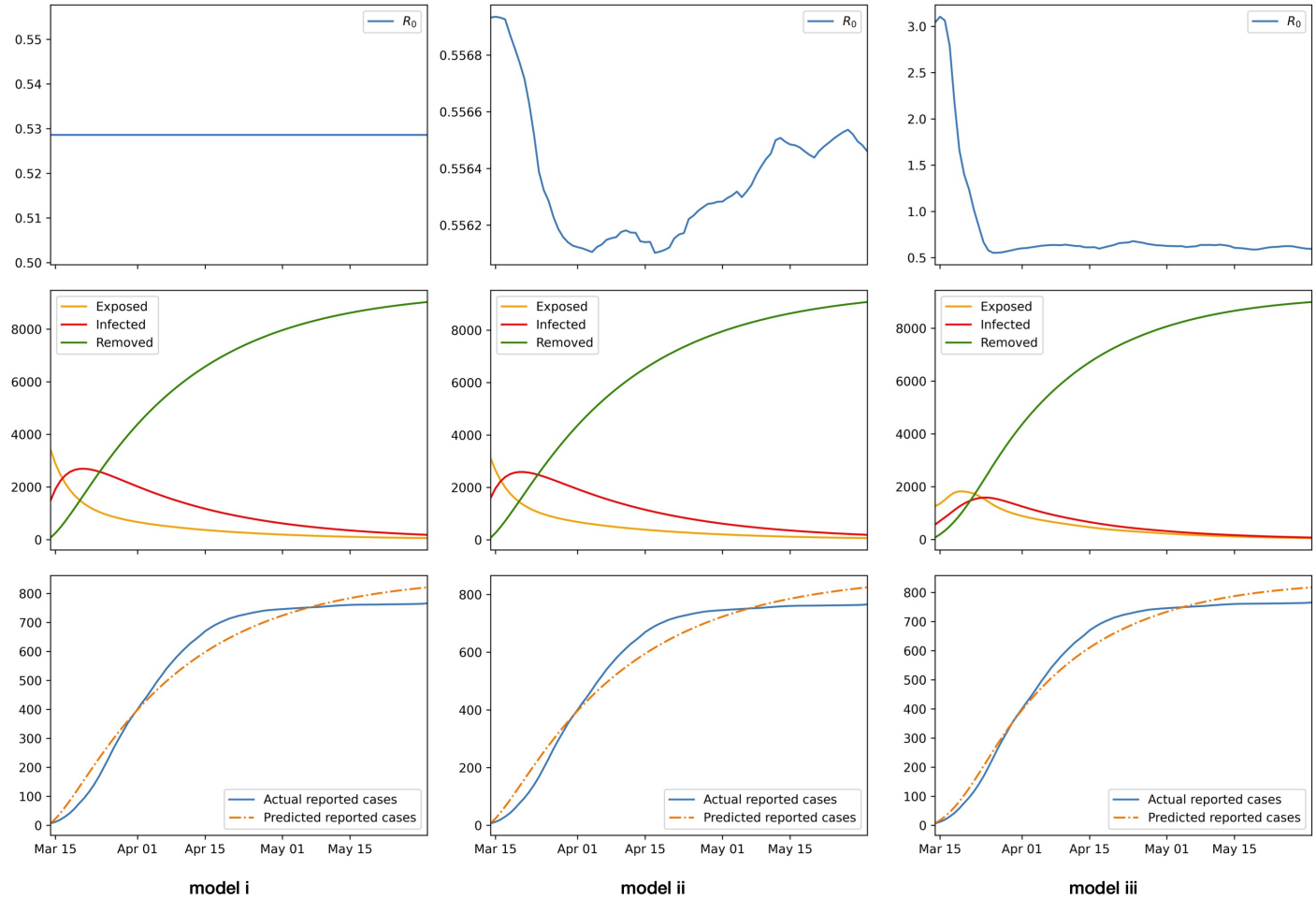

Fig A.5. Fit models. The time series values for (top) estimated R_0, (middle) estimated compartment populations, and (bottom) actual reported cases versus predicted reported cases. Plots are shown for each of the models: (left) model i is the baseline, dummy model with a constant transmission rate, (middle) model ii uses trips data, (right) model iii uses trips and entrances data.

parameter		model		
	description	i. constant β	ii. trips data	iii. trips & entrances data
N	Population. Set based on estimated population.	77,000	77,000	77,000
r	Reporting rate. Set based on serology and case reports data.	11^{-1}	11^{-1}	11^{-1}
d	Reporting delay. Set based on prior information.	7	7	7
σ	Latent rate. Set based on prior work.	5.2^{-1}	5.2^{-1}	5.2^{-1}
γ	Removal rate. Fit by model training.	0.10	0.10	0.18
$b0$	Transmission rate parameter. Fit by model training.	0.0529	0.0289	3.9126e-03
$b1$	Transmission rate parameter. Fit by model training.	-	0.027	7.69112e-02
$b2$	Transmission rate parameter. Fit by model training.	-	0.004	7.2082e-03
$b3$	Transmission rate parameter. Fit by model training.	-	-	2.6668e-06
$b4$	Transmission rate parameter. Fit by model training.	-	-	1.3168
E(0)	Initial exposed population. Fit by model training.	3408	3105	1261
I(0)	Initial infectious population. Fit by model training.	1488	1597	567
R(0)	Initial removed population. Set based on reported cases.	77	77	77
LL	Log-likelihood value for model fit over the training period	528.08	545.01	415.2

Table A.5. Parameters for fit models.

A.7 Time from exposure to reported case

The average delay in time from exposure to reported case in our model (i.e. the full transition through compartments E,I,R,C) is due to the average latent period, σ^{-1}, plus average infectious period, γ^{-1}, plus average reporting delay, d. σ^{-1} is set to 5.2 based on previous research [43], γ^{-1} is estimated via model training, and d is set to 7 based on related work [19, 44]. Reporting delays, d, can be due to the time it takes to seek a test, for the test to be processed, and then officially reported. Note that at the start of the pandemic in Andorra, tests were sent for processing to Spain, potentially adding extra time to reporting delays.

A study in Singapore from March 2020 estimated an average reporting delay of 6.4 days (95% CI 5.8, 6.9) [44]. A reporting delay of $d = 7$ was implicitly assumed by Arroyo-Marioli et al. [19]. They estimated time series values for the effective reproduction number for 124 countries across the world and validated their work by correlating their estimates of R_t to mobility data from the "COVID-19 Community Mobility Reports" collected by Google, where the lag between R_t and mobility was 14 days (2 weeks). They assumed an SIR model rather than an SEIR model, with time from exposure to removed, γ^{-1}, of 7 days, implying a reporting delay of 7 days $(14 - 7)$. We note that Arroyo-Marioli et al. produced time series estimates for R_t in Andorra. However their estimates are not comparable to the R_0 estimates in this work because (a) they did not correct for the reporting error that caused an influx of 78 additional cases June 1-10 [56], as was done in this work, and (b) their estimates were for the effective reproduction number versus the basic reproduction number. We also note that Google's "COVID-19 Community Mobility Reports" are not available for Andorra.

The best fit for the model estimated $\gamma^{-1} \sim 5.5$ days ($\gamma = 0.18$). Combined with $\sigma^{-1} = 5.2$, $d = 7$ results in a total estimated average delay from exposure to case report $\sim$17.7 days. This is consistent with previous work over a similar study period in Andorra that studied the correlation between mobility metrics and transmission rates over various lags and found the best correlations were with mobility metrics lagged by 18 days [23].

A.8 Robustness check

As a robustness check, we trained and tested all models over an additional set of training and testing periods that ended slightly earlier than those used for the main results. The training period for the robustness check was March 14 - May 14, 2020. The models were then tested on the period that directly followed this training period.

Table A.6 shows the comparison of MAPE values for each of the 3 models evaluated over the testing period. The model that uses the trips and entrances data consistently outperforms the other models.

	MAPE		
	model		
forecasting weeks	**i. constant β**	**ii. trips data**	**iii. trips & entrances data**
1	2.12	1.65	1.13
2	3.86	2.99	2.02
3	5.37	4.12	2.74
4	6.66	5.06	3.02
5	7.37	5.44	3.15
6	7.73	5.50	3.23
7	8.44	5.92	3.34
8	9.24	6.45	3.44
9	9.87	6.83	3.57
10	9.88	6.48	3.47

Table A.6. MAPE. MAPE values for the 3 models trained over the period used for a robustness check: March 14 - May 14, 2020.

Fig A.6 shows the resulting model time series values over the training period March 14 - May 14.

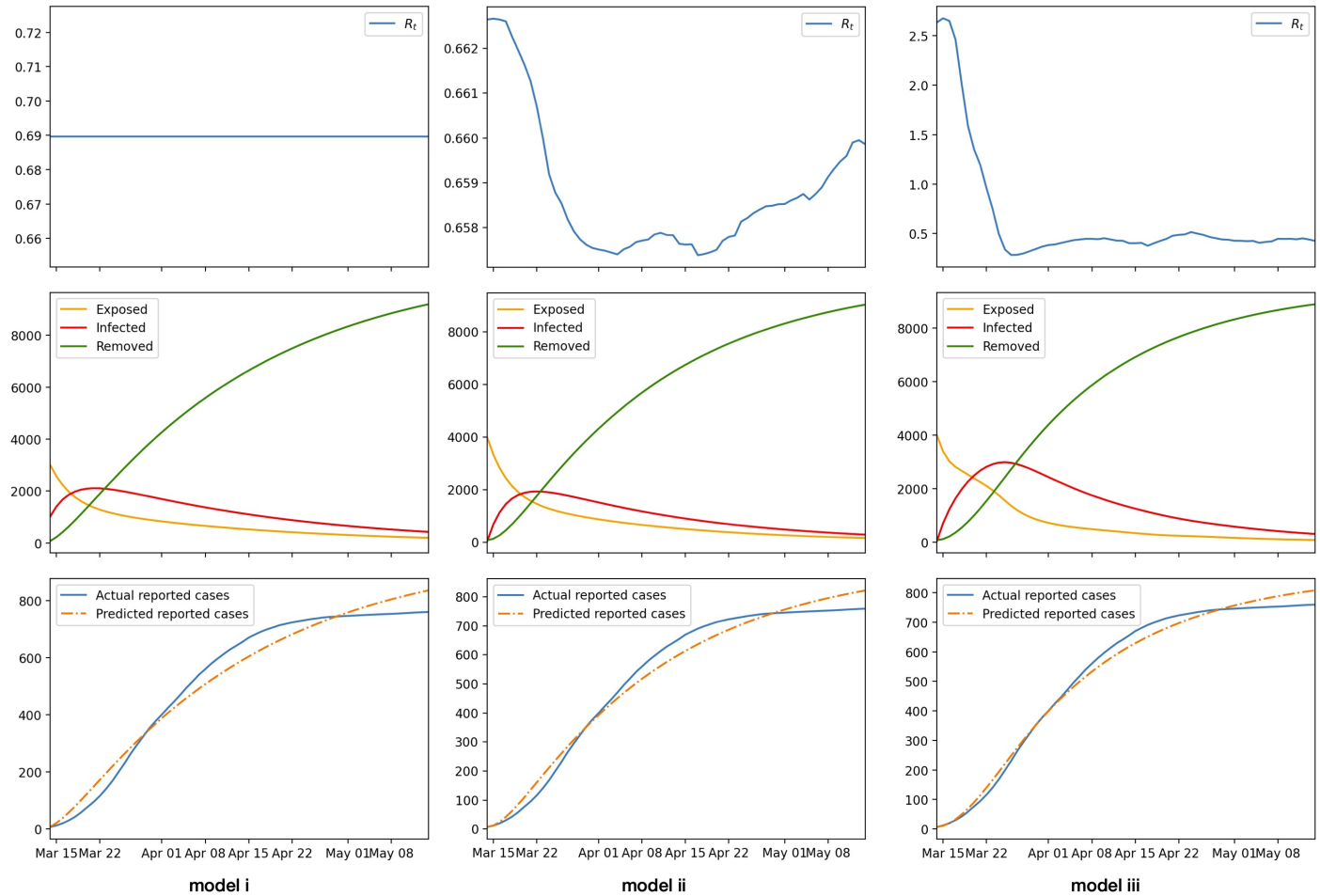

Fig A.6. Robustness check fit models. The time series values for the models fit for the robustness check. Values include (top) estimated R_0, (middle) estimated compartment populations, and (bottom) actual reported cases versus predicted reported cases. Plots are shown for each of the models: (left) model i is the baseline model and has a constant transmission rate, (middle) model ii uses trips data, (right) model iii uses trips and entrances data.