



Article

# Wastewater-Based Epidemiology for SARS-CoV-2 in Northern Italy: A Spatiotemporal Model

Matilde Fondriest <sup>1,\*</sup>, Lorenzo Vaccari <sup>1,†</sup>, Federico Aldrovandi <sup>2</sup>, Laura De Lellis <sup>3</sup>, Filippo Ferretti <sup>4</sup>, Carmine Fiorentino <sup>4</sup>, Erica Mari <sup>1,4</sup>, Maria Grazia Mascolo <sup>1</sup>, Laura Minelli <sup>5</sup>, Vincenza Perlangeli <sup>4</sup>, Giuseppe Bortone <sup>1</sup>, Paolo Pandolfi <sup>4</sup>, Annamaria Colacci <sup>1</sup> and Andrea Ranzi <sup>1</sup>

- <sup>1</sup> Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, 40139 Bologna, Italy; lvaccari@arpae.it (L.V.); ericamari1997@gmail.com (E.M.); mmascolo@arpae.it (M.G.M.); gbortone@arpae.it (G.B.); acolacci@arpae.it (A.C.); aranzi@arpae.it (A.R.)
- <sup>2</sup> Alma Mater Institute on Healthy Planet, Department of Biological, Geological and Environmental Sciences, University of Bologna, 40138 Bologna, Italy; federico.aldrovandi3@unibo.it
- <sup>3</sup> Hera Tech Srl, 40127 Bologna, Italy; laura.delellis@gruppohera.it
- <sup>4</sup> Local Health Authority of Bologna, Department of Public Health, 40124 Bologna, Italy; filippo.ferretti@ausl.bologna.it (F.F.); carmine.fiorentino@ausl.bologna.it (C.F.); vincenza.perlangeli@ausl.bologna.it (V.P.); paolo.pandolfi@ausl.bologna.it (P.P.)
- <sup>5</sup> Water Directorate, Hera SpA, 40127 Bologna, Italy; laura.minelli@gruppohera.it
- \* Correspondence: mfondriest@arpae.it
- † These authors contributed equally to this work.

**Abstract:** The study investigated the application of Wastewater-Based Epidemiology (WBE) as a tool for monitoring the SARS-CoV-2 prevalence in a city in northern Italy from October 2021 to May 2023. Based on a previously used deterministic model, this study proposed a variation to account for the population characteristics and virus biodegradation in the sewer network. The model calculated virus loads and corresponding COVID-19 cases over time in different areas of the city and was validated using healthcare data while considering viral mutations, vaccinations, and testing variability. The correlation between the predicted and reported cases was high across the three waves that occurred during the period considered, demonstrating the ability of the model to predict the relevant fluctuations in the number of cases. The population characteristics did not substantially influence the predicted and reported infection rates. Conversely, biodegradation significantly reduced the virus load reaching the wastewater treatment plant, resulting in a 30% reduction in the total virus load produced in the study area. This approach can be applied to compare the virus load values across cities with different population demographics and sewer network structures, improving the comparability of the WBE data for effective surveillance and intervention strategies.

**Keywords:** wastewater-based epidemiology; SARS-CoV-2; COVID-19; wastewater; biodegradation; sewer network; spatiotemporal model; public health; early-warning system



**Citation:** Fondriest, M.; Vaccari, L.; Aldrovandi, F.; De Lellis, L.; Ferretti, F.; Fiorentino, C.; Mari, E.; Mascolo, M.G.; Minelli, L.; Perlangeli, V.; et al. Wastewater-Based Epidemiology for SARS-CoV-2 in Northern Italy: A Spatiotemporal Model. *Int. J. Environ. Res. Public Health* **2024**, *21*, 741. <https://doi.org/10.3390/ijerph21060741>

Academic Editors: Christos Chalkias and Demosthenes Panagiotakos

Received: 4 April 2024

Revised: 23 May 2024

Accepted: 31 May 2024

Published: 6 June 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The global effort against the COVID-19 pandemic has underscored the need for innovative and robust surveillance methods to track the prevalence of SARS-CoV-2, the causative agent of COVID-19, within communities. Wastewater-based epidemiology (WBE) has emerged as a valuable tool for the proactive and scalable monitoring of viral presence in populations [1], and several countries have successfully implemented WBE as a surveillance system complementary to clinical testing [2,3]. This surveillance method not only aids in the early detection and prevention of disease outbreaks but also enables timely public health interventions such as targeted testing, contact tracing, and resource allocation [1,2]. Several studies [4,5] have emphasized the importance of environmental surveillance as an early warning system to detect the viral levels in the population and identify outbreaks before cases are reported to the healthcare system.

Recently, the application of this method as a complementary approach to clinical surveillance has become even more important given the reduction in recorded tests and the widespread occurrence of asymptomatic or mildly symptomatic cases [6]. Indeed, wastewater serves as a collective pool of genetic material shed by an entire community, providing insights into the prevalence of SARS-CoV-2, including asymptomatic and presymptomatic cases [1,2,7]. The advances in the monitoring of SARS-CoV-2 in wastewater have been significant, with molecular techniques such as quantitative reverse transcription—polymerase chain reaction (qRT-PCR) playing a pivotal role in the accurate quantification of viral RNA [1,3].

Italy, one of the hardest hit countries, has been monitoring urban sewage since July 2020 as part of a pilot study: the “SARI project” (Epidemiological Surveillance for SARS-CoV-2 in urban sewage in Italy), coordinated by the National Institute for Public Health and involving a network composed of regions, wastewater service providers, regional environmental protection agencies, and local health authorities [8].

On 17 March 2021, the European Commission recommended that EU member states begin monitoring SARS-CoV-2 in wastewater by 1 October 2021 [9]. Therefore, since October 2021, in Italy, existing research activities within the SARI project have been transformed into a surveillance system, with two main focuses: analyzing the trend of SARS-CoV-2 in urban wastewaters to predict epidemiological trends in the population and studying the spread of SARS-CoV-2 variants over time [8].

Following the principles of WBE and focusing on the surveillance objective defined by the SARI project, this study aimed to predict the temporal and spatial distribution of COVID-19 cases starting from SARS-CoV-2 detection in the main wastewater treatment plant (WWTP) of a medium-sized city in northern Italy.

This interest aligns with several studies conducted in many other countries, where various modeling techniques have been employed in the development of WBE for COVID-19 surveillance [10], including regression [5,11], artificial intelligence [12,13], and deterministic models [3,14–17].

In the present study, a deterministic model was applied: the equation upon which the model was based was initially proposed by Ahmed [14] and subsequently employed in various other studies [3,15–17]. The equation was adapted to integrate a spatial component, which enabled the calculation of the spatial and temporal distribution of COVID-19 cases by adjusting the SARS-CoV-2 RNA concentration data at the WWTP to take into account virus biodegradation effects.

The model accurately simulated the biodegradation of the virus along the sewer network via an approach similar to that employed by McCall et al. [18]. This estimate, together with the characteristics of the population and its geographical distribution, allowed us to determine the virus load produced in each zone within the study area and the corresponding number of new cases over time.

This approach proposes a methodology that can be effectively implemented in multiple cities where WBE is conducted. This methodology enables a reliable comparison of virus load values measured at WWTPs among cities characterized by different population demographics and sewer network structures.

The performance of the model was evaluated by comparing the estimated cases with those recorded by the regional healthcare system and considering various relevant health variables to interpret the results.

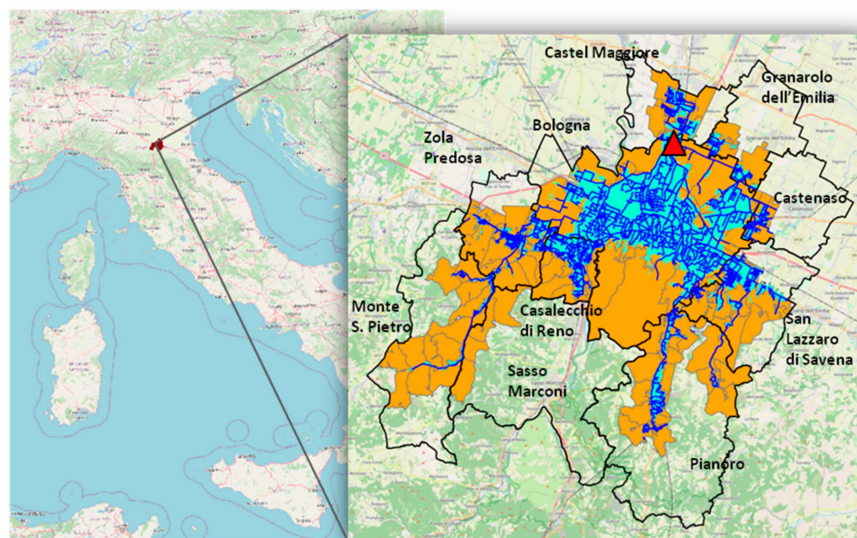
## 2. Materials and Methods

### 2.1. Study Area and Population Analysis

This study was conducted based on the measurements detected at the main WWTP in Bologna, the capital of the Emilia-Romagna region in northern Italy. The plant is located in the north of Bologna and serves the city and neighboring municipalities, catering to a capacity of 800,000 population equivalents.

The study area was defined as the area served by the WWTP. The study was conducted at the sub-municipal level; therefore, the municipalities were divided into census sections [19], which are already widely used for statistical analyses. Only Bologna was divided into statistical areas [20], as these areas are more extensive and provide more recent data than the census sections [19].

A three-step selection process was followed to define the areas (i.e., census sections and statistical areas) included in the study area. First, the study area was defined as the union of the areas encompassed by or intersecting the catchment area served by the sewer network. Additionally, the areas enclosed by the previously selected areas were included. Finally, those areas that were mainly connected to another treatment facility were excluded from the selected areas. The resulting study area included almost all the statistical areas of Bologna and a part of the census sections of the neighboring municipalities (Figure 1).



**Figure 1.** Map of the study area (orange) divided into census sections and statistical areas (delineated in gray). Representation of the wastewater treatment plant (WWTP) (red triangle), sewer network (blue), catchment area (light blue), names, and boundaries of the municipalities (black).

The study population is an estimate of the number of individuals domiciled in the study area, obtained by multiplying the number of residents [19,20] by a coefficient derived from the data in the Regional Assistance Registry, which represents the ratio of the domiciled population to the resident population at the municipal level (Table S1).

## 2.2. Virus Concentration and Wastewater Flow Rate at the WWTP

The concentration of SARS-CoV-2 RNA [GC/L] and the influent flow rate [ $\text{m}^3/\text{day}$ ] observed at the inlet of the WWTP were obtained from a data sharing platform within the SARI project network. The study period was chosen according to the availability of concentration data and ranged from 13 October 2021 (the first day of sampling) to 24 May 2023. Measurements were taken weekly or twice a week, with an irregular frequency, resulting in a total of 165 data points over the whole period.

Wastewater samples were collected and processed for RNA extraction and virus concentration determination according to the reference analytical protocol established for the SARI project [21]. Composite wastewater samples of 100 mL were collected over a 24 h period. A 50 mL aliquot was frozen and constituted the “archive sample” to be retained for any further determinations. The second 50 mL aliquot was processed immediately or stored at a refrigerated temperature ( $-20\text{ }^\circ\text{C}$ ) for a maximum of 24–48 h until analysis.

The wastewater sample concentration was achieved with a PEG/NaCl protocol according to the method published by Wu et al. [22]. The SARS-CoV-2 analysis included a first step of cell lysis with guanidine isothiocyanate, followed by extraction of the genetic material

based on the adhesion of nucleic acids to magnetic silica using the MiniMag/eGeneUP (bioMerieux Italia Spa, via di Campigliano, 58 | Loc. Ponte a Ema-50012 Bagno a Ripoli (FI), Italy) platform with a final volume of 100  $\mu$ L for RNA elution. The quantitative detection of SARS-CoV-2 was performed via real-time RT-qPCR in accordance with the EU 2021/472 Recommendation [9]. The quantitative determination of SARS-CoV-2 was based on ORF-1ab (nps 14), with Murine Norovirus used as a control virus during the analysis.

The mean ( $1.79 \times 10^5$  [GC/L]), median ( $0.83 \times 10^5$  [GC/L]), and maximum ( $1.10 \times 10^6$  [GC/L]) concentrations of SARS-CoV-2 RNA measured at the WWTP (Figure S2, Table S2) were within the range ( $10^2$ – $10^5$  [GC/L]) with maximum values above  $10^6$  [GC/L] found in the literature [4,5,23,24].

The influent flow rate at the WWTP inlet was measured using two electromagnetic flow meters (MG3100 and MG5100) installed on two pipelines (DN 1800 and DN 1400, respectively). Typically, only one pipeline was operational at a time. However, when the flow rate increased, the second pipeline was opened and the other electromagnetic flow meter was activated. The measurements represent the average daily flow rate. In the study period, the median flow rate was  $118.7 \times 10^3$  [m<sup>3</sup>/day], the minimum flow rate was  $96.8 \times 10^3$  [m<sup>3</sup>/day], and the maximum flow rate was  $432.9 \times 10^3$  [m<sup>3</sup>/day] (Figure S3, Table S3).

In combined sewer systems, such as the one described in this study, intense rainfall events can cause the significant dilution of the SARS-CoV-2 RNA concentration in the sewer, as well as fluctuations in flow velocity and other parameters. Therefore, the concentration data collected when the flow rate exceeded the 90th percentile of the values recorded in 2022 ( $156.5 \times 10^3$  [m<sup>3</sup>/day]) were treated as outliers and thus removed [11] (Table S3, Figure S3). In support of this method, the correlation between the influent flow values measured at the WWTP inlet and the cumulative daily precipitation values [25] measured by the monitoring stations in Bologna and neighboring municipalities was analyzed. The highest correlation (Pearson 0.88) was observed when considering the cumulative daily precipitation of the day preceding the flow measurement. This could be attributed to evening rains that reach the WWTP after several hours, contributing to the cumulative influent flow of the following day.

The daily load of SARS-CoV-2 RNA at the WWTP inlet [GC/day] was obtained by multiplying the observed values (after removing the outliers) of the SARS-CoV-2 RNA concentration by the influent flow rate. The historical time series of the daily SARS-CoV-2 RNA load was then processed using the locally estimated scatterplot smoothing (LOESS) method [26]. This was performed with a dual purpose: first, to transform the historical series into one with a daily regular data interval based on irregularly measured data points, and second, to smooth the values to obtain a signal purified from potential measurement errors (Figure S4). The LOESS method was implemented in Python [27], considering 11 neighboring data points for the local value estimation as in Rauch et al. [11].

### 2.3. Health Data

The Local Health Authority of Bologna provided daily COVID-19 cases within the statistical areas and census sections where the subjects were isolated.

For each patient, the following information was available: the census section or statistical area of isolation, the presence of symptoms and the date of onset of symptoms if symptomatic, the date of the swab test, diagnosis date (i.e., the date on which a positive case is officially recorded in the healthcare system), the admission date and discharge date in the case of hospitalization, the final outcome (i.e., whether the individual recovered or deceased), the final outcome date, and the cause of death if deceased.

The model presented in this article was implemented to estimate the new daily cases. Therefore, the data on the new daily cases reported from the healthcare system were selected to evaluate the performance of the model. The decision to estimate the new daily cases instead of the active cases was made in recognition that the duration of an infected individual's classification as an active case is likely to be more uncertain, varying

significantly from person to person, and influenced by the prevailing COVID-19 regulations (e.g., a confirmatory test was not always required to confirm the end of the infection period).

To reduce the individual variability, the date of symptoms onset (or the date of the swab test if asymptomatic) was taken as the reference date for defining a new case, hospitalization, or death reported by the healthcare system. Because of the correspondence between symptoms onset and the peak of SARS-CoV-2 shedding in feces, this was considered the best reference date for comparing predicted and reported cases [15,28,29].

The daily city-level data on the tests and vaccines (second and third doses) were provided by the Local Health Authority of Bologna. An analysis of these health variables was performed to interpret the model results.

#### 2.4. Model Equation

The model was based on an equation used in previous studies [3,14–17] to model the relationship between the SARS-CoV-2 concentration at WWTPs and the number of COVID-19 cases. This equation was adapted by integrating two different factors to consider the spatial component. This adjustment was made to obtain a more accurate estimate of virus biodegradation and, consequently, to derive the actual contributions of SARS-CoV-2 RNA from different areas within the catchment area and thus predict the corresponding number of COVID-19 cases.

A virus biodegradation factor and a population-based coefficient were introduced into the equation to differentiate the contribution of each area to the observed concentration at the WWTP. Both factors were calculated for each area.

The daily load of SARS-CoV-2 at the inlet of the WWTP [GC/day] was set equal to the sum of the daily virus load generated by each area, which was reduced to account for the virus biodegradation along the sewer network and refined by incorporating the population-based coefficient associated with the population of each area (Equation (1)).

$$C_W(t) \cdot Q_W(t) = S_h \cdot M_f \cdot IR(t) \cdot \sum_i \left( P_i \cdot e^{-k\theta_i} \cdot W_i \right) \quad (1)$$

where  $C_W(t)$  is the measured concentration of SARS-CoV-2 RNA at the inlet of the WWTP [GC/L],  $Q_W(t)$  is the wastewater flow rate at the inlet of the WWTP [L/day] ( $C_W(t)$  multiplied by  $Q_W(t)$  represents the daily load of SARS-CoV-2 RNA [GC/day] as derived in Section 2.2),  $S_h$  is the fecal shedding rate (i.e., the number of SARS-CoV-2 RNA genomic copies per gram of feces produced by each infected individual) [GC/g],  $M_f$  is the mass of feces produced per inhabitant per day [g/(inhabitant · day)],  $i$  represents the area within the municipality,  $P_i$  is the population in the area  $i$  [inhabitant],  $k$  is the biodegradation constant [ $\text{h}^{-1}$ ],  $\theta_i$  is the hydraulic residence time for the area  $i$  [h],  $W_i$  is the population-based coefficient characteristic of each area  $i$  [adim.], and  $IR(t)$  [adim.] is the average infection rate (i.e., the number of cases divided by population) across the entire study area, and it is the unknown variable of the equation (the methodology followed to derive the equation is detailed in the Supplementary Materials, Section S3). Once  $IR(t)$  was determined, the number of cases in each area over time ( $c_i$ ) was derived (Equation (2)).

$$c_i(t) = IR(t) \cdot P_i \cdot W_i \quad (2)$$

The model output was the number of predicted cases per day for each census section or statistical area included in the study area. However, when comparing the predicted and reported cases, the results were aggregated into grouped areas. The municipal boundaries were used to group the census sections, while for Bologna, a proximity criterion was chosen to group some statistical areas together. The aim of this approach was to increase the size of the population within each comparison zone, thereby enhancing the statistical significance when observing the spatial variability in the cases. The 88 statistical areas within the city of Bologna were therefore grouped into 59 areas, ensuring that none had less than 2000 inhabitants, while the aggregation of census sections by municipality resulted in zones with more than 2700 inhabitants (Figure S1, Table S12).

The infection rate for each grouped area over time  $IR_j(t)$  was then obtained (Equation (3)).

$$IR_j(t) = \frac{c_j(t)}{P_j} = IR(t) \cdot W_j \quad (3)$$

where  $c_j$  represents the number of cases in each grouped area,  $P_j$  is the total population of the grouped area, and  $W_j$  is the population-based coefficient for grouped areas (weighted average of  $W_i$  based on the population size).

#### 2.4.1. Population-Based Coefficient

The population-based coefficient ( $W_i$ ) allowed us to consider for each area the following population characteristics in the prediction of cases: age, sex, family size [19,20], and comorbidities. The comorbidity data were provided by the Local Health Authority of Bologna and were obtained from different databases related to the years 2021–2022 (hospital discharge forms [SDO], territorial pharmaceutical care [AFT], direct dispensing drugs [FED], and specific pathology).

The  $W_i$  coefficient, which is constant in time and variable across space, was defined as the ratio between a specific infection rate for area  $i$  ( $A_i$ ) and the infection rate for the entire study area ( $A_{TOT}$ ) (Equation (4)). To establish  $A_i$  and  $A_{TOT}$ , the Poisson regression was employed, and the regression coefficients ( $\beta_0, \beta_1, \dots, \beta_n$ ) were derived from a study conducted in the municipality of Bologna from February 2020 to November 2021 [30] (Table S4). The variables ( $x_1, x_2, \dots, x_n$ ) were calculated for each area and represent the fractions of the population belonging to the following classes: age (0–21, 21–65, >65), sex (M/F), family size (1, 2, 3, >3), and comorbidities (hypertension, diabetes, and the “other comorbidities” considered in the study [30]).

$$W_i = \frac{A_i}{A_{TOT}} = \frac{e^{\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n x_{ni}}}{e^{\beta_0 + \beta_1 x_{1TOT} + \beta_2 x_{2TOT} + \dots + \beta_n x_{nTOT}}} \quad (4)$$

#### 2.4.2. Fecal Shedding Rate and Mass of Feces

Two average values of the fecal shedding rate ( $S_i$ ) were calculated using the data from the Delta and Omicron variants in six different communities in the USA [31]. The population-weighted average shedding rates for the Delta and Omicron variants included in the model were  $10^{8.658}$  and  $10^{7.813}$ , respectively (Table S6).

Considering the evolution of the virus from the Delta variant to the Omicron variant in December 2021, the shedding rate for the Delta variant was used until 19 December, while that for the Omicron variant was applied from 3 January onwards, and a combination of the two values was used for the last two weeks of December (Table S8). The daily wet mass of feces ( $M_f$ ) produced by each individual was assumed to be 128 [g/(inhabitant · day)] [32].

#### 2.4.3. Biodegradation

The biodegradation of the virus along the sewer network was modeled using a first-order kinetics equation (Equation (S5)) [33–35]. The biodegradation constant ( $k$ ) was calculated on the basis of the values found in the experiments [4] conducted under the conditions most similar to those of the present study. The values were adjusted to the observed temperatures in the Bologna sewer network (12 °C, 16 °C, 20 °C) [36] according to a linear equation (Equation (5)) [33,35]. The linear coefficient ( $m$ ) and the intercept ( $q$ ) were calculated from the literature values of  $k$  and temperature ( $T$ ) [4], after which the  $k$  values corresponding to the observed temperatures in Bologna were determined. The three derived  $k$  values (0.097, 0.101, and 0.104 [h<sup>-1</sup>]) were included in the model, varying with the time of the year (Table S9).

$$\log_{10} k = m \cdot T + q \quad (5)$$

The hydraulic residence time ( $\theta_i$ ) for each area was estimated by dividing the effective distance ( $D_i$ ), i.e., the distance calculated along the network from the centroid of each area to the WWTP, by the wastewater flow velocity along the sewer network ( $v$ ).  $D_i$  was determined by applying the Dijkstra algorithm [37] to measure the shortest path along the sewer network from the centroid of each area to the WWTP (using the GeoPandas library (version 0.13.2) in Python).

The value of  $v$  was assumed to be constant and equal to the value (0.8 [m/s]) used for the Milan sewer network [38] due to the similarities between the two sewer systems. This value falls within the typical range of the average velocity (0.3–0.91) in the cross section of combined sewers [39].

### 2.5. Uncertainty and Sensitivity Analysis

The model uncertainty was calculated using the error propagation formula of Li et al. [3], where the uncertainty is expressed as the relative standard deviation (RSD). The formula includes the relative standard deviations of each model parameter. These uncertainty values were derived from the literature (Section S4).

The total number of cases is included in the formula because the uncertainty associated with the shedding rate and the mass of faces of each individual decreases as the number of infected individuals increases: beyond 10 cases, the impact of these uncertainty values on the total uncertainty becomes limited [3].

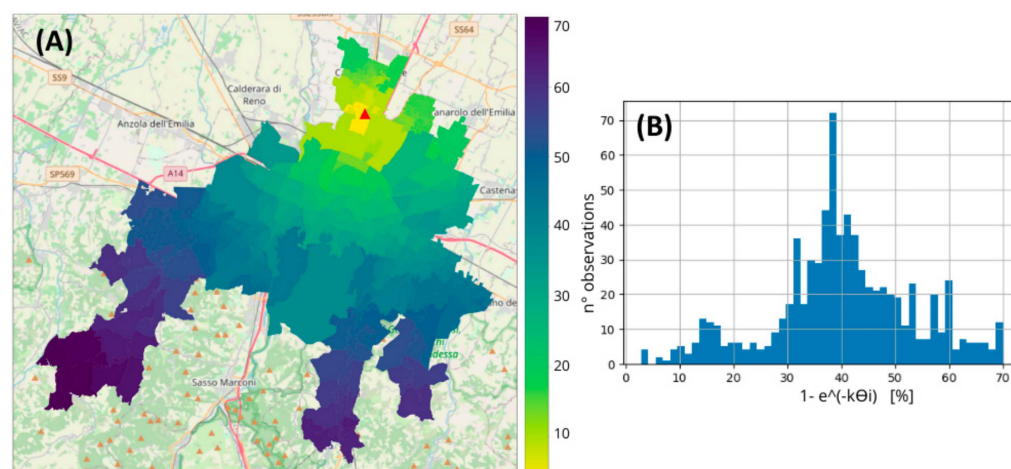
The uncertainty associated with the model results varied from 0.81 RSD to 0.33 RSD depending on the total number of infected individuals in the area.

The unknown variable ( $IR(t)$ ) exhibited a linear dependence on all the model parameters except  $k$  and  $\theta$ . Therefore, a 1% variation in any of the parameters resulted in a corresponding variation in  $IR(t)$  of  $\pm 1\%$ , while a 1% increase in  $k$  or  $\theta$ , resulted in a corresponding increase in  $IR(t)$  of 0.35% (Equation (S9)).

## 3. Results

### 3.1. Effect of Biodegradation

More than half of the areas considered were located within an effective distance range of 10 to 18 km from the WWTP (corresponding hydraulic residence time: 3.5–6 h). The virus load produced in these areas decreased by 30–45% along the path in the sewer network due to biodegradation. The virus loss was highly significant (60–70%) in the areas with the greatest effective distance (>30 km) from the WWTP (Table 1, Figure 2). However, the population density in these areas was very low.



**Figure 2.** (A) Map of biodegradation in the study area, the WWTP (red triangle). (B) Frequency distribution of biodegradation in the study area. Biodegradation is expressed as a  $1 - e^{-k\theta_i}$  percentage, calculated with  $k = 0.101$  ( $T = 16$  °C).

**Table 1.** Maximum, minimum, median, standard deviation, 25th percentile, and 75th percentile values of the following: linear distance of centroids from the WWTP [km], effective distance (calculated along the sewer network) of centroids from the WWTP ( $D_i$  [km]), hydraulic residence time ( $\theta_i$  [h]), and biodegradation ( $1 - e^{-k\theta_i}$  [%]) varying with the biodegradation constant  $k$ .

	Max	Min	Median	St.Dev.	25%	75%
Linear distance $i$ [km]	26.7	0.5	10.9	4.7	8.7	13.3
$D_i$ [km]	34.4	0.8	14.5	6.6	11.8	18.8
$\theta_i$ [h]	11.9	0.3	5.0	2.3	4.1	6.5
$1 - e^{-k\theta_i}$ [%], $k = 0.097$ [h <sup>-1</sup> ]	68.6	2.8	38.6	13.0	32.8	46.9
$1 - e^{-k\theta_i}$ [%], $k = 0.101$ [h <sup>-1</sup> ]	70.0	2.9	39.8	13.2	33.8	48.3
$1 - e^{-k\theta_i}$ [%], $k = 0.104$ [h <sup>-1</sup> ]	71.1	3.0	40.7	13.4	34.6	49.3

The ratio between the SARS-CoV-2 RNA load measured at the WWTP and that generated in the areas of the catchment before entering the sewer system and thus being biodegraded is independent of time and can be expressed as (Equation (6)):

$$\frac{\sum_i P_i \cdot e^{-k\theta_i} \cdot W_i}{\sum_i P_i \cdot W_i} \tag{6}$$

At 16 °C, the ratio was 0.694, and this value slightly decreased as  $k$  increased. Therefore, the virus load detected at the WWTP was approximately 70% of that actually produced in the study area.

Starting from the estimate of biodegradation, the contribution of each area to the virus load detected at the WWTP was evaluated, and indicators were created to better understand the significance of the biodegradation and virus production in the different areas (Section S5.1).

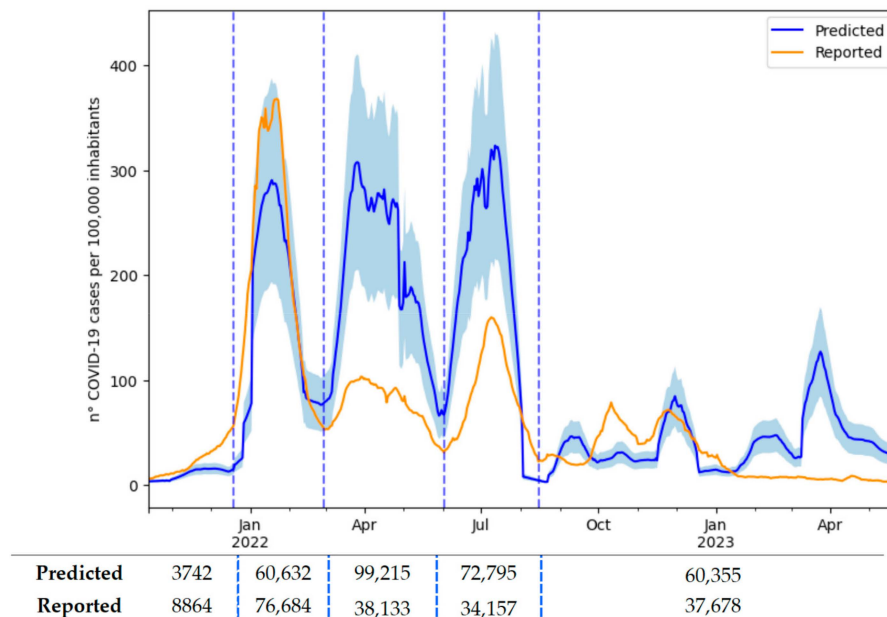
The contribution of hospitals to the viral load discharged into the sewer system and detected at the WWTP during the study period was evaluated based on the number of hospitalized individuals. While the ratio of the COVID-19 patients hospitalized to the total number of cases reported in the whole study area was very low (0.03), this value was much higher (i.e., 42, 14, and 4) when referring only to the cases reported in the three statistical areas of Bologna where hospitals were located. These coefficients were used to estimate the virus load generated in each area with hospitals and contributing to the virus load detected at the WWTP, taking into account biodegradation. For the area where the hospital with the highest number of COVID-19 patients was located, this predicted contribution ranged from 0.014%, without considering the hospitalized subjects, to 1%, considering the hospitalized subjects (Table S13).

### 3.2. Predicted and Reported Infection Rates in the Study Area

The predicted and reported daily  $IR(t)$  trends over the study area were compared throughout the entire study period (Figure 3). The reported  $IR(t)$  was obtained as the daily sum of the reported COVID-19 cases in the area relative to the date of symptom onset or the date of the swab test (Section 2.3), smoothed with a simple moving average over 7 days and divided by the total population within the study area. The mean absolute error between the predicted and reported  $IR(t)$  over the entire study period was 54 cases per 100,000 inhabitants, and the Spearman correlation coefficient was 0.599.

The time period was subdivided according to the trend of the reported cases in order to better compare the predicted and reported data. The transition dates were chosen to correspond to the relative minima of reported cases, except for the beginning of the first wave, for which the shift between the Delta and Omicron variants was chosen. This approach defined three waves and an “off-waves” period (Table 2).





**Figure 3.** Predicted (blue) and reported (orange) infection rate  $IR(t)$  across the entire study area and the uncertainty interval (RSD) (light blue) associated with the predicted  $IR(t)$ . Values of total predicted and reported COVID-19 cases over the single wave (delimited by dashed vertical lines) in the entire study area.

**Table 2.** Mean absolute error (MAE), mean absolute percentage error (MAPE), and the Spearman correlation coefficient between predicted and reported  $IR(t)$  in different time intervals (first, second, and third waves, off-waves period, and the entire study period).

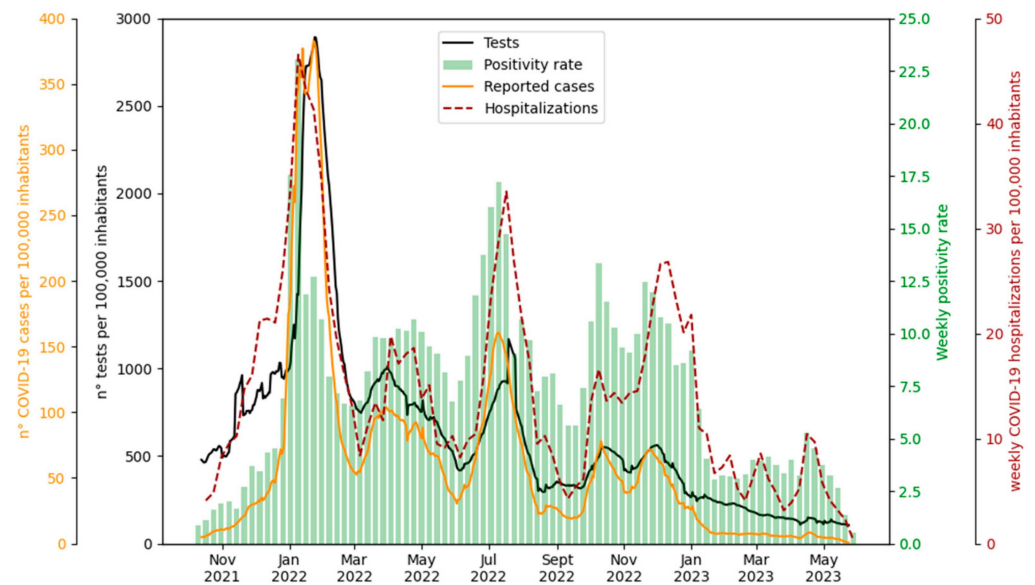
	1° Wave (19 December 2021–28 February 2022)	2° Wave (1 March 2022–2 June 2022)	3° Wave (3 June 2022–15 August 2022)	Off-Waves (13 October 2021–19 December 2021 and 15 August 2022–24 May 2023)	Entire Study Period (13 October 2021–24 May 2023)
MAE	47	120	109	26	54
MAPE	0.3	1.5	1.3	3.6	2.5
Spearman	0.858	0.954	0.897	-0.223	0.599

The correlation was consistently high across the three waves but significantly lower during the “off-waves” period. In the first wave, the predicted and reported cases were quite similar, with an underestimation within the uncertainty range (mean absolute percentage error of 0.3), whereas in the subsequent waves, an overestimation by the model was observed (possible explanations are provided in Section 4.3).

The time trends for the total tests and reported cases were very similar, showing a high correlation (Spearman 0.87). For both the tests and cases, the values in the first wave were significantly higher than those observed in the subsequent waves, whereas the positivity rate (i.e., the proportion of positive cases among the tests performed) showed a less pronounced difference between the first and subsequent waves, and its trend was more similar to that of the hospitalizations (Figure 4).

In addition, to evaluate the early warning capacity of the system, the time lag between the virus load at the WWTP and the number of active cases reported over time was investigated. The diagnosis date was taken as the reference date for the active cases, corresponding to the date on which a positive case was officially registered in the healthcare system. The maximum correlation (Spearman coefficient = 0.70) was observed when

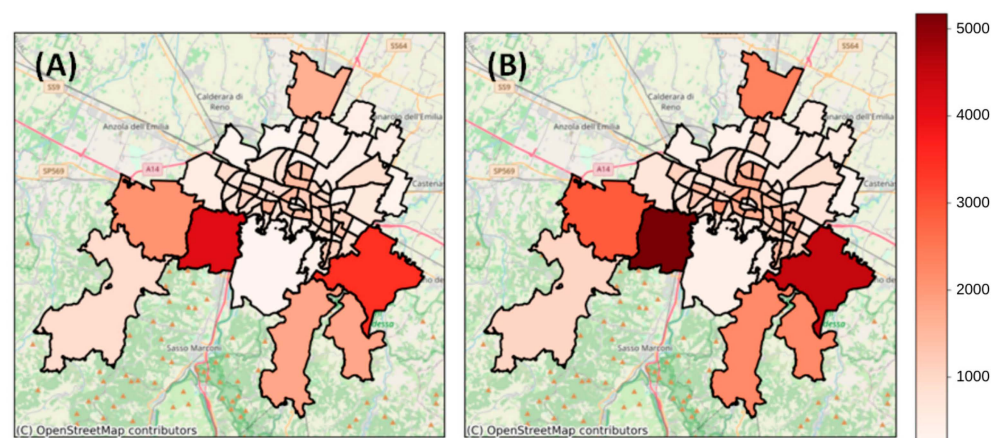
comparing the virus load data with the curve of the active cases recorded 9 days later, suggesting a time lag of 9 days.



**Figure 4.** Tests: Daily reported tests across the entire study area per 100,000 inhabitants averaged with a simple moving average (SMA) over 7 days. Positivity rate: Weekly average of positive cases among tests conducted. To calculate the positivity rate, the date on which the test was conducted was taken as the reference date for both tests and cases. Reported cases: reported daily COVID-19 cases over 100,000 inhabitants averaged with a simple moving average (SMA) over 7 days (reported  $IR(t)$ ). Hospitalizations: weekly sum of COVID-19 hospitalizations. Both reported cases and hospitalizations are represented based on the symptom onset date if symptomatic or the date of the test in asymptomatic cases.

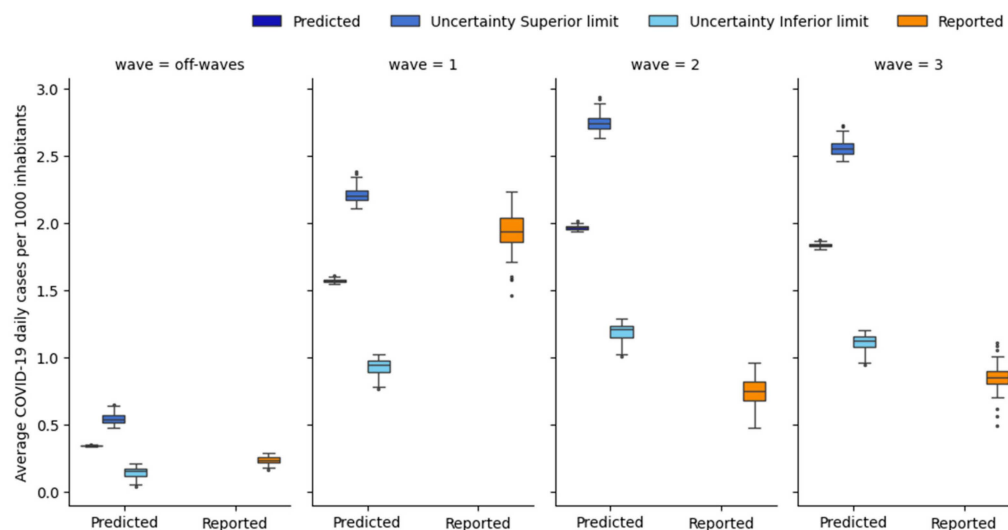
### 3.3. Spatial Comparison of the Predicted and Reported Cases in Each Grouped Area

The comparison between the predicted and reported cases in the grouped areas showed a medium/high accuracy of the model in estimating the spatial distribution of the cases (Figure 5). In particular, in the first wave, the error rate was low, with an average mean absolute percentage error across all the areas of 0.18 (min: 0.01, max: 0.20), while in subsequent waves, the error rate increased due to the greater gap between the predicted and reported daily  $IR(t)$  trends (Table S12).



**Figure 5.** Maps of predicted total cases (A) and reported total cases (B) during the first wave in the different grouped areas.

However, the model failed to accurately capture the spatial variation in  $IR_j(t)$  (Figure 6). In fact, the standard deviation of the predicted  $IR_j(t)$  among the different areas was an order of magnitude lower than that associated with the reported  $IR_j(t)$ . The minimal variations in the predicted  $IR_j(t)$  across the different grouped areas are due to the similarity in the values of the population-based coefficient  $W_j$  (max = 1.024, min = 0.983, standard deviation = 0.009) (Table S14).



**Figure 6.** Box plot of the daily average reported and predicted infection rates ( $IR_j$ ), along with the lower and upper uncertainty bounds associated with the daily average predicted  $IR_j$ . The  $IR_j$  is expressed as daily average cases per 1000 inhabitants over the wave and in the grouped area (for each box plot, one data point for each grouped area and for each wave is represented).

The standard deviation of  $W_j$  was low since the regression coefficients used to determine  $W_j$  were very low (max = 0.329, min =  $-0.008$ ) (Table S4). In addition, the population characteristics were very similar among the different areas (Table S5).

The uncertainty associated with the daily estimation of  $IR_j(t)$  in each area was calculated (Section 2.5). As the uncertainty depends on the number of cases (the lower the number of cases, the higher the uncertainty), its value varied from area to area and was higher than that calculated in the temporal analysis. Starting from the uncertainty value, the upper and lower limits associated with the estimation of  $IR_j(t)$  in each area were calculated (Figure 6).

Only during the first wave and off-wave periods did the reported values of  $IR_j(t)$  remain within the uncertainty range associated with the predicted  $IR_j(t)$ , while during the second and third waves, the reported  $IR_j(t)$  only partially overlapped with the uncertainty range.

The spatial variation in the reported  $IR_j(t)$  did not exhibit consistency over time; across the different areas, the  $IR_j(t)$  values deviated differently from the mean, showing both higher and lower values over time (Figure S7). Indeed, across the three different waves and in the off-waves period, only 21 out of 67 areas (31%) demonstrated consistent temporal variations. Of these, 11 areas consistently showed  $IR_j(t)$  values above the mean, while 10 areas consistently exhibited values below the mean. However, for 46 areas, the variation was not consistent.

In a comprehensive analysis of the three waves and the off-waves, considering the variation in  $IR_j(t)$  relative to the mean, 31 areas (46%) showed concordance between the predicted and reported variations in  $IR_j(t)$ , while this was not observed in 36 areas.

## 4. Discussion

### 4.1. Model Parameters

The fecal shedding rate ( $S_h$ ), the wet mass of feces ( $M_f$ ), and the biodegradation constant ( $k$ ) were assumed from the literature, as the site-specific data for these parameters were not available. The shedding rate showed significant variability, ranging from  $10^3$  to  $10^9$  [GC/g] [3,16,29,31]. This variability was attributed to the estimation method, which can be at the individual level [3,40] or at the population level (retrospectively determined based on the observed concentrations at WWTPs and the positive cases) [16,31]. In addition, shedding depends on factors such as population characteristics (e.g., sex, age, ethnic group [31], and health status [3]), viral variants or sub-variants [31], symptoms presented, and the stage of the disease [41,42].

Given the considerable variability in shedding values, two values were selected specifically for the variants prevalent during the study period (Delta and Omicron) [31]. Moreover, we averaged the shedding values calculated across diverse and numerous populations to account for the inter-individual shedding variability.

Prasek et al. [31] considered a six-day sum of reported cases to be appropriate to represent the number of infected individuals contributing to the daily virus load in wastewater. Therefore, the shedding coefficients derived from Prasek et al. [31] were included in the model as daily averages. This allowed us to consider that the cases contributing to the concentration at the WWTP were not only the new daily cases but also the active cases with a high shedding rate, i.e., the cases occurring within the six days surrounding the day of the measurement [15,43].

The value chosen for the daily wet mass of the feces produced by each individual represents the median of data collected across the various communities in many countries over a long period (1934–2011) [32].

The value of the biodegradation constant varies depending on the type of sewer system (combined or separate), wastewater temperature and pressure [18,23,44], the initial virus concentration [18,33] and structure [34], and the other influent parameters, such as suspended solids [45], biofilm [46], BOD, and pH. The values of  $k$  in the literature range from  $0.004$  [ $\text{h}^{-1}$ ] [33] to  $0.120$  [ $\text{h}^{-1}$ ] [4], as determined from the experiments conducted under specific conditions. In the present study, the value for the biodegradation constant was derived from an article [4] that closely matched our conditions: the  $k$  value was determined by analyzing the fluctuations in the virus concentration over short time intervals, and the virus was naturally present in the collected wastewater without any additional introduction. In fact, the SARS-CoV-2 shed in feces may contain intact virus, compromised capsid virus, and free nucleic acids; thus, the rapid decay of SARS-CoV-2 RNA observed in the study [4] reflects the decay of all three viral RNA sources typically found in naturally contaminated wastewater.

Notably, an improvement and possible future development of the model will involve the derivation of site-specific values for the described parameters.

### 4.2. Effect of the Biodegradation and Contribution of Each Area

The impact of biodegradation on reducing the viral load reaching the WWTP was assessed by estimating the reduction between the viral load generated from each area and that detected at the WWTP. The maximum reduction reached 70%, with an average reduction of approximately 30%, indicating the substantial influence of biodegradation on reducing the virus load in the sewer network.

In addition, the virus load [GC/day] produced in the entire study area over time, the removal of the effects of biodegradation, and the geographical distribution of the population were determined (Equation (6)) to allow the comparison of this parameter across different cities. According to the literature, the normalized value at the WWTP can be obtained by dividing the measured load by the total population served by the plant [47]. Our approach provides a more accurate normalization method that takes into account the actual distance of the population from the WWTP.

Furthermore, in the pursuit of the monitoring and early warning objectives, indicators were developed to assess the impact of the population density and biodegradation on the virus load detected at the WWTP. These indicators showed that densely populated areas, even those far from the WWTP, contributed significantly to the virus load at the WWTP. However, the contribution of distant areas was mitigated by the high levels of biodegradation. This led to the identification of potential monitoring sites.

In addition, the hospital contributions to the viral load during the study period were calculated retrospectively. In fact, the contribution of each area estimated by the model is based on the population domiciled in the area and does not take into account the possibility that the infected individuals may be located in different areas. The contribution of each hospital to the total viral load at the WWTP was considered negligible, whereas the individual contribution of each hospital was significant compared to the contribution of the area in which it was located.

#### 4.3. Health Variables Analysis to Compare the Predicted and Reported Infection Rates

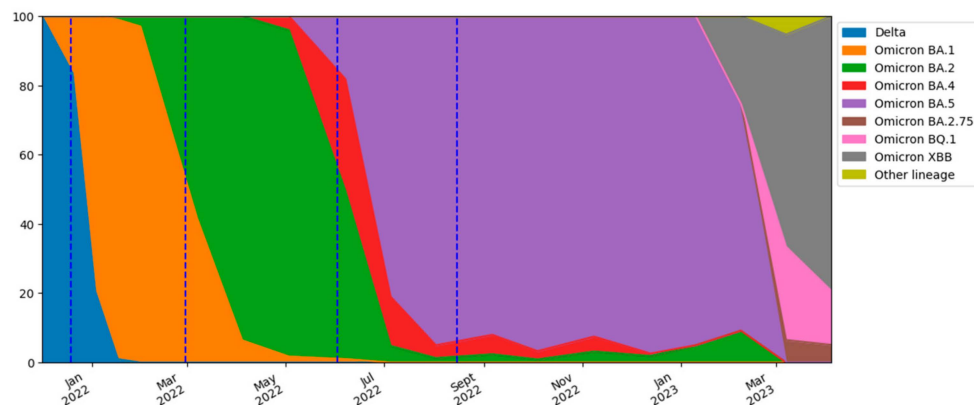
When interpreting the results of the comparison between the reported and predicted COVID-19 cases (Figure 3), it is important to consider that the reported number of positive cases in the health surveillance system was affected by an error due to the presence of many positive individuals who were not identified. This number of unreported cases varied over the study period due to a combination of factors including the changes in the number of tests carried out and in isolation protocols, the effectiveness of the vaccine in reducing symptoms, the milder variants, and the psychological impact on individuals [6].

Identifying an appropriate time period to compare the predicted and reported cases was challenging. Ideally, validation should be performed when the number of unreported cases is as low as possible. However, the quantitative data on underreporting were not available. Therefore, a qualitative analysis of health variables that may have influenced the number of reported and predicted cases over time was conducted to better understand the comparison between them: time trends in vaccination rates, testing, and sub-variants were examined across the entire population of the study area.

From October 2021 to March 2022, the number of individuals receiving the third dose increased from 0% to 60% of the total resident population, peaking in December 2021 and January 2022. The increase in individuals receiving the third dose might have played a role in the decrease in the number of reported cases in the waves following the first wave. In fact, about 15 days after the administration of the third dose, individuals might have been less susceptible to severe outcomes, potentially affecting the detection of cases.

The decrease in the number of tests recorded since the end of January 2022 (Figure 4) could be partly due to the introduction of rapid self-diagnostic tests in the Emilia-Romagna region from 19 January. From the positivity rate (Figure 4), it is clear that if the number of tests performed during the three waves had been equivalent, more cases would probably have been reported than were actually recorded in the second and third waves. This supports the hypothesis that the underreporting was greater in these waves. This hypothesis also seems to be supported by the hospitalization curve. In fact, the ratio of hospitalizations to new cases in January 2022 was the lowest of the entire period (2%), suggesting a greater number of unreported cases in the other waves.

Furthermore, a transition from the Delta variant to the Omicron variant (BA.1) was observed during the study period. According to the rapid regional surveys [48], the presence of the Omicron variant in the population increased exponentially in December 2021, overtaking the Delta variant between the penultimate week and last week of the month. This transition was followed by further variations within the Omicron lineage (BA.2, BA.4, BA.5, BA.2.75, BQ1, and XBB) (Figure 7). An increase in COVID-19 cases was observed concurrently with the variant or lineage variations (Figures 3 and 7), in agreement with other studies [6,49]. Furthermore, the milder fluctuations in the predicted and reported cases in the months from October to December 2022 coincided with the minor variations in the Omicron lineages.



**Figure 7.** Percentage of SARS-CoV-2 variant lineages over time obtained from regional surveys.

It is important to emphasize that the shedding values used in the model [31] were only related to the Delta and Omicron variants (mainly BA.1). Therefore, the model overestimation in the second and third waves could also be attributed to a change in shedding due to the shift from Omicron BA.1 to BA.2 and from Omicron BA.2 to BA.4/BA.5, respectively. In fact, the shedding associated with the BA.1 sub-variant is likely to be lower than that associated with Delta and Omicron BA.2 and BA.5 [49]. However, this overestimation does not fully justify the magnitude of the observed difference between the predicted and reported cases.

In conclusion, the most reliable period for the accurate validation of the model could be from October 2021 to the end of February 2022, when the number of underreported cases was likely to be the lowest. According to the reported and predicted cases during this period, the model underestimated the number of cases reported by the healthcare system. This discrepancy, although within the uncertainty interval, is still noteworthy and may be due to the use of the parameters from the literature, such as the fecal shedding rate.

#### 4.4. Spatial Comparison of the Predicted and Reported Cases in Each Grouped Area

The temporal inconsistency in the variation in the reported  $IR_j$  within the same area suggests that a random effect could have caused the increase in cases in one area rather than another. Furthermore, in the areas where the variation in the reported  $IR_j$  remained consistent over time, other factors such as the deprivation and obesity indices [50] may need to be considered to explain the variability in  $IR_j$ .

Additionally, a limitation of the proposed approach is the use of an estimate of the population domiciled in each area (Section 2.1) to calculate the reported  $IR_j$ . In summary, the model accurately predicted the spatial distribution of cases because the reported  $IR_j$  showed minimal variations among the different grouped areas. Consequently, the distribution of reported cases across the areas was primarily influenced by the population size of each area, which also determined the distribution of the predicted cases.

## 5. Conclusions

The model presented aimed to correlate the virus concentration at the WWTP with the number of COVID-19 cases in the catchment area. This approach was based on previous models that either omitted biodegradation [14–17] or treated it as a uniform factor across the catchment area [3]. Our approach included a detailed consideration of the biodegradation along the sewer system, using an approach similar to that of McCall et al. [18] but with an additional step: associating the biodegradation factor with the population distribution and characteristics.

By applying the model to the data from the main WWTP of Bologna, the daily COVID-19 cases were estimated for each census section or statistical area served by the WWTP. To evaluate the performance of the model, the predicted cases were compared with those reported by the healthcare system. The results were first analyzed by comparing the

predicted and reported infection rates over time, and then by examining the variation in the different areas.

The model performed satisfactorily in predicting the trend of reported cases over time and space. The correlation between the predicted and reported cases was consistently high across the three waves, demonstrating the model's ability to predict the relevant waves and the significant increase in the number of cases.

The model slightly underestimated the number of reported cases in the first COVID-19 wave and overestimated them in subsequent waves. The analysis of several health variables, in particular the tests of time trends, suggested a greater underreporting of cases by the healthcare system in the second and third waves. This could partly explain the greater difference between the predicted and reported cases in these waves and demonstrates the potential use of the model as a valuable tool for early warning systems, providing a complementary approach to clinical surveillance. In addition, a time lag of 9 days was observed between the variations in the SARS-CoV-2 load at the WWTP and changes in the trend of reported active cases.

The inclusion of biodegradation in the model equation significantly improved the case estimates and allowed the calculation of the decline in virus load along the sewer network from each area to the WWTP. Conversely, the addition of the population-based coefficient did not significantly alter the final result.

The model estimated that biodegradation significantly reduced the virus load reaching the WWTP, resulting in a 30% reduction in the total virus load produced in the study area. The approach used in this study allows for the comparison of data from WWTPs in different cities by normalizing the detected values for virus biodegradation and the geographical distribution of the population.

Notably, potential changes to the parameters can be easily implemented to improve the accuracy of the model and can be obtained by collecting additional site-specific information and measurements (e.g., sewer flow velocity and biodegradation constant). In particular, an important future development could be the calculation of the parameter of the fecal shedding rate in situ and over time to obtain the shedding values specific to the demographic characteristics of the population and to the virus sub-variants to which the model is applied.

In addition, the model could be optimized by placing sampling and measurement points along the sewer network. This would improve the accuracy and reliability of the case estimates for the different areas and could help to identify the local differences in infection rates. Additionally, the model can be easily integrated into an online dashboard, enhancing the accessibility and immediacy for public health officials. Finally, the model could be adapted for the surveillance of other viruses or chemicals, making it a versatile tool for monitoring and preventing the impact of different viruses or substances on the population.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijerph21060741/s1>, Table S1: Resident and domiciled population in the study area; Table S2: Statistics of SARS-CoV-2 RNA concentration at the WWTP; Table S3: Statistics of wastewater flow rate at the WWTP; Table S4: Variables and estimates from the Global Poisson regression model; Table S5: Statistics of population characteristics; Table S6: Shedding rate values; Table S7: Variation of the spread of the Delta and Omicron variants and sub-variants; Table S8: Shedding values depending on the period; Table S9: Biodegradation rate constant values varying with the temperature and time of year; Table S10: Parameters in the error propagation formula; Table S11: Uncertainty associated with  $IR(t)$  varying with the number of cases; Table S12: For each grouped area: population domiciled in the area, number of reported and predicted cases, and percentage contribution of the area to the detected value of SARS-CoV-2 at the WWTP; Table S13: Contribution of hospitals in terms of viral load detected at the WWTP during the study period; Table S14: Statistics of the daily average reported and predicted infection rate and of the population-based coefficient; Figure S1: Map of areas and grouped areas included in the study; Figure S2: SARS-CoV-2 RNA concentration at the WWTP; Figure S3: Wastewater flow rate at the WWTP; Figure S4: Daily load of SARS-CoV-2 RNA at the inlet of the WWTP; Figure S5: Maps of the three indicators; Figure S6: Maps

of the statistical areas of Bologna, considering hospitalized COVID-19 individuals; Figure S7: Infection rate variation across grouped areas during the study period. Equation (S1–S7): Equations used to derive the model equation; Equation (S8): Error propagation formula; Equation (S9): Normalized local sensitivity analysis; Equation (S10–S12): Equations of the three indicators; Equation (S13): Hospitals coefficient.

**Author Contributions:** Conceptualization, M.F., L.V., G.B., A.C. and A.R.; methodology, M.F. and L.V.; software, M.F.; validation, M.F. and L.V.; formal analysis, M.F. and L.V.; investigation, M.F., L.V., F.A., L.D.L., F.F., E.M., M.G.M. and L.M.; resources, L.D.L., M.G.M., L.M., V.P., G.B., P.P., A.C. and A.R.; data curation, M.F., L.V., L.D.L., F.F., E.M. and L.M.; writing—original draft preparation, M.F., L.V., M.G.M. and A.C.; writing—review and editing, M.F., L.V., F.F., C.F., E.M., M.G.M., A.C. and A.R.; visualization, M.F. and L.V.; supervision, C.F., M.G.M., V.P., G.B., P.P., A.C. and A.R.; project administration, M.G.M., A.C. and A.R.; funding acquisition, G.B. and A.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded within the EU Project “EC G.A. NO. 060701/2021/864481/SUB/ENV.C2 -Support to the member states to establish national systems, local collection points, and digital infrastructure for monitoring covid 19 and its variants in waste waters”, and the Project “Epidemiologia delle acque reflue: implementazione del sistema di sorveglianza per l’identificazione precoce di agenti patogeni, con particolare riferimento al SARS-CoV2”, promoted by the Italian Ministry of Health (PROGRAMMA CCM 2020).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Databases and information system: <https://www.istat.it/en/analysis-and-products/databases> (accessed on 27 May 2024); Statistical areas of Bologna: <https://opendata.comune.bologna.it/explore/dataset/aree-statistiche/information/> (accessed on 27 May 2024); dext3r: <https://simc.arpae.it/dext3r> (accessed on 27 May 2024).

**Conflicts of Interest:** Laura De Lellis was employed by the company Hera Tech Srl. Laura Minelli was employed by the company Water Directorate, Hera SpA. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Corchis-Scott, R.; Geng, Q.; Seth, R.; Ray, R.; Beg, M.R.; Biswas, N.; Charron, L.; Drouillard, K.D.; D’Souza, R.; Heath, D.D.; et al. Averting an Outbreak of SARS-CoV-2 in a University Residence Hall through Wastewater Surveillance. *Microbiol. Spectr.* **2021**, *9*, e00792-21. [[CrossRef](#)] [[PubMed](#)]
2. Bertels, X.; Demeyer, P.; Van Den Bogaert, S.; Boogaerts, T.; Van Nuijs, A.L.N.; Delputte, P.; Lahousse, L. Factors influencing SARS-CoV-2 RNA concentrations in wastewater up to the sampling stage: A systematic review. *Sci. Total Environ.* **2022**, *820*, 153290. [[CrossRef](#)]
3. Li, X.; Zhang, S.; Shi, J.; Luby, S.P.; Jiang, G. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. *Chem. Eng. J.* **2021**, *415*, 129039. [[CrossRef](#)] [[PubMed](#)]
4. Weidhaas, J.; Aanderud, Z.T.; Roper, D.K.; VanDerslice, J.; Gaddis, E.; Ostermiller, J.; Hoffman, K.; Jamal, R.; Heck, P.E.; Zhang, Y.; et al. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Sci. Total Environ.* **2021**, *775*, 145790. [[CrossRef](#)] [[PubMed](#)]
5. Robotto, A.; Lembo, D.; Quaglino, P.; Brizio, E.; Polato, D.; Civra, A.; Cusato, J.; Di Perri, G. Wastewater-based SARS-CoV-2 environmental monitoring for Piedmont, Italy. *Environ. Res.* **2022**, *203*, 111901. [[CrossRef](#)]
6. Baldovin, T.; Amoroso, I.; Fonzo, M.; Bertonecello, C.; Groppi, V.; Pitter, G.; Russo, F.; Baldo, V. Trends in SARS-CoV-2 clinically confirmed cases and viral load in wastewater: A critical alignment for Padua city (NE Italy). *Heliyon* **2023**, *9*, e20571. [[CrossRef](#)]
7. McMahan, C.S.; Self, S.; Rennert, L.; Kalbaugh, C.A.; Kriebel, P.D.; Graves, D.; Colby, C.; Deaver, B.J.A.; Popat, S.C.; Karanfil, P.T.; et al. COVID-19 wastewater epidemiology: A model to estimate infected populations. *Lancet Planet. Health* **2021**, *5*, e874–e881. [[CrossRef](#)]
8. La Rosa, G.; Iaconelli, M.; Veneri, C.; Mancini, P.; Ferraro, G.B.; Brandtner, D.; Lucentini, L.; Bonadonna, L.; Rossi, M.; Grigioni, M.; et al. The rapid spread of SARS-COV-2 Omicron variant in Italy reflected early through wastewater surveillance. *Sci. Total Environ.* **2022**, *837*, 155767. [[CrossRef](#)]
9. EUR-LEX—32021H0472—EN—EUR-LEX. Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021H0472&qid=1628798981209> (accessed on 27 May 2024).



10. Ciannella, S.; González-Fernández, C.; Gómez-Pastora, J. Recent progress on wastewater-based epidemiology for COVID-19 surveillance: A systematic review of analytical procedures and epidemiological modeling. *Sci. Total Environ.* **2023**, *878*, 162953. [CrossRef]
11. Rauch, W.; Schenk, H.; Insam, H.; Markt, R.; Kreuzinger, N. Data modelling recipes for SARS-CoV-2 wastewater-based epidemiology. *Environ. Res.* **2022**, *214*, 113809. [CrossRef]
12. Morvan, M.; Lo Jacomo, A.; Souque, C.; Wade, M.J.; Hoffmann, T.; Pouwels, K.B.; Lilley, C.; Singer, A.C.; Porter, J.; Evens, N.; et al. An analysis of 45 large-scale wastewater sites in England to estimate SARS-CoV-2 community prevalence. *Nat. Commun.* **2022**, *13*, 4313. [CrossRef]
13. Jiang, G.; Wu, J.; Weidhaas, J.; Li, X.; Chen, Y.; Mueller, J.F.; Li, J.; Kumar, M.; Zhou, X.; Arora, S.; et al. Artificial neural network-based estimation of COVID-19 case numbers and effective reproduction rate using wastewater-based epidemiology. *Water Res.* **2022**, *218*, 118451. [CrossRef]
14. Ahmed, W.; Angel, N.; Edson, J.; Bibby, K.; Bivins, A.; O'Brien, J.; Choi, P.M.; Kitajima, M.; Simpson, S.L.; Li, J.; et al. First confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci. Total Environ.* **2020**, *728*, 138764. [CrossRef] [PubMed]
15. Prasek, S.M.; Pepper, I.L.; Innes, G.K.; Slinski, S.; Ruedas, M.; Sánchez, A.C.; Brierley, P.; Betancourt, W.Q.; Stark, E.R.; Foster, A.R.; et al. Population level SARS-CoV-2 fecal shedding rates determined via wastewater-based epidemiology. *Sci. Total Environ.* **2022**, *838*, 156535. [CrossRef] [PubMed]
16. Schmitz, B.W.; Innes, G.K.; Prasek, S.M.; Betancourt, W.Q.; Stark, E.R.; Foster, A.R.; Abraham, A.G.; Gerba, C.P.; Pepper, I.L. Enumerating asymptomatic COVID-19 cases and estimating SARS-CoV-2 fecal shedding rates via wastewater-based epidemiology. *Sci. Total Environ.* **2021**, *801*, 149794. [CrossRef] [PubMed]
17. Wani, H.; Menon, S.; Desai, D.; Dsouza, N.; Bhatena, Z.; Desai, N.B.; Rose, J.B.; Shrivastava, S. Wastewater-Based Epidemiology of SARS-CoV-2: Assessing Prevalence and Correlation with Clinical Cases. *Food Environ. Virol.* **2023**, *15*, 131–143. [CrossRef]
18. McCall, C.; Fang, Z.; Li, D.; Czubai, A.J.; Juan, A.; LaTurner, Z.W.; Ensor, K.B.; Hopkins, L.; Bedient, P.B.; Stadler, L.B. Modeling SARS-CoV-2 RNA degradation in small and large sewersheds. *Environ. Sci.* **2022**, *8*, 290–300. [CrossRef]
19. Databases and Information System. Available online: <https://www.istat.it/en/analysis-and-products/databases> (accessed on 27 May 2024).
20. Statistical Areas of Bologna. Available online: <https://opendata.comune.bologna.it/explore/dataset/aree-statistiche/information/> (accessed on 27 May 2024).
21. La Rosa, G.; Bonadonna, L.; Suffredini, E. Protocollo della Sorveglianza di SARS-CoV-2 in reflui urbani (SARI)-rev. 3. *Zenodo* **2022**, *1*, 1–26. [CrossRef]
22. Wu, F.; Zhang, J.; Xiao, A.; Gu, X.; Lee, W.L.; Armas, F.; Kauffman, K.M.; Hanage, W.P.; Matus, M.; Ghaeli, N.; et al. SARS-CoV-2 Titers in Wastewater Are Higher than Expected from Clinically Confirmed Cases. *MSystems* **2020**, *5*, 10–1128. [CrossRef]
23. Ahmed, W.; Bertsch, P.M.; Bibby, K.; Haramoto, E.; Hewitt, J.; Huygens, F.; Gyawali, P.; Korajkic, A.; Riddell, S.; Sherchan, S.P.; et al. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environ. Res.* **2020**, *191*, 110092. [CrossRef]
24. La Rosa, G.; Mancini, P.; Ferraro, G.B.; Veneri, C.; Iaconelli, M.; Bonadonna, L.; Lucentini, L.; Suffredini, E. SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Sci. Total Environ.* **2021**, *750*, 141711. [CrossRef] [PubMed]
25. dext3r. Available online: <https://simc.arpae.it/dext3r> (accessed on 27 May 2024).
26. Cleveland, W.S.; Devlin, S.J. Locally Weighted Regression: An approach to regression analysis by local fitting. *J. Am. Stat. Assoc.* **1988**, *83*, 596–610. [CrossRef]
27. Cappellari, M.; McDermid, R.M.; Alatalo, K.; Blitz, L.; Bois, M.; Bournaud, F.; Bureau, M.; Crocker, A.F.; Davies, R.L.; Davis, T.A.; et al. The ATLAS3D project—XX. Mass–size and mass– $\sigma$  distributions of early-type galaxies: Bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *Mon. Not. R. Astron. Soc.* **2013**, *432*, 1862–1893. [CrossRef]
28. Benefield, A.E.; Skrip, L.A.; Clement, A.; Althouse, R.A.; Chang, S.; Althouse, B.M. SARS-CoV-2 viral load peaks prior to symptom onset: A systematic review and individual-pooled analysis of coronavirus viral load from 66 studies. *MedRxiv (Cold Spring Harb. Lab.)*, 2020, *1*, 1–26. [CrossRef]
29. Foladori, P.; Cutrupi, F.; Segata, N.; Manara, S.; Pinto, F.; Malpei, F.; Bruni, L.; La Rosa, G. SARS-CoV-2 from faeces to wastewater treatment: What do we know? A review. *Sci. Total Environ.* **2020**, *743*, 140444. [CrossRef] [PubMed]
30. Zeleke, A.J.; Miglio, R.; Palumbo, P.; Malaguti, E.; Chiari, L.; Due, U. Spatiotemporal heterogeneity of SARS-CoV-2 diffusion at the city level using geographically weighted Poisson regression model: The case of Bologna, Italy. *Geospat. Health* **2022**, *17*. [CrossRef] [PubMed]
31. Prasek, S.M.; Pepper, I.L.; Innes, G.K.; Slinski, S.; Betancourt, W.Q.; Foster, A.R.; Yaglom, H.; Porter, W.T.; Engelthaler, D.M.; Schmitz, B.W. Variant-specific SARS-CoV-2 shedding rates in wastewater. *Sci. Total Environ.* **2023**, *857*, 159165. [CrossRef] [PubMed]
32. Rose, C.; Parker, A.; Jefferson, B.; Cartmell, E. The Characterization of feces and Urine: A review of the literature to inform advanced Treatment technology. *Crit. Rev. Environ. Sci. Technol.* **2015**, *45*, 1827–1879. [CrossRef]

33. Bivins, A.; Greaves, J.; Fischer, R.J.; Yinda, K.C.; Ahmed, W.; Kitajima, M.; Munster, V.J.; Bibby, K. Persistence of SARS-CoV-2 in water and wastewater. *Environ. Sci. Technol. Lett.* **2020**, *7*, 937–942. [[CrossRef](#)] [[PubMed](#)]
34. Yang, S.; Dong, Q.; Li, S.; Zhao, C.; Kang, X.; Ren, D.; Xu, C.; Zhou, X.H.; Peng, L.; Sun, L.; et al. Persistence of SARS-CoV-2 RNA in wastewater after the end of the COVID-19 epidemics. *J. Hazard. Mater.* **2022**, *429*, 128358. [[CrossRef](#)]
35. Silverman, A.I.; Boehm, A.B. Systematic Review and Meta-Analysis of the persistence and disinfection of human coronaviruses and their viral surrogates in water and wastewater. *Environ. Sci. Technol. Lett.* **2020**, *7*, 544–553. [[CrossRef](#)]
36. Cipolla, S.S.; Maglionico, M. Heat Recovery from Urban Wastewater: Analysis of the Variability of Flow Rate and Temperature in the Sewer of Bologna, Italy. *Energy Procedia* **2014**, *45*, 288–297. [[CrossRef](#)]
37. Cormen, T.H.; Leiserson, C.E.; Rivest, R.L.; Stein, C. *Introduction to Algorithms*; MIT Press: Cambridge, MA, USA, 2001.
38. Compagni, R.D.; Polesel, F.; Von Borries, K.J.F.; Zhang, Z.; Turolla, A.; Antonelli, M.; Vezzaro, L. Modelling micropollutant fate in sewer systems—A new systematic approach to support conceptual model construction based on in-sewer hydraulic retention time. *J. Environ. Manag.* **2019**, *246*, 141–149. [[CrossRef](#)] [[PubMed](#)]
39. Bareš, V.; Jiráček, J.; Pollert, J. Spatial and temporal variation of turbulence characteristics in combined sewer flow. *Flow Meas. Instrum.* **2008**, *19*, 145–154. [[CrossRef](#)]
40. Zhang, Y.; Cen, M.; Hu, M.; Du, L.; Hu, W.; Kim, J.J.; Dai, N. Prevalence and persistent shedding of fecal SARS-CoV-2 RNA in patients with COVID-19 infection: A systematic review and meta-analysis. *Clin. Transl. Gastroenterol.* **2021**, *12*, e00343. [[CrossRef](#)] [[PubMed](#)]
41. Natarajan, A.; Zlitni, S.; Brooks, E.F.; Vance, S.E.; Dahlen, A.; Hedlin, H.; Park, R.; Han, A.X.; Schmidtke, D.T.; Verma, R.; et al. Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA suggest prolonged gastrointestinal infection. *Med* **2022**, *3*, 371–387.e9. [[CrossRef](#)] [[PubMed](#)]
42. Cheung, K.; Hung, I.; Chan, P.P.Y.; Lung, K.-C.; Tso, E.Y.; Liu, R.; Ng, Y.Y.; Chu, M.Y.; Chung, T.W.-H.; Tam, A.R.; et al. Gastrointestinal manifestations of SARS-CoV-2 infection and virus load in fecal samples from a Hong Kong cohort: Systematic review and meta-analysis. *Gastroenterology* **2020**, *159*, 81–95. [[CrossRef](#)]
43. Cavany, S.; Bivins, A.; Wu, Z.; North, D.; Bibby, K.; Perkins, A. Inferring SARS-CoV-2 RNA shedding into wastewater relative to the time of infection. *Epidemiol. Infect.* **2022**, *150*, e21. [[CrossRef](#)]
44. Hart, O.E.; Halden, R.U. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Sci. Total Environ.* **2020**, *730*, 138875. [[CrossRef](#)]
45. Petala, M.; Dafou, D.; Kostoglou, M.; Karapantsios, T.; Kanata, E.; Chatziefstathiou, A.; Sakaveli, F.; Kotoulas, K.; Arsenakis, M.; Roilides, E.; et al. A physicochemical model for rationalizing SARS-CoV-2 concentration in sewage. Case study: The city of Thessaloniki in Greece. *Sci. Total Environ.* **2021**, *755*, 142855. [[CrossRef](#)] [[PubMed](#)]
46. Li, J.; Ahmed, W.; Metcalfe, S.; Smith, W.; Choi, P.M.; Jackson, G.; Cen, X.; Zheng, M.; Simpson, S.; Thomas, K.V.; et al. Impact of sewer biofilms on fate of SARS-CoV-2 RNA and wastewater surveillance. *Nat. Water* **2023**, *1*, 272–280. [[CrossRef](#)]
47. Verani, M.; Federigi, I.; Muzio, S.; Lauretani, G.; Calà, P.; Mancuso, F.; Salvadori, R.; Valentini, C.; La Rosa, G.; Suffredini, E.; et al. Calibration of Methods for SARS-CoV-2 Environmental Surveillance: A Case Study from Northwest Tuscany. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16588. [[CrossRef](#)] [[PubMed](#)]
48. EpiCentro. Sorveglianza Genomica del Virus SARS-CoV-2 e delle sue Varianti di Interesse in Sanità Pubblica in Italia. Available online: <https://www.epicentro.iss.it/coronavirus/sars-cov-2-monitoraggio-varianti> (accessed on 27 May 2024).
49. Rector, A.; Bloemen, M.; Thijssen, M.; Delang, L.; Raymenants, J.; Thibaut, J.; Pussig, B.; Fondu, L.; Aertgeerts, B.; Van Ranst, M.; et al. Monitoring of SARS-CoV-2 concentration and circulation of variants of concern in wastewater of Leuven, Belgium. *J. Med. Virol.* **2023**, *95*, e28587. [[CrossRef](#)] [[PubMed](#)]
50. Ho, J.S.-Y.; Fernando, D.; Chan, M.Y.; Sia, C. Obesity in COVID-19: A Systematic review and Meta-analysis. *Ann. Acad. Med. Singap.* **2020**, *49*, 996–1008. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.