Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Associations between SARS-CoV-2 RNA concentrations in wastewater and COVID-19 rates in days after sampling in small urban areas of Seville: A time series study

Francisco Javier Rodríguez Rasero, Luis A. Moya Ruano, Pablo Rasero Del Real, Lucila Cuberos Gómez, Nicola Lorusso

HIGHLIGHTS

• Wastewater (WW) surveillance system for SARS-CoV-2 in a city of 700,000 inhabitants
• Data for small urban areas (average population of 20,000) collected during six months.
• COVID-19 cases were geolocated within urban areas.
• Significant associations SARS-CoV-2 in WW and COVID-19 rates on subsequent days

ABSTRACT

Wastewater surveillance systems for SARS-CoV-2 can be used to support public health decisions, complementary to clinical surveillance. We examined the lead-lag associations between SARS-CoV-2 RNA copies in wastewater and COVID-19 rates in relatively small urban areas of Seville, adjusting for internal mobility, temperature, and wastewater-related variables. The association COVID-19 rates-RNA copies were statistically significant from three to 27 days after sampling. Temperature is a confounding factor for both viral RNA counts and mobility. The model that best fitted data used cases six days after sampling. A logarithmic unit increase in viral RNA count in wastewater was associated with a 26.9% increase in COVID-19 rate per 100,000 inhabitants (95% CI: 13.1-42.4%), within the urban area, six days later. Surveillance system for SARS-CoV-2 in wastewater has great potential for public health. Knowing the specific association between SARS-CoV-2 RNA copies in wastewater and COVID-19 daily rates may help to improve its performance.

© 2021 Elsevier B.V. All rights reserved.

Keywords:
SARS-CoV-2 RNA
COVID-19
Wastewater epidemiology
Urban areas
Mobility
Vulnerable population

1. Introduction

Information on spatial and temporal trends of SARS-CoV-2 RNA in wastewater can be used to support public health decisions and manage
the response to the pandemic, complementary to clinical surveillance of COVID-19 (WHO Regional Office for Europe, 2021).

A wastewater surveillance system for SARS-CoV-2, that is, the systematic collection, analysis, and interpretation of data on SARS-CoV-2 RNA in wastewater, can be used in different phases of the COVID-19 pandemics. In pandemic phase, it may be useful for monitoring circulation of SARS-CoV-2 in population, checking effects of non-pharmaceutical interventions (NPIs), and monitoring situation in specific locations. In addition, the surveillance system can also be used to detect the appearance of new mutations and variants of SARS-CoV-2 (Medema et al., 2020a; WHO Regional Office for Europe, 2020).

Several settings have already deployed wastewater surveillance programs. In Europe, the European Commission launched in March 2021 a Recommendation to support Member States in establishing wastewater surveillance systems across the Union (European Commission, 2021). In Andalusia (Spain), the Regional Ministry of Health and Families established the Andalusian wastewater surveillance network (RAVAR) in July 2020 (Government of Andalusia, 2020).

The RAVAR aims at monitoring SARS-CoV-2 not only in wastewater treatment plants (WWTP), but in sewerage networks as well. A provincial steering committee, made up by the health authority and the water authority, must select the sampling sites based on common criteria, including location of deprived areas. The deployment of RAVAR started in Seville, with 691,395 inhabitants in 2020 (National Statistics Institute of Spain, 2021), where the public company responsible for wastewater sanitation and treatment (EMASESA) had begun to monitor SARS-CoV-2 in WWTP and sewerage network even before, in May 2020.

When monitoring circulation of SARS-CoV-2, finding an association between environmental SARS-CoV-2 data and COVID-19 reported patient number some days later can be a key element to improve the usefulness of a wastewater surveillance system, since it would allow us to estimate changes in COVID-19 prevalence some days ahead of positive test results. Due to differences in sewerage networks, sampling strategies and analysis conditions, inter alia, this association will be specific for each location.

Several studies have assessed the presence and trends of gene copies of SARS-CoV-2 in wastewater (Ahmed et al., 2020a; Baldwin et al., 2021; Gonzalez et al., 2020; Haramoto et al., 2020; Kocamemi et al., 2020; Kumar et al., 2020; La Rosa et al., 2020; Randazzo et al., 2020; Westhaus et al., 2021). Wastewater surveillance results are often interpreted only in a qualitative way, without relating viral counts in wastewater to COVID-19 rates. A semi-quantitative approach, capable of indicating relative levels of infection at local level, requires more research to develop, calibrate, and validate than this qualitative approach (Daughton, 2020). Thus, some studies have verified the relationship between SARS-CoV-2 RNA in wastewater and COVID-19 health-related outcomes, with samples taken at inlets to WWTP or primary sludge (Carceryen et al., 2021; Greenwald et al., 2021; Medema et al., 2020b; Nemudryi, 2020; Peccia et al., 2020; Tomasino et al., 2021; Weidhaas et al., 2021; Westhaus et al., 2021).

This approach may provide an estimation of changes in COVID-19 health-related outcomes, but only at level of the population served by the WWTP (i.e., at city or supra-city level). However, to our knowledge, quantitative associations between SARS-CoV-2 RNA copies in wastewater and COVID-19 rates in urban areas have not been established yet. A possible reason is that models may suffer greatly from the limited understanding and variability of some parameters, as well as the potential data noise (Kitajima et al., 2020; Zhu et al., 2021).

In this study, we aimed to create a reproducible methodology to investigate the association between SARS-CoV-2 RNA copies in wastewater and COVID-19 rates in small urban areas from seven days before sampling to 28 days after sampling. To this end, we carried out a pilot study using data from urban areas of Seville, from July 19, 2020, to February 11, 2021.

2. Material and methods

2.1. Selection of urban areas

The urban areas in Seville were shaped by the domestic wastewater basin upstream of a sampling point. The wastewaters collected at this point must be domestic sewage, without contributions (or meaningful influence) from hospitals, industries, and major transport infrastructures. Water Authority has georeferenced all the sewage network, including sewer pipes, sampling points and its associated basins. Data on population living within urban areas were retrieved from the Basic Spatial Data of Andalusia (Statistical and Carthography Institute of Andalusia, 2021a).

We delimited 28 urban areas in Seville that fulfilled the selection criteria. We selected eight of them (Table 1), covering 160,335 inhabitants (23.2% of the population), based on three factors: distribution within the city (urban areas should cover the largest number of municipal districts), population (urban areas should include as large population as possible), and inequities (some of the urban areas should include deprived areas). Fig. 1 shows the distribution of urban areas.

In October, we stopped virus determination to carry out an intermediate analysis of the results and optimize the sample management protocol, reducing delivery times and the limit of quantification. Consequently, we reduced the communication time of results from laboratory from 72 to 36 h after sampling, approximately.

2.2. Data collection

2.2.1. Viral counts

We used the number of SARS-CoV-2 gene copies per litre of wastewater (gc/L), as a proxy for the virus circulation in each urban area. We collected grab samples weekly at the peak hours of toilet flushing. We estimated these peak hours through water consumption data and treatment plant input flows, which showed that water usage peak hour consistently ranges between 9 and 11 am. At each sampling point, the sample has been taken at approximately the same time.

We performed the analysis of SARS-CoV-2 in wastewater following the Protocol published by Ministry for the Ecological Transition and Demographic Challenge of Spain (CSIC, UB, 2021).

The concentration method consisted of an aluminium chloride precipitation. 200 mL sample were transferred to 250 mL NUNC centrifuge bottles and inoculated with Mengovirus (strain VMC0 CECT 100000). The pH was measured and adjusted to 6.0 by adding HCl (2 N). Afterwards, 2 mL of an AlCl$_3$ solution (4%) was added and the pH was readjusted to 6.0 by adding NaOH (2 N). Then, the sample was mixed using an orbital shaker at 150 rpm for 15 min at room temperature. After that, the sample was centrifuged at 1700 × g for 20 min in a Thermo Scientific Heraeus Multifuge X3 centrifuge with Piramoon 6x250LE FiberLite Rotor. The supernatant was discarded, and the pellet was resuspended with 10 mL of beef extract (4%) using a vortex, shaken at 200 rpm for 20 min at room temperature and centrifuged at 1900g for 30 min. Subsequently, the supernatant was discarded, and the pellet was resuspended in 1 mL of PBS buffer. Samples were immediately extracted or stored at -20 °C freezer.

| Urban area | Population | Extent (ha) | Presence of deprived areas |
|------------|------------|-------------|---------------------------|
| S03        | 21,602     | 74.3        | No                        |
| S04        | 15,521     | 44.8        | No                        |
| S07        | 21,380     | 81.8        | Yes                       |
| S10        | 11,673     | 57.3        | Yes                       |
| S14        | 11,438     | 50.0        | No                        |
| S16        | 49,124     | 293.9       | Yes                       |
| S17        | 18,678     | 110.9       | Yes                       |
| S19        | 10,719     | 90.7        | No                        |

Table 1

Selected urban areas.
The viral RNA extraction was performed using the NucleoSpin RNA virus kit (Macherey-Nagel) in samples from July to October, and QIAamp Viral RNA (Qiagen) in samples from November to January, according to the manufacturer instructions. Previous laboratory tests showed that there were no statistically significant differences between the results of the two kits (data not shown.)

We performed the viral RNA detection using Applied Biosystems StepOnePlus Real-Time PCR System. The reaction was performed following the Protocol, that involves assays that target the IP4 gene (Institut Pasteur, 2020), N genes (Centers for Disease Control and Prevention, 2020) and the E gene (Corman et al., 2020) of SARS-CoV-2 and MgV as recovery process control (Foladori et al., 2020). Samples with >1% recovery were considered as valid.

Two wells of direct RNA and two wells of the 10-fold diluted RNA were analysed for each sample due to the possible presence of PCR inhibitors, every RT-qPCR assay included negative (nuclease-free water) and positive controls for each gene. The positive controls were used directly from the commercial material (genes N1 and N2: IDT 10006625; IP4: abm G633-16; gene E: IDT 10006896).

The estimation from Ct to gene copies per litre was performed using the following formula (no normalization was applied to data):

\[
\text{Number of gc/L} = \frac{10^{b+2} \times V_e \times D \times V_c \times 10^7}{V_P \times B_0 \times V_o}
\]

where \( b \) is the interception in calibration curve, \( m \) the slope in calibration curve, \( V_e \) the volume in mL of eluted RNA, \( V_P \) the volume in mL of samples used in the PCR (usually 5 mL), \( D \) the dilution in the PCR (1 with no dilution, and 10 if dilution was 1/10), \( V_c \) the volume in mL of the sample after concentration, \( B_0 \) the initial biomass in mL used in the extraction and \( V_o \) the initial volume in mL of the sample.

Wastewater network collects all sewage from basins. Thus, at the sampling point, there will be viral gene copies from all infected people that have excreted virus in this area (faeces, urine, and respiratory secretions), including asymptomatic carriers (Chen et al., 2020; Hong et al., 2021). Therefore, the total of infected people will be:

\[
\text{Infected people} = (AC)_r + (PA)_r + (PA)_nr
\]

where \((AC)_r\) is the number of active cases residents in the urban area, \((PA)_r\) the number of presymptomatic and asymptomatic residents in the area and \((PA)_nr\) the number of presymptomatic and asymptomatic non-residents in the area, such as people working in the area and residents (temporary or not) non-registered by the Andalusian Healthcare Service database.

2.2.2. Internal mobility

The number of close contacts per person has been related with reported contacts of cases and SARS-CoV-2 transmission (Badr et al., 2020; lacus et al., 2020; Ingelbeen et al., 2020). In COVID-19 transmission dynamics, internal mobility is more important than mobility across provinces (lacus et al., 2020). On the other hand, an increase or decrease in the internal mobility could also indicate a variation of \((PA)_r\) and \((PA)_nr\) terms, thus having an influence on viral counts. Therefore, internal mobility could be a confounding variable. Eventually, mobility data may also be influenced by summer decreases of population and by non-pharmaceutical interventions (NPIs) on Covid-19 incidence.

We used the internal mobility of Seville as an indicator of number of contacts in the population living in the urban area. Internal mobility data were retrieved from DataCOVID study, that provides daily relative mobility data, expressed as percentage of mobility versus a reference week (14-20 February 2020), and derived from aggregated and anonymised mobile phone data (State Secretariat for Digitalisation and Artificial Intelligence of Spain, 2021). We aimed at controlling mobility trends, so we used the 7-day average relative internal mobility.

Fig. 1. Mapping of selected urban areas of Seville (1:30,000 scale). Note: S19 is out of scale.
2.2.3. Ambient temperature

Temperature could be related to COVID-19 incidence, virus count in wastewater (e.g., through a dilution effect with high water consumption) and mobility (modifying behaviour patterns of the population). As we aimed at controlling trends, we used 7-day average of daily average temperature ($T_{\text{day}}$), estimated as $T_{\text{day}} = (T_{\text{max}} + T_{\text{min}})/2$, where $T_{\text{max}}$ is the daily maximum temperature and $T_{\text{min}}$ the daily minimum temperature, both obtained from the Meteorological Agency of Spain (AEMET, 2021).

2.2.4. Vulnerable people

We also investigated whether the presence of vulnerable people in an urban area was related with COVID-19 rates. To this end, we included a dummy variable indicating the presence or absence of people living in socioeconomic deprived areas, within the urban area. Locations of deprived areas were collected from the Andalusian Strategy on Social Inclusion (Regional Ministry for Equity and Social Policies of Andalusia, 2018).

2.2.5. Other variables

Finally, we considered variables related with wastewater sampling: pH, temperature (Tw), chemical oxygen demand (COD), conductivity, precipitations (sum of precipitations in sampling day and in day before sampling), day of the week, urban area, and sample group. The use of biocides such as bleach could modify pH, which could have an influence on RNA degradation. Sample group is a dummy variable with two categories: first group (samples from July to October) and second group (from November to January), because analysis conditions were different in these groups. In particular, samples from the first group were kept refrigerated longer than those from the second group, which were sent to the laboratory earlier in order to reduce the time for reporting results from the laboratory (see 2.1. Selection of urban areas). As mentioned above, the RNA extraction kit used was different (see 2.2.1. Viral counts). We also investigated the possible interaction between this variable and viral RNA counts.

2.3. Estimation of COVID-19 active cases, by urban area

We defined “active case” as case diagnosed with an Active Infection Diagnostic Test (AIDT) to detect the presence of active infection by SARS-CoV-2 and notified to Andalusian Epidemiological Surveillance System (SVEA).

To estimate the number active cases per urban area, we must identify the entry and exit data of every active case within SVEA:

- **Entry**: day on which a patient declared to have symptoms compatible with Covid-19 (registered in SVEA) or, in the absence of this information, day on which patient tested positive for the AIDT.
- **Exit**: entry day plus ten days, hospital admission date or death day (whichever occurs first, according to the SVEA criteria).

We estimated the number of active cases in a day $d$ as the sum of active cases that meet the following condition:

**Entry date ≤ day d ≤ Exit date**

The SVEA network collects all cases confirmed by AIDT from public and private laboratories. Active cases are geolocated by using personal address data from Andalusian Healthcare Service database.

For case mapping, we used a Geographical Information System, QGIS (QGIS Development Team, 2021), and georeferenced data to assign each case to an urban area. We carried out the data cleaning process with nordir and geodir (Statistical and Cartography Institute of Andalusia, 2021b) to normalise address and compare them to the Andalusia Digital Street Directory, respectively.

2.4. Statistical analysis

We performed a regression analysis using quasi-Poisson models, a design widely used in environmental epidemiology for modelling overdispersed counts (Bhaskaran et al., 2013). Quasi-Poisson models keep the Poisson variance function $V(\mu) = \mu$, but allow a general positive dispersion parameter $\phi$ (Dunn and Smyth, 2018). In a simplified way:

$$y \sim \text{Pois} (\mu_x), \quad \text{var}(y_i) = \phi \mu_i$$

$$\ln \left( \lambda_{u,d,i} \right) = \beta_0 + \beta_1 \log \, v_{u,d} + \beta_2 \, T_{d} + \beta_3 \, m_d + \beta_4 \, p_u + \ln \, P_u + \sum_t \beta_t \, n_{u,d,t}$$

where $\ln$ denotes natural logarithm and $log$ the base-10 logarithm, $d$ is the day of wastewater sampling, $\nu$ is the number of days backward or forward, in relation with the sampling day, that the model considers, and:

- $\lambda_{u,d,i}$ is the active case count in the urban area $u$ in a day $d + t$.
- $v_{u,d} \, is \, the \, viral \, RNA \, count \, in \, the \, urban \, area \, u \, in \, a \, day \, d, \text{in} \, \text{gc/L}.$
- $T_d$ is the environmental temperature in Seville in a day $d$, in °C.
- $m_d$ is the internal mobility in Seville in a day $d$, in percentage versus the mobility during the reference week (14-20 February 2020).
- $p_u$ is a dummy variable indicating whether there is a deprived area within the urban area $u$ or not.
- $n_{u,d,t}$ is a variable related with wastewater sampling in an urban area $u$ and a day $d$.
- $\ln \, P_u$ is an offset term (total population in urban area $u/100,000$).

By including the offset term, we modelled rates of active cases per 100,000 inhabitants, rather than active cases. The coefficient of interest $\beta_1$ measures the change in COVID-19 active cases (rate per 100,000 inhabitants) by the increase in a logarithmic unit in viral count. Results are expressed in terms of percentage of change in the COVID-19 rate per 100,000 inhabitants, estimated as $[\exp(\beta_1) - 1] \times 100$.

Data included in the Poisson model must fulfill the condition $v_{u,d} \geq \text{LoQ}$, where LoQ is the limit of quantification of viral RNA counts ($\text{LoQ} = 2.5 \times 10^4$ and $1.6 \times 10^6 \, \text{gc/L}$ for sample group 1 and 2, respectively). LoQ was calculated applying three times the standard deviation of the limit of detection (LoD). LoD was defined as the lowest amount of analyte detectable in 50% of cases (50% positives of the replicates at that specific concentration).

We fitted models using active case counts within the urban area from seven days before wastewater sampling ($t = -7$) to 28 days after wastewater sampling ($t = 28$). We will refer to these models as Mt, being Mt. the model using active cases from the day $t$.

To assess the inclusion of the rest of variables, we followed the GCV (Generalized Cross Validation) criterion. According to the results of preliminary analysis, we assumed a linear relationship between active cases and independent variables except for temperature, which was included using a natural cubic spline with four degrees of freedom.

In the sensitivity analysis, we checked a different exit period from the SVEA (entry day plus fourteen days, instead of ten days). We also considered all available data, substituting $v_{u,d}$ if $v_{u,d} < \text{LoQ}$ by zero. Likewise, we included viral counts through a standardized term (virus load per day), mobility using natural cubic splines, and temperature with threshold variables. The standardized term (in gene copies per person and day) was estimated as follows:

$$\text{standardized viral counts} = \log \left( \frac{\text{flow}}{\text{population}} \times \frac{\text{flow rate}}{0.1^3} \right)$$

where log is the base-10 logarithm, and flow is the sewage flow rate (m³/day).

We considered the association as statistically significant if the 95% confidence interval of the percentage of change in the COVID-19 rate
per 100,000 inhabitants does not include the null value (0%). In the cases of categorical variables (sample group and vulnerable people), we considered differences between groups as statistically significant if \( p < 0.05 \). We performed all statistical analysis using the R statistical software (R Core Team, 2021).

3. Results and discussion

3.1. Wastewater analysis

From July 19, 2020, to January 21, we took 199 samples in these areas. In wastewater samples, we also measured pH, Tw, chemical oxygen demand, and conductivity. Sampling works can be seen at https://www.youtube.com/watch?v=0KF7qBVa10I.

3.2. Estimation of COVID-19 active cases, by urban area

We selected active cases from 12 of July 2020 to 11 February 2021 (covering period of wastewater data collection plus lag and lead days). During this period, there were 37,022 total active cases in Seville (5355 cases per 100,000 people). We carried out the case mapping (Fig. 2) to sort COVID-19 cases by selected urban areas. There were 8210 cases within these areas during the study period.

3.3. Data transformation and time series

After removing data with viral RNA counts below the limit of quantification, we eventually used 164 observations. We created a tidy and join dataset containing a time series with daily active cases per urban area with lag and leads (-7 to 28 days of difference with respect to sampling), viral counts per sampling day and urban area, daily temperature, and internal mobility, as well as the rest of variables.

Fig. 3 shows the time series of daily active cases, viral RNA counts, average temperature, and mobility. The overall trend of daily cases (in natural logarithmic scale) and viral RNA counts (in base-10 logarithmic scale) are reasonably similar, as suggested in previous studies (Greenwald et al., 2021; Medema et al., 2020b; Nemudryi, 2020; Tomasino et al., 2021; Weidhaas et al., 2021; Westhaus et al., 2021). The average temperature decreases steadily throughout the study period, while the mobility (expressed as a percentage of mobility compared to the reference week, 14-20 February 2020) increases until November, when the COVID-19 incidence led to implement measures to limit the mobility.

Fig. 4 shows the scatter plots of the active cases (at different days after sampling) and the main variables (viral RNA counts, average temperature, and mobility). The relationship between the number of cases and viral RNA counts, as well as between cases and mobility, is roughly linear. Temperature seems to follow a non-linear relationship with the number of active cases.

3.4. Associations between viral RNA concentrations and COVID-19 rates

The association between active case rates and RNA viral counts was statistically significant for rates from three to 27 days after sampling. These results are consistent with the incubation period of COVID-19 (time between exposure to the virus and the development of the symptoms), which ranges from two to 14 days (Lauer et al., 2020; Zaki and Mohamed, 2021), and the duration of virus shedding in faeces, estimated to be 26.0 days (95% CI: 21.7–34.9) from symptom onset (Miura et al., 2021).

The model that best fitted data (lowest GCV value) used, as a response variable, cases six days after sampling. In this model (M6), a logarithmic unit increase in viral RNA counts in wastewater is associated with 26.2% (95% CI: 12.7%, 41.3%) increase in COVID-19 rate per
100,000 inhabitants, within the urban area six days later, with the other variables remaining constant. The highest association occurs at 11 days after sampling, with a 36.7% (95% CI: 12.7%, 41.3%) increase in COVID-19 rate per 100,000 inhabitants per a logarithmic unit increase in viral RNA counts (Fig. 5).

We also investigated potential differences between the two groups of samples of the study. Theoretically, different conditions within analysis could have an influence on the results, basically because of the different extent of the degradation of RNA in both groups. Results showed that there were some differences between both groups, but they were not significant (e.g., \( p = 0.257 \) in the model M6). We also investigated the interaction between virus concentration and the sample group. The interaction term was not significant either (e.g., \( p = 0.499 \) in the model M6). Therefore, the reduction in the time of analysis did not affect the results, and the decay of SARS-CoV-2 RNA in wastewater is therefore not relevant, consistently with previous studies (Ahmed et al., 2020b; Chin et al., 2020; Wu et al., 2021). Instead, with this reduction, we were able to record the results of each sample within 36 h of collection, within the maximum 48-hour timeframe recommended for early warning surveillance purposes (Gawlik et al., 2021).

In sensitivity analysis, we used the standardized SARS-CoV-2 RNA counts variable (gene copies per person and day), instead of gene copies per litre. This approach, using specific population and/or human faecal indicators, is often recommended (D’Aoust et al., 2021; European Commission, 2021; Medema et al., 2020a; Zhu et al., 2021). However, it did not improve the models, likely owing to the similar characteristics in composition of raw sewage, without influence of hospitals and industrial discharges, and in daily sewage flow at the sampling points.

We also performed the analysis with the all available data (replacing values < LoQ by zero) and results are slightly different, but also statistically significant (e.g., in M6, increase in COVID-19 rates per 1 logarithmic unit of SARS-CoV-2 RNA = 16.6%, 95% CI: 10.1%, 23.7%).

There were some uncertainties both in epidemiological and in wastewater data. The sensitivity of surveillance systems to detect new cases has been enhanced throughout the pandemic. After the first epidemic wave of March 2020, the updating of the protocols for surveillance and control of COVID-19 has made it possible to ensure the capacities for early detection and response for case and outbreak management (Board of Alerts and Preparedness and Response Plans, 2020). Moreover, the improvement in diagnostic tests, the deployment of NPIs, the increase in vaccinated people, and the emergence of new SARS-CoV-2 variants may influence in the detection of cases. All these factors could affect the association between COVID-19 case rates and SARS-CoV-2 in wastewater.
A key factor is how long a COVID-19 case is considered “active”. We used the SVEA criteria (10 days), which is consistent with the ECDC (European Centre for Disease Prevention and Control, 2020) and CDC (Centers for Disease Control and Prevention, 2021) criteria for home isolation for non-severe cases. Nevertheless, we also considered an exit from SVEA of active cases of 14 days after entry, instead of ten days. In this case, associations between virus and case rates were not significant, with broader confidence intervals (e.g., for M6, increase in COVID-19 rates per 1 logarithmic unit of SARS-CoV-2 RNA = 14.2%, 95% CI: -6.5%, 40.5%). It could be due to extra uncertainties introduced when we extended the period of active case.

The shedding profile of infected individuals is one of the most critical factors in wastewater-based epidemiology, leading to uncertainty in measures and a “background noise” in wastewater data (Zhu et al., 2021). We also assumed that the proportion between active cases contributing to viral counts and infected people is constant within an urban area, but it also introduces some uncertainty.

The sampling strategy could introduce some uncertainty as well. Sampling frequency depends on the intended use of the data. Daily sampling has been suggested as ideal, but in practice wastewater is being sampled around once or twice a week (WHO Regional Office for Europe, 2021). European Commission’s Recommendation suggests a minimum sampling frequency of two samples per week (taken at inlets to wastewater treatment plants or where relevant upstream at the wastewater collecting networks) although, based on the local epidemiological situation, this sampling frequency should be reduced to one.
sample per week. In other locations (e.g., main sewer catchments) the definition of the locations and of the sampling frequencies should be adapted to the local needs (European Commission, 2021). According to our results, if samples are taken in the sewage network and virus circulation is high, one sample per week would be sufficient to meet the objectives of a sewage surveillance system for SARS-CoV-2.

We have assumed that SARS-CoV-2 circulation is constant between a sampling and the next one. Therefore, a longer time between samples could decrease the effectiveness of the surveillance system since, during this time, the system would not detect variations in the circulation of the virus in the population. On the other hand, and depending on the number of urban areas monitored, taking two or more samples per week would entail an additional cost that could make the surveillance system inefficient.

These factors introduce a high variability in results. They may be the main responsible for the relatively broader confidence intervals of the association between COVID-19 rate and virus. Because of uncertainties, models should be continually updated with new information.

The association between relative internal mobility and viral counts was statistically significant for rates the fourth day onward after sampling. The highest association occurs at 12 days after sampling (Fig. 6). The lagged effects on COVID-19 rates in urban areas is consistent with published studies about mobility patterns and COVID-19 case rates (Askitas et al., 2021; Badr et al., 2020; Flaxman et al., 2020; Iacus et al., 2020; Liu et al., 2021). In the M6, a 10% increase in internal mobility was associated with 28.3% (95% CI: 10.2%, 49.4%) increase in COVID-19 rate per 100,000 inhabitants, within the urban area, six days later.

To assess daily and seasonal mobility, we used accessible and free data. In other settings with no available data on mobility, models should include one or more terms to consider the number of contacts and seasonal variations in population, taking also into account the discontinuous mobility restrictions stabilised during the period of study.

We have included the average temperature in the models because, according to our results, it seems to be a confounding factor. However, the influence of temperature on the spread of COVID-19 is still unclear, as some studies reported that lower temperatures are more favourable for the transmission of the virus, while several studies argue that there is either no correlation between temperature and COVID-19 confirmed cases, or temperature raise can even increase the rate of transmission (Bashir et al., 2020; Briz-Redón and Serrano-Aroca, 2020; Sharifi and Khavarian-Garmirs, 2020; To et al., 2021). The relationship could also be site-specific, owing for example to different behaviour patterns relating to ambient temperature.

The daily precipitations were also statistically significant. In the model M6, one litre per square metre is associated with 3.2% (95% CI: 0.54%, 6.0%) increase in COVID-19 rate, probably due to a dilution effect on wastewater. The rest of the wastewater-related variables were not statistically significant.

In this study we also addressed if the presence of vulnerable people living within the urban areas has an influence on the COVID-19 rates. Results showed a non-significant relationship between the presence of socioeconomic deprived areas within the urban areas and COVID-19 rates (e.g., $p = 0.218$ in M6). The relationship between presence of deprived areas and viral RNA counts is not significant either ($p = 0.962$).

However, several factors increase the exposure of people of low socio-economic status to COVID-19, such as poor housing conditions, overcrowding and fewer opportunities to work from home. Poverty may also reduce the immune system’s ability to combat the virus (Patel et al., 2020). In addition, the aggregation of COVID-19 and non-communicable diseases on a background of social and economic disparity exacerbates the adverse effects of each separate disease (Horton, 2020). Therefore, the non-significant relationship could be due to the relatively low number of COVID-19 cases reported in the study.

Previous studies have usually used correlation analysis and/or simple linear regression for describing the association between viral RNA and COVID-19 epidemiological parameters (Carcereny et al., 2021; Greenwald et al., 2021; Medema et al., 2020b; Nemurí, 2020; Tomasino et al., 2021; Weidhaas et al., 2021; Westhaus et al., 2021). However, because the response (active cases) is a count, a Poisson distribution is suitable for modelling the data, as previously has been suggested by Peccia et al. (2020). According to our findings, multivariate Poisson regression analysis has shown to be useful to estimate changes in COVID-19 rates in small urban areas from SARS-CoV-2 RNA copies in wastewater.

Likewise, the usefulness of spatial analysis in a wastewater surveillance system has already been highlighted (Kitajima et al., 2020; Medema et al., 2020a; Thompson et al., 2020; WHO Regional Office for Europe, 2021). Maps to spatially visualize the data are not yet widely used in this context, but georeferenced data can be used in concert with clinical results to help decision-making (Gonzalez et al., 2020).

4. Conclusions

The methodology used in our study can provide an indicative value of how much the rate is going to change in days after sampling once the viral RNA counts are known. This methodology allows estimating changes in small areas of interest incorporating temporal and spatial variability in the estimations, so it would enable to take localised interventions and test the effectiveness of these interventions. The availability of the results of wastewater analysis is a factor of utmost importance in the surveillance system, since reducing analysis time leads to increase the time to take actions (such as NPIs) in specific urban areas.

We have shown that a quantitative approach is also possible and contribute to improve the performance of a wastewater surveillance system for SARS-CoV-2. Collaborative work between water sanitation workers, epidemiologists and environmental health practitioners is essential to improve its usefulness. The methodology used in this study is reproducible and can be used for small urban areas within cities for supporting the decision-making process, but models must be tailored to each location.

CRediT authorship contribution statement

PR selected urban areas, designed sampling management protocol, and provided wastewater analysis data. LC designed SARS-CoV-2 RNA analysis in wastewater. NL defined the concept of “active case” and provided raw COVID-19 cases. LM did the cleaning data on cases and assigned cases to basins. FR coordinated the study, designed the methodology, estimated case counts by basins, created the time series dataset, did the statistical analysis, and wrote the drafts of the report. All authors reviewed the report, contributed to the discussion and conclusions, had full access to all the data in the study and had final responsibility for the decision to submit for publication.
Declaration of competing interest
FR, LM and NL are employees of the Regional Ministry of Health and Families of Andalusia. PR and LC are employees of EMASESA, the metropolitan Water Supply and Sanitation Company of Seville. All authors declare no competing interests.

Acknowledgments
The study was supported by EMASESA and the Regional Ministry of Health and Families of Andalusia. EMASESA is a public company funded by the City Council of Seville. We wish to express our appreciation of the input provided by Alberto Carmona Ubagó, Auxiliaridia Vergara Díaz, Cecília Martín Gandul, Eva Puerto Segura and Luis Ferreiro Almeida, from the Regional Ministry of Health and Families of Andalusia, and José A. González Carballo, from EMASESA, as well as the Health Unit of the Seville City Council, EMASESA staff and workers from the Andalusian Public Health System involved in generation of data for this study.

Appendix A. Supplementary material
Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.150573.

References
AEMET, 2021. AEMET OpenData [WWW Document]. AEMET OpenData. https://opendata.aemet.es/centrosdescargas/inscio/. Accessed (9 March 2021).
Ahmed, W., Angel, N., Edison, J., Bibby, K., Bivins, A., O'Brien, J.W., Chie, P.M., Kitajima, M., Simpson, S.L., Li, J., Tscharke, B., Verheugen, R., Smith, W.J.M., Auğen, J., Dieren, L., Hugenholtz, P., Thomas, K.V., Mueller, J.F., 2020. 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. 728. https://doi.org/10.1016/j.scitotenv.2020.110748.
Baldovin, T., Amoruso, I., Fonzo, M., Buja, A., Baldovin, V., Crocchi, S., Bertolacci, C., 2021. SARS-CoV-2 RNA detection and persistence in wastewater samples: an experimental network for COVID-19 environmental surveillance in Padua, Veneto Region (NE Italy). Sci. Total Environ. 760. https://doi.org/10.1016/j.scitotenv.2021.138833.
Bhadra, M.F., Ma, B., Bilal, Komal, B., Bashir, M.A., Tan, D., Bashir, M., 2020. Correlation between climate indicators and COVID-19 pandemic in New York, USA. Sci. Total Environ. 728. https://doi.org/10.1016/j.scitotenv.2020.138833.
Bhaskaran, K., Gasparini, A., Hajat, S., Smeeth, L., Armstrong, B., 2013. Time series regression studies in environmental epidemiology. Int. J. Epidemiol. 42, 1187–1195. https://doi.org/10.1093/ijey/dyt092.
Board of Alerts and Preparedness and Response Plans, 2020. Coordinated Response Actions to Control the Transmission of COVID-19. https://doi.org/10.1007/978-1-4419-0118-7.
Bric-Redón, Á., Serrano-Aroca, Á., 2020. A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. Sci. Total Environ. 728. https://doi.org/10.1016/j.scitotenv.2020.138811.
Carcenén, A., Martínez-Velázquez, A., Bosch, A., Allende, A., Truchado, P., Cascales, J., Romalde, J.L., Lois, M., Polo, D., Sánchez, G., Pérez-Cataluña, A., del Real, C., 2021. Detection of 2019 novel coronavirus (2019-ncov) by real-time RT-PCR. Eurosurveillance 26, 2000445. https://doi.org/10.2807/1660-7970.ES.2020.25.32.2000445.
D’Aoust, P.M., Mercier, E., Montpetit, D., jia, J.J., Alexander, I., Neault, N., Baig, A.T., Mayne, J., Zhang, X., Alain, T., Langlois, M.A., Servos, M.R., MacKenzie, M., Figvez, D., MacKenzie, AE., Graber, T.E., Delatolla, R., 2021. Quantitative analysis of SARS-CoV-2 RNA in wastewater communities with low COVID-19 incidence and prevalence. Water Res. 188, 116560. https://doi.org/10.1016/j.watres.2020.116560.
Dunn, P.K., Smyth, G.K., 2018. Generalized Linear Models With Examples in R. Springer. https://doi.org/10.1007/978-1-4419-0111-8.
European Centre for Disease Prevention and Control, 2020. Guidance for Discharge and Ending of Isolation of People With COVID-19 Scope of This Document Target Audience. Stockholm.
European Commission, 2021. Commission Recommendation of 17.3.2021 on a Common Approach to Establish a Systematic Surveillance of SARS-CoV-2 and Its Variants in Wastewaters in the EU. https://doi.org/10.2760/300580.
Fajardo, P., Combarro, S., Prado, M., de la Cueva, C., Alain, T., Langlois, M.A., Servos, M.R., MacKenzie, M., Figvez, D., MacKenzie, AE., Graber, T.E., Delatolla, R., 2021. Quantitative analysis of SARS-CoV-2 RNA in wastewater communities with low COVID-19 incidence and prevalence. Water Res. 188, 116560. https://doi.org/10.1016/j.watres.2020.116560.
Kitajima, M., Ahmed, W., Bibby, K., Carducci, A., Gerba, C.P., Hamilton, K.A., Haramoto, E., Malla, B., Thakali, O., Kitajima, M., 2020. First environmental surveillance for SARS-CoV-2 RNA detection and persistence in wastewater samples: an experimental network for COVID-19 environmental surveillance in Padua, Veneto Region (NE Italy). Sci. Total Environ. 760. https://doi.org/10.1016/j.scitotenv.2020.138833.
La Rosa, G., Iaconelli, M., Mancini, P., Bonanno Ferraro, G., Veneri, C., Bonadonna, L., Romalde, J.L., Lois, M., Polo, D., Sánchez, G., Pérez-Cataluña, A., Díaz-Reolid, A., 2021. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. Sci. Rep. 11, 1972. https://doi.org/10.1038/s41598-021-9442-x.
Yang, J., Ye, G., Cheng, Z., 2020. The presence of SARS-CoV-2 RNA in the feces of COVID-19 patients. J. Med. Virol. 92, 833–840. https://doi.org/10.1002/jmv.25825.
Zhu, H., Berah, T., Eaton, J.W., Monod, M., Ghani, A.C., Donnelly, C.A., Riley, S., Vollmer, M.A.C., Ferguson, N.M., Okell, L.C., Bhart, S., 2020. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nat. 584, 257–261. https://doi.org/10.1038/s41586-020-2405-7.
Liu, Y., Mengenstein, C., Kelly, J., Lowe, R., Jil, M., 2021. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. BMC Med. 191 (19), 1–12. https://doi.org/10.1186/s12916-020-01872-8.

Medena, G., Beene, F., Heijnen, L., Pettersson, S., 2020a. Implementation of environmental surveillance for SARS-CoV-2 virus to support public health decisions: opportunities and challenges. Curr. Opin. Environ. Sci. Heal. https://doi.org/10.1016/j.coesh.2020.09.006.

Medena, G., Heijnen, L., Elsinga, G., Italiaander, R., Brouwer, A., 2020b. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in the Netherlands. Environ. Sci. Technol. Lett. 7, 511–516. https://doi.org/10.1021/acs.estlett.0c00357.

Miura, F., Kitajima, M., Omori, R., 2021. Duration of SARS-CoV-2 viral shedding in faeces as a parameter for wastewater-based epidemiology: re-analysis of patient data using a shedding dynamics model. Sci. Total Environ. 769, 144549. https://doi.org/10.1016/j.scitotenv.2020.144549.

National Statistics Institute of Spain, 2021. Population figures and Demographic Censuses. Continuous Register Statistics [WWW Document]. https://www.ine.es/dyngs/INEBase/en/categoria.htm?c=Estadistica_P&cid=1254735572981.

Nemudryi, A., 2020. Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. https://doi.org/10.1016/j.jcrx.2020.100098.

Patel, J.A., Nielsen, F.B.H., Badiani, A.A., Assi, S., Unadkat, V.A., Patel, B., Ravindrane, R., Wardle, H., 2020. Poverty, inequality and COVID-19: the forgotten vulnerable. Public Health 183, 110. https://doi.org/10.1016/j.puhe.2020.05.006.

Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-Massana, A., Ko, A.I., Malik, A., Wang, D., Wang, M., Warren, J.L., Weinberger, D.M., Arnold, W., Omori, S.B., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat. Biotechnol. 38, 1164–1167. https://doi.org/10.1038/s41587-2020-0684-z.

QCIS Development Team, 2021. QCIS Geographic Information System. R Core Team, 2021. R: A Language and Environment for Statistical Computing. Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., Sánchez, G., 2020. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water Res. 181. https://doi.org/10.1016/j.watres.2020.115942.

Regional Ministry for Equity and Social Policies of Andalusia, 2018. Estrategia regional andaluza Para la cohesión e inclusión social. Intervención en zonas desfavorecidas. Seville (Spain).

Sharif, A., Khavarian-Garmsir, A.R., 2020. The COVID-19 pandemic: impacts on cities and major lessons for urban planning, design and management. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2020.142391.

State Secretariat for Digitalisation and Artificial Intelligence of Spain, 2021. DataCOVID study [WWW Document]. https://www.mitma.gob.es/ministerio/covid-19/evolucion-movilidad-big-data/movilidad-municipal. (Accessed 9 March 2021).

Statistical and Carthography Institute of Andalusia, 2021a. Basic spatial data of Andalusia (DERA) [WWW Document]. https://www.juntadeandalucia.es/institutoedestadisticaycartografiorera/DERA. (Accessed 10 January 2019).

Statistical and Carthography Institute of Andalusia, 2021b. IDEAndalucia - Nordir y Geodir.

Thompson, JR., Nancharalah, Y.V., Gu, X., Lee, W.L., Rajal, V.B., Haines, M.B., Girones, R., Ng, L.C., Alm, E.J., Wuertz, S., 2020. Making waves: wastewater surveillance of SARS-CoV-2 for population-based health management. Water Res. 184. https://doi.org/10.1016/J.WATRES.2020.116181.

To, T., Zhang, K., Maguire, B., Tereebsy, E., Fong, I., Parikh, S., Zhu, J., 2021. Correlation of ambient temperature and COVID-19 incidence in Canada. Sci. Total Environ. 750, 141484. https://doi.org/10.1016/j.scitotenv.2020.141484.

Tomasino, M.P., Semedo, M., Ferraz, E., Rocha, A., Carvalho, M.F., Magalhães, C., Mucha, A.P., Vieira e Moreira, P., 2021. SARS-CoV-2 RNA detected in urban wastewater from Porto, Portugal: Method optimization and continuous 25-week monitoring. Sci. Total Environ. 792, 148467. https://doi.org/10.1016/j.scitotenv.2021.148467.

Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Forsgren, K., Laan, J., VanderLaan, N., 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. Sci. Total Environ. 775, 145790. https://doi.org/10.1016/j.scitotenv.2021.145790.

Westhau, S., Weber, F.A., Schivy, S., Linnenmann, V., Brinkmann, M., Widera, M., Groves, C., Janke, A., Hollert, H., Wintgens, T., Geisek, S., 2021. Detection of SARS-CoV-2 in raw and treated wastewater in Germany – suitability for COVID-19 surveillance and potential transmission risks. Sci. Total Environ. 751, 141750. https://doi.org/10.1016/j.scitotenv.2021.141750.

WHO Regional Office for Europe, 2020. Status of Environmental Surveillance for SARS-CoV-2 Virus. Copenhagen, Denmark.

WHO Regional Office for Europe, 2021. Report Expert Consultation on Public Health Needs Related to Surveillance of SARS-CoV-2 in Wastewater. Summary Report. Copenhagen, Denmark.

Wu, J., Wang, Z., Lin, Y., Zhang, L., Chen, J., Li, P., Liu, W., Wang, Y., Yao, C., Yang, K., 2021. Technical framework for wastewater-based epidemiology of SARS-CoV-2. Sci. Total Environ. 791.

Zaki, N., Mohamed, A.A., 2021. The estimations of the COVID-19 incubation period: a scoping reviews of the literature. J. Infect. Public Health 14, 638–646. https://doi.org/10.1016/j.jiph.2021.01.019.

Zhu, Y., Oishi, W., Barros, C., Saito, M., Kitajima, M., Sano, D., 2021. Early warning – Early warning and intervention dynamics. Nat. Biotechnol. 38, 1167. https://doi.org/10.1038/s41587-2020-100098.