

Hydrological and physicochemical parameters associated with SARS-CoV-2 and pepper mild mottle virus wastewater concentrations for a large-combined sewer system

Luan Nguyen Thanh ^{a,*}, Mounia Hachad^a, Natasha McQuaid^a, Kateryna Krylova^a, Loan Nguyen Ha Thanh^a, Flavia Visentin^a, Jean-Baptiste Burnet^a, Fernando Sanchez Quete^b, Thomas Maere ^c, Alexandra Tsitouras^b, Peter Vanrolleghem^c, Dominic Frigon^b, Stephanie Loeb^b, Sarah Dorner^a and Eyerusalem Goitom^a

^a Département des génies civil, géologique et des mines, Polytechnique Montréal, Montréal, Canada

^b Department of Civil Engineering, McGill University, Montreal, QC, Canada

^c Department of Civil and Water Engineering, Université Laval, Quebec City, QC, Canada

*Corresponding author. E-mail: thanh-luan.nguyen@polymtl.ca

 LNT, 0009-0004-6592-0999; TM, 0000-0003-2885-0421

ABSTRACT

During COVID-19, surveillance of SARS-CoV-2 in wastewater has been a promising tool for tracking viral infection at the community level. However, in addition to the shedding rates within the community, SARS-CoV-2 concentrations in raw wastewater are influenced by several environmental factors. This study investigated the effects of wastewater characteristics on the viral quantification of SARS-CoV-2 and pepper mild mottle virus (PMMoV) for a large wastewater system with combined sewers. Principal component analysis illustrated that water temperature negatively correlates with SARS-CoV-2 and PMMoV in wastewater, but flow rate and EC are highly correlated with SARS-CoV-2 in spring and winter. The normalization using EC enhanced the correlation with clinical data compared to normalization using pH, flow rate, and raw SARS-CoV-2. The normalization using PMMoV reduced the correlation with clinical data. Multiple linear and random forest (RF) applied to predict the concentrations of SARS-CoV-2 in wastewater, given the confirmed cases and physicochemical parameters. RF regression was the best model to predict SARS-CoV-2 in wastewater ($R^2=0.8$), with the most important variables being the confirmed cases followed by water temperature. RF model is a potent predictor of the presence of SARS-CoV-2 in wastewater. This enhances the degree of reliability between community outbreaks and SARS-CoV-2 monitoring.

Key words: COVID-19, regression models, SARS-CoV-2 normalization, wastewater-based epidemiology (WBE)

HIGHLIGHTS

- SARS-CoV-2 in wastewater was affected by the physiochemical parameters.
- Normalization using EC slightly improved the correlation between wastewater-based epidemiology data and clinical data.
- Random forest is more powerful than a multiple linear regression model in the prediction of SARS-CoV-2 concentrations under various impacts of physiochemical parameters.
- The number of infected persons and temperatures are primary factors driving the concentration of SARS-CoV-2 in wastewater.

1. INTRODUCTION

In the early days of the COVID-19 pandemic in 2020, the high rate of COVID-19 infections worldwide was largely underestimated as many infected people did not have access to clinical testing (Kumblathan *et al.* 2021). Additionally, asymptomatic infected persons accounted for 15–30% of all positive cases (Pollan *et al.* 2020). In response to the pandemic, the global scientific community mobilized and implemented wastewater surveillance to determine the magnitude of infection rates and predict the pandemic's evolution at a local level (Haramoto *et al.* 2020; Pollan *et al.* 2020; Carrillo-Reyes *et al.* 2021; Hasan *et al.* 2021). Thus, earlier in the pandemic, SARS-CoV-2 Ribonucleic acid (RNA) was detected in human excreta (Izquierdo-Lara *et al.* 2021); and in wastewater in several countries such as the Netherlands (Medema *et al.* 2020), Australia (Ahmed *et al.* 2020a), United Arab Emirates (Hasan *et al.* 2021), and Mexico (Carrillo-Reyes *et al.* 2021). Canadian wastewater surveillance for SARS-CoV-2 RNA across several locations has resulted in a rich dataset that could potentially be

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

used in comprehensive retrospective, multidisciplinary analyses to improve our understanding of the dynamics of the subsequent waves of the COVID-19 pandemic (Hrudey *et al.* 2022). Wastewater monitoring for SARS-CoV-2 has demonstrated sufficient efficiency for the improvement of the health surveillance systems in Canada (Hrudey *et al.* 2022). SARS-CoV-2 surveillance in wastewater delivers information to policymakers and the community rapidly and cost-effectively (Gonzalez *et al.* 2020; Spurbeck *et al.* 2021).

Pepper mild mottle virus (PMMoV) is a promising indicator for monitoring human fecal loads in wastewater. The PMMoV, a plant pathogenic virus, is the most prevalent RNA virus in human feces (Kitajima *et al.* 2018) and was measured in high concentrations in fecal matter (10^9 viral copies per gram of dry weight (Zhang *et al.* 2006) in untreated and treated wastewater (10^5 – 10^6 gc/L, (Colson *et al.* 2010; Kitajima *et al.* 2018; Hamza *et al.* 2019; Bonanno Ferraro *et al.* 2021). Additionally, the occurrence of PMMoV is more persistent and abundant than commonly used microbial trackers (Kitajima *et al.* 2018; Hamza *et al.* 2019); therefore, PMMoV has often been used for SARS-CoV-2 and other pathogens in wastewater surveillance as a fecal biomarker (Wu *et al.* 2020; D'Aoust *et al.* 2021; Fores *et al.* 2021).

Many studies have compared trends of SARS-CoV-2 in wastewater with clinical cases (Ahmed *et al.* 2020a; Haramoto *et al.* 2020; Tomasino *et al.* 2021; Robotto *et al.* 2022). Several studies have considered the normalization of SARS-CoV-2 with a fecal biomarker to improve the correlation with the number of infections from the clinical data. However, oftentimes this normalization of SARS-CoV-2 using PMMoV, crAssphage, or other fecal markers minimally improves or reduces the correlation of SARS-CoV-2 with clinical data (Zhang *et al.* 2006; Nagarkar *et al.* 2022; Vadde *et al.* 2022). The presence of SARS-CoV-2 and fecal indicators can be affected by various factors in the wastewater sewer systems, including the seasonal, daily, and precipitation-driven hydraulic patterns in sewer flow, partitioning with the particulate phase (suspended solids), in addition to the population shedding rates. Therefore, further research is needed to investigate the impact of wastewater characteristics on SARS-CoV-2, PMMoV, and their implication in the association between SARS-CoV-2, and clinical data. The objectives of this research are:

- (i) To assess temporal variability of SARS-CoV-2 and PMMoV concentrations across the two main interceptor sewers of Montreal.
- (ii) To determine the potential impacts of hydrological (flow rate) and physicochemical wastewater parameters (total suspended solid (TSS), volatile suspended solid (VSS), electrical conductivity (EC)), and temperature on SARS-CoV-2 and PMMoV concentrations.
- (iii) To assess if normalizing SARS-CoV-2 concentration using various wastewater characteristics facilitates its comparison with clinical metrics.
- (iv) To statistically evaluate the effects of wastewater characteristics on the occurrence and fate of SARS-CoV-2 in the sewer and determine the correlation with clinical cases.

2. METHODOLOGY

2.1. Site description and sampling plan

The City of Montréal (Québec, Canada) is located on an island in the Saint-Lawrence River. Given that all water flows toward the North-East tip of the Island of Montreal, the Jean-R.-Marcotte wastewater treatment plant (WWTP) was designed to receive all wastewater from the sewer-connected population, making it one of the largest wastewater treatment plants in the world. The sewer network consists of a mixture of combined (rain and sanitary) sewers in the older, most densely populated sectors and sanitary or pseudo-sanitary sewers in the more recently developed lower-density sectors. In dry weather, the WWTP treats around 2.5 million m³ of water daily and up to 7.6 million m³ when it rains. Two large interceptor sewers convey wastewater and stormwater to the WWTP. The WWTP serves a total of approximately 1.763 million inhabitants, with the North interceptor serving 51% of the population and the South interceptor serving 49% of the population. Twenty-four-hour composite flow samples were collected from the influent interceptor sewers collocated with the WWTP's regular sampling points for monitoring and regulatory purposes and were collected daily and biweekly from February 2020 to December 2021. The samples were either stored in a freezer at -80 °C (frozen samples) or at 4 °C (fresh samples) depending on the period of collection (Table S1). Hydraulic and physicochemical parameters that are thought to play a role in the processes of viral dynamics and transport (e.g., the influence of the sorption of virus fragments onto suspended materials and decay rate) were measured by the City of Montreal using standard methods and instruments for their regular routine monitoring, except for EC, which was measured in the laboratory. The daily confirmed cases were reported by the National

Institute of Public Health of Quebec. COVID-19 infection waves were defined as Wave 1 (Wild type 25 February–11 July 2020), Wave 2 (Wild type, 23 August 2020–20 March 2021), Wave 3 (Alpha variant, 21 March–17 July 2021), Wave 4 (Delta variant, 18 July–4 December 2021), and Wave 5 (Omicron variant, 5 December 2021–12 March 2022). Although names of dominant variants in Canada for each wave are provided by the [Government of Canada \(2022\)](#), Waves 1–5 in Montréal consisted of mixtures of variants identified in the wastewater samples that were correlated with clinical cases ([N'Guessan *et al.* 2022](#); [Jayme *et al.* 2024](#)). The strongest signal with regard to variants of concern was the arrival of the Alpha variant at the end of 2020, becoming dominant in wastewater and clinical samples in March 2021.

2.2. SARS-CoV-2 RNA extraction and molecular analysis

2.2.1. Viral RNA concentration

Samples (200 mL) were homogeneously mixed and transferred into the falcon tubes using a graduated cylinder. The pH was adjusted to between 3.5 and 4.0 with HCl (1N), and MgCl_2 (final concentration = 25 mM) was added. Samples were centrifuged at 3,348 g for 10 min to separate the supernatant and the pellets. The supernatant was filtered using mixed cellulose ester membranes (0.45 μm pore size, 47 mm diameter). Then, the filters were cut into eight pieces to facilitate the subsequent digestion process and placed carefully into a 5 mL-Eppendorf tube using forceps. The pellets and filters were kept at -80°C prior to RNA extraction.

Bovine Respiratory Syncytial Virus (BRSV) in the form of Inforce 3, an intranasal bovine vaccine containing a live inactivated virus (Zoetis, Parsippany-Troy Hills, NJ, USA), was used as an external control in the RNA SARS-CoV-2 extraction process. BRSV was spiked at 10% v/v directly to Eppendorf tubes prior to extraction as an external process control.

2.2.2. Viral RNA extraction and RT-qPCR assay

Viral RNA was extracted using the RNeasy Power microbiome kit (Qiagen, Germantown, MD, USA) following the manufacturer's instructions with a slight modification (samples were incubated at 55°C and shaken at 300 rpm after adding BRSV and lysis buffer solutions) ([Sanchez-Quete *et al.* 2024](#)). Following the lysis step, samples were centrifuged at 2,500 g at room temperature for 2 min, and all the supernatant was transferred to clean 2 mL tubes. Afterwards, all samples were treated with 1,200 μL of phenol:chloroform:isoamyl (pH 6.5–8.0, Sigma-Aldrich, Saint Louis, USA) and 10% v/v β -mercaptoethanol in the lysis buffer.

The quantification of SARS-CoV-2 target gene N1, PMMoV, and BRSV RNA was performed by a Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) with the 5plex Rotor Gene-Q (Qiagen, Germantown, MD, USA) using primer and probe sets designed by the Centers for Disease Control and Prevention (Protocol S1). RT-qPCR reaction was performed in a total volume of 25 μL containing $2 \times$ TaqMan Fast Virus 1-Step Master mix (Applied Biosystems, Waltham, MA, USA), 500 nM of each primer, and 250 nM of the specific probe. Amplification consisted of reverse transcription at 95°C for 5 min, initial denaturation at 95°C for 20 s, followed by 40 cycles of 3 s at 95°C and 30 s at 60°C ([Sanchez-Quete *et al.* 2024](#)). Quality control measures were implemented for each RT-qPCR run, including no-template controls, positive controls, and internal controls. All no-template controls (NTC) samples yielded negative results. Runs were considered valid only if they met the following criteria: $R^2 > 0.98$, the slope between -3.1 and -3.6 , and efficiency between 90 and 110%.

2.3. Data analysis

The data from SARS-CoV-2, PMMoV, BRSV recovery, hydraulic, and physicochemical parameters were not normally distributed. We, therefore, used the Kruskal–Wallis test to determine the differences in SARS-CoV-2 and PMMoV across seasons (Table S2). The Kruskal–Wallis post hoc pairwise comparison test was performed to analyze the differences between each pair of seasons. Principal component analysis (PCA) was used to evaluate the impacts of physico-chemical parameters on SARS-CoV-2 and PMMoV concentrations in wastewater. The SARS-CoV-2 was normalized using multiplication or division with physicochemical parameters and PMMoV using pooled data to improve the understanding of the impacts of wastewater characteristics on the correlation between SARS-CoV-2 in wastewater and clinical data. Subsequently, the Spearman correlation test was performed with raw or normalized SARS-CoV-2 concentration in wastewater and with clinical data (including confirmed cases, positive tests, moving average of 7 days) across the full dataset as well as during each of the COVID-19 waves. We also investigated lag times of 0–8 days between SARS-CoV-2 in sewage and confirmed cases from 28 February to 30 June in 2021, with the high-frequency sampling period to examine its impact on correlations.

Multiple linear regression and random forest models were used to predict the occurrence of SARS-CoV-2 in wastewater, with confirmed cases and wastewater characteristics serving as predictors. All models were trained and tested using 80 and 20%, respectively, of the pooled data. SARS-CoV-2 in wastewater was the dependent variable. Next, the models were built with non-transformed (actual observations) and natural log-transformed values of SARS-CoV-2 quantification. Model results were evaluated using the correlation coefficient, mean absolute error, and root mean square error. Finally, the accuracy of predicted SARS-CoV-2 levels in wastewater from the model was compared with actual SARS-CoV-2 in wastewater.

All statistical analyses and data visualizations were executed with the R programming language (version 4.2.2) with $\alpha = 0.05$. The data were standardized using the *z*-score method prior to PCA. The PCA results were visualized using the 'ggbiplot' package. The 'ggplot2' and 'ggscatter' packages were used to visualize the results from multiple linear regression and random forest models. Multiple linear and random forest regression models were implemented by 10 *k*-fold cross-validations and the 'repeatedcv' method using the 'CARET' package (Kuhn *et al.* 2020).

3. RESULTS AND DISCUSSION

3.1. Seasonal variation of SARS-CoV-2 and fecal indicator (PMMoV) in wastewater

3.1.1. SARS-CoV-2 and PMMoV concentrations

Table S1 displays the mean, median, maximum, and minimum values for SARS-CoV-2, PMMoV, physicochemical, and hydrological parameters measured at the WWTP. SARS-CoV-2 and PMMoV concentrations varied considerably within and across interceptors over the entire study period (Figure 1).

The concentration of SARS-CoV-2 within interceptors across seasons and years ranged from 0.3 to 290.6 gc/mL. In 2020, the maximum concentration of SARS-CoV-2 was 167.4 gc/mL, whereas in 2021, it was 290.6 gc/mL. The mean SARS-CoV-2 concentrations in the North interceptor were 1.56 and 1.72 times higher than in the South interceptor for samples stored at -80 and 4 °C, respectively (Figure 1A, B, Appendix 1). The average increase of SARS-CoV-2 concentration at the North interceptor could be attributed to the different number of people served by each interceptor as well as the number of infected people (Amoah *et al.* 2022). The variations observed in the concentration of SARS-CoV-2 between and within the interceptors could be attributed to different factors, including the time travel within the wastewater infrastructure (Bitter *et al.* 2022; Mac Mahon *et al.* 2022). In fact, the total network of wastewater interceptors on the island of Montreal extends over 115 km, and the travel time of approximately 8 h varies depending on precipitation and snowmelt that influence flow rates and velocities. The size of the population served, the shedding of viral particles by infected individuals throughout various infection waves (Lescure *et al.* 2020), and dilution from non-viral sources, such as rainwater runoff, groundwater infiltration, and industrial effluent, can further significantly impact viral concentrations. These dilution rates can fluctuate between different areas and over time, especially during periods of heavy rainfall (Langeveld *et al.* 2023).

PMMoV concentrations ranged from 5.8 to 6,484.5 gc/mL, and they were considerably higher in 2021 (annual mean 1,116 gc/mL in the North interceptor and 937 gc/mL in the South interceptor) than in 2020 (annual mean 485 gc/mL in the North interceptor and 291 gc/mL in the South interceptor) in Appendix 1. Except for samples stored at -80 °C, there were no substantial differences in PMMoV concentrations observed between the North and South interceptors (Figure 1C, 1D, Appendix 1). PMMoV concentrations were significantly lower in samples preserved at -80 °C, which could be attributable to the fact that all samples taken in 2020 were kept in the freezer for longer periods before being processed due to quarantine and lockdown measures. In contrast, the 2021 samples were evaluated within 48 h (Figure 1(d)). Simpson *et al.* (2021) found that freezing samples for an extended period affects viral particle stability and lowers RT-qPCR quantification by about 60%. Other studies highlighted the importance of proper storage and processing techniques in maintaining the integrity of viral genetic material in wastewater samples (Ahmed *et al.* 2020b). Sanchez-Quete *et al.* (2024) examined the decay of RNA viruses, specifically SARS-CoV-2, in wastewater stored at 4 °C. It was found that the decay rate of the viral genome is influenced by the storage temperature, with a noted decrease in viral concentration over time when stored under these conditions, whereas Williams *et al.* (2024) found that samples stored at 4 and -80 °C generally maintained higher RNA quality and that historical wastewater samples archived at -80 °C can be reliably quantified even after 12 months of storage.

On the other hand, we tested whether SARS-CoV-2 and PMMoV were linked under both storage settings and found that SARS-CoV-2 and PMMoV were positively associated for samples held at 4 °C (Spearman correlation, $R = 0.43$, $p < 0.01$, Figure 1(f)) but not for samples stored at -80 °C (Figure 1(e)). Nevertheless, the correlation was weak at 4 °C. The observed

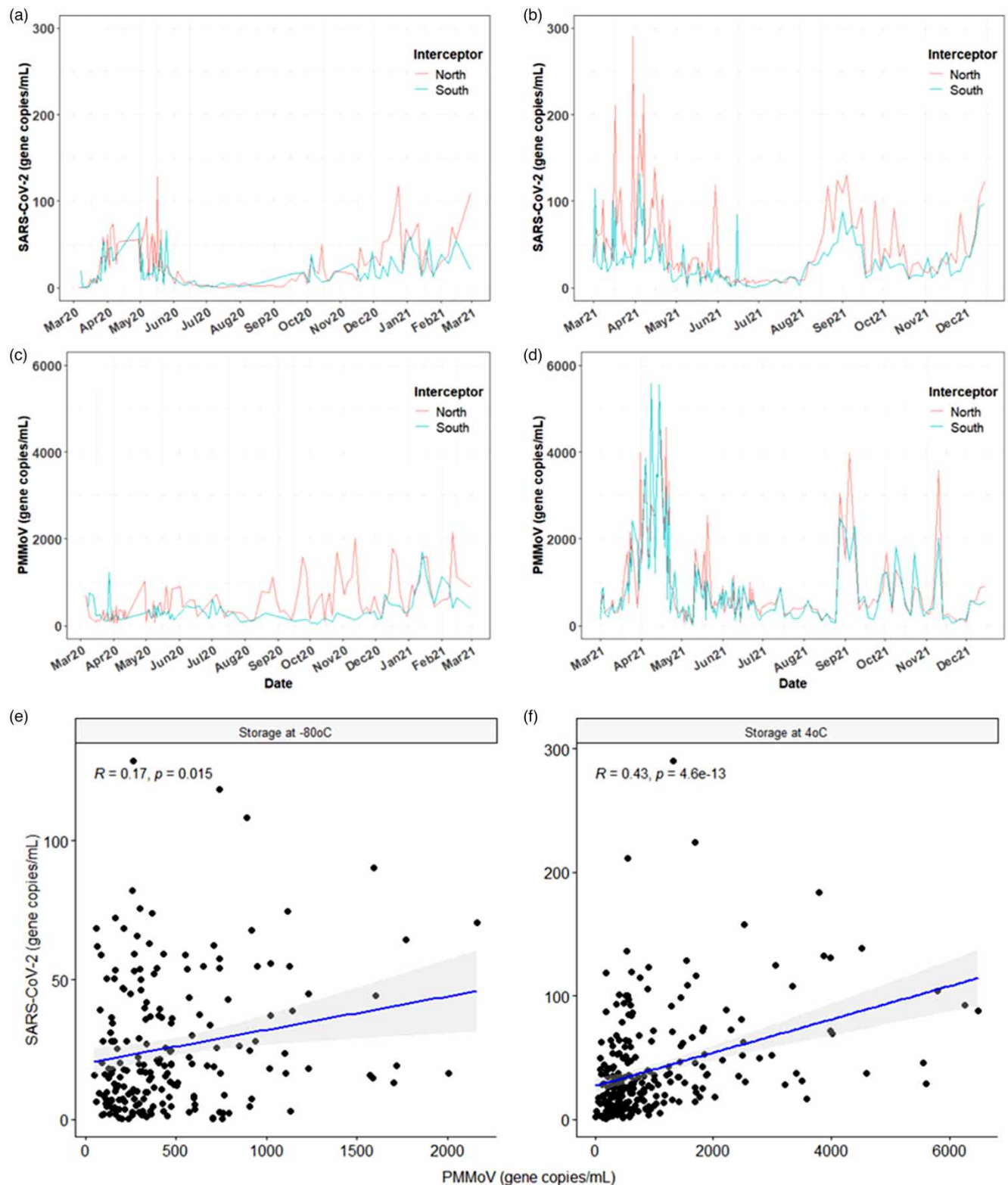


Figure 1 | Temporal series of SARS-CoV-2 (a and b) and PMMoV (c and d) concentrations in North and South interceptors from March 2020 to December 2021. Panels (a) and (c) show time series for samples stored at -80°C before analysis. Panels (b) and (d) show samples stored at 4°C . Panels (e) and (f) show the correlation of SARS-CoV-2 and PMMoV for samples stored at -80 and 4°C .

difference in the association between SARS-CoV-2 and PMMoV in fresh and frozen samples can be explained by the impact of storage conditions. Studies found that PMMoV is highly persistent in both particle and aqueous phases at various temperatures (Sala-Comorera *et al.* 2021; Burnet *et al.* 2023). As a result, sample storage conditions should be considered before using PMMoV as a normalizing biomarker. Sala-Comorera *et al.* (2021) found a difference in the persistence rate of SARS-CoV-2 RNA and PMMoV, implying that the use of PMMoV as a surrogate is restricted.

3.1.2. Seasonal variation of SARS-CoV-2 and PMMoV

The seasonal variation in SARS-CoV-2 concentration was observed across all the seasons (Figure 2). SARS-CoV-2 concentrations in summer were significantly lower than in the other seasons (Figure 2(a)). This difference can be explained by the lower number of infected persons leading to the lower concentration of SARS-CoV-2 in wastewater. Seasonal variation

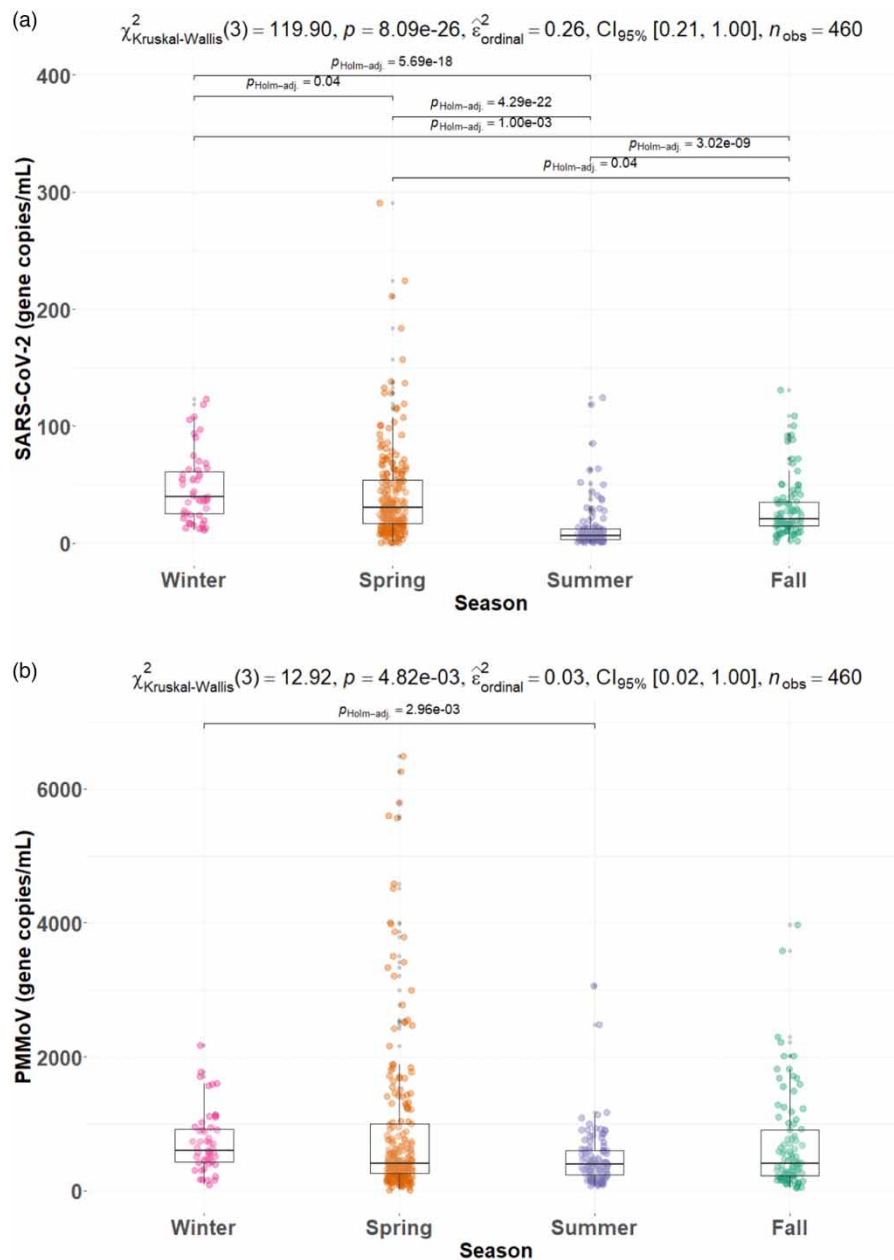


Figure 2 | The seasonal variation of SARS-CoV-2 and PMMoV in both sewer interceptors over a period of more than 1.5 years of monitoring.

for different respiratory viruses has been well-known for hundreds of years (Moriyama *et al.* 2020). Winter conditions or colder seasons, both indoor and outdoor, enhance the rapid spread of a range of respiratory virus infections, including SARS-CoV-2 (Molinari *et al.* 2007; Moriyama *et al.* 2020; Paules *et al.* 2020). On the other hand, PMMoV is constantly observed in wastewater and is less influenced by seasonality and temporal variation (Figure 2(b)). According to Kitajima *et al.* (2018), PMMoV is known to be found in various water sources worldwide without substantial seasonal changes. The variation of both PMMoV and SARS-CoV-2 in spring highly fluctuated in comparison to the other seasons. The snowmelt period might increase the peak transport of fecal contaminants, including pathogens, viruses, and bacteria, by reducing travel times and decay (Staley *et al.* 2017; Tolouei *et al.* 2019).

3.2. Relationship between physiochemical parameters and viral concentration of SARS-CoV-2 and PMMoV

The physiochemical parameters in both North and South interceptors at different storage temperatures are summarized in Table S3. Site-specific variations were observed in wastewater characteristics during the study period involving variations of SARS-CoV-2 gene target (N1) quantification and PMMoV. The pH measured at the WWTP (6.1–8.7) prior to pH adjustment for sample analysis had a maximum recorded value in the North interceptor in autumn 2020. The EC ranged from 108.8 to 2,830 ($\mu\text{S}/\text{cm}$), with the highest value measured in the North interceptor in winter 2020. The highest flow rate (7.5–37.3 m^3/s) was observed in the North interceptor in 2020 during spring associated with the snowmelt period. TSS (34–351 mg/L) and VSS (32–228 mg/L) exhibited variability across the years and interceptors. As much of the sewer network consists of combined sewers, overflows commonly occur during the spring snowmelt period and during intense precipitation. With higher flow, the travel time in the sewer network decreases, resuspension of sewer sediments increases, which can partly offset concentration decreases from the dilution from stormwater (Tolouei *et al.* 2019). The water temperature fluctuated between 5.4 and 25.1 $^{\circ}\text{C}$, whereas turbidity levels varied from 25.3 to 238 Nephelometric Turbidity Unit (NTU), with the highest concentration recorded in the North interceptor during the summer.

PC1 and PC2 accounted for 28.6 and 21.5% of the total variance, respectively (Figure 3). Forms of particulate matter (TSS, VSS, turbidity) have the most contribution in explaining the variability in the PCA dataset (Figure S2). The water temperature contributed to around 14% of the variation, while BRSV recovery had the lowest contribution in the variance captured by PC1 and PC2. The PCA biplot illustrates distinct clusters of physicochemical parameters, SARS-CoV-2, PMMoV, and seasons,

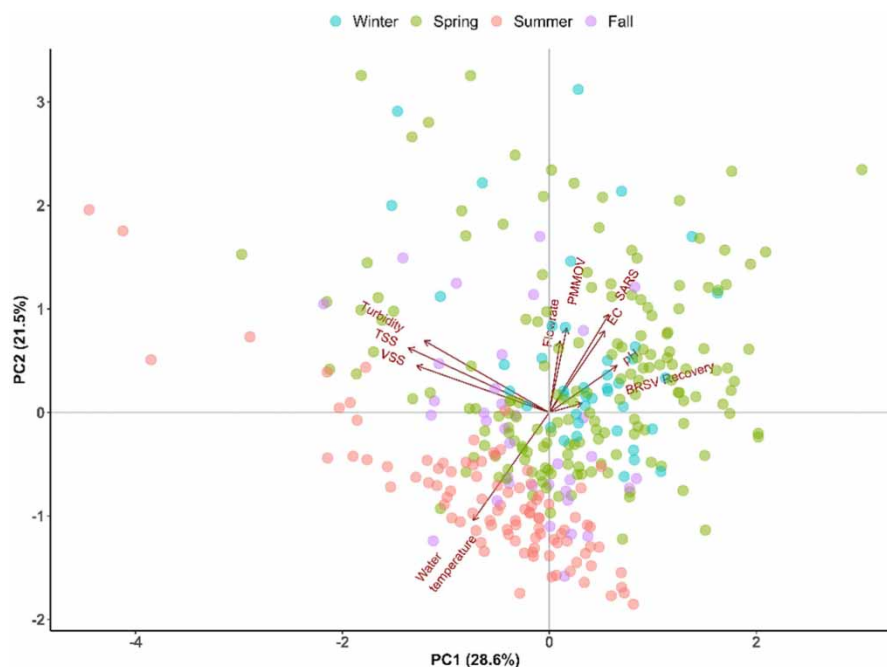


Figure 3 | PCA biplot of SARS-CoV-2, PMMoV, and wastewater parameters with a two-dimensional PCA biplot illustrating both the PC scores of samples (dots) and the loadings of variables (vectors). The dispersion of data points indicates the degree of correlation between the wastewater quality parameters and the virus's gene copies.

suggesting a relationship between wastewater composition and the prevalence of the viruses. There were no or weak associations between SARS-CoV-2 or PMMoV with TSS, VSS, and turbidity (evidenced by the 90° angle between the respective vectors in Figure 3). Both viruses were significantly correlated with the EC, flow rate, and, to a lesser extent, pH during the winter and spring seasons. In fact, pH, EC, and seasonal factors, such as precipitation or snowmelt, have been reported to increase the incidence of SARS-CoV-2 and PMMoV in wastewater (Amoah *et al.* 2022; Dhiyebi *et al.* 2023; Langeveld *et al.* 2023). EC is a parameter closely related to the degree of dilution of samples, whereas flow rate has a greater effect on the settling and resuspension of TSS and settled virus particles in combined sewers, especially during periods of snowmelt or heavy rain. Furthermore, the analyzed flow rates are the flow rates that enter the WWTP where samples are collected and do not include the portion of the total flow lost to receiving waters through combined sewer overflows.

The distinct obtuse angle and the summer points that cluster around the water temperature vector (Figure 3) show that the water temperature over 20 °C negatively correlates with SARS-CoV-2 and PMMoV. As reported by Dhiyebi *et al.* (2023), even though the temperature in sewers does not change as much as the air temperature, it can still affect how PMMoV is found in wastewater. Studies also showed that the decay rate of SARS-CoV-2 RNA increased with temperature (Ahmed *et al.* 2020b; Burnet *et al.* 2023). However, we would expect wastewater temperature to be a minor contribution relative to the lower shedding patterns of the SARS-CoV-2 and PMMoV in summer, as seen in our seasonal virus variation results presented in Section 4.1. Spring samples were generally positively correlated with EC and pH. Many of the spring SARS-CoV-2 samples from 2021 consisted of the Alpha variant of concern (B.1.1.7).

Variations in virus shedding rates and patterns, flow changes, and dilution events, along with the size of the population served in the sewershed, create challenges in directly comparing wastewater data with clinical testing data; normalizing data based on wastewater quality measurements would address these issues and enhance the utility of wastewater-based surveillance (WBS) in predicting COVID-19 incidence in the community. Utilizing multivariate analysis such as PCA, therefore, assists in identifying the most relevant wastewater parameters that could potentially enhance the correlation between WBS and clinical data.

3.3. Comparison between the correlations of SARS-CoV-2 normalization and clinical data

The associations between SARS-CoV-2 in sewage and clinical data were examined using Spearman's correlation test. We compared the correlation strengths between non-normalized and normalized SARS-CoV-2 data for gene targets (normalized meaning whether they were multiplied or divided by wastewater properties) and clinical data such as the confirmed cases, the 7-day moving average of confirmed cases, and positive tests. The data from the sewershed of both interceptors were pooled (Figure 4).

Overall, correlation coefficients were statistically significant ($p < 0.05$). Both unnormalized and normalized SARS-CoV-2 were generally strongly correlated with the confirmed cases, positive tests, and the 7-day moving average. All investigated lag-times (0–8 days) were found to be statistically significant (Table S4), and it is difficult to discern the optimal time shift (2–6 days, depending on the normalizing methods). The leading time results were comparable to those in Alberta (Canada) (Maal-Bared *et al.* 2023) and the meta-analysis by Bibby *et al.* (2021). The moving average of SARS-CoV-2 wastewater data provides a more robust estimate of trend directions than raw data in sewage surveillance (Ai *et al.* 2021). The incubation of COVID-19-infected persons is up to 5–6 days (Backer *et al.* 2020) and around 30–40% of asymptomatic infections occur during 14 quarantine days (Oran & Topol 2020; Rasmussen & Popescu 2021).

None of the normalization approaches resulted in statistically significant changes in correlations between WBS and clinical data. However, the normalization approach involving multiplying by measured wastewater parameters (except for PMMoV) improved correlations between all wastewater-based surveillance data and clinical data. The normalization of SARS-CoV-2 concentrations by EC (multiplication) had the highest correlation with clinical data (Table S4). EC-based normalization ($P = 0.69$ – 0.71) improved correlations by around 0.1 compared to unnormalized SARS-CoV-2 ($P = 0.61$ – 0.63) in Table S4. Normalization of SARS-CoV-2 concentrations by PMMoV ($P = 0.44$ – 0.55) did not enhance the correlations between WBS and clinical data; in fact, it resulted in even lower correlations compared to unnormalized SARS-CoV-2.

The correlation between unnormalized and normalized SARS-CoV-2 with confirmed cases was also investigated for each separate COVID-19 wave that covered our sampling period (Table S4). The main objective was to investigate if a period of higher incidence of SARS-CoV-2 and frequent wastewater sampling would lead to a better correlation between WBS and clinical data. The results show that the correlation for raw and most normalized SARS-CoV-2 data with clinical data was the highest for the third wave (21 March–17 July 2021).

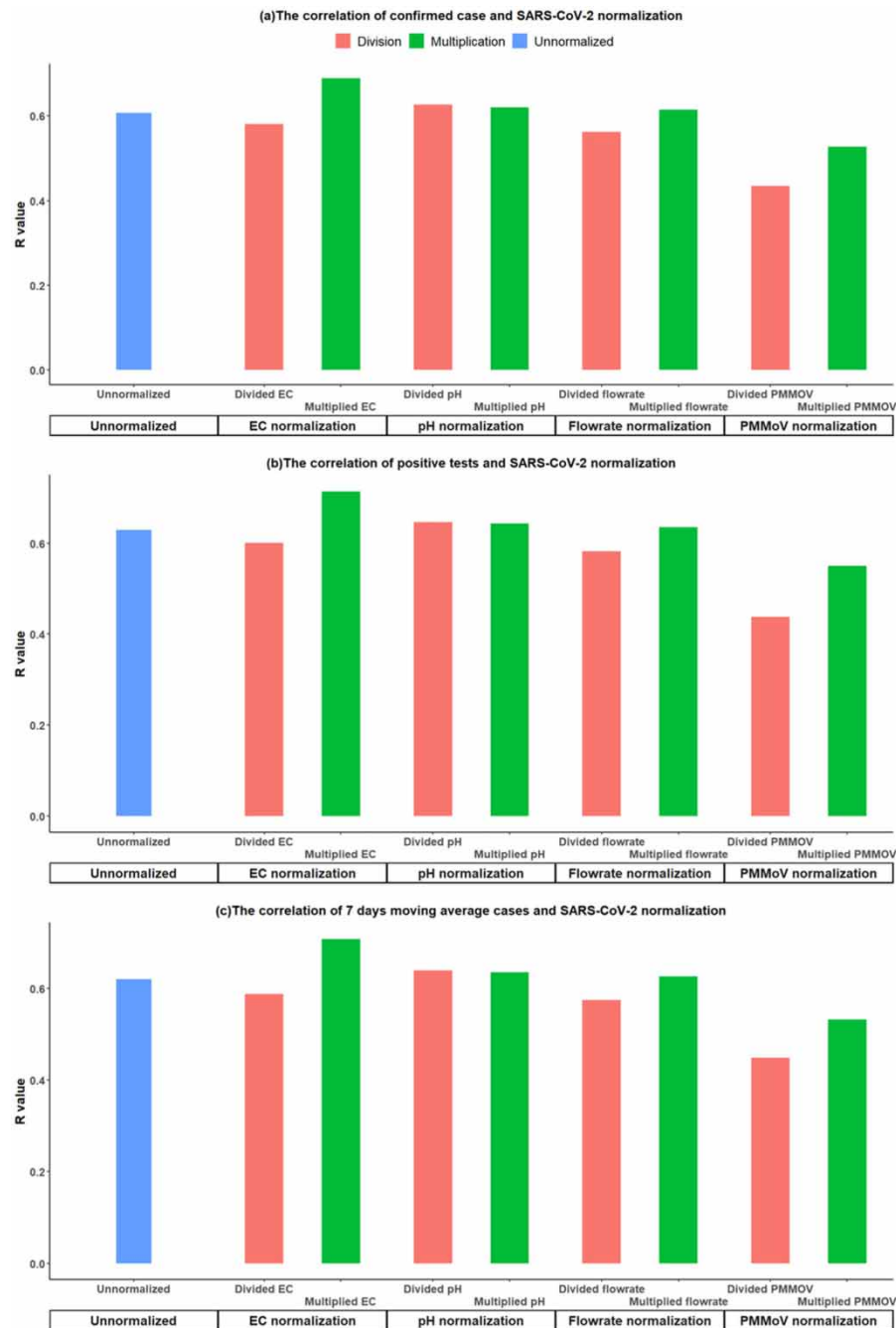


Figure 4 | The strength of association of unnormalized and normalized SARS-CoV-2 with different clinical data (the number of confirmed cases, positive tests, and moving average 7 days).

Numerous published studies have tried to identify the optimal method and most effective biomarker for normalizing SARS-CoV-2 levels in wastewater, aiming to improve COVID-19 surveillance data validation and interpretation (Greenwald *et al.* 2021). Among the various approaches explored, normalization by daily average flow rate has emerged as a widely utilized method in sewage surveillance. This approach helps elucidate fluctuations in SARS-CoV-2 levels in wastewater attributable to factors like rainfall, snowmelt, or groundwater inflow (Arabzadeh *et al.* 2021; Mazumder *et al.* 2022; Langeveld *et al.* 2023). PMMoV is commonly utilized as a biomarker to detect human fecal contamination and has also been suggested for standardizing SARS-CoV-2 levels in wastewater to better understand the variability in the virus concentrations due to

various sewer processes and population dynamics (Wu *et al.* 2020; D'Aoust *et al.* 2021). Nevertheless, like our findings, numerous studies have indicated that normalizing the data using PMMoV did not yield a substantial enhancement in the relationship between the viral load in wastewater and the incidence of COVID-19 cases. In some cases, PMMoV normalization even increases the noise and reduces the correlation with epidemiological metrics (Feng *et al.* 2021; Nagarkar *et al.* 2022; Maal-Bared *et al.* 2023). The seasonality of PMMoV, variations in local dietary habits, the types of matrices collected (such as primary sludge versus raw wastewater samples), the lab analysis method, distinctions in the protein-capsid RNA structure and partitioning characteristics of SARS-CoV-2 and PMMoV between liquid and solid phases were proposed as the primary factors contributing to the weak improvement of associations identified with clinical data (Zhang *et al.* 2006; D'Aoust *et al.* 2021; Mazumder *et al.* 2022).

Alternative normalization approaches have been proposed in recent studies. SARS-CoV-2 normalized by fecal makers such as crAssphage, fecal coliforms, HF183, and β -2 microglobulin resulted in low improvements in association with COVID-19 cases (Zhang *et al.* 2006; Nagarkar *et al.* 2022; Vadde *et al.* 2022). Whereas the SARS-CoV-2 normalized with chemical biomarkers such as caffeine, and the metabolite paraxanthine (Hsu *et al.* 2022) strongly improved the correlation with case numbers and population due to their high stability and constancy in the sewer systems (Choi *et al.* 2020; Hsu *et al.* 2022). The uncertainty and the knowledge gap in fecal shedding for SARS-CoV-2 and fecal biomarkers such as PMMoV and crAssphage, directly influence the accuracy of the wastewater-based epidemiology model (LaTurner *et al.* 2021; Arts *et al.* 2023). Additionally, the presence of these viruses in the sewer, as evidenced by their apparent decay rate (Ahmed *et al.* 2020b; Sala-Comorera *et al.* 2021) and the RNA viral distribution under various contributions (Bertels *et al.* 2022; Burnet *et al.* 2023) such as temperature storage condition and seasonal variability should be considered when choosing the normalization method. Another consideration is the additional cost and effort of measuring non-regulatory parameters when it is possible to make use of the routine monitoring parameters with data already collected by wastewater treatment plants. For health institutions in WBS programs, the normalization of SARS-CoV-2 by faster and more cost-effective, routinely measured wastewater characteristics (total nitrogen, total phosphorus, and ammonia nitrogen) may provide better results in comparison to the use of fecal biomarkers (Isaksson *et al.* 2022; Sakarovitch *et al.* 2022). EC sensor measurements have been widely used in monitoring sewage and sewage dilution, thereby making it an appropriate alternative to flow normalization in wastewater surveillance (Langeveld *et al.* 2023).

3.4. Predicting SARS-CoV-2 in wastewater using random forest regression models

The level of SARS-CoV-2 in sewers is not only influenced by the number of infected persons but potentially also by a variety of characteristics in the wastewater. To investigate our data more in-depth and to complement the PCA analysis, multiple linear and random forest regression models were used to identify variables (wastewater quality measurements and fecal biomarkers) that have strong predictive power for SARS-CoV-2 in wastewater. Based on the selected model's metrics performances, the correlation between measured and predicted SARS-CoV-2 wastewater concentrations was significantly higher using the random forest approach ($R^2 = 0.81$) than the multiple linear regression ($R^2 = 0.51$) (Figure S3). Thus, only the random forest regression model is presented and discussed (Figure 5).

Figure 5 illustrates variable importance and the correlation plots between measured and predicted log-transformed SARS-CoV-2 concentrations in the wastewater for the random forest regression model. The variable importance indicates that the number of confirmed cases is the most important explanatory variable predictor, as expected, and it is followed by water temperature, and to a lesser extent by PMMoV, BRSV recovery, pH, EC, and flow rate. The random forest regression results imply a high confidence in SARS-CoV-2 wastewater surveillance for future confirmed case prediction. Both the PCA and random forest findings indicated water temperature as an important factor contributing to SARS-CoV-2 variations in wastewater. It is to be investigated if the latter is merely an artifact of COVID-19 being a seasonal disease and thus correlated with temperature or if it is related to higher viral decay in sewer systems at higher temperatures or seasonal hydrological differences.

Overall, our measurement and prediction results over a 21-month sampling period suggest that daily quantification of SARS-CoV-2 RNA from 24-h composite samples can be useful for predicting infection trends in the community. The results confirm promising expectations toward the surveillance of other pathogens in wastewater for public health purposes. It should be noted that the proposed random forest model would need to be retrained for each new virus, each new virus variant, each new city, and when infection or sewer shed conditions in a city would change.

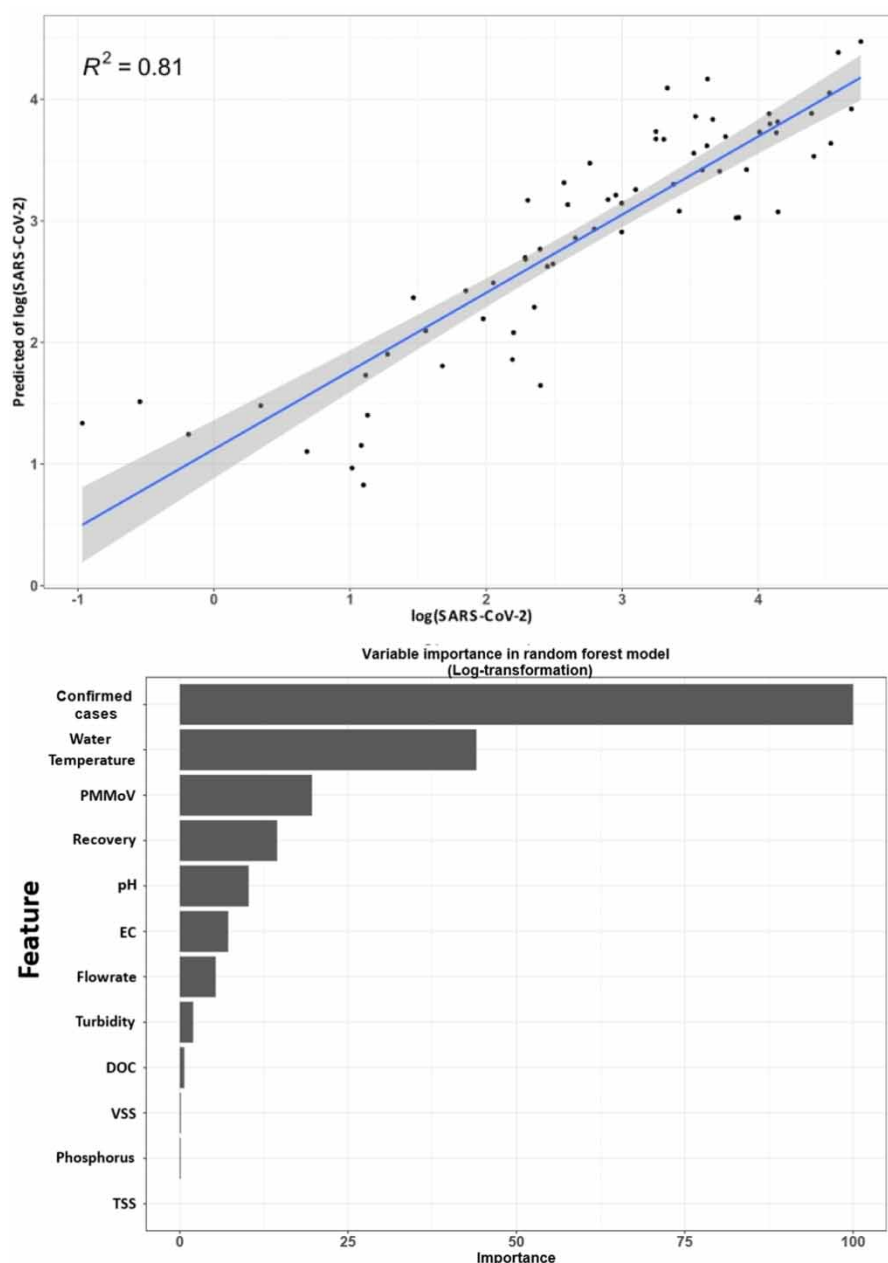


Figure 5 | Correlation of SARS-CoV-2 in wastewater and predicted SARS-CoV-2 using random forest with natural log-transformed SARS-CoV-2 data (top panel) and variable importance plot (bottom) for the random forest model. Training data = 272 observations and testing data = 67 from both storage conditions (-80 and 4 °C).

4. CONCLUSIONS

This study presents 21 months of SARS-CoV-2 wastewater data at the beginning of the pandemic in the City of Montreal, Quebec, for one of the largest wastewater collection systems in the world. The use of various physicochemical parameters and fecal biomarkers for the normalization and prediction of SARS-CoV-2 data was explored to understand impacts on the correlation of SARS-CoV-2 surveillance and clinical data. The following conclusions were drawn:

- SARS-CoV-2 and PMMoV in the North sewage system were higher than in the South sewage system. This difference can be explained by the different populations and number of infected persons in the two interceptors. On the other hand, storage condition also affects the correlation between SARS-CoV-2 and PMMoV in frozen ($R = 0.17$) and fresh ($R = 0.43$) samples.

- The SARS-CoV-2 concentration in summer is significantly lower than in the other seasons, whereas there is no significant difference in PMMoV concentration across the seasons.
- Normalization of SARS-CoV-2 using flow rate and pH did not improve the correlation with clinical data. Combined sewer systems lose a portion of the total flow to receiving waters during snowmelt and intense precipitation events. Normalization of SARS-CoV-2 using EC slightly improves the correlation with clinical data, and EC is a more robust indicator of dilution in large combined sewer systems. Monitoring EC is less labor-intensive than chemical biomarkers and fecal indicators. Besides, EC monitoring could be available for regulatory routine monitoring at the wastewater treatment plant. Therefore, EC normalization can be applied cost-effectively and efficiently to improve the correlation between surveillance and clinical data during COVID-19 and future pandemics.
- Random forest illustrates the high correlation ($R^2 = 0.81$) between predicted SARS-CoV-2 and real SARS-CoV-2 in wastewater. The random forest model is a powerful prediction tool for SARS-CoV-2 in wastewater based on the variation of environmental factors such as EC, flow rate, and water temperature. The number of confirmed cases and water temperature are the most important variables that drive the SARS-CoV-2 concentration in wastewater in long-term monitoring and the prediction of SARS-CoV-2 in wastewater.

ACKNOWLEDGEMENTS

The study was financially supported by the Québec Water Research Center (CentrEau), the Natural Sciences and Engineering Research Council (NSERC) Strategic Grant, the Fonds de Recherche du Québec (FRQ), the Trottier Family Foundation, and the Molson Foundation. We would like to acknowledge all staff of the City of Montreal and the Jean-R. Marcotte Wastewater Treatment Plant members for their valuable contribution to sampling and logistic support.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J. W., Choi, P. M., Kitajima, M., Simpson, S. L., Li, J., Tschärke, B., Verhagen, R., Smith, W. J. M., Zaugg, J., Dierens, L., Hugenholtz, P., Thomas, K. V. & Mueller, J. F. (2020a) *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**, 138764. <https://doi.org/10.1016/j.scitotenv.2020.138764>.
- Ahmed, W., Bertsch, P. M., Bibby, K., Haramoto, E., Hewitt, J., Huygens, F., Gyawali, P., Korajkic, A., Riddell, S., Sherchan, S. P., Simpson, S. L., Sirikanchana, K., Symonds, E. M., Verhagen, R., Vasani, S. S., Kitajima, M. & Bivins, A. (2020b) *Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology*, *Environ. Res.*, **191**, 110092. <https://doi.org/10.1016/j.envres.2020.110092>.
- Ai, Y., Davis, A., Jones, D., Lemeshow, S., Tu, H., He, F., Ru, P., Pan, X., Bohrerova, Z. & Lee, J. (2021) *Wastewater SARS-CoV-2 monitoring as a community-level COVID-19 trend tracker and variants in Ohio, United States*, *Sci. Total Environ.*, **801**, 149757. <https://doi.org/10.1016/j.scitotenv.2021.149757>.
- Amoah, I. D., Abunama, T., Awolusi, O. O., Pillay, L., Pillay, K., Kumari, S. & Bux, F. (2022) *Effect of selected wastewater characteristics on estimation of SARS-CoV-2 viral load in wastewater*, *Environ. Res.*, **203**, 111877. <https://doi.org/10.1016/j.envres.2021.111877>.
- Arabzadeh, R., Grunbacher, D. M., Insam, H., Kreuzinger, N., Markt, R. & Rauch, W. (2021) *Data filtering methods for SARS-CoV-2 wastewater surveillance*, *Water Sci. Technol.*, **84** (6), 1324–1339. <https://doi.org/10.2166/wst.2021.343>.
- Arts, P. J., Kelly, J. D., Midgley, C. M., Anglin, K., Lu, S., Abedi, G. R., Andino, R., Bakker, K. M., Banman, B., Boehm, A. B., Briggs-Hagen, M., Brouwer, A. F., Davidson, M. C., Eisenberg, M. C., Garcia-Knight, M., Knight, S., Peluso, M. J., Pineda-Ramirez, J., Diaz Sanchez, R., Saydah, S., Tassetto, M., Martin, J. N. & Wigginton, K. R. (2023) *Longitudinal and quantitative fecal shedding dynamics of SARS-CoV-2, pepper mild mottle virus, and crAssphage*, *mSphere*, **8** (4), e0013223. <https://doi.org/10.1128/msphere.00132-23>.
- Backer, J. A., Klinkenberg, D. & Wallinga, J. (2020) *Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020*, *Euro. Surveill.*, **25** (5), 2000062. <https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062>.
- Bertels, X., Demeyer, P., Van den Bogaert, S., Boogaerts, T., van Nuijs, A. L. N., Delpitte, P. & Lahousse, L. (2022) *Factors influencing SARS-CoV-2 RNA concentrations in wastewater up to the sampling stage: a systematic review*, *Sci. Total Environ.*, **820**, 153290. <https://doi.org/10.1016/j.scitotenv.2022.153290>.

- Bibby, K., Bivins, A., Wu, Z. & North, D. (2021) Making waves: plausible lead time for wastewater based epidemiology as an early warning system for COVID-19, *Water Res.*, **202**, 117438. <https://doi.org/10.1016/j.watres.2021.117438>.
- Bitter, L. C., Kibbee, R., Jimenez, G. C. & Ormeci, B. (2022) Wastewater surveillance of SARS-CoV-2 at a Canadian university campus and the impact of wastewater characteristics on viral RNA detection, *ACS ES T Water*, **2** (11), 2034–2046. <https://doi.org/10.1021/acsestwater.2c00060>.
- Bonanno Ferraro, G., Suffredini, E., Mancini, P., Veneri, C., Iaconelli, M., Bonadonna, L., Montagna, M. T., De Giglio, O. & La Rosa, G. (2021) Pepper mild mottle virus as indicator of pollution: assessment of prevalence and concentration in different water environments in Italy, *Food Environ. Virol.*, **13** (1), 117–125. <https://doi.org/10.1007/s12560-020-09458-6>.
- Burnet, J. B., Cauchie, H. M., Walczak, C., Goeders, N. & Ogorzal, L. (2023) Persistence of endogenous RNA biomarkers of SARS-CoV-2 and PMMoV in raw wastewater: impact of temperature and implications for wastewater-based epidemiology, *Sci. Total Environ.*, **857** (Pt 2), 159401. <https://doi.org/10.1016/j.scitotenv.2022.159401>.
- Carrillo-Reyes, J., Barragan-Trinidad, M. & Buitron, G. (2021) Surveillance of SARS-CoV-2 in sewage and wastewater treatment plants in Mexico, *J. Water Process Eng.*, **40**, 101815. <https://doi.org/10.1016/j.jwpe.2020.101815>.
- Choi, P. M., Li, J., Gao, J., O'Brien, J. W., Thomas, K. V., Thai, P. K., Jiang, G. & Mueller, J. F. (2020) Considerations for assessing stability of wastewater-based epidemiology biomarkers using biofilm-free and sewer reactor tests, *Sci. Total Environ.*, **709**, 136228. <https://doi.org/10.1016/j.scitotenv.2019.136228>.
- Colson, P., Richet, H., Desnues, C., Balique, F., Moal, V., Grob, J. J., Berbis, P., Lecoq, H., Harle, J. R., Berland, Y. & Raoult, D. (2010) Pepper mild mottle virus, a plant virus associated with specific immune responses, fever, abdominal pains, and pruritus in humans, *PLoS One*, **5** (4), e10041. <https://doi.org/10.1371/journal.pone.0010041>.
- D'Aoust, P. M., Mercier, E., Montpetit, D., Jia, J. J., Alexandrov, I., Neault, N., Baig, A. T., Mayne, J., Zhang, X., Alain, T., Langlois, M. A., Servos, M. R., MacKenzie, M., Figeys, D., MacKenzie, A. E., Graber, T. E. & Delatolla, R. (2021) Quantitative analysis of SARS-CoV-2 RNA from wastewater solids in communities with low COVID-19 incidence and prevalence, *Water Res.*, **188**, 116560. <https://doi.org/10.1016/j.watres.2020.116560>.
- Dhiyebi, H. A., Abu Farah, J., Ikert, H., Srikanthan, N., Hayat, S., Bragg, L. M., Qasim, A., Payne, M., Kaleis, L., Paget, C., Celmer-Repin, D., Folkema, A., Drew, S., Delatolla, R., Giesy, J. P. & Servos, M. R. (2023) Assessment of seasonality and normalization techniques for wastewater-based surveillance in Ontario, Canada, *Front. Public Health*, **11**, 1186525. <https://doi.org/10.3389/fpubh.2023.1186525>.
- Feng, S. C., Roguet, A., McClary-Gutierrez, J. S., Newton, R. J., Kloczko, N., Meiman, J. G. & McLellan, S. L. (2021) Evaluation of sampling, analysis, and normalization methods for SARS-CoV-2 concentrations in wastewater to assess COVID-19 burdens in Wisconsin communities, *Acs Es&T Water*, **1** (8), 1955–1965. <https://doi.org/10.1021/acsestwater.1c00160>.
- Fores, E., Bofill-Mas, S., Itarte, M., Martinez-Puchol, S., Hundesa, A., Calvo, M., Borrego, C. M., Corominas, L. L., Girones, R. & Rusinol, M. (2021) Evaluation of two rapid ultrafiltration-based methods for SARS-CoV-2 concentration from wastewater, *Sci. Total Environ.*, **768**, 144786. <https://doi.org/10.1016/j.scitotenv.2020.144786>.
- Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M. H., Yetka, K., Thompson, H., Keeling, D., Mitchell, J. & Gonzalez, D. (2020) COVID-19 surveillance in Southeastern Virginia using wastewater-based epidemiology, *Water Res.*, **186**, 116296. <https://doi.org/10.1016/j.watres.2020.116296>.
- Government of Canada (2022) *Federal, Provincial, Territorial Public Health Response Plan for Ongoing Management of COVID-19*. Ottawa: Government of Canada.
- Greenwald, H. D., Kennedy, L. C., Hinkle, A., Whitney, O. N., Fan, V. B., Crits-Christoph, A., Harris-Lovett, S., Flamholz, A. I., Al-Shayeb, B., Liao, L. D., Beyers, M., Brown, D., Chakrabarti, A. R., Dow, J., Frost, D., Koekemoer, M., Lynch, C., Sarkar, P., White, E., Kantor, R. & Nelson, K. L. (2021) Tools for interpretation of wastewater SARS-CoV-2 temporal and spatial trends demonstrated with data collected in the San Francisco Bay Area, *Water Res. X*, **12**, 100111. <https://doi.org/10.1016/j.wroa.2021.100111>.
- Hamza, H., Rizk, N. M., Gad, M. A. & Hamza, I. A. (2019) Pepper mild mottle virus in wastewater in Egypt: a potential indicator of wastewater pollution and the efficiency of the treatment process, *Arch. Virol.*, **164** (11), 2707–2713. <https://doi.org/10.1007/s00705-019-04383-x>.
- Haramoto, E., Malla, B., Thakali, O. & Kitajima, M. (2020) First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan, *Sci. Total Environ.*, **737**, 140405. <https://doi.org/10.1016/j.scitotenv.2020.140405>.
- Hasan, S. W., Ibrahim, Y., Daou, M., Kannout, H., Jan, N., Lopes, A., Alsafar, H. & Yousef, A. F. (2021) Detection and quantification of SARS-CoV-2 RNA in wastewater and treated effluents: surveillance of COVID-19 epidemic in the United Arab Emirates, *Sci. Total Environ.*, **764**, 142929. <https://doi.org/10.1016/j.scitotenv.2020.142929>.
- Hrudey, S. E., Bischel, H. N., Charrois, J., Chik, A. H., Conant, B., Delatolla, R., Dorner, S., Graber, T. E., Hubert, C. & Isaac-Renton, J. (2022) Wastewater surveillance for SARS-CoV-2 RNA in Canada, Canadian Science Publishing 1840 Woodward Drive, Suite 1, Ottawa, ON K2C 0P7. DOI: <https://doi.org/10.1139/facets-2022-0148>.
- Hsu, S. Y., Bayati, M., Li, C., Hsieh, H. Y., Belenchia, A., Klutts, J., Zemmer, S. A., Reynolds, M., Semkiw, E., Johnson, H. Y., Foley, T., Wieberg, C. G., Wenzel, J., Johnson, M. C. & Lin, C. H. (2022) Biomarkers selection for population normalization in SARS-CoV-2 wastewater-based epidemiology, *Water Res.*, **223**, 118985. <https://doi.org/10.1016/j.watres.2022.118985>.
- Isaksson, F., Lundy, L., Hedström, A., Székely, A. J. & Mohamed, N. (2022) Evaluating the use of alternative normalization approaches on SARS-CoV-2 concentrations in wastewater: experiences from two catchments in northern Sweden, *Environments*, **9** (3), 39. <https://doi.org/10.3390/environments9030039>.

- Izquierdo-Lara, R., Elsinga, G., Heijnen, L., Munnink, B. B. O., Schapendonk, C. M. E., Nieuwenhuijse, D., Kon, M., Lu, L., Aarestrup, F. M., Lycett, S., Medema, G., Koopmans, M. P. G. & de Graaf, M. (2021) [Monitoring SARS-CoV-2 circulation and diversity through community wastewater sequencing, The Netherlands and Belgium](#), *Emerg. Infect. Dis.*, **27** (5), 1405–1415. <https://doi.org/10.3201/eid2705.204410>.
- Jayme, G., Liu, J. L., Galvez, J. H., Reiling, S. J., Celikkol, S., N'Guessan, A., Lee, S., Chen, S. H., Tsitouras, A., Sanchez-Quete, F., Maere, T., Goitom, E., Hachad, M., Mercier, E., Loeb, S. K., Vanrolleghem, P. A., Dorner, S., Delatolla, R., Shapiro, B. J., Frigon, D., Ragoussis, J. & Snutch, T. P. (2024) [Combining short- and long-read sequencing technologies to identify SARS-CoV-2 variants in wastewater](#), *Viruses*, **16** (9), 1495–1510. 2024.2008. 2007.24311639. <https://doi.org/10.3390/v16091495>.
- Kitajima, M., Sassi, H. P. & Torrey, J. R. (2018) [Pepper mild mottle virus as a water quality indicator](#), *Npj Clean Water*, **1** (1), 1–9. <https://doi.org/10.1038/s41545-018-0019-5>.
- Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B. & Team, R. C. (2020) Package 'caret', *R J.*, **223** (7), 48.
- Kumblathan, T., Liu, Y., Uppal, G. K., Hrudey, S. E. & Li, X. F. (2021) [Wastewater-based epidemiology for community monitoring of SARS-CoV-2: progress and challenges](#), *ACS Environ. Au*, **1** (1), 18–31. <https://doi.org/10.1021/acsenvironau.1c00015>.
- Langeveld, J., Schilperoort, R., Heijnen, L., Elsinga, G., Schapendonk, C. E. M., Fanoy, E., de Schepper, E. I. T., Koopmans, M. P. G., de Graaf, M. & Medema, G. (2023) [Normalisation of SARS-CoV-2 concentrations in wastewater: the use of flow, electrical conductivity and crAssphage](#), *Sci. Total Environ.*, **865**, 161196. <https://doi.org/10.1016/j.scitotenv.2022.161196>.
- LaTurner, Z. W., Zong, D. M., Kalvapalle, P., Gamas, K. R., Terwilliger, A., Crosby, T., Ali, P., Avadhanula, V., Santos, H. H., Weesner, K., Hopkins, L., Piedra, P. A., Maresso, A. W. & Stadler, L. B. (2021) [Evaluating recovery, cost, and throughput of different concentration methods for SARS-CoV-2 wastewater-based epidemiology](#), *Water Res.*, **197**, 117043. <https://doi.org/10.1016/j.watres.2021.117043>.
- Lescure, F. X., Bouadma, L., Nguyen, D., Parisey, M., Wicky, P. H., Behillil, S., Gaymard, A., Bouscambert-Duchamp, M., Donati, F., Le Hingrat, Q., Enouf, V., Houhou-Fidouh, N., Valette, M., Mailles, A., Lucet, J. C., Mentre, F., Duval, X., Descamps, D., Malvy, D., Timsit, J. F., Lina, B., van-der-Werf, S. & Yazdanpanah, Y. (2020) [Clinical and virological data of the first cases of COVID-19 in Europe: a case series](#), *Lancet Infect. Dis.*, **20** (6), 697–706. DOI: [https://doi.org/10.1016/S1473-3099\(20\)30200-0](https://doi.org/10.1016/S1473-3099(20)30200-0).
- Maal-Bared, R., Qiu, Y., Li, Q., Gao, T., Hrudey, S. E., Bhavanam, S., Ruecker, N. J., Ellehoj, E., Lee, B. E. & Pang, X. (2023) [Does normalization of SARS-CoV-2 concentrations by pepper mild mottle virus improve correlations and lead time between wastewater surveillance and clinical data in Alberta \(Canada\): comparing twelve SARS-CoV-2 normalization approaches](#), *Sci. Total Environ.*, **856** (Pt 1), 158964. <https://doi.org/10.1016/j.scitotenv.2022.158964>.
- Mac Mahon, J., Criado Monleon, A. J., Gill, L. W., O'Sullivan, J. J. & Meijer, W. G. (2022) [Wastewater-based epidemiology \(WBE\) for SARS-CoV-2 – a review focussing on the significance of the sewer network using a Dublin city catchment case study](#), *Sci. Total Environ.*, **86** (6), 1402–1425. <https://doi.org/10.2166/wst.2022.278>.
- Mazumder, P., Dash, S., Honda, R., Sonne, C. & Kumar, M. (2022) [Sewage surveillance for SARS-CoV-2: molecular detection, quantification, and normalization factors](#), *Curr. Opin. Environ. Sci. Health*, **28**, 100363. <https://doi.org/10.1016/j.coesh.2022.100363>.
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R. & Brouwer, A. (2020) [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** (7), 511–516. <https://doi.org/10.1021/acs.estlett.0c00357>.
- Molinari, N. A., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E. & Bridges, C. B. (2007) [The annual impact of seasonal influenza in the US: measuring disease burden and costs](#), *Vaccine*, **25** (27), 5086–5096. <https://doi.org/10.1016/j.vaccine.2007.03.046>.
- Moriyama, M., Hugentobler, W. J. & Iwasaki, A. (2020) [Seasonality of respiratory viral infections](#), *Annu. Rev. Virol.*, **7** (1), 83–101. <https://doi.org/10.1146/annurev-virology-012420-022445>.
- Nagarkar, M., Keely, S. P., Jahne, M., Wheaton, E., Hart, C., Smith, B., Garland, J., Varughese, E. A., Braam, A., Wiechman, B., Morris, B. & Brinkman, N. E. (2022) [SARS-CoV-2 monitoring at three sewersheds of different scales and complexity demonstrates distinctive relationships between wastewater measurements and COVID-19 case data](#), *Sci. Total Environ.*, **816**, 151534. <https://doi.org/10.1016/j.scitotenv.2021.151534>.
- N'Guessan, A., Tsitouras, A., Sanchez-Quete, F., Goitom, E., Reiling, S. J., Galvez, J. H., Nguyen, T. L., Loan Nguyen, H. T., Visentin, F. & Hachad, M. (2022) [Detection of prevalent SARS-CoV-2 variant lineages in wastewater and clinical sequences from cities in Québec, Canada](#). medRxiv, 2022.2002. 2001.22270170.
- Oran, D. P. & Topol, E. J. (2020) [Prevalence of asymptomatic SARS-CoV-2 infection: a narrative review](#), *Ann. Intern. Med.*, **173** (5), 362–367. <https://doi.org/10.7326/M20-3012>.
- Paules, C. I., Marston, H. D. & Fauci, A. S. (2020) [Coronavirus infections-more than just the common cold](#), *JAMA*, **323** (8), 707–708. <https://doi.org/10.1001/jama.2020.0757>.
- Pollan, M., Perez-Gomez, B., Pastor-Barriuso, R., Oteo, J., Hernan, M. A., Perez-Olmeda, M., Sanmartin, J. L., Fernandez-Garcia, A., Cruz, I., Fernandez de Larrea, N., Molina, M., Rodriguez-Cabrera, F., Martin, M., Merino-Amador, P., Leon Paniagua, J., Munoz-Montalvo, J. F., Blanco, F. & Yotti, R. & ENE-COVID Study Group (2020) [Prevalence of SARS-CoV-2 in Spain \(ENE-COVID\): a nationwide, population-based seroepidemiological study](#), *Lancet*, **396** (10250), 535–544. [https://doi.org/10.1016/S0140-6736\(20\)31483-5](https://doi.org/10.1016/S0140-6736(20)31483-5).
- Rasmussen, A. L. & Popescu, S. V. (2021) [SARS-CoV-2 transmission without symptoms](#), *Science*, **371** (6535), 1206–1207. <https://doi.org/10.1126/science.abf9569>.

- Robotto, A., Lembo, D., Quaglino, P., Brizio, E., Polato, D., Civra, A., Cusato, J. & Di Perri, G. (2022) Wastewater-based SARS-CoV-2 environmental monitoring for Piedmont, Italy, *Environ. Res.*, **203**, 111901. <https://doi.org/10.1016/j.envres.2021.111901>.
- Sakarovich, C., Schlosser, O., Courtois, S., Proust-Lima, C., Couallier, J., Petrau, A., Litrico, X. & Loret, J. F. (2022) Monitoring of SARS-CoV-2 in wastewater: what normalisation for improved understanding of epidemic trends? *J. Water Health*, **20** (4), 712–726. <https://doi.org/10.2166/wh.2022.012>.
- Sala-Comorera, L., Reynolds, L. J., Martin, N. A., O'Sullivan, J. J., Meijer, W. G. & Fletcher, N. F. (2021) Decay of infectious SARS-CoV-2 and surrogates in aquatic environments, *Water Res.*, **201**, 117090. <https://doi.org/10.1016/j.watres.2021.117090>.
- Sanchez-Quete, F., Chiang, L., Caldera Sequeira, F., Goitom, E., Tsitouras, A., Maere, T., Vanrolleghem, P. A., Dorner, S., Frigon, D. & Loeb, S. K. (2024) Suspended solids and optimal RNase inhibitors impact the partitioning and decay of SARS-CoV-2 in wastewater, *ACS ES&T Water*, **4** (12), 5980–5988. <https://doi.org/10.1021/acsestwater.4c01074>.
- Simpson, A., Topol, A., White, B. J., Wolfe, M. K., Wigginton, K. R. & Boehm, A. B. (2021) Effect of storage conditions on SARS-CoV-2 RNA quantification in wastewater solids, *PeerJ*, **9**, e11933. <https://doi.org/10.7717/peerj.11933>.
- Spurbeck, R. R., Minard-Smith, A. & Catlin, L. (2021) Feasibility of neighborhood and building scale wastewater-based genomic epidemiology for pathogen surveillance, *Sci. Total Environ.*, **789**, 147829. <https://doi.org/10.1016/j.scitotenv.2021.147829>.
- Staley, Z. R., He, D. D. & Edge, T. A. (2017) Persistence of fecal contamination and pathogenic O157:H7 in snow and snowmelt, *J. Great Lakes Res.*, **43** (2), 248–254. <https://doi.org/10.1016/j.jglr.2017.01.006>.
- Tolouei, S., Burnet, J. B., Autixier, L., Taghipour, M., Bonsteel, J., Duy, S. V., Sauve, S., Prevost, M. & Dorner, S. (2019) Temporal variability of parasites, bacterial indicators, and wastewater micropollutants in a water resource recovery facility under various weather conditions, *Water Res.*, **148**, 446–458. <https://doi.org/10.1016/j.watres.2018.10.068>.
- Tomasino, M. P., Semedo, M., Vieira, E. M. P., Ferraz, E., Rocha, A., Carvalho, M. F., Magalhaes, C. & Mucha, A. 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>.
- Vadde, K. K., Al-Duroobi, H., Phan, D. C., Jafarzadeh, A., Moghadam, S. V., Matta, A. & Kapoor, V. (2022) Assessment of concentration, recovery, and normalization of SARS-CoV-2 RNA from two wastewater treatment plants in Texas and correlation with COVID-19 cases in the community, *ACS ES T Water*, **2** (11), 2060–2069. <https://doi.org/10.1021/acsestwater.2c00054>.
- Williams, R. C., Perry, W. B., Lambert-Slosarska, K., Fitcher, B., Pellett, C., Paterson, S., Grimsley, J. M., Wade, M. J., Weightman, A. J. & Jones, D. L. (2024) Examining the stability of viral RNA and DNA in wastewater: effects of storage time, temperature, and freeze-thaw cycles, *Water Res.*, **259**, 121879.
- Wu, F., Zhang, J., Xiao, A., Gu, X., Lee, W. L., Armas, F., Kauffman, K., Hanage, W., Matus, M., Ghaeli, N., Endo, N., Duvallet, C., Poyet, M., Moniz, K., Washburne, A. D., Erickson, T. B., Chai, P. R., Thompson, J. & Alm, E. J. (2020) SARS-CoV-2 titers in wastewater are higher than expected from clinically confirmed cases, *mSystems*, **5** (4), e00614–e00620. <https://doi.org/10.1128/mSystems.00614-20>.
- Zhang, T., Breitbart, M., Lee, W. H., Run, J. Q., Wei, C. L., Soh, S. W., Hibberd, M. L., Liu, E. T., Rohwer, F. & Ruan, Y. (2006) RNA viral community in human feces: prevalence of plant pathogenic viruses, *PLoS Biol.*, **4** (1), e3. <https://doi.org/10.1371/journal.pbio.0040003>.

First received 4 October 2024; accepted in revised form 13 January 2025. Available online 21 January 2025