Artificial neural network-based estimation of COVID-19 case numbers and effective reproduction rate using wastewater-based epidemiology

Guangming Jiang [a,b,*], Jiangping Wu [a], Jennifer Weidhaas [c], Xuan Li [a], Yan Chen [a], Jochen Mueller [d], Jiaying Li [d], Manish Kumar [e], Xu Zhou [f,g], Sudipti Arora [h], Eiji Haramoto [i], Samendra Sherchan [j], Gorka Orive [k,l], Unax Lertxundi [k,l], Ryo Honda [m], Masaaki Kitajima [n], Greg Jackson [o]

[a] School of Civil, Mining and Environmental Engineering, University of Wollongong, Australia
[b] Illawarra Health and Medical Research Institute (IHMRI), University of Wollongong, Wollongong, Australia
[c] University of Utah, Civil and Environmental Engineering, 110 Central Campus Drive, Suite 2000, Salt Lake City, UT, USA
[d] Queensland Alliance for Environmental Health Sciences, The University of Queensland, Australia
[e] Sustainability Cluster, School of Engineering, University of Petroleum & Energy Studies, Dehradun, Uttarakhand 248007, India
[f] Shenzhen Engineering Laboratory of Microalgal Bioenergy, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China
[g] State Key Laboratory of Urban Water Resource and Environment, School of Environment, Harbin Institute of Technology, Harbin 150090, China
[h] Dr. B. Lal Institute of Biotechnology, Jaipur, India
[i] Interdisciplinary Center for River Basin Environment, University of Yamanashi, Kofu, Japan
[j] Department of Environmental Health Sciences, Tulane University, New Orleans, LA, USA
[k] NanoBioCel Group, Laboratory of Pharmaceutics, School of Pharmacy, University of the Basque Country UPV/EHU, Paseo de la Universidad 7, Vitoria-Gasteiz 01006, Spain
[l] Biomedical Research Networking Centre in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Vitoria-Gasteiz, Spain
[m] Faculty of Geosciences and Civil Engineering, Kanazawa University, Kanazawa 920-1192, Japan
[n] Division of Environmental Engineering, Hokkaido University, Hokkaido 060-8628, Japan
[o] Queensland Alliance for Environmental Health Sciences (QAEHS), The University of Queensland, 4102, Brisbane, Australia

ARTICLE INFO

ABSTRACT

As a cost-effective and objective population-wide surveillance tool, wastewater-based epidemiology (WBE) has been widely implemented worldwide to monitor the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA concentration in wastewater. However, viral concentrations or loads in wastewater often correlate poorly with clinical case numbers. To date, there is no reliable method to back-estimate the coronavirus disease 2019 (COVID-19) case numbers from SARS-CoV-2 concentrations in wastewater. This greatly limits WBE in achieving its full potential in monitoring the unfolding pandemic. The exponentially growing SARS-CoV-2 WBE dataset, on the other hand, offers an opportunity to develop data-driven models for the estimation of COVID-19 case numbers (both incidence and prevalence) and transmission dynamics (effective reproduction rate). This study developed artificial neural network (ANN) models by innovatively expanding a conventional WBE dataset to include catchment, weather, clinical testing coverage and vaccination rate. The ANN models were trained and evaluated with a comprehensive state-wide wastewater monitoring dataset from Utah, USA during May 2020 to December 2021. In diverse sewer catchments, ANN models were found to accurately estimate the COVID-19 prevalence and incidence rates, with excellent precision for prevalence rates. Also, an ANN model was developed to estimate the effective reproduction number from both wastewater data and other pertinent factors affecting viral transmission and pandemic dynamics. The established ANN model was successfully validated for its transferability to other states or countries using the WBE dataset from Wisconsin, USA.

* Corresponding author at: School of Civil, Mining and Environmental Engineering, University of Wollongong, Australia
E-mail addresses: gjiang@uow.edu.au, gjiang@uow.edu.au (G. Jiang).

1. Introduction

Wastewater-based epidemiology (WBE) was mostly used to monitor the human use or exposure to chemicals by analyzing marker compounds in influents of wastewater treatment plants (WWTPs) or sewer catchments (Choi et al., 2020; EMCDDA, 2016; Gao et al., 2018; Gonzalez-Marino et al., 2020; He et al., 2021; Li et al., 2019; van Nuijs et al., 2011; Zheng et al., 2019; Zuccato E, 2005). It has also been developed and applied as a population-wide surveillance tool for estimating the prevalence of infectious diseases such as poliovirus and hepatitis A virus (Asghar et al., 2014; Hellmér et al., 2014). During the coronavirus disease 2019 (COVID-19) pandemic, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was shed into sewers from feces, urine, saliva, sputum, and other potential sources (Li et al., 2022; Li et al., 2021c; Pan et al., 2020; van Doorn et al., 2020). Thus, WBE found its ideal application as an efficient approach to identify COVID-19 prevalence in communities connected with sewer systems.

At different stages of the COVID-19 outbreak, WBE has been widely used to provide early warning or estimate the community prevalence or incidence of COVID-19 (Acosta et al., 2021; Agrawal et al., 2021; D'Aoust et al., 2021; Gibas et al., 2021; Kitamura et al., 2021; Li et al., 2021b; Medema et al., 2020; Nemudryi et al., 2020; Róka et al., 2021; Rusiñol et al., 2021; Scott et al., 2021; Sherchan et al., 2020; Westhaus et al., 2021). Through quantifying SARS-CoV-2 RNA in wastewater, estimation of individuals shedding the virus was conducted in different WWTPs, sewer catchments, university campuses, buildings, and transportation vessels (Ahmed et al., 2020a; Betancourt et al., 2021; Rusiñol et al., 2021; Wong et al., 2021; Zhang et al., 2022). As a cost-effective, objective and population-wide surveillance tool, the advantage of WBE is its capability in capturing asymptomatic and presymptomatic patients no matter whether they are tested for COVID-19 infection or not.

The quantification of COVD-19 case numbers through WBE involves five steps: 1) virus shedding by ill individuals; 2) evaluation of virus loss or dilution from in-sewer processes; 3) sample collection, transport, and storage; 4) analysis of SARS-CoV-2 RNA concentrations; and 5) back-estimation of case numbers. For the last step, it is usually conducted using equation 1 by taking into consideration of viral shedding, in-sewer decay, and wastewater flow.

$$P_{\text{COVID}} = \frac{C_{\text{RNA}} \times Q \times e^{k \times t}}{Pop \times C_{\text{S}} \times Q_{\text{S}} \times P_{\text{S}}} \tag{1}$$

where, P_{COVID} is the number of COVID-19 cases within the sewer or WWTP catchment boundary and C_{RNA} is viral RNA copies per liter of wastewater sample (gene copies/L). Among others, there are parameters related to shedding, in-sewer decay, and catchment properties. C_{S} is viral shedding per gram of feces (gene copies/g). Q_{S} is the daily shedding amount of feces of an individual (g/(day•person)); and P_{S} is the shedding probability in feces from an infected person (unitless). k is the in-sewer decay rate constant (day^{-1}) and t is the hydraulic retention time or in-sewer travel time (day). Q is the daily wastewater flow rate (L/day); and Pop is the population of the WWTP or sewer catchment (person). Some WBE programs reported the normalized viral concentration $\frac{C_{\text{RNA}} \times Q}{Pop}$ as the per capita viral load (gene copies/(day•person)).

The accuracy of WBE in quantifying COVID-19 case numbers can be assessed based on the correlation between C_{RNA} (or viral load) and clinically confirmed cases (Nemudryi et al., 2020; Róka et al., 2021). However, to inform the pandemic management and policymaking, different parameters of prevalence or incidence, i.e. daily new cases, future daily new cases, rolling average cases, future average cases, total active cases, and effective reproductive number (also known as effective reproduction rate, R_i) etc., were adopted and the WBE performance in estimating these epidemiological parameters varied greatly in different studies (D'Aoust et al., 2021; Huang et al., 2021; Huisman et al., 2021a; Huisman et al., 2021b; Róka et al., 2021; Weidhaas et al., 2021). The

usually conflicting and variable observations in literature imply extra factors need to be considered in addition to the C_{RNA} and those parameters as dictated in conventional WBE back-estimation (Eq.1).

Although there are advances in improving the sampling and storage, concentration, RNA extraction and qPCR analysis, the accurate back-estimation of COVID-19 case numbers from C_{RNA} or viral load is still a big challenge. Many WBE studies only captured a subset of all the parameters required for Eq. 1. Catchment parameters such as sewer hydraulic retention time (HRT) can have a big impact on the decay of SARS-CoV-2 and its RNA in wastewater, and hence on its concentration in the downstream sampling points (Ahmed et al., 2020b; Bivins et al., 2020; Shi et al., 2021). It is challenging to calculate the HRT of a catchment. Alternative metrics linked to the catchment size (and consequently HRT), such as population and daily wastewater flow rate, should be adopted in the WBE back-estimation of COVID-19 infections. With the fast evolution and mutation of SARS-CoV-2, the shedding pattern and magnitudes of distinct current variants can change as well (Despres et al., 2021). Meanwhile, the progressive expansion of vaccination coverage can also add uncertainty to the back-estimation. Environmental and weather conditions such as precipitation and air temperature have been shown to play a role in the WBE back-estimation of COVID-19 case numbers (Li et al., 2021b).

Previous WBE studies correlated C_{RNA} with the clinically confirmed COVID-19 case numbers, because the actual number of infections was not available. However, clinical testing is known to only capture a fraction of the infections as it is nearly impossible to achieve a 100% test rate (Fernandez-Cassi et al., 2021; Reese et al., 2020). This makes it challenging to validate the WBE estimated case numbers, which were intrinsically objective and not impacted by clinical testing rate. The clinical testing rate could not be included in the traditional WBE. There is an urgent need to improve the WBE back-estimation for the inclusion of more parameters, which are relevant and specific to the COVID-19 clinical testing, viral transmission, and vaccination rollout.

However, the current understanding about the impacts of these factors on the WBE estimation is limited. Our recent study identified and quantified various uncertainties associated with the WBE application for COVID-19 monitoring (Li et al., 2021c). Due to the lack of deterministic models to simulate the complex processes involved in WBE, black-box or data-driven models were proposed to be used to obtain estimates of COVID-19 case numbers with reasonable accuracy (Li et al., 2021b). The artificial neural network (ANN) modelling approach, inspired by biological neural systems, is an effective modelling tool (Krogh, 2008). ANN models are trained with past data to learn the patterns of the underlying process and generalize mathematical relationships between input and output data. It has the potential to predict any complex system with high precision provided its architecture and parameters are properly set.

Machine learning, including ANN models, has found many different applications in studies of the COVID-19 pandemic, including the clinical diagnosis using blood test and chest X-ray (Brinati et al., 2020; Brunese et al., 2020; Mohammad-Rahimi et al., 2021), interactions of human mobility (transportation), air quality and COVID-19 transmission (Asad et al., 2021; Rahman et al., 2021), the forecast or early detection of outbreaks and pandemic dynamics (Allam et al., 2020; Braga et al., 2021; Shawaqfah and Almomani, 2021; Wieczorek et al., 2020). The applications of machine learning and artificial intelligent for the COVID-19 pandemic were reviewed for their potentials in treatment, medication, screening, prediction, forecasting, contact tracing, clinical trials, and drug/vaccination process (Lalmuanawma et al., 2020; Mottaqi et al., 2021). A recent study used random forest method to predict the daily COVID-19 cases based on square root of viral concentration in wastewater, with or without normalization to chemical oxygen demand, for two wastewater treatment plants (Koureas et al., 2021). The machine learning models showed improvement in comparison to linear regression models, which also have high correlation coefficients between wastewater measurements and cumulative cases around 0.8-0.9 in this study. The wider uses and benefits of various machine learning

techniques in WBE was proposed in anticipation of providing accurate viral outbreaks detection and early warning of hotspots (Abdeldayem et al., 2022; Matheri et al., 2022). Among different machine learning techniques, ANN can provide both regression (continuous quantity) and classification (discrete class label) predictions. Thus, ANN regression models are suitable for the prediction of COVID-19 case numbers. However, there is still a lack of artificial intelligence models developed using large datasets obtained from a number of diverse WWTPs or catchments.

This study aims to use ANN as an alternative to the conventional WBE back-estimation, because no deterministic association can be established based on existing limited understanding of COVID-19. WBE has been widely implemented worldwide in many countries, and there is a large and exponentially increasing dataset available for the training of ANN-based WBE estimation models. However, not all WBE programs determine and record all crucial and relevant data required as inputs for the WBE back-estimation, including conventional WBE data and those linked to catchment, weather, clinical testing and vaccination. As a result, we chose to collect all relevant data over 1.5 years (May 2020 – December 2021) using the Utah state, USA as a case study. Subsequently, we utilized this extensive dataset to build ANN models for the accurate estimation of COVID-19 case numbers in a WBE approach. The performance and application of the proposed ANN model were also thoroughly evaluated for the estimation of different epidemiological parameters including prevalence rate, incidence rate, and effective reproduction rate.

2. Materials and methods

2.1. Wastewater monitoring program in Utah, USA

Wastewater samples were collected by Utah Department of Environmental Quality (DEQ) at 47 wastewater treatment plants (WWTPs) statewide in Utah (Table SI-1), representing approximately 80% of the state's population. These treatment plants were selected by Utah's DEQ and Department of Health (DOH) by considering the facility size, community susceptibility to new infections and the community Health Improvement Index. The number of treatment plants monitored has changed over time due to funding availability and willingness of the utilities to collect and ship samples. Some of the remote or less populated sites were dropped over time. In other cases, smaller communities were put back into the sampling plan if they were tourist destinations and during summer or winter vacation time.

Wastewater was collected weekly at 47 WWTPs as 24-hr flow weighted composite samples. If results indicated a sharp increase, additional follow-up samples were collected within the constraints of sample capacity. The SARS-CoV-2 RNA was extracted from 100 mL of pasteurized wastewater collected in May 2020 to July 2021 by a uniform method developed by the University of Utah, Utah State University, and Brigham Young University (Weidhaas et al., 2021). The SARS-CoV-2 RNA extracted from 40 mL of pasteurized wastewater collected from July 2021 to December 2021 by the Utah Department of Public Health Laboratory. The 40 mL wastewater samples were concentrated and extracted using the Promega Wizard ® Enviro TNA Kit following the manufacturer's instructions. All samples from May 2020 to December 2021 included, recovery controls added prior to RNA concentration, replicate sample extractions, replicate qPCR runs, and positive and negative qPCR controls. Additional details on quality control and quality assurance samples were previously reported (Weidhaas et al., 2021). Raw virus concentrations, after being normalized to the number of people living in the sewer catchments and wastewater flow rates, were reported as per capita viral load (SARS-CoV-2 gene copies per person per day).

2.2. Data collection for the ANN model development

The full dataset includes parameters related to viral load, catchment, weather, clinical testing, vaccination, case numbers and transmission as shown in Table 1. The catchment data (location, population, wastewater flow rates) of the sampling sites were provided on the SARS-CoV-2 Sewage Monitoring website by the Utah Department of Environmental Quality (https://deq.utah.gov/water-quality/sars-cov-2-sewage-monitoring). The viral load in wastewater together with the daily COVID-19 new case number were downloaded from the same website for each sampling site till the most recent date (6 December 2021). The rolling and future averages of 3-day, 7-day and 14-day were calculated for each wastewater sampling date, considering the case identification lead time by clinical test, wastewater sampling frequency and the duration of viral shedding. The daily new cases within the coming 7 days of the wastewater sampling date were also extracted to evaluate the early-warning capacity of WBE. The daily new cases and rolling or future averages represent the COVID-19 incidence (new cases) and prevalence (accumulated cases) rates of the sampled sewer catchments, respectively.

Corresponding to each wastewater sampling date, the historical weather data of the sampling location was obtained from https://www.wunderground.com/ to determine the precipitation and various air temperatures (daily average, monthly average, and yearly range). Using the air temperatures and wastewater flow rate, the wastewater temperature on the sampling day was calculated according to the method by Hart and Halden (2020). Briefly, the soil temperature was first calculated based on the air temperature and time of the year. Wastewater temperature was then calculated from soil and air temperature for an initial estimate of domestic wastewater discharge temperature (17.8–31.2°C) based on an assumed range of 25–75% hot water and temperatures of 13°C and 50°C for unheated and heated indoor water. The clinical test rates were obtained from Utah governmental COVID-19 data website (https://coronavirus.utah.gov/). Other COVID-19 related parameters for clinical test positive rate, vaccination ratios, and effective reproduction rate (effective reproduction rate) were extracted from https://covidactnow.org.

2.3. Statistical data analysis

The full dataset was first checked for its consistency based on general knowledge of wastewater systems and WBE. Then, it was analyzed for its basic statistics such as minimum, maximum, mean, and standard deviation to identify any suspicious data for each parameter listed in Table 1. Statistical Analysis of the data was performed using R (ver 4.1.2, http://www.R-project.org/). Histograms of each parameter were plotted to check the data distribution. Box plots of each parameter were plotted to check the symmetricity or skewness of the distribution of each parameter. Irregularly distributed parameters were identified as potential sources for low model quality. Following that, scatter plots between all feature and target parameters were plotted to identify the dependencies of the targets with the features as inputs.

The inputs and inputs-targets correlations were calculated to identify the relationship between all parameters. Specifically, the correlation coefficients between all inputs and all targets indicate dependencies between single input and single target in the dataset. Multiple linear regression analysis was then performed on the COVID-19 case numbers (ANN targets), with WBE relevant factors (ANN features). The coefficients for each of the ANN features were determined together with the standard error and significance value.

2.4. Artificial neural network models

The ANN model was designed with three layers: input, hidden layer and output layers. The neural network modelling process used in this study may be described in three steps: (i) pre-processing of the original data set (determination of test positive ratio, calculation of prevalence

Table 1
Parameters of collected COVID and WBE data from the Utah state (USA), and their use in the ANN models.

Category	Symbol	Type	ANN use[*]	Definition	Units
Wastewater analysis	Date	Date/time	-	Wastewater sampling date	dd/mm/yyyy
	VL	Numeric	F	SARS-CoV-2 viral load in wastewater	MGC/person/day
	ST	Categorical	-	Wastewater sampling technique	-
Catchment	Loc	Categorical	-	The sewer or wastewater treatment plant for wastewater sampling	-
	Pop	Numeric	F	population	person
	ADWF	Numeric	F	Average dry weather flow	ML/day
Weather	Prain	Numeric	F	Daily precipitation at the sampling location	mm
	Tair	Numeric	F	Average daily air temperature	ºC
	Twater	Numeric	F	Average daily wastewater temperature	ºC
Clinical test	TR	Numeric	-	Ratio of population being tested clinically on the sampling date [Δ]	-
	TPR	Numeric	F	Positivity ratio of clinical tests	-
Vaccination	Vcr	Numeric	F	The ratio of completed vaccination (2 injections)	-
	Vir	Numeric	F	The ratio of initiated vaccination (1 injection)	-
Case numbers	P1, P2, P3, …, P7	Numeric	T	Daily new cases per 100, 000 population for the WWTP on the 1st, 2nd, 3rd, …, 7th day since the wastewater sampling date	Case/100,000 person
	P3d, P7d, P14d, P3dF, P7dF, P14dF	Numeric	T	3-day, 7-day and 14-day rolling and future average of daily new cases per 100, 000 population of the wastewater sampling date	Case/100,000 person
Effective reproduction rate	R_i	Numeric	T	The COVID-19 effective reproduction rate, which represents how fast COVID is spreading in a given area by estimating the number of people that a newly infected person goes on to eventually infect.	-

[*] F and T indicate the data were used as features (input) and targets (response) of the ANN models, respectively. Some reserved parameters are indicated as "-".
[Δ] Two sets of test ratios and test positive ratios were collected, one for the Utah state and one for the counties. The state-level data is applied when the county-level data is not available. The clinical test ratio and positive ratio of specific wastewater catchment was determined by its overlapping with the county boundaries if possible.

rates and wastewater temperature from raw data, and identification of outliers); (ii) partitioning of the pre-processed data set into learning, validation, and test sets; (iii) ANN model architecture setting, training, testing and validation.

In the first stage the pre-processed data set was sorted by location and sampling date, and the training, validation, and test data were constructed using a randomization procedure. The percentage of observations per data set was assigned to be 70%, 15% and 15%, for the training, validation and test sets, respectively. The training data set was used to train the ANN in Matlab R2021b. The validation data set was used in conjunction with the learning data set to determine when to stop the training process such that the resulting model exhibited good generalization properties. The test data set allows the assessment of the prediction capabilities of the ANN model.

Step 3 involves the ANN architecture setting and optimization. The nodes/neurons for the ANN input and output layer were set by the number of inputs and targets (Table 1), respectively. The number of neurons in the hidden layer was established before the ANN model architecture was completed. The optimal number of neurons in the hidden layer was determined using the exhaustive search function with test error as the fitness criteria in Alyuda NeuroIntelligence ver 2.2. The best architecture was then constructed for the training and validation analysis using Matlab R2021b. The ANN model was systematically evaluated for their performance in predicting incidence, prevalence, and effective reproduction rate as described in Section 2.5.

2.5. Evaluation of ANN model performance and transferability

To determine how the ANN-based WBE back-estimation can predict the daily new cases (up to 7 days from the wastewater sampling date), 7 testing scenarios (ANN-IR, as shown in Table 2) for incidence rates, i.e., P1, P2, …, P7, respectively, were evaluated by comparing their ANN performance using the correlation coefficients (R) and mean squared error (MSE) between ANN model predictions and targets (i.e., clinical testing confirmed case numbers). The R and MSE were calculated for the whole dataset incorporating training and test groups, which showed similar results as overfitting was avoided by using the Bayesian regularization as the optimization procedure. As each training session starts with different initial weights and biases, and different divisions of data

Table 2
ANN model structures for the evaluation of different capacities in estimating the COVID-19 epidemiological parameters.

Group	No. of scenarios	ANN features	ANN targets
Prediction of incidence rate (ANN-IR)	7	Pop, ADWF, VL, Twater, Prain, Tair, TPR, Vcr, Vir	P1, P2, P3, …, P7
Prediction of prevalence rate (ANN-PR)	6	Pop, ADWF, VL, Twater, Prain, Tair, TPR, Vcr, Vir	P3d, P7d, P14d, P3dF, P7dF, P14dF
Prediction of effective reproduction rate (ANN-R_i)	1	Pop, ADWF, VL, Twater, Prain, Tair, TPR, Vcr, Vir	R_i
Contributions of inputs	8	All represents the complete sets of input data. Different combinations among the categories of weather (W), clinical testing (T), and vaccination (V). The eight scenarios are All, All-V, All-T, All-W, All-V-T, All-V-W, All-T-W, All-V-T-W.	Incidence, prevalence rate and R_i

into training, validation, and test sets, ten ANN models were trained to ensure a well generalized network was achieved. The ten networks were used to obtain the average and 95% confidence interval of R and MSE for each scenario. The same was applied for all the other evaluation scenarios described in Table 2.

Six more scenarios (ANN-PR) were tested for the ANN performance in predicting the 3d, 7d and 14d rolling average (P3d, P7d, P14d) and future average (P3dF, P7dF, P14dF) of daily new cases of the wastewater sampling date. These case numbers represent the COVID-19 prevalence rate. In addition, one ANN model (ANN-R_i) was trained for the prediction of effective reproduction rate. Finally, to determine the contribution of clinical test positive ratio, vaccination coverage and weather data to improve the ANN model for WBE back-estimation, eight more scenarios were tested to identify the improvement of ANN performance.

To demonstrate the transferability of the established ANN model, similar WBE dataset as that from Utah, USA was collected from Wisconsin, USA (https://www.dhs.wisconsin.gov/covid-19/wastewater.htm). The same ANN model was employed in estimating COVID-19 incidence, prevalence and effective production rate. The model performance was thus determined using the same procedures for Utah dataset, as described above. Also, the developed ANN models were further validated with literature data and some data collected from different countries through the WATMOC network (Asia-Pacific Network for Wastewater Monitoring of COVID-19, www.watmoc.com).

3. Results and Discussion

3.1. Statistical analysis of ANN inputs

Histograms and box plots of the Utah dataset, both ANN features and targets, were plotted to check the data distribution (Figure SI-1 and SI-2). Some of the feature data is non-symmetric, including Pop, Flow, VL, and case numbers (P1, P2, ..., P7 and P3d, P7d, P14d, P3dF, P7dF and P14dF). Their distributions are skewed with high distribution towards the low range and some outliers in the high range. In contrast, weather (Tair, Twater and Prain), clinical test positive ratio (TPR), vaccination rates (Vcr, Vir) and effective reproduction rate (R_i) showed more symmetric distributions. No outliers were eliminated as the distribution is intrinsic due to the nature of the data. Instead, the irregularly distributed parameters were identified as potential sources for reduced ANN model quality and poor estimates.

Scatter plots (Figure SI-3) and correlation matrix between all ANN features and representative targets (P4, P14d and R_i) were used to identify the dependencies of the targets with the inputs. There are some obviously high positive correlations between Flow and Pop, Twater and Tair, Vir and Vcr, and pairs of prevalence rates (P3d, P7d, P14d, P3dF, P7dF and P14dF) (Fig. 1). All correlations of single feature parameter with incidence rate (P4) are significant, and in the rank of TPR, Twater,

Fig. 1. Correlations between all ANN input features and targets (P4 and P14d, representing incidence and prevalence rate, respectively, and R_i) in the WBE datasets obtained in Utah, USA. The numbers and circle sizes indicate the correlation coefficient; and blank cells indicate insignificant correlations by a cut-off p=0.01.

Tair, VL, Vcr, Vir, Flow, Pop and Prain. For prevalence rate, all correlations of single features are significant except Prain, and the absolute correlation coefficients are in the same order as that for incidence rate. For the effective reproduction rate (R_i), vaccination rates showed the highest correlation, i.e., 0.51 and 0.5 for Vcr and Vir, respectively. The weather, i.e., Tair and Twater (as it is mostly determined by Tair), also showed correlation at 0.29 and 0.33, respectively. This confirmed the relationship of the COVID-19 dynamic of case numbers with vaccination rollout and the weather. The results confirmed the selected features are relevant and contribute to the prediction of targets. Thus, no feature selection or reduction is needed, and the ANN models were developed using the full set of features as inputs. Instead, a thorough evaluation of contribution from different categories of features was conducted as described in Section 2.5.

Multiple linear regression analysis indicated a limited performance for prevalence rate (r^2=0.63) and fairly poor performance for incidence rate and effective reproduction rate, with r^2 of 0.31 and 0.42, respectively (Table SI-2). The results imply that only 63%, 31% and 42% of the variability in incidence (P4), prevalence (P14d) and effective reproduction rate (R_i), respectively, could be captured and explained by the multiple linear models. The low r^2 values suggest that the relationship between the features and targets is unlikely linear.

3.2. Back-estimation of various COVID-19 case numbers using WBE

Following the statistical analysis of the COVID-19 dataset, including data of wastewater, catchment, weather, clinical testing and vaccination, ANN models were trained using the same dataset for its performance in back-estimate COVID-19 case numbers, including both prevalence and incidence rates. The final structure of the ANN model has the hidden layer with its number of neurons determined by architecture search (SI Table x). The activation functions for the hidden and output layers of the ANN model were hyperbolic tangent and logistic function, respectively. Sum of squares was used as the error function for the output layer. The training process was conducted using the Bayesian regularization as the optimization procedure to avoid overfitting. The ANN was trained to a converged state when the sum-squared error, the sum-squared weights, and the effective number of parameters reached constant values.

Using the full set of input parameters, ANN models showed acceptable performance in determining the COVID-19 incidence rate, i.e. daily new cases within one week of the wastewater sampling date. The correlation coefficients (R) and *MSE* between ANN estimates and clinical tested positive cases were between 0.61 and 0.79, 524 and 1056, respectively (Fig. 2A). Especially, it is clear that ANN-IR models gave the best estimates of case numbers for P4, with a high R and a low *MSE* among all scenarios of P1 to P7. This indicates that WBE approach likely allows about 4 days of early predication of future daily new cases, for the Utah sewage monitoring program. The reported early prediction capacity of WBE was between 2-24 days (Ai et al., 2021; Barrios et al., 2021; Nemudryi et al., 2020; Róka et al., 2021; Rusiñol et al., 2021; Sangsanont et al., 2021). However, the leading time of prediction identified in this study by a modelling approach is more reliable as it was determined based on nearly 1.5 years of data from a comprehensive sewage monitoring program covering sewer catchments of various sizes, in comparison to other reports based on very limited data from few catchments. It is also noted that the ANN estimations are relatively scattered due to the various conditions affecting the daily new cases reported by clinical testing (Fig. 2B). As shown in Fig. 3, the ANN-IR model mostly underestimated the case numbers when they are in the high ranges (peaks). In contrast, the estimations were close to clinically reported cases when the COVID-19 cases were in the medium or low range, i.e. <100 per 10,000 people. This also leads to the relatively limited performance of the ANN-IR model in predicting COVID-19 incidence rates. The reason is probably due to the intrinsically skewed input data of case numbers, as discussed in Section 3.1, being used to
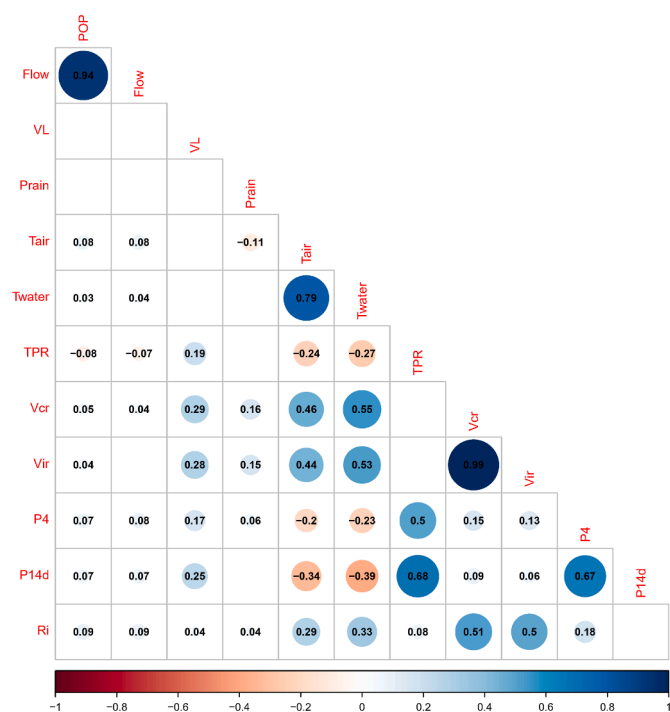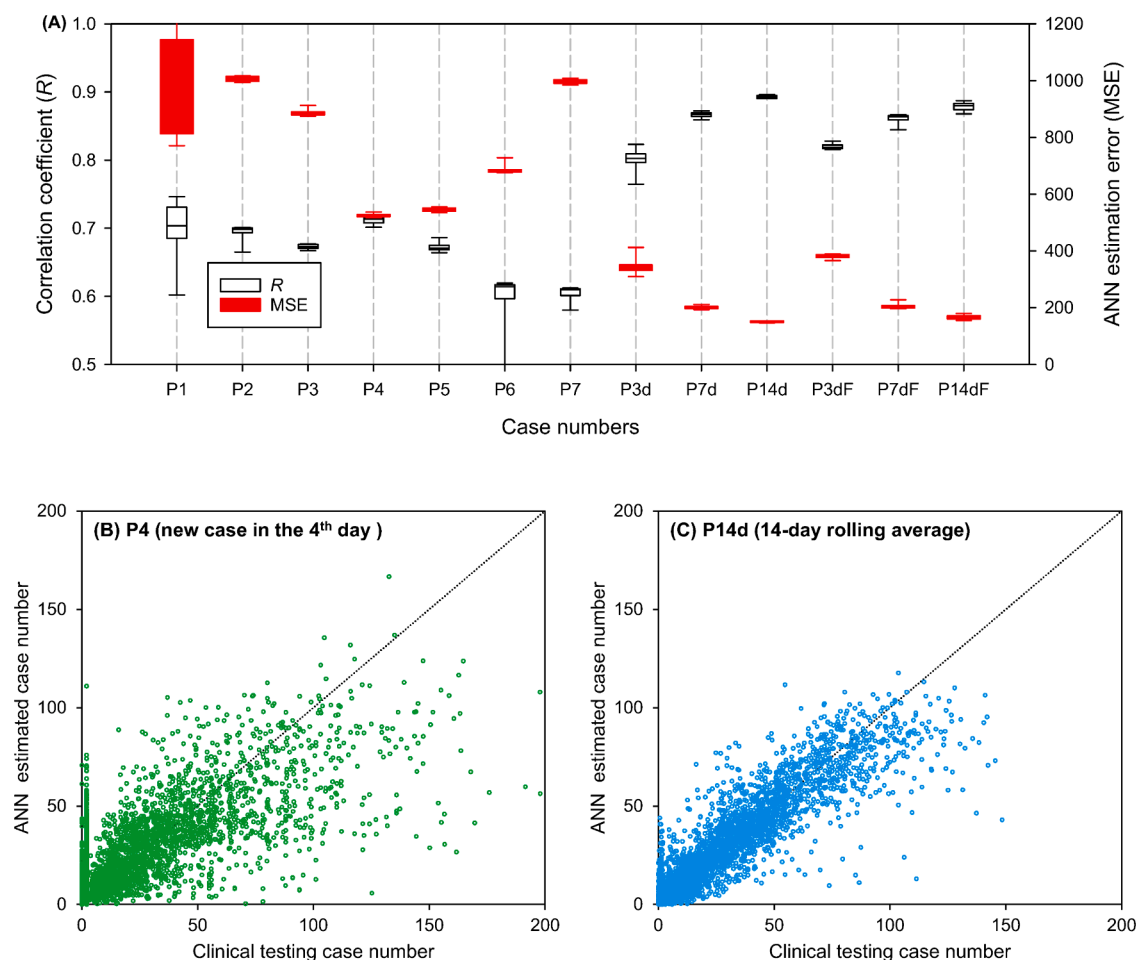
Fig. 2. (A) Box plots of the correlation coefficient (R) and estimation error (MSE) of ANN models being trained to predict incidence rates, i.e., P1, P2, ..., P7 (ANN-IR) and prevalence rates, i.e., P3d, P7d, P14d and P3dF, P7dF and P14dF (ANN-PR) in Utah, USA. (B) ANN outputs vs. clinical testing reported incidence rate, i.e. case numbers on the 4th day of the wastewater sampling date. (C) ANN outputs vs. the prevalence rate reported by clinical test, i.e., the 14-day running average of the wastewater sampling date.

train the ANN-IR model.

For the estimation of prevalence rates by ANN-PR models, P14d showed the highest R of 0.92 and lowest MSE of 150, among rolling averages (P3d, P7d, P14d) and future averages (P3dF, P7dF, P14dF) (Fig. 2A). The accuracy of ANN-PR models is higher than ANN-IR models, as confirmed by the lower scatteredness of ANN estimations (Fig. 2C). The range of R for prevalence estimations is between 0.81-0.92, which is much higher than the range of 0.61-0.78 for incidence estimations. As shown in Fig. 3, the ANN-PR model showed excellent performance in estimating 14-day rolling average to a very high accuracy, for different catchments with a population ranging from six thousand to half a million. The ANN-PR model captured the P14d trend accurately with only some peaks being underestimated. It is also noted that the accuracy for smaller WWTPs was lower due to the reduced representativeness of wastewater samples.

The back-estimation of COVID-19 case numbers using a WBE approach is largely based on the assumption of viral shedding into sewers. Viral shedding (via feces, saliva, and/or sputum) varies by individual and through time after infection, from the pre-symptom onset till after the recovery. The temporal dynamics of shedding, the total amount of virus shed by an infected individual and the ratio of shedding patients may be responsible for the observed different performances in estimating incidence or prevalence rates. Currently, excretions and bodily fluids including feces, urine, blood, saliva, serum, sputum, etc. are regarded as the major shedding source of SARS-CoV-2 RNA in wastewater (Kim et al., 2020; Lo et al., 2020; Peng et al., 2020). Our previous

meta-analysis revealed that the mean shedding magnitude was $10^{4.52}$ ± 0.13 gene copies/g feces, and the mean shedding probability was 0.54 ± 0.09 (Li et al., 2021c). It is also possible other shedding sources such as sputum may be a major contributor to the viral load in wastewater (Li et al., 2021a). It is known that COVID-19 patients continue to shed virus even after recovery of symptoms (Sun et al., 2020; Tao et al., 2021; Wu et al., 2020a). Thus, all active COVID-19 cases (prevalence), not just new cases (incidence), contribute to the measured viral load in wastewater.

The shedding load in sputum and throat swab samples peaked in the first week following the symptom onset and then decreased to 1% of the peak load after three weeks, while the fecal shedding loads remained high until five weeks after the symptom onset (Jones et al., 2020). Fecal shedding was reported to peak at 0.34 day after the symptom onset, with a shedding concentration about 10^3 times higher than the median concentration over the whole shedding period (Miura et al., 2021). As a result, our ANN models generated reasonable estimates of incidence rates, reflecting the fact that daily new cases are expected to play a major role in the total SARS-CoV-2 RNA loads from patients in the wastewater.

It is also observed that the ANN model performance was in the ascending order of P3d, P7d and P14d. This is likely related to the recovery time of COVID-19 patients. An Indian study in early 2020 (March to April) reported average recovery times of 221 patients as 25 days (95%CI: 16-34 days) (Barman et al., 2020). In New South Wales, Australia, among 2904 cases confirmed between January to May 2020, 20% recovered by 10 days, 60% by 20 days, and 80% by 30 days (Liu et al., 2021). In the UK, the mean disease duration in the first phase of
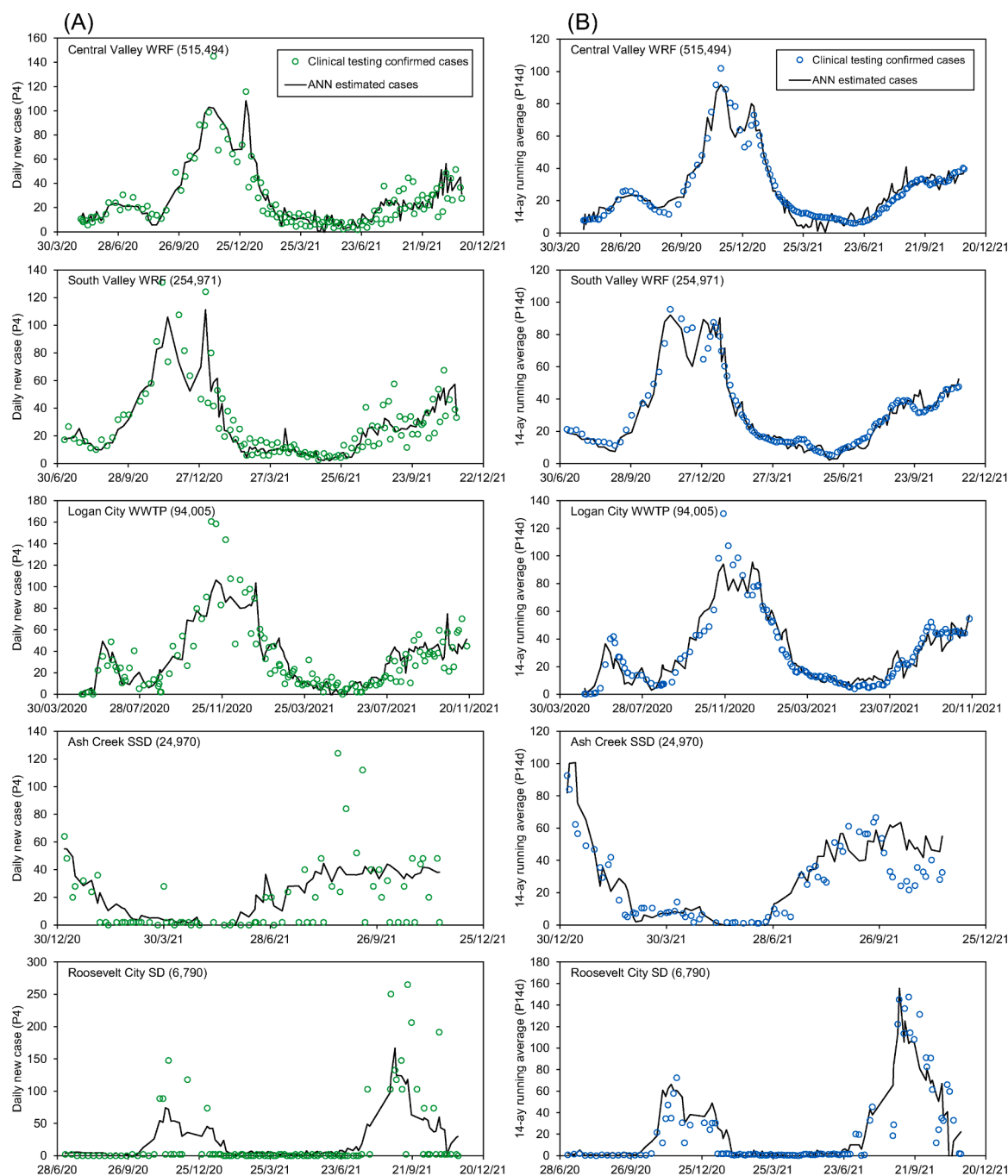
Fig. 3. Daily new cases (green circles) on the 4th day (A) and 14-day running average (B) case numbers of wastewater sampling date and ANN estimated cases (lines) for selected wastewater treatment plants with different populations, i.e., 500, 250, 100, 25, and 6 thousand people in Utah, USA.

the pandemic (January to June 2020) was estimated to be 23.5 ± 9.9 days (n=2045), while children have a significantly shorter recovery time compared to adults (p=0.04) (Mizrahi et al., 2020). Considering the ongoing shedding of viruses by COVID-19 patients, WBE approach tends to capture prevalence rates within the time window of the average recovery time, especially in the early stages of infection.

3.3. Estimation of COVID-19 effective reproduction number

For the prediction and management of COVID-19 outbreaks, effective reproduction rate (infection rate) is an important indicator showing the increasing (R_i>1) or decreasing (R_i<1) trend of case numbers. This study managed to collect the effective reproduction rate for all the Utah

WWTP catchments on wastewater sampling days. The data shows how fast COVID-19 was spreading in a WWTP catchment by estimating the number of people that a newly infected person goes on to eventually infect. The effective reproduction rate was initially calculated by a mathematical model which combines trends in daily new cases from approximately the last 14 days, with estimates for other variables, such as how many days on average occur between infection and transmission.

This study developed an ANN model (ANN-R_i) to predict the R_i with the same inputs used for the case number estimation. As shown in Fig. 4, the ANN-R_i model showed a good performance with a correlation coefficient R =0.72 ± 0.003 and MSE = 0.007 ± 0.0001 (values with 95% confidence interval) in predicting R_i. In general, the ANN-R_i model slightly underestimated the R_i, likely due to the viral load in wastewater
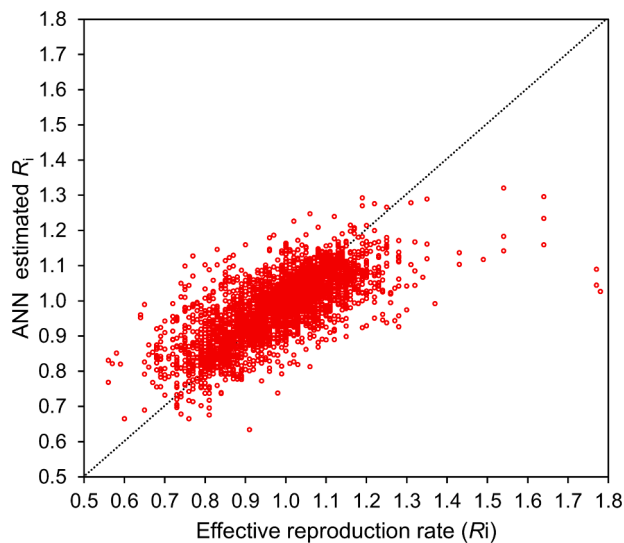
Fig. 4. The regression plot of the ANN estimated effective reproduction rate vs. the reported effective reproduction rate determined in conventional approach in Utah, USA.

reflects more about prevalence rate, instead of incidence rate while R_i was primarily calculated from new case numbers. For only a few days with high reported effective reproduction rate >1.2, ANN models tend to underestimate the R_i at a ceiling level around 1.1. This is likely caused by the limited amount of data during those days with R_i>1.2 to train the ANN model adequately. Other models suitable for small dataset might be employed to provide predictions in this range.

The effective reproductive number R_i of COVID-19 in each WWTP catchment is dynamic and affected by regional pandemic management, policies, and the effectiveness of interventions such as vaccination. A recent study reported a wastewater estimated R_i through optimizing the fit to the clinical case-based R_i by changing the viral shedding load distribution (Huisman et al., 2021a). The approach is based on the inferred COVID-19 incidence rate from viral RNA concentrations in wastewater. It is thus susceptible to the correlation between viral concentration and incidence rate, which was shown to be lower than that of the prevalence rate in this study. In comparison, our ANN approach combined the wastewater data with other factors such as weather, vaccination rates and clinical testing coverage to generate the estimation of R_i. To our knowledge, this is the first time R_i is estimated from both viral concentrations in wastewater and other factors affecting the viral transmission and pandemic dynamics.

3.4. Contributions of ANN model inputs: vaccination, clinical testing and weather

A systematic evaluation was conducted by developing ANN models with complete and partial input data to determine their capacity in predicting the COVID-19 incidence and prevalence rates. It is clear that partial inputs decreased the ANN performance to various degrees (Table 3). The full input always generated the highest R and lowest MSE for both the estimation of incidence and prevalence rates. Inputs without vaccine (V), or clinical testing (T) or weather (W) data slightly reduced the ANN model performance. The contribution of single input was in the order of weather > clinical testing > vaccination. When two types of parameters were removed from the inputs, the combination of vaccination and clinical testing data showed the highest contribution to P4 estimation, in comparison to the combination of clinical testing and weather data for P14d estimation. When only wastewater and catchment data were used to train the ANN model, it didn't give any reliable predictions of incidence rate (R=0.43 ± 0.04, MSE=877 ± 37) and a

Table 3
The correlation coefficient (R) and estimation error (MSE) of the ANN-IR, ANN-PR and ANN-R_i models being trained with complete or partial inputs to predict incidence rate P4, prevalence rate P14d and effective production rate R_i, respectively.

		R		MSE	
	Inputs	Average	95% CI	Average	95% CI
P4	All	0.72	0.002	525.7	3.02
	All-V	0.68	0.009	583.3	13.70
	All-T	0.69	0.003	572.9	4.61
	All-W	0.68	0.008	587.5	12.48
	All-V-T	0.55	0.021	765.4	25.24
	All-V-W	0.62	0.007	677.4	9.30
	All-T-W	0.60	0.021	699.0	31.96
	All-V-T-W	0.44	0.046	921.6	136.41
P14d	All	0.89	0.004	151.4	5.76
	All-V	0.87	0.005	184.6	6.92
	All-T	0.86	0.011	199.5	15.77
	All-W	0.85	0.005	209.6	6.76
	All-V-T	0.78	0.012	290.5	14.40
	All-V-W	0.82	0.004	243.0	4.51
	All-T-W	0.76	0.008	313.1	9.98
	All-V-T-W	0.58	0.027	501.5	34.72
R_i	All	0.85	0.003	0.0072	0.0001
	All-V	0.78	0.007	0.0104	0.0003
	All-T	0.83	0.016	0.0086	0.0008
	All-W	0.82	0.002	0.0090	0.0001
	All-V-T	0.72	0.025	0.0128	0.0010
	All-V-W	0.71	0.003	0.0134	0.0001
	All-T-W	0.78	0.004	0.0104	0.0002
	All-V-T-W	0.67	0.005	0.0148	0.0002

barely acceptable prevalence rate (R=0.57 ± 0.01, MSE=499 ± 19). Table 3A also shows that vaccination and weather are important input variables for the estimation of R_i, while clinical testing only slightly contributed to the better estimation. The R of the ANN model was only 0.31±0.05 when both vaccination and weather data were excluded as inputs. Overall, the analysis demonstrated that vaccination data contributed the most, followed by weather and clinical testing, to the ANN estimation of COVID case numbers and effective reproduction rate.

Our previous studies have identified that, in addition to C_{RNA}, the in-sewer decay of viral RNA and weather (temperature and precipitation) contribute to a better estimation of COVID-19 case numbers (Li et al., 2021b). This paper, for the first time, included vaccination and clinical testing coverage (positive ratio) in the WBE back-estimation of case numbers. The actual COVID-19 case number is usually unknown as clinical testing is known to only capture a portion of the total infections (Vallejo et al., 2020). The USA Centre for Disease Control and Prevention (CDC) estimated that only 1 in 4.3 (95% CI 3.7-5.0) of total COVID-19 infections were reported through clinical testing (Reese et al., 2020). Higher clinical testing coverage tends to capture a higher proportion of the total 'true' infections. This makes the clinical testing coverage critical for developing an ANN model to quantify the correlations between C_{RNA} and clinically confirmed case numbers.

COVID-19 vaccines are effective in alleviating the symptoms and the viral shedding from patients, which is a critical process in the WBE back-estimation (Eq. 1). Vaccination provides protection by preventing the infection of humans and minimizing transmission through reducing viral shedding (Bartsch et al., 2020). In preclinical nonhuman primate challenge experiments, several vaccines successfully prevented disease and prevented or reduced nasal shedding and virus replication in the lower respiratory tract depending on the vaccine dose (Corbett et al., 2020; van Doremalen et al., 2020). Also, many vaccine breakthrough infections (infection of a person after receipt of all recommended doses of authorized COVID-19 vaccine for ≥14 days) have been reported worldwide. However, the clinical report of viral shedding from fully or partially vaccinated COVID-19 patients is still lacking, thereby the potential impacts of vaccination on viral RNA shedding and WBE back-estimation remains unclear due to limited clinical data. The ANN

model addressed the issue by including both ratios of initiated (1st dose) and completed (both doses) vaccinated population as input parameters, which indeed improved the ANN performance.

Weather conditions (air temperature and precipitation) improved the ANN performance (Table 3). Our previous study also observed contributions of air temperature and daily precipitation on estimating COVID-19 prevalence using data-driven models (Li et al., 2021b). Generally, a higher air temperature is associated with a higher wastewater temperature (Hart and Halden, 2020), which facilitated the in-sewer decay of SARS-CoV-2 RNA, leading to a lower C_{RNA} in wastewater (Ahmed et al., 2020b; Bivins et al., 2020). Also, warmer weather likely reduced the transmission of COVID-19 as individuals were outside more than inside. Significant dilution of SARS-CoV-2 concentration in combined sewers has been observed due to the storm water or precipitation inflow (Chavarria-Miró et al., 2020). Furthermore, an increase of COVID-19 incidences in 4-8 days after flooding events were observed in some metropolitan regions (Han and He, 2021). This is likely related to increased chance of exposure to overflowed sewage, fresh human excreta, sewage-contaminated surfaces or aerosols, in areas with human activities and receiving sewage overflows during urban flooding events. In addition, studies from China and Bangladesh and a meta-study of 166 countries all revealed that higher temperature and relative humidity reduced the transmission of COVID-19 in the community, leading to lower incidences and deaths (Haque and Rahman, 2020; Qi et al., 2020; Wu et al., 2020b). Thus, the improved performance of ANN with the inclusion of weather conditions (air temperature and precipitation) might also be related to the variations of COVID-19 transmission caused by storm water and air temperature.

3.5. Transferability, implications, and limitations

The established ANN models established above was further assessed for its transferability to a different WBE program in Wisconsin, USA. The dataset was obtained from 66 WWTPs between August 2020 and December 2021. Both incidence and prevalence rates were successfully estimated by the models with excellent accuracy as shown in Fig. SI-4 and SI-5. However, the early warning lead time was determined as 2 days (Fig. SI-4A), with the highest R and lowest MSE. For prevalence estimation, the 3-day running average was shown with the best accuracy. These differences are likely due to the higher sampling frequency in Wisconsin (about twice a week) comparing to Utah (weekly). The effective reproduction rate R_i was also estimated with high accuracy (Fig. SI-6). Overall, the ANN models were successfully transferred to a different WBE program with the same inputs/outputs, neural network structure and training strategies.

For WBE sites or programs using consistent sampling and analytical procedure, the current ANN models can be successfully used to generate case numbers and R_i. However, input variables related to sampling or analytical approaches were not included in the ANN models. When trained ANN models were directly used to estimate case numbers in different WBE programs, the actual estimates would be shifted by different sampling and analytical practices. This is confirmed when ANN models trained with Utah dataset gave matching trends but not accurate estimates for WBE data collected from Canada, Japan, and India through the WATMOC network (www.watmoc.com).

Analytical approaches including sample preparation, concentration, RNA extraction, and RT-qPCR detection introduced a great amount of uncertainties to the C_{RNA}, where up to seven orders of magnitude difference were observed with the recovery efficiency using 36 standard operating procedures (Pecson et al., 2021). There was less than a 0.5-log difference in the interlaboratory and intra-laboratory results when the same procedure was applied to replicate samples. Conflicting results were also reported from different studies regarding the sampling techniques. Curtis et al. (2021) observed negligible impacts of sampling technique on C_{RNA}, where a good agreement was achieved between most grab samples and their respective composite. In contrast, Gerrity et al.

(2021) reported a 10-fold increase in C_{RNA} from composite samples than that of corresponding grab samples. To build a universal ANN model utilizing datasets from WBE campaigns using different sampling and analytical approaches, input variables representing the analytical and sampling approaches are essential.

During the wastewater surveillance in Utah, various variants of SARS-CoV-2 have been detected including Alpha, Beta, and Delta variants (University of Utah Health Communications, 2021). However, as the identification of the variants requires sequencing approaches, which are more costly and time-consuming than the current clinical testing approach (RT-qPCR), the proportion of each variant among the population was not determined. Recent clinical studies revealed higher viral titers in swab samples of the Delta than that of the Alpha and Beta variant of SARS-CoV-2 (Despres et al., 2021). To date, the shedding dynamics of these variants in feces, sputum, or other bodily fluids have not been reported yet. Our study revealed that the contribution of multiple strains/mutants can be handled together through ANN models for the incidence or prevalence estimations. However, extra attention shall be paid if a sudden outbreak or peak of a new variant with different shedding dynamics occurs. Furthermore, the recent emerging Omicron variant was also found to be more transmissible and partially resistant to existing vaccines based on currently limited knowledge (Karim and Karim, 2021). The clinical testing and coverage of vaccines played important roles in the incidence, prevalence, and Ri estimations using ANN models in this study. Emerging of variants with higher transmissivity and vaccine resistance might affect the accuracy of the ANN model. However, the architecture and inputs of the established ANN model can be retained but will need to be retrained when new data becomes available.

Although ANN or other machine learning models hold great promise for their application in WBE, only one previous study compared the effectiveness of random forest method with linear regression (Koureas et al., 2021). The study focused on case numbers and both methods provide very high-level accuracy, likely due to the small dataset obtained from two WWTPs. This study not only obtained a large dataset but critically identified features/inputs for such WBE models. Clinical testing, vaccination and weather data were indispensable for the prediction in diverse catchments. In addition, other machine learning or artificial intelligence models should be evaluated using the collected datasets for a comprehensive comparison of different data-driven techniques. Especially, the performance in predicting incidence rates needs to be improved by adopting other modeling algorithms.

4. Conclusions

This study for the first time developed an ANN-based back-estimation method for WBE in calculating the COVID-19 case numbers and effective reproduction rate from wastewater analysis and other supporting input data including vaccination rate, clinical testing positive rate and weather data. The conclusions are:

- Using a wastewater-based epidemiology approach, ANN models were shown to estimate the COVID-19 prevalence and incidence rates accurately in diverse sewer catchment. The prevalence rates were estimated more accurately than the incidence rates. WBE likely provided the upcoming COVID-19 incidence 2-4 days ahead the wastewater sampling date, depending on sampling frequency.
- The correlation between SARS-CoV-2 RNA concentration or viral load in wastewater with COVID-19 case numbers is limited. It needs to be used in conjunction with weather, clinical testing, and vaccination rates for accurate ANN predictions of COVID case numbers.
- A unique approach to estimate effective reproduction rate using both wastewater data and other parameters affecting viral transmission and pandemic dynamics was devised for the first time using an ANN model.

- ANN models have a huge potential and is practical with the current limited knowledge and many uncertainties to achieve the WBE back-estimation of COVID case numbers. The model is transferable for different WBE sites or programs.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Guangming Jiang reports financial support was provided by Australian Research Council. Guangming Jiang reports financial support was provided by Australian Academy of Science. Guangming Jiang reports financial support was provided by Australian Government Department of Industry Science Energy and Resources.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2022.118451.

References

Abdeldayem, O.M., Dabbish, A.M., Habashy, M.M., Mostafa, M.K., Elhefnawy, M., Amin, L., Al-Sakkari, E.G., Ragab, A., Rene, E.R, 2022. Viral outbreaks detection and surveillance using wastewater-based epidemiology, viral air sampling, and machine learning techniques: a comprehensive review and outlook. Sci. Total Environ. 803, 149834.

Acosta, N., Bautista, M.A., Hollman, J., McCalder, J., Beaudet, A.B., Man, L., Waddell, B. J., Chen, J., Li, C., Kuzma, D., Bhatnagar, S., Leal, J., Meddings, J., Hu, J., Cabaj, J.L., Ruecker, N.J., Naugler, C., Pillai, D.R., Achari, G., Ryan, M.C., Conly, J.M., Frankowski, K., Hubert, C.R., Parkins, M.D, 2021. A multicenter study investigating SARS-CoV-2 in tertiary-care hospital wastewater. viral burden correlates with increasing hospitalized cases as well as hospital-associated transmissions and outbreaks. Water Res. 201.

Agrawal, S., Orschler, L., Lackner, S., 2021. Long-term monitoring of SARS-CoV-2 RNA in wastewater of the Frankfurt metropolitan area in Southern Germany. Sci. Rep. 11 (1), 5372.

Ahmed, W., Bertsch, P.M., Angel, N., Bibby, K., Bivins, A., Dierens, L., Edson, J., Ehret, J., Gyawali, P., Hamilton, K., Hosegood, I., Hugenholtz, P., Jiang, G., Kitajima, M., Sichani, H.T., Shi, J., Shimko, K.M., Simpson, S.L., Smith, W.J.M., Symonds, E.M., Thomas DSC, K.V., Verhagen, R., Zaugg, J., Mueller, J.F, 2020a. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: a surveillance tool for assessing the presence of COVID-19 infected travelers. J. Travel Med.

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., Vasan, 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. -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.

Allam, Z., Dey, G., Jones, D.S, 2020. Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future urban health policy internationally. AI 1 (2), 156–165.

Asad, S.M., Dashtipour, K., Rais, R.N.B., Hussain, S., Abbasi, Q.H., Imran, M.A., 2021. Employing Machine Learning for Predicting Transportation Modes Under the COVID-19 Pandemic: A Mobility-Trends Analysis, pp. 235–240.

Asghar, H., Diop, O.M., Weldegebriel, G., Malik, F., Shetty, S., El Bassioni, L., Akande, A. O., Al Maamoun, E., Zaidi, S., Adeniji, A.J, 2014. Environmental surveillance for polioviruses in the Global Polio Eradication Initiative. J. Infect. Dis. 210 (suppl_1), S294–S303.

Barman, M.P., Rahman, T., Bora, K., Borgohain, C., 2020. COVID-19 pandemic and its recovery time of patients in India: a pilot study. Diabetes Metab. Syndr. 14 (5), 1205–1211.

Barrios, M.E., Díaz, S.M., Torres, C., Costamagna, D.M., Blanco Fernández, M.D., Mbayed, V.A, 2021. Dynamics of SARS-CoV-2 in wastewater in three districts of the Buenos Aires metropolitan region, Argentina, throughout nine months of surveillance: A pilot study. Sci. Total Environ. 800, 149578.

Bartsch, S.M., O'Shea, K.J., Ferguson, M.C., Bottazzi, M.E., Wedlock, P.T., Strych, U., McKinnell, J.A., Siegmund, S.S., Cox, S.N., Hotez, P.J, 2020. Vaccine efficacy needed for a COVID-19 coronavirus vaccine to prevent or stop an epidemic as the sole intervention. Am. J. Prev. Med. 59 (4), 493–503.

Betancourt, W.Q., Schmitz, B.W., Innes, G.K., Prasek, S.M., Pogreba Brown, K.M., Stark, E.R., Foster, A.R., Sprissler, R.S., Harris, D.T., Sherchan, S.P., Gerba, C.P., Pepper, I.L, 2021. COVID-19 containment on a college campus via wastewater-based epidemiology, targeted clinical testing and an intervention. Sci. Total Environ. 779.

Bivins, A., Greaves, J., Fischer, R., Yinda, K.C., Ahmed, W., Kitajima, M., Munster, V.J., Bibby, K., 2020. Persistence of SARS-CoV-2 in water and wastewater. Environ. Sci. Technol. Lett. 7 (12), 937–942.

Braga, M.d.B., Fernandes, R.d.S., Souza Jr., G.N.d., Rocha, J.E.C.d., Dolácio, C.J.F., Tavares Jr., I.d.S., Pinheiro, R.R., Noronha, F.N., Rodrigues, L.L.S., Ramos, R.T.J., Carneiro, A.R., Brito, S.R.d., Diniz, H.A.C., Botelho, M.d.N., Vallinoto, A.C.R, 2021. Artificial neural networks for short-term forecasting of cases, deaths, and hospital beds occupancy in the COVID-19 pandemic at the Brazilian Amazon. PLoS One 16 (3), e0248161.

Brinati, D., Campagner, A., Ferrari, D., Locatelli, M., Banfi, G., Cabitza, F., 2020. Detection of COVID-19 infection from routine blood exams with machine learning: a feasibility study. J. Med. Syst. 44 (8), 135.

Brunese, L., Martinelli, F., Mercaldo, F., Santone, A., 2020. Machine learning for coronavirus covid-19 detection from chest x-rays. Procedia Comput. Sci. 176, 2212–2221.

Chavarria-Miró, G., Anfruns-Estrada, E., Guix, S., Paraira, M., Galofré, B., Sáanchez, G., Pintó, R., Bosch, A., 2020. Sentinel surveillance of SARS-CoV-2 in wastewater anticipates the occurrence of COVID-19 cases. medRxiv.

Choi, P.M., Bowes, D.A., O'Brien, J.W., Li, J., Halden, R.U., Jiang, G., Thomas, K.V., Mueller, J.F, 2020. Do food and stress biomarkers work for wastewater-based epidemiology? A critical evaluation. Sci. Total Environ. 736, 139654.

Corbett, K.S., Flynn, B., Foulds, K.E., Francica, J.R., Boyoglu-Barnum, S., Werner, A.P., Flach, B., O'Connell, S., Bock, K.W., Minai, M., 2020. Evaluation of the mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. N. Engl. J. Med. 383 (16), 1544–1555.

Curtis, K., Keeling, D., Yetka, K., Larson, A. and Gonzalez, R. 2021. Wastewater SARS-CoV-2 RNA Concentration and Loading Variability from Grab and 24-Hour Composite Samples. medRxiv: the preprint server for health sciences, 2020.2007.2010.20150607.

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.

Despres, H.W., Mills, M.G., Shirley, D.J., Schmidt, M.M., Huang, M.-L., Jerome, K.R., Greninger, A.L. and Bruce, E.A. 2021. Quantitative measurement of infectious virus in SARS-CoV-2 Alpha, Delta and Epsilon variants reveals higher infectivity (viral titer:RNA ratio) in clinical samples containing the Delta and Epsilon variants. medRxiv: the preprint server for health sciences, 2021.2009.2007.21263229.

EMCDDA, 2016. Assessing Illicit Drugs in Wastewater Advances in Wastewater-Based Drug Epidemiology, Luxembourg.

Fernandez-Cassi, X., Scheidegger, A., Bänziger, C., Cariti, F., Tuñas Corzon, A., Ganesanandamoorthy, P., Lemaitre, J.C., Ort, C., Julian, T.R., Kohn, T., 2021. Wastewater monitoring outperforms case numbers as a tool to track COVID-19 incidence dynamics when test positivity rates are high. Water Res. 200.

Gao, J., Li, J., Jiang, G., Yuan, Z., Eaglesham, G., Covaci, A., Mueller, J.F., Thai, P.K, 2018. Stability of alcohol and tobacco consumption biomarkers in a real rising main sewer. Water Res. 138, 19–26.

Gerrity, D., Papp, K., Stoker, M., Sims, A., Frehner, W., 2021. Early-pandemic wastewater surveillance of SARS-CoV-2 in Southern Nevada: Methodology, occurrence, and incidence/prevalence considerations. Water Res. X 10, 100086.

Gibas, C., Lambirth, K., Mittal, N., Juel, M.A.I., Barua, V.B., Roppolo Brazell, L., Hinton, K., Lontai, J., Stark, N., Young, I., Quach, C., Russ, M., Kauer, J., Nicolosi, B., Chen, D., Akella, S., Tang, W., Schlueter, J., Munir, M., 2021. Implementing building-level SARS-CoV-2 wastewater surveillance on a university campus. Sci. Total Environ. 782.

Gonzalez-Marino, I., Baz-Lomba, J.A., Alygizakis, N.A., Andres-Costa, M.J., Bade, R., Bannwarth, A., Barron, L.P., Been, F., Benaglia, L., Berset, J.D., Bijlsma, L., Bodik, I., Brenner, A., Brock, A.L., Burgard, D.A., Castrignano, E., Celma, A., Christophoridis, C.E., Covaci, A., Delemont, O., de Voogt, P., Devault, D.A., Dias, M. J., Emke, E., Esseiva, P., Fatta-Kassinos, D., Fedorova, G., Fytianos, K., Gerber, C., Grabic, R., Gracia-Lor, E., Gruner, S., Gunnar, T., Hapeshi, E., Heath, E., Helm, B., Hernandez, F., Kankaanpaa, A., Karolak, S., Kasprzyk-Hordern, B., Krizman-Matasic, I., Lai, F.Y., Lechowicz, W., Lopes, A., Lopez de Alda, M., Lopez-Garcia, E., Love, A.S.C., Mastroianni, N., McEneff, G.L., Montes, R., Munro, K., Nefau, T., Oberacher, H., O'Brien, J.W., Oertel, R., Olafsdottir, K., Pico, Y., Plosz, B.G., Polesel, F., Postigo, C., Quintana, J.B., Ramin, P., Reid, M.J., Rice, J., Rodil, R., Salgueiro-Gonzalez, N., Schubert, S., Senta, I., Simoes, S.M., Sremacki, M.M., Styszko, K., Terzic, S., Thomaidis, N.S., Thomas, K.V., Tscharke, B.J., Udrisard, R., van Nuijs, A.L.N., Yargeau, V., Zuccato, E., Castiglioni, S., Ort, C., 2020. Spatio-temporal assessment of illicit drug use at large scale: evidence from 7 years of international wastewater monitoring. Addiction 115 (1), 109–120.

Han, J., He, S., 2021. Urban flooding events pose risks of virus spread during the novel coronavirus (COVID-19) pandemic. Sci. Total Environ. 755, 142491.

Haque, S.E., Rahman, M., 2020. Association between temperature, humidity, and COVID-19 outbreaks in Bangladesh. Environ. Sci. Policy 114, 253–255.

Hart, O.E., Halden, R.U. 2020. Modeling wastewater temperature and attenuation of sewage-borne biomarkers globally. Water Res. 172, 115473.

He, C., Li, J., Jiang, G., Chen, S., Niel, C., Yuan, Z., Mueller, J.F., Thai, P., 2021. Transformation of phthalates and their metabolites in wastewater under different sewer conditions. Water Res. 190, 116754.

Hellmér, M., Paxéus, N., Magnius, L., Enache, L., Arnholm, B., Johansson, A., Bergström, T., Norder, H., 2014. Detection of pathogenic viruses in sewage provided early warnings of a virus and norovirus outbreaks. Appl. Environ. Microbiol. 80 (21), 6771–6781.

Huang, Y., Johnston, L., Parra, A., Sweeney, C., Hayes, E., Hansen, L.T., Gagnon, G., Stoddart, A., Jamieson, R., 2021. Detection of SARS-CoV-2 in wastewater in Halifax, Nova Scotia, Canada, using four RT-qPCR assays. Facets 6, 959–965.

Huisman, J., Scire, J., Caduff, L., Fernandez-Cassi, X., Ganesanandamoorthy, P., Kull, A., Scheidegger, A., Stachler, E., Boehm, A., Hughes, B., Knudson, A., Topol, A., Wigginton, K., Wolfe, M., Kohn, T., Ort, C., Stadler, T. and Julian, T. 2021a Wastewater-based estimation of the effective reproductive number of SARS-CoV-2, medRxiv.

Huisman, J.S., Scire, J., Caduff, L., Fernandez-Cassi, X., Ganesanandamoorthy, P., Kull, A., Scheidegger, A., Stachler, E., Boehm, A.B., Hughes, B., Knudson, A., Topol, A., Wigginton, K.R., Wolfe, M.K., Kohn, T., Ort, C., Stadler, T. and Julian, T.R. 2021b. Wastewater-based estimation of the effective reproductive number of SARS-CoV-2. medRxiv, 2021.2004.2029.21255961.

Jones, D.L., Baluja, M.Q., Graham, D.W., Corbishley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H., Farkas, K., 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. Sci. Total Environ. 749, 141364.

Karim, S.S.A., Karim, Q.A, 2021. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. Lancet 398 (10317), 2126–2128.

Kim, J.-M., Kim, H.M., Lee, E.J., Jo, H.J., Yoon, Y., Lee, N.-J., Son, J., Lee, Y.-J., Kim, M. S., Lee, Y.-P., Chae, S.-J., Park, K.R., Cho, S.-R., Park, S., Kim, S.J., Wang, E., Woo, S., Lim, A., Park, S.-J., Jang, J., Chung, Y.-S., Chin, B.S., Lee, J.-S., Lim, D., Han, M.-G., Yoo, C.K, 2020. Detection and Isolation of SARS-CoV-2 in Serum, Urine, and Stool Specimens of COVID-19 Patients from the Republic of Korea. Osong Public Health Res. Perspect. 11 (3), 112–117.

Kitamura, K., Sadamasu, K., Muramatsu, M., Yoshida, H., 2021. Efficient detection of SARS-CoV-2 RNA in the solid fraction of wastewater. Sci. Total Environ. 763, 144587.

Koureas, M., Amoutzias, G.D., Vontas, A., Kyritsi, M., Pinaka, O., Papakonstantinou, A., Dadouli, K., Hatzinikou, M., Koutsolioutsou, A., Mouchtouri, V.A., Speletas, M., Tsiodras, S., Hadjichristodoulou, C., 2021. Wastewater monitoring as a supplementary surveillance tool for capturing SARS-COV-2 community spread. A case study in two Greek municipalities. Environ. Res. 200, 111749.

Krogh, A., 2008. What are artificial neural networks? Nat. Biotechnol. 26 (2), 195–197.

Lalmuanawma, S., Hussain, J., Chhakchhuak, L., 2020. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. Chaos Solitons Fractals 139, 110059.

Li, J., Gao, J., Thai, P.K., Shypanski, A., Nieradzik, L., Mueller, J.F., Yuan, Z., Jiang, G., 2019. Experimental investigation and modeling of the transformation of illicit drugs in a pilot-scale sewer system. Environ. Sci. Technol. 53 (8), 4556–4565.

Li, X., Kulandaivelu, J., Guo, Y., Zhang, S., Shi, J., O'Brien, J., Arora, S., Kumar, M., Sherchan, S.P., Honda, R., Jackson, G., Luby, S.P., Jiang, G., 2022. SARS-CoV-2 shedding sources in wastewater and implications for wastewater-based epidemiology. J. Hazard. Mater., 128667

Li, X., Kulandaivelu, J., Guo, Y., Zhang, S., Shi, J., O'Brien, J., Luby, S.P., Jackson, G., Jiang, G., 2021a. Potential role of sputum as a critical SARS-CoV-2 virus shedding source in the back-estimation of COVID-19 prevalence through wastewater-based epidemiology. EST Water. Under Review.

Li, X., Kulandaivelu, J., Zhang, S., Shi, J., Sivakumar, M., Mueller, J., Luby, S., Ahmed, W., Coin, L., Jiang, G., 2021b. Data-driven estimation of COVID-19 community prevalence through wastewater-based epidemiology. Sci. Total Environ. 789, 147947.

Li, X., Zhang, S., Shi, J., Luby, S.P., Jiang, G., 2021c. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. Chem. Eng. J. 415, 129039.

Liu, B., Jayasundara, D., Pye, V., Dobbins, T., Dore, G.J., Matthews, G., Kaldor, J., Spokes, P., 2021. Whole of population-based cohort study of recovery from COVID-19 in New South Wales Australia. Lancet Reg. Health West Pac. 12, 100193.

Lo, I.L., Lio, C.F., Cheong, H.H., Lei, C.I., Cheong, T.H., Zhong, X., Tian, Y., Sin, N.N, 2020. Evaluation of SARS-CoV-2 RNA shedding in clinical specimens and clinical characteristics of 10 patients with COVID-19 in Macau. Int. J. Biol. Sci. 16 (10), 1698.

Matheri, A.N., Belaid, M., Njenga, C.K., Ngila, J.C, 2022. Water and wastewater digital surveillance for monitoring and early detection of the COVID-19 hotspot: industry 4.0. Int. J. Environ. Sci. Technol.

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. Letters 7 (7), 511–516.

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.

Mizrahi, B., Shilo, S., Rossman, H., Kalkstein, N., Marcus, K., Barer, Y., Keshet, A., Shamir-Stein, N.a., Shalev, V., Zohar, A.E., Chodick, G., Segal, E., 2020. Longitudinal Symptom Dynamics of COVID-19 Infection, p. 6208.

Mohammad-Rahimi, H., Nadimi, M., Ghalyanchi-Langeroudi, A., Taheri, M., Ghafouri-Fard, S., 2021. Application of machine learning in diagnosis of COVID-19 through X-Ray and CT images: a scoping review. Front. Cardiovasc. Med. 8.

Mottaqi, M.S., Mohammadipanah, F., Sajedi, H., 2021. Contribution of machine learning approaches in response to SARS-CoV-2 infection. Inform. Med. Unlocked 23, 100526.

Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Cicha, C., Vanderwood, K.K., Wilkinson, R., Wiedenheft, B., 2020. Temporal detection and phylogenetic assessment of SARS-CoV-2 in municipal wastewater. Cell Rep. Med. 1 (6), 100098.

Pan, Y., Zhang, D., Yang, P., Poon, L.L., Wang, Q., 2020. Viral load of SARS-CoV-2 in clinical samples. Lancet Infect. Dis. 20 (4), 411–412.

Pecson, B.M., Darby, E., Haas, C., Amha, Y., Bartolo, M., Danielson, R., Dearborn, Y., Di Giovanni, G., Ferguson, C., Fevig, S., 2021. Reproducibility and sensitivity of 36 methods to quantify the SARS-CoV-2 genetic signal in raw wastewater: findings from an interlaboratory methods evaluation in the US. Environ. Sci.

Peng, L., Liu, J., Xu, W., Luo, Q., Chen, D., Lei, Z., Huang, Z., Li, X., Deng, K., Lin, B., 2020. SARS-CoV-2 can be detected in urine, blood, anal swabs, and oropharyngeal swabs specimens. J. Med. Virol.

Qi, H., Xiao, S., Shi, R., Ward, M.P., Chen, Y., Tu, W., Su, Q., Wang, W., Wang, X., Zhang, Z., 2020. COVID-19 transmission in Mainland China is associated with temperature and humidity: a time-series analysis. Sci. Total Environ. 728, 138778.

Rahman, M.M., Paul, K.C., Hossain, M.A., Ali, G.G.M.N., Rahman, M.S., Thill, J.-C, 2021. Machine learning on the COVID-19 pandemic, human mobility and air quality: a review. IEEE Access 9, 72420–72450.

Reese, H., Iuliano, A.D., Patel, N.N, Garg, S., Kim, L., Silk, B.J., Hall, A.J., Fry, A., Reed, C., 2020. Estimated incidence of coronavirus disease 2019 (COVID-19) illness and hospitalization—United States, February–September 2020. Clin. Infect. Dis.

Róka, E., Khayer, B., Kis, Z., Kovács, L.B., Schuler, E., Magyar, N., Málnási, T., Oravecz, O., Pályi, B., Pándics, T., Vargha, M., 2021. Ahead of the second wave: early warning for COVID-19 by wastewater surveillance in Hungary. Sci. Total Environ. 786.

Rusiñol, M., Zammit, I., Itarte, M., Forés, E., Martínez-Puchol, S., Girones, R., Borrego, C., Corominas, L., Bofill-Mas, S., 2021. Monitoring waves of the COVID-19 pandemic: Inferences from WWTPs of different sizes. Sci. Total Environ. 787.

Sangsanont, J., Rattanakul, S., Kongprajug, A., Chyerochana, N., Sresung, M., Sriporatana, N., Wanlapakorn, N., Poovorawan, Y., Mongkolsuk, S., Sirikanchana, K., 2021. SARS-CoV-2 RNA surveillance in large to small centralized wastewater treatment plants preceding the third COVID-19 resurgence in Bangkok, Thailand. Sci. Total Environ. 151169

Scott, L.C., Aubee, A., Babahaji, L., Vigil, K., Tims, S., Aw, T.G., 2021. Targeted wastewater surveillance of SARS-CoV-2 on a university campus for COVID-19 outbreak detection and mitigation. Environ. Res. 200.

Shawaqfah, M., Almomani, F., 2021. Forecast of the outbreak of COVID-19 using artificial neural network: case study Qatar, Spain, and Italy. Results Phys. 27, 104484.

Sherchan, S.P., Shahin, S., Ward, L.M., Tandukar, S., Aw, T.G., Schmitz, B., Ahmed, W., Kitajima, M., 2020. First detection of SARS-CoV-2 RNA in wastewater in North America: a study in Louisiana, USA. Sci. Total Environ. 743, 140621-140621.

Shi, J., Li, X., Zhang, S., Sharma, E., Sivakumar, M., Sherchan, S.P., Jiang, G., 2021. Enhanced decay of coronaviruses in sewers with domestic wastewater. Sci. Total Environ., 151919

Sun, J., Xiao, J., Sun, R., Tang, X., Liang, C., Lin, H., Zeng, L., Hu, J., Yuan, R., Zhou, P., Peng, J., Xiong, Q., Cui, F., Liu, Z., Lu, J., Tian, J., Ma, W., Ke, C., 2020. Prolonged persistence of SARS-CoV-2 RNA in body fluids. Emerg. Infect. Dis. 26 (8), 1834–1838.

Tao, W., Wang, X., Zhang, G., Guo, M., Ma, H., Zhao, D., Sun, Y., He, J., Liu, L., Zhang, K., Wang, Y., Weng, J., Ma, X., Jin, T., Zhu, S., 2021. Re-detectable positive SARS-CoV-2 RNA tests in patients who recovered from COVID-19 with intestinal infection. Protein Cell 12 (3), 230–235.

University of Utah Health Communications, U., 2021. Understanding Omicron and Other COVID-19 Variants.

Vallejo, J.A., Rumbo-Feal, S., Conde-Pérez, K., López-Oriona, Á., Tarrío, J., Reif, R., Ladra, S., Rodiño-Janeiro, B.K., Nasser, M., Cid, Á., Veiga, M.C., Acevedo, A., Lamora, C., Bou, G., Cao, R., Poza, M., 2020. Highly predictive regression model of active cases of COVID-19 in a population by screening wastewater viral load. medRxiv, 2020.2007.2002.20144865.

van Doorn, A.S., Meijer, B., Frampton, C.M., Barclay, M.L., de Boer, N.K., 2020. Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. Aliment. Pharmacol. Ther.

van Doremalen, N., Lambe, T., Spencer, A., Belij-Rammerstorfer, S., Purushotham, J.N., Port, J.R., Avanzato, V.A., Bushmaker, T., Flaxman, A., Ulaszewska, M., 2020. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. Nature 586 (7830), 578–582.

van Nuijs, A.L.N., Castiglioni, S., Tarcomnicu, I., Postigo, C., de Alda, M.L., Neels, H., Zuccato, E., Barcelo, D., Covaci, A., 2011. Illicit drug consumption estimations derived from wastewater analysis: a critical review. Sci. Total Environ. 409 (19), 3564–3577.

Weidhaas, J., Aanderud, Z.T., Roper, D.K., VanDerslice, J., Gaddis, E.B., Ostermiller, J., Hoffman, K., Jamal, R., Heck, P., Zhang, Y., Torgersen, K., Laan, J.V., LaCross, N., 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. Sci. Total Environ. 775, 145790.

Westhaus, S., Weber, F.A., Schiwy, S., Linnemann, V., Brinkmann, M., Widera, M., Greve, C., Janke, A., Hollert, H., Wintgens, T., Ciesek, 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.

Wieczorek, M., Siłka, J., Woźniak, M., 2020. Neural network powered COVID-19 spread forecasting model. Chaos Solitons Fract. 140, 110203.

Wong, J.C.C., Tan, J., Lim, Y.X., Arivalan, S., Hapuarachchi, H.C., Mailepessov, D., Griffiths, J., Jayarajah, P., Setoh, Y.X., Tien, W.P., Low, S.L., Koo, C., Yenamandra, S. P., Kong, M., Lee, V.J.M., Ng, L.C, 2021. Non-intrusive wastewater surveillance for monitoring of a residential building for COVID-19 cases. Sci. Total Environ. 786, 147419.

Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., Yin, H., Xiao, Q., Tang, Y., Qu, X., Kuang, L., Fang, X., Mishra, N., Lu, J., Shan, H., Jiang, G., Huang, X., 2020a. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. Lancet Gastroenterol. Hepatol. 5 (5), 434–435.

Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., Du, M., Liu, M., 2020b. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Sci. Total Environ. 729, 139051.

Zhang, S., Li, X., Shi, J., Sivakumar, M., Luby, S., O'Brien, J., Jiang, G., 2022. Analytical performance comparison of four SARS-CoV-2 RT-qPCR primer-probe sets for wastewater samples. Sci. Total Environ. 806, 150572.

Zheng, Q., Tscharke, B., O'Brien, J., Gerber, C., Mackie, R., Gao, J., Thai, P., 2019. Uncertainties in estimating alcohol and tobacco consumption by wastewater-based epidemiology. Curr. Opin. Environ. Sci. Health 9, 13–18.

Zuccato, E.C.C., Castiglioni, S, Calamari, D, Bagnati, R, Schiarea, S, Fanelli, R, 2005. Cocaine in surface waters- a new evidence-based tool to monitor community drug abuse. Environ. Health 4 (14), 7.