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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Viral outbreaks detection and surveillance using wastewater-based epidemiology, viral air sampling, and machine learning techniques: A comprehensive review and outlook

Omar M. Abdeldayem,⁎, Areeg M. Dabbish,✝, Mahmoud M. Habashy,✝, Mohamed K. Mostafa,✝, Mohamed Elhefnawy, Lobna Amin, Eslam G. Al-Sakkari, Ahmed Ragab, Eldon R. Rene

HIGHLIGHTS
• SARS-CoV-2 can be transmitted in different environments.
• Viral air surveillance provides insights for viral spread in highly infected areas.
• Wastewater-based epidemiology (WBE) is used widely for monitoring viral outbreaks.
• Artificial Intelligence (AI)-WBE framework could provide accurate viral outbreaks detection.
• AI synergized with WBE needs further investigation and research.

ABSTRACT
A viral outbreak is a global challenge that affects public health and safety. The coronavirus disease 2019 (COVID-19) has been spreading globally, affecting millions of people worldwide, and led to significant loss of lives and deterioration of the global economy. The current adverse effects caused by the COVID-19 pandemic demands finding new detection methods for future viral outbreaks. The environment’s transmission pathways include and are not limited to air, surface water, and wastewater environments. The wastewater surveillance, known as wastewater-based epidemiology (WBE), can potentially monitor viral outbreaks and provide a complementary clinical testing method. Another investigated outbreak surveillance technique that has not been yet implemented in a sufficient number of studies is the surveillance of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) in the air. Artificial Intelligence (AI) and its related machine learning (ML) and deep learning (DL) technologies are currently emerging techniques for detecting viral outbreaks using global data. To date, there are no reports that illustrate the potential of using WBE with AI to detect viral outbreaks. This study investigates the transmission pathways of SARS-CoV-2 in the environment and provides current updates on the surveillance of viral outbreaks using WBE, viral air sampling, and AI. It also proposes a novel...
1. Introduction

On 31 December 2019, the World Health Organization (WHO) office in China was informed by the Wuhan Municipal Health Commission of pneumonia cases of unknown etiology (Mostafa et al., 2021; World Health Organization, 2020c). On 9 January 2020, the Chinese Center for Disease Control and Prevention (China CDC) declared that a novel coronavirus (2019-nCoV) was detected in 15 cases who were suffering from pneumonia (European Centre for Disease Prevention and Control, 2020). On 11 March 2020, the WHO declared 2019-nCoV (later named COVID-19 or Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2)) a global pandemic after achieving a widespread in over 110 countries resulting in 118,319 confirmed cases (World Health Organization, 2020a).

On the 25th of May 2021, the total number of COVID-19 confirmed cases reached 167.2 million, while the total number of deaths exceeded 3.4 million people (World Health Organization, 2020b). The confirmed cases occurred globally, as follows: 39.6% in America, 18.3% in South-East Asia, 32.4% in Europe, 5.9% in Eastern Mediterranean, 2.1% in Africa, and 1.7% in Western Pacific, respectively (World Health Organization, 2020b). The quick spread of the disease resulted in a global lockdown to stop the virus spread (Al Huraimel et al., 2020; Lau et al., 2020). However, this action has adversely affected the economy, where the International Monetary Fund (IMF) anticipated a 3.5% recession in the global economy in 2020 due to COVID-19, which outweighs the 2008’s recession (International Monetary Fund, 2020; Wang and Su, 2020). However, the world output is expected to increase by 5.5% and 4.2% in 2021 and 2022, respectively (International Monetary Fund, 2021). It may be caused by reducing people and goods’ movement, especially in countries with the largest world economies (Baldwin and Mauro, 2020; Fernandes, 2020; Mostafa et al., 2021).

According to the International Civil Aviation Organization (ICAO), international air travel has dropped from 4.5 billion passengers taking flights in 2019 to 1.8 billion in 2020, leading to a drop of more than 50%. Hence, this staggered financial losses of more than $370 billion (Economic Development, 2021). Also, the sea freight volumes have decreased by 11% in the United States and by 4% in the EU27 between April and June 2020 compared to June 2008 (International Transport Forum, 2020). It is essential to find a way to provide an early warning for any future pandemic to avoid a deterioration in the health and economic sectors, as happened during the COVID-19 pandemic. The three main approaches that are applied to provide an early warning include sewage-based epidemiology, air biosensors, and computer techniques, such as artificial intelligence (AI) and its related Machine Learning (ML) technologies (Orive et al., 2020; Ribeiro et al., 2020a; Vaishya et al., 2020).

Wastewater-based epidemiology (WBE) is a commonly used approach to provide quantitative and qualitative information about inhabitants’ usage of drugs within a given wastewater catchment (Orive et al., 2020). Recently, this approach was proposed in the infectious diseases field (Al Huraimel et al., 2020; Orive et al., 2020). During the SARS-CoV-1 outbreak, between 16 and 37% of the patients suffered from diarrhea (Amoah et al., 2020; Yeo et al., 2020). Also, during the current COVID-19 pandemic, between 2 and 35% of the patients suffered from the same symptom (Wang et al., 2020a). Therefore, the study detected viral ribonucleic acid (RNA) in feces and sewage (Amoah et al., 2020; Pan et al., 2020). Many researchers have also detected infectious coronavirus virions in feces (Xiao et al., 2020). A recent study detected SARS-CoV-2 RNA in wastewater samples before reporting any case, which means that monitoring the virus could be possible before documenting it via the health surveillance system (Medema et al., 2020; Sherchan et al., 2020). In another study, researchers collected sewage samples in Paris during the pandemic, which showed a positive relationship between the SARS-CoV-2 genome units and the number of fatal cases registered regionally and nationally (Wurtzer et al., 2020).

Air biosensors are used as analytical devices to detect the respiratory virus by converting biological material, such as microorganisms, antibodies, tissues, biomimetic enzymes, or cell receptors, into measurable signals (Ribeiro et al., 2020a). The biological materials immobilize on a transducer surface, which produces a biochemical response by interacting with the analyte in the solution (Nguyen et al., 2019). The transducer then converts this biochemical response to a quantifiable signal measured using a digital detection module (Nguyen et al., 2019). The four main categories of biosensors are electrochemical biosensors, piezoelectric biosensors, thermal biosensors, and optical biosensors (Bukkitgar et al., 2020; Samson et al., 2020; Suleman et al., 2021). The types of biosensors used for the detection of emerging infectious diseases (EIDs) depends on two main factors: (1) the characteristics of the analyte (i.e., concentration, structure, and size); and (2) the matrix in which the analyte exists (i.e., liquid, air) (Pecjic et al., 2006). Different detected viruses include the influenza virus using electrochemical bio/immunosensor (Saylan et al., 2019). Also, the Middle respiratory syndrome coronavirus (MERS) was detected using optical bio/immunosensor (Ravina et al., 2020; Santiago, 2020) and SARS-CoV-2 using piezoelectric immunosensor or thermal biosensor (Lee et al., 2018; Woo et al., 2020).

AI and its related ML technologies are computer techniques and algorithms that turn the available data into valuable insights and knowledge to detect and diagnose the infection (Allam et al., 2020). It can also track the virus; thus, control its spread in real-time (Vaishya et al., 2020). For the current COVID-19 outbreak, an AI-driven algorithm detected and warned about it on 31 December 2019, seven days before the WHO’s official notice (Allam et al., 2020). Similarly, a company called “MetaBiota (USA)” which is concerned about epidemics monitoring, used a predictive tool to detect and warn some countries like Japan, Thailand, Taiwan, and South Korea of the coronavirus outbreak in their countries seven days before it was officially announced (Heilweil, 2020).

Neural networks were used as a tool for the prediction of the number of COVID-19 cases (Wiezorek et al., 2020a; Wiezorek et al., 2020b). The results were compared to the actual data obtained in different regions. For instance, in their first study, Wiezorek et al. (2020a) developed a complex artificial neural networks model composed of multiple layers to precisely predict the COVID-19 cases count and spread (Wiezorek et al., 2020a). In addition, this model considered the geographical conditions, i.e., location, latitude, and longitude, as input data. The training data was extracted from the actual number of cases obtained from each region over two weeks before developing the model. Although it is difficult to obtain accurate predictions using limited data, this model maintained a high accuracy of spread estimation. Moreover, in the second or further study, the authors introduced new techniques to their model to make it adaptive, i.e. adjust itself.
automatically according to the new real-time data and pattern (Wieczorek et al., 2020b). Fortunately, the observed overall prediction accuracy was as high as 88%, hitting the ceiling of 99% for some specific regions. Another factor considered by the authors was the time of prediction. Their concern was to minimize this time for a possible fast prediction. Thus, NAdam training model was utilized as it offers the shortest training time while guaranteeing the highest possible efficiency based on extensive tests conducted by the authors.

Machine learning was also used as a detection tool of the disease itself upon evaluating various diagnostic sources such as X-ray, computed tomography (CT), and ultrasound images (El-Rashidy et al., 2021). Convolutional neural network (CNN) based models have been commonly used in several studies for detection and prediction purposes (Jaiswal et al., 2020; Jin et al., 2020; Ozturk et al., 2020; Sedik et al., 2020). Furthermore, ML and deep learning (DL) models can help fast drug repurposing or even drug discovery for a novel disease, such as in COVID-19 (Ivanov et al., 2020; Jain et al., 2021; Mohanty et al., 2020; Zhou et al., 2020). These examples make it clear that utilizing ML and DL in real-world applications such as detecting individual cases, predicting numbers of cases, spread modeling, and drug development is a must and not a luxurious action.

This comprehensive study aims to review and analyze the primarily used techniques for viral early detection, focusing on the detection of the SARS-CoV-2 virus. The investigated methods include WBE, air biosensors, and artificial intelligence. This paper investigates the literature and analyzes the most commonly applied WBE techniques and the state-of-art technologies used for virus surveillance. Moreover, it reviews the applicability and potential of the AI and ML algorithms for viral outbreaks. Based on the existing literature reports, a novel framework combining WBE and AI/ML technologies has been proposed in this paper. This framework exploits an ensemble of ML and DL techniques that can help predict and forecast future viral outbreaks, in addition to recommending a set of actions to mitigate the spread of the pandemic. Finally, some remarks and recommendations are provided for applying the framework proposed in this work and the best AI technologies. Additionally, the importance of collaboration between researchers/scientific community and the stakeholders to address future challenges in viral outbreaks is also discussed. The proposed framework can be tested within a collaborative project using datasets collected from different organizations, and the results can be disseminated in a separate study.

2. Modes of environmental transmission

The COVID-19 pandemic is caused by SARS-CoV-2 (Foladori et al., 2020). Coronavirus have single-stranded RNA (Dabbish et al., 2021; Yeo et al., 2020) and usually infect humans and other animals’ respiratory, gastrointestinal, and nervous systems (Chen et al., 2020a). During the past decades, coronaviruses have caused many outbreaks. In 2002, the SARS-CoV started, and in 2012, the Middle East respiratory syndrome (MERS-CoV) took away many lives. The symptoms for SARS-CoV-2 are milder than the symptoms for SARS-CoV and MERS-CoV. However, SARS-CoV-2 is considered to be more infectious than the SARS-CoV variant (Chen et al., 2020a).

As SARS-CoV-2 is spreading at a faster rate, the mutations of the virus are also observed at a very high rate. To date, more than 1000 mutations for the SARS-CoV-2 have been recorded (Chen et al., 2020a). For the SARS-CoV-2 to infect someone, the virus has to interact with the spike glycoprotein (S protein) and the host angiotensin-converting enzyme 2 (ACE2) receptor while entering the host cells (Sasaki et al., 2021; Walls et al., 2020). The SARS-CoV-2 is becoming more and more infectious as it evolves because the new mutations increase the free binding energy with the ACE2 receptor (Chen et al., 2020a). In the United States alone, a study was done by Wang et al. (2021) showed that the SARS-CoV-2 has four sub-strains and 11 major mutations. The mutation 23403A>G-(D614G) is the second-highest strain found in the United States and is the most dominant strain worldwide (Guruprasad, 2021; Wang et al., 2021; Yurkovetskiy et al., 2020). Toyoshima et al. (2020) found that the two variants, which are ORF1ab 4715 L and S614G, are associated with high fatality rates in 28 countries. The conventional transmission routes for SARS-CoV-2 are respiratory droplets, direct and indirect contact with contaminated surfaces. However, there can be other possible transmission ways depending on the virus load and the surrounding environmental conditions. These possible transmission paths are as follows: (i) stool and sewage, (ii) natural water, and (iii) air environments, which can be indoor and outdoor. Fig. 1 explains how SARS-CoV-2 can be transmitted through different pathways.

2.1. SARS-CoV-2 in stool and sewage

The National Institute for Public Health and the Environment (RIVM), a Dutch research institute, detected the SARS-CoV-2 in wastewater treatment plants in the Netherlands. The SARS-CoV-2 was detected in wastewater from Amsterdam Schiphol airport and in a wastewater treatment plant in Kaatsheuvel using molecular Polymerase chain reaction (PCR) methods (National Institute for Public Health and the Environment, The Netherlands, 2020b). Usually, the RNA polymerase gene is amplified to detect the existence of SARS-CoV-2 using conventional PCR techniques (Lau et al., 2003). In a research study conducted by the RIVM, the SARS-CoV-2 virus was detected in 66% of stool samples taken from patients (National Institute for Public Health and the Environment, The Netherlands, 2020a). Wastewater monitoring is a useful approach for detecting viral infections among populations. The RIVM has used this approach to detect norovirus, poliovirus, and measles viruses in wastewater. The RNA of the SARS-CoV-2 can be detected in wastewater. These studies show that SARS-CoV-2 detection in wastewater indicates that the population is infected, however, they do not confirm that the virus particles present in the wastewater are contagious.

A study conducted by Wang et al. (2005) showed that SARS-CoV could survive in wastewater for 14 days at 4°C and for 2 days at 20°C. During these times, the SARS-CoV can be infectious by fecal-oral transmission. Moreover, the MERS-CoV RNA was detected in the feces of 14.6% of the patients. The MERS-CoV can remain viable in sewers at lower temperatures than the SARS-CoV-2 (Yeo et al., 2020). These results confirm that the SARS-CoV-2 and MERS-CoV can be infectious by fecal-oral transmission. In fact, the SARS-CoV-2 belongs to the same coronavirus family as SARS-CoV and MERS-CoV. It is very likely possible that SARS-CoV-2 can also be infectious by fecal-oral transmission. A study done by Zaneti et al. (2021) showed that wastewater treatment plants could be a transmission pathway for the SARS-CoV-2. However, more research is needed to assess the viability of the SARS-CoV-2 under different conditions. There is a high risk for transmission of SARS-CoV-2 through the air when wastewater operators clean the screens manually or when wastewater treatment tanks are not covered. The risk of transmission of SARS-CoV-2 for wastewater treatment workers increases as the population infected increases. The reason is that the lower the infected population is, the higher the dilution of the virus in wastewater (Zaneti et al., 2021).

2.2. SARS-CoV-2 in natural water

Since fecal-oral can be a pathway for SARS-CoV-2 transmission, surface water can be a potential transmission way for SARS-CoV-2, especially in places where there are inadequate sewage facilities and drinking water is contaminated with fecal sludge (Al Huraimel et al., 2020). SARS-CoV-2 pollutes the water bodies by several pathways. One way can be through bypassing the untreated wastewater, which is contaminated by SARS-CoV-2, to rivers, lakes, or other water streams. Another way can be via the treated wastewater effluents contaminated with SARS-CoV-2 due to the inefficient removal of viruses from the
wastewater (Bhowmick et al., 2020). The most likely pathway is the discharge of raw sewage directly into water bodies without treatment in areas with low basic sanitation (Al Huraimel et al., 2020). The SARS-CoV-2 has to be active before reaching the water bodies so that the water can be considered as a transmission pathway for the virus. As previously mentioned, the activity of the virus depends on the viral load and the surrounding environmental conditions (Wang et al., 2005). Therefore, the water bodies can be considered as a potential transmission pathway for SARS-CoV-2 when the prevailing environmental conditions can favorably maintain the activity of the infectious virus.

A study conducted on urban rivers of Quito, Ecuador, where wastewater was discharged directly to river streams without any treatment, evaluated the presence of SARS-CoV-2 in river streams. The samples were taken from three different locations along the Quito rivers during the peak of SARS-CoV-2 cases in Ecuador on June 5th, 2020. The results show that the SARS-CoV-2 nucleocapsid protein gene was present in the three sampled locations along the Quito rivers (Guerrero-Latorre et al., 2020). However, it is not confirmed yet that natural water is a transmission pathway for SARS-CoV-2. For SARS-CoV-2 to be transmissible via water, the virus needs to get concentrated enough to persist for long times in water bodies (Bilal et al., 2020). Moreover, the existing activated sludge wastewater treatment plants and drinking water treatment plants need to be more efficient in removing viruses, including the SARS-CoV-2. Treatments like ultraviolet (UV) disinfection and chlorination are expected to kill the SARS-CoV-2 present in contaminated water (Al Huraimel et al., 2020; Pecson et al., 2020). Chlorine doses for removal of bacteria and viruses (including SARS-CoV-2) will change depending on the contact time, chlorine demand, water characteristics, and discharge requirements. However, the chlorine dose often varies between 5 and 20 mg/L (Environmental Protection Agency, 1999). Disinfection using UV radiation is sufficient to inactivate the SARS-CoV-2. Published data suggests that UV doses less than 5 mJ/cm² are enough to ensure that the effluent water stream is free from coronaviruses (Pecson et al., 2020).

2.3. SARS-CoV-2 in the air environment (indoor and outdoor)

According to the WHO, the primary method for SARS-CoV-2 spread is direct contact with an infected person (Sohrabi et al., 2020). However, airborne transmission of SARS-CoV-2 is a possible pathway since the other coronaviruses like MERS-CoV and SARS-CoV were transmitted via airborne droplets (Hadei et al., 2020; Morawska and Cao, 2020). Several studies were conducted to assess the likelihood of the spreading of the SARS-CoV-2 via aerosol transmission. A study done by Faridi et al. (2020) did not detect any positive SARS-CoV-2 readings from the samples taken 2 to 5 m away from patients’ beds at the Imam Khomeini Hospital complex in Tehran, Iran. Another study conducted at the SARS-CoV-2 center in Singapore showed negative results for the aerosol transmission of the SARS-CoV-2 (Ong et al., 2020a).

On the other hand, at the University of Nebraska Medical Center (USA), thirteen SARS-CoV-2 patients were examined for viral shedding by collecting air samples from each individual. Viral contamination was detected in all samples showing that airborne transmission can be a pathway for SARS-CoV-2 spreading (Santarpia et al., 2020). Moreover, Kenarkooohi et al. (2020) examined the aerosol dispersion of SARS-CoV-2 in a hospital in Iran that treats patients diagnosed with SARS-CoV-2. Two viral RNA air samples were tested and detected positive out of six samples collected from the intensive care unit. Kenarkooohi et al. (2020) indicated that indoor environments can be a potential place for the aerosol transmission of the SARS-CoV-2. Aerosols transmission can be through coughing, sneezing, deep breathing, or talking. As the loudness of voice increases during talking, the rate of particle emission increases; therefore, the risk of aerosol transmission becomes higher (Santarpia et al., 2020).

In outdoor environments, apart from the human-to-human transmission, the SARS-CoV-2 can be transmitted through respiratory droplets of an infected person; while talking, sneezing, breathing, or coughing (Freeman and Eykelboosh, 2020). Particles emitted from the infected person can settle quickly if their size is larger than 5 μm;
however, if the particle size is smaller than 5 μm, it can stay suspended in the air. The aerosol transmission of large particles (> 5 μm) is limited to a distance of 1 to 2 m. On the other hand, the small particles (< 5 μm) can travel longer distances and get inhaled by another person (Bourouiba, 2020). The risk of infection by aerosol droplets in outdoor environments depends on the aerosols particle size and the virus’s ability to contain aerosols to be concentrated enough to cause an infection when inhaled by another person (Freeman and Eykelbosh, 2020). In general, outdoor environments have a lower risk of SARS-CoV-2 transmission than indoor environments (Weed and Foad, 2020).

2.4. Indirect effects on the air quality due to SARS-CoV-2 (indoor and outdoor)

Air pollution has been increasing worldwide, affecting public health (Amin, 2019). However, with the SARS-CoV-2 pandemic, people stayed in their homes, and businesses started to decrease their workload. As a result, air quality improved, and air pollution had reduced; however, this effect is expected to be only temporary (Zambrano-Monserrate et al., 2020). Mostafa et al. (2021) conducted a study to see how the lockdown measures for combating SARS-CoV-2 affected Egypt’s air quality, focusing on the two most populous cities Cairo and Alexandria, between the 15th of March and the 1st of May. They compared noise levels, nitrogen dioxide, carbon monoxide, greenhouse gases, and ozone levels before the pandemic (2015–2019) and during the pandemic (2020). The results showed a 75% decrease in the noise level, a 5% decrease in carbon monoxide emissions leading to an average concentration of 120 ppb. Greenhouse gases have decreased by 4%, while a 15% and 33% reduction in nitrogen dioxide was achieved lead to a concentration reduction estimated by 1.5 × 1015 molecules/cm² over Cairo and Alexandria. On the other hand, the results showed a 2% increase in the ozone levels over Cairo and Alexandria.

Another study conducted by Zambrano-Monserrate et al. (2020) also showed decreased greenhouse gas emissions, nitrogen dioxide emissions, and noise levels during the pandemic. The decrease of nitrogen dioxide emissions and the increase of ozone levels during the pandemic was also documented by Sicard et al. (2020) and Siciliano et al. (2020). Also, it was found that as a result of the lockdown in China, NO2 and PM2.5 emissions in Wuhan city were reduced by 22.8% (2020). Also, it was found that as a result of the lockdown in China, NO2 and PM2.5 emissions in Wuhan city were reduced by 22.8% (2020). The results showed a 75% decrease in the noise level, a 5% decrease in carbon monoxide emissions leading to an average concentration of 120 ppb. Greenhouse gases have decreased by 4%, while a 15% and 33% reduction in nitrogen dioxide was achieved lead to a concentration reduction estimated by 1.5 × 1015 molecules/cm² over Cairo and Alexandria. On the other hand, the results showed a 2% increase in the ozone levels over Cairo and Alexandria.

Another study conducted by Zambrano-Monserrate et al. (2020) also showed decreased greenhouse gas emissions, nitrogen dioxide emissions, and noise levels during the pandemic. The decrease of nitrogen dioxide emissions and the increase of ozone levels during the pandemic was also documented by Sicard et al. (2020) and Siciliano et al. (2020). Also, it was found that as a result of the lockdown in China, NO2 and PM2.5 emissions in Wuhan city were reduced by 22.8 μg/m³ and 1.4 μg/m³, respectively (Zambrano-Monserrate et al., 2020). As a conclusion of that study, a decrease in PM2.5 and NO2 emissions were observed during the pandemic in several countries, including Italy, Spain, China, Germany, France, and Egypt (Mostafa et al., 2021; Zambrano-Monserrate et al., 2020).

It is well-known that indoor air quality has a significant effect on human health. Each year, 3.8 million people die from diseases related to indoor air pollution (Nwanaji-Enwerem et al., 2020). The sources of indoor air pollution include, but are not limited to, cooking, cleaning, candle burning, smoking, insulation, and personal care products (Habre et al., 2014). During the pandemic, people stayed at their homes most of the time, which led to an increase in indoor air pollution. In houses, individuals were aware of the risk of SARS-CoV-2, increasing the usage of cleaning products, especially disinfectants with hazardous chemicals, which negatively affected indoor air quality in households (Nwanaji-Enwerem et al., 2020). The components of household cleaning products are terpenes, sodium hypochlorite, acetic acid, or ammonia. These chemicals can undergo various reactions leading to an additional pollution source at homes (Weschler and Carslaw, 2018). Furthermore, the prolonged stay at home increased the indoor pollutants associated with cooking (Nwanaji-Enwerem et al., 2020). These pollutants include nitrogen dioxide, nitric oxide, polycyclic aromatic hydrocarbons, acrolein, and nitrous acid (Weschler and Carslaw, 2018).

3. Wastewater-based epidemiology and surveillance

Surveillance and monitoring programs for COVID-19 at wastewater treatment plants are encouraged to be established to assess its abundance (Naddeo and Liu, 2020). Wastewater provides an opportunity for virus surveillance since the viruses excreted by patients can be traced in a wastewater treatment plant from a plant that serves a population. Additionally, this technique can be adequate to estimate the number of infected people during the early stage of the outbreak (Ahmed et al., 2021; Mallapaty, 2020; Wu et al., 2020b). The early tracing of the SARS-CoV-2 can help the health entities and politicians to adopt policies that can ensure the persistence of healthcare systems against its collapse and overwhelming (Al Huraimel et al., 2020). Wastewater plants usually serve large populations with more than one million inhabitants (Mallapaty, 2020). Hence, wastewater surveillance can act as a supplementary approach to clinical testing campaigns to estimate the spread of SARS-CoV-2 in a community (Wu et al., 2020b).

WBE can provide more opportunities than clinical testing, especially for developing countries with limited resources and a lack of clinical testing equipment. However, it still cannot take over the clinical testing, which is still the optimum choice for the infected patients’ determinations (Al Huraimel et al., 2020). Additionally, clinical testing is challenging, especially for large populations where testing is time and labor-intensive (Mao et al., 2020). Besides that, WBE can estimate the total infected citizens, including those asymptomatic and presymptomatic (Lodder and Husman, 2020; Mallapaty, 2020). One of the advantages of the WBE is that it can detect viruses at a very low level, which is essential during the beginning of an outbreak or the end of the stationary phase after the healthcare systems interventions (Ahmed et al., 2020).

It can be noticed that most of the viral detection techniques are molecular-based. Additionally, the time required, costs, and sensitivity of those techniques vary significantly. The abundance of coronavirus in a population is tracked by wastewater samples determined by the amount of viral RNA excreted in feces per capita. Therefore, this number can be extrapolated to determine the number of infected people using the concentrations of viral RNA in wastewater (Mallapaty, 2020). To determine the RNA concentration, the sampling preparation process demonstrated in Fig. 2.

The WBE has several challenges, especially when building quantitative predictions from the viral RNA, which leads to significant inaccuracies in estimating infected cases. These uncertainties are due but not limited to (i) the complexity of wastewater matrices and the dilute nature of biomarkers in wastewater; (ii) inabilities in pinpointing the suitable sample locations; (iii) non-adequacy of current sampling techniques; and (iv) the need for effective concentrating virus methods (Al Huraimel et al., 2020).

Table 1 demonstrates a summary of the commonly used methods used for viral detection in wastewater. The following subsections present these methods in more detail, along with the testing challenges. Besides, Table 2 illustrates the various approaches used by different countries to detect viruses in wastewater.

3.1. Real-time reverse transcriptase - polymerase-chain-reaction testing (RT-qPCR)

The Reverse Transcription Polymerase Chain Reaction (RT-PCR) is a technique that combines reverse transcription of RNA into DNA and amplifies a specific DNA part with a polymerase chain reaction (Freeman et al., 1999). Real-time PCR (RT-qPCR) is mainly used to quantify the amount of a specified RNA that can be performed by fluorescence monitoring of amplification reactions. The RT-qPCR was proved to be a sensibly and robust method for the early detection of SARS-CoV-2 during outbreaks (Randazzo et al., 2020). The genetic materials of SARS-CoV-2 have been detected in different wastewater samples worldwide, including Australia, China, France, Germany, Italy, India, Japan, Netherlands, Spain, United Arab Emirates, and the United States of America (Ahmed et al., 2020; Arora et al., 2020; Balboa et al., 2021; Fongaro et al., 2020; Haramoto et al., 2020; Hasan et al., 2021; Kumar et al., 2020b; La Rosa et al., 2020; Rimoldi et al., 2020; Westhaus et al., 2021; Wu et al., 2020a; Wu et al., 2020b;
The RT-PCR was used as a WBE tool and detected viral RNA on several occasions where the number of confirmed cases was low, as in the Netherlands and Spain (Lodder and Husman, 2020; Randazzo et al., 2020). It illustrates the surveillance system’s sensitivity and its capabilities as an early warning technique for viral outbreaks.

Kumar et al. (2020b) investigated the presence of the SARS-CoV-2 by using RT-qPCR in India at the Old Pirana wastewater treatment plant, which receives an influent from a hospital that accommodates COVID-19 patients. It was noticed that the quantity of the viral RNA in the wastewater corresponded to the number of declared infected patients at the hospitals. The viral concentration in the influent was ranging between $5.6 \times 10^1$ and $3.5 \times 10^2$ copies L$^{-1}$. Gonzalez et al. (2020) studied the frequency of detections and concentrations for twenty-one weeks in south-eastern Virginia, the USA using RT-qPCR. During this period, the concentration range was between $10^1$ and $10^4$ copies per 100 mL in samples where viral RNA was detected. The normalized loading rate fluctuations at the WWTP were aligned with the number of known outbreaks in the study.

### 3.2. Nucleic acid sequence-based amplification (NASBA)

Nucleic Acid Sequence Based Amplification (NASBA) is another method that can be used for virus detection in wastewater. The NASBA technique relies mainly on the detection of the RNA. NASBA uses several enzymes to amplify the numerous target nucleic acid sequences of the SARS-CoV-2. In contrast to the other nucleic acid amplification methods, NASBA can amplify directly from an RNA without the need for a reverse transcription step (Farkas et al., 2020). The three used enzymes, T7 RNA polymerase, reverse transcriptase, and RNase H, can amplify one strand template of RNA (Hryniszyn et al., 2013). NASBA has the potential to detect a low concentration of DNA or RNA sequences during 15–60 min and at a temperature of 37–65 °C (Farkas et al., 2020).

NASBA was used by Jean et al. (2001) to determine the concentration of Hepatitis A Virus (HAV) in wastewater from the Saint-Nicolas wastewater treatment plant in Canada. Even though there was significant bacterial contamination in the wastewater ($2 \times 10^5$ fecal coliforms/mL), the results showed high HAV detection. Hence, the bacterial presence doesn’t affect the analysis. NASBA was able to detect 0.4 ng of target RNA/mL compared with 4 ng/mL for RT-PCR, indicating NASBA’s strong potential (Jean et al., 2001). However, it is noteworthy to mention that several reasons limit the adoption of NASBA as a tool for viral detection in wastewater. This technique has difficulties in producing reliable and quantitative results. Moreover, its costs are relatively higher than the PCR test in terms of viral detection in water and the environment (Walker et al., 2017).

### 3.3. Biosensors

Biosensors are portable analytical instruments that can detect biological pathogens (i.e., SARS-CoV-2) and proteins with measurable physicochemical properties by converting biological reactions into measurable electrical signals (Goode et al., 2015; Mehrotra, 2016). Given that viral detection, the biological molecule could be the isolated microorganism, the viral nucleic acid, viral protein particles, or the host’s antibodies against the virus. A biochemical response is obtained specific to the viral particles or their genome. The transducer captures the recognition response converting it into a quantifiable signal computed by a digital detector, pursued by analyzing and presenting the detected signal via a processor (Samson et al., 2020; Yang et al., 2020). Biological responses can be transduced to measurable signals once they interacted with their specific targets (Neethirajan et al., 2017). Aptamer-based...
biosensors (Aptasensor) have a single-stranded DNA that can bind to the target protein or DNA with high accuracy, producing a detectable signal once it gets bound (Farkas et al., 2020).

Biosensors have been successfully used to detect norovirus in aquatic samples. The aptasensors used are based on electrochemical, fluorescence, colorimetric, and surface plasmon resonance detection platforms (Farkas et al., 2020; Weerathunge et al., 2019). Generally, biosensors are resistant to environmental inhibitors, allowing high recovery of it and allowing low detection limits (Schilling et al., 2018). Moreover, the biosensors can detect viruses in a short period of 7–16 min (Liu and Zhu, 2005). Furthermore, the interaction between molecules can be observed in real-time (Liu and Zhu, 2005). Qiu et al. (2020) succeeded in using a dual-function plasmonic biosensor that combines plasmonic photothermal effect and surface plasmon resonance to detect SARS-CoV-2 RNA. The biosensor showed high accuracy with a low detection limit of 0.22 pM. Nevertheless, considering the necessity of rapid detection of the infection status within a population and the viral presence in the environment, biosensors can serve as a potential tool in the surveillance and detecting SARS-CoV-2 (Lahrich et al., 2021).

3.4. Nano(bio)sensors and quantum dots

The development of nano(bio)sensors is essential due to their unique physicochemical and optical properties of materials of small sizes (Suleman et al., 2021). These properties give the nano-sensors high accuracy and sensitivity besides the fast detection of viral infections. In addition, merging the nano-sized sensors with the electrochemical detection methods offers the advantage of reduced viral detection time and the ability to detect low concentrations (Bukkitgar et al., 2020). This combination is considered a cost-effective method for fast detecting of coronavirus-induced infections (Lim and Bonanni, 2020).

In order to prepare the nano-sensors, a surface modification step of the parent nano-particles should be done. One method of surface modification or functionalization is the loading or doping with other active materials having high sensitivity to viruses and pathogens (Shetti et al., 2021). In this regard, various nano-sensors were developed, including nano-spherical particles, nano-islands and nano-wires of gold, and graphene nano-particles to detect various viral infections, including SARS-CoV-2 (Antiochia, 2020; Asif et al., 2020). For instance, Au fiber Bragg grating FBG probe decorated with graphene oxide nanoparticles was used as a sensor to rapidly detect COVID-19 from patients’ saliva (Samavati et al., 2020). Additionally, Vadlamani et al. (2020) prepared a low-cost and highly sensitive sensor consisting of nano-cobalt supported on titania nano-tubes to detect the SARS-CoV-2 virus (Vadlamani et al., 2020). Fortunately, this detection method was relatively quick because of the use of an electrochemical sensor that detects the spike protein on the virus’s surface. In a recent study, a new dual-functional plasmonic biosensor containing two-dimensional gold nano-islands decorated with DNA receptors was tested as a promising COVID-19 virus analyzer (Qiu et al., 2020). It is worth noting that this

### Table 1
A summary of the commonly used methods for viral detection in wastewater.

| Method      | Sensitivity | Time required | LOD (SARS-COV2) | Specificity | Advantages                                                                 | Disadvantages                                                                 | References                                                                 |
|-------------|-------------|---------------|-----------------|-------------|----------------------------------------------------------------------------|------------------------------------------------------------------------------|---------------------------------------------------------------------------|
| RT-PCR      | 89-95%      | 2 h           | 10 copies/µL    | 93%         | Highly sensitive towards specific organisms                               | Requires expensive instruments; Requires the knowledge of the detected organisms; Requires concentration of samples; Highly susceptible to contamination by inhibitors; Have low limit of detection; Cannot differentiate between dead and alive organisms; Interference of humic acids; Relatively slow detection time | (Asif et al., 2021; Nassir et al., 2020; Pilevar et al., 2020; Rabiee et al., 2020; Russo et al., 2020) |
| NASBA       | Up to 95%   | 60-90 mins    | 1 copy/reaction | 98.9%       | Highly sensitive and specificity; Simple procedure; Relatively fast; Sample contamination resistance | Expensive kits; Few available essays for environmental sampling; False positive has higher chances over false negative | (Fakruddin et al., 2012; Lahrich et al., 2021; Matou et al., 2019; Chartratta et al., 2004) |
| Biosensors  | Up to 96.7% | 15 mins       | 0.22 pM         | Up to 100%  | Affordable cost; High sensitivity; Rapid real time detection; Simple procedure; Portability; Requires small samples | Antibody binding is influenced by environmental parameters such as pH, temperature, etc.; Binding of antibody-antigen can be influenced by reagents, solvents, or radiation; Transcription regulation is a complex process | (Choi, 2020; Ejeian et al., 2018; Lahrich et al., 2021; Qiu et al., 2020) |
| Flow cytometry (FCM) | NA         | 15-45 mins    | NA              | NA          | Rapid detection; High throughput; High accuracy; Ease of use | Requires fresh samples; Only detect alive cells | (Brown et al., 2015; Brussaard, 2004; Ma et al., 2013) |
| ELISA       | 20-80%     | 2 h           | 1 ng/mL         | >98%        | Affordable costs; Relatively simple procedure | Low specificity; Concentration of sample is required; Time consuming; Low limit of detection | (Feng et al., 2020; Pilevar et al., 2020; Sakamoto et al., 2018; Saville et al., 2001; Strelecki, 2020) |
| PFCE        | NA         | 24-26 h       | NA              | NA          | Very sensitive to genetic differences; High throughput | Too sensitive to differentiate between sources; A prior database is required; Time consuming | (Meays et al., 2004) |
Table 2
Viral detection and quantification in wastewater and sludge by various countries.

| Country            | Sample source            | Sample type | Sampling pretreatment | Concentration method | Sample volume (mL) | Concentration (copies/L) | Recovery (%) | Positive percentage (%) | Detection technique | Sampled virus          | Reference                  |
|--------------------|--------------------------|-------------|-----------------------|----------------------|--------------------|-------------------------|--------------|------------------------|----------------------|------------------------|--------------------------|
| Australia          | Raw sewage               | Composite   | → Centrifugation       | → Ultrafiltration     | 100                | 0-120                   | NA           | 22%                    | RT-qPCR              | SARS-CoV-2             | (Ahmed et al., 2020)    |
| Canada             | Primary sludge           | Composite   | → Settled at 4°C for an hour | → Precipitated with polyethylene glycol → Centrifugation | 250                | 1.7 × 10^7-3.8 × 10^3  | 9.3 ± 4.9%   | 90%                    | RT-qPCR              | SARS-CoV-2             | (D'Aoust et al., 2021)  |
| Chile              | Raw sewage               | Grab        | NA                    | → Ultrasoundcentrifugation → Phosphate buffered saline → pH 7.4 | 36                 | 4.4 × 10^3-2.3 × 10^7  | 47%          | 81%                    | RT-qPCR              | polyomavirus JC        | (Levican et al., 2019)  |
| Czech Republic     | Raw sewage               | Composite   | → Storage at 5°C       | → Flocculation → Bee extract solution → Centrifugation | 500                | NA                     | 35.5 ± 13%   | 11-27%                 | RT-PCR               | SARS-CoV-2             | (Mlejnkova et al., 2020) |
| Ecuador            | River water              | Grab        | → 1 N HCl → Storage at 4°C | → Skimmed Milk Flocculation method | 2000               | 2.07 × 10^8-3.19 × 10^9 | NA           | 100%                  | RT-qPCR              | SARS-CoV-2             | (Guerrero-Latorre et al., 2020) |
| Egypt              | Raw sewage               | Grab        | NA                    | → Viradel protocol → Filtration → Organic flocculation | 5000               | 3.9 × 10^4 to 3.3 × 10^8 | NA           | 94%                    | RT-PCR               | Pepper mild mottle virus | (Hamza et al., 2019)   |
| Egypt              | Raw sewage               | Grab        | NA                    | → Viradel protocol → Filtration → Organic flocculation | 5000               | 1.5 × 10^4 and 1.5 × 10^7 | NA           | 94%                    | RT-qPCR              | Adenovirus             | (Hamza et al., 2019)   |
| France             | Raw sewage               | Composite   | NA                    | → Centrifugation      | 11                 | 5 × 10^4-3 × 10^6     | NA           | 100%                  | RT-qPCR              | SARS-CoV-2             | (Wurter et al., 2020)  |
| Germany            | Raw sewage               | Composite   | → Centrifugation       | → Ultrafiltration → Centrifugation | 45                 | 3 × 10^7 -2 × 10^8   | NA           | 100%                  | RT-qPCR              | SARS-CoV-2             | (Westhaus et al., 2021) |
| India              | Raw sewage               | Composite   | → Filtration → Centrifugation | → PEG precipitation | 50                 | 3.5 × 10^7           | 57%          | 0-100%                | RT-PCR               | SARS-CoV-2             | (Kumar et al., 2020b)  |
| Italy              | Raw sewage               | Composite   | → Viral inactivation treatment at 56°C | → Two-phase (PEG-dextran) separation method → Centrifugation | 250                | 2.4 × 10^4 - 5.6 × 10^4 | 2.04 ± 0.70% | 45-65%                | RT-qPCR              | SARS-CoV-2             | (La Rosa et al., 2021) |
| Italy              | Raw sewage               | Composite   | → PEG-dextran → Filtration | 250                 | NA                 | NA                     | NA           | 100%                  | RT-PCR               | SARS-CoV-2             | (La Rosa et al., 2020) |
| Japan              | Raw sewage               | Grab        | NA                    | → PEG-dextran precipitation → Centrifugation | 100                | 1.2 × 10^4-3.5 × 10^7 | 45%          | 47%                    | RT-qPCR              | SARS-CoV-2             | (Hata et al., 2020)    |
| Netherlands        | Raw sewage               | Composite   | → Centrifugation       | → Ultrafiltration     | 250                | 7.9 × 10^3-2.2 × 10^6 | 73±3%        | NA                    | RT-qPCR              | SARS-CoV-2             | (Medema et al., 2020)  |
| Spain              | Raw sewage               | Grab        | → Centrifugation       | → Adsorption precipitation | 200                | 0-5 × 10^3           | NA           | 11%                    | RT-PCR               | SARS-CoV-2             | (Randazzo et al., 2020) |
| Spain              | Raw sewage               | Composite   | → Centrifugation       | → Ultrafiltration → PEG-dextran precipitation | 100                | NA                     | NA           | 100%                  | RT-qPCR              | SARS-CoV-2             | (Balboa et al, 2020)   |
| UAE                | Raw sewage               | Composite   | → Thermal deactivation | → Ultrafiltration → Centrifugation | 250                | 2.86 × 10^2-2.90 × 10^3 | NA           | 85%                    | RT-qPCR              | SARS-CoV-2             | (Hasan et al., 2021)   |
| USA                | Primary sewage sludge    | Composite   | → Storage at -80°C     | → RNeasy PowerSoil Total RNA Kit | 40                 | 1.7 × 10^6 - 4.6 × 10^9 | NA           | 20-60%                | RT-qPCR              | SARS-CoV-2             | (Pecia et al., 2020)   |
| USA                | Raw sewage               | Composite and grab | → Centrifugation      | → Ultrafiltration → Electro-negative membrane | 1000               | 0-7.5 × 10^3           | 54-56%       | 29% (2/7)             | RT-PCR               | SARS-CoV-2             | (Sherchan et al., 2020) |
analytical device combines both the plasmonic photothermal (PPT) effect and localized surface plasmon resonance (LSPR) sensing transduction.

In addition, other forms of carbon-based nano-sensors are deemed to be excellent candidates for viral detection, such as carbon nanotubes and dots (Wang and Dai, 2015), in addition to biochars (Spanu et al., 2020). This is owed to their excellent electrochemical and optical sensing characteristics. From an economic and environmental point of view, biomass-derived from waste matrices (e.g., agricultural residues) has received much attention to prepare these carbonaceous materials (Bhat et al., 2020). Moreover, there are other types of bio-sensors such as electro-spin nano-fibers, 1-D carbon nano-tubes, and quantum dots that are small in size and have high activity and accuracy towards viral detection (Castillo-Henriquez et al., 2020; Shetti et al., 2021).

Quantum dots (QDs) are considered very promising sensors as they are highly selective (Castillo-Henriquez et al., 2020). Besides, they can detect a diverse range of substances/compounds in a wide range of clinical conditions rapidly and without the requirement of further laboratory tests. The fluorescence and photo-electrochemical features of the developed QDs enable them to detect various microRNAs, DNAs, and proteins which are evidence of their ability to detect a wide variety of viruses (Ma et al., 2018). Therefore, they were successfully used for the detection of circulating cell-free miRNAs in lung carcinogenesis (Singh et al., 2018), plant viruses (Hong and Lee, 2018), and most recently, the COVID-19 virus (Manivannan and Ponnuchamy, 2020). They were used not only for the detection but also for combating the virus by hampering the genomic replication of the viral RNA.

3.5. Nucleic acid staining with fluorescent dyes (NASFD)

Highly fluorescent nucleic acid dyes can be used to detect the SARS-CoV-2 virus in wastewater, wherein the wastewater samples pass via a filter with a pore size of 0.22 μm (Maestre-Carballa et al., 2019). After that, the nucleic acids in the SARS-CoV-2 virus particles get stained by a fluorescent dye that allows the formation of fluorescent dots with dimensions more significant than the actual virus particles. These fluorescent dots appear after the excitation of the fluorescent dye bounded with the nucleic acids. Thus, the fluorescent dots will be counted as SARS-CoV-2 virus particles by viewing the sample under an epifluorescence microscope (Corpuz et al., 2020).

One of the most common dyes that have been used widely in recent research is SYBR Green I, whose blue emission is observed with the help of an epifluorescence microscope (Tonkrongjan et al., 2019). Wu and Liu (2009) used SYBR Green I to count viruses from influent, effluent, and sludge samples. Also, Otawa et al. (2007) adopted a direct count method with SYBR Green I as a staining agent to count the number of virus particles in mixed liquor-activated sludge samples.

One of the pros of using this method is that counting stained virus particles even under lower magnifications is possible; thus, the need for transmission electron microscopy (TEM) instrumentation can be avoided (Ortmann and Suttle, 2009). Also, there is no need to cultivate the samples in the laboratories since the virus particles can still be counted using this method. Moreover, it will be easier to differentiate the virus particles with nucleic acids from virus-like particles without nucleic acids through DNase treatment (Forterre et al., 2013).

3.6. Flow cytometry (FCM)

Another method employed to count SARS-CoV-2 virus in wastewater samples is by combining fluorescent nucleic acid-specific dyes with flow cytometry (FCM). Firstly, in the FCM unit, wastewater samples are diluted with the buffer solutions, stained with the fluorescent dyes (e.g., propidium iodide, thiazole orange, SYBR Green I, SYBR Green II, and SYBR Gold), and injected into the flow cytometer. After that, the SARS-CoV-2 virus particles enter a stream in a single file under the hydrodynamics effect of the surrounding sheath fluid. Then, the individual particles intersect with a monochromatic light beam, usually from an argon-ion laser. Finally, each particle’s interactions with the incident laser beam produce scattered fluorescence particles that can be collected by detectors and get analyzed as scatter and fluorescence intensity, respectively (Corpuz et al., 2020).

Furthermore, in the FCM method, the wastewater samples are stained with a fluorescent dye that binds selectively either to the DNA or RNA. Hence, the intensity of the DNA/dye and RNA/dye complexes’ fluorescence correlates with the sample’s DNA/RNA content (Adan et al., 2017). The positives about using FCM are its high accuracy and quantification speed (Brown et al., 2019). The FCM approach was used by Huang et al. (2016) to quantify the virus particles in activated sludge and effluent samples from three wastewater reclamation plants. According to another study done by Brown et al. (2019) on determining the number of virus particles from activated sludge samples using FCM, it was shown that the FCM method has a higher sensitivity than that of epifluorescence microscopy (EFM). Consequently, it can be said that FCM has a higher sensitivity and quantification speed in counting virus particles compared to the EFM method.

3.7. In situ fluorescence

A different method that can only be performed in specialized laboratories is fluorescence-based analysis. The conceptual design of a fluorescence instrument has been presented by Pollard (Pollard, 2012) that in-situ and online monitoring of viruses’ numbers in varying water matrices, including effluent wastewater, can be done. The proposed fluorescence device has an inline filter to remove any bacteria present in the wastewater sample. After this filter, a mixing coil is placed, where DNase and the fluorescent probe SYBR-Gold are added to the sample to form the DNA/RNA-SYBR viral complex. After that, this mixture is directed to reverse osmosis (RO) filter to concentrate the viruses. The RO concentrate, stained viral particles, passes across a unit to measure the fluorescence signal. While the permeate becomes the background fluorescence that is subtracted from the sample viral fluorescence.

Despite that the proposed device is still under the conceptual stage, this method’s tests have shown that the excitation-emission matrix (EEM) fluorescence intensity has a linear correlation ($r^2 = 0.97$) with the viral count.

3.8. Immunofluorescence assay (IFA)

Immunofluorescence (IFA) is a method that can be used to quantify and analyze the infectivity of SARS-CoV-2 virus. In this method, an infected cell culture sample is adsorbed on a microscopic slide. Next, the fixed sample is incubated sequentially with a specific antibody and a fluorescent chemical-conjugated secondary antibody that recognizes the former to detect the viral protein antigen. Finally, the fluorophore-conjugated antibody is fluoresced under optical excitation, where the antigen-antibody complex appears as a fluorescent particle under a fluorescence microscope (Corpuz et al., 2020; Im et al., 2019).

Calgua et al. (2011) used this method successfully to quantify HAdV and JCPyV in raw sewage inflowing into wastewater treatment plants. The results showed that IFA has one higher order of magnitude in sensitivity compared with other cell culture methods used in this study. Schlindwein et al. (2010) used the IFA method to verify human adenovirus 2 (HAdV 2) viability in an activated sludge and effluent samples from a wastewater treatment plant.

3.9. Enzyme-linked immunosorbent assay (ELISA)

The enzyme-linked immunosorbent assay (ELISA) is a method utilized to detect the presence of microbial antigens in various matrices. ELISA depends on the antigen-binding principle to its specific antibody, causing a change in color or fluorescence due to the resultant enzyme activity.
This method works as follows: firstly, antigen-binding at a specific antibody is immobilized on a surface, usually in a set of 96-well microtiter plates. Secondly, a second enzyme-linked antibody, which is specified for the same antigen, is utilized to form an antibody-antigen-antibody sandwich. Then, the enzyme-coupled antibody is reacted with a substrate that alters its color when modified by the enzyme. Finally, the color change or fluorescence is correlated with the probed antigens’ concentration in the sample (Corpuz et al., 2020). This method was utilized successfully by Atabakhsh et al. (2020) to detect rotavirus in influent and effluent samples of wastewater treatment plants. In that study, ELISA methods helped determine the removal efficiencies of rotavirus from urban wastewater treatment systems.

3.10. Pulsed-field gel electrophoresis (PFGE)

Pulsed-field gel electrophoresis (PFGE) is a method that uses a pulsating electric field that enables the separation of high molecular weight DNA fragments of the SARS-CoV-2 virus according to their molecular sizes (Corpuz et al., 2020; Nassonova, 2008). In this method, two separate electrodes are used to generate alternating electric fields that cause reorientation of the molecules periodically to align to the imposed electric field (Le et al., 2017). Molecular sizes and DNA molecules’ charges affect their ability to re-orient themselves and respond to the set modulated electric fields. Smaller DNA molecules usually take a shorter time to re-orient themselves by migrating through the gel matrix pores towards the new anodes; while, larger molecules take a longer time. Also, the larger DNA molecules, which migrate slower than the set pulse time, tend to migrate as one band via the gel matrix. Le et al. (2017); Lopez-Canovas et al. (2019). Furthermore, the band pattern, which is formed by a viral community works as its fingerprint, where the number of the formed bands estimates the number of different viruses (diversity) in a sample (Corpuz et al., 2020).

The PFGE method was applied by Otawa et al. (2007) to estimate the diversity of viruses in activated sludge samples from 14 different wastewater treatment plants. The results showed that the prevailing sizes of viral DNAs in the samples were in the range of 40–70 kb. The PFGE method was also able to demonstrate the mutual similarity in the detected viral species from the different activated sludge wastewater treatment plants based on the similarity of the band patterns. Another study was done by Wu and Liu (2009), where they used PFGE to determine the virus diversities in influent, primary settling tank, effluent, and sludge samples from a municipal wastewater treatment plant. The results revealed that the activated sludge, anaerobic digestion sludge, and effluent had the largest number of bands; besides, there were similar band patterns between the influent and primary settling tank samples. The results also showed that the dominant sizes of viral DNAs were between the ranges of 30–80 kb and 200–350 kb.

3.11. Transmission electron microscopy (TEM)

Optical microscopes cannot detect viruses. However, Electron microscopes can identify and determine viruses (Corpuz et al., 2020). Transmission electron microscopy is one of the earliest techniques to identify and classify viruses based on morphology (Roingeard et al., 2019). The working principle of a TEM depends on negative staining, which affects the ability to re-orient themselves and respond to the set modulated electric field (Le et al., 2017). Molecular sizes and DNA molecules’ charges affect their ability to re-orient themselves by migrating through the gel matrix pores towards the new anodes; while, larger molecules take a longer time. Also, the larger DNA molecules, which migrate slower than the set pulse time, tend to migrate as one band via the gel matrix. The examination showed differences depending on the sampling location. For example, a wide diversity of Eukarya viral particles and bacteriophages were detected in the influent. Similarly, the same observations were noticed in the WWTP effluent and sludge. The examination of wastewater samples by TEM also elaborated the diversity of viral morphologies with filamentous viral particles. Moreover, the TEM examination showed that 300 nm length of viral filamentous particles are typical for tobamoviruses (Bacnik et al., 2020).

3.12. Testing challenges in wastewater

The testing for SARS-CoV-2 in wastewater is not fully standardized yet and still lacks a fully standardized protocol for accurate detection and determination (Sherchan et al., 2020). Al Huraimel et al. (2020) illustrated that quantitative determination of infected citizens by SARS-CoV-2 from wastewater is a challenging process due to several reasons such as (i) the amount of viral RNA exists in feces, (ii) the lack of representative samples along the periods, (iii) the detection ability of kits at low concentrations, (iv) the variability of viral RNA excretion rates between inhabitation, (v) the vague links between the predictions of the number of cases and the amount of detection RNA concentration in wastewater, (vi) the necessity to evolve envelop virus concentrating methods (Ahmed et al., 2020; Mallapaty, 2020).

4. Coronaviruses air sampling and detection

4.1. Coronaviruses air sampling

An essential step in microbial air sampling is to eliminate any damage that could affect the microbial cell by limiting the microbial stress (e.g., long sampling time and dryness effect) during the sampling procedure. Furthermore, preserving the cell viability is not as important as ensuring the nucleic acid integrity, especially for RNA viruses, as the RNA possesses high instability (Verreault et al., 2008). Maintaining optimum sampling conditions (sampling volume, flow rate, and time) also plays a major role in precise sampling and targeted virus detection. Multiple factors can affect the indoor sampling of SARS-like viruses, including air movement, ventilation, humidity, and temperature (Guo et al., 2020). For hospital air sampling, patients’ numbers inside the room, occupancy, distances from the patient’s bed, and activity during sampling could influence the results (Cox et al., 2020). Therefore, surrounding factors such as sampling environmental conditions and sampler type should be assessed upon standardizing a protocol for SARS-CoV-2 air sampling (Rahmani et al., 2020).

Bioaerosols samplers are classified into four categories, as follows: liquid impingers, filters, impactors, and cyclones (Borges et al., 2021; Sung et al., 2017). For indoor air sampling, biosamplers are usually located 40–50 cm (Dubuis and Duchaine, 2021) and up to a distance of 1.5–2 m (Dumont-Leblond et al., 2021; Dumont-Leblond et al., 2020) from the resident for collecting the optimum and representative air sampling volume.

Liquid impinge-type biosamplers collect/absorb the bioaerosols into a liquid medium. Most of the recent studies have recommended using the biosampler for air sampling of SARS-like viruses (Rahmani et al., 2020). The biosampler is normally equipped with a high flow rate vacuum pump ranging from 10 to 60 L/min to collect the air samples (Rahmani et al., 2020). Due to the use of a high flow rate vacuum pump, this sampling tool may be more appropriate for outdoor air sampling rather than for collecting samples from indoor environments. However, Fardil et al. (2020) have applied this impinger technique for the sampling of SARS-Cov-2 in an indoor environment (Imam Khomeini
Hospital Complex) after replacing the high volume vacuum pump by an average flow-rate pump, although the limited flow rate capacity negatively impacted the overall effectiveness of the viral detection. It has been reported that liquid impingers are the most widely used tool for sampling viruses in the air because they help to maintain the viability and integrity of the viral particles during sampling (Rahmani et al., 2020). Additionally, the viral analyses can be conducted directly without the need for viral extraction from the medium and/or a filter surface (Pan et al., 2019; Rahmani et al., 2020).

Moreover, filters are extensively used for viral air sampling because of the high efficiency in capturing virus-containing aerosols having a particle size ~ 500 nm. The four basic filtration mechanisms are interception, diffusion, inertial impaction, and electrostatic attraction. Different materials for the membrane filters have been tested in sampling airborne viruses, such as polytetrafluoroethylene (PTFE), cellulose, tightly packed cotton, gelatin, Teflon, and polycarbonate (PC) (Pan et al., 2019; Rahmani et al., 2020). Gelatin and PTFE are the most widely used filters for air sampling of SARS-like viruses (Booth et al., 2005; Fabian et al., 2009). The advantage of using the gelatin filter over the other commonly used filters is their ability to dissolve in the liquid medium that facilitates the enumeration of microbial or viral particles in cell cultures without any harmful impact on the viability of the viruses (Pan et al., 2019). Besides, gelatin filters have been commonly used in modern air sampling instruments, such as MD-8 Airscan sampler (Rahmani et al., 2020). Zhao et al. (2014) have reported that the MD-8 Airscan sampler with a gelatin filter enhanced the efficiency of collecting RNA viruses, i.e., ~100%. Similarly, PTFE filter has been recommended by the National Institute for Occupational Safety and Health (NIOSH) for sampling airborne viruses (Rahmani et al., 2020), owing to its unique structure, which offers an advantage of easy elution of the target viruses from the membrane without any interaction between them (Lindsley et al., 2017).

Although filters are very effective in collecting virus-containing particles, they also have some drawbacks, including the possibility of viral destruction in certain conditions, such as sampling of large air volumes (Pan et al., 2019; Rahmani et al., 2020). High volume sample pumps require a longer sampling time, resulting in filter dryness and thus causing damage to the viruses (Rahmani et al., 2020). Other drawbacks that were reported for gelatin filters include: (i) dissolution of gelatin filters may occur in case of sampling at high relative humidity or if high volume sample pumps are used, (ii) desiccation of viruses may take place in the case of sampling at low relative humidity (Fabian et al., 2009; Pan et al., 2019; Rahmani et al., 2020). Impactors like the Andersen 6-stage sampler and the slit sampler collect airborne particles into a liquid medium, where a vacuum pump is used to withdraw the aerosol. The collection media is then recovered, and aliquots are used for viral isolation (Pan et al., 2019). The use of a high-resolution slit-sampler in the collection of SARS coronaviruses was also addressed in a previous study (Booth et al., 2005). However, negative results were reported for all the viral cultures, although 2 out of the 10 samples were tested positive using RT-PCR. This could be attributed to the low concentrations of the viruses in the air or that these instruments are inappropriate for sampling aerosolized viruses (Pan et al., 2019).

Cyclone sampler exerts centrifugal forces on particles in order to be separated from the air flow. This sampler has been used for sampling influenza A virus (IAV) and SARS-like viruses (Pan et al., 2019; Rahmani et al., 2020). The NIOSH has developed a multistage cyclone sampler that operates at 3.5 L/min. The first stage consists of a 15 mL tube that captures aerosol particles larger than 4 μm. The second stage consists of a 1.5 mL tube that captures aerosol particles between 1 and 4 μm. While the third stage consists of a polytetrafluoroethylene (PTFE) filter that collects aerosol particles smaller than 1 μm (Chia et al., 2020). However, the cyclone sampler was not efficiently successful in collecting SARS-CoV-2 viruses present in indoor air of a hospital complex (Chia et al., 2020a).

Table 3
Summary of the commonly used air samplers for coronavirus sampling.

| Sampler | Target virus | Sampling time | Sampling flow rate (sampling volume) | Limitations | References |
|---------|--------------|---------------|-------------------------------------|-------------|------------|
| Filter: MD-8 airscan sampler | MERS-CoV | 20 min | 50 L/min (1000 L) | Dehydration of viruses during sampling | (Kim et al., 2016) |
| Filter: MD-8 airscan sampling device (Sartorius) and sterile gelatin filters | MERS-CoV | 20 min | 50 L/min (1000 L) | Dehydration of viruses during sampling; Possibility of inactivation of some fraction of the collected viruses during the extraction process of the filters | (Agranovski et al., 2004) |
| Filter: High-efficiency particulate air (HEPA) filters installed in the pipeline connecting sampler and vacuum pump | SARS coronaviruses | 4 h | 4 L/min (960 L) | May require high volume sample pumps | (Booth et al., 2005) |
| Filter: PTFE membrane filter with a pore size of 0.3 μm in a closed-face, 3-piece disposable plastic cassette attached to a personal sample pump | SARS coronaviruses | 10.5 – 13 h | 2 L/min (1260–1560 L) | Desiccation of viruses may occur in case of sampling at low relative humidity; Dissolution of gelatin filters may occur in case of sampling at high relative humidity and high sampling volume | (Ong et al., 2020a) |
| Filter: Sterilized gelatin filters with pore size 3 μm placed in styrene filter cassette | SARS-CoV-2 | 1 h | 5 L/min (300 L) | The limited flow rate capacity; Antifoam is required to overcome production of foam in the culture medium | (Faridi et al., 2020) |
| Impinger: attached to a personal sample pump with average flow | SARS-CoV-2 | 1 h | 1 L/min (60 L) | Limit the smallest cut-off size to 0.2–0.3 μm | (Booth et al., 2005) |
| Impactor: High-resolution slit-sample system | SARS coronaviruses | 18 min | 30 L/min (540 L) | Degradation of viral RNA may occur during the collection | (Booth et al., 2005) |
| Cyclone: SASS 2300 wetted wall cyclone sampler | SARS-CoV-2 | 30 min | 300 min (9000 L) | Degradation of viral RNA may occur during the collection | (Borges et al., 2021; Guo et al., 2020) |
| Cyclone Bioaerosol Sampler | SARS-CoV-2 | 4 h | 3.5 L/min (840 L) | Degradation of viral RNA may occur during the collection | (Chia et al., 2020) |
| Cyclone and filter: Cyclone sampler and 37-mm filter cassettes and 0.3-μm polytetrafluoroethylene filters | SARS-CoV-2 | 4 h | 100 L/min (400 L) | Degradation of viral RNA may occur during the collection | (Ong et al., 2020a) |
et al., 2020). It was reported that physical damage might have occurred to the viruses due to the impaction forces created during the sampling process (Blachere et al., 2009; Brown et al., 2013; Cao et al., 2011; Pan et al., 2019).

Table 3 summarizes the commonly used air samplers for coronaviruses sampling and the limitations of each sampler. Several precautions must be taken in consideration during the selection step of the air sampler for both indoor and outdoor environments. For example, it is recommended not to operate the cyclone sampler with an air-flow rate greater than 200 L/min to avoid the degradation of viral RNA in the collected samples (Borges et al., 2021). Additionally, it is recommended not to use high-volume sample pumps with filter samplers to evade viral dehydration (Robotto et al., 2021).

4.2. Coronaviruses detection techniques

Following the sampling phase, the viral RNA is extracted from the sample matrices using the viral RNA/DNA isolation kit and stored in an elution buffer for further analysis (Azhar et al., 2014; Faridi et al., 2020). Once the viral genomic material is extracted, several examinations and identification techniques are performed, as demonstrated in Fig. 3 (Corman et al., 2020; Drosten et al., 2002; Shetti et al., 2021). Nowadays, multiple identification techniques are modified to detect viruses in air samples with high efficiency as shown in Table 4. Moreover, Cheng et al. (2020) studied SARS-CoV-2 air sampling in an isolation room and highlighted that positive environmental air samples were detected only if the patient’s viral load within the clinical samples exceeded 1000 copies/mL. They also reported that the SARS-CoV-2 viral load ranging from 110 to 9400 copies/mL resulted in a positive SARS-CoV-2 in the environmental air samples (Cheng et al., 2020), with a distance ≥ 2 m between the air sample and the patient (Lednicky et al., 2020).

4.3. Reverse transcriptase real-time polymerase chain reaction testing (RT-qPCR)

The gold standard technique for SARS-CoV-2 detection in air samples is RT-qPCR owing to its high sensitivity and accuracy (Lednicky et al., 2020; Lei et al., 2020; Rahmani et al., 2020), as previously mentioned. Although RT-qPCR can precisely identify SARS-CoV-2 in several clinical specimens including, blood, saliva, sputum, stool, and urine (Peng et al., 2020; Wang et al., 2020b), various studies have proved the effectiveness of RT-qPCR in detecting SARS-CoV-2 in the air as well. This technique can predict thousands of viral copies present in 1 mL of liquid aerosol within the proper limit of detection (LOD) (Ong et al., 2020b).

Therefore, a key step in the air monitoring system is to enrich the viral concentrations, which are distributed in the environment, into smaller volumes of liquid to exceed the LOD, in conjunction with increasing the sample volume. Guo et al. (2020) have examined the indoor environment within hospitals using RT-qPCR, and they detected the presence of SARS-CoV-2 in patients’ rooms, around air outlets, and in physician’s office areas. A large sample volume (~ 9000 L) was collected using the SASS 2300 wetted wall cyclone sampler to achieve positive results (Guo et al., 2020). The advantages of using RT-qPCR are the high detection sensitivity, ease of use, relative rapid results, and the simultaneous diagnosis of multiple respiratory agents.

4.4. Reverse transcription-loop-mediated isothermal amplification (RT-LAMP)

Another technique for SARS-CoV-2 identification based on nucleic acid amplification is the reverse transcription-loop-mediated isothermal amplification (RT-LAMP), which offers an advantage of accurate and sensitive detection of SARS-CoV-2 in clinical samples (Yu et al., 2020). This novel technique is used to overcome the constraints of qPCR as long processing time. Unlike RT-qPCR, RT-LAMP is processed under the isothermal condition at constant temperature (65 °C), thus reducing the time consumption as temperature changes; therefore, the reaction lasts only 20 min (Park et al., 2020). Additionally, the LOD of RT-LAPM is only 80 copies of RNA/mL of samples, which is considered lower than the limit detected by qPCR (Huang et al., 2020). The concept of nucleic acid detection relies on magnesium pyrophosphate production upon amplification resulting in turbidity that could be seen by the naked eye and quantified using a turbidimeter (Hong et al., 2004).

Fig. 3. SARS-CoV-2 air surveillance, sampling, and identification.
Based on the LAMP technique, Zhu et al. (2020) have developed nanoparticle-based biosensors (NBS) to achieve the benefits of rapid and precise detection of SARS-CoV-2. In RT-LAMP-NBS one-step assay, the LAMP primers are designed to detect and amplify the opening reading frame fragment 1a/b sequence and nucleoprotein sequence of the SARS-CoV-2 genome in one tube. The NBS detects and analyzes the results. With the aid of this biosensor, it can predict very low concentrations of SARS-CoV-2 (12 copies) per reaction (Zhu et al., 2020). Therefore, it overcomes non-specific amplification of other non-SARS-CoV-2 genomic sequences, increasing the specificity and reducing false-positive outcomes. RT-LAMP was widely used for SARS-CoV-2 detection in clinical samples. Nevertheless, it has not been applied yet for SARS-CoV-2 detection in air samples (Thi et al., 2020).

4.5. Transmission electron microscope (TEM)

As previously discussed in an earlier section, TEM is beneficially used to recognize and characterize viruses along with various matrices. However, viral identification in air samples using TEM is challenged by the viral concentration enrichment within the sample to reach a detectable limit (Richert-Poggeler et al., 2018). Subsequently, previous studies on different coronaviruses relied primarily on large sample volumes. Kim et al. (2016) have reported MERS-CoV existence, and Agranovski et al. (2004) had positive results for SARS-CoV in air samples when the volume of the sample in both studies exceeded 1000 L (Agranovski et al., 2004; Kim et al., 2016). Another pre-requisite step to increase the viral load in the collected sample is through viral cultures. The tubing used to interconnect the sampler with the vacuum pump are nect the sampler with the vacuum pump are

The collected sample is through viral cultures. The tubing used to interconnect the sampler with the vacuum pump are

Table 4
Summary of the commonly used techniques for viral detection in air.

| Method       | Mechanism of detection | Time required* | Advantages                                                                 | Disadvantages                                                                                   | References |
|--------------|-------------------------|----------------|-----------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------------|
| RT-qPCR      | Nucleic acid amplification | 2 h            | → Sensitivity  
→ Accuracy  
→ Detects thousands of viral copies in 1 mL liquid aerosol  
→ Simultaneous diagnosis of multiple respiratory agents | → Time consuming  
→ Requires viral enrichment to reach limit of detection | (Guo et al., 2020; Ong et al., 2020b) |
| RT-LAMP      | Nucleic acid amplification | 20 min         | → Rapid results  
→ High sensitivity  
→ High accuracy  
→ Detects 12 of viral copies per reaction  
→ Greater amplification products yield  
→ Less sensitive to the presence of contaminants in the sample | → Alteration in the sample PH affects the assay performance | (Park et al., 2020; Thi et al., 2020; Yu et al., 2020; Zhu et al., 2020) |
| TEM          | Viral morphological examination | Few min       | → Effective in confirmation of viral identity | → Requires large sample volume | (Agranovski et al., 2004; Kim et al., 2016) |
| CRISPR-Cas   | Nucleic acid identification – gene editing | 40 min       | → High sensitivity  
→ High specificity  
→ Short detection time | → Susceptible for false results if a mutation took place in the target sequence | (Abduljaili, 2020; Broughton et al., 2020; Jia et al., 2020) |
| Biosensors   | Measuring biological response | 15 min       | → Detect low viral concentrations in the sample  
→ Different viral bio-recognition molecules are accessible for use (viral genome, spike protein, antibody) | → High cost and long processing time are required for target and biological matrix preparation | (Izvani, 2020; Kim et al., 2020; Qiu et al., 2020; Samson et al., 2020) |

4.6. Clustered regularly interspaced short palindromic repeats (CRISPR) – Cas (CRISPR associated) technique

Introduction of modern gene-editing CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) - Cas (CRISPR associated) technique was extremely vital in the field of viral detection (Bayat et al., 2018). This technology targets viral RNA degradation; thus, it controls viral replication within the host cell and reduces transmission (Brough et al., 2008). CRISPR-Cas system was discovered firstly as part of natural bacterial and archaeal immune responses against viruses. Afterward, scientists repurposed its use for gene-editing in mammalian cells (Jinek et al., 2012). Recently, the CRISPR-Cas system is a potential tool for COVID-19 molecular diagnosis and therapeutic approaches (Kumar et al., 2020c; Zuo et al., 2017).

Applying the CRISPR-Cas system for SARS-CoV-2 detection improved the sensitivity, specificity, and detection time (30–60 min) over traditional PCR techniques (Jia et al., 2020). Broughton et al. (2020) applied CRISPR-Cas12-based lateral flow assay to diagnose SARS-CoV-2 infection. They took swabs from the nasopharyngeal or oropharyngeal in a consistent period (around 40 min). Whereas, Hajian et al. (2019) combined CRISPR-Cas technology with FET to develop a new biological sensor. The biosensor is composed of a CRISPR-chip connected to graphene-based FET. With the advantage of skipping the amplification process, it managed to detect a 1.7 femtomolar concentration of nucleic acid in only 15 min (Hajian et al., 2019).

4.7. Biosensors for SARS-CoV-2 detection

Several biosensors were developed to detect coronaviruses such as MERS and SARS-CoV-1 to diagnose the infection rapidly. Most of the biosensors were designed based on optical systems. Moreover, the piezoelectric mechanism was used for SARS-CoV-1 detection, compared to an electrochemical biosensor for MERS identification. Huang et al. (2009) introduced a new biosensor relying on a fiber-optical mechanism to detect SARS-CoV-1 nucleocapsid protein (Huang et al., 2009). Similarly, Roh and Jo (2011) utilized the same biorecognition molecule to detect 0.1 pg/mL of SARS-CoV-1 through a quantum dots-coupled RNA aptamer chip (Roh and Jo, 2011). However, nanoparticle-based biosensors were used widely in MERS detection. Teengam et al. (2017) created pyrrolidinyl peptide nucleic acid probes via nanoparticle aggregation to develop an optical biosensor to determine MERS. These nucleic acid probes were synthesized to effectively hybridize with the complementary sequence of MERS RNA (Teengam et al., 2017).
Moreover, Layqah and Eissa (2019) utilized gold nanoparticles to modify carbon electrodes carrying MERS antigen, then quantifying the specific antibodies in the samples volumetrically (Layqah and Eissa, 2019). Since the biorecognition material, whether it is a nucleic acid, antigen, antibody, or PCR product, is typically similar across the species, the previously developed biosensors can be modified to determine SARS-CoV-2.

A recognition target is identified to design a biosensor specific for SARS-CoV-2 detection, such as any region of SARS-CoV-2 RNA, its proteins, viral proteases, or human immunoglobulins as a part of the human immune response towards the virus (Escors et al., 2001; Huang et al., 2009; Tian et al., 2020; Xu et al., 2020). The recognition method is then chosen either through nucleic acid hybridization, receptor-ligand interaction, antigen–antibody binding, or enzymatic reaction. An additional component included for the environmental sampling of SARS-CoV-2 directly from the ambient air is an air sampler. Given the fact that the airborne SARS-CoV-2 concentration in the air is relatively low (103–104 of viral RNA copies per 1 m3 of air during the COVID-19 pandemic), a complex air sampling procedure is essential. Accordingly, this implements point-of-care risk assessment, which boosts viral concentration to a detectable limit (Yang et al., 2011). Therefore, a proper aerosol to hydrosol enrichment capacity equivalent to 105–106 folds (obtained from 103 to 106 of SARS-CoV-2 genome copies per 1 m3 of air) is necessary (Kim et al., 2020). Since the particles carrying the SARA-CoV-2 virus are loaded with other compounds (e.g., proteins, salts, organic and inorganic substances, in addition to other viruses), the viral-containing particles are larger than the size of the virus. Therefore, the sampler used for airborne viral detection should be designed with filters of larger pore sizes than the actual viral size (25–400 nm) (Hogan et al., 2005).

A novel biosensor was introduced by Qiu et al. (2020) after studying the RNA sequences of both SARS-CoV viruses. They designed a specific oligonucleotide merging the sequencing of both SARS-CoV-1 and SARS-CoV-2 and measured the hybridization effect via dual plasmonic-photothermal mechanisms. The biosensor effectively detected 0.22 pM concentrations of SARS-CoV-2 (Qiu et al., 2020). Seo et al. (2020) have developed a sensor-based technique for the detection of SARS-CoV-2. The sensor was designed based on field-effect transistor (FET) biosensing ability, in which a FET graphene film is coated with a specific antibody targeting SARS-CoV-2 spike protein (Seo et al., 2020). This method provided high accuracy and sensitivity in detecting low concentrations of SARS-CoV-2 in nasopharyngeal swabs of infected patients without the need for sample pretreatment. FET biosensor was proved to detect 1 fg/mL concentration of SARS-CoV-2 spike proteins if it dissolved in a phosphate buffer saline and a concentration of 100 fg/mL. Also, if it is present in clinical transport medium (Samson et al., 2020). Although it has not been applied, it is considered a potential method for the sensitive detection of SARS-CoV-2 in the air. Extensive research is invested nowadays to reduce the drawbacks and improve these biosensors’ performance for an efficient and more reliable diagnosis of SARS-CoV-2 in different matrices.

5. Role of computational modeling in viral outbreaks

Besides theory, experiment, and simulation, computational modeling has quickly evolved to become one of the main pillars of science (Chang and Makatsoris, 2001). The advancement in computational modeling tools and technologies and their versatile adaptation in many fields has led to a significant quality improvement of the executed tasks. Among those tasks are the early detection of disease vector outbreaks and the risk map assessments (Agbehadj et al., 2020; Simsek and Kantarcı, 2020). The availability of massive global data (health and environmental) led to the advancement in computing capacities and the emergence of technologies such as Big Data, AI, ML, Internet of Things (IoT), and Cloud Computing (Allam et al., 2020). Notably, this allowed massive amounts of data collected from different sources to be retrieved and analyzed in real-time to provide informative insights and predictions, as presented in Fig. 4. Hence, the availability of predictive modeling techniques should be expected to contribute remarkably to detect future viral outbreaks using the current available data, including environmental and health datasets. Currently, the detection of outbreaks using computational methods such as AI is focused mainly on identifying the virus spread between humans using data collected from infected humans. Using the same technology, the AI can also detect outbreaks using WBE.

5.1. Artificial intelligence and machine learning for detection of outbreaks of infected persons

AI and its related ML technologies can accurately track and survey the virus spread, alarming the patients and healthcare authorities, based on the availability of representative data (Ahmad et al., 2020; Kumar et al., 2020a). These technologies can contribute to providing fast, reliable, and cost-effective tools to help manage viral outbreaks. Moreover, they can rapidly identify high-risk patients and provide risk maps and valuable insights to control infection in real-time. Table 5 summarizes the computational modeling methods, mainly AI/ML techniques used recently in different jurisdictions for forecasting the status of the SARS-CoV-2 pandemic.

Moreover, advanced ML techniques such as deep learning (DL) networks are used to detect and predict the SARS-CoV-2 pandemic. A deep learning network (nCOVnet) has also been proposed for the fast detection of the pandemic using chest X-ray images (Panwar et al., 2020). As a transfer learning approach, the authors of another study (Nayak et al., 2021) compared the performance of different pre-trained convolutional neural network (CNN) models such as AlexNet (Krizhevsky et al., 2012), VGG-16 (Simonyan and Zisserman, 2014), SqueezeNet (Iandola et al., 2016), and ResNet-54 (He et al., 2016) for detecting the pandemic using chest X-ray images. Another approach was recently proposed by Zheng et al. (2020) to predict the probability of the pandemic infection. This approach exploits the 3D chest CT volumes through the segmentation of the 3D lung region using pretrained U-Net (Ronneberger et al., 2015) architecture: after that, the segmented lung is fed into a 3D DL network for prediction.

Outside China, the AI-driven algorithms were able to detect the SARS-CoV-2 even before the WHO was informed. Consequently, several alarms were triggered to warn travelers of the risks of traveling to specific locations. On the other hand, Chinese companies did not trigger those alarms, indicating that the data-sharing restriction might have led to these severe consequences (Allam et al., 2020). BlueDot was one of the companies with a successful history of detecting outbreaks (Dananjayan and Raj, 2020). It had previously detected the outbreak of the Zika virus in Florida in addition to spotting the COVID-19 pandemic wave 9 days before the WHO released the alerting statement of the emergence of coronavirus (Bogoch et al., 2016). BlueDot used the available data from news reports, airline flights, and previous outbreaks in predicting the SARS-CoV-2 emerging from Chinese regions. In recent work, eight of ten cities that could be potentially affected by the new coronavirus were successfully predicted (Bogoch et al., 2020). Metabiota is another company that successfully used a similar approach by using Big Data analytics to exploit flight data to forecast infections, several days before the first reported case, in countries at risks such as Japan, Thailand, Taiwan, and South Korea (Allam et al., 2020). Companies such as BlueDot (Netherlands), Metabiota (USA), and others that can exploit the breakthrough of AI technologies and big data analytics can help the healthcare sectors, governments, and businesses minimize the risk resulting from unexpected viral outbreaks. Accordingly, various AI techniques, big data analytics, and their applicability in contact tracing, detection of viral outbreaks, and recommending control strategies are discussed in the following subsections.

5.2. Contact tracing

As the SARS-CoV-2 is spread mainly from person to person through droplets, discharges the nose, or contact transmission, therefore, contact
tracing can be an appropriate method for monitoring and control the spread of the outbreak (World Health Organization, 2020a). The contact tracing process relies on using infected patients’ data to forecast and predict the virus spread. Multiple COVID-19 infected countries used digital contact tracing, mainly smartphone applications synergized with different technologies such as Bluetooth, Global Positioning System (GPS), Social media, contact details, network-based APIs, mobile tracking data, card transaction data, and system physical address (Lalmuanawma et al., 2020). Such data from different sources change over time and grows quickly; hence, it is called “big data” (Devakunchari, 2014). Many platforms can handle big data challenges in terms of volume, velocity, and variety. Storm (Apache Storm), S4 (Neumeyer et al., 2010), Kafka (Apache Kafka), Apache Spark (Apache Spark™ - Unified Analytics Engine for Big Data), and Flink (Apache Flink: Stateful Computations over Data Streams) can handle data stream processing that can process large volumes of data in a relatively short time span. The digitalization of contract tracing and the exploitation of big data analytics are superior by their virtual real-time fast abilities compared to the non-digital process. The collected individual personal data through digital applications are then used by ML and AI techniques to track citizens who are vulnerable to the viral outbreak due to their recent contact with infected persons. Lalmuanawma et al. (2020) listed over 36 countries that have successfully used digital contact tracing by using either centralized, decentralized, or hybrid techniques.

Digital contact tracing is limited mainly by privacy and control over data and data security breaches. According to Sohrabi et al. (2020), several contact tracing applications, including those produced by countries, violated privacy laws and were reported to be unsafe. Several countries already have their contact tracing applications to track the spread of viruses; however, to fight the viral spread, a global action should be performed for a better tracing. Wen et al. (2020) have analyzed the documentation of 41 released contracted tracing applications for COVID-19 contacts tracing. The results showed that several applications could expose the identity information to enable fingerprinting and tracking specifying users. This addresses issues regarding the question of who should be capable and has the right to trace people. Confidential and publically available data should be addressed appropriately. Restrictions of using or sharing data would also be a matter of concern.

5.3. Accurate detection of outbreaks: motivations to combine WBE and AI

Water is a vital resource for society, the economy, and the environment (Abdeldayem et al., 2020), especially during a pandemic where hygiene is critical for spreading the virus. During the last decade, the water sector has been navigating parallelly into two transformations (Garrido-Baserba et al., 2020). The first transformation is the fourth water revolution, “Water 4.0,” which aims to more sustainable and rational water resources management (Sedlak, 2014). The second one is the digitalization of the water sector and its empowerment with AI and Big Data to initiate new functionalities in the water and wastewater management sectors (Poch et al., 2020; Ramírez Calderón et al., 2020). In the context of viral detection, AI-based tools that exploit the data collected from smart city sources have the potential to monitor and control viral outbreaks.

AI has been conventionally used for controlling state variables in wastewater treatment plants or outbreaks detection from Big Data.
As previously mentioned, WBE can provide great opportunities to detect viral outbreaks, especially for countries with limited resources to carry out clinical tests. However, upon conducting an extensive search, it was found that no approach in the scientific literature combines AI with WBE for the detection of viral outbreaks. Hence, there is a clear need to fill this gap and propose new frameworks for such purposes. In response to this need, this paper proposes a framework for using AI to predict viral outbreaks using WBE. The framework is presented in the following subsection.

5.4. The proposed WBE-AI framework

The proposed framework depends on the concept of data-driven modeling that exploits an ensemble of diversified types of ML methods; predictive, descriptive, and prescriptive. This is because no single method can work perfectly in all situations and all types of data. These methods use various algorithms from statistical learning, AI, and combinatorial optimization. It is essential to combine human expertise with additional knowledge extracted by the AI/ML methods (Ragab et al., 2019). The role of human expertise is central to the framework proposed in this work. First, the expert quantifies the data quality by using an assessment technique consistent with his/her subjective knowledge. This is to provide a clear assessment of the data uncertainty. The second task is to perform the verification and validation procedures to ensure the trustworthiness of the resulting predictive/descriptive/prescriptive models. This is based on his/her subjective knowledge and enables models to be updated with minimal efforts from human experts.

For the case of viral outbreaks, the inputs to the framework consist of different types of data, including images, reports, tabular data, IoT, and others. These data are processed, and a number of parameters are then extracted, such as the concentration of viruses in wastewater samples, the official number of infected people in the region of the wastewater plant, population density, and wastewater flow/generation rate. Those parameters are potentially able to complement each other and enable models to be updated with minimal efforts from human experts.
infected citizens and mortality, vulnerable regions of the outbreaks, and others. Moreover, the framework provides the decision-maker with control strategies that should be followed to minimize any expected risks. The proposed framework is illustrated in Fig. 5, while the framework’s building blocks are also presented in more detail in the following sections.

5.4.1. Predictive machine learning and deep learning methods
The general concept of predictive ML methods is to learn specific properties from a training dataset to build a model capable of making predictions (Alpaydin, 2014). Predictive methods can be divided into two groups: regression and pattern classification. Machine learning regression methods are based on analyzing existing relationships between the variables and trends in the data to make predictions about numerical variables, e.g., the number of infected citizens. In contrast to regression methods, the task of pattern classification is to assign discrete class labels to particular observations as outcomes of a prediction. To go back to the above example: a pattern classification task could predict a high or low infection. The most common ML predictive methods are the support vector machine (SVM) (Bishop, 2006), Bayesian networks (Jensen and Nielsen, 2007), Quadratic discriminant analysis (QDA) (Friedman, 1989), and the artificial neural network (ANN) (Sundui et al., 2021).

Although traditional ML methods have been proven to be effective for predictive analysis, these methods cannot be successfully applied in some cases. Therefore, advanced DL methods such as convolutional neural networks (CNN) (LeCun et al., 1989), recurrent neural networks (RNN) (Hochreiter and Schmidhuber, 1997; Medsker and Jain, 2001), stacked automatic encoders (SAE) (Kingma et al., 2014), deep belief networks (DBN) (Hinton, 2009) should be added to the list to ensure building accurate and representative predictive models. Although DL algorithms are powerful, the functions they create during the training procedure are sophisticated and opaque. Humans who would like to use these models should have the basic ability to understand and trust them before making important decisions.

5.4.2. Descriptive machine learning methods
Descriptive machine learning methods do precisely what the name implies. They describe or summarize raw data and make it interpretable by humans (Novak et al., 2009). The main advantage of descriptive methods over predictive ones is that they enable and allow decision-makers to analyze past behaviors and understand how they might influence future outcomes. Among the descriptive ML methods, the decision trees (DT) (Quinlan, 1986), rough set theory (RST) (Pawlak, 1998), and Logical Analysis of Data (LAD) (Crama et al., 1988) are the most popular ones.

These descriptive methods represent a cornerstone in the framework since decision-makers always seek accurate predictions and meaningful insights that help them take the right actions very rapidly. In the proposed framework, the decision trees are suggested as one of the most widely used and practical descriptive methods and is quite popular and ranked as number 1 in the pre-eminent paper entitled “Top 10 Algorithms in Data Mining” (Wu et al., 2008). It is a flow-chart-like structure, represented as sets of “if” and “then” rules that are readable (interpretable) by humans. LAD is a ML classification technique based on the theories of Boolean algebra and combinatorial optimization (Boros et al., 2000). The fundamental concept of LAD is the extraction of interpretable patterns from the data to discover hidden knowledge from a set of training observations. The extracted patterns are represented as sets of “if” and “then” rules that are easy to interpret (Ragab et al., 2018).

5.4.3. Prescriptive machine learning methods
Prescriptive learning methods aim to recommend actions, control strategies, and policies for optimizing a target process. This is an advantageous characteristic over the predictive and descriptive methods reported in the literature. Providing the decision-makers with recommended control strategies through the proposed framework can help in minimizing the spread of the viral outbreak. This can be done through predicting and analyzing the current situation through predictive and descriptive methods and then select from many prescriptions the best one that mitigates the outbreak of the infection. Several prescriptive ML techniques can be adopted and used in the framework, such as deep reinforcement learning (RL) (Mousavi et al., 2018), fuzzy petri nets (Liu et al., 2017), and evolutionary-based methods (Yu and Gen, 2010).
6. Discussions: remarks, recommendations, and future research perspectives

Several approaches for the detection of viral outbreaks have been presented in this article. However, there are still many aspects that require further attention to obtain a more accurate viral outbreaks surveillance approach. This prioritized need of viral outbreaks surveillance tools is attributed to the fact that the world has been in several dangerous situations over the last two decades, primarily due to the spread of several infectious diseases that could even last for the coming few years. Therefore, the different scientific research communities (biologists, computer scientists, doctors, engineers, statisticians, etc.) should strongly collaborate and identify the existing knowledge gaps in outbreaks surveillance tools and develop tools and techniques that will rapidly and efficiently protect the safety of communities and its people. With the ongoing COVID-19 pandemic, which has left behind many deserted city centers worldwide, we must ask ourselves an important question concerning the future vitality of this pandemic. The authors are deeply convinced that the strength of any city center is that it is alive, with its residents, workers, students, restaurants, shops, schools, shows, museums, lively streets, green spaces, and public squares. In this line of progressive research, AI and related digitalization technologies can play a major role in achieving that goal and making cities more inclusive and sustainable. In this section, several remarks are discussed on the existing viral detection tools and recommendations on the exploitation of advanced AI technologies are provided in order to achieve reliable and robust outbreak monitoring and control strategies.

Viral air surveillance is a complicated process demanding advanced technologies and governmental approvals for widespread applications. The development of these technologies requires proper standardization techniques, relevant calibration, and efficient reporting methods. Furthermore, air sampling methodology design must consider the choice of the sampling location, sample treatment before analysis, boosting the viral load concentration, physical and biological viability of the virus after sampling. Nevertheless, efforts have to be exerted to select the air sampling method compatible with the downstream analytical procedures. Further research is required to design novel biosensors for viral detection in air samples, in addition to the identification in clinical samples. This study shows that the WBE can have great potential for the detection of viral outbreaks. However, some challenges are still faced while applying it and need to be further addressed. Firstly, viral source concentration should be accurately identified from exogenous sources to avoid overestimating or underestimating the real number of patients. Secondly, it is challenging to account for temporal fluctuations needed to estimate the accurate population size. The standardization of methods should be considered to measure the microbial factors in wastewater. A standardized approach used by all stakeholders is not provided yet. Hence, most of the current approaches are resulting from diligent efforts by researchers. Therefore, it would be fundamentally essential to have standardized methods for future research in WBE. Besides, existing data analytics techniques need to be further improved to avoid inaccuracies related to a false positive or false negative. Moreover, the cross-validation of data and interdisciplinary collaboration between different agencies would be vital for better WBE.

The ML and DL have recently contributed to the improvement of screening, prediction, and forecasting tools for viral outbreaks with reasonable accuracy and reliability. However, the urgent associated with outbreaks mandates top-notch performance in terms of precisely predicting and screening, which has not yet been achieved. Hence, the continuous upgrading of the current models is mandatory for better forecasting in the future. An important aspect that needs to be addressed is the data privacy associated with specific ML that requires private data as an input for viral outbreak detection, especially for technologies such as contact tracing. The use of ML and DL for detecting outbreaks using environmental parameters such as WBE is essential. To the authors’ current knowledge, no employed ML/DL is used for early outbreak detection using WBE. Consequently, using these AI technologies for viral detection using WBE needs to be explored because it has a vast potential in improving viral outbreak detection and mapping.

As previously mentioned, several attempts were done to predict how the viral spread took place. Consequently, the number of cases was very close to the observed ones, with a deviation of only 2 to 10 cases over the actual reported cases (De Simone and Piangerelli, 2020; Namasudra et al., 2021). As an additional real-world application of ML usage, diagnosis of the cases proved its wide range coverage and multiple capabilities (Ahmad et al., 2021; Fan et al., 2020). For example, Liang et al. (2021) proposed a deep learning framework to analyze COVID-19 cases using medical images to increase testing sensitivity. Fortunately, the developed CNN model showed high detection accuracy after being trained by multiple X-ray and computed tomography images. The statistical indices of the developed CNN were calculated and had the values of F1 score > 97% and specificity >99%. In another study, deep CNN was developed and utilized to detect COVID-19 using medical images (Gaur et al., 2021). In this case, the overall model accuracy was around 92%, with a sensitivity of >94%. These techniques were applied in other studies where their accuracy and statistical indices were relatively high (Pham, 2020; Shah et al., 2020).

All these reported studies represent strong evidence that these techniques should be applied on a large scale in different areas, especially the medical sector and environmental monitoring for viral detection. However, one obstacle for this to happen on a large scale is the lack of data transfer. Collaborations and official agreements between different sectors to provide the needed data to the research community will serve significantly in overcoming this obstacle. Open-source data sets (Shuja et al., 2020), alongside rigorous analysis based on sophisticated models, lead to suitable decisions during hard times. Our proposed framework offers a good way for any viral outbreak, especially COVID-19, surveillance in wastewater or air based on different data sources and forms inspired by the aforementioned techniques. Besides detecting the viral outbreak, there is an indispensable need for a fast drug discovery process. Drug discovery is an exhausting process due to the very-high dimensional search space of molecules that can reach up to 10^{20} small drug-like molecules or peptides. As a result, it may take years to develop an effective and certified drug for commercial use. Fortunately, ML can play a key role in accelerating the drug discovery process, not only tracking the spread of the contagious virus. Prescriptive ML techniques such as RL can be used for training a search agent. This search agent can introduce good candidates of protein sequences according to an evaluation function. As a result, the number of lab experiments and clinical trials needed for a final vaccine can gradually decrease.

The buzz around AI and related ML/DL technologies reflects its enormous disruptive potential based on key tangible benefits over conventional modeling methods. An important question is how to fully and better exploit the available historical data as essential assets to build accurate and versatile outbreak descriptive/predictive/prescriptive models that will help to address the above-mentioned challenges. The aim is to predict and describe how various components and phenomena interact and influence each other to achieve accurate monitoring and control of different pandemic-like scenarios. The spectacular success of advanced AI methods in a wide range of data analytics applications offers them as appropriate/best available candidates. These advanced AI methods, combined with innovative processing techniques and human expertise, have shown great success in large-scale industrial applications where the data have complex and ill-defined distributions (Elhefnawy et al., 2021; Ragab et al., 2019). Over the past few years, AI has been exploited in different cases, driven by advanced and relatively inexpensive online computing infrastructures. The multitude of diversified ML and DL has proven to be promising for many data-driven modeling applications. This is one of the main motivations of the proposed WBE-AI framework.
The proposed framework aims to deliver timely and highly reliable insight into the viruses’ presence for accurately monitoring and controlling their outbreaks. Confidently, the proposed framework can help to protect people and communities while maintaining business and operational continuity. Facilities and stakeholders such as hospitals, manufacturing/industrial, and healthcare facilities can benefit from the proposed framework.

Researchers need flexible and generic tools that can be easily adapted to meet the new needs to gain the true rewards and values from existing data. To do so, here are some remarks and recommendations on the proper use and applicability of this framework.

6.1. Hyperparameters optimization and AutoML

In order to get the best performance of the ML or DL models, hyperparameter optimization is highly recommended. For example, in case of the existing ML methods, the optimization algorithm, its learning rate, regularization parameters, etc. are among the hyperparameters that need to be tuned within a search space. In the case of the DL method(s), other hyperparameters should be optimized, such as the number of layers (fully connected or convolutional), the activation function, the number of filters in each convolutional layer in the case of convolutional neural networks, the batch size and the number of epochs for training, etc.

For automating the design of ML models and deep neural networks, automated machine learning tools such as Auto-Keras (Jin et al., 2019), H2OAutoML (LeDell and Poirier, 2020), Auto-sklearn (Feurer et al., 2019), Auto-WEKA (Thornton et al., 2013), and Auto-Pytorch (Zimmer et al., 2021), in addition to neural architecture search (NAS) techniques (Wistuba et al., 2019) are commonly used instead of the exhausting task of manual design of such models with numerous hyperparameters. However, these techniques have some limitations due to their highly demanding computational power to obtain the best performing model in a large search space.

6.2. Human-centric and Data-centric AI

In an analogy to the existing practices on continuous software engineering “DevOps” (Bass et al., 2015), the practices for continuous delivery of ML applications are called “MLOps” (Mäkinen et al., 2021). The ML/DL technique used for implementing the predictive/descriptive/prescriptive models is just the first step within the whole cycle of MLOps. There is an urgent need to highlight the importance of data processing techniques for successfully deploying any robust AI products in the commercial market. Using the same AI technique while working on the data engineering process is a potential research direction under the “Data-centric AI” umbrella. One of the reasons behind that is the availability of several sophisticated ML/DL techniques in the absence of representative and informative datasets and the need to monitor and update the deployed models over time based on their performance and incoming data.

The proposed computational modeling workflow is done under the full supervision of a human expert. The involvement of the human expert within our proposed AI framework is an indispensable need for the emerging “Human-centric AI”. Despite AI technologies’ reliability and unprecedented performance, these technologies cannot completely replace human expertise and knowledge. Human expertise is essential due to the following reasons: (i) assess the quality of data, and (ii) select and leverage the suitable ML/DL method that fits the need. However, the combination of “Data-centric AI” & “Human-centric AI” can result in an augmented intelligence that may facilitate the adoption of AI in various complex applications such as the one addressed in this article.

6.3. Transfer learning: from COVID-19 to other viruses

To solve the problem of the insufficiently labeled SARS-COV-2 data, a common and effective method is transfer learning, which realizes the diagnosis of any target virus by simply transferring the information of other viruses to the target COVID-19 case. Moreover, generative learning AI techniques can predict and simulate new virtual viral strains based on the historical viral data. Accordingly, these potential upcoming strains can be confronted and dealt with based on the previously mentioned prescriptive learning techniques that can accelerate the drug discovery process. Furthermore, different simulation and AI techniques can be used to design new adaptive and accurate viral detectors for rapid monitoring and control of potential viral outbreaks. The proposed framework exploits the deep transfer learning methods and combines the advantages of DL and transfer learning and their excellent feature extraction capabilities.

Aside from the use of AI techniques to simulate and design new detectors, the scale-up of producing highly detective nano (bio)sensors, including quantum dots, in the future is a must despite being challenging due to techno-economic considerations and royalty issues. However, they possess some appealing features, such as high sensitivity, which means working with very low concentrations of pathogens, high accuracy, and fast detection. These advantages make them excellent candidates to replace the traditional detection and diagnosis methods currently being used to overcome their drawbacks. Hence, future research should focus on feasibility studies in real environments, scale-up and commercial-scale production of nano-sensors.

Finally, the effects of cross talk and viral defense signaling between the viruses and the host should also be investigated, especially in the newly mutated viruses such as SARS-Cov-2. Such investigations should also address the existing knowledge gaps in detection using the currently applied techniques, identify the effects of transmission in the environment and their spreading abilities.

7. Conclusions

Viral outbreaks are matters of great concern for human survival. Therefore, understanding its transmission means and its early detection is a matter of concern at various levels. WBE has been a growing technique for viral outbreak detection using different monitoring techniques. Besides that, surveillance of viruses in indoor areas has been performed; however, this study shows that the extension and its applicability on higher scales still need further studies. AI and its related ML and DL technologies are promising for viral outbreak detection, risk map formation, and decision support. These technologies augment researchers and decision-makers with multiple perspectives, which can address the outbreaks monitoring and surveillance challenges. Synergizing WBE and AI tools/techniques will significantly contribute to the surveillance of viral outbreaks. In this work, a novel framework has been proposed that combines an ensemble of ML and DL techniques that can be used for different “fit-the-purpose” and by using various types of monitored data sets. This ensemble includes descriptive, predictive and prescriptive techniques that can help decision-makers to accurately detect and forecast future viral outbreaks in addition to recommending versatile control strategies for preventing the spread of the pandemic. Additionally, remarks and recommendations are provided for the successful implementation/development of the proposed framework, including hyperparameter optimization of ML and DL techniques, augmentation of human-centric & data-centric AI, and the exploitation of transfer learning for the detection of current viral outbreaks using the past experiences in this field.

CRediT authorship contribution statement

Omar M. Abdeldayem: Data collection and analysis, Writing - original draft, Visualization, Conceptualization, Methodology, Project administration, Writing - review & editing
Areeg M Dabbish: Writing - original draft, Methodology, Visualization, Data collection and analysis, Writing - review & editing
Mahmoud M. Habashy: Writing - original draft, Methodology, Visualization, Data collection and analysis, Writing - review & editing
Mohamed K. Mostafa: Conceptualization, Methodology, Visualization, Writing - original draft, Resources, Writing - review & editing, Supervision
Mohamed Elhefnawy: Data collection and analysis, Writing - original draft, Visualization
Lobna Amin: Data collection and analysis, Writing - original draft, Visualization
Esma G. Al-Sakkari: Data collection and analysis, Writing - review & editing
Ahmed Ragab: Conceptualization, Methodology, Writing - review & editing, Supervision
Eldon R. Rene: Conceptualization, Writing - review & editing, Supervision

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors dedicate this paper to the memorial of their colleague Omar Eldaghar who passed away in Uganda while serving the privileged communities to have access to basic human needs. We hope that we can carry on with his legacy and ethics in making this world a better place for everyone "Shoulder to shoulder for those who lack basic human needs."

E.R. Rene thanks IHE Delft for providing staff time support under the project “Support to Society” to co-work with researchers from Egypt and students of the ERASMUS+ IMETE programme.

References

Abdeldayem, O.M., Eldaghar, O., Mostafa, M.K., Habashy, M.M., Hassan, A.A., Mahmoud, H., et al., 2020. Mitigation plan and water harvesting of flashflood in and rural communities using modelling approach: a case study in Alouma village. Egypt. Water 12, 2565.
Abduljalil, J.M., 2020. Laboratory diagnosis of SARS-CoV-2: available approaches and limitations. New Microbes New Infect. 36, 100713.
Adan, A., Alizada, G., Kiraz, Y., Baran, Y., Nalbant, A., 2017. Flow cytometry: basic principles and applications. Crit. Rev. Biotechnol. 37, 163–176.
Agbejade, I.E., Awozie, B.O., Ngowi, A.B., Jones, D., 2020. Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future urban health policy implications for reducing transmission of COVID-19. JAMA 323, 1472–1477.
Allam, Z., Dey, G., Jones, D., 2020. Review of big data analytics, and students of the ERASMUS+ IMETE programme. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors dedicate this paper to the memorial of their colleague Omar Eldaghar who passed away in Uganda while serving the privileged communities to have access to basic human needs. We hope that we can carry on with his legacy and ethics in making this world a better place for everyone “Shoulder to shoulder for those who lack basic human needs.”

E.R. Rene thanks IHE Delft for providing staff time support under the project "Support to Society" to co-work with researchers from Egypt and students of the ERASMUS+ IMETE programme.

References

Abdeldayem, O.M., Eldaghar, O., Mostafa, M.K., Habashy, M.M., Hassan, A.A., Mahmoud, H., et al., 2020. Mitigation plan and water harvesting of flashflood in and rural communities using modelling approach: a case study in Alouma village. Egypt. Water 12, 2565.
Abduljalil, J.M., 2020. Laboratory diagnosis of SARS-CoV-2: available approaches and limitations. New Microbes New Infect. 36, 100713.
Adan, A., Alizada, G., Kiraz, Y., Baran, Y., Nalbant, A., 2017. Flow cytometry: basic principles and applications. Crit. Rev. Biotechnol. 37, 163–176.
Agbejade, I.E., Awozie, B.O., Ngowi, A.B., Jones, D., 2020. Artificial intelligence (AI) provided early detection of the coronavirus (COVID-19) in China and will influence future urban health policy implications for reducing transmission of COVID-19. JAMA 323, 1472–1477.
Fan, Z., Jamil, M., Sadiq, M.T., Huang, X., Yu, X., 2020. Exploiting multiple optimizers with transfer learning techniques for the identification of COVID-19 patients. J. Healthc. Inform. Res. 17, 898812.

Fardis, N., Sadiq, K., Sadeghi, K., Kaddafi, X., Yavarian, J., Shamsipur, M., et al., 2020. A field indoor air measurement of SARS-CoV-2 in the patient rooms of the largest hospital in Iran. Sci. Total Environ. 725, 138401.

Farkas, I., Mannson, F., Hillier, C., Su, J., Malm, S.K., Walker, D.J., 2020. Emerging technologies for the rapid detection of enteric viruses in the aquatic environment. Curr. Opin. Environ. Sci. 16, 1–6.

Feng, W., Newbigging, A.M., Le, C., Pang, B., Peng, H., Cao, Y., et al., 2020. Molecular diagnosis of COVID-19: challenges and research needs. Anal. Chem. 92, 10196–10209.

Fernandes, N., 2020. Economic Effects of Coronavirus Outbreak (COVID-19) on the World Economy.

Feuer, M., Klein, A., Eggespenger, K., Springenberg, J.T., Blum, M., Hutter, F., 2019. Auto- skim: efficient and robust automated machine learning. Automated Machine Learning. Springer, Cham, pp. 423–437.

Fokalid, P., Cutrupi, F., Segata, N., Manara, S., Pinto, F., Malpel, F., et al., 2020. SARS-CoV-2 from faeces to wastewater treatment: what do we know? A review. Sci. Total Environ. 743, 140444.

Fongaro, G., Stoco, P.H., Gisrad, E.C., Magri, M.E., Rogovski, P., Rodríguez-Lázaro, D., 2021. The presence of SARS-CoV-2 RNA in human sewage in Santa Catarina, Brazil, November 2019. Sci. Total Environ. 778, 146108.

Forterre, P., Soler, N., Krupovic, M., Marguet, E., Ackermann, H.W., 2013. Fake virus particles trick scientists. Nat. Biomed. Eng. 3, 427–437.

Freeman, S., Eyalbosh, A., 2020. COVID-19 and outdoor safety: considerations for use of outdoor recreational spaces. National Collaborating Centre for Environmental Health 329.

Freeman, M.W., Walker, S.J., Vrana, K.E., 1999. Quantitative RT-PCR: pitfalls and potential. BioTechniques 26, 122–125.

Friedman, J.H., 1989. Regularized discriminant analysis. J. Am. Stat. Assoc. 84, 165–175.

García-Fontana, C., Rodríguez-Sanchez, A., Muñoz-Palazon, B., Gonzalez-Martinez, A., Velázquez-Medina, C., 2020. Profile of the spatial distribution patterns of the human and bacteriophage virome in a wastewater treatment plant located in the south of Spain. Water 12, 2136.

Garrido-Basera, M., Corominas, L., Cortes, U., Rosso, D., Poch, M., 2020. The fourth revolution in the water sector examines the digital revolution. Environ. Sci. Technol. 54, 4698–4705.

Gaur, L., Bhatia, U., Janhji, N.Z., Muhammad, G., Masud, M., 2021. Medical image-based detection of COVID-19 using deep convolution neural networks. Multimedia Systems 1–10.

Gonzalez, R., Curtis, K., Bivins, A., Bibby, K., Weir, M.H., Yetka, K., et al., 2020. COVID-19 surveillance in southeastern Virginia using wastewater-based epidemiology. Water Res. 186, 112692.

Gooze, J.A., Rushworth, J.V., Millner, P.A., 2015. Biosensor regeneration: a review of common techniques and outcomes. Langmuir 31, 6267–6276.

Guo, Z., Wang, Z.Y., Zhang, S.F., Li, H., Li, C., et al., 2020. Aerosol and surface distribution of severe acute respiratory syndrome coronavirus 2 in hospital wards, Wuhan, China. 2020. Emerg. Infec. Dis. 26, 1583–1591.

Guruprasad, L., 2021. Human SARS-CoV-2 spike protein mutations. Proteins 89, 567–576.

Hajian, H., Coull, B., Mosher, E., Godbold, J., Grun, A., et al., 2014. Sources of indoor air pollution in New York City residences of asthmatic children. J. Expo. Sci. Environ. Epidemiol. 24, 269–278.

Hades, M., Hopke, P.K., Jonidi, A., Shashavani, A., 2020. A letter about the airborne transmission of SARS-CoV-2 based on the current evidence. Aerosol Air Qual. Res. 20, 911–914.

Hajian, R., Balderston, S., Tran, T., deboer, T., Etienne, J., Sandhu, M., et al., 2019. Detection of unamplified target genes via CRISPR-Cas9 immobilized on a graphene-oxidized quartz microelectrode in the field. Anal. Chem. 91, 270–2713.

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, 2707–2713.

Haramoto, E., Malla, B., Takah, O., Kitajima, M., 2021. First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. Sci. Total Environ. 737, 140405.

Hasan, S.W., Ibrahim, Y., Dasa, M., Kannott, H., Jan, N., Lopes, A., et al., 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.

Hata, A., Honda, R., Hara-Yamamura, H., Meuch, Y., 2020. Detection of SARS-CoV-2 in Wastewater in Japan by Multiple Molecular Assays Implication for Wastewater-based Epidemiology (WBE).

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Heuwell, R., 2020. How Al is Battling the Coronavirus Outbreak. Vox.

Hinton, G.E., 2009. Deep belief networks. Scholarpedia 4, 5947.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780.

Hogan Jr., C.J., Kettleson, E.M., Lee, M.H., Ramaswami, B., Angenent, L.T., Biswas, P., 2005. Sampling methodologies and doseage assessment techniques for ultrafine and infratine aerosol particles. J. Appl. Microbiol. 99, 1442–1453.

Hong, S., Lee, C., 2018. The current status and future outlook of quantum dot-based biosensors for plant virus detection. Plant Pathol. 34, 85–92.
Lalmuannawma, S., Hussain, J., Chhalachhauk, L., 2020. Applications of machine learning and artificial intelligence for Covid-19 (SARS-COV-2) pandemic: a review. Chaos Solitons Fractals 131, 110019.

Lau, L.T., Fung, W.Y., Wong, F.P., Lin, S.S., Wang, C.R., Li, H.L., et al., 2003. A real-time PCR for SARS coronavirus incorporating target gene pre-amplification. Biochem. Biophys. Res. Commun. 312, 1290–1296.

Lau, H., Khorasipour, V., Kashi, P., Mikolajczyk, A., Schubert, J., Bania, J., et al., 2020. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J. Travel Med. 7, 1–7.

Laue, M., 2010. Electron microscopy of viruses. Methods Cell Biol. 96, 1–20.

Layqah, L.A., Eissa, S., 2019. An electrochemical immunosensor for the corona virus associated with the Middle East respiratory syndrome using an array of gold nanoparticle-modified carbon electrodes. Microchim. Acta. 186, 224.

Le, T., Chun-Xiao, W., Shu-Lin, L. 2017. Pulsed field gel electrophoresis. Reference module. Life Sci. https://doi.org/10.1016/B978-0-12-805633-8.00068-0.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., et al., 1989. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1, 541–551.

LeDell, E., Posier, S., 2020. H2O autoai: scalable automatic machine learning. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1498–1507.

Ledesnky, J.A., Lavauzard, M., Fan, Z.H., Jutta, A., Tilly, T.B., Gangwar, M., et al., 2020. Viable SARS-COV-2 in the air of a hospital room with COVID-19 patients. Int. J. Infect. Dis. 100, 476–482.

Lee, T., Adhikari, Park, S.Y., Kim, G.H., Kim, J., Kim, T.H., et al., 2018. Recent advances in AV biosensors composed of nanobio hybrid material. Micromachines 9, 651.

Lee, H., He, F., Liu, Y., Huang, Z., Ling, S., Jiang, Z., et al., 2020. SARS-COV-2 environmental contamination associated with persistently infected COVID-19 patients. Influenza Other Respir. Viruses 14, 688–699.

Levician, J., Levin, A., Ampuero, M., Caggo, A., 2013. JC polyomavirus circulation in one-year surveillance in wastewater in Santiago, Chile. Infect. Genet. Evol. 17, 151–158.

Li, S., Jiang, L., Yu, G., Guo, X., Li, H., Li, L., et al., 2021. Fast automated detection of COVID-19 using machine learning framework for its images. Infect. Genet. Evol. 63, 104191.

Lim, R.R.X., Bonanno, A., 2020. The potential of electrochemistry for the detection of coronavirus-induced infections. Trends Anal. Chem. 133, 116081.

Lindesley, William G., Green, Brett J., Blachere, Françoise M., Martin, Stephen B., Law, Brandon F., Jensen, Paul A., et al., 2017. Sampling and characterization of bioaerosols. Manual of Analytical Methods (NMAM).

Liu, W.T., Zhu, L., 2005. Environmental microbiology-on-a-chip and its future prospects. Trends Biotechnol. 23, 174–179.

Liu, H.C., You, J.X., Li, Z.W., Tan, J.C., 2017. Fuzzy petri nets for knowledge representation: a literature review. Eng. Appl. Artif. Intell. 60, 45–56.

Liu, Y., Ning, Z., Chen, Y., Guo, M., Liu, Y., Gali, N.K., et al., 2020. Aerodynamic analysis of SARS-COV-2 in two Wuhan hospitals. Nature 582, 557–560.

Lodder, W., Husman, A.M., & L., 2020. SARS-COV-2 in wastewater: potential human health risk, but also data source. Lancet Gastroenterol. Hepatol. 5, 533–534.

Lopez-Canovas, L., Martinez Benitez, M.B., Herrera Isidron, J.A., Flores, S., E, T., 2019. Pulsed field gel electrophoresis: past, present, and future. Anal. Chem. 573, 17–29.

Ma, L., Mao, G., Liu, J., Yu, H., Gao, G., Wang, Y., 2013. Rapid quantification of bacteria and viruses in influent, settled water, activated sludge and effluent from a wastewater treatment plant using flow cytometry. Water Sci. Technol. 68, 1763–1769.

Ma, F., Li, C.C., Zhang, C.Y., 2018. Development of quantum dot-based biosensors: principles and applications. J. Mater. Chem. B 6, 6713–6730.

Maestro, S., Magallón, L., Luengas, E., García-Nava, N., García-Heredia, I., Martinez-Hernandez, F., Martinez-Garcia, M., 2019. Insights into the antibiotic resistance dissemination in a wastewater effluent microbiome: bacteria, viruses and vesicles matter. Environ. Microbiol. 21, 4582–4596.

Mäkinen, S., Skogström, H., Atalik, Z., Lehtonen, K., Mäkinen, T., 2021. Who Needs MDLSp? What Data Scientists Seek to Accomplish and How Can MDLSp Help?

Mallapaty, S., 2020. How sewage could reveal true scale of coronavirus outbreak. Nature 580, 176–177.

Manivannan, S., Ponnuchamy, K., 2020. Quantum dots as a promising agent to combat COVID-19. Appl. Organomet. Chem. 34, e5887.

Mao, K., Zhang, H., Yang, Z., 2020. Can a paper-based device trace COVID-19 sources with the most accuracy? A comparison of different testing methods. Sens. Actuators B Chem. 308, 127708.

Marcondes, E.C., Schwan, T.K., 2020. Identification of bacteria, viruses and drug-resistant mechanisms in wastewater effluent: a review. Front. Cell. Infect. Microbiol. 10, 576875.

Matsuo, E., Mogasa, C.M., Ekanpu, R.A., Debogbrasova, S., Lugeba, G.W., Laurent, T., et al., 2010. Phase II evaluation of sensitivity and specificity of PCR and NASBA followed by oligochromatography for diagnosis of human african trypanosomiasis in clinical samples from D.R. Congo and Uganda. PLoS Negl. Trop. Dis. 4, e737.

Mejías, C.I., Breeuwer, K., Nordeng, A., 2004. Source tracking fecal bacteria in water: a critical review of current methods. J. Environ. Manag. 75, 71–79.

Medema, G., Heijnen, L., Elisiga, 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, 511–516.

Medsker, L.R., Jain, L.C., 2001. Recurrent neural networks. Des. Appl. S. Mehrpouro, P., 2016. Biosensors and their applications - a review. J. Oral Biol. Craniofac. Res. 6, 55–59.

Mljekovska, H., Sovova, K., Vasikova, P., Ocasnovsky, V., Jasikova, J., Juranova, E., 2020. Preliminary study of SARS-COV-2 occurrence in wastewater in the Czech Republic. Int. J. Environ. Res. Public Health 17, 5508.

Mohanty, S., Harun Al Rashid, M., Mridul, M., Mohanty, C., Swayamidda, S., 2020. Applications of artificial intelligence for COVID-19 drug repurposing. J. Diabetes Metab. Syndr. Clin. Res. Rev. 14, 1027–1031.

Morawska, L., Cao, J., 2020. Airborne transmission of SARS-CoV-2: the world should face the reality. Environ. Int. 135, 105730.
Substituting the input text with a natural representation:

O.M. Abdeldayem, A.M. Dabbish, M.M. Habashy et al. Science of the Total Environment 803 (2022) 149834

Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V., Pachori, R.B., 2021. Application of deep learning in the detection of COVID-19. Int. J. Med. Inf. 146, 103265.

Mousavi, S.S., Schukat, M., Lednicky, J.A., Wu, C.Y., 2019. Collection, particle sizing and detection of airborne viruses. J. Hazard. Mater. 124656.

Orive, G., Lertxundi, U., Barcelo, D., 2020. Early SARS-CoV-2 outbreak detection by inverse gel electrophoresis: theory, instruments and applications. Cell Tissue Res. 2, 557–565.

Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V., Pachori, R.B., 2021. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. Signal Process. Control 64, 102365.

Nassonova, E.S., 2008. Pulsed electroemission: theory, instruments and applications. J. Phys. D Appl. Phys. 41, 184001.

Neumeyer, L., Robbins, B., Nair, A., Kesari, A., 2010. S4: distributed stream computing platform. ACM SIGMOD Rec. 39, 617–628.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.

Nguyen, H.H., Lee, S.H., Lee, U.J., Fermin, C.D., Kim, M., 2019. Immobilized enzymes in biosensor applications. Mater. Sci. Eng. C 96, 467–476.
