Water Reuse

Water Reuse Vol 00 No 0, 1 doi: 10.2166/wrd.2024.121

Synergizing cybersecurity in healthcare with novel bioprocessing for sustainable energy-centric water remediation

K. Sita Kumari[a], Vasundhara Ghorpade[b], Fatima Moayad Sami[c], Sulaima Lebbe Abdul Haleem[d], Suresh Babu Kondaveeti[e] and Sherzod Kiyosov[f,*]

[a] Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India
[b] Department of PSM, Krishna Vishwa Vidyapeeth, Karad, MS 415539, India
[c] Department of Medical Laboratory Technics, Al-Noor University College, Nineveh, Iraq
[d] Faculty of Technology, South Eastern University of Sri Lanka, Sri Lanka
[e] Department of Biochemistry, Symbiosis Medical College for Women, Symbiosis International (Deemed University), Pune, India
[f] Tashkent Institute of Finance, Tashkent, Uzbekistan
*Corresponding author. E-mail: kiyosov_sherzod@tfi.uz

ABSTRACT

The introduction of several novel chemicals, materials, and processes with varied levels of complexity in recent decades has been a result of the incredibly rapid technological advancement that has characterised those decades. This in turn has led to an increase in the number of pollutants released into the environment, necessitating their effective removal. The results of environmental monitoring reveal that several pollutants are exceeding the limit in ground water, which raises questions about the efficacy of the wastewater treatment methods now in use. This research proposes a novel method for sustainable smart grid (SG)-based energy analysis in water remediation and network cybersecurity analysis for healthcare application. Here the water-caused damages have been analysed based on a healthcare application using SG energy analysis and the network cyber security analysis is carried out using the federated blockchain model (SGEA_FB). Experimental analysis is carried out in terms of network integrity, throughput, scalability, and training accuracy. According to an analysis of the exergy destruction rates in every method component, power generation subsystem has the greatest exergy destruction rate at 15.4 MW. A training accuracy of 96%, throughput of 86%, network integrity of 76%, and scalability of 73% were all achieved by the suggested method.

Key words: cybersecurity analysis, energy analysis, healthcare application, sustainable smart grid, water remediation

HIGHLIGHT

- This research proposes a novel technique for water remediation and cybersecurity analysis for a healthcare application based on the sustainable energy management system. Here the smart grid-based network energy management is carried out and the blockchain federated model has been used for cybersecurity analysis.

1. INTRODUCTION

Water, which covers more than two-thirds of the planet's surface, is essential to all forms of life. Although water is abundant, there is a limited supply (Hossain *et al.* 2019). Water is crucial to life on Earth in many ways. Since it is so commonplace, we often fail to see its essence. Water is full of life and character. Over time, both the chemical and physical qualities might change. Water is essential to all forms of life. Each live cell is composed of a solution, suspension, or emulsion that is between 25 and 85% water (Zhang *et al.* 2023). Young tissues have a greater capacity for water storage than their older counterparts. In general, plants cannot survive without water. They deplete soils of moisture and useful minerals. While plants are composed of 50–75% water, humans are 60–65% water for men and 50–60% water for women (Siddiqui & Dincer 2021). It plays a role in the production and breakdown of many chemical compounds and has a slight effect on operating systems. Natural water has a variety of components, including minerals, organic pollutants, and dissolved gases like air and carbon dioxide. Multiple illnesses have been shown to spread via water. That is why it is so important to keep an eye on the quality of water in real time. Water quality (WQ) may be altered by both natural and manmade processes under certain circumstances. Some of these processes are to blame for the changes in WQ (Kiehbadroudinezhad *et al.* 2023), including

increasing temperatures, increased biological oxygen demand, higher chemical oxygen utilisation, and increased nitrogen as well as phosphorus combinations. Important causes such as population expansion, automation, urbanisation, and others have contaminated water supply and quality. Because of pollution from industry, cities, and farms, the quality of Kosovo's surface water has lately deteriorated. It is common to classify water as either 'surface water' or 'ground water.' The potential for pollution from domestic, industrial, and agricultural practices necessitates the discharge of such water.

With regard to the method-training stage, machine learning (ML) refers to the intelligent methods that can modify their behaviour in response to newly presented data, while artificial intelligence (AI) refers to the notion of providing methods with the capacity to do tasks and draw conclusions that would require an intelligent person in the same position (Vivek *et al.* 2021). Because it is a bottom-up mathematical approach that makes use of historical knowledge, ML excels at data correlation and can draw inferences from systems with sophisticated mathematical underpinnings that are difficult even for individuals to comprehend and apply. When it comes to water applications, smart technologies and developing AI and ML are filling a hole that has been left unfulfilled by conventional methods and ideas. It is estimated that by 2030, the water industry would have spent $10 billion on AI, out of a total of more than $90 billion. There has been extensive use of ML in throughput/water-based agriculture, monitoring of natural systems, and water and wastewater treatment, to name a few water-related applications. Applications of hybrid approaches that combine two ML techniques, such as ANN-RF and SVM-RF, are examples of this kind. Research shows that ML and AI may improve modelling processes for use with water-related applications (Fouladi *et al.* 2022).

2. RELATED WORKS

Environmental preservation has become increasingly important in recent years for human survival worldwide. Wastewater gathered from both industrial and home purposes waste pollutes the environment. To protect our environment, wastewater treatment has been implemented with several influent indications. Therefore, wastewater treatment facilities clean up contaminated water, maintain a clean environment, and correctly restore the environment for human observation. With its porous, thermostable structure and wide range of uses, including the elimination of contaminants, the odour associated with medical treatments, the purification of gaseous phases, and wastewater treatment, activated carbon is a popular adsorption material (Sayed *et al.* 2022). Through the use of activated carbon, different pollutant particles can be quickly and affordably absorbed in order to remove colours from wastewater (Wang *et al.* 2021). A number of ML algorithms, including neural networks, SVM, and tree-based methods, were used to predict WWTP using activated carbon (Kowthaman *et al.* 2022). For forecasting the activated carbon-based methylene blue number and iodine number, this (Bhavani *et al.* 2023) paper offered different ML techniques, including support vector regression (SVR), linear regression (LR), and random forest regression (RFR). There are some studies in the literature that look at the process of treating industrial wastewater using renewable energy. A solar photovoltaic-based system for brackish water treatment was studied by Choi *et al.* (2022). The method of membrane capacitive deionisation was used. For the experiment, a pilot water treatment facility with a 1 kW capacity was created. According to reports, the proposed technology is cost-effective when compared with traditional water treatment methods. However, the system's overall performance in terms of efficiency was not looked into. An integrated system for producing power and treating wastewater was subjected to an energy study by Danaeifar *et al.* (2023). In the study, sewage sludge digestion was used to generate electricity from the municipal wastewater. According to their findings, wastewater flow involves a sizable amount of energy, and the use of major resources throughout the treatment process is justifiable. A fresh approach to wastewater reclamation put emphasis on coastline wastewater treatment facilities (Alawad *et al.* 2023). Energy was produced by using salinity gradient between reclaimed water and the salty seawater. Panagopoulos (2021) looked at a high salinity solar-based brine wastewater treatment method. The decontamination procedure used was electro-oxidation (Caldera & Breyer 2019). The solar-based system had an energy consumption of 96.3 kWh/m^3 and a requirement for UV sun radiation of 232 kJ/L. Solar energy was used to treat industrial brine through an evaporative method. Additionally, glass–metal vacuum tubes were incorporated into the compound solar collectors to help them operate at high temperatures. Ahmed *et al.* (2022) created a two-stage dynamic thresholds technique based on a genetic algorithm for detecting pollution events in water distribution. Regression trees were employed in a different study by Nahar & Sunny (2020) to identify outliers in measurements of WQ. In a study (Mollahosseini *et al.* 2019), the categorisation of contamination events was accomplished by comparing the weighted support vector machine (SVM) classifier with the minimum volume ellipsoid (MVE), after a sequence analysis. To discover outliers in a time series and detect events,

Anyaoha & Zhang (2022) utilised a binomial event discriminator. A technique for locating contamination events was utilised by Maggio *et al.* (2022) and was based on the DempstereShafer evidence theory. Using recent observations from WQ sensors, the technique initially employs an autoregressive model to forecast future WQ parameters.

3. SYSTEM MODEL

The blockchain-based paradigm of an intelligent peer-to-peer energy trading platform is shown in Figure 1. The real-time and day-ahead energy trading based on pre-processed data as well as short-term energy projection are two separate modules that make up smart contract-enabled intelligent energy trading. Data about energy trade are processed and stored on each node of the suggested platform. In this analysis, dispatchable loads like ESS, shapeable loads like electric vehicles, and solar energy generation (photovoltaics) have all been taken into account. Utilising various ML as well as data mining techniques, energy consumption data from these sources are analysed to extract useful time-series patterns as well as hidden knowledge that may be used to satisfy future energy demand. Pre-processed data are utilised to control as well as schedule distributed energy resources in real time as well as a day in advance. Additionally, ML models are utilised to anticipate the short-term energy consumption utilising discovered time-series properties, such as hourly, daily, weekly, annual, and seasonally. Every exchange that takes place between nodes that serve as consumers and prosumers is recorded as an energy trading transaction (ETT) in the state database. Through the client application, which is utilised for secure energy trading, players such as prosumers, consumers, and utility operators can communicate with the method.

With the help of the smart grid (SG), grid operators can look deeper into the system and have more leeway in dealing with the intermittent nature of renewable energy sources. In turn, this overcomes a key hurdle, speeding up the deployment of wind and solar energies, and increasing their share of the energy market. Safety and dependability, self-healing, efficiency, intelligence, adaptability, environmentally friendly, quality, and stability are only few of the benefits of smart grids that have been proposed.

Dataset characteristics: These were determined by investigation on the information included in the water_potability.csv file [37]. Approximately 3,276 records' worth of metrics are in the file. To tidy up the dataset, all rows with empty cells in any input columns were removed. As a result, there were fewer records in the dataset, but the ones that were still there were useful for the research. The following metrics relating to WQ were measured and noted:

1. The pH value finds acid–base balance and the acidity or alkalinity of the water. The pH range that is advised is 6.5–8.5.
2. The ability of water to precipitate soap due to calcium as well as magnesium is used to measure hardness.
3. Total dissolved solids (TDS), a measure of solids, show how mineralised the water is. High mineralisation is indicated by a higher TDS value. It is expressed in mg/L.

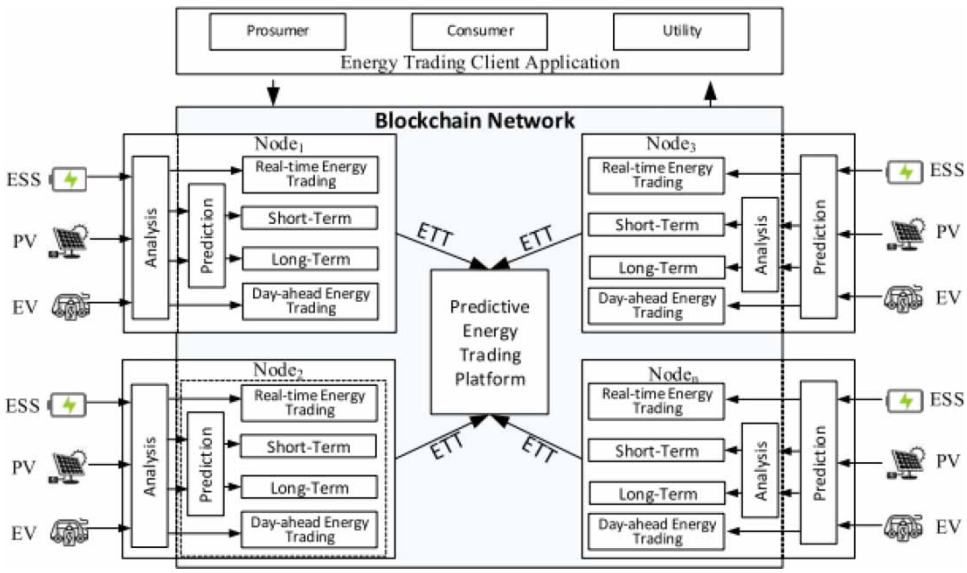

Figure 1 | Smart grid energy trading platform – conceptual scenario.

4. When chlorine and ammonia are utilised to treat drinking water, chloramines are created. The units are mg/L or ppm.
5. The chemical industry is the main user of sulphate, which is found in ambient air, groundwater, plants, and food. It is expressed in mg/L.
6. A solution's ionic process for transporting current is measured by conductivity. Microsiemens per centimetre, or S/cm, is the unit of measurement.
7. Decomposing natural organic matter and artificial sources are sources of organic carbon in source waters. It calculates the overall carbon content of organic molecules in pure water. It is expressed in mg/L.
8. By measuring its concentration, trihalomethane in drinking water can be identified. The amount of organic matter, the chlorine content, and the temperature of the water all affect it differently. The nephelometric turbidity unit is used to measure it.
9. The turbidity scale measures how well water emits light. The test reveals the colloidal matter-related waste discharge quality. It is expressed as ppm.
10. Water that is potable is safe for ingestion by people. Its value is either 0 or 1. The 'potability' characteristic is the output of the proposed method in this study, whereas the first nine features are its inputs.

As a result, the data are divided into two categories for classification purposes: 'Potable' and 'Not Potable.' Each of these two classifications is assigned a one-digit code, with '1' denoting potability and '0' denoting non-potability. Table 1 shows the statistical analysis of the dataset's characteristics.

Data processing

Data processing happens after data collection. Data processing thus begins with the gathering of primary data and involves the gradual transformation of that data into useful information. The different steps for data processing are listed as follows:

- To create data with a comprehensible format, raw data are pre-processed after data collection.
- The key characteristics are chosen. All measurable characteristics are chosen for this dataset.
- Data are divided into two subsets, one of which is utilised as a sample of training data and the other for testing purposes.
- The data subsets are subjected to ML techniques.
- The findings are gathered and contrasted.

SG-based energy analysis

A sophisticated system of systems (SOS), the foundation of a SG, emphasises ESS technology to improve asset utilisation while preserving dependable system operation, demand responsiveness, and environmental protection by utilising multiple generating sources. The main grid of the SG incorporates numerous micro-grids, making it a complicated system. The electric networks that correlate to home power systems and neighbourhoods, respectively, are called nano-grids and micro-grids. Additionally, they are linked to either another micro-grid or the grid that distributes electricity. Different technologies are used in the micro-grid, which includes loads, distributed generators, distributed systems, EVs, ESSs, and power electronic

Table 1 | Dataset features with their statistical analysis values, including standard deviation (StdDev) and interquartile range (IQR)

Feature	Min	Max	Mean	Median	StaDev	Mode	IQR
Solids	320.94	56,459	21,917	20,934	8,642.2	320.94	11,560
Handness	73.492	31,734	195.57	197.19	32,635	73.492	39.692
pH value	0.2275	14	7.086	7.0023	1.5733	0.2275	1.9635
Chloramines	1.3909	13.127	7.1343	7.1439	1.5845	1.3909	1.9716
Organic carbon	2.2	27.009	14.358	14.322	3.25	2.2	4.5611
Conductivity	201.62	753.34	42,653	423.46	80,713	201.62	115.74
Sulfate	129	481.03	333.22	323.23	41.205	129	51.647
Turbidity	1.45	64,947	3.9697	3.9682	0.78035	1.45	1.0704
Trihalomethanes	8.577	124	66.401	66.542	16.077	8.577	21.399

methods. Additionally, these methods are required to link DER to nano-grids utilising PV inverters or batteries to achieve the necessary frequency.

This work suggests an SG architecture, which is shown in Figure 2, to address the difficulties of integrating RESs and nano/micro-grids. Every LC receives set points from the relevant SG Controller during a centralised operation. This sort of control, however, has redundancy and poor reliability [38]. Each local controller makes decisions locally in a decentralised system [39]. A hybrid method is used in the current research to define both the centralised and decentralised activities within the main grid.

The three levels that make up the suggested SOS architecture are macro-grid (MG) level, also known as the main grid, MG level, and the nano-grid level. To transfer energy in various forms, such as gas or electricity, a transmission system operator (TSO) is also employed. A service provider is utilised to interface with operators as well as to ensure the SG's proper operation. When considering micro-grid operation modes, two objectives are distinguished at the micro-grid level. The grid-connected mode is used to conduct financial transactions through interactions with the wholesale markets. However, in island mode, the micro-grid's energy and financial flows are controlled by a neighbourhood market called the retail market, and energy efficiency services are offered. The micro-grid ensures the network's stability and security in real time for both operating modes. Additionally, dynamic pricing can be used to close the gap between supply and generation and implement demand response. Power consumption, generation, and voltage measurements are read from and sent via the smart metre. It engages in a two-way dialogue with shops and controllers and supports a number of billing methods, including time of use (TOU) and real-time pricing (RTP).

Federated blockchain model-based cybersecurity analysis

Figure 3 depicts the suggested smart energy trading mechanism using blockchain technology. The proposed infrastructure is a technological framework for delivering blockchain services to end users using distributed ledger technology (DLT) and smart contracts. The DLT relies on a network of peers, or nodes, as depicted in Figure 1. Each node maintains its own copy of the ledger to guarantee data integrity. The distributed ledger (DL) maintains a running chain of blocks that store the unalterable energy trading transactions. The suggested system data lake, by contrast, manages and stores data pertaining to distributed energy sources, system users, and energy trading operations. The data lake is regarded as an off-chain database that stores information about transactions involving the trading of energy and is also utilised by data analytics methods. Additionally, the proposed method offers RESTfull API functionality to support front-end energy trading client application with back-end services. To carry out real-time as well as day-ahead energy trading, energy consumption data are analysed and employed

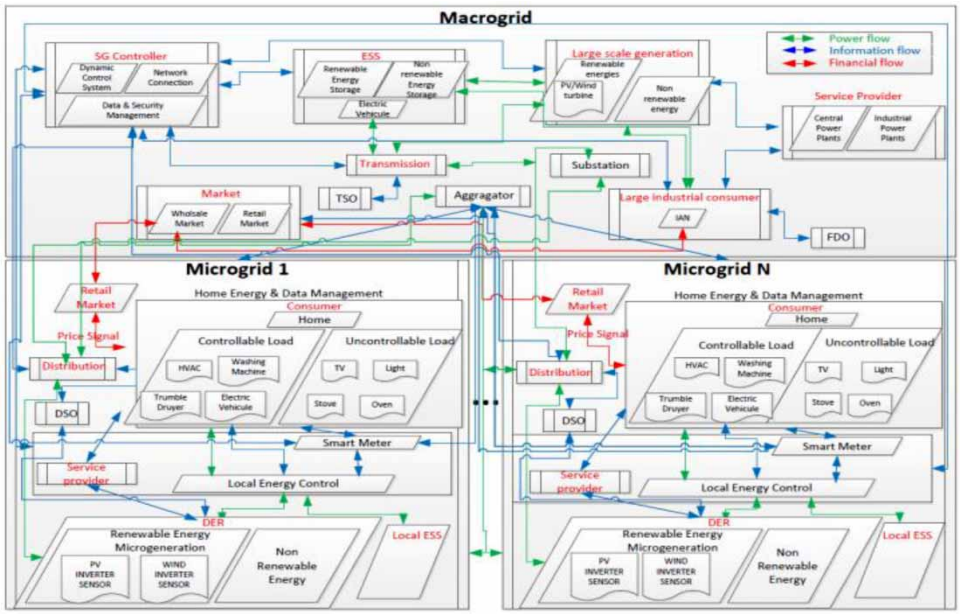

Figure 2 | The proposed Smart Grid Architecture.

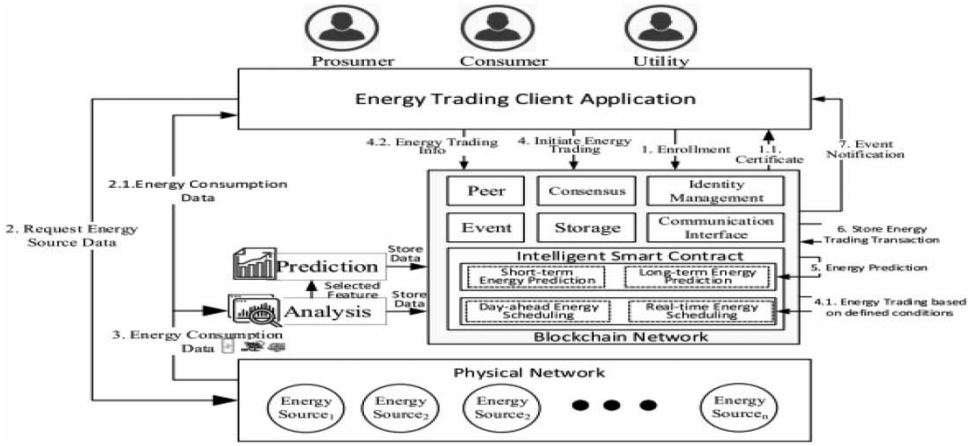

Figure 3 | System workflow of the intelligent blockchain-based energy trading platform.

in an intelligent smart contract. Similarly to this, a ML algorithm forecasts future energy demand using the analysed data. Data from every energy trade transaction are kept on the blockchain. Upon the completion of the successful ETT, the notification is provided to the relevant user.

The proposed blockchain-enabled predictive energy trading platform has a modular architecture with independent layers that make it simple for developers to add or remove new components without affecting other layers. A bus transmission line connects several dispersed energy resources, such as solar energy as well as dispatchable loads, in the distribution grid network model. Every node in the distributed grid network is outfitted with distributed energy resources, including solar energy and shapeable and dispatchable loads, which may both produce as well as consume energy. Finally, every bus is utilised to link a node to grid to transfer as well as consume energy. A number of blockchain capabilities, including identity management, API interfaces, distributed ledgers, peer-to-peer (P2P) communication, and consensus managers are provided by the blockchain-based energy trading service layer. The DL is made up of synchronised, shared, and replicated digital data that are dispersed throughout the blockchain network, with each network participant maintaining a copy of the ledger. The existing privacy-preserving techniques for federated learning of deep neural networks with billions of model parameters are inadequate. Although homomorphic encryption (HE)-based approaches offer solid privacy safeguards, they are practically useless in practise due to their extraordinarily large computational and communication overheads (Panagopoulos 2021; Ahmed *et al.* 2022) [32]. It is simple for attackers to deduce data $x = g\ 1\ (G)$ from the disclosed information if $g()$ is an invertable function. For example, in the context of federated deep neural network learning, $g(x)$ is the exchanged neural network parameters or its gradients based on input training data x. In general, $g()$ is NOT invertable. The NN model parameters G are sought for classification tasks by minimising a loss function established on training data x by the following equation

$$G = \min \mathrm{LCE}(G, x) \tag{1}$$

Even if $g()$ in (1) is NOT invertable, opponents can still use x to estimate x in the sense of a Bayesian restoration by minimising $|g(x)G|$. Each of the chosen participants updates their model, then uses it to train local data. Following the local training process, each participant uploads the updated models, which are then averaged and added to the central model as it is today. The central model's updating process is shown in the following equation.

$$M_{t+1} = M_t + \frac{1}{m_t}\sum_{k=1}^{m_t} u_t^k \tag{2}$$

where u_t^k stands for method updates submitted by k-th participant, and M_t represents the current global model at t-th iteration. We now present our fresh approach to the supervised FL model training by the following equation

$$\min_{x_1,\ldots\ldots,x_s \in \mathrm{R}^d} \{F(x) := f(x) + \lambda\psi(x)\}$$

$$f(x) := \frac{1}{n}\sum_{i=1}^{n} f_i(x_i),\ \psi(x) := \frac{1}{2n}\sum_{i=1}^{n}\|x_i - \bar{x}\|^2 \tag{3}$$

Due to the f_i assumptions we will make, F is extremely convex; therefore, the following equation only has one solution, which we denote as

$$x(\lambda) := (x_1(\lambda),\ \ldots, x_n(\lambda)) \in \mathbb{R}^{nd} \tag{4}$$

We further let $\bar{x}(\lambda) := \frac{1}{n}\sum_{i=1}^{n} x_i(\lambda)$. Intuitively, this limit situation should compel best local methods mutually similar while minimising loss f. This limit case will resolve in particular by the following equation

$$\min\{f(x) : x_1,\ \ldots, x_n \in \mathbb{R}^d, x_1 = x_2 = \ldots = x_n\}$$

$$\psi(x(\lambda)) \le \frac{f(x(\infty)) - f(x(0))}{\lambda}$$

$$f(x(\lambda)) \le f(x(\infty)) \tag{5}$$

For every $\lambda > 0$ and $1 \le i \le n$, we have by the following equation

$$x_i(\lambda) = \bar{x}(\lambda) - \frac{1}{\lambda}\nabla f_i(x_i(\lambda)) \tag{6}$$

Further, we have $\sum_{i=1}^{n} \nabla f_i(x_i(\lambda)) = 0$. By deducting a multiple of local gradient from the average model, best local models (5) are found. Let $P(z) := \frac{1}{n}\sum_{i=1}^{n} f_i(z)$. Then, $x(\infty)$ is a unique minimiser of P and we have by the following equation

$$\| \nabla P(\bar{x}(\lambda)) \|^2 \le \frac{2L^2}{\lambda}(f(x(\infty)) - f(x(0))) \tag{7}$$

If $\alpha \le 1\ 2L$, then

$$\mathbb{E}\left[\|x^k - x(\lambda)\|^2\right] \le \left(1 - \frac{a\mu}{n}\right)^k \|x^0 - x(\lambda)\|^2 + \frac{2na\sigma^2}{\mu}$$

where $\mathcal{C} := \frac{1}{n}\max\left\{\frac{L}{1-p}, \frac{\lambda}{p}\right\}$ and

$$\sigma^2 := \frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{1}{1-p}\|\nabla f_i(x_i(\lambda))\|^2 + \frac{\lambda^2}{p}\|x_i(\lambda) - \bar{x}(\lambda)\|^2\right) \tag{8}$$

Three neural networks will be employed instead of the conventional two models that are trained as the generator and discriminator. In a modified method, two NN will function as a generator and a third as a discriminator. There will be three NN:

(a) Encryptor: Input is a shared key and plain text in binary order, which will produce the encrypted text.
(b) Decryptor: Shared key as well as encrypted text will be inputs, and the output will be the decrypted text.
(c) Eavesdropper: It will only accept input in the form of encrypted text, intercepting the text and performing decryption operations without using the shared key.

The generative adversarial network uses the following layers: The architecture of all three NNs is the same and consists of the following layers:

(a) Fully connected dense layer
(b) Convolutional layer
(c) Flatten layer

As a result, we utilised four convolutional layers, one flattening layer, and one thick layer. Strided convolution is used in place of the pooling layer: Between the convolution layers, the input matrix MaxPooling layer is utilised for down sampling, which reduces the overall number of training specifications in the network. But in this case, strided convolutions are utilised in place of a direct down sampling to enable the network to learn its own spatial sampling.

Utilised activation mechanisms: Binary sequences 0 and 1 are used for the encryption process. So, with the exception of the final layer, where sigmoid activation was utilised, we used tan$h$ activation to normalise the output of each layer in [1, 1]. Generator G receives random noise z as input, and the resulting sample distribution is defined as $Pg = G(z)$. The performance of the generator G is then gradually improved by using a discriminator D to continuously separate the generated samples from real data. The model's objective function is provided by the following equation:

$$\min_G \max_D = E_{x\sim P_2}[(\log(D(x)] + E_{z-P_7}[(1 - D(G(z)))] \tag{9}$$

The distribution and look of a steg image resemble those of the cover image when a discriminator is added to the steganography system. The loss function can be found in the following equation:

$$L_{\text{disc}} = \mathbb{E}_{c\sim P_c}[\log D(c)] + \mathbb{E}_{c\sim P_s, s\to P_s}[\log(1 - D(H(c, s))] \tag{10}$$

The training method for the hiding and extracting networks is improved to reduce the number of false positives in the reconstruction of the steganographic cover picture c_0 and the actual secret image, respectively. Mean square error (MSE), a popular choice for cost functions, only penalises pixel disparities between two pictures while ignoring their basic properties. To achieve this, we propose a new cost function that is minimised to improve the method by the following equation:

$$\underset{R_e}{\text{argmin}} \frac{1}{n}\sum_{i=1}^{n}(1 - \text{SS IM}(c_i, H_\theta(c_i, s_i)) \tag{11}$$

where s stands for the steg image, also known as the rebuilt secret picture, c for the cover image, also known as original secret image, and n for quantity of training samples.

4. RESULTS AND DISCUSSION

The machine utilised for the experiment has the following components: an Intel Core i5 7200 U processor, 8 GB of RAM, a 1 TB hard drive, and NVIDIA GTX 760MX graphics. Python 3.5 environments were utilised to evaluate how the suggested method works in practise. To establish results of the offered method, we carried out a statistical analysis by calculating expected performance.

Dataset description: The datasets were composed of the following variables: feed flow rate ($F = 400$–600 L/h), permeate flux (Pflux (L/h m^2)), condenser inlet temperature, evaporator inlet temperature, and feed salt content. The training division is typically used to get the method parameters. While the testing division checks its performance, the validation division assesses the training accuracy of continuing training to prevent overfitting.

Table 2 shows the results of a study using salt water as the comparative medium. $S = 35$–140 g/L, $T_{\text{cond}} = 20$–30 °C, and $T_{\text{evap}} = 60$–80 °C were the results of an examination of the water's salinity and temperature. In this research, we focus on training precision, throughput, network reliability, and scalability.

Water salinity analysis is shown in Figure 4(a)–4(f). The range of interest is from 35 to 140 g/L. This method achieves 86% accuracy in training, 74% throughput, 72% network integrity, and 63% scalability, whereas the conventional RL method achieves 82% accuracy in training, 70% throughput, 65% network integrity, and 58% scalability. The SVM method achieves 84% accuracy in training, 72% throughput, 68% network integrity, and 62% scalability.

For $T_{\text{cond}} = 20$–30 °C, the salinity of the water is compared in Figure 5(a)–5(f). We compare the results of the proposed method to those of the conventional RL (85% accuracy, 73% throughput, 66% network integrity, 64% scalability) and SVM (88 accuracy, 76% throughput, 69% network integrity, 66% scalability).

Figure 6(a)–6(f) shows a comparison with $T_{\text{evap}} = 60$–80 °C depending on the salinity of the water. In comparison to conventional RL training methods (93% accuracy, 82% throughput, 69% network integrity, and 65% scalability) and SVM

Table 2 | Analysis based on various water analysis

Water analysis	Techniques	Training accuracy	Throughput	Network integrity	Scalability
$S = 35–140$ g/L	RL	82	70	65	58
	SVM	84	72	68	62
	SGEA_FB	86	74	72	63
$T_{cond} = 20–30$ °C	RL	85	73	66	64
	SVM	88	76	69	66
	PRE_SWA_DL	91	80	74	71
$T_{evap} = 60–80$ °C	RL	93	82	69	65
	SVM	94	84	72	69
	PRE_SWA_DL	96	86	76	73

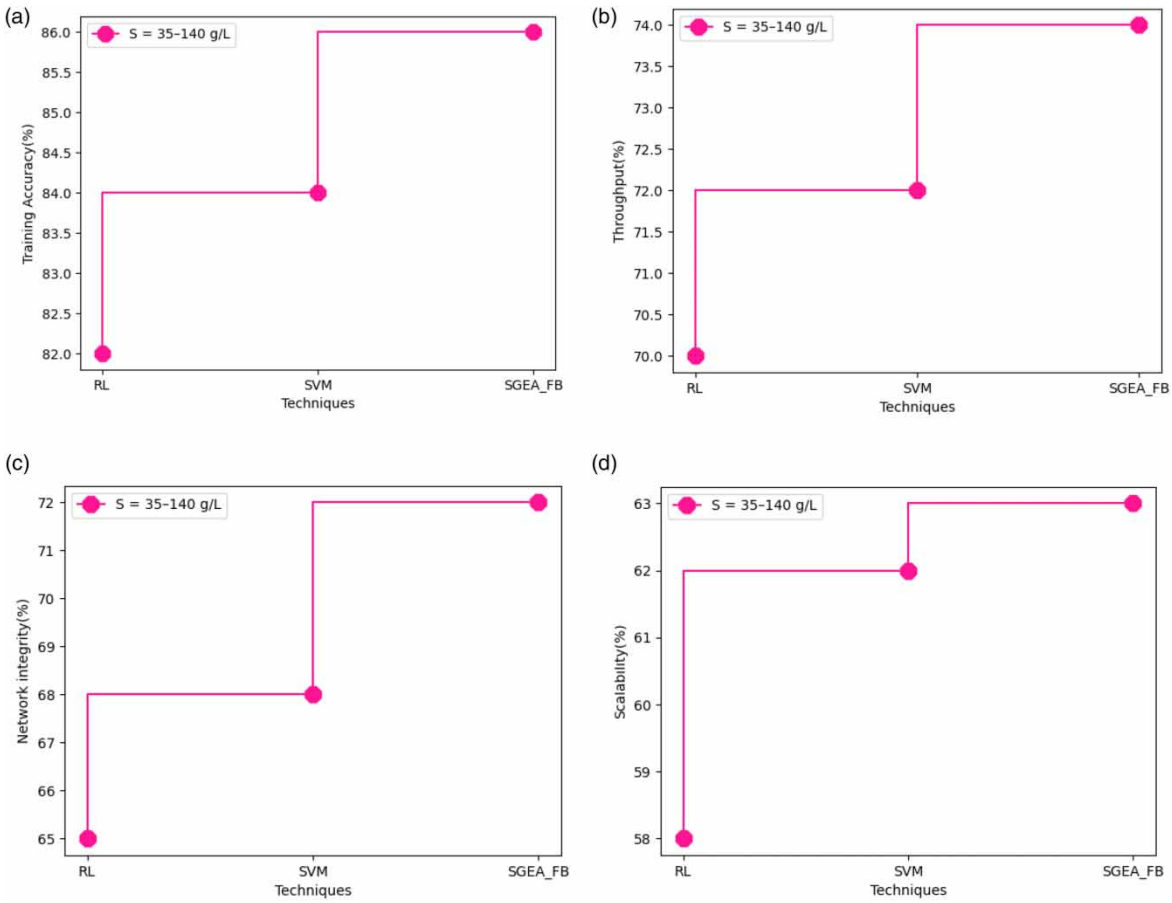

Figure 4 | Comparative for $S = 35–140$ g/L based on water salinity in terms of (a) training accuracy, (b) throughput, (c) network integrity, and (d) scalability.

training methods (94% accuracy, 84% throughput, 72% network integrity, and 69% scale), the suggested method boasts 96% accuracy, 86% throughput, 76% network integrity, and 73% scalability. For each modelling dataset, 70% was set aside for training and 30% was put aside for testing. Model performance was measured using (i) training data, (ii) testing data, and (iii) complete data. Each of the 'goodness-of-fit' metrics was applied to the results produced by many DL and ML techniques.

The findings indicate that weekly container cleaning is the most popular option, whereas most people wash their hands before handling water. They also utilised river water, although some of them never washed the container. In addition, just 27% of the family washed their hands after defaecating while 72% utilised soap to take a bath. Additionally, it was discovered that while 100% of users select tap water for cooking, only 73 and 49% of users, respectively, choose river water for bathing

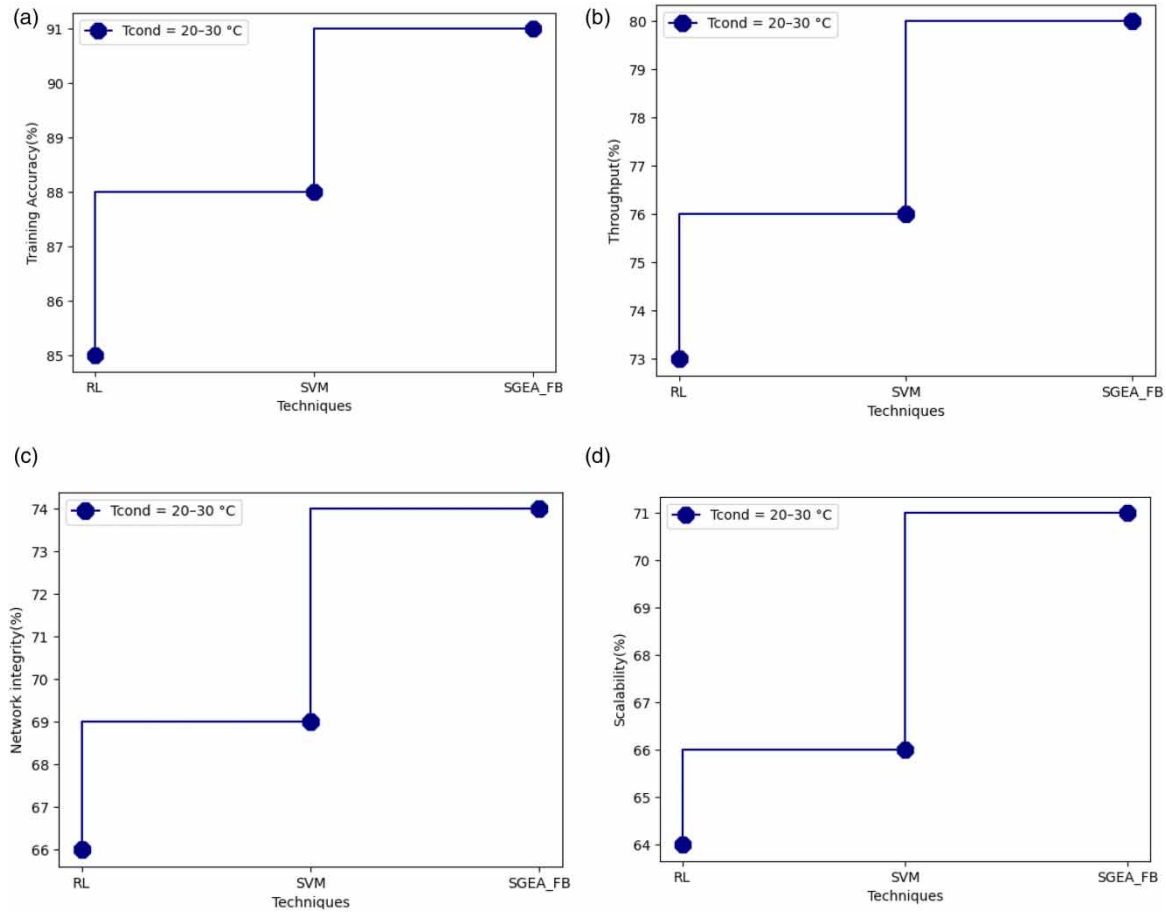

Figure 5 | Comparative for $T_{cond} = 20–30$ °C based on water salinity in terms of (a) training accuracy, (b) throughput, (c) network integrity, (d) scalability.

and cleaning. About 27% of the people urinate in the river since they do not have a toilet at home. These preferences for various water sources may be brought on by the following factors: (i) consumers demand high-quality water for consumption; (ii) water bills are expensive; (iii) it takes more time and effort to gather water from rivers or public taps; and (iv) there are often disruptions. No matter the source, the consumers never perform any kind of home treatment. The age of the water collector is classified into three groups by the simple factor analysis: 11–22: 45%, 22–33: 40%, and 33–44: 14%, respectively. About 90% of the users collect water twice daily, and average daily time spent collecting water was determined to be 44 min. All users physically gather water. The study also predicts that for the selected families, the average monthly income is less than 1,000 ETB for 56% of users, or 2,327 ETB, according to the factor analysis, while for 18% it is from 1,000 to 2,000, for 9% from 2,000 to 4,000, for 3% from 4,000 to 6,000, for 1.5% from 6,000 to 8,000, and for 14% from 8,000 to 10,000. This indicates that 74% of the customers come from low-income families, which may help to explain why the majority have selected dual water sources and find it difficult to pay their average monthly water cost of 62 ETB. Given the socioeconomic makeup of the research area, income is likely to have an impact on water preferences. The study's population has an education level that is below the 10th standard in 55% of the cases, the 12th standard in 3% of the cases, and only 21% of the people are graduates, according to the data on education. Additionally, the majority of low-income families have members with education levels below the 10th grade. In terms of health problems, diarrhoea (75%) and influenza (98%) are relatively prevalent diseases, as are other seasonal water-related illnesses like malaria and cholera (20%).

5. CONCLUSION

This study, by proposing a sustainable energy management system, also suggests a unique approach to water remediation and cybersecurity analysis for healthcare applications. The blockchain federated paradigm has been implemented for cyber

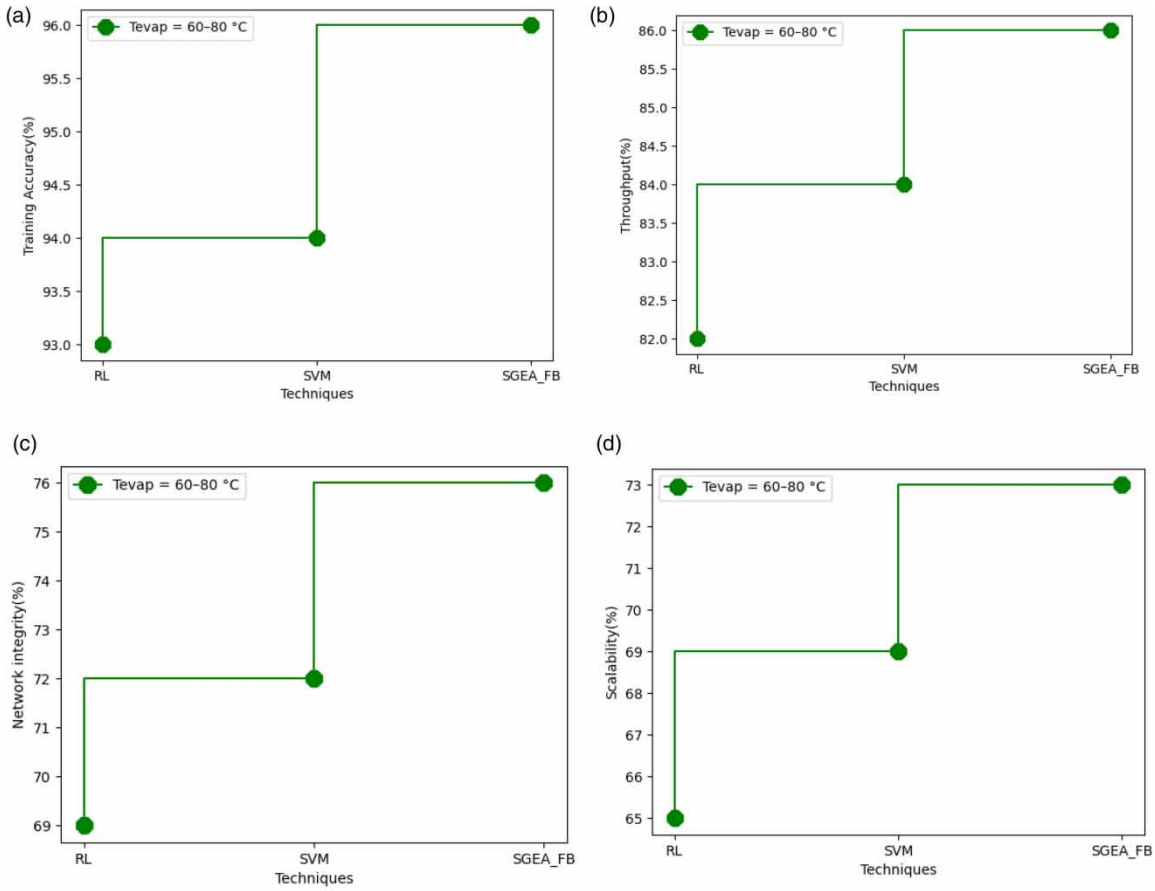

Figure 6 | Comparative for T_{evap} = 60–80 °C based on water salinity in terms of (a) training accuracy, (b) throughput, (c) network integrity, and (d) scalability.

analysis, and SG-based network energy management is implemented here. This study verified the reliability of multi-model ensemble prediction of chlorine residual concentrations in potable water. To choose the most applicable model, it is essential to carefully examine the available options. The models should be compared using a variety of efficiency indicators. Then, under different conditions of operation, this system is thermodynamically analysed by measuring its energy and exergy. The average annual energy efficiency is predicted to be 11.1%, while the average annual exergy efficiency is predicted to be between 12.5 and 15.9%. Furthermore, under optimal conditions, the energy efficiency of the membrane-based wastewater treatment subsystem is 73.3%. The proposed strategy was able to improve training accuracy by 96%, throughput by 86%, network integrity by 76%, and scalability by 73%.

ETHICAL APPROVAL

This paper does not contain any studies with animals performed by any of the authors.

DATA AVAILABILITY STATEMENT

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

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Ahmed, F., Aziz, M. S. A., Palaniandy, P. & Shaik, F. 2022 A review on application of renewable energy for desalination technologies with emphasis on concentrated solar power. *Sustainable Energy Technologies and Assessments* **53**, 102772.

Alawad, S. M., Mansour, R. B., Al-Sulaiman, F. A. & Rehman, S. 2023 Renewable energy systems for water desalination applications: A comprehensive review. *Energy Conversion and Management* **286**, 117035.

Anyaoha, K. E. & Zhang, L. 2022 Transition from fossil-fuel to renewable-energy-based smallholder bioeconomy: Techno-economic analyses of two oil palm production systems. *Chemical Engineering Journal Advances* **10**, 100270.

Bhavani, N. P. G., Harne, K., Singh, S., Abdukarimovich, O. A., Balaji, V., Singh, B., … & Mohanty, S. N. 2023 Economic analysis based on saline water treatment using renewable energy system and microgrid architecture. *Water Reuse*.

Caldera, U. & Breyer, C. 2019 Assessing the potential for renewable energy powered desalination for the global irrigation sector. *Science of the Total Environment* **694**, 133598.

Choi, W., Dong, F., Hatzell, M. & Mul, G. 2022 Solar energy utilization and photo (electro) catalysis for sustainable environment. *ACS ES&T Engineering* **2** (6), 940–941.

Danaeifar, M., Ocheje, O. M. & Mazlomi, M. A. 2023 Exploitation of renewable energy sources for water desalination using biological tools. *Environmental Science and Pollution Research* **30** (12), 32193–32213.

Fouladi, J., AlNouss, A., Bicer, Y. & Al-Ansari, T. 2022 Thermodynamic analysis of a renewable energy-water-food nexus: A trade-off analysis of integrated desalination, gasification and food systems. *Case Studies in Thermal Engineering* **34**, 102024.

Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S. & Sunny, M. S. H. 2019 Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access* **7**, 13960–13988.

Kiehbadroudinezhad, M., Merabet, A., Hosseinzadeh-Bandbafha, H. & Ghenai, C. 2023 Environmental assessment of optimized renewable energy-based microgrids integrated desalination plant: Considering human health, ecosystem quality, climate change, and resources. *Environmental Science and Pollution Research* **30** (11), 29888–29908.

Kowthaman, C. N., Kumar, P. S., Selvan, V. A. M. & Ganesh, D. 2022 A comprehensive insight from microalgae production process to characterization of biofuel for the sustainable energy. *Fuel* **310**, 122320.

Maggio, G., Squadrito, G. & Nicita, A. 2022 Hydrogen and medical oxygen by renewable energy based electrolysis: A green and economically viable route. *Applied Energy* **306**, 117993.

Mollahosseini, A., Abdelrasoul, A., Sheibany, S., Amini, M. & Salestan, S. K. 2019 Renewable energy-driven desalination opportunities – a case study. *Journal of Environmental Management* **239**, 187–197.

Nahar, K. & Sunny, S. A. 2020 Duckweed-based clean energy production dynamics (ethanol and biogas) and phyto-remediation potential in Bangladesh. *Modeling Earth Systems and Environment* **6** (1), 1–11.

Panagopoulos, A. 2021 Water-energy nexus: Desalination technologies and renewable energy sources. *Environmental Science and Pollution Research* **28** (17), 21009–21022.

Sayed, E. T., Olabi, A. G., Elsaid, K., Al Radi, M., Alqadi, R. & Abdelkareem, M. A. 2022 Recent progress in renewable energy based-desalination in the Middle East and North Africa MENA region. *Journal of Advanced Research*.

Siddiqui, O. & Dincer, I. 2021 Development and analysis of a new renewable energy-based industrial wastewater treatment system. *Journal of Environmental Management* **290**, 112564.

Vivek, C. M., Ramkumar, P., Srividhya, P. K. & Sivasubramanian, M. 2021 Recent strategies and trends in implanting of renewable energy sources for sustainability – a review. *Materials Today: Proceedings* **46**, 8204–8208.

Wang, F., Xu, J., Liu, L., Yin, G., Wang, J. & Yan, J. 2021 Optimal design and operation of hybrid renewable energy system for drinking water treatment. *Energy* **219**, 119673.

Zhang, Y., Anoopkumar, A. N., Aneesh, E. M., Pugazhendhi, A., Binod, P., Kuddus, M., … & Sindhu, R. 2023 Advancements in the energy-efficient brine mining technologies as a new frontier for renewable energy. *Fuel* **335**, 127072.

First received 19 September 2023; accepted in revised form 22 November 2023. Available online 2 February 2024