A Two-Layer Water Demand Prediction System in Urban Areas Based on Micro-Services and LSTM Neural Networks

AHMED ABDEL NASSER[ID][1]**, MAGDI Z. RASHAD**[ID][1]**, AND**
SHERIF E. HUSSEIN[ID][2]**, (Member, IEEE)**
[1]Faculty of Computers and Information Systems, Mansoura University, Mansoura 35516, Egypt
[2]Department of Computer Engineering and Control Systems, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt

Corresponding author: Sherif E. Hussein (sherif_hussein@mans.edu.eg)

ABSTRACT In recent years, scarce water resources became one of the main problems that endanger human species existence and the advancement of any nation. In this research, smart water meters were implemented, distributed, and installed in a regional area in Cairo while data were collected at uniform intervals then sent to the cloud instantly. The solution paradigm uses an Internet of Things (IoT) based on micro-services and containers. The design incorporates real-time streaming and infrastructure performance optimization to store data. A second layer to analyze the acquired data was used to model water consumption using Long Short-Term Memory (LSTM). The designed LSTM is validated and tested to be utilized in the forecast of future water demand. Moreover, two alternative machine learning methods, namely Support Vector Regression and Random Forest commonly utilized in time series forecasting applications, were used for a comparative analysis of which LSTM has proven to be superior. The proper integration of the system elements is the key to the proposed system success. Based on the success of the designed system, it can be applicable on a national scale. That can enable the optimal management of consumers' demand and improve water infrastructure utilization. The proposed paradigm presents a testbed for various scenarios that can be used in water resources management.

INDEX TERMS Water smart meters, time series, resources management, IoT, LSTM, micro-services.

I. INTRODUCTION

The smart water metering systems have just begun to gain momentum as water utilities started to use real-time data acquisition that can be stored and used in data analytics to save the scarce water resources in an optimal way [1]. One of the most crucial research directions supporting that trend is advanced metering infrastructure (AMI) that can offer a remote connection between water utilities [2]. However, the communication itself can take many forms, such as power fiber optics, cellular transmission, and broadband communication, among others [3]. In a smart water metering system, data can communicate between smart meters and water utilities with the support of analytical software architecture to take proper decision regarding certain actions to monitor and control the water supply in the system or to issue appropriate alerts to warn consumers or guide them to reduce their consumptions [4]. It can also predict distinctive patterns in

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Fazlul Kader[ID].

water consumption for a future forecast. Due to the complex nature of the water system which includes pumping stations, reservoirs, and consumer services. Accurate prediction could help manage water utilities to avoid problems that arise in the times of peak consumption or water leakage [5].

In deploying a smart water metering communications network, proper technology for data transmissions must be used. As there is a diverse number of options that differ in cost, popularity, reliability, scalability, and security, among other indicators, choosing the proper architecture and communication technologies can present a barrier to water utilities. Water utilities integration needs to be planned carefully to ensure the durability of the communications network [2]. Extensive research has been done for electricity consumption modelling, which has significant differences [6]–[8]. However, electricity metering can offer measurements of higher granularity and accuracy compared to water metering devices [9]. Moreover, energy consumption patterns are much more recognized and sometimes can be fixed, compared to the variable water patterns [10]. For example, the energy consumption of

most household devices can be directly calculated through their technical characteristics, which is not possible for showers as an example, even in the same household [11], [12].

The technology of the Internet of Things (IoT) has been utilized in various applications and is identified as one of the main factors of success for Smart Cities. Nowadays, in the IoT, data and its understanding are getting more important and remain the main concern rather than the objects that generate these data. To achieve data understanding, exchanging, and sharing for both information and knowledge, these objects need a lightweight and novel platform for the future provisioning of IoT services. The Web of Objects (WoO) is supported by inter-operable micro-services and the granularity of heterogeneous objects as well as virtualization through virtual objects composites. To implement the IoT of cross-domain applications, Jarwar *et al.* introduced a WoO enabled inter-operable micro-services architecture and demonstrated the implementation using a use case [13]. Moreover, the IoT dynamic environment behavior requires them to be able to evolve and scale over time, adopting novel technologies and various requirements. Micro-service architecture style has recently gained significant popularity in many fields due to the challenges in building large-scale, complex, and distributed applications and platforms on the Web. Krylovskiy *et al.* applied the micro-service architecture paradigm to design a Smart City IoT platform [14]. They suggested various benefits using their paradigm as compared to the other architectures' approaches.

Pau *et al.* showed that the power systems evolution using the smart grid paradigm is highly dependent on the distribution grids modernization [15]. They suggested that using new technologies, infrastructures, and applications is crucially required. Their research presented a smart metering infrastructure with a large set of possible services directed to the management and automation of distribution grids. Their architecture was based on a cloud solution, which facilitates the communication between the smart meters and the distribution grid services interface. Because a large number of applications can be implemented on the cloud, the focus was on enabling the automatic reconfiguration of the grid using a real-time distributed state-estimation algorithm. Kamienski *et al.* studied the Irrigation for agriculture which is considered the main consumer of fresh water worldwide [16]. The intensive use of technology could be useful to optimize the use of water, improve the crops quality and reduce the energy consumption. Kamienski and his colleagues claimed that even though, the IoT and other associated technologies are the normal choices for smart water management applications, there is still a debate about how appropriate those choices are in real world scenarios with the on-site pilots' deployment. Also, the platforms development of IoT-based applications should be suitable to different climates, crops, and countries. They proposed IoT based approaches and methods for smart water management in the precision irrigation domain and demonstrated their use in Spain, Brazil, and Italy. They presented two pilot studies based on a proposed architecture and scenario development process.

An additional research by Jarwar *et al.*, discussed the objects from data management infrastructure, various energy generation, and consumption terminals [17]. However, they emphasized that the acquired data is only useful when it is available on-time for services that extract meaningful information to achieve intelligent decisions. The micro-services-based data analysis, data caching, data processing, data virtualization, and data ingestion methods can be applied to enhance energy efficiency, management services provisioning, and data availability across different buildings. WoO offers mechanisms for data aggregation, abstraction, and ingestion with virtual objects and composite virtual objects using scalability, ontologies and availability of services with micro-services. Their research proposed the utilization of data processing micro-services modelling to improve data availability while exposing services capabilities with micro-services. They presented a semantic web agent based on an ontology for linking, availability, re-usability, and enhancement of services, data-objects, and micro-services. The authors presented a use case to evaluate their paradigm, which included data collection from different sources and processing and provision of various BEEMS. They mimicked the enhanced data availability for BEEMS using a use case scenario.

Moreno *et al.* emphasized that one of the most frequent and costly natural disasters that affect humankind is flooding [18]. Their developed architecture was based on the Message Queuing Telemetry Transport (MQTT) protocol, with security and encryption mechanisms, to send real-time data packages from fixed nodes to a server. The accessibility of data could be managed through graphical representations and customizable queries, to allow for flood analysis and prediction systems. Kamienski *et al.* explained that fresh water smart management for precision irrigation is essential to increase crop yield and decrease the involved costs while contributing to environmental sustainability [19]. The technologies utilization offers a means to provide the precise amount of water needed for plant irrigation. They demonstrated the suitability of the IoT for smart water management applications, although the different technologies integration is needed for a successful system. The developed project using smart water management platform based on IoT for precision irrigation in agriculture applied in four pilot studies in Europe and Brazil. The results showed that the system requires the re-engineering of some components and specially designed configurations to provide higher scalability with less computational resources. Therefore, the amount of forward-planning is considered the main obstacle in the adoption of many water utilities' solutions. While the ultimate objective is to focus on predicting the aggregate water consumption of urban areas populations.

In this research, the problem of data acquisition for short- and long-term water consumption was studied to help in forecasting and management. Moreover, a smart water meter

with the ability to send data to the cloud was designed, while a complete architecture solution for data acquisition in real-time was proposed. Data were aggregated to be analyzed further based on micro-services and machine learning techniques in an expandable and secure manner with high performance, considering the big data involved in the proposed system. The proposed system was divided into two layers; the first layer is for data acquisition and aggregation, while the second layer is for water demand forecasting using machine learning techniques that predicts water demand for different regions of households. The rest of this article is organized as: In section II, a related work is presented with recent developments in the field of water demand prediction based on AI techniques and cloud services. In section III, a full description of data acquisition, the proposed LSTM prediction, and the micro-services architecture are briefly explained. In section IV, the solution architecture design and evaluation are presented. Section V explains the results achieved by aggregating data collected from distributed smart meters, design and evaluation of the water consumption prediction model using LSTM. While section VI presents the conclusion and the potentials of the proposed paradigm for water consumption acquisition, water demand prediction, and water demand management.

II. RELATED WORK

Water development and saving efforts have lately focused on increasing user-consumer awareness by devising intervention scenarios that aim at educating users about their consumer behavior and guide them to reduce their consumptions. Research on data mining and machine learning techniques were recently used for short-term water consumption prediction and pattern recognition and on intervention methods that exploit these techniques to inform consumers and stimulate behavioral changes [20]. Besides, companies are currently investing in water monitoring devices that are installed on household bathroom faucets and measuring real-time water consumption, for online statistics and sending alerts to the consumers [9]. These interventions and alerts require that individual short-term consumption to be predicted as accurately as possible and in real-time so that they can be compared with future planned consumptions [21].

Several Artificial Intelligence techniques have been utilized by water demand forecasting over the last decades. In a recent study by Ghalehkhondabi *et al.*, they have investigated the research done during the period between 2005 and 2015 related to water demand forecasting based on Artificial Intelligence. They found that Fuzzy models, metaheuristic optimization, Artificial Neural Networks, and Support Vector Machines were the most commonly used techniques. They postulated that Artificial Neural Network was the most prominent method used in water demand forecasting. However, they concluded that it is still difficult to choose any method as the winner among other methods [22]. Even though, Artificial Intelligence methods and their hybrid were applied in water demand forecasting, researchers indicated

that further contribution is yet to be made to achieve a better water demand forecasting [23].

Muhammad, and Feng investigated several artificial intelligence techniques such as support vector machine, artificial neural networks, fuzzy logic, and extreme learning machines as well as hybrid models and Autoregressive Integrated Moving Average (ARIMA) in urban water demand forecasting. They concluded that artificial intelligence methods showed superiority especially Artificial Neural Networks for short-term water demand forecasting [24].

Papageorgiou *et al.* proposed a time series prediction hybrid approach based on Fuzzy Cognitive Maps and Artificial Neural Networks. Their proposed method aimed to select the interconnections and attributes for time series prediction following the training stage. They compared the proposed approach prediction with real data of daily water demand to validate the model performance [25].

Shabani *et al.* proposed a Support Vector Machine model based on the polynomial kernel function to predict monthly water demand in a use case city in Canada. They aimed to assess phase space reconstruction before the input variables combination design. They concluded that their approach could achieve satisfactory lag time which in turns improve the support vector machine model performance [26].

Recently, cloud computing services became an integral part of any modern system among both corporations and individuals because of its vast and flexible facilities. Therefore, the huge computing demand can only be met by the cloud computing infrastructure which can lead to an ever-growing complexity to meet both quality of service and service level agreement [27]. Narayanan *et al.* proposed an underground water distribution system based on an IoT architecture that is integrated with Fog computing. To achieve that design in a smart city, the authors forecasted the customers water demand. They used ARIMA to predict the daily demand for a period of three months in their case study. Afterwards, the water distribution system based on an IoT architecture was designed using hydraulic engineering to distribute water with minimal losses [28].

The related work even though offers many soft computing methods for predicting water demand, it lacked the possibility to accurately model short-term water demand. Moreover, many of the current research aim to forecast the overall future water consumption instead of granular water demands. The literature failed to offer means for fully integrated systems with online training and the possibility to incorporate cloud services that can manage water demands based on regions and using advanced and flexible methods that can adapt to the ever-changing water demand behavior of individual users. Therefore, Proof of concept needs to be proposed for a national scale system that offers a layered infrastructure that can be expanded to utilize the full ICT capabilities.

III. METHODOLOGY

This section describes the general methods and techniques used in the model design, data acquisition, preparation, and

evaluation of the solution architecture. The steps used to describe the proposed two-layer water demand prediction system are shown in figure 1. While the next subsection explains in detail the design and implementation of the smart water meter, followed by an explanation of time series prediction techniques, the proposed neural network model, alternative machine learning techniques, regression accuracy metrics, and the benefits of micros-services architecture to provide useful possible processing and management solutions.
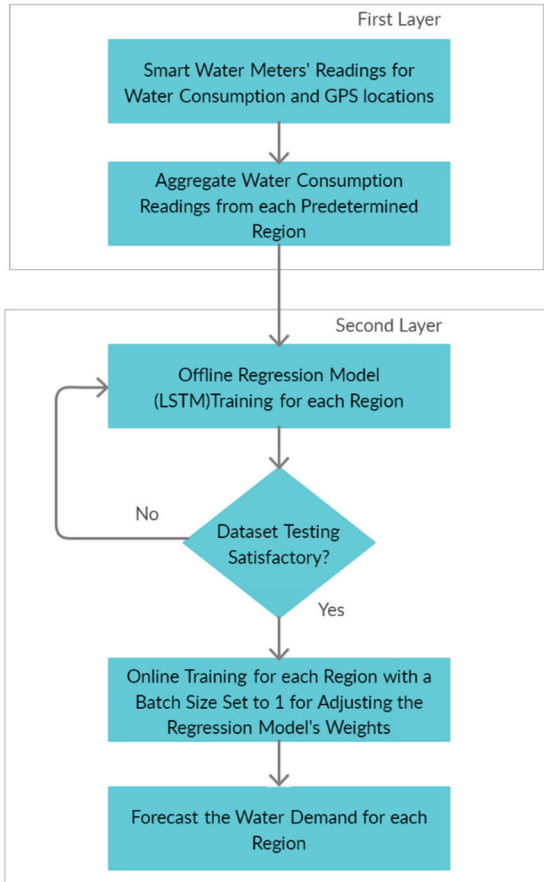

FIGURE 1. A flowchart that describes the steps of the proposed two-layer water demand prediction system.

A. SMART WATER METERING

The methodology is focused on an autonomous measuring unit that is used with sensors to monitor water consumption along with GPS information. The design prototype for the smart water meter shown in figure 2, is composed of a microcontroller connected to the internet through a Wi-Fi module, a water flow sensor and a GPS sensor [29], [30].

Station, Wi-Fi access point, and microcontroller are all features that can be found in the nodeMCU Dev board. That combination of features made the development board a versatile tool for both IoT applications and Wi-Fi networking. Moreover, it can also be used as an access point, station host a webserver, or upload and fetch data to MQTT brokers through the internet. Therefore, it was chosen to interact with both the
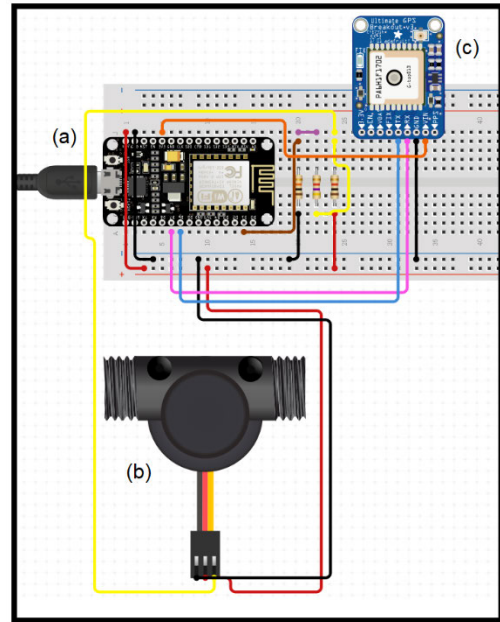

FIGURE 2. Detailed Design for the smart water flow meter; (a) ESP8266 12E NodeMCU, (b) water flow sensor G1/2″, and (c) adafruit ultimate GPS.

GPS module and the water flow sensor. The water flow sensor consists of a water rotor, a hall-effect sensor, a water rotor, and a plastic valve body. The sensor operates when water flows through the sensor rotor causing it to roll. The rotor speed changes with the water flow and the corresponding hall effect sensor accordingly changes the output. The water sensor can sense a range of water starting from 1 m^3/min up to 29 m^3/min with a sensitivity of 1%. The GPS module is built using the MTK3339 chipset which can track up to 22 satellites with a built-in antenna and a receiver sensitivity of -165 dBm tracking). While the GPS module can also make 10 location updates per second suitable for high sensitivity and high-speed tracking or logging applications. Also, it has a very low power consumption of only 20 mA when the update rate ranges from 1 to 10 Hz with a position accuracy less than 3 m. The water smart meter was validated by measuring the water flow of several predetermined water quantities with the corresponding time duration to determine the accuracy for each meter which was found to be below 2% of the measured quantity.

Because this research is based on collecting water consumption data from different households distributed among a neighborhood in an urban area, a selection for suitable points to install the smart meters were predetermined to evenly cover as many houses as possible [1]. Therefore, distributed smart meters were installed across 20 households in a region located in Maadi district in Cairo, according to the map shown in figure 3 with blue markers.

B. TIME SERIES REGRESSION

Water demand prediction is considered a use case from time series prediction. Time series modelling is an active area of

FIGURE 3. Smart water flow meters' distribution in the pilot study indicated by blue markers and generated by Scribble Maps (http://www.scribblemaps.com).

research that has attracted a lot of attention recently. The main objective of time series modelling is to collect and analyze past time series observations to develop a suitable model which describes the basic pattern of the time series. The model is used to predict future values for the time series. The Autoregressive Integrated Moving Average (ARIMA) is considered one of the most popular and frequently used stochastic time series models that captures a suite of different standard temporal structures in time series data [31], [32]. ARIMA model has subclasses of other models, such as Moving Average (MA) which uses the dependency between an observation and a residual error from a moving average model applied to lagged observations [33], the Autoregressive (AR) which uses the dependent relationship between an observation and some number of lagged observations [34], and Autoregressive Moving Average (ARMA) model that combines both MA and AR [35]. The adoption of the ARIMA model is due to the simplicity to represent varieties of time series as well as the possibility to associate the Box-Jenkins methodology that suggests an iterative three-stage approach to estimate ARIMA model's numerous parameters and hyper-parameters for optimally building the model [36], [37]. However, these models are assumed to be in a linear form, which is not suitable for many situations. To overcome this limitation, a few non-linear stochastic models have been

proposed [38], [39]; however, the implementation process is not simple or straight forward as the ARIMA models. On the other hand, Holt Winters extended the idea of simple exponential smoothing by comprising the forecast equation and three smoothing equations; one for the level, one for the trend, and one for the seasonal component, with corresponding smoothing parameters which results in accurate predictions for univariate time series data [40]. Recently, the use of artificial neural networks (ANNs) in the domain of time series forecasting has attracted increasing attention [41]. The main benefit of ANNs is their capability of non-linear modelling when applied to time series prediction, without any a prior knowledge about data statistical distribution [42]. The time series model is formed based on the given dataset using adaptive techniques. Due to these features, ANNs are naturally self-adaptive and data-driven [43].

A breakthrough in time series forecasting occurred with the recent advances in cloud computing and the ability to solve very complex mathematical formulations over many servers as well as streaming and storing data across multiple locations which opened the way for Deep Learning Neural Networks (DLNNs) to be practically used to solve highly complex problems. DLNNs can be used to solve pattern classification problems and can be applied to other fields such as regression, function estimation, signal processing,

and time series forecasting problems [44], [45]. The main advantage of DLNN is the ability to achieve better training data generalization. DLNN adds the ability to model the sequence dependence complexity among the input variables compared to regression predictive modeling. A special type of DLNN called recurrent neural networks is designed to handle sequence dependence. To elaborate on the deep learning methods for tuning the coefficients involved in the Holt's Winter method, Recurrent Neural Network (RNN) is able to learn prediction from sequences of data and a variance of RNN called Long Short-Term Memory (LSTM) is able to learn from even longer sequences of data. Others have used SVM regression as an alternative machine learning technique for time series forecast [46]. However, a few recent comparative research studies have favored LSTM over SVM regression regarding the accuracy of both methods [47]–[49].

C. LONG SHORT-TERM MEMORY

RNN is a category of Artificial Neural Networks that can learn long term dependencies that is useful when the network needs to retain information over long time periods. That means it can handle successive sequence of events in which the understanding of each even is based on previous events. Moreover, the deepest the RNN, the longer the memory period and consequently better capabilities can be achieved. However, RNN has its limitation because of the vanishing gradient problem due to its architecture restriction to long term memory capabilities. Therefore, a special type of RNN namely LSTM are designed to solve those problems to allow it to retain information for longer periods of times.

LSTMs have the ability to maintain a constant error that allows them to recursively learn through both time and layers. Additionally, as seen in figure 4, LSTMs use a special type of cells called gated cells that can store information in a different way compared to the RNN and allow to read from them. Each cell can make a decision by its own regarding the information while closing and opening their cells to execute those decisions. The LSTMs architecture are like chains allowing them to contain information over long time periods to solve problems that RNN might fail to solve.

LSTM consists of three main parts including; a type of gates called input gates that can add information to the cells; a type of gates called forget gates that allow to remove information when they are not necessary anymore; and a third type of gates called output gates responsible for selecting and outputting the necessary information. The compact forms of the LSTM unit equations for the forward pass are:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \tag{1}$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{3}$$

$$\widetilde{c}_t = \sigma_h(W_c x_t + U_c h_{t-1} + b_c) \tag{4}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ + \widetilde{c}_t \tag{5}$$

$$h_t = o_t \circ \sigma_h(c_t) \tag{6}$$

where $h_o = 0$, $c_o = 0$ the initial values, the subscript t is the time step, and the operator o represents the Hadamard product.

x_t is the LSTM unit input vector; f_t is the LSTM unit forget gate's activation vector; o_t is the LSTM unit output gate's activation vector; c_t is the cell state vector; i_t is the LSTM unit input/update gate's activation vector; h_t is the output vector of the LSTM unit; $\widetilde{c}_t$ is the cell input activation vector; and W is the bias vector parameters and weight matrices which need to be learned during training. The Activation functions σ_g and σ_h are sigmoid function and hyperbolic tangent function respectively.

The LSTM neural network uses deep learning to address the problems associated with the time series complexity in large architectures [50].

D. ALTERNATIVE ML REGRESSION METHODS

There are several alternative machine learning methods that are commonly used in time series forecasting as reported in the recent related literature. Among those methods is Support Vector Regression that is proposed for estimating the continuous function of training datasets. It is able to model complex nonlinear relationships by using an appropriate kernel function that maps the input into higher dimensional feature space and transforms the nonlinear relationships into linear forms Since previous studies endorsed the significance of the RBF kernel, it was used also in this work for the development of the SVR [51].

Random Forest is another successful regression technique. It uses multiple learning algorithms for forecasting both classification and regression problems. RF combines the results of decision trees trained by the "bagging" method. RF is one of the most successful Artificial Intelligence techniques among the current algorithms that use decision tree methods. It can handle large number of input variables [52].

As the forest building progresses, it estimates the generalization error. Moreover, it is a superior method in estimating the missing data while maintaining good accuracy. Besides, it is a relatively fast method that can produce a forest of decision trees for both regression and classification use cases.

E. REGRESSION MODEL ACCURACY METRICS

Root Mean Square Error (RMSE) is one of the main accuracy measures which can estimate how accurate the model can predict a certain response in regression problems. The RMSE is the calculation of the square root of the residuals' variance. It can indicate the model fitting to the data and how close those data to the model's predicted values. The lower the RMSE the better model accuracy.

On the other hand, Mean Absolute Error (MAE) represents another accuracy measure specially if there are outliers in the time series. It is the absolute value of the difference between the actual value and the forecasted value. Therefore, MAE estimates expected error from the forecast on average.

Another accuracy metric called the Mean Absolute Percentage Error (MAPE) is a widely used forecast accuracy
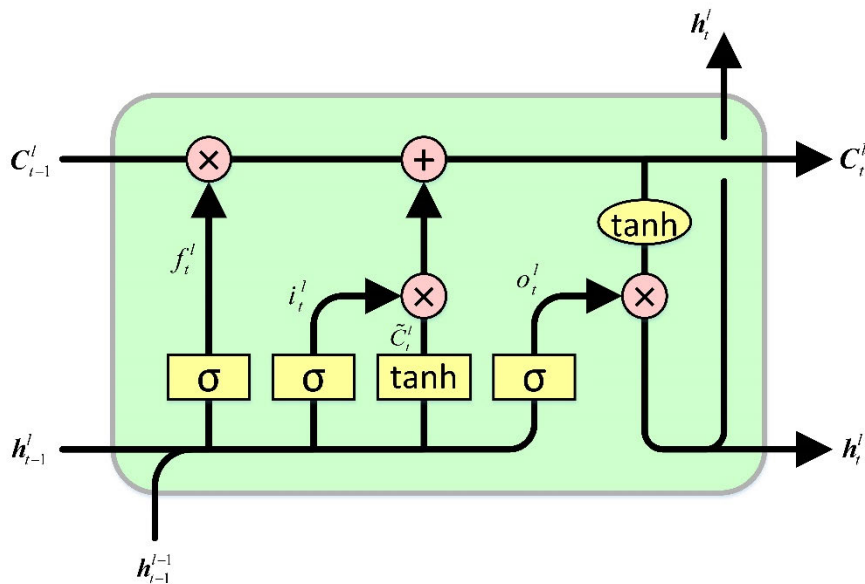
FIGURE 4. LSTM neural network architecture.

metric, because of its benefits regarding interpretability and scale-independency. However, MAPE has its limitation as it produces undefined or infinite values if the time series have zero or close-to-zero actual values. To solve that issue, an alternative forecast accuracy measure called the Mean Arctangent Absolute Percentage Error (MAAPE) is used in this research. MAAPE has been developed to correlate with MAPE. Hence, MAAPE rely on the slope as an angle, while MAPE relies on the slope as a ratio. Therefore, MAAPE can inherently preserve the MAPE philosophy and at the same time overcome the problem that might be caused from the division by zero using bounded influences for outliers in a fundamental manner. That could be achieved through considering the ratio as an angle instead of a slope [53]. Therefore, the results can be verified in a quantitative way from the performance metrics of RMSE, MAE, and MAAPE.

F. MICRO-SERVICES MOTIVATION

The use of cloud computing is an essential constituent of IoT as it is an IT paradigm that offers the ability of ubiquitous access to shared pools of configurable system resources and provides higher-level services that can be provisioned with minimal management effort and time, often over the Internet. Moreover, cloud computing is based on resources sharing to achieve economies and coherence of scale, like public utility [54].

On the other hand, when millions of objects communicate and exchange information between IoT applications the single business logic will result in a highly complex system. When the system is broken into small parts with micro-services architecture; it can dispel the complexity of the system [12]. Micro-services offer rapid development, loose coupling, lightweight, scalability, Interoperability, Single Task-Oriented, Broken Object Avoidance, Load Balancing, Strong Modularization, Plug & Play, Decentralized Governance, and Decomposability. Therefore, microservices are considered one of the most promising modern technologies that can improve the cloud processing capabilities. Containers is considered an efficient way to develop and deploy micro-services which can be thought-out as an operating system virtualization in which workloads can share operating systems resources. Even though, they have been in used just recently, they are widely adopted with an impressive acceptance among business executives and IT professionals who are already using containers in mission critical workloads. While the rest of the business executives and IT professionals are making plans to incorporate the technology in their future systems. Moreover, containers can succeed in services that virtual machines can fail to do in the development environment [55]. Some of their distinctive advantages are their ability to be launched or abandoned instantly. Besides, they do not need an operating system overhead in the container environment as opposite to the virtual machine environment. Therefore, containers can be considered a milestone that can play a vital role to simplify the development transfer from one platform or environment t another. However, as found in all technologies, there are challenges associated with the container's technology as most container-based applications are stateless [56]. Although, this is an issue for stateful applications, there are workarounds to solve that problem. One possible solution is to provide the reliable storage necessary to support stateful applications [57]. In addition, as with all platforms that deal with big data, data security can be a major concern. However, many attempts were proposed to solve those issues when containers are deployed in critical areas [58].

One of the most important advantages of cloud computing is its developer productivity. As mentioned above,

developers can instantly start up their own cloud instances, provision the component they want, and scale down and up easily [59]. Moreover, containers can be an ideal technology when developers need to shift their utilities between private cloud, on-premises, and public cloud architectures. Containers can be moved quickly and with minimal disruption because they are independent of the underlying operating system and infrastructure. Many organizations use multiple public clouds providers and can shift their workloads back and forth, depending upon certain performance criteria such as price special offers from the service providers. Therefore, containers make this process simple, reliable, and economical specially for building IoT applications [60].

IoT devices are composed of various sensors that can generate many data points, which can be acquired at a high rate. A simple temperature sensor may generate a few bytes of data per minute, while a complex assembly station might generate gigabytes of data in just few seconds. These huge amount datasets are ingested into the data processing pipeline for transformation, storage, querying, processing, and analysis [61]. Each dataset is comprised of multiple data points that represent specific measures. For example, a connected ventilation, heating, and air conditioning system would provide desired temperature, ambient temperature, air quality, humidity, load, energy consumption, and blower speed measures.

In a large shopping mall, these data points are collected frequently from hundreds of appliances. Since these devices may lack the power to run the full TCP networking stack, they may use other protocols like ZigBee and Z-Wave to send the data to a gateway that can aggregate the data points and process in the system [62].

MQTT is one of the most popular connectivity protocols in IoT that is used in this research. MQTT is a very lightweight messaging protocol that can operate with a constrained resource such as low memory, bandwidth links, and processing capability for IoT devices. MQTT has been utilized in various fields, including energy monitoring, smart cities, healthcare, and so on. MQTT protocol is built on top of TCP/IP protocol enabling IoT devices to connect to the Internet. MQTT is a Client-Server messaging protocol. MQTT consists of three components: publisher, subscriber, and a broker [63].

IV. THE INFRASTRUCTURE ARCHITECTURE SOLUTION

The proposed solution starts with a Smart Water Meter as the IoT device in the proposed paradigm which "talk" to the cloud to send the water flow measurements and the GPS information. When data is already in the cloud, the software processes it and decides whether to perform an action without the need for user intervention.

The IoT gateway plays an important role in the translation between sensors protocols, sensor data aggregation, and sensor data processing before to be sent onward. Because there can be several connectivity models, protocols, and energy profiles associated with the dispersed nature of the

IoT systems, gateways are the means to control and manage these complex environments.

However, for a higher throughput and lower latency, an MQTT proxy was used to communicate with Cloud IoT Core and publish telemetry events on behalf of bound devices as was the scenario in this research.

The MQTT proxy pushes the collected data of water consumption to an Apache Kafka cluster in docker containers, where data can take multiple paths. Kafka is a distributed, reliable and fault tolerant streaming platform which is best suited for the proposed infrastructure. It is followed by Apache Spark that consumes data from Kafka to perform some analytics and build predictive models. Data points that need to be processed in real-time go into the hot path in which an LSTM neural network was previously taught to predict the water demand. At the same time, water consumption can be analyzed after acquiring them over a certain period. These data points are collected and analyzed through a process that takes the data processing pipeline cold path. The data points are fed through the cold path for online training of the LSTM neural network to update its parameter in an offline manner. In this paradigm, it is important to track water smart meter readings in real-time to correct the measured data. These data points go through an Apache Spark cluster for almost real-time processing, as shown in figure 5.

No matter which path that the data points will pursue, they will finally be ingested by the Spark ML Pipeline interface into the system. Apache Kafka is considered a high-performance data ingestion layer dealing with huge datasets. While, the data processing pipeline components responsible for cold path and hot path analytics will act as subscribers of Apache Kafka.

V. RESULTS AND DISCUSSION

A. DATA ACQUISION AND AGGREGATION

The main purpose of data collection through the implemented IoT system is to acquire enough data for further "machine learning" processing stages. In the backend, the measured data is being evaluated using a big data engine. This is necessary since the amount of data is increasing enormously, and there must be a backend with a large amount of processing power and memory to process and correlate the various measured quantities. The data in this research were collected with consent from all the tenants and are installed in a neighborhood in a superb located in Maadi district. The acquired data contains drinking water consumption measurements collected from 20 residential houses along with their GPS information. The measured data included water consumption information from distributed locations, where water smart meters were installed. Data were collected during the years 2017 and 2018 to be used in the offline stage of inspection to discard erroneous measurements from the datasets. As a result, there were 20 datasets taken over 12 months in such a way that avoided any inconsistencies in the measurements. The sampling rate of which data was
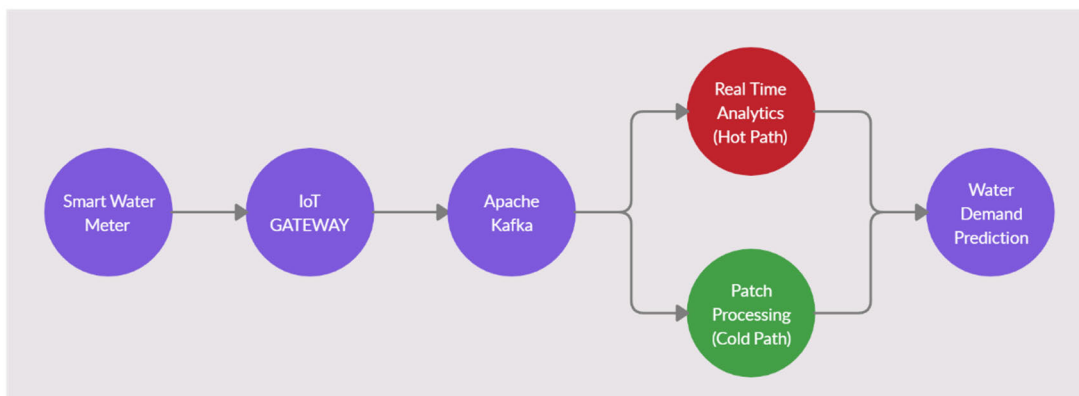
FIGURE 5. Abstract architecture for water consumption data prediction.

sampled was kept constant. Measurements were collected every 10 min. That rate was synchronized among all smart water meters to be aggregated every 10 mins or multiple of that time duration. Therefore, the resulted data covered a duration of 12 month-long, with 10 mins resolution that reveals volumetric water consumption at participating households. Moreover, the aggregate dataset can be used to reveal other features necessary for water demand management. Figure 6 shows the aggregated water consumption for two weeks, collected from different numbers of households. It is obvious that the more households participating in the data acquisition stage the better pattern of repeatability in the aggregated data. We have used Apache Spark that is good for finding some unexpected correlations in the acquired data sets and can stream them simultaneously for machine learning and batch processing. Moreover, it has an in-built interactive mode and the execution occupations of 10 to 100 times quicker than Hadoop MapReduce. In addition, Spark uses Resilient Distributed Datasets [63], which is the reason behind its higher computational performance than Hadoop. In addition, Spark can achieve real-time analytics because of its streaming module which is known as Spark streaming [64].

B. WATER DEMAND PREDICTION

The LSTM network is a recurrent neural network that is trained using Backpropagation. LSTM is used to address difficult sequence problems in machine learning to achieve optimal results. Instead of neurons, LSTM networks have layers of memory blocks. A block consists of components that make it outperforms a traditional neuron combined with a memory for recent sequences. The block has gates that manage both the block's state and the output. It operates on an input sequence and each gate within a block uses activation units to control its triggering state, making the change of state and addition of information flowing through the block to be conditional [65]. LSTM can achieve adequate learning and memory from one layer of LSTMs. Therefore, the use of higher-order abstractions can be layered with multiple of such layers to achieve better performance [66].

LSTMs are sensitive to the scale of the input data, specifically when the sigmoid or tanh activation functions are used. So, data were rescaled to the range of 0-to-1, prior to be trained and tested [67]. With time series data, the sequence of values is important. Therefore, the ordered dataset was split into training and testing datasets with 70% of the observations used to train the model, leaving the remaining 30% for the model testing.

The optimal batch size depends on the task as it limits the number of samples to be shown to the network before weight is updated. This same limitation is imposed when making predictions with the fitting model. One solution to this problem is to fit the model using online learning. This can be achieved by setting the batch size to a value of 1 while updating the network weights after each training example. This can have the effect of faster learning but can also add instability to the learning process as the weights widely vary with each batch. Therefore, we optimized both the number of neurons in the hidden layer and the batch size in the offline training stage then kept the number of neurons in the hidden layer while selecting a batch size of 1 in the online training and prediction stage. A mean squared error optimization function is used for this regression problem with the Adam optimization algorithm. The Adam optimization algorithm is an extension to stochastic gradient descent that combines both the root mean square propagation, and the adaptive gradient algorithm [68]. The LSTM parameters, namely the number of neurons in the hidden layer and the batch size, were found to be 10 and 6, respectively using 10-fold cross-validation. Therefore, the LSTM parameters chosen in the online training and prediction stage are set to an input layer with 1, 2, or 3 inputs, a hidden layer with 10 LSTM neurons or blocks, and an output layer with a single value prediction. The LSTM blocks used sigmoid activation functions and a batch size of 1 while the number of epochs was limited to 300 to decrease the training time as there was no significant increase in the model regression accuracy beyond that value.

Once the model has been trained using the training dataset, the performance of the model could be estimated to give metrics suitable for comparison. Then, the predictions were
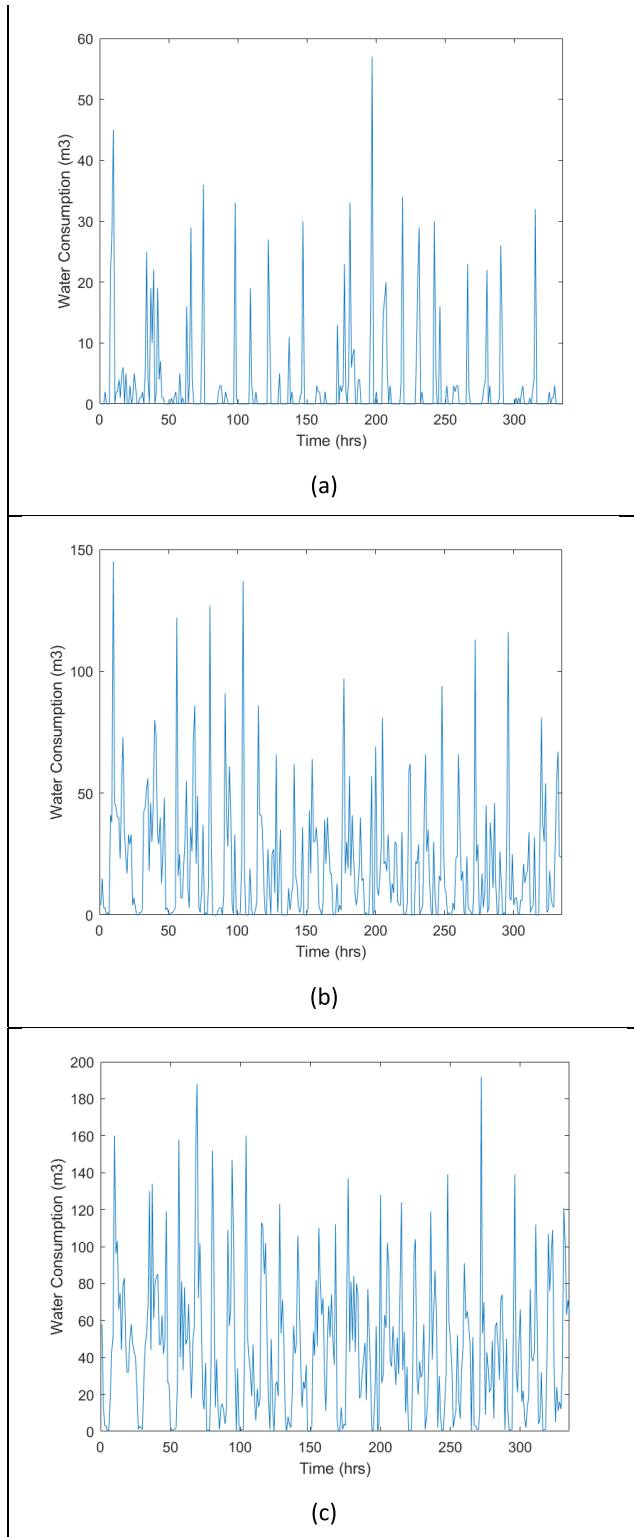
FIGURE 6. Aggregated Water consumption for two weeks based on:
(a) Two households, (b) 10 Households, (c) 20 Households.

inverted before calculating error scores to ensure that performance is reported in the same units as the original data (m³).

Predictions are generated using the LSTM model and compared with the testing dataset to get an indication of the

model performance. The predictions were shifted so that they align on the x-axis with the original dataset. Figure 7 is showing the original dataset in blue, the predictions for the training dataset in orange, and the predictions on the unseen testing dataset in green. The real aggregated datasets for 2, 10, and 20 households were used to model the water demand prediction with 3 different LSTM architectures for each. The first architecture uses one recent time step to make the prediction for the next time step. The second architecture uses two recent time steps while the third architecture uses three recent time steps to make the prediction. It can be noticed that the greater the number of aggregated households' datasets, the better the periodicity that can be captured by the LSTM model. In addition, the third architecture with three recent time steps performed a better job capturing the relation of water consumption prediction in m³ and time in hrs. as will be furtherly evaluated and assessed using the suitable accuracy metrics.

Support Vector Regression and Random Forest were trained on the same datasets while their corresponding models' performance metrics were evaluated for the sake of completeness. SVR was chosen for the comparative study because of its popularity in the water demand forecasting as reported in the related literature. While RF proved to be successful in several time series applications. The SVR has used a radial basis function kernel with a resulted accuracy performance comparable to what has been achieved using LSTM. While RF was outperformed by both LSTM and SVR methods. However, LSTM always consider long term dependencies and evaluate new value after understanding the whole series pattern. Whereas SVR and RF consider each row as a sample for training data and predict the outcome accordingly and will not consider the previous patterns. Therefore, LSTM can be superior in its deep learning capabilities while using large sizes of datasets.

Support vector regression modelling was applied to the dataset using the Gaussian RBF kernel. The three associated hyperparameters are the penalty factor C, the insensitivity parameter ε, and the Gaussian RBF function parameter σ. The value of C acts as a regularization parameter such that a very small C means a negligible penalty, while for a large C, a penalty gets more important and SVR tries to fit the data. The influence of ε affects the model complexity as for a very small value of ε there is not enough margin to include the data points and the SVR function tries to fit the data, but for a large value of ε there is enough margin causing a tendency for the model to get flat. On the other hand, a very small value of σ means the kernel is more localized resulting in a tendency to overfit, while a large value of σ makes ε less flexible. The optimal hyperparameters used in this research to reduce the 10-fold cross-validation loss, were found using the Bayesian optimization algorithm. The hyperparameters, namely, ε, and σ were found after 1000 epochs to be 36.51, 0.021, and 0.083, respectively.

On the other hand, the RF regression hyperparameters includes the depth of the trees in the forest. Deep trees tend
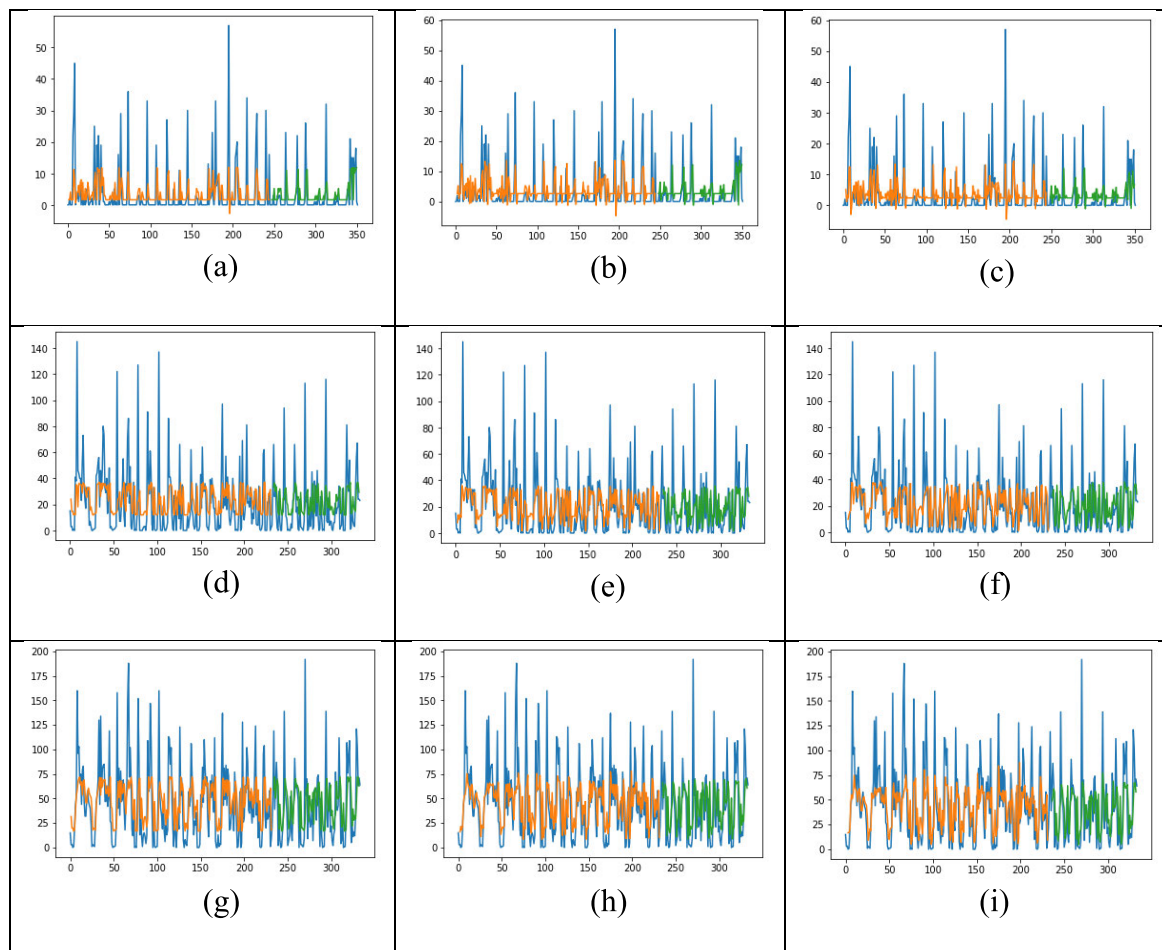
FIGURE 7. LSTM prediction for different number of aggregated households water consumption over two weeks based on: (a). (b), (c) Two households for 1, 2, and 3 inputs respectively, (d), (e), (f) 10 Households for 1, 2, and 3 inputs respectively, (g), (h), (i) 20 Households for 1, 2, and 3 inputs respectively. The blue curve represents the actual dataset while the orange curve is the model response based on the training dataset and the green curve is the model response based on the testing dataset.

TABLE 1. The RMSE in m^3 for different LSTM neural networks that has 1, 2, and 3 inputs applied to aggregated datasets from 2, 10, and 20 households as compared to support vector regression with the RBF kernel and random forrest.

	2 Households		10 Households		20 Households	
	RMSE (Testing)	RMSE (Training)	RMSE (Testing)	RMSE (Training)	RMSE (Testing)	RMSE (Training)
1 input	6.07	8.25	22.60	24.63	35.36	34.71
2 inputs	6.09	8.15	22.35	24.59	34.80	34.63
3 inputs	6.08	8.14	22.07	24.59	**34.32**	34.76
SVR	7.32	9.64	23.51	25.72	**37.11**	36.82
RF	8.41	10.72	27.14	29.33	**41.51**	40.71

to overfit, but shallow trees tend to underfit. When growing the trees, the number of predictors to sample at each node can range from 1 to all the predictors. Because the ensembles with more learners are more accurate, the number of trees

in the ensemble needs to be tuned due to the tendency of the Bayesian optimization to choose random forests containing many trees. Therefore, models containing many learners were penalized, as the available computation resources is

TABLE 2. The MAE in m³ for different LSTM neural networks that has 1, 2, and 3 inputs applied to aggregated datasets from 2, 10, and 20 households as compared to support vector regression with the RBF kernel and random forrest.

	2 Households		10 Households		20 Households	
	MAE (Testing)	MAE (Training)	MAE (Testing)	MAE (Training)	MAE (Testing)	MAE (Training)
1 input	1.24	3.69	4.86	11.95	7.84	18.17
2 inputs	1.07	3.29	4.83	12.14	7.62	18.04
3 inputs	1.07	3.30	4.53	11.57	**7.38**	17.71
SVR	1.37	4.13	5.71	13.86	**8.41**	19.92
RF	1.46	4.42	6.33	15.42	**8.81**	21.56

TABLE 3. The MAAPE for different LSTM neural networks that has 1, 2, and 3 inputs applied to aggregated datasets from 2, 10, and 20 households as compared to support vector regression with the RBF kernel and random forrest.

	2 Households	10 Households	20 Households
1 input	1.21	0.90	0.72
2 inputs	1.18	0.90	0.70
3 inputs	1.18	0.89	**0.69**
SVR	1.33	1.05	**0.92**
RF	1.42	1.11	**0.98**

a consideration. To find the model achieving the minimal, penalized, out-of-bag quantile error with respect to tree complexity and number of predictors to sample at each node, Bayesian optimization and 10-fold cross-validation were used. The hyperparameters, namely the number of decision trees, minimum sample split, maximum depth, maximum leaf node, minimum samples leaf, and bootstrap sample fraction, were found after 1000 epochs to be 127, 71, 14, 36, 224, and 0.17, respectively.

The different LSTM architectures, the SVR, and the RF were evaluated using three different datasets (the aggregated data of water consumption from 2, 10, and 20 households). The models evaluation against RMSE, MAE, and MAAPE accuracy metrics are summarized in table 1, table 2, and table 3, respectively. The RMSE is noticed to be with higher values than those of the MAE for the corresponding LSTM, SVR, and RF models. However, both table 1 and table 2 could support that LSTM and SVR were comparable in their performance. However, when it comes to all the architecture models using all datasets, MAAPE is the right metric to use for the overall comparison. The values of the MAAPE reflects better performance with aggregated data from more households. While LSTM with three inputs is a better architecture choice that outperforms other models including LSTM with one and two inputs as well as the SVR and the RF models.

VI. CONCLUSION
By leveraging Artificial Intelligence and Machine Learning, governments can forecast the networks' and customers'

needs, automate preventative actions, and tailor their services and products based on quantitative and qualitative measures. Moreover, the IoT business opportunities are limitless as grids and smart meters optimize resources, and remote monitoring solutions increase the efficiency of water network. In addition, analytics is considered an essential component of every successful IoT application. Therefore, IoT technology can provide insights in real time and empower intelligent, data-driven decisions that improve the national welfare.

The smart water meters were installed to cover a neighborhood that can represent water consumption in the pilot study. Real-time streaming is critical for the system solution for further processing and possible prediction necessary for critical management situations. The paradigm used in this research takes advantage of containers and micro-services opposite to virtual machines cloud architecture to increase performance and decrease cost on the national scale IoT system used to collect water consumption data in a suburb in Cairo as a pilot study. The proposed system can be expanded to cover the whole country with sub models for different regions and represent the first and second stage in a smart water management system.

The pilot study offers a testbed for water consumption to be incorporated in a water demand management system that can be scaled up on a national scale with integrated services taking into account security, cost, and scalability.

The main advantage of the two-layer paradigm is to collect aggregated water consumption from different regions to be used to achieve an offline consumption model based on

time and region. That is followed by real-time prediction over time for the water demand with an adaptive machine learning paradigm. Based on the water demand prediction a number of scenarios for both water utilities management and consumer behavior management can be incorporated for the ultimate goal of reduced water consumption. Also, this research suggests a management system that needs to offer quantitative measures for water demand reduction in peak times, better water demand distribution, and lower water consumption. In addition, it needs to measure the effect of planned city development and expansion imposed on water network infrastructure and performance.

Future directions need to tackle accurate simulation for performance metrics related to the IoT cloud in order to optimize the microservices integration for a better performance. Another direction will be to add a monitoring service to continuously measure the LSTM neural network performance and any failing component in the system before it can cause significant performance degradation. Moreover, several recent meta-heuristic techniques can be combined with LSTMs to optimize their hyperparameters to achieve a higher performance.

DATA AVAILABILITY

All the acquired and analyzed water demand data collected during this study are included in this article. The data are accessible through the IEEE DataPort Open Access Data Platform. The generated datasets of this study are available from the corresponding author on request.

REFERENCES

[1] T. Kumura, N. Suzuki, M. Takahashi, S. Tominaga, S. Morioka, and S. Ivan, "Smart water management technology with intelligent sensing and ICT for the integrated water systems," *NEC Tech. J.*, vol. 9, no. 1, pp. 103–106, 2015.

[2] E. Vinothini and N. Suganya, "Automated water distribution and performance monitoring system," *Int. J. Eng. Innov. Technol. (IJEIT)*, vol. 3, no. 8, Feb. 2014, pp. 30–32.

[3] A. Whittle, M. Allen, A. Preis, and M. Iqbal, "Sensor networks for monitoring and control of water distribution systems," in *Proc. Int. Soc. Struct. Health Monit. Intell. 6th Int. Conf. Struct. Health Monit. Intell. Infrastruct.*, Hong Kong, Dec. 2013, pp. 9–11.

[4] J. Shri, A. Jagadeesan, and A. Lavanya, "Theft identification and automated water supply system using embedded technology," *Int. J. Adv. Res. Electr., Electron. Instrum. Eng.*, vol. 2, no. 8, Aug. 2013, pp. 3727–3733.

[5] A. Piasecki, J. Jurasz, and R. Skowron, "Forecasting surface water level fluctuations of lake Serwy (Northeastern Poland) by artificial neural networks and multiple linear regression," *J. Environ. Eng. Landscape Manage.*, vol. 25, no. 4, pp. 379–388, Dec. 2017.

[6] B. P. Esther and K. S. Kumar, "A survey on residential demand side management architecture, approaches, optimization models and methods," *Renew. Sustain. Energy Rev.*, vol. 59, pp. 342–351, Jun. 2016.

[7] M. E. Bildirici, "Economic growth and electricity consumption: MS-VAR and MS-granger causality analysis," *OPEC Energy Rev.*, vol. 37, no. 4, pp. 447–476, Dec. 2013.

[8] C. Kang, Y. Wang, Y. Xue, G. Mu, and R. Liao, "Big data analytics in China's electric power industry: Modern information, communication technologies, and millions of smart meters," *IEEE Power Energy Mag.*, vol. 16, no. 3, pp. 54–65, May 2018.

[9] J. Campillo, E. Dahlquist, F. Wallin, and I. Vassileva, "Is real-time electricity pricing suitable for residential users without demand-side management?" *Energy*, vol. 109, pp. 310–325, Aug. 2016.

[10] A. Grønhøj and J. Thøgersen, "Feedback on household electricity consumption: Learning and social influence processes," *Int. J. Consum. Stud.*, vol. 35, no. 2, pp. 138–145, Mar. 2011.

[11] B. Lin and X. Liu, "Electricity tariff reform and rebound effect of residential electricity consumption in China," *Energy*, vol. 59, pp. 240–247, Sep. 2013.

[12] M. Jaradat, M. Jarrah, A. Bousselham, Y. Jararweh, and M. Al-Ayyoub, "The Internet of energy: Smart sensor networks and big data management for smart grid," *Procedia Comput. Sci.*, vol. 56, pp. 592–597, 2015.

[13] M. A. Jarwar, S. Ali, and I. Chong, "A microservices model to enhance the availability of data for building energy efficiency management services," *Energies*, vol. 12, no. 3, p. 360, Jan. 2019.

[14] A. Krylovskiy, M. Jahn, and E. Patti, "Designing a smart city Internet of Things platform with microservice architecture," in *Proc. 3rd Int. Conf. Future Internet Things Cloud*, Aug. 2015, pp. 25–30.

[15] M. Pau, E. Patti, L. Barbierato, A. Estebsari, E. Pons, F. Ponci, and A. Monti, "A cloud-based smart metering infrastructure for distribution grid services and automation," *Sustain. Energy, Grids Netw.*, vol. 15, pp. 14–25, Sep. 2018, doi: 10.1016/j.segan.2017.08.001.

[16] C. Kamienski, J. Soininen, T. Cinotti, and R. Maia, "SWAMP: An IoT-based smart water management platform for precision irrigation in agriculture," in *Proc. Global Internet Things Summit (GIoTS)*, Jun. 2018, pp. 1–6.

[17] M. A. Jarwar, S. Ali, M. G. Kibria, S. Kumar, and I. Chong, "Exploiting interoperable microservices in Web objects enabled Internet of Things," in *Proc. 9th Int. Conf. Ubiquitous Future Netw. (ICUFN)*, Jul. 2017, pp. 49–54, doi: 10.1109/icufn.2017.7993746.

[18] C. Moreno, R. Aquino, J. Ibarreche, I. Pérez, E. Castellanos, E. Álvarez, R. Rentería, L. Anguiano, A. Edwards, P. Lepper, R. M. Edwards, and B. Clark, "RiverCore: IoT device for river water level monitoring over cellular communications," *Sensors*, vol. 19, no. 1, p. 127, Jan. 2019.

[19] C. Kamienski, J.-P. Soininen, M. Taumberger, R. Dantas, A. Toscano, T. Salmon Cinotti, R. Filev Maia, and A. Torre Neto, "Smart water management platform: IoT-based precision irrigation for agriculture," *Sensors*, vol. 19, no. 2, p. 276, Jan. 2019.

[20] I. Sedliacik and J. Dado, "Unintended consequences of interventions in electricity production and consumption," *New Trends Issues Proc. Humanities Social Sci.*, vol. 3, no. 4, pp. 16–22, Mar. 2017.

[21] R. Kasperowicz, "Electricity consumption and economic growth: Evidence from poland," *J. Int. Stud.*, vol. 7, no. 1, pp. 46–57, May 2014.

[22] I. Ghalehkhondabi, E. Ardjmand, W. A. Young, and G. R. Weckman, "Water demand forecasting: Review of soft computing methods," *Environ. Monitor. Assessment*, vol. 189, no. 7, Jul. 2017.

[23] B. M. Brentan, E. Luvizotto Jr., M. Herrera, J. Izquierdo, and R. Pérez-García, "Hybrid regression model for near real-time urban water demand forecasting," *J. Comput. Appl. Math.*, vol. 309, pp. 532–541, Jan. 2017.

[24] A. U. Muhammad, X. Li, and J. Feng, "Artificial intelligence approaches for urban water demand forecasting: A review," in *Machine Learning and Intelligent Communications—MLICOM*, vol. 294. Cham, Switzerland: Springer, 2019, pp. 595–622. [Online]. Available: https://doi.org/10.1007/978-3-030-32388-2_51

[25] E. I. Papageorgiou, K. Poczeta, and C. Laspidou, "Hybrid model for water demand prediction based on fuzzy cognitive maps and artificial neural networks," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE)*, Jul. 2016, pp. 1523–1530.

[26] S. Shabani, P. Yousefi, and G. Naser, "Support vector machines in urban water demand forecasting using phase space reconstruction," *Procedia Eng.*, vol. 186, pp. 537–543, 2017.

[27] S. Jeddi and S. Sharifian, "A water cycle optimized wavelet neural network algorithm for demand prediction in cloud computing," *Cluster Comput.*, vol. 22, no. 4, pp. 1397–1412, Dec. 2019.

[28] L. Narayanan and S. Sankaranarayanan, "IoT-based water demand forecasting and distribution design for smart city," in *Journal Of Water And Climate Change*. London, U.K.: IWA Publishing, 2019, doi: 10.2166/wcc.2019.019.

[29] V. Ejiofor and O. Oladipo, "Microcontroller based automatic water level control system," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 1, no. 6, pp. 1390–1396, Aug. 2013.

[30] S. Ria, K. Manjit, and L. Hemant, "Design and development of automatic water flow meter," *Int. J. Comput. Sci., Eng. Appl.*, vol. 3, no. 3, pp. 49–59, Jun. 2013.

[31] C. Yuan, S. Liu, and Z. Fang, "Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model," *Energy*, vol. 100, pp. 384–390, Apr. 2016, doi: 10.1016/j.energy.2016.02.001.

[32] B. Baruque, S. Porras, E. Jove, and J. L. Calvo-Rolle, "Geothermal heat exchanger energy prediction based on time series and monitoring sensors optimization," *Energy*, vol. 171, pp. 49–60, Mar. 2019, doi: 10.1016/j.energy.2018.12.207.

[33] H. Bonakdari, H. Moeeni, I. Ebtehaj, M. Zeynoddin, A. Mahoammadian, and B. Gharabaghi, "New insights into soil temperature time series modeling: Linear or nonlinear?" *Theor. Appl. Climatol.*, vol. 135, nos. 3–4, pp. 1157–1177, Feb. 2019, doi: 10.1007/s00704-018-2436-2.

[34] M. R. Kazemi and M. R. Mahmoudi, "A note on limiting distribution of the sample auto-covariance function for the first-order autoregressive (AR(1)) model," *Commun. Statist.-Theory Methods*, vol. 48, no. 15, pp. 3876–3883, Aug. 2019, doi: 10.1080/03610926.2018.1481980.

[35] K. Chan and G. Goracci, "On the ergodicity of first-order threshold autoregressive moving-average processes," *J. Time Ser. Anal.*, vol. 40, no. 2, pp. 256–264, Mar. 2019, doi: 10.1111/jtsa.12440.

[36] M. Becerra, A. Jerez, B. Aballay, H. O. Garcés, and A. Fuentes, "Forecasting emergency admissions due to respiratory diseases in high variability scenarios using time series: A case study in chile," *Sci. Total Environ.*, vol. 706, Mar. 2020, Art. no. 134978, doi: 10.1016/j.scitotenv.2019.134978.

[37] S. M. Idrees, M. A. Alam, and P. Agarwal, "A prediction approach for stock market volatility based on time series data," *IEEE Access*, vol. 7, pp. 17287–17298, 2019, doi: 10.1109/access.2019.2895252.

[38] F. Tronarp and S. Särkkä, "Iterative statistical linear regression for Gaussian smoothing in continuous-time non-linear stochastic dynamic systems," *Signal Process.*, vol. 159, pp. 1–12, Jun. 2019, doi: 10.1016/j.sigpro.2019.01.013.

[39] T. Bessaoudi, F. B. Hmida, and C.-S. Hsieh, "Robust state and fault estimation for non-linear stochastic systems with unknown disturbances: A multi-step delayed solution," *IET Control Theory Appl.*, vol. 13, no. 16, pp. 2556–2570, Nov. 2019, doi: 10.1049/iet-cta.2018.5213.

[40] C. C. Holt, "Forecasting seasonals and trends by exponentially weighted moving averages," *Int. J. Forecasting*, vol. 20, no. 1, pp. 5–10, Jan. 2004.

[41] S. Smyl, "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting," *Int. J. Forecasting*, vol. 36, no. 1, pp. 75–85, Jan. 2020, doi: 10.1016/j.ijforecast.2019.03.017.

[42] R. Laubscher, "Time-series forecasting of coal-fired power plant reheater metal temperatures using encoder-decoder recurrent neural networks," *Energy*, vol. 189, Dec. 2019, Art. no. 116187, doi: 10.1016/j.energy.2019.116187.

[43] A. Borovykh, C. W. Oosterlee, and S. M. Bohté, "Generalization in fully-connected neural networks for time series forecasting," *J. Comput. Sci.*, vol. 36, Sep. 2019, Art. no. 101020, doi: 10.1016/j.jocs.2019.07.007.

[44] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Appl. Energy*, vol. 251, Oct. 2019, Art. no. 113315, doi: 10.1016/j.apenergy.2019.113315.

[45] P. Yu and X. Yan, "Stock price prediction based on deep neural networks," *Neural Comput. Appl.*, vol. 32, no. 6, pp. 1609–1628, Mar. 2020, doi: 10.1007/s00521-019-04212-x.

[46] A. Candelieri, I. Giordani, F. Archetti, K. Barkalov, I. Meyerov, A. Polovinkin, A. Sysoyev, and N. Zolotykh, "Tuning hyperparameters of a SVM-based water demand forecasting system through parallel global optimization," *Comput. Oper. Res.*, vol. 106, pp. 202–209, Jun. 2019.

[47] D. Karmiani, R. Kazi, A. Nambisan, A. Shah, and V. Kamble, "Comparison of predictive algorithms: Backpropagation, SVM, LSTM and Kalman filter for stock market," in *Proc. Amity Int. Conf. Artif. Intell. (AICAI)*, Feb. 2019, pp. 228–234, doi: 10.1109/AICAI.2019.8701258.

[48] S. Harbola and V. Coors, "Comparative analysis of LSTM, RF and SVM architectures for predicting wind nature for smart city planning," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 4, no. 9, pp. 65–70, 2019.

[49] S. K. Lakshminarayanan and J. McCrae, "A comparative study of SVM and LSTM deep learning algorithms for stock market prediction," in *Proc. AICS Conf.*, vol. 2563, 2019, pp. 446–457.

[50] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D: Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306, doi: 10.1016/j.physd.2019.132306.

[51] H. Zhong, J. Wang, H. Jia, Y. Mu, and S. Lv, "Vector field-based support vector regression for building energy consumption prediction," *Appl. Energy*, vol. 242, pp. 403–414, May 2019, doi: 10.1016/j.apenergy.2019.03.078.

[52] W. Yuchi, E. Gombojav, B. Boldbaatar, J. Galsuren, S. Enkhmaa, B. Beejin, G. Naidan, C. Ochir, B. Legtseg, T. Byambaa, P. Barn, S. B. Henderson, C. R. Janes, B. P. Lanphear, L. C. McCandless, T. K. Takaro, S. A. Venners, G. M. Webster, and R. W. Allen, "Evaluation of random forest regression and multiple linear regression for predicting indoor fine particulate matter concentrations in a highly polluted city," *Environ. Pollut.*, vol. 245, pp. 746–753, Feb. 2019.

[53] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *Int. J. Forecasting*, vol. 32, no. 3, pp. 669–679, Jul. 2016.

[54] S. E. Hussein and S. M. Badr, "Healthcare cloud integration using distributed cloud storage and hybrid image compression," *Int. J. Comput. Appl.*, vol. 80, no. 3, pp. 9–15, Oct. 2013.

[55] D. Bernstein, "Containers and cloud: From LXC to docker to kubernetes," *IEEE Cloud Comput.*, vol. 1, no. 3, pp. 81–84, Sep. 2014.

[56] P. D. Diamantoulakis, V. M. Kapinas, and G. K. Karagiannidis, "Big data analytics for dynamic energy management in smart grids," *Big Data Res.*, vol. 2, no. 3, pp. 94–101, Sep. 2015.

[57] R. Desai, "Real time analysis using Hadoop," *Int. J. Eng. Comput. Sci.*, 2016, doi: 10.18535/ijecs/v5i6.52.

[58] I. Y. I. A. N. Hashem, S. Mokhtar, A. Gani, and S. U. Khan, "The rise of 'big data' on cloud computing: Review and open research issues," *Inf. Syst.*, vol. 47, pp. 98–115, Jan. 2015.

[59] A. Alharthi, V. Krotov, and M. Bowman, "Addressing barriers to big data," *Bus. Horizons*, vol. 60, no. 3, pp. 285–292, May 2017, doi: 10.1016/j.bushor.2017.01.002.

[60] D. S. Linthicum, "Moving to autonomous and self-migrating containers for cloud applications," *IEEE Cloud Comput.*, vol. 3, no. 6, pp. 6–9, Nov. 2016.

[61] T. Robles, R. Alcarria, and D. M. de Andrés, "An IoT based reference architecture for smart water management processes," *J. Wireless Mobile Netw., Ubiquitous Comput., Dependable Appl.*, vol. 6, no. 1, pp. 4–23, Mar. 2015.

[62] M. R. Bendre and V. R. Thool, "Analytics, challenges and applications in big data environment: A survey," *J. Manage. Anal.*, vol. 3, no. 3, pp. 206–239, Jul. 2016, doi: 10.1080/23270012.2016.1186578.

[63] M. Dias de Assunção, A. da Silva Veith, and R. Buyya, "Distributed data stream processing and edge computing: A survey on resource elasticity and future directions," *J. Netw. Comput. Appl.*, vol. 103, pp. 1–17, Feb. 2018, doi: 10.1016/j.jnca.2017.12.001.

[64] A. Bansal, R. Jain, and K. Modi, "Big data streaming with spark," in *Big Data Processing Using Spark in Cloud* (Studies in Big Data), vol. 43, M. Mittal V. Balas, L. Goyal, and R. Kumar, Eds. Singapore: Springer, 2019.

[65] S. Han, Y.-H. Qiao, J. Yan, Y.-Q. Liu, L. Li, and Z. Wang, "Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network," *Appl. Energy*, vol. 239, pp. 181–191, Apr. 2019, doi: 10.1016/j.apenergy.2019.01.193.

[66] J. Ma, Z. Li, J. C. P. Cheng, Y. Ding, C. Lin, and Z. Xu, "Air quality prediction at new stations using spatially transferred bi-directional long short-term memory network," *Sci. Total Environ.*, vol. 705, Feb. 2020, Art. no. 135771, doi: 10.1016/j.scitotenv.2019.135771.

[67] S. Mouatadid, J. F. Adamowski, M. K. Tiwari, and J. M. Quilty, "Coupling the maximum overlap discrete wavelet transform and long short-term memory networks for irrigation flow forecasting," *Agricult. Water Manage.*, vol. 219, pp. 72–85, Jun. 2019, doi: 10.1016/j.agwat.2019.03.045.

[68] Y. Yu and F. Liu, "Effective neural network training with a new weighting mechanism-based optimization algorithm," *IEEE Access*, vol. 7, pp. 72403–72410, 2019.

AHMED ABDEL NASSER received the B.S. degree (Hons.) in computer science from the Modern Academy for Computer Science and the M.S. degrees in computer science and information system from the Arab Academy for Science, Technology, and Maritime Transport and the Sadat Academy for Management Sciences, Egypt, respectively. He works as a Governmental Consultant and has introduced the Smart Card for optimizing the distribution of a few strategic resources in Egypt. He was awarded a prize from The American University in Cairo in programming ACM. He was the former Head of research and development of a political party in Egypt. He is also a Microsoft Certified Professional.

MAGDI Z. RASHAD received the Ph.D. degree in computer science from the Faculty of Engineering, Cairo University, Egypt. He is currently a Professor of computer science with Mansoura University, Egypt. He has served as the Head of the Computer Science Department and the Vice Dean of the Faculty of Computers and Information Systems, Mansoura University. He is the author of over 160 articles published in peer-reviewed international scientific journals. He has also served as a reviewer for various international journals, such as the IEEE INTERNET OF THINGS (IoT). His research interests include artificial intelligence, pattern recognition, machine learning, image processing, cloud computing, and the IoT.

SHERIF E. HUSSEIN (Member, IEEE) received the B.S. and M.S. degrees (Hons.) in computer engineering from Mansoura University, Egypt, the Ph.D. degree from the University of Strathclyde, Glasgow, U.K., in 2003, and the M.B.A. degree (Hons.) from Cardiff Metropolitan University, Cardiff, U.K. He was a Postdoctoral Researcher at VNIIM, Saint Petersburg, Russia. He is currently an Associate Professor with the Department of Computer Engineering and Control Systems, Mansoura University; a Computer Technical Consultant for the Ministry of Higher Education, Egypt; and an author of about 40 publications. His research interests include cloud computing, data science, image and signal processing, modeling, and artificial intelligence. He is a member of ICMI and the Institute of Engineering and Technology (IET). He received the ORS Award, the International Scholarship Award, and the University of Strathclyde Award.

● ● ●