

Computer Networks

Real-time and Intelligent Flood Forecasting using UAV-assisted Wireless Sensor Network --Manuscript Draft--

Manuscript Number:	COMNET-D-20-00619
Article Type:	Research Paper
Keywords:	Unmanned Aerial Vehicles; wireless sensor networks; River flow Prediction
Corresponding Author:	Mohammad Hossein Anisi, PhD University of Essex UNITED KINGDOM
First Author:	Shidrokh Goudarzi, PhD
Order of Authors:	Shidrokh Goudarzi, PhD Mohammad Hossein Anisi Domenico Ciunzo, PhD Seyed Ahmad Soleymani, PhD Salwani Abdullah, PhD Muhammad Khurram Khan, PhD
Abstract:	<p>The Wireless Sensor Network (WSN) is a promising technology that could be used to monitor rivers' water levels for early warning flood detection in the 5G context. However, during a flood, sensor nodes may be washed up or become faulty, which seriously affects network connectivity. To address this issue, Unmanned Aerial Vehicles (UAVs) could be integrated with WSN as routers or data mules to provide reliable data collection and flood prediction. In light of this, we propose a fault-tolerant multi-level framework comprised of a WSN and a UAV to monitor river levels. The framework is capable to provide seamless data collection by handling the disconnections caused by the failed nodes during a flood. Besides, an algorithm hybridized with Group Method Data Handling (GMDH) and Particle Swarm Optimization (PSO) is proposed to predict forthcoming floods in an intelligent collaborative environment. The proposed water-level prediction model is trained based on the real dataset obtained from the Selangor River in Malaysia. The performance of the work in comparison with other models has been also evaluated and numerical results based on different metrics such as coefficient of determination (R^2), correlation coefficient (R), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and BIAS are provided.</p>
Suggested Reviewers:	<p>Mahdi Zareei, PhD Tecnologico de Monterrey: Instituto Tecnológico y de Estudios Superiores de Monterrey m.zareei@tec.mx</p> <p>Mojtaba Alizadeh, PhD Lorestan University mojtaba2023@gmail.com</p>
Opposed Reviewers:	

Real-time and Intelligent Flood Forecasting using UAV-assisted Wireless Sensor Network

Sh. Goudarzi^a, M. H. Anisi^{b,*}, D. Ciunzo^c, S. A. Soleymani^d, S. Abdullah^a and M. K. Khan^e

^aCentre for Artificial Intelligent (CAIT), Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia.

^bSchool of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, United Kingdom.

^cDepartment of Electrical Engineering and Information Technologies, University of Naples "Federico II", 80125 Naples, Italy.

^dSchool of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, Johor, Malaysia.

^eCenter of Excellence in Information Assurance (CoEIA), King Saud University, Saudi Arabia.

ARTICLE INFO

Keywords:

Unmanned Aerial Vehicles
Wireless Sensor Networks
Group Method Data Handling
Particle Swarm Optimization
River flow
Prediction

ABSTRACT

The Wireless Sensor Network (WSN) is a promising technology that could be used to monitor rivers' water levels for early warning flood detection in the 5G context. However, during a flood, sensor nodes may be washed up or become faulty, which seriously affects network connectivity. To address this issue, Unmanned Aerial Vehicles (UAVs) could be integrated with WSN as routers or data mules to provide reliable data collection and flood prediction. In light of this, we propose a fault-tolerant multi-level framework comprised of a WSN and a UAV to monitor river levels. The framework is capable to provide seamless data collection by handling the disconnections caused by the failed nodes during a flood. Besides, an algorithm hybridized with Group Method Data Handling (GMDH) and Particle Swarm Optimization (PSO) is proposed to predict forthcoming floods in an intelligent collaborative environment. The proposed water-level prediction model is trained based on the real dataset obtained from the Selangor River in Malaysia. The performance of the work in comparison with other models has been also evaluated and numerical results based on different metrics such as coefficient of determination (R^2), correlation coefficient (R), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and BIAS are provided.

1. Introduction

Effective river flow prediction is required to reduce the damage caused by potential surges. Various techniques have been proposed such as surge forecasting, river training (i.e. taking structural measures to reduce the flood flow velocity), real-time alerts, stormwater predictions, and emergency management [16].

The 5G network provides high peak data rates with low latency and massive network capacity that would be very useful in flood management. In this regard, a great deal of attention has been paid to the use of Wireless Sensor Network (WSN), one of the enabling technologies in 5G networks, for river monitoring and flow predictions. However, there are some key shortcomings in the standalone use of WSN[16]. The main concern is that some nodes could be destroyed or become damaged due to the success of the flood.

Hence, given the multi-hop nature of WSNs, such failure could put an end to the whole routing process if the failed nodes are network bottlenecks. Alternatively, such failures could also result in poor Quality of Service (QoS) and/or increased energy consumption due to increased re-transmission of unsuccessful packets.

Therefore, due to the above issues as well as a limited coverage and computation capability of WSNs, standalone WSNs are progressively merged with in-

terconnected dynamic nodes such as Unmanned Aerial Vehicles (UAVs) that are enabled by the Internet of things (IoT) technology. This development could be also empowered by forwarding collected data to the cloud through UAVs, and managing it using Software Defined Network (SDN) to quickly and reliably detect and locate unexpected events. In order to minimize manufacturing costs, wireless sensor nodes are generally used in monitoring areas and are self-organized into WSNs to collect environmental data. However, the awareness of location information in the cloud is important for real-time event detection. Hence, the UAVs can serve as mobile anchors to assist localization, and as mobile relays to transfer data from sensor nodes to the cloud. Several works have recently been proposed in the literature to address data collection from sensor nodes using UAV [5, 34].

The advantages of integrating UAVs with WSNs for flood prediction are highlighted as follows: (i) Each node can send its data to the associated UAV by a single-hop transmission to the base station which reduces the energy consumption of the WSN; (ii) the accuracy and timeliness of the river flow predictions could be improved by the use of UAVs to provide real-time location information; (iii) the scalability of the network, limited by the low energy of the nodes, would be enhanced and the nodes could be distributed over long distances to cover the river.

UAVs were used as relays [5, 11] to improve communication in WSNs. UAVs were leveraged as mobile

*Corresponding author

✉ m.anisi@essex.ac.uk (M.H. Anisi)

sinks within ubiquitous sensor networks to improve the connectivity of the ground sensor nodes. The deployment of wireless sensor nodes was investigated in post-disaster environments [34] using a quadrotor equipped with Inertial Navigation System (INS) and Global Positioning System (GPS) sensors.

There are others researches focused on UAV-assisted WSN for developing new sensing applications [35, 6] and for modeling effective mobility patterns for data collection [8, 13, 33]. The goal of another study [8] was to deploy micro-UAVs at various locations in a disaster area to rapidly generate communication networks for search and rescue operations. In this work, UAVs will fly close to the ground to capture high-resolution images of disaster sites. Interface protocols have also been established to easily manage large groups of micro-UAVs. Although various works such as [35, 6, 8, 13, 33] have discussed the integration of WSNs and UAVs, to the best of our knowledge, no study considers UAV as a gateway or data mule for WSN to optimize flood prediction [3]. On the other hand, previous works have ignored the impact of the dynamic topology of UAVs on integration with WSNs, which is difficult to control during the deployment stage. Therefore, this paper aims to design a real-time UAV-assisted WSN model to mitigate the number of packets lost to destroyed/faulty nodes during the flood and to provide accurate flood predictions. In the proposed air-ground network model, the WSN monitors the river and reports the water level to the central processing unit (i.e. the base station). When a node fails to transmit its data via multi-hop communication, a UAV is called to bridge the communication and send the data to the base station. Wireless sensor nodes are deployed at the edge of the urban river to monitor water flow behavior during times of flood or prolonged rainfall, and UAVs are adapted for wireless data collection from the sensors. In order to optimally use the UAV and efficiently control its topology, the disaster area is divided into several sub-regions by the cloud and the center of each sub-region is considered as the hovering point of the UAV. Then, the sensor nodes are grouped into these sub-regions according to the received signal strength (RSS) of the detected beacons. In each sub-region, packets are forwarded based on a random walk process to collect the data of sensor nodes. If the packet returns to the starting node with an expected time of t , it can be determined that there is a failure and that the UAV is functioning as a relay and forwards the next packet to the cloud.

The main contributions of this paper are threefold:

- We propose a framework for real-time data collection based on a multi-hop WSN and a UAV in which the UAV as a router relays the data packets of the sensor nodes when they fail to find any available node as the next hop.

- We integrate cloud and SDN to manage network connectivity across the data center and simplify the dynamic programming process. we divide the disaster area into several sub-regions, and the random walk model is used by the UAV to collect data of each sub-region, including nodes IDs and neighbor tables in sub-regions. Then, the collected data will be forwarded to the cloud empowered by SDN for flood prediction.
- We propose a novel prediction model for predicting floods. Once river flow data is transmitted to the central prediction unit, integrated Group Method Data Handling (GMDH) with Particle Swarm Optimization (PSO) is used to forecast floods.

The rest of the paper is structured as follows. Section 2 discusses the related works on the topic. In Section 3 we provide a statement for the considered problem, whereas in Section 4 we outline a multi-level network model. Section 5 presents the prediction model, whereas Section 6 explains the results. Section 7 illustrates the discussion. Finally, conclusions and future directions of the paper are given in Section 8.

2. RELATED WORK

Although several works integrated UAVs and WSNs, it should be stressed that none of them make use of UAVs to enable higher-resilience WSNs during flood prediction or make evaluations based on real data.

Concerning quick learning for UAV navigation tasks, some previous works typically emphasize accurate methods for components such as perception and relative pose estimation [3] or trajectory optimization and control [28]. UAVs can support various wireless communication protocols. For example, UAVs can communicate with WSNs in a self-organized way by ZigBee modules [30], and have the ability to serve as relays to forward data to the cloud [21]. In [7, 25, 37], researchers explain that there are multiple ways to manage UAVs to form a network to cover all sensors, whether the sensors are directly connected to the base station, or whether the relays can be used to increase the range of communication. But there is no explicit analysis when the sensors are disconnected from the base station for some time.

Recently, machine learning models have been used to analyze and assess linear scattering coefficients as strong alternatives to traditional regression models in natural rivers. These models include Artificial Neural Networks (ANNs), Genetic Programming (GP), Adaptive NeuroFuzzy Inference Systems (ANFIS), and Support Vector Machines (SVM) to evaluate the longitudinal dispersion constant [2]. Among these techniques is the GMDH that is a self-organizing method with non-linear network models. It uses a combination of a

quadratic polynomial in a multi-layer procedure [23]. Many recent algorithms such as GMDH networks have been able to perform accurate predictions, especially the river water stage prediction. The GMDH networks were a quick learning machine planned by Ivakhnenko in the 1960s. This method is a heuristic self-organizing algorithm that develops inductive self-organizing strategies to solve a lot of complicated and sensible issues. The GMDH networks have been applied in engineering sciences, fault detection, ranking, and selection modeling, and chemical action [15, 22].

The GMDH networks provide effective and efficient technical performance in various engineering fields [22], but their training suffers from certain disadvantages such as local minimum and slow convergence. Therefore, selecting an applicable training model is one of the paramount steps within the development of a data-driven model. This study adopted the PSO technique [15] to train GMDH networks for river prediction models. The developed model is a hybrid method for one-day-ahead prediction of river water where a non-linear regression approach is adopted due to the complex process of river flow prediction in natural rivers. It is evaluated in simulated networks in Malaysia, where some other neural network-based models, including DE, GA, and ANN, are also tested for comparison. Recently, GMDH networks were used to forecast scour depth around hydraulic structures [22]. It has outperformed standard non-linear regression methods in prediction processes. They build analytical functions inside feedforward networks within quadratic polynomials. The weight coefficients of these quadratic polynomials are the result of regression technique exploitation [14]. The effective forecasting technique for river water stages would minimize losses from flooding exploitation due to the prediction of what people close to the river need [18, 26]. Some limitations of the GMDH technique include slow convergence in training, imprecision in parameter assessment, overfitting, the partition of information, and low accuracy. Therefore, a hybrid version of GMDH was planned to considerably boost its performance. Robinson and colleagues [27] presented a Multi-Objective GMDH (MOGMDH) algorithm within a consistency criterion that used three different selectors within the choice procedure. This significantly improved the performance of the GMDH algorithmic program. Hiassat et al. [9] proposed the Genetic Programming-GMDH algorithmic program, which applies genetic programming to discover the simplest functions that can map inputs to outputs for every layer of the GMDH algorithmic program, and they presented a model that achieves better results than the standard GMDH algorithm in time series predictions using financial and weather information. Genetic Algorithms (GAs) have recently attracted attention in feedforward self-organizing networks. In this study, neuron connections are controlled

to adjacent layers [24].

The lack of effective training algorithms for training multi-layer perceptron is an important issue in GMDH networks. In recent years, some data-driven improvements to training algorithms such as Back Propagation (BP) [10], Levenberg-Marquardt procedure [17], and scaled conjugate gradient procedure [20] have been used to perform training tasks. Usually, gradient-based methods have some drawbacks, such as slow speed convergence during training and getting trapped in local minimums. The PSO algorithm can perform training tasks for ANNs that may be applied in hydrologic environments. The PSO method optimizes complicated numerical functions and was initially developed to model social behavior [15]. Additionally, it is recognized as an evolutionary method under the field of computational intelligence [4, 31]. So far, several prediction approaches have been proposed. However, none of these approaches has taken into consideration the effect of data collection by UAVs for river flow prediction along with the PSO algorithm for training the GMDH model. We made a comparison to prove the novelty of the proposed model. The comparison with the state of the art is provided in I. The table presents the proposed models that used UAVs for data collection from the sensor nodes using UAVs.

3. Problem Statement

3.1. Communication gap caused by faulty nodes

In WSN-based flood monitoring approaches, nodes might be destroyed or get faulty during a flood that seriously affects the network connectivity. To overcome this issue, UAVs could be deployed to act as routers or data mules to fill the network communication gap caused by the inactive nodes. UAVs relay packets from the isolated nodes and enable continuous flood monitoring. In our UAV-assisted data collection mechanism, the WSN is modeled as an undirected graph as follows:

Let $G = (V; E)$ be a simple connected and undirected graph, where V and E represent the vertex and the edge set, respectively. In the WSN, the sensor nodes n and the wireless communication links m are modeled as vertices and edges, respectively. The set of vertices is represented as $V = \{v_1, v_2, v_3, \dots, v_n\}$ and set of edges $E = \{e_1, e_2, e_3, \dots, e_m\}$ is expressed as the wireless communication links. The $degree(v_i)$ represents the degree of a vertex and shows the number of valid neighbors of a sensor node. The value of $degree(v_i)$ may change during the flood prediction process, due to the destroyed nodes. Also, the Valid neighbors are defined as nodes with valid wireless communication capability. Furthermore, we assume that each node possesses the information of its neighbors in a table that includes the connectivity status, neighbor node

Table 1
Comparison with the state of the art.

Model	Faulty nodes mgt	Data collection	Topology mgt	Data prediction
De Freitas <i>et al.</i> [5]	○	●	○	○
Tuna <i>et al.</i> [34]	○	●	○	○
Valente <i>et al.</i> [35]	○	●	○	○
Erdelj <i>et Natalizio</i> [6]	○	●	○	○
Hauert <i>et al.</i> [8]	○	●	○	○
Jawhar <i>et al.</i> [13]	○	●	○	○
Sun <i>et al.</i> [33]	○	●	○	○
Skoczylas [31]	○	●	○	○
Shaw <i>et Mohseni</i> [30]	○	●	●	○
Mozaffari <i>et al.</i> [21]	○	●	●	○
Goddemeier <i>et al.</i> [7]	○	●	●	○
Pinto <i>et al.</i> [25]	○	●	●	○
Yu <i>et al.</i> [37]	○	●	●	○
Najafzadeh <i>et Sattar</i> [23]	○	○	○	●
Najafzadeh <i>et al.</i> [22]	○	○	○	●
Kalantary <i>et al.</i> [14]	○	○	○	●
Proposed model	●	●	●	●

IDs and the radio signal strength indicator (RSSI) between the nodes. Matrix C shows the connectivity status between the nodes as follows:

$$C_{v_i, v_j} = \begin{cases} 1 & \text{if } e(v_i, v_j) \text{ exists} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

According to the matrix, if $e(v_i, v_j)$ exists, the vertices v_i and v_j can communicate. Otherwise, If there is no possibility for wireless communication between v_i and v_j , the UAV nodes is called to collect data from the node.

To solve this problem, the sensor nodes are grouped into N sub-regions by the cloud using a number of beacons with known locations. Each sensor node records all the detected beacons, and selects the certain sub-region based on the highest RSS of the beacons' signal. Then, the random walk process is applied for propagating data on a connected graph with n vertices and m edges at the sub-regions. Given K sensor nodes in a sub-region, the distance matrix is defined as D and the location of the first UAV hovering point is expressed as ml_f and while the UAV moves to the m th location, the distance matrix is defined as ml_m . With m recorded locations, the collected data of K sensor nodes can be predicted through the proposed flood prediction model in Section V.

3.2. Flood prediction

More formally, we can denote an environment and the sensors nodes within it by a tuple $\varepsilon = (S, G, P, T)$, where $S = \{S_i \mid i = 1, \dots, K\}$ is the set of K sensor nodes, $G = (V; E)$ where V and E represent the vertex and the edge set, respectively and encodes the layout of the physical environment between locations V . To each location $v \in V$ we assign two spatial coordinates (x_v, y_v) . So, we can denote the distance between two locations

V, v' as $|v - v'|$, P is a spatial phenomena that is monitored by the sensors in S , and $T = \{t_1, t_2, \dots\}$ models time as a sequence of discrete timesteps of unknown length. To measure the performance of the sensors, we utilise the root mean squared error (RMSE) of their predictions. In order to do this we denote the predictions that the prediction unit makes based on collected data from sensors at time t by $P_t = \{p_t^v \mid v \in V\}$. Suppose the actual measurements made at those locations are $A_t = \{a_t^v \mid v \in V\}$, then the objective of the prediction model is to coordinate all nodes movements so as to minimise for all timesteps $t \in T$:

$$RMSE(P_t, A_t) = \sqrt{\frac{\sum_{v \in V} (a_t^v - p_t^v)^2}{V}} \quad (2)$$

Formally, the RMSE can be considered of as a quality measure of the situational awareness the sensors nodes achieve, since it measures the difference between the measurements of the sensors, and the actual values of the phenomena in the environment. We chose this measure because it has been used in related work to ascertain the accuracy of sensor predictions.

4. Proposed multi-level architecture

Here, the details of the suggested architecture are explained. The proposed network model is an adaptable and scalable model with multiple applications. The model was designed with three layers. In the cloud-SDN layer, a centralized SDN controller was defined as the main control entity and the central processing unit for action predictions. The SDN controller linked the ground WSN and UAV. The second layer included UAVs operated on-demand, with progressive sensors and communication. The third layer covered ground

WSNs with scalar sensors such as rainfall sensors and water level sensors. Fig. 1 shows the network model and the key components of the cloud-SDN, UAVs, and sensors. The main components of the suggested framework are presented in detail.

4.1. Cloud-SDN

A communication network is an important component of the flood control system. With the integration of advanced technologies and applications for achieving smarter controlling of rivers, a vast amount of data from different locations will be generated for analysis, update, control, and real-time flood predicting methods. Thus, the management of these networks is the main challenge due to the scale. Moreover, the equipment may not be able to exchange information due to heterogeneous devices and applications. Hence, it is a vital issue to find the best communication infrastructure to control and manage all devices throughout the total system, considering the real-time constraint. In this model, cloud computing-based SDN is a good solution to the aforementioned problems, thanks to the following advantages. Cloud technology offers high computing capacity to flood prediction utilities. Moreover, flexible per-flow routing is possible using SDN and the flow can be defined across multiple network layers. Also, a logically centralized controller can improve the service efficacy of flood prediction. Also, due to the programmability of SDN, the network is made more active and an appropriate radio access interface can be selected for data delivery. Last but not least, quick-response cloud service is essential for river monitoring on the basis of the real-time road conditions.

4.2. UAVs

Generally, UAVs as aerial agents refer to active objects with behavior, state, and location, which are autonomous and mobile. They can move freely with state and code in execution without suspending services, provide better asynchronous interaction, reduce communication cost, and enhance flexibility. For greater geographical distances where ground nodes are infeasible, UAV-based systems can be integrated. UAVs collected data from the sensing targets and transmitted the collected data to the ground control station or terrestrial user equipment. Various reasons have been provided for the use of UAVs in the proposed network model. The main reason is that the employment of UAVs will lead to lower traffic over the wireless channel. Also, in comparison to traditional network forwarding, the reliability of the path will be significantly improved as the numbers of hops will be reduced where packets are diffused in the network over multiple hops. The direct communication, where the UAV collects data from each sensor node, is used for data acquisition.

4.3. WSNs

The ground control station was configured for data analysis and to control management operations. Ground data was distributed between ground control stations and UAV communication nodes. Sensor nodes are flexible network elements that deliver (real-time) collected water level data to the central processing unit. However, considering the extremely large area and numerous working scenarios involved in flood control, it is impossible to manage floods without using UAVs as detection tools. These were generally controlled from the ground control station.

5. Prediction Model

In this section, the methodology for flood prediction using UAVs along with a PSO algorithm for training the GMDH model is described.

5.1. GMDH Approach

The GMDH method has various stages. The first stage involves partitioning data into training data and testing data. This division is based on consecutive heuristic selection points in the data set. Also, this partitioning is obtained by calculating the variance of data from the mean value. Points should have high variance and be employed in the testing data set for model checking, outside of the data in the training set. In the second step, input data for the input matrix was chosen in pairs and, between each pair, a quadratic polynomial was taken with the corresponding output. The least-square fitting [36] is used to set the polynomial coefficients. To verify polynomial's suitability, the outputs of the polynomials were evaluated using data points in the testing data. Mostly, Mean Squared Error (MSE) was used to select suitable polynomials for the next layer. Finally, this process was repeated until the smallest MSE was higher than the previous layer. A suitable data model was obtained by tracing back the polynomial path with the smallest MSE in each layer. The GMDH method relies on self-organizing methods for the assessment and estimation of recording machine models with uncertain variable relationships. GMDH networks use a regression based on the Ivakhnenko polynomial [12] as follows:

$$y = a_0 + \sum_{i=1}^M a_i x_i + \sum_{i=1}^M \sum_{j=1}^M a_{ij} x_i x_j + \sum_{i=1}^M \sum_{j=1}^M \sum_{k=1}^M a_{ijk} x_i x_j x_k \quad (3)$$

where M is the number of input variables, $(x_1, x_2, x_3, \dots, x_M)$ are the input variables; and $(a_0, a_i, a_{ij}, a_{ijk}, \dots)$ are the coefficients. Generally, Eq.(1) is the quadratic form of the two variables shown in Eq.(2):

$$y = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (4)$$

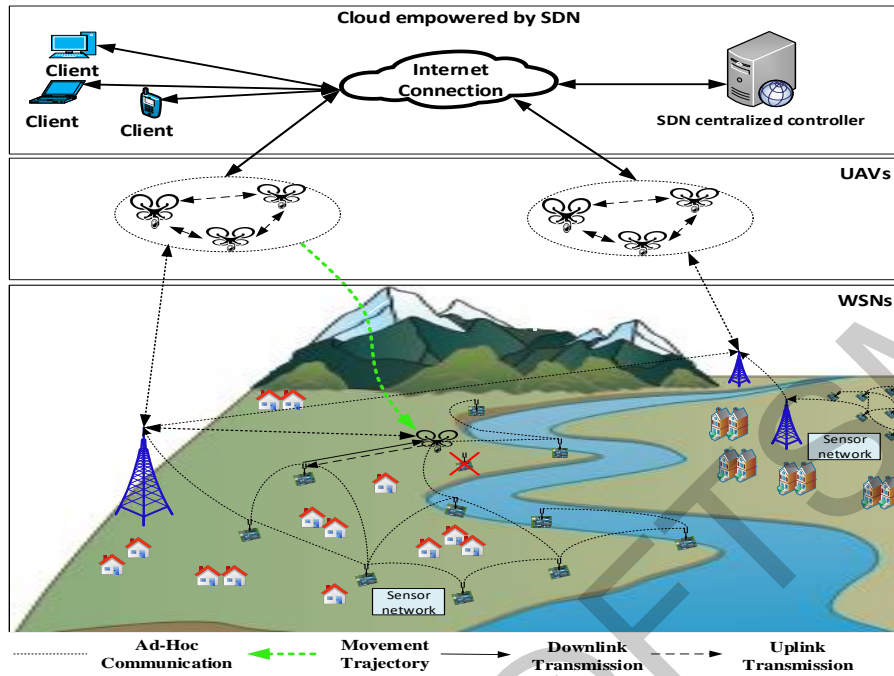


Figure 1: Proposed Multi-level Architecture.

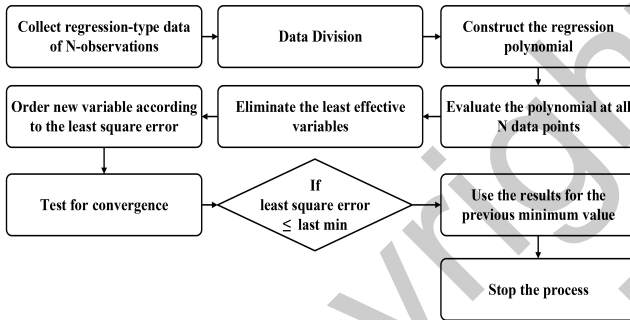


Figure 2: GMDH structure.

The configuration of the GMDH model employed in this study is presented in Fig. 2.

5.2. PSO Algorithm

The goal of the PSO scheme designed by Kennedy and Eberhart was to rehabilitate social interactions between agents in order to solve non-linear continuous optimization issues [15]. PSO is an advanced search technique for non-linear function optimization that self-resembles the motion of agents in a bird flock. Particles refer to the available solutions to a problem. Particles have efficient cooperation and competition in their generational evolution. In this context, all particles control fly concerning the flying experiences of themselves and their neighbors. PSO is capable of discovering the global optimum with a high convergence rate. The PSO trainer used in this paper originated from the idea that PSO can boost the convergence behavior of

particles by accurately selecting weighting parameters instead of randomly choosing them within a typical interval. Thus, PSO was adopted for training multi-layer perceptrons where matrix learning problems are restricted.

5.3. The proposed hybrid GMDH-PSO algorithm

The usual version of GMDH has some shortcomings that need to be addressed: (i) how to train two-layered high-precision networks; (ii) how to specify the best number of input variables; (iii) how to choose a polynomial order to form a vector solution in every node; and (iv) how to select input variables. This study focused on these issues using the proposed GMDH-PSO model.

5.3.1. Using PSO in the training process

It is apparent from previous sections that the GMDH method has some limitations in the training process. Hybridization of the PSO model with standard GMDH can solve this problem. In this application, a three-layered perceptron was chosen. PSO was used to train the GMDH network. Initially, the fitness function of every particle was determined. The error function at current particle positions was evaluated to determine the fitness value of every swarm particle. Also, fitness values were determined on the basis of the particle position vectors corresponding to the network weight matrix. In this hybrid technique, all training data was set to the GMDH network. Then, the weights of each data set were updated such that the size of the train-

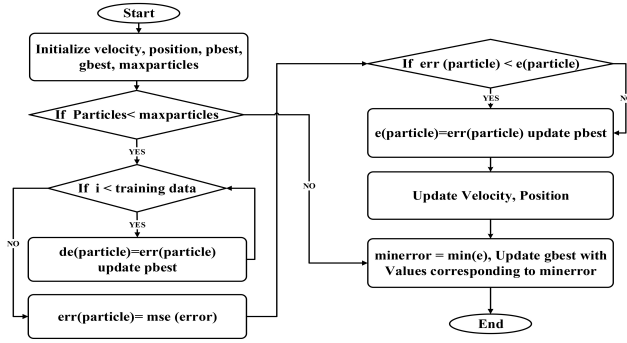


Figure 3: Architecture of the proposed SDN-EC framework.

ing set was equal to the number of updated weights. The vector of each particle was selected to show their error vector. This vector stored the minimum errors encountered by each particle due to their input patterns. This value shows the Mean Square Error (MSE) during training. The flowchart procedure for training a GMDH network using PSO is given in Fig. 3.

Weight training was used for the following reason: W^1 shows the weight matrix between the input layer and the hidden layer; W^2 denotes the weight connection matrix between the hidden layer and the output layer. The i -th particle of a PSO in multi-layer perceptron training is denoted as follows:

$$W_i = (W_i^1, W_i^2) \quad (5)$$

For every particle, the former best fitness value was defined to present the position of the particle as follows:

$$P_i = (P_i^1, P_i^2) \quad (6)$$

The best particle index among all the particles in the population is shown by b and the best matrix is presented by:

$$P_b = (P_b^1, P_b^2) \quad (7)$$

The particle velocity i is denoted by

$$V_i = (V_i^1, V_i^2) \quad (8)$$

The formula for particle manipulation in each iteration is presented as follows:

$$V_i^{[j]}(m, n) = V_i^{[j]}(m, n) + \left(r\alpha [P_i^{[j]}(m, n) - W_i^{[j]}(m, n)] + s\beta [P_b^{[j]}(m, n) - W_b^{[j]}(m, n)] \right) / t \quad (9)$$

where m and n denote matrix rows and columns, respectively; r and s are positive constants; t is the time step between observations and is commonly taken as unity; α and β are random numbers from 0 to 1; V and W refer to the new values.

$$W_i^{[j]} = W_i^{[j]} + V_i^{[j]}t \quad (10)$$

where $j = 1, 2$; $m = 1, \dots, M_j$; $n = 1, \dots, N_j$; M_j and N_j are the rows and column sizes of the matrices W , P , and V . Eq.7. was utilized to compute new particle velocities based on its previous velocity and the distance of its current position from its best experience and the best experience of its group. Then, a new position according to the new velocity is determined using Eq.7. Also, Eq.8 was used to determine the fitness of the i -th particle in terms of an output mean squared error of the neural network as follows:

$$f(W_i) = \frac{1}{S} \sum_{k=1}^S \left[\sum_{l=1}^O \{t_{kl} - p_{kl}(W_i)\}^2 \right] \quad (11)$$

In the above equation, the fitness value is f ; the target output is t_{kl} ; the number of output neurons is O ; the predicted output according to W_i is p_{kl} ; the number of training set is S .

5.3.2. Region and Data Description

Hydrographs offered daily water level records from Selangor River via <http://infobanjir.water.gov.my>. The Selangor River is the main river in Selangor, Malaysia. It runs from Kuala Kubu Bharu in the east and flows into the Straits of Malacca at Kuala Selangor in the west. The data presented through this website were suitable indicators of potential flooding or landslides. This study utilized the data from this website with discretion. This study extracted online hydrograph data for three stations—Selangor, Selayang, and Bernam—on the Selangor River. According to the existing hydrographs on 27 December 2018, the average water level measured by Station1 was about 48.72. These values were about 37, 44, and 21 for Station1, Station2, and Station3, respectively.

5.3.3. Data normalization

The water levels data set at the Selangor river were predicted over one and two days based on measured daily levels. Data normalization was done to avoid false patterns that can be created by inconsistencies. The dataset had some variations because the collection devices were located in different time zones and geographical locations. Data were normalized by dividing the total daily water levels by the number of hours within that day. The normalized data series were computed as:

$$D_t' = \frac{D_t}{H_t} \quad (12)$$

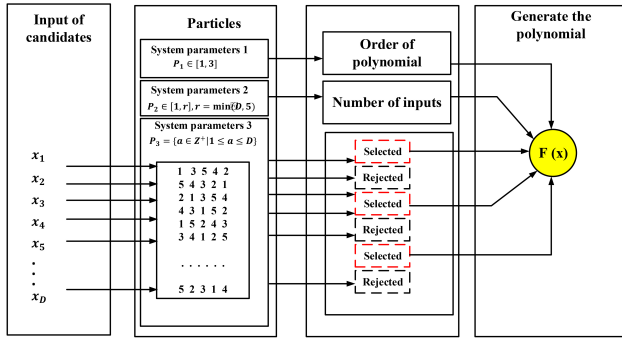


Figure 4: Construction of polynomials by PSO.

where D_t is total daily water level, D_t^i is the normalized data, and H_t is the number of hours in the i -th day.

5.3.4. Construction of polynomials by PSO

Particles were used as search agents in the PSO. The grouping of input variables from the previous layer was determined on the basis of the position of each particle. This data was then moved to the next layer. Every particle contained three main parameters: P_1 , P_2 , and P_3 . P_1 was defined as a polynomial order. In this context, the polynomial order was created from the previous layers and generated randomly. For simplicity, this study took 2 in each layer.

However, this value can be either 2 or 3. The number of input variables was generated randomly and was obtained from the previous layer. We defined D and $r = 2$ as the width of the input dataset and the default lower bound, respectively. The number of input variables was $P_2 \in [1, r]$, where $r = \min(D, 5)$.

The position of every particle representing tall candidates in the current layer of the network was $P_3 = \{a \in Z^+ | 1 \leq a \leq D\}$, which is a sequence of integers. These three parameters were used to arrange nodes to move to the next layer.

P_1 , P_2 , and P_3 were used to determine the polynomial order, the number of node groupings, and the whole sequence, respectively. Fig. 4 shows the procedure for the three defined parameters used to form the polynomial. In our hybrid model, three parameters were used to create the polynomial and all particles consisted of separate parameter sets. Generated polynomials were employed as an objective function for PSO.

5.3.5. Framework of the GMDH-PSO

The GMDH-PSO framework is comprised of six main steps:

First, the input variables of the system were determined. The primary population of PSO structures and corresponding learning parameters c_1 and c_2 were created. The input variables of the model were defined as x_i ($i = 1, 2, 3, \dots, n$) and were related to output vari-

able y . Then, the normalization of input data was completed. In both experiments, the original data needed to be normalized to generate equivalent water level data. In the second phase, training data for PSO and testing data was formed. The input-output data set $(x_i, y_i) = (x_{1i}, x_{2i}, \dots, x_{ni}, y_i)$, $i = 1, 2, 3, \dots, n$ was divided into a training and testing dataset. The size of the training and testing dataset were represented by n_{tr} and n_{te} , respectively, where $n = n_{tr} + n_{te}$. The training dataset was employed to construct the GMDH-PSO model. The testing dataset was utilized to evaluate model quality.

In the third phase, the primary information that would be used to construct the GMDH-PSO structure was determined. Note that the previously mentioned process determined the model's structural optimization by PSO variation operators. In this context, we defined the maximum number of generations as the termination method to balance model accuracy and complexity. The maximum number of input variables was used for every node in each layer. Moreover, the value of the weighting factor was determined for the aggregate objective function.

In the fourth phase, the Polynomial Neuron (PN) structure was determined using the PSO algorithm. The least-square technique was used for parameter optimization through multiple-regression analysis. This technique was used to provide the formula to compute coefficients. The objective function, which was the main instrument used to control evolutionary searches in the solution space, was defined based on the following generated polynomial:

$$f(x_1, x_2) = a_1 + a_2x_1 + a_3x_2 + a_4x_1x_2 + a_5x_1^2 + a_6x_2^2 \quad (13)$$

where a_1, a_2, \dots, a_6 are the constants assessed using the training dataset. The formula used to compute coefficients was obtained using the least-square method in the following formula: $a = (x^t x)^{-1} x^t y$.

In the fifth phase, if the current structure was the best, the model proceeded to phase 6, otherwise it returned to phase 3. This procedure was repeated for all nodes at all layers (from the input layer to the output layer).

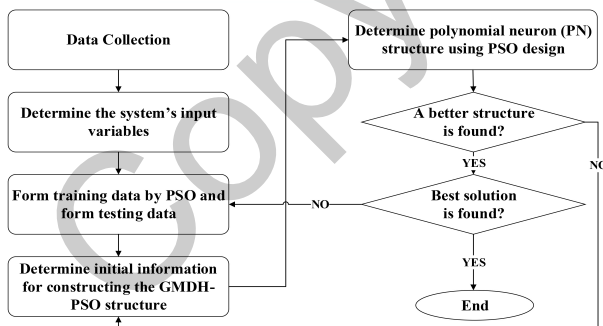
In the sixth phase, if an acceptable solution was obtained, then the algorithm was stopped, otherwise the model returned to step 2. The GMDH-PSO algorithm was carried out by consecutively repeating steps 2–6. When the termination condition was met, one solution vector with the optimum performance was selected in the last population generation as a solution vector and all remaining solution vectors were rejected. The pseudocode of GMDH-PSO is represented in Algorithm 1. Also, Fig. 5 shows the GMDH-PSO model.

Algorithm 1: Pseudo-code for GMDH-PSO

```

1 Input: Maximum number of input variables,
  number of particles, original data,
  size of training dataset, the maximum number of
  generations, value of the weighting factor
  and number of iterations;
  Output: Output variable, predicted data
2 Create the primary population of PSO
3 Create learning parameters C1, C2
4 Normalization (input-data)
5 Training-dataset =
   $\{(x_i, y_i) = (x_{1i}, x_{2i}, \dots, x_{ni}, y_i)\}, i =$ 
   $1, 2, \dots, n_{tr}$ 
6 Testing-dataset =
   $\{(x_j, y_j) = (x_{1j}, x_{2j}, \dots, x_{nj}, y_j)\}, j =$ 
   $1, 2, \dots, n_{te}$ 
7  $n_{tr}$  = size of training dataset
8  $n_{te}$  = size of testing dataset
9  $n = n_{tr} + n_{te}$ 
10 Define Particles =
   $\{par_1(P_{11}, P_{12}, P_{13}), \dots, par_m(P_{m1}, P_{m2}, P_{m3})\}$ 
11 Generate Polynomial by (PSO(Particles))
12 if the current structure is the best then
13   | go to next step
14 else
15   | go to Line 11
16 if solution is acceptable then
17   | go to End
18 else
19   | go to Line 6
20 End

```

**Figure 5:** Construction of polynomials by PSO.**6. Results**

The GMDH-PSO network was compared with earlier models such as DE [19], GA [29], and ANN [1] and the results are presented in this section. In these comparisons, the main indicators for prediction errors were calculated for model evaluation [32]. Table II presents 48 river level values over 24 hours from these

Table 2

Sample of observed data of water level

No	Time	Selangor (Station 1)	Selayang (Station 2)	Bernam (Station 3)
1	00:00	31.67	44.40	21.58
2	00:30	31.65	44.76	21.85
3	01:00	31.66	44.85	22.05
4	01:30	31.66	45.10	22.15
5	02:00	31.75	45.15	22.10
6	02:30	31.75	45.40	22.09
7	03:00	31.74	45.20	22.10
8	03:30	31.55	45.70	22.10
9	04:00	33.65	45.90	22.10
10	04:30	31.70	46.15	22.10
11	05:00	31.70	46.18	22.10
12	05:30	33.65	46.23	22.10
13	06:00	31.64	46.20	22.08
14	06:30	31.62	46.18	22.10
15	07:00	31.58	46.24	22.12
16	07:30	31.57	46.10	22.16
17	08:00	31.55	46.12	22.17
18	08:30	31.58	46.14	22.17
19	09:00	31.62	46.19	22.19
20	09:30	31.68	46.25	22.20
21	10:00	31.65	46.30	22.25
22	10:30	31.70	46.25	22.20
23	11:00	31.90	46.15	22.27
24	11:30	32.35	46.02	22.33
25	12:00	32.40	45.85	22.38
26	12:30	32.36	45.50	22.40
27	13:00	32.05	45.63	22.40
28	13:30	32.45	45.40	22.44
29	14:00	32.85	45.20	22.50
30	14:30	33.15	45.05	22.50
31	15:00	33.17	45.20	22.49
32	15:30	33.15	45.20	22.45
33	16:00	33.15	45.22	22.45
34	16:30	33.05	45.22	22.39
35	17:00	32.80	45.25	22.35
36	17:30	32.75	45.22	22.33
37	18:00	32.53	45.20	22.30
38	18:30	32.45	45.20	22.30
39	19:00	32.50	45.25	22.25
40	19:30	32.20	45.20	22.20
41	20:00	32.10	45.18	22.18
42	20:30	31.70	45.10	22.10
43	21:00	31.50	44.90	22.10
44	21:30	31.95	44.85	21.90
45	22:00	32.15	45.70	21.85
46	22:30	32.40	45.55	21.78
47	23:00	32.80	45.60	21.70
48	23:30	32.02	45.55	21.68

three stations. This study used the correlation coefficient R , $RMSE$, and $BIAS$ data for accuracy evaluation in the training and testing stages as follows:

$$R = \frac{\sum_{i=1}^M A \cdot B}{\sqrt{\sum_{i=1}^M A^2 \cdot \sum_{i=1}^M B^2}} \quad (14)$$

Table 3
Accuracy of the obtained results

Models	R	RMSE	BIAS
GMDH-PSO, training phase	0.964	0.167	4.82
GMDH-PSO, testing phase	0.895	0.240	4.11
GA [29]	0.876	0.356	-2.76
DE [19]	0.868	0.378	-3.56
FFBP-NN [1]	0.789	0.465	-3.83

$$RMSE = \left[\frac{\sum_{i=1}^M (Y_{i(Model)} - Y_{i(Original)})^2}{M} \right]^{\frac{1}{2}} \quad (15)$$

$$BIAS = \frac{\sum_{i=1}^M (Y_{i(Model)} - Y_{i(Original)})}{M} \quad (16)$$

where M refers to total events, $A = Y_{i(Original)} - \bar{Y}_{i(Original)}$, $B = Y_{i(Model)} - \bar{Y}_{i(Model)}$, and $Y_{i(Model)}$ are network outputs that refer to the predicted values, $\bar{Y}_{i(Model)}$ refers to the mean of the predicted values, $Y_{i(Original)}$ was target data that showed the observed values, and $\bar{Y}_{i(Original)}$ was the mean of the observed values.

The results for the statistical parameters show that the GMDH-PSO model was in good agreement with previously published methods. Also, the model obtained a precise prediction for the training phase. The values of R and $RMSE$ were equal to 0.96 and 0.167, respectively. $BIAS$ showed good precision in GMDH-PSO training equal to 4.82.

The R and $RMSE$ values were equal to 0.96 and 0.167, respectively. $BIAS$ (equal to 4.82) showed good precision in GMDH-PSO training. The obtained values ($R = 0.89$, $RMSE = 0.24$, and $BIAS = 4.11$) proved the high performance and efficiency of the proposed model during testing. Table V displays the results of the proposed model for the training and testing phases. In comparison, the R values from the testing phase were not significantly different from those for the training stage. On the contrary, $BIAS$ and $RMSE$ were noticeably improved parameters. Table III indicates the accuracy of the obtained results.

A comparison of the GA and DE models was performed for GMDH-PSO. The results showed that the appearance of the GA technique was more accurate than that of the DE model. The values for $RMSE$ and $BIAS$ were equal to 0.356 and -2.76, respectively, for the GA model, and 0.378 and -3.56, respectively, for the DE model. The $BIAS$ for the DE model is smaller than the GA model, but the $RMSE$ for the DE model is very similar to the GA model. Our proposed model showed better results in terms of $BIAS$ compared to the DE and GA models. The GMDH-PSO model exhibited $RMSE$ values approximately 20% lower than

the DE and GA models. The ANN-based model was used for evaluation analysis. This study developed a Feedforward Backpropagation (FFBP) model for water level prediction. In this model, it was assumed that the proposed network had three hidden layers and that each layer had four neurons. This study found that the FFBP-NN technique had more accurate predictions with $RMSE = 0.465$ and $BIAS = -3.83$ compared to the DE model. Although, the FFBPNN model, compared to the GA model, obtained more errors. In total, the GMDH-PSO model showed slightly better performance than the GA model in terms of accuracy. The predicted and measured data of Station 1, Station 2, and Station 3 for the proposed models are shown, respectively, in Figures 6, 7, and 8.

6.1. Performance of GMDH network using BP training and PSO training

This section outlines the performance evaluation of GMDH-PSO and GMDH-BP during the training and testing phases. Model evaluation statistics were $MAPE$ (Mean Absolute Percentage of Error), R , $RMSE$:

$$MAPE = \frac{1}{M} \left[\frac{\sum_{i=1}^M |Y_{i(Model)} - Y_{i(Original)}|}{\sum_{i=1}^M Y_{i(Original)}} \times 100 \right] \quad (17)$$

whereas predicted value network output is $Y_{i(Model)}$, average predicted values are $\bar{Y}_{i(Model)}$, observed values are $Y_{i(Original)}$, average observed values are $\bar{Y}_{i(Original)}$, and M is the total number of events. The training phase shows that the use of the PSO model as a trainer of the GMDH network provided better performance than the use of a back-propagation algorithm. The R values for the GMDH-PSO and GMDHBP models were 0.97 and 0.86, respectively. The $RMSE$ for GMDH-PSO and GMDH-BP were 0.167 and 0.24, respectively. The $MAPE$ values were 0.113 for the GMDH-PSO model and 0.215 for the GMDH-BP model. The GMDH-GP model had a complicated structure because of the creation of tree structures in each neuron, making this procedure very time-consuming.

The testing phase indicated that GMDH-PSO had better performance than GMDH-BP in terms of accuracy. This study used the remaining data sets for testing performances. The R value for GMDH-PSO was 0.96 and the GMDH-GP model was 0.87. The $RMSE$

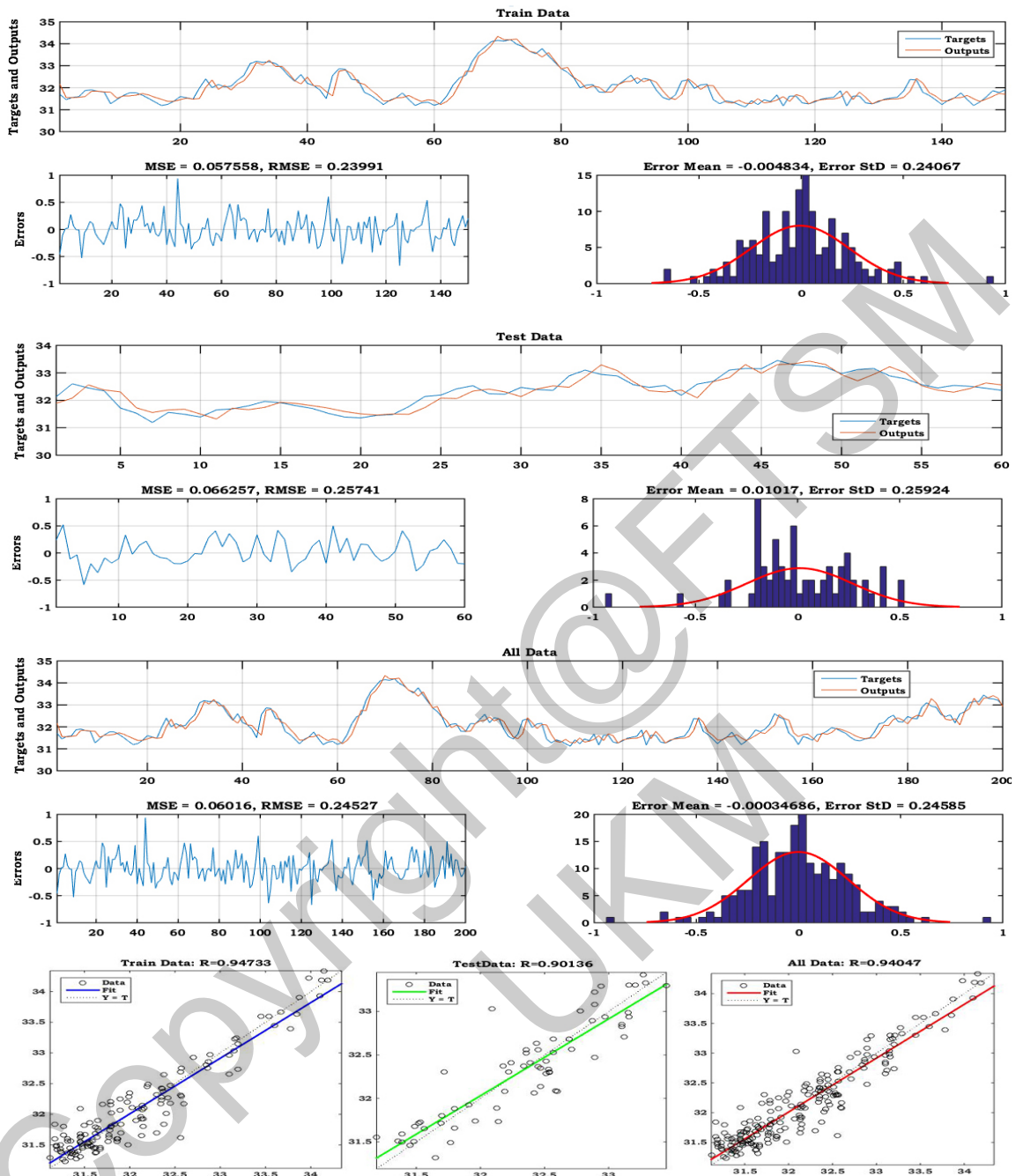


Figure 6: The plots of GMDH-PSO model predicted versus actual values for training, testing and all data sets for Station 1.

and *MAPE* for the GMDH-PSO model were smaller than the GMDH-BP model. Table III shows the comparison results.

7. Simulation validation

In this section, we discussed some experiments conducted to demonstrate the accuracy of our proposed model, and the obtained results were analyzed. These experiments were conducted to implement self-developed

UAV-WSN modules, which were simulated with the OMNET++ tool. In our system, each sensor directly communicates with the UAVs to save energy and decrease the end to end communication delays. This study assumed that the active sensor nodes would communicate with UAVs if they were within the range of the beacon signal. Furthermore, the slept sensor node did not communicate if the beacon signal was weaker than the threshold or the beacon signal was not available. During data collection, this study assumed that each

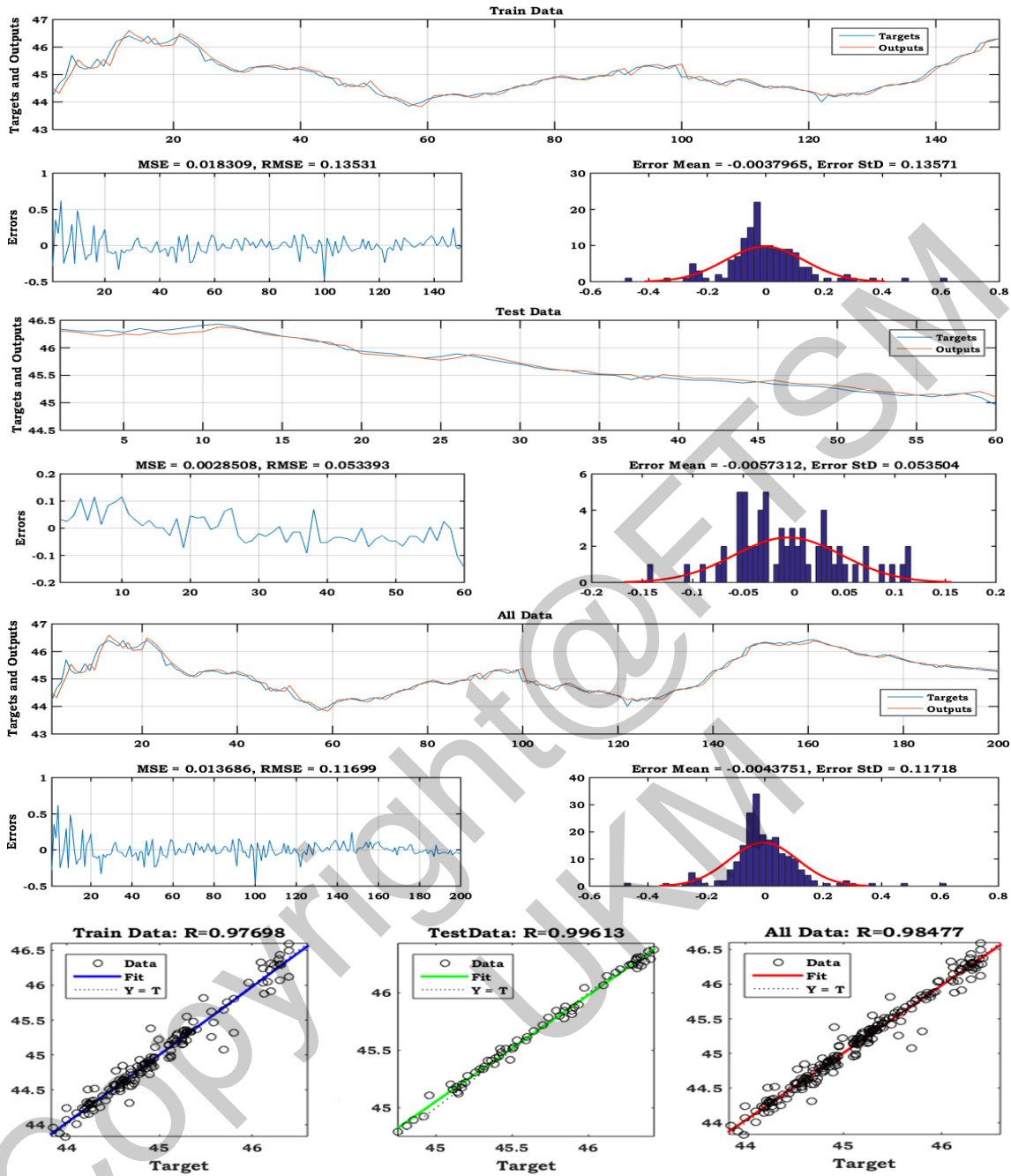


Figure 7: The plots of GMDH-PSO model predicted versus actual values for training, testing and all data sets for Station 2.

active sensor could periodically transmit sensing data to the UAVs. Table IV shows all the parameters used in our simulations and two sets of valuations.

To perform competition experiments, this study carried out different experiments under different experimental conditions. In the first experiment, every sensor node always transmitted a packet between the client and the server. In the second experiment, UAV carried out routing and packet switching between the source node and destination node. In this experiment, if a

sensor node failed, former sensor nodes could not send their data to the sink node. By employing UAVs, the data collection was possible throughout the WSN, which sent data to the central processing unit for river prediction. In this context, performance evaluations were evaluated with two response variables: Round-Trip Time (RTT) delay and packet loss rate. RTT refers to how long it took for a packet to be sent back and forth from the source to the destination. Packet loss rate refers to the ratio of packets lost in the test to the data

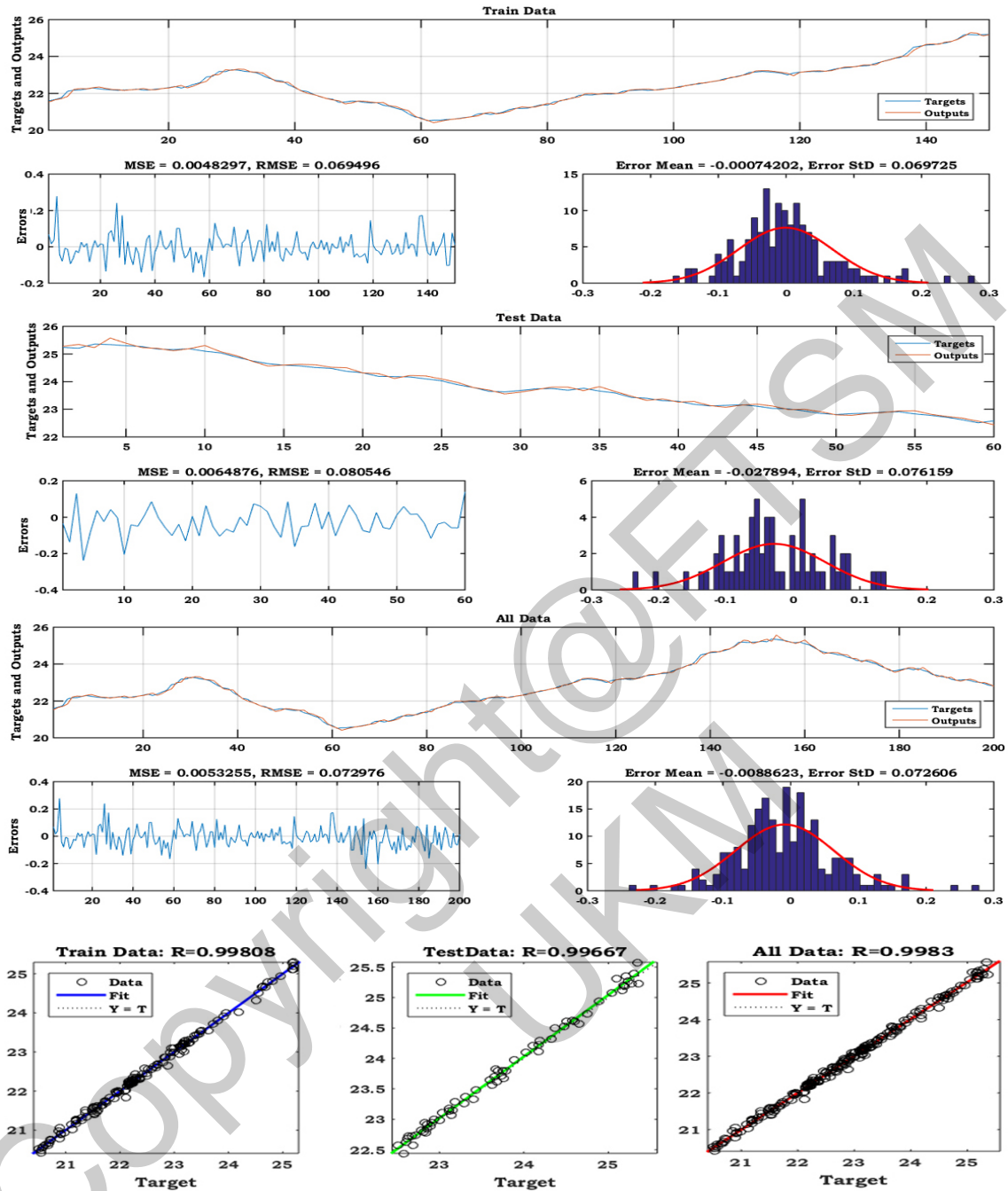


Figure 8: The plots of GMDH-PSO model predicted versus actual values for training, testing and all data sets for Station 3.

groups sent during transmissions. Besides, each experimental result is the average of the 30 runs for each simulation scenario. The 95% confidence interval (CI) has been calculated for the collected performance metrics. To this end, the parameter values used in this study are shown in Table V. These values were carefully selected to reflect realistic scenarios.

7.1. Experiment 1: Without UAVs

In this experiment, data packets were spread across multiple hops to the sink node, meaning that river data and WSN were sent to the base station directly or via other sensors. In this experiment, communication performance was better for direct communication. It was assumed that the communication range of sensor nodes was less than the threshold distance. Each sensor node could decode a direct message from its nearest neighbor in the second time slot [12].

Table 4
Comparison results using GMDH-PSO and GMDH-BP

Models	Training			Testing		
	R	RMSE	MAPE	R	RMSE	MAPE
GMDH-PSO	0.97	0.167	0.113	0.96	0.240	0.215
GMDH-BP	0.87	0.21	0.25	0.87	0.26	0.56

Table 5
Summary of simulation parameters

Parameter	Value
Area of the sensor	400x4000 (m ²)
The UAV's altitude	50-200 m
Number of sensors	500-4000
RF frequency	2 (GHz)
SNR	15 (dB)
Speed of UAV	10 (m/s)
Transmission bit rate of sensor	25 kbps
Path loss between the UAV and sensors	2
Packet for channel access	5 bytes
Data packet length	5-200 byte
Transmission range of sensor, R	20 m
Data transmission interval	2 s
Channel access	Frame random access
Number of particles	100
Number of variables	7
Omega	0.05-0.09
Maximum iteration	500
Error	0.00001
Weighting coefficients C1 and C2	0-2 2.5

Table 6
Set of primary factors in two experiments

Experiments	Antenna gain	Distance
Without UAV	3 dBi	40
	10dBi	40
	3 dBi	60
	10 dBi	60
	3 dBi	90
With UAV	10 dBi	90
	3 dBi	40
	10dBi	40
	3 dBi	60
	10 dBi	60
	3 dBi	90
	10 dBi	90

7.2. Experiment 2: With UAVs

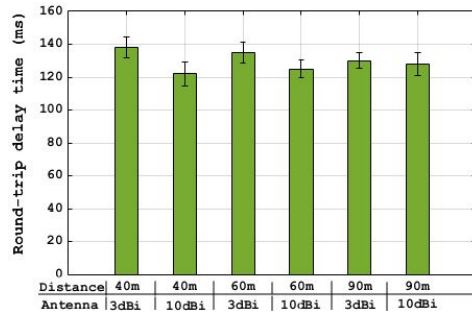
In this experiment, each sensor directly communicated with UAVs. These sensors were found to save energy. Moreover, the delay in end-to-end communication was reduced. This study constructed a module in the central processing unit to manage communications between UAVs and WSNs. In our prototype, UAVs acted as a bridge for routing messages to increase signal range. Consequently, the fixed sensor

node problem was addressed. In this study, UAVs were used as data mules that collected data from the WSN and sent it to the processing unit. The experiment used UAVs flying at distances between 40 m and 90 m with various antenna gains (3 and 10 dBi) and a set communication rate (9600 b/s) and transmission power (10 dBm). Figures 9 and 10 show the results of these two experiments. The two sets of experiments were simulated thirty-five times, and the Shapiro–Wilk normality test was used to test the normality of experiment sets. This study's simulations were organized based on channel access time, which is when sensor nodes can only access one channel at a time. Channel access was conducted in consecutive time frames to minimize packet loss and each time frame was composed of a specific number of timeslots. The size of each timeslot varied and was sufficient for a sensor to transmit its information and be acknowledged by the UAV. In each frame, the sensor randomly chose a timeslot to negotiate with the UAV. Essentially, every node randomly chose a timeslot and asked for authorization from the UAV. If a sensor node received a positive reply, then it waited for its transmission turn. In this case, the node stopped demanding timeslots. The response was comprised of the exact time when the node had connectivity to the UAV. When all the sensors had a given timeslot and completed their channel access, the UAV started the communication process.

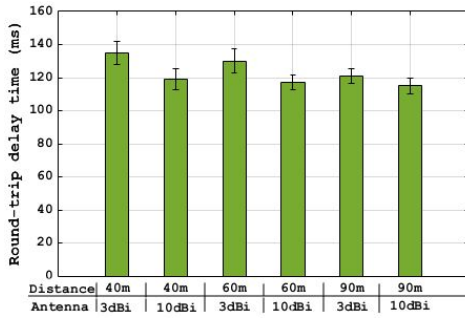
The experiment results showed that UAVs can improve the data collection and provide a reasonably well depiction of remotely sensed environments. Compared with the existing efforts [19, 29, 1, 32], the main advantage of this study is to design a UAV-WSN model for river flow prediction.

8. Conclusions and Future directions

This study used UAV remote sensing for scenarios where a sensor node is unable to send data packets in multi-hop communications to provide robust WSNs. The usage of UAVs can improve the accuracy of water level predictions to prevent floods. Experiments tested data collection performance with and without UAVs for river monitoring. This study's UAV-WSN model proposed the hybridization of the PSO and GMDH models for water level predictions. To validate the precision of the developed GMDH-PSO model, its performance was compared to the DE, GA, and ANN mod-



(a) Without UAVs



(b) With UAVs

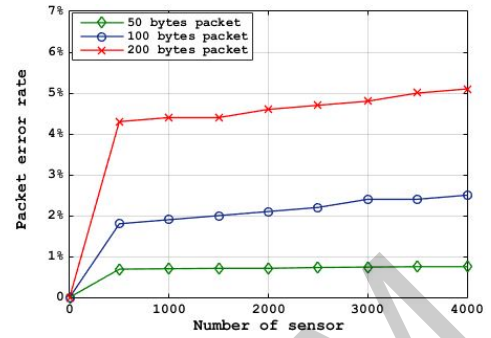
Figure 9: Results (a) and (b) of round-trip time (RTT) delay

els. The GMDH-PSO method outperformed the other models. The statistical indicators used for the performance evaluation of the proposed model indicated lower *RMSE* and higher *R* and *BIAS* compared to the GA and DE models for all nodes. Also, this study compared GMDH-PSO and GMDH-BP during the training and testing stages. The outcomes showed that *MAPE* was lower in the GMDH-PSO model. Results underlined the ability of GMDH-PSO to predict non-linear time series data. For future works, this study recommends the use of other techniques to predict river water levels such as reinforcement learning.

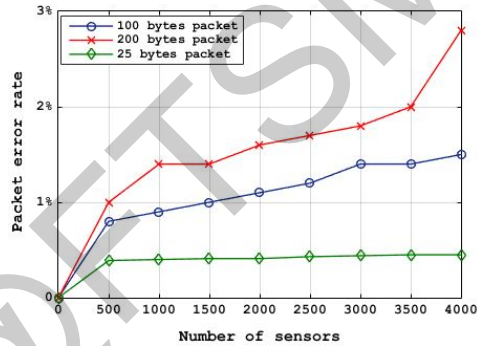
In future research, to improve the computation services while reducing the latency, we plan to apply edge computation (EC). Additionally, we will consider forecast of different environmental phenomena, such as urban underground drainage or rainfall-flow.

References

- [1] Anele, A.O., Hamam, Y., Abu-Mahfouz, A.M., Todini, E., 2017. Overview, comparative assessment and recommendations of forecasting models for short-term water demand prediction. *Water* 9, 887.
- [2] Azamathulla, H.M., Wu, F.C., 2011. Support vector machine approach for longitudinal dispersion coefficients in natural streams. *Applied Soft Computing* 11, 2902–2905.
- [3] Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A.I., Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences* 21, 6201–6217.
- [4] Clerc, M., Kennedy, J., 2002. The particle swarm-explosion, sta-



(a) Without UAVs



(b) With UAVs

Figure 10: Results (a) and (b) of the packet loss rate.

bility, and convergence in a multidimensional complex space. *IEEE transactions on Evolutionary Computation* 6, 58–73.

- [5] De Freitas, E.P., Heimfarth, T., Netto, I.F., Lino, C.E., Pereira, C.E., Ferreira, A.M., Wagner, F.R., Larsson, T., 2010. Uav relay network to support wsn connectivity, in: *International Congress on Ultra Modern Telecommunications and Control Systems*, IEEE. pp. 309–314.
- [6] Erdelj, M., Natalizio, E., 2016. Uav-assisted disaster management: Applications and open issues, in: *2016 international conference on computing, networking and communications (ICNC)*, IEEE. pp. 1–5.
- [7] Goddemeier, N., Daniel, K., Wietfeld, C., 2012. Role-based connectivity management with realistic air-to-ground channels for cooperative uavs. *IEEE Journal on Selected Areas in Communications* 30, 951–963.
- [8] Hauer, S., Leven, S., Zufferey, J., Floreano, D., 2015. The swarming micro air vehicle network (smavnet) project. Retrieved from .
- [9] Hiassat, M., Abbod, M., Mort, N., 2003. Using genetic programming to improve the gmdh in time series prediction. *Statistical data mining and knowledge discovery*, 257–268.
- [10] Hirose, Y., Yamashita, K., Hijiya, S., 1991. Back-propagation algorithm which varies the number of hidden units. *Neural networks* 4, 61–66.
- [11] Hughes, D., Ueyama, J., Mendiondo, E., Matthys, N., Horré, W., Michiels, S., Huygens, C., Joosen, W., Man, K.L., Guan, S.U., 2011. A middleware platform to support river monitoring using wireless sensor networks. *Journal of the Brazilian Computer Society* 17, 85–102.
- [12] Ivakhnenko, A., 1971. Polynomial theory of complex systems: *Ieee transactions on systems, man and cybernetics*.
- [13] Jawhar, I., Mohamed, N., Al-Jaroodi, J., 2015. Uav-based data communication in wireless sensor networks: Models and strategies, in: *2015 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE. pp. 687–694.

- [14] Kalantary, F., Ardalan, H., Nariman-Zadeh, N., 2009. An investigation on the su–nspt correlation using gmdh type neural networks and genetic algorithms. *Engineering Geology* 104, 144–155.
- [15] Kennedy, J., Eberhart, R.C., 1997. A discrete binary version of the particle swarm algorithm, in: 1997 IEEE International conference on systems, man, and cybernetics. *Computational cybernetics and simulation*, IEEE. pp. 4104–4108.
- [16] Khatibi, R., Sivakumar, B., Ghorbani, M.A., Kisi, O., Koçak, K., Zadeh, D.F., 2012. Investigating chaos in river stage and discharge time series. *Journal of Hydrology* 414, 108–117.
- [17] Lera, G., Pinzolas, M., 2002. Neighborhood based levenberg-marquardt algorithm for neural network training. *IEEE transactions on neural networks* 13, 1200–1203.
- [18] Li, J., Tan, S., 2015. Nonstationary flood frequency analysis for annual flood peak series, adopting climate indices and check dam index as covariates. *Water Resources Management* 29, 5533–5550.
- [19] Li, X., Liu, H., Yin, M., 2013. Differential evolution for prediction of longitudinal dispersion coefficients in natural streams. *Water resources management* 27, 5245–5260.
- [20] Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks* 6, 525–533.
- [21] Mozaffari, M., Saad, W., Bennis, M., Debbah, M., 2017. Mobile unmanned aerial vehicles (uavs) for energy-efficient internet of things communications. *IEEE Transactions on Wireless Communications* 16, 7574–7589.
- [22] Najafzadeh, M., Barani, G.A., Hessami Kermani, M., 2013. Aboutment scour in live-bed and clear-water using gmdh network. *Water Sci Technol* 67, 1121–1128.
- [23] Najafzadeh, M., Sattar, A.M., 2015. Neuro-fuzzy gmdh approach to predict longitudinal dispersion in water networks. *Water Resources Management* 29, 2205–2219.
- [24] Nariman-Zadeh, N., Darvizeh, A., Ahmad-Zadeh, G., 2003. Hybrid genetic design of gmdh-type neural networks using singular value decomposition for modelling and prediction of the explosive cutting process. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 217, 779–790.
- [25] Pinto, L.R., Moreira, A., Almeida, L., Rowe, A., 2017. Characterizing multihop aerial networks of cots multirotors. *IEEE Transactions on Industrial Informatics* 13, 898–906.
- [26] Qi, H., Qi, P., Altinakar, M., 2013. Gis-based spatial monte carlo analysis for integrated flood management with two dimensional flood simulation. *Water Resources Management* 27, 3631–3645.
- [27] Robinson, C., 1998. Multi-objective optimisation of polynomial models for time series prediction using genetic algorithms and neural networks. Ph.D. thesis. University of Sheffield.
- [28] Rucco, A., Sujit, P., Aguiar, A.P., De Sousa, J.B., Pereira, F.L., 2017. Optimal rendezvous trajectory for unmanned aerial-ground vehicles. *IEEE Transactions on Aerospace and Electronic Systems* 54, 834–847.
- [29] Sahay, R.R., Dutta, S., 2009. Prediction of longitudinal dispersion coefficients in natural rivers using genetic algorithm. *Hydrology Research* 40, 544–552.
- [30] Shaw, A., Mohseni, K., 2010. A fluid dynamic based coordination of a wireless sensor network of unmanned aerial vehicles: 3-d simulation and wireless communication characterization. *IEEE Sensors Journal* 11, 722–736.
- [31] Skoczylas, M., 2014. Vision analysis system for autonomous landing of micro drone. *acta mechanica et automatica* 8, 199–203.
- [32] Soleymani, S.A., Goudarzi, S., Anisi, M.H., Hassan, W.H., Idris, M.Y.I., Shamsirband, S., Noor, N.M., Ahmedy, I., 2016. A novel method to water level prediction using rbf and ffa. *Water Resources Management* 30, 3265–3283.
- [33] Sun, P., Boukerche, A., Wu, Q., 2017. Theoretical analysis of the target detection rules for the uav-based wireless sensor networks, in: 2017 IEEE International Conference on Communications (ICC), IEEE. pp. 1–6.
- [34] Tuna, G., Mumcu, T.V., Gulez, K., Gungor, V.C., Erturk, H., 2012. Unmanned aerial vehicle-aided wireless sensor network deployment system for post-disaster monitoring, in: International Conference on Intelligent Computing, Springer. pp. 298–305.
- [35] Valente, J., Sanz, D., Barrientos, A., Cerro, J.d., Ribeiro, Á., Rossi, C., 2011. An air-ground wireless sensor network for crop monitoring. *Sensors* 11, 6088–6108.
- [36] Ya'acob, N., Yusop, N., Shariff, K.K.M., Yusof, A.L., Ali, M.T., Salleh, M.K.M., 2012. Observation of tweek characteristics in the mid-latitude d-region ionosphere, in: 2012 IEEE Symposium on Wireless Technology and Applications (ISWTA), IEEE. pp. 27–31.
- [37] Yu, T., Wang, X., Jin, J., McIsaac, K., 2018. Cloud-orchestrated physical topology discovery of large-scale iot systems using uavs. *IEEE Transactions on Industrial Informatics* 14, 2261–2270.

Shidrokh Goudarzi received her Ph.D. degree in communication system and wireless network from Malaysia-Japan International Institute of Technology (MJIT), Universiti Teknologi Malaysia (UTM). In 2014, She received three-year full scholarship to study Ph.D. at (UTM). Then, She joined the Department of Advanced Informatics School at Universiti Teknologi Malaysia as a Postdoctoral Fellow from 2018 to 2019. Currently, she is a senior lecturer at Universiti Kebangsaan Malaysia (UKM). She also serves as reviewer for Canadian Journal of Electrical and Computer Engineering, KSII Transactions on Internet and Information Systems Journal, Journal of Engineering and Technological Sciences, Mathematical Problems in Engineering and IEEE Access. Her research interests are in wireless networks, artificial intelligence, machine learning, next generation networks, Internet of Things (IoT) and Mobile/distributed/Cloud Computing.

Mohammad Hossein Anisi is an Assistant Professor at the School of Computer Science and Electronic Engineering, University of Essex and head of Internet of Everything Laboratory. Prior to that, he worked as a Senior Research Associate at University of East Anglia, UK and Senior Lecturer at University of Malaya, Malaysia where he received 'Excellent Service Award' for his achievements. His research has focused specifically on real world application domains such as energy management, transportation, healthcare and other potential life domains. As a computer scientist, he has designed and developed novel architectures and routing protocols for Internet of Things (IoT) enabling technologies including wireless sensor and actuator networks, vehicular networks, heterogeneous networks, body area networks and his research results have directly contributed to the technology industry. He has strong collaboration with industry and been working with several companies in the UK with the focus on monitoring and automation systems based on IoT concept capable of reliable and seamless generation, transmission, processing and demonstration of data. He has published more than 90 articles in high quality journals and several conference papers and won two medals for his innovations from PECIPTA 2015 and IIDEX 2016 expositions. He has received several International and national funding awards for his fundamental and practical research as PI and Co-I. Dr Anisi is an associate editor of a number of journals including 'IEEE Access', 'Ad Hoc and Sensor Wireless Networks', 'IET Wireless Sensor Systems', 'International Journal of Distributed Sensor Networks', 'KSII Transactions on Internet and Information Systems journals' and 'Journal of Sensor and Actuator Networks'. He has been guest editor of special issues of the journals and Lead organizer of special sessions and workshops at IEEE conferences such as IEEE CAMAD, IEEE PIMRC, IEEE VTC and etc. He has been also serving as executive/technical committee member of several conferences. Hossein is Fellow of Higher Education Academy and Senior Member of IEEE. He is also a technical committee member of Finnish-Russian University Cooperation in Telecommunications (FRUCT), Senior Member of Institute of Research Engineers and Doctors (the IRED), Member of ACM, IEEE Council on RFID, IEEE Sensors Council, IEEE Systems Council and International Association of Engineers (IAENG).

Domenico Ciuonzo is currently an Assistant Professor at University of Napoli Federico II, Italy. He holds a Ph.D. in Electronic Engineering from University of Campania "L. Vanvitelli", Italy and, from 2011, he has held several visiting researcher appointments. Since 2014 he is (senior) editor of several IEEE, IET and ELSEVIER journals. His research interests include data fusion, statistical signal processing, wireless sensor networks, Internet of Things, traffic analysis and machine learning.

Seyed Ahmad Soleymani received his Ph.D. degree in computer science from faculty of engineering, Universiti Teknologi Malaysia (UTM). He received his M.S degree from the Department of Computer Engineering, Islamic Azad University, Iran and B.S. degree from the Department of Computer Engineering, Sadjad University, Iran. His research interests are in the area of Wireless Sensor Network (WSN), Mobile Ad Hoc Network (MANET), Vehicular Ad Hoc Network (VANET), Internet of Things and Nano Things (IoT and IoNT), Visible Light Communication (VLS), Intelligent Algorithms (IAs), Big Data and Machine Learning.

Salwani Abdullah She obtained her BSc in Computer Science from Universiti Teknologi Malaysia and her master degree specializing in Computer Science from Universiti Kebangsaan Malaysia (UKM). She did her PhD in Computer Science at The University of Nottingham, United Kingdom. Now, she is a Professor in Computational Optimization at the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Her research interest falls under Artificial Intelligence and Operation Research, particularly computational optimization algorithms (heuristic and meta-heuristic, evolutionary algorithms, local search) that involve different real world applications for single and multi-objective continuous and combinatorial optimization problems such as timetabling, scheduling, routing, nurse rostering, dynamic optimization, data mining problems (feature selection, clustering, classification, time-series prediction), intrusion detection and search based software testing.

Muhammad Khurram Khan is currently working as a Professor of Cybersecurity at the Center of Excellence in Information Assurance (CoEIA), King Saud University, Kingdom of Saudi Arabia. He is one of the founding members of CoEIA and has served as Manager R&D from March 2009 until March 2012. He, along with his team, developed and successfully managed Cybersecurity research program at CoEIA, which turned the center as one of the best centers of excellence in Saudi Arabia and in the region. He is founder and CEO of the 'Global Foundation for Cyber Studies and Research' (<http://www.gfcyber.org>), an independent, non-profit, and non-partisan cybersecurity think-tank in Washington D.C. USA. Prof. Khurram is the Editor-in-Chief of a well-reputed International journal 'Telecommunication Systems' published by Springer-Nature with its recent impact factor of 1.707 (JCR 2019). Furthermore, he is the editor of several international journals, including, IEEE Communications Surveys & Tutorials, IEEE Communications Magazine, IEEE Internet of Things Journal, IEEE Transactions on Consumer Electronics, Journal of Network & Computer Applications (Elsevier), IEEE Access, IEEE Consumer Electronics Magazine, PLOS ONE, Electronic Commerce Research, IET Wireless Sensor Systems, Journal of Information Hiding and Multimedia Signal Processing, and International Journal of Biometrics, etc. He has also played role of the guest editor of several international journals of IEEE, Springer, Wiley, Elsevier Science, and Hindawi. Moreover, he is one of the organizing chairs of more than 5 dozen international conferences and member of technical committees of more than 10 dozen international conferences. In addition, he is an active reviewer of many international journals as well as research foundations of Switzerland, Italy, Saudi Arabia and Czech Republic. Prof. Khurram is an honorary Professor at IIIRC, Shenzhen Graduate School, China and an adjunct professor at Fujian University of Technology, China. He has secured an outstanding leadership award at IEEE international conference on Networks and Systems Security 2009, Australia. He

has been included in the Marquis Who's Who in the World 2010 edition. Besides, he has received certificate of appreciation for outstanding contributions in 'Biometrics & Information Security Research' at AIT international Conference, June 2010 at Japan. He has been awarded a Gold Medal for the 'Best Invention & Innovation Award' at 10th Malaysian Technology Expo 2011, Malaysia. Moreover, in April 2013, his invention has got a Bronze Medal at '41st International Exhibition of Inventions' at Geneva, Switzerland. In addition, he was awarded best paper award from the Journal of Network & Computer Applications (Elsevier) in Dec. 2015. Prof. Khurram is the recipient of King Saud University Award for Scientific Excellence (Research Productivity) in May 2015. He is also a recipient of King Saud University Award for Scientific Excellence (Inventions, Innovations, and Technology Licensing) in May 2016. He has published more than 360 papers in the journals and conferences of international repute. In addition, he is an inventor of 10 US/PCT patents. He has edited 9 books/proceedings published by Springer-Verlag, Taylor & Francis and IEEE. He has secured several national and international competitive research grants with an amount of over USD 3 Million in the domain of Cybersecurity. Prof. Khurram has played a leading role in developing 'BS Cybersecurity Degree Program' and 'Higher Diploma in Cybersecurity' at King Saud University. In 2019, he has played an instrumental role as a cybersecurity subject expert for a USD 6 Million series B investment in a South Korean startup 'SecuLetter', which has received a corporate valuation of USD 38 Million. His research areas of interest are Cybersecurity, digital authentication, IoT security, biometrics, multimedia security, cloud computing security, cyber policy, and technological innovation management. He is a fellow of the IET (UK), fellow of the BCS (UK), fellow of the FTRA (Korea), senior member of the IEEE (USA), senior member of the IACSIT (Singapore), member of the IEEE Consumer Electronics Society, member of the IEEE Communications Society, member of the IEEE Computers Society, member of the IEEE Technical Committee on Security & Privacy, member of the IEEE IoT Community, member of the IEEE Smart Cities Community, and member of the IEEE Cybersecurity Community. He is also the Vice Chair of IEEE Communications Society Saudi Chapter. He is a distinguished Lecturer of the IEEE. His detailed profile can be visited at <http://www.professorkhurram.com>



Copyright@FTSM
UKM



Copyright@FTSM
UKM

Dear Editor-in-Chief,

We would like to confirm that this manuscript has not been published or submitted for consideration elsewhere. We also declare that there is no conflict of interests to disclose.

Thank you for your time and consideration!

Sincerely,

The authors

Copyright@FTSM
UKM