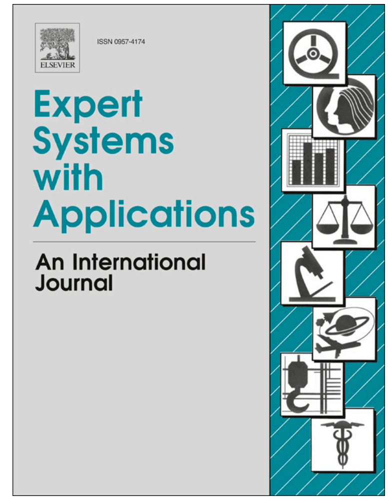
Journal Pre-proofs



Real-time Water demand pattern estimation using an Optimized Extended Kalman Filter

Fatemeh Attarzadeh, Ali Naghi Ziaei, Kamran Davary, Esmaeil Fallah Choulabi

PII: S0957-4174(23)02584-8
DOI: https://doi.org/10.1016/j.eswa.2023.122082
Reference: ESWA 122082

To appear in: *Expert Systems with Applications*

Received Date: 10 October 2022
Revised Date: 24 September 2023
Accepted Date: 7 October 2023

Please cite this article as: Attarzadeh, F., Naghi Ziaei, A., Davary, K., Fallah Choulabi, E., Real-time Water demand pattern estimation using an Optimized Extended Kalman Filter, *Expert Systems with Applications* (2023), doi: https://doi.org/10.1016/j.eswa.2023.122082

Real-time Water demand pattern estimation using an Optimized Extended Kalman Filter

Fatemeh Attarzadeh [a], Ali Naghi Ziaei [b,*], Kamran Davary [c], Esmaeil Fallah Choulabi [d]

[a] Department of Water Science and Engineering, College of Agriculture, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. f.attarzade@yahoo.com

[b] Department of Water Science and Engineering, College of Agriculture, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. an-ziaei@um.ac.ir (Corresponding Author)

[c] Department of Water Science and Engineering, College of Agriculture, Ferdowsi University of Mashhad (FUM), Mashhad, Iran. kamdav@um.ac.ir

[d] Department of Electrical Engineering, Faculty of Engineering, University of Guilan, Rasht, Iran. fallah_e@guilan.ac.ir

Declaration of interests

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

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

CRediT author statement

Fatemeh Attarzadeh: Analysis and/or interpretation of data, Code development, Data Acquisition, Drafting and revising the manuscript

Ali Naghi Ziaei.: Conception and design of study, Analysis and/or interpretation of data, Writing and revising the manuscript

Kamran Davary: Supervision

* Corresponding author

Esmaeil Fallah Choulabi: Supervision, Writing- Reviewing and Editing

Abstract

This study presents a hybrid approach for the estimation of real-time water demand multipliers using the Kalman filter (KF) and extended Kalman filter (EKF). Multiple Linear Regression (MLR) and Nonlinear Regression (NLR) models were applied to predict water demand multipliers at each time step with historical flow data. The estimation performance of EKF is highly affected by the state noise covariance matrix (Q) and the measurement noise covariance matrix (R). An inappropriate value of Q and R significantly degrades the EKF's performance and makes the filter diverge. So, the particle swarm optimization algorithm (PSO) was used to tune the noise covariance matrices Q and R at each time step of EKF. Then the optimal values of noise covariance matrices are inserted in the real-time water demand multiplier estimation process. To find the optimal values of Q and R, the mean absolute percentage error (MAPE) between measured and simulated pressure was minimized. The proposed method was evaluated in Net1 and Net3 benchmark networks. The root means square error (RMSE) of EKF-PSO estimated water demand multiplier for Net1 and Net3 were 0.063 and 0.198, respectively. The simulation results indicated that the EKF-PSO algorithm was more accurate than the conventional EKF algorithm. Moreover, the KF-PSO performed poorly when dealing with nonlinear hydraulic systems.

Keywords: Demand multiplier, PSO, Online Calibration, Water Distribution Networks, EPANET

1. Introduction

Hydraulic analysis models are widely used by water utilities and others who are involved in the analysis, design, operation, and maintenance of water distribution networks (WDNs). The estimation of water demand patterns to achieve the optimal operation of water distribution networks is crucial. (Abu-Mahfouz et al., 2019). Estimation of water demand patterns is one of the critical steps in the planning and design of water distribution systems. Due to economic limitations, it is practically impossible to measure real-time nodal demand for all water consumers (Zhao et al., 2018), but the monitoring data (e.g., measured pressure data from Supervisory Control and Data Acquisition or SCADA) and water distribution model (EPANET) can be assimilated simultaneously to give a real-time water demand pattern. Generally, the data assimilation (DA) technique updates the model states or parameters using real-time observations.

Data assimilation is widely used in many subjects, including electric power systems (Blood et al., 2008; Netto et al., 2016), petroleum engineering (Kang et al., 2017), prediction of soil moisture and temperature

(Liu et al., 2010; Yu et al., 2014), estimation of soil hydraulic parameters (Liu et al., 2021), prediction of groundwater contaminant concentration (Assumaning and Chang, 2016), river water temperature (Rajesh and Rehana 2021), river flood forecasting (Li et al., 2014), management of water resources (Kurtz et al., 2017), hydrology (Wang, 2009), and meteorological sciences (Pelosi et al., 2017).

Several studies have integrated online measurements into hydraulic state estimation models. For example, Preis et al. (2011) employed genetic algorithms (GA) to update the predictions of the water demand multiplier based on online measurements. Nasseri et al (2012) proposed a hybrid model that combines the Extended Kalman Filter (EKF) and Genetic Programming (GP) for forecasting water demand. Vassiljev and Koppel (2015) applied the Levenberg-Marquardt algorithm (LMA) and the Genetic algorithm (GA) to estimate real-time demands in a water distribution system. Their results demonstrated that the LMA works much faster than the GA. Do et al. (2016) used the genetic algorithm (GA) to calibrate the predicted demand multiplier factors. Their results showed that GA has a high computational cost. Salloom et al. (2021) proposed a novel deep neural network for real-time water demand forecasting. This model depends on water demand history, making it susceptible to abnormalities in water demand. Zhang et al. (2023) proposed a deep fuzzy mapping nonparametric model (DFM) to estimate real-time nodal demands in water distribution systems. The DFM approach includes a unique analytical solution derived through mathematical theory. The results showed that this method is more accurate and computationally efficient compared to traditional calibration methods. Although the capabilities of artificial neural networks (ANNs) can be improved using larger training datasets, it would be computationally expensive and impracticable, especially for large WDSs (Garzón et al., 2022). One practical advantage of data assimilation models is that there is no need for historical data, so it is appropriate for real-time forecasting problems.

Some researchers have studied the issue of near-real-time demand estimation using data assimilation methodologies based on the Kalman filter. For example, Shang et al. (2006) used an extended Kalman filter (EKF), a predictor-corrector method, to estimate water demand patterns. In this paper, water demand patterns were predicted by a seasonal autoregressive integrated moving average (ARIMA) time-series model. Kang and Lansey (2009) applied two real-time methods for the demand estimation problem, the Kalman filter (KF) and the tracking state estimator (TSE). The results showed that KF performed poorly in a looped network. They also stated that pipe flow data are significantly more effective than pressure measurements in estimating reliable demands. Okeya et al. (2014b) applied DA methods to estimate unmetered domestic demands of a WDN and demonstrated that the Ensemble Kalman Filter (EnKF) performed well compared to KF in terms of updating water demand model parameters. However, KF is less time-consuming than EnKF. Jung et al. (2016) proposed an optimal node grouping model to improve real-time demand estimation. They linked the Kalman filter-based demand estimation and a genetic algorithm for node group optimization. According to their results, more pipe flow sensors can enhance demand estimation accuracy. Do et al. (2017) employed the particle filter method for the estimation of near-real-time demand multipliers. In their presented method, the nodal water demand is predicted by a nonlinear model, and the prediction is corrected by real-time pressure measurements. Zhou et al. (2018) proposed a self-adaptive method based on KF for dual calibration of pipe roughness and nodal demands. They aimed to assimilate online pressure data from pressure sensors in a water hydraulic model (EPANET) to estimate the real-time water demand. In most nonlinear systems, the EnKF is favored over than EKF. Nonetheless, the EKF is used to increase the effectiveness of the estimation because it is brief and explicit (Chen et al., 2019). Compared to other data assimilation methods, EKF is straightforward to implement, but it suffers from the costly calculation of state and measurement noise covariance matrices (Sun et al., 2016; Chen et al., 2021). One of the main challenges of the EKF method is finding the optimal values of tuning parameters such as covariance matrices Q and R of state and measurement noises. This issue has rarely been addressed in the literature.

The optimal performance of the Kalman filter depends on the quality of prior assumptions of the process noise covariance matrix, Q and the measurements noise covariance matrix, R (Mohamed and Schwarz, 1999).

In the EKF process, the values of Q and R matrices have a significant impact on the convergence rate and estimation error (Wang and Mu, 2019). The improper value of Q and R may significantly demote the EKF's performance and even make the filter diverge (Akhlaghi et al., 2017). If R and/or Q are too small in the estimation process, the uncertainty around the true value will reduce and a biased solution will be made. If R and/or Q are selected too large, the filter may diverge (Mohamed and Schwarz, 1999). For instance, when R rises, Kalman gain reduces, which makes a diverging estimate (Maheshwari and Nageswari, 2022). So the noise covariance matrices Q and R should be obtained by taking into account the stochastic properties of the corresponding noises (Laamari et al., 2015); but, more often these are unknown, in most cases, the covariance matrices are employed as free tuning parameters. In several cases, these matrices were adjusted by trial and error approaches which suffer from large time consumption. To overcome this problem, evolutionary algorithms were used to find the optimal values of the two matrices Q and R (Shi et al., 2002). Kaba and Kıyak (2020) proposed an evolutionary algorithm based on the Kalman filter (EA–KF) for tuning the noise covariance matrices to simulate quadrotors. Rossi et al. (2022) presented a hybrid method that combines an extended Kalman filter (EKF) with a genetic algorithm for the estimation of Li-Ion cell parameters. They tuned the covariance matrices of the EKF by using genetic algorithms (GA). Jatoth and Kumar (2009) investigated the tuning of Unscented Kalman filters (UKF) using Particle Swarm Optimization (PSO) and Bacterial Foraging Optimization (BFO). Their result demonstrated that UKF-PSO is superior to UKF-BFO.

In the present study, the particle swarm optimization (PSO) method is combined with KF and EKF to estimate the water demand pattern based on measured pressure values. In other words, to optimize the process noise covariance Q and observation noise covariance R, an evolutionary optimization algorithm is used to reduce the estimation error of KF and EKF. The time interval between two successive measurements in the water demand estimation problem allows us to run PSO in each estimation step and achieve optimal values of Q and R. This approach in water demand estimation is used for the first time in this paper. Therefore, this investigation aims to highlight the impact of a hybrid approach (EKF-PSO) to estimate real-time water demand multipliers in WDNs.

2. Methodology

The proposed approach was established in two steps. First, KF/EKF was used and the covariance matrices of state noise and measurement noise were optimized by particle swarm optimization (PSO) at each time step of the EKF. Then the optimal values of these covariance matrices were applied in a real-time water demand multiplier estimation loop.

The proposed approach was evaluated in two modified benchmark networks (i.e., Net1 and Net3). A detailed description of the KF and EKF methodologies that are applied for water demand-state estimation in WDN is presented first.

2.1. Water Demand Forecasting Model (WDFM)

WDFM is a vital part of the real-time hydraulic model. A variety of methods has been used to forecast water demand patterns including regression analysis.

In this study, the WDFMs are based on Multiple Linear Regression (MLR) and Nonlinear Regression (NLR). MLR models are relatively simple to implement. Nevertheless, they are limited to the high-nonlinearity system. A multiple linear regression equation is defined as follows [Montgomery et al., 2021]:

$$d_t = \sum w_{t-i} d_{t-i}$$
$$d_t = w_{t-1}d_{t-1} + w_{t-6}d_{t-6} + w_{t-12}d_{t-12} + w_{t-18}d_{t-18} + w_{t-24}d_{t-24} + w_{t-30}d_{t-30} + w_{t-36}d_{t-36} + w_{t-42}d_{t-42} + w_{t-48}d_{t-48} \quad (1)$$

where d_t and d_{t-i} are the water demand multiplier from time step t and $t-i$ respectively and w_{t-i} is the associated weight for d_{t-i} (e.g., t-6 indicates 6 hours before the current time when the time step is 1 hour). Additionally, a nonlinear regression model (NLR) was used to linearize the complex system which is essential for EKF. Nonlinear regression is a form of regression to find nonlinear models between a dependent variable and a set of independent variables. The nonlinear equation was considered as follows:

$$d_t = \alpha_1 d_{t-1}^{\alpha_2} + \alpha_3 d_{t-6}^{\alpha_4} + \alpha_5 d_{t-12}^{\alpha_6} + \alpha_7 d_{t-18}^{\alpha_8} + \alpha_9 d_{t-24}^{\alpha_{10}} + \alpha_{11} d_{t-30}^{\alpha_{12}} + \alpha_{13} d_{t-36}^{\alpha_{14}} + \alpha_{15} d_{t-42}^{\alpha_{16}} + \alpha_{17} d_{t-48}^{\alpha_{18}} \quad (2)$$

where $\alpha_1, \alpha_2, ..., \alpha_{18}$ are the parameters for the nonlinear regression model. All of the regression models were trained using 80 percent of the total data and then tested using the rest of 20 percent.

2.2 WDN hydraulics equation

The steady-state hydraulic relationships in WDN can be presented by the nodal flow continuity and pipe head loss equations (Bhave and Gupta, 2006). For steady incompressible flow, for each node, the algebraic sum of inflow and outflow must be zero. Thus,

$$\sum_{i,j=1}^{N} Q_{ij} + q_i = 0 \quad (3)$$

in which q_i is nodal demand, Q_{ij} is pipe flows and N is the number of nodes. Pipe head loss relationship for all pipes in the network can be expressed as follows,

$$h_f = He_i - He_j = R_{ij} Q_{ij}^{\;n} \quad (4)$$

where h_f is the head loss in a pipe; He_i, He_j are nodal head for node i and node j, the exponent n is equal to 1.852 for Hazen–Williams formula and R_{ij} is resistance constant of pipe that obtained as follows,

$$R = \frac{10.68 \times L}{C_{HW}^{\ 1.852} \times D^{4.87}} \tag{5}$$

where C_{HW} is the Hazen-Williams coefficient, L is pipe length in meters, and D is pipe diameter in meters. Eq. (4) can also be expressed as:

$$Q_{ij} = \left(\frac{He_i - He_j}{R_{ij}} \right)^{1/n} \tag{6}$$

Using Equation (6), Equation (3) can be rewritten as:

$$\sum_{i,j=1}^{N} \left(\frac{He_i - He_j}{R_{ij}} \right)^{1/n} + q_j = 0 \tag{7}$$

The relation between the demand pattern and nodal head can be obtained by substituting $q_j = q_{base_j} \times q_{pattern}$ in Eq. (7).

2.3. Evolutionary algorithm (EA)

Evolutionary algorithms (EAs) are stochastic optimization techniques and population-based that mimic natural evolution. A lot of swarm intelligence optimization algorithms, such as Particle Swarm Optimization (PSO) (Eberhart and Kennedy, 1995), Ant Colony Optimization (ACO) (Dorigo et al., 1996), Brain Storm Optimization (BSO) (Shi, 2011), Invasive Weed Optimization (IWO) (Mehrabian and Lucas, 2006), Imperialist Competitive Algorithm (ICA) (Atashpaz Gargari and Lucas, 2007), Bacterial Foraging Optimization (BFO) (Passino, 2012), Grey Wolf Optimization (GWO) (Mirjalili et al., 2014), Orthogonal Learning framework for Brain Storm Optimization (OLBSO) (Ma et al., 2020), have been proposed to tackle complex optimization problems.

In the last decade, evolutionary algorithms (EAs) have been applied for the optimization of WDNs (Jung and Karney, 2008; Dini and Tabesh, 2014; Do et al, 2016). Attarzadeh et al. (2022) and Attarzadeh et al. (2023) applied six evolutionary algorithms (EAs) to calibrate the pipe roughness coefficient and water demand coefficient in water distribution systems. These EAs include the genetic algorithm (GA), the particle swarm optimization (PSO), the Gray Wolf Optimization algorithm (GWO), the invasive weed optimization (IWO), the Imperialist Competitive Algorithm (ICA), and the Simulated Annealing (SA). The results showed that all six algorithms can decrease the difference between observed and simulated pressure and flow data after calibration. The results obtained for the Apulian network indicate that the performance of the PSO is superior in terms of accuracy, the number of objective function evaluations (NFE), run time, and convergence rate.

PSO is one of the most well-known and widely used swarm intelligence algorithms and metaheuristic techniques, because of its simplicity and ability to be used in a wide range of applications. Compared to other heuristic algorithms, PSO does not need to learn many parameters. It is suitable for multidimensional engineering problems and is capable of finding optimal solutions quickly (Sun et al., 2021; Song and Rahmat-Samii, 2021). Yarat et al. (2021) also reported its capability in various application fields. In this study, the PSO is implemented to tune the parameters of the Extended Kalman Filter to improve its performance.

2.3.1. Particle Swarm Optimization (PSO) algorithm

The PSO is a population-based optimization technique first proposed by Eberhart and Kennedy (1995). This algorithm is inspired by the behavior of a flock of birds or fish and applies swarm intelligence to find optimal solutions. The process starts with a set of particles (P_0). In PSO, each bird is represented by a particle, and a collection of birds is identified as a swarm. Each particle has a fitness value, which is based on the objective function value. In each iteration, the particles move from their current position (X_i^t) to their new position (X_i^{t+1}) according to Eq. (8):

$$X_i^{t+1} = X_i^t + V_i^{t+1} \tag{8}$$

The particle velocity (V_i^{t+1}) can be calculated using the following formula (Eq. (9)):

$$V_i^{t+1} = wV_i^t + c_1 r_1 (\mathrm{Pbest}_i - X_i^t) + c_2 r_2 (\mathrm{Gbest}_i - X_i^t) \tag{9}$$

where Pbest_i is the best particle position, Gbest_i is the global best particle position, $c_1 = 2$ and $c_2 = 2$ are acceleration constants which were found by trial and error that works well for almost all applications, r_1 and r_2 are random numbers between 0 and 1 (Babu and Vijayalakshmi, 2012; Moghaddam et al., 2018), w is the inertia weight that represents the exploration and exploitation ability of the algorithm for which the allowable value changes in the range of 0.4 to 0.9. In this algorithm, a reduction coefficient called w_{damp} is used to increase the exploration characteristics in the final steps (see Eq. 10). The value of w_{damp} factor considered in this paper is 0.998 (Zaji and Bonakdari, 2014; Moghaddam et al., 2018).

$$w^{t+1} = w^t \times w_{damp} \tag{10}$$

2.4. Objective Function

In this study, the covariance matrices of state noise and measurement noise are modified simultaneously to achieve the optimal solution. The objective function of the model is to minimize the mean absolute percentage error (MAPE) between the measured and estimated nodal head values. The objective function is given as Eq. (11):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{He_{obs} - He_{sim}}{He_{obs}}\right| \times 100 \tag{11}$$

where n is the number of nodal head measurement locations in the network, He_{obs} is the observed nodal head (m), and He_{sim} is the estimated nodal head (m).

2.5. Data assimilation

The Kalman Filter (KF) is an optimal recursive algorithm to estimate the state of a process which was introduced and developed by Kalman (1960). The general framework of the KF consists mainly of two steps including forward prediction and measurement correction. In the state equation, the priori state estimation of the water demand multiplier x_t^- is computed as below:

$$x_t^- = Ax_{t-1}^+ + \omega_t \tag{12}$$

where x_{t-1}^+ is the state vector at time $t-1$; A is a transition matrix that converts the state vector from time $t-1$ to the next time t; ω_t is the error vector and the superscripts – and + represent the predicted variable and the updated variable, respectively. The prior error covariance matrix P_t^- at time t has been shown in Eq. (13).

$$P_t^- = A_t P_{t-1}^+ A_t^T + Q \tag{13}$$

where P_{t-1}^+ is the posterior error covariance at time $t-1$. In the simulation, the error covariance matrix P_0^- of EKF is initially set as a unit matrix (Shi et al., 2002). Q is the state noise covariance which should be tuned. If the number of state estimation (water demand multiplier) is equal to n, the value of measurement noise covariance matrix will be $Q = Q_{ini} \times I_n$, where I_n is the unit matrix with dimension $n \times n$ and Q_{ini} is the initial value for Q. The posterior state estimate of the water demand multiplier is given by:

$$x_t^+ = x_t^- + K_t(z_t - H_t x_t^-) \tag{14}$$

where H_t is the measurement operator matrix which converts the model states (i.e., demands) to the WDN observations (i.e., pressure); z_t is measurement variable (i.e., pressure); K_t is Kalman gain matrix at time step t calculated with Eq. (15).

$$K_t = P_t^- H_t^T (H_t P_t^- H_t^T + R)^{-1} \tag{15}$$

where H_t^T is the transposed measurement operator, and R is the covariance matrix of measurement noise. If the number of pressure measurement points is equal to m, the value of measurement noise covariance matrix will be $R = R_{ini} \times I_m$, where I_m is the unit matrix with dimension $m \times m$ and R_{ini} is the initial value for R. The posterior error covariance matrix is calculated as (Sen et al., 2004):

$$P_t^+ = P_t^- - K_t H_t P_t^- \tag{16}$$

According to Eq. (7) the relationship between the water demand multiplier and the nodal head is as follows:

$$F = \frac{1}{q_{base_j}} \sum_{i,j=1}^{N} \left(\frac{He_i - He_j}{R_{ij}} \right)^{\frac{1}{n}} - q_{pattern} = 0 \tag{17}$$

where the measurement operator matrix, H_t can be obtained by inverting the Jacobian matrix J as follows:

$$J = \begin{bmatrix} \frac{\partial F_1}{\partial He_1} & \frac{\partial F_1}{\partial He_2} & \cdots & \frac{\partial F_1}{\partial He_J} \\ \frac{\partial F_2}{\partial He_1} & \frac{\partial F_2}{\partial He_2} & \cdots & \frac{\partial F_2}{\partial He_J} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_J}{\partial He_1} & \frac{\partial F_J}{\partial He_2} & \cdots & \frac{\partial F_J}{\partial He_J} \end{bmatrix} \tag{18}$$

$$H_t = [\mathrm{J}\ ]^{-1}_{J \times J} \tag{19}$$

In the extended Kalman Filter (EKF), the estimated state is defined as:

$$x_t^- = f(x_{t-1}^+) + \omega_t \quad (20)$$

where f is the nonlinear function of x_{t-1}^+ and the linearized system dynamics can be written as:

$$x_t^- = Ax_{t-1}^+ + \omega_t \quad (21)$$

A is a nonlinear model operator of the partial differential function which is written as follows:

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \quad (22)$$

2.6. Proposed Framework

As discussed before, the parameters to be modified are the noise covariance matrices Q and R. To improve the performance of EKF, the PSO algorithm was used to optimize these parameters. The framework of the EKF-PSO is illustrated in Fig.1. In the first step, the initial state of the system (x_0^-), the covariance of the estimation error (P_0^-), the process noise covariance Q and observation noise covariance R are manually chosen to provide good performance. Then water demand forecasting models are developed to predict water demand multipliers based on historical flow data. These models are used as an estimate of water demand multipliers in the previous time step. In the second step, the state vector (x_t^-) and the covariance error (P_t^-) are predicted. Then, Kalman gain is obtained by considering the measurement operator matrix H_t which relates the true model state to the observations. After that, the error covariance matrix (P_t^+) is computed. When the SCADA measurements of the current time step (z_t) or nodal head measurements are available, the state estimation vector (x_t^+) is corrected. So the EKF method gives an estimation of the water demand multipliers by using measured nodal heads and the water demand forecasting model. Then the EPANET is executed to simulate the nodal head. The MAPE criterion is calculated as an objective function between the measured and simulated nodal head. If the value of MAPE is not acceptable, the covariance matrices Q and R are optimized by the PSO algorithm. The new updated Q and R are then used for the adaptation of the EKF for the next iteration until the criterion (number of iterations) is established. This step works at each time step of the EKF. In the third step, obtained values Q and R from the previous step are inserted into EKF to estimate the real-time water demand multipliers.

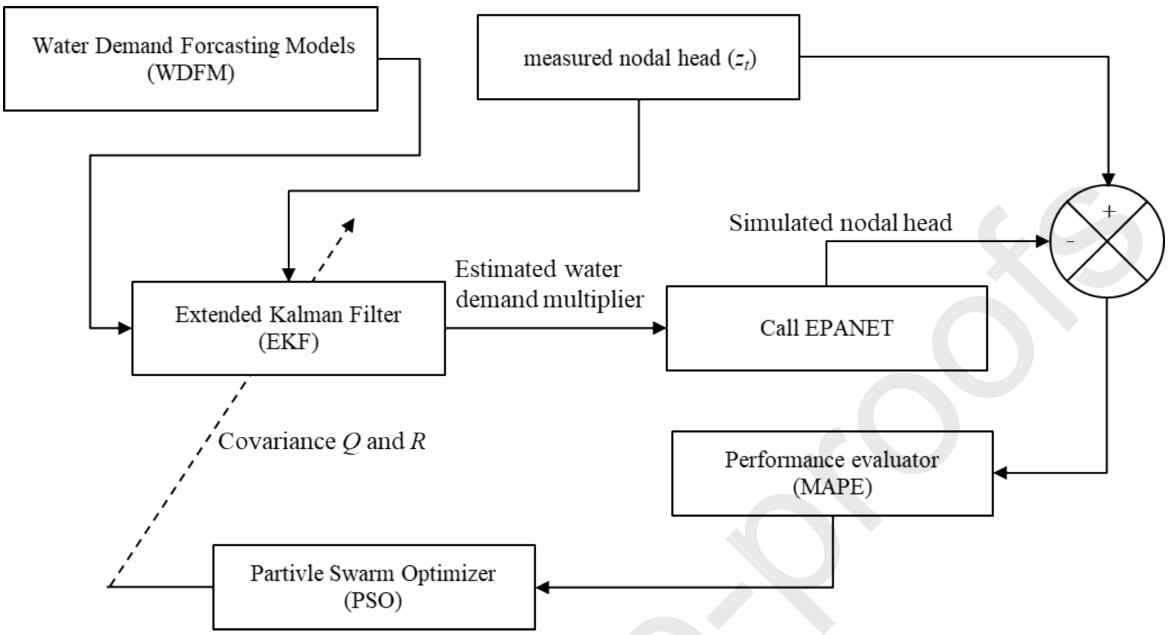


Fig. 1. Block chart of the Extended Kalman Filter tuning procedure based on PSO

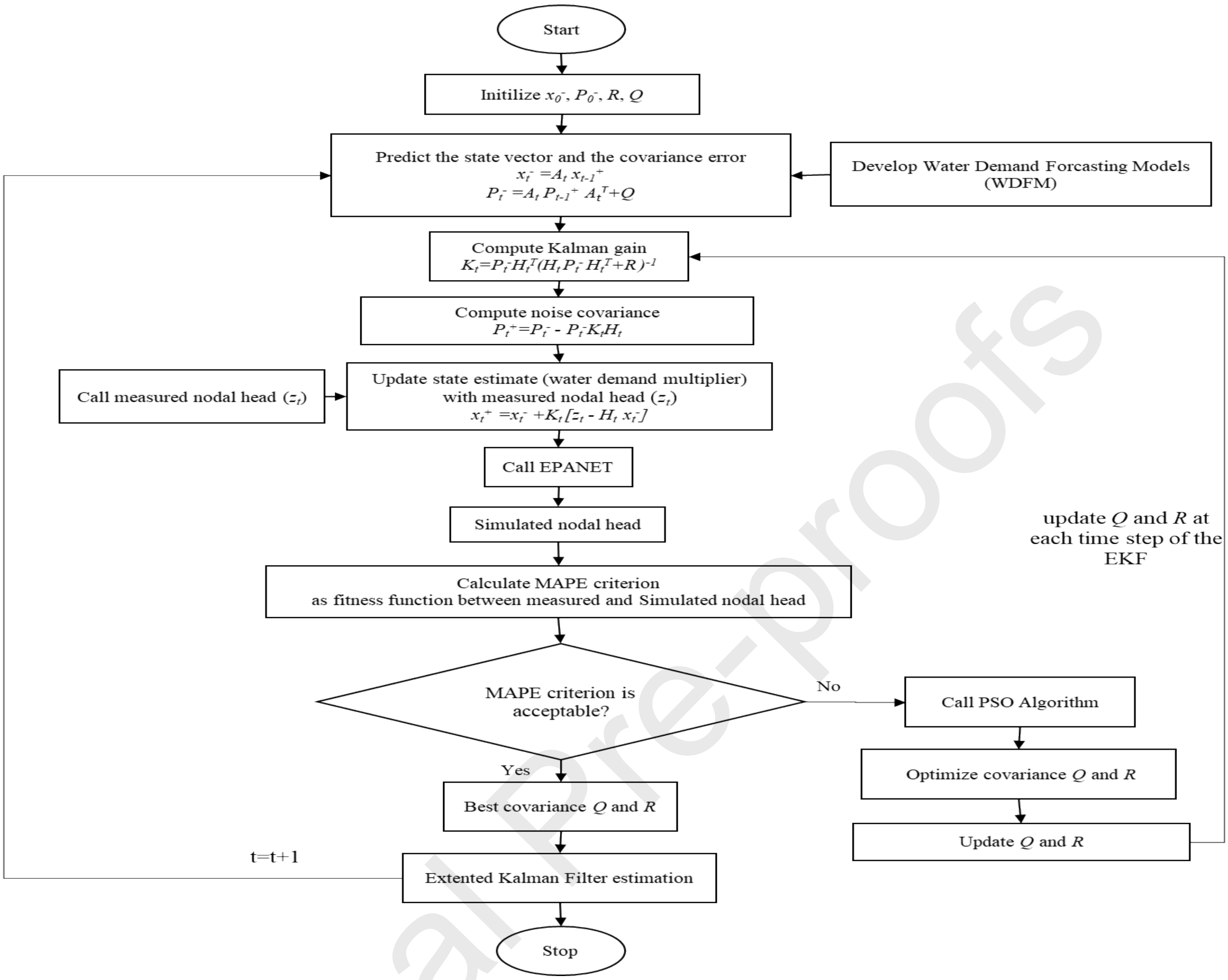


Fig. 2. The flowchart of the EKF-PSO algorithm

The data assimilation algorithms were implemented using the EPANET toolkit in MATLAB version R2018b. These algorithms were evaluated in two modified examples of the EPANET software, i.e., the Net1 and Net3 networks.

2.7. Case study 1: Net1

The features of the first case study were illustrated in Fig. 3 (Chu et al., 2021). This network has 8 nodes, 11 pipes, and 1 reservoir. Three hypothetical pressure gages were considered in Nodes 3, 5, and 7. A single demand pattern was considered as an actual pattern for all nodal demands that varies every 15 min (see Fig. 4). To create noisy measured pressure data, the network model was run with EPANET using the assumed demand pattern, and pressure values in these nodes were obtained. Then a random noise with a normal distribution ($\mu = 0$, $\sigma = 0.1$) was added to the simulated pressure values.

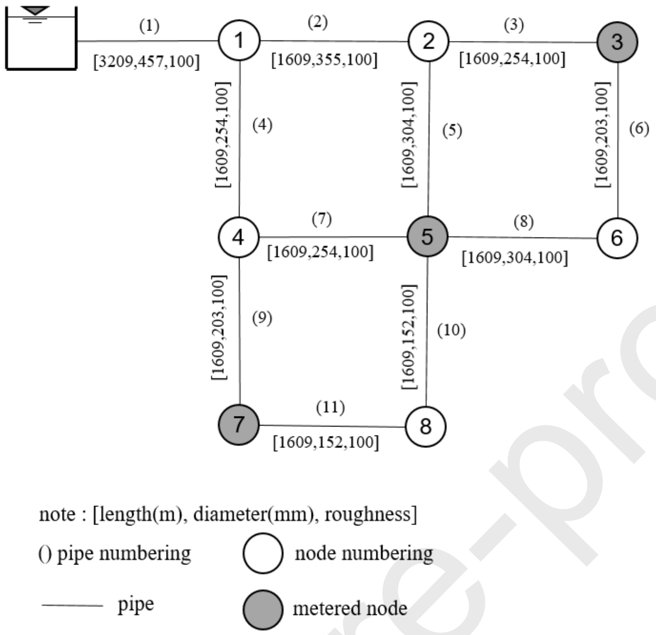


Fig. 3. Schematic view of Net1

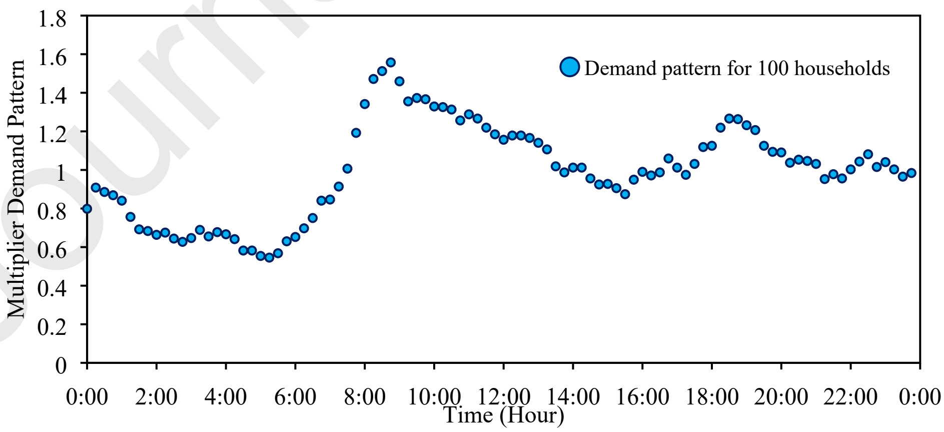


Fig. 4. Assumed demand pattern for Net1 (Do et al.,2017)

2.8. Case study 2- Net3

To assess the performance of the proposed method in a WDN with multiple demand patterns, the Net3 network was selected (see Fig. 5). This network consists of 92 nodes, 3 tanks, 2 reservoirs, 117 pipes, and 2 pumps. The nodal demands were categorized into four groups based on the magnitudes of the base demand: nodes with base demands less than 10 L/s are DMP[1]1, nodes with base demands from 10 L/s to 20 L/s are DMP2, nodes with base demands from 20 L/s to 30 L/s are DMP3 and nodes with base demands larger than 30 L/s are DMP4. The nodal pressure has been reported in 12 nodes and three measured tank levels were available (see Fig. 6). According to Do et al. (2017), four actual demand patterns have been obtained by adding a random noise with a normal distribution ($\mu = 0$, $\sigma = 0.15$). A set of pressure measurement data are simulated by EPANET for 48 hours. Then a random error ($\Delta_{measurement}$ = ±1.0 m.) is added to provide noisy nodal pressure.

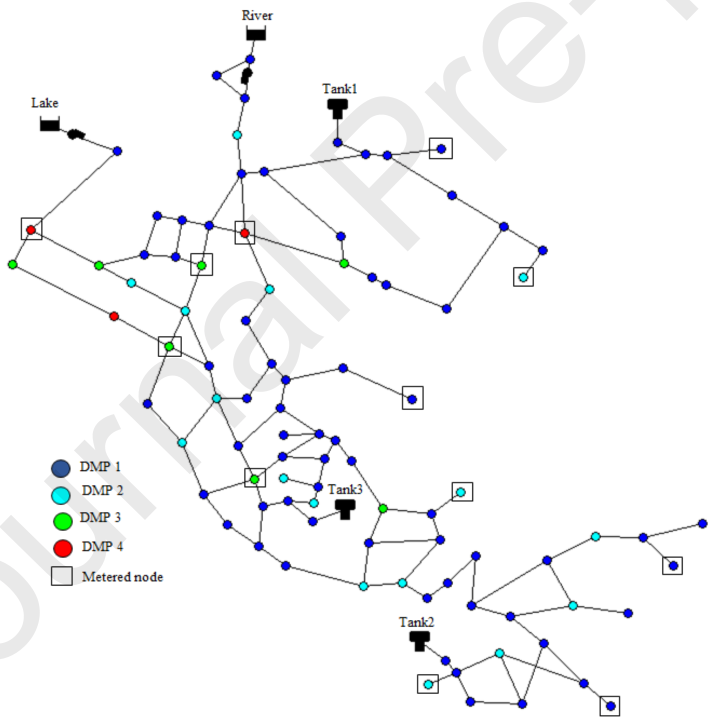


Fig. 5. Schematic view of Net 3

[1] Demand Multiplier Pattern (DMP)

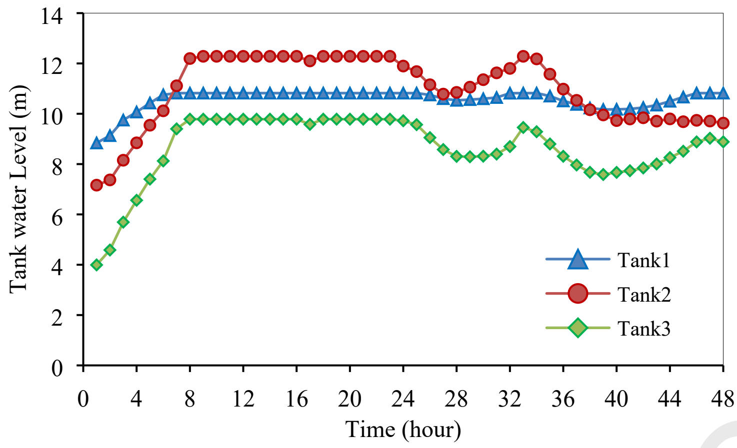


Fig. 6. The measured water level for three tanks

2.9. Model performance

The proposed models were evaluated with five different performance indices, including coefficient of determination (R^2), mean absolute percentage error (MAPE), root mean square error (RMSE), normalized root mean squared error (NRMSE), and Nash-Sutcliffe efficiency (NSE), which are calculated as follows (Chen et al.,2019):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \quad (22)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (23)$$

$$NRMSE = \frac{RMSE}{\overline{y}} \quad (24)$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (25)$$

where n represents the number of data, $\bar{y}$ is the data set average, $\hat{y}_i$ and y_i denote the predicted and the observed water demand pattern, respectively. The NSE value ranges from -∞ to 1. The NSE< 0 implies that the mean observed value is better than the predicted value.

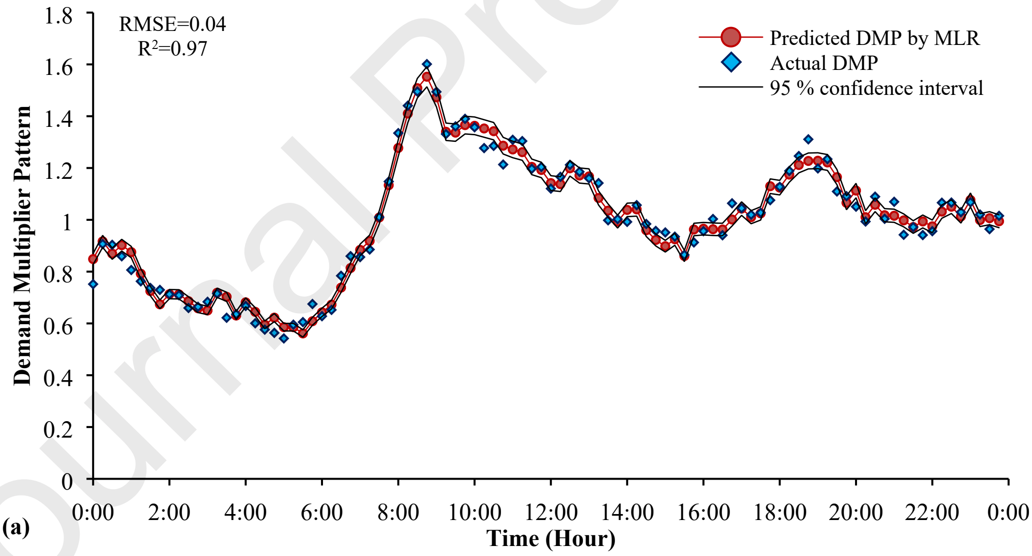
3. Results and Discussions

In this section, the results obtained for each case study are shown and discussed. For all simulations, an Intel (R) Core (TM) i7-8550U CPU 1.99 GHz processor with 8 GB of RAM on a 64-bit Windows operating system was used. The water demand forecasting models based on MLR and NLR for both case studies 1 and 2 are presented in Table 1. The predicted demand multipliers and the confidence intervals for both case studies have been plotted in Fig. 7 and Fig. 8. These figures confirm a good agreement between the actual and estimated demand patterns by MLR and NLR. However, 26, and 23% of actual demand multipliers are out of the range of 95 % confidence intervals for MLR and NLR, which shows the better performance of NLR. In Fig. 8 the forecasted demand multipliers approximately yield a good match with the actual demand multipliers for four demand patterns over 48 hours.

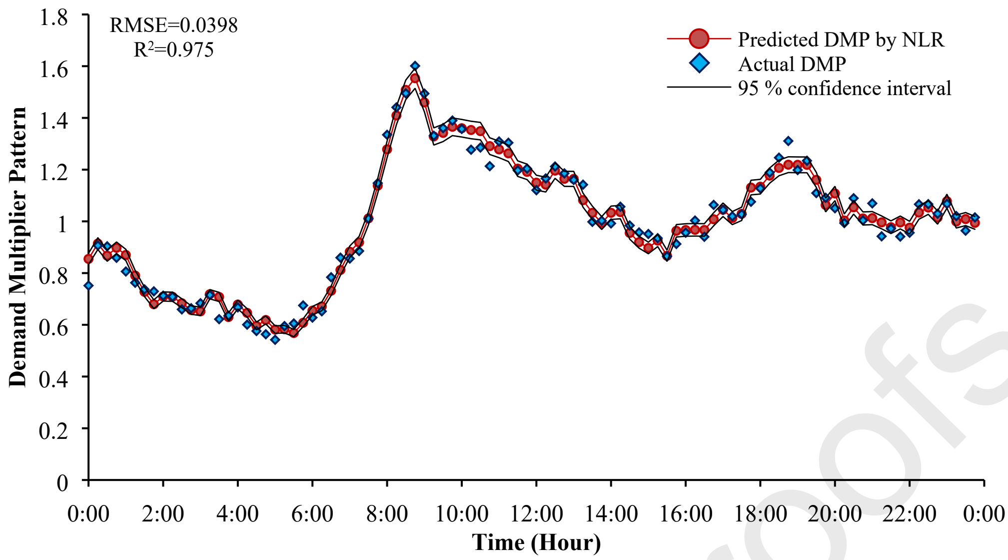
Table 1. Water demand forecasting models for both case studies 1 and 2

Net 1	DMP1	MLR	$d_t = 0.178 d_{t-1} + 0.003 d_{t-6} + 0.00015 d_{t-12} + 0.0045 d_{t-18} + 0.82 d_{t-24}$	(26)
		NLR	$d_t = 4.589 d_{t-1}^{0.035} + 1.926 d_{t-6}^{-0.016} - 8.834 d_{t-12}^{0.003} + 2.641 d_{t-18}^{-0.007} + 0.675 d_{t-24}^{1.192}$	(27)
Net 3	DMP1	MLR	$d_t = 0.113 d_{t-1} - 0.036 d_{t-6} + 0.008 d_{t-12} + 0.012 d_{t-18} + 0.029 d_{t-24} + 0.021 d_{t-30} + 0.005 d_{t-36} - 0.004 d_{t-42} + 0.849 d_{t-48}$	(28)
		NLR	$d_t = 3.119 d_{t-1}^{0.017} + 2.534 d_{t-6}^{-0.019} + 2.025 d_{t-12}^{-0.013} - 0.292 d_{t-18}^{0.034} - 0.987 d_{t-24}^{-0.018} + 1.261 d_{t-30}^{-0.0021} - 0.0000007 d_{t-36}^{-7.102} - 0.0583 d_{t-42}^{2.222} - 6.777 d_{t-48}^{-0.067}$	(29)
	DMP2	MLR	$d_t = 0.066 d_{t-1} - 0.033 d_{t-6} - 0.014 d_{t-12} - 0.004 d_{t-18} + 0.019 d_{t-24} + 0.012 d_{t-30} + 0.022 d_{t-36} + 0.019 d_{t-42} + 0.908 d_{t-48}$	(30)
		NLR	$d_t = -1.081 d_{t-1}^{-0.007} + 15.412 d_{t-6}^{-0.004} + 17.142 d_{t-12}^{-0.003} + 15.282 d_{t-18}^{-0.002} + 15.149 d_{t-24}^{-0.0009} - 0.041 d_{t-30}^{5.752} - 0.213 d_{t-36}^{25.925} - 5.8281 d_{t-42}^{0.0096} - 55.268 d_{t-48}^{-0.0093}$	(31)

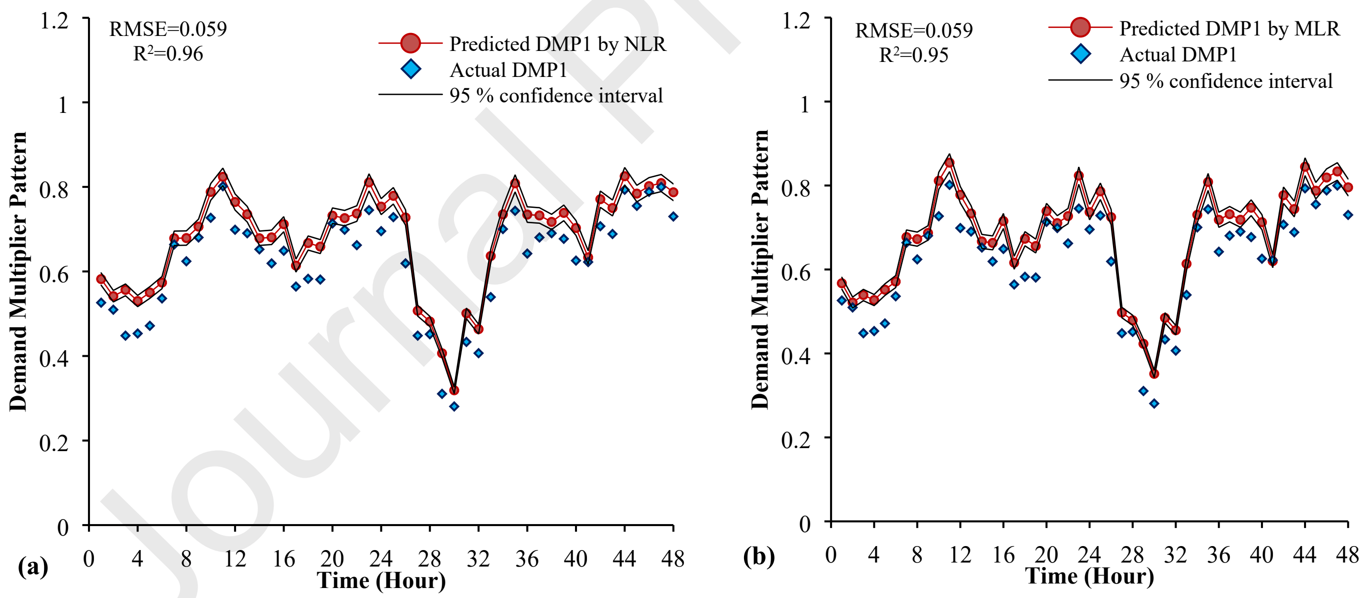
DMP3	MLR	$d_t = 0.004d_{t-1} + 0.009d_{t-6} - 0.005d_{t-12} - 0.0019d_{t-18} - 0.0017d_{t-24} + 0.0016d_{t-30} + 0.0033d_{t-36} - 0.0077d_{t-42} + 0.996d_{t-48}$	(32)
	NLR	$d_t = -6.199d_{t-1}^{-0.006} + 58.599d_{t-6}^{0.0006} - 0.956d_{t-12}^{-0.022} + 56.717d_{t-18}^{-0.0001} + 51.177d_{t-24}^{0.0017} + 57.735d_{t-30}^{-0.0001} + 59.186d_{t-36}^{0.0005} + 60.243d_{t-42}^{-0.0005} - 335.427d_{t-48}^{-0.0023}$	(33)
DMP4	MLR	$d_t = 0.004d_{t-1} - 0.002d_{t-6} + 0.002d_{t-12} + 0.007d_{t-18} + 0.0002d_{t-24} - 0.0003d_{t-30} - 0.002d_{t-36} + 0.009d_{t-42} + 0.996d_{t-48}$	(34)
	NLR	$d_t = 0.015d_{t-1}^{3.532} + 2.189d_{t-6}^{-0.0058} + 3.141d_{t-12}^{0.0641} + 1.247d_{t-18}^{-0.005} - 0.756d_{t-24}^{-0.0466} + 1.623d_{t-30}^{0.027} - 0.053d_{t-36}^{1.923} + 1.922d_{t-42}^{0.047} - 8.11d_{t-48}^{-0.088}$	(35)



(a)



(b)**Fig. 7.** Predicted demand patterns for Net 1 using (a) MLR and (b) NLR



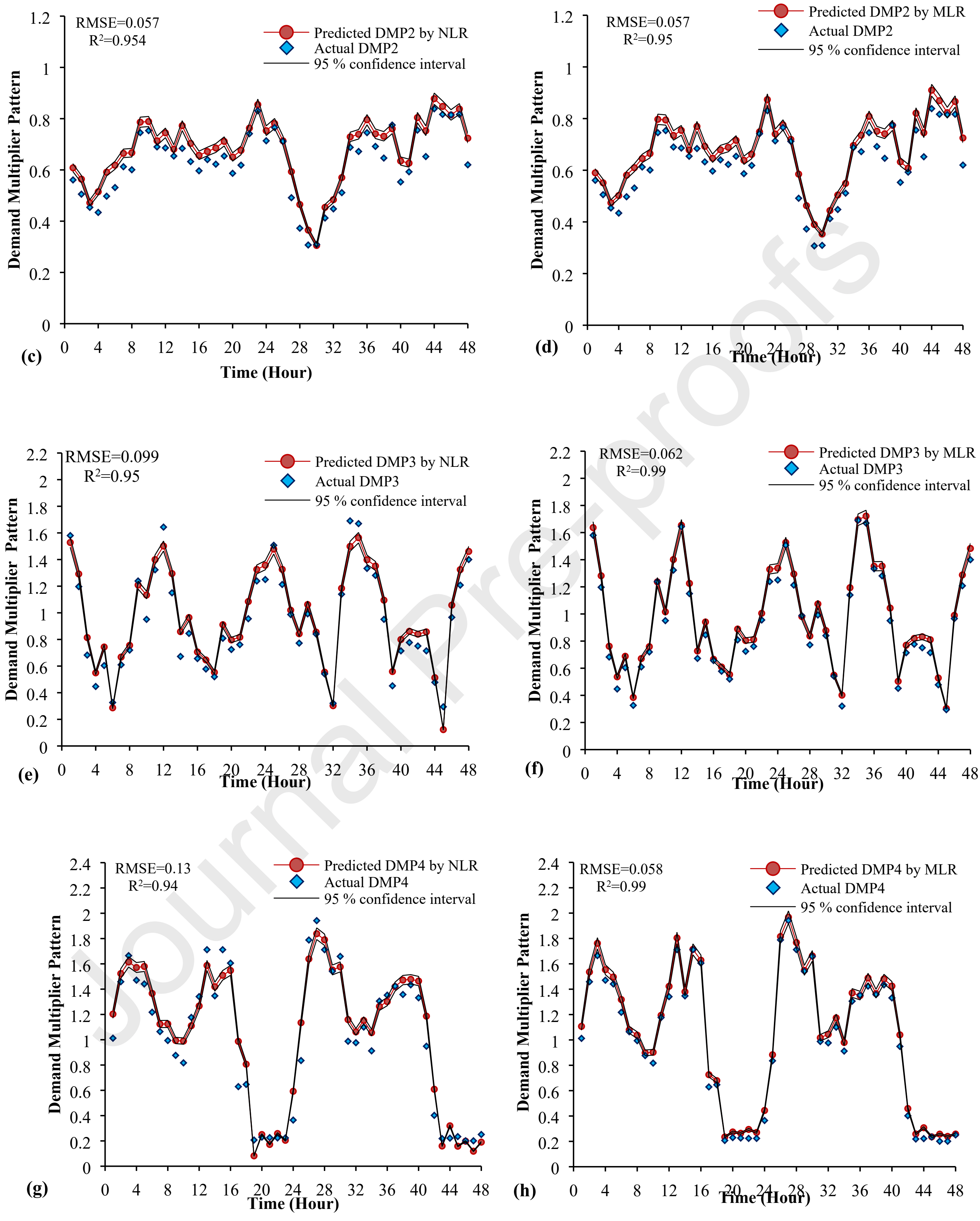

RMSE=0.057
R²=0.954
Predicted DMP2 by NLR
Actual DMP2
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(c)
RMSE=0.057
R²=0.95
Predicted DMP2 by MLR
Actual DMP2
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(d)
RMSE=0.099
R²=0.95
Predicted DMP3 by NLR
Actual DMP3
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(e)
RMSE=0.062
R²=0.99
Predicted DMP3 by MLR
Actual DMP3
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(f)
RMSE=0.13
R²=0.94
Predicted DMP4 by NLR
Actual DMP4
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(g)
RMSE=0.058
R²=0.99
Predicted DMP4 by MLR
Actual DMP4
95 % confidence interval
Demand Multiplier Pattern
Time (Hour)
(h)

Fig. 8. Four predicted demand patterns for Net 3 using MLR and NLR and their confidence interval

The values of noise covariance matrices *Q* and *R* have an important effect on the KF and EKF performance. The manually varied covariance matrices *Q* and *R* of EKF with their corresponding performances (MAPE and RMSEs) are presented in Table 2. The manual adjustment of *Q* and *R* values is very time-consuming and improper values of these covariances lead to imprecise estimates (cases 1 and 5 in Table 2). It is difficult to figure out a relationship between the value of the covariance matrices and the best-simulated demand pattern. Therefore, to achieve the optimal covariance matrices, the PSO algorithm was applied. The results of KF-PSO and EKF-PSO are shown in Table 3.

Table 2. The EKF performances for different *Q* and *R* values

Network	Case	*Q*	*R*	MAPE (%)	RMSE	Qualification
Net 1	1	1e-8	0.01	12.375	0.196	Very poor
	2	1e-6	0.01	11.527	0.179	Poor
	3	1e-7	0.001	8.329	0.127	Good
	4	0.001	0.01	6.15	0.076	Very Good
Net 3	5	1e-9	1e-8	662.525	2.722	Very poor
	6	0.001	0.01	61.876	0.390	Poor
	7	0.0001	0.001	51.694	0.34	Good
	8	1e-6	1e-9	49.891	0.293	Very Good

Table 3. Comparison of KF-PSO and EKF-PSO results for the optimal R and Q

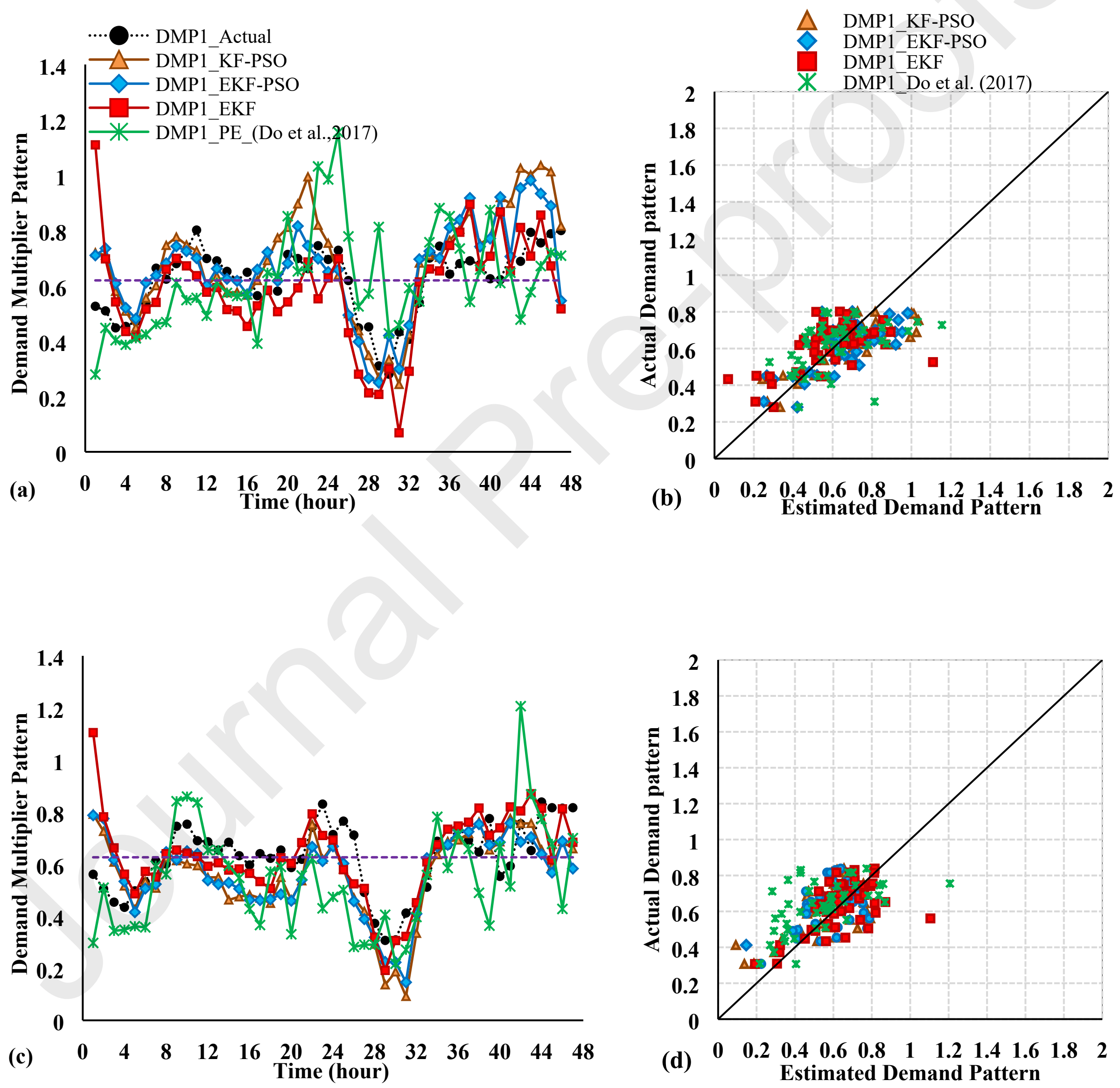
Network	Methods	MAPE (%)	RMSE	NSE	Optimization time (s)
Net 1	KF-PSO	5.945	0.064	0.93	638
	EKF-PSO	5.927	0.063	0.93	639
Net 3	KF-PSO	33.869	0.229	0.14	8401
	EKF-PSO	25.822	0.198	0.3	8538

As seen in Table 3, the EKF-PSO showed better performance, especially for Net 3. The Nash-Sutcliffe efficiency coefficient (NSE) is obtained at 0.93 and 0.14 for Net 1 and Net 3 respectively. The MAPE of EKF-PSO (25.02%) is smaller than KF-PSO (30.46%) while the execution time of KF-PSO is less than EKF-PSO. Since the EKF estimates the nonlinear states, the EKF computational cost is a little more than the KF. For example, in Net 3, the time required by the EKF algorithm at each step was 38 seconds. For the EKF-PSO algorithm, the computation time is related to the population of the swarm size and the iterations. In this study, the swarm size and iteration were set at 30 and 20, respectively, and the computation time of EKF-PSO at each time step was 177 seconds. Although the hybrid of PSO into EKF slows the computational speed, it improves the performance of state estimation. Since the time interval between two successive head measurements was one hour, there was enough time to optimize the covariance matrices in each time step. This optimization can be implemented even in smaller time steps. Moreover, The optimization time depends on the complexity of the water distribution network. In most real water distribution networks, the time interval between two pressure measurements ranges from 30 minutes to one hour. So, there was enough time to optimize the covariance matrices in each time step. On the other hand, the manual adjustment of EKF needs great effort and it is not possible to easily find the best values of the covariance matrices by trial and error method. It is noteworthy that KF-PSO gives an improper result for Net 3 due to larger nonlinearity compared to Net 1. The setting parameters of the PSO algorithm were selected by sensitivity analysis and they have been listed in Table 4.

Table 4. Particle swarm optimization algorithm parameters.

Iteration	20
Swarm size	30
Inertia weight (0–1)	0.729

Inertia weight damping ratio	0.99
Personal learning coefficient (0–2)	2
Global learning coefficient (0–2)	2



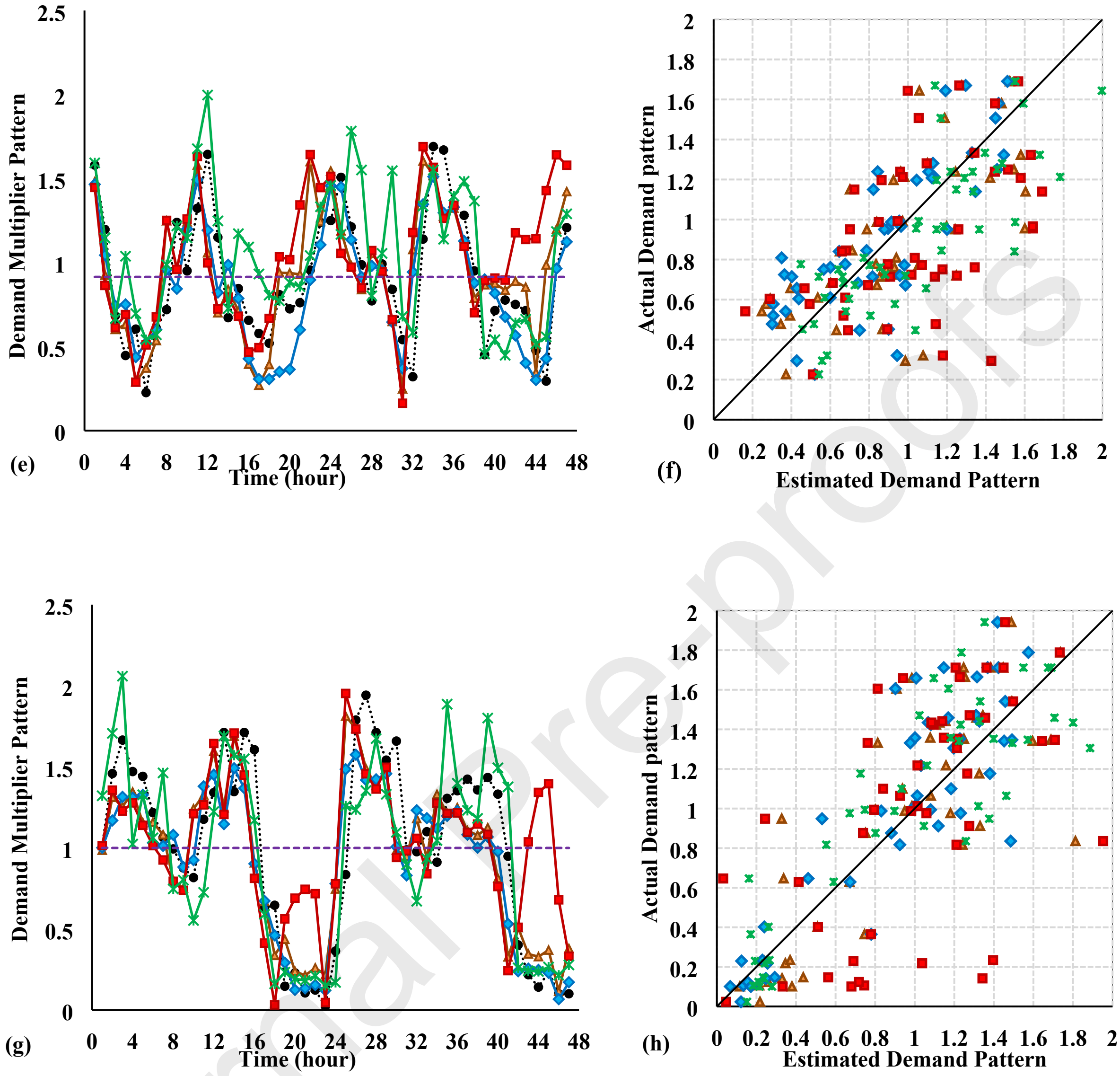


Fig. 9. Hourly estimated demand pattern curves (a, c, e, g) and Scatter grams (b, d, f, h) for estimated demand pattern in Net 3.

In the right graphs of Fig. 9, the scattergrams of the four estimated demand patterns by all methods were compared with Do et al. (2017) results. The results demonstrate that the EKF-PSO method provided good accuracy of demand multiplier pattern estimation compared to the conventional EKF method and KF-PSO method. On the other hand, the estimated demand patterns derived from the particle filter (PF) method in Do et al. (2017) deviated significantly from the actual value. The estimated DMPs were derived from all methods by using a measurement error of ±1.0 m.

Table 5. Error indices for different methods in Net3

Methods	EKF					EKF-PSO				
DMPs	1	2	3	4	Ave.*	1	2	3	4	Ave.
MAPE (%)	20.556	16.53	50	112.47	49.89	17.788	20.145	29.992	35.181	25.759
RMSE	0.157	0.136	0.399	0.480	0.293	0.128	0.14	0.241	0.281	0.198
NRMSE	0.254	0.218	0.436	0.480	0.347	0.207	0.224	0.264	0.281	0.244
NSE	-0.590	-0.037	-0.142	0.269	-0.125	-0.045	-0.088	0.583	0.75	0.3
R^2	0.35	0.34	0.23	0.31	0.307	0.49	0.36	0.63	0.78	0.565
Methods	KF-PSO					Do et al. (2017)				
DMPs	1	2	3	4	Ave.	1	2	3	4	Ave.
MAPE (%)	19.015	21.025	36.721	58.335	33.774	23.307	23.051	29.449	43.270	29.769
RMSE	0.144	0.142	0.305	0.325	0.229	0.169	0.186	0.277	0.285	0.229
NRMSE	0.233	0.226	0.333	0.325	0.279	0.273	0.297	0.303	0.285	0.29
NSE	-0.326	-0.113	0.334	0.665	0.14	-0.821	-0.919	0.448	0.744	-0.137
R^2	0.59	0.418	0.447	0.67	0.531	0.19	0.329	0.63	0.76	0.477

*Average

Table 5 shows the Performance indicators for four methods including the EKF, the KF-PSO, the EKF-PSO, and the particle filter (PF) used by Do et al. (2017). The root mean square errors between the estimated demand multipliers and the actual demand multipliers indicate that the EKF-PSO method (RMSE=0.198) obtained relatively better results than the EKF method (RMSE=0.480), the KF-PSO method (RMSE=0.229) and PF method (RMSE=0.229). The NSE values for DMP1 and DMP2 (estimated by all models) are negative but for DMP3 and DMP4 are positive. The DMP3 and DMP4 have a significantly greater mean and variance (Figs.

9 (e) and 9 (g)). To properly compare the methods' performance, each DMP was separately evaluated, i.e., For DMP4, the EKF-PSO method has the highest NSE (0.75), R^2 (0.78), and the lowest RMSE (0.281). Thus, the EKF-PSO method was the best. The performances of KF-PSO, EKF, and PF (Do et al., 2017) are all inferior to the EKF-PSO method.

To validate the correctness of the results, the estimated demand multipliers were substituted as inputs in EPANET then simulated nodal pressures were compared with the measured value. As depicted in Fig. 10, the EKF-PSO method can reasonably capture the measured pressure but the curve of the EKF method deviates significantly from the measured value, especially from t=40 to t=46 hours. On the other hand, the base demand of node 26 and node 19 are 8.85 L/s and 36.48 L/s respectively. This could be one of the reasons for the differences between simulated and measured pressure at node 19. Any small changes in the water demand pattern can cause a large change in the demand, which in turn will increase the pressure. As shown in Table 6, in terms of the assessment factors, RMSE, NRMSE, NSE, and R2, the EKF-PSO produced better results compared with the other models. The overall results proved that the simulated pressure corresponds very well with the measured pressure

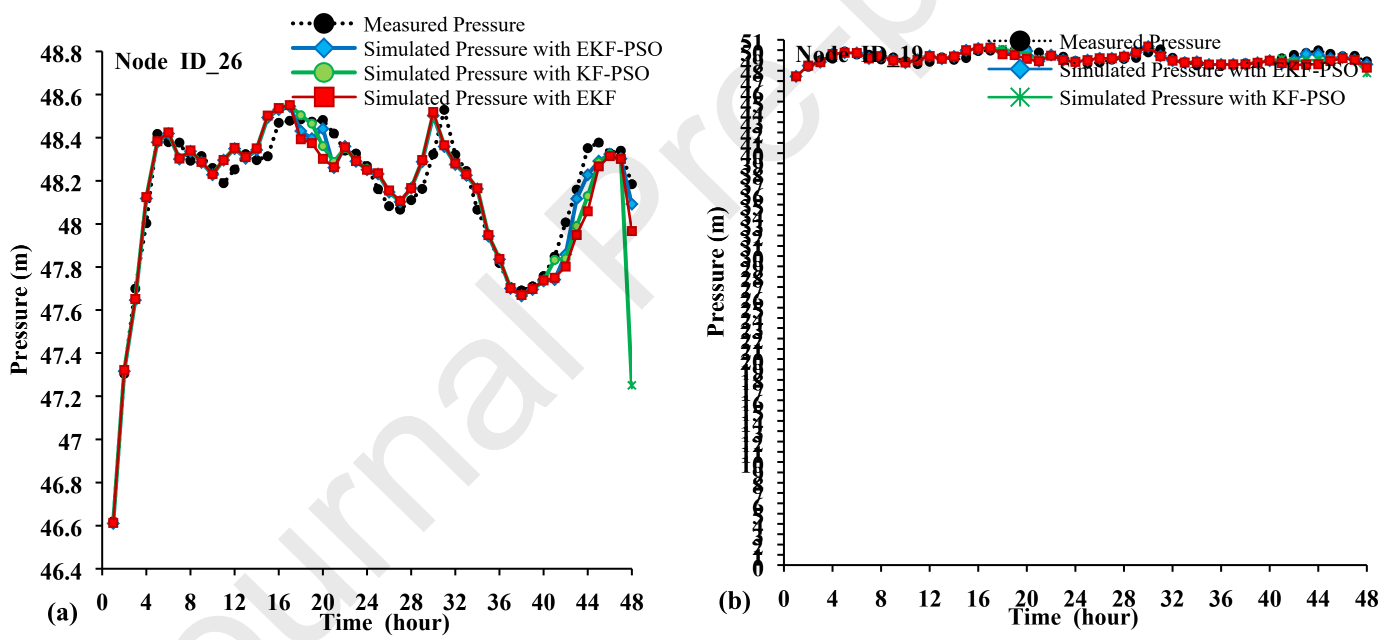


Fig. 10. Comparison of simulated nodal pressures with those measured in Net 3

Table 6. Error indices for simulated nodal pressures with those measured in Net3

Node	19	26

Methods	EKF-PSO	KF-PSO	EKF	EKF-PSO	KF-PSO	EKF
MAPE (%)	0.55	0.62	0.72	0.13	0.17	0.16
RMSE	0.33	0.39	0.46	0.07	0.15	0.10
NRMSE	0.006	0.008	0.009	0.001	0.003	0.002
NSE	0.69	0.57	0.40	0.96	0.86	0.94
R^2	0.71	0.62	0.50	0.97	0.87	0.94

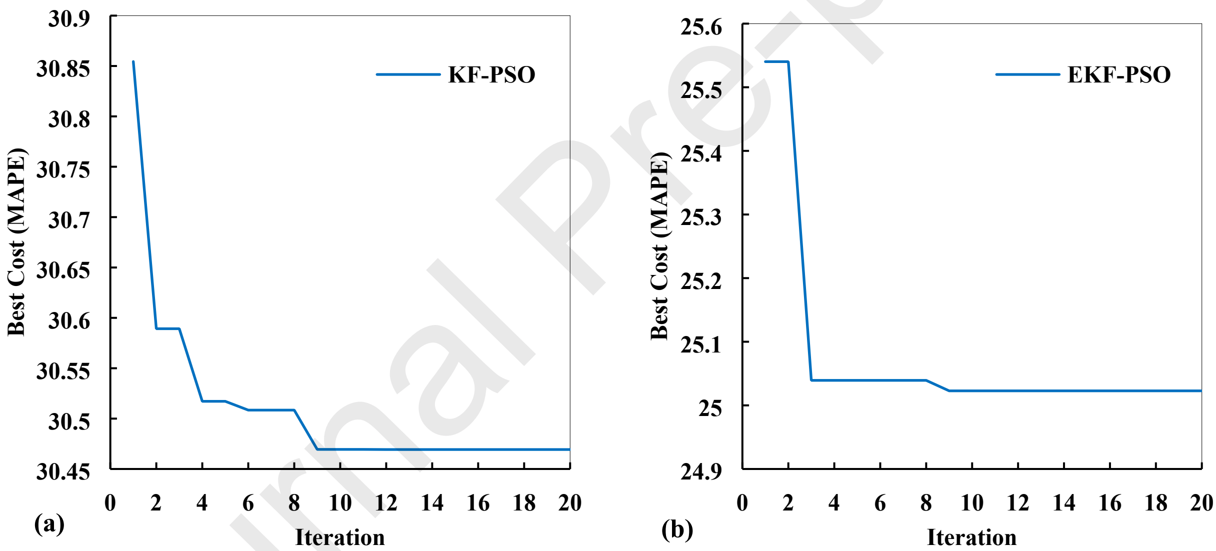


Fig. 11. Evaluation of best cost of proposed methods

Fig. 11 represents the convergence rate of proposed methods in reaching the minimum MAPE in Net3. The EKF-PSO method has achieved the optimum solution (MAPE=25.02%) with 9 iterations. Fig. 12 shows the time variation of the noise covariance matrices *Q* and *R*. As can be seen, in the early time step, the value of covariance matrices *Q* and *R* shows hourly fluctuation, but it will gradually tend to a certain value. This shows that in the long run, the EKF-PSO computation cost is gradually decreased.

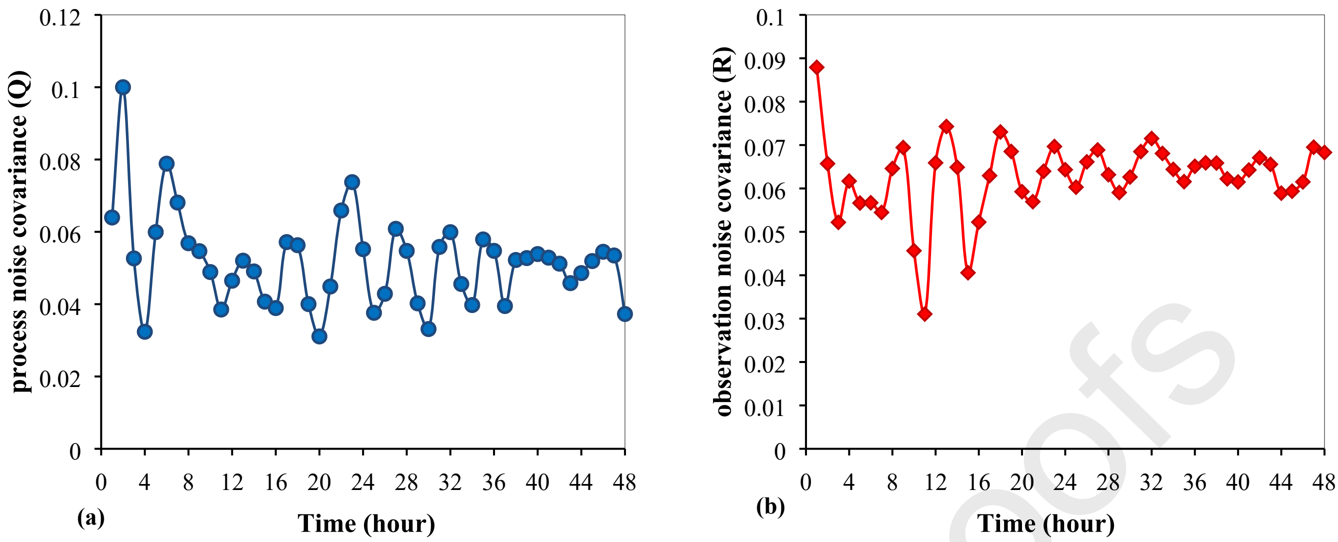


Fig. 12. Time variation of the noise covariance matrices (a)Q and (b) R

4. Conclusions

In this study, a predictor-corrector approach has been adopted by a hybrid technique. Multiple Linear Regression (MLR) and Nonlinear Regression (NLR) models were applied to predict water demand multipliers at each time step with past historical data. A series of online pressure observations were used to correct the prediction. This paper has developed a combination of the Kalman filter methods with the particle swarm optimization (PSO) algorithm to achieve high performance in the estimation of real-time demand multipliers. The output of Kalman filter methods strongly depends on the state noise covariance matrix (Q) and the measurement noise covariance matrix (R). The PSO algorithm was applied for the optimization of these matrices at each time step of the EKF. Then the optimal values of noise covariance matrices are used in the real-time water demand multiplier estimation loop. The performance of the proposed methodologies was tested and validated in two WDNs, two modified examples of the EPANET software, the Net1 and Net3 networks. The results indicated that the KF-PSO method works poorly due to the nonlinearities of hydraulic systems. The comparison of the results demonstrated that obtained estimations by the EKF-PSO method are more precise than the conventional EKF method.

Declaration of competing interest

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

References

Abu-Mahfouz, A. M., Hamam, Y., Page, P. R., Adedeji, K. B., Anele, A. O., & Todini, E. (2019). Real-time dynamic hydraulic model of water distribution networks. *Water (Switzerland)*, *11*(3), 470. https://doi.org/10.3390/w11030470

Akhlaghi, S., Zhou, N., & Huang, Z. (2017, July). Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation. In *2017 IEEE power & energy society general meeting* (pp. 1-5). IEEE. https://doi.org/10.1109/PESGM.2017.8273755

Assumaning, G. A., & Chang, S.-Y. (2016). Application of Sequential Data-Assimilation Techniques in Groundwater Contaminant Transport Modeling. *Journal of Environmental Engineering*, *142*(2), 04015073. https://doi.org/10.1061/(asce)ee.1943-7870.0001034

Atashpaz-Gargari, E., & Lucas, C. (2007, September). Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition. In *2007 IEEE congress on evolutionary computation* (pp. 4661-4667). Ieee. https://doi.org/10.1109/CEC.2007.4425083

Attarzadeh, F. (2023). *Real-time Calibration of demand pattern and Hazen-Williams coefficient in water distribution networks*. [Doctoral dissertation, Ferdowsi University of Mashhad, Mashhhad, Iran].

Attarzadeh, F., Ziaei, A. N., Davari, K., & Fallah Choulabi, E. (2022). *Comparison of Five Evolutionary Algorithms for Calibration of Water Distribution Networks*. Journal of Hydraulics, 17(2), 21-45. https://doi.org/10.30482/jhyd.2021.295023.1539

Babu, K. J., & Vijayalakshmi, D. P. (2013). Self-adaptive PSO-GA hybrid model for combinatorial water distribution network design. *Journal of pipeline systems engineering and practice*, *4*(1), 57-67. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000113

Blood, E. A., Krogh, B. H., & Ilic, M. D. (2008, July). Electric power system static state estimation through Kalman filtering and load forecasting. In *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century* (pp. 1-6). IEEE. https://doi.org/10.1109/PES.2008.4596742

Chen, G., Long, T., Bai, Y., & Zhang, J. (2019). A forecasting framework based on Kalman filter integrated multivariate local polynomial regression: application to urban water demand. *Neural Processing Letters*, *50*(1), 497-513. https://doi.org/10.1007/s11063-019-10001-3

Chen, J., Hu, S., Ye, Y., Huang, H., Langari, R., & Tang, C. (2021). A cascaded scheme for high-performance estimation of vehicle states. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, *235*(8), 2101-2113. https://doi.org/10.1177/0954407021993240

Chu, S., Zhang, T., Yu, T., Wang, Q. J., & Shao, Y. (2021). A noise adaptive approach for nodal water demand estimation in water distribution systems. *Water Research*, *192*, Article 116837. https://doi.org/10.1016/j.watres.2021.116837

Dini, M., & Tabesh, M. (2014). A new method for simultaneous calibration of demand pattern and Hazen-Williams coefficients in water distribution systems. *Water Resources Management*, *28*, 2021-2034. https://doi.org/10.1007/s11269-014-0592-4

Do, N. C., Simpson, A. R., Deuerlein, J. W., & Piller, O. (2016). Calibration of water demand multipliers in water distribution systems using genetic algorithms. *Journal of water resources planning and management*, *142*(11), 04016044. https://doi.org/10.1061/(asce)wr.1943-5452.0000691

Do, N. C., Simpson, A., Deuerlein, J., & Piller, O. (2017). Particle filter-based model for online estimation of demand multipliers in water distribution systems under uncertainty. *Journal of Water Resources Planning and Management*, *143*(11), 04017065-15. https://doi.org/10.1061/(asce)wr.1943-5452.0000841

Dorigo, M., Maniezzo, V., & Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *26*(1), 29-41. https://doi.org/10.1109/3477.484436

Eberhart, R., & Kennedy, J. (1995, October). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the sixth international symposium on micro machine and human science* (pp. 39-43). Ieee. https://doi.org/10.1109/mhs.1995.494215

Garzón, A., Kapelan, Z., Langeveld, J., & Taormina, R. (2022). Machine Learning-Based Surrogate Modeling for Urban Water Networks: Review and Future Research Directions. *Water Resources Research*, 58(5), e2021WR031808. https://doi.org/10.1029/2021WR031808

Gupta, R., Bhave, P. R. (2006). *Analysis of Water Distribution Networks*. India: Alpha Science International.

Jung, D., Choi, Y. H., & Kim, J. H. (2016). Optimal node grouping for water distribution system demand estimation. *Water*, *8*(4), 160. https://doi.org/10.3390/w8040160

Jatoth, R. K., & Kumar, T. K. (2009). Swarm intelligence based tuning of unscented Kalman filter for bearings only tracking. *International Journal of Recent Trends in Engineering*, *2*(5), 177. https://doi.org/10.1109/ARTCom.2009.109

Jung, B. S., & Karney, B. W. (2008). Systematic exploration of pipeline network calibration using transients. *Journal of Hydraulic Research*, *46*(sup1), 129-137. https://doi.org/10.1080/00221686.2008.9521947

Kaba, A., & Kıyak, E. (2020). Optimizing a Kalman filter with an evolutionary algorithm for nonlinear quad rotor attitude dynamics. Journal of Computational Science, 39, Article 101051. https://doi.org/10.1016/j.jocs.2019.101051

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering, Transactions of the ASME*, *82*(1). https://doi.org/10.1115/1.3662552

Kang, B., Yang, H., Lee, K., & Choe, J. (2017). Ensemble Kalman filter with principal component analysis assisted sampling for channelized reservoir characterization. *Journal of Energy Resources Technology*, *139*(3). https://doi.org/10.1115/1.4035747

Kang, D., & Lansey, K. (2009). Real-time demand estimation and confidence limit analysis for water distribution systems. *Journal of Hydraulic Engineering*, *135*(10), 825-837. https://doi.org/10.1061/(asce)hy.1943-7900.0000086

Kurtz, W., Lapin, A., Schilling, O. S., Tang, Q., Schiller, E., Braun, T., ... & Brunner, P. (2017). Integrating hydrological modelling, data assimilation and cloud computing for real-time management of water resources. *Environmental modelling & software*, *93*, 418-435. https://doi.org/10.1016/j.envsoft.2017.03.011

Laamari, Y., Chafaa, K., & Athamena, B. (2015). Particle swarm optimization of an extended Kalman filter for speed and rotor flux estimation of an induction motor drive. *Electrical Engineering*, *97*(2), 129-138. https://doi.org/10.1007/s00202-014-0322-1

Li, X. L., Lü, H., Horton, R., An, T., & Yu, Z. (2014). Real-time flood forecast using the coupling support vector machine and data assimilation method. *Journal of Hydroinformatics*, *16*(5), 973-988. https://doi.org/10.2166/hydro.2013.075

Liu, D., Yu, Z. B., & Hai-shen, L. (2010). Data assimilation using support vector machines and ensemble Kalman filter for multi-layer soil moisture prediction. *Water Science and Engineering*, *3*(4), 361-377. https://doi.org/10.3882/j.issn.1674-2370.2010.04.001

Liu, K., Huang, G., Šimůnek, J., Xu, X., Xiong, Y., & Huang, Q. (2021). Comparison of ensemble data assimilation methods for the estimation of time-varying soil hydraulic parameters. *Journal of Hydrology*, *594*, Article 125729. https://doi.org/10.1016/j.jhydrol.2020.125729

Ma, L., Cheng, S., & Shi, Y. (2020). Enhancing learning efficiency of brain storm optimization via orthogonal learning design. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *51*(11), 6723-6742. https://doi.org/10.1109/TSMC.2020.2963943

Maheshwari, A., & Nageswari, S. (2022). Effect of Noise Covariance Matrices on State of Charge Estimation Using Extended Kalman Filter. *IETE Journal of Research*, 1-12. https://doi.org/10.1080/03772063.2022.2055657

Mehrabian, A. R., & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. *Ecological informatics*, *1*(4), 355-366. https://doi.org/10.1016/j.ecoinf.2006.07.003

Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in engineering software*, *69*, 46-61. https://doi.org/10.1016/j.advengsoft.2013.12.007

Moghaddam, A., Alizadeh, A., Faridhosseini, A., Ziaei, A. N., & Heravi, D. F. (2018). Optimal design of water distribution networks using simple modifed particle swarm optimization approach. *Desalination and water treatment*, *104*, 99-110. https://doi.org/10.5004/dwt.2018.21911

Mohamed, A. H., & Schwarz, K. P. (1999). Adaptive Kalman filtering for INS/GPS. *Journal of geodesy*, *73*, 193-203. https://doi.org/10.1007/s001900050236

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

Nasseri, M., Moeini, A., & Tabesh, M. (2011). Forecasting monthly urban water demand using extended Kalman filter and genetic programming. *Expert Systems with Applications*, *38*(6), 7387-7395. https://doi.org/10.1016/j.eswa.2010.12.087

Netto, M., Zhao, J., & Mili, L. (2016, July). A robust extended Kalman filter for power system dynamic state estimation using PMU measurements. In *2016 IEEE Power and Energy Society General Meeting (PESGM)* (pp. 1-5). IEEE. https://doi.org/10.1109/PESGM.2016.7741374

Okeya, I., Kapelan, Z., Hutton, C., & Naga, D. (2014). Online modelling of water distribution system using data assimilation. *Procedia Engineering*, *70*, 1261-1270. https://doi.org/10.1016/j.proeng.2014.02.139.

Passino, K. M. (2012). Bacterial foraging optimization. In *Innovations and Developments of Swarm Intelligence Applications* (pp. 219-234). IGI Global. https://doi.org/10.4018/978-1-4666-1592-2.ch013

Pelosi, A., Medina, H., Van den Bergh, J., Vannitsem, S., & Chirico, G. B. (2017). Adaptive Kalman filtering for postprocessing ensemble numerical weather predictions. *Monthly Weather Review*, *145*(12), 4837-4854. https://doi.org/10.1175/MWR-D-17-0084.1

Preis, A., Whittle, A., & Ostfeld, A. (2009). On-line hydraulic state prediction for water distribution systems. https://doi.org/10.1061/41036(342)32

Rajesh, M., & Rehana, S. (2021). Prediction of river water temperature using machine learning algorithms: a tropical river system of India. *Journal of Hydroinformatics*, *23*(3), 605-626. https://doi.org/10.2166/HYDRO.2021.121

Rossi, C., Falcomer, C., Biondani, L., & Pontara, D. (2022). Genetically Optimized Extended Kalman Filter for State of Health Estimation Based on Li-Ion Batteries Parameters. Energies, 15(9), 3404. https://doi.org/10.3390/en15166006

Salloom, T., Kaynak, O., & He, W. (2021). A novel deep neural network architecture for real-time water demand forecasting. Journal of Hydrology, 599, 126353. https://doi.org/10.1016/j.jhydrol.2021.126353

Sen, Z., Altunkaynak, A., & Özger, M. (2004). Sediment concentration and its prediction by perceptron Kalman filtering procedure. *Journal of Hydraulic Engineering*, *130*(8), 816-826. https://doi.org/10.1061/(asce)0733-9429(2004)130:8(816)

Shang, F., Uber, J. G., van Bloemen Waanders, B. G., Boccelli, D., & Janke, R. (2008). Real time water demand estimation in water distribution system. In *Water Distribution Systems Analysis Symposium 2006* (pp. 1-14). https://doi.org/10.1061/40941(247)95

Shi, K. L., Chan, T. F., Wong, Y. K., & Ho, S. L. (2002). Speed estimation of an induction motor drive using an optimized extended Kalman filter. *IEEE Transactions on Industrial Electronics*, *49*(1), 124-133. https://doi.org/10.1109/41.982256

Shi, Y. (2011). Brain storm optimization algorithm. In *Advances in Swarm Intelligence: Second International Conference, ICSI 2011, Chongqing, China, June 12-15, 2011, Proceedings, Part I 2* (pp. 303-309). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21515-5_36

Song, L., & Rahmat-Samii, Y. (2021, August). Hybridizing Particle Swarm and Brain Storm Optimizations for Applications in Electromagnetics. In 2021 XXXIVth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS) (pp. 1-4). IEEE. https://doi.org/10.23919/URSIGASS51995.2021.9560569

Sun, B., Zhou, Y., Wang, J., & Zhang, W. (2021). A new PC-PSO algorithm for Bayesian network structure learning with structure priors. Expert Systems with Applications, 184, 115237. https://doi.org/10.1016/j.eswa.2021.115237

Sun, L., Seidou, O., Nistor, I., & Liu, K. (2016). Review of the Kalman-type hydrological data assimilation. *Hydrological Sciences Journal*, *61*(13), 2348-2366. https://doi.org/10.1080/02626667.2015.1127376

Vassiljev, A., & Koppel, T. (2015). Estimation of real-time demands on the basis of pressure measurements by different optimization methods. *Advances in Engineering Software*, *80*, 67-71. https://doi.org/10.1016/j.advengsoft.2014.09.023

Wang, D., Chen, Y., & Cai, X. (2009). State and parameter estimation of hydrologic models using the constrained ensemble Kalman filter. *Water resources research*, *45*(11). https://doi.org/10.1029/2008WR007401

Wang, W., & Mu, J. (2019). State of charge estimation for lithium-ion battery in electric vehicle based on Kalman filter considering model error. *Ieee Access*, *7*, 29223-29235. https://doi.org/10.1109/ACCESS.2019.2895377

Yarat, S., Senan, S., & Orman, Z. (2021). A Comparative Study on PSO with Other Metaheuristic Methods. *Applying Particle Swarm Optimization*, 49-72. https://doi.org/10.1007/978-3-030-70281-6_4

Yu, Z., Fu, X., Lü, H., Luo, L., Liu, D., Ju, Q., ... & Wang, Z. (2014). Evaluating ensemble Kalman, particle, and ensemble particle filters through soil temperature prediction. *Journal of Hydrologic Engineering*, *19*(12), 04014027. https://doi.org/10.1061/(asce)he.1943-5584.0000976

Zaji, A. H., & Bonakdari, H. (2014). Performance evaluation of two different neural network and particle swarm optimization methods for prediction of discharge capacity of modified triangular side weirs. *Flow Measurement and Instrumentation*, *40*, 149-156. https://doi.org/10.1016/j.flowmeasinst.2014.10.002

Zhang, Q., Yang, J., Zhang, W., Kumar, M., Liu, J., Liu, J., & Li, X. (2023). Deep fuzzy mapping nonparametric model for real-time demand estimation in water distribution systems: A new perspective. Water Research, 120145. https://doi.org/10.1016/j.watres.2023.120145

Zhou, X., Xu, W., Xin, K., Yan, H., & Tao, T. (2018). Self-Adaptive Calibration of Real-Time Demand and Roughness of Water Distribution Systems. Water Resources Research, 54(8), 5536-5550. https://doi.org/10.1029/2017WR022147