IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

SPECIAL SECTION ON ADVANCED SENSOR TECHNOLOGIES ON WATER MONITORING AND MODELING

Received July 25, 2018, accepted September 14, 2018, date of publication September 28, 2018, date of current version October 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2872506

Sensor Failure Detection and Faulty Data Accommodation Approach for Instrumented Wastewater Infrastructures

KARTHICK THIYAGARAJAN, (Member, IEEE), SARATH KODAGODA, (Member, IEEE), LINH VAN NGUYEN, (Member, IEEE), AND RAVINDRA RANASINGHE, (Member, IEEE)

Centre for Autonomous Systems, University of Technology Sydney, Sydney, NSW 2007, Australia

Corresponding author: Karthick Thiyagarajan (karthick.thiyagarajan@uts.edu.au)

This work was supported by the Data Analytics on Sewers Project through in part by the Sydney Water Corporation, in part by the Melbourne Water Corporation, in part by the Water Corporation (WA), and in part by the South Australian Water Corporation.

ABSTRACT In wastewater industry, real-time sensing of surface temperature variations on concrete sewer pipes is paramount in assessing the rate of microbial-induced corrosion. However, the sensing systems are prone to failures due to the aggressively corrosive environmental conditions inside sewer assets. Therefore, reliable sensing in such infrastructures is vital for water utilities to enact efficient wastewater management. In this context, this paper presents a sensor failure detection and faulty data accommodation (SFDFDA) approach that aids to digitally monitor the health conditions of the sewer monitoring sensors. The SFDFDA approach embraces seasonal autoregressive integrated moving average model with a statistical hypothesis testing technique for enabling temporal forecasting of sensor variable. Then, it identifies and isolates anomalies in a continuous stream of sensor data whilst detecting early sensor failure. Finally, the SFDFDA approach provides reliable estimates of sensor data in the event of sensor failure or during the scheduled maintenance period of sewer monitoring systems. The SFDFDA approach was evaluated by using the surface temperature data sourced from the instrumented wastewater infrastructure and the results have demonstrated the effectiveness of the SFDFDA approach and its applicability to surface temperature monitoring sensor suites.

INDEX TERMS Anomalies detection, faulty data accommodation, forecasting, SARIMA model, sensor failure detection, sensor monitoring, sewer corrosion, time series modeling, wastewater infrastructure.

I. INTRODUCTION

Sensors are essential constituents of any critical infrastructure monitoring system. They play an important role in maintaining the system safety and reliability [1], [2]. However, in real-time systems, sensors can provide spurious data owing to different erratic factors including the exposure of the sensor to a harsh environment and inherent sensing malfunctions [3]. Spurious data emanating from the sensors can be momentary or long-lasting. Momentary faulty data are likely to happen randomly due to changes in sensor characteristics and electronics [4]. Those temporary data should not be attributed to sensor failures. Instead, they need to be isolated as anomalies. However, the continuous spurious data is probably an indication of a sensor failure and results in downgrading the performance of an entire monitoring system. Therefore, early sensor failure detection is significant for pertinent intervention strategies while monitoring the environmental phenomena of critical infrastructure assets.

An urban sewerage system is an ideal example of a critical underground infrastructure system. Presently, the concrete pipes in sewer systems are suffering from a higher rate of structural degradation due to hydrogen sulphide-induced concrete corrosion [5]. Because of the increasing corrosion in sewers, the wastewater industry around the globe incurs economic repercussions that are estimated to be in the order of billion dollars [6]. In order to manage the sewer infrastructure effectively, the water utilities rely on the sensor monitoring systems that acquire information-rich data about the corrosion. In this context, the temperature on the concrete surface was identified as an important observation that can provide vital data to the models predicting the rate of sewer corrosion [7], [8].

Currently, there is no reliable sensor system available to monitor temporal dynamics of surface temperature under aggressively corrosive environmental conditions of the sewer. For that reason, this collaborative research between the University of Technology Sydney and four government-owned water utilities aims to develop an advanced sensor suite for monitoring surface temperature variations in sewer pipes. In this regard, a sensor suite was developed and deployed in sewer systems for over three months between 3rd November 2016 and 07th February 2017 in the municipal sewer of the Sydney city in Australia. The field testing campaign has demonstrated that the sensor suite is robust and capable of monitoring for long-term in sewer conditions. However, the confined sewer systems are typically hostile environments to both sensor monitoring and human inspections. Since the sensor system cannot be physically monitored every time, automatic sensor failure detection approaches become a salient need of the sewer monitoring system.

The collaborative research reported in this paper focuses on a SFDFDA approach for a sewer monitoring application. The work motivates the development of a SFDFDA approach that possesses the following three properties:

- Forecasting: Forecasting is a process of predicting the future trends of data based on the collected historical data trends by using a mathematical model [9]. The surface temperature data measured by the sensor deployed inside the sewer pipe is represented as a time-series data. By using the past temporal dynamics of the surface temperature variable, the future trends will be foreseen by using a mathematical model. The forecast data will be acting as a virtual sensor to compare against the actual upcoming sensor data from the sewer systems for detecting anomalies and sensor failures. Also, in the event of scheduled sensor maintenance, the forecast data will be potentially used to replace the actual measurements.
- Anomalies detection and isolation: Anomalies are unexpected patterns in the data that do not comply with the normal behavioral trends [10]. So, the sensor data that suddenly deviates or rare occurrences from the normal pattern is flagged as an anomaly [11]. Hence, it is important to detect and isolate anomalies.
- Sensor failure detection and accommodation: Sensors are prone to fail over time. Detecting early sensor failure will enhance the present sewer monitoring capabilities for effective management of sewer infrastructure assets. Also, it prevents the faulty data to train the forecasting model. Once the sensor failure has been detected, the faulty sensor data needs to be accommodated with the predicted data [12], [13] and this process will be continued until operator addresses the issue.

In this paper, SFDFDA approach using Seasonal Autoregressive Integrated Moving Average (SARIMA) model is proposed with sewer monitoring system as the application domain. The major contributions of our proposed scheme are enumerated as follows:

1) SARIMA model based forecasting for the surface temperature variable from sewer pipes was implemented to form a framework for detecting anomalies and sensor failure.
2) By using the statistical approach, anomalies present in the sensor data were detected and isolated.
3) Sensor failure detection model was implemented by using forecasting technique and the faulty data is accommodated by using the forecast values.

The remainder of this paper is organized as follows: Section II presents the brief review of related work to the SFDFDA approach. Section III describes the sensor suite. Section IV presents the SARIMA model for forecasting the surface temperature variable. Section V illustrates the methodology for the proposed SFDFDA approach. Section VI evaluates the SFDFDA approach and finally, Section VII concludes the paper.

II. BRIEF REVIEW OF RELATED WORK

The state-of-the-art method for implementing the SFDFDA approach through hardware redundancy has been proposed in [14] and [15], where the measurement variable is obtained by using several identical sensors. Then, a voting logic is used to detect the faulty sensor [16], [17]. In the sensing applications where the sensors are expensive, analytical redundancy approaches are popular [16]. In this approach, the signal between the sensor model and sensor is compared to generate the residual error. Then, the sensor failure is detected by setting a threshold logic for the generated residual error values. In the event of detecting the sensor failure, the predicted data is used for sensor failure accommodation [18].

Computational modeling using artificial neural networks is widely used for detecting sensor failures mainly because of its adaptability to dynamic environments [19], [20]. This method works by training the neurons and developing a structure based on the training data for comparing with the sensor measurements to detect the sensor failure [21]. On the other hand, time series based forecasting models that represent sensor data as a linear time series was used for detecting early sensor failure [22]. There are several time series forecasting techniques available in the literature like Random Walk (RW) method, Simple Exponential Smoothing (SES) method and Autoregressive Moving Average (ARMA) Model [23]. In the RW model, the variable value takes the independent random step. This method takes an assumption that past data is not informative and only the present observation is useful [24]. The SES model is used in applications of forecasting seasonal data. However, it is not an appropriate model in applications where the data has trends [25]. The ARMA method is an important method in time series forecasting [26]. This method is a stationary stochastic process that combines the Autoregressive (AR) model and Moving Average (MA) model.

G.E.P. Box and G.M. Jenkins extended the ARMA model to the ARIMA model, which integrates the AR and MA parts of the model with differencing [27], [28]. Among the

FIGURE 1. Field deployment of sensor suite displaying the sensing system installed inside the sewer pipe near the head-space and access station constructed outside the sewer pipe for data monitoring and acquisition.

time series forecasting methods, the ARIMA model has been widely used over the last two decades for forecasting applications [29]. The difference between the ARMA and the ARIMA model is that the ARIMA model converts the non-stationary data into stationary data for predicting the linear time series [26]. For the data that shows seasonal trends, the ARIMA model is extended to the SARIMA model [30]–[32]. Although ARIMA and SARIMA models have been used for forecasting applications in different sectors [33]–[37], their application in forecasting variables inside sewer has not been reported. This work utilizes the SARIMA for forecasting the surface temperature variable.

In order to achieve better forecasting results, the forecast model needs to be provided with anomaly-free data during model training. Therefore, anomaly detection is vital for the application motivated in this work. Methods based on clustering, support vector machine and kernel functions are used for anomalies detection [38]–[40]. However, those approaches are dependent on static routing trees or assigning threshold values to the data streams [3]. In contrast, our work focuses on detecting anomalies through statistical techniques for each sensor measurements. By using the stochastic time series models like SARIMA, anomalies can be detected in the data streams [41], [42]. Once the anomaly is detected, the faulty sensor reading is isolated. Then, the faulty information needs to be accommodated with the reliable value [43].

III. SENSOR SUITE

For monitoring the diurnal variations of surface temperature in sewer pipes, a custom-built sensor suite meeting the requirements specified by the sewer operators was designed and developed using an infrared radiometer sensor. This sensor was housed in a tailor-made enclosure and installed near the head-space of the confined concrete sewer pipe. The sensor measures the thermal radiations of the exposed sewer surface and produces an output in the form of an electrical signal. This signal is transmitted to the access station constructed outside the sewer pipe for processing. Fig. 1 shows the installation of temperature sensor inside the sewer and access station constructed outside the sewer. To the best of our knowledge, the sensor suite used in this work is the first one to demonstrate long-term temporal monitoring of surface temperature dynamics inside aggressively corrosive sewer conditions through non-contact measurements. The SFDFDA approach proposed in this work is applied to the aforementioned sensor suite.

IV. FORECASTING SURFACE TEMPERATURE DATA USING SARIMA MODEL

This section elaborates the forecasting technique employed in this paper by using the surface temperature data sourced from the instrumented wastewater infrastructure.

A. SURFACE TEMPERATURE DATA FROM THE SENSOR SUITE

The surface temperature data coming from the sensor suite can be observed as a time series S_t, where the values of the data are at equally spaced times $t, t-1, t-2, \ldots$ by $S_t, S_{t-1}, S_{t-2}, \ldots$. The time interval between the two sensor measurements is one hour.

B. FORMULATION OF THE SARIMA MODEL

The ARIMA model is a combination of two independent models namely AR model and MA model with finite differencing of data points. Mathematically, the AR part of the ARIMA model can be defined as in (1), which is an autoregressive process of order p. It can be succinctly expressed as $AR(p)$. This $AR(p)$ regresses the evolving variable against its prior values in the series.

$$AR(p)_t = c + \phi_1 \widetilde{S}_{t-1} + \phi_2 \widetilde{S}_{t-2} + \ldots + \phi_p \widetilde{S}_{t-p} + \varepsilon_t \quad (1)$$

where $AR(p)_t$ is the actual value of $AR(p)$ at time period t, $\phi_1, \phi_2, \ldots, \phi_p$ are the finite set of weight parameters of the $AR(p)$ with c as a constant and p as the order of the model $AR(p)$ with $\widetilde{S}_{t-1}, \widetilde{S}_{t-2}, .., \widetilde{S}_{t-p}$ as previous deviations from the mean value. The ε_t is the random shock and it is assumed to be a white noise process [44]. The ε_t is identically distributed i.e. $\varepsilon_t \sim IN(\mu, \sigma^2)$, where the mean $\mu = 0$ and a constant variance σ^2 [45]. The MA part of the ARIMA model is mathematically defined in (2) and it can be called as a moving average process of order q. It can be expressed as $MA(q)$. This $MA(q)$ model uses its past errors as the explanatory variables.

$$MA(q)_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \quad (2)$$

where $MA(q)_t$ is the actual value of $MA(q)$ at time period t, $\theta_1, \theta_2, \ldots, \theta_q$ are the finite set of weight parameters of the $MA(q)$ with c as a constant and q as the order of the model $MA(q)$. Similar to $AR(p)$, the ε_t of $MA(q)$ is assumed to be a white noise process with identically distributed random variables with zero mean and constant variance. Both $AR(p)$ and $MA(q)$ are combined together to form an ARMA model. The model is mathematically defined in (3) and it can be expressed as $ARMA(p, q)$.

$$\begin{aligned} AR(p)_t + MA(q)_t &= c + \phi_1 \widetilde{S}_{t-1} + \phi_2 \widetilde{S}_{t-2} \\ &\quad + \ldots + \phi_p \widetilde{S}_{t-p} + \theta_1 \varepsilon_{t-1} \\ &\quad + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3) \end{aligned}$$

where the predictor is $AR(p)_t + MA(q)_t$ of the $ARMA(p, q)$ as it includes the prior values of $AR(p)$ and past errors of $MA(q)$. The $AR(p)_t + MA(q)_t$ expression can be denoted as $\widetilde{S}_t$. In $ARMA(p, q)$, the order value p of $AR(p)$ model and q of $MA(q)$ model are not greater than 2 [46]. Upon simplification (3) is reduced to (4). The constant term c is omitted for simplicity [47] and (4) is rearranged to (5), where the order of the models p and q denotes the p autoregressive term and q moving average term.

$$\widetilde{S}_t = c + \varepsilon_t + \sum_{n=1}^{p} \phi_n \widetilde{S}_{t-n} + \sum_{m=1}^{q} \theta_m \varepsilon_{t-m} \quad (4)$$

$$\widetilde{S}_t - \sum_{n=1}^{p} \phi_n \widetilde{S}_{t-n} = \varepsilon_t + \sum_{m=1}^{q} \theta_m \varepsilon_{t-m} \quad (5)$$

In time series data, the backshift operator B governs a value in the series to produce its prior value [48]. Mathematically, it is defined in (6), where k is the time series backward observation of the time period.

$$B^k \widetilde{S}_t = \widetilde{S}_{t-k} \quad (6)$$

Generally, the $ARMA(p, q)$ model is manipulated by using (6). By using the lag operator, the $ARMA(p, q)$ model equation in (5) can be expressed as in (7).

$$\left(1 - \sum_{n=1}^{p} \phi_n B^n\right) \widetilde{S}_t = \left(1 + \sum_{m=1}^{q} \theta_m B^m\right) \varepsilon_t \quad (7)$$

The $ARMA(p, q)$ model is suitable only for stationary time series data. However, the sensor data emerging from the sewer pipes possess non-stationary behavior. In order to process the non-stationary nature of sewer data, the ARIMA model is proposed for the application reported in this work to forecast surface temperature measurements. This model obtains homogeneous non-stationary behavior by supposing a suitable d^{th} difference value of the process to the stationary $ARMA(p, q)$. The differencing is mathematically defined in (8), where the $(1-B)^d = \Delta^d$. Mathematically, the general form of ARIMA model can be defined as in (9) and it can expressed as $ARIMA(p, d, q)$.

$$\widetilde{S}_t = \Delta^d \widetilde{S}_t \quad (8)$$

$$\left(1 - \sum_{n=1}^{p} \phi_n B^n\right) \Delta^d \widetilde{S}_t = \left(1 + \sum_{m=1}^{q} \theta_m B^m\right) \varepsilon_t \quad (9)$$

where the p, d and q are the integers referring to the order of autoregressive, integrated and moving average parts of the $ARIMA(p, d, q)$ model. The integer d governs the level of differencing. SARIMA is employed in applications where the time series data presents seasonal changes [48]. The SARIMA is denoted as $SARIMA(p, d, q)(P, D, Q)_{S_p}$, where P is the seasonal autoregressive parameter, D is the degree of seasonal differencing parameter, Q is the seasonal moving average parameter and the subscript S_p denotes the seasonal period this stochastic model. The forecasts of sewer surface temperature variable by using $SARIMA(p, d, q)(P, D, Q)_{Sp}$ is given by (10), where the Φ and Θ are the weight parameters of seasonal autoregressive term and seasonal moving average term respectively.

$$\begin{aligned} &\left(1 - \sum_{n=1}^{p} \phi_n B^n\right)\left(1 - \sum_{n=1}^{P} \Phi_n B^{S_p}\right)(\Delta)^d (\Delta^{S_p})^D \widetilde{S}_t \\ &\quad = \left(1 + \sum_{m=1}^{q} \theta_m B^m\right)\left(1 + \sum_{m=1}^{Q} \Theta_m B^{S_p}\right) \varepsilon_t \quad (10) \end{aligned}$$

C. AUTOMATIC SELECTION OF ARIMA MODEL PARAMETERS

For the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model, the order parameters p, d, q, P, D and Q are automatically determined by using Hyndman and Khandakar algorithm [49] [50]. The differencing terms d and D are computed by performing the unit root test such as the Kwiatkowski Phillips Schmidt Shin (KPSS) test. If the values of differencing parameters d and D are known, then the algorithm [49] selects the values for p, q, P and Q through minimization of an Akaike information criterion (AIC) given in (11).

$$AIC = -2\log(L) + 2(p + q + P + Q + K_n) \tag{11}$$

where L is the maximized likelihood of the forecasting model $SARIMA(p, d, q)(P, D, Q)_{S_p}$, fitted to the differenced data $(\Delta)^d(\Delta^{S_p})^D\tilde{S}_t$ and K_n is the number of parameters estimated to compute one-step ahead forecasts.

D. COMPUTING PREDICTION INTERVALS OF THE FORECASTS AT ANY LEAD TIME

A prediction interval is an estimate of an upper and lower bound of an interval in which the observable variable of the future is expected to lie with a specified probability based on the past observed values [50], [51]. Considering that $g's$ are Gaussian distribution with standard deviation σ_g, then the probability distribution $(S_{t+f}|S_t, S_{t-1}, S_{t-2}, \ldots)$ of a future observable value S_{t+f} of the process will be normal with mean $\hat{S}(f)$ and standard distribution is given in (12) [46].

$$\sigma(f) = \left(1 + \sum_{j=1}^{f-1} \psi_j^2\right)^{1/2} \sigma_g \tag{12}$$

The variate $\left[\left[S_{t+f} - \hat{S}_t(f)\right]/\left[\sigma(f)\right]\right]$ will posses a unit normal distribution. Therefore, for S_{t+f}, $\hat{S}_f \pm \mu_{\lambda/2}\sigma(f)$ will provide the bounds of the prediction interval with probability $(1-\lambda)$. $\mu_{\lambda/2}$ is the deviate transcended by a proportion of $\lambda/2$ of the unit normal distribution. Mathematically, the prediction interval for the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model can be computed by using (13) [46].

$$\hat{S}_{t+f}(\pm) = \hat{S}_t(f) \pm \mu_{\lambda/2}\left(1 + \sum_{j=1}^{f-1} \psi_j^2\right)^{1/2} \sigma_g \tag{13}$$

where $\mu_{\lambda/2}$ are the percentiles of the standard normal distribution. In this paper $\mu_{\lambda/2} = 95\%$. The forecast value $\hat{S}_{t+f}$ coming from the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model with the probability of $1 - \lambda$ will lie between the upper interval $\hat{S}_{t+f}(+)$ and lower interval $\hat{S}_{t+f}(-)$, i.e. Probability $\left\{\hat{S}_{t+f}(-) < \hat{S}_{t+f} < \hat{S}_{t+f}(+)\right\}$.

V. SFDFDA APPROACH

The SFDFDA approach proposed in this work presents advanced data analytics solution by combining predictive analytics and diagnostic analytics methods. The predictive analytics component of the SFDFDA approach features $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model for forecasting the surface temperature variable. The forecast data $\hat{S}_{t+f}$ will function as a virtual sensor. The forecasting process of the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model involves three main steps. The first step of the forecasting model uses the time-series data to provide the surface temperature sensor data from 4th November 2016 to 10th November 2016 for training the forecast model. The presence of anomalies will downgrade the prediction performance of the forecast model. Therefore, it is important to detect the anomalies and isolate them before supplying the data for training the forecast model. In this work, we compared the trends of sensor data with the benchmark sensor measurements before supplying the data for training. The initial training data contains 168 data points. Then, the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model parameters are determined in the second step by invoking the Hyndman and Khandakar algorithm for automatic selection of p, d, q, P, D and Q, and then building a forecast model using the determined parameters. Finally in the third step, the model forecasts the surface temperature data $\hat{S}_{t+f}$ for the next day containing 24 data points with $\hat{S}_{t+f}(-)$ and $\hat{S}_{t+f}(+)$ values. The pseudocode for SARIMA forecast process is presented in Algorithm 1.

Algorithm 1 Pseudocode for the SARIMA Forecast

```
for all i ∈ 1 : length[R(t)_i] do
    Computing p, d, q, P, D and Q
    Forecasting [Ŝ_{t+f}]_i
    Computing [Ŝ_{t+f}(−)]_i and [Ŝ_{t+f}(+)]_i
    i = i + 24
end for
```

The diagnostic analytics component of the SFDFDA approach presents statistical techniques for detecting sensor failure and anomalies using the data from the surface temperature sensor and $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model forecasts. This component employs statistical hypothesis testing for computing probability value (p-value) to detect anomalies and sensor failure. A p-value is obtained by performing Pearson's chi-squared test, denoted as χ^2. It determines the divergence of the observed sensor data from the values that would be forecasted using $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model under the null hypothesis of no association. The chi-squared distribution χ^2_{df} is used in the χ^2 for goodness of fit of the observed sensor data distribution to a distribution of $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model data. The χ^2_{df} is characterized by degrees of freedom df, whose value is one less than the number of total data points in the data set used to compute one χ^2 measure. A sliding window mechanism [16], [21] is incorporated within the SFDFDA framework to provide a set of data for computing p-value. This mechanism is illustrated in Fig. 2, where the sliding window of size W_L data points keeps moving as the time t progresses. In the proposed SFDFDA approach, $W_L = n$. So, the χ^2_{df} of observed sensor data and the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model data takes n data points for computing χ^2. Therefore, the df of that χ^2

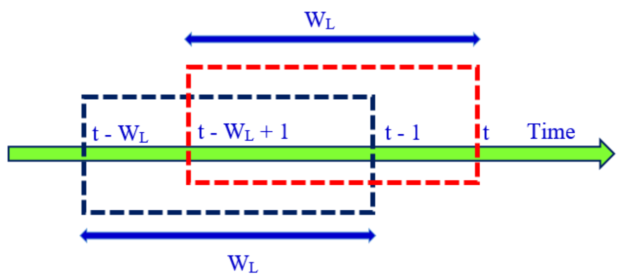


FIGURE 2. Illustration of the sliding window mechanism at time period *t*.

Algorithm 2 Pseudocode - Function CHI CALCULATOR

/* FUNCTION PARAMETERS */
$[R(t)]_i$, $[\hat{S}_{t+f}]_i$, W_L
/* INITIALIZATION OF VARIABLES */
$Total = 0$

for all $i \in 1 : (1 + W_L)$ **do**
  $Total = Total + \left([R(t)]_i - [\hat{S}_{t+f}]_i\right)^2 / [\hat{S}_{t+f}]_i$
end for
$return(Total)$

will be $W_L - 1$ for all sliding windows. The χ^2 measure for the testing dataset $\sum$ of size W_L is measured using (14):

$$\chi^2 = \sum_{i=1}^{i=W_L} \frac{\left[(R_t)_i - (\hat{S}_{t+f})_i\right]^2}{(\hat{S}_{t+f})_i} \tag{14}$$

where i is the instantaneous time, χ^2 is the cumulative statistic of Pearson's chi-squared test, R_t is the observed surface temperature sensor data, $\hat{S}_{t+f}$ is the expected data resulting from the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model. The pseudocode for determining χ^2 of each sliding window is presented in Algorithm 2.

Since the value of W_L is static for all the computations, df will be same for all the computations as well with the value of $df = n - 1$. After computing χ^2 and df, it is therefore important to set a critical significance level to determine the p-value for each sliding window. The critical significance level is denoted as α. Typically, in the proposed SFDFDA approach α is 5% i.e., $\alpha = 0.05$. Given the α and df, the contingency table that shows multivariate frequency distribution of the variables will be referred. This table provides values with respect to df and α. Then, by comparing the measures of χ^2 with χ^2_{df} p-value is given by (15).

$$P - value = P(\chi^2_{df,\alpha} \geq \chi^2) \tag{15}$$

For the critical level of α, the statistical hypothesis testing provides significant value only if the is $\chi^2_{df,\alpha}$ greater than the χ^2. In case of χ^2 being greater than the $\chi^2_{df,\alpha}$, the statistical hypothesis testing provides a non-significant value. The determination of significant and non-significant value plays a paramount role in the SFDFDA approach for detecting the anomalies and sensor failure.

Algorithm 3 Pseudocode - Function P-Value CALCULATOR

/* FUNCTION PARAMETERS */
$[R(t)]_i$, $[\hat{S}_{t+f}]_i$, W_L
/* INITIALIZATION OF VARIABLES */
$df = W_L - 1$

$\chi^2 = chiCalculator(R(t)]_i, [\hat{S}_{t+f}]_i, W_L)$
$P - value = 1 - chi2Cdf(\chi^2, df)$
$return(P - value)$

In a sliding window, if the p-value is a significant value i.e., p-value > 0.95, then the sensor data measurements of that particular sliding window will be pushed to the training dataset of the $SARIMA(p, d, q)(P, D, Q)_{S_p}$ model for forecasting future values. Consequently, the sliding window progresses to the next window for performing statistical testing. This process iterates as long as the surface temperature provides measurements.

In the case of sliding window producing a non-significant p-value i.e., p-value < 0.95, then the sensor data measurements of that particular window is examined to check the presence of anomalies or any indications of early sensor failure. Each surface temperature measurement within that sliding window is evaluated with the prediction intervals. Precisely, the condition defined in (16) is examined.

$$\left[\left(\hat{S}_{t+f}(-)\right)_i\right] < \left[\left(R_t\right)_i\right] < \left[\left(\hat{S}_{t+f}(+)\right)_i\right] \tag{16}$$

where i is the instantaneous time. If the condition in (16) is not satisfied, then there arise three scenarios. In all the scenarios, the SFDFDA approach will look for continuity of individual data of sensor measurements present outside of $\hat{S}_{t+f}(-)$ or $\hat{S}_{t+f}(+)$. The three scenarios are as follows:

- In the first scenario where one or two $(R_t)_i$ present outside of $\hat{S}_{t+f}(-)$ or $\hat{S}_{t+f}(+)$, then that respective sensor data is regarded as an anomaly. Subsequently, the SFDFDA approach performs data accommodation process, where the value of $(R_t)_i$ is accommodated by $\left(\hat{S}_{t+f}\right)_i$.
- In the second scenario where there are three or more $(R_t)_i$ present outside of $\hat{S}_{t+f}(-)$ or $\hat{S}_{t+f}(+)$ and their continuity is less than three successive times, $(R_t)_i$ is still flagged as an anomaly. However, a sensor failure warning will be issued for inspection. In addition to the warning signal, the SFDFDA approach undergoes data accommodation process to replace the faulty $(R_t)_i$ with respective $\left(\hat{S}_{t+f}\right)_i$.
- Finally, in the third scenario where there are more than one $(R_t)_i$ present outside of $\hat{S}_{t+f}(-)$ or $\hat{S}_{t+f}(+)$ and their continuity is three or more successive times in one sliding window, then a signal of early sensor failure is issued. In this scenario, the SFDFDA approach will also invoke data accommodation process to replace the faulty $(R_t)_i$ with respective $\left(\hat{S}_{t+f}\right)_i$. The SFDFDA approach

Algorithm 4 Psuedocode - Detecting Sensor Failure and Anomalies, Data Accommodation Process

```
/* FUNCTION PARAMETERS */
P − value, [R(t)]_i, [Ŝ_{t+f}(−)]_i, [Ŝ_{t+f}]_i, [Ŝ_{t+f}(+)]_i, W_L
Timestamp
/* INITIALIZATION OF VARIABLES */
COUNTERS : Total = 0; Warning_Count = 0;
Failure_Count = 0; Previous_ID = 0;
FLAG : Failed = FALSE

if (!(P − value > α)) then
  for all i ∈ 1 : (1 + W_L) do
    if (![R(t)]_i ≥ [Ŝ_{t+f}(−)]_i&&[R(t)]_i ≤ [Ŝ_{t+f}(+)]_i)
    then
      /* Evaluating Sensor Failure Condition*/
      if (i == Previous_ID + 1) then
        (Failure_Count = Failure_Count + 1)
        if (Failure_Count >= 3) then
          Message: Sensor_Failure_(Timestamp)
          (Failed = TRUE)
        end if
      end if
      Previous_ID = i
      /* Evaluating Sensor Warning Condition */
      if (!Failed) then
        (Warning_Count = Warning_Count + 1)
        if (Warning_Count >= 3) then
          Message: Sensor_Warning_(Timestamp)
        end if
      end if
      /* Anomaly Detection & Data Accommodation */
      Message: Anomaly_Detected_(Timestamp)
      Data Accommodation: [R(t)]_i = [Ŝ_{t+f}]_i
    end if
  end for
end if
return([R(t)]_i)
```

iterates the data accommodation process until $(R_t)_i$ is present within the $\hat{S}_{t+f}(-)$ and $\hat{S}_{t+f}(+)$.

VI. EXPERIMENTAL EVALUATION

This section evaluates the proposed SFDFDA approach by using the surface temperature sensor data sourced from the instrumented wastewater infrastructure. During the course of the sewer monitoring campaign, the surface temperature sensor demonstrated robust behavior and did not generate prolonged spurious data. However, the sensor has produced some spurious data in the interim of laboratory evaluation. So, for evaluating the SFDFDA approach, we have injected anomalies based on the lab data to the time series data observed during the field testing. In addition, we have implanted continuous spurious data on 22nd to 23rd December 2016 and 5th to 6th February 2017 to simulate sensor failure. The size of the sliding window was heuristically chosen as 6 based on the knowledge of sensor characteristics and therefore, each window takes 6 sensor measurements.

Figure 3 and Fig. 4 illustrates the results of the experiments of SFDFDA approach to demonstrate sensor failure detection, anomaly detection and isolation, and data accommodation process from 11th November 2016 to 23rd December 2016 and from 24th December 2016 to 6th February 2017 respectively. The first plot displays the model forecast within the prediction interval and the second plot shows the sensor data with some random and continuous spurious measurements. By using the forecast and sensor data, the p-value is determined and presented in the third plot along with the critical value of 0.95. For the random and continuous spurious data, the p-value was observed to be lower than the critical value and finally in the last plot, the data accommodation process illustrating the replacement of faulty data is shown.

To measure the detection performance of random and continuous spurious data, we use successful detection rate (SDR) [3] as our metric of accuracy. The SDR for the injected and observed anomalies of the periods from 11th November 2016 to 23rd December 2016 and from 24th December 2016 to 6th February 2017 were 100% and 93.34% respectively. The reason for the slightly lower value of SDR in the latter period is due to the closeness of anomaly to the forecast value. In this case, the p-value remains to be higher than the critical value. The SFDFDA approach successfully detected anomalies including the two successive ones and thereby reported the anomalies with time-stamp and accommodated the corresponding forecast data. In the case of more frequent anomalies present in a single window frame, the SFDFDA approach issues sensor failure warning for early intervention. In this case, the p-value remains lower than the critical value. Long-term monitoring of sensor data for detecting sensor failure is displayed in Fig. 5. Overall for the entire time period, the SDR is 97.06%. For the continuous spurious data on 22nd to 23rd December 2016 and 5th to 6th February 2017, SDR were 100% for each period, indicating the efficacy of sensor failure detection.

The forecasting performance of ARIMA model was evaluated and compared with the other models, where ARIMA outperforms the other stochastic models [52]. In order to evaluate the forecasting performance of the SARIMA model, we compared the forecasting data of two different periods with the anomalies-free sensor measurements. The first period is from 25th to 30th November 2016 and the second period is from 11th to 16th January 2017. Fig. 6 presents the temporal profile of forecasted and sensor measurements data, where it can be observed that the profiles tend to follow a similar pattern to each other. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) were used as statistical metrics to evaluate the forecasting performance of the SARIMA model. The MAE, MAPE and RMSE for the period from 25th to 30th November 2016 and from 11th to 16th January 2017 were 0.18°C,

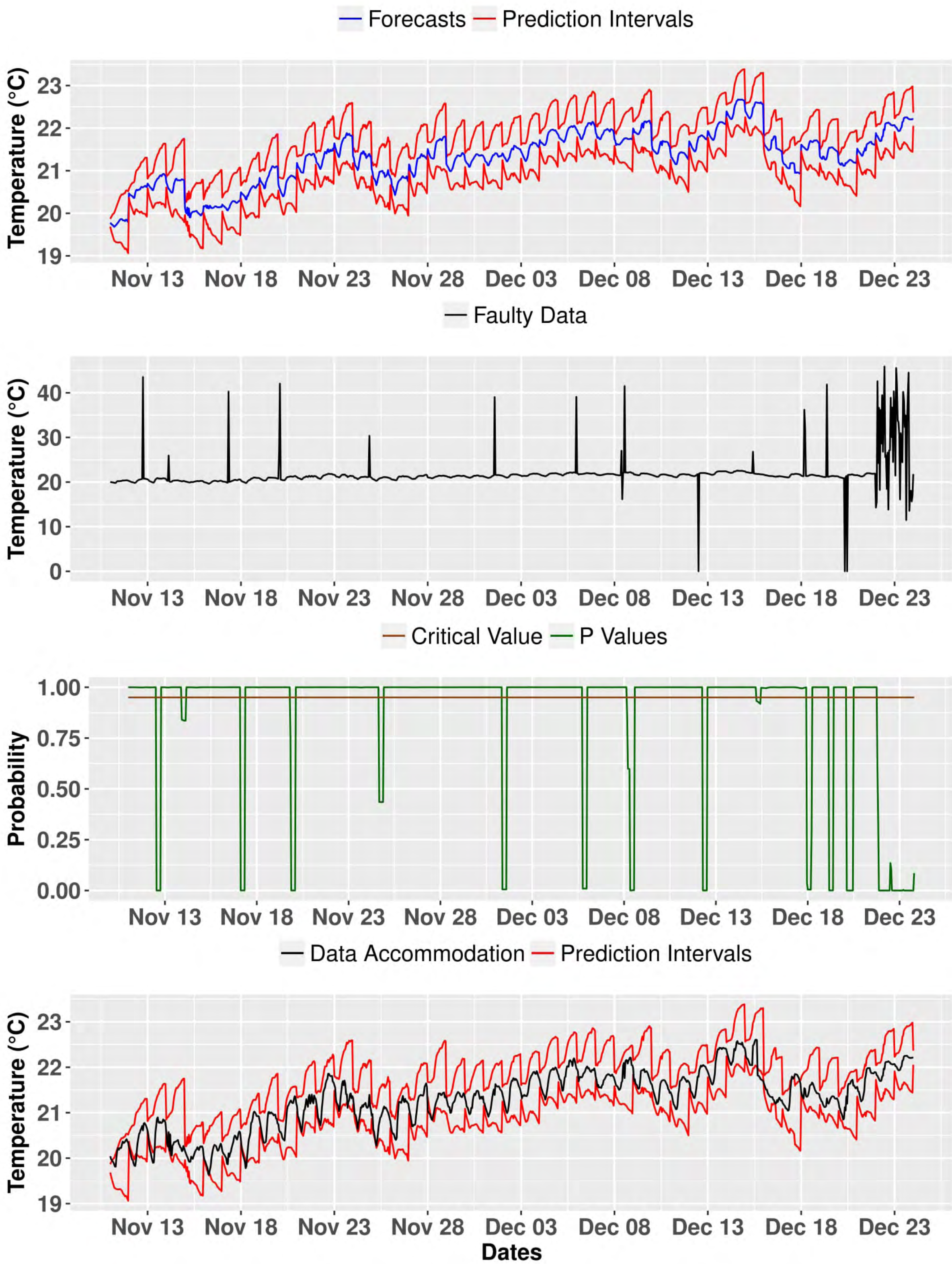


FIGURE 3. **Experiment-1: Evaluation of SFDFDA approach.**

0.0086%, 0.24°C and 0.15°C, 0.0062%, 0.17°C respectively. These statistical metrics demonstrate the higher accuracy of the forecast model employed in this work.

As in Fig. 7, the data accommodation occurs once the sensor failure is detected and the corresponding forecast data is utilized to replace the continuous spurious data.

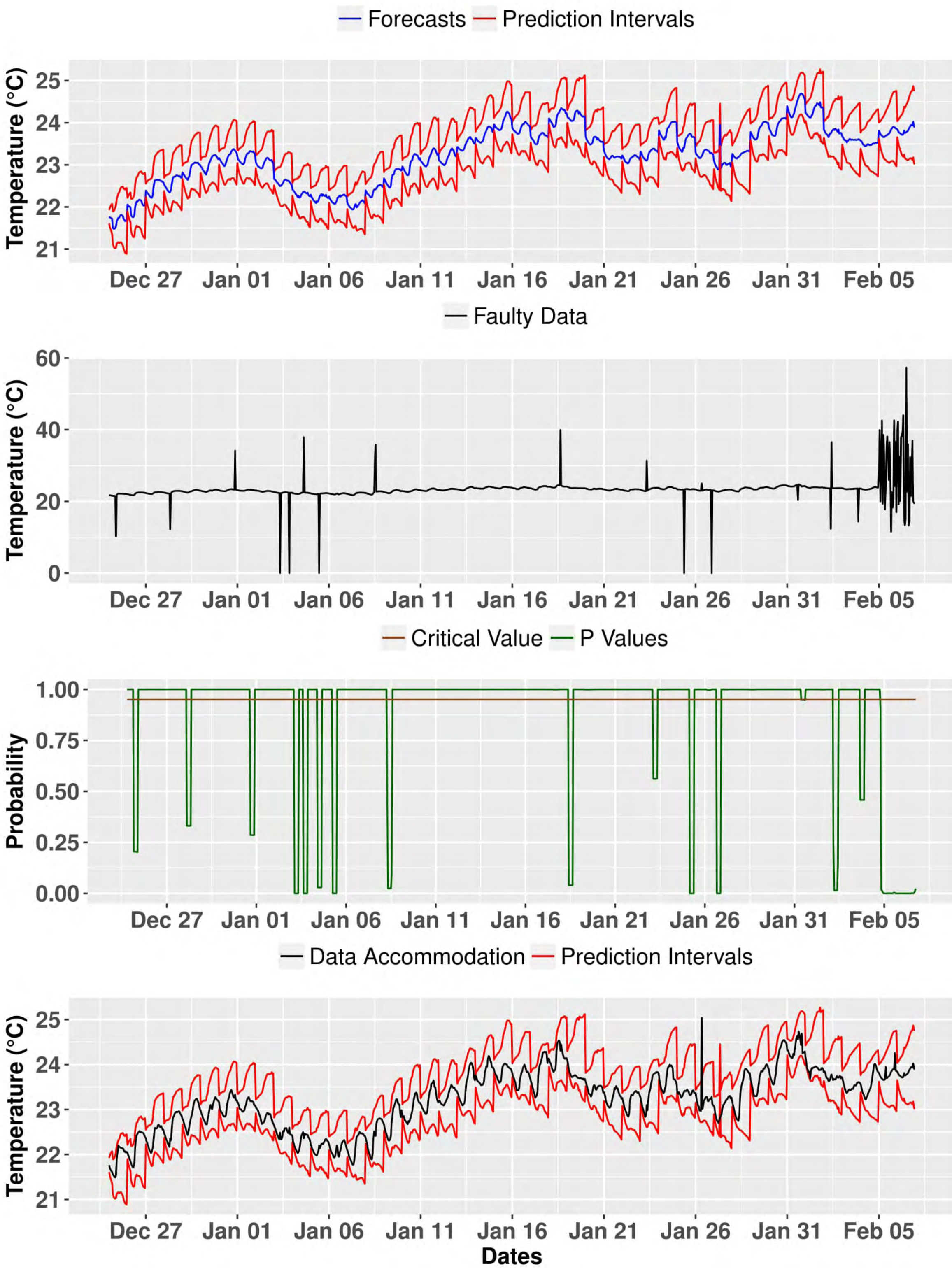


FIGURE 4. Experiment-2: Evaluation of SFDFDA approach.

Performance analysis on data accommodation process was carried out to determine MAE, MAPE and RMSE on the failure period from 22nd to 23rd December 2016 and from 5th to 6th February 2017, and were determined to be 0.28°C, 0.01%, 0.31°C and 0.46°C, 0.02%, 0.53°C respectively. These results show that the data accommodation process provides satisfactory prediction when the sensor generates spurious data.

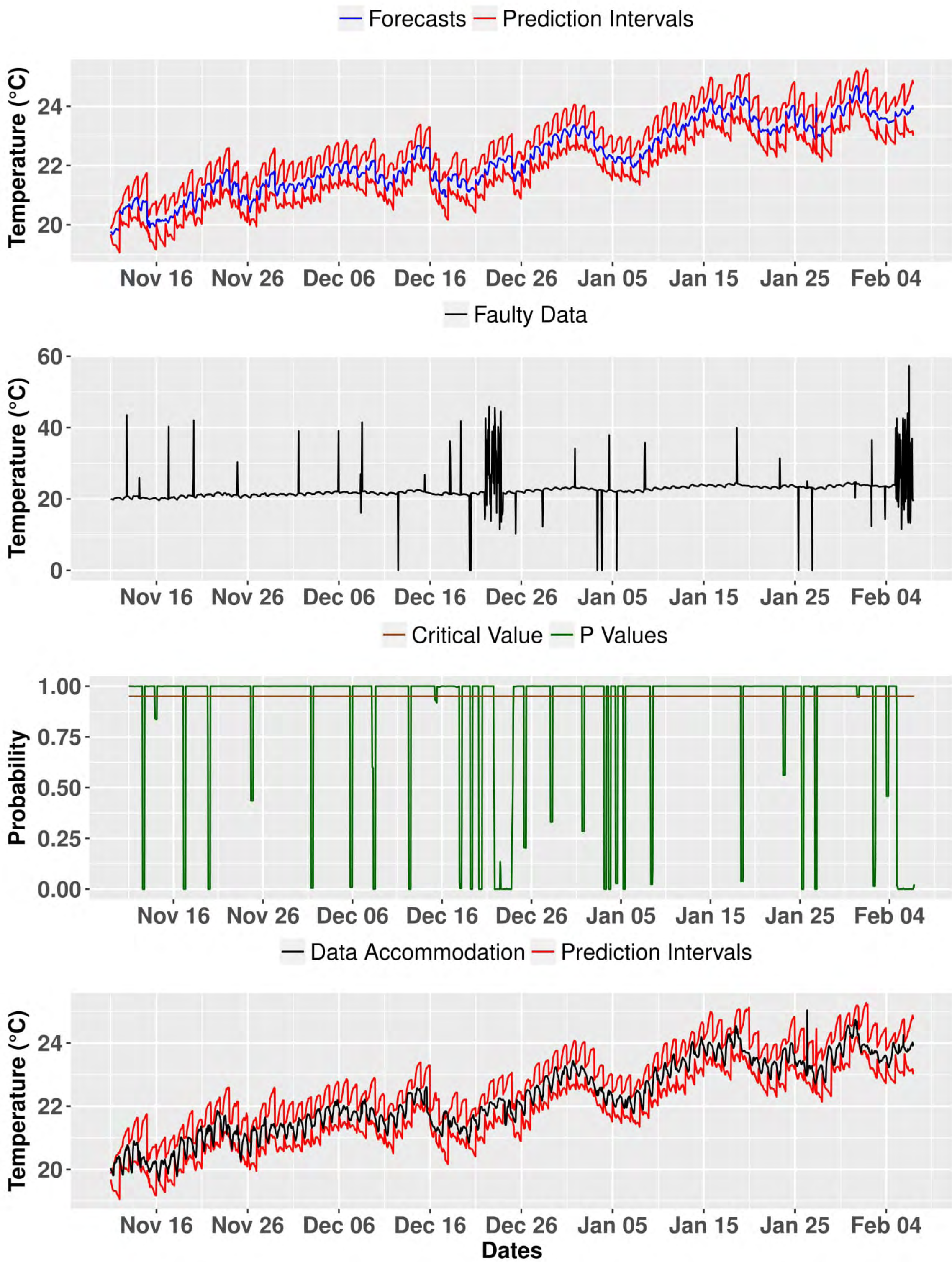


FIGURE 5. Long-term evaluation of SFDFDA approach.

VII. DISCUSSION

In this paper, we proposed a SFDFDA approach, and its effectiveness was evaluated by utilizing the surface temperature sensor measurements. The proposed approach has the potential to be used for detecting failures of different sensors, which are monitoring other essential infrastructure parameters in distribution networks such as pH, resistivity, conductivity etc. Besides detecting the sensor failures and anomalies, the proposed approach can be used for forecasting the sensor parameters.

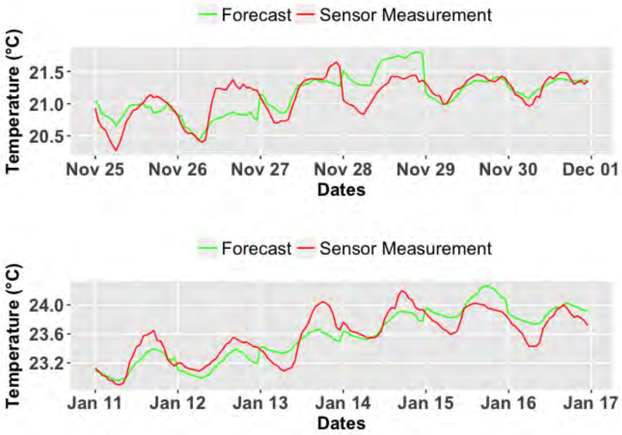


FIGURE 6. **Illustration of SFDFDA approach forecasting performance.**

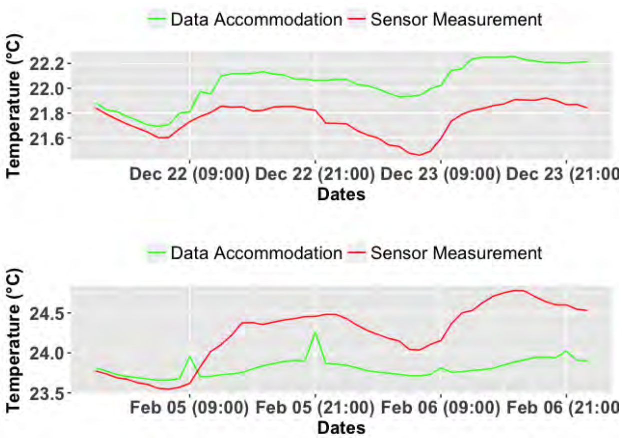


FIGURE 7. **Illustration of the data accommodation during sensor failure.**

The SFDFDA approach simultaneously check for anomalies and sensor failures. In the broader context, surface temperature data has the potential to be used in sewer corrosion predicting models. Presence of anomalies in the observed data can affect the prediction accuracy of those models. Hence, in the proposed approach anomaly detection and isolation are carried out simultaneously to the process of detecting sensor failure. The sensor failure detection of the SFDA algorithm is set to a heuristic criterion, which is based on the repeated observation of faulty data. The criterion was based on the sensor characteristics, where three or more consecutive faulty data means a sensor failure. The size of the sliding window was heuristically chosen as 6, based on the knowledge of sensor characteristics.

The data collection was organised during the summer season of the Sydney city in Australia. Although the ambient temperature of the field location was high, the sewer air temperature was between 20°C and 26°C. This can be due to the thickness of the concrete layer and the soil layer above the sewer pipe providing a thermal barrier. More details about the design and development of the surface temperature sensing suite, and field experimentation will be published subsequently in a journal.

VIII. CONCLUSION

This paper presented an approach called SFDFDA, for detecting early sensor failure based upon the real-time operational data sourced from an urban sewer. The SFDFDA approach utilizes SARIMA model for forecasting the sensor data to comprehend the temporal dynamics of the variable. This forecasting mechanism is used as a framework to provide an alternate measure to physical sensor measurements. The forecast data from the SARIMA model was used as a reference measure in the SFDFDA approach to perform anomaly detection, early sensor failure detection and data accommodation. The SFDFDA approach integrates the forecasting mechanism with statistical diagnostic method. In the event of detecting anomalies, the algorithm isolates the spurious data and accommodates the data with the corresponding forecast value. Further, in case of a continuity of faulty data, the early sensor failure is detected and data accommodation process is invoked to provide predicted measures. Experimental evaluation demonstrates that SFDFDA approach can be used for surface temperature monitoring in-sewer application with high detection accuracy and efficiency.

ACKNOWLEDGMENT

The research participants are Data61 - Commonwealth Scientific and Industrial Research Organization (CSIRO), University of Technology Sydney (UTS) and The University of Newcastle (UoN).

REFERENCES

[1] Y. Zhang, Y. Fu, Z. Wang, and L. Feng, "Fault detection based on modified kernel semi-supervised locally linear embedding," *IEEE Access*, vol. 6, pp. 479–487, 2017.

[2] S. Yin, S. X. Ding, and D. Zhou, "Diagnosis and prognosis for complicated industrial systems—Part I," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2501–2505, Apr. 2016.

[3] P.-Y. Chen, S. Yang, and J. A. McCann, "Distributed real-time anomaly detection in networked industrial sensing systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3832–3842, Jun. 2015.

[4] W. He, P.-L. Qiao, Z.-J. Zhou, G.-Y. Hu, Z.-C. Feng, and H. Wei, "A new belief-rule-based method for fault diagnosis of wireless sensor network," *IEEE Access*, vol. 6, pp. 9404–9419, 2018.

[5] G. Jiang, J. Keller, and P. L. Bond, "Determining the long-term effects of H2S concentration, relative humidity and air temperature on concrete sewer corrosion," *Water Res.*, vol. 65, pp. 157–169, Nov. 2014.

[6] M. P. H. Brongers, P. Y. Virmani, and J. H. Payer, "Drinking water and sewer systems in corrosion costs and preventative strategies in the United States," Federal Highway Admin. Publ., U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. FHWA-RD-01-156, 2002.

[7] K. Thiyagarajan, S. Kodagoda, and J. K. Alvarez, "An instrumentation system for smart monitoring of surface temperature," in *Proc. 14th Int. Conf. Control, Autom., Robot. Vis. (ICARCV)*, Nov. 2016, pp. 1–6.

[8] K. Thiyagarajan, S. Kodagoda, and N. Ulapane, "Data-driven machine learning approach for predicting volumetric moisture content of concrete using resistance sensor measurements," in *Proc. IEEE 11th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2016, pp. 1288–1293.

[9] L. Yi, X. Ke, and S. Junde, "Research on forecasting and early-warning methods," in *Proc. IEEE 9th Int. Conf. Mobile Ad-Hoc Sensor Netw., MSN*, Dec. 2013, pp. 570–576.

[10] A. Patcha and J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Comput. Netw.*, vol. 51, no. 12, pp. 3448–3470, Aug. 2007.

[11] W. Yuan, L. Zhou, D. Guan, G. Han, and L. Shu, "Anomaly detection for civil aviation pilots using step-sensors," *IEEE Access*, vol. 5, pp. 11236–11243, 2017.

[12] I. Samy, I. Postlethwaite, and D.-W. Gu, "Detection and accommodation of sensor faults in UAVs—A comparison of NN and EKF based approaches," in *Proc. IEEE Conf. Decision Control*, Dec. 2010, pp. 4365–4372.

[13] C. Aubrun and C. Leick, "Sensor fault accommodation: Application to an activated sludge process," *IFAC Proc. Volumes*, vol. 38, no. 1, pp. 371–375, 2005.

[14] I. Samy, I. Postlethwaite, and D. Gu, "Neural network based sensor validation scheme demonstrated on an unmanned air vehicle (UAV) model," in *Proc. 47th IEEE Conf. Decision Control*, Dec. 2008, pp. 1237–1242.

[15] M. R. Napolitano, Y. An, and B. A. Seanor, "A fault tolerant flight control system for sensor and actuator failures using neural networks," *Aircraft Design*, vol. 3, no. 2, pp. 103–128, 2000.

[16] I. Samy, I. Postlethwaite, and D.-W. Gu, "Survey and application of sensor fault detection and isolation schemes," *Control Eng. Pract.*, vol. 19, no. 7, pp. 658–674, 2011.

[17] P. Goupil, "AIRBUS state of the art and practices on FDI and FTC in flight control system," *Control Eng. Pract.*, vol. 19, no. 6, pp. 524–539, 2011.

[18] Y. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annu. Rev. Control*, vol. 32, no. 2, pp. 229–252, 2008.

[19] S. Toma, L. Capocchi, and G.-A. Capolino, "Wound-rotor induction generator inter-turn short-circuits diagnosis using a new digital neural network," *IEEE Trans. Ind. Electron.*, vol. 60, no. 9, pp. 4043–4052, Sep. 2013.

[20] R. Isermann and P. Ballé, "Trends in the application of model-based fault detection and diagnosis of technical processes," *Control Eng. Pract.*, vol. 5, no. 5, pp. 709–719, 1997.

[21] S. Hussain, M. Mokhtar, and J. M. Howe, "Sensor failure detection, identification, and accommodation using fully connected cascade neural network," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1683–1692, Mar. 2015.

[22] T. Kuzin and T. Borovicka, "Early failure detection for predictive maintenance of sensor parts," in *Proc. ITAT*, 2016, pp. 123–130.

[23] N. S. Nalawade and M. M. Pawar, "Forecasting telecommunications data with autoregressive integrated moving average models," in *Proc. 2nd Int. Conf. Recent Adv. Eng. Comput. Sci. (RAECS)*, Dec. 2015, pp. 1–6.

[24] L. R. F. Simmons, "M-competition—A closer look at NAIVE2 and median APE: A note," *Int. J. Forecasting*, vol. 2, no. 4, pp. 457–460, 1986.

[25] E. S. Gardner, "Exponential smoothing: The state of the art—Part II," *Int. J. Forecasting*, vol. 22, no. 4, pp. 637–666, Oct./Dec. 2006.

[26] P. Mondal, L. Shit, and S. Goswami, "Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices," *Int. J. Comput. Sci., Eng. Appl.*, vol. 4, no. 2, p. 13, 2014.

[27] G. E. P. Box and G. C. Tiao, "Intervention analysis with applications to economic and environmental problems," *J. Amer. Statist. Assoc.*, vol. 70, no. 349, pp. 70–79, 1975.

[28] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Seres Analysis: Forecasting and Control* (Wiley Series in Probability and Statistics), 4th ed. 2013.

[29] Z. Qiao, J. Zhou, F. Si, Z. Xu, and L. Zhang, "Fault diagnosis of slurry pH data base on autoregressive integrated moving average and least squares support vector machines," in *Proc. 9th Int. Conf. Natural Comput. (ICNC)*, Jul. 2013, pp. 141–145.

[30] C. S. Luo, L. Y. Zhou, and Q. F. Wei, "Application of SARIMA model in cucumber price forecast," *Appl. Mech. Mater.*, vols. 373–375, pp. 1686–1690, 2013.

[31] T.-M. Choi, Y. Yu, and K.-F. Au, "A hybrid SARIMA wavelet transform method for sales forecasting," *Decision Support Syst.*, vol. 51, no. 1, pp. 130–140, 2011.

[32] H. A. Mombeni, S. Rezaei, S. Nadarajah, and M. Emami, "Estimation of water demand in Iran based on SARIMA models," *Environ. Model. Assessment*, vol. 18, no. 5, pp. 559–565, 2013.

[33] K. Yunus, T. Thiringer, and P. Chen, "ARIMA-based frequency-decomposed modeling of wind speed time series," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2546–2556, Jul. 2016.

[34] A. K. Rajeevan, P. V. Shouri, and U. Nair, "ARIMA based wind speed modeling for wind farm reliability analysis and cost estimation," *J. Elect. Eng. Technol.*, vol. 11, no. 4, pp. 869–877, 2016.

[35] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[36] A. Geetha and G. M. Nasira, "Time series modeling and forecasting: Tropical cyclone prediction using ARIMA model," in *Proc. 3rd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, Mar. 2016, pp. 3080–3086.

[37] R. N. Calheiros, E. Masoumi, R. Ranjan, and R. Buyya, "Workload prediction using ARIMA model and its impact on cloud applications' QoS," *IEEE Trans. Cloud Comput.*, vol. 3, no. 4, pp. 449–458, Oct./Dec. 2015.

[38] S. Rajasegarar, C. Leckie, and M. Palaniswami, "Hyperspherical cluster based distributed anomaly detection in wireless sensor networks," *J. Parallel Distrib. Comput.*, vol. 74, no. 1, pp. 1833–1847, Jan. 2014.

[39] S. Subramaniam, T. Palpanas, D. Papadopoulos, V. Kalogeraki, and D. Gunopulos, "Online outlier detection in sensor data using non-parametric models," in *Proc. 32nd Int. Conf. Very Large Data Bases (VLDB)*, 2006, pp. 187–198.

[40] J. Branch, C. Giannella, B. Szymanski, R. Wolff, and H. Kargupta, "In-network outlier detection in wireless sensor networks," *Knowl. Inf. Syst.*, vol. 34, no. 1, pp. 23–54, 2013.

[41] M. R. Saybani, T. Y. Wah, A. Amini, and S. R. A. S. Yazdi, "Anomaly detection and prediction of sensors faults in a refinery using data mining techniques and fuzzy logic," *Sci. Res. Essays*, vol. 6, no. 27, pp. 5685–5695, 2011.

[42] H. Z. Moayedi and M. A. Masnadi-Shirazi, "Arima model for network traffic prediction and anomaly detection," in *Proc. Int. Symp. Inf. Technol. (ITSim)*, vol. 4, Aug. 2008, pp. 1–6.

[43] X. He, Z. Wang, Y. Liu, L. Qin, and D. Zhou, "Fault-tolerant control for an Internet-based three-tank system: Accommodation to sensor bias faults," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2266–2275, Mar. 2017.

[44] K. W. Hipel and A. I. McLeod, *Time Series Modelling of Water Resources and Environmental Systems*, vol. 45. Amsterdam, The Netherlands: Elsevier, 1994.

[45] J. C. Paul, S. Hoque, and M. M. Rahman, "Selection of best ARIMA model for forecasting average daily share price index of pharmaceutical companies in Bangladesh: A case study on square pharmaceutical Ltd," *Global J. Manage. Bus. Res.*, vol. 13, no. 3-C, Mar. 2013.

[46] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.

[47] R. Adhikari and R. K. Agrawal. (2013). "An introductory study on time series modeling and forecasting." [Online]. Available: https://arxiv.org/abs/1302.6613

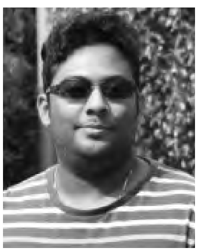
[48] I. I. Permanasari, A. E. Hidayah, and I. A. Bustoni, "SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence," in *Proc. Int. Conf. Inf. Technol. Elect. Eng. (ICITEE)*, Oct. 2013, pp. 203–207.

[49] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for R," *J. Statist. Softw.*, vol. 27, no. 3, pp. 1–22, 2008.

[50] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*. Otexts.org., 2014. [Online]. Available: https://robjhyndman.com/uwafiles/fpp-notes.pdf

[51] R. D. Snyder, J. K. Ord, and A. B. Koehler, "Prediction intervals for ARIMA models," *J. Bus. Econ. Statist.*, vol. 19, no. 2, pp. 217–225, 2001.

[52] K. Thiyagarajan, S. Kodagoda, and L. Van Nguyen, "Predictive analytics for detecting sensor failure using autoregressive integrated moving average model," in *Proc. 12th IEEE Conf. Ind. Electron. Appl.*, Jun. 2017, pp. 1923–1928.



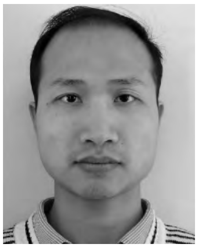
KARTHICK THIYAGARAJAN (M'18) received the B.E. degree in electronics and instrumentation engineering from Anna University, Chennai, India, in 2011, the M.Sc. degree in mechatronics from the University of Newcastle Upon Tyne, Newcastle Upon Tyne, U.K., in 2013, and the Ph.D. degree in smart sensor technologies from the University of Technology Sydney, Sydney, Australia, in 2018.

He is currently a Research Fellow with the Centre for Autonomous Systems, University of Technology Sydney. His current research interests includes sensing technologies, predictive analytics, and infrastructure robotics. His Ph.D. research work was awarded with the Student Water Prize 2018 from the NSW–Australian Water Association.

SARATH KODAGODA (M'10) received the B.Sc. (Eng.) degree (Hons.) in electrical engineering from the University of Moratuwa, Sri Lanka, in 1995, and the M.Eng. and Ph.D. degrees in robotics from Nanyang Technological University, Singapore, in 2000 and 2004, respectively.

He was a Design Engineer in a reputed multinational company. He is currently an Associate Professor, the Deputy Director of teaching and research integration with the Centre for Autonomous Systems, the Founder of the iPipes Lab, and a Program Coordinator of the Degree in mechanical and mechatronics engineering with the University of Technology Sydney, Ultimo, NSW, Australia. His current research interests include infrastructure robotics, sensors and perception, machine learning, and human robot interaction.

Dr. Kodagoda has served as a keynote speaker, a general chair, an associate editor, and a program committee member in number of top robotic conferences. He is currently serving as the Secretary to the Australian Robotics and Automation Association and a Co-Chair of the 'A Robotic Roadmap for Australia.'



LINH VAN NGUYEN (M'15) received the Ph.D. degree in robotics from the University of Technology Sydney, Ultimo, NSW, Australia, in 2015.

He was with the University of Technology Sydney as a Post-Doctoral Research Associate until 2015. In 2016, he was as a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He rejoined the Centre for Autonomous Systems, University of Technology Sydney, as a Research Fellow, in 2016. His research interest includes robotics, Internet of Things, sensor placement, artificial intelligence, machine learning, signal processing, embedded systems, and non-destructive testing.

RAVINDRA RANASINGHE (M'97) received the B.Sc. (Eng.) degree (Hons.) in computer science and engineering from the University of Moratuwa, Moratuwa, Sri Lanka, in 1995, and the Ph.D. degree in wireless communication protocols from the University of Melbourne, Parkville, VIC, Australia, in 2002.

Before joining the Centre for Autonomous Systems, University of Technology Sydney, Ultimo NSW, Australia, he was in several technology startup companies in the USA, Australia, and Sri Lanka. He is currently a Senior Research Fellow with the Centre for Autonomous Systems, University of Technology Sydney. His current research interests include perception for robotic systems, robotics and autonomous systems, machine learning, mobile robotics networks, and sensor network.