Received: 20 February 2022 | Revised: 16 April 2022 | Accepted: 23 April 2022

DOI: 10.1049/wss2.12039

IET Wireless Sensor Systems

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

LoRaWAN-implemented node localisation based on received signal strength indicator

Ibrahim Aqeel[1,2] | Ephraim Iorkyase[1,3] | Hussein Zangoti[2,4] | Christos Tachtatzis[1] | Robert Atkinson[1] | Ivan Andonovic[1]

[1]Electronic and Electrical Engineering Department, University of Strathclyde, Glasgow, UK

[2]Computer Engineering and Networks Department, Jazan University, Jazan, Saudi Arabia

[3]Electronic and Electrical Engineering Department, University of Agriculture Makurdi, Makurdi, Nigiria

[4]Department of Computing and Information Science, Florida International University, Miami, Florida, USA

Correspondence

Ephraim Iorkyase, Electronic and Electrical Engineering Department, University of Strathclyde, 11200 SW 8th St, Glasgow G1 1XW, UK.
Email: ephraim.iorkyase@uam.edu.ng

[Corrections added on 24-September-2022, after first online publication: The last author's name was corrected from Ivan Aondonovic to Ivan Andonovic.]

Abstract

Long Range Wireless Area Network (LoRaWAN) provides desirable solutions for Internet of Things (IoT) applications that require hundreds or thousands of actively connected devices (nodes) to monitor the environment or processes. In most cases, the location information of the devices arguably plays a critical role and is desirable. In this regard, the physical characteristics of the communication channel can be leveraged to provide a feasible and affordable node localisation solution. This paper presents an evaluation of the performance of LoRaWAN Received Signal Strength Indicator (RSSI)-based node localisation in a sandstorm environment. The authors employ machine learning algorithms, Support Vector Regression and Gaussian Process Regression, which turn the high variance of RSSI due to frequency hopping feature of LoRaWAN to advantage, creating unique signatures representing different locations. In this work, the RSSI features are used as input location fingerprints into the machine learning models. The proposed method reduces node localisation complexity when compared to GPS-based approaches whilst provisioning more extensive connection paths. Furthermore, the impact of LoRa spreading factor and kernel function on the performance of the developed models have been studied. Experimental results show that the SVR-enhanced fingerprint yields the most significant improvement in node localisation performance.

1 | INTRODUCTION

The decreasing cost and increasing processing capabilities of computing and communication technologies have fuelled the exponential increase in the number of interconnected devices, commonly referred to as the Internet of Things (IoT) [1]. The deployment of extensive IoT implementations is more often than not subject to the fundamental design constraints of limited resources in terms of low power consumption and low processing capabilities. The availability of Low-Power Wide Area Network (LPWAN) technologies has provisioned the characteristics aligned with the needs of these applications. Amongst the range of options, LoRa [2–4] has been widely adopted owing to an advantageous combination of features: low-cost and low power consumption with long range wireless connectivity.

Accurate node localisation is central to many beneficial applications within extensive IoT networks [5, 6]. The bulk of existing applications and services harness mature Global Navigation Satellite Systems such as the Global Positioning System (GPS) [7] or Global Navigation Satellite System [8]. Although these platforms provide accurate location estimations, their implementations are relatively expensive and more importantly in the context of IoT, prohibitively power hungry [9]. For instance, the GPS consumes between 30 and 50 mA acquiring a GPS fix which can take tens of seconds [10], largely attributable to the necessary exchanges of data. A more aligned approach is to develop network-based localisation techniques, which harness a suitable parameter inherent to data transmission as the foundation. The paper details the development and performance evaluation of a Long Range Wireless Area Network (LoRaWAN) [11, 12] enabled location estimation scheme taking advantage of the flexibility provisioned by the LoRa spectrum modulation scheme.

Inherent features of radio signals that characterise wireless networks can be used to determine the location of a node in the network [13, 14]. Signal propagation is dynamic and models the environment in which electromagnetic signals propagate is challenging. Measurable quantities such as the Received Signal Strength Indicator (RSSI) [15, 16] vary with the position of the transmitters and are time dependent. A number of reported network-based location estimations based on RSSI indicate that deterministic solutions lack accuracy because of the temporal dependence of the measurements.

Here, in order to maintain simplicity of implementation whilst meeting the needs of extensive low power deployments, a technique is developed and evaluated extending reported research that relies on the inherent relationship between RSSI and physical distance; the relationship will in turn be the basis to estimate node location from the new, unknown node RSSI value. We approach the evaluation of the proposed procedure for the estimation of node location in sandstorm environment through field trial and using LoRaWAN as an example technology. The aim of the evaluation is threefold. First, we aim at demonstrating the feasibility of transformed RSSI-based node localisation using machine learning algorithms. We do that by showing that transformed RSSI (RSSI ratio) estimation outperforms the estimation of node location based on absolute RSSI benchmark. Second, we aim at demonstrating the impact of Spreading Factor (SF) and kernel function on estimation of node location. This is done by using different SFs to gather different datasets and different kernels to evaluate the kernelised algorithms. The results demonstrate the impact of these parameters on the performance of node localisation models. Third, we aim at demonstrating the consistency of the best performing technique across different scenarios by applying a public dataset (Antwerp dataset).

The physical locations and relative elevations of nodes with respect to receivers within the operational environment are central to the solution. Thus, the approach adopts a 'fingerprinting' methodology that models transmission for the network under the environment that governs the coverage area; in this case, the 'fingerprint' is established from the comprehensive mapping of RSSI values. A number of node localisation techniques developed using RSSI-based fingerprinting for the estimation of node location in LoRaWAN and SigFox settings have been reported and summarised in Table 1. The base models have been enhanced through the use of machine learning in order to increase node location accuracy.

In Ref. [13], the fingerprint was enhanced through K-Nearest Neighbour methods to achieve a mean distance error of 689 m (Sigfox) and 398 m (LoRaWAN) from 84 to 68 base stations, respectively. However, neither the location of the base stations was provided nor the SF was used in the development of LoRaWAN-based technique. Authors in Ref. [14] focussed on the development of an outdoor parking positioning system for a restricted coverage area (340 × 340 m) utilising 4 LoRaWAN base stations transmitting at an SF of 7. The Maximum Likelihood method achieved a mean distance error of 24 m. Gaussian Process Regression (GPR)-based fingerprinting for localisation [17] achieved a mean distance error of 25 m in a campus outdoor area (150 × 250 m), utilising 10 LoRaWAN base stations transmitting at an SF of 12. The latter studies focussed on solutions for relatively modest outdoor coverage areas.

In summary, previous studies have used empirical RSSI measurement for node localisation in moderate coverage areas. However, the idea of transforming RSSI measurement into average RSSI ratio by pairs of gateways as input fingerprint, combined kernel function and high SF can be harnessed to optimise and improve estimation of node location in more extensive coverage areas (in the order of kms).

The remainder of the paper is organised as follows; Section 2 details the data gathering infrastructure, measurement methodology, the coverage area under consideration and environmental conditions; Section 3 describes the establishment of the 'fingerprint'; Section 4 details the enhancement to the fingerprinting technique owing to the application of two kernel-based machine learning techniques, Support Vector Regression (SVR) and Gaussian Process Regression (GPR); Section 5 presents an evaluation of the node location performance of the proposed approaches; finally, Section 6 draws conclusions.

2 | MEASUREMENT METHODOLOGY

Measurements are executed in Jazan City in Saudi Arabia to capture the radio propagation of LoRa nodes in sandstorm condition. Figure 1 shows the map illustrating the location of the gateways, deployed across a semi-uniform grid given that the terrain is characterised by buildings and natural obstacles such as trees. Gateways/receivers are positioned on the outskirts of the Jazan City on four elevated structures with their respective elevations provided in Table 2. The four gateways (Figure 1; black circles) were located at points 4–7 km around the coverage area containing the transmitter nodes at varying locations. Although we used only four gateways to demonstrate our proposed method. This can be easily scaled to any number of gateways.

Gateways are placed at elevated positions to extend the range of the network that would otherwise be impaired due to buildings and natural obstacles. The transmitter node is fixed when taking measurements and moved between grid points within the coverage area. Measurements were taken from 150 locations (viz. grid points) using different SFs. The distance between grid points is approximately 100 m. The closest measurement is taken at a distance of 4 km and the furthest 7 km. 20 RSSI packets are recorded from each measurement location at different SFs, referred to as SF9, SF10, SF11 and SF12, respectively; a total of 3000 measurements were acquired for each SF. Each packet comprised GPS location coordinates as a payload with gateways issuing an acknowledgement on the successful receipt of the payload. The measured RSSIs at each gateway were uploaded to The Things Network (TTN) server along with the payload information.

The data acquisition system consisted of four LoRaWAN transceiver gateways and transmitter accessing the Internet

TABLE 1 Related work in location fingerprinting

Reference	Model	Test environment	Technology	No. Of (GWs)	Spread factor	Mean error (m)
(Aernouts, et al., 2018) [13]	KNN	Outdoor (52 km^2)	SigFox	84	---	689
(Aernouts, et al., 2018) [13]	KNN	Outdoor (52 km^2)	LoRaWAN	68	SF = 7–12	398
(Choi, et al., 2018) [14]	Maximum Likelihood	Outdoor parking (340 × 340 m)	LoRaWAN	4	SF = 7	24
(Zhe, et al., 2019) [17]	Gaussian process	Outdoor (150 × 250 m)	LoRaWAN	10	SF = 12	25

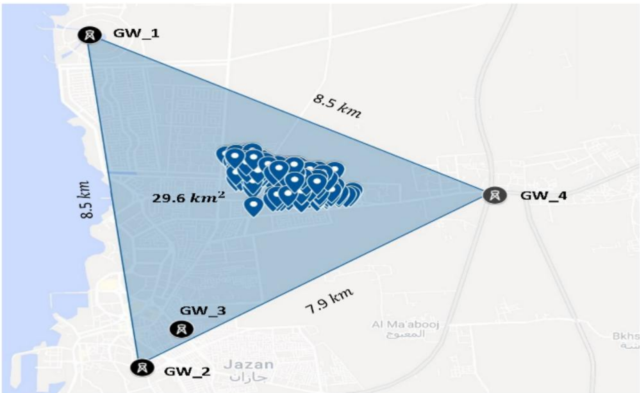


FIGURE 1 Map illustrating the location of gateways (black circles)

TABLE 2 Location of gateways

Gateway	Location	Height (m)
GW1	Top of University tower	100
GW2	Communication tower (1)	90
GW3	Communication tower (2)	70
GW4	Top of Water tower	40

through laptops. The gateways comprise iC880ASPI LoRaWAN 868 MHz concentrators connected to a Wi-Fi enabled host (Raspberry Pi 3 Model B SBC platform with 16 GB micro-SD card) via a SubMiniature Version A antenna of 2 dBi and are housed in Acrylonitrile butadiene Styrene Enclosures with mains electrical power supply. The enclosures are designed to guarantee operation between −5°C and +55°C, meeting the requirements of the operational environmental conditions. Gateways are the data collectors of the architecture utilising 868 MHz channels for data transmission. Packets can be received from different nodes with different SFs, up to 8 channels in parallel. Gateways are also equipped with an external control microprocessor and an RPi 3 unit is connected to the Institute for Mobile & Satellite Communication Technology concentrator via the Serial Peripheral Interface bus. The RPi 3 is Wi-Fi enabled and connected to 4G connectors in order to receive and transmit data to the server ('TTN' server). Transmitter nodes are a Sodaq One v2 LoRaWAN device with an 868 MHz antenna of 3dBi connected to a GPS module (Ublox Eva 7 M). The node consists of an RN2483 transceiver with 14 dBm transmission power and a bandwidth of 125 kHz powered by an 800 mAh lithium battery.

2.1 | Testbed environment

Two test beds were designed for this measurement campaign to capture the radio propagation of LoRa nodes in sandstorm condition for node localisation as a function of SF. In the first instance, RSSI measurements were taken to characterise the propagation of LoRa nodes in sandstorm environment as compared to clear sky. Figure 2 shows the two environmental conditions: clear sky and sandstorm in the city of Jazan. In this testbed, measurements were taken at locations positioned 100 m away from each other up to 3 km. At each location, the transmitter transmits more than 10 packets and were received by the gateway at a SF of 7. The gateway was placed on the roof of a stationary car (approximately 2 m above the ground). The location of the transmitter node was taken with reference to the gateway. The transmitter node was placed in a car (approximately 1 m above ground level) and moved to the pre-defined locations until all measurements were taken. Second, measurements were taken to validate the proposed node localisation technique. In this case, four gateways are located on the outskirts of the urban area, and the transmitters were located in the rural environment as shown in Figure 1. The propagation path between the test area and the gateways is characterised by buildings of different elevations (9–30 m). In fact, the experimental environment (sandstorm) in this work can be termed a semi-urban environment.

Both measurements were acquired during the monsoon winds in the months of July and August. The wind speed is the most important environmental factor that impacts signal propagation in this context. It is reasonable to expect that as the strength of the wind increases, the density of the perturbed and sand particles increases and the impact on the propagation of the radio signals becomes more significant and time dependent. Table 3 represents the weather conditions in the month of July and August, when measurements were taken. The most challenging season is characterised by dust, high temperature and humidity. Apart from the climatic factors, the radio environment is also characterised by trees and buildings, which can create challenges.

2.2 | LoRa transmitter distance estimation

In order to evaluate the performance of LoRa link for long range transmission and transmitter distance estimation for device location in sandstorm environment, the Two Ray Ground Reflection Model is used to estimate transmitter-receiver distance from known measured LoRaWAN RSSI. The two-ray ground reflection model is used in this work because it provides better prediction at long distances compared to other ready-made models [6]. The average RSSI values are used as representative samples to estimate the distance between the transmitter and the receiver. Figure 3 shows the variation in the actual and estimated distances with respect to measured RSSI. As can be seen, there is significantly more attenuation to the signal strength in sandstorm conditions compared to clear sky. However, in both situations, the signal attenuation appears to plateau within a certain range of greater distances.

For clear condition, the RSSI sensitivity value is approximately −106dBm at 600 m and greater, and for sandstorm condition, RSSI sensitivity value is approximately −112dBm at 900 m and greater. Consequently, a large proportion of estimated distances for clear and sandstorm conditions 'bunch up' at 600–700 m and 900–1000 m, respectively.

However, in sandstorm condition, the model could produce inaccurate estimates even at shorter distances of 200 m. The estimated distances later bunch up and are significantly greater or less than actual distances with significant error. In clear condition, reasonable estimates are obtained only up to 600 m after which the estimated distances seem to cluster around 650 m, which is far less than the actual distances. It can be concluded that the use of ready-made propagation model with RSSI measurements to determine distances, and hence position of LoRa devices under clear sky condition is grossly inadequate for the long-range application of LoRaWAN for IoT. Therefore, we will investigate the use of location fingerprint for node location in sandstorm environment in the next section.


(a) clear (b) sandstorm

FIGURE 2 The two environmental conditions. (a) clear and (b) sandstorm

3 | DATA PREPARATION AND FINGERPRINTS

During the experiment, 20 packets were transmitted from each of the 150 designated locations and were expected to be received by the LoRaWAN gateways deployed in the vicinity of the experimental environment. The vector of absolute RSSI values received by the LoRaWAN gateways in the experimental environment is used to develop node localisation models. The calculated average values of RSSI of the 20 received packets at each grid location represent the fingerprint of each location. Figure 4 shows the RSSI pattern at various points in the radio environment for each of the LoRa gateways. The figures reveal the complexity of the radio environment (sandstorm), which does not fit any well-known propagation model. The complexity of the signal attenuation with distance is a result of noise and distortions. In case of a missing RSSI value in an observation, we substitute the missing value by using mean imputation method [18, 19], which increases the amount of information that can be used, and hence, improves the performance of node localisation models (as discussed in Section IV). The procedure for mean imputation method used in substituting missing RSSI values is as follows:

- Separation of each group of 20 packets by location.
- examination of all data for each gateway.
    - In the case of loss of all RSSI data at a specific gateway (referred to as 'Monotone') with the same location, the missing RSSI values are replaced with a specified value.
    - In the case of missing data at a specific gateway (referred to as 'Non-Monotone') with the same location, the mean of measured RSSI values (not Null) of the G_j for each location is calculated, where G_j denotes the number of the gateways. The missing values are therefore replaced with the mean value in G_j for each location.

The resulting data contains six columns [RSSI_Gateway1, RSSI_Gateway2, RSSI_Gateway3, RSSI_Gateway4, Longitude, and Latitude] and 3000 rows of absolute RSSI values (20 packets at 150 locations).

In a challenging radio environment, characterised by reflections and obstructions such as the one under consideration, the dynamic variations of absolute RSSI values with time introduce noise in the fingerprints and may impair the performance of node localisation. In this paper, we propose to derive robust fingerprints by taking ratios of RSSI values between gateway pairs in order to mitigate the variations in absolute RSSI values.

	Humidity (%)	Temperature (°C)	Wind speed (km/h)
Dust sky	60–85	35–45	5–10
Sandstorm sky	60–85	35–45	13–27
Strong sandstorm	60–85	35–45	37

TABLE 3 Testbed environmental condition

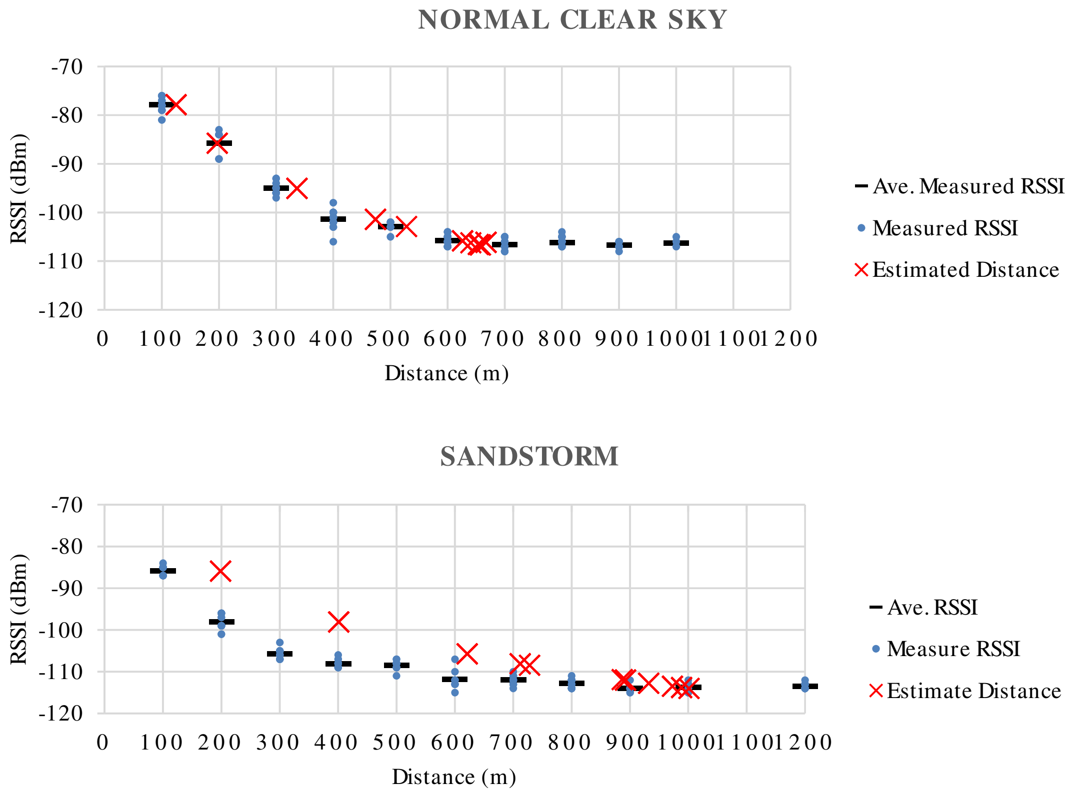


FIGURE 3 Measured and estimated transmitter distance

Assume $G = \{g_1, g_2, \ldots, g_n\}$ is a set of gateways deployed in the area under consideration, and $L = \{l_1, \ldots, l_m\}$ represents the reference node locations. The location feature space, l_i, can then be represented by gateways and measured absolute RSSI values $r \in R$ where $R = \{r_1, r_2, \ldots, r_n\}$. The RSSI ratio is defined at each location for a unique pair of gateways. The received signal strength ratio for the gateways g_i and g_j can be computed for measurement taken at location l = [(g_i; r_i); (g_j; r_j)] as in Equation (1).

$$RSSI_{\text{ratio}}(g_i, g_j) = \frac{r_i}{r_j} \tag{1}$$

With $i < j$ for uniqueness, where r is absolute RSSI value.

The total samples of the $RSSI_{ratio}(g_i, g_j)$ is 3000 × 8 (6 columns for the $RSSI_{ratio}$ between pair of gateways and 2 for location coordinates).

The mean of the RSSI ratios for each location is computed as given in Equation (2).

$$Mean\ RSSI_{ratio} = \frac{\sum_{i,j=1}^{n=20} \frac{r_i}{r_j}}{n} \tag{2}$$

where $g_{i,j}$ denotes the number of unique pair of gateways that measures the signal strength of the node at location l_i.

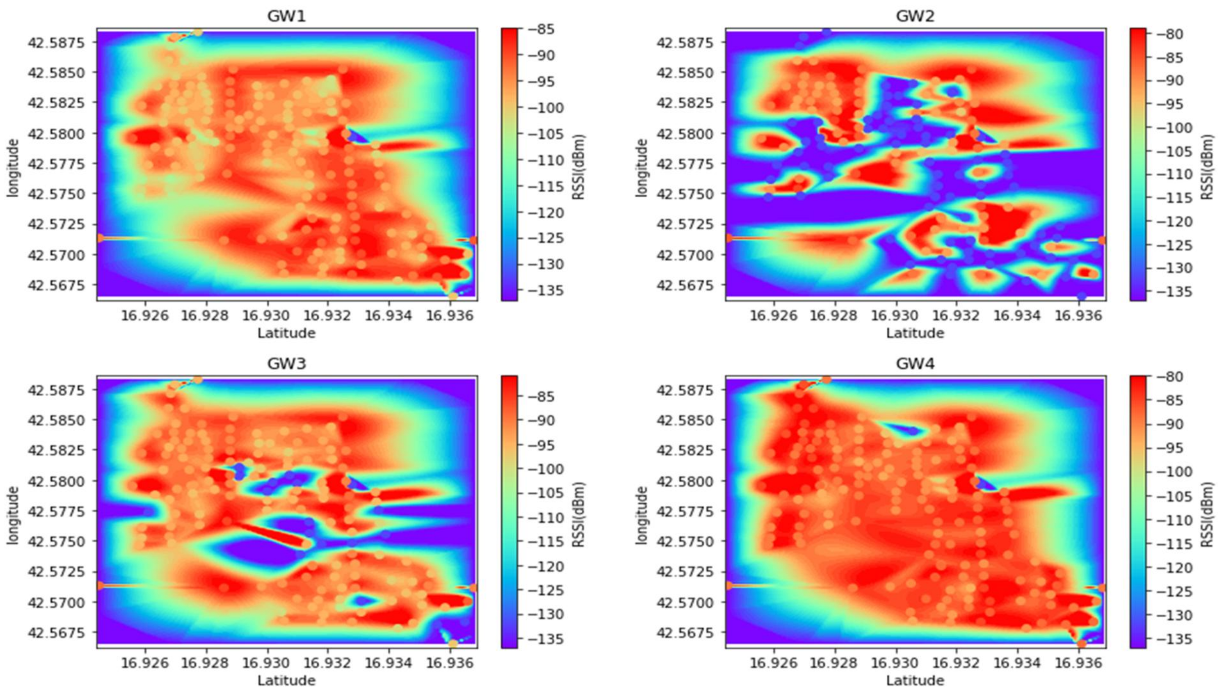
Mean RSSI ratios will be used in the subsequent analysis. The proposed node localisation technique is shown in Figure 5. It is important to note that in machine learning technique, separate datasets are needed to train and validate the model. Here, the RSSI_Ratios/location data collected during experiment from 150 locations is randomly divided into training and test sets. A total of 120 × 8 (the RSSI_Ratios between pairs of gateways) randomly selected RSSIs with reference locations are used for training the models and 30 × 6 remaining RSSI_Ratios without reference locations are used to validate the developed models.

4 | KERNEL-BASED LOCALISATION

4.1 | Support Vector Regression

Support Vector Regression (SVR) [16, 21, 22], dedicated to regression problems, is a variant of the well-known Support Vector Machine (SVM) technique. SVR uses the same principle as SVM [23, 24] for classification, mapping the data into a high dimensional feature space using non-linear transformations; linear regression is then executed in this space. Kernel functions perform the non-linear transformation of the data into higher dimensional feature space that then enables the linear separation. Effectively, linear regression in a high dimensional space corresponds to non-linear regression in the low-dimensional input space [25]. Invariably, regression methods derive a function, say $f(x)$, with the least deviation between predicted and observed output for all training data. Further, SVR minimises the influence of the error in the observed data by establishing boundary margins around the

(a) RSSI map for SF9



(b) RSSI map for SF10

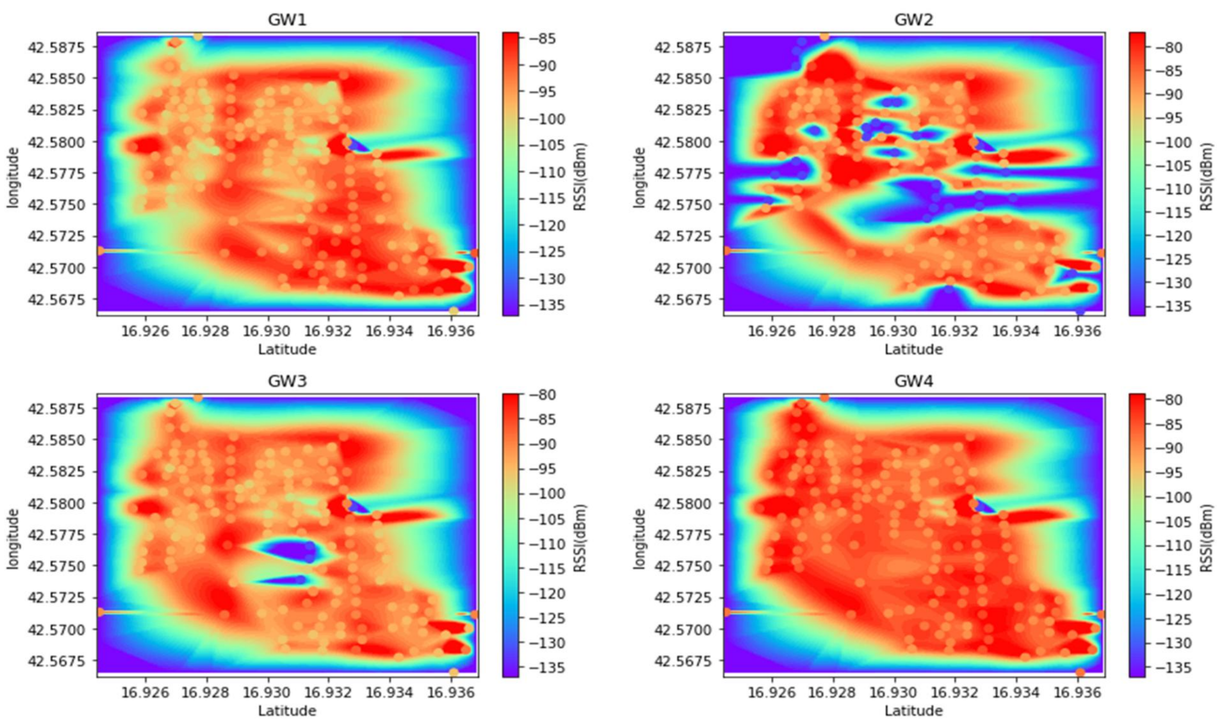


FIGURE 4 Spatial spread of Received Signal Strength Indicator (RSSI) for (a) SF9, (b) SF10, (c) SF11 and (d) SF12

hyperplane outside which data is not considered for regression. The prediction becomes challenging given that the SVR output is a real number. Consequently, a tolerance margin, epsilon is set in approximation to the SVM.

Two basic types of SVR are applied—epsilon-SVR and nu-SVR [26–28]—differentiated by the manner the parameters therein are managed. The main attribute of SVR is the use of a non-linear kernel transformation to map the input variables

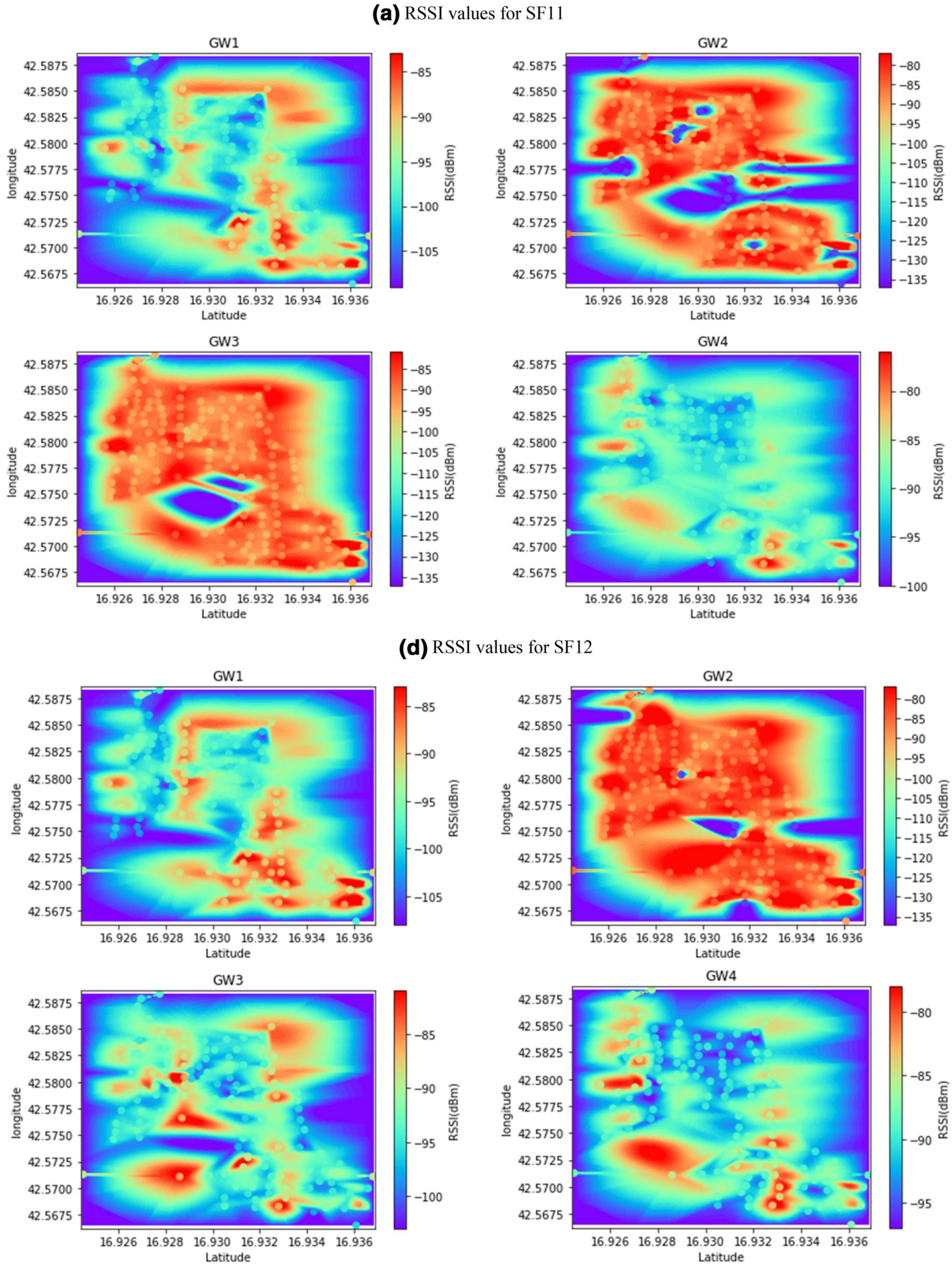


FIGURE 4 (Continued)

into a feature space such that the relation with the output variable becomes linear in the transformed space. Second, SVR's excellent generalisation capabilities result from the use of non-linear kernels with good approximation. Also, SVR does not suffer from local minima problem because it possesses convex optimisation formulation. It can better solve small samples and non-linear dimensional problems. The linear case of SVR is modelled as given in Equation (3).

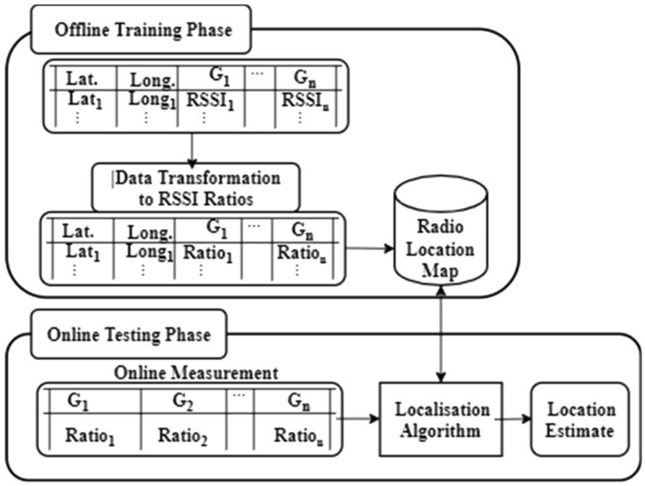


FIGURE 5 Node localisation technique

$$f(x) = \langle w \cdot x \rangle + b \tag{3}$$

SVR can be formulated as a convex optimisation as given in Equation (4).

$$\text{Minimise} \quad \frac{1}{2}\|w\|^2$$
$$\text{Subject to} \quad \{y_i - \langle w \cdot x_i \rangle - b \le \varepsilon \; \langle w \cdot x_i \rangle + b - y_i \le \varepsilon \tag{4}$$

ε is the acceptable deviation of estimated locations from actual location.

An implicit assumption is that the function $f(x)$ can approximate all input pairs (x_i, y_i) with precision ε , that is, it is assumed that optimisation is feasible. Therefore, in order to accommodate errors, slack variables ξ_i, ξ_i^* are introduced to cope with otherwise infeasible optimisation constraints given in Equation (4) [29], where the constant $C > 0$ determines the degree to which deviations larger than ξ are tolerated with l being the number of samples as in Equation (5).

$$\text{Minimise} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$
$$\text{Subject to} \quad \{y_i - \langle w \cdot x_i \rangle - b \le \varepsilon + \xi_i \; \langle w \cdot x_i \rangle + b$$
$$- y_i \le \varepsilon + \xi_i^* \; \xi_i, \xi_i^* \ge 0 \tag{5}$$

ξ_i, ξ_i^* are the slack variables that make allowance for the localisation errors to exist up to the value of ξ_i and ξ_i^* without degrading performance. C is the box constraint, a positive numeric value that controls the penalty imposed on data points that lie outside the ε margin and helps to prevent overfitting.

A standard dualisation method with Lagrange multipliers α_i, α_i^* is used [30] to solve Equation (5), with ω expressed as in Equation (6).

$$\omega = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)x_i \tag{6}$$

where $\alpha_i \ge 0$ and $\alpha_i^* \ge 0$.

Substituting Equation (6) into Equation (3) and Equation (5) produces Equation (7).

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\langle x_i \cdot x \rangle + b \tag{7}$$

The dot product of the input vectors can be replaced with their non-linear transformation, the kernel function, represented by $k(x_i, x)$ to form the non-linear solutions given in Equation (8).

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)k(x_i, x) + b \tag{8}$$

Kernel functions make the SVR applicable to both linear and non-linear approximations. SVRs yield an acceptable generalisation performance as only the support vectors are used for prediction and are based on structural risk minimisation that seeks to minimise the generalisation rather than the training error [31].

4.2 | Gaussian Process Regression

The Gaussian Process (GP) is a probabilistic kernel-based technique that has been applied in many practical problems including estimation, classification, prediction, and prognosis due to its advantage of being flexible, probabilistic, and non-parametric [17, 26]. A GP can model any system or process according to a normal or Gaussian distribution, where the mean and covariance function depend on the training data; the process is a collection of random variables with a joint Gaussian distribution [32]. Thus, any function sample has a Gaussian distribution defined by its mean function $m(x)$ and covariance function $k(x, x')$.

The model assumes that the output is a realisation of a GP with joint probability density function as given in Equation (9).

$$f(x) \sim GP(m(x), k(x, x')) \tag{9}$$

where, $m(x) = E(f(x))k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))] \;\; k(x, x') = cov(f(x), f(x'))$

Here, the GP method is applied to a regression problem.

Assuming $X = [x_1, x_2, \ldots x_N]$ represents N by 6-dimensional RSSI ratio input vectors, and the corresponding outputs are $y = [y_1, y_2, \ldots y_N]$, representing the dual location coordinates. When a new input vector x^* is given, the goal is to predict the corresponding output y^* (unknown location coordinates). The spatial relationship between the input variable and the expected output can be modelled as a GP by Equation (10).

TABLE 4 Optimal parameters used for epsilon_SVR algorithm

Epsilon_SVR kernels				
Spreading factors	**Parameter tuning**	**RBF+Matren**	**RQ+Matern**	**Exp.+Matern**
SF9 RSSI Ratio	Epsilon	1.00E-03	1.00E-04	1.00E-03
	Nu_Matern	10	1.5	1
	Median	444	453	477
SF10 RSSI Ratio	Epsilon	1.00E-03	1.00E-10	1.00E-05
	Nu_Matern	4	8.5	8
	Median	359	381	410
SF11 RSSI Ratio	Epsilon	1.00E-03	1.00E-03	1.00E-03
	Nu_Matern	1	1	1
	Median	313	303	326
SF12 RSSI Ratio	Epsilon	1.00E-06	1.00E-06	1.00E-03
	Nu_Matern	3	14.5	8.5
	Median	323	346	329

TABLE 5 Optimal parameters used for Gaussian Process Regression (GPR) algorithm

GPR kernels				
Spreading factors	**Parameter tuning**	**RBF+Matren**	**RQ+Matern**	**Exp.+Matern**
SF9 RSSI Ratio	Alpha	1.00E-05	1.00E-06	1.00E-05
	Nu_Matern	1.5	4	2.5
	Median	431	423	430
SF10 RSSI Ratio	Alpha	1.00E-06	1.00E-08	1.00E-05
	Nu_Matern	8	11.5	6
	Median	378	388	380
SF11 RSSI Ratio	Alpha	1.00E-09	1.00E-08	1.00E-09
	Nu_Matern	1	11	1
	Median	392	387	392
SF12 RSSI Ratio	Alpha	1.00E-07	1.00E-07	1.00E-05
	Nu_Matern	9	10	8
	Median	317	361	379

TABLE 6 Optimal parameters used for Nu-SVR algorithm

Nu_SVR kernels				
Spreading factors	**Parameter tuning**	**RBF+Matren**	**RQ+Matern**	**Exp.+Matern**
SF9 RSSI Ratio	Nu_SVR	0.44	0.45	0.09
	Nu_Matern	2	7	1
	Median	410	399	425
SF10 RSSI Ratio	Nu_SVR	0.97	0.27	0.22
	Nu_Matern	3	11	9
	Median	357	353	366
SF11 RSSI Ratio	Nu_SVR	0.75	0.39	0.47
	Nu_Matern	1	1	1
	Median	336	338	338
SF12 RSSI Ratio	Nu_SVR	0.55	0.86	0.78
	Nu_Matern	1	1.5	1.5
	Median	320	324	309

TABLE 7 Statistical performance of the epsilon-SVR

		Exp.+Matren	RBF+Matern	RQ+Matern
Absolute RSSI_ SF 9	Min (m)	121	145	130
	Median (m)	487	442	421
	Mean (m)	536	493	483
	RMSE (m)	603	556	542
Absolute RSSI_SF10	Min (m)	123	109	93
	Median (m)	541	506	428
	Mean (m)	551	559	513
	RMSE (m)	619	656	607
Absolute RSSI_SF11	Min (m)	83	117	127
	Median (m)	520	410	426
	Mean (m)	539	575	537
	RMSE (m)	609	660	613
Absolute RSSI_ SF12	Min (m)	110	128	80
	Median (m)	508	336	344
	Mean (m)	537	441	392
	RMSE (m)	604	508	446

Abbreviation: RMSE, Root Mean Square Error.

TABLE 8 Kernel functions

Kernel	Formula
Radial basis function (RBF)	$k(\boldsymbol{x}, \boldsymbol{x}') = e^{\left(-\frac{\boldsymbol{x}-\boldsymbol{x}'^2}{2\sigma^2}\right)}$
Rational Quadratic (RQ)	$k(\boldsymbol{x}, \boldsymbol{x}') = \left(1 + \frac{d(\boldsymbol{x},\boldsymbol{x}')^2}{2\alpha l^2}\right)^{-\alpha}$
Matern	$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma^2 \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\gamma\sqrt{2\nu d}\left(\frac{\boldsymbol{x}}{l}, \frac{\boldsymbol{x}'}{l}\right)\right)^{\nu} k_v\left(\gamma\sqrt{2\nu d}\left(\frac{\boldsymbol{x}}{l}, \frac{\boldsymbol{x}'}{l}\right)\right)$
ExpSineSquared	$k(\boldsymbol{x}, \boldsymbol{x}') = \exp\left(-2\left(\sin\left(\frac{\pi}{p} * d(\boldsymbol{x}, \boldsymbol{x}')\right)/l^2\right)^2\right)$

TABLE 9 Performance of different algorithms based on Rational ExpSineSquared + Matern Kernel Function

Models		SF9 RSSI Ratio (m)	SF10 RSSI Ratio (m)	SF11 RSSI Ratio (m)	SF12 RSSI Ratio (m)
Epsilon-SVR	min	90	108	84	97
	median	477	410	326	329
	mean	637	498	453	393
	RMSE	767	559	588	447
Nu-SVR	min	140	56	111	116
	median	425	366	338	309
	mean	541	432	474	404
	RMSE	609	498	612	459
GPR	min	36	59	32	62
	median	430	380	392	379
	mean	503	450	488	424
	RMSE	576	508	595	481

Abbreviation: RMSE, Root Mean Square Error.

TABLE 10 Performance of different algorithms based on Radial basis function (RBF) + Matern Kernel Function

Models		SF9 RSSI Ratio (m)	SF10 RSSI Ratio (m)	SF11 RSSI Ratio (m)	SF12 RSSI Ratio (m)
Epsilon-SVR	min	75	45	84	42
	median	444	359	313	323
	mean	532	449	453	385
	RMSE	629	507	583	440
Nu-SVR	min	55	64	129	83
	median	410	357	336	320
	mean	518	452	454	396
	RMSE	600	514	567	443
GPR	min	40	45	32	49
	median	431	378	392	317
	mean	502	454	488	425
	RMSE	572	526	595	491

Abbreviations: GPR, Gaussian Process Regression; RMSE, Root Mean Square Error; SVR, Support Vector Regression

TABLE 11 Performance of different algorithms based on Quadratic + Matern Kernel Function

Models		SF9 RSSI Ratio (m)	SF10 RSSI Ratio (m)	SF11 RSSI Ratio (m)	SF12 RSSI Ratio (m)
Epsilon-SVR	min	59	66	85	70
	Median	453	381	303	346
	mean	571	451	451	378
	RMSE	694	509	573	431
Nu-SVR	min	119	57	109	121
	Median	399	353	338	324
	mean	508	440	463	391
	RMSE	575	503	555	444
GPR	min	49	52	35	22
	Median	423	388	387	361
	mean	489	447	490	385
	RMSE	560	506	678	433

Abbreviations: GPR, Gaussian Process Regression; RMSE, Root Mean Square Error; SVR, Support Vector Regression

$$y_i = \varphi(x_i; W) + \epsilon, \ \epsilon \sim N(0,), \ i = 1, \ldots, N \tag{10}$$

where φ is a function parameterised by vector W; ϵ is assumed to be the noise caused by perturbations represented by a distributed Gaussian distribution N with zero mean and variance σ_n^2.

The prior probability on y is given by Equation (11).

$$\begin{aligned} E[y] &= E[\varphi(x; W) + \epsilon] = 0 \\ cov[y] &= K(X, X) + \sigma_n^2 I \end{aligned} \tag{11}$$

where E is the mean function, and cov is the variance function. The distribution with the new input can be expressed by the function in Equation (12).

$$[y \, y^*] \sim GP\left(0, \left[K(X, X) + \sigma_n^2 I \, K(X, x^*) \, K(X, x^*)^T \, K(x^*, x^*)\right]\right) \tag{12}$$

where $K(X, x^*) = [k(x_1, x^*), \ldots k(x_N, x^*)]$ can be written as k^*. The prediction can be presented by Equations (13) and (14).

$$E(y^*) = k^{*T}\left(K + \sigma_n^2 I\right)^{-1} y^T \tag{13}$$

$$cov[y^*] = K(x^*, x^*) - k^{*T}\left(K + \sigma_n^2 I\right)^{-1} K^* \tag{14}$$

5 | EXPERIMENTAL RESULTS

The experimental results are presented in this section. The offline measurements were taken in the suburb region of Jazan City in Saudi Arabia. The testbed considered is an environment characterised with sandstorms, tall buildings, masts and towers. The testbed area was divided into a semi-uniform grid with side

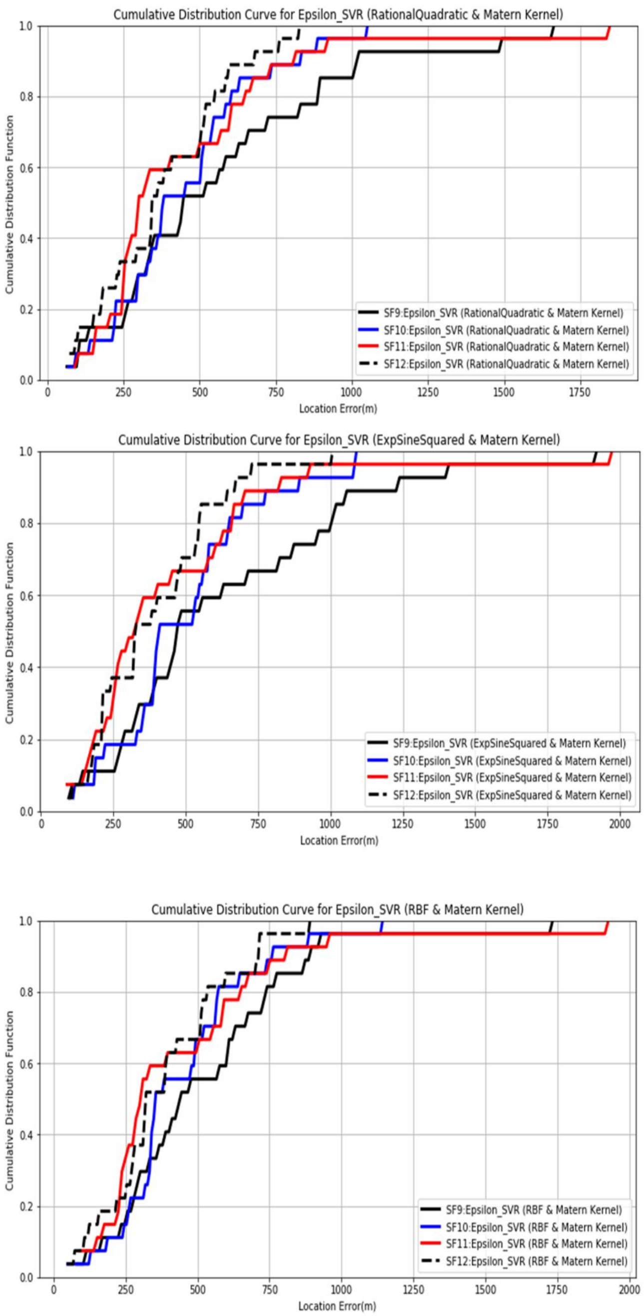


FIGURE 6 Cumulative Distribution Function (CDF) for epsilon_SVR models using combined kernels

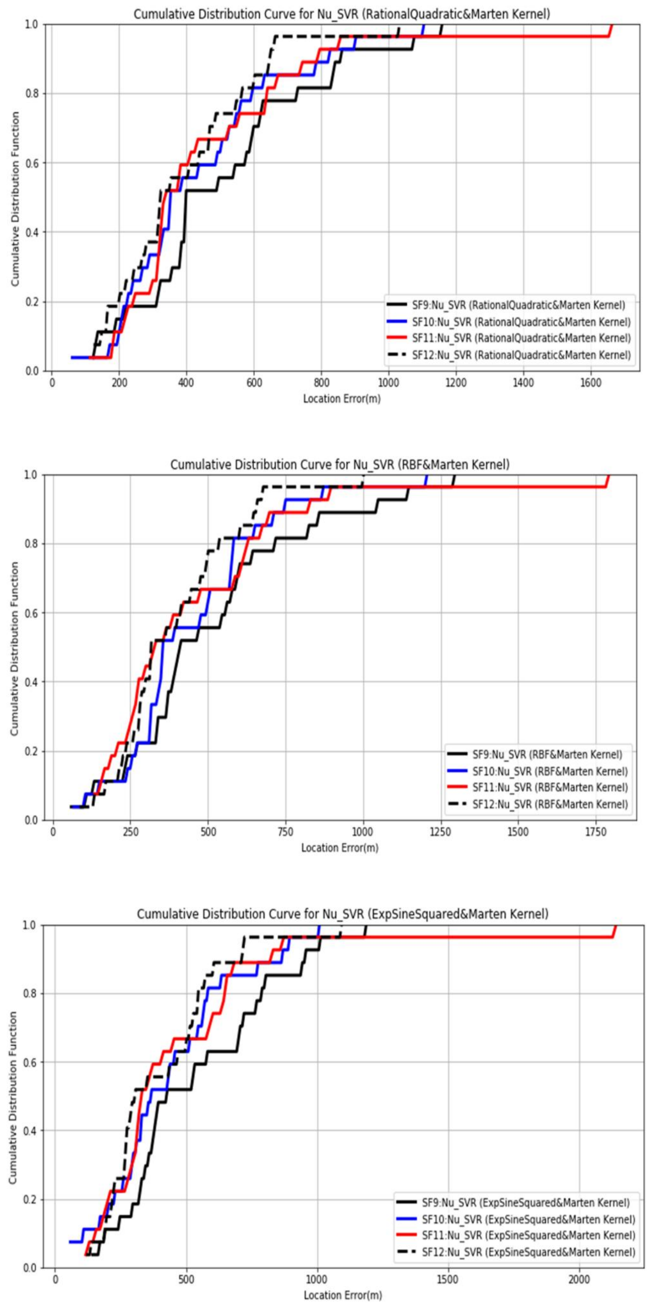


FIGURE 7 Cumulative Distribution Function (CDF) for Nu_SVR models using combined kernels

measurement of 100 m to form 150 measurement locations. The training data was measured at 130 randomly selected locations where 20 time samples of RSSI were measured at each measurement location. For locations with null readings, a default RSSI value of $-132dB$ was used as valid data. The data collected was used to analyse the localisation performance of the developed models. The impact of the kernel functions, SF and transformed RSSI features was analysed. In addition, the time complexity of the models was investigated. The following metrics are used in result analysis: Haversian distance metric, Root Mean Square Error and Cumulative Distribution Function (CDF).

5.1 | Parameter tuning

The hyperparameters associated with the machine learning algorithms impact the overall performance of models; thus, central is the tuning of parameters to optimise their accuracy.

Hyperparameters are tuned for each dataset namely RSSI ratios of SF9, SF10, SF11, and SF12. The optimal model hyperparameters are unique to a single dataset. A random search method is used to select the optimal parameters of the epsilon-SVR, nu-SVR and GPR algorithms. A grid of hyperparameters values is established, and a random combination of the values is selected to train the model. Moreover, for SVR,

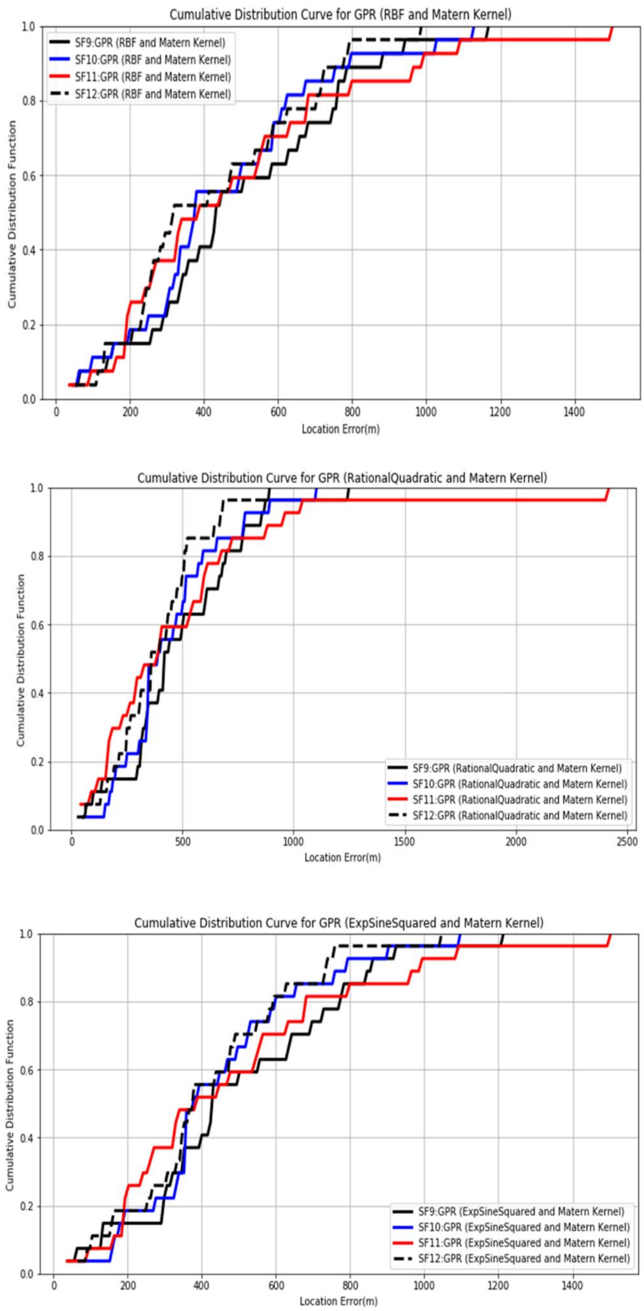


FIGURE 8 Cumulative Distribution Function (CDF) for Gaussian Process Regression (GPR) models using combined kernels

TABLE 12 Antwerp city dataset

Gateways	SF9	SF10	SF11	SF12
No. of packets	6254	4559	2708	1430

hyper-parameter C, regularisation constant, epsilon and nu for nu-SVR are also optimised using the same methodology. For GPR, the only hyperparameter to be tuned is alpha. Some kernels such as 'Matern' have optimised parameters. The summary of the optimal parameters used in each algorithm for each dataset is given in Table 4, Table 5, and Table 6.

TABLE 13 Model performance on Antwerp data (ratios)

	Model	Median (m)	Mean (m)	Time
SF 9_Ratio RSSI	RBF + Matern	708	871	199 min
	RQ + Matern	707	872	198 min
SF 10_Ratio RSSI	RBF + Matern	684	859	105 min
	RQ + Matern	688	858	108 min
SF 11_Ratio RSSI	RBF + Matern	614	784	38.5 min
	RQ + Matern	616	786	38.3 min
SF 12_Ratio RSSI	RBF + Matern	433	660	10 min
	RQ + Matern	447	660	10.5 min

5.2 | Impact of transformed RSSI features

To evaluate the impact of the transformed fingerprints (RSSI ratio), the SVR methods are first evaluated using absolute RSSI values for all spreading factors (SF9, SF10, SF11, and SF12). Table 7 shows the statistical performance of the models when absolute RSSI data is used. The median localisation error using absolute RSSI features with SF12 is 336 m. On the other hand, epsilon_SVR with SF11 using RSSI_ratio provides the best median localisation error of 303 m as shown in Table 7, enhancing precision by 28.8% over using absolute RSSI with SF11. The transformed data (RSSI_Ratio) have shown to improve the accuracy of the node localisation. We believe that the improved performance of the developed node localisation model using transformed data (RSSI ratio) is because the average RSSI ratio reduces the noise in the absolute RSSI data.

5.3 | Impact of kernel functions

In machine learning, a kernel is used to transform linearly inseparable data to linearly separable data. In effect, kernel functions compute similarities between samples in the data. A range of kernel functions are used in the establishment of SVR- and GPR-enhanced localisation models. In addition, different kernels are combined in order to further investigate the effect of kernel functions on the performance of the models. The kernels used in the evaluation are given in Table 8 [32] [33] [34].

RSSI ratio data and the corresponding location coordinates are used as training inputs to the algorithm. Whilst the data used for training remained constant, the kernel function was varied in order to test its impact on performance. Results for each algorithm, epsilon-SVR, nu-SVR and GPR with combined kernel functions are shown in Table 9, Table 10 and Table 11. It is evident that the combined kernel functions outperformed the commonly used kernels on the same dataset for all three algorithms. More specifically, the Rational Quadratic + Matern kernel has the lowest median error of 303 m with the epsilon-SVR algorithm; in other words, the model locates the node with error less than 303 m for 50% of the time. The median location error is 309 m for

TABLE 14 Model performance on Antwerp data (absolute)

	Model	Median (m)	Mean (m)	Time
SF 9__Absolute RSSI	RBF + Matern	808	895	155 min
	RQ + Matern	781	894	157 min
SF 10__Absolute RSSI	RBF + Matern	776	951	32 min
	RQ + Matern	750	924	31 min
SF 11__Absolute RSSI	RBF + Matern	666	837	32 min
	RQ + Matern	625	807	32.89 min
SF 12__Absolute RSSI	RBF + Matern	580	735	8.27 min
	RQ + Matern	541	750	8.73 min

TABLE 15 Model performance of support vector regression (SVR) on Jazan data (ratios)

	Models	Median (m)	Mean (m)	Time
SF 9 Ratio RSSI	RBF + Matern	444	532	3.74 min
	RQ + Matern	453	571	3.21 min
SF 10 Ratio RSSI	RBF + Matern	359	449	3.4 min
	RQ + Matern	381	451	3.45 min
SF 11 Ratio RSSI	RBF + Matern	313	453	3.31 min
	RQ + Matern	303	451	3.28 min
SF 12 Ratio RSSI	RBF + Matern	323	385	3.38 min
	RQ + Matern	346	378	3.34 min

ExpSineSquared + Matern kernel in nu-SVR. In GPR, Radial basis function (RBF) + Matern kernel gives a median error of 317 m.

5.4 | Impact of spreading factor

The impact of SF on the performance of the models developed is evaluated using average RSSI ratios at different SFs (9, 10, 11 and 12) as input fingerprints to SVR and GPR. The results presented in Table 9, Table 10 and Table 11 indicate that higher spreading factors (SF11 and 12) yielded improved node localisation performance compared to lower spreading factors (SF9 and 10). SF11 and SF12 derived models produce the highest level of consistency irrespective of the combined kernels used. More specifically, epsilon-SVR at SF11 provides a median error of 303 m, a 30% improvement in precision compared to the performance at SF9 (453 m). The significant improvement at higher SF could be attributed to the quality of data collected. It has been observed in the reported experiment that the quality of data is a function of the SF used. Whilst we experienced significant loss of packets at SF9 and SF10, there was little or negligible loss of data at SF11 and 12. It should be noted that at higher SFs, latency is a consideration as the transfer of packets is subject to significant delays. However, the trade-off between latency and accuracy in this application may be a design option. Shadowing and reflections are more likely to impact reception at low SF values.

5.5 | Accuracy

Here, the accuracy of the models is measured as the average Haversian distance metric between the estimated and true location of a node as given in Equation (15):

$$d = 2r\left(\sqrt{\left(\frac{\varphi - \varphi_0}{2}\right) + cos\, cos\,(\varphi_0)\, cos\, cos\,(\varphi)\left(\frac{\lambda - \lambda_0}{2}\right)}\right) \tag{15}$$

where, $\varphi_0 = latitude\ of\ real\ location$, $\varphi = latitude\ of\ estimated\ location$, $\lambda_0 = longitude\ of\ real\ location$, and $\lambda = longitude\ of\ estimated\ location$

RSSI ratio data and the optimised kernel are used in order to evaluate and compare the performance of the three algorithms. RSSI ratio features (120 × 6) and their corresponding location coordinates were used to train the algorithms; the data from the remaining 30 locations were used as test data. The overall performance of the three models is captured by CDFs of the localisation error as shown in Figure 6, Figure 7 and Figure 8. Each model provides a localisation accuracy with a median error of less than 400 m. Epsilon-SVR has the lowest median error of 303 m compared to 309 and 317 m for nu-SVR and GPR, respectively. SVR outperformed the GPR model in terms of the overall accuracy.

5.6 | Analysis of Antwerp dataset

To further demonstrate the feasibility and consistency of the developed method for LoRaWAN localisation, we explore a public dataset of LoRaWAN messages obtained in the city centre of Antwerp. It holds 123,529 messages which were collected over a 3-week period. City of Things hardware and a Firefly ×1 GPS receiver was mounted on 20 cars of Antwerp's postal service, which drove around in the city centre while continuously acquiring the current latitude and longitude of the car as well as the Horizontal Dilution of Precision (HDOP) of the GPS signal. The acquired location information is sent in a LoRaWAN message via the IM880 B-L radio module. With HDOP, messages with poor GPS signal quality could be removed. Information stored in the dataset include 68

TABLE 16 Model performance of support vector regression (SVR) on Jazan data (absolute)

	Model	Median (m)	Mean (m)	Time
SF 9 Absolute RSSI	RBF + Matern	442	493	3.18 min
	RQ + Matern	421	483	3.5 min
SF 10 Absolute RSSI	RBF + Matern	506	559	3.23 min
	RQ + Matern	428	513	3.43 min
SF 11 Absolute RSSI	RBF + Matern	410	575	3.27 min
	RQ + Matern	426	537	3.4 min
SF 12 Absolute RSSI	RBF + Matern	336	441	3.23 min
	RQ + Matern	344	392	3.44 in

LoRaWAN base stations, the receiving time, SF, HDOP, latitude and longitude. For an unbiased comparison, we preprocessed the Antwerp city LoRaWAN data before applying our method. The preprocessing and result of the work done with Antwerp dataset is captured as follows:

1. Four gateways (GW_9, GW_14, GW_27, and GW_69) were selected out of the 68 gateways in the Antwerp city dataset. This is done in order to have the same number of gateways as in our Jazan city dataset.
2. The four selected gateways should have the same SF as the Jazan city dataset, which are SF9, SF10, SF11, and SF12, in order to guarantee the Antwerp city dataset has the same features as Jazan dataset as shown in Table 12.
3. The same data transformation technique (RSSI ratio) applied on Jazan city data is used on Antwerp city data.
4. The same kernel functions are used as with the Jazan city dataset, which are (a) Matern with RBF and (b) Matern with RQ.
5. The tables below show the results for both RSSI ratios and Absolute RSSI for Antwerp city dataset.

As can be seen in both Tables 13 and 14, the RSSI ratio data transformation technique helped improve the node localisation performance. This proves the consistency in our technique.

5.7 | Runtime analysis

The time complexity of the methods used in this work have been analysed. Tables 13–16 show the different models and the run time for each algorithm. There is little or negligible difference in the run time of the algorithms for the Jazan city data irrespective of the SF. However, when the same analysis is conducted using the Antwerp city dataset, there is significant change in the run time of the algorithm. This indicates that the run time of our algorithms is data dependent.

6 | CONCLUSION

An investigation into the use of kernelised learning methods in tandem with RSSI ratios for node localisation in LPWAN settings has been detailed. Specifically, epsilon and nu-Support Vector Regression and GP Regression have been used to develop localisation models, capturing the dynamic relationship between RSSI ratios and node location estimation.

A combination of four gateway RSSI ratio pairs formed inputs to the models during training and the optimum kernel function defined through an evaluation of the effect of kernels on the performance of the estimate of location. Each model provides a localisation accuracy with median error of less than 400 m, using combined kernels. Epsilon-SVR yields the lowest median error of 303 m compared to 309 and 317 m for nu-SVR and GPR, respectively. SVR outperformed the GPR model in terms of overall accuracy. The result demonstrates that the combination of different kernel functions can enhance localisation accuracy. The analysis of SF and transformed RSSI (RSSI ratio) shows the significance of SF and RSSI ratio on the performance of the node localisation models. The evaluation with High SFs (SF11 and SF12) outperform low SFs (SF9 and SF10). The RSSI ratio shows improved localisation accuracy over absolute RSSI. The evaluation of the proposed method on the Antwerp city data provides evidence on the consistency of our findings.

ACKNOWLEDGMENTS

Ibrahim Aqeel expresses his gratitude to the Emirate of Jazan Province, the Jazan Municipality and the Ministry of Environment Water and Agriculture in Jazan City for facilitating the data gathering phases through access to the environment and support. This research study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

All authors have no conflicts of interest.

DATA AVAILABILITY STATEMENT

Participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

ORCID

Ephraim Iorkyase https://orcid.org/0000-0003-4354-8247

REFERENCES

1. Lueth, K.L.: IoT 2019 in Review: The 10 Most Relevant IoT Developments of the Year (2020). January 2020. [Online]. https://iot-analytics.com/iot-2019-in-review/
2. Mekkia, K., et al.: A comparative study of LPWAN technologies for large-scale IoT. ICT Express. 5(1), 1–7 (2019). https://doi.org/10.1016/j.icte.2017.12.005
3. Andrade, R.O., Yoo, S.G.: A comprehensive study of the use of LoRa in the development of smart cities. Appled Sciences. 9(4753), 1–39 (2019). https://doi.org/10.3390/app9224753
4. Migabo, E., et al.: A comparative survey study on LPWA networks: LoRa and NB–IoT. ICT Express. 3(1), 14–21 (2017)
5. Tanaka, M., et al.: A study of bus location system using LoRa: bus location system for community bus notty. In: The 6th Global Conference on Consumer Electronics (GCCE), Nagoya (2017)
6. Wan, J., et al.: Wearable IoT enabled real-time health monitoring system. EURASIP J. Wirel. Commun. Netw. 298(1) (2018). https://doi.org/10.1186/s13638-018-1308-x
7. Kaplan, E.D., Hegarty, C.J.: Understanding GPS: Principles and Applications, 2nd ed. Artech House, Norwood (2005)
8. Bernhard, H.-W., Herbert, L., Elmar, W.: GNSS – Global Navigation Satellite Systems,GPS, GLONASS, Galileo, and More, 1 ed. Springer-Verlag Wien (2008)
9. Raskovic, D., Giessel, D.: Battery-aware embedded GPS receiver node. In: 4th Annu. Int. Conf. Mobile Ubiquitous Syst Netw. Ser- vices (MobiQuitous), Philadelphia (2007)
10. Fargas, B., Petersen, M.: GPS-free geolocation using LoRa in low-power WANs. In: 2017 Global Internet of Things Summit (GIoTS). Geneva (2017)
11. Workgroup, T.M.: LoRaWAN™ what Is it? A Technical Overview of LoRa® and LoRaWAN™ (2015). [Online]. https://lora-alliance.org/sites/default/files/2018-04/what-is-lorawan.pdf. Accessed February 2020
12. Haxhibeqiri, J., et al.: A survey of LoRaWAN for IoT: from Technology to application. sensors Review. 18(3995), 1–38 (2018). https://doi.org/10.3390/s18113995
13. Aernouts, M., et al.: Sigfox and LoRaWAN datasets for fingerprint localization in large urban and rural areas. Data. 313(2) (2018). https://doi.org/10.3390/data3020013
14. Wongeun, C., et al.: Low-power LoRa signal-based outdoor positioning using fingerprint algorithm. SPRS International Journal of Geo-Information. 7(11), 1–15 (2018)
15. Iorkyase, E., et al.: Low complexity wireless sensor system for partial discharge localisation. IET Wirel. Sens. Syst. 9(3), 158–165 (2019). https://doi.org/10.1049/iet-wss.2018.5075
16. Lemic, F., et al.: Regression-based estimation of individual errors in fingerprinting localisation. IEEE Access. 7, 33652–33664 (2019). https://doi.org/10.1109/access.2019.2903880
17. Zhe, H., et al.: Enhanced Gaussian process-based localization using a low power wide area network. 164 IEEE COMMUNICATIONS. 23(1), 164–167 (2019). https://doi.org/10.1109/lcomm.2018.2878704
18. Schafer, J.L., Graham, J.W.: Missing data: our view of the state of the art. Psychol. Methods. 7(2), 147–177 (2002). https://doi.org/10.1037/1082-989x.7.2.147
19. Kim, W., et al.: A Comparison of the Effects of Data Imputation Methods on Model Performance. Korea (South) (2019)
20. Ke, S., et al.: Support vector regression based indoor location in IEEE 802.11 environments. Mobile Inf. Syst., 1–4 (2015). https://doi.org/10.1155/2015/295652
21. Clarke, S.M., Griebsch, J.H., Simpson, T.W.: Analysis of support vector regression for approximation of complex engineering analyses. J. Mech. Des. 127(6), 1077–1087 (2005). https://doi.org/10.1115/1.1897403
22. Livinsa, Z.M., Jayashri, S.: Localization with beacon based support vector machine in Wireless Sensor Networks. In: 2015 International Conference on Robotics, Automation, Control and Embedded Systems (RACE), pp. 18–20. Chennai (2015)
23. Zhang, L., et al.: A comprehensive study of bluetooth fingerprinting-based algorithms for localization. In: 2013 27th International Conference on Advanced Information Networking and Applications Workshops. Barcelona (2013)
24. Rahul, K., Mariette, A.: Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers, 1st ed, pp. 39–80. Apress (2015)
25. Ashourloo, D., et al.: An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 9(9), 4344–4351 (2016). https://doi.org/10.1109/jstars.2016.2575360
26. Omidi, A., Barakati, S., Tavakoli, S.: Application of nusupport vector regression in short-term load forecasting. In: 2015 20th Conference on Electrical Power Distribution Networks Conference (EPDC), Zahedan (2015)
27. Iorkyase, E., et al.: Radio location of partial discharge sources: a support vector regression approach. IET Sci. Meas. Technol. 12(2), 230–236 (2018). https://doi.org/10.1049/iet-smt.2017.0175
28. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
29. Shi, K., et al.: Support vector regression based indoor location in IEEE 802.11 environments. Mobile Inf. Syst., 1–14 (2015). https://doi.org/10.1155/2015/295652
30. Kai, C., et al.: Mixed kernel function support vector regression for global sensitivity analysis, vol. 96, pp. 201–214. Elsevier Ltd (2017)
31. Rasmussen, C., Williams, C.: Gaussian Processes for Machine Learning, p. 13. MIT Press (2006)
32. Basak, D., Pal, S., Patranabis, D.: Support vector regression. Neural Information Processing-Letters and Reviews. 11(10), 203–224 (2007)
33. Tan, Y., Wang, J.: A support vector machine with a hybrid kernel and minimal Vapnik-chervonenkis dimension. IEEE Trans. Knowl. Data Eng. 16(4), 385–395 (2004). https://doi.org/10.1109/tkde.2004.1269664

How to cite this article: Aqeel, I., et al.: LoRaWAN-implemented node localisation based on received signal strength indicator. IET Wirel. Sens. Syst. 1–16 (2022). https://doi.org/10.1049/wss2.12039