



Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

Nazmat Toyin Surajudeen-Bakinde¹, Frederick Ojiemhende Ehiagwina^{1,2,*},
Akindele Segun Afolabi¹ & Ayinde Mohammed Usman^{1,3}

¹Department of Electrical and Electronics Engineering, University of Ilorin, PMB 1515
Tanke, Ilorin, Kwara State, Nigeria

²Department of Electrical/Electronic Engineering, The Federal Polytechnic Offa,
Osogbo-Ajesse-ipo Road, PMB 420, Offa, Kwara State, Nigeria

³Department of Electrical & Electronic Engineering, Stellenbosch University, Private
Bag X1, Matieland, 7602, Stellenbosch Central, South Africa

*E-mail: Frederick.ehiagwina@fedpoffaonline.edu.ng

Highlights:

- The Holt-Winters (HW) method and the GA-Holt-Winters technique were showed good forecast behavior for GSM 900 RL, GSM 1800 RL, and GSM 1800 FL.
- There was a decrease of about 16% in the mean square error (MSE) of the prediction of the GA-Holt-Winters technique compared to the Holt-Winters method for GSM 900 RL in both locations.
- The MSE prediction values for GSM 1800 RL in locations 1 and 2 decreased by 22% and 45% for the GA-Holt-Winters technique compared to the Holt-Winters method.
- There was a decrease of about 28% and 8% in the MSE of the prediction of the GA-Holt-Winters technique compared to the Holt-Winters method for GSM 900 RL in locations 1 and 2 respectively.

Abstract. In this research, the suitability of a genetic algorithm (GA) modified Holt-Winters (HW) exponential model for the prediction of spectrum occupancy data was investigated. Firstly, a description of spectrum measurement that was done during a two-week duration at locations (8.511 °N, 4.594 °E) and (8.487 °N, 4.573 °E) of the 900 MHz and 1800 MHz bands is given. In computing the spectrum duty cycle, different decision thresholds per band link were employed due to differing noise levels. A frequency point with a power spectral density less than the decision threshold was considered unoccupied and was assigned a value of 0, while a frequency point with a power spectral density larger than the decision threshold was considered occupied and was assigned a value of 1. Secondly, the spectrum duty cycle was used in the evaluation of the forecast behavior of the forecasting methods. The HW approach uses exponential smoothing to encode the spectrum data and uses them to forecast typical values in present and future states. The mean square error (MSE) of prediction was minimized using a GA by iteratively adjusting the HW discount factors to improve the forecast accuracy. A decrease in MSE of between 8.33 to 44.6% was observed.

Keywords: *cognitive radio network; genetic algorithm; Holts-Winters exponential smoothing; spectrum measurement; spectrum occupancy; spectrum prediction.*

1 Introduction

The cognitive radio networks (CRNs) paradigm has been proposed as a viable solution to spectrum allocation inefficiencies. Under this paradigm, a cognitive user (CU) is able to utilize a vacant channel opportunistically, without preventing licensed users from gaining access to the channel when desired. This demands a comprehensive understanding of the spectrum utilization profile and the dynamic behavior of licensed users in a realistic scenario via spectrum measurement [1-3]. However, continuous spectrum measurement is expensive and time-consuming. This necessitates spectrum prediction. Spectrum prediction uses historically observed data from spectrum sensing to forecast future channel states of licensed channels [4,5]. Some techniques that have been used for spectrum prediction include time series techniques, artificial neural network (ANN), Markov model, Bayesian inference, k-nearest neighbors (KNN), etc. [6-10]. ANNs have a number of advantages, including the ability to detect complex nonlinear relationships between dependent and independent variables without requiring formal statistical training, the ability to detect all possible interactions between the predictor variables, and the availability of multiple training algorithms. Some of the drawbacks of ANNs are their black box character, higher computing cost, proclivity to overfitting, and the empirical nature of model creation.

The Markov model's main advantages are its simplicity and out-of-sample forecasting accuracy. Unfortunately, the Markov model is ineffective in explaining occurrences and in most situations cannot be considered an accurate representation of the underlying reality. Using the Bayesian inference approach has certain drawbacks as well. It gives false results when caution is not exercised. It has the potential to generate posterior distributions that are highly impacted by priors. Although the KNN technique is simple to construct, its efficiency or speed decreases rapidly as the data set expands. This is because when the input variables are limited, it works well, but as the number of variables increases, it fails to anticipate the output of additional data points. Owing to its tendency to choose neighbors based on distance criteria, the KNN method is particularly sensitive to outliers. Selected literature on spectrum prediction and associated issues is reviewed next. Predicting the behavior of CUs in wireless networks was the goal of the authors in [6], who gave a comparison of two time series models – autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA) – that can forecast primary user behavior as well as spectral opportunities for wireless networks in the Wi-Fi frequency band. The results revealed that the SARIMA model had the best overall performance. This is because the SARIMA algorithm has higher precision in terms of availability and occupancy times, it performs better, and is more convenient for cognitive radio networks. As a result, the spectrum efficiency is

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

enhanced, and interference and collisions between licensed users and CUs are decreased. Furthermore, a two-stage data-driven spectrum estimation (SE) technique for a CRN was proposed in [11], comprising a null-hypothesis approach with strong Chi-Square Goodness of Fit for validation of the intended signal. The key contribution of this study was the development of an optimized scalable ARIMA model for frugal SE with minimal response time in terms of data length and lag order.

In most cases, CRNs use random spectrum sensing of the channels. As such, the possibility of a channel being sensed as active is quite high, resulting in considerable data loss and a reduction in the network's effective throughput. The difficult challenges of choosing a channel for spectrum sensing following spectrum prediction were exploited in [12] to boost the CU's throughput. In addition, a new technique for increasing CU throughput by utilizing underlay communication in the spectrum sensing and prediction phases was described. The results demonstrated a considerable increase in throughput when this method was used.

Kumar, *et al.* [13] noted that due to the wait state problem in interweave mode, picking the licensed spectrum for sensing at random in a high traffic intensity network reduces the throughput of cognitive users. However, dual spectrum access overcomes the wait state problem, resulting in increased spectrum usage throughput. Furthermore, sensing decisions made without cooperation increase sensing inaccuracy and have an impact on sensing performance, resulting in a drop in throughput. Therefore, the authors offered a new strategy for cognitive users called hybrid spectrum access with spectrum prediction and cooperation, in which spectrum prediction, cooperation, and hybrid spectrum access are combined to increase the system's sensing performance and throughput. These mitigate unintended licensed-user and CU collisions.

The influence of the collision factor was addressed in [14]. It was discovered that without prediction, the throughput drops as the collision factor grows. To lower the prediction error, the authors ensured that the network was trained using an NN based on the multi-layer perceptron model. Following the spectrum prediction, the acquired data demonstrated an increase in CU throughput. Extending the concept of spectrum prediction further, the authors in [15] studied the overlap between spectrum monitoring and spectrum prediction. On the basis of received packet characteristics such as the receiver error count, the CU detects the appearance of a licensed user during data transmission through spectrum monitoring. To improve the performance of cognitive radio networks by recognizing the formation of PU promptly and properly, spectrum prediction and monitoring techniques are used concurrently for spectrum mobility. Spectrum mobility strategies are divided into two categories: reactive and proactive. The

findings of both the spectrum monitoring and spectrum prediction algorithms were combined using the AND and OR fusion rules.

A proposed backpropagation training model was used in neural system-based spectrum prediction in [16]. A genetic algorithm (GA) and a hybrid combination of shuffled frog-leaping algorithm (SFLA) was developed to improve the structure of the neural system and reduce the forceful weight auxiliary pattern. The GA was used in this case to avoid capturing over-fitted solutions. Randomness was created using the selection, crossover, and mutation processes, which spreads out the populace to unify to the set that holds the global optimum solution. Simulation results demonstrated that by enhancing the system, the GA-SFLA-based hybrid algorithm improved the outcomes of finding the optimal weights; also, the suggested conspire results showed great forecast accuracy.

The authors of [17] investigated multi-step-ahead spectrum prediction for CRNs with many future states using the support vector machine (SVM) technique. The scenario was based on slots. The goal was to see if multi-step-ahead spectrum prediction outperformed short-term prediction in terms of reduced channel switching and enhanced network throughput. Sensing was done with a traditional energy detector. In addition, the authors proposed new closed-form formulas for detection probability in additive white Gaussian noise (AWGN) and Rayleigh fading channels. The SVM method exhibited low prediction error rates, and multi-step-ahead idle-channel scheduling by the SU resulted in a 51% decrease in channel switching. For multi-step-ahead prediction with three future states, a 4% improvement in throughput was reported. Based on the literature review above, it can be summarized that spectrum predictions have been based on any of the following broad techniques: time series-based approaches, static machine learning (ML)-based methods, and hybrids of the previous two.

One of the most important advantages of ML is its capacity to analyze enormous amounts of data and spot patterns without errors associated with human analysis. It has the capacity to improve with time as well. Due to the ever-increasing volumes of data handled, the application of ML technology often enhances efficiency and accuracy. This provides additional experience to the algorithm or software, which may then be utilized to make better judgments or predictions. Without the requirement for human interaction, ML enables immediate adaptability. Because ML is automated, it can save time and money by allowing developers and analysts to focus on higher-level activities, which a machine cannot do [18-20]. An error in an ML interface can wreak disaster since all following occurrences may be faulty, biased, or just unpleasant. Owing to the fact that ML develops over time as a consequence of exposure to large data sets, there may be a point when the algorithm or interface is not quite ready for your needs. In other words, ML takes time. Handling massive amounts of data and running

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

computer models consumes a lot of computational power, which may be expensive [18,19]. Conversely, models based on time series are reliant on their own past experiences. As a result, they are unable to show anything that has not occurred previously. Due to the inherent unpredictability of these systems, forecasting time series can be a difficult undertaking. It is also uncertain how much a non-linear deterministic process keeps its qualities when it is distorted by noise. Even if the system's equations are deterministic, noise can alter it in a variety of ways. Hence, the need to optimize time series models such as ARIMA and HW [21,22]. Therefore, this research investigated the impact of GA optimization of the HW technique in the prediction of spectrum occupancy data sets. To the best of our knowledge, this is the first instance in which the Holt-Winters times series technique has been optimized using a genetic algorithm for the prediction of GSM spectrum data. This supports the creation of the cognitive radio network's spectrum choice framework, a crucial component of the dynamic spectrum assessment paradigm.

The rest of this paper is organized as follows. In Section 2, a general overview of the proposed technique is presented. In Section 3, the spectrum data collection and processing procedures are described. Additionally, the GA-Holt-Winters prediction model implementation procedure is framed, along with a description of the software and machine configurations, and GA parameterization used in this study. The results obtained are reported and discussed in Section 4, where, in terms of observed mean square error (MSE), the GA-Holt-Winters approach was found to have better forecast behavior with the data set than the Holt-Winters method. The conclusions reached are presented in Section 5.

2 Proposed Method

HW exponential smoothing can effectively predict periodic series with only a few training samples. The need for this technique arose owing to the inherent periodicity and noisiness associated with the 900 MHz and 1800 MHz communication data set, which serves as the basis of this spectrum prediction. Measured spectrum occupancy data are processed to extract spectral duty cycles as indication of spectral vacancies. An enhanced time series forecasting model of spectrum duty cycles based on this strategy is hereby presented. The optimal smoothing settings for the HW exponential smoothing are chosen using the GA optimization approach. As observed in [23], hybrid models are more promising than stand-alone models due to their superior capacity to describe nonlinear and unpredictable elements. In addition, they have a much lower training complexity than static ML-based techniques.

Winters [24] extended the Holt [25] technique to express periodicity in time series data. The HW technique is also known as the triple exponential smoother

(TES). It has two variations, depending on the type of the seasonal constituent; these are the additive and the multiplicative method. The additive method is the smoother choice when the seasonal fluctuations are fairly the same over the entire time series. On the other hand, the multiplicative Holt-Winters technique is favorable when the seasonal changes depend on the level of the series. The seasonal component is expressed in absolute terms at the scale of the observed series in the additive method, and the series is seasonally adjusted in the level equation by subtracting the seasonal component. Hence, for each cycle, the seasonal components add up to approximately zero. The seasonal component is expressed in relative terms (percentages) with the multiplicative method. By dividing with the seasonal component, the series is seasonally adjusted. Therefore, for each cycle, the seasonal components add up to $\approx m$, the frequency of the seasonality [26]. The HW seasonal technique is made up of four equations. Three for smoothing and one for forecasting. These are described in Section 3. Many researchers have applied it in forecasting in several areas, including air traffic [27,28], infant mortality [29], electricity consumption [30], aircraft failure rate [31], and so on.

GAs are optimization algorithms based on the principle of Darwinian natural selection. They mimic the evolution process in searching for a solution. The solution is encoded as a string of binary digits or real numbers. Genetic operators such as selection, mutation and crossover are used to produce improved solutions with each iteration or generation. The algorithm ends when a preset criterion is met. Therefore, with each iteration, members of the old string set with the highest fitness function are utilized in generating a different group of strings. Since its inception, the GA has been effectively employed in scientific and engineering applications to find near-optimal solutions to a range of problems [32-36]. The GA or its variant extension of the Holt-Winters method have been used by different authors in forecasting in diverse fields. Peng, *et al.* [37] used a niching GA and HW method to predict mining subsidence crucial in engineering construction over underground mines. The relative prediction errors were less than 2% and the mean error was -0.18%. The authors reported that this was a better performance than the SVM-based predictive technique. Amzi [38] used a GA to estimate the HW parameters when forecasting tourist arrival data in Langkawi, Malaysia. The results of the GA outperformed the conventional optimization approaches in terms of mean average percentage error.

The HW method permits one to correctly predict seasonal series with comparatively small training samples. With the use of this technique, a hybrid predictive model to forecast spectrum occupancy is proposed here. The GA is used to optimize the smoothing parameters for the HW technique. To the best of our knowledge, no study has previously reported the application of GA-HW to wireless spectrum occupancy data sets covering 900/1800 MHz.

3 Methodology

3.1 The Genetic Algorithm

The GA starts with a string population and consequently produces other generations of populations of strings in accordance with specific nature-inspired operations of reproduction, crossover, and mutation. The reproduction operation allows retaining the parent chromosome and the transfer of same to the offspring, producing a set of improved solutions. In this case, there is no change to the chromosome. That is, the output of this process is the same as the input. This usually leads to a local optimum [37,39]. In the Crossover operation, two chromosomes are concatenated to generate two new chromosomes through the process of gene switching for a simple one-point crossover operation on a binarized population. For example, if two strings in the current population P are I and I' , then [37,39],

$$I = \{x_1, \dots, x_j, \dots, x_n\} \quad (1)$$

$$I' = \{x'_1, \dots, x'_j, \dots, x'_n\} \quad (2)$$

The crossover point is fixed through the random generation of an integer j from 1 to n . The resulting cross indexes are Eqs. (3) and (4) [35]:

$$I = \{x_1, \dots, x_{j-1}, x'_j, \dots, x'_n\} \quad (3)$$

$$I' = \{x'_1, \dots, x'_{j-1}, x_j, \dots, x_n\} \quad (4)$$

The Mutation operation, in contrast to Reproduction and Cross-over, involves reversing the value of one gene of a chromosome in a random manner, resulting in a different but mutated output. Let s_j be randomly selected, which mutate into x'_j , if $x_j = 1$ then $x'_j = 0$ and if $x_j = 0$ then $x'_j = 1$. This GA operation creates a completely new species; in so doing, it helps to get out of local optimums by the creation of an arbitrary locus [37,39,40].

The flowchart for the implementation of the GA is shown in Figure 1. The GA's implementation is summarized in Figure 1. First, the algorithm generates a random initial population. After that, the algorithm generates a series of new populations. The program creates the future population using the individuals in the current generation at each step. The algorithm follows these stages to generate a new population. The algorithm calculates the fitness value of each member of the current population, providing raw fitness scores. The raw fitness ratings are scaled to create a more useful range of values. Members, referred to as 'parents',

are chosen based on their scaled fitness score. Some of the contemporary population's fitness-challenged individuals are designated as 'elite'.

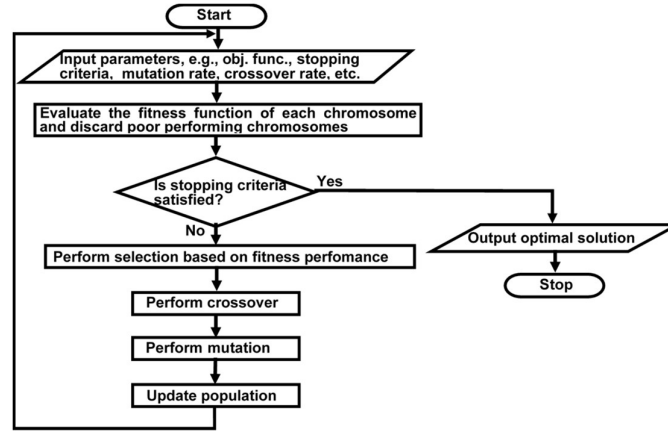


Figure 1 Flowchart of the genetic algorithm.

The following generation inherits these outstanding individuals. Children are created by mixing the vector entries of two parents (crossover) or by applying random modifications to a single parent (mutation). To generate the next generation, the present population is replaced by children. When the stopping criterium is satisfied, i.e., when the value of the fitness function for the best point in the current population is less than or equal to the fitness limit or the maximum number of evolutions (1,000) is reached, the algorithm comes to a halt. If not, the algorithm continues evolution by crossover and mutation to create a new population [41].

3.2 Holt-Winters Method

It is assumed that the seasonal time series model is:

$$y_t = L_T + \zeta_t + \varepsilon_t \quad (5)$$

where L_t = the linear trend component, which can be represented by:

$$\beta_0 + \beta_1 t : \beta_0, \beta_1 \in \mathbf{Z} \quad (6)$$

ζ_t = the seasonal adjustment with $\zeta_t = \zeta_{t+m} = \zeta_{t+2m} = \dots$ for $t = 1, \dots, m-1$

where m = the period length of each cycle; and the error ε_t is taken to be uncorrelated with zero mean and constant variance σ^2 . As mentioned in Section 1, the seasonal components add up to zero during one cycle, that is,

Genetic Algorithm-Holt-Winters Based Minute Spectrum
Occupancy Prediction: An Investigation

$$\sum_{t=1}^m \varsigma_t \quad (7)$$

The HW method calculates dynamic estimates for three components, namely, level (L_t), trend (T_t), and seasonality (ς_t), either using an additive model or a multiplicative model [24,38,42]. The procedure for updating the parameter estimates once the current observation y_t is obtained, is as shown in Eqs. (8)-(15).

The additive model is as shown in Eqs. (8)-(10) [28]:

$$L_t = \alpha(y_t - \varsigma_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (8)$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1} \quad (9)$$

$$\varsigma_t = \delta(y_t - L_{t-1} - T_{t-1}) + (1 - \delta)\varsigma_{t-m} \quad (10)$$

The forecast value is given in Eq. (11):

$$\hat{y}_t = L_{t-1} + T_{t-1} + \varsigma_{t-m} \quad (11)$$

The multiplicative model is as shown in Eqs. (12)-(14):

$$L_t = \alpha(y_t / \varsigma_{t-m}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (12)$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1} \quad (13)$$

$$\varsigma_t = \delta(y_t / (L_{t-1} + T_{t-1})) + (1 - \delta)\varsigma_{t-m} \quad (14)$$

The forecast value is given as:

$$\hat{y}_t = (L_{t-1} + T_{t-1})\varsigma_{t-m} \quad (15)$$

where α, γ and δ are discount factors ranging from [0,1]. An initial value of 0.2 is selected for all three discount factors.

In this study, various values for the seasonal period were investigated. The period that resulted in the least mean square deviation (MSD) was selected for further optimization using the GA. The expression for computing the MSD is given in Eq. (16).

$$MSD = \sum_{t=1}^n \frac{|y_t - \hat{y}_t|^2}{n} \quad (16)$$

where y_t = the true value, $\hat{y}_t$ = the corresponding fitted/predicted one with n observed samples.

3.3 Genetic-Holt-Winters Algorithm

The genetic algorithm is used to determine the optimal values of α , γ and δ such that the forecast errors are minimized. If α , γ and δ are too small, over-smoothing takes place. If they are closer to one, no smoothing takes place. The procedure followed are itemized as follows [37]:

1. Initialization of the GA parameters, e.g., population size and evolution number.
2. Initialization and interval selection of the HW smoothing parameters α , γ and δ .
3. Evaluation of the objective function, e.g. MSE between the predicted data and the spectrum occupancy, is done as shown in Eq. (17).
4. Evaluation of the fitness of each individual in the niche, followed by selection of the best one in the niche for the next generation.
5. Optimization of the parameters by minimizing the MSE using the GA operations described in Section 3.1.

$$f = \min \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (17)$$

Eqs. (18)-(20) are used to determine the initial values of the levels, trend, and seasonality index [42].

$$L_0 = n^{-1} \sum_{t=1}^n y_t \quad (18)$$

$$T_0 = N \quad (19)$$

$$S_0 = \frac{\bar{y}_m}{L_0} \quad (20)$$

where N is a selected integer, and $\bar{y}_m$ is the average of the selected period samples.

The implementation framework is shown in Figure 2. The prediction accuracy of the HW model is enhanced by using the GA to minimize the MSE via iterative adjustment of parameters α , γ and δ .

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

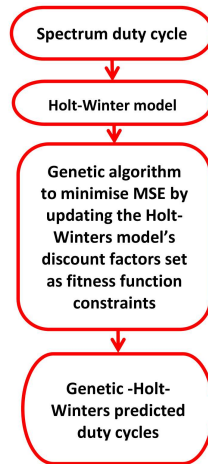


Figure 2 GA-Holt-Winters model implementation framework.

3.4 Spectrum Occupancy Dataset

A selected portion of the spectrum band (885 to 1880 MHz) was measured for spectrum occupancy in a few chosen places in Ilorin, a city in the North Central zone of Nigeria. The data collection points were categorized by the degree of urbanization with the details of the locations shown in Table 1.

Table 1 Details of selected measurement locations.

Locations	Category	GPS Coordinates
Post Office-GRA area (LOC1)	Urban	8.511 °N, 4.594 °E
Sango-Basin area (LOC2)	Sub-urban	8.487 °N, 4.573 °E

3.4.1 Spectrum Band

The cellular bands are highlighted. The cellular wireless communication services that were taken into account in this research are listed in Table 2, along with the associated allocated frequency bands.

Table 2 Wireless services in the chosen spectrum band (885 to 1880 MHz).

Wireless Services	Allocated Band (MHz)	Bandwidth (MHz)
GSM 900 forward link (FL)	880 – 915	35
GSM 900 reverse link (RL)	925 – 960	35
GSM 1800 forward link (FL)	1710 – 1785	75
GSM 1800 reverse link (RL)	1805 – 1880	75

3.4.2 Measurement Equipment Setup

A sensitive spectrum analyzer is required since the measuring equipment must be able to detect both weak and strong signals over the specified frequency range. An energy detector, a field strength analyzer (BK PRECISION 2640) with a frequency range of 100 kHz to 2.0 GHz, an omnidirectional antenna, a GPS-based mobile phone, and a high-capacity storage device integrated into a laptop made up the experimental setup. The coordinates of each location's spectrum measurement could be found thanks to the GPS. Table 3, taken from the BK PRECISION 2640 user manual, summarizes the main features of the field strength analyzer. The setup of the spectrum measurement tools and accoutrements is shown in Figure 3.

Table 3 Selected specifications of the BK PRECISION 2640 field strength analyzer.

Parameter	Value
Frequency range	100 kHz to 2.0 GHz
Resolution	3.125 Hz
Resolution bandwidth	Variable
Input impedance	50 Ω
Sweep time	Min. 500 ms
Measurement amplitude range	-45 dBm to -110 dBm
Average noise level	-110 dBm max
Input sensitivity	@35 MHz -2,000 MHz: 150 mVrms
Frequency selection mode	Centre, Start/Stop, Span
Reference level accuracy	± 3.0 dB @ < 600 kHz; ± 2.0 dB @ ≥ 600 kHz
Log scale	0.2 dB/div minimum in 0.25 dB span

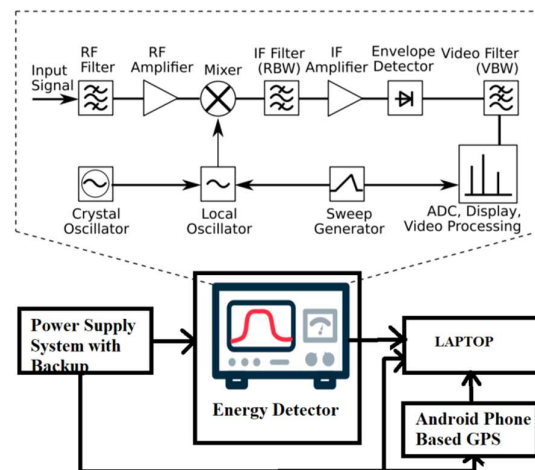


Figure 3 Spectrum evaluation equipment and accessories set up.

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

Due to its simplicity, the measurement technique is based on energy detection. Aside from the measuring power, this was preferred because not much information was previously available.

3.4.3 Data Collection and Processing

The power spectral densities (PSD) were measured by the spectrum analyzer. The method used to process the measured data was then indicated. If the PSD for a certain channel was greater than a predetermined threshold, the channel was considered occupied; otherwise, the channel was considered unoccupied. The collected data of the field measurement are represented in matrix $\mathbf{Y}$ (Eq. (21)), where each element is the received signal power ($\mathbf{P}$), dependent on both time t_i and frequency f_j . The power density, temporal variations and duty cycle are usually of interest. The duty cycle may be defined empirically as a ratio of the time a frequency fragment is declared occupied. In other words, it is the probability that a communication channel is busy [43].

$$\mathbf{Y} = [\mathbf{P}(t_i, f_j)] \quad (21)$$

Every element $\mathbf{P}(t_i, f_j)$ of the matrix represents a sample of the received power observed by the energy detector at the instant of time $t_i, \forall i = 1, 2, \dots, N_t$, with N_t being the maximum number of time points in the measurement, and each frequency slot $f_j, \forall j = 1, 2, \dots, N_f$, with N_f being the maximum number of bins under consideration.

Spectrum occupancy DC is computed using a specific set of measurements from the spectrum analyser over a range of frequencies and a time interval. When the binary hypothesis is used, the indicator function $\theta_D(t_i, f_j)$ with a value of 1 is assigned for signal power above the decision threshold λ_k and a value of 0 is assigned for signal power below the decision threshold. This is shown in Eq. (22) [43]:

$$\theta_D(t_i, f_j) = \begin{cases} 0, & \text{if } \mathbf{P}(t_i, f_j) < \lambda_k \\ 1, & \text{if } \mathbf{P}(t_i, f_j) \geq \lambda_k \end{cases} \quad (22)$$

For frequency slot f_j , the measured DC $\Delta(t_i, f_j)$ is given as,

$$\Delta(t_i, f_j) = \frac{1}{N_t} \sum_{i=1}^{N_t} \theta_D(t_i, f_j) \quad (23)$$

The average duty cycle is evaluated from the mean of individual $\Delta(t_i, f_j)$ for all frequency slots N_f and is expressed as,

$$\Delta_{i,j}(avg) = \frac{1}{N_f} \sum_{i=1}^{N_f} \Delta(t_i, f_j) \quad (24)$$

The measurement resolution is chosen as 250 kHz, slightly wider than the 200 kHz bandwidth of the service bands. The number of frequency readings per time slot n_f is given by Eq. (25),

$$n_f = \frac{f_{stop} - f_{start}}{R_m} \quad (25)$$

The total number of time slots for each location is defined by Eq. (26),

$$n_{ts} = \frac{T_m}{\text{sweep time}} \quad (26)$$

where T_m is a function of the time spent in measurement. For each service band, two weeks (20,160 minutes) is spent on either the forward link or the reverse link. Thus, there are $[n_f \times n_{ts}]$ data points to be captured during measurement per location.

3.4.4 Frequency Bin Size and Resolution Bandwidth

The measurement technique takes into account the link between the frequency resolution and the transmitted signal's bandwidth. The bin size is necessary to ascertain this relationship. The service bandwidth is taken into account while choosing the bin size. If the bin size stays somewhat smaller than the signal bandwidth, the spectrum occupancy prediction will be more accurate. In an area with low PU activity, a bin size greater than the signal bandwidth causes an overestimation of channel usage [44].

Additionally, by reducing the resolution bandwidth (RBW), the system is better able to detect low-power signals at the expense of longer measurement times. This is due to the fact that a smaller RBW improves the system's ability to resolve the signal frequency, lowering the noise floor. The RBW was 12.5 kHz. This is a suitable compromise between measurement duration, as represented by the average sweep time of the spectrum analyzer and the detection capabilities, as shown by the observed duty cycle [44].

3.4.5 Setting of Decision Threshold

The decision threshold chosen has an impact on the spectrum duty cycle, as shown in [45]. Overestimation is caused by noise samples that are above the decision threshold while the decision threshold is noticeably low. However, because of the incorrect detection of faded primary transmissions, a very high decision threshold causes underestimation of the real spectrum occupancy level. Variable m_i -dB criteria are used in setting the decision threshold. Here, the threshold was placed m_i -dB above the average noise level determined by a matched load placed across the SA, depending on the band. Variable m_i -dB criteria were used since the variance of noise $\sigma^2 X(f)$ and the noise levels may

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

change from one band to another depending on the measurement setup, and as such a fixed m-dB threshold across the whole frequency range being measured is unsuitable. The decision threshold λ_k in dB is expressed in Eq. (27):

$$\lambda_k = \mu X(f) + m_i \quad (27)$$

where $\mu X(f)$ is the average noise level.

3.5 Experimental Setup

A presentation of the design and settings of the machine used in the experiment is given in this section. The fitness function adopted in this study, with justification, is explained. Furthermore, the GA-Holt-Winters model parameters are initialized and the running of the proposed GA-Holt-Winters model and the comparison method, i.e., the unmodified HW model, are discussed.

3.5.1 Settings and Design of Experimentation Machine

The GA-Holt-Winters model advanced in this research is used to predict spectrum occupancy data. The result of the GA-Holt-Winters model was compared with that of the unmodified HW model. The GA-HW technique was implemented in Microsoft Excel 2016 Solver and MATLAB R2018a on a machine configured as follows: Intel Core i5, CPU 2.88 GHz, RAM 8 GB, 64-bit operating system. The MSE was used as the objective function to determine the GA-Holt-Winters model's accuracy in the forecasting of spectrum occupancy. The MSE provides an estimate of the error between the true spectrum occupancy and the forecast by the GA-Holt-Winters model. The closer MSE is to 0, the more accurate the prediction model.

The MSE is computed using Eq. (28).

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^N (y_t - \hat{y}_t)^2 \quad (28)$$

3.5.2 Initialization of the GA-Holt-Winters Model Parameter Settings

The frequency of crossing is determined by the crossover probability. If there is no crossover, the offspring will be identical replica of their parents. If there is crossover, the offspring is made up of chromosomes from both parents. If the likelihood of crossover is 100 percent, then crossing produces all offspring. If it is 0 percent, a new generation is created using exact copies of the chromosomes from an older population. The number of chromosomes in a population is measured by its size (in one generation). If there are not enough chromosomes, the GA only has a few crossover options and only a small portion of the search space is examined. On the other hand, the GA slows down when there are too

many chromosomes. According to previous research, increasing the population size after a certain point (which depends primarily on encoding and the problem) is ineffective because it does not speed up the problem-solving process.

Crossover is done with the assumption that the new chromosomes will have better bits of the old chromosomes and thus are improved. However, it is preferable to leave a portion of the population to the following generation. The chance of chromosomal sections mutating is expressed as a percentage. If there is no mutation, the offspring is taken without any changes following crossover. When mutation occurs, a portion of the chromosome is altered. If the mutation probability is 100%, the entire chromosome is altered; if it is 0%, nothing is altered. The GA is mutated to avoid it from slipping into a local optimum – although this should not happen very frequently – as the GA will then revert to random search.

In this work, the population size = 20, crossover rate = 100%, mutation rate = 10% and the number of evolutions = 1000. The initial HW discount factors were set at 0.3 each.

4 Results and Discussion

The results obtained from the models presented in Section 3 are reported and analyzed in this section.

4.1 Selection of the Choice of Period

To determine the number of periods best suited for the forecast, the MSDs for when the number of periods is 3, 6, 9 and 12 were evaluated. The MSDs at 3 periods were consistently the lowest in each band for the links considered, as can be seen from Table 4.

Table 4 Mean absolute deviation of spectrum duty cycle of selected links.

Bands/Links	Mean Absolute Deviation at selected period number			
	3	6	9	12
GSM 900 RL LOC1	4.00	4.23	4.40	4.56
GSM 900 RL LOC2	1.96	2.08	2.14	2.21
GSM 1800 RL LOC1	6.32	6.89	7.24	7.32
GSM 1800 RL LOC2	4.44	4.71	4.73	4.85
GSM 1800 FL LOC1	3.97	10.63	49.97	86.03
GSM 1800 FL LOC1	0.67	0.71	0.70	0.77

The MSD for all links for when the number of periods was 3 and 12 varied by up to 12% for GSM 900 RL in both locations; 16% for GSM 1800 RL location 1; and 9% for GSM 1800 RL location 2. There was a wide deviation in the case of

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

GSM 1800 FL location 1 and 15% in the case of GSM 1800 FL location 2. Therefore, the number of periods selected was 3.

4.2 Prediction Results

A plot of the prediction results for HW and GA-HW for GSM 900 RL for both locations are shown in Figures 4 and 5 respectively.

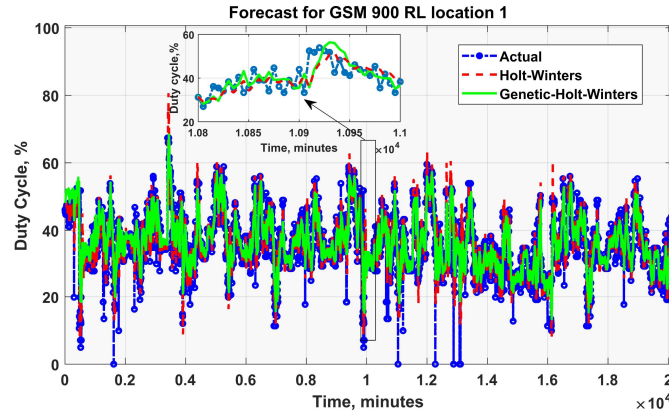


Figure 4 Forecast for GSM 900 RL location 1.

The MSE of the HW was 36.34 while that of GA-HW was 30.51 for the data set of GSM 900 RL location 1.

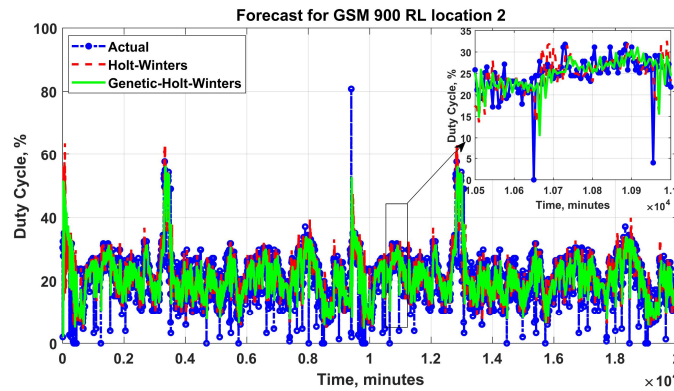


Figure 5 Forecast for GSM 900 RL location 2.

The MSE of HW was 8.91 while that of GA-HW was 7.56 for the data set of GSM 900 RL location 2.

A plot of the prediction results for HW and GA-HW for GSM 1800 RL for both locations are shown in Figures 6 and 7 respectively.

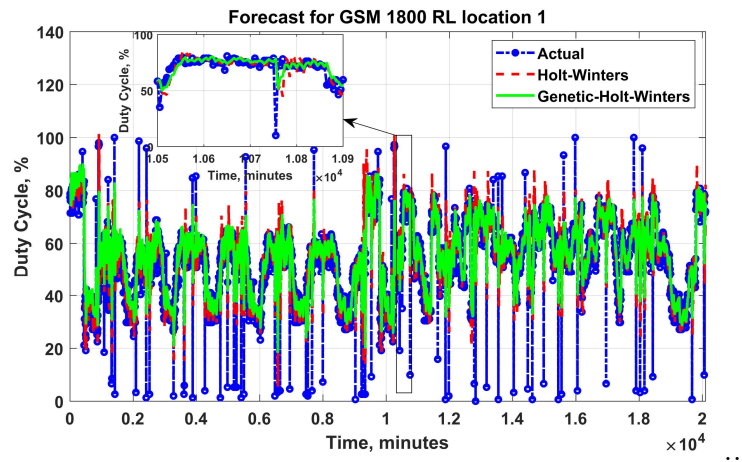


Figure 6 Forecast for GSM 1800 RL location 1.

The MSE of HW was 121.97 while that of GA-HW was 95.58 for the data set of GSM 1800 RL location 1.

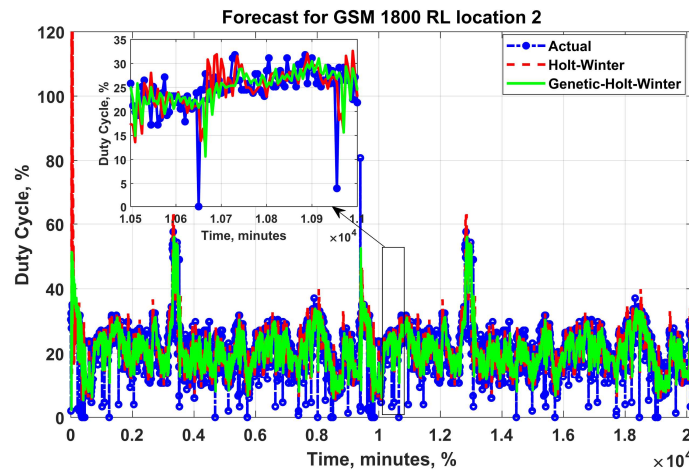


Figure 7 Forecast for GSM 1800 RL location 2.

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

The MSE of HW was 60.96 while that of GA-HW was 33.77 for the data set of GSM 1800 RL location 2. A plot of the prediction results for HW and GA-HW for GSM 1800 FL for both locations are shown in Figures 8 and 9 respectively.

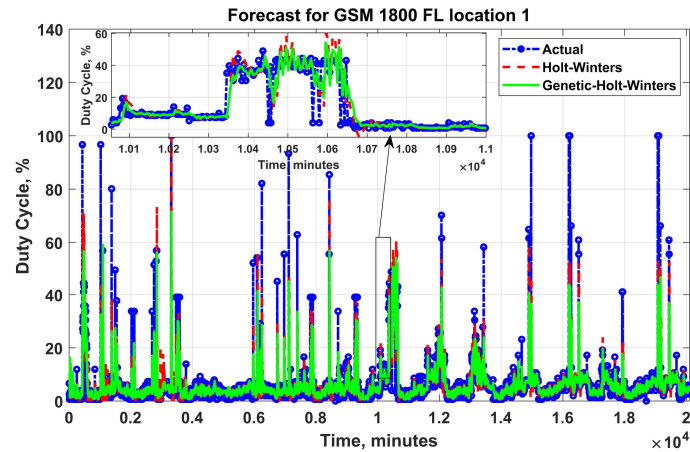


Figure 8 Forecast for GSM 1800 FL location 1.

The MSE of HW was 103.65 while that of GA-HW was 74.86 for the data set of GSM 1800 FL location 1.

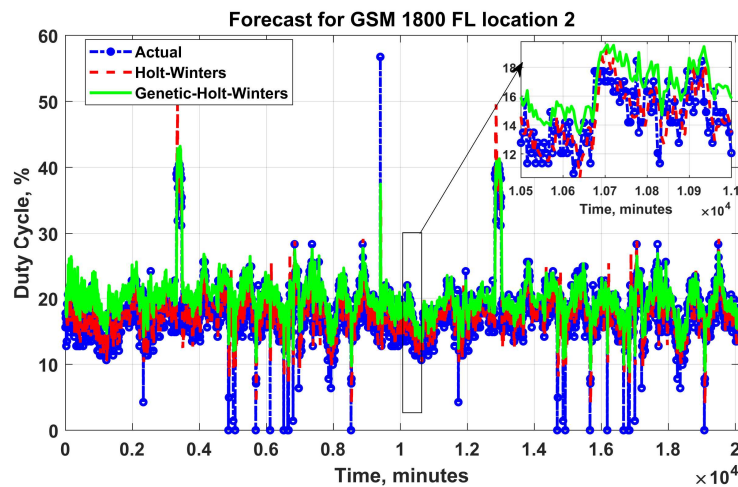


Figure 9 Forecast for GSM 1800 FL location 2.

The MSE of HW was 0.84 while that of GA-HW was 0.77 for the data set of GSM 1800 FL location 1. A good forecast of the spectrum duty cycle was obtained for the links considered. This was due to these spectrum links being characterized with several vacancies. For other links, the MSE obtained are shown in Table 5. Consistently, the MSE values obtained by the GA-Holt-Winters technique were lower than those obtained by the Holt-Winters method. The optimized discount factors, which were modified in the implementation of the GA-Holt-Winters approach are shown in Table 6. The initial values used were 0.30 each. For GSM RL 900 RL location 2, GSM 1800 RL location 1 and GSM 1800 FL location 1, there were no trend components.

Table 5 Comparison of MSE for selected links.

Bands/Links	Forecast Mean Square Error		Percentage Decrease in MSE
	Holt-Winters	GA-Holt-Winters	
GSM 900 RL LOC1	36.34	30.51	16.04
GSM 900 RL LOC2	8.91	7.56	15.15
GSM 1800 RL LOC1	121.97	95.58	21.64
GSM 1800 RL LOC2	60.96	33.77	44.60
GSM 1800 FL LOC1	103.65	74.86	27.78
GSM 1800 FL LOC2	0.84	0.77	8.33

Table 6 Obtained discount factors for minimum MSE of forecast links.

Bands/Links	Discount Factors		
	α	γ	δ
GSM 900 RL LOC1	0.31	0.00	0.11
GSM 900 RL LOC2	0.47	0.00	0.01
GSM 1800 RL LOC1	0.38	0	0.07
GSM 1800 RL LOC2	0.02	0.06	0.48
GSM 1800 FL LOC1	0.38	0.00	0.24
GSM 1800 FL LOC2	0.24	0.23	0.11

In Table 7, the results obtained from GA-enhanced Holt-Winters (GHW) prediction were compared with those of Holt-Winters (HW), ARMA, and ARIMA. GHW consistently performed better than the other techniques.

Table 7 Comparison of prediction results using MSE.

	HW	GHW	ARMA	ARIMA
900 RL	22.629	19.100	25.110	24.920
1800 RL	91.470	64.674	125.530	86.732

5 Conclusion

Spectrum prediction is motivated by the knowledge that continuous spectrum measurement is expensive and time-consuming. In spectrum prediction, historical

Genetic Algorithm-Holt-Winters Based Minute Spectrum Occupancy Prediction: An Investigation

data from spectrum sensing are used in forecasting future spectrum states. In this research, the suitability of the GA modified Holt-Winters exponential model in the prediction of spectrum occupancy data was investigated. The Holt-Winters method and the GA-Holt-Winters technique were observed to show good forecast behavior for GSM 900 RL, GSM 1800 RL, and GSM 1800 FL. There was a decrease of about 16% in the MSE of the GA-Holt-Winters technique compared to the Holt-Winters technique for GSM 900 RL in both locations.

The MSE values for GSM 1800 RL in locations 1 and 2 decreased by 22% and about 45% for the GA-Holt-Winters technique compared to the Holt-Winters technique. Finally, there was a decrease of about 28% and 8% in the MSE of the GA-Holt-Winters technique compared to the Holt-Winters technique for GSM 900 RL in locations 1 and 2 respectively. However, the Holt-Winters method and the GA-Holt-Winters technique could not be applied to GSM 900 FL in both locations owing to the presence of large amounts of unoccupied spectrum slots. A more robust methodology such a neural network-based technique will be used in the future to predict spectrum duty cycle for all links.

Acknowledgments

The Tertiary Education Trust Fund (TETFund), Nigeria is appreciated for providing funding for this research under the Academic Staff Training and Development (AST&D) and Institution Based Research (IBR) programs. We also appreciate the Department of Electrical and Electronics Engineering, University of Ilorin for providing the measurement instrument used in this study. Finally, we are grateful to the Research and Innovation Unit of the Federal Polytechnic Offa, Kwara State in Nigeria.

References

- [1] Faruk, N., Bello, O.W., Sowande, O.A., Onidare, S.O., Muhammad, M.Y. & Ayeni, A.A., *Large Scale Spectrum Survey in Rural and Urban Environments within the 50 MHz - 6 GHz Bands*, Measurement, **92**, pp. 228-238, 2016.
- [2] Ali, S., Chen, Z. & Yin, F., *Spectrum Occupancy of Cellular Networks in Pakistan for Cognitive Radio-Measurements using Spectrum Analyzer*, International Journal of Information and Electronics Engineering, **6**(1), pp. 26-31, 2016.
- [3] Melo, M.A.S., Arroyave, G.A.R. & Quijano, J.L.A., *Spectrum Occupation Assessment in Bogota and Opportunities for Cognitive Radio Systems*, IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC), pp. 932-935, 2018

- [4] Xing, X., Jing, T., Cheng, W., Huo, Y. & Cheng, X., *Spectrum Prediction in Cognitive Radio Networks*, IEEE Wireless Communications, **20**(2), pp. 90-96, 2013.
- [5] Yang, J. & Zhao, H., *Enhanced Throughput of Cognitive Radio Networks by Imperfect Spectrum Prediction*, IEEE Communications Letters, **19**(10), pp. 1738-1741, 2015.
- [6] Hernandez, C., Giral, D. & Martinez, F., *Radioelectric Spectrum Prediction Based in Arima and Sarima Time Series Models*, International Journal of Applied Engineering Research, **13**(22), pp. 15688-15695, 2018.
- [7] Elias, F.G., Fernández, E.M. & Reguera, V.A., *Multi-Step-Ahead Spectrum Prediction for Cognitive Radio in Fading Scenarios*, Journal of Microwaves, Optoelectronics and Electromagnetic Applications, **19**, pp. 457-484, 2020.
- [8] Jain, S., Goel, A. & Arora, P., *Spectrum Prediction Using Time Delay Neural Network in Cognitive Radio Network*. In Smart innovations in Communication and Computational Sciences, Springer, Singapore, pp. 257-269, 2019.
- [9] Tuberquia-David, L.M., Cruz, L. & Hernández, C., *Spectral Prediction: Approaches in Cognitive Radio Networks*, International Journal of Applied Engineering, **13**(10), pp. 8051-8063, 2018.
- [10] Shaghluf, N. & Gulliver, T.A., *Spectrum and Energy Efficiency of Cooperative Spectrum Prediction in Cognitive Radio Networks*, Wireless Networks, **25**(6), pp. 3265-3274, 2019.
- [11] Chakraborty, D. & Sanyal, K.S., *Testbed Implementation of a Scalable ARIMA Model for Spectrum Estimation in Cognitive Radio - A Null Hypothesis Approach*, IETE Journal of Research, 2021.
- [12] Thakur, P. & Singh, G., *Effect of Spectrum Prediction on Cognitive Radio Networks*, in *Spectrum Sharing in Cognitive Radio Networks: Towards Highly Connected Environments*, Wiley, pp. 77-96, 2021.
- [13] Kumar, A., Thakur, P., Pandit, S. & Singh, G., *HSA-SPC: Hybrid Spectrum Access with Spectrum Prediction and Cooperation for Performance Enhancement of Multiuser Cognitive Radio Network*, Computer Networks, **203**, 108596, 2022.
- [14] Singh, A.K. & Ranjan, R., *Multi-Layer Perceptron Based Spectrum Prediction in Cognitive Radio Network*, Wireless Pers Commun, **123**, pp. 3539-3553, 2022.
- [15] Thakur, P. & Singh, G., *Spectrum Mobility in Cognitive Radio Networks Using Spectrum Prediction and Monitoring Techniques*, in *Spectrum Sharing in Cognitive Radio Networks: Towards Highly Connected Environments*, Wiley, pp. 147-166, 2021.
- [16] Supraja, P., Babu, S., Gayathri, V.M. & Divya, G., *Hybrid Genetic and Shuffled Frog-leaping Algorithm for Neural Network Structure Optimization and Learning Model to Predict Free Spectrum in Cognitive*

Genetic Algorithm-Holt-Winters Based Minute Spectrum
Occupancy Prediction: An Investigation

- Radio*, International Journal of Communication Systems, **34**(6), e4532, 2021.
- [17] Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K. & Wang, X. *Deep Learning-based Weather Prediction: A Survey*, Big Data Research, **23**, 100178, 2021.
 - [18] Neu, D.A., Lahann, J. & Fettke, P., *A Systematic Literature Review on State-of-the-art Deep Learning Methods for Process Prediction*, Artificial Intelligence Review, pp. 1-27, 2021.
 - [19] Aghbashlo, M., Peng, W., Tabatabaei, M., Kalogirou, S.A., Soltanian, S., Hosseinzadeh-Bandbafha, H., Omid, M. & Lam, S.S., *Machine Learning Technology in Biodiesel Research: A Review*, Progress in Energy and Combustion Science, **85**, 100904, 2021.
 - [20] Radhwan, A., Kamel, M., Dahab, M.Y. & Hassanien, A.E., *Forecasting Exchange Rates: A Chaos-based Regression Approach*, International Journal of Rough Sets and Data Analysis (IJRSDA), **2**(1), pp. 38-57, 2015.
 - [21] Soto, J., Melin, P. & Castillo, O., *A New Approach for Time Series Prediction Using Ensembles of IT2FNN Models with Optimization of Fuzzy Integrators*, International Journal of Fuzzy Systems, **20**(3), pp. 701-728, 2018.
 - [22] Jiang, W., Wu, X., Gong, Y., Yu, W. & Zhong, X., *Holt–Winters Smoothing Enhanced by Fruit Fly Optimization Algorithm to Forecast Monthly Electricity Consumption*, Energy, **193**, 116779, 2020.
 - [23] Holt, C.E. *Forecasting Seasonals and Trends by Exponentially Weighted Averages*. Carnegie Institute of Technology, Pittsburgh USA. 1957.
 - [24] Winters, P.R., *Forecasting Sales by Exponentially Weighted Moving Averages*. Management Science, **6**, pp. 324-342, 1960.
 - [25] Hyndman, R.J. & Athanasopoulos, G., *Forecasting: Principles and practice*, 2nd Edition, OTexts: Melbourne, Australia, 2018
 - [26] Dantas, T.M., Oliveira, F.L.C. & Repolho, H.M.V., *Air Transportation Demand Forecast Through Bagging Holt Winters Methods*, Journal of Air Transport Management, **59**, pp. 116-123, 2017.
 - [27] Okafor, E.G., Buraimoh, O.C., Uhuegho, K.O. & Soladoye, M.A. *Propensity to Fly in Nigeria: A Forecast of Domestic Air Passenger Traffic Flow in Some Nigeria Airports*, Nigerian Journal of Technology, **38**(2), pp. 456-464, 2019.
 - [28] Kurniasih, N., Ahmar, A.S., Hidayat, D.R., Agustin, H. & Rizal, E., *Forecasting Infant Mortality Rate for China: A Comparison between α -Sutte Indicator, ARIMA, and Holt-Winters*, Journal of Physics: Conference Series, **1028**(1), 012195, IOP Publishing, 2018.
 - [29] Jiang, W., Wu, X., Gong, Y., Yu, W. & Zhong, X., *Holt-Winters Smoothing Enhanced by Fruit Fly Optimization Algorithm to Forecast Monthly*, Energy, **193**, 116779, 2020.

- [30] Yang, Y.M., Yu, H. & Sun, Z., *Aircraft Failure Rate Forecasting Method Based on Holt-Winters Seasonal Model*, IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 520-524, 2017.
- [31] Holland J.H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbe, MI, 1975.
- [32] Enitan, A.M. & Adeyemo, J., *Food Processing Optimization Using Evolutionary Algorithms*, African Journal of Biotechnology, **10**(72), pp. 16120-16127, 2011.
- [33] Sana, S.S., Ospina-Mateus, H., Arrieta, F.G. & Chedid, J.A., *Application of Genetic Algorithm to Job Scheduling under Ergonomic Constraints in Manufacturing Industry*, Journal of Ambient Intelligence and Humanized Computing, **10**(5), pp. 2063-2090, 2019.
- [34] Mirjalili, S., Dong, J.S., Sadiq, A.S. & Faris, H., *Genetic Algorithm: Theory, Literature Review, and Application in Image Reconstruction*, Nature-Inspired Optimizers, pp. 69-85, 2020.
- [35] Roetzel, W., Luo, X. & Chen, D., *Optimal Design of Heat Exchanger Networks*, Design and Operation of Heat Exchangers and their Networks, pp. 231-317, Academic Press, an Imprint of Elsevier, 2020.
- [36] Peng, S., Qin, S. & Li, G., *Predicting Expressway Subsidence Based on Niching Genetic Algorithm and Holt-Winters Model*, Arabian Journal of Geosciences, **12**(354), pp. 1-10, 2019.
- [37] Amzi, N.L.B.M., *Parameters Estimation of Holt-Winter Smoothing Method Using Genetic Algorithm*, Mater of Science Dissertation, Faculty of Science, Universiti Teknologi Malaysia, 2013.
- [38] Metawa, N., Hassan, M.K. & Elhoseny, M., *Genetic Algorithm Based Based Model for Optimizing Bank Lending Decisions*, Expert Systems with Applications, **80**, pp. 75-82, 2017.
- [39] Chiroma, H., Abdul-kareem, S., Noor, A.S.M., Abubakar, A.I., Safa, N.S., Shuib, L., Hamza, M.F., Gital, A.Y. & Herawan, T., *A Review on Artificial Intelligence Methodologies for the Forecasting of Crude Oil Price*, Intelligent Automation & Soft Computing, pp. 1-14, 2016.
- [40] Mathswork, *How the Genetic Algorithm Works?* Accessed from: <https://www.mathworks.com/help/gads/how-the-genetic-algorithm-works.html#243>. (17 June 2022)
- [41] Montgomery, D.C., Jennings, C.L. & Kulahci, M., *Introduction to Time Series Analysis and Forecasting*, Wiley Series in Probability and Statistics. Hoboken. New Jersey, John Wiley & Sons Publication, 2008.
- [42] Faruk, N., Bello, O.W., Sowande, O.A., Onidare, S.O., Muhammad, M.Y. & Ayeni, A.A., *Large Scale Spectrum Survey in Rural and Urban Environments within the 50 MHz - 6 GHz Bands*, Measurement, **91**, pp.228-238, 2016.

Genetic Algorithm-Holt-Winters Based Minute Spectrum
Occupancy Prediction: An Investigation

- [43] Mehdawi, M., Riley, N., Paulson, K., Fanan, A. & Ammar, M., *Spectrum Occupation Survey in Hull-UK for Cognitive Radio Applications: Measurement & Analysis*. International Journal of Scientific & Technology Research, **2**(4), pp. 231-236, 2013.
- [44] Lopez-Benitez, M., Umberto, A. & Casadevall, F., *Evaluation of Spectrum Occupancy in Spain for Cognitive Radio Application*, IEEE in Vehicular Technology Conference (VTC Spring), 69th, Barcelona: IEEE, 2009.