Comparative Modelling and Artificial Neural Network Inspired Prediction of Waste Generation Rates of Hospitality Industry: The Case of North Cyprus

Soolmaz L. Azarmi [1,2,*], **Akeem Adeyemi Oladipo** [3,*] , **Roozbeh Vaziri** [3] **and Habib Alipour** [1]

[1] Faculty of Tourism, Eastern Mediterranean University, Famagusta, TRNC via Mersin 10, Turkey; habib.alipour@emu.edu.tr
[2] Faculty of Tourism, Cyprus Science University, Girne, TRNC via Mersin 10, Turkey
[3] Faculty of Engineering, Cyprus Science University, Girne, TRNC via Mersin 10, Turkey; roozbehvaziri@csu.edu.tr
[*] Correspondence: solmaz.azarmi59@gmail.com (S.L.A.); akeemoladipo@csu.edu.tr (A.A.O.)

check for updates

Abstract: This study was undertaken to forecast the waste generation rates of the accommodation sector in North Cyprus. Three predictor models, multiple linear regression (MLR), artificial neural networks (ANNs) and central composite design (CCD), were applied to predict the waste generation rate during the lean and peak seasons. ANN showed highest prediction performance, specifically, lowest values of the standard error of prediction (SEP = 2.153), mean absolute error (MAE = 1.378) and highest R^2 value (0.998) confirmed the accuracy of the model. The analysed waste was categorised into recyclable, general waste and food residue. The authors estimated the total waste generated during the lean season at 2010.5 kg/day, in which large hotels accounted for the largest fraction (66.7%), followed by medium-sized hotels (19.4%) and guesthouses (2.6%). During the peak season, about 49.6% increases in waste generation rates were obtained. Interestingly, 45% of the waste was generated by British tourists, while the least waste was generated by African tourists (7.5%). The ANN predicted that small and large hotels would produce 5.45 and 22.24 tons of waste by the year 2020, respectively. The findings herein are promising and useful in establishing a sustainable waste management system.

Keywords: urban waste; hospitality sector; waste generation rates; artificial neural network prediction; sustainable waste management

---

1. Introduction

In recent decades, tourism has become one of the most important and significant sectors in the economies of many countries. In fact, the sector accounts for 10–12% of the world's Gross Domestic Product (GDP) and approximately 14% of total employment [1]. Even though tourism can sustain high levels of employment, the sector is a source of environmental impacts with consequent public health concerns [2]. One of the most significant impacts of tourism is the generation of municipal waste, which increases as the seasonal population of tourists rises [3,4]. The heterogeneous nature of hospitality waste poses a significant risk to water and air quality, and is generally liable to cause various health hazards if not properly managed.

From a sustainability perspective, one approach to the reduction of the threatening environmental and health impacts from hospitality generated wastes is the conversion to useful value-added or alternative products. In this regard, several governments have launched policies to promote the

conversion of municipal waste to green precursors or products, many with a specific focus on green fertilizers, bioelectricity, biofuels or bioadsorbents [5–9].

However, to manage and reutilize hospitality sector waste (HSW) in a sustainable way, accurate prediction of HSW generation rate and composition is important [10–12]. A failure to make accurate HSW predictions and assessments could lead to several widespread problems in waste management systems and the environment, including irrelevant policies, increased environmental impacts as well as inadequate or overestimated capacity of disposal infrastructures. Inefficient disposal or waste management infrastructure can cause serious impacts on health [1,4]. Specifically, improperly managed, designed and operated recycling/incineration plants cause air pollution or spread of disease. For instance, hotel kitchen waste ferments after a short time, creating conditions favourable to the growth and survival of microbial pathogens and resulting in the spread of infectious diseases. Also, spent cooking oil is a common hospitality waste; note that unattended spent cooking oil attracts flies, vermin and rats, which could create a health hazard and pest control problem.

Cyprus, politically partitioned into two main parts (south and north), is a major tourist destination in the Mediterranean region. Comparatively speaking, recent years have seen tourism growing at a faster rate in North Cyprus (formally the Turkish Republic of North Cyprus (TRNC)). Meanwhile, in TRNC, the available statistical information regarding waste generation from hospitality industry demonstrates a lack of sufficient reliable data per hospitality facility; hence, it is difficult to develop accurate forecasting systems.

According to the TRNC Hoteliers Association, there was nearly an 83% and 68% bed occupancy rate in the peak and lean seasons, respectively, in 2014–2016. The increasing inflow of tourists in the first quarter of 2017 indicated that the occupancy rate is expected to increase by 6–8% in the peak season of 2017, subsequently leading to more HSW. Of concern is the lack of studies that quantify the magnitude of waste generated in the accommodation sector of TRNC and the subsequent effect of this problem on the environment.

To mitigate the impact of HSW on the ecosystem, we need reliable data concerning HSW generation. Meanwhile, the process of predicting HSW generation is challenging and often intensified by uncontrollable parameters [10,13]. In recent years, various conventional, regression, non-algorithm and descriptive statistical methods of forecasting municipal solid waste (MSW) generation have been reported [13–16].

However, there are limited data concerning the forecasting of HSW generation in the peak and lean seasons as well as an optimal prediction model for this purpose. Hence, this paper tries to contribute to filling the mentioned gaps in the HSW generation rates, specifically in TRNC. The outcome of this research is expected to help policymakers and accommodation sector owners to initiate sustainable waste management practices.

In this study, multiple linear regression (MLR), central composite design (CCD) and artificial neural network (ANN) models were applied in predicting the rate of hospitality sector waste generation. Among these methods, MLR is widely applied to forecasting waste generation due to its simple algorithm and well-developed statistical theory [15]. However, MLR can neither adapt to new situations nor learn from new data; its precision is poor when imprecise data are utilised and it rarely considers all factors affecting waste generation [12,17,18].

CCD under response surface methodology is a combination of a statistical and mathematical technique for empirical modelling of complex problems in which the response of interest is influenced by several independent variables. CCD considers the interaction effects between the operational parameters to produce high prediction accuracy on complex nonlinear systems [19]. To the best of the authors' knowledge, there are no reported data on the application of CCD to forecasting waste generation. ANN is a brain neuron-inspired data-driven technique that can directly learn linear and nonlinear relationships between variables from a set of data compared to the conventional forecasting techniques [20–22].

The strengths and weaknesses of the proposed models were elucidated and an optimal prediction model was established based on conformity with the actual dataset and sensitivity analyses. To date, most studies in this field have focused specifically on the prediction of the total municipal solid waste (MSW) generation rate without considering the interactive effects of the influencing factors (viz., waste management practices, nationality of tourists, nature of waste generated and actual sources of waste in the hospitality facility) to manage HSW sustainably. This paper is written under the belief that the prediction of the amount of HSW produced will be helpful in the stages of transportation, storage, disposal and reutilization and, thus contribute to a sustainable tourism management.

2. Research Methodology

2.1. Research Area and Dataset

Given that the purpose of this research is to predict waste generation rates in the accommodation sector and explore the effects of variables contributing to the waste generation rates, a quantitative approach was employed. Three districts, Nicosia, Famagusta and Girne, were selected to assess the waste generation in the accommodation sectors of TRNC according to the concentrated tourism activities in these districts.

A total of 22 accommodation options, including non-starred guesthouses and large, medium and small hotels, were investigated in this study. Seventy-five percent of these facilities are situated in Girne (a tourism hub), 18% in Nicosia (the capital city) and 7% in Famagusta (a port and student city). The investigated facilities were composed of 36%, 30%, 27% and 7% small, large and medium hotels and guesthouses, respectively. The tourism activities in TRNC remain active seasonally, with the fewest activities taking place in winter (the lean season) and most taking place in summer (the peak season).

A pilot study was conducted to minimise ambiguity in the sampling questions. Also, prior to data collection, the management of the accommodations were assured of confidentiality to minimise the social desirability bias and ensure the accuracy and credibility of the sample data. The data from daily waste generated were collected randomly over a specified period of each month of the lean and peak tourism seasons. We calculated the average daily and yearly generation rate per room and sub-units of the accommodation.

2.2. Model Development and Description of the Input Parameters

The MLR, CCD and ANN as linear, quadratic and non-algorithmic models, respectively, were used to predict the waste generation rates in the hospitality sector of TRNC. To train and test the models, a 3-fold cross-validation procedure was employed to avoid any possible desirability bias. Hence, the average results of three different simulations were compared with the actual data and reported herein.

Among the different parameters that affect the generation rate of hospitality waste, five independent parameters were selected as the most effective ones, including the nationality of tourists visiting the investigated facilities, the nature of waste management practices in each facility, the type of waste generated, the seasonal flow and the type of the accommodation. These parameters were encoded as presented in Table 1.

Table 1. Coding of sub-class of the independent parameters.

Nationality	Season	Accomm. Type	Waste Type	Waste Manage. Practice
British (1)	Lean (0.5)	Guesthouse (0.5)	Food waste (1)	None (2.5)
Russian (2)	Mid (1.5)	Small hotel (2.5)	Glass (2)	Recycle (5.5)
Asian (3)	Peak (2.5)	Medium hotel (4.5)	Cooking oil (3)	Reuse (9.0)
Scandinavian (4)		Large hotel (8.5)	Garden waste (4)	Landfill (12.5)
German (5)			Aluminium (5)	Incineration (16.5)
French (6)			Organic waste (6)	
Turkish (7)			Wood (7)	
Arabs (8)			Paper (8)	
African (9)			Plastic (9)	

The numbers in the parentheses represent the coded value.

2.3. Multiple Linear Regression Analysis

The multiple linear regression (MLR) as a predictive analysis, attempts to explain the relationship between a dependent variable and two or more explanatory variables. The MLR model for predicting the HSW generation can be described as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 +, \ldots, + \beta_n x_n \tag{1}$$

The predicted value of HSW generated is represented by the dependent variable y, $x_1, \ldots, x_n$ represent the five independent variables in this study, and $\beta_0, \ldots, \beta_n$ denote the impact of each independent variable on the response variable.

2.4. Principle of Central Composite Design

Central composite design (CCD) is an efficient approach for modelling complex problems in which the responses are influenced by various independent variables. Hence, we can minimise time consumption and reduce experimental complexities [14]. Herein, the SigmaXL software (Ver 7.0, Ontario, Canada) was employed to generate 5-level-5-factors CCD) matrix. Five independent variables, viz., nationality (A), accommodation type (B), season (C), type of waste (D), and waste management practice (E), were selected based on pilot studies and literature reports to assess their effects on the waste generation rates (WGR).

The independent variables were coded into two levels, low (-1) and high ($+1$), and the axial points are coded as ($+\alpha$) and ($-\alpha$). The total number of experimental data runs generated from the CCD is 44, obtained according to Equation (2):

$$N = 2^x + 2x + x_r = 2^5 + (2 \times 5) + 2 = 44 \tag{2}$$

where N is the total number of runs required, x is the number of variables and x_r is the repeated runs.

The range of the chosen independent variables, with actual and coded levels, is presented in Table 2, where only the most influential runs were selected out of 44 experimental runs. The factorial design comprises 32 full factorials, 10 axial points and two repeated runs, which resulted in an orthogonal distribution of 44 experiments. The experiments were run randomly to minimise errors due to the systematic trends in the factors. A quadratic polynomial regression model was recognised to evaluate and quantify the influence of the variables on the responses obtained from the experiments: The data obtained from the experimental design were utilised to generate a polynomial equation that was analysed to quantify the influence of the variables on the waste generation rates (%).

The results thereafter were subjected to analysis of variance (ANOVA). The ANOVA was applied to evaluate and model the relationship between the response variable (waste generation rates (WGR (%)) and the independent variables, also to test the significance and the adequacy of the model. The efficiency of the quadratic polynomial model was articulated based on coefficients of determination

(R^2), predicted R^2 and adjusted R^2. The statistical significance of the model was verified with Fisher variation ratio (*F*-value), the probability value (Prob > *F*) with 95% confidence level and adequate precision.

Table 2. : Design matrix containing coded values, actual and predicted WGR (%).

Independent Variables	Symbol			Level of Factors				
			$-\alpha$ (−2)	−1	0	1	α (2)	
Nationality	A		1.0	4.0	7.0	11.0	14.0	
Accommodation type	B		0.5	2.5	4.5	6.5	8.5	
Season	C		0.5	1	1.5	2	2.5	
Type of waste	D		1.0	4.0	7.0	11.0	14.0	
Waste manag. practice	E		1.5	5.5	9.0	12.5	16.5	
Run order	**A**	**B**	**C**	**D**	**E**	**Waste generation rate (%)**		
						CCD	ANN	MLR	Actual
1	0	0	0	0	0	73.41	68.32	89.23	61.34
2	0	0	−1	0	0	69.23	89.32	78.01	92.56
3	0	2	0	0	0	67.67	52.11	50.98	56.41
4 *	**1**	**0**	**0**	**0**	**0**	**49.89**	**47.55**	**46.89**	**48.55**
5	−2	0	0	0	0	75.01	94.99	89.23	96.19
6	2	0	0	0	0	83.41	90.11	90.89	89.33
7	0	2	−2	0	0	71.13	82.88	72.88	92.01
8 *	**0**	**−2**	**2**	**0**	**0**	**96.63**	**95.66**	**93.66**	**94.23**
9	0	0	2	−2	0	76.29	75.99	75.23	75.11
10	−2	0	0	2	0	78.46	89.23	64.99	88.21
11 *	**0**	**2**	**0**	**0**	**−2**	**78.22**	**77.01**	**75.55**	**77.46**
12	2	2	2	0	2	74.21	87.98	79.98	87.21
13	1	−1	1	1	1	96.99	76.89	86.89	97.99
14 *	**−1**	**1**	**1**	**1**	**1**	**90.04**	**88.11**	**91.11**	**90.67**
15	1	1	−1	1	1	75.66	88.66	80.23	87.12
16	−1	1	−1	1	1	75.23	83.41	96.19	81.23
17	1	0	1	−1	1	86.44	89.23	59.99	90.11
18	−1	2	1	−1	1	78.01	87.67	92.01	88.11
19	−1	−1	−1	−1	1	95.66	79.89	94.23	80.05
20 *	**−1**	**1**	**−1**	**−1**	**1**	**54.99**	**55.81**	**54.11**	**56.99**

* **Bold** indicates the run orders with desirability function greater than 0.9 and standard deviation less than 0.3.

2.5. Principle of Artificial Neural Network Model

In the late 1990s, the ANN methodology was introduced to tourism forecasting [21]. ANN is a bio-inspired computational processing system akin to the vast network of brain neurons [7]. Lately, research activities in forecasting with ANN have indicated that it can be a promising substitute for conventional linear methods. ANN is highly attractive due to its remarkable characteristics, pertinent particularly to noise and fault tolerance, high parallelism, learning and generalisation capabilities, and nonlinearity [19–23].

The typical ANN architecture is organised in three distinct layers (input, one or more hidden, and an output) containing nodes that are interconnected by weighted synapses. The network structure changes based on the input and output information that flows through it. The independent problem variables are represented in the input layer nodes; the nodes in the hidden layer add an internal representation of non-linear data to the network and the output layer of the ANN is the solution to the problem [21,24].

The relationship between the output (O) and the inputs (I1, I2, I3 . . . ., Ip) is represented mathematically as follows [25]:

$$O_x = b_{x,i} + \sum_{j=1}^{n} w_j f \left(\sum_{i=1}^{m} W_{ji} I_i + B_{oj} \right) \tag{3}$$

where O_x (x = 1, 2, 3, 4, . . .) is the output variable; w_j and W_{ji} (j = 1, 2, 3, . . . , n; i = 0, 1, 2, 3, . . . , m) are connection weights; m and n represent the number of input and hidden nodes, respectively. The f corresponds to the sigmoidal activation function; $b_{x,i}$ and B_{0j} represent the bias terms associated with each input, output and hidden layer nodes, respectively.

In this study, MATLAB R2017a software (MathWorks, Inc., Natick, MA, USA) was utilised to predict the waste generate rates (WGR) of various classes of accommodation sectors in TRNC. A multilayer ANN architecture was utilized and bias neurons were added to each layer to avoid network collapse. The connecting weights were randomly chosen and changed through the training procedure to obtain the minimised mean squared error (MSE). The developed ANN architecture was utilised to investigate the association between inputs and output (waste generation rates), as depicted in Figure 1.
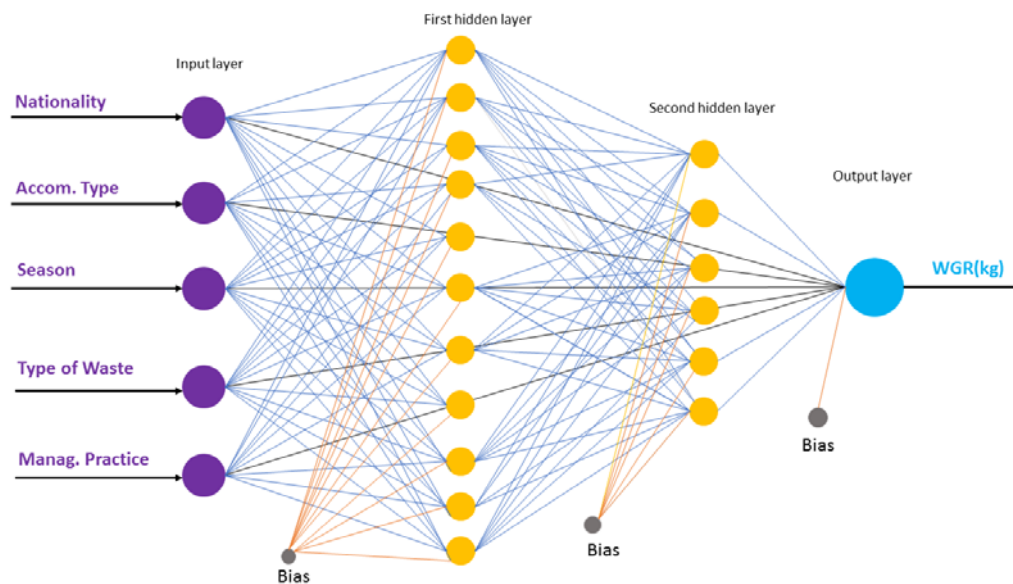

Figure 1. Optimised ANN structure of 5-11-6-1 selected for forecasting WGR in accommodation sectors.

2.6. Model Performance Evaluation

To evaluate the prediction performance of the models, four statistical indices were applied; the hybrid fractional error function (HYBRID), standard error of prediction (SEP), mean absolute error (MAE) and correlation coefficient (R^2) values were derived using the following equations:

$$\text{HYBRID} = \frac{100}{n-p} \sum_{i=1}^{n} \left| \frac{(w_o(t) - w_p(t))^2}{w_o(t)} \right| \tag{4}$$

$$\text{SEP} = \sqrt{\frac{\sum_{i=1}^{n} (w_o(t) - w_p(t))^2}{n-1}} \tag{5}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |w_o(t) - w_p(t)| \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(w_o(t) - w_p(t)\right)^2}{\sum_{i=1}^{n} \left(w_o(t) - w_o'(t)\right)^2} \tag{7}$$

where n is the number of observations, w_o is the observed values of rate of waste generation for type t, p is the number of independent parameters, w_o' is the average of HSW generation and w_p is the predicted value of HSW generation for type t. R^2 measures the closeness of the observed data to the predicted data, MAE is a statistical quantity that measures how close predictions are to the eventual outcomes, and SEP is a measure of the accuracy of the predictions. The smaller the value of the error indices for a specified model, the higher the prediction performance of the model [20,21].

3. Results and Discussion

3.1. Results of MLR

The MLR analysis herein was performed using stepwise regression (SPSS 17.0, Chicago, IL, USA). Its p-value was calculated for each input variable and the significant variable was identified following the criteria of having a p-value ≤ 0.05. The MLR was subjected to 3-fold cross-validation procedures, and the averages of the variable estimates were utilised to obtain the regression equation for predicting generation rate of HSW. Multicollinearity was avoided in the final regression equation using a tolerance filter of 0.5:

$$\text{Total WGR} = -89.561 + 3.5781(\text{A}) + 13.112(\text{B}) - 4.9871(\text{C}) - 11.916(\text{D}) + 7.7812(\text{E}) \tag{8}$$

where A, ... , E symbolise the input parameters described in Table 2. The statistical characteristics of the regression equation are represented in Table 3. As obtained, the MLR model indicates that both waste management practice (E) and accommodation type (B) are highly significant parameters, followed by season (C), which is slightly significant and influenced the generated waste quantity with an α-level less than 0.1.

Table 3. Statistical characteristics of the developed MLR model.

Predicted WGR	Parameter	Parameter Coefficient	T-Value	α-Level	Standard Error
Total	Intercept	−89.56	−1.41	0.32	63.7
	A	3.578	1.89	0.56	1.89
	B	13.11	0.95	0.04	13.8
	C	−4.987	−1.25	0.08	−3.98
	D	−11.92	3.41	0.16	3.5
	E	7.781	−0.88	0.02	−8.89
Peak season	Intercept	−29.56	−0.84	0.18	34.8
	A	2.254	1.63	0.96	1.38
	B	6.89	1.16	0.00	5.96
	C	1.855	0.77	0.07	2.39
	D	−6.365	−0.31	0.28	20.3
	E	4.332	−0.38	0.05	−11.3
Lean season	Intercept	−59.89	5.04	0.68	−11.89
	A	1.324	1.14	0.16	1.16
	B	6.21	1.71	0.03	3.63
	C	−3.134	−0.34	0.07	9.32
	D	−5.555	1.14	0.32	−4.89
	E	3.449	0.055	0.00	62.8

T indicates the statistical significance of the model parameter; Acceptable α-level = 0.1.

3.2. Analysis of AAN Model Results

In the current study, 138 experimental datasets were fed into the ANN network and randomly classified. Of these, 68% of the datasets were trained, 18% were tested and the remaining 14% were validated. The three-layered feedforward neural network herein consists of the logsig transfer function at hidden layer and a linear transfer function (purelin) at the output layer. To minimise network error, numerical overflows and achieve higher homogeneous results, the model inputs and output were normalised and scaled in the rank of 0.1 to 0.9 using Equation (9):

$$X^* = 0.1 + \frac{X - X_{\min}}{X_{\max} - X_{\min}} 0.8 \tag{9}$$

where the normalised value of the output variable is X^*, and X, X_{max} and X_{min} represent their actual, maximum and minimum values, respectively.

The connection weights of the trained ANN with the corresponding bias terms were employed to estimate the relative significance of each independent variable (I_i) on waste generation rate (kg/day) as given in Equation (10):

$$I_i = \frac{\sum_{m=1}^{m=N_h} \left(\left(\left| W_{jm}^{ih} \right| / \sum_{k=1}^{N_i} \left| W_{km}^{ih} \right| \right) \times \left| W_{mn}^{ho} \right| \right)}{\sum_{k=1}^{k=N_i} \left\{ \sum_{m=1}^{m=N_h} \left(\left| W_{km}^{ih} \right| / \sum_{k=1}^{N_i} \left| W_{km}^{ih} \right| \right) \times \left| W_{mn}^{ho} \right| \right\}} \tag{10}$$

where I_i is the relative importance of the *i*th input variable on the response; W, N_i and N_h represent the connection weights, numbers of input and hidden neurons, respectively. The subscripts '*k*', '*m*' and '*n*' is the input, hidden and output neuron, while the superscripts '*i*', '*h*' and '*o*' represent the input, hidden and output layers, respectively.

3.2.1. Selection of Backpropagation (BP) Training Algorithm

Ten BP algorithms were investigated to select the best BP training algorithm, as depicted in Table 4. The highest value of the degree of correlation (R^2) and the least mean square error (MSE) were used as the yardstick to select the best BP. Of all the BP algorithms examined, the Levenberg–Marquardt (LMA) BP algorithm specifically resulted in the least mean square error (0.0014) and its R^2 value (0.989) is closest to unity. Hence, LMA was selected as the training algorithm in this research.

Table 4. Comparison of backpropagation algorithms.

Backpropagation (BP) Algorithm	MSE	Epoch	R^2	Best Linear Equation
Fletcher–Reeves conjugate gradient BP	0.0323		0.867	y = 0.0671x + 0.0678
Batch gradient descent	0.0098	1000	0.918	y = 0.6612x + 0.0893
Scaled conjugate gradient BP	0.0167	78	0.456	y = 0.7905x + 0.0698
One step secant backpropagation	0.0088	32	0.789	y = 0.0622x + 0.0256
Powell–Beale conjugate gradient BP	0.0433	56	0.908	y = 0.5623x + 0.0998
* Levenberg-Marquardt backpropagation	**0.0014**	**11**	**0.989**	**y = 0.9017x + 0.0219**
BFGS quasi-Newton backpropagation	0.0093	30	0.965	y = 0.7221x + 0.0083
Variable learning rate backpropagation	0.007	178	0.671	y = 0.9323x + 0.0044
Polak–Ribi'ere conjugate gradient BP	0.0091	32	0.379	y = 0.9011x + 0.0391
Batch gradient descent with momentum	0.0205	1000	0.881	y = 0.8312x + 0.6733

Bold indicates the best and selected BP.

3.2.2. Optimisation of Neuron Number

To obtain an optimised ANN structure, robust networks were constructed by varying the iteration, hidden neurons and learning rates. Our networks were trained perfectly with over 1000 iterations and the optimal learning rate was 0.2, as listed in Table 5.

Table 5. The parameters of the optimised ANN model used in this study.

Parameter	Value
Input layer neurons	5
Output layer neurons	1
Hidden layers	2
Hidden layer neurons	17
Training method	Levenberg–Marquardt backpropagation
Error goal	0.015%
Epochs	1000
Data division	Random
Momentum (mu)	0.001
Transfer function of hidden layer	Logsig
Learning rate	0.2

The assessment of MSE during training and testing for an optimum number of neurons in the hidden layers is presented in Table 6. As seen in the training set, the MSE was 0.0923 when 13 neurons were used and decreased to 0.0131 when 17 neurons were utilised. The MSE reached a minimum level and increasing the number of neurons beyond 17 does not decrease the MSE further. Hence, 17 neurons were chosen as optimum for the developed ANN topology shown in Figure 1.

The optimised neural network model was used to predict the amount of waste generated in two different seasons (peak and lean) by considering a different type of waste. The comparison between the ANN predictive values, CCD, MLR and the actual values is shown in Figure 2.

Table 6. Optimisation of neuron number at hidden layer, using testing and training data set.

Number of Neurons	Training Set			Testing Set		
	MAE	MSE	R^2	MAE	MSE	R^2
1	0.1278	0.3219	0.7865	0.0899	0.7012	0.6821
2	0.1806	0.3131	0.6679	0.0986	0.6131	0.6178
3	0.1311	0.3094	0.9012	0.0911	0.5694	0.8611
4	0.1155	0.2308	0.8694	0.0888	0.5308	0.7389
5	0.0969	0.1396	0.6311	0.0814	0.4388	0.6398
6	0.0889	0.1131	0.8332	0.0768	0.4098	0.8798
7	0.0835	0.1094	0.5694	0.0732	0.3994	0.6873
8	0.0811	0.1088	0.6377	0.0711	0.3768	0.7997
9	0.0678	0.1046	0.8296	0.0689	0.3288	0.7654
10	0.0561	0.0991	0.7156	0.0605	0.3109	0.5679
11	0.0458	0.0934	0.8967	0.0598	0.2987	0.6899
12	0.0449	0.0808	0.7855	0.0534	0.2855	0.6656
13	0.0328	0.0623	0.8987	0.0511	0.2616	0.8202
14	0.0389	0.0431	0.0934	0.0109	0.0934	0.8144
15	0.0356	0.0334	0.0557	0.0098	0.0557	0.6098
16	0.0298	0.0308	0.9131	0.0082	0.0198	0.7813
17	**0.0211**	**0.0131**	**0.9933**	**0.0067**	**0.0125**	**0.8989**
18	0.0469	0.0134	0.8131	0.0096	0.0139	0.5199
19	0.0209	0.0308	0.8694	0.0734	0.0394	0.7656
20	0.0668	0.0396	0.7366	0.0611	0.0108	0.8088

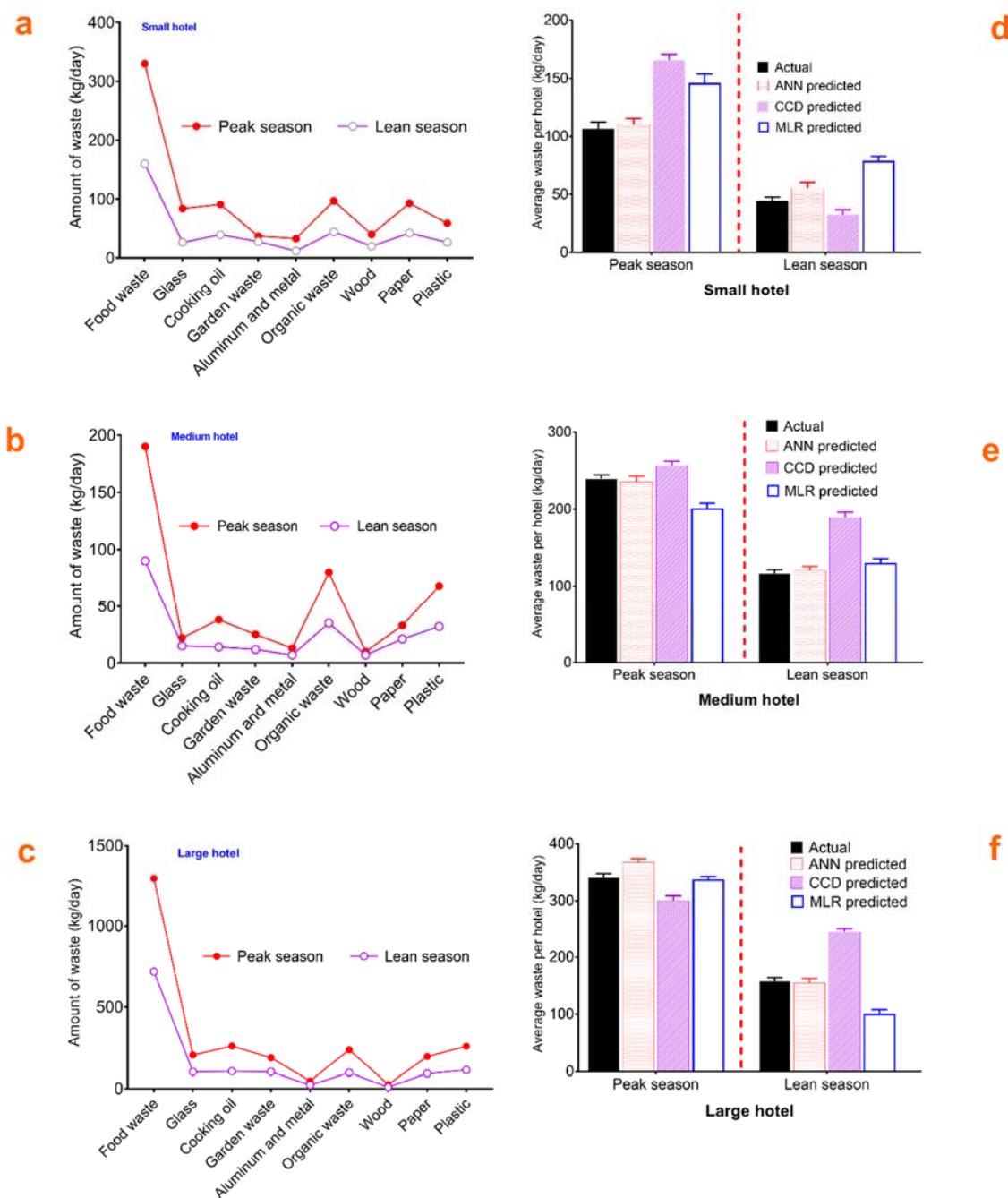**Bold** show optimum number of hidden layer neuron.

Figure 2. Type and amount of waste generated per day in each season (**a**–**c**), and actual and predicted average daily waste generated per (**d**) small hotel, (**e**) medium hotel, and (**f**) large hotel.

As shown in Figure 2a–c, food waste is the most generated waste in all the investigated facilities. In the lean season, a total of 970 kg of food waste was generated per day, while this figure increased by almost 1.9% in the peak season. Organic waste (vegetables, milk, bread, etc.) is also commonly generated in all facilities. A total of 415 kg/day of organic waste is generated during the peak season, when the large hotels account for 57%, small hotels 23.4% and medium hotels 19.6%. During the lean season, the organic waste generated decreased to 178.5 kg/day; small hotels generated the least (19.7%). The least generated waste in all the facilities appears to be wood. Only 61 kg/day of wood was generated during the peak season and 28 kg/day was generated during the lean season.

Figure 2d–f, represents the average waste generated per day. During the peak season, the average waste generated was 106.5, 239.5 and 340.9 kg/day by small, medium and large hotels, respectively. Meanwhile, the average waste generated decreased to about 63% during the lean season. As mentioned, the type of waste management practices in each facility has a significant influence on the waste generation rate, as does the nationality of the tourists visiting the facilities.

The nationality influences the environmental performance in each facility [26]. For instance, based on our research, hotels that have tourists from Arab countries have high food wastage, while those with guests from Britain presented high water consumption and more organic waste, which could be attributed to lifestyle. Regarding the nationality of most tourists visiting the investigated facilities, about 33% British visited the facilities followed by Russian (18%), Turkish (16%), Scandinavian (10%), German (8%), Arab (7%), French (5%) and African (3%).

Figure 3a–c shows the observed waste generation rate (%) per day based on the tourist nationality considering peak season. In a small hotel, 19.21% waste is generated by British tourists per day (85.6 kg), Turkish tourists generated 65.8 kg waste, which is 16.29 WGR (%)/day, and the least waste was generated by French tourists (5.67%) per day. A similar pattern is observed in medium-sized hotels; however, Arabs generated the least waste, accounting for 2.2% WGR per day. In contrast, Turkish tourists generated the most waste in large hotels (19.16%), followed by the British (15.1%), and the least waste was generated by Asians (6.9%) in large hotels. This research has helped us to understand the pattern of visitors in each facility and their range of waste generating practices.

The waste generation rate per type of facility based on the type of waste management practices is shown in Figure 3d–f. Many hotel facilities take very little action to reduce their environmental impact; specifically, small hotels regard their environmental responsibility as a secondary objective [27]. In most cases, small hotels commonly generate low quantities of waste that are unattractive to waste recycling firms since they often require specific quantities of waste to be collected [27,28]. In this research, the WGR/day of small hotel firms without any waste management practices is 30%, while the WGR of those that engaged in landfill practice is 40% per day. In medium-sized and large hotels without waste management practices, 50% and 70% of WGR, respectively, was observed per day. This is largely due to their increased room occupancy and higher range of waste generating services compared with small hotels.

The ANN prediction appears to be in reasonable agreement with the observed data. More deviation between the actual (residuals) and predicted values was observed in the CCD and MLR models than in the ANN model. Hence, the higher predictive capacity of the ANN can be attributed to its universal capability to approximate complex nonlinear systems, whereas CCD is effective if the system is restricted to second-order polynomial regression [22,29].

3.3. Analysis of CCD

The correlation between the response (WGR) and independent factors (Table 2) was developed using the CCD of the SigmaXL software (ver. 7.0, Ontario, Canada). The standard deviation and correlation coefficient were utilised to evaluate the fitness of the models developed. The smaller the standard deviation and the closer the R^2 value to unity, the better the model is at forecasting the response [19]. Table 7 indicates that the quadratic model was not aliased and has a comparatively low standard deviation of 3.361 and relatively high R^2 value of 0.9985, which is in reasonable agreement with the predicted R^2 (0.9966). Also, the PRESS of the quadratic equation is low (169.23), which revealed the reliability and better precision of the experimental results. Hence, the results indicate that the quadratic model can be used to describe the relationship between the response (WGR) and the interacting variables. Hence, the codified quadratic equation after eliminating the insignificant terms is shown in Equation (11):

$$\text{Total WGR} = 129.5 - 8.78(A) + 45.3(B) - 0.92(C) - 1.92(D) + 58.1(E) + 0.78AB \\ + 3.86BE - 0.89A^2 + 2.98B^2 - 2.56C^2 + 5.87E^2 \tag{11}$$
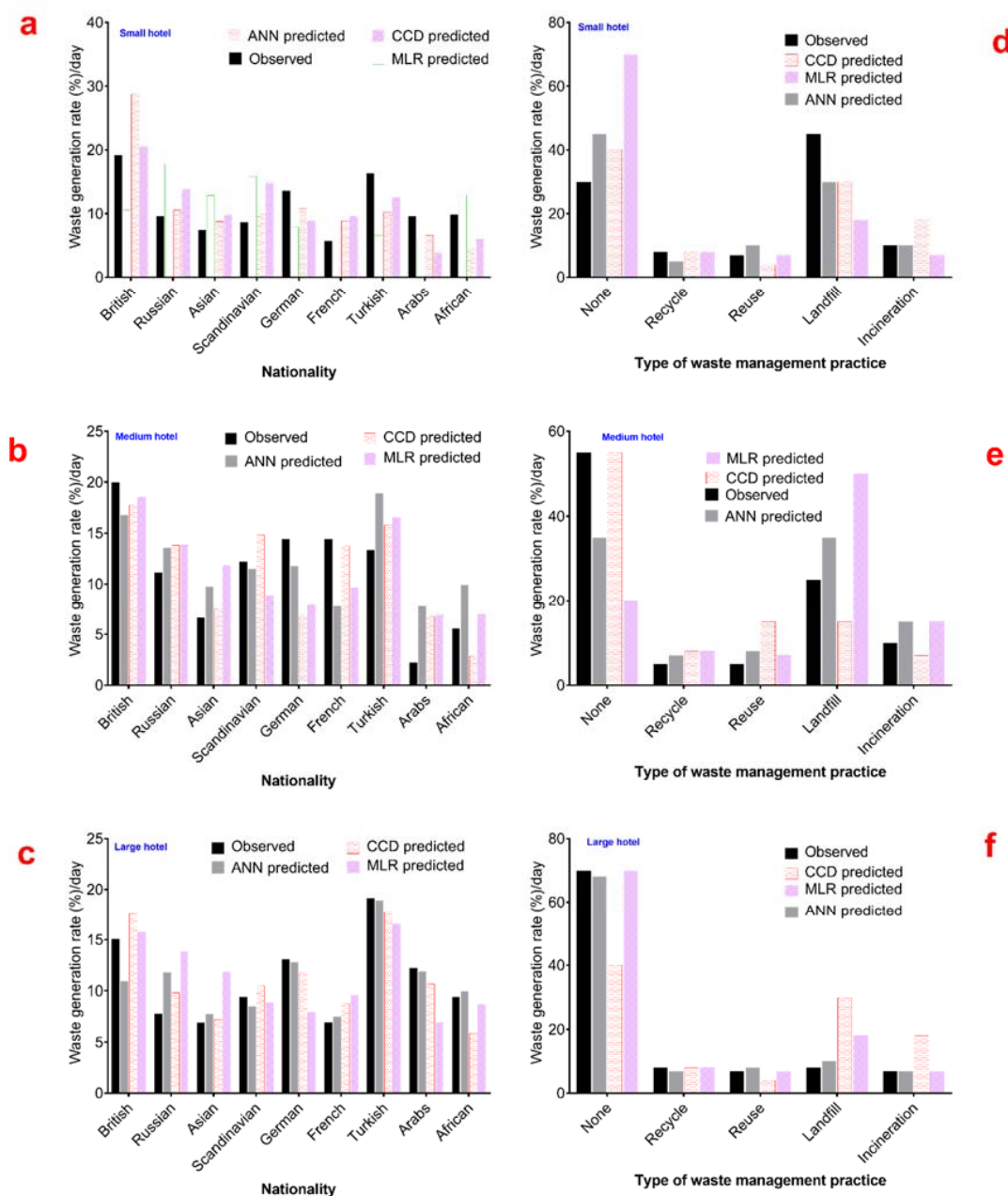
Figure 3. Waste generation rate (%)/day based on tourist nationality (**a**–**c**) and based on the type of waste management practice in each facility (**d**–**f**) during the peak season.

The obtained quadratic equation was further evaluated using ANOVA. As tabulated, the quadratic model for HSW generation rate has an *F*-value of 329.55 and a *p*-value of 0.0012, implying that the model is significant. For the model terms, the largest *F*-value signifies the most significant effect on the response variable and the model term with a *p*-value less than 0.05 is significant [30]. In this case, the significant model terms are A, B, C, E, AB, BE, A^2, B^2 and E^2, while C^2, BC and AE are insignificant. The model term having the most significant effect on the response is B with an *F*-value of 679.87. The "Lack of Fit" *F*-value of 1.76 signifies that it is not significant relative to the pure error and there is a 65.88% chance that its *F*-value being this large could be due to noise [22]. The non-significant "Lack of fit" for WGR indicated the good predictability of the model.

The variability of the independent variables was evaluated based on the results of Equation (11); the negative value of the coefficient A (-8.78) indicates that as the nationality of the tourists

changes from 1 to 9, as coded in Table 1, the HSW generation rate decreased. We inferred that British tourists generate more waste than Arab tourists. The positive values of the coefficients indicate that these parameters had a positive effect on the HSW generation rate. For instance, the smaller hotels generate less waste compared to the large hotels with a range of services and greater room occupancy. The interactive influence of the independent variables was investigated and depicted via three-dimensional (3D) response surface plots. The 3D surface plots are very effective at observing complex systems in which two or more variables are significant [31–33].

As shown in Figure 4a, the interaction between the season (C) and the type of waste management practice (E) in each facility indicates increased HSW generation rates as C tended towards the peak season (2.5) with accommodation sectors without waste management practices. Figure 4b indicates that the HSW generation rate increased from 65% to 85% in a facility without proper waste management practices, specifically with tourists in the range 1−4 (Table 1). Figure 4c shows that about 44% WGR was observed in the lean season for tourists of nationality 7−9; however, the WGR increased beyond 80% in the lean season when tourists of nationality 1−5 occupied these facilities. Figure 4d depicts the interactive influence of accommodation type (B) and tourists' nationality (A). According to Equation (10), AB is statistically significant ($p = 0.0002 < 0.05$) and the WGR increased as the accommodation type changed from small to large hotels as the tourist nationality increased from 1 to 9. This indicates that the WGR is higher at large hotels irrespective of the nationality of the tourists.

Table 7. Model Summary statistics and ANOVA for the regression model for WGR.

Source	Std Dev.	R^2	Adj. R^2	Pred. R^2	PRESS	Remark
Linear	20.93	0.5617	0.6541	0.6045	6489.65	
2FI	9.956	0.9014	0.7681	0.8988	5109.99	
Quadratic	3.361	0.9985	0.9864	0.9966	169.23	Suggested
Cubic	2.951	0.8941	0.8679	0.7899	3705.82	Aliased

Source	Sum of squares	d*f*	Mean square	*F*-value	Prob > F	
Model	16283.54	9	1809.282	329.55	0.0012	Significant
A	738.38	1	738.38	6.5161	<0.0001	
B	10544.33	1	10,544.33	679.87	<0.0001	
C	6318.32	1	6318.32	14.788	<0.0001	
D	307.43	1	307.43	2.0871	0.0494	
E	189.66	1	189.66	298.34	<0.0001	
AB	987.11	1	987.11	679.87	<0.0001	
AC	133.68	1	133.68	14.788	0.0496	
AD	1569.56	1	1569.56	24.661	0.0586	
AE	1875.55	1	1875.55	98.343	0.3556	
BC	334.38	1	334.38	19.878	0.0001	
BD	564.99	1	564.99	3.7881	0.0589	
BE	68.55	1	68.55	4.6619	<0.0001	
CD	348.99	1	348.99	8.3466	0.1558	
CE	167.57	1	167.57	6.8745	0.2596	
DE	1038.44	1	1038.44	1.7889	0.3856	
A^2	568.66	1	568.66	4.0911	<0.0001	
B^2	568.77	1	568.77	8.3421	<0.0001	
C^2	38.38	1	38.38	9.8731	0.6018	
D^2	138.55	1	138.55	14.788	0.2429	
E^2	1705.38	1	1705.38	150.88	<0.0001	
Residual	1239.345	11	112.667			
Lack of fit	678.4321	8	84.804	1.758	0.6588	Not significant
Pure error	560.9128	5	112.183			
Total	17522.89	20				

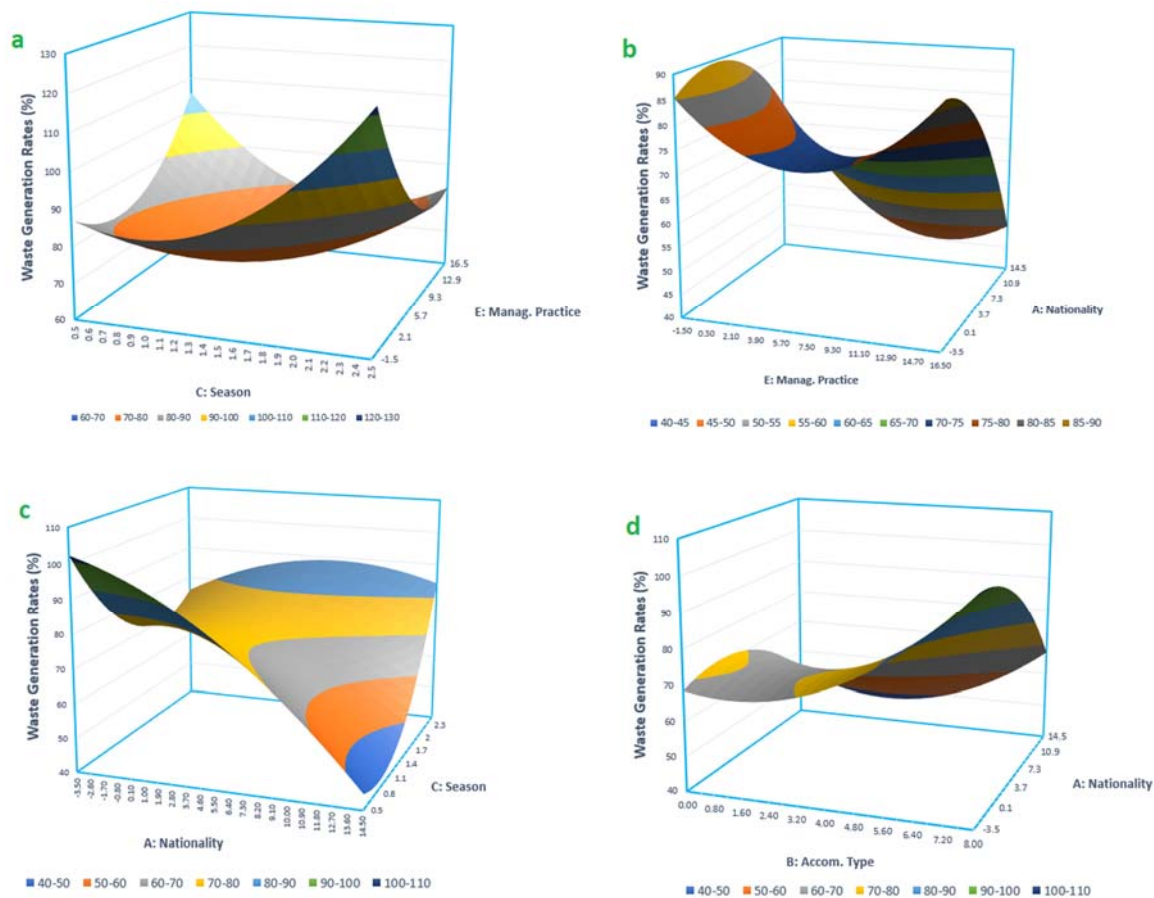PRESS: predicted residual error sum of squares; df: degree of freedom.

Figure 4. Response surface plots for interactive influence of independent parameters on the WGR. Waste generation rate (%) as a function of (**a**) season and management practice (**b**) management practice and tourist nationality (**c**) season and nationality (**d**) accommodation type and nationality

3.4. Estimation of Waste Generated and Comparison of Predictive Performance of Models

Forecasting of HSW generation can be classified into short-term (ranging from days to few months), mid-term (few months to 2–4 years) and long-term forecasting (> 5 years). Table 8 summarises the estimated total waste generated by each facility investigated and comparative predictive performance of each model. Herein, statistical analysis was performed to compare the constructed ANN, CCD and MLR models, in terms of their predictive performance, using HYBRID, R^2, MAE and SEP (Equations (4)–(7); Table 8).

The ANN model shows the lowest error values and highest R^2 compared to the CCD and MLR models. Based on the obtained results, the ANN architecture is more reliable and accurate in terms of predictive capability and fitting to the non-linear relationship between the variables and HSW generation rate. On the other hand, one of the most significant advantages of the CCD-based model is its ability to clarify the interactive effect of the variables on the response (WGR), which highlights its usefulness in predicting the rate of HSW generation. Hence, combining the abilities of CCD and ANN models in a hybrid fashion could result in powerful modelling and predictive models.

Table 8. Estimated total waste generated and comparison of predictive performance of models

Accommodation Type	Season	Per Day	Predicted for Next 3 Years (kg)		
		Observed (kg)	ANN	CCD	MLR
Small hotel	Peak	864.1	3092.3	2292.3	2679.8
	Lean	399.5	1848.5	1679.6	1799.5
Medium hotel	Peak	479.5	1870.1	1822.1	1987.9
	Lean	233.7	934.8	978.9	698.7
Large hotel	Peak	2727.8	12275.1	11891.7	12098.5
	Lean	1377.3	7898.9	9873.9	88761.8
Guesthouse	Peak	88.5	389.8	334.9	278.6
	Lean	52.9	198.6	160.7	256.8
Statistical parameters	**ANN**	**CCD**	**MLR**		
R^2	0.9982	0.8982	0.9054		
MAE	1.378	1.469	3.981		
SEP	2.153	4.71	9.891		
HYBRID	98.781	103.4	145.9		

3.5. Sensitivity Analysis and Relative Importance of Input Variables

Figure 5 shows the relative significance of the independent variables on the response (WGR) obtained from Equation (10). It is important to stress that all the input variables had an impact on the waste generation rates. However, the accommodation type (B) seemed to be the most influential variable on WGR, while the second most influential variable was the season (C), followed by waste management practice (E), nationality (A) and type of waste (D), according to ANN. Meanwhile, the CCD and MLR tend to deviate slightly from the ANN predicted data, as shown in Figure 5. The desirability function (D) was applied to select the acceptable ranking and the minimum, middle and maximum values of desirability were configured as D = 0.0, 0.5 and 1.0, respectively. A desirability value closer to 1.0 means that the corresponding sensitivity analysis is able to represent the actual scenario. Hence, the ranking based on the models is as follows; ANN (D = 0.99) > CCD (D = 0.78) > MLR (0.65). It is inferred that the ANN can be employed to simulate and predict the complex independent variable behaviour in any form of non-linearity and can effectively overcome the limitation of quadratic correlation assumed in CCD and MLR.
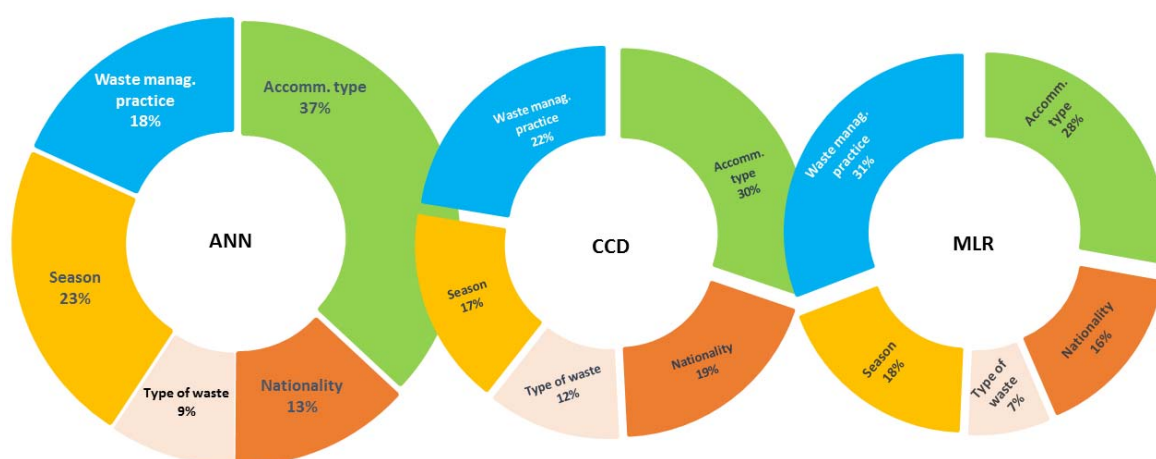
Figure 5. Relative importance of input variables on WGR forecasted by ANN, CDD and MLR.

4. Conclusions

The accommodation sector is an essential component of the tourism and travel business. It is worth mentioning that increases in hospitality sector operations result in increased quantities of

municipal waste, constituting ecosystem damage and a significant increase in the environmental footprint. To curtail the ugly face of tourism activities, precise prediction of the quantity of hospitality waste generated is required to enable the development of an integrated waste management and reutilization system. Note that inaccurate prediction of hospitality waste generated may result in a negative impact on the environment.

For the first time, this study has shown that municipal waste from hospitality facilities can be forecasted by considering measurable and effective parameters via an artificial neural network-inspired forecasting model. The hospitality waste generation rates were analysed based on three categories: recyclable, general waste and food residue. ANN, CCD and MLR were employed to predict the average HSW generation rate using nationality, type of waste, season, accommodation type, and type of waste management practices as predictors. These predictors were selected based on the correlation test and Cronbach's alpha of 0.93. The results showed that 4159.9 kg (recyclable: 58.5%, general waste: 23.6% and food residue: 17.9%) and 2063.4 kg (recyclable: 33.6%, general waste: 18.5% and food residue: 47.9%) of waste were generated during the peak and lean season from the 22 hospitality facilities investigated, respectively.

Importantly, the use of the ANN model to predict the average HSW generation rate led to reliable results and the difference between the observed and predicted values was not statistically significant. However, the MLR model demonstrated lower prediction accuracy compared to CCD. It was found that Turkish tourists generated more waste (19.16% WGR/day) in large hotels compared with the British (15.1% WGR/day), and Asians generated the least average waste (20.96%) in all the facilities investigated. The findings of this study imply the need for further research to investigate the possible sources of the waste and factors limiting hotels from managing the waste effectively. In conclusion, the results herein are promising and would be useful in establishing a sustainable waste management plans.

References

1. Çiçek, D.; Zencir, E.; Kozak, N. Women in Turkish tourism. *Waste Manag.* **2017**, *31*, 228–234. [CrossRef]
2. Mateu-Sbert, J.; Ricci-Cabello, I.; Villalonga-Olives, E.; Cabeza-Irigoyen, E. The impact of tourism on municipal solid waste generation: The case of Menorca Island (Spain). *Waste Manag.* **2013**, *33*, 2589–2593. [CrossRef] [PubMed]
3. Shamshiry, E.; Nadi, B.; Mokhtar, M.B.; Komoo, I.; Hashim, H.S.; Yahaya, N. Integrated Models for Solid Waste Management in Tourism Regions: Langkawi Island, Malaysia. *J. Environ. Public Health* **2011**, *2011*, 709549. [CrossRef] [PubMed]
4. Arbulú, I.; Lozano, J.; Rey-Maquieira, J. Tourism and solid waste generation in Europe: A panel data assessment of the Environmental Kuznets Curve. *Waste Manag.* **2015**, *46*, 628–636. [CrossRef] [PubMed]
5. Hanifzadeh, M.; Nabati, Z.; Longka, P.; Malakul, P.; Apul, D.; Kim, D.S. Life cycle assessment of superheated steam drying technology as a novel cow manure management method. *J. Environ. Manag.* **2017**, *199*, 83–90. [CrossRef] [PubMed]
6. Molino, A.; Larocca, V.; Chianese, S.; Musmarra, D. Biofuels Production by Biomass Gasification: A Review. *Energies* **2018**, *11*, 811. [CrossRef]
7. Oladipo, A.A.; Ifebajo, A.O. Highly efficient magnetic chicken bone biochar for removal of tetracycline and fluorescent dye from wastewater: Two-stage adsorber analysis. *J. Environ. Manag.* **2018**, *209*, 9–16. [CrossRef] [PubMed]

8.　Kim, D.S.; Hanifzadeh, M.; Kumar, A. Trend of biodiesel feedstock and its impact on biodiesel emission characteristics. *Environ. Prog. Sustain. Energy* **2018**, *37*, 7. [CrossRef]

9.　Oladipo, A.A.; Ifebajo, A.O.; Nisar, N. High-performance magnetic chicken bone-based biochar for efficient removal of rhodamine-B dye and tetracycline: Competitive sorption analysis. *Water Sci. Technol.* **2017**, *76*, 373–385. [CrossRef] [PubMed]

10.　Beigl, P.; Lebersorger, S.; Salhofer, S. Modelling municipal solid waste generation: A review. *Waste Manag.* **2008**, *28*, 200–214. [CrossRef] [PubMed]

11.　Batinic, B.; Vukmirovic, S.; Stanisavljevic, N.; Ubavin, D.; Vukmirovic, G. Using ANN model to determine future waste management targets-case study of Serbia. *J. Sci. Ind. Res.* **2011**, *70*, 513–518.

12.　Intharathirat, R.; Salam, P.A.; Kumar, S.; Untong, A. Forecasting of municipal solid waste quantity in a developing country using multivariate grey models. *Waste Manag.* **2015**, *39*, 3–14. [CrossRef] [PubMed]

13.　Abbasi, M.; Hanandeh, A.E. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. *Waste Manag.* **2016**, *56*, 13–22. [CrossRef] [PubMed]

14.　Sha'Ato, R.; Aboho, S.Y.; Oketunde, F.O.; Eneji, I.S.; Unazi, G.; Agwa, S. Survey of solid waste generation and composition in a rapidly growing urban area in Central Nigeria. *Waste Manag.* **2007**, *27*, 352–358. [CrossRef] [PubMed]

15.　Xu, L.; Gao, P.; Cui, S.; Liu, C. A hybrid procedure for MSW generation forecasting at multiple time scales in Xiamen City, China. *Waste Manag.* **2013**, *33*, 1324–1331. [CrossRef] [PubMed]

16.　Denafas, G.; Ruzgas, T.; Martuzevicius, D.; Shmarin, S.; Hoffmann, M.; Mykhaylenko, V.; Ogorodnik, S.; Romanov, M.; Neguliaeva, E.; Chusov, A.; et al. Seasonal variation of municipal solid waste generation and composition in four East European cities. *Resour. Conserv. Recycl.* **2014**, *89*, 22–30. [CrossRef]

17.　Noori, R.; Karbassi, A.; Sabahi, M.S. Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *J. Environ. Manag.* **2010**, *91*, 767–771. [CrossRef] [PubMed]

18.　Constantino, H.A.; Fernandes, P.O.; Teixeirac, J.P. Tourism demand modelling and forecasting with artificial neural network models: The Mozambique case study. *Tékhne-Rev. Appl. Manag. Stud.* **2016**, *14*, 113–124. [CrossRef]

19.　Oladipo, A.A.; Gazi, M. Nickel removal from aqueous solutions by alginate-based composite beads: Central composite design and artificial neural network modeling. *J. Water Process. Eng.* **2015**, *8*, e81–e91. [CrossRef]

20.　Jahandideh, S.; Jahandideh, S.; Asadabadi, E.B.; Askarian, M.; Movahedi, M.M.; Hosseini, S.; Jahandideh, M. The use of artificial neural networks and multiple linear regression to predict rate of medical waste generation. *Waste Manag.* **2009**, *29*, 2874–2879. [CrossRef] [PubMed]

21.　Palmer, A.; Montano, J.J.; Sesé, A. Designing an artificial neural network for forecasting tourism time series. *Tour. Manag.* **2006**, *27*, 781–790. [CrossRef]

22.　Zhang, G.P.; Qi, M. Neural networks forecasting for seasonal and trend time series. *Eur. J. Oper. Res.* **2005**, *160*, 501–514.

23.　Zhang, G.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62. [CrossRef]

24.　Zorpas, A.A.; Lasaridi, K.; Voukkali, I.; Loizia, P.; Inglezakis, V.J. Solid waste from the hospitality industry in Cyprus. *W.I.T. Trans. Ecol. Environ.* **2012**, *166*, 41–49.

25.　Radwan, H.R.; Jones, E.; Minoli, D. Managing solid waste in small hotels. *J. Sustain. Tour.* **2010**, *18*, 175–190. [CrossRef]

26.　Maclaren, V.W.; Yu, C.C. Solid waste recycling behavior of industrial-commercial-institutional establishments. *Growth Chang.* **1997**, *28*, 93–110. [CrossRef]

27.　Azadi, S.; Karimi-Jashni, A. Verifying the performance of artificial neural network and multiple linear regression in predicting the mean seasonal municipal solid waste generation rate: A case study of Fars province, Iran. *Waste Manag.* **2016**, *48*, 14–23. [CrossRef] [PubMed]

28.　Dutta, M.; Ghosh, P.; Basu, J.K. Application of artificial neural network for the decolorization of direct blue 86 by using microwave assisted activated carbon. *J. Taiwan Inst. Chem. Eng.* **2012**, *43*, 879–888. [CrossRef]

29.　Azarmi, S.L.; Alipour, H.; Oladipo, A.A. Using artificial neural network and desirability function to predict waste generation rates in small and large hotels during peak and lean seasons. In Proceedings of the 7th Advances in Hospitality & Tourism Marketing & Management (AHTMM 2017) Conference, Famagusta, Cyprus, 10–15 July 2017; pp. 539–547.

30. Oladipo, A.S.; Ajayi, O.A.; Oladipo, A.A.; Azarmi, S.L.; Nurudeen, Y.; Atta, A.Y.; Ogunyemi, S.S. Magnetic recyclable eggshell-based mesoporous catalyst for biodiesel production from crude neem oil: Process optimization by central composite design and artificial neural network. *C. R. Chim.* **2018**, *21*, 684–695. [CrossRef]

31. Aslani, M.A.A.; Celik, F.; Yusan, S.; Aslani, C.K. Assessment of the adsorption of thorium onto styrene–divinylbenzene-based resin: Optimization using central composite design and thermodynamic parameters. *Process. Saf. Environ. Prot. J.* **2017**, *109*, 192–202. [CrossRef]

32. Teixeira, J.P.; Fernandes, P.O. Tourism time series forecast with artificial neural networks. *Tékhne-Rev. Appl. Manag. Stud.* **2014**, *12*, 26–36. [CrossRef]

33. Oladipo, A.A.; Gazi, M. Targeted boron removal from highly-saline and boron-spiked seawater using magnetic nanobeads: Chemometric optimisation and modelling studies. *Chem. Eng. Res. Des.* **2017**, *121*, 329–338. [CrossRef]