Article

Profiling of Secondary Metabolites of Optimized Ripe Ajwa Date Pulp (*Phoenix dactylifera* L.) Using Response Surface Methodology and Artificial Neural Network

Fanar Alshammari [1], Md Badrul Alam [1,2], Marufa Naznin [3], Ahsan Javed [1], Sunghwan Kim [3,4,*] and Sang-Han Lee [1,2,*]

1 Department of Food Science and Biotechnology, Graduate School, Kyungpook National University, Daegu 41566, Republic of Korea
2 Food and Bio-Industry Research Institute, Inner Beauty/Antiaging Center, Kyungpook National University, Daegu 41566, Republic of Korea
3 Department of Chemistry, Kyungpook National University, Daegu 41566, Republic of Korea
4 Mass Spectrometry Converging Research Center and Green-Nano Materials Research Center, Daegu 41566, Republic of Korea
* Correspondence: sunghwank@knu.ac.kr (S.K.); sang@knu.ac.kr (S.-H.L.); Tel.: +82-053-950-7754 (S.-H.L.)

Abstract: The date palm (*Phoenix dactylifera* L.) is a popular edible fruit consumed all over the world and thought to cure several chronic diseases and afflictions. The profiling of the secondary metabolites of optimized ripe Ajwa date pulp (RADP) extracts is scarce. The aim of this study was to optimize the heat extraction (HE) of ripe Ajwa date pulp using response surface methodology (RSM) and artificial neural network (ANN) modeling to increase its polyphenolic content and antioxidant activity. A central composite design was used to optimize HE to achieve the maximum polyphenolic compounds and antioxidant activity of target responses as a function of ethanol concentration, extraction time, and extraction temperature. From RSM estimates, 75.00% ethanol and 3.7 h (extraction time), and 67 °C (extraction temperature) were the optimum conditions for generating total phenolic content (4.49 ± 1.02 mgGAE/g), total flavonoid content (3.31 ± 0.65 mgCAE/g), 2,2-diphenyl-1-picrylhydrazyl (11.10 ± 0.78 % of inhibition), and cupric-reducing antioxidant capacity (1.43 μM ascorbic acid equivalent). The good performance of the ANN was validated using statistical metrics. Seventy-one secondary metabolites, including thirteen new bioactive chemicals (hebitol II, 1,2-di-(syringoyl)-hexoside, naringin dihydrochalcone, erythron-guaiacylglycerol-β-syringaresinol ether hexoside, erythron-1-(4′-*O*-hexoside-3,5-dimethoxyphenyl)-2-syrngaresinoxyl-propane-1,3-diol, 2-deoxy-2,3-dehydro-N-acetyl-neuraminic acid, linustatin and 1-deoxynojirimycin galactoside), were detected using high-resolution mass spectroscopy. The results revealed a significant concentration of phytoconstituents, making it an excellent contender for the pharmaceutical and food industries.

Keywords: Ajwa dates; antioxidant; artificial neural network; polyphenolics; response surface methodology

check for updates

1. Introduction

Extraction is the first and most significant step in recovering and purifying bioactive chemicals from plant sources, and it is often accomplished using maceration, distillation, or Soxhlet reflux extraction, but long extraction times and low extraction effectiveness limit these approaches [1]. For large-scale extraction, the best approach should provide high efficiency with minimal processing time. In general, either empirical or statistical methods can be used to optimize a process [2,3]. The one-factor-at-a-time technique is an empirical system that involves changing one element at a time while maintaining all other variables constant [4]. The main limitation of this strategy is that it neglects the interplay between the factors, and hence, it cannot account for all the effects of a parameter on the response. One

more burden is that it requires numerous trials to finish the study, thereby increasing the time, costs, and reagent and material consumption [5]. To solve this challenge, multivariate statistical approaches were adopted for optimizing analytical procedures. One of the most prominent multivariate strategies used in analytical optimization is the response surface methodology (RSM) [4]. RSM is a collection of statistical and mathematical methodologies for constructing, developing, and adjusting procedures in which many variables impact a desired response, with the purpose of optimizing that response. It can be used to develop, formulate, and create new products, as well as refining the design of current ones. It describes how the independent variables influence the processes, either singly or together. This experimental approach provides a mathematical model that illustrates the chemical or biological procedures in addition to assessing the impact of independent factors [5,6]. However, any nonlinear relationship between the variables may reduce the forecast accuracy of RSM [7]. Artificial neural networks (ANNs) are rapidly being utilized as prediction tools in a variety of disciplines, including engineering, owing to their capacity to use learning algorithms to identify input–output links for nonlinear complex systems [8].

The date palm (*Phoenix dactylifera* L., Arecaceae family) is a popular edible fruit consumed all over the world. The Ajwa date is only cultivated in Medina, Saudi Arabia. It is one of the most expensive and valued cultivars on the market owing to religious and ethnomedical beliefs regarding its health-promoting qualities [9]. The beneficial properties of Ajwa dates are mentioned in the Old Testament, "Hadith", and Islamic literature, and eating these dates is thought to cure several chronic diseases and afflictions [9]. It is regarded to have cardioprotective [10], hepatoprotective [11], nephroprotective [12], and constipation-relieving [13] properties and antioxidant, anti-inflammatory [14], anticancer [15], antifungal, antibacterial, and antiviral activities [16]. The fruit is rich in dietary fiber, minerals, organic acids, and vitamins and has great nutritional and therapeutic value [13]. Carbohydrates constitute more than 70% of the fruit [17]. In addition, it contains abundant bioactive components such as polyphenols, including phenolic acids, flavonoids, and lignans [13].

The optimization of the extraction of various dates from Tunisia, Algeria, Egypt, and other locations and their polyphenolic content as well as antioxidant activities were described in the prior literature; however, no systematic statistical technique was applied [1,18,19]. Additionally, the majority of the optimization of the extraction process was performed solely using RSM methodology, but the illustrious scientists made no attempt to compare the efficacy of predictive modeling with alternative, more effective methods such as ANN. Furthermore, to the best of our knowledge, this is the first study to use heat extraction (HE) with RSM and ANN to enhance the polyphenolic components and antioxidant activity of ripe Ajwa date pulp (RADP) extracts. The aim was to use the RSM central composite design (CCD) tool to investigate and optimize extraction parameters such as extraction temperature, time, and ethanol concentration to acquire the maximum polyphenolic content and antioxidant potentiality from RADP. We argue that the projected values generated by the RSM-CCD approach correspond to the actual results and that this statistical tool is an effective model to optimize RADP polyphenolic compound extraction and antioxidant activity. The estimating capabilities and modeling effectiveness of the RSM and ANN models were also statistically examined. Additionally, we have also profiled the secondary metabolites of RADP using high-resolution mass spectrometry analysis.

2. Results and Discussion

2.1. Fitting of the RSM and ANN Models

Table 1 summarizes the experimental parameters and findings of 20 extraction situations. All response variables were transformed into second-order quadratic polynomial equations to account for the difference in the various replies as functions of the extraction factors. The statistical significance of the fitted second-order quadratic model equations were determined using ANOVA (Table 2). To improve the model fit and prediction, non-significant terms ($p > 0.05$) were removed from the models. The p-values were used to determine the importance of each coefficient. The model terms were considered significant,

very significant, and strikingly significant when the *p*-values were less than 0.05, 0.01, and 0.001, respectively; the model terms were not significant when the *p*-values were greater than 0.05 [20].

Table 1. Central composite design (CCD) for independent variables and corresponding response values (experimental).

Run	Independent Variables			Responses			
	(X_1)	(X_2)	(X_3)	TPC (Y_1)	TFC (Y_2)	DPPH (Y_3)	CUPRAC (Y_4)
1	75	4	70	4.30 ± 0.68	3.16 ± 0.73	10.81 ± 0.72	1.38 ± 0.05
2	25	2	50	3.81 ± 0.83	2.55 ± 0.52	10.25 ± 0.39	1.47 ± 0.02
3	75	2	50	3.90 ± 0.34	2.55 ± 0.25	9.75 ± 0.37	1.16 ± 0.16
4	100	3	60	4.30 ± 0.26	3.44 ± 0.65	10.11 ± 0.24	1.12 ± 0.10
5	75	2	70	3.50 ± 0.41	2.56 ± 0.32	9.95 ± 1.00	1.17 ± 0.07
6	50	3	60	4.53 ± 0.82	3.10 ± 0.21	10.45 ± 0.10	1.40 ± 0.08
7	0	3	60	3.01 ± 1.02	2.52 ± 0.25	9.57 ± 0.56	0.91 ± 0.06
8	75	4	50	4.31 ± 0.49	2.95 ± 0.45	9.88 ± 0.26	0.88 ± 0.04
9	50	3	60	4.30 ± 0.42	3.15 ± 0.54	10.40 ± 0.95	1.46 ± 0.08
10	50	3	60	4.51 ± 0.23	3.09 ± 0.98	10.42 ± 0.35	1.43 ± 0.04
11	50	3	80	3.78 ± 0.62	2.26 ± 0.10	9.50 ± 0.26	1.11 ± 0.02
12	50	3	60	4.46 ± 0.06	3.01 ± 1.02	10.48 ± 0.33	1.45 ± 0.06
13	50	3	60	4.48 ± 0.04	3.12 ± 0.56	10.41 ± 0.53	1.43 ± 0.16
14	25	4	70	3.15 ± 0.08	2.16 ± 0.46	9.60 ± 0.35	1.00 ± 0.13
15	50	3	60	4.51 ± 0.06	2.95 ± 0.29	10.20 ± 0.15	1.45 ± 0.19
16	50	3	40	3.85 ± 0.24	2.48 ± 0.37	9.60 ± 0.60	0.92 ± 0.15
17	25	2	70	3.25 ± 0.68	2.15 ± 0.19	9.36 ± 0.72	1.12 ± 0.07
18	50	1	60	2.87 ± 0.68	2.03 ± 0.49	9.80 ± 0.39	1.32 ± 0.10
19	50	5	60	2.90 ± 0.64	2.50 ± 1.05	9.90 ± 0.72	0.86 ± 0.02
20	25	4	50	2.96 ± 0.80	2.38 ± 1.06	9.53 ± 0.39	0.75 ± 0.08

X_1: ethanol concentration (%); X_2: time (h); X_3: temperature (°C). TPC: total phenolic content (mg gallic acid equivalent/g dry weight extract); TFC: total flavonoid content (mg catechin equivalent/g dry weight extract); DPPH: DPPH radical scavenging activity (% inhibition); CUPRAC: cupric-reducing antioxidant capacity (µM ascorbic acid equivalent).

The statistical test known as the F-value, which compares the source mean square to the residual square, is frequently employed in conjunction with the *p*-value. The better the F-value, the more significant the model [21]. Table 2 outlines the model F-value range, which runs from 34.85 to 154.89. Adjusted R^2 adjusts for the number of terms in the model, whereas projected R^2 measures the variation in new data explained by the model. The various R^2 values represent the degree of interpretation concerning the mean that the model can explain. For a fair level of agreement, the maximum expected difference between adjusted R^2 and predicted R^2 is 0.2; otherwise, it may indicate a problem with the model or the experimental data used to create the model [22]. The model is more accurate if adjusted R^2 and projected R^2 values are near 1. The constructed regression models, therefore, have a high level of statistical significance, as shown by their R^2 values (0.9691–0.9929), adjusted R^2 value (0.9413–0.9865), and predicted R^2 value (0.8432–0.9580). It is important to note that R^2 is ineffective for establishing the effectiveness or validity of a nonlinear model [23]. Contrarily, the most widely used metrics among the numerous model selection techniques are the predicted residual sum of squares (PRESS), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC) [24]. The PRESS measures a model's propensity to predict; hence, the lower the PRESS score, the greater the model's prediction [25]. The PRESS value range in the current study is between 0.0464 and 0.5369, which supports the claim. Furthermore, it is generally known that Schwarz's Bayesian information criterion (BIC) and Akaike information criterion (AIC) are both penalized-likelihood information criteria. In contrast to BIC, which is an estimate of the posterior probability of a model being authentic in a specific Bayesian setup, AIC is an estimate of

a constant plus the distance between the fitted likelihood function of the model and the unknown actual likelihood function of the data [26]. The lower AICc and BIC model is typically considered the "best" model. AICc and BIC values, in this case, vary from -55.62 to -4.85 and -70.11 to -19.33, respectively, and indicate that the intended responses can be predicted with great accuracy. Furthermore, the appropriate precision value is an indicator of the signal-to-noise ratio. It is preferable to have a ratio >4 [27]. Here, the ratio was between 18.9896 and 34.9882, suggesting a sufficient signal, indicating that the model is suitable for this procedure. The coefficient of variation (CV) is a measure of a model's reproducibility; in general, if CV < 10%, the model is reproducible [28]. In this study the CV values range from 0.9981 to 3.44. The modified R^2 ($R^2 \geq 0.80$) was well within acceptable limits in this study, showing that the experimental data fits second-order polynomial equations satisfactorily [29]. To demonstrate the interactions between the independent variables, 3D surfaces and contour plots were constructed using multiple linear regression equations. The main and cross-product effects of the independent variables on the response variables are more easily understood from these 3D charts (Figure 1A–D).

There is growing body of evidence that ANN modeling is superior and more sophisticated than RSM and that ANNs are an emerging viable alternative to the RSM system for complicated nonlinear multivariate modeling. In terms of fitting experimental responses, prediction, and modeling of biological processes, ANNs are more precise than RSM [30,31]. The experimental values were subjected to ANN modeling for further model verification. The predicted values obtained after training the ANN model are given in supplementary data Table S1.

Table 2. ANOVA for quadratic model (transfer function: none).

Source	RC	SS	DF	MS	F-Value	*p*-Value	
ANOVA for Quadratic Model for TPC							
Model		7.18	9	0.7975	81.83	<0.0001	Significant
Intercept	4.45						
Linear terms							
X_1	0.3400	1.85	1	1.85	189.79	<0.0001	Significant
X_2	0.0200	0.0064	1	0.0064	0.6567	0.4366	Non-significant
X_3	-0.0575	0.0529	1	0.0529	5.43	0.0421	Significant
Interaction terms							
X_1X_2	0.2700	0.5832	1	0.5832	59.84	<0.0001	Significant
X_1X_3	-0.0050	0.0002	1	0.0002	0.0205	0.8889	Non-significant
X_2X_3	0.1425	0.1625	1	0.1625	16.67	0.0022	Significant
Quadratic terms							
X_1^2	-0.2089	1.10	1	1.10	112.55	<0.0001	Significant
X_2^2	-0.4001	4.03	1	4.03	413.02	<0.0001	Significant
X_3^2	-0.1676	0.7064	1	0.7064	72.48	<0.0001	Significant
Lack of Fit		0.0617	5	0.0123	1.73	0.2819	Non-significant
Pure error		0.0358	5	0.0072			
R^2							0.9866
Adjusted R^2							0.9745
Predicted R^2							0.9262
Adeq Precision							23.5003
C.V. %							2.58
PRESS							0.5369
BIC							-19.77
AICc							-5.28

Table 2. *Cont.*

ANOVA for quadratic model for TFC							
Model		3.11	9	0.3454	39.85	<0.0001	significant
Intercept	3.05						
Linear terms							
X_1	0.2388	0.9120	1	0.9120	105.24	<0.0001	Significant
X_2	0.1113	0.1980	1	0.1980	22.85	0.0007	Significant
X_3	−0.0525	0.0441	1	0.0441	5.09	0.0477	Significant
Interaction terms							
X_1X_2	0.1450	0.1682	1	0.1682	19.41	0.0013	Significant
X_1X_3	0.1050	0.0882	1	0.0882	10.18	0.0097	Significant
X_2X_3	0.0475	0.0180	1	0.0180	2.08	0.1795	Non-significant
Quadratic terms							
X_1^2	−0.0328	0.0271	1	0.0271	3.13	0.1073	Non-significant
X_2^2	−0.2116	1.13	1	1.13	129.89	<0.0001	Significant
X_3^2	−0.1853	0.8637	1	0.8637	99.66	<0.0001	Significant
Lack of Fit		0.0585	5	0.0117	2.07	0.2213	Non-significant
Pure error		0.0282	5	0.0056			
R^2							0.9729
Adjusted R^2							0.9485
Predicted R^2							0.8432
Adeq Precision							21.4961
C.V. %							3.44
PRESS							0.5009
BIC							−22.11
AICc							−7.63
ANOVA for quadratic model for DPPH							
Model		3.12	9	0.3470	34.85	<0.0001	Significant
Intercept	10.39						
Linear terms							
X_1	0.1706	0.4658	1	0.4658	46.77	<0.0001	Significant
X_2	0.0444	0.0315	1	0.0315	3.16	0.1057	Non-significant
X_3	0.0069	0.0008	1	0.0008	0.0759	0.7885	Non-significant
Interaction terms							
X_1X_2	0.1837	0.2701	1	0.2701	27.12	0.0004	Significant
X_1X_3	0.2437	0.4753	1	0.4753	47.73	<0.0001	Significant
X_2X_3	0.2112	0.3570	1	0.3570	35.85	0.0001	Significant
Quadratic terms							
X_1^2	−0.1399	0.4920	1	0.4920	49.40	<0.0001	Significant
X_2^2	−0.1374	0.4746	1	0.4746	47.65	<0.0001	Significant
X_3^2	−0.2124	1.13	1	1.13	113.88	<0.0001	Significant
Lack of Fit		0.0505	5	0.0101	1.03	0.4887	Non-significant
Pure error		0.0491	5	0.0098			
R^2							0.9691
Adjusted R^2							0.9413
Predicted R^2							0.8504
Adeq Precision							18.9896
C.V. %							0.9981
PRESS							0.4822
BIC							−19.33
AICc							−4.85

Table 2. *Cont.*

ANOVA for quadratic model for CUPRAC							
Model		1.10	9	0.1218	154.89	<0.0001	Significant
Intercept	1.43						
Linear terms							
X_1	0.0419	0.0281	1	0.0281	35.68	0.0001	Significant
X_2	−0.1144	0.2093	1	0.2093	266.19	<0.0001	Significant
X_3	0.0481	0.0371	1	0.0371	47.13	<0.0001	Significant
Interaction terms							
X_1X_2	0.0963	0.0741	1	0.0741	94.25	<0.0001	Significant
X_1X_3	0.0762	0.0465	1	0.0465	59.15	<0.0001	Significant
X_2X_3	0.1363	0.1485	1	0.1485	188.87	<0.0001	Significant
Quadratic terms							
X_1^2	−0.1074	0.2899	1	0.2899	368.74	<0.0001	Significant
X_2^2	−0.0886	0.1975	1	0.1975	251.22	<0.0001	Significant
X_3^2	−0.1086	0.2967	1	0.2967	377.37	<0.0001	Significant
Lack of Fit		0.0055	5	0.0011	2.37	0.1827	Non-significant
Pure error		0.0023	5	0.0005			
R^2							0.9929
Adjusted R^2							0.9865
Predicted R^2							0.9580
Adeq Precision							34.9882
C.V. %							2.36
PRESS							0.0464
BIC							−70.11
AICc							−55.62

RC: regression coefficient; SS: sum of squares; MS: mean square.

The ANN model predicted values that were reasonably close to the observed values, indicating the appropriateness of our model. The nonlinear correlations between the extraction parameters (X_1, X_2, and X_3) and response variables (Y_1, Y_2, Y_3, and Y_4) were predicted using the ANN model. By comparing the error between network training and testing, the number of neurons in the hidden layer was modified using the hit-and-trial strategy. During development, the input layer was not triggered using the transfer function, but the output and hidden layers were activated using pure line (purelin) and tangent sigmoid transfer (tansig) functions in MATLAB. In the experiment, the lowest feasible error between training and testing was evaluated, and the minimum number of epochs to prevent model overfitting was considered; the results were consistent with the findings of previous works [32]. Training the network with the Levenberg–Marquardt technique yielded the best validation performance for all dependent variables: TPC (Y_1), TFC (Y_2), DPPH (Y_3), and CUPRAC (Y_4) (supplementary data Figures S1–S4).

2.2. Comparison of the Prediction Abilities of the RSM and ANN Models

The predictive and estimate capabilities of RSM and ANN models were compared. Comparative similarity plots were used to examine the projected values of the four target responses (Y_1, Y_2, Y_3, and Y_4) obtained using the ANN model. The results indicate that compared to the RSM model, the ANN model was more correct, precise, and capable of assessment in terms of fitting the experimental data to all target responses (supplementary data Table S1). The RSM model had a greater degree of divergence between the projected and actual data than the ANN model, whereas the ANN model had stable residuals with low variation.
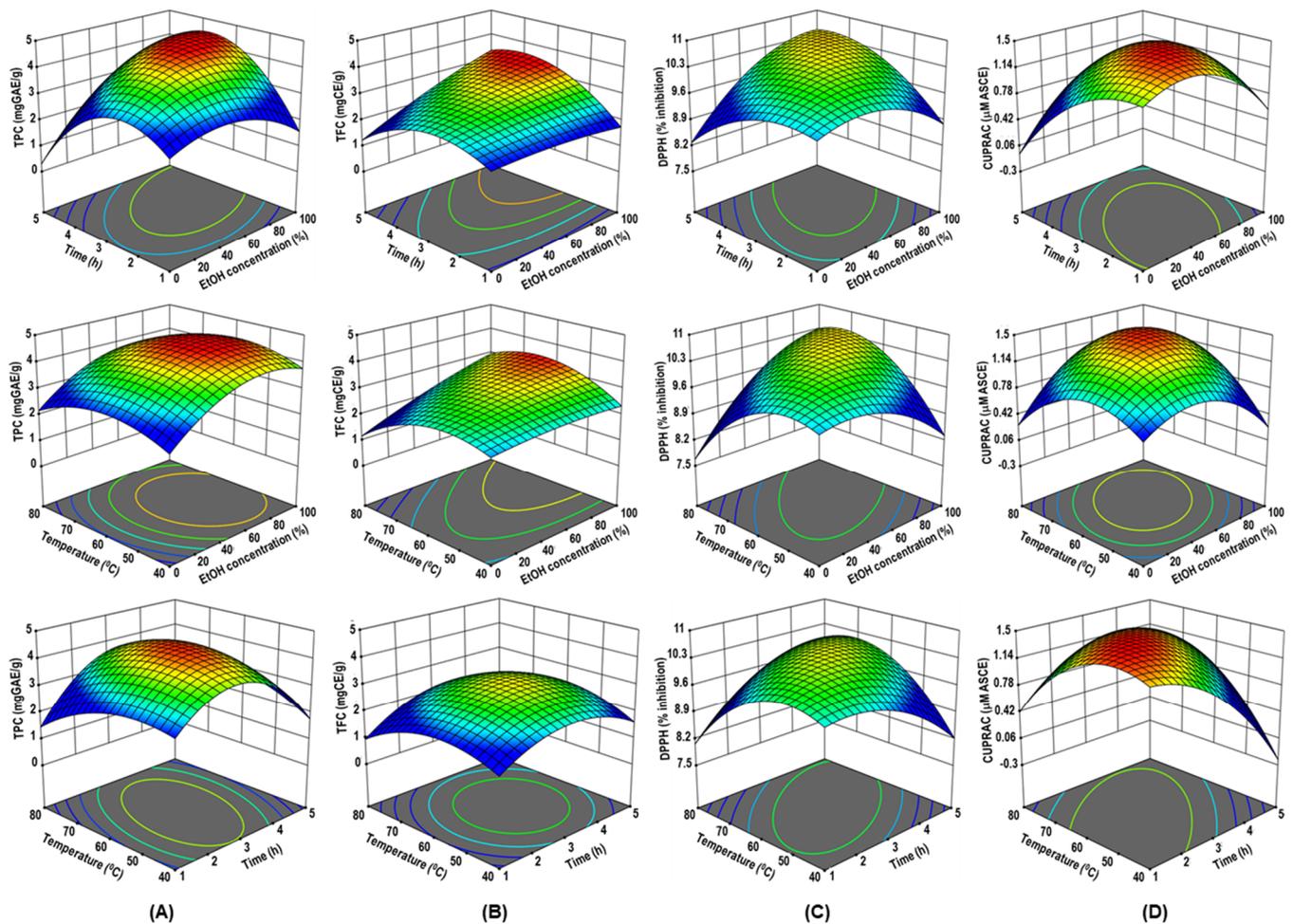
Figure 1. The three-dimensional (3D) response surface plots of RADP extraction on ethanol concentration, time, and temperature for TPC (**A**), TFC (**B**), DPPH radical scavenging activity (**C**), and CUPRAC (**D**) as a function of significant interaction factors for RSM.

The values of the correlation coefficient (R^2), root-mean-square error (RMSE), absolute average deviation (AAD), and standard error of prediction (SEP) were calculated to compare the RSM and ANN models (Table 3). Better modeling requires lower RMSE, AAD, and SEP, and greater R^2. The calculated R^2 values of the trained ANN model were higher than those of the RSM model, indicating the superiority of the ANN model in terms of predicting all four dependent variables. The AAD is a measure of the projected data's divergence from the actual data, whereas the RMSE numbers demonstrate the model's absolute fit. The ANN model performed better than the RSM model, as seen by the former's lower AAD and RMSE values. In addition, the ANN model had low SEP values, in the range of 0.0449–0.2903. The ANN model has a substantially stronger predictive capacity than the RSM model because the former can approximate nonlinear systems, but the latter is only successful if the system is constrained to second-order polynomial regression. The ANN model is likewise unaffected by experimental design, and it is faster at calculating many responses in a single run than the RSM model, which requires multiple runs for multi-response optimization in a typical experimental design [33]. Unlike the RSM model, building an ANN model requires many iterative calculations and long computation time. Moreover, ANNs are also useful as they are flexible for the addition of new experimental data for model generation [33].

Table 3. Comparison of the prediction abilities of the RSM and ANN models.

Parameters	TPC		TFC		DPPH		CUPRAC	
	RSM	ANN	RSM	ANN	RSM	ANN	RSM	ANN
R^2	0.9866	0.9899	0.9729	0.9859	0.9691	0.9828	0.9929	0.9954
RMSE	0.8053	0.4475	0.1663	0.1134	0.2586	0.1635	0.1063	0.0903
AAD (%)	4.079	2.052	6.011	3.753	0.8815	0.7011	6.764	3.738
SEP (%)	0.2225	0.1518	0.4591	0.2903	0.0529	0.0449	0.3710	0.2728

2.3. Effect of HE Parameters on the TPC and TFC

The TPC and TFC in the RADP extracts varied from 2.87 to 4.53 mgGAE/g and 2.03 to 3.44 mgCAE/g, respectively (Table 1). The linear effect of ethanol concentration (X_1), temperature (X_3), and the quadratic component of (X_1^2), (X_2^2), and (X_3^2) as well as the interaction of (X_1X_3) and (X_2X_3) were significant for TPC. The linear effect of (X_1), (X_2), and (X_3); the quadratic component of (X_2^2) and (X_3^2); and the interaction of (X_1X_2) and (X_1X_3) were significant for TFC (Table 2). The following second-order polynomial equations shown in Equations (1) and (2) demonstrate the relationships among TPC, TFC, and their variables.

$$TPC(Y_1) = 4.45 + 0.3400X_1 + 0.0200X_2 - 0.0575X_3 - 0.2089X_1^2 - 0.4001X_2^2 - 0.1676X_3^2 + 0.2700X_1X_2 \\ -0.0050X_1X_3 + 0.1425X_2X_3 \tag{1}$$

$$TFC(Y_2) = 3.05 + 0.2388X_1 + 0.1113X_2 - 0.0525X_3 - 0.0328X_1^2 - 0.2116X_2^2 - 0.1853X_3^2 + 0.1450X_1X_2 \\ +0.1050X_1X_3 + 0.0475X_2X_3 \tag{2}$$

The lack of fit values for TPC and TFC (F = 1.73 and 2.07, respectively) were non-significant, indicating that the model accurately predicted R^2 = 0.9866 (TPC) and 0.9729 (TFC) and Adj.R^2 = 0.9745 (TPC) and 0.9485 (TFC) (Table 2). The RSM model adequately predicted the effects of parameters on the TPC and TFC of the RADP extract. As shown in Figure 1A,B, when the temperature (X_3) was fixed (60 °C), the maximum TPC and TFC were achieved when the ethanol concentration (X_1) was 50%, and time was 3 h. This could be because a medium concentration of ethanol may make the solvent more polar and dissolve more polyphenols, both polar and moderately polar ones [1]. Experiments in a previous comparative study revealed that the extraction of polyphenols from green tea leaves using a high hydrostatic pressure procedure augmented as the percentage of ethanol in the solvent, which peaked at 50% ethanol and dropped after that [34]. Hence, extraction of polyphenols in hydroalcoholic solution is highly efficient as the polyphenols are highly soluble in these solutions [1]. Furthermore, when ethanol is present at a moderate quantity in water, it can disrupt and break the architecture and structure of phospholipids that make up the lipid bilayer of membranes, affecting the penetrability of plant cells, thereby allowing better extraction and diffusion of the polyphenolic compounds [34].

2.4. Effect of HE Parameters on the In Vitro Antioxidant Capacity (AC)

DPPH radical scavenging activity and CUPRAC analysis revealed a linear significant effect of ethanol concentration (X_1) and the quadratic effect of temperature (X_3) on antioxidant activity. The fitted second-order polynomial equations for DPPH (% inhibition) and CUPRAC (ascorbic acid equivalent µM) are shown in Equations (3) and (4):

$$DPPH(Y_3) = 10.39 + 0.1706X_1 + 0.0444X_2 - 0.0069X_3 - 0.1399X_1^2 - 0.1374X_2^2 - 0.2124X_3^2 + 0.1837X_1X_2 \\ +0.2437X_1X_3 + 0.2112X_2X_3 \tag{3}$$

$$CUPRAC(Y_4) = 1.43 + 0.0419X_1 - 0.1144X_2 + 0.0481X_3 - 0.1074X_1^2 - 0.0886X_2^2 - 0.1086X_3^2 \\ +0.0963X_1X_2 + 0.0762X_1X_3 + 0.1363X_2X_3 \tag{4}$$

The AC values ranged from 9.36% to 10.81% inhibition of DPPH and from 0.75 to 1.47 µM ascorbic acid equivalent (Table 1). The ANOVA results show that the data fitted

the model results for DPPH (R^2 = 0.9691 and Adj. R^2 = 0.9413) and CUPRAC response (R^2 = 0.9929 and Adj. R^2 = 0.9865), and the lack of fit was non-significant (F = 1.03 for DPPH and 2.37 for CUPRAC) (Table 2). From the 3D response surface plots shown in Figure 1C,D, both DPPH radical scavenging activity and CUPRAC activity increase with increasing ethanol concentration to reach the maximum values at 100% and 75% ethanol, respectively. This means that the greater the quantity of the organic solvent, the better the electron and proton donation capacities are. This result is consistent with the prior finding for TFC that 100% ethanol is required for maximum extraction [34]. Increasing the ethanol concentration allows for the extraction of significant polyphenolics in RADP, both in terms of quality and quantity. There is increasing evidence that the quality and quantity of polyphenolic compounds and antioxidant activity differ with the ethanol concentrations [35,36].

2.5. Model Validation

The desirability function optimizes responses for polyphenolic content (TPC and TFC) and antioxidant activity (DPPH and CUPRAC) concurrently. The parameters were forecasted using Derringer's desirability function, allowing a multivariate analysis to discover the ideal level for all responses in a single extraction [37]. In this study, the following conditions were used to achieve the maximal overall desirability D = 0.934 (on a scale of 0 to 1): X_1, 75%; X_2, 3.7 h; and X_3, 67 °C. Figure 2 shows the contour plot as a function of ethanol concentration, extraction time, and temperature at optimum condition. To verify the sufficiency of the model equations, a duplicate experiment was conducted in the optimal conditions predicted by Derringer's desire model. The following results were obtained: TPC = 4.49 ± 1.02 mgGAE/g, TFC = 3.31 ± 0.65 mgCAE/g, % inhibition of DPPH = 11.10% ± 0.78%, and μM ascorbic acid equivalent CUPRAC = 1.43 ± 0.43. As stated in Table 4, the relative standard deviations (RSDs) of TPC, TFC, DPPH, and CUPRAC showed that the predicted values for all groups were very similar to the experimental results. This result is also supported by prior research [29].

Table 4. Experiment data of the validation of predicted values at optimal extraction conditions of RADP.

Response	Exp.	Pred.	Std	RSD (%)
TPC (mgGAE/g)	4.49 ± 1.02	4.53	0.028	0.006
TFC (mgCAE/g)	3.31 ± 0.65	3.30	0.007	0.002
DPPH (% inhibition)	11.10 ± 0.78	10.69	0.290	0.027
CUPRAC (μM ASCE)	1.43 ± 0.43	1.41	0.014	0.010

Optimal condition: ethanol concentration (%), 75 %; time (h), 3.7; temperature (°C), 67. Exp.: experimental value; Pred.: predicted value; Std: standard deviation; RSD: relative standard deviation.

2.6. Identification of Secondary Metabolites in RADP with High-Resolution Mass Spectrometry

Secondary metabolites in the RADP extracts were identified using ESI-MS/MS in the positive and negative ionization modes. As indicated in Table 5, 71 compounds were identified in the negative mode using MS^n data from the mass of the precursor ion, fragments, recognized fragmentation patterns for the given classes of compounds, and neutral mass loss, and from comparison with the existing literature and searches in online databases. Furthermore, the significance of these results was determined by finding the confidence level. Level 3 denotes a tentative candidate, whereas level 2 indicates the probable structure of the identified compound [38].

Table 5. List of possible identified compounds of the optimized extract of ripe Ajwa date pulp (RADP) by electrospray ionization mass spectrometry (ESI-MS)/MS.

	No.	Compound Name	EF	OM $(m/z)^{-/+}$	CM $(m/z)^{-/+}$	MS/MS (Negative Mode)	CE	CL
Phenolic acids and derivatives	1	Coumalic acid [#]	$C_6H_4O_4$	139.0050	139.0031	111.01, 95.01	20	3
	2	Hydroxybenzoylhexose	$C_{13}H_{16}O_8$	299.0876	299.0884	281.06, 237.04, 179.03, 163.06, 137.02	20	2
	3	Coumaroyl hexose	$C_{15}H_{18}O_8$	325.0929	325.0923	163.03, 147.04	10	2
	4	Caffeoylshikimic acid	$C_{16}H_{16}O_8$	335.0776/337.0932	335.0772	179.01, 161.03, 155.03, 137.05	20	2
	5	Caffeic acid hexoside	$C_{15}H_{18}O_9$	341.1100	341.0872	215.03, 179.06, 161.04	20	2
	6	Caffeic acid derivatives	$C_{18}H_{18}O_9$	377.0885	377.0878	341.10, 215.03, 179.06, 161.04	10	2
	7	Dicaffeoyl shikimic acid	$C_{22}H_{26}O_{13}$	497.1297	497.1295	335.01, 178.02, 135.02	20	2
	8	Hebitol II [#]	$C_{21}H_{30}O_{14}$	505.1603	505.1557	341.08,325.09, 179.03, 163.03	30	3
	9	1,2-di-(syringoyl)-hexoside [#]	$C_{24}H_{28}O_{14}$	539.1377/541.1533	539.1401	359.09, 341.08, 197.04, 153.05	30	3
Flavonoids and derivatives	10	Chrysoeriol	$C_{13}H_{16}O_8$	299.0561/301.0717	299.0555	285.03, 255.02, 153.01, 147.04, 135.03, 125.03	20	2
	11	Quercetin	$C_{15}H_{10}O_7$	301.0354/303.0510	301.0348	273.02, 257.03, 229.05, 179.01, 151.01	20	2
	12	Dihydrokaempferol hexoside	$C_{21}H_{22}O_{11}$	449.1089	449.1083	287.04, 269.05, 259.06, 169.01, 151.01	20	2
	13	Isoquercitrin	$C_{21}H_{20}O_{12}$	463.0878	463.0876	301.05, 268.01, 179.02, 151.01	20	2
	14	Isorhamnetin hexoside	$C_{22}H_{22}O_{12}$	477.1035/479.1191	477.1033	315.05, 300.01, 271.02, 255.05, 179.05, 151.02	20	2
	15	Luteolin hexosyl sulfate	$C_{21}H_{20}O_{14}S$	527.0491/529.0647	527.0495	447.05, 285.01, 241.06	20	2
	16	Chrysoeriol hexosyl sulfate	$C_{22}H_{22}O_{14}S$	541.0645/543.0801	541.0652	299.05, 284.05, 241.02	20	2
	17	Isoquercitrin sulfate	$C_{21}H_{20}O_{15}S$	543.0441/545.0597	543.0444	463.05, 301.01, 268.01, 179.02, 151.01	20	2
	18	Apigenin-8-C-(pentosyl) hexoside	$C_{26}H_{28}O_{14}$	563.1655	563.1400	473.01, 443.02, 413.05, 340.08, 311.02	30	2
	19	Naringin dihydrochalcone [#]	$C_{27}H_{34}O_{14}$	581.1863	581.1870	436.13, 274.08, 167.03, 149.06, 133.06,	30	3
	20	Afzelin gallate	$C_{28}H_{24}O_{14}$	583.1093	583.1087	297.05, 285.04, 169.01	20	2
	21	Luteolin rhamnosyl hexoside	$C_{27}H_{30}O_{15}$	593.1507	593.1506	447.09, 285.03, 153.01, 135.04	20	2
	22	Chrysoeriol rhamnosyl hexoside	$C_{28}H_{32}O_{15}$	607.1669	607.1663	461.10, 299.05, 284.03, 153.01, 149.05	20	2
	23	Quercetin rhamnosyl glucoside	$C_{27}H_{30}O_{16}$	609.1459	609.1455	463.08, 447.09, 301.02, 151.04	20	2
	24	Isorhamnetin rhamnosyl glucoside	$C_{28}H_{32}O_{16}$	623.1617/625.1773	623.1612	477.10, 315.05, 299.05, 165.05	20	2
	25	Quercetin diglucoside	$C_{27}H_{30}O_{17}$	625.1410/627.1566	625.1404	463.08, 301.01	20	2
	26	Isorhamnetin diglucoside	$C_{28}H_{32}O_{17}$	639.1563/641.0612	639.1561	447.01, 315.01	20	2
	27	Luteolin dihexosyl sulfate	$C_{27}H_{30}O_{19}S$	689.1029	689.1023	519.11, 489.10, 471.09, 399.07, 369.06, 339.05	20	2
	28	Luteolin rhamnosyl dihexoside	$C_{33}H_{40}O_{20}$	755.2046	755.2034	709.16, 593.10, 575.05, 285.01	20	2
	29	Chrysoeriol rhamnosyl dihexoside	$C_{34}H_{42}O_{20}$	769.2189	769.2191	623.16, 461.10, 299.05, 284.03, 153.02	20	2

Table 5. *Cont.*

	No.	Compound Name	EF	OM (*m/z*)$^{-/+}$	CM (*m/z*)$^{-/+}$	MS/MS (Negative Mode)	CE	CL
Lignans	30	Erythro-guaiacylglycerol-β-syringaresinol ether hexoside [#]	$C_{38}H_{48}O_{17}$	775.2821	775.2813	613.22, 417.14, 181.05, 151.03	30	3
	31	Erythro-1-(4″-glucoside-3,5-dimethyoxyphenyl)-2-syringaresinoxyl-propane-1,3-diol [#]	$C_{39}H_{50}O_{18}$	805.2926	805.2919	643.23, 417.14, 181.05, 151.03	30	3
Sialic acids	32	2-Deoxy-2,3-dehydro-N-acetylneuraminic acid [#]	$C_{11}H_{17}NO_8$	290.0879	290.0876	230.06, 200.05, 171.01, 128.07	20	3
	33	N-acetyl-α-neuraminic acid [#]	$C_{11}H_{19}NO_9$	308.0987/310.1143	308.0987	290.09, 219.06, 200.05, 146.08, 128.07	20	3
	34	6′-Sialyllactose [#]	$C_{23}H_{39}NO_{19}$	632.2039	632.2044	290.09, 200.05, 128.07	30	3
Amino acids	35	L-proline	$C_5H_9NO_2$	114.0570	114.0561	70.06	10	2
	36	Pyroglutamic acid	$C_5H_7NO_3$	128.0360/130.0416	128.0353	82.3, 71.9	10	2
	37	L-aspartic acid	$C_4H_7NO_4$	132.0329	132.0302	116.03, 88.04	10	2
	38	Allysine [#]	$C_6H_{11}NO_3$	144.0682	144.0666	127.04, 126.05, 100.07	20	3
Sugar molecules	39	Ribonic acid	$C_5H_{10}O_6$	165.0421	165.0418	149.04, 105.01, 87.00, 75.00	10	2
	40	L-Galactose	$C_6H_{12}O_6$	179.0572	179.0561	161.04, 143.03, 113.02, 101.02,	10	2
	41	Mannitol	$C_6H_{14}O_6$	181.0725	181.0718	165.01, 147.03, 129.05, 111.00	20	2
	42	Gluconic acid	$C_6H_{12}O_7$	195.0522	195.0504	177.05, 159.02, 129.05, 98.90	10	2
	43	Sedoheptulose	$C_7H_{14}O_7$	209.0679	209.0680	191.05, 179.05, 149.04,	20	2
	44	Hexose derivative	$C_{12}H_{19}O_{10}$	323.0977	323.0978	179.05, 161.04, 143.03, 113.02, 101.02	10	2
	45	Maltitol	$C_{12}H_{24}O_{11}$	343.1255	343.1240	283.10, 265.09, 179.05, 161.04, 143.03	20	2
	46	Unsaturated digalacturonate [#]	$C_{12}H_{16}O_{12}$	351.0574/353.0730	351.0569	291.07, 273.06, 175.02, 131.03	20	3
	47	Xylosmaloside [#]	$C_{18}H_{20}O_9$	379.1027/381.1183	379.1029	343.08, 217.05, 179.05, 161.04	20	3
Carboxylic acids	48	Fumaric acid	$C_4H_4O_4$	115.0050	115.0037	71.01	10	2
	49	Glutaconic acid	$C_5H_6O_4$	129.0203	129.0203	111.00, 85.02	10	2
	50	Glutaric acid	$C_5H_8O_4$	131.0355	131.0350	113.00, 87.02	10	2
	51	3-Methylglutaconic acid	$C_6H_8O_4$	143.0367	143.0361	99.03	20	2
	52	Methyl glutaric acid	$C_6H_{10}O_4$	145.0521	145.0506	127.02, 101.02	10	2
	53	2-Hydroxyglutaric acid	$C_5H_8O_5$	147.0301	147.0299	129.01, 99.03	10	2
	54	Hydroxymethyl glutaric acid	$C_6H_{10}O_5$	161.0459	161.0455	143.03, 117.05, 99.04	10	2
	55	Citric acid	$C_6H_8O_7$	191.0197/193.0353	191.0197	173.00, 129.01, 111.00	20	2

Table 5. *Cont.*

	No.	Compound Name	EF	OM (*m/z*)$^{-/+}$	CM (*m/z*)$^{-/+}$	MS/MS (Negative Mode)	CE	CL
Fatty acids	56	Palmitic acid	$C_{16}H_{32}O_2$	255.2330	255.2330	237.23, 211.24, 197.22	20	2
	57	Linolenic acid	$C_{18}H_{30}O_2$	277.2165	277.2169	259.20, 233.22, 205.21, 179.25, 165.23	10	2
	58	α-Linoleic acid	$C_{18}H_{32}O_2$	279.2331	279.2330	261.22	10	2
	59	Oleic acid	$C_{18}H_{34}O_2$	281.2487	281.2486	263.25, 181.21, 127.25	10	2
	60	Hydroxy octadecatrienoic acid [#]	$C_{18}H_{30}O_3$	293.2120	293.0216	275.22	20	3
	61	Hydroxy octadecadienoic acid	$C_{18}H_{32}O_3$	295.2276	295.2273	277.23	20	2
	62	Hydroxy octadecenoic acid	$C_{18}H_{34}O_3$	297.2433	297.2429	279.23	20	2
	63	Dihydroxy octadecadienoic acid	$C_{18}H_{32}O_4$	311.2246/313.2402	311.2239	293.22, 275.23	20	2
	64	Dihydroxy octadecenoic acid	$C_{18}H_{34}O_4$	313.2381/315.2537	313.2378	295.23, 277.25, 183.32	20	2
	65	Dihydroxy octadecanoic acid	$C_{18}H_{36}O_4$	315.2538/317.2694	315.2535	297.23, 279.25	20	2
	66	Trihydroxy octadecadienoic acid	$C_{18}H_{32}O_5$	327.2176	327.2171	309.23, 291.25, 273.23	20	2
	67	Trihydroxy octadecenoic acid	$C_{18}H_{34}O_5$	329.2346/331.2502	329.2333	311.25, 293.26, 275.23	20	2
Others	68	Linustatin [#]	$C_{16}H_{27}NO_{11}$	408.151/410.1666	408.1506	318.11, 246.09, 228.08, 214.07	20	3
	69	Norbellidifodin [#]	$C_{13}H_8O_6$	259.024/261.0396	259.0248	241.01, 215.12, 187.05, 171.03	30	3
	70	1-Deoxynojirimycin hexoside [#]	$C_{12}H_{23}NO_9$	324.1293	324.1295	161.04, 144.06, 143.03, 113.02	30	3
	71	Oxycoumarin-4-acetic acid methyl ester hexoside [#]	$C_{18}H_{20}O_{10}$	395.0962	395.0978	233.04, 205.05, 161.02, 133.02	30	3

EF: elemental formula; OM: observed mass; CM: calculated mass; CE: Collision Energy (eV); CL: confidence level; $^{-/+}$: negative mode/positive mode. [#] First time identification in Ajwa date fruits.
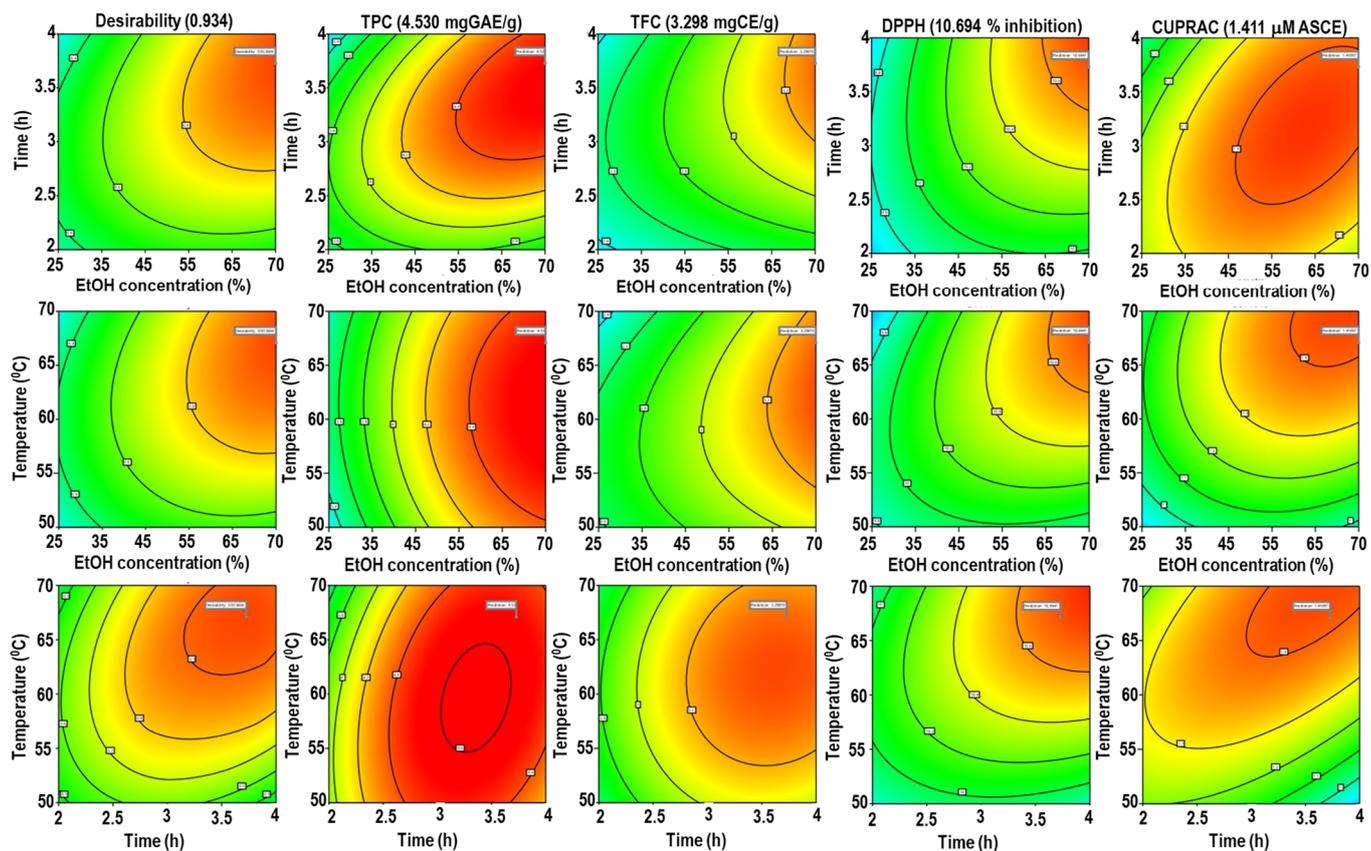
Figure 2. Desirability surface plot as a function of ethanol concentration, extraction time, and temperature for TPC, TFC, DPPH, and CUPRAC.

2.6.1. Phenolic Acids

A phenolic acid may lose its methyl (15 Da), hydroxyl (18 Da), or carboxyl (44 Da) moiety to form a specific fragment ion. The fragmentation of a phenolic acid glycoside begins with the cleavage of the glycosidic link to yield the m/z of the phenolic acid and the corresponding loss of the sugar molecule (neutral mass loss of 162 Da). Thus, compounds **2**–**7** were tentatively identified as hydroxybenzoylhexose, coumaroyl hexose, caffeoylshikimic acid, caffeic acid hexoside, caffeic acid derivatives, and dicaffeoyl shikimic acid, respectively [17,39,40]. Among these, coumaroyl hexose (IC_{50} value: 37.7 μM), caffeoylshikimic acid (IC_{50} value: 2.9 μM), caffeic acid hexoside (IC_{50} value: 0.55 μM), and dicaffeoyl shikimic acid (IC_{50} value: 0.38 μM) were found to have DPPH-scavenging activity [41–43]. Additionally, coumaroyl glucoside severely inhibits glycogen phosphorylase, a crucial target for producing antihyperglycemic medicines [44]. The formation of advanced glycation end products (AGEs) and the activity of the angiotensin-converting and acetylcholinesterase enzymes are all inhibited by caffeine acid glucoside [45]. In addition, compound **1** yielded a precursor ion [M–H]⁻ at m/z 139.0050, generated characteristic ions at m/z 111.01 and 95.01 by loss of neutral molecules of CO and COO, respectively, and was provisionally identified as coumalic acid; this is the first time that this compound was identified in Ajwa dates (supplementary data Figure S5A). Furthermore, compound **8** produced a monoisotopic ion [M–H]⁻ at m/z 505.1603 and yielded fragment ions at m/z 325.09 [M-H-181 Da], m/z 341.08 [M-H-165 Da], and MS^2 ion at m/z 341.08. Further loss of glucosyl moiety (162 Da) led to the generation of an ion at m/z 179.03, and the following loss of water molecules yielded an ion at m/z 161.02, thereby confirming the presence of caffeoyl hexoside. Based on the fragmentation behavior, compound **8** was tentatively confirmed as hebitol II, which has been identified in RADP for the first time (Figure 3A). According to Wang et al., hebitol II prevented the growth of *Plasmodium falciparum* 3D7,

a malarial parasite line [46]. Moreover, compound **9** was tentatively identified as 1,2-di-(syringoyl)-hexoside with molecular formula $C_{24}H_{28}O_{14}$, which yielded a deprotonated ion at *m/z* 539.1377 and generated the following fragment ions: *m/z* at 359.09 ([M-H-syringoyl moiety]), 341.08 ([M-H-syringoyl moiety-H_2O]), 197.04 (syringic acid), and 153.05 because of the loss of water molecule from ion *m/z* 197.04 (Figure 3B). Compound **9** too has been identified for the first time in RADP.
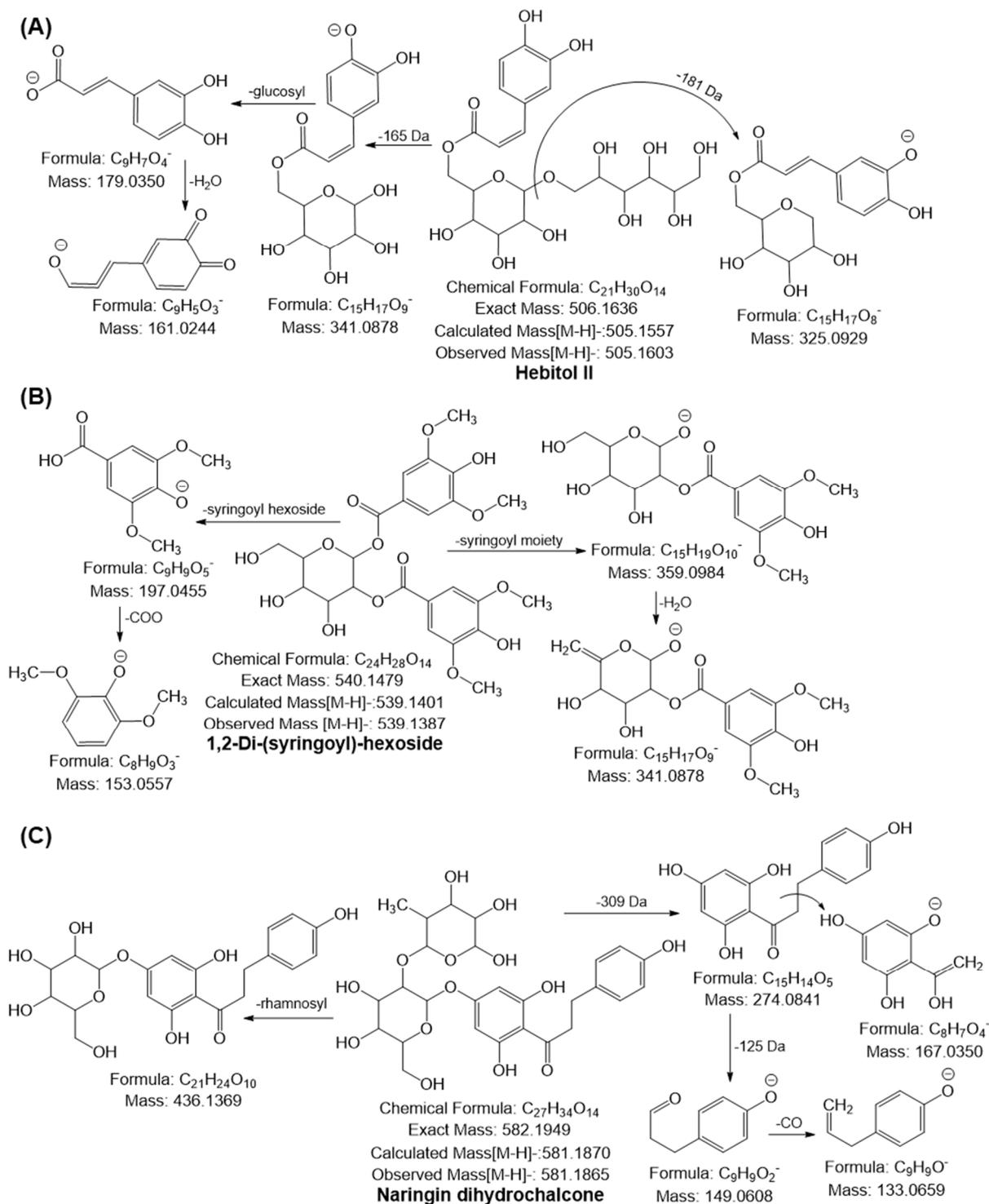
Figure 3. Possible mass fragmentation behavior of identified compounds in RADP. (**A**) Hebitol II, (**B**) 1,2-di-(syringoyl)-hexoside, and (**C**) naringin dihydrochalcone.

2.6.2. Flavonoids

A previous review demonstrated that each subgroup of flavonoids exhibits a different fragmentation behavior in MS^2 analysis. The cleavage of the C-ring bonds (retro-Diels-Alder, i.e., RDA mechanism) produces ions with the A- or B-ring and some part of the C-ring, which is the most common fragmentation of flavonoids [47]. Notable losses of small neutral molecules, such as CO (28 Da), C_2H_2O (42 Da), COO (44 Da), 2CO (56 Da), CO + COO (72 Da), and 3CO (84 Da), may also occur [48]. Based on a comparison of the fragmentation patterns with those previously published in the literature, compounds **10**, **11**, and **20** were identified as chrysoeriol, quercetin, and afzelin gallate, respectively [1,17,40]. Along with their antioxidant properties, quercetin, and chrysoeriol have shown potent anticancer, anti-inflammatory, antidiabetic, and antihyperlipidemic effects [49,50]. Flavonoids are frequently glycosylated. The typical fragmentation of *O*-glycosides produces neutral species corresponding to sugar units (hexoses, 162 Da; deoxyhexoses, 146 Da; pentoses, 132 Da) and an aglycone ion. Conversely, C-glucosides produce a sequence of fragments because of the cleavage of the C-C bonds with the sugar moiety; examples of such fragments are $[M-H-60]^-$, $[M-H-90]^-$, and $[M-H-120]^-$, which serve as the hallmark diagnostic ions of glycone [51,52]. Compounds **12–18** and **21–29** were identified as dihydrokaempferol hexoside (aromadendrin hexoside), isoquercitrin, isorhamnetin hexoside, luteolin hexosyl sulfate, chrysoeriol hexosyl sulfate, isoquercitrin sulfate, apigenin-8-C-(pentosyl)-hexoside (vitexin rhamnoside), luteolin rhamnosyl hexoside, chrysoeriol rhamnosyl hexoside (chrysoeriol rutinoside), quercetin rhamnosyl glucoside (rutin), isorhamnetin rhamnosyl glucoside (isorhamnetin rutinoside), quercetin diglucoside, isorhamnetin diglucoside, luteolin dihexosyl sulfate, luteolin rhamnosyl dihexoside (luteolin rutinoside), and chrysoeriol rhamnosyl dihexoside (chrysoeriol rutinoside), respectively, based on the similarities observed in the comparison of their fragmentation behaviors and with the behaviors reported in the literature [1,13,17,40]. The flavonoid glycosides' DPPH-scavenging action is as follows: isorhamnetin rutinoside > isoquercitrin > isorhamnetin hexoside > chrysoeriol rutinoside > rutin > vitexin rhamnoside $\cong$ luteolin rutinoside [53–56]. The deprotonated molecular ion $[M–H]^-$ at m/z 581.1863 exhibited MS^2 fragment ions at m/z 436.13 and 274.08 by loss of rhamnosyl (145 Da) and rhamnosyl-glucosyl moieties (309 Da). The ion at m/z 274.08 further yielded MS^3 ion at m/z 167.03 and 149.06 through cleavage of the α-β bond and loss of the phloroglucinol moiety (125 Da); hence, compound **19** was tentatively identified as naringin dihydrochalcone, which has been identified for the first time in RADP (Figure 3C). Naringin dihydrochalcone has more excellent antioxidant activity than naringin [57]. It was proposed as a potential therapeutic agent for Alzheimer's disease treatment due to its ability to combat several effects that increase neurogenesis, suppress neuroinflammation, and decrease Aβ levels [58].

2.6.3. Lignans

The monoisotopic mass $[M–H]^-$ at m/z 775.2821 yielded a characteristic fragment ion at m/z 613 by loss of hexosyl moiety (162 Da) and m/z 417.155, confirming the presence of syringaresinol. The MS^2 ion at m/z 417.15 underwent further cleavage of the furfuran ring followed by the loss of two molecules of CH_3 generated ion at m/z 181.05 and 151.00, respectively; compound **30** was tentatively confirmed as a furfuran-type lignan compound, namely erythro-guaiacylglycerol-β-syringaresinol ether hexoside (Figure 4A). Furthermore, the deprotonated ion $[M–H]^-$ of compound **31** was obtained at m/z 805.2926, and it yielded a fragment ion at m/z 643.23 by the loss of the hexosyl moiety, which was 31 Da higher than that of compound **30**. This suggests that the compound has one more methoxy group. The other fragmentation behavior observed was similar to that of the quasimolecular ion of compound **30**, suggesting that compound **31** was erythro-1-(4'-*O*-hexoside-3,5-dimethoxyphenyl)-syringaresinoxyl-propane-1,3-diol (supplementary data Figure S5B). Both lignans were identified for the first time in RADP. The biological activities of these lignans remain unknown [59,60].
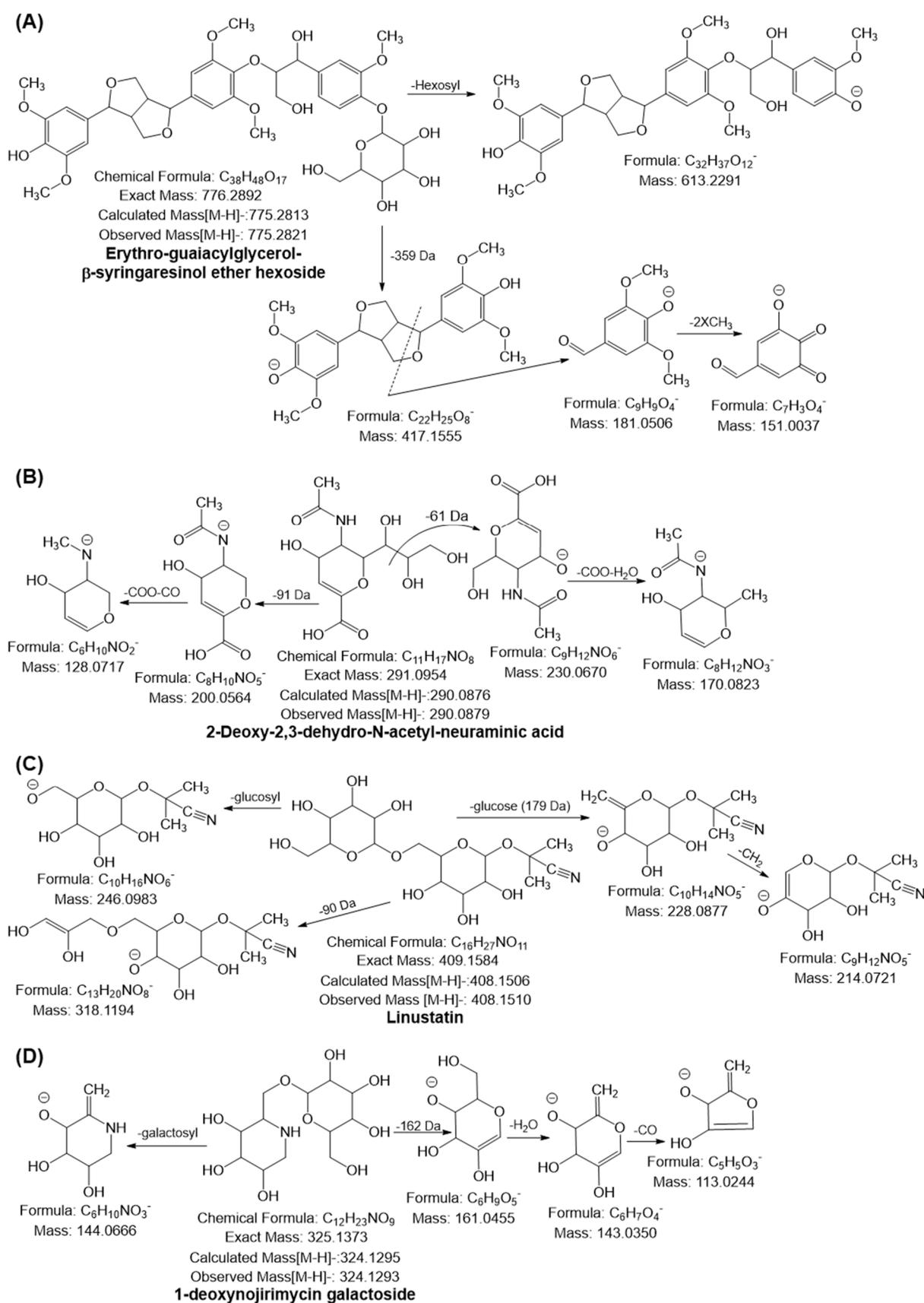
Figure 4. Possible mass fragmentation behavior of identified compounds in RADP. (**A**) erythron-guaiacylglycerol-β-syringaresinol ether hexoside, (**B**) 2-deoxy-2,3-dehydro-N-acetyl-neuraminic acid, (**C**) linustatin, and (**D**) 1-deoxynojirimycin galactoside.

2.6.4. Sialic Acid and Derivatives

The deprotonated ion [M–H]$^-$ of No. **32** was observed at m/z 290.0879, and diagnostic ions were generated at m/z 230.06 and 200.05 by the loss of 61 and 91 Da, respectively. Both MS^2 ions further generated fragment ions at m/z 146.08 ([M-H-230.06-COO-H$_2$O]) and 128.07 ([M-H-200.05-COO-CO]), suggesting that the compound was 2-deoxy-2,3-dehydro-N-acetyl-neuraminic acid (DANA) (Figure 4B). The monoisotopic mass [M–H]$^-$ at m/z 308.0987 and 632.2039 with molecular formulas C$_{11}$H$_{19}$NO$_9$ and C$_{23}$H$_{39}$NO$_{19}$ suggested that compounds **33** and **34** were N-acetyl-neuraminic acid and 3′-sialyllactose (SL), respectively, as their mass fragmentation behaviors were quite similar to the behavior of the quasimolecular ion of compound **32** (supplementary data Figure S5C,D). This is the first time that sialic acid has been identified in RADP. DANA is a glucose-dependent potentiator of insulin secretion [61]. The positive effects of SL on inflammation and immunological homeostasis by altering the gut microbiota profile have been documented in numerous investigations. It also slows the progression of rheumatoid arthritis by reducing chemokines and cytokines, and it relieves osteoarthritis by promoting cartilage regeneration and preventing cartilage from being destroyed [62–64].

2.6.5. Amino Acids, Carboxylic Acids, and Fatty Acids

From comparisons of the mass and the fragmentation behaviors of the precursor ion based on mass spectroscopic analysis reported in the literature and various online databases, compounds **35–38**, **48–55**, and **56–67** were identified as amino acids, carboxylic acids, and fatty acids, respectively (Table 5) [17,40,47,48,65–67]. Pyroglutamic acid is utilized as a humectant in products for dry skin and hair and has potential anti-angiotensin-converting enzyme activity [68].

2.6.6. Sugar Molecules

From the monoisotopic ion [M–H]$^-$ at m/z 351.0569, the diagnostic ions yielded at m/z 291.07 ([M-H-COO-H$_2$O]), 273.06 ([M-H-291.07-H$_2$O]), 175.02 ([M-H-glucoronyl]), and 131.03 ([M-H-175.02-COO]). Compound **46** was tentatively identified as unsaturated digalacturonate; this is the first time that this compound was found in RADP (supplementary data Figure S6A). Furthermore, compound **47** was tentatively identified as xylosmaloside with the molecular formula C$_{18}$H$_{20}$O$_9$, and this compound generated the deprotonated ion [M–H]$^-$ at m/z 379.1027 and the following mass fragmentation pattern: m/z 343.08 ([M-H-36 Da]), 217.05 ([M-H-162 Da]), 179.05 (xylose), and 161.04 ([M-H-179.05-18 Da]) (supplementary data Figure S6B). This compound too was identified for the first time in RADP. Compounds **39–45** were confirmed as sugar molecules from comparison of their deprotonated ion mass and fragmentation behaviors with those reported in the literature and online databases [47,48,65–67,69]. Maltitol, one of the sugar molecules, inhibits *Streptococcus mutans* DMST 18777, and xylosmaloside has greater antioxidant properties than ascorbic acid [70,71].

2.6.7. Others

Compound **68** yielded a monoisotopic mass [M–H]$^-$ at m/z 408.1510 and generated a diagnostic ion at m/z 246.09 by loss of the glucosyl moiety (162 Da). Furthermore, a prominent ion peak was observed at m/z 228.08 with the neutral loss of glucose moiety. In the MS^3 spectra of m/z 228, ions at m/z 214.07 were observed owing to the loss of CH$_2$, suggesting that the compound was linustatin, a cyanogenic glucoside, which has also been identified for the first time in RADP (Figure 4C). The deprotonated ion [M–H]$^-$ at m/z 324.1293 and the characteristic ions at m/z 144.06 ([M-H-galactosyl moiety]) and 161.04 ([M-H-162 Da]) were observed. The fragment ion at m/z 161.04 further yielded ions at m/z 143.03 and 113.02 through the consecutive loss of H$_2$O and CO. Based on the mass fragmentation behavior, compound **69** was tentatively identified, for the first time in RADP, as 1-deoxynojirimycin galactoside (Figure 4D). Although 1-deoxynojirimycin galactoside's pharmacological role is unknown, its aglycone component (1-deoxynojirimycin) has a

powerful antihyperglycemic and anti-obesity activity [72]. Compound **70** was tentatively identified as oxycoumarin-4-acetic acid methyl ether hexoside, with the molecular formula $C_{18}H_{20}O_{10}$, and it yielded a deprotonated ion $[M–H]^-$ at m/z 395.0968. It gave a prominent peak at m/z 233.04, owing to the loss of galacotosyl moiety (162 Da). In the MS^3 spectra of 233.04, ions at m/z 205.05 and 161.02 were generated through the loss of CO and acetic acid methyl ether moiety (supplementary data Figure S6C).

3. Materials and Methods

3.1. Sample Collection and Preparation

Ripe Ajwa dates were collected from an Ajwa date farm located in Medina, Saudi Arabia, and were identified by a scientific officer at the National Herbarium and Genebank of Saudi Arabia. A voucher specimen (No. NHG005) has been stored in our laboratory for further investigation. Heat extraction (HE) was performed in a temperature-controlled heater using a round-bottom flask with a condenser to prevent solvent evaporation (Soxlet water bath C-WBS-D6, Changshin Science, Seoul, Republic of Korea). The dry powder samples (10.0 g) were extracted thrice with 100 mL solvent according to the investigational design shown in supplemental data Tables S1 and S2. After extraction, the samples were filtered using Whatman No. 1 filter paper (Schleicher & Schuell, Keene, NH, USA), concentrated to dryness in a rotary evaporator (Tokyo Rikakikai Co. Ltd., Tokyo, Japan) at 40 °C and 50 rpm, and then lyophilized using a freeze-dryer (Il-shin Biobase, Goyang, Republic of Korea). The RADP extract was kept at −20 °C before subsequent experiments.

3.2. Antioxidant Activities

The total phenolic content (TPC) and total flavonoid content (TFC) in ripe Ajwa date pulp (RADP) extracts were determined using the Folin–Ciocalteu test and the aluminum chloride colorimetric method, respectively [45]. The TPC (y = 0.0512x + 0.0018; r^2 = 0.9835) and TFC (y = 0.014x + 0.0021; r^2 = 0.9994) were determined using the corresponding regression equations for the calibration curves. The TPC was expressed in terms of the gallic acid equivalent (mg)/dry weight sample (g), and the TFC in terms of the catechin equivalent (mg)/dry weight sample (g). The DPPH radical scavenging assay and the cupric-reducing antioxidant capacity (CUPRAC) assay were applied to assess the antioxidant activity of RADP extracts [28].

3.3. Experimental Design of Response Surface Methodology (RSM) for the Extraction Process

The RSM model was developed for the extraction of phenolic compounds from RADP; the independent process factors were ethanol concentration (X_1), time (X_2), and temperature (X_3), and the response variables were TPC (Y_1), TFC (Y_2), DPPH-scavenging activity (Y_3), and CUPRAC (Y_4). The extractions were carried out using a central composite design (CCD) with three components and five layers. The response variables were fitted to the second-order polynomial model equation shown in Equation (5), which specifies the relationship between the independent variables and responses.

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \sum_{i=1}^{n} \beta_{ii} X_{ii}^2 + \sum_{i}^{n-1} \sum_{j}^{n} \beta_{ij} X_{ij} \tag{5}$$

where **Y** is the response variable, and X_i and X_j are the independent coded variables; β_0 denotes the constant coefficient, and β_i, β_{ii}, and β_{ij} denote the coefficients of linear, quadratic, and interaction effects, respectively. Design Expert 11 was used for the RSM analysis and multiple linear regression (Stat-Ease, Minneapolis, MN, USA). To select the best model, the function parameter was selected as "None" and various R^2 (coefficient of determination) values, PRESS (the predicted residual sum of squares), BIC (the Bayesian information criterion), and AICc (the Akaike information criterion–second-order) were considered. The F-values with $p < 0.05$ indicated statistical significance. The interaction outcome of each factor on the response value was represented in the form of three-dimensional (3D) surface plots.

3.4. Artificial Neural Networks (ANN) Modeling

An ANN was used to predict a nonlinear relationship between RADP's input parameters (X_1, X_2, and X_3) and response variables (Y_1, Y_2, Y_3, and Y_4). The Neural Network ToolboxTM of MATLAB was used for the nonlinear analysis, and the multilayer perceptron (MLP) was utilized to map layers of independent and response variables using a back-propagation feed-forward ANN model [28]. The hit-and-trial method was used to calculate the required number of neurons in the hidden layer, which ranged from 1 to 15, to minimize discrepancies between the predicted and experimental results of RADP. Figure 5 depicts the three-layered fundamental architecture (MLP topology) of the ANN model. The model's fundamental architecture consists of three layers: a three-neuron input layer representing the three independent factors (ethanol concentration, extraction time, and temperature); a six-neuron hidden layer; and an output layer representing the four response variables (TPC, TFC, DPPH-scavenging activity, and CUPRAC value) of RADP. The investigational dataset used to create the RSM model was also used to develop an ANN model: network training (70%: 20 points), validation (15%: 5 points), and network testing (15%: 5 points). In the ANN design, the output signal was made by sending the weighted sum of the input variables to each neuron through an activation function. This function was usually nonlinear and was represented by the hidden layer.
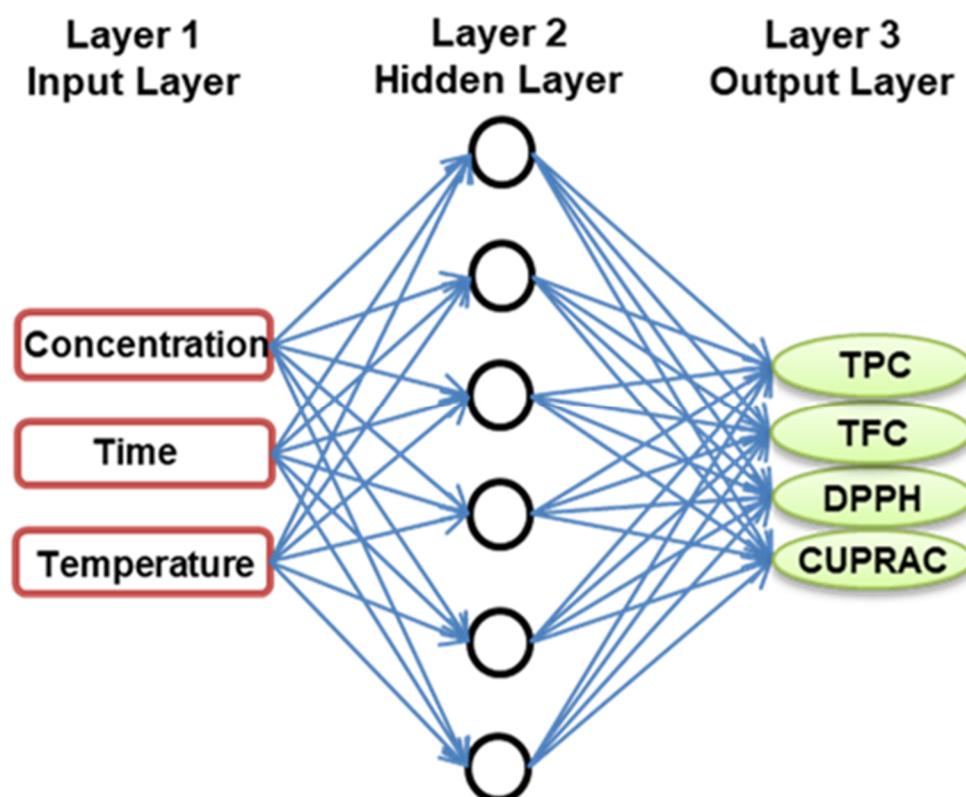
Figure 5. Optimal topology of a developed ANN model.

3.5. Comparison of the Prediction Ability of the RSM and ANN Models

The following equations (Equations (6)–(9)) were used to calculate various statistical parameters, including the correlation coefficient (R^2), root-mean-square error (*RMSE*), absolute average deviation (*AAD*), and standard error of prediction (*SEP*), to compare the estimation capabilities of RSM and ANN for improving the extraction procedure with better TPC, TFC, and antioxidant potential of RADP.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_p - Y_e)^2}{\sum_{i=1}^{n}(Y_m - Y_e)^2} \qquad (6)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_p - Y_e)^2}{n}} \tag{7}$$

$$AAD = \left[\frac{\sum_{i=1}^{n}(|Y_p - Y_e|/Y_e)}{n}\right] \times 100 \tag{8}$$

$$SEP = \frac{RMSE}{Y_m} \times 100 \tag{9}$$

where Y_p is the predicted response; Y_e is the observed response; Y_m is the average response variable; n is the number of experiments.

3.6. Validation of the Model

Derringer's desire function was used to find the ideal conditions for maximizing all replies in a single experiment. Each response is turned into a unique desirability function ranging from 0 to 1 during this procedure (in the order of the lowest to the highest desirability). The component functions are then combined to create a total desirability function. The total desirability function is constructed using the following equation [1].

$$D = \left(d_1^{w1} d_2^{w2} \ldots . d_n^{wn}\right)^{1/\sum wi} \tag{10}$$

Response surface and desirability function analyses were used to determine the optimal RADP extraction parameters, which were 75% ethanol concentration, 3.7 h of extraction time, and 67 °C temperature with D = 0.934. A triple experiment was conducted under the best possible circumstances to validate the existing model. The average experimental results were then compared to the predicted results. The experimental data were also compared to the values that the model indicated. To compare the observed and anticipated outcomes of RADP, Equation (11) was utilized to calculate the relative standard deviation (*RSD*).

$$RSD\ (\%) = \frac{Standard\ deviation\ between\ predicted\ and\ experimental\ values}{Mean\ values\ between\ predicted\ and\ experimental\ values} \times 100 \tag{11}$$

The resulting data were analyzed and optimized for all response circumstances when the *RSD*% was <10. Additionally, the electrospray ionization mass spectrometry (ESI-MS)/MS profiles of bioactive compounds were found under ideal circumstances of RADP.

3.7. Analysis of Chemical Compounds by ESI-MS/MS

The Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific INC., San Jose, CA, USA) was used to conduct the positive (+) and negative (−) mode ESI-MS investigations of RADP. A 500 μL graded syringe (Hamilton Company Inc., Reno, NV, USA) and a 15 μL/min syringe pump (Model 11, Harvard, Holliston, MA, USA) were used to immerse the RADP in the ESI source. The normal negative mode ESI-MS conditions were as follows: mass resolution of 140,000 (full width at half maximum, FWHM), sheath gas flow rate of 5, seep gas flow rate of 0, auxiliary gas flow rate of 0, spray voltage of 4.20 kV, capillary temperature of 320 °C, S-lens Rf level, and automatic gain control of 5 E 6. The MS/MS experiments were conducted on the same instrument utilizing three unique stepwise normalized collision energies (10, 30, and 40) to produce better fragmentations of each peak, hence facilitating the confirmation of RADP for each compound [48].

The Xcalibur 3.1 with Foundation 3.1 (Thermo Fisher Scientific Inc. Rockford, IL, USA) was used to process the collected mass spectral data. The m/z peaks were tentatively identified by comparing their calculated (exact) masses of protonated and/or deprotonated (M–H) adducts with the m/z values and ESI-MS/MS fragmentation patterns from the in-house MS/MS database and online databases such as METLIN [65], CFM-ID 4.0 [66] and FooDB [67]. The chemical structure was drawn using ChemDraw Professional 15.0 (PerkinElmer, Waltham, MA, USA).

3.8. Statistical Analysis

The experimental results were statistically analyzed using Design Expert 11 and MATLAB software. All data were reported as the mean $\pm$ standard deviation of at least three independent experiments (n = 3), each with three sample replicates. Differences were considered significant at $p < 0.001$, $p < 0.01$, and $p < 0.05$.

4. Conclusions

This work was the first to use two modeling methodologies (RSM and ANN) to optimize the heat extraction (HE) conditions for achieving the highest amount of TPC, TFC, and antioxidant potential of RADP. The ANN model is more accurate and reliable for optimizing the extraction conditions of RADP for improved attainment of TPC, TFC, and antioxidant characteristics, as evidenced by the greater R^2 and lower RMSE, AAD, and SEP compared to the RSM model. The optimal conditions (75% ethanol, extraction time of 3.7 h, and extraction temperature of 67 °C) were determined to maximize the TPC, TFC, and antioxidant potential of RADP. Under optimum conditions, TPC, TFC, DPPH radical scavenging effect, and ascorbic acid equivalent CUPRAC value were obtained as 4.53 mg GAE/g, 3.30 mg CAE/g, 10.69% inhibition, and 1.41 µM, respectively. RADP contains phenolic acids, flavonoids, sialic acids, lignans, etc., as determined by high-resolution mass spectrometry analysis. Sialic acids and lignans were described for the first time in RADP as bioactive substances. Based on these data, we can infer that RADP, a viable candidate for an antioxidant functional food, could be used extensively in the nutraceutical and pharmaceutical industries.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ph16020319/s1. Figure S1: The R^2 value of training, validation, test, and overall for TPC during ANN. Figure S2: The R^2 value of training, validation, test, and overall for TFC during ANN. Figure S3: The R^2 value of training, validation, test, and overall for DPPH during ANN. Figure S4: The R^2 value of training, validation, test, and overall for CUPRAC during ANN. Figure S5: Possible mass fragmentation behavior of identified compounds in RADP. (A) Coumalic acid, (B) Erythro-1-(4″-O-β-D-glucopyranoside-3,5-dimethyoxyphenyl)-2-syringaresinoxyl-propane-1,3-diol, (C) N-acetyl-α-neuraminic acid, and (D) 6′-sialyllactose. Figure S6: Possible mass fragmentation behavior of identified compounds in RADP. (A) Unsaturated digalacturonate, (B) Xylosmaloside, and (C) Oxycoumarin-4-acetic acid methyl ester hexoside. Table S1: Central composite design (CCD) for the independent variables and corresponding response value in RSM and ANN (predicted). Table S2: Independent process variables with experimental ranges and levels for heat reflux extraction of ripe Ajwa date pulp (RADP).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sedraoui, S.; Badr, A.; Barba, M.G.M.; Doyen, A.; Tabka, Z.; Desjardins, Y. Optimization of the Ultrahigh-Pressure–Assisted Extraction of Phenolic Compounds and Antioxidant Activity from Palm Dates (*Phoenix dactylifera* L.). *Food Anal. Methods* **2020**, *13*, 1556–1569. [CrossRef]

2. Liyana-Pathirana, C.; Shahidi, F. Optimization of extraction of phenolic compounds from wheat using response surface methodology. *Food Chem.* **2005**, *93*, 47–56. [CrossRef]

3. Tabaraki, R.; Nateghi, A. Optimization of ultrasonic-assisted extraction of natural antioxidants from rice bran using response surface methodology. *Ultrason. Sonochemistry* **2011**, *18*, 1279–1286. [CrossRef] [PubMed]

4. Kusuma, H.S.; Amenaghawon, A.N.; Darmokoesoemo, H.; Neolaka, Y.A.B.; Widyaningrum, B.A.; Onowise, S.U.; Anyalewechi, C.L. A comparative evaluation of statistical empirical and neural intelligence modeling of Manihot esculenta-derived leaves extract for optimized bio-coagulation-flocculation of turbid water. *Ind. Crops Prod.* **2022**, *186*, 115194. [CrossRef]

5. Bezerra, M.A.; Santelli, R.E.; Oliveira, E.P.; Villar, L.S.; Escaleira, L.A. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. *Talanta* **2008**, *76*, 965–977. [CrossRef]

6. Baş, D.; Boyacı, İ.H. Modeling and optimization I: Usability of response surface methodology. *J. Food Eng.* **2007**, *78*, 836–845. [CrossRef]

7. Desai, K.M.; Survase, S.A.; Saudagar, P.S.; Lele, S.S.; Singhal, R.S. Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. *Biochem. Eng. J.* **2008**, *41*, 266–273. [CrossRef]

8. Zobel, C.W.; Cook, D.F. Evaluation of neural network variable influence measures for process control. *Eng. Appl. Artif. Intell.* **2011**, *24*, 803–812. [CrossRef]

9. Yasin, B.R.; El-Fawal, H.A.; Mousa, S.A. Date (Phoenix dactylifera) Polyphenolics and Other Bioactive Compounds: A Traditional Islamic Remedy's Potential in Prevention of Cell Damage, Cancer Therapeutics and Beyond. *Int. J. Mol. Sci.* **2015**, *16*, 30075–30090. [CrossRef]

10. Al-Yahya, M.; Raish, M.; AlSaid, M.S.; Ahmad, A.; Mothana, R.A.; Al-Sohaibani, M.; Al-Dosari, M.S.; Parvez, M.K.; Rafatullah, S. 'Ajwa' dates (*Phoenix dactylifera* L.) extract ameliorates isoproterenol-induced cardiomyopathy through downregulation of oxidative, inflammatory and apoptotic molecules in rodent model. *Phytomedicine Int. J. Phytother. Phytopharm.* **2016**, *23*, 1240–1248. [CrossRef]

11. Almatroodi, S.A.; Khan, A.A.; Aloliqi, A.A.; Ali Syed, M.; Rahmani, A.H. Therapeutic Potential of Ajwa Dates (*Phoenix dactylifera*) Extract in Prevention of Benzo(a)pyrene-Induced Lung Injury through the Modulation of Oxidative Stress, Inflammation, and Cell Signalling Molecules. *Appl. Sci.* **2022**, *12*, 6784. [CrossRef]

12. Hassan, S.M.A.; Aboonq, M.S.; Albadawi, E.A.; Aljehani, Y.; Abdel-Latif, H.M.; Mariah, R.A.; Shafik, N.M.; Soliman, T.M.; Abdel-Gawad, A.R.; Omran, F.M.; et al. The Preventive and Therapeutic Effects of Ajwa Date Fruit Extract Against Acute Diclofenac Toxicity-Induced Colopathy: An Experimental Study. *Drug Des. Devel. Ther.* **2022**, *16*, 2601–2616. [CrossRef]

13. Khalid, S.; Khalid, N.; Khan, R.S.; Ahmed, H.; Ahmad, A. A review on chemistry and pharmacology of Ajwa date fruit and pit. *Trends Food Sci. Technol.* **2017**, *63*, 60–69. [CrossRef]

14. Rahmani, A.H.; Aly, S.M.; Ali, H.; Babiker, A.Y.; Srikar, S.; Khan, A.A. Therapeutic effects of date fruits (*Phoenix dactylifera*) in the prevention of diseases via modulation of anti-inflammatory, anti-oxidant and anti-tumour activity. *Int. J. Clin. Exp. Med.* **2014**, *7*, 483–491. [PubMed]

15. Siddiqui, S.; Ahmad, R.; Khan, M.A.; Upadhyay, S.; Husain, I.; Srivastava, A.N. Cytostatic and Anti-tumor Potential of Ajwa Date Pulp against Human Hepatocellular Carcinoma HepG2 Cells. *Sci. Rep.* **2019**, *9*, 245. [CrossRef] [PubMed]

16. Boulenouar, N.; Marouf, A.; Cheriti, A. Antifungal activity and phytochemical screening of extracts from *Phoenix dactylifera* L. cultivars. *Nat. Prod. Res.* **2011**, *25*, 1999–2002. [CrossRef] [PubMed]

17. Nematallah, K.A.; Ayoub, N.A.; Abdelsattar, E.; Meselhy, M.R.; Elmazar, M.M.; El-Khatib, A.H.; Linscheid, M.W.; Hathout, R.M.; Godugu, K.; Adel, A.; et al. Polyphenols LC-MS2 profile of Ajwa date fruit (*Phoenix dactylifera* L.) and their microemulsion: Potential impact on hepatic fibrosis. *J. Funct. Foods* **2018**, *49*, 401–411. [CrossRef]

18. Almusallam, I.A.; Mohamed Ahmed, I.A.; Babiker, E.E.; Al Juhaimi, F.Y.; Fadimu, G.J.; Osman, M.A.; Al Maiman, S.A.; Ghafoor, K.; Alqah, H.A.S. Optimization of ultrasound-assisted extraction of bioactive properties from date palm (*Phoenix dactylifera* L.) spikelets using response surface methodology. *LWT* **2021**, *140*, 110816. [CrossRef]

19. Benkerrou, F.; Bachir bey, M.; Amrane, M.; Louaileche, H. Ultrasonic-assisted extraction of total phenolic contents from Phoenix dactylifera and evaluation of antioxidant activity: Statistical optimization of extraction process parameters. *J. Food Meas. Charact.* **2018**, *12*, 1910–1916. [CrossRef]

20. Dahmoune, F.; Remini, H.; Dairi, S.; Aoun, O.; Moussi, K.; Bouaoudia-Madi, N.; Adjeroud, N.; Kadri, N.; Lefsih, K.; Boughani, L.; et al. Ultrasound assisted extraction of phenolic compounds from P. lentiscus L. leaves: Comparative study of artificial neural network (ANN) versus degree of experiment for prediction ability of phenolic compounds recovery. *Ind. Crops Prod.* **2015**, *77*, 251–261. [CrossRef]

21. Aung, T.; Kim, S.J.; Eun, J.B. A hybrid RSM-ANN-GA approach on optimisation of extraction conditions for bioactive component-rich laver (*Porphyra dentata*) extract. *Food Chem* **2022**, *366*, 130689. [CrossRef] [PubMed]

22. Qadir, R.; Anwar, F.; Naseem, K.; Tahir, M.H.; Alhumade, H. Enzyme-Assisted Extraction of Phenolics from Capparis spinosa Fruit: Modeling and Optimization of the Process by RSM and ANN. *ACS Omega* **2022**, *7*, 33031–33038. [CrossRef]

23. Spiess, A.N.; Neumeyer, N. An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacol.* **2010**, *10*, 6. [CrossRef] [PubMed]

24. Gu, Y.; Wei, H.-L.; Balikhin, M.M. Nonlinear predictive model selection and model averaging using information criteria. *Syst. Sci. Control Eng.* **2018**, *6*, 319–328. [CrossRef]

25. Williams, A.O.F.; Akanbi, O.D. Statistical modeling and optimization of the bleachability of regenerated spent bleaching earth using response surface methodology and artificial neural networks with genetic algorithm. *Chem. Prod. Process Model.* 2022; *in press*. [CrossRef]

26. Vrieze, S.I. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychol. Methods* **2012**, *17*, 228–243. [CrossRef]

27. Rebollo-Hernanz, M.; Cañas, S.; Taladrid, D.; Segovia, Á.; Bartolomé, B.; Aguilera, Y.; Martín-Cabrejas, M.A. Extraction of phenolic compounds from cocoa shell: Modeling using response surface methodology and artificial neural networks. *Sep. Purif. Technol.* **2021**, *270*, 118779. [CrossRef]

28. Choi, H.-J.; Naznin, M.; Alam, M.B.; Javed, A.; Alshammari, F.H.; Kim, S.; Lee, S.-H. Optimization of the extraction conditions of Nypa fruticans Wurmb. using response surface methodology and artificial neural network. *Food Chem.* **2022**, *381*, 132086. [CrossRef]

29. Xu, S.; Li, X.; Liu, S.; Tian, P.; Li, D. *Juniperus sabina* L. as a Source of Podophyllotoxins: Extraction Optimization and Anti-cholinesterase Activities. *Int. J. Mol. Sci.* **2022**, *23*, 10205. [CrossRef]

30. Huang, S.-M.; Kuo, C.-H.; Chen, C.-A.; Liu, Y.-C.; Shieh, C.-J. RSM and ANN modeling-based optimization approach for the development of ultrasound-assisted liposome encapsulation of piceid. *Ultrason. Sonochemistry* **2017**, *36*, 112–122. [CrossRef]

31. Kuo, C.-H.; Liu, T.-A.; Chen, J.-H.; Chang, C.-M.J.; Shieh, C.-J. Response surface methodology and artificial neural network optimized synthesis of enzymatic 2-phenylethyl acetate in a solvent-free system. *Biocatal. Agric. Biotechnol.* **2014**, *3*, 1–6. [CrossRef]

32. Javed, A.; Naznin, M.; Alam, M.B.; Fanar, A.; Song, B.-R.; Kim, S.; Lee, S.-H. Metabolite Profiling of Microwave-Assisted Sargassum fusiforme Extracts with Improved Antioxidant Activity Using Hybrid Response Surface Methodology and Artificial Neural Networking-Genetic Algorithm. *Antioxidants* **2022**, *11*, 2246. [CrossRef]

33. Aklilu, E.G.; Adem, A.; Kasirajan, R.; Ahmed, Y. Artificial neural network and response surface methodology for modeling and optimization of activation of lactoperoxidase system. *South Afr. J. Chem. Eng.* **2021**, *37*, 12–22. [CrossRef]

34. Xi, J.; Wang, B. Optimization of Ultrahigh-Pressure Extraction of Polyphenolic Antioxidants from Green Tea by Response Surface Methodology. *Food Bioprocess Technol.* **2013**, *6*, 2538–2546. [CrossRef]

35. Do, Q.D.; Angkawijaya, A.E.; Tran-Nguyen, P.L.; Huynh, L.H.; Soetaredjo, F.E.; Ismadji, S.; Ju, Y.H. Effect of extraction solvent on total phenol content, total flavonoid content, and antioxidant activity of Limnophila aromatica. *J. Food Drug Anal.* **2014**, *22*, 296–302. [CrossRef] [PubMed]

36. Zhu, K.-X.; Lian, C.-X.; Guo, X.-N.; Peng, W.; Zhou, H.-M. Antioxidant activities and total phenolic contents of various extracts from defatted wheat germ. *Food Chem.* **2011**, *126*, 1122–1126. [CrossRef]

37. Derrien, M.; Badr, A.; Gosselin, A.; Desjardins, Y.; Angers, P. Optimization of a green process for the extraction of lutein and chlorophyll from spinach by-products using response surface methodology (RSM). *LWT—Food Sci. Technol.* **2017**, *79*, 170–177. [CrossRef]

38. Schymanski, E.L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H.P.; Hollender, J. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* **2014**, *48*, 2097–2098. [CrossRef] [PubMed]

39. Ostrowski, W.; Wojakowska, A.; Grajzer, M.; Stobiecki, M. Mass spectrometric behavior of phenolic acids standards and their analysis in the plant samples with LC/ESI/MS system. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2014**, *967*, 21–27. [CrossRef]

40. Najm, O.A.; Addnan, F.H.; Mohd-Manzor, N.F.; Elkadi, M.A.; Abdullah, W.O.; Ismail, A.; Mansur, F.A.F. Identification of Phytochemicals of *Phoenix dactylifera* L. Cv Ajwa with UHPLC-ESI-QTOF-MS/MS. *Int. J. Fruit Sci.* **2021**, *21*, 848–867. [CrossRef]

41. Zeng, X.; Wang, Y.; Qiu, Q.; Jiang, C.; Jing, Y.; Qiu, G.; He, X. Bioactive phenolics from the fruits of Livistona chinensis. *Fitoterapia* **2012**, *83*, 104–109. [CrossRef]

42. Li, F.; Yan, T.T.; Fu, Y.Y.; Zhang, N.L.; Wang, L.; Zhang, Y.B.; Du, J.; Liu, J.F. New phenylpropanoid glycosides from Illicium majus and their radical scavenging activities. *Chem. Biodivers.* **2021**, *18*, e2001012. [CrossRef] [PubMed]

43. Nam, S.-H.; Kim, Y.-M.; Walsh, M.K.; Wee, Y.-J.; Yang, K.-Y.; Ko, J.-A.; Han, S.; Thanh Hanh Nguyen, T.; Kim, J.Y.; Kim, D. Synthesis and functional characterization of caffeic acid glucoside using Leuconostoc mesenteroides dextransucrase. *J. Agric. Food Chem.* **2017**, *65*, 2743–2750. [CrossRef] [PubMed]

44. Tsagkarakou, A.S.; Chasapi, S.A.; Koulas, S.M.; Tsialtas, I.; Kyriakis, E.; Drakou, C.E.; Kun, S.; Somsák, L.; Spyroulias, G.A.; Psarra, A.-M.G.; et al. Structure activity relationship of the binding of p-coumaroyl glucose to glycogen phosphorylase and its effect on hepatic cell metabolic pathways. *Eur. J. Med. Chem. Rep.* **2021**, *3*, 100011. [CrossRef]

45. Błaszczak, W.; Jeż, M.; Szwengiel, A. Polyphenols and inhibitory effects of crude and purified extracts from tomato varieties on the formation of advanced glycation end products and the activity of angiotensin-converting and acetylcholinesterase enzymes. *Food Chem.* **2020**, *314*, 126181. [CrossRef]

46. Wang, H.; Zhao, W.; Choomuenwai, V.; Andrews, K.T.; Quinn, R.J.; Feng, Y. Chemical investigation of an antimalarial Chinese medicinal herb Picrorhiza scrophulariiflora. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 5915–5918. [CrossRef]

47. Naznin, M.; Badrul Alam, M.; Alam, R.; Islam, S.; Rakhmat, S.; Lee, S.-H.; Kim, S. Metabolite profiling of Nymphaea rubra (Burm. f.) flower extracts using cyclic ion mobility–mass spectrometry and their associated biological activities. *Food Chem.* **2023**, *404*, 134544. [CrossRef] [PubMed]

48. Alam, M.B.; Naznin, M.; Islam, S.; Alshammari, F.H.; Choi, H.J.; Song, B.R.; Kim, S.; Lee, S.H. High Resolution Mass Spectroscopy-Based Secondary Metabolite Profiling of *Nymphaea nouchali* (Burm. f) Stem Attenuates Oxidative Stress via Regulation of MAPK/Nrf2/HO-1/ROS Pathway. *Antioxidants* **2021**, *10*, 719. [CrossRef]

49. Salehi, B.; Machin, L.; Monzote, L.; Sharifi-Rad, J.; Ezzat, S.M.; Salem, M.A.; Merghany, R.M.; El Mahdy, N.M.; Kılıç, C.S.; Sytar, O.; et al. Therapeutic Potential of Quercetin: New Insights and Perspectives for Human Health. *ACS Omega* **2020**, *5*, 11849–11872. [CrossRef]

50. Aboulaghras, S.; Sahib, N.; Bakrim, S.; Benali, T.; Charfi, S.; Guaouguaou, F.E.; Omari, N.E.; Gallo, M.; Montesano, D.; Zengin, G.; et al. Health Benefits and Pharmacological Aspects of Chrysoeriol. *Pharmaceuticals* **2022**, *15*, 973. [CrossRef]

51. Kachlicki, P.; Piasecka, A.; Stobiecki, M.; Marczak, Ł. Structural Characterization of Flavonoid Glycoconjugates and Their Derivatives with Mass Spectrometric Techniques. *Molecules* **2016**, *21*, 1494. [CrossRef]

52. Vukics, V.; Guttman, A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. *Mass Spectrom. Rev.* **2010**, *29*, 1–16. [CrossRef] [PubMed]

53. Li, X.; Jiang, Q.; Wang, T.; Liu, J.; Chen, D. Comparison of the Antioxidant Effects of Quercitrin and Isoquercitrin: Understanding the Role of the 6″-OH Group. *Molecules* **2016**, *21*, 1246. [CrossRef] [PubMed]

54. Choi, J.S.; Jung, M.J.; Park, H.J.; Chung, H.Y.; Kang, S.S. Further isolation of peroxynitrite and 1,1-diphenyl-2-picrylhydrazyl radical scavenging isorhamnetin 7-O-glucoside from the leaves of *Brassica juncea* L. *Arch. Pharmacal Res.* **2002**, *25*, 625–627. [CrossRef] [PubMed]

55. Delazar, A.; Sabzevari, A.; Mojarrab, M.; Nazemiyeh, H.; Esnaashari, S.; Nahar, L.; Razavi, S.M.; Sarker, S.D. Free-radical-scavenging principles from Phlomis caucasica. *J. Nat. Med.* **2008**, *62*, 464–466. [CrossRef]

56. Hyun, S.K.; Jung, Y.J.; Chung, H.Y.; Jung, H.A.; Choi, J.S. Isorhamnetin glycosides with free radical and ONOO-scavenging activities from the stamens of Nelumbo nucifera. *Arch. Pharmacal Res.* **2006**, *29*, 287–292. [CrossRef]

57. Nakamura, Y.; Watanabe, S.; Miyake, N.; Kohno, H.; Osawa, T. Dihydrochalcones: Evaluation as novel radical scavenging antioxidants. *J. Agric. Food Chem.* **2003**, *51*, 3309–3312. [CrossRef]

58. Nguyen, T.H.H.; Woo, S.-M.; Nguyen, N.A.; Cha, G.-S.; Yeom, S.-J.; Kang, H.-S.; Yun, C.-H. Regioselective Hydroxylation of Naringin Dihydrochalcone to Produce Neoeriocitrin Dihydrochalcone by CYP102A1 (BM3) Mutants. *Catalysts* **2020**, *10*, 823. [CrossRef]

59. Kim, S.; Lee, E.Y.; Hillman, P.F.; Ko, J.; Yang, I.; Nam, S.J. Chemical Structure and Biological Activities of Secondary Metabolites from *Salicornia europaea* L. *Molecules* **2021**, *26*, 2252. [CrossRef]

60. Limongelli, F.; Crupi, P.; Clodoveo, M.L.; Corbo, F.; Muraglia, M. Overview of the Polyphenols in Salicornia: From Recovery to Health-Promoting Effect. *Molecules* **2022**, *27*, 7954. [CrossRef]

61. Minami, A.; Fujita, Y.; Shimba, S.; Shiratori, M.; Kaneko, Y.K.; Sawatani, T.; Otsubo, T.; Ikeda, K.; Kanazawa, H.; Mikami, Y.; et al. The sialidase inhibitor 2,3-dehydro-2-deoxy-N-acetylneuraminic acid is a glucose-dependent potentiator of insulin secretion. *Sci. Rep.* **2020**, *10*, 5198. [CrossRef]

62. Kang, L.-J.; Oh, E.; Cho, C.; Kwon, H.; Lee, C.-G.; Jeon, J.; Lee, H.; Choi, S.; Han, S.J.; Nam, J.; et al. 3′-Sialyllactose prebiotics prevents skin inflammation via regulatory T cell differentiation in atopic dermatitis mouse models. *Sci. Rep.* **2020**, *10*, 5603. [CrossRef] [PubMed]

63. Perdijk, O.; Van Baarlen, P.; Fernandez-Gutierrez, M.M.; Van den Brink, E.; Schuren, F.H.; Brugman, S.; Savelkoul, H.F.; Kleerebezem, M.; Van Neerven, R.J. Sialyllactose and galactooligosaccharides promote epithelial barrier functioning and distinctly modulate microbiota composition and short chain fatty acid production in vitro. *Front. Immunol.* **2019**, *10*, 94. [CrossRef] [PubMed]

64. Jeon, J.; Kang, L.J.; Lee, K.M.; Cho, C.; Song, E.K.; Kim, W.; Park, T.J.; Yang, S. 3′-Sialyllactose protects against osteoarthritic development by facilitating cartilage homeostasis. *J. Cell. Mol. Med.* **2018**, *22*, 57–66. [CrossRef] [PubMed]

65. Domingo-Almenara, X.; Guijas, C.; Billings, E.; Montenegro-Burke, J.R.; Uritboonthai, W.; Aisporna, A.E.; Chen, E.; Benton, H.P.; Siuzdak, G. The METLIN small molecule dataset for machine learning-based retention time prediction. *Nat. Commun.* **2019**, *10*, 5811. [CrossRef] [PubMed]

66. Wang, F.; Liigand, J.; Tian, S.; Arndt, D.; Greiner, R.; Wishart, D.S. CFM-ID 4.0: More Accurate ESI-MS/MS Spectral Prediction and Compound Identification. *Anal. Chem.* **2021**, *93*, 11692–11700. [CrossRef]

67. Naveja, J.J.; Rico-Hidalgo, M.P.; Medina-Franco, J.L. Analysis of a large food chemical database: Chemical space, diversity, and complexity. *F1000Research* **2018**, *7*, 993. [CrossRef]

68. Šudomová, M.; Hassan, S.T.S.; Khan, H.; Rasekhian, M.; Nabavi, S.M. A Multi-Biochemical and In Silico Study on Anti-Enzymatic Actions of Pyroglutamic Acid against PDE-5, ACE, and Urease Using Various Analytical Techniques: Unexplored Pharmacological Properties and Cytotoxicity Evaluation. *Biomolecules* **2019**, *9*, 392. [CrossRef]

69. Alam, M.B.; Ahmed, A.; Islam, S.; Choi, H.J.; Motin, M.A.; Kim, S.; Lee, S.H. Phytochemical Characterization of Dillenia indica L. Bark by Paper Spray Ionization-Mass Spectrometry and Evaluation of Its Antioxidant Potential Against t-BHP-Induced Oxidative Stress in RAW 264.7 Cells. *Antioxidants* **2020**, *9*, 1099. [CrossRef]

70. Swapana, N.; Noji, M.; Nishiuma, R.; Izumi, M.; Imagawa, H.; Kasai, Y.; Okamoto, Y.; Iseki, K.; Singh, C.B.; Asakawa, Y.; et al. A New Diphenyl Ether Glycoside from Xylosma longifolium Collected from North-East India. *Nat. Prod. Commun.* **2017**, *12*, 1934578X1701200832. [CrossRef]

71. Haewpetch, P.; Rudeekulthamrong, P.; Kaulpiboon, J. Enzymatic Synthesis of Maltitol and Its Inhibitory Effect on the Growth of Streptococcus mutans DMST 18777. *Biomolecules* **2022**, *12*, 167. [CrossRef]

72. Gao, K.; Zheng, C.; Wang, T.; Zhao, H.; Wang, J.; Wang, Z.; Zhai, X.; Jia, Z.; Chen, J.; Zhou, Y.; et al. 1-Deoxynojirimycin: Occurrence, Extraction, Chemistry, Oral Pharmacokinetics, Biological Activities and In Silico Target Fishing. *Molecules* **2016**, *21*, 1600. [CrossRef] [PubMed]