

## Article

# Chemical Characterization of *Marrubium vulgare* Volatiles from Serbia

Milica Aćimović <sup>1,\*</sup>, Stefan Ivanović <sup>2</sup>, Katarina Simić <sup>2</sup>, Lato Pezo <sup>3</sup>, Tijana Zeremski <sup>1</sup>, Jelena Ovuka <sup>1</sup> and Vladimir Sikora <sup>1</sup>

- <sup>1</sup> Institute of Field and Vegetable Crops Novi Sad, Maksima Gorkog 30, 21000 Novi Sad, Serbia; tijana.zeremski@ifvcns.ns.ac.rs (T.Z.); jelena.ovuka@ifvcns.ns.ac.rs (J.O.); vladimir.sikora@ifvcns.ns.ac.rs (V.S.)
- <sup>2</sup> Institute of Chemistry, Technology and Metallurgy, University of Belgrade, 11000 Belgrade, Serbia; stefan.ivanovic@ihtm.bg.ac.rs (S.I.); katarina.simic@ihtm.bg.ac.rs (K.S.)
- <sup>3</sup> Institute of General and Physical Chemistry, University of Belgrade, 11000 Belgrade, Serbia; latopezo@yahoo.co.uk
- \* Correspondence: milica.acimovic@ifvcns.ns.ac.rs

**Abstract:** *Marrubium vulgare* is a cosmopolitan medicinal plant from the Lamiaceae family, which produces structurally highly diverse groups of secondary metabolites. A total of 160 compounds were determined in the volatiles from Serbia during two investigated years (2019 and 2020). The main components were *E*-caryophyllene, followed by germacrene D,  $\alpha$ -humulene and  $\alpha$ -copaene. All these compounds are from sesquiterpene hydrocarbons class which was dominant in both investigated years. This variation in volatiles composition could be a consequence of weather conditions, as in the case of other aromatic plants. According to the unrooted cluster tree with 37 samples of *Marrubium* sp. volatiles from literature and average values from this study, it could be said that there are several chemotypes: *E*-caryophyllene,  $\beta$ -bisabolene,  $\alpha$ -pinene,  $\beta$ -farnesene, *E*-caryophyllene + caryophyllene oxide chemotype, and diverse (unclassified) chemotypes. However, occurring polymorphism could be consequence of adaptation to grow in different environment, especially ecological conditions such as humidity, temperature and altitude, as well as hybridization strongly affected the chemotypes. In addition, this paper aimed to obtain validated models for prediction of retention indices (RIs) of compounds isolated from *M. vulgare* volatiles. A total of 160 experimentally obtained RIs of volatile compounds was used to build the prediction models. The coefficients of determination were 0.956 and 0.964, demonstrating that these models could be used for predicting RIs, due to low prediction error and high  $r^2$ .



**Citation:** Aćimović, M.; Ivanović, S.; Simić, K.; Pezo, L.; Zeremski, T.; Ovuka, J.; Sikora, V. Chemical Characterization of *Marrubium vulgare* Volatiles from Serbia. *Plants* **2021**, *10*, 600. <https://doi.org/10.3390/plants10030600>

Academic Editor: J sus Pal -P ul

Received: 28 February 2021

Accepted: 16 March 2021

Published: 23 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:**   2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** horehound; GC–MS; retention indices; QSRR; boosted trees regression model

## 1. Introduction

*Marrubium vulgare* L., also known as white horehound, is a perennial species from the Lamiaceae family. It is indigenous to the region between the Mediterranean Sea and Central Asia; however, today it is found worldwide, apart from the coldest regions and high altitudes [1]. This plant is highly resistant to drought and due to this it grows well in semiarid areas [2]. Additionally, as it is a moderate salt-tolerant species this medicinal plant could be grown on saline soil [3]. The surface of *M. vulgare* vegetative and generative organs is densely clothed with glandular and nonglandular trichomes which accumulate secondary metabolites [4]. *M. vulgare* produces structurally highly diverse groups of secondary metabolites, thus represents a valuable source of bioactive compounds and preparations with health-promoting effects: antioxidant, hepatoprotective, antiproliferative, anti-inflammatory, antidiabetic, and antimicrobial [5]. The use of this herb in traditional medicine is recorded worldwide for ameliorating chronic cough and cold, numerous conditions related to skin, liver, gastric, heart, and immune system [6]. Generally, *M. vulgare* is poor in essential oil, and the major compounds are diverse [1,7–17].

This proves that there are various chemotypes of *M. vulgare*. The lack of information in this field is pointed out by Yabrir [1,18]. The studies about genus *Marrubium* are mainly focused on taxonomical, morphological, and genetic diversity [4,19–26].

The main aim of this investigation was to determine volatiles composition of *M. vulgare* grown in Serbia during two years and to compare its chemical composition with literature data not only of *M. vulgare* but with other species from this genus as well (*M. anisodon*, *M. aschersonii*, *M. astracanicum*, *M. crassidens*, *M. deserti*, *M. duabense*, *M. parviflorum*, *M. peregrinum*, *M. persicum*, *M. propinquum*, *M. velutinum*). Another goal was to establish the new quantitative structure retention relationship (QSRR) models for anticipating the retention indices (RIs) of certain compounds in *M. vulgare* volatiles obtained by GC–MS chromatography utilizing the genetic algorithm (GA) variable selection method and the boosted trees regression. Furthermore, we gather information about the volatile compounds of species from *Marrubium* genus in order to classify the chemotype of *M. vulgare* from this study according to unrooted cluster tree.

## 2. Results

The main components in *M. vulgare* volatiles were *E*-caryophyllene with 24.6% and 23.0%, followed by germacrene D with 9.6% and 17.0%,  $\alpha$ -humulene with 5.2% and 5.3% as well as  $\alpha$ -copaene with 3.3% and 6.1% in 2019 and 2020, respectively. All these compounds are from the sesquiterpene hydrocarbons class which was dominant in both years of the investigation, 52.0% in 2019 and 67.8% in 2020. This variation in volatiles composition could be a consequence of weather conditions, as in case of other aromatic plants [27–33].

However, some of the components detected in *M. vulgare* volatiles during the two-year research have not been detected yet in this species, while other components have not been detected in other species of this genus. ScienceDirect Elsevier, SpringerLink, PubMed, Scopus, Scifnder, Web of Science, Wiley Online, and Google Scholar databases were reviewed and scientific publications from 1990 until 2020 that deal with chemical composition of volatiles species from genus *Marrubium* were summarized and shown in Table 1.

**Table 1.** Chemical composition of *Marrubium vulgare* during two years (2019 and 2020).

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
1	2E-Hexenal O	Train	892.915	-	-	847	0.2	<i>M. aschersonii</i> [34], <i>M. deserti</i> [35], <i>M. peregrinum</i> [36], <i>M. vulgare</i> [10,12,15,16,34]
2	Furan, 2,5-diethyltetrahydro O	Validation	853.684	-	-	897	0.1	
3	1-Octen-3-ol O	Validation	965.818	976	0.2	974	0.6	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [39], <i>M. deserti</i> [35], <i>M. duabense</i> [40], <i>M. parviflorum</i> [41,42], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [45], <i>M. propinquum</i> [41], <i>M. velutinum</i> [44], <i>M. vulgare</i> [7,8,10,15]
4	2-Pentyl furan O	Train	1059.803	-	-	989	0.1	
5	3-Octanol O	Test	962.233	-	-	992	0.1	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [46], <i>M. duabense</i> [40], <i>M. peregrinum</i> [36,44], <i>M. velutinum</i> [44]
6	Linalool OMN	Train	1106.041	1102	0.1	1098	0.1	<i>M. aschersonii</i> [34], <i>M. astracanicum</i> [46], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [45], <i>M. velutinum</i> [44], <i>M. vulgare</i> [8,10,12,17,34,36,47,49]
7	n-Nonanal O	Train	1078.484	-	-	1102	0.1	<i>M. aschersonii</i> [34], <i>M. deserti</i> [35], <i>M. duabense</i> [40], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [45], <i>M. velutinum</i> [44], <i>M. vulgare</i> [34]
8	E-Thujone OMN	Train	1118.307	-	-	1114	0.1	<i>M. peregrinum</i> [43], <i>M. vulgare</i> [8,15]
9	NI-1	-	-	-	-	1132	0.1	-

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
10	Geijerene O	Train	1192.301	1143	0.1	1139	0.6	<i>M. incanum</i> [50,51], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [43], <i>M. vulgare</i> [50]
11	2E-Nonen-1-al O	Validation	1097.602	-	-	1156	0.1	
12	$\beta$ -Cyclocitral O	Train	1216.889	-	-	1219	0.1	<i>M. peregrinum</i> [44], <i>M. velutinum</i> [44], <i>M. vulgare</i> [10]
13	Cogeijerene O	Train	1203.235	-	-	1283	0.1	
14	Pregeijerene O	Train	1149.857	1290	0.1	1287	0.2	<i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [43]
15	Thymol AR	Test	1209.017	1292	0.3	-	-	<i>M. deserti</i> [52], <i>M. vulgare</i> [7,8,10,15,50]
16	2-Undecanone O	Train	1269.228	1295	0.1	1292	Trace	<i>M. vulgare</i> [15]
17	Carvacrol AR	Validation	1179.072	1302	0.1	-	-	<i>M. duabense</i> [40], <i>M. incanum</i> [50], <i>M. parviflorum</i> [42], <i>M. peregrinum</i> [43], <i>M. vulgare</i> [7,8,10,49,50]
18	$\delta$ -Elemene ST	Test	1512.436	-	-	1336	0.1	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. deserti</i> [35,40], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [47], <i>M. peregrinum</i> [44], <i>M. persicum</i> [53], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [47,50]
19	$\alpha$ -Cubebene ST	Train	1491.202	-	-	1348	0.1	<i>M. astracanicum</i> [38], <i>M. crassidens</i> [39], <i>M. deserti</i> [35,40], <i>M. duabense</i> [40], <i>M. parviflorum</i> [47], <i>M. peregrinum</i> [44], <i>M. persicum</i> [45], <i>M. vulgare</i> [8,47]
20	Eugenol AR	Train	1372.624	-	-	1357	0.4	<i>M. aschersonii</i> [34], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [53], <i>M. velutinum</i> [44], <i>M. vulgare</i> [10,12,34,36,47]
21	$\alpha$ -Copaene ST	Train	1475.878	1377	3.3	1377	6.1	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. deserti</i> [35], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [42,47,48], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [53], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [8–13,36,47,50]
22	$\beta$ -Bourbonene ST	Train	1487.976	1385	0.8	1384	1.2	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. deserti</i> [35,52], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [45,53], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [9,10,13,50]
23	NI-2	-	-	-	-	1388	0.1	-
24	$\beta$ -Cubebene ST	Test	1475.610	1390	0.1	1389	0.2	<i>M. aschersonii</i> [34], <i>M. deserti</i> [35], <i>M. peregrinum</i> [43,44], <i>M. parviflorum</i> [42], <i>M. velutinum</i> [44], <i>M. vulgare</i> [12,13,34,47]
25	$\beta$ -Elemene ST	Train	1475.506	1392	0.4	1391	1.0	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. deserti</i> [35,52], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [44], <i>M. persicum</i> [53], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [47,50]
26	Z-Caryophyllene ST	Train	1463.161	1407	0.1	1406	0.2	
27	$\alpha$ -Z-Bergamotene ST	Train	1428.215	1416	0.2	-	-	
28	E-Caryophyllene ST	Validation	1463.161	1422	24.6	1423	23.0	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. astracanicum</i> [38,46], <i>M. crassidens</i> [38,39], <i>M. deserti</i> [35,52], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [45,53], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [7–13,16,17,34,36,47,49,50]
29	$\beta$ -Copaene ST	Test	1459.623	1430	0.4	1430	1.3	<i>M. incanum</i> [50], <i>M. vulgare</i> [50]

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
30	α-E-Bergamotene ST	Train	1428.215	1436	0.1	1435	0.1	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [46], <i>M. crassidens</i> [38], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [44], <i>M. velutinum</i> [44], <i>M. vulgare</i> [47]
31	NI-3	-	-	1445	0.2	1444	0.6	-
32	α-Humulene ST	Validation	1503.999	1454	5.2	1455	5.3	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. astracanicum</i> [38,46], <i>M. crassidens</i> [38,39], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [36,43], <i>M. persicum</i> [45], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [8–10,12,13,15,34,36,47,50]
33	Sesquisabinene ST	Train	1442.740	-	-	1457	0.9	
34	E-β-Farnesene ST	Test	1431.419	1457	1.3	-	-	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. crassidens</i> [39], <i>M. parviflorum</i> [41,42,47], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [45], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [8,10,12,16,17,34,36,47]
35	C16H34 A	Train	1573.436	1462	1.5	1462	0.2	
36	NI-4	-	-	-	Trace	-	0.2	-
37	Z-Muurola-4(14),5-diene ST	Train	1482.433	-	-	1466	0.1	
38	NI-5	-	-	1469	0.1	-	-	-
39	NI-6	-	-	1472	0.1	-	-	-
40	E-Cadina-1(6),4-diene ST	Train	1481.465	-	-	1475	Trace	<i>M. vulgare</i> [15]
41	γ-Muurolole ST	Test	1450.203	1479	0.1	-	-	<i>M. incanum</i> [50], <i>M. peregrinum</i> [43,44], <i>M. parviflorum</i> [42], <i>M. velutinum</i> [44]
42	Germacrene D ST	Test	1450.188	1483	9.6	1487	17.0	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38,39], <i>M. deserti</i> [35,52], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [45,53], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [9–13,15–17,34,36,47,50]
43	E-β-Ionone O	Test	1471.735	1486	0.4	1489	Trace	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. duabense</i> [40], <i>M. incanum</i> [51], <i>M. parviflorum</i> [42], <i>M. peregrinum</i> [43,44], <i>M. thessalum</i> [54], <i>M. vulgare</i> [12]
44	NI-7	-	-	-	-	1489	0.1	-
45	epi-Cubebol OST	Train	1622.285	-	-	1495	0.2	
46	Viridiflorene ST	Validation	1507.447	1497	0.1	-	-	
47	Bicyclogermacrene ST	Validation	1493.697	1498	0.2	1498	0.2	<i>M. astracanicum</i> [38], <i>M. crassidens</i> [38,39], <i>M. deserti</i> [35,52], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [36,43,44], <i>M. persicum</i> [45], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [10,11,17,50]
48	NI-8	-	-	-	-	1499	0.7	-
49	Pentadecane A	Test	1486.884	1500	0.2	1500	Trace	
50	α-Muurolole ST	Train	1465.650	1501	0.1	1501	0.2	<i>M. aschersonii</i> [34], <i>M. deserti</i> [35,52], <i>M. incanum</i> [51], <i>M. peregrinum</i> [43], <i>M. velutinum</i> [44], <i>M. vulgare</i> [12,13,34]
51	Germacrene A ST	Train	1450.188	1508	0.1	1506	0.1	<i>M. incanum</i> [50], <i>M. parviflorum</i> [47,48]

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
52	β-Bisabolene ST	Validation	1425.139	1511	0.2	1507	0.2	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. crassidens</i> [38], <i>M. parviflorum</i> [47], <i>M. peregrinum</i> [44], <i>M. persicum</i> [45], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [11–13,17,34,47,49]
53	γ-Cadinene ST	Test	1450.203	1513	0.2	1515	0.4	<i>M. deserti</i> [52], <i>M. incanum</i> [50], <i>M. parviflorum</i> [47,48], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [53], <i>M. velutinum</i> [44], <i>M. vulgare</i> [7,10,15,47]
54	δ-Cadinene ST	Test	1475.070	1523	4.7	1528	9.7	<i>M. deserti</i> [52], <i>M. incanum</i> [50], <i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [43,44], <i>M. persicum</i> [53], <i>M. velutinum</i> [44], <i>M. vulgare</i> [7,10,15,47]
55	E-Cadina-1,4-diene ST	Train	1471.521	1533	0.1	1533	0.1	<i>M. vulgare</i> [15]
56	α-Cadinene ST	Train	1465.650	-	-	1537	0.1	<i>M. peregrinum</i> [43,44], <i>M. velutinum</i> [44], <i>M. vulgare</i> [47]
57	α-Calacorene ST	Train	1540.123	-	-	1543	0.1	<i>M. deserti</i> [52], <i>M. vulgare</i> [12,15]
58	NI-9	-	-	1555	0.2	1552	0.2	-
59	E-Nerolidol OST	Validation	1567.136	1561	3.5	1564	1.5	<i>M. anisodon</i> [37], <i>M. deserti</i> [52], <i>M. parviflorum</i> [42], <i>M. peregrinum</i> [43,44], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [9,36]
60	NI-10	-	-	-	-	1571	0.1	-
61	NI-11	-	-	1577	0.2	1575	0.9	-
62	NI-12	-	-	-	-	1582	0.3	-
63	Caryophyllene oxide OST	Test	1636.612	1580	1.0	1583	1.8	<i>M. anisodon</i> [37], <i>M. astracanicum</i> [46], <i>M. crassidens</i> [38,39], <i>M. deserti</i> [52], <i>M. duabense</i> [40], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [41,42,47,48], <i>M. peregrinum</i> [36,43], <i>M. persicum</i> [45,53], <i>M. propinquum</i> [41], <i>M. thessalum</i> [54], <i>M. velutinum</i> [44], <i>M. vulgare</i> [8–10,12,36,47,50]
64	NI-13	-	-	-	-	1587	0.1	-
65	Viridiflorol OST	Validation	1573.436	1597	0.1	-	-	<i>M. aschersonii</i> [34], <i>M. astracanicum</i> [38], <i>M. crassidens</i> [38], <i>M. incanum</i> [51], <i>M. parviflorum</i> [47], <i>M. peregrinum</i> [43], <i>M. vulgare</i> [10,12,34,47]
66	Hexadecane A	Train	1594.576	1602	0.1	-	-	<i>M. duabense</i> [40], <i>M. velutinum</i> [44]
67	Humulene epoxide II OST	Train	1626.959	1607	0.2	1607	0.2	<i>M. anisodon</i> [37], <i>M. incanum</i> [51], <i>M. thessalum</i> [54], <i>M. vulgare</i> [10]
68	Muurola-4,10(14)-dien-1-β-ol OST	Train	1605.330	-	-	1627	0.3	
69	NI-14	-	-	1628	0.1	-	-	-
70	4,4-dimethyl-Tetracyclo [6.3.2.0(2,5).0(1,8)] tridecan-9-ol O	Validation	1605.030	-	-	1631	0.2	
71	NI-15	-	-	1632	0.1	-	-	-
72	Caryophylla-4(12), 8(13)-dien-5-α-ol OST	Train	1605.030	1636	0.1	1635	0.3	
73	epi-α-Muurolol (=tau-muurolol) OST	Test	1605.030	1642	0.2	1641	0.6	<i>M. astracanicum</i> [38], <i>M. deserti</i> [35], <i>M. incanum</i> [51], <i>M. parviflorum</i> [42], <i>M. peregrinum</i> [44], <i>M. velutinum</i> [44]
74	α-Muurolol (=Torreyol) OST	Train	1652.148	-	-	1645	0.1	

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
75	$\alpha$ -Cadinol OST	Train	1682.934	1654	0.3	1654	0.9	<i>M. crassidens</i> [38], <i>M. deserti</i> [35,52], <i>M. incanum</i> [50,51], <i>M. parviflorum</i> [42], <i>M. persicum</i> [45], <i>M. vulgare</i> [12,50]
76	NI-16	-	-	1658	0.2	1656	0.2	-
77	NI-17	-	-	1662	0.1	1662	0.1	-
78	E-Calamenen- 10-ol OST	Train	1608.844	-	-	1669	0.1	-
79	NI-18	-	-	1668	0.2	-	-	-
80	NI-19	-	-	-	-	1670	0.2	-
81	8-Heptadecene O	Train	1607.164	-	-	1673	0.2	-
82	1-Tetradecanol O	Train	1702.771	1675	0.1	-	-	-
83	Germacra-4(15),5, 10(14)-trien-1- $\alpha$ -ol OST	Train	1700.003	1682	0.1	1685	0.2	-
84	Heptadecane A	Validation	1726.886	1696	0.3	1696	0.2	<i>M. anisodon</i> [37], <i>M. parviflorum</i> [42,47], <i>M. vulgare</i> [10,47]
85	Pentadecanal O	Validation	1581.928	1710	0.1	1711	0.1	<i>M. anisodon</i> [37]
86	Mint sulfide ST	Train	1778.777	1733	0.1	1736	0.1	-
87	NI-20	-	-	1734	0.1	-	-	-
88	NI-21	-	-	1742	0.1	-	-	-
89	NI-22	-	-	1743	0.4	1744	0.1	-
90	E-3-Octadecene O	Train	1722.391	-	-	1777	0.1	-
91	n-Pentadecanol O	Train	1787.022	1778	0.1	-	-	<i>M. parviflorum</i> [42]
92	NI-23	-	-	-	-	1782	0.1	-
93	Octadecane A	Validation	1950.093	1796	0.1	-	-	<i>M. parviflorum</i> [47], <i>M. peregrinum</i> [43], <i>M. vulgare</i> [47]
94	NI-24	-	-	1819	0.1	-	-	-
95	6,10,14-trimethyl- 2-Pentadecanone O	Train	1915.818	1844	4.8	1842	0.5	<i>M. peregrinum</i> [44], <i>M. velutinum</i> [44], <i>M. vulgare</i> [10]
96	NI-25	-	-	1849	0.1	-	-	-
97	NI-26	-	-	1853	0.2	-	-	-
98	NI-27	-	-	1888	0.1	-	-	-
99	NI-28	-	-	1891	0.8	1891	0.1	-
100	Nonadecane A	Test	1869.346	1897	0.2	1897	0.2	<i>M. duabense</i> [40], <i>M. parviflorum</i> [47], <i>M. peregrinum</i> [43], <i>M. vulgare</i> [10,15,47]
101	NI-29	-	-	1904	0.1	1906	Trace	-
102	5E,9E-Farnesyl acetone OST	Train	1956.289	1916	0.3	1915	Trace	<i>M. thessalum</i> [54], <i>M. vulgare</i> [15]
103	NI-30	-	-	1918	Trace	1917	Trace	-
104	NI-31	-	-	1924	0.1	-	-	-
105	NI-32	-	-	-	-	1926	Trace	-
106	NI-33	-	-	1925	0.1	-	-	-
107	NI-34	-	-	1929	0.1	-	-	-
108	NI	-	-	1938	0.1	1940	Trace	-
109	Hexadecanoic acid O	Validation	1995.491	1960	3.9	-	-	<i>M. parviflorum</i> [42], <i>M. peregrinum</i> [36], <i>M. vulgare</i> [36,47]
110	NI-35	-	-	1973	0.1	1974	Trace	-
111	Eicosane A	Train	2034.560	1997	0.2	1994	0.1	<i>M. parviflorum</i> [48]
112	NI-36	-	-	2001	0.1	-	-	-

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
113	E,E-Geranyl linalool OD	Train	2028.645	2027	1.6	-	-	<i>M. aschersonii</i> [34], <i>M. parviflorum</i> [42], <i>M. vulgare</i> [12,34]
114	3,7,11,15-tetramethyl-(E,E)-1,6,10,14-Hexadecatetraen-3-ol OD			-	-	2028	0.9	
115	Manool OD	Train	2064.196	2057	0.3	-	-	
116	NI-37	-	-	2061	0.1	-	-	-
117	NI-38	-	-	2067	0.1	-	-	-
118	NI-39	-	-	2084	0.1	-	-	-
119	NI-40	-	-	2096	0.1	-	-	-
120	Heneicosane A	Train	2120.284	2101	1.6	2100	1.3	<i>M. parviflorum</i> [42,47], <i>M. peregrinum</i> [43], <i>M. propinquum</i> [41], <i>M. vulgare</i> [10]
121	NI-41	-	-	2108	0.3	2105	0.2	-
122	NI-42	-	-	2112	0.2	2110	0.3	-
123	Phytol OD	Test	2124.818	2116	1.4	2113	0.4	<i>M. anisodon</i> [37], <i>M. incanum</i> [51], <i>M. parviflorum</i> [41,42], <i>M. peregrinum</i> [36], <i>M. propinquum</i> [41], <i>M. vulgare</i> [10,15]
124	NI-43	-	-	2131	0.2	-	-	-
125	NI-44	-	-	2143	0.2	2143	0.1	-
126	NI-45	-	-	2147	0.2	-	-	-
127	NI-46	-	-	2164	0.2	2160	0.2	-
128	NI-47	-	-	2167	0.1	2172	0.3	-
129	NI-48	-	-	2175	0.6	2176	0.4	-
130	NI-49	-	-	2181	0.9	2179	0.2	-
131	NI-50	-	-	2183	0.4	-	-	-
132	NI-51	-	-	2198	2.4	2195	2.4	-
133	Docosane A	Validation	2194.421	2205	0.9	2198	0.6	<i>M. crassidens</i> [39], <i>M. parviflorum</i> [47]
134	NI-52	-	-	-	-	2201	0.1	-
135	NI-53	-	-	-	-	2209	0.3	-
136	NI-54	-	-	2215	0.3	-	-	-
137	NI-55	-	-	2225	0.3	2221	0.1	-
138	NI-56	-	-	2246	0.3	-	-	-
139	NI-57	-	-	2258	0.2	2253	0.3	-
140	NI-58	-	-	2270	0.1	2265	0.1	-
141	NI-59	-	-	2277	0.2	2274	0.2	-
142	NI-60	-	-	2293	3.8	2288	1.7	-
143	Tricontane A	Train	2381.642	2305	3.6	2302	2.6	
144	NI-61	-	-	2309	0.2	2305	0.2	-
145	NI-62	-	-	2344	0.2	2341	0.3	-
146	NI-63	-	-	2380	0.1	2377	0.1	-
147	NI-64	-	-	2383	0.1	2382	0.1	-
148	Tetracosane A	Train	2493.491	2401	0.3	2395	0.2	<i>M. deserti</i> [52], <i>M. parviflorum</i> [41,42,47], <i>M. propinquum</i> [41]
149	NI-65	-	-	2454	0.4	2447	0.2	-
150	NI-66	-	-	2488	0.2	2483	0.2	-
151	Pentacosane A	Test	2510.087	2503	0.8	2497	0.6	<i>M. anisodon</i> [37], <i>M. parviflorum</i> [42,47]

Table 1. Cont.

No	Compound/Class	Cycle	RI <sub>pred.</sub>	2019		2020		Reference
				RI <sup>a</sup>	%	RI <sup>a</sup>	%	
152	Heptacosane A	Train	2730.537	2702	0.6	2696	0.5	<i>M. anisodon</i> [37], <i>M. aschersonii</i> [34], <i>M. incanum</i> [51], <i>M. parviflorum</i> [42], <i>M. vulgare</i> [34]
153	NI-67	-	-	-	-	2766	0.1	-
154	Octacosane A			2801	0.1	2791	Trace	<i>M. crassidens</i> [39], <i>M. parviflorum</i> [42], <i>M. persicum</i> [45]
155	Squalene T	Train	2870.673	2835	0.1	2823	0.1	
156	NI-68	-	-	2868	0.1	2855	0.1	-
157	Nonacosane A	Test	2930.732	2905	0.7	2892	0.6	<i>M. anisodon</i> [37], <i>M. crassidens</i> [39], <i>M. persicum</i> [45]
158	Untriacontane A	Validation	3150.673	3105	0.4	3095	0.3	
159	NI-69	-	-	-	-	3212	0.1	-
160	Trtriacontane A	Train	3319.753	3300	0.1	3301	Trace	
	Oxygenated monoterpenes OMN				0.1		0.2	
	Sesquiterpene hydrocarbons ST				52.0		67.8	
	Oxygenated sesquiterpenes OST				5.8		6.2	
	Oxygenated diterpenes OD				3.3		1.3	
	Triterpene T				0.1		0.1	
	Aromatics AR				0.4		0.4	
	Alkanes A				11.7		7.4	
	Other O				9.9		3.4	
	NI				16.7		12.5	
	Total				100		99.3	

RI<sub>pred.</sub>—BRT calculated retention index; RI<sup>a</sup>—retention index experimentally obtained on a HP-5MS column; Other—aliphatic hydrocarbons, aliphatic aldehydes and alcohols, aliphatic acids, their esters and aldehydes, aromatic esters with aliphatic acids, alkyl-aromatic alcohols, or aryl esters of aromatic acids; NI—not identified compound.

The predicted RIs are shown in Table 1, and confirm the good quality of the constructed BRT model by showing the relationship between the predicted and experimental RI values. Graphical comparison between experimentally obtained RIs of *M. vulgare* volatiles composition (RI<sup>a</sup>), the retention time indices found in NIST database (RI<sup>b</sup>) and the retention time indices predicted by the two BRT models (RI<sub>pred.</sub>) are presented in Figure 1.

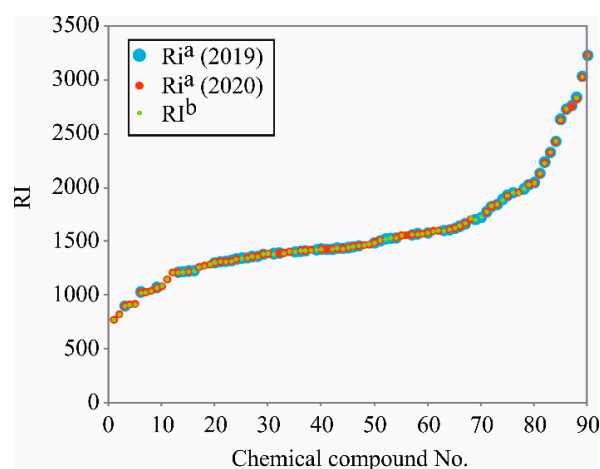
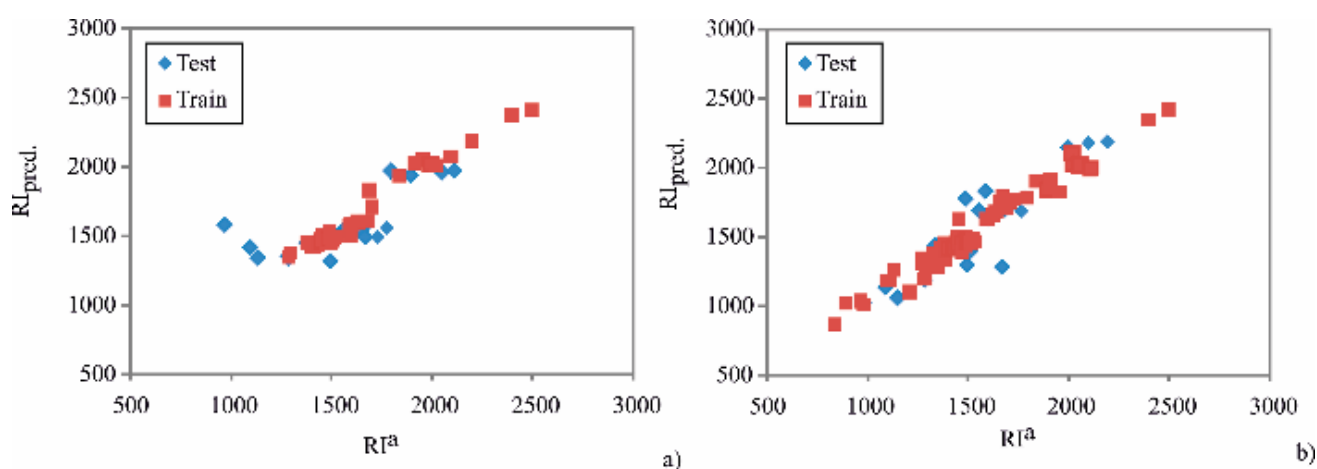


Figure 1. Retention indices (RIs) of the *M. vulgare* volatiles composition, from experimentally obtained GC–MS data on a HP-5MS column (RI<sup>a</sup>) and NIST database (RI<sup>b</sup>).



In order to calculate the molecular descriptors, the PaDel-descriptor was used in this investigation. Due to a great amount of data that was obtained, it was required to select the most important set of descriptors to build the adequate model which would be able to predict the RIs [55]. The factor analysis was done before the GA calculation, and only ca. 320 uncorrelated descriptors remained in the GA calculation [56,57]. The seven most significant molecular descriptors chosen by GA are as follows: four 2D autocorrelation descriptors (AATSC4e, AATSC2p, GATS6v and MATS5v), two Barysz matrix descriptors (VR1\_Dzs and SM1\_Dzv) and Vertex adjacency information (magnitude) descriptor (VAdjMat).

The predicted RIs and molecular descriptors are presented in Table 1. Seven molecular descriptors were utilized for predictions of RIs in the two BRT models. The predicted RIs are presented in Figure 2, and visually confirm the adequate prediction capabilities of the constructed BRT by showing the relationship between the predicted and experimental retention values.



**Figure 2.** Comparison of experimentally obtained RIs on a HP-5MS column (RIa) with BRT pre-dicted values (RIpred.) in 2019 (a) and 2020 (b).

Separation of compounds in GC–MS and their RIs is linked to their affinity towards mobile and stationary phases. Affinity and solubility of separated molecules directly depend on their chemical structure and physico-chemical properties, which could be expressed by molecular descriptors. According to Pearson’s correlation coefficients, there was a rather poor correlation between all 3D autocorrelation descriptors (Table 2). Therefore, utilized molecular descriptors were appropriate to predict RIs of compounds in *M. vulgare* by the two multivariate BRT models [58].

**Table 2.** The correlation coefficient matrix for the selected descriptors by GA.

	AATSC2p	MATS5v	GATS6v	SM1_Dzv	VR1_Dzs	VAdjMat
AATSC4e	0.031	−0.138	−0.135	0.030	0.205	0.224
AATSC2p		−0.265	−0.131	0.036	−0.231	−0.010
MATS5v			0.212	−0.008	0.066	0.109
GATS6v				0.072	0.131	0.214
SM1_Dzv					0.058	0.084
VR1_Dzs						2.339

Detailed explanations about the descriptors were found in the Handbook of Molecular Descriptors [59]. These descriptors encode different aspects of the molecular structure and were applied to develop the QSRR model. According to Pearson’s correlation, there was a rather poor correlation between all molecular descriptors. Hence, utilized descriptors were appropriate to predict RIs of compounds isolated from *M. vulgare* volatiles by the two multivariate BRT models. The calibration and predictive capability of a QSRR model

should be tested through model validation. The most widely used squared correlation coefficient ( $r^2$ ) can provide a reliable indication of the fit of the model, thus, it was employed to validate the calibration capability of a QSRR model.

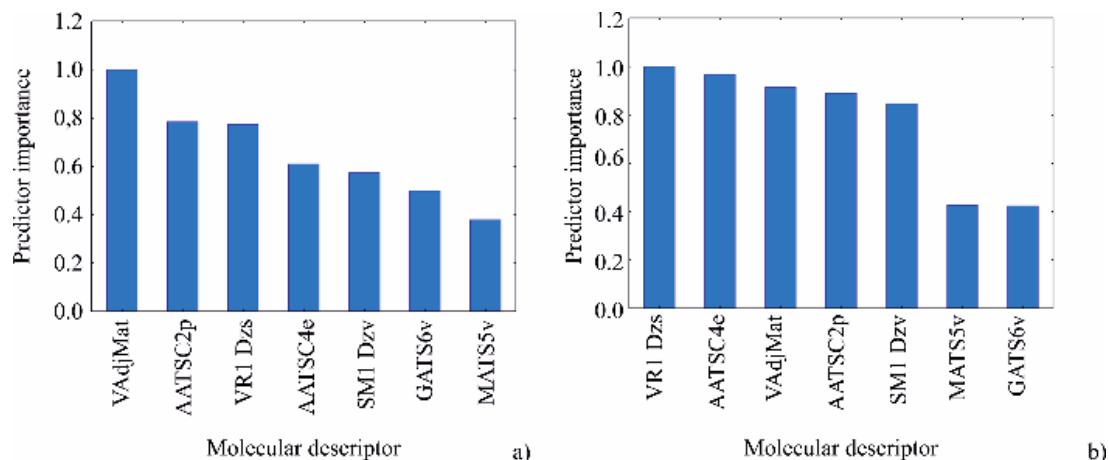
In order to explore the nonlinear relationship between RIs and the descriptors selected by GA, BRT technique was used to build the two predictive models. Two BRT models were constructed to predict the retention time of compounds isolated from *M. vulgare* volatiles, respectively. The coefficients of determination were 0.956 and 0.964, respectively, indicating that these models could be used for prediction of RIs, due to low prediction error and high  $r^2$ . The tests of the two BRT models fit (2019 and 2020) are shown in Table 3, with the higher  $r^2$  values and lower  $\chi^2$ , MBE, RMSE, and MPE values showing the better fit to the experimental results [60,61].

**Table 3.** The “goodness of fit” tests for the developed BRT model.

Boosted Tree Model	$\chi^2$	RMSE	MBE	MPE	$r^2$
2019	4455.272	66.160	−13.063	3.285	0.956
2020	3975.751	62.581	−7.698	3.241	0.964

$\chi^2$ —reduced chi-square, MBE—mean bias error, RMSE—root mean square error, MPE—mean percentage error.

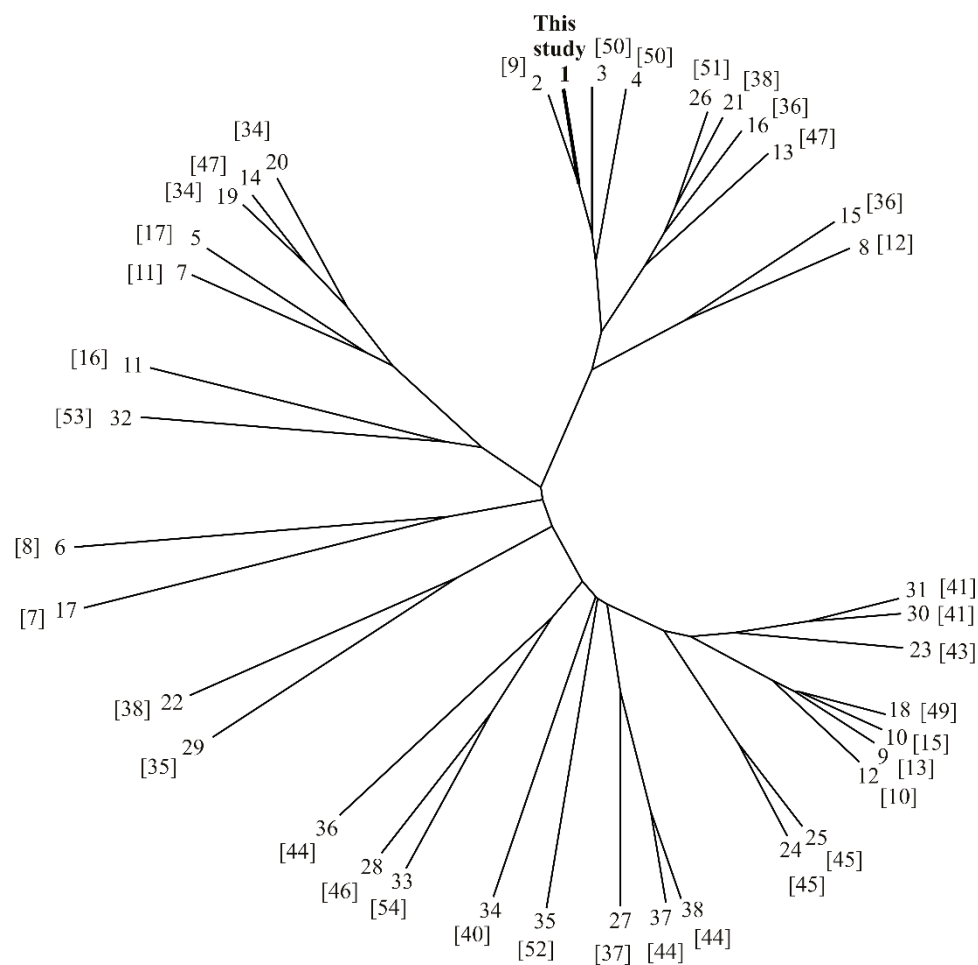
Obtained results reveal the reliability of the BRT models for predicting the RIs of compounds in *M. vulgare* volatiles obtained by GC–MS analysis. The influence of the seven most important molecular descriptors, identified by using genetic algorithm on the RIs was studied in this section. According to the Figure 3, VAdjMat was the most important molecular descriptor for chemical compounds’ RIs calculation in *M. vulgare*, during 2019, while VR1 Dzs was the most important variable during 2020.



**Figure 3.** Predictor importance of the molecular descriptors on RI in 2019 (a) and 2020 (b).

### 3. Discussion

According to the cluster analysis (unrooted cluster tree) with 37 samples of *Marrubium* sp. volatiles from literature and average values from this study (Figure 4), it could be said that there are several chemotypes, but only *E*-caryophyllene chemotype [9,12,36,38,47,50,51] is clearly segregated. However, these are samples of *M. vulgare*, *M. incanum*, *M. parviflorum*, *M. peregrinum*, and *M. crassidens* grown in Serbia, Poland, Slovakia, Egypt, and Iran. This indicated that genus *Marrubium* is very diverse in the case of volatiles composition.



**Figure 4.** Unrooted cluster tree for different *Marrubium* samples.

In addition, *E*-caryophyllene is a compound which is occurring in all samples (except *M. vulgare* from Eastern Algeria [15]), but in *E*-caryophyllene chemotype its content ranges between 15.6% and 45.8%. Other chemotypes can be classified as  $\beta$ -bisabolene (13.1–28.3%) [11,17,34,47],  $\alpha$ -pinene (21.5–28.9%) [16,53],  $\beta$ -farnesene (20.2–24.2%) [37,44], *E*-caryophyllene + caryophyllene oxide chemotype [44,46,54], and diverse (unclassified) chemotypes [7,8,10,13,15,35,38,40,41,43,45,49,52].

Occurring polymorphism could be a consequence of adaptation to grow in different environments [19,62], especially ecological conditions such as humidity, temperature and altitude [22] as well as hybridization [20] strongly affected the chemotypes, as well as biotic and abiotic stresses (including temperature, light, water, salt, and oxidative stresses) [63].

Detected compounds in *M. vulgare* volatiles obtained by GC–MS analysis were used for QSRR analysis. The following seven molecular descriptors that characterize the RIs of obtained compounds were suggested by the genetic algorithm. Selected molecular descriptors were not autocorrelated which was suggested by a correlation coefficient matrix; thus, descriptors were suitable for QSRR analysis. These descriptors were utilized as inputs for the boosted trees regression models, for estimating the RIs using a set of GC–MS data from a series of 160 compounds found in *M. vulgare* volatiles. Statistical models that quantify the relation between the structure of molecules and their chromatographic RIs were represented by the quantitative structure retention relationship (QSRR) model [58,64]. Numerous publications related to the QSRR analysis in plants from Lamiaceae family could be found in the literature: *Thymus vulgaris* [65], *T. serpyllum* [66], *Satureja kitaibelii* [55], *Salvia officinalis* [67], as well as *Stachys* sp. [68]. The connection between the molecular

descriptors and the retention time can be established by artificial neural network, machine learning algorithms [69–73] or by boosted trees regression (BRT) [74].

## 4. Materials and Methods

### 4.1. Plant Material

*M. vulgare* was grown in the Institute of Field and Vegetable Crops (IFVCNS) collection garden of medicinal and aromatic plants in Bački Petrovac (45°21' N; 19°35' E), confirmed by Milica Rat, PhD, and deposited at the Herbarium BUNS (Department of Biology and Ecology, Faculty of Sciences, University of Novi Sad) under the voucher number 2-1409. After seed maturation (August 2018), it was collected and sown in field conditions in September 2018 and 2019. The experimental plot was 10 m long and 5 m wide, with a 70 cm spacing between rows. From the seven rows, only three central rows were used for collecting plant material to avoid edge effects (one row one sample).

### 4.2. Volatiles Isolation and Analysis

Flowering aerial parts of *M. vulgare* (*Marrubii herba*) were collected during July 2019 and 2020, dried in a solar dryer at a temperature of 40° with air circulation. After drying, plant material was fragmented, and volatiles was extracted by Clevenger apparatus. Taking in account that *M. vulgare* produces trace amounts of essential oil, it was trapped in n-hexane. This process was performed in tree repetition for both years, as well as analysis of chemical composition.

GC–MS analysis was carried out using an Agilent 7890A apparatus equipped with a 5975 C MSD, FID and a nonpolar HP-5MS fused-silica capillary column (30 m × 0.25 mm, film thickness 0.25 µm). The carrier gas was helium, and its inlet pressure was 19.6 psi and linear velocity of 1 mL/min at 210 °C. The injector temperature was 250 °C, injection volume was 1 µL, split ratio, 10:1. Mass detection was carried out under source temperature conditions of 230 °C and interface temperature of 315 °C. The EI mode was set at electron energy, 70 eV with mass scan range of 40–600 amu. Temperature was programmed from 60 to 300 °C at a rate of 3 °C/min. The components were identified based on their linear retention index relative to C8–C32 n-alkanes, compared with data reported in the literature (Adams4 and NIST11 databases). The relative percentage of the oil constituents was expressed as percentages by FID peak area normalization.

### 4.3. QSRR Analysis

PaDel-descriptor software was used to calculate specified molecular descriptors [75], as described in our previous investigation [66]. Factor analysis and genetic algorithm (GA) were used to determine the most important descriptors [76,77]. The relationship between the chosen descriptors was examined and collinear descriptors were excluded. Statistica 10 software was used for the statistical investigation of the data [78].

### 4.4. BRT Model

In order to relate and to predict categorical or continuous dependent variables the BRT model could be used [79,80], as it does not require transformation or outliers [81]. The BRT method calculation is connected to the boosting methods enforced to regression trees [82]. The main idea is to calculate a set of simple trees, where each successive tree is built for the prediction residuals of the preceding tree [83]. This method builds binary trees such as partition the data into two samples at each split node [78].

The decision trees are combined through a cross-validation or “boosting” procedure in order to acquire the single computational model [84]. BRT modeling consists of the following steps: (a) an initial regression tree is defined according to a minimum loss function; (b) the other trees are engaged in the iterative process in which several new regression trees were developed and selected to the subsequent according to the StatSoft Statistica’s recommendation—the least square error (LSE); (c) step (b) is repeated until a stopping criterion is reached (for instance, the value of LSE).

In this study, several regularization parameters were set in order to optimize the fit between experimental results and computing model: the number of trees (between 100 and 1000), learning rate (between 0.0005 and 0.1), random test data proportion (0.1–0.9) and subsample proportion (0.1–0.9). According to Statistica's manual, prior to computation, a subsample of data is created, according to random test data proportion of the cases, and these data are treated as test samples used to evaluate the appropriate fit of the model. The remaining set of data is used for the analyses via stochastic gradient boosting (for the selection of samples for consecutive boosting steps).

#### 4.5. Cluster Analysis

The cluster analysis (CA) was used to evaluate intra- and interpopulation variability and differentiation of volatile constituents of *Marrubium* samples collected in different locations and/or taken from literature reports. The phylogenetic tree diagram for *Marrubium* samples was calculated and plotted using R software 4.0.3 (64-bit version). The R package "ape" (Analysis of Phylogenetics and Evolution) was used for calculation, applied as a graphical tool to represent the arrangements of similar volatiles concentration (evaluated in the cluster analysis). The obtained experimental results were collected in the matrix, after which the hierarchical cluster analysis was performed. The distance matrix was determined using the Euclidean method, while the cluster analysis was performed using the "complete" method.

## 5. Conclusions

The main components in *M. vulgare* volatiles were *E*-caryophyllene (24.6% and 23.0%), followed by germacrene D (9.6% and 17.0%),  $\alpha$ -humulene (5.2% and 5.3%) and  $\alpha$ -copaene (3.3% and 6.1%) in 2019 and 2020, respectively. All these compounds are from sesquiterpene hydrocarbons class, which was dominant in both years of the investigation, 52.0% in 2019 and 67.8% in 2020.

The results demonstrated that the boosted trees regression models were adequate in predicting the RIs of the compounds in *M. vulgare* volatiles obtained by GC–MS analysis on a HP-5MS column. The coefficients of determination were 0.956 and 0.964 (for compounds found in *M. vulgare* volatiles, during the years 2019 and 2020, respectively), which is a good indication that these models could be used as a fast mathematical tool for prediction of RIs, due to low prediction error and moderately high  $r^2$ . Suitable models with high statistical quality and low prediction errors were derived, and it could be further used for estimation of RIs of newly detected compounds.

According to the unrooted cluster tree with 37 samples of *Marrubium* sp. volatiles from literature and average values from this study, it could be said that there are several chemotypes: *E*-caryophyllene,  $\beta$ -bisabolene,  $\alpha$ -pinene,  $\beta$ -farnesene, *E*-caryophyllene + caryophyllene oxide chemotype, and diverse (unclassified) chemotypes. However, occurring polymorphism could be a consequence of adaptation to grow in different environments, especially ecological conditions such as humidity, temperature and altitude, as well as hybridization which strongly affected the chemotypes. Further research on *M. vulgare* chemotypes needs to be focused on genetic markers, because evaluation of genetic diversity has key importance in improving the quality of raw material used for industrial purposes.

**Author Contributions:** Conceptualization, M.A. and J.O.; methodology, S.I. and K.S.; software, L.P.; validation, S.I., K.S. and T.Z.; formal analysis, L.P.; investigation, M.A.; resources, M.A. and V.S.; data curation, L.P.; writing—original draft preparation, M.A.; writing—review and editing, T.Z. and V.S.; visualization, L.P.; supervision, J.O.; project administration, T.Z.; funding acquisition, M.A., T.Z., J.O. and V.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, grant number: 451-03-9/2021-14/200032.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Vele Tešević, Marina Todosijević, Jovana Stanković Jeremić, Jovana Ljujić, and Mirjana Cvetković for participating in this research.

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

## References

1. Yabrir, B. Essential oil of *Marrubium vulgare*: Chemical composition and biological activities. A review. *Nat. Prod. Sci.* **2019**, *25*, 81–91. [\[CrossRef\]](#)
2. Lippai, A.; Smith, P.A.; Price, T.V.; Weiss, J.; Lloyd, C.J. Effects of temperature and water potential on germination of horehound (*Marrubium vulgare*) seeds from two Australian Localities. *Weed Sci.* **1996**, *44*, 91–99. [\[CrossRef\]](#)
3. Nedjimi, B.; Souissi, Z.E.; Guit, B.; Daoud, Y. Differential effects of soluble salts on seed germination of *Marrubium vulgare* L. *J. Appl. Res. Med. Aromat. Plants* **2020**, *17*, 100250. [\[CrossRef\]](#)
4. Dmitruk, M.; Haratym, W. Morphological differentiation of non-glandular and glandular trichomes on *Marrubium vulgare* L. *Mod. Phytomorphol.* **2014**, *6*, 85.
5. Aćimović, M.; Jeremić, K.; Salaj, N.; Gavarić, N.; Kiprovska, B.; Sikora, V.; Zeremski, T. *Marrubium vulgare* L.: A phytochemical and pharmacological overview. *Molecules* **2020**, *25*, 2898. [\[CrossRef\]](#)
6. Lodhi, S.; Vadnere, G.P.; Sharma, V.K.; Usman, M.R. *Marrubium vulgare* L.: A review on phytochemical and pharmacological aspects. *J. Intercult. Ethnopharmacol.* **2017**, *6*, 429–452. [\[CrossRef\]](#)
7. Mahmoud, A.A.; Gendy, A.S.H.; Said-Al Ahl, H.A.H.; Grulova, D.; Astatkie, T.; Abdelrazik, T.M. Impact of harvest time and water stress on the growth and essential oil components of horehound (*Marrubium vulgare*). *Sci. Hortic.* **2018**, *232*, 139–144. [\[CrossRef\]](#)
8. Said-Al Ahl, H.A.H.; Gendy, A.S.H.; Mahmoud, A.A.; Mohamed, H.F.Y. Essential oil composition of *Marrubium vulgare* L. cultivated in Egypt. *Int. J. Plant. Sci. Ecol.* **2015**, *1*, 138–141.
9. El-Leithy, A.S.; El-Hanafy, S.H.; Omer, E.A.; El-Sayed, A.A.A. Effect of nitrogen and potassium biofertilization on growth, yield and essential oil production of the white horehound, *Marrubium vulgare* L. plant. *J. Hortic. Sci. Ornament. Plants* **2013**, *5*, 46–59.
10. Weel, K.G.C.; Venskutonis, P.R.; Pukalskas, A.; Gruzdiene, D.; Linssen, J.P.H. Antioxidant activity of horehound (*Marrubium vulgare* L.) grown in Lithuania. *Lipid* **1999**, *101*, 395–400. [\[CrossRef\]](#)
11. Morteza-Semnani, K.; Saeedi, M.; Babanezhad, E. The essential oil composition of *Marrubium vulgare* L. from Iran. *J. Essent Oil Res.* **2008**, *20*, 488–489. [\[CrossRef\]](#)
12. Golparvar, A.R.; Hadipanah, A.; Mehrabi, A.M.; Armin, A. Essential oil composition of *Marrubium vulgare* L. from Iran. *Flavour Herb. Drug* **2015**, *6*, 1–5.
13. Zarai, Z.; Kadri, A.; Chobba, I.B.; Mansour, R.B.; Bekir, A.; Mejdoub, H.; Gharsallah, N. The *in-vitro* evaluation of antibacterial, antifungal and cytotoxic properties of *Marrubium vulgare* L. essential oil grown in Tunisia. *Lipids Hum. Health Dis.* **2011**, *10*, 161. [\[CrossRef\]](#)
14. Kadri, A.; Zarai, Z.; Bekir, A.; Gharsallah, N.; Damak, M.; Gdoura, R. Chemical composition and antioxidant activity of *Marrubium vulgare* L. essential oil from Tunisia. *Afr. J. Biotechnol.* **2011**, *10*, 3908–3914.
15. Abadi, A.; Hassani, A. Essential oil composition and antioxidant activity of *Marrubium vulgare* L. growing wild in Eastern Algeria. *Int. Lett. Chem. Phys. Astron.* **2013**, *9*, 17–24. [\[CrossRef\]](#)
16. Bayir, B.; Gunduz, H.; Usta, T.; Sahin, E.; Ozdemir, Z.; Kayir, O.; Sen, O.; Aksit, H.; Elmastas, M.; Erenler, R. Chemical composition of essential oil from *Marrubium vulgare* L. leaves. *J. New Results Sci.* **2014**, *6*, 44–50.
17. El-Hallous, E.I.; Alsanie, W.F.; Ismail, I.A.; Dessoky, E.S. Utilization of *Marrubium vulgare* extract as a therapeutic to hepatic damage induced by Carbon Tetrachloride in rats. *Int J. Pharm. Res. Allied Sci.* **2018**, *7*, 168–178.
18. Yabrir, B. Chemical composition and biological activities of some *Marrubium* species essential oil: A review. *Chem. J. Mold.* **2018**, *13*, 8–23. [\[CrossRef\]](#)
19. Salehi, N.; Kharazian, N.; Shiran, B. Genetic diversity of *Marrubium* species from Zagros Region (Iran), using inter simple sequence repeat molecular marker. *J. Sci. Islamic Repub. Iran.* **2018**, *29*, 7–19.
20. Kharazian, N.; Hashemi, M. Chemotaxonomy and morphological studies in five *Marrubium* L. species in Iran. *Iran. J. Sci. Technol. Trans. Sci.* **2017**, *41*, 17–31. [\[CrossRef\]](#)
21. Akgul, G.; Ketenoglu, O.; Pinar, N.M.; Kurt, L. Pollen and seed morphology of the genus *Marrubium* (Lamiaceae) in Turkey. *Ann. Bot. Fenn.* **2008**, *45*, 1–10. [\[CrossRef\]](#)
22. Ahvazi, M.; Jamzad, Z.; Balali, G.R.; Saeidi, H. Trichome micro-morphology in *Marrubium* L. (Lamiaceae) in Iran and the role of environmental factors on their variation. *Iran. J. Bot.* **2016**, *22*, 39–58. [\[CrossRef\]](#)
23. Ahvazi, M.; Balali, G.R.; Jamzad, Z.; Saeidi, H. A taxonomical, morphological and pharmacological review of *Marrubium vulgare* L., an old medicinal plant in Iran. *J. Med. Plants* **2018**, *17*, 7–24.
24. Martin, E.; Cetin, O.; Akgul, G.; Hilal, A.Y. New chromosome numbers in the genus *Marrubium* (horehound) from Turkey. *Biol. Divers. Conserv.* **2011**, *4*, 185–188.
25. Buyukkartal, H.N.; Colgecen, H.; Akgul, G. Comparative leaf, stem and root anatomies of taxa *Marrubium bourgaei* and *Marrubium heterodon* (Lamiaceae). *Aust. J. Crop Sci.* **2016**, *10*, 1516–1522. [\[CrossRef\]](#)
26. Marzouk, R.I.; El-Darier, S.M.; Nour, I.H.; Kamal, S.A. Numerical taxonomic study of *Marrubium* L. (Lamiaceae) in Egypt. *Catrina* **2015**, *13*, 25–35.

27. Sgarbossa, J.; Schmidt, D.; Schwerz, F.; Schwerz, L.; Prochnow, D.; Caron, B.O. Effect of season and irrigation on the chemical composition of *Aloysia triphylla* essential oil. *Rev. Ceres* **2019**, *66*, 85–93. [[CrossRef](#)]
28. Castelo, A.V.M.; Del Menezzi, C.H.S.; Resck, I.S. Seasonal variation in the yield and the chemical composition of essential oils from two Brazilian native arbustive species. *J. Appl. Sci.* **2012**, *12*, 753–760. [[CrossRef](#)]
29. Yavari, A.; Nazeri, V.; Sefidkon, F.; Hassani, M.E. Influence of some environmental factors on the essential oil variability of *Thymus migricus*. *Nat. Prod. Comm* **2010**, *5*, 943–948. [[CrossRef](#)]
30. Zouari-Bouassida, K.; Trigui, M.; Makni, S.; Jlaiel, L.; Tounsi, S. Seasonal variation in essential oils composition and the biological and pharmaceutical protective effects of *Mentha longifolia* leaves grown in Tunisia. *Biomed. Res. Int.* **2018**, 7856517. [[CrossRef](#)]
31. Gosztola, B.; Sarosi, S.; Nemeth, E. Variability of the essential oil content and composition of chamomile (*Matricaria recutita*) affected by weather conditions. *Nat. Prod. Commun.* **2010**, *5*, 465–470. [[CrossRef](#)]
32. Acimovic, M.; Cvetkovic, M.; Stankovic, J. Effect of weather conditions, location and fertilization on coriander fruit essential oil quality. *Essent Oil Bear Plants* **2016**, *19*, 1208–1215.
33. Aziz, E.E.; Badawy, E.M.; Zheljzkov, V.D.; Nicola, S.M.; Fouad, H. Yield and chemical composition of essential oil of *Achillea millefolium* L. as affected by harvest time. *Egypt J. Chem.* **2019**, *62*, 533–540. [[CrossRef](#)]
34. Hamdaoui, B.; Wannas, W.A.; Marrakchi, M.; Brahim, N.B.; Marzouk, B. Essential oil composition of two Tunisian horehound species: *Marrubium vulgare* L. and *Marrubium aschersonii* Magnus. *J. Essent Oil Bear Plants* **2013**, *16*, 608–612. [[CrossRef](#)]
35. Laouer, H.; Yabrir, B.; Djeridane, A.; Youfsi, M.; Beldovini, N.; Lamamra, M. Composition, antioxidant and antimicrobial activities of the essential oil of *Marrubium deserti*. *Nat. Prod. Commun.* **2009**, *4*, 1133–1138. [[CrossRef](#)] [[PubMed](#)]
36. Nagy, M.; Svajdlenka, E. Comparison of essential oils from *Marrubium vulgare* L. and *M. peregrinum* L. *J. Essent Oil Res.* **1998**, *10*, 585–587. [[CrossRef](#)]
37. Kirimer, N.; Kurkcuoglu, M.; Akgul, G.; Baser, K.H.H.; Mahmoud, A.A. Composition of the essential oil of *Marrubium anisodon* C. Koch of Turkish origin. *Rec Nat. Prod.* **2015**, *9*, 234–236.
38. Teimori, M.; Khavari-Nejad, R.A.; Yassa, N.; Nejadساتari, T. Analysis of the essential oil of *Marrubium crassidens* Bioos. and *M. astracanicum* Jacq. *J. Appl. Sci.* **2008**, *8*, 1793–1795. [[CrossRef](#)]
39. Hamedeyazdan, S.; Asnaashari, S.; Fathiazad, F. Characterization of non-terpenoids in *Marrubium crassidens* Bioss. essential oil. *Adv. Pharm. Bull.* **2013**, *3*, 429–432.
40. Golmakani, H.; Rabbani Nasab, H.; Sharifan, M.; Kamali, H.; Yadollahi, A. The essential oil composition and antibacterial activity of *Marrubium dubanese* Murata from North Khorassan Province, Iran. *J. Essent Oil Bear Plants* **2016**, *19*, 963–971. [[CrossRef](#)]
41. Hamedeyazdan, S.; Zarei, M.; Salem, A.; Asnaashari, S.; Fathiazad, F. Essential oil of two Iranian horehound species: *Marrubium propinquum* and *Marrubium parviflorum*. *Pharm. Sci.* **2017**, *23*, 143–149. [[CrossRef](#)]
42. Bal, Y.; Kaban, S.; Kirimer, N.; Baser, K.H.C. Composition of the essential oil of *Marrubium parviflorum* Fisch. et Mey. subsp. *oligodon* (Boiss.) Seybold. *J. Essent Oil Res.* **1999**, *11*, 300–302. [[CrossRef](#)]
43. Kaurinovic, B.; Vlaisavljevic, S.; Popovic, M.; Vastag, D.; Djurendic-Brensel, M. Antioxidant properties of *Marrubium peregrinum* L. (Lamiaceae) essential oil. *Molecules* **2010**, *15*, 5943–5955. [[CrossRef](#)]
44. Lazari, D.; Skaltsa, H.; Constantinidis, T. Essential oils of *Marrubium velutinum* Sm. and *Marrubium peregrinum* L., growing wild in Greece. *Flavour Fragr. J.* **1999**, *14*, 290–292. [[CrossRef](#)]
45. Hamedeyazdan, S.; Fathiazad, F.; Asnaashari, S. Chemical composition of the essential oil from *Marrubium persicum* C.A. Mey. (Lamiaceae). *Pharm. Sci.* **2013**, *19*, 35–38.
46. Nik, B.Z.; Mirza, M. Composition of the essential oil of *Marrubium astracanicum* Jacq. *J. Essent Oil Res.* **2003**, *15*, 342–343. [[CrossRef](#)]
47. Khanavi, M.; Ghasemian, L.; Motlagh, E.H.; Hadjikhondi, A.; Shafiee, A. Chemical composition of the essential oils of *Marrubium parviflorum* Fisch. & C.A. Mey. and *Marrubium vulgare* L. from Iran. *Flavour Fragr. J.* **2005**, *20*, 324–326.
48. Sarikurcu, C.; Ozer, M.S.; Calli, N.; Popović-Djordjević, J. Essential oil composition and antioxidant activity of endemic *Marrubium parviflorum* subsp. *oligodon*. *Ind. Crops Prod.* **2018**, *119*, 209–213. [[CrossRef](#)]
49. Miloudi, K.; Hamimed, A.; Benmimoun, Y.; Bellebna, Y.; Taibi, A.; Tilmatine, A. Intensification of essential oil extraction of the *Marrubium vulgare* using pulsed electric field. *J. Essent Oil Bear Plants* **2018**, *21*, 811–824. [[CrossRef](#)]
50. Zawislak, G. Comparison of chemical composition of the essential oil from *Marrubium vulgare* L. and *M. incanum* Desr. during the second year of cultivation. *Acta Agrobot.* **2015**, *68*, 59–62. [[CrossRef](#)]
51. Petrovic, S.; Pavlovic, M.; Maksimovic, Z.; Milenkovic, M.; Couladis, M.; Tzakou, O.; Niketic, M. Composition and antimicrobial activity of *Marrubium incanum* Desr. (Lamiaceae) essential oil. *Nat. Prod. Commun.* **2009**, *4*, 431–434. [[CrossRef](#)] [[PubMed](#)]
52. Chemsal, A.E.; Zellaoui, A.; Ozturk, M.; Erol, E.; Ceylan, O.; Duru, M.E.; Gherraf, N. Antibiofilm formation, antioxidant and anticholinesterase activities of essential oil and methanol extract of *Marrubium deserti* de Noe. *J. Mater. Env. Sci.* **2016**, *7*, 993–1000.
53. Mohammadhosseini, M. First report of screening of the profiles of the essential oils and volatiles from the aerial parts of *Marrubium persicum* using classical and advanced methods prior to Gas Chromatographic Mass Spectrometric determination. *J. Med. Plants Prod.* **2016**, *2*, 169–180.
54. Agryropoulou, C.; Skaltsa, H. Identification of essential oil components of *Marrubium thessalum* Boiss. & Heldr., growing wild in Greece. *Nat. Prod. Res.* **2012**, *26*, 1–7.
55. Aćimović, M.; Pezo, L.; Tešević, V.; Čabarkapa, I.; Todosijević, M. QSRR Model for predicting retention indices of *Satureja kitaibelii* Wierzb. Ex Heuff. essential oil composition. *Ind. Crops Prod.* **2020**, *154*, 112752. [[CrossRef](#)]

56. Nekoei, M.; Salimi, M.; Dolatabadi, M.; Mohammadhosseini, M. Prediction of antileukemia activity of berbamine derivatives by genetic algorithm–multiple linear regression. *Mon. Chem.* **2011**, *142*, 943. [CrossRef]
57. Nekoei, M.; Mohammadhosseini, M.; Pourbasheer, E. QSAR study of VEGFR-2 inhibitors by using genetic algorithm–multiple linear regressions (GA-MLR) and genetic algorithm–support vector machine (GA-SVM): A comparative approach. *Med. Chem. Res.* **2015**, *24*, 3037–3046. [CrossRef]
58. Azar, P.A.; Nekoei, M.; Riahi, S.; Ganjali, M.R.; Zare, K. A quantitative structure-retention relationship for the prediction of retention indices of the essential oils of *Anmnoides atlantica*. *J. Serb. Chem Soc.* **2011**, *76*, 891–902. [CrossRef]
59. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors, Methods and Principles in Medicinal Chemistry*; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2000.
60. Bakić, V.V.; Pezo, M.L.; Jovanović, M.P.; Turanjanin, V.M.; Vučićević, B.S.; Mirkov, N.S. Technical analysis of photovoltaic/wind systems with hydrogen storage. *Science* **2012**, *16*, 865–875.
61. Arsenović, M.; Pezo, L.; Stanković, S.; Radojević, Z. Factor space differentiation of brick clays according to mineral content: Prediction of final brick product quality. *Appl. Clay. Sci.* **2015**, *115*, 108–114. [CrossRef]
62. Boulila, A.; Sanaa, A.; Salem, I.B.; Rokbeni, N.; Mrabet, Y.; Hosni, K.; Fernandez, X. Antioxidant properties and phenolic variation in wild populations of *Marrubium vulgare* L. (Lamiaceae). *Ind. Crops Prod.* **2015**, *76*, 616–622. [CrossRef]
63. Mahdavi, A.; Moradi, P.; Mastinu, A. Variation in terpene profiles of *Thymus vulgaris* in water deficit stress response. *Molecules* **2020**, *25*, 1091. [CrossRef] [PubMed]
64. Noorizadeh, H.; Farmany, A.; Noorizadeh, M. Quantitative structure-retention relationships analysis of retention index of essential oils. *Química Nova* **2011**, *34*, 242–249. [CrossRef]
65. Driouche, Y.; Messadi, D. Quantitative structure-retention relationship model for predicting retention indices of constituents of essential oils of *Thymus vulgaris* (Lamiaceae) (Short communication). *J. Serb. Chem Soc.* **2019**, *84*, 405–416. [CrossRef]
66. Acimovic, M.; Pezo, L.; Stankovic Jeremic, J.; Cvetkovic, M.; Rat, M.; Cabarkapa, I.; Tesevic, V. QSRR model for predicting retention indices of geraniol chemotype of *Thymus serpyllum* essential oil. *J. Essent Oil Bear Plants* **2020**, *23*, 464–473. [CrossRef]
67. Pavlič, B.; Teslić, N.; Kojić, P.; Pezo, L. Prediction of the GC-MS retention time for terpenoids detected in sage (*Salvia officinalis* L.) essential oil using QSRR approach. *J. Serb. Chem. Soc.* **2020**, *85*, 9–23. [CrossRef]
68. Noorizadeh, H. Linear and nonlinear quantitative structure linear retention indices relationship models for essential oils. *Eurasian J. Anal. Chem* **2013**, *8*, 50–63.
69. Héberger, K. Quantitative structure–(chromatographic) retention relationships. *J. Chromatogr A* **2007**, *1158*, 273–305. [CrossRef]
70. Kalisz, R.; Bączek, T.; Buciniński, A.; Buszewski, B.; Sztupecka, M. Prediction of gradient retention from the linear solvent strength (LSS) model, quantitative structure-retention relationships (QSRR), and artificial neural networks (ANN). *J. Sep. Sci* **2003**, *26*, 271–282. [CrossRef]
71. Khodadoust, S.; Ghaedi, M.; Hadjmohammadi, M.R. Dispersive nano solid material-ultrasound assisted microextraction as a novel method for extraction and determination of bendiocarb and promecarb: Response surface methodology. *Talanta* **2013**, *116*, 637–646. [CrossRef]
72. Wolfender, J.L.; Martia, G.; Thomas, A.; Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* **2015**, *1382*, 136–164. [CrossRef] [PubMed]
73. Zisi, C.; Sampsonidis, I.; Fasoula, S.; Papachristos, K.; Witting, M.; Gika, H.G.; Nikitas, P.; Pappa-Louisi, A. QSRR modeling for metabolite standards analyzed by two different chromatographic columns using multiple linear regression. *Metabolites* **2017**, *7*, 7. [CrossRef]
74. Deconinck, E.; Zhang, M.H.; Petit, F.; Dubus, E.; Ijjaali, I.; Coomans, D.; Vander Heyden, Y. Boosted regression trees, multivariate adaptive regression splines and their two-step combinations with multiple linear regression or partial least squares to predict blood–brain barrier passage: A case study. *Anal. Chim. Acta* **2008**, *609*, 13–23. [CrossRef]
75. Yap, C.W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32*, 1446–1474. [CrossRef]
76. Goldberg, D.E. *Genetic Algorithms in Search, Optimisation and Machine Learning*; Addison-Wesley: Boston, MA, USA, 1989.
77. Gramatica, P. Principles of QSAR models validation: Internal and external. *Qsar Comb. Sci* **2007**, *26*, 694–701. [CrossRef]
78. Statistica 10 Software (StatSoft, Inc. STATISTICA, ver. 10, Data Analysis Software System). Available online: <https://www.statsoft.de/en/software/tibco-statisticatm> (accessed on 15 December 2018).
79. Foroughia, M.; Hossein, M.; Azqhandic, A.; Kakhki, S. Bio-inspired, high, and fast adsorption of tetracycline from aqueous media using Fe<sub>3</sub>O<sub>4</sub>-g-CN@PEI-β-CD nanocomposite: Modeling by response surface methodology (RSM), boosted regression tree (BRT), and general regression neural network (GRNN). *J. Hazard. Mater.* **2020**, *388*, 121769. [CrossRef]
80. Cheong, Y.L.; Leitão, P.J.; Lakes, T. Assessment of land use factors associated with dengue cases in Malaysia using Boosted Regression Trees. *Spat. Spatiotemporal Epidemiol.* **2014**, *10*, 75–84. [CrossRef] [PubMed]
81. Froeschke, J.T.; Froeschke, B.F. Spatio-temporal predictive model based on environmental factors for juvenile spotted seatrout in Texas estuaries using boosted regression trees. *Fish. Res.* **2011**, *111*, 131–138. [CrossRef]
82. Sut, N.; Simsek, O. Comparison of regression tree data mining methods for prediction of mortality in head injury. *Expert Syst. Appl.* **2011**, *38*, 15534–15539. [CrossRef]



- 
83. Mazaheri, H.; Ghaedi, M.; Azqhandi, M.A.; Asfaram, A. Application of machine/ statistical learning, artificial intelligence and statistical experimental design for the modeling and optimization of methylene blue and Cd (II) removal from a binary aqueous solution by natural walnut carbon. *Phys. Chem. Chem. Phys.* **2017**, *19*, 11299–11317. [[CrossRef](#)]
  84. Salonen, J.S.; Luoto, M.; Alenius, T.; Heikkilä, M.; Seppä, H.; Telford, R.J.; Birks, H.J. Reconstructing palaeoclimatic variables from fossil pollen using boosted regression trees: Comparison and synthesis with other quantitative reconstruction methods. *Quat Sci. Rev.* **2014**, *88*, 69–81. [[CrossRef](#)]