

Article

# Weighted Wiener Indices of Molecular Graphs with Application to Alkenes and Alkadienes

Simon Brezovnik <sup>1,2</sup>, Niko Tratnik <sup>1</sup>  and Petra Žigert Pleteršek <sup>1,3,\*</sup>

<sup>1</sup> Faculty of Natural Sciences and Mathematics, University of Maribor, 2000 Maribor, Slovenia; simon.brezovnik2@um.si (S.B.); niko.tratnik@um.si (N.T.)

<sup>2</sup> Faculty of Education, University of Maribor, 2000 Maribor, Slovenia

<sup>3</sup> Faculty of Chemistry and Chemical Engineering, University of Maribor, 2000 Maribor, Slovenia

\* Correspondence: petra.zigert@um.si

**Abstract:** There exist many topological indices that are calculated on saturated hydrocarbons since they can be easily modelled by simple graphs. On the other hand, it is more challenging to investigate topological indices for hydrocarbons with multiple bonds. The purpose of this paper is to introduce a simple model that gives good results for predicting physico-chemical properties of alkenes and alkadienes. In particular, we are interested in predicting boiling points of these molecules by using the well known Wiener index and its weighted versions. By performing the non-linear regression analysis we predict boiling points of alkenes and alkadienes.

**Keywords:** weighted Wiener index; alkenes; alkadienes; boiling point



**Citation:** Brezovnik, S.; Tratnik, N.; Pleteršek, P.Ž. Weighted Wiener Indices of Molecular Graphs with Application to Alkenes and Alkadienes. *Mathematics* **2021**, *9*, 153. <https://doi.org/10.3390/math9020153>

Received: 1 November 2020

Accepted: 7 January 2021

Published: 12 January 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/>).

## 1. Introduction

The *Wiener index* is a graph invariant based on distances in a graph. It is denoted by  $W(G)$  and defined as the sum of distances between all pairs of vertices in  $G$ :

$$W(G) = \sum_{\{u,v\} \subseteq V(G)} d_G(u,v). \quad (1)$$

The name *Wiener index* is usual in chemical literature, since Harold Wiener [1] in 1947 seems to be the first who considered it. His original definition was slightly different—yet equivalent—to (1); the definition of the Wiener index in terms of distances between vertices of a graph, such as in Equation (1), was first given by Hosoya [2].

Wiener stated that the boiling points of organic compounds as well as their physical properties depend on the number and structural arrangement of atoms in a molecule. In chemistry, an organic compound is generally any chemical compound that contains carbon-C and hydrogen-H atoms. Molecules composed only of this two types of atoms are called hydrocarbons. The graph representation of a hydrocarbon with only single bonds (saturated hydrocarbon) is a simple graph where vertices represent carbon atoms, two vertices being adjacent if there is a CC bond between the corresponding C atoms. Since the valency of a carbon atom equals four, the number of hydrogen atoms adjacent to each C atom is determined and can therefore be omitted in the graph representation. A comparative study between the Wiener index and several other topological indices as predictors for the boiling points of saturated hydrocarbons called alkanes was done in Reference [3], showing that best results are obtained in 3- or 4-parameters relationship, both of them including the Wiener index.

In the case of hydrocarbons with double or triple bonds – unsaturated hydrocarbons, the graph representation should be a multigraph. Such graphs can be described with the vertex-adjacency matrix  $A$  for which  $(A)_{ij} = m_{ij}$  where  $m_{ij}$  is the multiplicity (i.e., bond order) of the edge between vertices  $v_i$  and  $v_j$  (see Reference [4]). But there is some ambiguity in the mathematical modelling of unsaturated hydrocarbons. In one of the seminal papers

on the prediction of physico-chemical properties from topological indices by Basak et al. [5] three different clusters of hydrocarbons were investigated: alkanes, alkylbenzenes and polycyclic aromatic hydrocarbons, all of them modelled with a simple graph although the last two groups of molecules contain double CC bonds. In the past few years much attention is given to the unsaturated hydrocarbons called carbon nanotubes, especially to fullerenes (cubic 3-connected planar graphs with only pentagonal and hexagonal faces), which are modelled with simple graphs. For some latest result on topological indices (parameters) of this graphs see References [6–8]. The Wiener index of some alkenes was calculated in Reference [9] with the use of the eccentricity of a graph and the double bonds were treated the same way as single bonds. Moreover, the Wiener index of unsaturated hydrocarbons was considered in Reference [10]. It was shown that the Wiener index  $W$  of a molecular graph can be written as  $W = W_s + W_d + W_t + W_a$ , where  $s$ ,  $d$ ,  $t$ , and  $a$  refer to the contributions of single, double, triple, and aromatic bonds, respectively. Then, multiple linear regression analysis was performed and a good correlation between the logarithms of n-octanol-water partition coefficients of the molecules and the mentioned components was established ( $R = 0.95$ ). In this investigation, the considered collection of molecules included 71 saturated and unsaturated hydrocarbons. Furthermore, in Reference [11] the authors distinguish between double and single bonds of molecules with the use of quite complex charge density of an atom (CMI—Charge-related Molecular Index).

On the other hand, some authors introduced the concept of a weighted graph to consider multiple bonds in unsaturated hydrocarbons. The approach of searching the optimal weight  $x$  of a double bond in alkenes was addressed by Randić and Pompe in Reference [12]. They selected a set of 39 alkenes and chose as a property a molar refraction number and then use the statistical analysis in optimizing  $x$  so that the standard error of the regression was the smallest. These two authors et al. in Reference [13] performed a QSPR study for predicting gas-phase reaction rate constants of unsaturated organic compounds with OH radical. Two approaches were taken; the first one was an optimized 6-parameters MLR regression model originated with a large pool of topological descriptors and the second one focused only on one topological index where the weights of the atoms are variables that are to be determined by the optimization procedure in the best fit to the considered physico-chemical property. Grossman et al. [14] used the concept of weighted paths in the characterizations of heteroatomic molecules and their activity.

Moreover, in Reference [15] authors defined a novel distance between two adjacent atoms in a molecular graph which corresponds to the length of the bond. Next, they introduced so-called valence overall Wiener index and found out that it has a good correlation with molecular volume ( $R = 0.9955$ ), boiling point ( $R = 0.9906$ ), partition coefficients ( $R = 0.9280$ ), molecular refractions ( $R = 0.9946$ ), critical temperature ( $R = 0.9906$ ), and critical pressures ( $R = 0.9307$ ) for various hydrocarbons with different types of bonds. Furthermore, a very similar approach was used in References [16,17] for the hyper-Wiener index. More precisely, in Reference [16] a novel type of hyper-Wiener index for unsaturated hydrocarbons was defined by taking into account the relative distances. Then, three physical properties were considered. After applying multiple regression analysis for 41 unsaturated hydrocarbons, good correlations with the following three properties were obtained: boiling point ( $R = 0.9820$ ), molecular volumes ( $R = 0.9810$ ), heats of atomic formation ( $R = 0.9775$  in one case and  $R = 0.9999$  in another case). On the other hand, the authors of Reference [17] used the relative distances to define a novel overall hyper-Wiener index. They performed multiple regression analysis in a QSAR modelling of six physical properties (boiling point, molar volumes, partition coefficient, molecular refractions, critical temperature, and critical pressures) of 42 alkanes and unsaturated hydrocarbons with one to six carbons. It was concluded that the novel overall hyper-Wiener index has good relationships with molecular properties of these hydrocarbons with different types of multiple bonds.

We are interested in the correlation between the weighted Wiener index and the boiling points of alkenes and alkadienes. The aim of this paper is to mathematically

model an unsaturated hydrocarbon with an edge-weighted graph, then calculate the corresponding distances in the graph and obtain a new Wiener index. With the use of the QSPR (Quantitative Structure Property Relationships) we show that the new Wiener index gives very good predictions of the boiling points of alkenes and alkadienes. This approach can be applied in the calculation of different distance-based topological indices for unsaturated hydrocarbons as well as for organic compounds in general. At the end, we compare our method for a group of considered molecules with the method used in References [15–17].

## 2. Graph Theory Preliminaries

A graph  $G$  is an ordered pair  $G = (V, E)$  of a set  $V$  of vertices and a set  $E$  of edges, which are 2-element subsets of  $V$ . The edge  $e = \{u, v\}$  between vertices  $u$  and  $v$  will be also denoted as  $e = uv$ . All the basic concepts from graph theory can be found in Reference [18]. Having a molecule, if we represent atoms by vertices and bonds between them by edges, we obtain a *molecular graph*.

If  $G$  is a connected graph, then a function  $w : E(G) \rightarrow \mathbb{R}^+$  is called an *edge-weight* of  $G$ . The pair  $(G, w)$  is known as the *edge-weighted graph*.

A *path of length  $n - 1$*  between vertices  $v_1$  and  $v_n$  in a connected graph  $G$  is a sequence of vertices  $P = (v_1, v_2, \dots, v_n)$  such that  $v_i$  is adjacent to  $v_{i+1}$  for any  $i \in \{1, 2, \dots, n - 1\}$ . The *weight of a path  $P = (v_1, v_2, \dots, v_n)$*  in an edge-weighted graph  $(G, w)$  is defined as

$$w(P) = \sum_{i=1}^{n-1} w(v_i v_{i+1}).$$

A *shortest path* between vertices  $u$  and  $v$  of  $(G, w)$  is a path with the minimum weight  $w(P)$  among all possible paths between  $u$  and  $v$ . The *distance* between  $u$  and  $v$ ,  $d_{(G,w)}(u, v)$ , is the weight of any shortest path  $P$  between  $u$  and  $v$ , that is,  $d_{(G,w)}(u, v) := w(P)$ . Obviously, if  $w(e) = 1$  for any  $e \in E(G)$ , then  $d_{(G,w)}(u, v)$  is the standard graph distance denoted simply as  $d_G(u, v)$ .

Let  $(G, w)$  be an edge-weighted graph. The *Wiener index* of  $(G, w)$  [19] is defined as

$$W(G, w) = \sum_{\{u,v\} \subseteq V(G)} d_{(G,w)}(u, v).$$

Obviously, if  $w \equiv 1$ , then  $W(G, w)$  is the standard Wiener index  $W(G)$ . The Wiener index of a weighted graph can be called the *weighted Wiener index*.

Let  $G$  be a connected graph and  $D \subseteq E(G)$  a subset of its edges. If  $e \in D$ , then  $e$  will be called a *double edge* of  $G$ . In this paper, we denote a connected graph  $G$  with the set of double edges  $D$  as  $G^D$ . Note that in figures the double edges are depicted with two lines between corresponding vertices and that  $G^D$  can be represented as a multigraph.

If  $a \in \mathbb{R}^+$  is a positive constant and  $G^D$  a connected graph with double edges, then the weight  $w' : E(G) \rightarrow \mathbb{R}^+$  is defined in the following way:

- (i) if  $e \in E(G) \setminus D$ , then  $w'(e) = 1$ ,
- (ii) if  $e \in D$ , then  $w'(e) = a$ .

Consequently, the Wiener index of  $G^D$  with respect to  $a$ , denoted by  $W_a(G^D)$ , is defined as the Wiener index of  $(G, w')$ , that is,

$$W_a(G^D) := W(G, w').$$

Finally, we define the valence Wiener index [15] by replacing the classical distances with the relative distances. Obviously, the distances in a molecule are determined by chemical bonds, so the distances between two adjacent vertices in a molecular graph should be correlated with the bond lengths between two atoms. As a consequence, the relative distance defines the distance of the CC single bond ( $sp^3 - sp^3$  type) as 1. Then the

relative distances of other types of chemical bonds are defined as the ratio of their bond lengths and the bond length of CC single bond ( $sp^3 - sp^3$  type), which equals 1.544. The bond lengths and relative distances of some types of chemical bonds [20] are shown in Table 1.

**Table 1.** Bond types with their lengths and relative distances.

Bond Type	Bond Length	Relative Distance
$sp^3 - sp^3$	1.544	1.000
$sp^3 - sp^2$	1.501	0.972
$sp^2 - sp^2$	1.483	0.960
$sp^2 - sp^2 + \pi$	1.339	0.867
$sp^2 - sp + \pi$	1.309	0.848

If  $G^D$  is a connected graph with double edges, then we define edge-weight  $w''$  so that  $w''(e)$  is the relative distance of the corresponding chemical bond. Hence, the *valence Wiener index*  $W_v$  is

$$W_v(G^D) := W(G, w'').$$

### 3. Algorithms for Computing the Wiener Indices

In this section, we present two algorithms that were used for computing the Wiener index of an edge-weighted graph  $(G, w)$ . Throughout the section, the  $(i, j)$  entry of a matrix  $A$ ,  $(A)_{ij}$ , will be denoted as  $A[i][j]$ .

Let  $V(G) = \{1, \dots, n\}$  be the vertex set of  $(G, w)$ . We represent the graph by a  $n \times n$  array (matrix), such that if  $e = \{i, j\}$  is an edge, then the  $(i, j)$  entry and the  $(j, i)$  entry of the array are both equal to  $w(e)$ , all diagonal elements are set to 0, and all the other entries of the array are initialized to infinity. Then, we perform the well-known Floyd–Warshall algorithm [21] (Algorithm 1) to compute the distance matrix  $D$  of  $(G, w)$ . More precisely, for the *distance matrix*  $D$  it holds  $D[i][j] = D[j][i] = d_{(G,w)}(i, j)$  for any  $i, j \in V(G)$ . In our version of the algorithm, we consider only undirected graphs, but Algorithm 1 can be easily adopted so that it can be used also on directed graphs. It is easy to see that it can be implemented in  $O(n^3)$  time.

When the distance matrix of  $(G, w)$  is computed, we can easily calculate the Wiener index  $W(G, w)$  by the procedure described in Algorithm 2. It is obvious that this algorithm correctly computes the Wiener index of an edge-weighted graph and has the time complexity  $O(n^2)$ .

For the clarity, we present the described procedure for computing the index  $W_a$  of the graph  $G^D$  with label 6D11 from Figure 1. The obtained distance matrix is

$$\begin{bmatrix} 0 & a & a+1 & a+2 & a+2 & a+2 \\ a & 0 & 1 & 2 & 2 & 2 \\ a+1 & 1 & 0 & 1 & 1 & 1 \\ a+2 & 2 & 1 & 0 & 2 & 2 \\ a+2 & 2 & 1 & 2 & 0 & 2 \\ a+2 & 2 & 1 & 2 & 2 & 0 \end{bmatrix}.$$

Hence, we obtain

$$W_a(G^D) = 5a + 23.$$

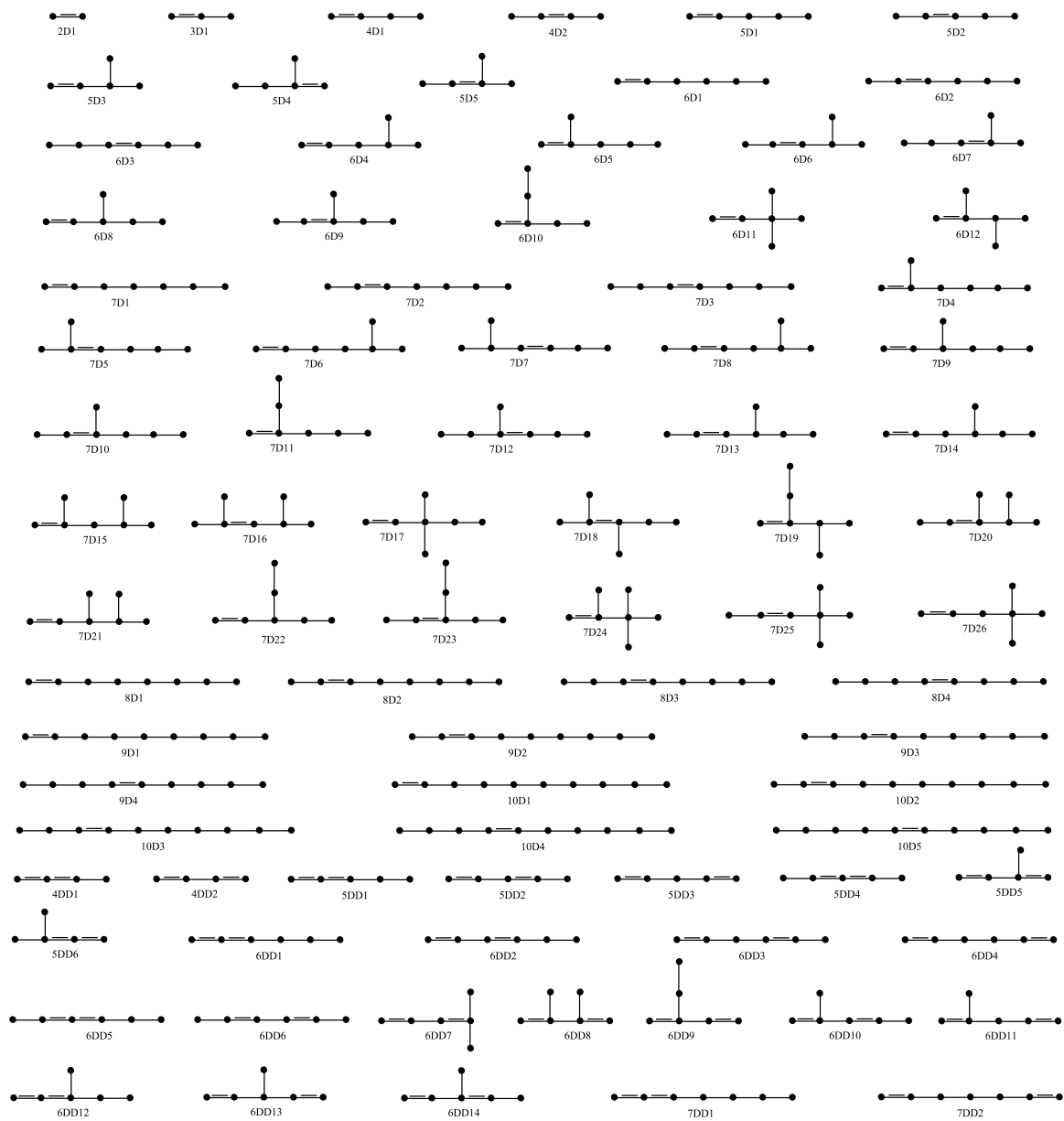


Figure 1. Graphs of considered molecules with their labels.

**Algorithm 1:** Floyd-Warshall algorithm.

---

**Input:** An edge-weighted graph  $(G, w)$  with  $n$  vertices  $(1, 2, \dots, n)$ , represented by the  $n \times n$  array.  
**Output:** The distance matrix  $D$ , that is,  $n \times n$  array of weighted distances between all pairs of vertices.

```

1  $D = n \times n$  array with all entries initialized to infinity
2 for each edge  $e = \{i, j\}$  do
3   |  $D[i][j] := w(e)$ 
4   |  $D[j][i] := w(e)$ 
5 end
6 for each vertex  $i = 1, \dots, n$  do
7   |  $D[i][i] := 0$ 
8 end
9 for  $k = 1, \dots, n$  do
10  | for  $i = 1, \dots, n$  do
11    | for  $j = 1, \dots, n$  do
12      | if  $D[i][j] > D[i][k] + D[k][j]$  then
13        | |  $D[i][j] := D[i][k] + D[k][j]$ 
14      | end
15    | end
16  | end
17 end
18 return  $D$ 

```

---

**Algorithm 2:** Wiener index of an edge-weighted graph.

---

**Input:** Distance matrix  $D$  for an edge-weighted graph  $(G, w)$ .  
**Output:** The Wiener index of  $(G, w)$ , that is,  $W(G, w)$ .

```

1  $W := 0$ 
2 for  $i = 1, \dots, n - 1$  do
3   | for  $j = i + 1, \dots, n$  do
4     |  $W := W + D[i][j]$ 
5   | end
6 end
7 return  $W$ 

```

---

#### 4. Experimental Data

In this section we present the data that was used. Firstly, we consider the alkenes (with one double bond). We wanted to obtain the boiling points at normal pressure for unbranched alkenes with at most ten carbon atoms and also all branched alkenes with at most seven carbon atoms (the number of branched alkenes increases very quickly with the number of carbon atoms). Table 2 presents all such alkenes for which the data was available in Reference [22] (only some alkenes, for which we did not find the data, were omitted).

In our model, different stereoisomers are represented by the same graph with double edges. Therefore, it can not distinguish between E and Z isomer molecules (molecular descriptors that can distinguish between such isomers were considered, for example, in Reference [23]). However, the difference in the boiling points between E and Z isomers is usually small. Moreover, E isomers are commonly more stable and lower in energy, therefore at normal conditions compound contains mostly those types of geometric isomers. As a consequence, we considered only the data for E isomers.

To every alkene we assign a label (see the second column in Table 2) in which the first number represents the number of carbon atoms and the number of double bonds corresponds to the number of times letter D appears in the label. The last number in the label is simply the consecutive number. The alkenes from Table 2, together with their labels, are depicted in Figure 1.

**Table 2.** Alkenes, their boiling points, and Wiener indices.

Name	Label	Boiling Point at n.p. [K]	W	W <sub>2</sub>	W <sub>1/2</sub>	Test Set
ethene	2D1	169.41	1	2	0.5	
prop-1-ene	3D1	225.46	4	6	3	x
but-1-ene	4D1	266.92	10	13	8.5	
but-2-ene	4D2	274.03	10	14	8	
pent-1-ene	5D1	303.11	20	24	18	
pent-2-ene	5D2	309.49	20	26	17	x
3-methyl-1-butene	5D3	293.21	18	22	16	x
2-methyl-1-butene	5D4	304.3	18	22	16	
2-methyl-2-butene	5D5	311.7	18	24	15	
heks-1-ene	6D1	336.63	35	40	32.5	
heks-2-ene	6D2	341.03	35	43	31	
heks-3-ene	6D3	340.24	35	44	30.5	
4-methyl-1-pentene	6D4	327.02	32	37	29.5	
2-methyl-1-pentene	6D5	335.28	32	37	29.5	x
4-methyl-2-pentene	6D6	331.76	32	40	28	
2-methyl-2-pentene	6D7	340.46	32	41	27.5	
3-methyl-1-pentene	6D8	327.33	31	36	28.5	
3-methyl-2-pentene	6D9	343.59	31	39	27	x
2-ethyl-1-butene	6D10	337.83	31	36	28.5	
3,3-dimethyl-1-butene	6D11	314.4	28	33	25.5	
2,3-dimethyl-1-butene	6D12	328.77	29	34	26.5	
hept-1-ene	7D1	366.79	56	62	53	x
hept-2-ene	7D2	371.1	56	66	51	
hept-3-ene	7D3	368.82	56	68	50	
2-methyl-1-hexene	7D4	365.15	52	58	49	x
2-methyl-2-hexene	7D5	368.56	52	64	46	
5-methyl-1-hexene	7D6	358.46	52	58	49	
2-methyl-3-hexene	7D7	359.15	52	64	46	
5-methyl-2-hexene	7D8	361.26	52	62	47	x
3-methyl-1-hexene	7D9	357.05	50	56	47	
3-methyl-2-hexene	7D10	368.33	50	60	45	
2-ethyl-1-pentene	7D11	367.15	50	56	47	
3-methyl-3-hexene	7D12	366.69	50	62	44	x
4-methyl-2-hexene	7D13	360.71	50	60	45	
4-methyl-1-hexene	7D14	359.88	50	56	47	
2,4-dimethyl-1-pentene	7D15	354.76	48	54	45	
2,4-dimethyl-2-pentene	7D16	356.45	48	60	42	x
3,3-dimethyl-1-pentene	7D17	350.63	44	50	41	
2,3-dimethyl-2-pentene	7D18	370.55	46	58	40	
2-ethyl-3-methyl-1-butene	7D19	359.51	46	52	43	

Table 2. Cont.

Name	Label	Boiling Point at n.p. [K]	W	W <sub>2</sub>	W <sub>1/2</sub>	Test Set
3,4-dimethyl-2-pentene	7D20	364.65	46	56	41	x
3,4-dimethyl-1-pentene	7D21	353.95	46	52	43	
3-ethyl-1-pentene	7D22	357.26	48	54	45	
3-ethyl-2-pentene	7D23	369.16	48	58	43	
2,3,3-trimethyl-1-butene	7D24	351.04	42	48	39	x
4,4-dimethyl-2-pentene	7D25	349.89	46	56	41	
4,4-dimethyl-1-pentene	7D26	345.67	46	52	43	
okt-1-ene	8D1	394.44	84	91	80.5	
okt-2-ene	8D2	398.08	84	96	78	x
okt-3-ene	8D3	396.4	84	99	76.5	
okt-4-ene	8D4	395.41	84	100	76	
non-1-ene	9D1	420.03	120	128	116	
non-2-ene	9D2	423.19	120	134	113	x
non-3-ene	9D3	421.29	120	138	111	
non-4-ene	9D4	420.87	120	140	110	
dek-1-ene	10D1	443.34	165	174	160.5	
dek-2-ene	10D2	446.44	165	181	157	x
dek-3-ene	10D3	444.55	165	186	154.5	
dek-4-ene	10D4	443.94	165	189	153	
dek-5-ene	10D5	443.47	165	190	152.5	

However, considering the alkenes with more than one double bond is very demanding due to their number. Therefore, in Table 3 we included all the alkenes with two double bonds (alkadienes) up to six carbon atoms and two additional alkadienes with seven carbon atoms for which the data was available (again, a label was assigned to every alkene from Table 3). These alkadienes are depicted in the lower part of Figure 1.

Table 3. Alkadienes, their boiling points, and Wiener indices.

Name	Label	Boiling Point at n.p. [K]	W	W <sub>2</sub>	W <sub>1/2</sub>	Test Set
1,2-butadiene	4DD1	284	10	17	6.5	x
1,3-butadiene	4DD2	268.74	10	16	7	
1,2-pentadiene	5DD1	318.01	20	30	15	
1,3-pentadiene	5DD2	315.18	20	30	15	
1,4-pentadiene	5DD3	299.12	20	28	16	x
2,3-pentadiene	5DD4	321.41	20	32	14	
2-methyl-1,3-butadiene	5DD5	307.22	18	26	14	x
3-methyl-1,2-butadiene	5DD6	314	18	28	13	
1,2-heksadiene	6DD1	349.15	35	48	28.5	
1,3-heksadiene	6DD2	346.4	35	49	28	x
1,4-heksadiene	6DD3	337.52	35	48	28.5	x
1,5-heksadiene	6DD4	332.62	35	45	30	



Table 3. Cont.

Name	Label	Boiling Point at n.p. [K]	$W$	$W_2$	$W_{\frac{1}{2}}$	Test Set
2,3-heksadiene	6DD5	341.15	35	52	26.5	x
2,4-hexadiene	6DD6	353.15	35	51	27	
1,1-dimethylbutadiene	6DD7	349.45	32	46	25	
2,3-dimethyl-1,3-butadiene	6DD8	341.66	29	39	24	
2-ethyl-1,3-butadiene	6DD9	348.15	31	41	26	x
2-methyl-1,3-pentadiene	6DD10	349.15	32	45	25.5	
2-methyl-1,4-pentadiene	6DD11	329.15	32	42	27	
3-methyl-1,2-pentadiene	6DD12	343.15	31	44	24.5	
3-methyl-1,4-pentadiene	6DD13	328.15	31	41	26	x
3-methyl-1,3-pentadiene	6DD14	350.15	31	44	24.5	
1,2-heptadiene	7DD1	376.9	56	72	48	
1,6-heptadiene	7DD2	362.7	56	68	50	

## 5. Results

In this section, we present the results obtained by the QSPR analysis. We divided the data into the training and the test set (see “x” in the last column in Tables 2 and 3). The correlation between the boiling points and weighted Wiener indices  $W = W_1, W_2$ , and  $W_{\frac{1}{2}}$  was investigated. Weighted Wiener indices are also calculated in Tables 2 and 3.

### 5.1. Alkenes

After performing regression analysis on all three weighted Wiener indices, it was established that the best results are obtained with the index  $W_2$ . Therefore, we only present the results of the regression between the boiling points and  $W_2$ . The obtained logarithmic function based on the training set can be seen in Figure 2. The predicted boiling point ( $\widehat{BP}$ ) can now be computed as

$$\widehat{BP} = 63.071 \ln(W_2) + 106.69. \quad (2)$$

The correlation is very good since  $R^2 = 0.9847$  (for the comparison, the correlation with the Wiener index  $W$  is slightly lower, that is,  $R^2 = 0.9725$ ).

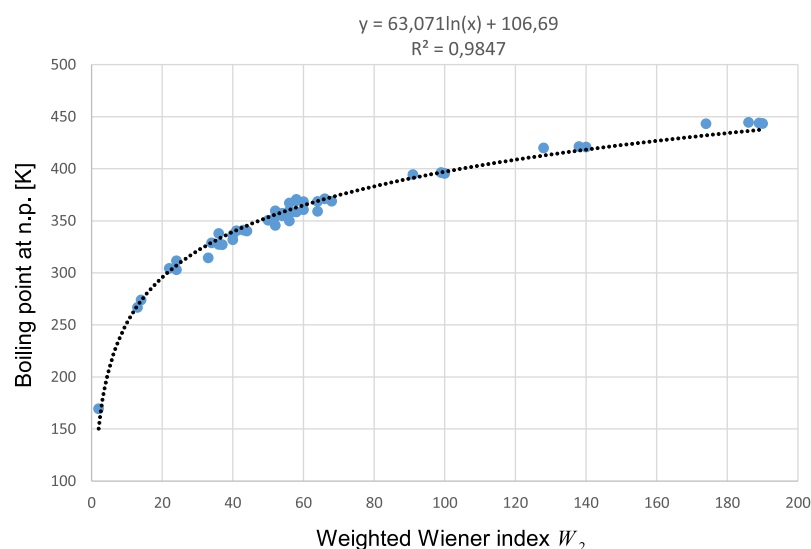


Figure 2. Weighted Wiener indices  $W_2$  and boiling points for the alkenes in the training set.

By applying Equation (2) we calculate the predicted boiling points and the corresponding residuals on the test set, see Table 4. Moreover, in Table 5 the mentioned data is provided for all considered alkenes. Finally, the graphical representation of boiling points and predicted boiling points is shown in Figure 3.

In order to obtain the best possible correlation, we consider separately also the unbranched alkenes. The correlation is slightly better, since  $R^2 = 0.9937$  (see Figure 4).

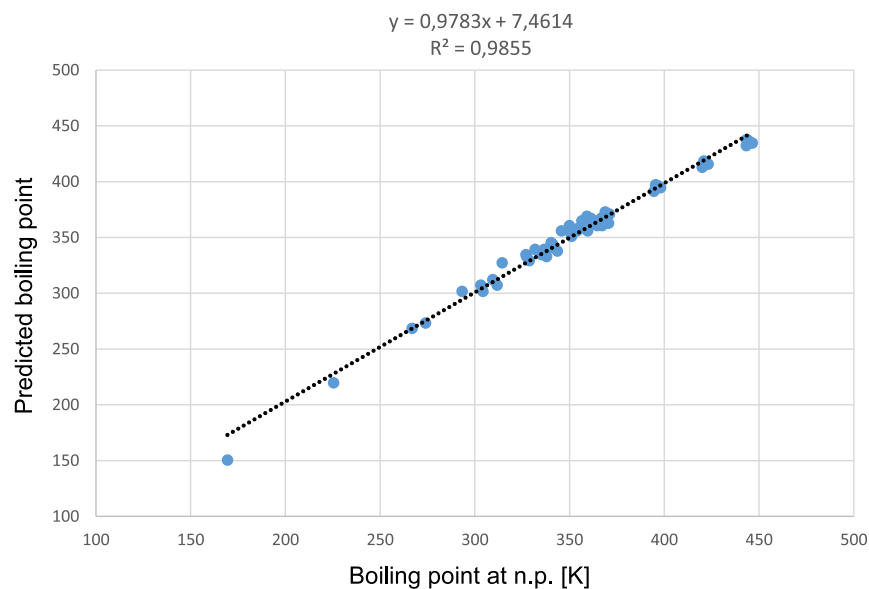


Figure 3. Boiling points and predicted boiling points for all considered alkenes.

Table 4. Data for the alkenes in the test set.

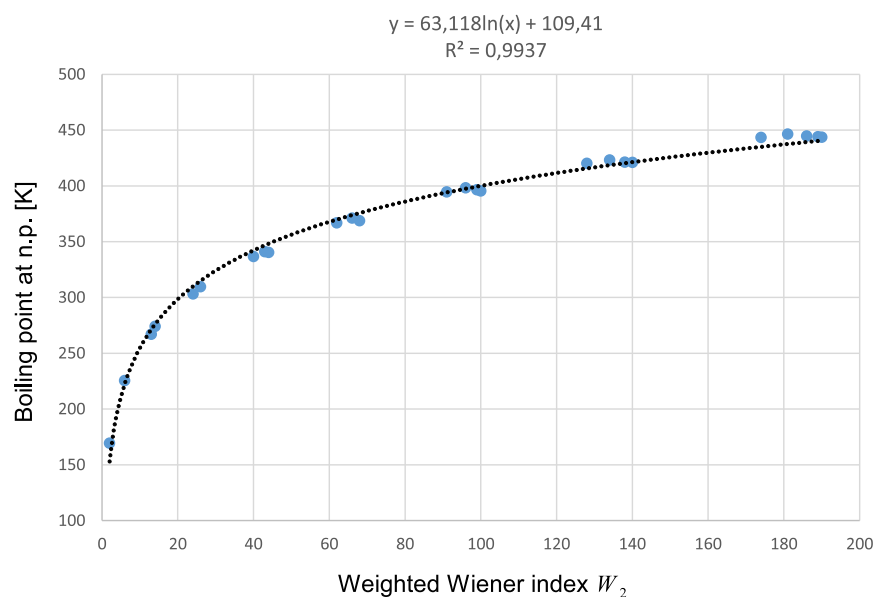
Name	Label	BP [K]	$\widehat{BP}$	Residual	% Residual
prop-1-ene	3D1	225.46	219.698	5.762	2.556
pent-2-ene	5D2	309.49	312.181	2.691	0.870
3-methyl-1-butene	5D3	293.21	301.645	8.435	2.877
2-methyl-1-pentene	6D5	335.28	334.434	0.846	0.252
3-methyl-2-pentene	6D9	343.59	337.754	5.836	1.698
hept-1-ene	7D1	366.79	366.992	0.202	0.055
2-methyl-1-hexene	7D4	365.15	362.786	2.364	0.647
5-methyl-2-hexene	7D8	361.26	366.992	5.732	1.587
3-methyl-3-hexene	7D12	366.69	366.992	0.302	0.082
2,4-dimethyl-2-pentene	7D16	356.45	364.924	8.474	2.377
3,4-dimethyl-2-pentene	7D20	364.65	360.573	4.077	1.118
2,3,3-trimethyl-1-butene	7D24	351.04	350.851	0.189	0.054
okt-2-ene	8D2	398.08	394.568	3.512	0.882
non-2-ene	9D2	423.19	415.602	7.588	1.793
dek-2-ene	10D2	446.44	434.564	11.876	2.660
average					1.301

**Table 5.** Data for all considered alkenes.

Name	Label	BP [K]	$\widehat{BP}$	Residual	% Residual
ethene	2D1	169.41	150.407	19.003	11.217
prop-1-ene	3D1	225.46	219.698	5.762	2.556
but-1-ene	4D1	266.92	268.464	1.544	0.578
but-2-ene	4D2	274.03	273.138	0.892	0.326
pent-1-ene	5D1	303.11	307.133	4.023	1.327
pent-2-ene	5D2	309.49	312.181	2.691	0.870
3-methyl-1-butene	5D3	293.21	301.645	8.435	2.877
2-methyl-1-butene	5D4	304.3	301.645	2.655	0.872
2-methyl-2-butene	5D5	311.7	307.133	4.567	1.465
heks-1-ene	6D1	336.63	339.351	2.721	0.808
heks-2-ene	6D2	341.03	343.913	2.883	0.845
heks-3-ene	6D3	340.24	345.363	5.123	1.506
4-methyl-1-pentene	6D4	327.02	334.434	7.414	2.267
2-methyl-1-pentene	6D5	335.28	334.434	0.846	0.252
4-methyl-2-pentene	6D6	331.76	339.351	7.591	2.288
2-methyl-2-pentene	6D7	340.46	340.909	0.449	0.132
3-methyl-1-pentene	6D8	327.33	332.706	5.376	1.642
3-methyl-2-pentene	6D9	343.59	337.754	5.836	1.698
2-ethyl-1-butene	6D10	337.83	332.706	5.124	1.517
3,3-dimethyl-1-butene	6D11	314.4	327.218	12.818	4.077
2,3-dimethyl-1-butene	6D12	328.77	329.101	0.331	0.101
hept-1-ene	7D1	366.79	366.992	0.202	0.055
hept-2-ene	7D2	371.1	370.936	0.164	0.044
hept-3-ene	7D3	368.82	372.819	3.999	1.084
2-methyl-1-hexene	7D4	365.15	362.786	2.364	0.647
2-methyl-2-hexene	7D5	368.56	368.995	0.435	0.118
5-methyl-1-hexene	7D6	358.46	362.786	4.326	1.207
2-methyl-3-hexene	7D7	359.15	368.995	9.845	2.741
5-methyl-2-hexene	7D8	361.26	366.992	5.732	1.587
3-methyl-1-hexene	7D9	357.05	360.573	3.523	0.987
3-methyl-2-hexene	7D10	368.33	364.924	3.406	0.925
2-ethyl-1-pentene	7D11	367.15	360.573	6.577	1.791
3-methyl-3-hexene	7D12	366.69	366.992	0.302	0.082
4-methyl-2-hexene	7D13	360.71	364.924	4.214	1.168
4-methyl-1-hexene	7D14	359.88	360.573	0.693	0.193
2,4-dimethyl-1-pentene	7D15	354.76	358.279	3.519	0.992
2,4-dimethyl-2-pentene	7D16	356.45	364.924	8.474	2.377
3,3-dimethyl-1-pentene	7D17	350.63	353.425	2.795	0.797
2,3-dimethyl-2-pentene	7D18	370.55	362.786	7.764	2.095

Table 5. Cont.

Name	Label	BP [K]	$\widehat{BP}$	Residual	% Residual
2-ethyl-3-methyl-1-butene	7D19	359.51	355.899	3.611	1.004
3,4-dimethyl-2-pentene	7D20	364.65	360.573	4.077	1.118
3,4-dimethyl-1-pentene	7D21	353.95	355.899	1.949	0.551
3-ethyl-1-pentene	7D22	357.26	358.279	1.019	0.285
3-ethyl-2-pentene	7D23	369.16	362.786	6.374	1.727
2,3,3-trimethyl-1-butene	7D24	351.04	350.851	0.189	0.054
4,4-dimethyl-2-pentene	7D25	349.89	360.573	10.683	3.053
4,4-dimethyl-1-pentene	7D26	345.67	355.899	10.229	2.959
okt-1-ene	8D1	394.44	391.194	3.246	0.823
okt-2-ene	8D2	398.08	394.568	3.512	0.882
okt-3-ene	8D3	396.4	396.509	0.109	0.027
okt-4-ene	8D4	395.41	397.143	1.733	0.438
non-1-ene	9D1	420.03	412.712	7.318	1.742
non-2-ene	9D2	423.19	415.602	7.588	1.793
non-3-ene	9D3	421.29	417.457	3.833	0.910
non-4-ene	9D4	420.87	418.364	2.506	0.595
dek-1-ene	10D1	443.34	432.077	11.263	2.541
dek-2-ene	10D2	446.44	434.564	11.876	2.660
dek-3-ene	10D3	444.55	436.283	8.267	1.860
dek-4-ene	10D4	443.94	437.292	6.648	1.497
dek-5-ene	10D5	443.47	437.625	5.845	1.318
average					1.433

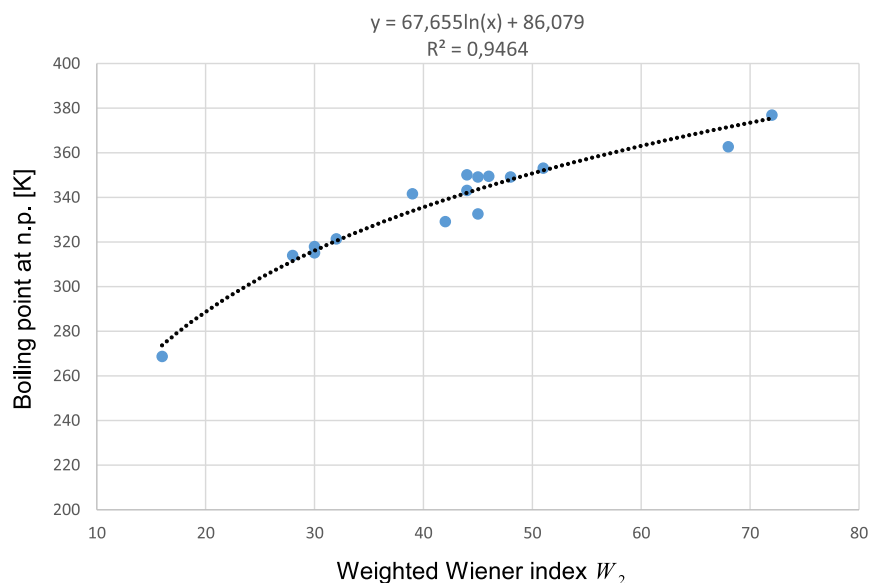
Figure 4. Weighted Wiener indices  $W_2$  and boiling points for the unbranched alkenes.

### 5.2. Alkadienes

Again, we only present the results of the regression between the boiling points and  $W_2$  since the best results are obtained with this index. Similarly, the logarithmic function gives the best fit on the training set, see Figure 5:

$$\widehat{BP} = 67.655 \ln(W_2) + 86.079. \quad (3)$$

The correlation is very good also in this case since  $R^2 = 0.9464$  (the correlation with the Wiener index  $W$  is weaker, that is,  $R^2 = 0.9059$ ).



**Figure 5.** Weighted Wiener indices  $W_2$  and boiling points for the alkadienes in the training set.

After applying Equation (3) on the test set and on all considered alkadienes, the calculated predicted boiling points and the corresponding residuals are shown in Tables 6 and 7, respectively. The boiling points and the predicted boiling points are shown in Figure 6.

As in the case of alkenes, we separately consider all 14 unbranched alkadienes. Again, the correlation is improved with  $R^2 = 0.9600$  (see Figure 7).

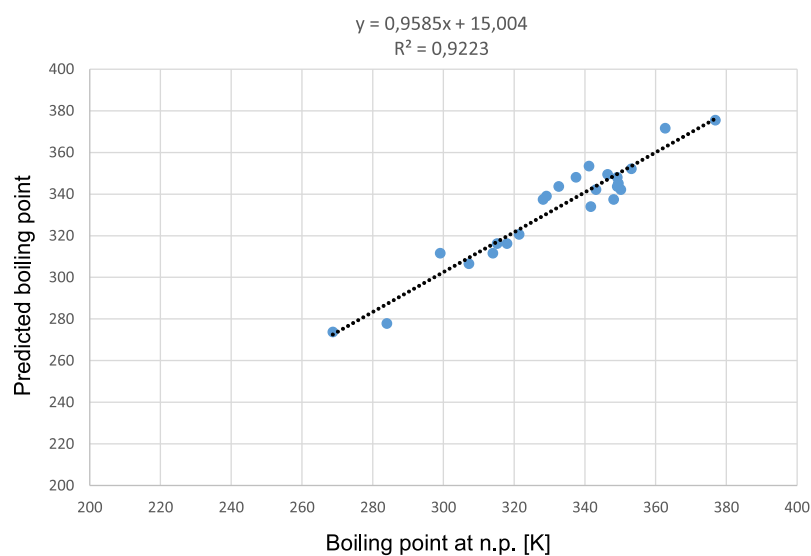
For the unbranched alkadienes we compare our method to the method used in References [15–17], where the valence Wiener index was applied. The correlation between the boiling points and  $W_v$  is weaker with  $R^2 = 0.9242$  (see Figure 8).

**Table 6.** Data for the alkadienes in the test set.

Name	Label	BP [K]	$\widehat{BP}$	Residual	% Residual
1,2-butadiene	4DD1	284	277.760	6.240	2.197
1,4-pentadiene	5DD3	299.12	311.519	12.399	4.145
2-methyl-1,3-butadiene	5DD5	307.22	306.506	0.714	0.233
1,3-heksadiene	6DD2	346.4	349.380	2.980	0.860
1,4-heksadiene	6DD3	337.52	347.985	10.465	3.101
2,3-heksadiene	6DD5	341.15	353.400	12.250	3.591
2-ethyl-1,3-butadiene	6DD9	348.15	337.321	10.829	3.111
3-methyl-1,4-pentadiene	6DD13	328.15	337.321	9.171	2.795
average					2.504

**Table 7.** Data for all considered alkadienes.

Name	Label	BP [K]	$\widehat{BP}$	Residual	% Residual
1,2-butadiene	4DD1	284	277.760	6.240	2.197
1,3-butadiene	4DD2	268.74	273.658	4.918	1.830
1,2-pentadiene	5DD1	318.01	316.187	1.823	0.573
1,3-pentadiene	5DD2	315.18	316.187	1.007	0.320
1,4-pentadiene	5DD3	299.12	311.519	12.399	4.145
2,3-pentadiene	5DD4	321.41	320.553	0.857	0.267
2-methyl-1,3-butadiene	5DD5	307.22	306.506	0.714	0.233
3-methyl-1,2-butadiene	5DD6	314	311.519	2.481	0.790
1,2-heksadiene	6DD1	349.15	347.985	1.165	0.334
1,3-heksadiene	6DD2	346.4	349.380	2.980	0.860
1,4-heksadiene	6DD3	337.52	347.985	10.465	3.101
1,5-heksadiene	6DD4	332.62	343.619	10.999	3.307
2,3-heksadiene	6DD5	341.15	353.400	12.250	3.591
2,4-hexadiene	6DD6	353.15	352.087	1.063	0.301
1,1-dimethylbutadiene	6DD7	349.45	345.106	4.344	1.243
2,3-dimethyl-1,3-butadiene	6DD8	341.66	333.937	7.723	2.260
2-ethyl-1,3-butadiene	6DD9	348.15	337.321	10.829	3.111
2-methyl-1,3-pentadiene	6DD10	349.15	343.619	5.531	1.584
2-methyl-1,4-pentadiene	6DD11	329.15	338.951	9.801	2.978
3-methyl-1,2-pentadiene	6DD12	343.15	342.098	1.052	0.306
3-methyl-1,4-pentadiene	6DD13	328.15	337.321	9.171	2.795
3-methyl-1,3-pentadiene	6DD14	350.15	342.098	8.052	2.299
1,2-heptadiene	7DD1	376.9	375.417	1.483	0.394
1,6-heptadiene	7DD2	362.7	371.550	8.850	2.440
average					1.719



**Figure 6.** Boiling points and predicted boiling points for all considered alkadienes.

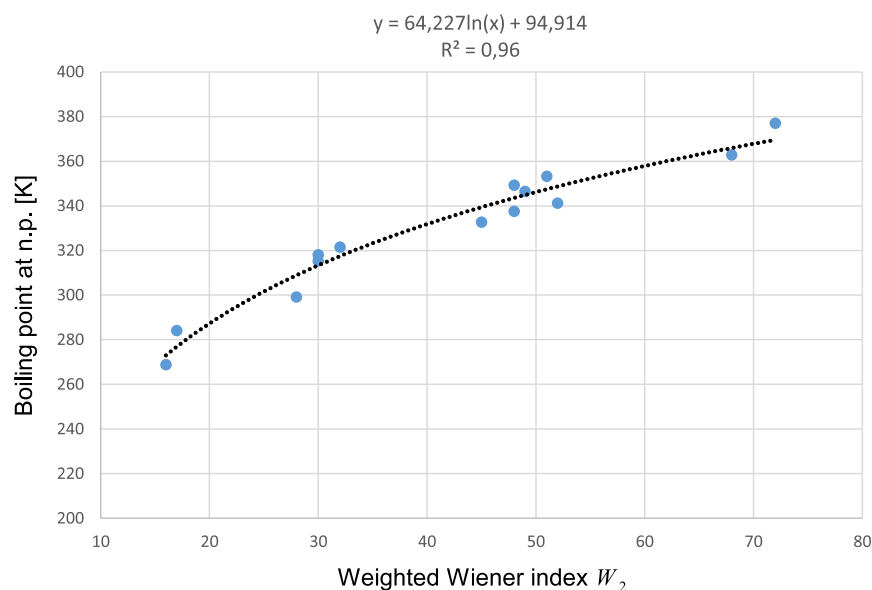


Figure 7. Weighted Wiener indices  $W_2$  and boiling points for the unbranched alkadienes.

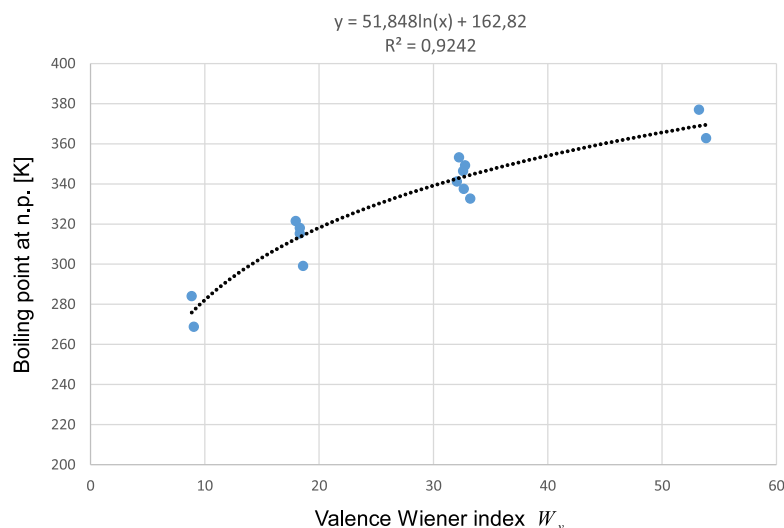


Figure 8. Valence Wiener indices  $W_v$  and boiling points for the unbranched alkadienes.

## 6. Conclusions

In this paper, we have modelled unsaturated hydrocarbons (alkenes and alkadienes) by edge-weighted graphs. In general, a weight of a double edge can be any real number; we have decided to use the weight equal to the number of bonds, that is, 2. On the other hand, it seems reasonable to use a weight less than 1 (for example 1/2) since multiple bonds are shorter than single bonds. Therefore, we have performed statistical analysis for both weights on double edges, that is, 2 and 1/2. We were also interested whether the weighted Wiener index will outdo the usual Wiener index (which does not distinguish between alkanes and alkenes) in the prediction of boiling points, so we have used weight 1 as well. At the end, we investigate correlation between the valence Wiener index and the boiling points for the unbranched alkadienes. However, the correlation for these molecules is better with the weighted Wiener index  $W_2$ .

It also turns out that the correlation coefficient in regression analysis between weighted Wiener indices and boiling points is the best for weight 2 and hence only these results are presented. Consequently, logarithmic models for predicting boiling points of alkenes and alkadienes are deduced. It would be interesting to investigate the physical meaning of

the obtained correlations. Moreover, similar approach can be applied also for unsaturated hydrocarbons with multiple bonds and for other physico-chemical properties. Furthermore, an open problem is to upgrade the described model in such a way that it would be possible to distinguish between different stereoisomers.

**Author Contributions:** All authors contributed equally to this work. Conceptualization, P.Ž.P.; methodology, S.B., N.T. and P.Ž.P.; software, N.T.; formal analysis, S.B., N.T. and P.Ž.P.; investigation, S.B., N.T. and P.Ž.P.; resources, S.B., N.T. and P.Ž.P.; data curation, S.B., N.T. and P.Ž.P.; writing—original draft preparation, S.B., N.T. and P.Ž.P.; writing—review and editing, S.B., N.T. and P.Ž.P.; visualization, S.B.; supervision, N.T. and P.Ž.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Slovenian Research Agency, grant number P1-0297 (N.T. and P.Ž.P.), J1-9109 (N.T., P.Ž.P. and S.B.) and J1-2452 (S.B.).

**Data Availability Statement:** Not applicable.

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

## References

- Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20. [[CrossRef](#)] [[PubMed](#)]
- Hosoya, H. Topological index. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *4*, 2332–2339. [[CrossRef](#)]
- Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener index: Development and applications. *Croat. Chem. Acta* **1995**, *68*, 105–129.
- Janežič, D.; Miličević, A.; Nikolić, S.; Trinajstić, N. *Graph Theoretical Matrices in Chemistry*; University of Kragujevac and Faculty of Science Kragujevac: Kragujevac, Serbia, 2007; pp. 22–24.
- Basak, S.C.; Niemi, G.J.; Veith, G.D. Predicting properties of molecules using graph invariants. *J. Math. Chem.* **1991**, *7*, 243–272. [[CrossRef](#)]
- Sabirov, D.S.; Ori, O.; László, I. Isomers of the C84 fullerene: A theoretical consideration within energetic, structural, and topological approaches. *Fuller. Nanotub. Car. Nanostruct.* **2018**, *26*, 100–110. [[CrossRef](#)]
- Dobrynin, A.A.; Vesnin, A.Y. On the Wiener Complexity and the Wiener Index of Fullerene Graphs. *Mathematics* **2019**, *7*, 1071–1088. [[CrossRef](#)]
- Andova, V.; Orlić, D.; Škrekovski, R. Leapfrog fullerenes and Wiener index *Appl. Math. Comput.* **2017**, *309*, 281–288.
- Yamuna, M.; Divya, T. Boiling point of alkanes and alkenes—from graph eccentricity. *Res. J. Pharm. Technol.* **2018**, *11*, 1962–1970. [[CrossRef](#)]
- Lukovits, I. Wiener indices and partition coefficients of unsaturated hydrocarbons. *Quant. Struct. Act. Relat.* **1990**, *9*, 227–231. [[CrossRef](#)]
- Bangov, I.P.; Moskovkina, M.; Patleeva, A. Charge-related molecular index (CMI)—A novel descriptor for quantitative structure/property relationship (QSPR) models. I. General considerations. *Bulg. Chem. Commun.* **2010**, *42*, 338–342.
- Randić, M.; Pompe, M. On characterization of the CC double bond in alkenes. *SAR QSAR Environ. Res.* **1999**, *10*, 451–471. [[CrossRef](#)]
- Pompe, M.; Veber, M.; Randić, M.; Balaban, A.T. Using variable and fixed topological indices for the prediction of reaction rate constants of volatile unsaturated hydrocarbons with OH radicals. *Molecules* **2004**, *9*, 1160–1176. [[CrossRef](#)] [[PubMed](#)]
- Grossman, S.; Jerman-Blažič, B.; Randić, M. A graph theoretical approach to quantitative structure-activity relationship. *Int. J. Quant. Chem. Quant. Biol. Symp.* **1986**, *12*, 123–139. [[CrossRef](#)]
- Li, X.; Lin, J. The valence overall Wiener Index for unsaturated hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1358–1362. [[CrossRef](#)] [[PubMed](#)]
- Li, X.H. The valence hyper-Wiener index for unsaturated hydrocarbons of chemical trees. *Chem. Phys. Lett.* **2002**, *357*, 255–258. [[CrossRef](#)]
- Li, X.; Hu, M.; Xiao, H. A novel definition of the overall hyper-Wiener index for unsaturated hydrocarbons. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 508–514. [[CrossRef](#)] [[PubMed](#)]
- West, D.B. *Introduction to Graph Theory*; Prentice Hall: Upper Saddle River, NJ, USA, 1996.
- Klavžar, S.; Nadjafi-Arani, M.J. Wiener index versus Szeged index in networks. *Discrete Appl. Math.* **2013**, *161*, 1150–1153. [[CrossRef](#)]
- Yang, P.; Gao, X.H. *Chemical Bonding and Structure-Property Relation*; Higher Education Press: Beijing, China, 1987.
- Floyd, R.W. Algorithm 97: Shortest Path. *Commun. ACM* **1962**, *5*, 345. [[CrossRef](#)]
- Dykyj, J.; Svoboda, J.; Wilhoit, R.C.; Frenkel, M.; Hall, K.R. *Vapor Pressure and Antoine Constants for Hydrocarbons, and Sulfur, Selenium, Tellurium, and Halogen Containing Organic Compounds*; Springer: Berlin, Germany, 1999; pp. 14–71.
- Bonchev, D. Kolmogorov's information, Shannon's entropy, and topological complexity of molecules. *Bulg. Chem. Commun.* **1995**, *28*, 567–582.