On variable Wiener index

Bono Lučić, Ante Miličević, Sonja Nikolić & Nenad Trinajstić*

The Rugjer Bošković Institute, P.O. Box 180, HR-10002 Zagreb, Croatia

Received 16 January 2003

A variable Wiener index \( W_v \), applicable only to acyclic structures, is defined as

\[ W_v = \sum_{e} [n_1(e) n_2(e)]^x \]

where \( n_1(e) \) and \( n_2(e) \) are the numbers of vertices lying on the two sides of the edge \( e \) and the summation is over all edges of the studied tree \( T \).

Recently we proposed a modification of the Wiener index\(^1\), applicable only to acyclic structures. The Wiener index in its original formulation\(^2\) gave greater weights to inner (interior) carbon-carbon (CC) bonds and smaller weights to the outer (terminal) CC bonds of alkanes. This opposes chemists' intuitive reasoning that the outer, more exposed bonds should have greater weights than inner bonds because the outer bonds are associated with larger parts of molecular surface and consequently are expected to make a greater contribution to physico-chemical properties\(^3\). One way to amend this is to modify the Wiener index.

The Wiener index is still attracting mathematical and computational chemists\(^4\) in spite of being the oldest topological index (introduced in 1947) since it is continuously used in QSAR and QSPR modeling\(^9\)-\(^15\). Thus, further studies on this index are warranted.

**Wiener index and its modification**

Originally the Wiener index \( W \) was used for studying properties of alkanes\(^2\). Alkanes can be represented graph-theoretically by trees\(^16\). These trees depict carbon skeletons of alkanes and the maximal vertex-degree is 4.

For alkanes-trees, Wiener's method for computing his index can be formalized by expression (1)

\[ W = \sum_{e} n_1(e) n_2(e) \]

where \( n_1(e) \) and \( n_2(e) \) are the numbers of vertices lying on the two sides of the edge \( e \) and the summation is over all edges of the studied tree \( T \).

Note that \( n_1 + n_2 = n \), where \( n \) is the number of vertices in \( T \).

The modified Wiener index \( mW \) is given by\(^1\):

\[ mW = \sum_{e} [n_1(e)/n_2(e)] \]

Properties of the modified Wiener index have already been studied\(^17\). Some of their results may be summarized as: (i) \( mW \) is not integer-valued in contrast to all other Wiener-type indices\(^3,14\); (ii) \( mW \) is an additive function of edge-contributions; and (iii) \( mW \) correctly accounts the extent of branching of alkanes-trees, that is, it satisfies the inequality:

\[ mW(P_n) < mW(T_n) < mW(S_n) \text{ for all } n \geq 5 \]

where \( P_n \) is any \( n \)-vertex tree, different from the \( n \)-vertex path \( P_n \) and the \( n \)-vertex star \( S_n \). Note that the inverse of the above inequality holds for the Wiener number\(^18,19\):

\[ W(P_n) > W(T_n) > W(S_n) \text{ for all } n \geq 5 \]

Both inequalities hold not just for alkanes-trees but also for all trees.

**Variable Wiener index**

Equations (1) and (2) can be collectively expressed as:

\[ W = \sum_{e} [n_1(e) n_2(e)]^\lambda \]
where \( W \) denotes the variable Wiener number. For \( \lambda = 1 \), Eq. (5) reduces to Eq. (1) and for \( \lambda = -1 \), it reduces to Eq. (2). This report describes our search for the optimal value of \( \lambda \).

The above defined variable Wiener index, Eq. (5), was first time considered by Gutman, Vukičević & Žerovnik. The same authors established various mathematical properties of this novel class of topological indices, but did not examine its applicability in QSPR/QSAR studies.

Since the starting point for this work was our earlier paper, we used the same set of alkane isomers (octanes) and their boiling points. The two-parameter structure-boiling point relationship (6) was employed:

\[
\text{bp} = a W^{1/3} + b P_3 + c \quad \ldots \quad (6)
\]

where \( \text{bp} \) stands for the boiling points of octanes and \( P_3 \) is the polarity number. The same type relationship with the Wiener number was already used by Lukovits and Linert, but their equation is based on the Wiener number and they have defined the polarity number as the number of vertices between graph-theoretical distance equal to 3 and denoted by \( P_2 \). We varied \( W \) from \( \lambda = 1 \) to \( \lambda = -1 \) in order to detect that value of \( \lambda \) which gives the least value of the standard error of estimate \( s \). This is the origin of the term variable in the name of \( W \).

**Results and Discussion:**

The structure-boiling point modeling was based on the CROMRsel procedure. The quality of models is expressed by fitted (descriptive) statistical parameters: the correlation coefficient \( r_{\text{fit}} \), the standard error of estimate \( s_{\text{fit}} \) and \( F \), the result of Fisher's test. In addition, the models were cross(externally)-validated using the one-leave-out procedure. Statistical parameters for the cross-validated models are denoted as \( r_{\text{cv}} \) and \( s_{\text{cv}} \), where \( \text{cv} \) stands for the cross-validation procedure.

The model with the best statistical parameters is obtained for \( \lambda = 1/3 \). This model is explicitly given below:

\[
\text{bp} = 20.315 (\pm 0.643) W^{1/3} + 5.001 (\pm 0.194) P_3 - 222.039 (\pm 10.589)
\]

\[
\text{N}=18 \quad r_{\text{fit}}=0.993 \quad s_{\text{fit}}=0.72 \quad r_{\text{cv}}=0.989 \quad s_{\text{cv}}=0.87 \quad F=503 \quad \ldots \quad (7)
\]

where is \( W^{1/3} \) the variable Wiener number \( W \) with exponent \( \lambda = 1/3 \) and \( N \) is equal to the number of octanes. Note \( s_{\text{fit}} \) was computed with \( N \) in the denominator. Since the values of \( r_{\text{cv}} \) and \( s_{\text{cv}} \) are close to values \( r_{\text{fit}} \) of and \( s_{\text{fit}} \), our model exhibits stability. Note also that \( W^{1/3} \) is not integer-valued and is an edge-contribution-additive function, but it obeys inequality (4) and the outer edges contribute more than inner edges to the value of the Wiener number.

We also give below the model with exponent \( \lambda = -1/3 \):

\[
\text{bp} = -88.981 (\pm 3.255) W^{-1/3} + 4.572 (\pm 0.213) P_3 + 380.072 (\pm 9.887)
\]

\[
\text{N}=18 \quad r_{\text{fit}}=0.990 \quad s_{\text{fit}}=0.83 \quad r_{\text{cv}}=0.986 \quad s_{\text{cv}}=0.98 \quad F=377 \quad \ldots \quad (8)
\]

where \( W^{-1/3} \) denotes the variable Wiener number \( W \) with exponent \( \lambda = -1/3 \). Both models (7) and (8) possess rather close statistical parameters; the model (7) being somewhat better. However, \( W^{1/3} \) satisfies the inequality (3) and the outer edges contribute more than inner edges to the value of the modified Wiener index. From this point of view, model (8) is preferred. It should be noted that in our earlier paper, the model based on \( \lambda = 1 \) had also better statistical parameters \( (r_{\text{fit}}=0.989, \quad s_{\text{fit}}=0.96^\circ C) \) than the model based on \( \lambda = -1 \) \( (r_{\text{fit}}=0.986, \quad s_{\text{fit}}=1.1^\circ C) \) that is intuitively favoured.

In Figs 1 and 2 we give plots between the experimental and calculated values of octane boiling points for fit and cross-validated models (7) and (8).
In Table 1 we give fitted statistical parameters of a number of structure-boiling point models for octanes, taken from the literature. However, these models were not (cross-)validated.

Of all models reported in Table 1, a five-parameter model based on the overall connectivity indices \( 3^{1} \) is better than our 2-parameter models, reported above, judging by the standard error of estimate. A four-parameter model also based on the overall connectivity indices and a three-parameter model based on \( W^{1/4} \), \( P_{3} \) and total walk count \( ^{31} \) are close to our models. Three two-parameter models based on \( W \) and \( P_{3} \), \( mW \) and \( P_{3} \), \( W \) and \( P_{2} \) (the number of vertices between which is the distance equal to three) \( ^{35} \) and the seven-parameter model based on the set of the first seven connectivity indices are also close to our models. But, this last model has too many parameters for modeling a set of only 18 samples. Interestingly, most of acceptable models in this case involve either the Wiener number or some of its variants. Advantage of our models is in the number of parameters — they rest on only two parameters, whilst all other good models needed more parameters.

Acknowledgement

This work was supported by the Ministry of Science and Technology of the Republic of Croatia through grant 0098034.

References

34 Randić M, J Am chem Soc, 97 (1975) 6609.