

Water Activity Prediction

Subjects: Engineering, Chemical

Contributor: Antonio ZUORRO

Water activity is one of the most important factors influencing the quality and stability of food, cosmetic, and pharmaceutical products.

Keywords: water activity ; sugars ; polyols ; Norrish model

1. Introduction

Water activity is one of the most important factors influencing the quality and stability of food, cosmetic, and pharmaceutical products ^{[1][2][3]}. Most of the studies on water activity have been carried out on food products, since this quantity has a significant effect on microbial stability, shelf life, and organoleptic characteristics ^[4].

Formal recognition of the importance of water activity dates back to the early 1950s, when Scott conducted a pioneering study showing that microbial growth and toxin production in food products were dependent on the activity, not the content, of water ^[5].

Although water activity can be rigorously defined in thermodynamic terms ^[6], its molecular origin is still far from being understood, as evidenced by the existence of different and sometimes conflicting explanations ^[7]. The most popular is that of “free water”, according to which water activity reflects its availability as a solvent or reagent, which results from the interactions between water molecules ^[8]. Another explanation attributes its origin to the structuring or ordering of water molecules induced by a solute ^[9]. In particular, a solute can behave as a “structure maker” or a “structure breaker”, depending on its ability to enhance or weaken the hydrogen-bonded water network. A further interpretation is based on the concepts of solute clustering and hydration number, that is, the number of water molecules close to the solute ^[10].

Despite the molecular significance of water activity remaining somewhat elusive, the importance of evaluating this quantity for the systems of interest is evident. For this purpose, many empirical and semi-empirical models have been developed ^{[8][11][12]}. One of the most used is the Norrish model, which provides a good compromise between accuracy and simplicity ^[13]. For a single-solute system, this model contains only one parameter, known as the Norrish constant, which can be easily determined from experimental data. However, in many situations, it may be necessary to estimate water activity in the absence of experimental information.

Molecular descriptors are numeric quantities associated with some structural feature or property of a molecule ^[14]. The use of these descriptors for property prediction has its basis in the principle of similarity, according to which similar molecular structures have similar chemical properties, just as different molecular structures have different chemical properties. Over the years, thousands of molecular descriptors have been used to predict the properties of various substances ^[15]. They can be classified into the following main categories ^{[16][17]}: constitutional; topological; geometrical; and quantum–chemical descriptors. Constitutional descriptors of a compound are based on the number and types of atoms, bonds, rings, etc., in its molecule. Topological descriptors are related to the two-dimensional structure of the molecule, which is regarded as a graph, with vertices representing atoms and edges representing bonds. Geometrical descriptors are derived from the three-dimensional structure of the molecule and consider different molecular features, such as molecular volume, total surface area, and solvent-accessible surface area. Finally, quantum–chemical descriptors are obtained from quantum–mechanical calculations aimed at characterizing the electronic properties of the molecule.

2. Correlation of Water Activity Data

For a single-solute system, the Norrish model provides the following expression for the dependence of water activity (a_w) on composition:

$$a_w = x_w e^{k_N x_s^2},$$

where k_N is the Norrish constant, x_w is the mole fraction of water, and x_s is the mole fraction of the solute. Equation (1) can be derived rigorously from the Kirkwood–Buff theory of solutions, where the thermodynamic meaning of k_N can also be deduced.

The correlation of water activity data by Equation (1) was performed by minimization of the following objective function:

$$\Phi(k_N) = \sum_{i=1}^n (a_{wi,exp} - a_{wi,calc})^2,$$

where n is the number of points of each data set, and the subscripts *exp* and *calc* indicate experimental and calculated values. Since the latter depend on the Norrish constant, it follows that $\Phi = f(k_N)$. The results of the estimation procedure are summarized in **Table 1**, where the mean absolute error, defined as:

$$\varepsilon = \frac{1}{n} |a_{wi,exp} - a_{wi,calc}|,$$

is also reported. The excellent agreement between experimental and calculated results ($2.71 \times 10^{-4} \leq \varepsilon \leq 1.35 \times 10^{-3}$) clearly attests the suitability of the Norrish model to describe the activity of water in the investigated systems.

Table 1. Estimated Norrish constant values (Δx_s : experimental range of solute mole fractions; k_N : Norrish constant; Φ_{min} : minimum value of the objective function; ε : mean absolute error).

Solute	Δx_s	k_N	Φ_{min}	ε
Glucose	0.023–0.111	−2.920	1.71×10^{-5}	1.00×10^{-3}
Fructose	0.026–0.142	−2.351	6.76×10^{-6}	6.64×10^{-4}
Xylose	0.020–0.059	−2.196	1.30×10^{-6}	2.71×10^{-4}
Sucrose	0.018–0.098	−6.777	6.87×10^{-6}	5.85×10^{-4}
Sorbitol	0.018–0.113	−2.494	4.18×10^{-5}	1.35×10^{-3}
Xylitol	0.020–0.112	−2.221	1.12×10^{-6}	7.46×10^{-4}
Glycerol	0.010–0.312	−0.908	1.65×10^{-5}	9.70×10^{-4}
Erythritol	0.007–0.069	−0.950	4.28×10^{-6}	7.08×10^{-4}

3. Use of Molecular Descriptors for the Prediction of the Norrish Constant

Different molecular descriptors were examined for their ability to predict the Norrish constant, with a focus on the classes of constitutional and topological indices.

Constitutional indices are zero-dimensional descriptors. They are the simplest and most used descriptors of a molecule, since they relate to easily determinable molecular features, such as the type of atoms, functional groups, bonds, or number of rings. In this study, the total information index on atomic composition (I_{AC}) was selected to describe the constitutional properties of the solutes. I_{AC} provides information about the type of atoms present in the molecule. This was the first theoretical information index introduced by Dancoff and Quastler in the early 1950s [18]. Since then, an increasing number of studies has revealed the importance of this and other composition-related descriptors in the development of structure–property relationships [19].

Topological indices are two-dimensional descriptors derived from the topological representation of a molecule. In recent decades, molecular topology has emerged as a powerful approach to evaluate structure–activity relationships [20], especially in the fields of pharmacology and toxicology [21][22]. According to this approach, molecular structures are described in terms of the mathematical properties of their associated graphs. For organic compounds, H-depleted graphs (i.e., graphs not including hydrogen atoms) are usually considered, due to the supposed limited contribution of these atoms to molecular connectivity. In this study, the first Zagreb index (Z_1) was selected as a measure of molecular connectivity. Z_1 belongs to the class of the Kier–Hall connectivity indices, and is one of the oldest and most studied topological descriptors [23]. It is related to the concept of vertex valency, and therefore characterizes the degree of atomic branching in the molecule.

With respect to the solutes examined here, it is interesting to consider that glucose and fructose, being isomers, have the same chemical formula. Accordingly, they are characterized by the same I_{AC} value. However, their molecular connectivity shows some differences, which are reflected in the different values of Z_1 . In other words, contrary to the information index on atomic composition, the Zagreb index allows for discrimination between the two isomers. This is an important point to highlight since, as can be seen from the experimental activity data and the estimated Norrish constants (**Table 1**), the two solutes affect the activity of water differently.

To express the dependence of the Norrish constant on the selected descriptors, two empirical models were initially used: the linear model and the exponential model. They are described, respectively, by Equations (4) and (5):

$$k_N = a_{11} + a_{12}I_{AC} + a_{13}Z_1,$$

$$k_N = a_{21}I_{AC}^{a_{22}}Z_1^{a_{23}}.$$

Both models contain three parameters, which were estimated by the least-squares procedure, yielding the results presented in **Table 2**. The quality of correlation was evaluated by calculating the following quantity:

$$\varphi = \frac{\Theta}{n - p},$$

where Θ is the sum of squared errors between experimental and calculated Norrish constants, n is the number of data points, and p is the number of model parameters. It can be noticed that ϕ represents an estimate of the model variance, and can therefore be related to the predictive accuracy of the model.

Table 2. Estimated parameters of the models for predicting the Norrish constant (k_N) from information indices (I_{AC} : information index on atomic composition; Z_1 : first Zagreb index; G : global information index; Θ : sum of squared errors; ϕ : statistical quantity defined by Equation (6)).

Model	Parameters	Θ	ϕ
$k_N = a_{11} + a_{12}I_{AC} + a_{13}Z_1$	$a_{11} = 1.333$ $a_{12} = -8.39 \times 10^{-2}$ $a_{13} = -1.87 \times 10^{-2}$	4.15×10^{-1}	8.29×10^{-2}
$k_N = a_{21}I_{AC}^{a_{22}}Z_1^{a_{23}}$	$a_{21} = -1.44 \times 10^{-2}$ $a_{22} = 1.026$ $a_{23} = 3.81 \times 10^{-1}$	2.67×10^{-1}	5.35×10^{-2}
$k_N = b_{11} + b_{12}G$	$b_{11} = 8.90 \times 10^{-1}$ $b_{12} = -3.98 \times 10^{-2}$	4.98×10^{-1}	8.29×10^{-2}
$k_N = b_{21}G^{b_{22}}$	$b_{21} = -7.65 \times 10^{-3}$ $b_{22} = 1.295$	3.35×10^{-1}	5.58×10^{-2}
$k_N = b_{31}G + b_{32}G^2$	$b_{31} = -2.18 \times 10^{-2}$ $b_{32} = -7.52 \times 10^{-5}$	3.10×10^{-1}	5.17×10^{-2}

References

- Jayaprakasha, H.M.; Jayaraj Rao, K.; Lokesh Kumar, W.A. Studies on the influence of water activity (aw) on the stability of foods—A critical appraisal. *J. Food Sci. Technol.* 1997, **34**, 273–285.
- Kerdudo, A.; Fontaine-Vive, F.; Dingas, A.; Faure, C.; Fernandez, X. Optimization of cosmetic preservation: Water activity reduction. *Int. J. Cosmet. Sci.* 2015, **37**, 31–40.
- Cundell, T. The role of water activity in the microbial stability of non-sterile drug products. *Eur. Pharm. Rev.* 2015, **20**, 58–63.
- Escobedo-Avellaneda, Z.; Rodríguez-Martínez, V.; Serment-Moreno, V.; Velázquez, G.; Welti-Chanes, J.; Antonio Torre, J. Selected Applications of Water Activity Management in the Food Industry. In *Water Activity in Foods: Fundamentals and Applications*, 2nd ed.; Barbosa-Cánovas, G.V., Fontana, A.J., Jr., Schmidt, S.J., Labuza, T.P., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2020; pp. 465–482.
- Scott, W.J. Water relations of *Staphylococcus aureus* at 30 °C. *Aust. J. Biol. Sci.* 1953, **6**, 549–564.
- Reid, D.S. Water Activity: Fundamentals and Relationships. In *Water Activity in Foods: Fundamentals and Applications*, 2nd ed.; Barbosa-Cánovas, G.V., Fontana, A.J., Jr., Schmidt, S.J., Labuza, T.P., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2020; pp. 13–26.
- Maneffa, A.J.; Stenner, R.; Matharu, A.S.; Clark, J.H.; Matubayasi, N.; Shimizu, S. Water activity in liquid food systems: A molecular scale interpretation. *Food Chem.* 2017, **237**, 1133–1138.

8. Labuza, T.P.; Altunakar, B. Water Activity Prediction and Moisture Sorption Isotherms. In *Water Activity in Foods: Fundamentals and Applications*, 2nd ed.; Barbosa-Cánovas, G.V., Fontana, A.J., Jr., Schmidt, S.J., Labuza, T.P., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2020; pp. 161–205.
9. Caurie, M. Water activity of multicomponent mixture of solutes and non-solutes. *Int. J. Food Sci. Technol.* 2005, 40, 295–303.
10. Gharsallaoui, A.; Rogé, B.; Génotelle, J.; Mathlouthi, M. Relationships between hydration number, water activity and density of aqueous sugar solutions. *Food Chem.* 2008, 106, 1443–1453.
11. Chirife, J.; Fontana Jr, A.J. Introduction: Historical Highlights of Water Activity Research. In *Water Activity in Foods: Fundamentals and Applications*, 2nd ed.; Barbosa-Cánovas, G.V., Fontana, A.J., Jr., Schmidt, S.J., Labuza, T.P., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2020; pp. 1–11.
12. Sereno, A.M.; Hubinger, M.D.; Comesaña, J.F.; Correa, A. Prediction of water activity of osmotic solutions. *J. Food Eng.* 2001, 49, 103–114.
13. Norrish, R.S. An equation for the activity coefficients and equilibrium relative humidities of water in confectionery syrups. *Int. J. Food Sci. Technol.* 1966, 1, 25–39.
14. Mauri, A.; Consonni, V.; Todeschini, R. Molecular Descriptors. In *Handbook of Computational Chemistry*; Leszczynski, J., Kaczmarek-Kedziera, A., Puzyn, T., Papadopoulos, M.G., Reis, H., Shukla, M.K., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 2065–2093.
15. Grisoni, F.; Consonni, V.; Todeschini, R. Impact of molecular descriptors on computational models. *Methods Mol. Biol.* 2018, 1825, 171–209.
16. Sahoo, S.; Adhikari, C.; Kuanar, M.; Mishra, B.K. A short review of the generation of molecular descriptors and their applications in quantitative structure property/activity relationships. *Curr. Comput.-Aided Drug Des.* 2016, 12, 181–250.
17. Grisoni, F.; Ballabio, D.; Todeschini, R.; Consonni, V. Molecular descriptors for structure–activity applications: A hands-on approach. *Methods Mol. Biol.* 2018, 1800, 3–53.
18. Bonchev, D.; Trinajstić, N. Chemical information theory: Structural aspects. *Int. J. Quantum Chem.* 1982, 16, 463–480.
19. Barigye, S.J.; Marrero-Ponce, Y.; Pérez-Giménez, F.; Bonchev, D. Trends in information theory-based chemical structure codification. *Mol. Divers.* 2014, 18, 673–686.
20. Stanton, D.T. On the importance of topological descriptors in understanding structure–property relationships. *J. Comput. Aided Mol. Des.* 2008, 22, 441–460.
21. Bueso-Bordils, J.I.; Alemán-López, P.A.; Suay-García, B.; Martín-Algarra, R.; Duarte, M.J.; Falcó, A.; Antón-Fos, G.M. Molecular topology for the discovery of new broad-spectrum antibacterial drugs. *Biomolecules* 2020, 10, 1343.
22. Basant, N.; Gupta, S.; Singh, K.P. Predicting toxicities of diverse chemical pesticides in multiple avian species using tree-based QSAR approaches for regulatory purposes. *J. Chem. Inf. Model.* 2015, 55, 1337–1348.
23. Gutman, I. Degree-based topological indices. *Croat. Chem. Acta* 2013, 86, 351–361.