Computers and Chemical Engineering 83 (2015) 186–202

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Robust chemical product design via fuzzy optimisation approach

Lik Yin Ng, Nishanth G. Chemmangattuvalappil*, Denny K.S. Ng

Department of Chemical and Environmental Engineering/Centre of Excellence for Green Technologies, The University of Nottingham Malaysia Campus, Broga Road, 43500 Semenyih, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 11 October 2014
Received in revised form
11 December 2014
Accepted 8 January 2015
Available online 17 January 2015

Keywords:
Product design
Inverse design techniques
Fuzzy optimisation
Property prediction uncertainties

ABSTRACT

Traditionally, the design of new chemical products for specific applications is done by using a combination of design heuristics, experimental studies and expert judgements. In addition to the conventional methods, chemical products can also be designed by using computer-aided molecular design (CAMD) techniques. Based on CAMD, optimal chemical products can be designed by identifying the molecule with the best properties that correspond with the target functionalities of the product. In general, the optimality of product property (termed as property superiority) is the only factor considered while designing optimal products by using CAMD techniques. However, it is noted that property prediction models are developed with certain accuracy and uncertainties. As the accuracy of property prediction models (termed as property robustness) can affect the effectiveness of CAMD techniques in predicting the product property, the effects of property prediction uncertainty have to be considered while applying CAMD techniques. This paper presents a systematic fuzzy optimisation based molecular design methodology. The methodology is developed for the design of optimum molecules used in chemical processes by considering and optimising both property superiority and robustness. Property superiority is quantified by property optimality. Meanwhile, property robustness is expressed by the standard deviation of the property prediction model, which is a measure of average variation between the experimental data and estimated values of product property using property prediction model. Fuzzy optimisation approach is extended in this work to address and trade off property superiority and robustness simultaneously. Molecular design technique is adapted in this work to identify the optimal molecular structure which satisfies multiple product specification. To illustrate the proposed method, a case study is presented where optimal solution is selected based on how much the solution satisfied the criteria of property superiority and robustness.

1. Introduction

1.1. Chemical product design

A chemical product can be defined as a system consists of different chemical substances which is designed and manufactured for one or more purposes (Cisternas and Gálvez, 2006). According to Cussler et al. (2010), chemical products can be generally categorised into three types. The first type is commodity chemicals such as acids and alcohols, for which the design goal is to manufacture these chemicals at minimum cost. The second type of chemical product is molecular products, such as pharmaceutical drugs. For this type of product, the speed in discovering and developing the products is more vital than the manufacturing cost of the products. The third type of chemical product is performance products. The value of this product is dependent on its functions, which are normally defined by the structure of the products. Although these distinct types of chemical products differ in their design key step and selling point, the design procedures for these products are identical and similar. According to Cussler and Moggridge (2001), the main purpose of chemical product design is to identify optimal products to be made for a specific application. Moggridge and Cussler (2000) proposed that the entire chemical product design process can be represented by four principal steps as follows:

(1) Identify customers' needs
(2) Generate ideas to meet the needs
(3) Select among ideas
(4) Manufacture products

Traditionally, the practise in searching for new chemical products with optimal performance is usually based on design

* Corresponding author. Tel.: +60 3 8924 8122; fax: +60 3 8924 8017.
E-mail address: Nishanth.C@nottingham.edu.my (N.G. Chemmangattuvalappil).

http://dx.doi.org/10.1016/j.compchemeng.2015.01.007
0098-1354/

Nomenclature

CI	connectivity index
GC	group contribution
QSPR/QSAR	quantitative structure property/activity relationships
TI	topological index
χ	connectivity index
ε	edge adjacency index
G	molecular sub-graph
h	height of signature
N_i	number of occurrence of first order group of type-i
M_j	number of occurrence of second order group of type-j
O_k	number of occurrence of third order group of type-k
C_i	contribution of the first order group of type-i
D_j	contribution of the second order group of type-j
E_k	contribution of the third order group of type-k
x	number of signatures
Ω_p	normalised property operator for target property p
v_p^{LL}	lower lower limit for target property p
v_p^{L}	lower limit for target property p
v_p^{LU}	lower upper limit for target property p
v_p^{UL}	upper lower limit for target property p
v_p^{U}	upper limit for target property p
v_p^{UU}	upper upper limit for target property p
V_p	value for target property p
λ_p	degree of satisfaction for target property p
λ_p^{r}	degree of satisfaction for property robustness for target property p
λ_p^{s}	degree of satisfaction for property superiority for target property p
λ_p^{*}	degree of satisfaction for target property p in two-phase approach
σ_p	standard deviation for property prediction model for target property p
R^2	coefficient of determination
$\log V_E$	affinity of fungicide
$\log \mu$	mobility of fungicide
$\log[R/(1-R)]$	retention
LC_{50}	lethal concentration

heuristics, experimental studies and expert judgements or experiences (Odele and Macchietto, 1993). These methods start from the identification of molecule from raw material, and search for the required and preferred properties from the molecule identified. As they are mainly based on trial and error approaches, these traditional methods are intrinsically inefficient, time consuming and costly (Venkatasubramanian et al., 1994). Furthermore, as these approaches are largely dependent on the available information and knowledge, it is challenging and difficult to search for new chemical products which possess optimal properties without systematic selection tools (Churi and Achenie, 1996). On the other hand, chemical product design can be done through reverse engineering approaches, where the design process begins with the identification of needs to fulfil, and search for the molecule that possesses properties which can meet the needs (Gani et al., 1991). In most cases, suitability and performance of a product are usually defined in terms of physical properties rather than chemical structure of the product. Hence, chemical product design problem can be considered as an inverse property prediction problem where the preferred attributes of the product are represented in terms of physical target properties, and the objective of the problem is to determine the molecular structure that matches these properties (Gani and O'Connell, 2001). As product specifications are often extracted from customer needs, it is required to translate qualitative attributes into quantitative parameters in order to design a product (Achenie et al., 2003). The process of representing product attributes by using measurable product properties is often done by computer-aided molecular design (CAMD) techniques.

1.2. Computer-aided molecular design (CAMD)

CAMD techniques are important for chemical product design for their ability in predicting, estimating and designing molecules with a set of predefined target properties (Harper and Gani, 2000). Property estimation is normally done by utilising property prediction models in predicting molecular properties from structural descriptors (Gani and Pistikopoulos, 2002). Some of the commonly used structural descriptors to quantify molecular structure include chemical bonds and molecular geometry (Randić et al., 1994). As mentioned previously, CAMD techniques employ property prediction models in inverse property prediction problems to estimate the molecular structure from a set of target properties. Currently, most of the CAMD techniques use property prediction models based on group contribution (GC) methods to verify and ensure that the generated molecules possess the specified set of desirable properties (Harper et al., 1999). By utilising molecular groups as structural descriptors, GC methods estimate the property of the molecule by summing up the contributions from the molecular groups in the molecule according to their appearance frequency (Ambrose, 1978). Property prediction models based on GC methods are widely used for property estimation because these models are simple to apply yet provide reasonably accurate predictions for many properties. Moreover, they can provide quick property estimations without significant errors and expensive computational effort (Constantinou et al., 1993). However, the early GC property prediction models become less reliable as the complexity of the molecule increases. As molecular groups were assumed to be independent and non-overlapping, resonance, conjugation and proximity effects were not taken into account (Mavrovouniotis, 1990). Hence, the models cannot differentiate between isomers and capture the interactions among different molecular groups. Constantinou and Gani (1994) presented an improved GC approach by defining the molecular groups as first and second order molecular groups. The basic level is called as first order molecular groups while the next higher level is known as second order molecular groups. Second order molecular groups are developed and defined by having the first order molecular groups as their building blocks. These second order molecular groups represent different types of interactions and the effect of these interactions among the first order molecular groups. Hence, isomers and compounds with functional groups can be distinguished. GC methods are later further extended by Marrero and Gani (2001) by identifying and incorporating third order molecular groups into the property prediction model. The formation of third order molecular groups is analogous to the second order molecular groups, but their contribution have been correlated to focus on molecular fragments or compounds whose description is insufficient through first and second order molecular groups. These include polyfunctional and structural groups such as multi-ring compounds, fused ring compounds and compounds which consist of various functional groups. A general representation of property model by using GC methods can be shown with the following equation:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (1)$$