INTERNATIONAL JOURNAL OF HYDROGEN ENERGY 39 (2014) 5733–5746

Available online at www.sciencedirect.com

ScienceDirect

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

Knowledge extraction for water gas shift reaction over noble metal catalysts from publications in the literature between 2002 and 2012

Çağla Odabaşı [a], M. Erdem Günay [b], Ramazan Yıldırım [a,*]

[a] Department of Chemical Engineering, Boğaziçi University, 34342 Bebek, Istanbul, Turkey
[b] Department of Energy Systems Engineering, Istanbul Bilgi University, 34060 Eyup, Istanbul, Turkey

ARTICLE INFO

Article history:
Received 18 November 2013
Received in revised form
21 January 2014
Accepted 24 January 2014
Available online 20 February 2014

Keywords:
Water gas shift reaction
Data mining
Knowledge extraction
Artificial neural networks
Decision trees
Support vector machines

ABSTRACT

In this work, a database (containing 4360 experimental data points) on water gas shift reaction (WGS) over Pt and Au based catalysts was constructed using the data obtained from the published papers between the years 2002 and 2012. Then, the database was analyzed using three data mining tools to extract knowledge in three areas: Decision trees to determine the empirical rules and conditions that lead to high catalytic performance (high CO conversion); artificial neural networks (ANNs) to determine the relative importance of various catalyst preparation and operational variables and their effects on CO conversion; support vector machines (SVMs) to predict the outcome of unstudied experimental conditions. It was concluded that, all three models were quite successful and they complement each other to extract knowledge from the past published works and to deduce useful trends, rules and correlations, which are not easily comprehensible by the naked eyes.

1. Introduction

Researches for more efficient and environmentally friendly energy technologies have increased in recent years as a result of continuous increase in world energy demand and growing concerns for environment. The fuel cell technology, which electrochemically converts hydrogen and oxygen to electricity and water, seems to be one of the most promising energy conversion systems for the future. Due to the technical difficulties in hydrogen storage, on-site production of hydrogen from a hydrocarbon using a fuel processor is considered to be the most feasible solution for small size proton exchange membrane (PEM) fuel cell applications in the near future. However, the hydrogen stream produced by the reforming process contains CO, CO_2 and H_2O. Especially, CO is known to be harmful to the anode catalyst of the PEM fuel cell even in trace amounts; hence, it must be totally eliminated [1]. It was suggested that water–gas shift (WGS) reaction can remove most of the CO coming from reformer (also producing some additional hydrogen), followed by the preferential CO oxidation (PROX) for the complete elimination of the remaining about 1% CO [2].

WGS is a commercially well-established two-step process: high temperature shift (573–723 K) over Fe–Cr based catalysts and low temperature shift (473–523 K) over Cu–Zn catalysts to

* Corresponding author. Tel.: +90 212 359 7248; fax: +90 212 287 2460.
E-mail address: yildirra@boun.edu.tr (R. Yıldırım).

0360-3199/$ – see front matter
http://dx.doi.org/10.1016/j.ijhydene.2014.01.160

reduce the amount of CO further. However, these catalysts are not suitable for fuel cell applications due to their low activity and stability as well as their special pretreatment and regeneration requirements [2]. Instead, noble metal catalysts (such us Pt, Rh, Pd, Ru and Au) over various supports (such as CeO_2, Al_2O_3, ZnO and MgO) have been investigated extensively for fuel cell applications in recent years.

WGS activity of noble metal catalysts seems to depend on very large number of variables and the complex interactions among them. The catalyst preparation variables such as the type and the amount of noble metal, type of the support, type and the amount of promoter (if it is used) and the choice of the catalyst preparation method have significant influence on the catalytic activity. The same is true for the operational variables such as the reaction temperature, the space time and the feed composition. For example, the reaction temperature plays a crucial role on the CO conversion since WGS reaction is equilibrium limited and mildly exothermic; at low temperatures, the reaction is kinetically controlled whereas it is limited by thermodynamic equilibrium at high temperatures [3]. Similarly, the feed composition (especially CO/H_2O ratio) has a substantial impact on the catalytic performance.

It is quite clear that the most suitable catalyst preparation and operational conditions cannot be determined in a single experimental work considering the presence of large number of variables in wide experimental ranges; that is why researchers usually survey the literature first and focus their efforts on the experimental region that has not been explored yet and seemingly promising for the good catalytic performance. However, simple manual review of literature may not be sufficient anymore to find the right answers because the publications are too large in number (i.e. hundreds for WGS reactions), scattered over various sources, and containing significant amount of non-uniformities, gaps, mismatches and even conflicts. Hence, we need more systematic and effective approaches and tools to extract knowledge from the published literature. If we can do that, this massive accumulation of experience will become an invaluable source; the databases created from these works can be easily used to develop models and deduce patterns and rules. This may be considered as analogous to interpret the results of a statistical experimental design. Although the data points in the literature are not created statistically and not performed under the same conditions as a statistical experimental design, the existence of large number of data points in the entire experimental spectra may still allow us to capture some useful information by using some effective data mining tools, which are readily available thanks to the fast developments in computational resources in the recent years.

Data mining is a field of computer science to extract knowledge in a database. It helps to spot the patterns that are too hard to detect with the naked eyes, and use these patterns to derive conclusions or make predictions using classification (like decision trees and support vector machines (SVM)), clustering (such as k-means clustering and Kohonen networks) and estimation (such as artificial neural networks (ANN) and SVM) techniques [4]. These techniques have become quite widespread in many fields in the last two decades.

ANNs are one of the most common tools used for approximation [4,5]; various successful applications of ANN in chemistry and chemical engineering as well as in the field of catalysis were discussed in details elsewhere [6]. Decision trees, on the other hand, can help to derive simple but valuable rules, such as finding the variables leading to high or low catalytic performance levels. Although there are various applications of this tool in diverse areas, their use is quite new in the field of catalysis [7–9]. Similarly, the SVMs, which can be used both for classification and estimation, had some limited application on catalysis such as the work of Baumes et al. [10] and Chae et al. [11] although it is widely used in other fields.

In all the works referenced above, the data mining tools are employed to analyze the data produced in single set of experimental results generated by the same group. To the best of our knowledge, there are only four works (three of them are ours) that involve the analysis of the entire literature in the field of catalysis [12–14].

Table 1 – Input variables and their ranges.

Input variable	Range (for continuous variables) or identity (for categorical variables)
Base metal type and weight %	Pt (0–27), Au (0–8), Ru (0–3), Rh (0–3), Ir (0–3), Cu (0–40), Pd (0–3)[a]
Preparation method	incipient to wetness impregnation (IWI), wet impregnation (WI), co-impregnation (CI), sequential impregnation (SI), sol–gel precipitation (SGP), co-precipitation (CP), homogenous deposition precipitation (HDP), Urea Gelation co-precipitation (UGC), solution combustion technique (SCT), flame spray pyrolysis (FSP), micro emulsion (ME), deposition precipitation (DP)
Calcination temperature (°C)	200–700
Calcination time (h)	1–10
Support type	Al_2O_3, MgO, CeO_2, TiO_2, MnO, Y_2O_3, ZrO_2,Tb_4O_7, HfO_2, La_2O_3, Co_3O_4, ThO_2, SiO_2, Fe_2O_3, Sm_2O_3, Gd_2O_3, Yb_2O_3,CaO, zeolite, hydroxyapatite, activated carbon
Promoter type and weight %	Li (0–1), Ce (0–79), Co (0–20), Mg (0–10), Fe (0–10), Mn (0–10), Zr (0–1), K (0–4), Ni (0–15), Ca (0–10), Cs (0–9), V (0–6), Rb (0–14), Y (0–10), Na (0–10), La (0–12), Gd (0–10), Yb (0–10), Zn (0–10), Re (0–1), Ti (0–10), Cr (0–10), Ho (0–10), Nd (0–10), Tm (0–10), Sm (0–10), Er (0–10),Sr (0–2), YSZ (0–1)[a]
Reaction temperature (°C)	5–810
H_2 vol.% in the feed	0–60
O_2 vol.% in the feed	0–1.4
CO vol.% in the feed	0.2–99.5
H_2O vol.% in the feed	0–93
CO_2 vol.% in the feed	0–25
CH_4 vol.% in the feed	0–51
time on stream (min)	0–5900
W/F ($cm^3\ min^{-1}\ g^{-1}$)	0.006–40

[a] Weight percent of the metals in the catalyst.