

PHOTOCATALYTIC HYDROGEN PRODUCTION OVER IONIC LIQUID COATED
SEMICONDUCTORS

by

Elif Can

B.S., Chemical Engineering, Boğaziçi University, 2013

M.S., Chemical Engineering, Boğaziçi University, 2015

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Chemical Engineering
Boğaziçi University

2021

To my family

From one to sixty-two

ACKNOWLEDGEMENTS

This PhD thesis would not have been possible without moral and material support of many people, and I would like to start mentioning them.

First, I would like to express my sincere gratitude to my thesis supervisor, Prof. Ramazan Yildirim, for his great enthusiasm and guidance throughout of this thesis. His training in all aspects of being a good academic researcher with an endless patience and encouragement has been appreciated greatly and will not be forgotten. Besides his academic wisdom, I enjoy and learn a lot from our gourmet talks about any kind of food.

I would like to acknowledge my thesis committee, Prof. Ahmet Erhan Aksoylu, Prof. Ahmet Kerim Avcı, Prof. Deniz Üner, and Assoc. Prof. Alper Uzun for their valuable contribution and guidance in the progression of this dissertation. I am also grateful to all ChE Faculty Members, I learned a lot from their experience and knowledge.

In this long PhD journey, I am very grateful for having such lovely friends of Meltem Baysal, Asuman Okur, Elif F. Yener, Kübra Yazıcı, Hatice K. Anakök, Elif Erdinç, and Burcu Acar. Very special and sincere thanks to Elif Gençtürk, without her generous support and great friendship, this journey would be unbearable.

I would like to thank Serhat Erşahin, Pınar Özdemir, Burcu Oral, and all other SOLENCAT members for their support and friendship. I greatly appreciate the help and contribution of Dr. Ahsan Jalal, İrem Gülçin Zırhlıoğlu, and Dr. Betül Uralcan to my thesis from different aspects. Also, many thanks to Yakup Bal for his technical assistance and kindness to me.

Last but not least, I would like to express my deepest thanks to my family members Yılmaz Can, Gülsüm Can, İpek Okutan, Z. Betül Can, Tunahan Okutan, E. Zülal Okutan, and S. Agah Okutan. I always feel their endless love and incredible support with me no matter the distance and the time we are apart.

The financial support provided by The Scientific and Technological Research Council of Turkey (TUBITAK, Project Number 217M528) and Bogazici University Scientific Research Projects (BAP, Project Number 14002) is gratefully acknowledged.

ABSTRACT

PHOTOCATALYTIC HYDROGEN PRODUCTION OVER IONIC LIQUID COATED SEMICONDUCTORS

The objective of this thesis is to develop active and stable dye sensitized and ionic liquid (IL) encapsulated photocatalysts for hydrogen production, and to investigate the roles of photocatalysts constituents on hydrogen production rate. First, comprehensive experimental datasets for photocatalytic (PWS) and photoelectrochemical (PECWS) water splitting were constructed from literature while a large dataset for water solubility in ILs were created computationally using DFT and COSMO-RS. Then, these datasets were analyzed by machine learning (ML); association rule mining (ARM), decision tree (DT), random forest (RF) and deep learning (DL) were implemented in R and Python environment. In PWS analysis, the clear trends and the high fitness of the models constructed, especially those involved the band gap, indicates that PWS over perovskites is predictable but the sound solutions like ion doping, use of co-catalyst, or use of sacrificial donors did not help as much as desired. The prediction accuracies of DL and classification accuracy of DT models for the water solubility in ILs were also high, and the properties of anionic parts of ILs are more influential for water solubility. The band gap of semiconductors used in PECWS were also predicted successfully even though the prediction accuracy of photocurrent density was not satisfactory. In the experimental part of the thesis, the enhancement of photocatalytic hydrogen production over 1 wt% Pt/TiO₂ sensitized by N719 dye and coated by a thin layer of ionic liquid ([BMIM][BF₄]) has been investigated; SEM-EDAX, CTEM, FTIR, XRD, and UV-Vis characterization of the photocatalysts, and electrochemical analysis of the respective photoelectrodes were performed. The IL coating increased the performance of 1%Pt/TiO₂ by providing better charge transfer between the photocatalyst and the aqueous reaction medium while simultaneously preventing the recombination of photogenerated electron-hole pairs; the improvement was much higher with the use of IL and N719 together. The performance of 1% Pt/SrTiO₃ was also tested but the reproducible results could not be obtained. Finally, the factors related to the structure of the reactor and operational conditions (like dead volume, gas-liquid interfacial area, sweep gas flowrate and stirring speed) were also found to be influential over the hydrogen production rate.

ÖZET

İYONİK SIVI KAPLI YARI İLETKENLER KULLANILARAK FOTO KATALİTİK HİDROJEN ÜRETİMİ

Bu tez çalışmasının amacı fotokatalitik hidrojen üretiminde kullanmak üzere boya ile duyarlılaştırılmış ve iyonik sıvı (IL) ile kaplanmış, aktif ve dayanıklı fotokatalizörler üretmek ve onu oluşturan parçaların hidrojen üretim hızına etkilerini incelemektir. Öncelikle, suyun fotokatalitik (PWS) ve fotoelektrokimyasal (PECWS) ayrıştırılması alanında literatürde bulunan deneysel sonuçlardan ve suyun IL içindeki çözünürlüğü ile ilgili ise DFT COSMO-RS kullanılarak oluşturulan üç farklı veri tabanı yapay öğrenme yöntemleri ile analiz edilmiştir. Bunun için birliktelik kuralları analizi (ARM), karar ağacı (DT), rassal orman (RF) ve derin öğrenme (DL) gibi yöntemler kullanılmış, analizler R ve Python ortamında yapılmıştır. PWS analizinde modellerin veriye yüksek uyumu, perovskitler ile yürütülen PWS deney sonuçlarının tahmin edilebilir olduğunu göstermiş, ancak verimi arttırmak için yaygın olarak uygulanan iyon doplama, eş-katalizör veya yapay donör kullanmanın istenildiği kadar etkili olmadığı ortaya konulmuştur. Suyun IL içinde çözünürlüğü DL ile başarılı bir şekilde tahmin edilebilirken, DT ile yapılan sınıflandırma çalışmaları da başarılı olmuştur; bu modeller suyun çözünürlüğü için iyonik sıvıların anyonik kısmına ait özelliklerinin daha belirleyici olduğunu göstermiştir. PECWS’de kullanılan yarı iletkenlerin enerji bant aralıkları yüksek doğruluk payı ile modellenmiş olsa da fotoakım yoğunluk değerleri iyi tahmin edilememiştir. Tezin deneysel kısmında ise, N719 boya ile duyarlılaştırılmış ve iyonik sıvı ([BMIM][BF₄]) ile kaplanmış %1 Pt/TiO₂’nin fotokatalitik hidrojen üretimine sağladığı iyileştirme araştırılmıştır. Üretilen fotokatalizörlerin SEM-EDAX, CTEM, FTIR, XRD, UV-Vis ve elektrokimyasal analizleri yapılmıştır. İyonik sıvının hem fotokatalizör ve reaksiyon ortamı arasında daha iyi yük transferi sağlayarak hem de foton yoluyla üretilmiş e⁻ ve deşğin yeniden birleşimini engelleyerek verimi arttırdığı gözlemlenmiştir. İyileşme, IL ve N719’un birlikte kullanıldığı durumlarda daha yüksek olmuştur. Aynı deneyler %1 Pt/SrTiO₃ katalizörü ile de yapılmış ancak tekrarlanabilir sonuçlar alınamamıştır. Son olarak reaktörün yapısal özellikleri ve çalışma koşullarına ait faktörlerin de (reaktördeki boşluk hacmi, gaz-sıvı ara yüzey alanı, taşıyıcı gaz debisi ve karıştırma hızı) hidrojen üretimi üzerinde etkili oldukları görülmüştür.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT.....	v
ÖZET	vi
LIST OF SYMBOLS	xxiv
LIST OF ACRONYMS/ABBREVIATIONS	xxv
1.INTRODUCTION	1
2. LITERATURE SURVEY	6
2.1. Photocatalytic Hydrogen Production.....	6
2.2. Water Splitting in Electrochemical Cell.....	10
2.3. Dye Sensitization for Visible Light Harvesting	14
2.4. Ionic Liquids and Their Applications in Chemical Engineering.....	16
2.5. Machine Learning in Catalysis, Photocatalysis, and Ionic Liquids	21
3. MATERIALS AND METHODS.....	26
3.1. Computational Details	26
3.1.1. Computational Details for Photocatalytic Water Splitting Over Perovskites	28
3.1.2. Computational Details for Photoelectrochemical Water Splitting.....	32
3.1.3. Computational Details for Water Solubility in ILs.....	36
3.2. Details About Experimental Work	43
3.2.1. Materials	43
3.2.2. Catalyst and Photoelectrode Preparation	44
3.2.3. Characterization	47
3.2.4. Photocatalytic Measurements	47
3.2.5. Photoelectrochemical Measurements.....	48
3.2.6. Critical Factors in Reactor Design and Operations.....	48
4. RESULTS and DISCUSSION	51
4.1. Machine Learning in Photocatalytic Hydrogen Production Over Perovskites Literature	51

4.1.1 Perovskite Materials.....	51
4.1.2. Ion Doping for Band Gap Modification.....	54
4.1.3. Preparation Methods for Perovskites	62
4.1.4. Co-Catalyst for Charge Separation	65
4.1.5. Thermal Treatment.....	69
4.1.6. Sacrificial Agents and Additives	71
4.1.7. Light Sources	73
4.1.8. Stability of Perovskites	74
4.1.9. Association Rule Mining for Factor Effects on Hydrogen Production...	74
4.1.10. Decision Tree Analysis to Determine Conditions for High Hydrogen Production	81
4.1.11. Random Forest Analysis to Predict Hydrogen Production	82
4.1.12. Analysis And Modeling of Band Gaps	83
4.2. Machine Learning in Photoelectrochemical Water Splitting Literature	84
4.2.1. Materials for PECWS	86
4.2.2. Methods For Semiconductor Synthesis and Electrode Fabrication	91
4.2.3. Light Properties.....	95
4.2.4. Electrolyte Properties.....	95
4.2.5. Association Rule Mining	98
4.2.6. Band Gap Prediction by Random Forest Algorithm.....	100
4.2.7. Band Gap Classification by Decision Tree Algorithm	101
4.2.8. Photocurrent Density at Bias of 1 V Classification by Decision Tree Algorithm.....	104
4.2.9. Photocurrent Density Classification	104
4.3. Machine Learning for Predicting Water Solubility in Ionic Liquids	107
4.3.1. Pre-analysis of Data	107
4.3.2. Analysis of Descriptor Effects by Association Rule Mining.....	108
4.3.3. Heuristics for Cation/Anion Pairing by Decision Tree.....	111
4.3.4. Predictive Models for Water Solubility by Deep Learning	118
4.3.5. Verification of Models Using Experimental Data from Literature.....	124

4.4. Experimental Studies Using SrTiO_3 as Photocatalyst in PWS Reaction	126
4.4.1. Structural Modifications of SrTiO_3	129
4.4.2. The Effect of Synthesis Temperature for SrTiO_3	132
4.4.3. The Effect of Calcination Temperature After Pt Impregnation onto SrTiO_3	133
4.4.4. The PWS Experiments with Au As Promoter.....	134
4.4.5. The Effect of Production Method for SrTiO_3	135
4.4.6. The Characterization of SrTiO_3 Prepared by Solid State Reaction	137
4.5. Experimental Studies Using TiO_2 as Photocatalyst in PWS Reaction.....	139
4.5.1. Physical and Chemical Characterization	141
4.5.2. Photocatalytic Activity of Pt/TiO_2 , Dye:Pt/TiO_2 , IL-Dye:Pt/TiO_2	146
4.5.3. Electrochemical Analysis of TiO_2 , IL:TiO_2 and IL-Dye:TiO_2 Photoelectrodes	150
4.6. Study Of Critical Factors Affecting the Hydrogen Evolution Via Mass Transfer Limitations.....	151
4.6.1. Effect of Dead Volume	154
4.6.2. Effect of Liquid to Gas Mass Transfer Area.....	155
4.6.3. Effect of Sweep Gas Flow Rate	157
4.6.4. Effect of Stirring Speed	159
5. CONCLUSIONS AND RECOMMENDATIONS	162
5.1. Conclusions	162
5.2. Recommendations	168
REFERENCES	170
APPENDIX A: ADDITIONAL INFORMATION FOR ANALYSIS OF PWS DATABASE	220
APPENDIX B: ADDITIONAL INFORMATION FOR ANALYSIS OF PECWS DATABASE	228
APPENDIX C: ADDITIONAL INFORMATION FOR ANALYSIS OF WATER SOLUBILITY IN ILs DATABASE	232

LIST OF FIGURES

Figure 2.1.	Schematic for photocatalytic water splitting reaction.	7
Figure 2.2.	The basic structures for photocatalytic water splitting (Schwarze <i>et al.</i> , 2013).	8
Figure 2.3.	The real-life examples of PWS reaction systems	9
Figure 2.4.	Number of studies published vs. reported PEC lifetime (Ager <i>et al.</i> , 2015).	12
Figure 2.5.	The classification of PEC systems based upon their electrode configurations (Minggu <i>et al.</i> , 2010).	13
Figure 2.6.	The classification of PEC systems based on their biasing methods (Minggu <i>et al.</i> , 2010).	14
Figure 2.7.	The basic diagram of dye-sensitization in PWS reaction.....	16
Figure 2.8.	The image for ionic liquid [BMIM][BF ₄] (purchased from Sigma Aldrich).	17
Figure 2.9.	The illustration of possible combinations between support, ionic liquid, and catalyst (Van Doorslaer <i>et al.</i> , 2010).	19
Figure 2.10.	The basic diagram of dye-sensitization in PWS.	23

Figure 2.11.	The visual demonstration for algorithms of a) KNN, b) Decision tree, c) Deep learning (Can <i>et al.</i> , 2021).	24
Figure 3.1.	Diagram for application of machine learning tools.	27
Figure 3.2.	Different steps of TiO ₂ based photoelectrode preparation a) Coating of FTO with Ti-nanoxide D/SP paste with doctor blade method, b) After coating, c) After drying at 475 °C, d) Dye sensitized and plain TiO ₂ electrodes in [BMIM][BF ₄]: acetonitrile solution, e) IL coated TiO ₂ electrode.	46
Figure 3.3.	The simple illustration of our photocatalytic reaction system.	48
Figure 3.4.	The pictures of reactors R1, R2, R3, R4, R5, R6, R7, and R8.	49
Figure 4.1.	Average hydrogen production rates based on a) common B-site elements and b) common A-site elements (solid balls are for UV; dashed balls are for visible light).	54
Figure 4.2.	The frequencies and average hydrogen production rates for a) A-site doped NaTaOs, b) B-site doped NaTaOs c) A-site doped SrTiOs, and d) B-site doped SrTiOs.	57
Figure 4.3.	Effects of doping on band gap of most commonly used perovskites.	62
Figure 4.4.	Properties of perovskites prepared by the most commonly used perovskites average a) hydrogen production rates, b) surface area, c) band gap, and d) crystal structure.	67

Figure 4.5.	Promoters used in the database vs. number of instances.	68
Figure 4.6.	Effects of inorganic sacrificial agents on average hydrogen production rates for visible and UV light conditions.	73
Figure 4.7.	Stability of some tantalates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order; *duration of test; **rate in first run.	75
Figure 4.8.	Stability of some titanates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order;* duration of test;**rate in first run.	76
Figure 4.9.	Stability of some niobates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order;* duration of test;**rate in first run.	77
Figure 4.10.	Stability of some indium-based perovskites. For each perovskite, bottom bar is the first run, the others are repeats in increasing order;* duration of test;**rate in first run.	78
Figure 4.11.	The optimal decision tree for visible light data.	82
Figure 4.12.	Predicted versus actual hydrogen production for random forest model for visible light data a) training b) testing.	83
Figure 4.13.	Results of the analysis of band gaps a) decision tree constructed for band gaps b) training results of random forest analysis c) testing results of random forest analysis.	85

Figure 4.14.	Effect of doping elements for TiO_2 on average band gap.	87
Figure 4.15.	The effect of top layer for some of the frequently used materials in middle layer a) ZnO , b) BiVO_4 , c) WO_3 , and d) Fe_2O_3	92
Figure 4.16.	Effect of co-catalyst using on the photoelectrochemical performance of frequently used bottom materials.	92
Figure 4.17.	The effect of synthesis method on the crystal structure and band gap energies are given for the most frequent bottom materials, a) TiO_2 , b) WO_3 , c) Fe_2O_3 , d) BiVO_4 , and e) ZnO	97
Figure 4.18.	The effect of electrolytes for different semiconductors, a) TiO_2 and b) WO_3	98
Figure 4.19.	The effect of bottom layer preparation method on the band gap value of that semiconductor.	100
Figure 4.20.	Predicted versus actual band gap values by using random forest model for a) validation set, b) testing set.	101
Figure 4.21.	The input variables vs a) the reduction in sum of squared errors (IncNodePurity) and b) the percentage increase in mean square error (%IncMSE).	102
Figure 4.22.	The decision tree for band gap classification.	103
Figure 4.23.	Predicted versus actual photocurrent values by using random forest model for a) validation sets, b) testing set.	105

Figure 4.24.	The decision tree for current density classification.	106
Figure 4.25.	Effects of cation descriptors on average water solubility for various IL groups (See Table C.3 in Appendix for the range of bins used for the average values of descriptors).	109
Figure 4.26.	Distribution of anions leading various solubility classes. Upper panel shows high water solubility class (<i>class C</i>) while middle and lower panels indicate intermediate (<i>class B</i>) and low water solubility (<i>class A</i>) classes respectively (<i>A#</i> represents the individual anions; the list of anions corresponding the index number # is given in Appendix C).	115
Figure 4.27.	Optimum decision tree structure for imidazolium dataset.	120
Figure 4.28.	Optimum decision tree structure for entire dataset.	121
Figure 4.29.	Testing results of optimum deep learning model developed for imidazolium dataset. (a) Predicted versus computed solubility for testing, (b) relative importance of descriptors (see Section 3.1.3 for model hyperparameters).	123
Figure 4.30.	Testing results of optimum deep learning model developed for entire dataset. (a) Predicted versus computed solubility for testing, (b) relative importance of descriptors ((see Section 3.1.3 for model hyperparameters).	124
Figure 4.31.	Experimental versus computed (by COSMO) and predicted (by multilayer fully connected network) solubility for various ILs. (see Appendix C for the experimental dataset).	126

Figure 4.32.	The hydrogen production vs time for ionic liquid coated and plain 1%Pt/SrTiO ₃ (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).	128
Figure 4.33.	The images of darkening spot on the reactor through reaction.	128
Figure 4.34.	The effect of ratio between ionic liquid and semiconductor on hydrogen production (catalyst: Pt/SrTiO ₃ , ionic liquid; [BMIM][BF ₄], preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).	129
Figure 4.35.	The comparison of Pt/Sr ₄ Ti ₆ O _y and Pt/SrTiO ₃ semiconductors in PWS reaction (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).	130
Figure 4.36.	The performance of calcined and not calcined Pt/Sr ₄ Ti ₆ O _y semiconductors in PWS reaction (catalyst: Pt/SrTiO ₃ , preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt).	131
Figure 4.37.	The performance of Pt/Sr ₄ Ti ₆ O _y , Pt/Al:Sr ₄ Ti ₆ O _y , and Pt/La:Sr ₄ Ti ₆ O _y semiconductors in PWS reaction (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, calcination: 400°C and 4 hours).	131
Figure 4.38.	The performance of Pt/SrTiO ₃ synthesized at two different temperatures (1000°C and 1100°C) in PWS reaction (catalyst:	132

	Pt/SrTiO ₃ , preparation method: SSR, Pt precursor: Pt nitrate salt, calcination: 400°C and 4 hours).	
Figure 4.39.	The performance of Pt/SrTiO ₃ calcined at different conditions (catalyst: Pt/SrTiO ₃ , preparation method: SSR, synthesis temperature: 1000°C, Pt precursor: Pt nitrate salt, IL: [BMIM][BF ₄]).	133
Figure 4.40.	The performance of Pt/SrTiO ₃ calcined at different conditions (catalyst: Pt/SrTiO ₃ , preparation method: SSR, synthesis temperature: 1000°C, Pt precursor: Pt hydrate salt, IL: [BMIM][BF ₄]).	134
Figure 4.41.	The performance of Au/SrTiO ₃ and IL: Au/SrTiO ₃ in PWS reaction (catalyst: Au/SrTiO ₃ , preparation method: SSR, synthesis temperature: 1000°C, Au precursor: Au chloride salt, IL: [BMIM][BF ₄]).	135
Figure 4.42.	The performance of Pt/SrTiO ₃ produced with two different methods (catalyst: Pt/SrTiO ₃ , SSR temperature: 1000°C, Pt precursor: Pt nitrate salt).	136
Figure 4.43.	The performance of plain and IL coated Pt/SrTiO ₃ in PWS reaction (catalyst: Pt/SrTiO ₃ , synthesis method: hydrothermal, Pt precursor: Pt nitrate salt).	137
Figure 4.44.	SEM images of SrTiO ₃ photocatalysts produced with solid state reaction a) Pt/SrTiO ₃ b) [BMIM][BF ₄]+Pt/ SrTiO ₃	138
Figure 4.45.	Band gap energies of Pt/SrTiO ₃ , [EMIM][N(Tf) ₂]+Pt/ SrTiO ₃ , [BMIM][BF ₄]+Pt/ SrTiO ₃ , and [BMIM][PF ₆]+Pt/ SrTiO ₃	140

	calculated from UV-Vis absorbance data using Kubelka-Munk function.	
Figure 4.46.	FTIR analysis of semiconductors before and after reaction for a-b) Pt/SrTiO ₃ , c-d) [EMIM][N(Tf) ₂]+Pt/ SrTiO ₃ , and e-f) [BMIM][BF ₄]+Pt/ SrTiO ₃ , respectively.	140
Figure 4.47.	XRD analysis of semiconductors a) Pt/SrTiO ₃ before reaction, b) Pt/SrTiO ₃ after reaction, c) IL coated semiconductors before reaction, and d) IL coated semiconductors after reaction.	142
Figure 4.48.	a) CTEM of IL-Dye:Pt/TiO ₂ , b) SEM-EDAX analysis of IL-Dye:Pt/TiO ₂	143
Figure 4.49.	XRD of TiO ₂ , Pt/TiO ₂ , Dye:Pt/TiO ₂ , IL:Pt/TiO ₂ and IL-Dye:Pt/TiO ₂	143
Figure 4.50.	a) UV-Vis absorbance spectra of N719 dye, b) Kubelka-Munk diffuse reflectance absorption spectra of TiO ₂ , Pt/TiO ₂ , Dye:Pt/TiO ₂ , IL:Pt/TiO ₂ , IL-Dye:Pt/TiO ₂ , c) Band gap energies TiO ₂ , Pt/TiO ₂ , Dye:Pt/TiO ₂ , IL:Pt/TiO ₂ , IL-Dye:Pt/TiO ₂	145
Figure 4.51.	FTIR of (blue) TiO ₂ , (red) Pt/TiO ₂ , (pink) Dye, (yellow) Dye:Pt/TiO ₂ , (purple) IL:Pt/TiO ₂ , and (green) IL-Dye:Pt/TiO ₂	146
Figure 4.52.	Hydrogen production of Pt/TiO ₂ , Dye:Pt/TiO ₂ , IL:Pt/TiO ₂ , and IL-Dye:Pt/TiO ₂ under 300W solar light illumination in TEOA solution 10% (600 ml reaction solution, 2g active catalyst/L).	149

Figure 4.53.	Stability test for hydrogen production of IL-Dye:Pt/TiO ₂ under 300W solar light illumination in TEOA solution 10% (600 ml reaction solution, 2g active catalyst/L).	150
Figure 4.54.	a) Photoelectrochemical profile of the TiO ₂ , IL:TiO ₂ and IL-Dye:TiO ₂ photoelectrodes (cyclic voltammetry profiles are recorded in a TEOA/Na ₂ SO ₃ aqueous solution), b) variation of the photocurrent density as a function of measured potential, c) photoconversion efficiency as a function of measured potential.	152
Figure 4.55.	a) The effect of dead volume in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and dead volumes (mass transfer area: 72 cm ² , catalyst concentration: 2g/L, stirring speed: 420 rpm, N ₂ flow rate: 30 ml/min, under Science-Tech solar simulator).	156
Figure 4.56.	a) The effect of liquid to gas mass transfer area at two different catalyst concentration in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and liquid to gas mass transfer area (dead volume: 150 ml, stirring speed: 420 rpm, N ₂ flow rate: 30 ml/min, under Science-Tech solar simulator).	158
Figure 4.57.	a) The effect of sweep gas flow rate in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and sweep gas flow rate (dead volume: 75 ml, mass transfer area: 30 cm ² , catalyst concentration: 2g/L, N ₂ flow rate: 30 ml/min, stirring speed: 420 rpm, under Science-Tech solar simulator).	160

Figure 4.58.	a) The comparison of hydrogen production at different stirring rates with repeated experiments, b) The relationship between cumulative hydrogen production and stirring speed (dead volume: 150 ml, mass transfer area: 72 cm ² , catalyst concentration: 2g/L, N ₂ flow rate: 30 ml/min, under Science-Tech solar simulator).	161
Figure A.1.	Optimal decision tree for UV light data.	228
Figure A.2.	Predictions by random forest model for UV a) testing set and b) training set.	229
Figure C.1.	The distribution of input variables for cations in the entire dataset.	246
Figure C.2.	Optimum decision tree structure for ammonium dataset.	250
Figure C.3.	Optimum decision tree structure for functionalized imidazolium dataset.	251
Figure C.4.	Optimum decision tree structure for phosphonium dataset.	252
Figure C.5.	Optimum decision tree structure for piperidinium dataset.	253
Figure C.6.	Optimum decision tree structure for pyridinium dataset.	254
Figure C.7.	Optimum decision tree structure for pyrrolidinium dataset.	255
Figure C.8.	Optimum decision tree structure for sulfonium dataset.	256

LIST OF TABLES

Table 3.1.	Input variables included into the dataset and their ranges.	30
Table 3.2.	Number of instances in each class of visible dataset (in the case of PWS).	31
Table 3.3.	Number of instances in each class of UV dataset (in the case of PWS).	31
Table 3.4.	Number of instances in each class of bandgap dataset (in the case of PWS).	31
Table 3.5.	The details for input and output variables in PECWS dataset.	33
Table 3.6.	Number of instances in each class of PEC dataset.	36
Table 3.7.	Number of instances in each class of bandgap dataset (in the case of PEC).	36
Table 3.8.	Summary of ionic liquid groups and their representation in dataset.	38
Table 3.9.	The cation descriptors and their ranges for each dataset.	39
Table 3.10.	The anion descriptors and their ranges for each dataset.	40
Table 3.11.	Number of instances in class A, class B, and class C for each dataset.	41

Table 3.12.	The hyperparameters of the ultimate DL model for both imidazolium and entire dataset.	42
Table 3.13.	The details of chemicals used.	43
Table 3.14.	Eight different custom-made glass reactors used.	50
Table 4.1.	Frequency of A and B doping for the most commonly used perovskites (i.e., complete list of dopants for most common perovskites).	55
Table 4.2.	The average calcination temperature of the samples prepared using different methods*.	70
Table 4.3.	Major factors leading high hydrogen production rate.	80
Table 4.4.	Confusion matrix for band gap data.	84
Table 4.5.	The confusion matrix for decision tree model for photocurrent density prediction.	105
Table 4.6.	Association rules leading class A (low solubility) for imidazolium dataset.	112
Table 4.7.	Association rules leading class A (low solubility) for entire dataset.	114
Table A.1.	Distribution of data points among the performance classes for significant variables (visible light dataset).	222

Table A.2.	Distribution of data points among the performance classes for significant variables (UV dataset).	223
Table A.3.	Results of association rule mining for visible light data for production rate class high.	226
Table A.4.	Results of association rule mining for UV light data for production rate class high.	227
Table A.5.	Confusion matrix for visible light data*.	229
Table B.1.	Discretization of the numeric input in ten levels.	230
Table B.2.	The confusion matrix for sample subset.	232
Table B.3.	The confusion matrix for bandgap testing subset (PECWS).	232
Table B.4.	The confusion matrix for photocurrent density testing and validation subsets (PECWS).	233
Table C.1.	Full names of cations.	234
Table C.2.	Full names of anions.	240
Table C.3.	Discretization of the numeric input in ten levels.	244
Table C.4.	Confusion matrix for imidazolium dataset.	247

Table C.5.	Confusion matrix for ammonium dataset.	247
Table C.6.	Confusion matrix for functionalized imidazolium dataset.	247
Table C.7.	Confusion matrix for phosphonium dataset.	248
Table C.8.	Confusion matrix for piperidinium dataset.	248
Table C.9.	Confusion matrix for pyridinium dataset.	248
Table C.10.	Confusion matrix for pyrrolidinium dataset.	248
Table C.11.	Confusion matrix for sulfonium dataset.	249
Table C.12.	Detailed information about the 49 ILs.	257

LIST OF SYMBOLS

cat	Catalyst
$E_{\text{H}_2/\text{H}^+}$	Energy of half reaction between hydrogen and H^+ ion
$E_{\text{O}_2/\text{H}_2\text{O}}$	Energy of half reaction between oxygen and water
h	Hour
J_{op}	Operational photocurrent density
M	Molarity
P_{in}	Incident irradiance
PM3	Parametric method 3
R	Free software environment for statistical computing and graphics
R^2	R-squared
wt	Weight

LIST OF ACRONYMS/ABBREVIATIONS

AIL	Acidic Ionic Liquid
ALRED	Al Reduction
AM	Air Mass
ANN	Artificial Neural Network
ARM	Association Rule Mining
AVG	Average
AZO	Aluminum Doped Zinc Oxide
BORATEBUF	Borate Buffer
CB	Conduction Band
CHEMDEP	Chemical Deposition
CHEMGRO	Chemical Growth
CHEMSOLR	Chemical Solution Route
COMM	Commercial
CO-PRE	Co-Precipitation
COSMO	Conductor-Like Screening Model
COSMO-RS	Conductor-Like Screening Model for Realistic Solvents
CP	Complexity Parameter
CPK	Corey-Pauling-Koltun
CSD	Cambridge Structure Database
CTEM	High Contrast Transmission Electron Microscopy
CV	Cyclic Voltammetry
CVAL	Cross Validation
DCM	Dichloromethane
DEP	Deposition
DIPC	Dip Coating
DL	Deep Learning
DOCTOR	Doctor Blade
DRIFT	Diffuse Reflectance for Infrared Fourier Transform
DROPCAST	Drop Casting
DSSC	Dye Sensitized Solar Cell

DT	Decision Tree
EDAX	Energy Dispersive X-Ray Analysis
ELECTRODEP	Electrodeposition
ES	Electro Spinning
FLAMEOXI	Flame Oxidation
FTIR	Fourier Transform Infrared Spectrophotometry
FTO	Fluorine Doped Tin Oxide
GC	Gas Chromatography
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HM	Hummer's Method
HOMO	Highest Occupied Molecular Orbital
HT	Hydrothermal
IAE	International Energy Agency
ICSD	Inorganic Crystal Structure Database
IDAC	Infinite Dilution Activity Coefficient
IE	Ion Exchange
IL	Ionic Liquid
IMDB	Internet Movie Database
IMP	Impregnation
ITO	Indium Doped Tin Oxide
IUPAC	International Union of Pure and Applied Chemistry
KNN	K Nearest Neighbour
LHS	Left Hand Side
LUMO	Lowest Unoccupied Molecular Orbital
MAE	Mean Absolute Error
MBE	Molecular Beam Epitaxy
MFC	Mass Flow Controller
ML	Machine Learning
MOD	Metal Organic Decomposition
MOF	Metal-Organic Framework
MTOE	Million Tonnes of Oil Equivalent
MW	Molecular Weight

NCS	Thiocyanato
NHE	Normal Hydrogen Electrode
NP	Not Provided Information
PC	Polymerized Complex
PEC	Photoelectrochemical
PECWS	Photoelectrochemical Water Splitting
PHOSPHATEBUF	Phosphatebuf
PHOTODEP	Photodeposition
PHOTORED	Photoreduction
PIL	Protic Ionic Liquid
PTP	Pechini Type Process
PV	Photovoltaic
PWS	Photocatalytic Water Splitting
RAM	Random Access Memory
RELU	Rectified Linear Unit
RF	Random Forest
RHE	Reversible Hydrogen Electrode
RHS	Right Hand Side
RMSE	Root Mean Squared Error
RTIL	Room Temperature Ionic Liquid
SCE	Saturated Calomel Electrode
SCILL	Solid Catalyst Ionic Liquid Layer
SEM	Scanning Electron Microscopy
SG	Sol-Gel
SILAR	Successive Ionic Layer Adsorption and Reaction
SILP	Supported Ionic Liquid Phase
SOLVOT	Solvothermal
SPINC	Spin Coating
SPRAYC	Spray Coating
SSR	Solid State Reaction
SSTEEL	Stainless Steel
STH	Solar to Hydrogen
TBA	Tetrabutylammonium

TEOA	Triethanolamine
UNIFAC	Universal Quasi-Chemical Functional Group Activity Coefficients
UV	Ultraviolet
UV-VIS	Ultraviolet-Visible Spectroscopy
VB	Valence Band
VD	Vapor Deposition
WT	Weight
XRD	X-Ray Powder Diffraction
ZPE	Zero-Point Energy

1. INTRODUCTION

Global energy demand has been increasing continuously with the increase in world's population, and this demand is expected to be doubled by 2050. In 2018 the total energy supply was reported as 14,282 million tonnes of oil equivalent (MTOE) by International Energy Agency (IAE) (Birol, 2020b). Unfortunately, great amount of the energy demand was met by fossil fuels such as coal, natural gas, and oil, which are limited resources. In addition to their non-renewability, energy production from fossil fuel causes to the emission of greenhouse gas such as carbon dioxide, which has been emitted at the level of 33,513 Mt at 2018 (Birol, 2020b).

Hydrogen has been considered as a promising energy carrier due to its high energy density (140 MJ kg^{-1}), which is superior to fossil fuels such as gasoline (46.4 MJ kg^{-1}) and coal (24.0 MJ kg^{-1}). Hydrogen is also abundant, light, and simple element in earth (Acar and Dincer, 2015) but does not exist in its molecular form; the high cost of hydrogen production methods such as steam reforming, partial oxidation, coal gasification and water electrolysis and associated CO_2 emission in fossil fuel-based methods (Mansilla *et al.*, 2018) prevents countries from the attempt of mass production. Today, most of the hydrogen demand of oil refining and chemical synthesis has been supplied by fossil fuel-based methods although they cause approximately 830 Mt of CO_2 per year (Birol, 2020a), which corresponds the annual CO_2 emissions of the United Kingdom and Indonesia together.

Photocatalytic hydrogen production from water is considered as a potentially cost-efficient method to produce hydrogen fuel with little impact on the environment. The primary purpose is to carry out the reaction via solar energy because it is renewable, sustainable, and free source although most of the semiconductors investigated in the literature are UV-light sensitive for now. In a simple photocatalytic water splitting (PWS) system, a semiconducting material absorbs light from light source, if that photon energy exceeds the band gap of semiconductor, it leads photogenerated electrons and holes which act as reducing and oxidizing agent to produce hydrogen from H^+ and oxygen from O^{2-} . TiO_2 is the first examined semiconductor in the pioneering work of Fujishima and Honda (Fujishima and Honda, 1972). Since then, titanium-based materials have been the most preferred semiconductors in PWS reaction due to its reliable activity, high applicability, and

relatively long-term stability. Besides TiO_2 , over 140 metal oxides, nano-oxides, perovskites, and other semiconductor materials are known to catalyze the PWS reaction, that variety creates a massive accumulation of research in literature.

Photoelectrochemical (PEC) water splitting to H_2 and O_2 is another attractive way of producing large scale of pure hydrogen with theoretical efficiency of 30.7% (J. J. Park *et al.*, 2014; Ullah *et al.*, 2020; Zhang, *et al.*, 2014), but typically efficiencies are below 10% in practical applications (Kado *et al.*, 2012). For an effective PECWS reaction, without the need of an external bias, semiconductor materials should have good chemical stability, low cost, rapid charge transfer without recombination, suitable band gap and suitable band edge positions for water splitting (i.e., conduction band more negative than reduction potential (0V vs NHE) whereas valence band more positive than oxidation potential (1.23 V vs NHE)) (Das *et al.*, 2011; Kumar *et al.*, 2018). If the band edges of semiconducting material do not overlap with the water splitting potential, an external bias should be applied to split water (Kim *et al.*, Shin *et al.*, 2011). In addition, since visible light is a large portion of the solar spectrum, the photo electrodes should be able to utilize visible light rather than UV (i.e., have low band gap) (Cho *et al.*, 2011; Hill *et al.*, 2013) as in the case of PWS. To use the energy from sun efficiently, optimization of the photoelectrochemical cell components (i.e., photo electrodes and electrolyte) is crucial. The technology is still at its infancy; hence large number of studies have been performed to find a way to overcome the problems that limits the performance.

One of the mutual problems of PWS and PECWS reactions that best performed semiconductors (such as TiO_2) are photoactive mostly under UV irradiation, which is about 4% of solar radiation. Thus, the possible modifications of semiconductors such as metal doping, (Jia *et al.*, 2019; Kochuveedu *et al.*, 2013; Q. Wang and Domen, 2020) surface modification (Chen *et al.*, 2010; Qin *et al.*, 2012) and dye sensitization to improve their visible light activity have been also investigated. Especially organic but also inorganic dye molecules work as light absorber which enables visible light harvesting and also make easier charge separation. Dye sensitization is widely used for various solar driven energy applications like solar cells (Hagfeldt *et al.*, 2010), photoelectrochemical water splitting (photoanode modification) (Z. Yu, Li, and Sun, 2015) and photocatalytic water splitting (Wang and Domen, 2020; Watanabe, 2017; Zhang *et al.*, 2016) to improve the visible light

sensitivity of semiconductors; the ruthenium complex dyes are the most common materials used for this purpose while other alternatives have been also tested (Zhang *et al.*, 2016).

Another challenge of PWS and PECWS systems is the high possibility of the recombination of photo excited electrons and holes which reduces the hydrogen yield. After significant amount of research, it was clearly shown that depositing promoter (co-catalyst) on semiconductor (Gómez-Solís *et al.*, 2014; Su, Wang, *et al.*, 2013; Tapan *et al.*, 2016; Zhang *et al.*, 2014) or doping element into crystal structure (Fu *et al.*, 2014; Iwase *et al.*, 2009; Li *et al.*, 2015; Senthil *et al.*, 2016) increases the efficiency by inhibiting recombination and favouring hydrogen production. It is also shown that electrolytes and additives (with water) in the reaction solution may suppress the back reaction and increase the hydrogen yield; the common additives for PWS or PECWS reactions are methanol (Pinheiro *et al.*, 2014) ethanol (Saadetejad and Yıldırım, 2018), alkali solutions (Cai *et al.*, 2016; Yang Liu *et al.*, 2008), solid additives (Cho *et al.*, 2011; F. Li *et al.*, 2014b), and ionic liquids (De Souza *et al.*, 2006; Opallo and Lesniewski, 2011; Qi *et al.* 2013; Snyder *et al.*, 2010).

The use of ionic liquids (ILs) in various fields of research has drawn great attention in recent years. They are basically organic salts with a variety of intermolecular interactions ranging from dipole-dipole interactions to hydrogen bonding, setting their affinities for hosting different guest molecules (Hayes *et al.*, 2015; Olivier-Bourbigou *et al.*, 2010). Such high degree of diversity in intermolecular interactions is originated from the availability of many cation/anion combinations with different physical and chemical properties (Gorman, 2001; Weingärtner, 2008). The ILs have been used as catalyst, solvent, lubricant, binder, surface modifier, and electrolyte owing to their desired properties such as low volatility, low flammability, and the selective solubility/capacity for various liquids (Amarasekara, 2016; Khan *et al.*, 2013; Zhang and Etzold, 2016). They also provide high ionic conductivity, and offer wide electrochemical windows, which make them good candidates for electrochemical processes as solvent, catalyst, or electrolyte (Amarasekara *et al.*, 2019; Endres, 2010). The water solubility of ILs become crucial since they extensively used as electrolyte in various processes, such as water electrolysis (De Souza *et al.*, 2006), photoelectrochemical water splitting (Opallo and Lesniewski, 2011; Snyder *et al.*, 2010), electrochemical or photoelectrochemical CO₂ reduction in water (Yang *et al.*, 2020), and dye degradation (Qi *et al.*, 2013).

By considering all these attempts to develop an efficient semiconductor it can be concluded that numerous ways are available to improve photocatalytic and photoelectrochemical hydrogen production, hence large number of papers has been published in recent years creating a huge accumulation of data in literature. Physical, chemical, optical, and electronic structure of semiconductors, synthesis method, properties of reaction solutions, and light specifications are the major elements that affect energy conversion efficiency from solar to hydrogen. Each component of the reaction is needed to be balanced to achieve the highest possible hydrogen yield but that requires great time, labor, and money. At this point, machine learning (ML) tools which were developed to extract non-trivial, previously unknown, and potentially useful knowledge from large and complex literature databases, can be used for this purpose. In addition to this, almost endless number of ILs can be created considering that the large number of potential anions and cations are available (or synthesizable). Thus, machine learning tools can be also applied to synthetic dataset developed computationally to predict any property of an IL which is existing or non-existing yet. In the literature, there are already examples of ML algorithm applications on the fields of PWS, PECWS, IL, or other chemical engineering topics (Amarasekara *et al.*, 2019; Can *et al.*, 2021; Can and Yildirim, 2019a; Günay and Yıldırım, 2021; Jalal *et al.*, 2019; Odabaşı *et al.*, 2014)

In this thesis, the subject of photochemistry is studied from two different complementary aspects. The first goal is to analyze the experimental or computational data in the field of photocatalysis, photo electrochemistry, and ionic liquids to gain fundamental knowledge and to reveal critical factors by using machine learning tools. Several algorithms such as association rule mining (ARM), decision tree (DT), random forest (RF) and deep learning (DL) were implemented on the datasets developed in R and Python environment for this purpose. The enhancement of photocatalytic hydrogen production over 1 wt% Pt/TiO₂ particulates sensitized with N719 dye (di-tetrabutylammonium cis-bis(isothiocyanato)bis(2,2'-bipyridyl-4,4'-dicarboxylato) ruthenium(II)) and coated by a thin layer of ionic liquid [BMIM][BF₄] (1-butyl-3-methylimidazolium tetrafluoroborate) has been also investigated experimentally. The Pt was loaded onto TiO₂ surface by incipient to wetness impregnation while dye and ionic liquid were loaded by solvent evaporation method in the presence of ethanol and acetone, respectively. The photocatalytic tests were performed in a semi-batch glass reactor under visible light irradiation provided by a solar

simulator at room conditions. SEM-EDAX, CTEM, FTIR, XRD, and UV-Vis characterization of the photocatalysts, and electrochemical analysis of the respective photoelectrodes were performed. Dye sensitization or ionic liquid coating were also applied SrTiO₃ perovskite semiconductor, but the results were not as encouraging as those of TiO₂ experiments due to the reproducibility problem. Although the effect of several factors such as precursors and precursor compositions of SrTiO₃, doping element, promoter, preparation method, and post-treatment temperature were examined in detail, the problem could not be solved. As the final step, the effects of reactor structure and operational conditions, which may affect the results through mass transfer limitations and inconsistencies in measurement and creates difficulties in comparison of results from different works, were also investigated experimentally.

In the next chapter, a literature survey about photocatalytic water splitting, photoelectrochemical water splitting, dye sensitization, ionic liquid applications and machine learning algorithms are explained. Chapter 3 includes all necessary information about computational works and experimental studies performed. The results obtained from machine learning applications and experiments were presented and discussed in Chapter 4. The main conclusions are declared and recommendations for future work are offered in the last chapter.

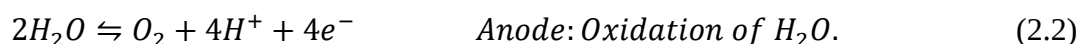
2. LITERATURE SURVEY

2.1. Photocatalytic Hydrogen Production

Hydrogen is considered as a promising energy carrier due to its high energy density, abundance in nature and cleanness; it can be used as a fuel or utilized in synthetic fuel production through Fisher-Tropsch type processes. However, the hydrogen is mostly produced via steam reforming or partial oxidation of fossil fuels (natural gas is the largest resource with 48%) (Dincer and Acar, 2014; Kothari *et al.*, 2008). Since these processes significantly contribute to the greenhouse gas emissions, new and cleaner technologies like semiconductor based solar water splitting have been extensively studied in recent years. The ultimate goal is to use solar energy, which is clean, renewable and free of charge. Fujishima and Honda split water under UV as the first time in 1972; they used a photoelectrochemical cell made of a semiconductor photoanode, cathode and electrolyte solution (Fujishima and Honda, 1972). A few years later, the studies on photocatalytic water splitting (PWS) over particulate semiconductors also started (Y. Ma *et al.*, 2014).

Photocatalytic water splitting (PWS) can be investigated in 3 steps; (i) absorption of photons with energies exceeding the semiconductor band gap and leading to the generation of electron and hole pairs in the semiconductor; (ii) migration of these photo-generated particles (e^- and h^+) which resulted in charge separation; (iii) surface chemical reactions between these carriers and present compound (e.g., water, water-alcohol solution) (Ismail and Bahnemann, 2014).

The reaction mechanism for photocatalytic water splitting can be described as



Water splitting reaction needs the standard Gibbs free energy change of 237 kJ/mol or 2.46 eV/molecule. The basic steps of water splitting process are shown in Figure 2.1 schematically. The yellow ball stands for semiconducting material and the grey dots on the semiconductor represent the promoter (co-catalyst). To use promoters for reduction and/or

oxidation sites are not essential; some semiconductors may also provide active sites for hydrogen and oxygen production by themselves (without promoter). The important point here is to develop not a photocatalyst to reach the maximum attainable efficiency with visible light (sunlight) irradiation; UV-light requiring catalyst is not practical for industrial applications. There are several approaches for modifying the electronic structure of a semiconductor such as metal and non-metal doping, changing preparation conditions, or dye-sensitization to make it visible light sensitive (detailed explanation in 2.3). On the other hand, the recombination of photo-excited electrons and holes before splitting water is a big challenge for PWS, and co-catalysts (or promoters) were used to inhibit this (Fukuzumi, Hong, and Yamada, 2013); Pt, Ru, Ir, and NiO_x and Rh are the most common co-catalysts. Methanol, ethanol, or some other sacrificial agents are also added to the reaction solution to suppress recombination reaction. Combining semiconductors (Z-scheme), modification of crystal structure and changing morphology of semiconductors are also other attempts which hinder recombination reaction.

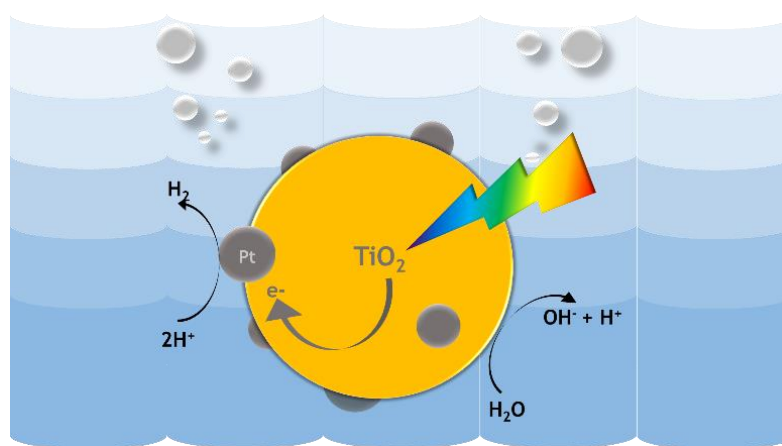


Figure 2.1. Schematic for photocatalytic water splitting reaction.

Over 140 metal oxides, nano-oxides, perovskites, and other semiconductor materials are known to catalyze the photocatalytic water-splitting reaction (Jiangtian Li and Wu, 2015). TiO_2 is the first, and still the most common semiconductor due to its low cost, high stability, nontoxicity, and environmental friendliness (Ge *et al.*, 2017; Y. Ma *et al.*, 2014). However, its band gap is suitable for UV light, which is only 4% of sunlight (NREL, 2020); it must be modified by ion doping and sensitization for the visible light (43% of sunlight) to be used under visible or solar light illumination as it is mentioned before. Perovskites have

also gained significant attention in recent years because they can be not only synthesized in numerous formulations and structures, but also easily modified by various methods. Perovskites are compounds in a crystal structure of ABX_3 ; A and B are two cations that often in different atomic sizes and X stands for anion which is generally oxygen. The detailed literature survey about the perovskites in photocatalytic water splitting reaction will be explained in the following chapter.

Significant progress has been made in recent years about PWS efficiency. In 2019, Li *et al.*, produced O_2 and H_2 (1:2, molar ratio) with a H_2 production rate of $11,000 \mu\text{mol g}^{-1} \text{h}^{-1}$ by using Au/N-doped TiO_2/MgO (111) as photocatalyst without any sacrificial reagents at 270°C ; its external quantum efficiency was measured as 81.8% at 437 nm (Yiyang Li *et al.*, 2019). Then, Takata *et al.*, claimed that they conducted the PWS reaction with an external quantum efficiency up to 96% at 350-360 nm by using aluminum doped $SrTiO_3$ cocatalysts and selectively photo depositing cocatalyst such as Rh, Cr_2O_3 , and $CoOOH$ onto the photocatalyst (Takata *et al.*, 2020). The three basic structures of photocatalytic water splitting reaction systems are given in Figure 2.2. and the real-life examples of those systems used in laboratories are presented in Figure 2.3.

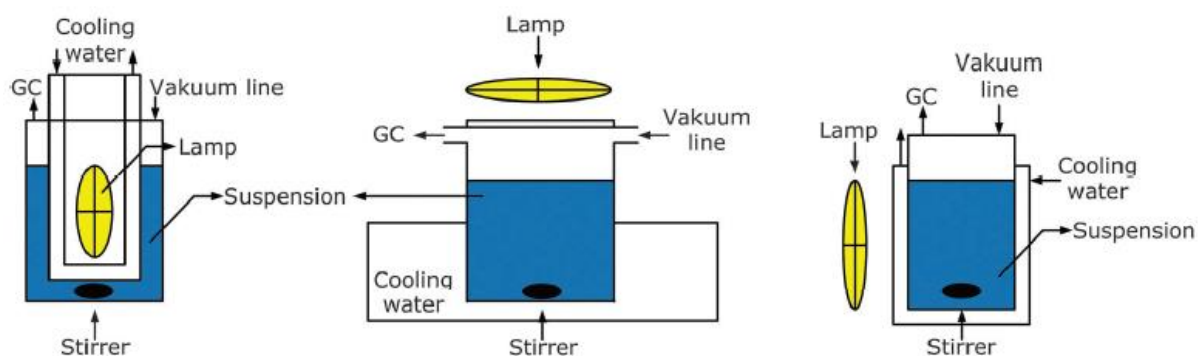


Figure 2.2. The basic structures for photocatalytic water splitting (Schwarze *et al.*, 2013).



Figure 2.3. The real-life examples of PWS reaction systems.

Nevertheless, despite the significant progress in the field, the photoconversion efficiency and hydrogen yield remains low even though large number of pristine or modified semiconductors, co-catalysts, sacrificial donor, and other elements of the process have been tested. Additionally, the progress in the field, in terms of the power conversion efficiency, hydrogen yield or any other performance measures used in the field, cannot be precisely assessed or compared with the previous works due to the lack of standard testing and reporting protocols as have been attempted in closely related field of solar cells (Khenkin *et al.*, 2020; Reese *et al.*, 2011).

One of the most important and well-accepted sources of noncomparable results in the field is the uncertainties associated with the energy transferred to the photocatalytic systems. First, the light sources (even the same type of source with the same power like 300 W xenon arc lamp) are not standard; the frequency interval and distribution as well as the intensity may differ for different brands and models. Even if the light intensity of source is measured and reported properly, the amount of energy absorbed by the photocatalyst is not always clear (Z. Chen *et al.*, 2010; L. Yang and Liu, 2007). The solid photocatalyst, liquid reaction medium and glass reactor may scatter and reflect the light (Qureshi and Takanabe, 2017) in a degree that depends on the reactor material, reactor geometry, position of light source, and physical/optical characteristics of photocatalyst and reaction solution (Ballari *et al.*, 2010; Braslavsky *et al.*, 2011; Kisch, 2013; Kisch and Bahnemann, 2015; Soares *et al.*, 2007).

Another problem that may prevent the comparison of results obtained in different works is the differences in the external (among the solid photocatalyst, liquid reaction solution and gaseous phase) and internal (within the solid particles) mass transfer limitations in different experimental systems. The solid, liquid, and gas phase coexist in most of the

photocatalytic systems and slight changes in experimental set-up may create significant deviation in the performance. For instance, when the rates of mass transfer between those solid-liquid-gas interfaces are faster than the reaction, the yield is limited due to the fast recombination of photogenerated electron and holes and the back reaction to water from produced hydrogen and oxygen (Ipek and Uner, 2020). Indeed, these issues have been discussed for more than three decades as Escudero *et al.* pointed out the importance of mass transfer limitations in photocatalysts in 1988, (Escudero *et al.*, 1989). Then Armenante and Kirwan characterized the mass transfer in microparticles with extensive experimental measurements in 1989 (Armenante and Kirwan, 1989). These works were followed by several studies focused on the limitations in heterogeneous photocatalysis trying to develop valid explanations and solutions for these limitations (Ballari *et al.*, 2008; Chen *et al.*, 2000; Ipek and Uner, 2020); the effects of catalyst concentration, particle size and stirring rate, have been discussed.

However, the effects of mass transfer limitations could not be easily counted when the results from two different laboratories are compared because no one is fully reporting the properties of reaction system used (like the size of gas-liquid interfacial area between the reaction medium and empty space in reactor). In papers mentioned, the effects of stirring speed, catalyst concentration, particle size, and even flow rate of sweep gas on mass transfer are presented. Therefore, while offering an efficient and reproducible photocatalyst for photocatalytic hydrogen production, the reaction system and catalyst properties should be well defined in a standardized manner. Although it is known that the major uncertainty arises from the incomplete description for the light adsorption by the photocatalyst (in terms of both frequency and intensity), the information that should be needed to account the mass transfer limitations have not been fully provided either.

2.2. Water Splitting in Electrochemical Cell

In a photo electrochemical cell (PEC), the light absorption, charge separation and migration, recombination rate, and surface reactions are still key processes as in the case of PWS. In PEC, there is an n-type semiconductor as the anode and a p-type semiconductor as the cathode; the reduction and oxidation reactions occur on cathode and anode separately facilitating the collections of reaction products in separate chambers (Ager *et al.*, 2015). The semiconducting materials are placed on a conductive substrate, and it forms the photo anode

or photo cathode for PEC water splitting (Li and Wu, 2015). The minimum potential required to split the water is 1.23 eV at room temperature. When the possible losses are taken into considerations the practical potential of semiconductor should be between 1.6 and 2.0 eV. The detailed band edges requirements are: the conduction band edge of semiconductor should be lower than the E_{H_2/H^+} and valence band edge of semiconductor should be higher than E_{O_2/H_2O} . The performance of a PEC is defined with solar-to-hydrogen (STH) conversion efficiency (η) and the formulation is given as

$$\eta = \frac{(1.23 \text{ V})(J_{op})}{P_{in}} \quad (2.3)$$

where J_{op} is the operational photocurrent density in mA cm^{-2} and P_{in} is the incident irradiance in mW cm^{-2} . Since the STH efficiency is directly related with the band gap of semiconductor used, the wide light absorption spectral range and suitable band energetic of the semiconductor are significant parameters for STH efficiency of a PEC system. On the other hand, high charge mobility and long charge carrier diffusion length are required to suppress charge recombination rate as much as possible (Li and Wu, 2015). Besides the activity of the redox reactions in PECs, the stability is also a critical challenge for PEC devices as it was mentioned before. As it can be seen from the Figure 2.4 in most of the scientific papers it was revealed that lifetimes for overall PWS devices are 24 hours or less (Ager *et al.*, 2015). That's why increasing device longevity is a critical issue and it should be enhanced in PEC devices.

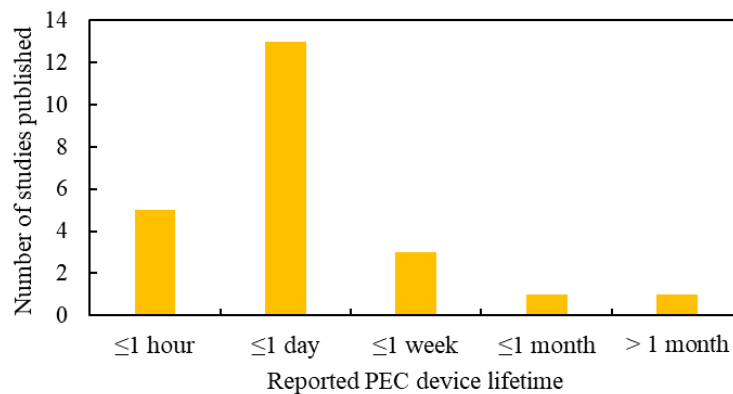


Figure 2.4. Number of studies published vs. reported PEC lifetime (Ager *et al.*, 2015).

As it was mentioned before, Fujishima and Honda have opened a new window in the field of solar fuel production via their light driven PEC water splitting system (Fujishima and Honda, 1972). As recent developments in literature are considered, the STH efficiencies for inorganic semiconductors are about 2.5% under AM 1.5G solar simulation due to their poor charge transfer abilities (Kim *et al.*, 2016; Kim and Lee, 2019; Pihosh *et al.*, 2015). On the other hand, the theoretical maximum STH efficiency for organic semiconductors in PEC water splitting is over 30% by considering their optical characteristics (i.e., band gap or band positions). However, the major drawback of organic semiconductors is their low stability in aqueous solutions. Yu *et al.*, produced an organic photoelectrode and achieved solar-to-hydrogen conversion efficiency of 4.33% with an onset potential of 0.55 V vs. RHE (reversible hydrogen electrode) under AM 1.5 G solar simulated light and after 10 h of stability test the electrode retained 90% of its initial photocurrent density (J. M. Yu *et al.*, 2020). To be industrially applicable, the cost of the hydrogen produced by any PEC should be competitive with fossil fuels. Significant progress is required to accomplish this aim, but all the scientific papers and the developments show us that it is possible.

The basic components of a PEC are anode (photo) electrode, cathode (photo) electrode, aqueous electrolyte solution, wire connections, and light source; however, these components can be arranged various ways. Minggu *et al.*, (Minggu, Wan Daud, and Kassim, 2010) divided the PEC systems into six groups according to their photo-electrode configurations as represented in Figure 2.5. If photo-electrode is n-type (electron donor), the hydrogen is produced over the metal electrode while the hydrogen is produced on photo-electrode if it is p-type (electron acceptor). To maximize the light absorption, several n-type or p-type photo electrodes can be layered together, and it is shown in Figure 2.5 (right bottom). Another configuration is stacking electrodes together and to build a connection between them as it is given in Figure 2.5 (left bottom).

Another classification system of PEC devices is based on their biasing methods. The detailed diagram of that classification is given by Figure 2.6. Although the PEC system with electrical bias is the easiest one to produce hydrogen, they are not feasible because the electricity used is provided by fossil fuels. As can be seen in Figure 2.6, another option is to create pH difference between anode and cathode chambers; a unit pH difference causes 0.06 V for that PE device. In general, acid-alkali pairs are used in pH biased devices and the chambers are separated by ion-exchange membrane. However, the production of acid-alkali

pairs also based on fossil fuels and a raw material is required. Therefore, the chemical (pH) biased systems are also not favourable. The remaining methods are PV-biased, DSSC-biased, or internal-biased systems, which are cheaper and more environmentally friendly than grid-biased or pH-biased systems (Minggu *et al.*, 2010). The PV, DSSC, and internal based ones are trending topics in the field of PEC, but they require more scientific studies to be viable for energy.

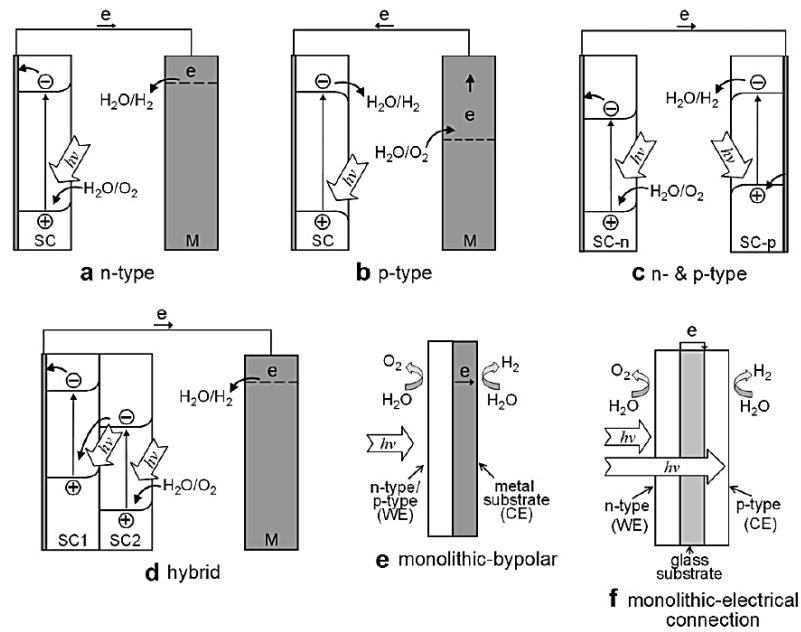


Figure 2.5. The classification of PEC systems based upon their electrode configurations (Minggu *et al.*, 2010).

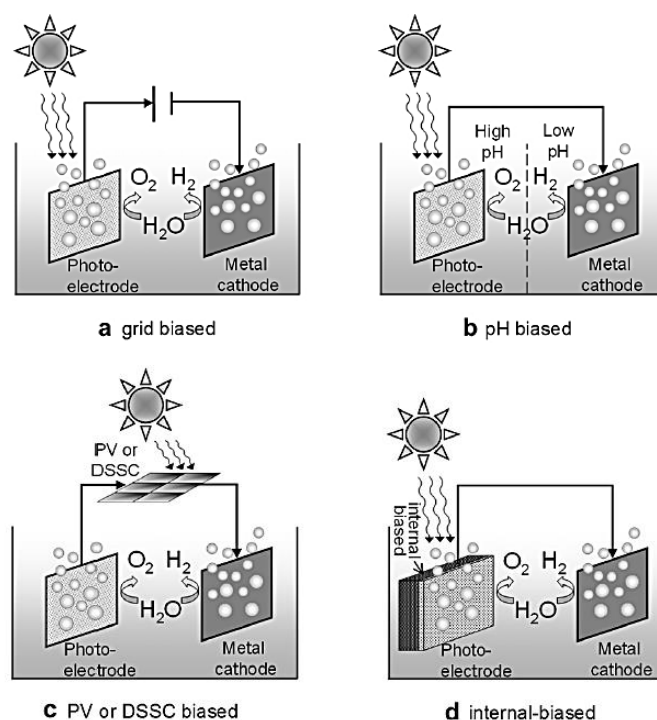


Figure 2.6. The classification of PEC systems based on their biasing methods (Minggu *et al.*, 2010).

2.3. Dye Sensitization for Visible Light Harvesting

The band structure, bandgap value, and the positions of valence or/and conduction band (VB or/and CB) of a semiconductor are the key parameters that affect the photocatalytic activity. In theory, as it was already explained, when the energy of incident light exceeds the band gap value of a semiconductor, the valence electrons of that semiconductor move to the conduction band and leave holes in the valence band. Those photogenerated electron-hole pairs initiate the process of photocatalytic water splitting. That's why the parameters affecting the light absorption first step has a significant role on the entire process and consequently on the hydrogen yield.

As previously mentioned, the primary goal of hydrogen production by photocatalytic water splitting is to use sun as light source. Since sunlight is mostly in visible range, the attempts for visible light harvesting gain importance in that field. Dye sensitization is one of the effective methods in the literature to improve visible light driven photocatalytic activity.

The general steps of photocatalytic water splitting with a dye-sensitized semiconductor can be divided into four steps as follows: (i) excitation of dye molecule by the incident light and the migration of photogenerated electrons from highest occupied molecular orbital (HOMO) to lowest unoccupied molecular orbital (LUMO) of the dye; (ii) the transition of those electrons from LUMO of the dye to the CB of the semiconductor; (iii) the reduction of water to form hydrogen in CB of the semiconductor (or transferring the electrons to co-catalyst for water reduction); (iv) the electron transfer from an electron donor to the dye for regeneration (Zhang *et al.*, 2016).

The basic diagram of dye-sensitization in PWS system is given in Figure 2.7. For an efficient cyclic process, LUMO level of the dye should be lower than CB of the semiconductor. Also, stability and cost of the dye molecules used are significant factors when designing a dye-sensitized PWS system.

Dye sensitization first studied extensively in solar cell technology (Zhang *et al.*, 2016). The pioneering work was performed in 1980 by Graetzel and his group. They produced Pt and RuO₂ co-deposited and Nb-doped titanium particles to be used in PWS. They conducted the experiments in the presence of [Ru(bpy)₃]²⁺ and methyl viologen [MV²⁺] and observed an enhancement in photocatalytic activity. Although ruthenium-based dyes are the most common alternatives in the literature, organic (Sharma *et al.*, 2018) and metal complex dyes (Sharma *et al.*, 2018) are also used for sensitization of semiconductors and particularly of TiO₂. Dye-sensitized TiO₂ particles can be used in solar cells (Hagfeldt *et al.*, 2010), in photoelectrochemical water splitting as photoanode (Z. Yu *et al.*, 2015), and in photocatalytic water splitting as semiconductor (Watanabe, 2017; Zhang *et al.*, 2016). Up to now, the Gratzel *et al.* reported the highest conversion efficiency from solar to electrical energy of 11.2% with N719 dye (Grätzel, 2005). Although bis(tetrabutylammonium)cis-bis(thiocyanato)bis(2,2'-bipyridine-4,4' dicarboxylato) ruthenium (II) (i.e. N719 dye) is one of the most high performance dye produced to be used in solar cells, its excellent contribution to the photocatalytic hydrogen production was also proven in several scientific papers (Huang *et al.*, 2020; Watanabe, 2017; Zhang *et al.*, 2016).

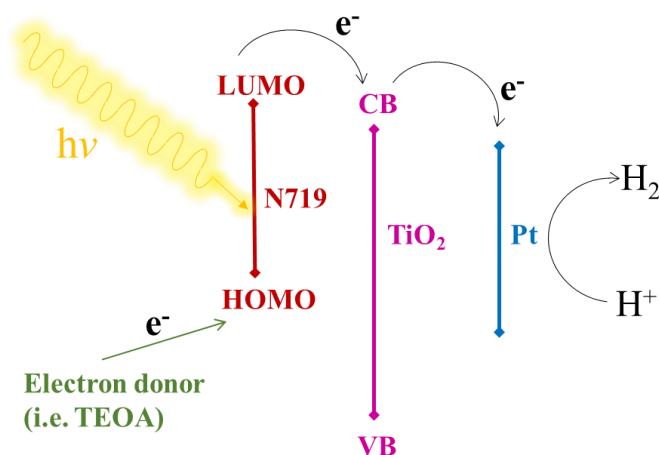


Figure 2.7. The basic diagram of dye-sensitization in PWS reaction.

2.4. Ionic Liquids and Their Applications in Chemical Engineering

Ionic liquids (ILs) are low melting point (generally $<100^{\circ}\text{C}$) salts. Information about the general properties of many ILs can be found in ILThermo database which is managed by the US National Institute of Standards and Technology. ILs have anionic and cationic parts and it is possible to design new ILs with desired properties by introducing new structural functionalities on those parts. By making different anion-cation combinations hydrophobicity, hydrophilicity, flexibility, viscosity, and stability of ILs can be changed (Olivier-Bourbigou *et al.*, 2010). That's why they are called as tuneable and multipurpose materials in the literature. It is claimed that origin of those unique characteristics of ILs stems from the various interactions between ions such as coulombic force, hydrogen bond, and intermolecular force (Zhang and Etzold, 2016). To provide a better visualization, images for a commercial ionic liquid ([BMIM][PF₆]) is given in Figure 2.8.

Many scientific papers revealed that ILs have good thermal and chemical stability (Amarasekara, 2016; A. Khan *et al.*, 2013; Zhang and Etzold, 2016) due to their low volatility and low flammability. ILs have also low mass transport resistance and high solubility of several reactant gases which are desirable properties to promote the efficiency of fuel cells (Zhang and Etzold, 2016). In the article of Khan (Khan *et al.*, 2013) oxygen solubility and diffusion coefficient of room temperature ILs (RTILs) have been found by using cyclic voltammetry and chronoamperometry. Their result supported the argument that

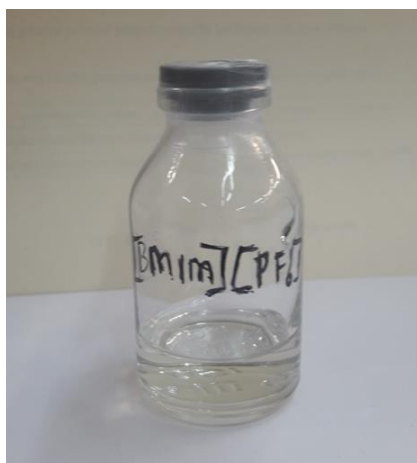


Figure 2.8. The image for ionic liquid [BMIM][PF₆] (purchased from Sigma Aldrich).

ILs have high solubility of oxygen (Kernchen *et al.*, 2007; Macfarlane *et al.*, 2014). ILs provide high ionic conductivity and a wide electrochemistry window so they are good candidates to be used in electrochemical processes as solvent or even as catalyst (Amarasekara *et al.*, 2019; Endres, 2010; Macfarlane *et al.*, 2014). Electrochemical window is simply related with the resistance of the cation to reduction and the resistance of the anion to the oxidation (Endres, 2010). As they compared with other possible electrolytes, it was observed that their electrochemical window (which is between 4.5-5 V) is higher than aqueous electrolytes but lower than organic solvents (G. R. Zhang and Etzold, 2016). It was also observed that conductivity of ILs have a negative relation with their viscosity so higher conductivity values can be reached by reducing their viscosities (Olivier-Bourbigou *et al.*, 2010). By considering all those IL properties it can be said that ILs may have an effect on reaction pathway, thermodynamic properties of reaction, mass transportations, and surface characteristics (Zhang and Etzold, 2016). It has been also proven that ILs are not easily biodegradable, and scientists have not come to an agreement whether ILs can be used in industry or not due to their possible harm to environment (Olivier-Bourbigou *et al.*, 2010).

Acidic ionic liquids (AILs), also called as protic ionic liquids (PILs) in literature, provide a great potential for proton transfer and they can form hydrogen bonds (proton donor or proton acceptor) (Olivier-Bourbigou *et al.*, 2010). Due to their electrochemical properties mentioned, AILs are often preferred in electrocatalytic reaction systems such as PECs (G. R. Zhang and Etzold, 2016).

As mentioned before, ionic liquids have many advantages since they are open to fine-tuning their structure in a way to enhance the selectivity, productivity, or both. However, handling them in an industrial application may be very difficult and catalytic reactions can be limited by diffusion due to their high viscosity. Also, a commercialized catalytic process may require a large amount of IL and their production costs are quite high for now (Skoda-Földes, 2014). Supported ionic liquid phase (SILP) attracts attention as a solution to those problems. In the production process of SILPs, a thin layer of ionic liquid and a homogeneous catalyst are dissolved together and then that mixture is applied to the internal surface of a porous inert solid (support) by physisorption (Kernchen *et al.*, 2007). That method was applied by many researchers and the SILPs generated used in hydrogenation reactions (Kernchen *et al.*, 2007). The critical point here is to determine the right amount of ionic liquid because mass transfer limitation can be problem as the IL film gets thicker. SILP can be produced in several ways such as impregnation of IL into support, creating covalent bonds between IL and solid, confinement of IL in the support or copolymerisation of the IL (Skoda-Földes, 2014). Solid catalyst with ionic liquid layer (SCILL) concept is also appealing. In that approach again a porous solid is coated with ionic liquid layer except that the solid is not inert but catalyst and the reaction system is heterogeneous. The pioneering work in the field of SCILL belongs to Erlebacher *et al.*, and they produced Pt/Ni-ionic liquid composite electrocatalysts to be used in oxygen reduction reaction (Zhang and Etzold, 2016). In many scientific articles, it was observed that to use SILP and SCILL promote the effectiveness of photocatalyst (Kernchen *et al.*, 2007). There are some possible explanations of that improvement: The IL on catalyst particle may act as co-catalyst and creates more active sites for reaction. Also, it may be derived from the solubility of several liquid and gaseous reactants in IL, which change the effective concentrations of those intermediates (Kernchen *et al.*, 2007). The possible configurations between support, ionic liquid, and catalyst are illustrated in Figure 2.9.

Solubility can be calculated by estimating the activity coefficients at infinite dilution (IDAC) (Marciniak, 2010). The experimental estimation of IDAC is commonly carried out by gas/liquid chromatography (Everett, 1965; Letcher and Moollan, 1995; Soares, 2011). Because this approach is time consuming and it does not offer any potential for screening many ILs, several computational models such as Universal quasi-chemical Functional Group Activity Coefficients (UNIFAC), modified UNIFAC, and CONductor-like Screening Model

for Realistic Solvents (COSMO-RS) are also used for the prediction of IDAC (Diedenhofen and Klamt, 2010). Among these methods, COSMO-RS is considered to provide more accurate results compared to the others (Diedenhofen *et al.*, 2003). COSMO-RS is the variant of dielectric continuum solvation method, which accounts for the electrostatic theory of interacting molecular surface descriptors with exact statistical thermodynamics methodology to predict properties (Klamt and Eckert, 2000).

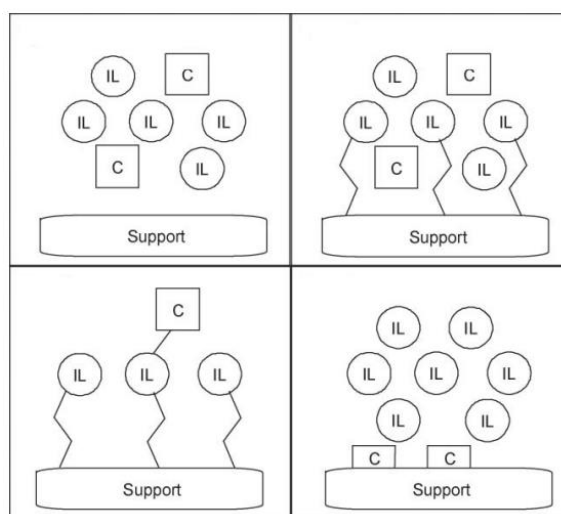


Figure 2.9. The illustration of possible combinations between support, ionic liquid, and catalyst (Van Doorslaer *et al.*, 2010).

The details of COSMO-RS theory can be found in literature (Eckert and Klamt, 2002; Gillham *et al.*, 2006; Matheswaran *et al.*, 2016). Because of its robust methodology, COSMO-RS is widely used for predicting the solubility of different components in different solvents (Diedenhofen *et al.*, 2003). For instance, Matheswaran *et al.* (Matheswaran *et al.*, 2016) studied IDAC of thiophene in ILs and showed that the COSMO-RS is able to predict the IDAC within an acceptable error range. Similarly, Fallanza *et al.* (Fallanza *et al.*, 2013), Klamt *et al.* (Klamt, 2003), and Lee *et al.* (Lee and Lin, 2015) performed COSMO-based calculations and concluded that COSMO-RS is robust in predicting the solubilities of hydrocarbons both qualitatively and quantitatively. Thus, the COSMO-RS calculations can serve as a versatile platform for quickly screening a large set of ILs (Gao *et al.*, 2015; Lee and Lin, 2015; T. Zhou *et al.*, 2012).

Their high ionic conductivity and wide electrochemical window makes the ionic liquids good candidates for electrochemical processes as solvent, catalyst, or electrolyte (Amarasekara *et al.*, 2019; Endres, 2010; Macfarlane *et al.*, 2014). The IL-water interaction plays a key role in most of the applications; for example, the ILs have been extensively used as electrolyte in various processes, such as water electrolysis (De Souza *et al.*, 2006), photoelectrochemical water splitting (Opallo and Lesniewski, 2011; Snyder *et al.*, 2010), electrochemical or photoelectrochemical CO₂ reduction in water (D. Yang *et al.*, 2020), and dye degradation (Qi *et al.*, 2013). *The solubility of ILs in water* is an important criterion for the selection of IL in these processes. However, there are also systems, in which the *water solubility in ILs* is more relevant; for instance, even a trace amount of water in the reaction medium or moisture in the air may be dissolved in IL layer coated over a solid catalyst (Skoda-Földes, 2014) or IL modified electrode (Safavi *et al.*, 2006; Wadhawan *et al.*, 2000; P. Yu *et al.*, 2005) and degrades its performance significantly. The same is also true for the processes in which the ILs are used as the solvent or reaction medium (Wender *et al.*, 2011). Additionally, the ILs have been used as solvent or agent during the synthesis of solid materials as electrode (Safavi *et al.*, 2006; Wadhawan *et al.*, 2000; P. Yu *et al.*, 2005), catalyst and semiconductor catalyst (Gołabiewska *et al.*, 2018; Hu *et al.*, 2010; Jing *et al.*, 2015; Ravishankar *et al.*, 2019; Ravishankar *et al.*, 2017; Tan *et al.*, 2019; Zwara *et al.*, 2019). The water (especially moisture from the air) dissolved in ILs may have significant adverse effects on such synthesis processes, and consequently on the final properties of the material synthesized (O'Mahony *et al.*, 2008). Hence, the predictive models, decision rules or any other form of knowledge that can be used to estimate the water solubility in ILs or to determine the suitable cation and anion combinations for the desirable water solubility may make significant contributions to the research in these fields.

The water solubility in ILs has been studied both experimentally (Boruń, 2019; Marcinkowska *et al.*, 2019; Martins *et al.*, 2015; Rieland and Love, 2020) and computationally (Freire *et al.*, 2010; Paduszyński, 2017; T. Zhou *et al.*, 2012) by various investigators. It was generally found that the affinity or the capacity of an IL for water was strongly dependent upon its structures and chemistry of constituent cation and anion as well as the temperature, relative humidity of air and the presence of impurities (Cao *et al.*, 2012; Freire *et al.*, 2007; Ranke *et al.*, 2009). Cao *et al.* (Cao *et al.*, 2012) studied the water solubility in 18 ILs experimentally and reported that the length of alkyl chain attached to the

imidazolium and pyrrolidinium cations was strongly related to hydrophilicity of the IL. The cations containing short length alkyl chains, which are more polar, have higher solubility of water while the long alkyl chain attached to these cations promote hydrophobicity. They also noticed that ILs containing halides offered a higher solubility of water compared to others.

Similar results were also obtained in the work of Huddleston *et al.* (Huddleston *et al.*, 2001); the halide containing ILs were miscible in any proportion; however, their properties changed significantly after the removal of water. The case was also similar for all hydrophilic ILs. On the other hand, the properties of hydrophobic ILs did not change significantly upon removal of water. In another work, the solubility of water in 1500 ILs were studied with COSMO-RS at room temperature (Zhou *et al.*, 2012); it was found that the strength of hydrogen bond was crucial in determining water solubility. The ILs having halides in their anion part were found to be highly hydrophilic as it was also observed in experimental works mentioned above. However, as electronegativity of halide decreases, the solubility of water also decreases as $[F]^- > [Cl]^- > [Br]^- > [I]^-$. It was also found that the ILs with sulfate anion alone exhibited a high capacity of water; however, if the sulfate anion had an attached atom or functional group such as hydrogen sulfate or methyl sulfate or octyl sulfate etc. the capacity of water tended to decrease. Here, they also noticed that length of alkyl chain attached to sulfate ion also influences the capacity of water. For example, ILs with methyl sulfate as anion had a higher capacity than ILs containing octyl sulfate, and this difference was rationalized by the weakening of electrostatic forces due to delocalization of charge (Kölle and Dronskowski, 2004).

2.5. Machine Learning in Catalysis, Photocatalysis, and Ionic Liquids

The studies involving various materials, methods, and modifications in the field of photocatalytic or photoelectrochemical water splitting reactions are summarized in the above sections of thesis; these combinations constitute an extensive domain for PWS or PEC research; consequently, large number of papers has been published over years creating a massive accumulation of experience in literature. However, this accumulation cannot be effectively utilized with conventional approaches because it is too big, heterogeneous, and distributed among the large number of sources. Instead, the machine learning and machine learning tools, which were developed to extract non-trivial, previously unknown, and

potentially useful knowledge from large and complex databases, can be used for this purpose.

Machine learning is a new branch of technology that help researchers focus on the most important information in the data they have collected. From industrial companies to the researchers in various scientific fields such as chemistry, medicine, biology and genetic, the machine learning has been used extensively to extract knowledge from the data in recent years. The machine learning tools can be divided into two main categories such as supervised and unsupervised learning. The detail of categorization is given in Figure 2.10, and it should be noted that different categorization for ML methods and several other ML tools (not shown in Figure 2.10) are also available. In supervised learning, algorithm creates paths connecting set of inputs with output and tries to predict output. By considering type of problem, the supervised learning can be also examined under two titles such as classification and regression methods. In real life, the issue of determining an email as spam or not spam and detecting an electronic fund transfer as fraud or not fraud are the most common application areas of classification ML algorithms. Logistic regression is one of the simplest classification methods which can be used to assign output into binary classes based on a statistical approach while support vector machine is a more powerful way to classify outputs based on their geometrical features. Decision tree (DT) and random forest (RF) algorithms can be used for both classification and regression problems. DT develops selection rules and heuristics for desired output variable by creating several if/then statements using input variables. Moreover, in RF technique, a large number of decision trees are created, and the output variable is predicted by majority voting (if classification) or by taking average (if regression). On the other hand, numerous options of ML algorithms such as linear regression, random forest, artificial neural network (ANN), and deep learning (DL) methods are available for regression problem. ANN is inspired by the animal neural processing; it can be considered as a simplified model of the brain. In this model, the inter-connected nodes and weighted links are summed up to reach output value. Then it compares the output node with a defined threshold value, as long as the output node is bigger than that threshold value, iteration continues. Deep learning is also a subfield of ANN which differentiates from ANN by its number of nodes or depth. Deep learning algorithm is generally used to construct predictive models for output variable due to its strength in analysis of large datasets with continuous numerical descriptors.

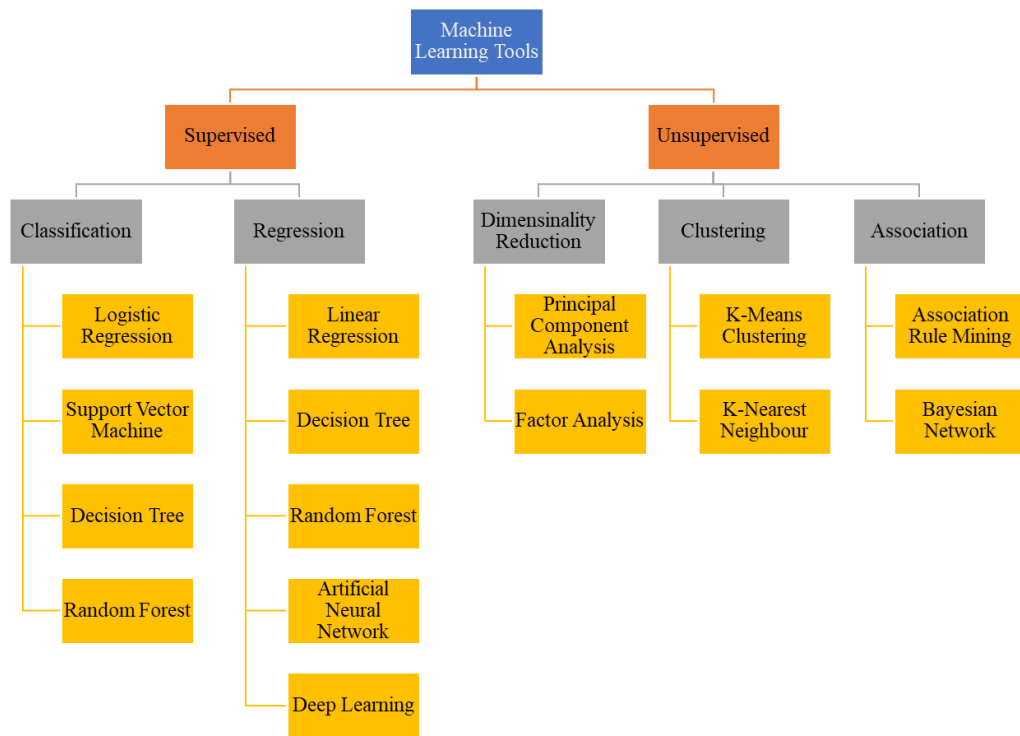


Figure 2.10. The basic diagram of dye-sensitization in PWS.

In the case of unsupervised learning, the algorithms can be divided into three categories such as dimensionality reduction, clustering, and association detection. In general, the dimensionality reduction is used in data pre-processing step of ML to reduce large number of input variables (dimensions) by retaining meaningful features of original data. Image compression and noise elimination from large dimensional images are some scopes of application for that algorithm. K-nearest neighbour (KNN) or k-means algorithms can be preferred in clustering problems. For example, IMDB website have an n dimensional space of movies and when someone clicks a specific movie page it offers some “related” movies which have common or almost common properties (cast, director, genre, or production year) with the main movie. It is possible to create such a recommendation machine for users with KNN algorithm. Finally, the most common algorithm for association detection is association rule mining (ARM) method that determines the effects of input factors on the output variable considering their restricted conditional probability distributions. The most known example of ARM is about customer buying behaviour; for example, it is discovered that men who purchase diapers for their children are also buy beers, so that type of knowledge may create a positive difference in supermarket shelf design. Bayesian network model is another tool to detect association between input variables, which also provide a graphical representation of

joint probability distribution. The simple visual demonstrations for algorithms mentioned above are given in Figure 2.11.

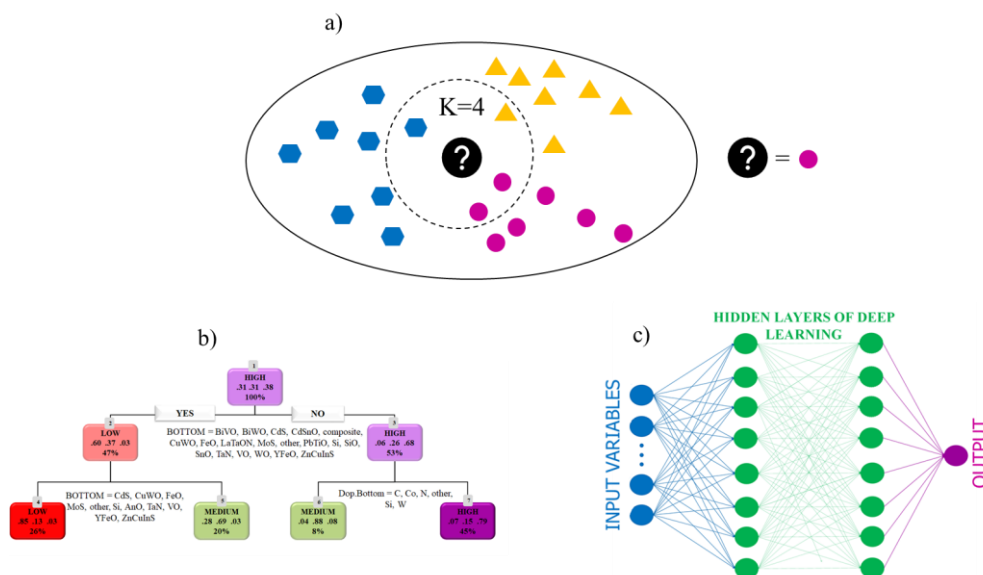


Figure 2.11. The visual demonstration for algorithms of a) KNN, b) Decision tree, d) Deep learning (Can et al., 2021).

As mentioned above these tools have been implemented in numerous subjects from economy and health to astronomy and biology in recent years. Up to now, the most preferred pattern recognition and prediction models in chemical engineering are artificial neural network, decision tree, random forest, and association rule mining (Günay and Yıldırım, 2021). For example, some groups (including ours) constructed and analyzed databases from past publications on catalytic CO oxidation (Günay and Yildirim, 2010), water gas shift reaction (Odabaşı *et al.*, 2014), steam and CO₂ reforming of methane (Baysal *et al.*, 2017; Şener *et al.*, 2018), direct alcohol fuel cells (Tapan *et al.*, 2016), photocatalysis over titania (Chesterfield and Adesina, 2009) oxidative coupling of methane (Kondratenko *et al.*, 2015), algal biomass productivity (Coşgun *et al.*, 2021), and electrodeposition parameters and materials (Kilic *et al.*, 2021). There are some works reporting the simultaneous use of molecular modeling and machine learning; the examples are study of structure-activity relationship for CO adsorption over Au clusters (Davran-Candan *et al.*, 2010), screening of perovskites for thermochemical water splitting (Emery *et al.*, 2016), predicting the crystal structure of materials (Graser *et al.*, 2018), CH₄ uptake over MOFs (Gülsoy *et al.*, 2019),

MOFs for gas storage and separation (Altintas *et al.*, 2021). The ML algorithms are also supportive tools to work through a specific subject in a statistical way; for example, a critical review of perovskite solar cells (Yılmaz and Yıldırım, 2021) or reproducibility, hysteresis, and stability analysis of perovskite photovoltaics (Khenkin *et al.*, 2020; Odabaşı and Yıldırım, 2020) are done prosperously with the help of ML tools.

In addition to this, the ML has been employed on various dataset involving the use of ILs recently. For instance, the solubility of H₂S in imidazolium ILs (Amedi *et al.*, 2016; Shafiei *et al.*, 2014), the solubility of CO₂ in ILs at different temperatures and pressures (Mesbah *et al.*, 2018; Ouaer *et al.*, 2020) and the melting points of pyridinium in ILs (Bini *et al.*, 2008) were predicted by using artificial neural network in various works. In addition, Venkatraman and Alsberg predicted CO₂ capture of 185 ILs at different conditions using various ML tools and compared their results with calculations based on CONductor-like Screening MOdel for Realistic Solvents (COSMO-RS) predictions (Venkatraman and Alsberg, 2017). The most effective structural descriptors for imidazolium ILs on the solubility of C₄-hydrocarbons was also studied via machine learning (Jalal *et al.*, 2019). Very recently, Benimam *et al.* reported a work involving the prediction of experimental activity coefficients for 53 ILs using various sML tools (Benimam *et al.*, 2020).

3. MATERIALS AND METHODS

The materials and methods used in the computational and experimental studies of this thesis will be given in sections 3.1 and 3.2, respectively. The subsections under 3.1 and 3.2 cover all necessary details.

3.1. Computational Details

The aim of the computational part of thesis was to extract previously unknown, and potentially useful knowledge from three different databases by using machine learning algorithms. The details about the databases are given in the following sections. Four different machine learning algorithms were examined, and the application was carried out in two environments: R, which is a free software for statistical computing and Python, which provides more general approach to machine learning. Association rule mining, decision tree, random forest algorithms were implemented with *arules*, *rpart*, and *randomForest* packages of R, respectively and *Keras* framework of Python was selected to perform deep learning algorithm. Some other packages of R were also used as complementary to machine learning algorithms. For instance, root mean squared error (RMSE) and mean absolute error (MAE) of models were evaluated using *rmse* and *mae* functions of *hydroGOF* package in R (Zambrano-Bigiarini, 2020) or *rmse* and *mae* functions of *Metrics* package in R (Frasco *et al.*, 2018). The *xlsx* package of R was used to import the matrices from R to Excel. To check whether multicollinearity among the input variables was a problem or not, *corrplot* package of R was used to determine and display the correlations (Wei *et al.*, 2021). To visualize decision trees better, *fancyRpartplot* function of the package *rattle* was preferred (Williams *et al.*, 2020) while *arulesviz* (Hahsler *et al.*, 2021) package of R was chosen for ARM visualization.

The diagram, which is given in Figure 3.1 summarizes the common steps of the procedure that we followed in machine learning analysis. The further details about the application of the specific ML tools onto the specific databases are explained in the corresponding section.

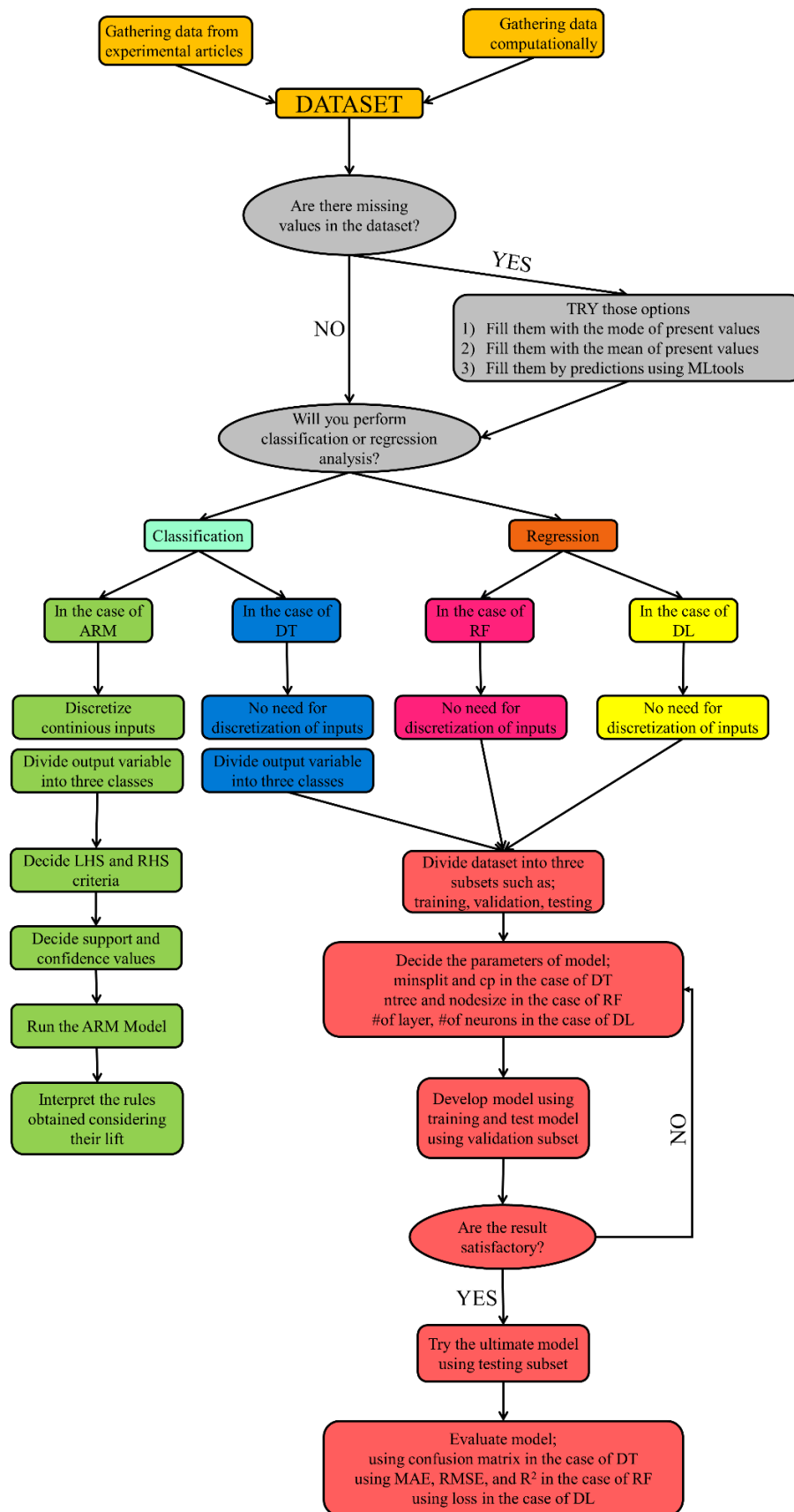


Figure 3.1. Diagram for application of machine learning tools.

As mentioned before, the evaluation parameters of the machine learning models used are MAE, RMSE, R^2 , and loss (as given in Figure 3.1). Their equations are represented below.

Mean absolute error (MAE) can be expressed as

$$\frac{1}{N} \sum_{i=1}^N |r_i - \hat{p}|. \quad (3.1)$$

Root mean square error (RMSE) can be expressed as

$$\sqrt{\frac{1}{N} \sum_{i=1}^N (r_i - \hat{p})^2}. \quad (3.2)$$

R-square (R^2) can be expressed as

$$1 - \frac{\sum (r_i - \hat{p})^2}{\sum (r_i - \bar{p})^2}. \quad (3.3)$$

Loss can be expressed as

$$\frac{\sum \frac{|A - F|}{A} \times 100}{N}. \quad (3.4)$$

Where r_i is real value, $\hat{p}$ is predicted value, $\bar{p}$ is mean value of r , A is for actual value, F is for forecast, and N is for number of observations.

3.1.1. Computational Details for Photocatalytic Water Splitting Over Perovskites

The database was constructed from 151 experimental papers on photocatalytic water splitting over perovskites published in 2005–2017; online sources including Elsevier, Wiley Online Library, American Chemical Society and Royal Society of Chemistry were reviewed (top hits in relevance sorting) and 540 instances (216 for visible; 324 for UV light) were extracted to form the database. The materials and methods used for photocatalyst preparation, properties of photocatalysts, and properties of reaction medium and light sources were used as input (predictor) variables while the hydrogen production rate was used

as the output (details in Table 3.1). In most of the articles, the hydrogen production ($\mu\text{mol g-cat}^{-1}$) was reported as a function time to show the performance. However, the hydrogen production in all these papers increased almost linearly with time, leading a high linear correlation between time and hydrogen production. Consequently, the hydrogen production rate ($\mu\text{mol g-cat}^{-1}\text{h}^{-1}$) used as the input variable to eliminate time from the input list. The multicollinearity, which may be exist among the catalyst preparation variables and measured characteristic, were checked, and no significant correlation were detected.

Datasets for experiments illuminated with visible and UV light were analyzed separately because their hydrogen production rates are in different range (obvious in Table 3.2). On the other hand, the band gap (eV) of perovskite semiconductors was also predicted since the dataset includes predictive variables for band gap; mainly perovskite material, promoter and/or doping element, perovskite preparation method, and post-treatment conditions. Hence band gap can be seen as both input and output in Table 3.2. For association rule mining and decision tree classification, each dataset (UV-light, visible-light, and band gap) was categorized in three classes as explained in Table 3.2, 3.3, and 3.4 below. This way, the number of data points in classes become approximately equal to prevent class imbalance problem. Class imbalance implies that the number of instances in one class is much higher than that of in another class (Drummond *et al.*, 2003). If the problem exists between the classes generated, the training algorithm will be biased towards the class with higher number of instances, so the model will fail to predict the output class correctly in testing step. Therefore, the number of instances in each class, as can be seen from tables below, were adjusted to prevent that issue.

The 4-fold cross validation (CVal) procedure was applied in both decision tree and random forest analysis; the dataset was divided into four smaller sets, and the model was built with three subsets (training) and tested with the remaining subset (testing). This procedure rotated four times and the average testing error was used to evaluate the fitness of the model (successful in classifying or predicting the unseen data, and valid for the entire range). For each method, the root mean square error for testing (RMSE) was calculated for various models with varying complexity and the model giving the minimum average RMSE for testing was selected as the best model.

Table 3.1. Input variables included into the dataset and their ranges.

Variable Name	Ranges for continuous or identities for discrete variables
Input Variables	
Elements Used in A site of Perovskite (including both host and doped elements) (ABX₃)	Ag, Ba, Bi, Ca, Ce, Co, Cr, Cu, Ga, H, In, K, La, Li, Mn, Na, Ni, Sb, Sm, Sn, Sr, Ta, Y, Yb, Zn
Elements Used in B site of Perovskite (including both host and doped elements) (ABX₃)	Bi, Cd, Co, Cr, Cu, Fe, Ga, In, K, Nb, Ni, Rh, Sb, Sn, Ta, Ti, Zr
Elements Used in X site of Perovskite (including both host and doped elements) (ABX₃)	N, O, S
Weight Percent of A site	0.070 - 0.770
Weight Percent of B site	0.050 - 0.770
Weight Percent of X site	0.053 - 0.540
Perovskite Preparation Methods	Co-precipitation (co-pre), electrospinning (ES), flux-type, hydrothermal (HT), ion-exchange (IE), novel, polymerized complex (PC), Pechini-type process (PTP), sol-gel, (SG), solid state reaction (SSR)
Calcination Temperature (K)	maximum 1773
Calcination Time (hour)	maximum 72
Promoter Loading Methods	Impregnation, photo-deposition (PD)
Elements Used as Promoters	Ag, Au, Cu, IrO ₂ , Ni, NiO, Pt, Rh, Ru, RuO ₂
Weight Percent of Promoter	maximum 8
Crystal Structure	amorphous, cubic, hexagonal, monoclinic, octahedral, orthorhombic, polyhedron, rhombohedral, tetragonal
Band Gap (eV)	1.27 -5.85
Surface Area (m²/g)	0.3 -206.5
Catalyst Weight (g/L)	0.2 -12.5
Percent of Alcohol in Reaction Solution	0 -100
Sacrificial Agents Used in Reaction Solution	AgNO ₃ , CH ₂ O, CH ₃ COOH, C ₃ H ₈ O CH ₄ , FeCl ₃ , H ₂ SO ₄ , NaI, KNO ₃ , HCOOH, KOH, Na ₂ S, Na ₂ SO ₃ , Na ₂ SO ₄ , NaNO ₂ , NaOH
Light Types	halogen, Hg lamp, UV, visible, Xe lamp
Power of Light (W)	8 - 1000
Minimum Wavelength of Light (nm)	200 - 440
Output Variables	
Band Gap (eV)	1.27 -5.85
Hydrogen Production Rate (μmol g-cat⁻¹h⁻¹)	0.001-36750

Table 3.2. Number of instances in each class of visible dataset (in the case of PWS).

Visible	Lower Limit ($\mu\text{mol gcat}^{-1} \text{ h}^{-1}$)	Upper Limit ($\mu\text{mol gcat}^{-1} \text{ h}^{-1}$)	Number of Instances
low	0.08	30	73
medium	30	350	71
high	350	21318	72

Table 3.3. Number of instances in each class of UV dataset (in the case of PWS).

UV	Lower Limit ($\mu\text{mol gcat}^{-1} \text{ h}^{-1}$)	Upper Limit ($\mu\text{mol gcat}^{-1} \text{ h}^{-1}$)	Number of Instances
low	0.001	130	109
medium	130	550	106
high	550	36750	109

Table 3.4. Number of instances in each class of bandgap dataset (in the case of PWS).

Band Gap	Lower Limit (eV)	Upper Limit (eV)	Number of Instances
visible	1.27	3.1125	143
UV_A	3.1125	3.89	120
UV_BC	3.89	6.3	109

As mentioned before the package of *rpart*, which can handle both categorical and continuous variables, was used to construct classification trees (Therneau *et al.*, 2018). The complexity of the decision tree models was adjusted by changing *minspl*, that is the allowable minimum number of instances in a node before splitting, and/or *cp* which is the *complexity parameter* (*cp*) which is used to control the size of the tree by evaluating the cost of adding another branch or depth to decision tree, if that cost is above the value of *cp*, then algorithm stops building tree. The trees became bigger, and model became more complex with decreasing *minspl* number (or *cp*). For this dataset, the *minspl* number and *cp* values arranged in a way that the DT model with the highest accuracy for *high* and *low* classes and with the simplest structure (easy to follow heuristics for future works) could be obtained.

The accuracy of high class is important to learn what to do while that of low class serves to determine what to avoid for high hydrogen production rates.

Since random forest creates decision trees using variables and instances randomly selected, it is not possible to produce exactly same model in even consecutive runs. Therefore *set.seed* function in R was used for pseudo-random number generation and reproduce same models (Lawson *et al.*, 2021). The complexity of the random forest models was adjusted by changing *ntree* (the number of trees in the forest) and *nodesize* (the minimum allowable number of instances in a terminal node). As *ntree* increases and/or *nodesize* decreases, model become more complex. The testing and training errors were evaluated with the increasing algorithm complexity (decreasing *nodesize* and/or increasing *ntree* number) to prevent over-prediction, but it was checked, and over-prediction was not observed in the scanned interval of model parameters. Therefore, the *ntree* and *nodesize* values which gave the highest R^2 and lowest rmse were recorded for each dataset and then the ultimate models were developed by using those *ntree* and *nodesize* values.

3.1.2. Computational Details for Photoelectrochemical Water Splitting

A comprehensive database was constructed by scanning hundreds of scientific papers on photoelectrochemical water splitting published between 2007 and 2020 at 13 different online sources including Elsevier, American Chemical Society, Royal Society of Chemistry, Wiley Online Library, Springer etc. After a detailed review (top hits in relevance sorting) 10560 instances (584 experiments different from each other) from 180 experimental articles were included in the database.

The chemical compounds forming electrodes and their preparation process, specifications of electrochemical cell, details about reaction solution and light source, were chosen as input variables while band gap (eV) of the working electrode and photocurrent density obtained (mA/cm^2) were defined as the output variables. In the prediction of photocurrent density, band gap was also used as input variable. The complete list of input and output variables with their ranges (for numeric variables) and sub-categories (for categorical variables) can be found in Table 3.5.

Association rule mining (ARM) was used to identify the effective input variables on the band gap of working electrode and high photocurrent density. To apply ARM algorithm,

the numeric input variables were discretized in ten levels (for details see Appendix B). The output variables of band gap and photocurrent density were also divided into three classes in a way that all classes have approximately equal number of instances to be used in ARM; the details are given in Table 3.6 and Table 3.7. The collected data from literature have some missing (78 out of 479 non-duplicated instances) band gap values; those missing values were predicted using random forest (RF) algorithm with a very high accuracy from the given values in the dataset (details in Results section).

Table 3.5. The details for input and output variables in PECWS dataset.

Variable Names	Ranges for numeric or sub-categories for categoric variables
Input Variables	
BOTTOM	BiVO, BiWO, CdS, CdSnO, composite, CuWO, FeO, g-C3N4, GaN, GaO, InO, LaFeO, LaTaON, MoS, NaNbO, NbO, other, PbTiO, SbO, Si, SiO, SnO, SrTiO, TaN, TaO, TiO, VO, WO, YFeO, ZnCuInS, ZnO
Dop.Bottom	B, C, Cl, Co, Cr, Cs, double dop, Fe, K, Li, Mn, Mo, N, Nb, other, P, Pt, Sb, Si, Sn, Ta, Ti, undoped, W
Co.cat.Bottom	Au, CdS, CePi, Co(OH), CoPi, Ni/NiO, NiO, NP, other, Pt, RuO
MIDDLE	BiOCl, BiOI, BiS, BiVO, CdSe, CuInS, CuO, ErO, FeO, FeOH, g-C3N4, InGaN, LaCoO, LaFeO, NiOH, NP, other, RGO, SnO, TiO, ZnO, ZnSe
Dop.Middle	Co, Ga, Mo, NP, Ti
TOP	Au, BiVO, CQD, FeOH/NiOOH, NP, other
Method.I	AlRed, anodization, chemDep, chemGro, chemSolR, comm, electrodep, etching, flameOxi, HT, hydrolysis, MBE, MOD, NP, other, polymerization, pyrolysis, SG, solvoT, spinC, sprayC, sputtering, SSR, VD
CalcT.I	25-1300°Celsius
CalcTime.I	0-30 hours
CalcCond.I	air, Ar/H ₂ , H ₂ , N ₂ , NH ₃ , NP, O ₂ , other
Method.II	AlD, chemDep, chemSolR, comm, dipC, electrodep, HM, HT, imp, NP, other, photoDep, photoRed, SG, SILAR, solvoT, spinC, sputtering
CalcT.II (°Celsius)	25-623
CalcTime.II (hours)	0-4

Table 3.5. The details for input and output variables in PECWS dataset. (cont.)

Variable Names	Ranges for numeric or sub-categories for categoric variables
Input Variables	
CalcCond.II	air, N ₂ , NP
Method.III	chemDep, chemRed, electrodep, NP, other, photoRed, solvoT
CE	Pt, Pt-Ti, Pt coil, Pt foil, Pt gauze, Pt plate, Pt wire, unknown
RE	Ag/AgCl, Hg/HgO, RHE, SCE, unknown
WEarea (cm²)	0.014-14
CEarea (cm²)	0.005-15
Substrate	AZO, FTO, ITO, Nb, other, Pt/Ti, Si, Ssteel, Ta, Ti, W, Zn
CoatingMeth.I	dep, dipC, doctorB, dropCast, electrodep in-situ, spinC, sprayC, sputtering
CoatingMeth.II	dep, dipC, dropCast, electrode, imp, in-situ, NP, other, spinC
CoatingMeth.III	dropCast, in-situ, NP, spinC
LasCalcT (°Celcius)	25-900
LastCalcTime (hours)	0-24
LastCalcCond	Air, Ar, N ₂ , NH ₃ , NP, O ₂
Type	Hg, Laser, Led, NP, W, Xe
Power (W)	35-1000
Wavelength (nm)	299-520
Intensity (mW/cm²)	1-1000
AM1.5G	0 and 1
ElectrolyteType	B ₄ K ₂ O ₇ , borateBuf, H ₂ SO ₄ , H ₂ SO ₅ , H ₃ PO ₄ , HBr, HClO ₄ , K ₂ HPO ₄ , K ₃ PO ₄ , KBr, KH ₂ PO ₄ , KHCO ₃ , KOH, Na ₂ S, Na ₂ SO ₃ , Na ₂ SO ₄ , NaCl, NaClO ₄ , NaNO ₃ , NaOH, NaS, other, phosphateBuf
E.Molarity (M)	0.01-5
Additive	C ₂ H ₅ OH, C ₂ H ₆ O ₂ , CH ₃ OH, H ₂ O ₂ , H ₃ PO ₄ , Na ₂ SO ₃ , Na ₂ SO ₄ , NaOH, NP, other, PBS, phosphateBuf
A.Molarity (M)	0-5
pH	0-14
Crystal.I.1	anatase, cubic, hexagonal, monoclinic, NP, orthorhombic, rhombohedral, rutile, tetragonal, triclinic

Table 3.5. The details for input and output variables in PECWS dataset. (cont.)

Variable Names	Ranges for numeric or sub-categories for categoric variables
Input Variables	
Crystal.I.2	hexagonal, monoclinic, NP, orthorhombic, rutile, tetragonal
Crystal.II	Amorphous, anatase, cubic, hexagonal, monoclinic, NP, rhombohedral, tetragonal
Crystal.III	Hexagonal, monoclinic, NP, rutile
Bias (V)	-1.2 - +2.5
Output Variables	
Band Gap (eV)	1.16 -4.37
Photocurrent Density (mA/cm2)	0-47.9

In the prediction process of band gap by using RF, 5-fold CVal was applied and the instances with unknown band gap values excluded from the procedure. The remaining 401 instances was divided into three parts as *testing*, *training*, and *validation* data, 100 instances (%25, 100 out of 401) were determined as *testing* subset. Then the 80% of remaining (60%, 241 out of 401) was used to construct the model which is called as *training* subset and 20% of remaining (15%, 60 out of 401) of remaining was used for *validation* in five rotations. Then the model with optimum parameters, which was determined from the results of 5-fold CVal procedure, was tested with *testing* data. In the dataset, photocurrent density values were collected at various bias values, therefore in *testing-training-validation* set separation, all the data coming from the same experiment were grouped together in order to prevent data leakage. The evaluation criteria for RF algorithm were selected as the mean absolute error (MAE), root mean squared error (RMSE), r-squared (R^2) as in the previous case. The predictions of band gap values in testing set were recorded and used for the ARM and DT analysis of photocurrent density.

In addition to prediction of band gap and photocurrent density, decision tree (DT) analysis was applied to extract heuristic rules leading high photocurrent density and extract unknown relations between input and output variables of both band gap and photocurrent density. A wide range of the parameters of *ntree* (number of trees in forest) and *nodesize* (number of instances in a terminal node of a tree) were scanned and the results were

evaluated to find the model with minimum complexity and highest prediction accuracy. The same limits for classes mentioned above for ARM were also used in DT analysis. The *complexity parameter* (explained in 3.1.2.) was chosen as the optimization criteria in DT algorithm; the decision trees with the highest accuracy and minimum complexity were determined as optimum. At this point accuracy can be associated with the purity of terminal nodes.

Table 3.6. Number of instances in each class of PEC dataset.

PEC	Lower Limit (mA/cm ²)	Upper Limit (mA/cm ²)	Number of Instances
A	1.23	47.9	3521
B	0.80	1.23	3382
C	0.00	0.80	3656

Table 3.7. Number of instances in each class of bandgap dataset (in the case of PEC).

Bandgap	Lower Limit (eV)	Upper Limit (eV)	Number of Instances
3high	3.00	4.37	135
2medium	2.40	3.00	138
low	1.16	2.40	128

3.1.3. Computational Details for Water Solubility in ILs

The dataset consisting of 16137 ILs were constructed from the most encountered 163 different cations (in nine groups) and 99 different anions by Ahsan Jalal (PhD in Chemical Engineering, Koc University, Istanbul); the water solubility/capacity, which were used interchangeably. The water solubility of ILs were estimated by COSMO-RS calculations using COSMOTermX version C30_1601 (“COSMOlogic GmbH and Co. KG,” 2021) at 298 K with TZVP parameterization. For this purpose, the cations, anions, and water molecule were obtained from COSMO-RS default database and COSMObaseIL. For the calculations of solubility, the Ionic Liquid Screening module was employed. Solvent capacity was calculated as

$$C^{\infty} = \frac{1}{\gamma_{i,j}^{\infty}} \quad (3.5)$$

where C^{∞} was the solvent solubility/capacity at infinite dilution for water in the corresponding IL while γ^{∞} was the activity coefficient of water at infinite dilution in the IL. No scaling was applied. The structure of database is summarized in Table 3.8. The structural descriptors were calculated through geometrical optimization using Spartan'14 Parallel Suite (Wavefunction Inc., Version 1.1.4) installed on a computer equipped with an Intel Core i5-4310UX CPU @ 2 GHz and 8 GB of RAM. The geometries of cations and anions were optimized separately using the default convergence criteria set for PM3 semi-empirical method. The details of ten cation descriptors and ten anion descriptors for each dataset are given in Table 3.9 and Table 3.10, respectively (further explanation in Appendix C). They were used in all analyses without any attempt to eliminate correlations or reduce the dimensionality because the set of 20 descriptors is sufficiently small for more than 16000 data, and all of three ML technique used here are robust against cross-correlations. The cations and anions are respectively labelled as $[C\#]^+$ and $[A\#]^-$, where # represents the index number of the corresponding component and given in the Appendix C.

The data subsets for the cation groups, as well as the entire dataset as whole, were analyzed using ARM, DT, RF, and deep learning (DL). ARM algorithm was used to identify the anion and cation properties for low water solubility. Since our descriptor variables are numerical, they were discretized in ten levels (see Appendix C). The datasets were divided into three classes in terms of output variable (solubility) as *A* (low solubility), *B* (intermediate solubility), and *C* (high solubility) in a way that all classes included approximately equal number of ILs to avoid problem of class imbalance (Drummond *et al.*, 2003). Consequently, the *class A* contained the ILs with the solubility lower than or equal to 1.5 mol/mol, while *class C* had the ILs with a water solubility of higher than 15 mol/mol; the ILs with the solubility values within these limits were placed in *class B*. The number of ILs in each class are given in Table 3.11.

Decision tree analysis, aimed to extract heuristic rules leading to the cation and anion properties for low solubility. The output variable (solubility) was divided into three classes; the same limits mentioned above for ARM were also used in DT analysis to have the consistency in the meaning of low, intermediate, and high solubility in the results obtained

in two methods. Again, the *complexity parameter* was used to optimize the tree structures for individual cation groups and entire dataset; if the extracted knowledge from decision trees with decreasing *minsplit* (increasing complexity) was not changed, the decision tree with higher *minsplit* (i.e., lower complexity) was selected. Additionally, the presence of branches with high number of ILs and high fraction of *class A*, was also taken into account in selection because the accuracy of *class A* is more critical than the other two (*class C*, i.e., high solubility, may be also important in some applications as well). To perform the decision tree analysis, as similar to previous cases, each dataset studied was divided into subsets; first, 25% of data were randomly selected and separated from the rest as *testing* data. Then the remaining 75% was used to construct the model by employing 5-fold cross validation; 60% of data was used for *training*, while the remaining 15% is used for *validating* in five rotations. The optimum model, which was selected from the results of 5-fold cross validation procedure, was tested with 25% data already separated for testing at the beginning. MAE, RMSE, R^2 , and class accuracies of validation subsets were used as the evaluation criteria of the models. The equations for MAE, RMSE, and R^2 are given in Section 3.1.

Table 3.8. Summary of ionic liquid groups and their representation in dataset.

Cation groups	Number of Cations	Number of Anions	Number of Instances (IL)	Range of Solubility (mol/mol at 293K)
Ammonium	33	99	3267	0.02 - 7989.21
Choline	1	99	99	0.44 - 1198.12
Functionalized Imidazolium	21	99	2079	0.03 - 2189.68
Imidazolium	41	99	4059	0.02 - 4195.33
Phosphonium	19	99	1881	0.01 - 17004.40
Piperidinium	7	99	693	0.02 - 4970.40
Pyridinium	21	99	2079	0.03 - 3247.40
Pyrrolidinium	15	99	1485	0.02 - 7079.86
Sulfonium	5	99	495	0.03 - 3693.42
Total	163	99	16137	0.01 - 17004.40

Table 3.9. The cation descriptors and their ranges for each dataset.

		Cation Molecular Wt. (amu)	Cation E HOMO (eV)	Cation E LUMO (eV)	Cation Dipole (debye)	Cation CPK Area (Å ²)	Cation CPK Ovality	Cation Polarizability	Cation HBD Count	Cation HBA Count	Cation ZPE (kJ/mol)
Ammonium	avg	174.7	-14.4	-4.4	8.3	253.4	1.4	57.4	0.2	1.4	878.4
	max	368.7	-12.3	-4.1	33.4	533.5	1.8	78.8	2.0	3.0	1954.3
	min	62.1	-18.5	-5.1	0.0	99.6	1.2	44.8	0.0	1.0	289.1
Choline	avg	104.2	-15.7	-4.4	2.2	153.5	1.2	49.3	1.0	2.0	494.7
	max	104.2	-15.7	-4.4	2.2	153.5	1.2	49.3	1.0	2.0	494.7
	min	104.2	-15.7	-4.4	2.2	153.5	1.2	49.3	1.0	2.0	494.7
Func.Imidazolium	avg	202.7	-13.3	-4.9	13.8	274.9	1.5	58.8	0.6	3.1	784.9
	max	325.5	-12.3	-4.6	37.1	449.1	1.7	71.4	2.0	4.0	1395.9
	min	122.2	-15.3	-5.4	3.5	163.8	1.3	50.3	0.0	2.0	349.8
Imidazolium	avg	183.0	-13.7	-5.0	15.3	267.1	1.4	58.2	0.3	2.0	811.4
	max	363.7	-12.0	-4.7	49.1	521.5	1.8	77.3	1.0	2.0	1746.9
	min	83.1	-15.2	-5.2	0.3	120.5	1.2	47.0	0.0	2.0	282.5
Phosphonium	avg	346.3	-13.5	-4.4	6.5	487.3	1.7	75.3	0.0	0.1	1681.8
	max	511.9	-12.1	-3.6	20.6	730.4	2.0	93.5	0.0	1.0	2592.7
	min	217.4	-15.1	-4.6	0.4	306.3	1.5	61.5	0.0	0.0	1020.3
Piperidinium	avg	155.0	-14.8	-4.3	3.5	209.6	1.3	54.4	0.1	1.7	729.3
	max	172.3	-13.5	-4.2	5.0	231.3	1.4	56.3	1.0	2.0	807.5
	min	139.2	-15.9	-4.8	2.2	181.9	1.2	52.3	0.0	1.0	572.8
Pyridinium	avg	149.1	-14.5	-5.8	8.2	206.0	1.3	54.0	0.1	1.3	569.9
	max	224.3	-12.7	-5.6	24.7	315.8	1.5	62.5	1.0	3.0	987.8
	min	80.1	-15.8	-6.2	0.4	114.0	1.1	46.9	0.0	1.0	257.9
Pyrrolidinium	avg	151.0	-14.7	-4.2	4.2	213.5	1.3	54.5	0.1	1.3	741.9
	max	240.5	-13.2	-4.1	14.4	342.3	1.5	64.4	1.0	2.0	1237.4
	min	100.2	-16.0	-4.9	0.0	150.6	1.2	49.3	0.0	1.0	500.4
Sulfonium	avg	121.6	-14.9	-5.7	2.9	168.4	1.3	51.0	0.0	1.0	451.6
	max	153.3	-13.8	-5.4	3.8	200.1	1.3	53.8	0.0	1.0	513.1
	min	91.2	-15.8	-5.9	2.2	139.1	1.2	48.2	0.0	1.0	366.0

Table 3.10. The anion descriptors and their ranges for each dataset.

		Anion Molecular Wt. (amu)	Anion E HOMO (eV)	Anion E LUMO (eV)	Anion Dipole (debye)	Anion CPK Area (Å ²)	Anion CPK Ovality	Anion Polarizability	Anion HBD Count	Anion HBA Count	Anion ZPE (kJ/mol)
Ammonium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Choline	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Func.Imidazolium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Imidazolium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Phosphonium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.1	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Piperidinium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Pyridinium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Pyrrolidinium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.0	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0
Sulfonium	avg	172.2	-5.5	4.7	7.4	157.4	1.3	49.5	0.2	3.1	251.4
	max	745.1	0.4	13.5	41.8	447.6	1.7	73.0	3.0	9.0	1250.2
	min	19.0	-9.9	-0.9	0.0	28.8	1.0	41.6	0.0	0.0	0.0

Table 3.11. Number of instances in *class A*, *class B*, and *class C* for each dataset.

Datasets	Number of Instances in class A	Number of instances in class B	Number of instances in class C
All	5747	5299	5091
Ammonium	1085	1160	1022
Choline	32	33	34
Functionalized Imidazolium	737	828	514
Imidazolium	1499	1288	1272
Phosphonium	805	411	665
Piperidinium	236	213	244
Pyridinium	670	771	638
Pyrrolidinium	513	442	530
Sulfonium	170	153	172

A multilayer fully connected neural network, as the deep learning algorithm, was used to construct the predictive models for water solubility in ILs considering its strength in analysis of large datasets with continuous numerical descriptors. As mentioned before the analysis was performed on *Python* by using *Keras* (Nain *et al.*, 2021) as the deep learning framework while *tensorflow* (Abadi *et al.*, 2016) was used for backend. Before modelling, the data were standardized first to prevent any inverse effects of the large differences in the numerical ranges of descriptors, and then, they were shuffled for randomization to prevent the processing of similar materials consequently. The number of hidden layer and number of neurons were determined by minimizing the validation loss using greedy search while the grid search was used to decide the activation function. Various optimizers were also tested through grid search, and *Adamax* (Kingma and Ba, 2015) was chosen for the final model, which had three hidden layers (with 4096, 128, 4096 neurons in three hidden layers respectively for both imidazolium and entire dataset); the activation functions were also chosen as *ReLU* (*rectified linear unit*) (Nwankpa *et al.*, 2018) in the first two hidden layers and *SOFTMAX* (Nwankpa *et al.*, 2018) in the last hidden layer while *tanh* (Nwankpa *et al.*, 2018) was utilized in the output layer. The optimum learning rate was found to be 0.01 while number of epochs was chosen as of 8000. *Dropout Regulation Technique* (Mele and Altarelli, 1993) was used in each layer to prevent over-fitting. As the search procedure, two hidden layer network structures were constructed using various combinations of

hyperparameters, and then, the options for the third hidden layer were searched based on six two-layer structures having the smallest (happen to be also very similar) validation lost. The optimum model hyperparameters found are given in below in Table 3.12.

Table 3.12. The hyperparameters of the ultimate DL model for both imidazolium and entire dataset.

Learning rate	Number of nodes for the 2nd hidden layer
0.01	4096
Number of epochs	The activation function for input layer
8000	relu
Number of hidden layer	The activation function for the 1st hidden layer
3	relu
Number of nodes for input layer	The activation function for the 2nd hidden layer
4096	softmax
Number of nodes for the 1st hidden layer	The activation function for the output layer
128	tanh

3.2. Details About Experimental Work

3.2.1. Materials

The details of the chemicals used in this work were given in Table 3.13. All materials were used without further treatment except [BMIM][BF₄], [BMIM][PF₆], and [EMIM][N(Tf)₂] which were kept under vacuum at 60 °C for one day.

Table 3.13 The details of chemicals used.

Chemical	Formula	Source
1-Butyl-3-methylimidazolium tetrafluoroborate	[BMIM][BF ₄]	Sigma-Aldrich
1-Ethyl-3-methyl-bis(trifluoromethylsulfonyl)imide	[EMIM][N(Tf) ₂]	Sigma-Aldrich
1-Butyl-3-methylimidazolium hexafluorophosphate	[BMIM][PF ₆]	Sigma-Aldrich
Gold (III) chloride trihydrate	HAuCl ₄ .3H ₂ O	Sigma-Aldrich
Aluminium oxide	Al ₂ O ₃	Merck Millipore
Acetone (for analysis, ≥99.8%)	C ₃ H ₆ O	Merck Millipore
Dichloromethane (DCM)	CH ₂ Cl ₂	Merck Millipore
Ethanol (for analysis, ≥99.9%)	C ₂ H ₅ OH	Merck Millipore
Ethylene glycol	C ₂ H ₆ O ₂	Merck Millipore
Fluorine doped tin oxide glass	FTO	Teknoma
Methanol	CH ₃ OH	Zag-Kimya
N719 dye	C ₅₈ H ₈₆ N ₈ O ₈ RuS ₂	Ossila
Sodium sulfide	Na ₂ SO ₃	Zag-Kimya
Strontium carbonate	SrCO ₃	Alfa Aesar
Strontium carbonate	SrCO ₃	Sigma-Aldrich
Stronsiyum nitrat	Sr(NO ₃) ₂	ACROS
Tetraaminplatinum(II) hydroxide hydrate	Pt(NH ₃) ₄ (OH) ₂ ·xH ₂ O	Sigma-Aldrich
Tetraaminplatinum(II) nitrate(≥50.0% Pt)	[Pt(NH ₃) ₄](NO ₃) ₂	Sigma-Aldrich
Ti-nano oxide D/SP paste	TiO ₂	Solaronix
Titanium (IV) dioxide	TiO ₂	Honeywell
Titanium (IV) dioxide (≥99.5%)	TiO ₂	Sigma-Aldrich
Titanium butoxide	Ti(OCH ₂ CH ₂ CH ₂ CH ₃) ₄	Sigma-Aldrich
Triethanolamine (extra pure)	C ₆ H ₁₅ NO ₃	Tekkim Kimya

3.2.2. Catalyst and Photoelectrode Preparation

3.2.2.1. Preparation of Pt/SrTiO₃ and [BMIM][BF₄]:Pt/SrTiO₃ with solid state reaction

Two preparation methods with several different parameters were tried to produce SrTiO₃ with the aim of finding the best semiconductor structure for photocatalytic water splitting reaction. In the first method, which is solid state reaction, the SrCO₃ and TiO₂ were mixed and then kept at 1000°C for 10 hours. After SrTiO₃ cooled down to room temperature, 1 wt% Pt/SrTiO₃ photocatalyst was prepared by simple incipient to wetness impregnation. 1 g of SrTiO₃ was mixed ultrasonically under vacuum for 15 mins while 0.02 g of tetraammineplatinum(II) nitrate was dissolved in 1 ml of distilled water by using ultrasonic mixer. Then, the solution was impregnated onto the strontium titanate, and it was rested for 15 mins under vacuum. The sample was dried at 100 °C overnight.

The ionic liquid coating onto the Pt/SrTiO₃ was carried out by simple mixing method. Two grams of ionic liquid was dissolved in five ml of dichloromethane (DCM) and that solution was dropped slowly onto the five grams of Pt/SrTiO₃ particles. During ionic liquid deposition, solid Pt/SrTiO₃ particles were vacuumed and mixed ultrasonically. After deposition, that sludge which involves ionic liquid, DCM, and Pt/SrTiO₃ was kept under dark and air free conditions for one week. In that way it was aimed that ionic liquid could penetrate the pores of Pt/SrTiO₃. As the last step, the ionic liquid coated Pt/SrTiO₃ was dried at 60°C for (at least) 6 hours to evaporate and remove the remaining DCM solution or any moisture from photocatalyst.

3.2.2.2. Preparation of Pt/SrTiO₃ and [BMIM][BF₄]:Pt/SrTiO₃ with hydrothermal method

As the second method, the procedure in the paper of Zwara et. al (Zwara *et al.*, 2019) was employed and SrTiO₃ was synthesized by hydrothermal reaction. In this method two precursor mixtures were prepared. One of them comprises of 6.8 gr titanium butoxide and 100 ml of ethylene glycol, while the other one contains 4.2 gr of Sr(NO₃)₂ and 40 ml of distilled water. Those solutions were mixed at a medium stirring speed separately for about 15 minutes and then they were mixed in one beaker. If it was aimed to produce ionic liquid coated SrTiO₃, the step ionic liquid should be also added to the beaker at this step by ensuring the molar ratio of 1:2 between ionic liquid and titanium butoxide is realized. Since the final solution becomes a very viscous fluid within seconds, all solutions (titanium butoxide

solution, strontium nitrate solution and ionic liquid if it is the case) should be mixed simultaneously and quickly. That viscous mixture was placed in an autoclave reactor and that autoclave was kept at 180°C for 24 hours. After 24 hours, the autoclave was stored in a safe place until it reaches to room temperature. Then, the mixture was washed several times with pure ethanol and distilled water, and it was maintained at 60°C until it was fully dried. The procedure for 1%Pt impregnation onto the SrTiO₃ was exactly the same as that explained in previous section. if 1%Pt was deposited onto the [BMIM][BF₄]/SrTiO₃ semiconductor, the amount of Pt nitrate salt was measured by considering only the amount of SrTiO₃ in [BMIM][BF₄]/SrTiO₃; the weight of [BMIM][BF₄] was not considered because it does not act as a semiconductor. The drying temperature after Pt impregnation was also set to 60°C for [BMIM][BF₄]/SrTiO₃ (not to 100°C as for SrTiO₃ alone) in considering that the high temperature might cause to the thermal degradation of ionic liquid.

3.2.2.3. Preparation of Pt/TiO₂

1 wt% Pt/TiO₂ photocatalyst was prepared by simple incipient to wetness impregnation method. 1 g of TiO₂ was mixed ultrasonically under vacuum for 15 mins while 0.02 g of tetraammineplatinum(II) nitrate was dissolved in 2.5 ml of distilled water by using ultrasonic mixer. Then, the solution was impregnated onto the titanium dioxide, and it was rested for 15 mins under vacuum. The sample was dried at 100 °C overnight.

3.2.2.4. Dye sensitization of Pt/TiO₂

The Pt/TiO₂ samples were ground before the dye sensitization process. 1 g of Pt/TiO₂ was suspended into 0.1 mM N719 dye solution (10 mL of ethanol and 0.001188 g of dye). The mixture was stirred gently for 72h at dark condition and then dried at 60 °C overnight. The dye sensitized particles are referred to as Dye:Pt/TiO₂.

3.2.2.5. Ionic liquid coating of dye sensitized Pt/TiO₂

The ionic liquid deposition was carried out in glove box under nitrogen atmosphere; the procedure was similar to the study of Babucci *et al.* (Babucci *et al.*, 2015) 1 g of Pt/TiO₂ photocatalyst (dye sensitized or not) and 0.2 g of ([BMIM][BF₄]) were dissolved in 4 grams of acetone. They were mixed for 1 hour (at dark if dye-sensitized samples were used) and then dried in vacuum oven for one day at 60 °C. Pt/TiO₂ and dye sensitized Pt/TiO₂ particles coated with ionic liquid are referred to as IL:Pt/TiO₂ and IL-Dye:Pt/TiO₂, respectively.

3.2.2.6. TiO_2 , IL: TiO_2 , and IL-Dye: TiO_2 photoelectrode preparation

FTO glass substrates were washed several times with distilled water and dried before deposition. Ti-nanoxide D/SP paste was stirred well before use to obtain a smooth layer on FTO. The titania paste was deposited on FTO glass by simple doctor blade method and then substrates were heated to 475 °C with a ramping rate of 30°C/min and they were kept at 475 °C for 30 mins. The calcined substrates were left in furnace to cool down to 120°C to avoid thermal shock before any further treatment. The pictures belong to the different steps of photoelectrode preparation was given below by Figure 3.2. To prepare dye sensitized photoanode, the TiO_2 coated FTO glasses were submersed into N719 solution (0.5mM ethanol sol.) while they were still warm (ca. 60 °C) and kept at dark for 24 h. IL coating was performed by solvent evaporation method. The TiO_2 or dye sensitized TiO_2 photoanodes were immersed in [BMIM][BF₄]: acetonitrile solution (1:20, v: v) and the solvent was evaporated under vacuum at 60°C for 12-15 hours.

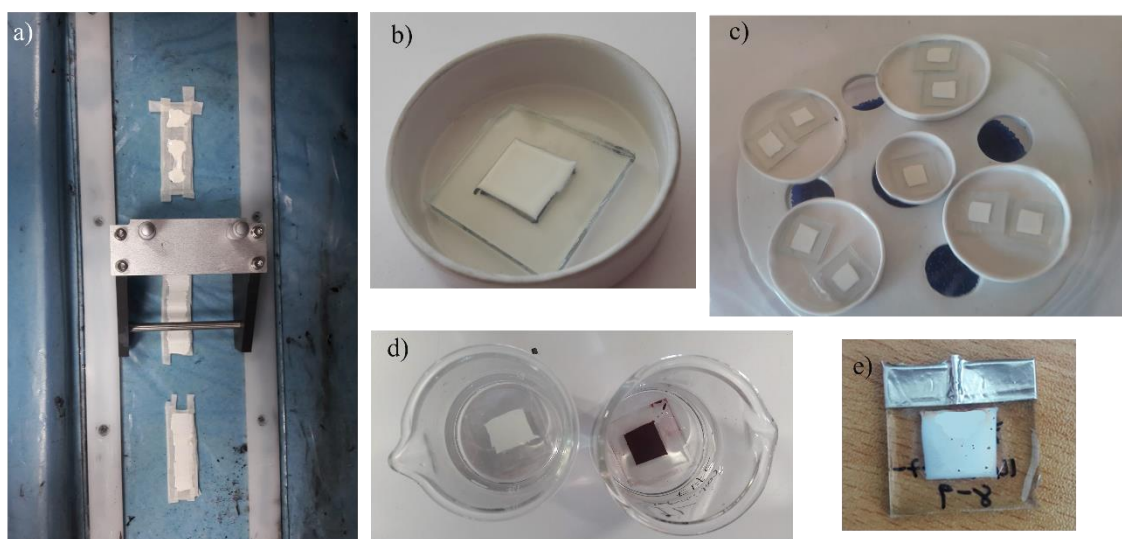


Figure 3.2. Different steps of TiO_2 based photoelectrode preparation a) Coating of FTO with Ti-nanoxide D/SP paste with doctor blade method, b) After coating, c) After drying at 475 °C, d) Dye sensitized and plain TiO_2 electrodes in [BMIM][BF₄]: acetonitrile solution, e) IL coated TiO_2 electrode.

3.2.3. Characterization

X-Ray powder diffraction analysis of photocatalysts were carried out by Rigaku D/MAX-Ultima+ with monochromated Cu-K α radiation (40 kV, 40 mA). The morphologies, particle sizes, and particle distributions of samples were analyzed by high contrast transmission electron microscopy (CTEM, FEI Tecnai G2 Spirit BioTwin) and scanning electron microscope (FEI Philips XL30 ESEM-FEG) equipped with energy dispersive X-ray analysis unit (SEM-EDAX). Infrared spectra of the samples were observed with a DRIFT FT-IR Jasco 6100 spectrometer with a resolution of 1cm⁻¹. The samples were mixed with dry KBr in the ratio of 1:1 (w:w) prior to analysis. UV-Vis diffuse reflectance spectra of the samples were obtained using Shimadzu UV-2600 with integrating sphere attachment unit. The reflectance and absorbance data were processed by the Kubelka-Munk function to obtain band gap values of particles. XRD and SEM-EDAX measurements were performed by Boğaziçi University Advanced Technologies Research Center while CTEM analysis was carried out by Middle East Technical University Central Laboratory; the other tests were performed in-house.

3.2.4. Photocatalytic Measurements

A simple illustration for the photocatalytic reaction system was given in Figure 3.3. The photocatalytic activities of Pt/TiO₂, IL:Pt/TiO₂, Dye:Pt/TiO₂, and IL-Dye:Pt/TiO₂ were tested by measuring hydrogen production using gas chromatography in a semi-batch reaction system. A custom-made 620 mL glass volumetric flask was used, and the dead volume was kept constant in each experiment. The ratio of photocatalyst weight to reaction solution was constant at 2 g/L in all experiments; 1.2 g of photocatalyst was suspended in 600 mL reaction solution which contained 10% triethanolamine (TEOA) and 90% distilled water (v/v). Since IL coated particles involve 0.2 g of IL per one g of active Pt/TiO₂, 1.44 g of photocatalyst (assumed to contain 0.24 g of IL and 1.2 g of Pt/TiO₂) was used in the experiments that contain IL; the contribution of dye to the weight of dye sensitized particles were negligible (0.0014 gr dye per 1.2 gr Pt/TiO₂). The reaction tests were conducted at room conditions. Prior to irradiation, the system was bubbled by Ar gas at 30 mL/min for an hour; the Ar flow also continued during the experiment under light irradiation with the same flowrate to carry the hydrogen produced to GC. ScienceTech IPS-300 (300 W) solar simulator was used as the light source. Shimadzu GC-2014 gas chromatography equipped with a thermal

conductivity detector and Teknokroma TR-GC1462010 packed column (4.6 m x 1/8'' up to 330°C) were used under the flow of Argon as the carrier gas. The amount of hydrogen produced was measured at one-hour intervals.

3.2.5. Photoelectrochemical Measurements

The photoelectrochemical experiments were performed in a three electrode assembly where TiO_2 , IL: TiO_2 , or IL-Dye: TiO_2 photoelectrode, platinum counter electrode, and saturated calomel reference electrode (SCE, 3M KCl) were used. The active area of the working electrode was kept constant at 1cm^2 . Aluminum wire for connection was attached to the edge of FTO substrate (outside the photoactive area) with conductive aluminum tape and sealed with epoxy resin. A 0.1 M Na_2SO_3 solution (in 10%TEOA aqueous solution by volume) was used as electrolyte. The electrochemical tests were carried out using a potentiostat (Biologic Sp-300 Science Instruments) in both dark and illuminated condition. ABET 1.5 G solar simulator was used as the light source with the intensity of 100 W/m^2 . Cyclic voltammetry (CV) measurements were performed in the potential range -1 V to 1 V at a scan rate of 20 mV/s as three sets of five CV cycles with ~30 min waiting time between sets.

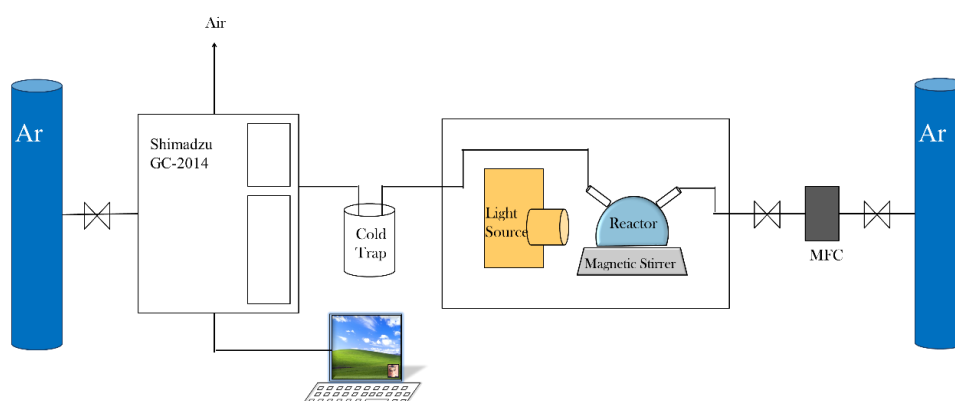


Figure 3.3. The simple illustration of our photocatalytic reaction system.

3.2.6. Critical Factors in Reactor Design and Operations

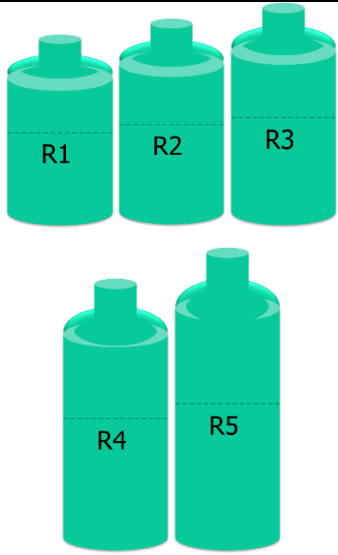
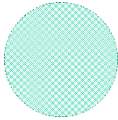
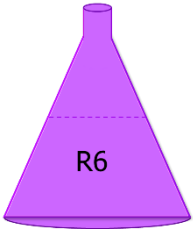
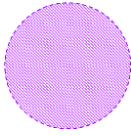
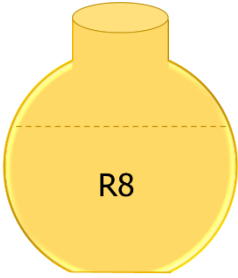
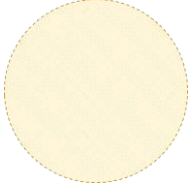
In the study of critical factors which are effective on photocatalytic water splitting through mass transfer limitations and measurement errors, Pt/TiO_2 was used as semiconductor and the preparation procedure was given in 3.2.2.3.

The basic structure of the glass reactor and reaction system in general are presented in Table 3.14 and the pictures of reactors are given in Figure 3.4. To test the effect of dead volume and gas-liquid interfacial area eight custom-made glass reactor configuration were used to be able to keep one these two variables constant while changing the other (see Table 3.14. The catalyst concentration (gr catalyst per liter of reaction solution) was also kept constant at 2 g/L while the experiments for interfacial area were also repeated at 3g/L to see the effect of catalyst concentrations. The reaction solution contained 20% methanol (MeOH) and 80% distilled water (v/v) in all experiments and conducted under room condition and the rotating speed of 420 rpm; the effect of rotation speed was also analyzed by changing its value as 200, 300, 420 and 450 rpm. Prior to illumination, the system was purged by argon gas for an hour; the Ar flow also continued during the experiment. The Ar flow rate was kept constant at 30ml/min while the effects of other variables were investigated; however, it is the effects of Ar flowrate was also investigated (performing tests at 15 ml/min, 30 ml/min and 45 ml/min). Shimadzu GC-2014 gas chromatography equipped with a thermal conductivity detector and 4.6 m x 1/8'' Teknokroma TR-GC1462010 packed column was used to analyze the product gases; the argon was used as the carrier gas. ScienceTech IPS-300 Class ACA solar simulator (250-2000 nm, 1 Sun, without UV filter) was used as the light source. The experiments were conducted for six hours, and hydrogen production data was recorded at one-hour intervals.



Figure 3.4. The pictures of reactors R1, R2, R3, R4, R5, R6, R7, and R8.

Table 3.14. Eight different custom-made glass reactors used.

Form of Reactor and Name of Reactor	Cross-sectional Area of Reactor (cm ²)	Total Volume of Reactor (ml)
	 30 cm ² (same through the tube for all of them)	R1: 745 ml R2: 510 ml R3: 385 ml R4: 350 ml R5: 330 ml
	 42 cm ² (while dead volume is 150 ml)	540 ml
	 72 cm ² (while dead volume is 150 ml)	475 ml
	 90 cm ² (while dead volume is 150 ml)	620 ml

4. RESULTS and DISCUSSION

The computational and experimental results are given and discussed in this section, respectively. Machine learning analysis of each dataset with a detailed review of the corresponding topic is given under separate subsections. The results of experiments are discussed chronologically to establish a comprehensible cause effect relationship between our actions.

4.1. Machine Learning in Photocatalytic Hydrogen Production Over Perovskites

Literature

Scientific articles included in the dataset are reviewed in detail before the using machine learning tools analysis to determine some previously unknown relations between input variables of PWS using simple statistical tools; this way the interpretation of ML results may become easier. The results examined in this section 4.1 was published as a paper (Can and Yildirim, 2019b) and revised for this thesis.

4.1.1 Perovskite Materials

The perovskites are the metal oxides with the general formula of ABX_3 ; A and B symbolize large and small cation respectively while X usually denotes oxygen (or occasionally halogens and S). The classification of perovskites by B-site is a common practice and it is also suitable for PWS (X. Chen *et al.*, 2010; Liu *et al.*, 2007). The hydrogen evolution performances of the most common perovskites (based on B-sites) are given in Figure 4.1a. The bubble sizes (and nearby numbers) indicate the frequency of perovskites in database containing those materials in their B-sites while the y-axis represents their average hydrogen evolution rate. The data obtained with UV (blue balls) and visible (red balls) light are presented separately. Since the number of datapoints in database is sufficiently large, the analyses presented below should represent the entire literature reasonably well.

Tantalates (Ta in B-site) are the most common perovskites in database; about 40 % of data (216 of 541 cases) were in this class. They are followed by titanates (Ti containing), which forms 24 % data (128 cases), niobates (Ni containing), which constitutes 18% of data

(96 cases) and indium containing perovskites (8% corresponding to 42 cases). As expected, the hydrogen production rates are higher under UV. The tantalates and niobates perform much better under UV while the visible light performance of tantalates is also high. The average visible light activity of In catalysts is higher than their average UV performance because S containing perovskites (AInS) are used for visible light while AInOs are used under UV.

Figure 4.1b shows the frequency and performance of the most common perovskites according to A-sites; only perovskites having more than 10 cases with same A are presented in Figure 4.1b for better generalizations. If A and B sites are brought together, it can be easily seen that the most common five perovskites are NaTaOs (92 cases), SrTiOs (71 cases), BiTaOs (33 cases), NaNbOs (33 cases) and KTaOs (25 cases). Here, ABO symbolizes the general structure (not stoichiometric formula) including all bare and doped ABO_3 as well as the layered perovskites. More than two third of NaTaOs (66 of 92 points) were tested under UV and resulted much higher average hydrogen production rate than that under visible light. For SrTiOs, however, 47 of 71 cases were obtained under visible light (only 24 UV cases). No visible light data was available for BiTaOs while NaNbOs were investigated under both UV (11 cases) and visible light (22 cases).

NaTaOs are not only the most common but also one of the most effective perovskites in the database (Figure 4.1b). As the most common material of the group, NaTaO₃ has a cubic crystal structure, which is believed to generate nano-step morphology and enhances the hydrogen production (Gómez-Solís *et al.*, 2014; Liu *et al.*, 2007) while its high crystallinity improves PWS by inhibiting electron-hole recombination (Gómez-Solís *et al.*, 2014). Additionally, its high surface area contributes to PWS activity by increasing active sites (Gómez-Solís *et al.*, 2014). Their bond angle of 160°-180° may be also increasing the effectiveness of tantalates because the ideal bond angle of B-O-B in perovskite was found to be 180°, which is the best for its luminescent property (Li *et al.*, 2008; Liu *et al.*, 2007; Obee and Brown, 1995). SrTaO₃, BiTaO₃ and KTaO₃ follow NaTaO₃ in hydrogen production with decreasing order. These perovskites have been also modified with ion doping (Iwase *et al.*, 2009; Li and Wu, 2015; Nikolaidis and Poullikkas, 2017). They also have been modified with use of co-catalyst (Hu and Teng, 2010; Iwase *et al.*, 2009; Li and Wu, 2015; Nikolaidis and Poullikkas, 2017; Zielińska *et al.*, 2008) and modifying the preparation methods (Gómez-Solís *et al.*, 2014; Li *et al.*, 2008) as discussed below.

SrTiOs are overwhelmingly the most common (55%) titanates in PWS; although their performances are not comparable to NaTaOs, they were frequently studied probably because Ti is more abundant. Besides, the band gap, surface area, particle size and crystallinity of SrTiOs can be easily tuned (Yang Liu *et al.*, 2008). For example, Puangpetch *et al.* obtained a red-shift in absorption spectra of SrTiO₃ with co-catalysts like Au, Ag, Pt, Ni, Fe, and Ce and enhanced hydrogen yield (Au was the best) under UV (Puangpetch, *et al.*, 2010). Au/SrTiO₃ was also found to be effective (under visible light) by Saadetnejad and Yildirim (Saadetnejad and Yildirim, 2018). CaTiOs, ZnTiOs, SnTiOs, KTiOs, BiTiOs and less common titanates like A₂Ti₂O₇ (A=Ga, La, Y) (Abe, Higashi, Sayama, Abe, and Sugihara, 2006b) were also investigated. Inserting nitrogen to oxygen site of ATiO₃ was also tested by Naik *et al.* with a red-shift in absorption spectrum (Naik *et al.*, 2011).

As tantalates, the niobates also have high conduction band levels, which were reported to improve hydrogen production by facilitating reduction (Gómez-Solís *et al.*, 2014; Iwase *et al.*, 2009; Liu *et al.*, 2007). The band gaps of niobates were lower than those of tantalates; hence, they are more suitable for visible-light harvesting (Liu *et al.*, 2007). Na is the most frequently used A-site element even though the hydrogen production rates over KNbOs and LaNbOs are higher. The indium based photocatalysts also have considerable hydrogen production rates although the number of articles involving this class is rather limited. ZnInSs are the most common (and effective) In perovskites in database with low toxicity, high visible-light activity, and stability. For example, Li *et al.*, synthesized and tested ZnIn_{1.5}S_{3.2575}, which kept its activity for 150 hours; Ag doping to A-site (up to certain loading) further improved the activity under visible light (Li *et al.*, 2014b). These results were also verified over ZnIn_{2.3}S₄ (Song *et al.*, 2015).

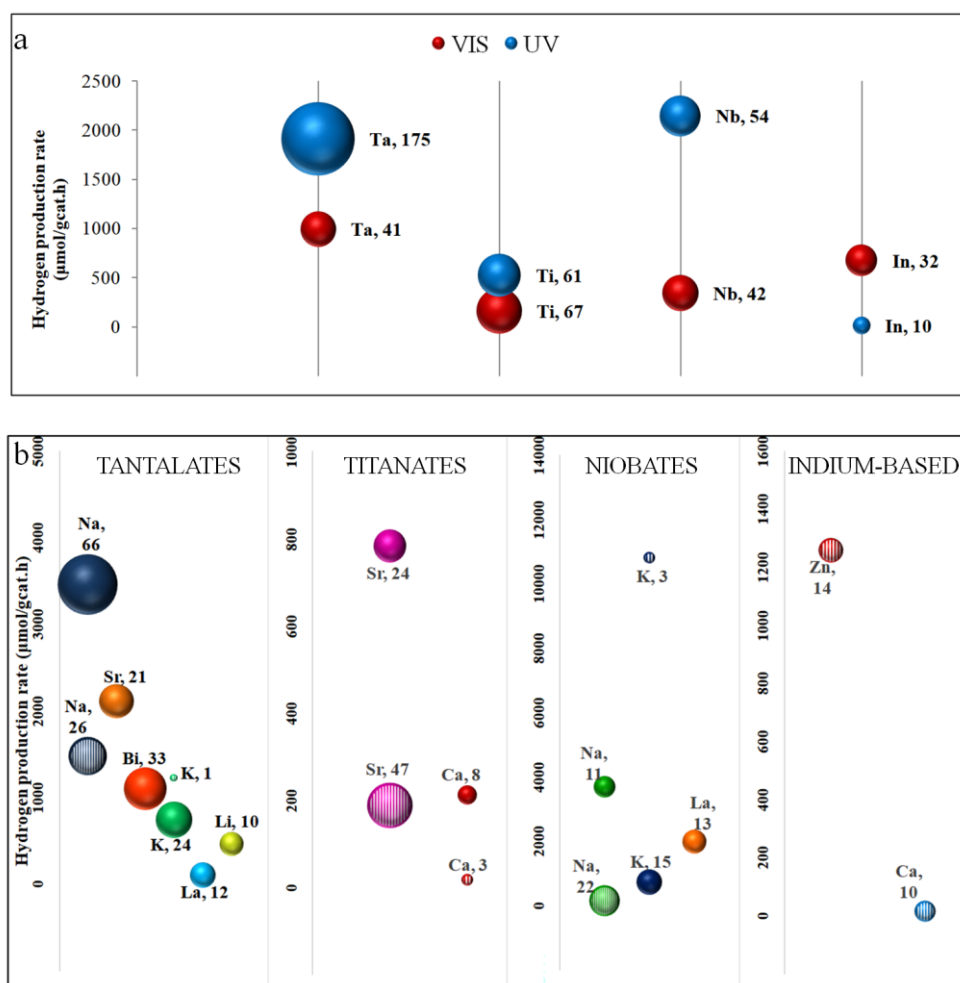


Figure 4.1. Average hydrogen production rates based on a) common B-site elements and b) common A-site elements (solid balls are for UV; dashed balls are for visible light).

4.1.2. Ion Doping for Band Gap Modification

Ion doping is a common and effective way to tune the band gap of perovskites. Doping changes the valence or conduction bands so that visible light harvesting may be possible (Chen *et al.*, 2010). Borgarello *et al.* studied ion doping in 1982 for the first time and found that Cr doped TiO₂ produced more hydrogen under visible light than bare TiO₂ (Chen *et al.*, 2010). Since then, the metal and non-metal doping to semiconductors have been extensively studied. Indeed, 260 data points out of 540 (48%) in database were obtained over doped perovskites (31% A-site, 27% B-site). There were also works involving both A and B sites doping (about 10%) and doping one site with two elements; X-site doping (for oxygen) were also recorded.

4.1.2.1. A-site doping

For A-site doping, 18 elements were identified in the database; La (38 times) and Ca (32 times) were the most common elements followed by H (20 times), K (13 times), Br and Sr (10 times each). Ca, La, K, and H were the most frequent choices for tantalates; for example, 25 out of 32 cases of Ca doping involved tantalates while each of La, K and H appeared 12 times. However, La was overwhelmingly the most preferred element (20 times in database) for titanates while the other dopants were rarely used. H, Ca and La doped niobates were almost equally present (8, 7 and 6 times respectively); we had only Ag doping examples for In perovskites. Further details for the doping to common perovskites are given in Table 4.1.

Table 4.1. Frequency of A and B doping for the most commonly used perovskites (i.e., complete list of dopants for most common perovskites).

	Total	A-Single Doped	A-Double Doped	B-Single Doped	B-Double Doped	A and B Co-doped	X-Single Doped
Tantalates	200	43	23	32	0	17	7
BaTaO	6	Ni (1)	0	Zr (1)	0	0	0
BiTaO	33	Ca (6), Ba (6), Sr (1)	K-La (5)	Cu (4)	0	0	0
CaTaO	8	0	0	Zr (7)	0	0	N (6)
InTaO	3	Ni (1)	0	0	0	0	N (1)
KTaO	25	0	Ca-Sr (18)	Zr (1)	0	La-Bi (1)	0
LaTaO	12	K (1), H(1) Ba (1), Bi (1), Ca (1), Ce (2), K (6), La (9), Sr (2), Y (2), Yb (2)	0	Al (7)	0	0	0
NaTaO	92		0	Bi (6), Nb (6)	0	La-Co (1), La- Cr (1)	0
SrTaO	21	Bi (1), H (1)	0	0	0	H-Nb (10), Bi- Ce (4)	0
Titanates	109	13	0	17	7	22	2
BaTiO	3	0	0	0	0	La-Cr (2)	0
BiTiO	7	Na (6)	0	0	0	0	N (1)
CaTiO	11	0	0	0	0	La-Cr (8)	0
CrTiO	1	0	0	Ta (1)	0	0	0
KTiO	7	Sn (1)	0	0	0	0	0
NaTiO	8	La (1)	0	Zr (2)	0	0	0
SmTiO	1	0	0	0	0	0	S (1)
SrTiO	71	Ag (1), Cr (2), La (1), Na (1)	0	Al (13), Fe (1)	Rh-Sb (7)	La-Cr (8), Cr- Sb (1), Bi-Fe (1), Cr-Ta (1), Ni-Ta (1)	0
Niobates	89	20	0	25		11	0
BaNbO	6	Ni (1), Zn (2)	0	Co (1), Ta (1)	0	0	0

Table 4.1. Frequency of A and B doping for the most commonly used perovskites (i.e., complete list of dopants for most common perovskites). (cont.)

	Total	A-Single Doped	A-Double Doped	B-Single Doped	B-Double Doped	A and B Co-doped	X-Single Doped
BiNbO	9	Ba (1), Ca (1), Sr (1)	0	0	0	La-Al (6)	0
CoNbO	3	Ba (1), Ca (1), Sr (1)	0	0	0	0	0
InNbO	3	Ba (1), Ca (1), Sr (1)	0	0	0	0	0
KNbO	18	0	0	Zr (6)	0	0	0
LaNbO	13	H (1)	0	Ta (5)	0	H-In (5)	0
NaNbO	33	Ca (4)	0	In (11)	0	0	0
SrNbO	4	H (3)	0	Al (1)	0	0	0
Indium-based	22	5	0	6	0	0	0
BaInO	1	0	0	Ta (1)	0	0	0
CuInS	7	0	0	Ga (5)	0	0	0
ZnInS	14	Ag (5)	0	0	0	0	0
Others							
ZnCdS	15	Cu (5)	0	0	0	0	0
SrSnO	9	Ba (1)	0	0	0	0	0

Although NaTaOs are the most effective photocatalysts for PWS, they lose their activity fast; hence, they are usually doped to suppress deactivation. About 30% of NaTaOs (28 out of 92) were doped in A-site with La and K or other less frequent elements. Various investigators reported that nickel co-catalyst and La dopant, together or separately, decreased electron-hole recombination and prolonged life of NaTaOs by creating new active sites (Husin *et al.*, 2017; Husin *et al.*, 2011; Iwase *et al.*, 2009; Kato *et al.*, 2003; Kudo and Kato, 2000). Doping generally expands surface area (more active sites) and decreases particle size (shorter distance for photogenerated electrons to travel). For example, $\text{La}_x\text{Na}_y\text{TaO}_3$ resulted in much higher hydrogen production rate than bare NaTaO_3 under UV light (Husin *et al.*, 2017, 2011; Kato *et al.*, 2003; Kudo and Kato, 2000). Kudo *et al.* and Kato *et al.* also observed that solid-state reaction method created unique step structures on $\text{La}_x\text{Na}_y\text{TaO}_3$ particles and promoted hydrogen yield by generating new active sites (Kato *et al.*, 2003; Kudo and Kato, 2000). The same group also studied Ca, Ba and Sr doped NaTaO_3 perovskites and obtained similar results (Iwase *et al.*, 2009). Jana *et al.* doped rare elements (Y, La, Ce, and Yb) on NaTaO_3 , and found that Y: NaTaO_3 and La: NaTaO_3 had best performances (Y: NaTaO_3 was better) (Jana *et al.*, 2014a). The performances of A-site doped NaTaOs are shown in Figure 4.2a as an example. La doping slightly improved average UV and visible light performances; however, Ba and Ca were more effective. The performance

with other dopants was close to that of bare NaTaOs. The other tantalates are also doped frequently. For example, almost 50% of BiTaOs, the second most common tantalates, were mostly doped with Ba, Ca, and K; however, only the less frequently used Sr produced better result than undoped BiTaOs (Chen *et al.*, 2014); the average performances of Ba, Ca and K doped, and undoped BiTaOs were almost the same.

Double metal or nonmetal doping to tantalates were also investigated. For example, Takayama *et al.*, tested $\text{KCaSrTa}_5\text{O}_{15}$ with several promoters (NiO, Ni, Cu, Ru, Rh, Ag, Pt, and Au), and obtained the highest performance with NiO under UV (Takayama *et al.*, 2014); Chen *et al.* reached to a similar result over $\text{K}_{0.5}\text{La}_{0.5}\text{Bi}_2\text{Ta}_2\text{O}_9$ (Chen *et al.*, 2012). Shimizu *et al.* compared the activity of $\text{H}_2\text{La}_{2/3}\text{Ta}_2\text{O}_7$, $\text{K}_2\text{La}_{2/3}\text{Ta}_2\text{O}_7$, and $\text{H}_2\text{SrTa}_2\text{O}_7$ with KTaO_3 ; H doping improved hydrogen production by facilitating the transfer of photogenerated electrons and holes (Shimizu *et al.*, 2005). Similarly, Huang *et al.* compared the efficiency of $\text{HSr}_2\text{TaNb}_2\text{O}_{10}$ with the perovskites doped with K; although they had nearly the same surface area, band gap and interlayer distance, H-doped perovskites provided better production rates (Huang *et al.*, 2011).

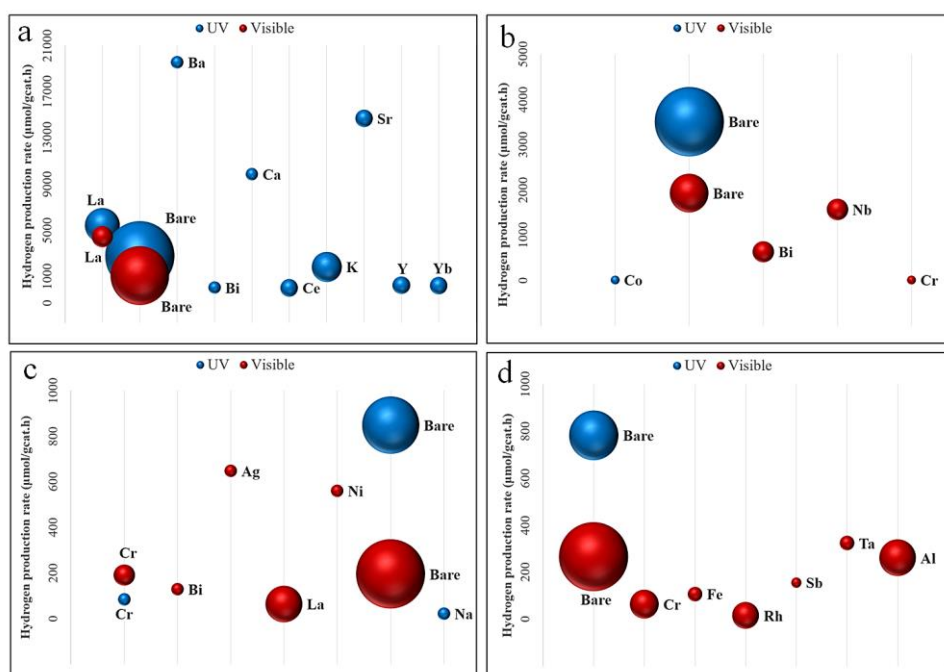


Figure 4.2. The frequencies and average hydrogen production rates for a) A-site doped NaTaOs, b) B-site doped NaTaOs c) A-site doped SrTiOs, and d) B-site doped SrTiOs.

A-site doping for SrTiOs was generally used to make these perovskites visible light active. However, the doping seemed to fail to improve activity in average; only few examples involving Ag and Ni produced better results than bare perovskites (Figure 4.2c). For example Ag doping produced the hydrogen production rate of $648 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ compared to the average of $243 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ for bare SrTiO₃ (Ishii *et al.*, 2004). However, La doping, which appeared in nine cases from three articles (Cai *et al.*, 2015; Lu *et al.*, 2017; Luo *et al.*, 2008) in database, did not produce the same improvement. In one of these works (Cai *et al.*, 2015), Luo *et al.* doped La to A-site of SrTiO₃ and obtained the hydrogen production rate of $67 \mu\text{mol gcat}^{-1} \text{h}^{-1}$; the other two groups (Lu *et al.*, 2017; Luo *et al.*, 2008) were also doped Cr in B site but they could not obtain improvement either. The hydrogen production rate over Cr (Wang *et al.*, 2010) and Na (Takata and Domen, 2009) doped SrTiO₃ was also lower than the bare catalyst under UV.

The other titanates were also doped by their A-sites. For example, Li *et al.* synthesized Sn:KTiO and achieved a red shift in absorbance spectrum (the band gap was 2.8 eV while it was 3.5 eV for bare KTiO); the average hydrogen production rate was $257 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ (Li *et al.*, 2011). A-site of NaTiO₃ was also doped with La and its performance was compared with La and Cr co-doped (A and B site) NaTiO perovskite (Shi *et al.*, 2011). La:NaTiO did not results in significant hydrogen production rate; however, unlike to La,Cr:SrTiO discussed above (Lu *et al.*, 2017; Luo *et al.*, 2008), a sharp increase in the hydrogen production rate was observed when Cr is co-doped to B site of La:NaTiO₃.

It was difficult to make generalization for the A-site doping for niobates because the number of cases is limited (12 out of 96) while the number of dopants is high (nine different elements) in the database. The notable ones are H doping to SrNbO₃ and LaNbO₃ (Li *et al.*, 2009; Natarajan *et al.*, 2017) and Ca doping to NaNbO₃, BiNbO₃, Co_xNb_yO₃, and InNbO₃ perovskites (Li *et al.*, 2009; Takayama *et al.*, 2014; Wei *et al.*, 2009). H_{2.33}Sr_{0.67}Nb₅O_{14.335} performed better than H₂LaNb₂O₇ (even better than H-In co-doped LaNbO₃) under UV. In both of those works (Li *et al.*, 2009; Wei *et al.*, 2009), Pt was used as promoter, and their band gap values are similar to each other but H_{2.33}Sr_{0.67}Nb₅O_{14.335} had larger surface area, which may be the reason of its better activity. In another work, Yin *et al.* tested Ca, Sr, and Ba doping to In_{0.5}Nb_{0.5}O₃ and Co_{1/3}Nb_{2/3}O₃ under UV in two separate articles (Yin *et al.*, 2003a, 2003b); the performances of doping agent were in the order of Ca > Sr > Ba for InNbO₃ while the order were as Ba > Ca > Sr for Co_xNb_yO₃. Li *et al.* prepared Ca, Sr, Ba,

and $\text{K}_{0.5}\text{La}_{0.5}$ doped $\text{Bi}_2\text{Nb}_2\text{O}_9$, and observed a red-shift in the doped catalysts compared to bare $\text{Bi}_2\text{Nb}_2\text{O}_9$. Sr-doped catalyst was the best while Ca and Ba-doped had worse performances under UV (Li *et al.*, 2010).

4.1.2.2. B-site doping

Examples containing 16 metals and transition metals doped for B-site were recorded in database. Doping of transition metal was reported to change donor-acceptor levels of perovskites so that band gap tuning was possible (Chen *et al.*, 2010). The tantalates were mostly doped by Nb (16) while Cr (18 cases) and Al (13 cases) doping were more common in titanates. The dominant choice for niobates was In with 16 cases even though Al, Ta and Zr were also used. Indium containing catalysts were B-doped less frequently (only 7 out of 42 data points).

NaTaOs were the most commonly B-doped tantalates in the database. For example, Bi doped Pt/NaTaO_3 was investigated by two groups in 2012. Kang *et al.* prepared Pt/Bi:NaTaO_3 by spray pyrolysis method while Kanhere *et al.* used solid state reaction (SSR) (Kang *et al.*, 2012; Kanhere *et al.*, 2012). The former group claimed that unique surface morphology and large surface area increased the photocatalytic activity under visible light. The latter group, however, observed that their synthesis condition increased hydrogen yield by narrowing the bandgap (Kang *et al.*, 2012; Kanhere *et al.*, 2012). Wang *et al.* doped Nb to B-site of NaTaO_3 at various ratios (M. Wang *et al.*, 2017), and $\text{NaNb}_{0.5}\text{Ta}_{0.5}\text{O}_3$ showed 3.8 times higher hydrogen production than bare NaTaO_3 under visible light; this difference attributed to the lower band gap and defects generated by Nb. However, despite of these positive results, the average performance of Bi and Nb doped NaTaOs were worse than the bare perovskites (Figure 4.2b); this was also true for the other dopants.

B doping was also applied to other tantalates. For example, Huang *et al.* produced $\text{HSr}_2\text{TaNb}_2\text{O}_{10}$ with SSR and obtained 11 times higher hydrogen production rate than TiO_2 . They argued that $\text{HSr}_2\text{TaNb}_2\text{O}_{10}$ possessed a stable and special structure, which enabled separation of photogenerated electron-hole pairs (Huang *et al.*, 2011). Senthil *et al.* studied Bi (A-site) and Ce (B-site) doped SrTaO_3 , and found that, with Bi^{3+} and Ce^{4+} doping, the dipole moment was created; this improved the photocatalytic activity by enhancing charge separation (Senthil *et al.*, 2016). Li *et al.* prepared $\text{La}_2\text{AlTaO}_7$ by SSR and studied under UV; Al doping lowered the conduction band, and the hydrogen production was succeeded

even without a co-catalyst (Li *et al.*, 2009b). Zhang *et al.* prepared $\text{BiTa}_{1-x}\text{Cu}_x\text{O}_4$ with different amount of Cu (Zhang *et al.*, 2009); the best performance under UV light was obtained over $\text{BiTa}_{0.98}\text{Cu}_{0.02}\text{O}_4$. Since all particles showed similar light diffuse reflectance spectra, the differences in activity were attributed to surface area change (Zhang *et al.*, 2009). Ishihara *et al.*, investigated the effects of Zr doping to KTaO_3 , and obtained much higher activity than bare KTaO_3 under UV (Ishihara *et al.*, 1999).

SrTiO_3 is the most commonly B-doped perovskites, not only in titanates but also in the entire database (33 data points out of 71). Asai *et al.*, studied Rh-doped and Rh-Sb co-doped SrTiO_3 in the presence of Ir co-catalyst, and they obtained the highest hydrogen yield over Rh,Sb: SrTiO_3 under UV. They argued that the doped metal ions worked as recombination centers unless another metal cation was co-doped to maintain the charge balance (Asai *et al.*, 2014). Recently, Ham *et al.* synthesized Al: SrTiO_3 by flux-mediated method and SSR. For pure SrTiO_3 , the higher surface areas and enhanced crystallinity were obtained by flux-mediated method; the activity was also more than that obtained with SSR. However, in both case, the activity first increased with the increasing Al loading and then decreased with further increase of Al content (Ham *et al.*, 2016). Luo and Maggard synthesized Fe doped SrTiO_3 with hydrothermal method (HT); however, they obtained no significant improvement compared to the bare SrTiO_3 (Luo and Maggard, 2006).

Co-doping of two elements to A and B sites of SrTiO_3 was also tested; Kang *et al.* co-doped Ta to B-site and Cr or Ni to A-site; they obtained very large surface area and unique surface morphology that enhanced the photocatalytic activity under visible light (Ni gave better result); it was attributed that Ni and Ta ions created a charge balance around Ti and made the surface morphology more suitable for PWS (Kang *et al.*, 2012; Kang and Park, 2011). In another work, Kato and Kudo studied Cr doping to A-site and Sb doping to B-site of SrTiO_3 ; they achieved much higher visible light activity than those over single doped or undoped catalysts (Kato and Kudo, 2002b). However, despite all these examples, the overall average performances of B doped SrTiO_3 were not better than plain perovskite (Figure 4.2d).

The B-site doping was also implemented to other titanates. Huerta *et al.* observed an improvement in the activity of Zr doped $\text{Na}_2\text{Ti}_6\text{O}_{13}$ compared to bare catalyst and attributed this to the different bond lengths and charge distribution of doped and bare samples. They also argued that Zr acted as an electron trap, which inhibited the recombination of electron-

hole pairs (Huerta-Flores *et al.*, 2017). Wang *et al.* and Lu *et al.* synthesized La, Cr co-doped CaTiO_3 , and both observed red shift in the band gap making this perovskite suitable for visible light (Lu *et al.*, 2017; Rajeshwar *et al.*, 2015).

The niobates were also B-doped (but less frequently than tantalates and titanates). For example, Lv *et al.* produced In doped NaNbO_3 particles with co-precipitation and observed surface area increase; however, the hydrogen production showed a plateau behavior indicating that improvement may be due to the better charge separation, not the surface area (Lv *et al.*, 2010). Iwakura *et al.* doped B-site of SrNbO_3 by Al, and changed the ordered and disordered material ratio in $\text{SrAl}_{0.5}\text{Nb}_{0.5}\text{O}_3$ by changing calcination time; the disordered $\text{SrAl}_{0.5}\text{Nb}_{0.5}\text{O}_3$ gave better hydrogen production rates under UV than ordered perovskites (Iwakura *et al.*, 2010); Li *et al.* investigated Cu doping to LaNiO_3 ; they claimed that Cu doping created a redox cycle of $\text{Cu}^{2+}/\text{Cu}^+$ over the pre-reduced $\text{Cu}:\text{LaNiO}_3$, and promoted photocatalytic activity. They also found that pre-reduction treatment affected activity by changing electron-hole recombination rate, interfacial charge transfer and electronic structure (Li *et al.*, 2010).

Co-doping in niobates was also investigated. For example, Li *et al.* co-doped La on A-site and Al on B-site of Bi_2NbO_7 (Li *et al.*, 2009); they observed that La doping enhanced the activity under UV by changing electronic structure and optical absorption. Similarly, Wei *et al.* compared the performances of co-doped; they observed that $\text{H,In}:\text{LaNbO}_3$ was one of the best due to its larger surface area and lower band gap (Wei *et al.*, 2009). In another work, Li *et al.* studied $\text{ALaTa}_{x/3}\text{Nb}_{2-x/3}\text{O}_7$ ($\text{A} = \text{K, H}; x = 0, 2, 3, 4 \text{ and } 6$), and obtained the best performance over $1\%\text{Pt}/\text{HLaTa}_{2/3}\text{Nb}_{4/3}\text{O}_7$. They claimed that Ta caused a new and more stable HLaNb_2O_7 form and created an electronic structure that helped the separation of photo-generated electron-hole pairs (Li *et al.*, 2010).

4.1.2.3. Doping effects on band gap

As presented above, there are numerous successful applications of doping in the literature; however, the average performance does not change significantly for most dopants. This can happen only if there are both high and low performances with the same dopant indicating that the other factor may be inflating or overshadowing the doping effects. In successful applications, the improvements were usually attributed to the shift in band gap, increase in surface area and decrease in particle size. Consequently, the effects of doping on

these properties (if there are any) were also analyzed. No definitive conclusion could be drawn for the surface area and the particle size; however, there seems to be some patterns for band gap (Figure 4.3) even though they cannot be generalized because they are unique for the perovskite and the dopants.

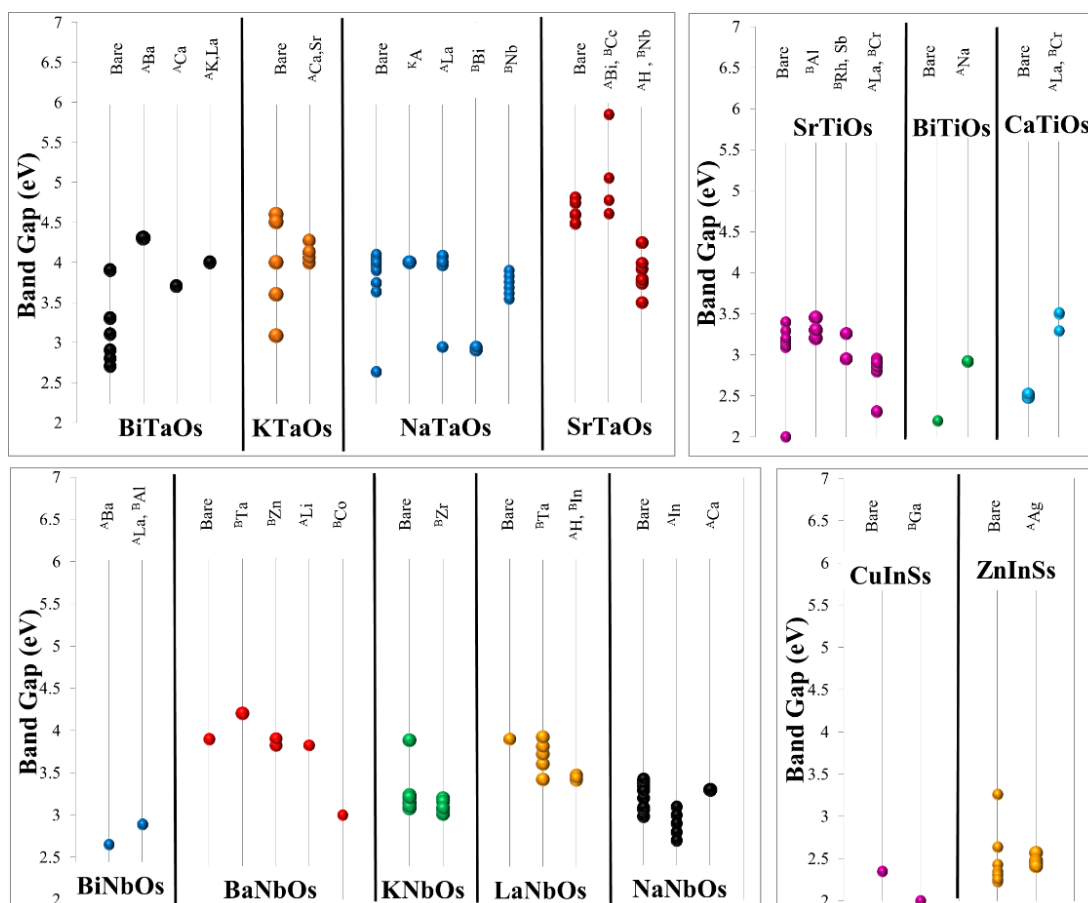


Figure 4.3. Effects of doping on band gap of most commonly used perovskites.

4.1.3. Preparation Methods for Perovskites

The solid-state reaction (SSR) is the most common perovskite preparation method (188 cases in database); this followed by hydrothermal synthesis (HT) with 148 cases. Methods like polymerized-complex (PC), sol-gel (SG), ion-exchange (IE), electro-spinning (ES), Pechini-type process (PTP), flux-type and co-precipitation are also utilized. Figure 4.4a shows the frequency distribution of cases for the preparation methods together with the average hydrogen production. There are also novel and rarely used methods that are not included to Figure 4.4 (58 cases). Data for UV and visible light are presented separately;

although the light source has no direct effects on catalyst preparation, it influences the method selection due to the band gap considerations.

Only HT and IE outperform SSR in average hydrogen production; although the superiority of HT against SSR is known (Li *et al.*, 2008), having similar performance with SSR, SG and PC is surprising because the last two are also known to be better. For example, Jeong *et al.* reported that $\text{NiO/Sr}_3\text{Ti}_2\text{O}_7$ prepared by PC gave 1.33 times higher yield than SSR catalyst under UV due to its higher surface area and purity; PC method also prolonged the lifetime of catalyst (Jeong *et al.*, 2006). Hu and Teng produced orthorhombic and cubic NaTaO_3 using SSR and SG respectively; SG-catalyst provided 23 times higher hydrogen production rates, and this was attributed to higher surface area, smaller particle size and high crystallinity due to lower calcination temperatures of SG (about 500°C against 1200°C for SSR) (Hu and Teng, 2007). The results for NiO/InTaO_4 are also consistent with the results of previously mentioned article (Chiou *et al.*, 2009); SG catalysts had smaller particle sizes, higher surface areas and enhanced photocatalytic activities. These examples indicate that SG and PC methods are superior as long as they provide the desired physical properties described above, and apparently, this does not always happen as will be discussed later.

The HT method was also compared with SSR and PC in production of NaNbO_3 (Li *et al.*, 2008) HT provided twice-higher yield than SSR while PC resulted in more than six times higher yield under UV if the calcination temperature was 600°C ; however, the yield was only 50% higher when the calcination temperature was 900°C . The authors claimed that smaller catalyst size by PC (at 600°C) may have shortened the diffusion distances of photogenerated electrons and increased the number of electrons reached to the active sites. Additionally, HT caused rectangular and cubic samples while SSR method resulted in spherical particles. The authors argued that the edges of HT catalysts may act as active sites. Hu *et al.* observed that UV activity of KTaO_3 prepared by SG was twice higher than that prepared by HT (Hu *et al.*, 2013). The higher surface area and advantageous crystalline structure of SG catalyst facilitated charge separation via more flattened Ta-O-Ta bond. New approaches were also tested and compared with conventional methods. For example, Liu *et al.* produced SrTiO_3 using PC, SSR and milling assistant method (Li *et al.*, 2008); PC catalysts were the best due to their smaller size and uniformity of their components while the milling assistant method caused catalyst agglomeration and degradation. Similarly, Lee *et al.* synthesized SrSnO_3 by simple wet chemical reaction and compared the results with

SSR; the catalysts obtained via the former method resulted in smaller particle size, better crystalline structure and five to ten times higher hydrogen production rate under UV (Lee *et al.*, 2012).

As clearly seen in Figure 4.4a, the average production rates are not in agreement with the examples above. HT delivered slightly higher average yield than SSR under UV while its visible light performances were much higher. Surprisingly, the average UV and visible light performances of HT are also close to each other.

Researchers generally explain the effects of preparation methods through changes in particle size, surface area, crystal structure or band gap. Consequently, the dependence of those properties on the preparation methods was analyzed in more details. The surface area distributions with the most common methods are given in Figure 4.4b. (IE was also added due to its high performance). SSR clearly produced surface areas in the lower side of distribution while SG did just the opposite; however, this difference did not seem to affect PWS activity. On the other hand, HT, which provides better hydrogen production rates, exhibits no clear trend in surface area. Only IE and PC exhibit some correlation between the surface area and performance; both surface area and hydrogen production rates are high for IE while both are relatively low for PC.

To analyze the effects of preparation method on band gaps, the data were classified into three groups by this variable as 1.27-3.11 eV (700-400 nm) representing visible, 3.11-3.89 eV (400-320 nm) representing UV-A and 3.89-6.30 eV (320-290 nm) for UV-B and UV-C. The correlation between the band gap and preparation methods was undisputable for SSR and HT, (Figure 4.4c); SSR usually produces UV sensitive perovskites while HT resulted in lower band gaps explaining its higher visible light activity. The results are not conclusive for SG and PC while IE resulted in higher bandgaps, which agrees with the fact that IE catalysts are tested under UV.

Figure 4.4d presents the crystal structures obtained by different methods. Both SSR and HT produce mostly orthorhombic and cubic structures while SSR also result in octahedral perovskites. To check the effect of this difference, the average performance of octahedral and the remaining SSR catalysts (orthorhombic and cubic) were compared. Indeed, the hydrogen production rate on octahedral structure (from 13 articles) was $866.5 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ while it was $1079 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ for remaining SSR data (orthorhombic and

cubic from 38 articles). The crystal structure may also explain the results reported for PC perovskites. As Figure 4.4d shows, significant portion of PC perovskites are tetragonal although the remaining cases are cubic and orthorhombic as HT and SSR suggesting that the low average performance of PC may be due to tetragonal samples. Again, the average hydrogen production rate was found to be $256 \mu\text{mol gcat}^{-1} \text{ h}^{-1}$ for tetragonal samples while it was $1650 \mu\text{mol gcat}^{-1} \text{ h}^{-1}$ for the remaining data. However, it should be note that the tetragonal samples are from the same source even though they contain various perovskites ($\text{Bi}_2\text{Ta}_2\text{O}_9$, $\text{Bi}_2\text{K}_{0.5}\text{La}_{0.5}\text{Ta}_2\text{O}_9$, $\text{Bi}_2\text{CaTa}_2\text{O}_9$ and $\text{Bi}_2\text{BaTa}_2\text{O}_9$) (Chen *et al.*, 2014).

In summary, there are indeed correlations between the preparation methods and the perovskite properties like surface area, crystal structure and band gap. However, the direct impact of these properties on the average hydrogen production could not be established in this section but some patterns emerged in more detailed analyses in Section 4.

4.1.4. Co-Catalyst for Charge Separation

One of the most effective ways to achieve charge separation in PWS is the use of co-catalysts; 424 instances out of 540 in the database were involved the use of a co-catalyst, which are usually transition metals and their oxides (Figure 4.5) as it also stated in various reviews (Kato *et al.*, 2002; Kudo and Miseki, 2009; Ran *et al.*, 2014; Shi and Guo, 2012). Pt is overwhelmingly the most common co-catalyst (178 data points), which was used with tantalates, niobates and titanates, and reported to improve hydrogen production (Chen *et al.*, 2010). However, Pt increased the hydrogen production up to a certain loading, and then it reduced photocatalytic activity due to shielding effect (Chen *et al.*, 2014; Kanhere *et al.*, 2012a; Wang *et al.*, 2012; Zhang *et al.*, 2013); the optimal loading is depended on the perovskite. In general, Pt does not change morphological properties such as XRD pattern, surface area, pore characteristics and crystallite size (Chen *et al.*, 2014; Kanhere *et al.*, 2012; Kato and Kudo, 2002b; Li *et al.*, 2009; Wang and Wang, 2012); it can slightly alter the band gap (Li *et al.*, 2008; Li *et al.*, 2009; Puangpetch *et al.*, 2010a). The photo deposition is the most common method (124 data points) for Pt loading; this is followed by impregnation (45 data points). Arney *et al.* compared these two methods and concluded that, while impregnation increased the activity by increasing surface area, the photo deposition contributed to the activity by smaller particle size and larger number of stepped surfaces (Arney *et al.*, 2013). In another work, the hydrogen production over the loaded (by photo

deposition) and bare NaTaO_3 particles were analyzed, and it was noticed a delay in hydrogen production with the presence of Pt. It was suggested that the photogenerated electrons were used for the Pt photo deposition first and then PWS starts (Jana *et al.*, 2014b). The common result from all mentioned articles was that Pt enhanced the photocatalytic activity; it enabled charge separation, suppressed recombination, and provided active sites for hydrogen formation. However, the results in Figure 4.5 do not support these results; average performance with Pt co-catalyst is lower than the average performance of bare semiconductors; apparently there are also low performance cases with Pt co-catalyst.

Rh was also used as co-catalyst. For example, Ham *et al.* impregnated Rh on SrTiO_3 and tested for PWS under visible light (Ham *et al.*, 2016); they proposed that Rh blocked the reverse reaction. Yoshida *et al.* studied Rh/ KTiO_3 under visible light and found that different Rh loading was required for different preparation methods (Yoshida *et al.*, 2014). However, Takayama *et al.* studied several co-catalysts on $\text{KCaSrTa}_5\text{O}_{15}$, and they found that Rh was one of the less effective under UV (Takayama *et al.*, 2014). As Pt, the average performance of Rh is also lower than the average obtained without a co-catalyst (Figure 4.5).

Ru is a co-catalyst that was used with perovskite like SrTiO_3 , $\text{ZnIn}_{1.5}\text{S}_{3.2575}$, Ag:ZnInS , or $\text{KCaSrTa}_5\text{O}_{15}$ (Asai *et al.*, 2014; F. Li *et al.*, 2014a; Takayama *et al.*, 2014); it was loaded by photo deposition in all those cases. Ru had the highest visible light activity in the database (Figure 4.5); however, there are also works that showed its ineffectiveness. For example, Takayama *et al.* tested Rh, NiO, Ag, Au, and Pt as co-catalyst and Ru gave the minimum activity under UV (Takayama *et al.*, 2014).

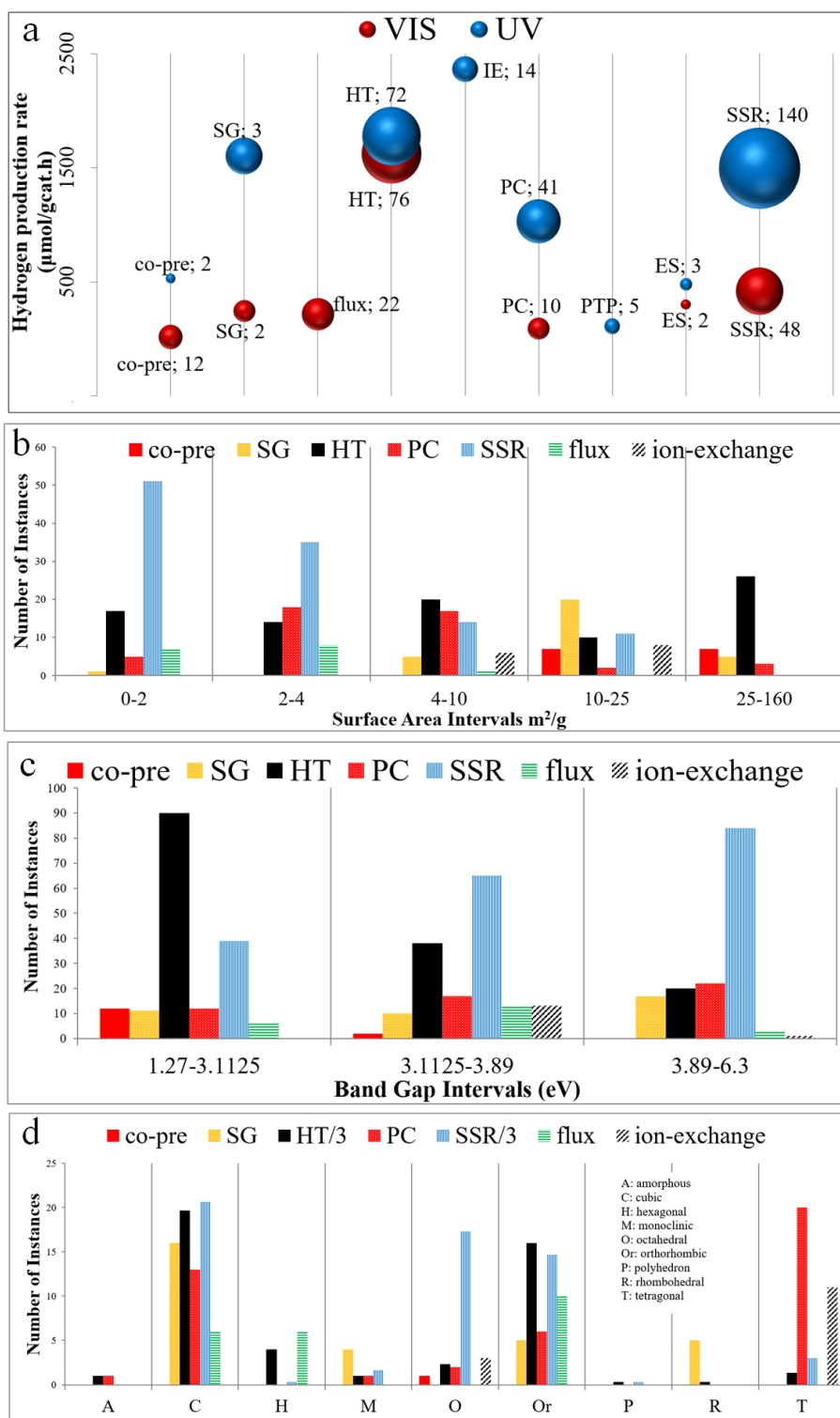


Figure 4.4. Properties of perovskites prepared by the most commonly used perovskites average a) hydrogen production rates, b) surface area, c) band gap, and d) crystal structure.

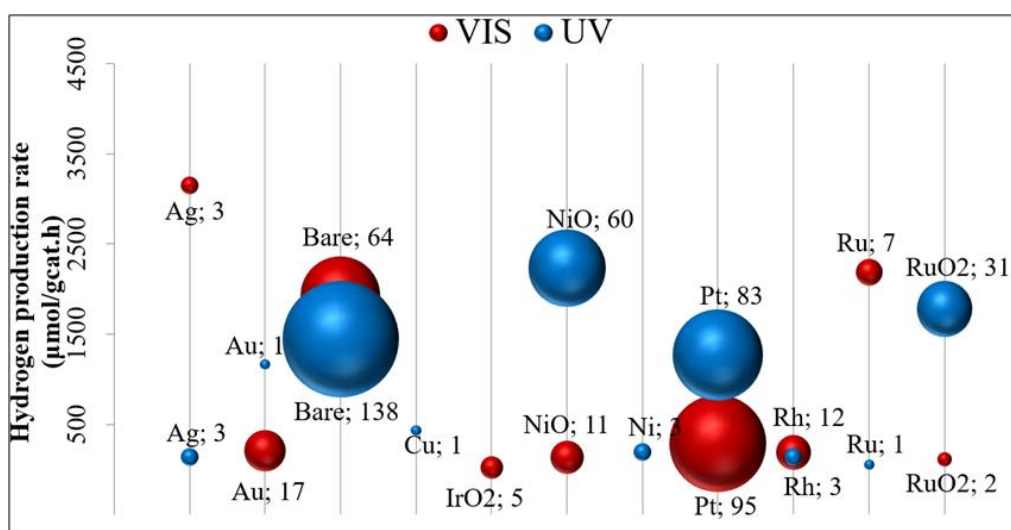


Figure 4.5. Promoters used in the database vs. number of instances.

Although there are only few cases in the database, Ag seems to be effective as co-catalyst. Shi and Guo photo deposited Ag over KTaO_3 and NaTaO_3 and found that the behavior of Ag was very similar to Pt; it did not create changes in XRD pattern, particle size or surface area but broadened the light absorption capabilities. While hydrogen production rate was directly proportional to Ag loading until some point, the excessive loading lowered the catalytic activity (Shi and Guo, 2012). There were also examples of Au, Cu, and Ni loading on several photocatalysts (Puangpetch *et al.*, 2010; Takayama *et al.*, 2014; Xu *et al.*, 2015).

NiO_x is the most preferred co-catalyst among transition metal oxides (71 cases; second after Pt). It was used with several perovskites and worked particularly well with tantalates and niobates (Chen *et al.*, 2014; Kato *et al.*, 2008; Lee *et al.*, 2012; Li *et al.*, 2009b, 2009a; Liu *et al.*, 2008). It has the highest average hydrogen production in the database (under UV). Since conduction band level of tantalates and niobates are generally higher than that of NiO_x , the photogenerated electrons in the conduction band of host perovskite can easily move to NiO_x providing charge separation (Asai *et al.*, 2014; Huang *et al.*, 2011; Ran *et al.*, 2014; Shi and Guo, 2012). Like Pt, NiO_x loading versus hydrogen production rate generally exhibited a volcanic shape; initial increase in rate is followed by a decline (Asai *et al.*, 2014; Dincer and Acar, 2014; Gómez-Solís *et al.*, 2014; Kothari *et al.*, 2008; Takayama *et al.*, 2014; Zhang *et al.*, 2009; Zhang *et al.*, 2016; Zhou *et al.*, 2009). Li *et al.* explained this as the fact that the perovskite and NiO_x formed p-type and n-type semiconductors

respectively; they had to be in optimum ratio to operate properly (Li *et al.*, 2009a). When NiO was loaded to tantalates, some Ni particles were reported to settle on the surface creating a core (Ni)-shell (NiO) structure providing selective diffusion of hydrogen and hydroxide ions. While hydrogen ion diffused to NiO and reduced, hydroxide could not; this prevents recombination of photogenerated oxygen and hydrogen (Zhou *et al.*, 2009). In most cases, there was no obvious change in surface area, crystal size and absorption band with NiO_x loading (Lee *et al.*, 2012; Li and Wu, 2015; Zhou *et al.*, 2009).

RuO_x was also used due to its lower conduction band with respect to H⁺/H energy level (Arai *et al.*, 2008); it works as electron trap to inhibit the recombination of photogenerated electron-hole pairs, facilitate oxygen production and suppress back reactions under visible light (Gómez-Solís *et al.*, 2014; Kim *et al.*, 2012; Zhang *et al.*, 2009). The activity generally increased with RuO_x loading first and reached a plateau; further increase of Ru loading gradually decreased the UV activity (Gómez-Solís *et al.*, 2014; Torres-Martínez *et al.*, 2010; Zhang *et al.*, 2009). Torres-Martínez *et al.* had similar results with those of previously mentioned articles; they synthesized La:NaTaO₃ with SG method, which provided high surface area, and loaded with ruthenium oxide nanoparticles. High specific surface area enhanced RuO₂ dispersion and increased UV activity (Torres-Martínez *et al.*, 2010). However, the surface area and crystal structure did not significantly change by RuO₂ loading itself (Li *et al.*, 2008).

4.1.5. Thermal Treatment

Calcination is a common thermal treatment that may affect morphological properties; it was applied in 351 data points (out of 540) in database. Various researchers showed that calcination temperature and time had a direct influence on particle size and surface area (Abe *et al.*, 2006a; Ham *et al.*, 2016; Hu and Teng, 2007; Iwakura *et al.*, 2010; Jeong *et al.*, 2006; Li *et al.*, 2008; Zielińska *et al.*, 2012). Generally, the particle size increases while the surface area decreases with increasing calcination temperature. The increase in particle size was attributed to the sintering (Zeng *et al.*, 2015). Abe *et al.* found that La₃TaO₇ prepared by PC had a volcanic relation between its surface area and calcination temperature (Abe *et al.*, 2006a). Crystallinity, purity, and sometimes band gap was also influenced by calcination temperature; usually, increasing calcination temperature increased the crystallinity and purity (Abe *et al.*, 2006a; Hu and Teng, 2007; Jeong *et al.*, 2006; Zielińska *et al.*, 2012). The

crystallinity affected electrochemical conductivity and the mobility of photogenerated electron-hole pairs; the higher mobility implied higher probability of those pairs reaching to active sites (Abe *et al.*, 2006a). Consequently, high calcination temperature may enhance the activity through increasing crystallinity (Iwakura *et al.*, 2010; Li *et al.*, 2008; Obee and Brown, 1995; Zeng *et al.*, 2015). However, particle size, surface area, crystallinity, and band gap act on together; hence, there is an optimum combination of calcination temperature and time for high photocatalytic activity (Iwakura *et al.*, 2010).

No clear patterns identified to link the calcination temperature to the surface area, particle size and crystal structure in database. However, the band gap seems to be correlated with calcination temperature; to see that, the band gap is divided into three classes (visible, UV-A and UV-B + UV-C) as in Section 3.1.3, and the average calcination temperature of perovskites, which have the band gap in these intervals, were calculated based on the method used during their preparation. The band gaps of perovskite prepared by HT, PC and SG were increased with increasing average calcination temperature; there was no clear trend for SSR. For example, the average calcination temperatures were respectively 993.8 K, 1123 K, and 1225.3 K for *visible*, *UV_A* and *UV_BC* class perovskites prepared by PC method. The class division and results for other methods are presented in Table 4.2.

Table 4.2. The average calcination temperature of the samples prepared using different methods*.

	Average Calcination Temperature			Number of Instances		
	Visible	UV_A	UV_BC	Visible	UV_A	UV_BC
	1.27-3.11 eV	3.11-3.89 eV	3.89-6.3 eV	1.27-3.11 eV	3.11-3.89 eV	3.89-6.3 eV
SSR	1284.6	1166.2	1297.3	43	61	84
HT	273.28	401.29	534.6	88	38	20
PC	993.83	1123	1225.3	12	17	22
SG	984.82	1008	1075.9	11	10	17

*Calcination temperature was taken as 273 K if no calcination was employed.

4.1.6. Sacrificial Agents and Additives

Some chemicals are added to the reaction medium to improve hydrogen yield by acting as electron donors or acceptors and reducing back reaction of photogenerated electrons (Chen *et al.*, 2010). There are both organic and inorganic sacrificial agents; some are photodegradable while the others are not. Photodegradable industrial pollutants (like oxalic acid, formic acid and formaldehyde) are clearly more preferred to clean the environment while producing hydrogen (Ni *et al.*, 2007); however, if this is not intended, the sacrificial agent should be non-photodegradable for a long time.

4.1.6.1. Organic sacrificial agents

Methanol is the most common sacrificial agent (60% of data in the database) (Li *et al.*, 2008; Saadetnejad and Yildirim, 2018; Shibli *et al.*, 2015; Singh *et al.*, 2008; Wu *et al.*, 2014; Zeng *et al.*, 2015); there were also examples involving ethanol, propanol, formic acid and formaldehyde. Puangpetch *et al.* compared the performances of pure water, solution of methanol, ethanol, and propanol for PWS; methanol provided highest enhancement under UV (Puangpetch *et al.*, 2010). The improvement was significant with 0-20 % methanol; then, it slowed down at higher percentages. Saadetnejad and Yildirim also verified this over Au/SrTiO₃; however, isopropyl alcohol was more effective if Al was doped on the catalyst (Saadetnejad and Yildirim, 2018).

Methanol degradation was also checked by previously mentioned articles, and found to be negligible (Li *et al.*, 2008; Li *et al.*, 2009b; Puangpetch *et al.*, 2010; Shibli *et al.*, 2015; Z. Wu *et al.*, 2014; Zeng *et al.*, 2015); this confirmed that methanol acted as hole scavenger and recombination blocker, but not hydrogen source (Puangpetch *et al.*, 2010). There were also other articles verifying this result (Arai *et al.*, 2008; Kim *et al.*, 2012; Li *et al.*, 2008; Li *et al.*, 2009b; Takata and Domen, 2009).

Methanol, propanol, formic acid, acetic acid, and formaldehyde were tested with SrTiO₃ under UV (Hu and Teng, 2010); and the formic acid had the best performance. Stepwise reactions involving intermediates occurred with all sacrificial agents except formic acid; the competition between those intermediates with valence band holes inhibited hydrogen evolution. However, the reaction occurred in single step with formic acid promoting hydrogen evolution. The same group also found that average hydrogen production increases with formic acid concentration. There are also other works in which

formaldehyde, formic acid, acetic acid and propanol were used as sacrificial agents indicating that the process could be also used to clean environment (Li, *et al.*, 2010; Juanjuan *et al.*, 2010; Zhou *et al.*, 2009; Zielińska *et al.*, 2012).

The data indicated that the alcohol (mostly methanol) concentration used with UV and visible light were quite similar; the ratio of the data with 20 % or below is 85 % in visible data while it was 83 % for UV.

4.1.6.2. Inorganic sacrificial agents

Inorganic compounds were also used as sacrificial agents, additives to aqueous methanol solution, cut-off filters for light or to adjust pH of solution. Na_2S - Na_2SO_3 pair was commonly used due to their good hole scavenger performances (Fan *et al.*, 2010; Li *et al.*, 2014a; Song *et al.*, 2015; Yu *et al.*, 2014; Zhang *et al.*, 2009). NaNO_2 was also used frequently together with the other agents to cut-off UV light (Huang *et al.*, 2011; Naik *et al.*, 2011; Zhang *et al.*, 2009; Zhao *et al.*, 2014). Xu *et al.* used NaOH to adjust the pH; the hydrogen evolution under visible light was the best in neutral solution (Lu *et al.*, 2017). Kato and Kudo utilized H_2SO_4 and NaOH to arrange pH; they obtained the maximum UV activity with NaOH (pH was 10.5) (Kato and Kudo, 2001). Lee *et al.* used KOH , NaOH , H_2SO_4 , CH_3COOH and Na_2SO_4 as sacrificial agents; the UV performance of CH_3COOH was much lower than the others (Lee *et al.*, 2013); the alkali solutions fastened the electron transfer between valence and conduction band (Liu *et al.*, 2008). There were also works in which AgNO_3 (Ham *et al.*, 2016; Jana *et al.*, 2014b; Jeong *et al.*, 2006; Kang *et al.*, 2012; Yang Liu *et al.*, 2008; Puangpetch *et al.*, 2010; Shi *et al.*, 2012), FeCl_3 (Jana *et al.*, 2014b), KNO_3 (Kato and Kudo, 2002a), KOH (Saadetnejad and Yıldırım, 2018), or NaI (Li *et al.*, 2009) were used.

As different from the organics, the inorganic additives were usually used for visible light experiments; while 37% of visible light instances in the database were obtained with inorganic additives, this number dropped to 8% for UV. Indeed, Na_2S - Na_2SO_3 , which was the most common choice, improved the average visible light activity from about 400 to 3000 $\mu\text{mol gcat}^{-1} \text{h}^{-1}$ while NaOH was effective under UV. The results for other common inorganic agents are given in Figure 4.6.

4.1.7. Light Sources

The main purpose of PWS is to utilize the sunlight; hence, the use of UV light may not be rational because it constitutes only 4% of sunlight. However, most of the well-established photocatalysts (like TiO_2) work better under UV because of their high band gap. Although these materials are usually modified for visible light harvesting, they are also studied with UV as part of the efforts to understand PWS.

About 66 % of data (324 data points) in the database were obtained under UV (mostly Hg lamp); the remaining 34 % (158 data points) were generated under visible light (mostly Xe arc lamp). The average hydrogen production rate under UV was $1532.6 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ while it was $846.6 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ with visible light. Additionally, the hydrogen production rate was lower than $100 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ in 46% of instances for visible light whereas this ratio was only 28% for UV. The ratios of data with the rate higher than $1000 \mu\text{mol gcat}^{-1} \text{h}^{-1}$ were 15 % and 24 % for visible and UV lights respectively.

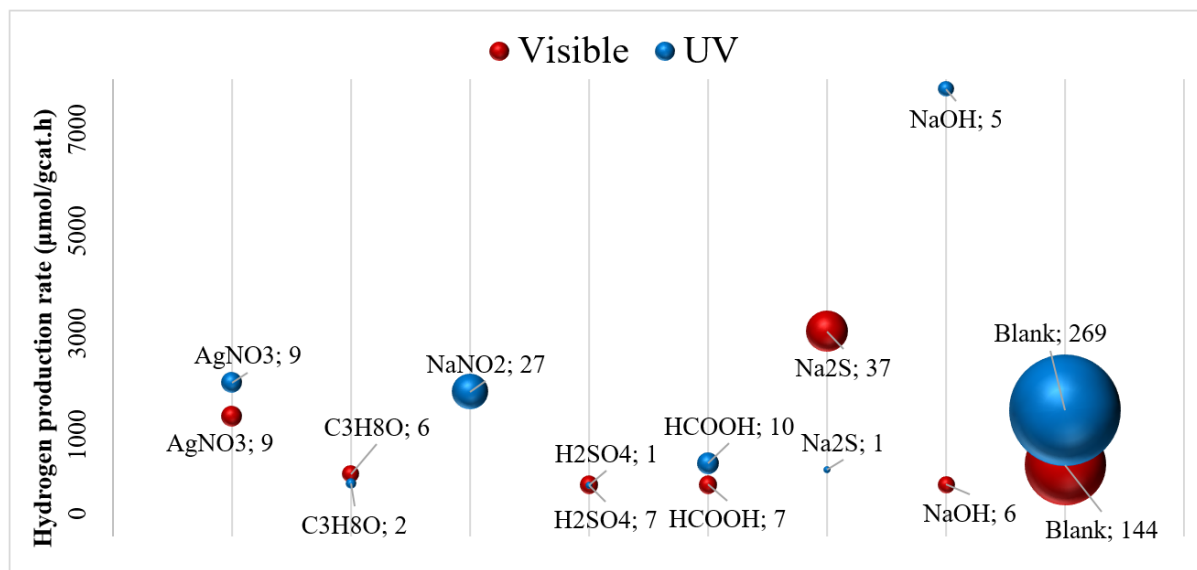


Figure 4.6. Effects of inorganic sacrificial agents on average hydrogen production rates for visible and UV light conditions.

4.1.8. Stability of Perovskites

The stability of the perovskites used in PWS was also analyzed using the data from 53 papers reporting stability tests. Although, most of the factors discussed may affect the stability, only the effects of perovskites and dopants could be analyzed using available data. Even this could be done partially due to non-standard testing procedures; in these tests, the experiment is usually shut down after certain run time and re-started for a new run for the same time interval several times; the change of activity compared to the first run is used as the indicator of stability. Since the initial activity and duration of experiment changes from paper to paper, only *the percent activity loss* could be compared. Figure 4.7 shows the results for tantalates; similar results for other perovskites are presented in Figure 4.8, 4.9, and 4.10. Clearly, these results are not sufficient to deduce definitive conclusions; however, they provide some general idea for the effects of doping. For example, as two works at the bottom of Figure 4.7 indicate, bare NaTaO₃ had significant loss in its stability after the first run, and Sr doping (third example from bottom) did not provide enhancement. However, La doping (fourth, five and sixth entries from bottom) improved the stability significantly (La and Cr co-doping even increased the performance in the second and third runs).

4.1.9. Association Rule Mining for Factor Effects on Hydrogen Production

Association rule mining was used to determine the variables that may have significant effects on hydrogen production. The dataset (UV or visible) was divided into three performance classes (*low*, *medium* and *high* performance) as explained in Section 3.1.1. The analyses were restricted to the one-factor associations to determine simple and easy to follow deductions. Table 4.3 shows the top 10 factors (based on *lift* values) leading to high hydrogen production; the list of remaining important factors is presented in Appendix A (Table A.1, A.2, A.3, and A.4).

The highest *lift* ratio (3.00) for *high* hydrogen production class for visible light belongs to ZnCdSs indicating that this is the most definitive way to have *high* hydrogen production rate. However, all 15 cases with ZnCdSs are already in *high* class; hence this factor is not suitable to appreciate the benefit of the method; instead, a more complicated case, for which the data distributed among all three classes, would be more suitable.

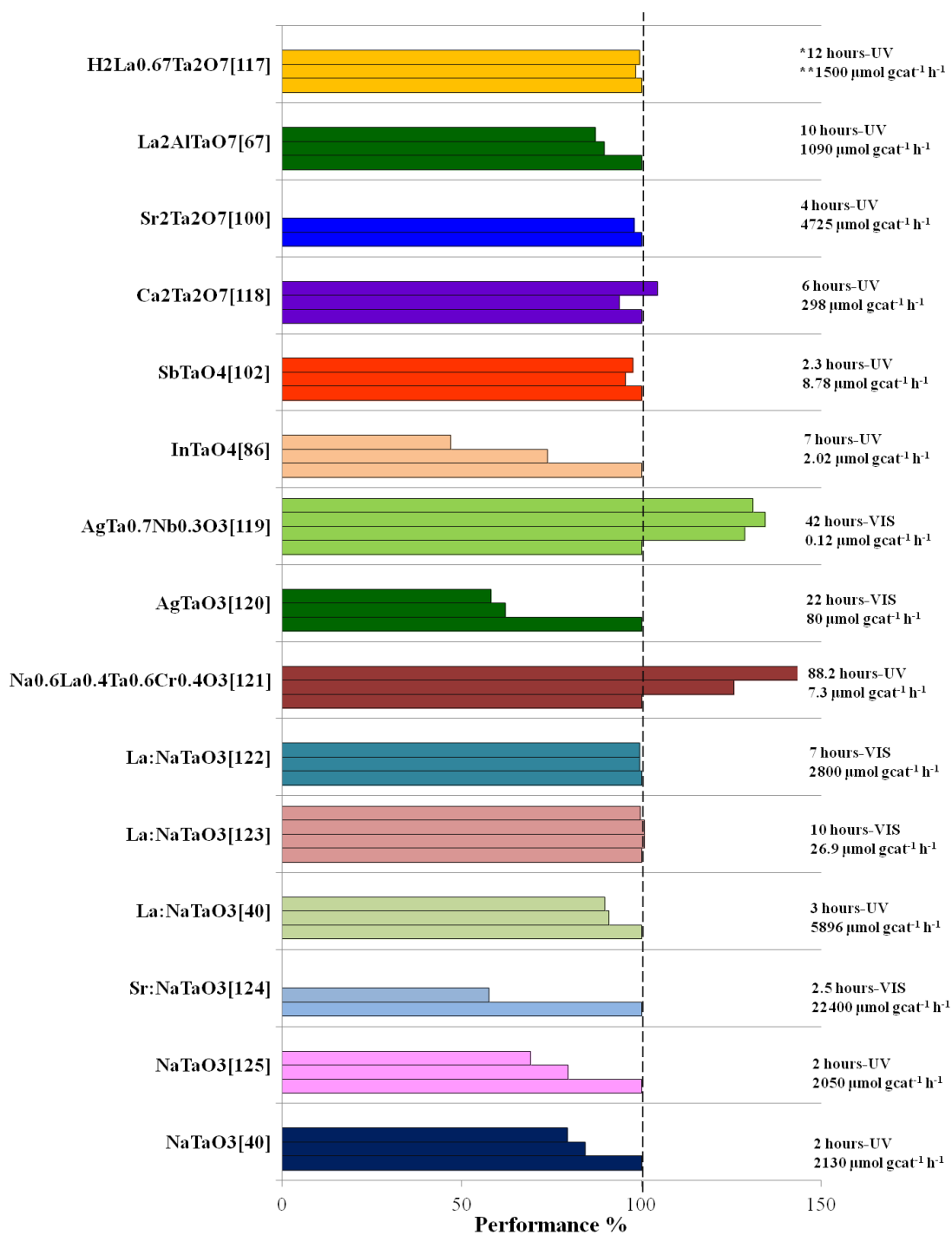


Figure 4.7. Stability of some tantalates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order; *duration of test; **rate in first run (Can and Yildirim, 2019).

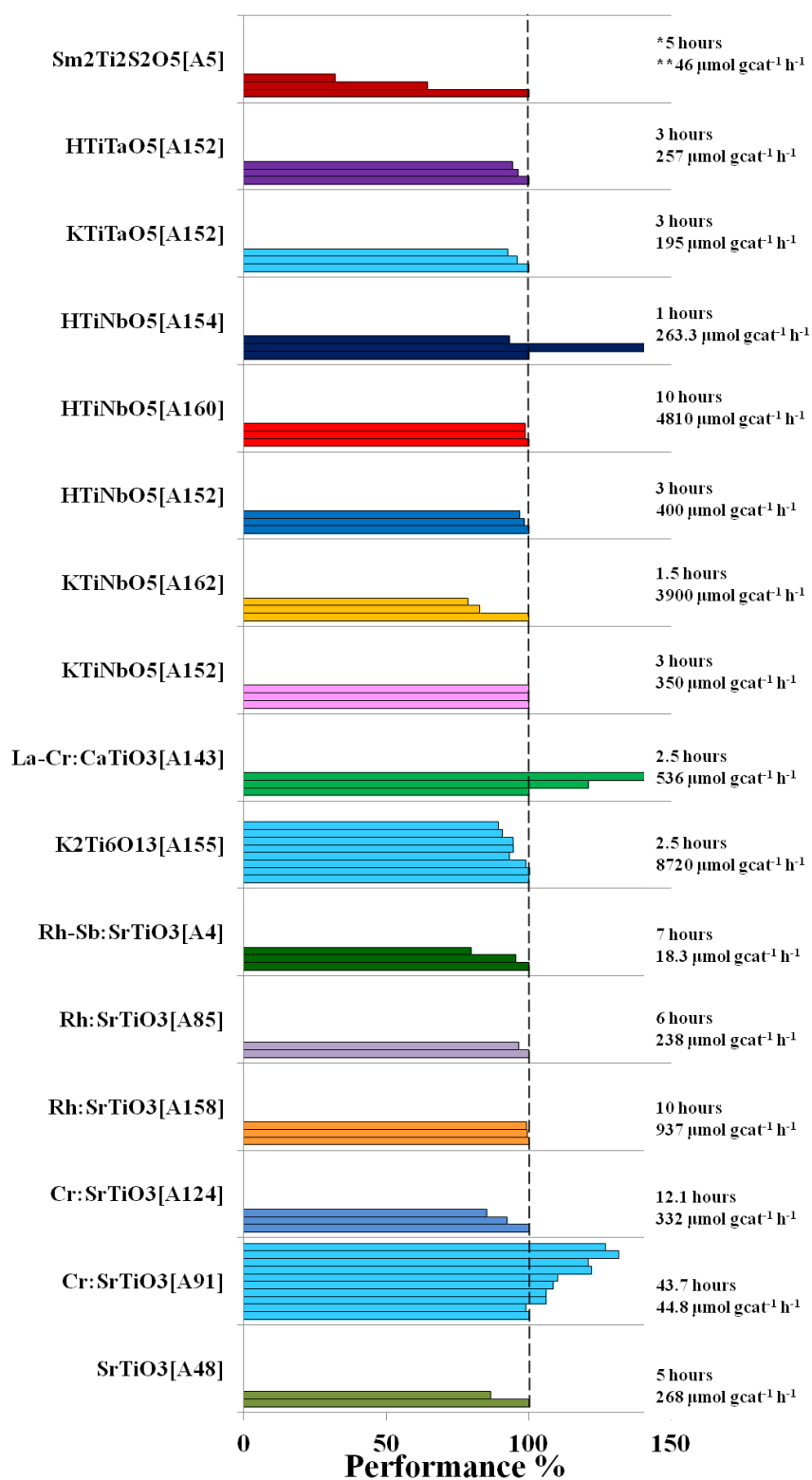


Figure 4.8. Stability of some titanates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order; * duration of test; ** rate in first run (Can and Yildirim, 2019).

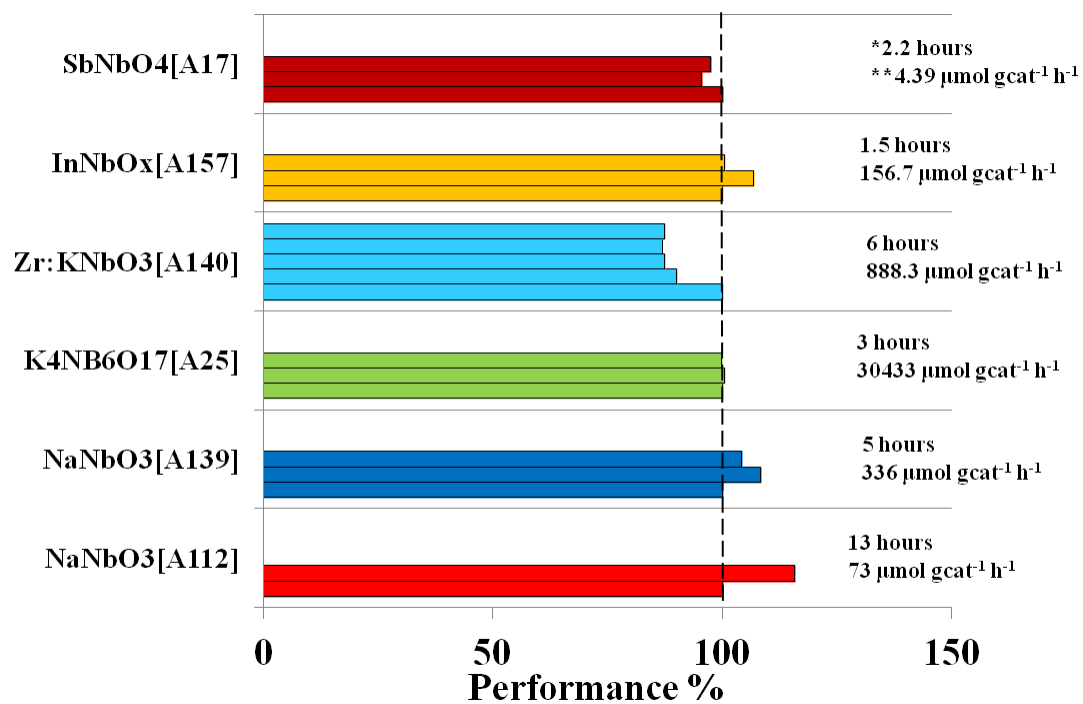


Figure 4.9. Stability of some niobates. For each perovskite, bottom bar is the first run, the others are repeats in increasing order;* duration of test;**rate in first run (Can and

Yildirim, 2019).

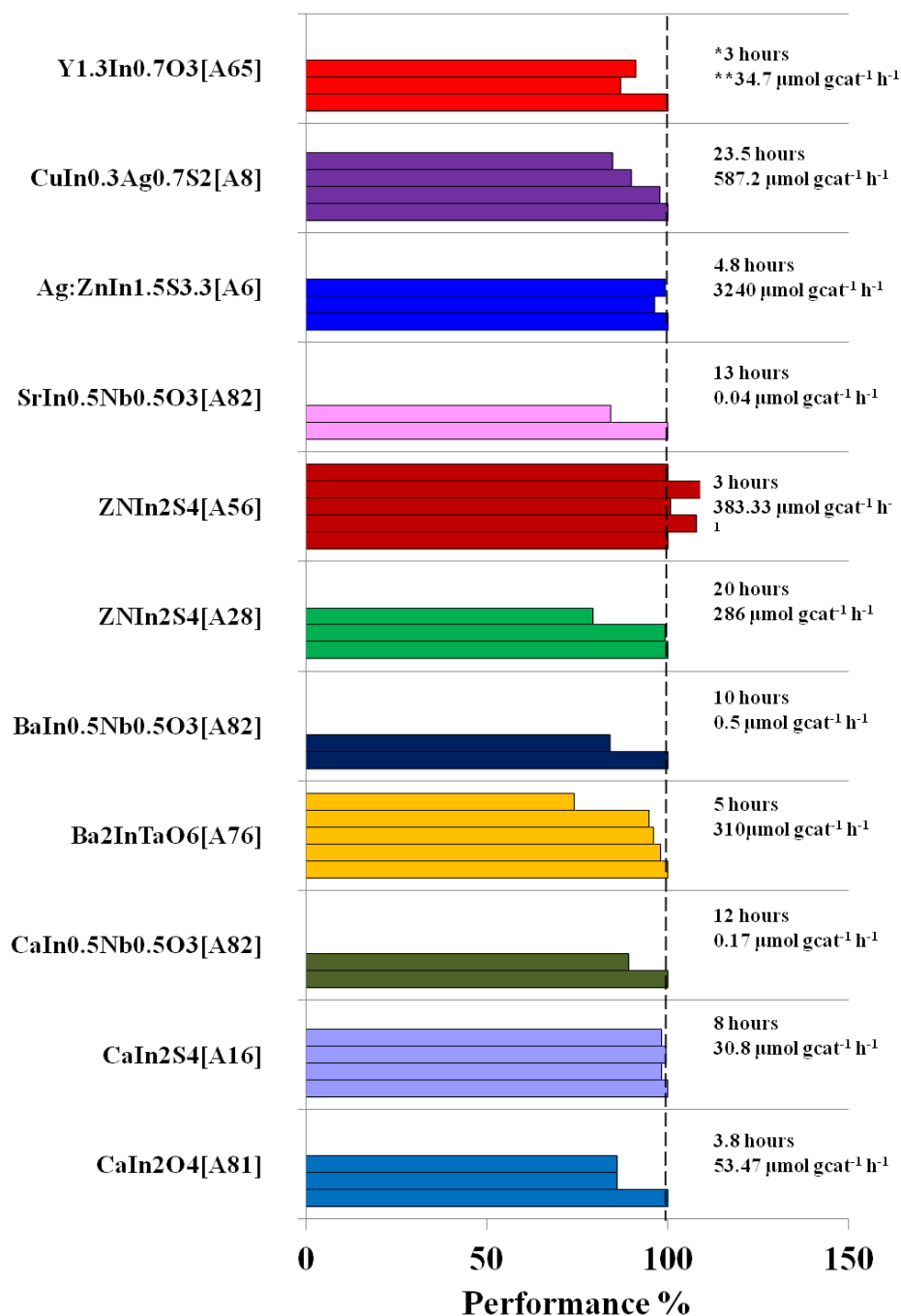


Figure 4.10. Stability of some indium-based perovskites. For each perovskite, bottom bar is the first run, the others are repeats in increasing order;* duration of test;**rate in first run (Can and Yildirim, 2019).

For example, the *support* of 0.13 (or 13 %) for Na₂S as sacrificial agent indicates that the fraction of data *used Na₂S and resulted high hydrogen production* under visible light is 28, which is 13% of total data (28/216). There are total 72 data points with *high* hydrogen production; hence the *confidence* is $28/72=0.39$; i.e. 39% of data having *high* hydrogen production rate involve Na₂S. If this ratio was also 39% in entire database, the use of Na₂S would not favor any class. However, there are 37 cases with Na₂S in entire data base; the ratio is $37/212=0.17$, which is much lower than 0.39 % for *high* class; then the lift is $0.39/0.17=2.27$. This can be also stated as *fraction of data with Na₂S in high class is 2.29 times higher than the entire database*; this is a strong indicator for the benefit of Na₂S.

As the above example shows, higher *lift* indicates higher probability for intended result while higher *support* shows higher reliability because the number of data supporting this argument is large; we considered the results with the support higher than or equal to 0.05, which corresponds more than 11 and 16 data point for visible and UV datasets, for better generalization. Similar analysis can be also done for the other factors. However, *the novel preparation methods* (second in Table 4.3) denote the collection of more than one method; hence, the rule should be interpreted as *the new approaches for catalyst preparation are somehow paid off*.

The distribution of a data among three classes for a given alternative of a factor can be also determined using association rule mining; this can provide further evidence for the analysis performed in Section 4.1.1. For example, such an analysis for NaTaOs indicates that 63% of data containing these perovskites have *high* hydrogen production under visible light while this value is 30 % and 7 % for *medium* and *low* classes respectively. This clearly shows that NaTaOs are good choices for visible data. Similar results for the common factors are presented in are Appendix (Table A.1, A.2, A.3, and A.4).

Table 4.3. Major factors leading high hydrogen production rate.

LHS <i>Production rate level</i>		RHS <i>This feature was used in PWS belonging to production rate levels in LHS</i>	Support <i>This is fraction of all data belonging to class in LHS and has feature in RHS</i>	Confidence <i>This fraction of data in LHS has feature in RHS</i>	Lift <i>(This fraction of data in LHS has feature in RHS)/ (this fraction of data in all data has feature in RHS)</i>
VISIBLE	350 ($\mu\text{mol gcat}^{-1} \text{h}^{-1}$) and above	Undoped perovskite name =ZnCdS	0.07	0.21	3.00
		Preparation method =novel	0.13	0.39	2.33
		Sacrificial agent =Na ₂ S +Na ₂ SO ₃	0.13	0.39	2.27
		Surface Area= between 50 and 200 m ² /g	0.07	0.22	2.09
		Calcination time =2 hours	0.05	0.15	1.94
		Undoped perovskite name =NaTaOs	0.08	0.24	1.89
		Promoter = no promoter	0.17	0.51	1.73
		Calcination time =no calcination	0.19	0.57	1.60
		Crystal =hexagonal	0.06	0.18	1.56
		Calcination temperature = no calcination	0.14	0.43	1.48
UV	550 ($\mu\text{mol-1 gcat}^{-1} \text{h}^{-1}$) and above	Sacrificial Agent =NaNO ₂	0.06	0.19	2.33
		Alcohol%= between 10 and 20 %	0.14	0.43	1.83
		Surface area= between 10 and 15 m ² /g	0.05	0.16	1.46
		Undoped perovskite name =NaTaOs	0.10	0.29	1.45
		Calcination temperature =between 1400 and 1500 K	0.08	0.25	1.42
		Calcination time = no calcination	0.07	0.22	1.36
		Calcination temperature = no calcination	0.07	0.22	1.36
		Crystal structure =octahedral	0.07	0.20	1.22
		Calcination time = between 7 and 10 hours	0.06	0.17	1.21
		Calcination time = 2 hours	0.07	0.20	1.20

4.1.10. Decision Tree Analysis to Determine Conditions for High Hydrogen

Production

Decision tree classification (DT) was applied to UV and visible light data to develop potential guidelines and heuristics for future works. The model for visible light is discussed below while the model for UV is presented in Appendix (Figure A.1).

The classification accuracy of the model (Figure 4.11) was found to be 82%, which is quite high: the model correctly classified 82 % of data (177 out of 216). The accuracy for *high*, *medium* and *low* classes were 76%, 82% and 88% respectively. More detailed information on accuracies is given in Appendix (Table A.5). As it will be apparent soon, the accuracy of individual branches is also important for rule deduction.

DT should be read from top to the terminal nodes through branches; the data satisfying condition stated in the node are sent to the left branch while the remaining data go to right. In Figure 4.11, DT first divided the dataset by calcination time. Then the data with low calcination time were divided again by catalyst preparation method creating the leftmost terminal nodes containing 17% of data (37 out of 216) labeled as *high* with 86% accuracy; 32 of these 37 data indeed belong to *high* class (covers 44% of entire data in *high* class). This simple and statistically reliable result should be considered as a heuristic for the future studies. Appearance of novel methods in the conditions leading high production rate was expected since a similar conclusion was also obtained with association rule mining above.

If SSR or novel methods were not used for perovskite preparation, the tree further divided the data by promoter suggesting that Au and Ag promoted catalyst would perform well. Then the divisions continued with A site elements and calcination conditions.

On the right side, the tree divided the data by A-site of perovskite followed by crystal structure and reached the rightmost terminal node containing 24 % data (52 out of 216), and labeled it as *low* performance with a remarkably high accuracy of 92 % (48 of these 52 data belongs to *low* class). This is again a very strong sign that may be used as a heuristic for future studies. Similar analysis can be performed for the branches leading to the other terminal nodes with sufficiently large number of data points and accuracy.

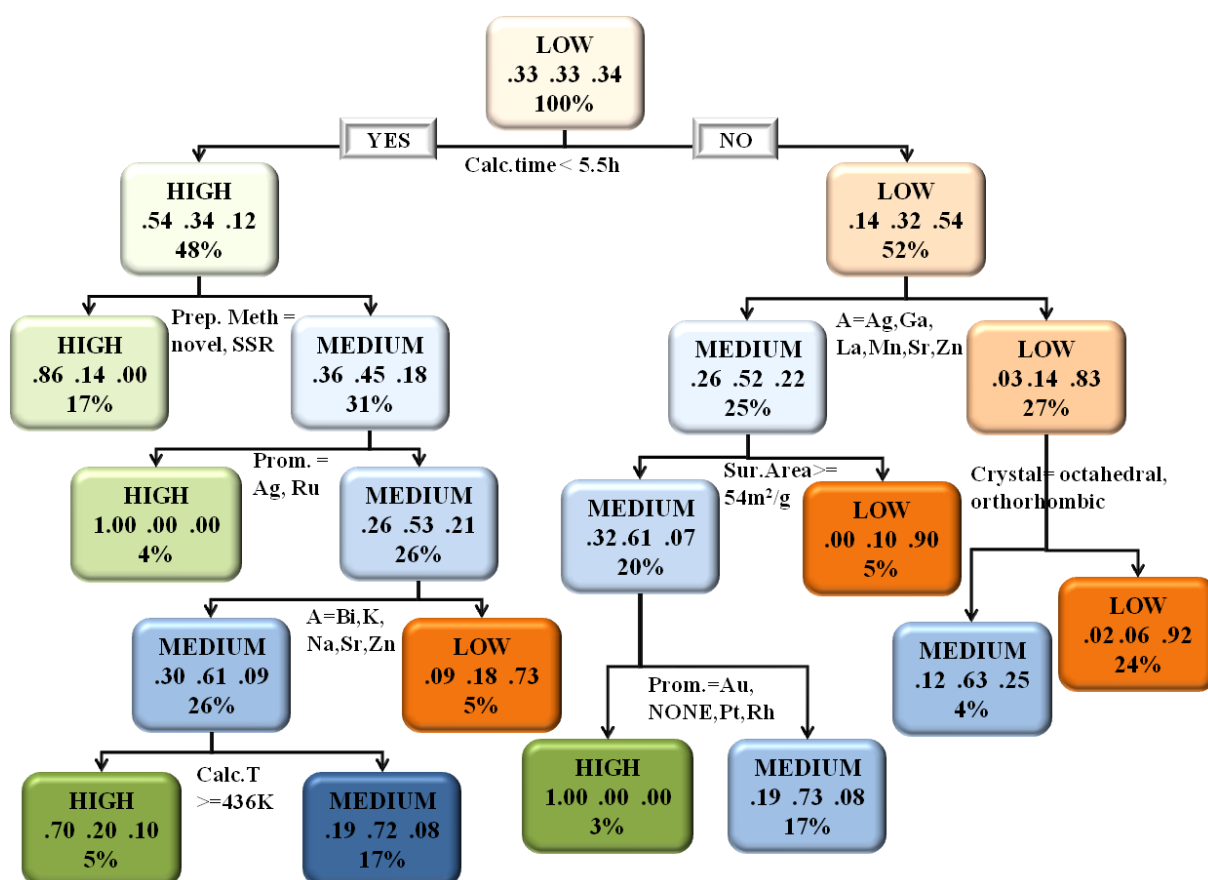


Figure 4.11. The optimal decision tree for visible light data.

4.1.11. Random Forest Analysis to Predict Hydrogen Production

Random forest (RF) regression was used to predict the hydrogen production rate, and the predicted versus real rates for training and testing for visible light are presented in Figure 4.12a and 4.12b respectively; each color in testing plot represents one set in 4-fold cross validation. The RMSE and R^2 for training were 558.1 and 0.95 for training. However, the true indicators for the predictive power are RMSE (1194) and R^2 (0.79) for testing, in which the model is forced to predict the data not seen before; these values are quite high for a dataset constructed from 151 different sources. Although such a model may not be sufficient for high accuracy predictions, it can be used for initial assessments. The model for UV data was also given in Appendix (Figure A.2).

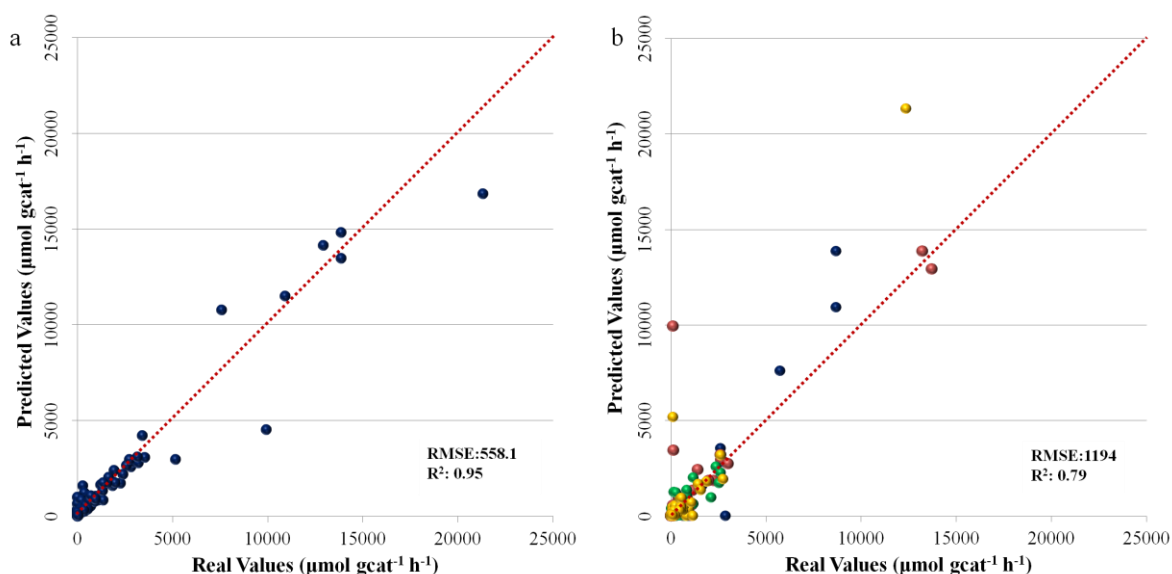


Figure 4.12. Predicted versus actual hydrogen production for random forest model for visible light data a) training b) testing.

4.1.12. Analysis And Modeling of Band Gaps

A new dataset (372 data points) was constructed from papers containing the band gaps of semiconductors as the output and the predictive variables affecting the material only; then it was analyzed using decision tree and random forest methods.

For decision tree analysis, the dataset was divided into three classes by band gaps as *visible*, *UV_A* and *UV-BC*. The confusion matrix is given in Table 4.4 while the optimum tree found is in Figure 4.13. The classification accuracy was 78 %, which is quite good, while it was 89 %, 56 % and 75% for *low*, *medium*, and *high* classes respectively. Details for accuracy levels are also presented in Table 4.4 below.

The data was first divided by B-site of perovskite, and then the left branch was divided by A-site reaching to the leftmost terminal node containing 20% of the data (48% of visible light data). Both the number of data and the purity of the node (92%) are remarkably high so that this branch can be used as a heuristic for material selection. After the first division in the left, the model used B-side doping, and generates the second leftmost terminal node with 100% purity covering 5 % of the data; the rule described in this branch can be also safely used for the future works.

Table 4.4. Confusion matrix for band gap data.

Actual Data		Predictions			Prediction
Class	# of data	VISIBLE	UV_A	UV_BC	Accuracy %
VISIBLE	143	127	9	7	88.8%
UV_A	120	30	69	21	57.5%
UV_BC	109	4	23	82	75.2%

On the right, the tree proceeds with further refinement with B-side doping (with type and then loading). The first division separated the *visible* region from the others while the second divided *UV_A* and *UV_BC*. The purity (79 %) and the reliability (30 % of data, 77 % of *UV_BC* data) of the rightmost node is also significantly high.

Random forest regression was used to predict the band gap from input variables; the predicted versus real band gap plots for the optimal model are presented in Figure 4.13b and 4.13c for training and testing respectively (colors characterizes subsets in cross validation). The RMSE and R^2 values for training were 0.24 and 0.94 respectively while the average results for the testing were 0.51 for RMSE and 0.61 for R^2 . Even though model may not be sufficient for precise predictions, it has some prediction ability, which can be further improved with the addition of new data and used for initial assessment in band gap modifications (Irfan *et al.*, 2017; Modak and Ghosh, 2016; Molak and Pilch, 2016; Takata and Domen, 2017).

Models correlating the band gap and the hydrogen production rate were also tested but the model fitness was not sufficient.

4.2. Machine Learning in Photoelectrochemical Water Splitting Literature

Similar to machine learning analysis of PWS, a comprehensive review of articles about photoelectrochemical water splitting was also done by using ML tools. In the following sections (from 4.2.1 to 4.2.4), a literature survey for the effective variables in PECWS reaction, statistical analysis of data using association rule mining (4.2.5), and the random forest prediction of band gap (4.2.6) for semiconducting materials used in PEC are presented (BU PhD student Burcu Oral also contributed to the data collection and ML analysis).

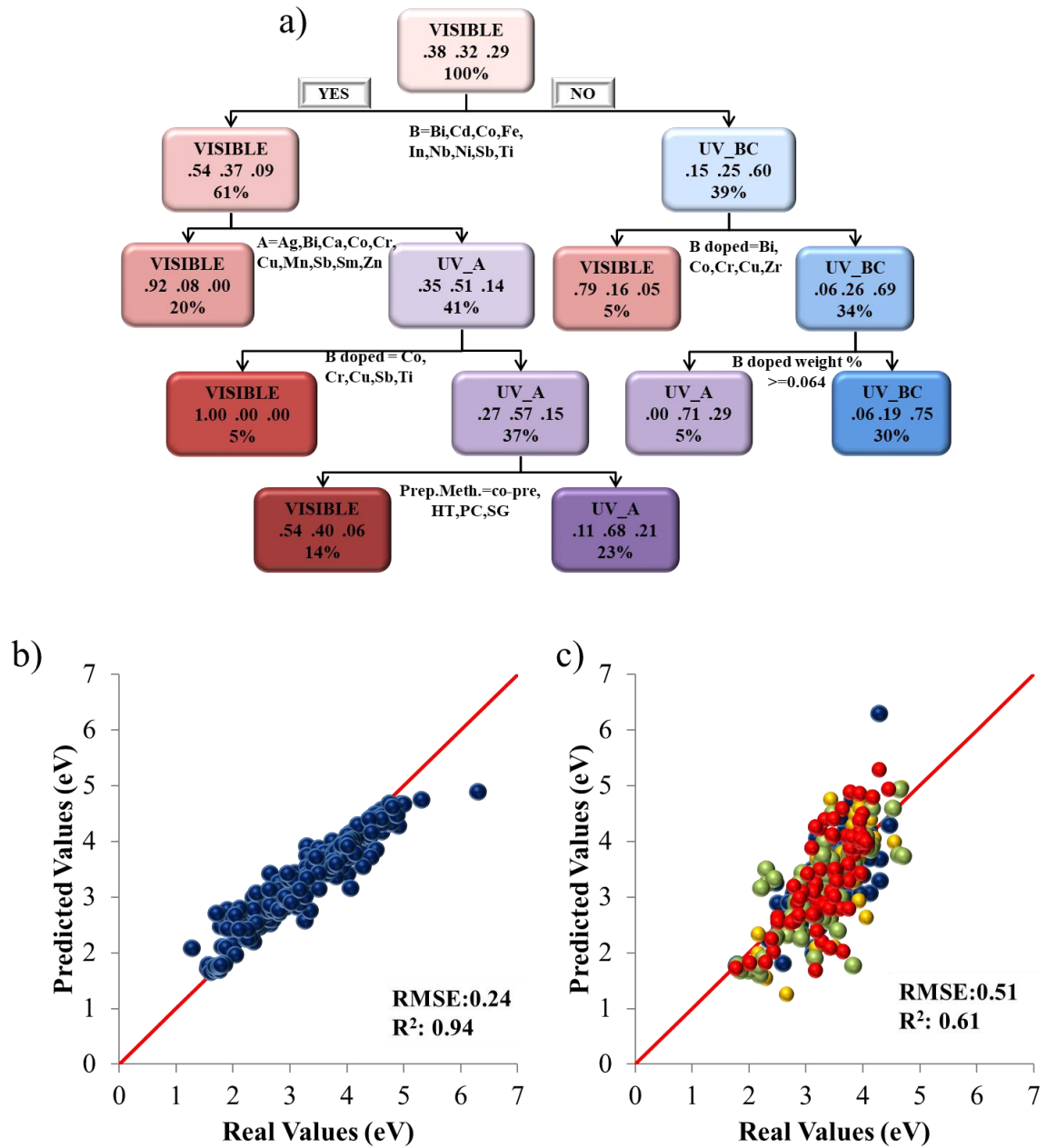


Figure 4.13. Results of the analysis of band gaps a) decision tree constructed for band gaps, b) training results of random forest analysis, c) testing results of random forest analysis.

4.2.1. Materials for PECWS

For photoelectrochemical water splitting, light active materials should be used as electrodes. Semiconductors are employed as photo electrode materials for photoelectrochemical activity. The position of the valence and the conduction bands of the semiconductor determines whether it can be used in reduction or oxidization reactions (Zhang *et al.*, 2018).

Semiconductors can be classified as n-type or p-type according to their majority carrier types; p-type semiconductors are hole abundant and generally suitable for reduction of water (hydrogen generation) whereas electron abundant n-type semiconductors are suitable for oxidation (oxygen generation) (Kment *et al.*, 2017). In our dataset, only the studies on n-type materials (photo-anodes) were used.

To enhance the water splitting performance of photo-electrodes, ion-doping, co-catalyst loading, nanoparticle/quantum dot sensitization, heterojunction structures and thermal treatments are some methods used to change the band gap, electron transport properties, crystal structure or band edge structure of the semiconductors (H. Li *et al.*, 2017).

Metal oxides are the most studied photoanodes due to their good chemical stability and low cost (Hassan *et al.*, 2012; Hu *et al.*, 2015; Wang *et al.*, 2013). Aside being the first semiconductor to split water photo electrochemically, TiO₂ has been the most studied photo anode throughout history (Das *et al.*, 2011). In our dataset 47 different articles, around 29% of data (167 instances out of 584), studied TiO₂. It is environmentally friendly, abundant, and chemically stable but its large band gap (around 3.2 eV) makes it UV active and the band edge structure of TiO₂ is not suitable for both hydrogen and oxygen evolution; therefore, it requires an external bias to perform overall water splitting (Iwase and Kudo, 2010). In order to increase the TiO₂ activity for photoelectrochemical water splitting; doping, co-catalyst loading, sensitization, thermal treatment, composite structures and layer structures have been studied. In our dataset, TiO₂ is mostly doped with carbon that lowered the band gap to 2.7 eV and a mid-gap around 1.6 eV was observed (Shaban and Khan, 2007, 2008; Xu *et al.*, 2007). Rani *et al.* doped TiO₂ by Fe, Ce and La and showed that the band gap of ~2.65 eV could be achieved by Fe doping but the highest improvement in

photocurrent density was due to La doping. Other doped ions are Cr (Momeni and Ghayeb, 2016; Tsai and Teng, 2008), Li (Sánchez-Tovar *et al.*, 2017), Ta (Altomare *et al.*, 2013), Si (Dong *et al.*, 2018), W (Jia *et al.*, 2019) and Nb (Das *et al.*, 2011), and they resulted in band gap reduction and performance enhancement. Hydrogen doping in the study of Xu *et al.* (Chen Xu *et al.*, 2013), resulted in reduction of TiO_2 and created oxygen vacancies, therefore improvement in the photocurrent density was observed. In Figure 4.14, it is seen that especially carbon and chromium doping had a significant effect on the band gap and photocurrent density values on TiO_2 compared to undoped TiO_2 .

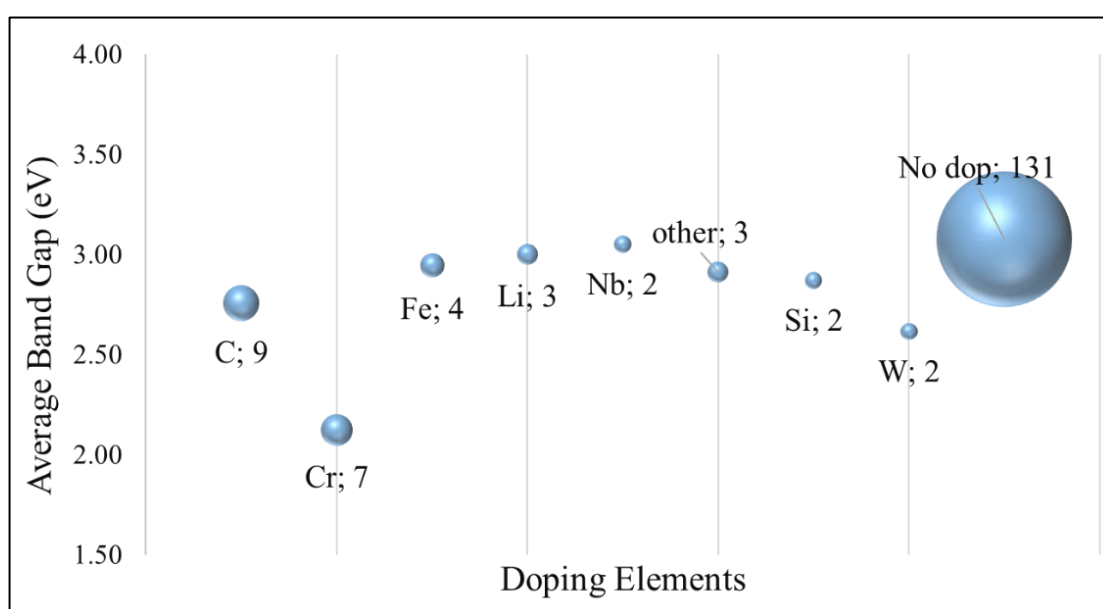


Figure 4.14. Effect of doping elements for TiO_2 on average band gap.

As the co-catalyst, Au (Jiang *et al.*, 2018; Wang, *et al.*, 2013; Subramanian *et al.*, 2017; Zhang *et al.*, 2014), Pt (Ahn *et al.*, 2018; Wang *et al.*, 2012) and Ru (Roy *et al.*, 2011) are often used with TiO_2 in order to promote charge separation. Graphene oxide (GO) (Subramanian *et al.*, 2017), CdS (Su *et al.*, 2013; Wang *et al.*, 2011) CdSe (Ji *et al.*, 2012) and g- C_3N_4 (Liu *et al.*, 2018) nanoparticle decoration of TiO_2 was also proved to be effective for band gap reduction and better charge separation. It was reported that, when TiO_2 was coupled with BiVO_4 to form a composite structure, the performance was improved by 5 folds in the visible region due to effective charge separation (Ho-Kimura *et al.*, 2014). Hybrid nanostructures were also shown to improve the performance (Lee *et al.*, 2014). ZnO layer over TiO_2 resulted in lowering the band gap (Momeni and Ghayeb, 2016) and high efficiency

due to better charge separation (Ji *et al.*, 2012; Momeni and Ghayeb, 2016; Sánchez-Tovar *et al.*, 2020). W-doped TiO₂ coupled with BiVO₄ (Jia *et al.*, 2019) showed visible light activity and high photoelectrochemical performance due to better charge separation; when BiOCl was used, the stability and the performance of TiO₂ also increased due to formation of p-n junction (Fan *et al.*, 2014). In addition, different phases of TiO₂, anatase and rutile, were coupled in a core-shell structure in literature, where almost 20 times increase in the photocurrent compared to anatase phase was observed due to enhanced charge separation (verma Atul, 2018). Thermal treatment and treatment atmosphere play important role in the performance of TiO₂ photo electrodes. Annealing at the temperatures above 400°C (Altomare *et al.*, 2013; Das *et al.*, 2011; Mohajernia *et al.*, 2017; Zhen *et al.*, 2013) showed increased conductivity and lower band gap due to rutile presence in the structure. In literature, effect of annealing on the crystal structure and the photoelectrochemical performance was investigated and it was reported that bi-phased TiO₂ nanotube arrays increased the performance (Chahrour *et al.*, 2020). Thermal treatment under argon and hydrogen (Mohajernia *et al.*, 2017) or ammonia (G. Liu *et al.*, 2012) flow increased light absorption and resulted in better PEC performance due to reduction of TiO₂. Other treatments such as Al reduction (Cui *et al.*, 2014), H₂O₂ treatment (Jiang *et al.*, 2015) and NaBH₄ treatment (Q. Kang *et al.*, 2013) showed improvement in light absorption and charge carrier properties of TiO₂ anodes.

Fe₂O₃ is the second most studied semiconductor in our dataset by 17% data (98 instances out of 584 from 33 different articles). Its band gap is suitable for visible light absorption (2.1 eV) but it has slow charge transfer and high recombination rate (Jun *et al.*, 2012; Mao, Han, and Park, 2010). In order to overcome the charge transfer problems, Al (Fu, Jiang, Zhang, *et al.*, 2014), P (Bu, Gao, Zhang, and Tian, 2019), Pt (Yong-sheng Hu *et al.*, 2008), Sn (J. Cai *et al.*, 2016; Ling, Wang, Wheeler, Zhang, and Li, 2011; Natarajan *et al.*, 2017), Ti (Fu, Jiang, Liu, *et al.*, 2014; Qian Li *et al.*, 2015; Stanescu *et al.*, 2020; Q. Sun, Cheng, Liu, and Qi, 2020; Z. Sun *et al.*, 2020; Yuan *et al.*, 2016) and Zn (W. He, Wu, Li, Chen, and Lu, 2020) ions were doped to Fe₂O₃. Heterostructure formation is another approach to increase the charge transfer properties of Fe₂O₃. SiO₂ under layer in Fe₂O₃ photo anode suppressed electron hole recombination hence improved the performance (Kang and Kang, 2015). Coating of Fe₂O₃ by Au (Carraro *et al.*, 2017), FeOOH, LaFeO₃ (Fang *et al.*, 2017), NiO (Bemana and Rashid-Nadimi, 2019), NiOOH (Qiu *et al.*, 2020) and TiO₂

(Atabaev and Atabaev, 2016; Carraro *et al.*, 2017) layers showed improved charge separation and photoelectrochemical performance. As in TiO_2 , thermal treatment temperature and atmosphere affects the performance of Fe_2O_3 photo anode (Bosso *et al.*, 2021; Ling *et al.*, 2011; Ramachandran *et al.*, 2020; Wang *et al.*, 2013).

ZnO is another widely studied semiconductor for PEC water splitting. It has large band gap like TiO_2 , but it has suitable band edge and properties for overall water splitting (Ye *et al.*, 2014). It has been studied in 22 different articles corresponding to 10% of the data (60 instances out of 584). It is mostly used in heterojunction structures rather than being doped or loaded by co-catalyst. However, doping by carbon (Kochuveedu *et al.*, 2013), cobalt (Khan *et al.*, 2019) and lithium (Lee *et al.*, 2016) ions reduced its band gap to visible light region as well, and decreased rate of combination due to separation of charges. Sensitization by carbon quantum dots (Xiao *et al.*, 2017), Au nanoparticles (Zhang *et al.*, 2018) and AgSbS_2 (Han *et al.*, 2015) also increased the PEC performance due to change in band gap and charge separation properties. Thermal treatment is also important for ZnO performance (Sharma *et al.*, 2014); Liu *et al.*, investigated the effect of annealing between 350-550 C, and showed that annealing at 450°C could increase the efficiency up to 7.5 times compared to unannealed ZnO (Liu *et al.*, 2019). In addition to thermal treatment, the treatment environment is also important for developing high performance ZnO photo electrodes; it is reported that the structure of ZnO changes with addition of nitrogen gas to argon environment during synthesis. Heterojunction by CuO and Ag, Au metals (Kwon *et al.*, 2018), core-shell structure with TiO_2 (Ji *et al.*, 2012), ZnFe_2O_4 (H. Jiang *et al.*, 2020), or Ga-ZnO (Xiao *et al.*, 2017) were employed to increase charge separation and visible light absorption of ZnO photo anodes.

BiVO_4 is another semiconductor of interest in PEC water splitting applications. Our dataset contains 21 articles and 58 instances (10% of total data) of BiVO_4 photoelectrode. Like ZnO, it is mostly used in heterostructure forms. It is used with TiO_2 (Ho-Kimura *et al.*, 2014; Jia *et al.*, 2019), WO_3 (Saito *et al.*, 2012) and reduced graphene oxide (N. A. Mohamed *et al.*, 2020; Subramanyam, Khan, Neeraja Sinha, Suryakala, and Subrahmanyam, 2020) to overcome charge transfer problems and increase PEC performance. It is mostly doped with higher valence metal ions like Mo (Berglund *et al.*, 2012; Luo *et al.*, 2013; Park *et al.*, 2014; Pilli *et al.*, 2011) and W (Berglund *et al.*, 2012; W. Luo *et al.*, 2013) to improve the charge separation. Luo *et al.* compared doping to Bi^{3+} and V^{5+} sites of BiVO_4 and reported that

doping to V^{5+} site improved the performance better than doping to Bi^{3+} site (Luo *et al.*, 2013).

WO_3 is another attractive candidate for PEC water splitting due to its stability and low band gap (~ 2.6 eV) but it suffers from fast recombination of charges (Kalanur *et al.*, 2013; Marsen *et al.*, 2007). In our dataset 20 different articles, with 55 instances (9% of the data), reported WO_3 as photo anode. It is reported that, in high concentration of nitrogen doping, PEC performance of WO_3 increased (Marsen *et al.*, 2007). Co-catalyst loading of PtO_x and RuO_2 were shown to offer more reduction and oxidation sites for ions and hence increase the PEC performance (Ma *et al.*, 2012). Heterojunction composite formation was used in order to enhance PEC performance of WO_3 by lowering the need for external bias and increasing the charge transfer properties. Other than doping, co-catalyst loading (Ahn *et al.*, 2018; Han *et al.*, 2015); treatment of WO_3 with hydrogen (Wang *et al.*, 2012) and alcohol vapour (Hsiao *et al.*, 2011) increased photocurrent density due to formation of oxygen vacancies and increased charge transfer rate. WO_3 is mostly coupled with $BiVO_4$ since $BiVO_4$ can absorb more portion of the solar spectrum and photo generated charge transfer from $BiVO_4$ to WO_3 increased the PEC efficiency (Hong *et al.*, 2011; Saito *et al.*, 2012; Zhang *et al.*, 2012). WO_3 was also coupled with reduced graphene oxide (Lin *et al.*, 2013), g- C_3N_4 (Li *et al.*, 2017), TiO_2 (Liu *et al.*, 2013) and $BiOI$ (Shi *et al.*, 2018). Optimization of thermal treatment temperature was proved to be effective in increasing the photocurrent density (Ding and Kim, 2016; Kalanur *et al.*, 2013; Marsen *et al.*, 2007; Qin *et al.*, 2012).

Besides these photo electrode materials, $\sim 25\%$ of data is composed of various semiconductors such as oxides (Qamar *et al.*, 2015; Weng *et al.*, 2014), nitride and oxynitride compounds (Alotaibi *et al.*, 2013), perovskites (Ahn *et al.*, 2018; Pinheiro *et al.*, 2014) and chalcogenides (Bosso *et al.*, 2021; Li *et al.*, 2009; Lu, *et al.*, 2013).

In Figure 4.15, the effect of top layer for some of the frequently used bottom materials is investigated. The average current densities values are obtained from 1 V vs NHE external bias condition since it is the most frequently used bias in our dataset. For Fe_2O_3 , it is seen that, using TiO_2 as a second layer showed improved photocurrent densities compared to other types of materials. Similarly, rGO seems to increase the performance when it is used with $BiVO_4$, but the data set contains only one case, which is not sufficient to make reliable conclusions. Instead of depositing another layer, some researchers employed co-catalyst for

performance enhancement. Semiconductor and co-catalyst pairing with their average photocurrent density comparison is given in Figure 4.16. It is seen that the co-catalysts do not seem to be effective for TiO_2 and WO_3 but provide some increase in performance of FeO and ZnO . RuO_2 and some other co-catalyst types are effective for ZnO whereas Au increased the photocurrent density for Fe_2O_3 .

4.2.2. Methods For Semiconductor Synthesis and Electrode Fabrication

The preparation method affects physical properties of the semiconductors such as surface area, band gap, crystal structure etc. The effect of synthesis methods on the band gap is analysed by ARM in section 3.2. In our dataset, hydrothermal method is most used semiconductor preparation method with 29% data (168 instance out of 584) followed by anodization 13% (77 instance), electrodeposition 13% (76 instance), solvothermal method 7%, sol-gel synthesis 6%, and many other methods with less than 5% data. By hydrothermal synthesis, it is possible to obtain nanostructures of TiO_2 (Fei *et al.*, 2010; X. Huang *et al.*, 2020) and porous WO_3 (Hou *et al.*, 2016). Qin *et al.*, synthesized WO_3 by hydrothermal synthesis, and by varying reaction time, they obtained different length and thickness WO_3 -flakes; they observed that thicker flakes showed stronger visible light absorption (Qin *et al.*, 2012). Optimization of hydrothermal time is important to synthesize the best performing semiconductor; some novel materials such as $\text{Cu}_2\text{In}_2\text{ZnSn}_3$ (Hong *et al.*, 2016), CaBi_2O_4 (Wang *et al.*, 2017) and NaInSn_2 (Chen *et al.*, 2018) were synthesized this way and showed remarkable PEC performance.

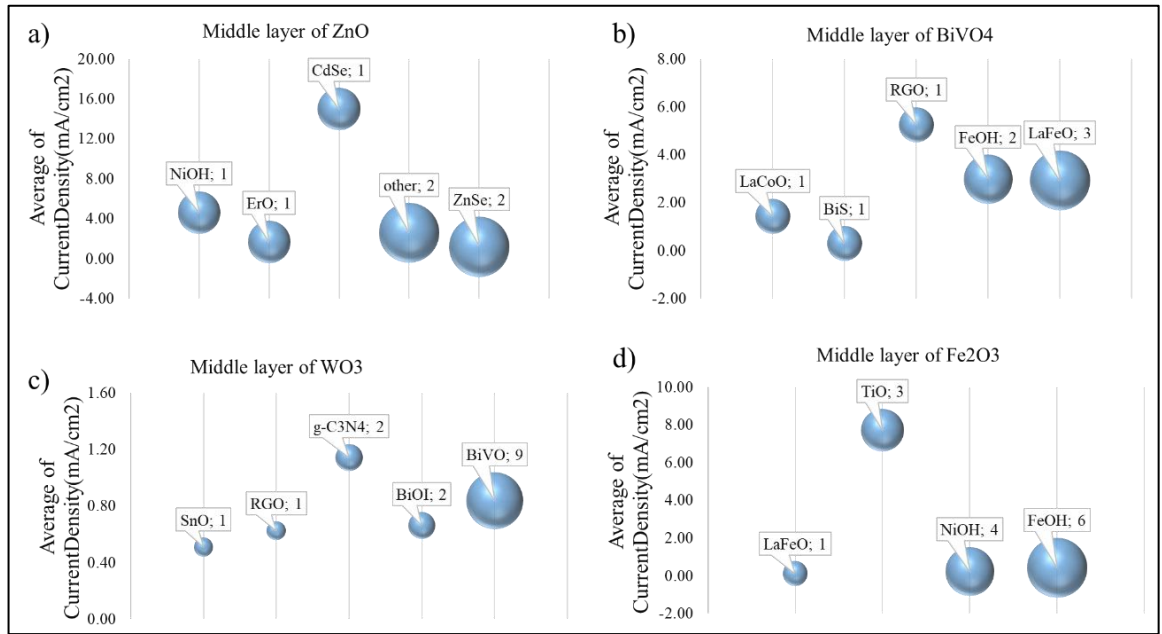


Figure 4.15. The effect of top layer for some of the frequently used materials in middle layer a) ZnO, b) BiVO₄, c) WO₃, and d) Fe₂O₃.

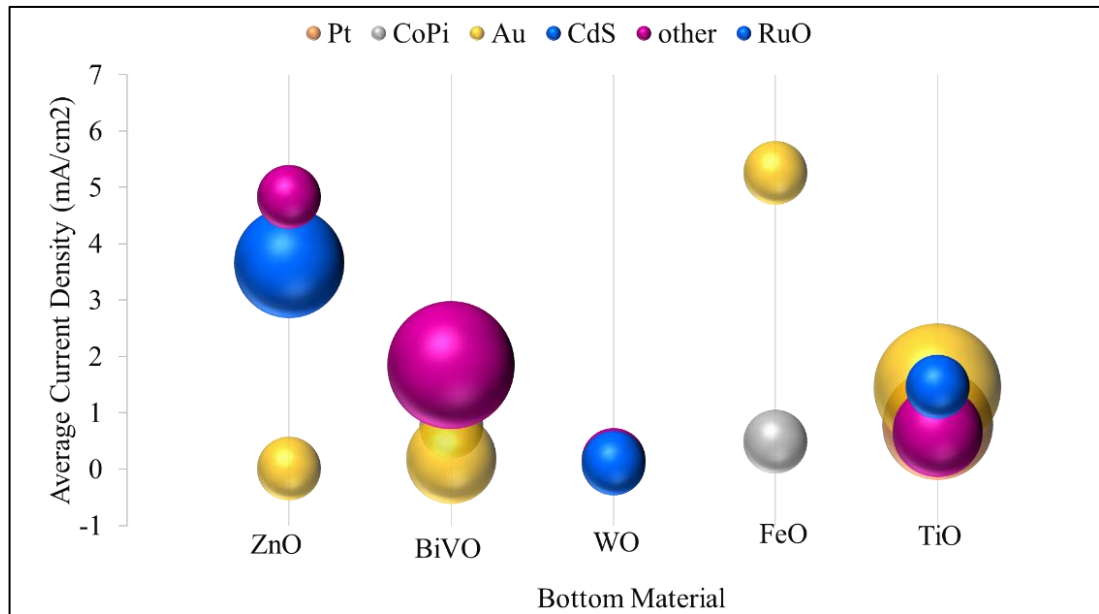


Figure 4.16. Effect of co-catalyst using on the photoelectrochemical performance of frequently used bottom materials.

Anodization is used to obtain metal oxide structures from metal foils in our dataset, and it is used to prepare Fe₂O₃, TiO₂, Ta₂O₅ (Zhang *et al.*, 2020) and NbO (Hu *et al.*, 2016;

Shaheen *et al.*, 2016) semiconductors. Self-assembled titania nanotube arrays (Altomare *et al.*, 2013; Mahajan *et al.*, 2008) and nanotube photonic crystal TiO_2 (Liu *et al.*, 2017) synthesized by anodization showed enhanced PECWS performance and good stability. Lucas-Ganados *et al.* showed that other than anodization time, rotation speed of the electrode was important for the thickness and structure of Fe_2O_3 (Lucas-Granados *et al.*, 2017).

Electrodeposition/electrochemical deposition is similar to anodization, but this time precursor solution is deposited on a substrate by applying a potential difference. Chou *et al.* compared the performance of Fe_2O_3 prepared by electrodeposition and Fe-oxidation. They observed that electrodeposition resulted in hematite structure with good light absorption but low photocurrent density due to slow charge transfer; however, the oxidation resulted in better photocurrent density while its light absorption capacity was low because of the magnetite structure, (Chou *et al.*, 2013). Shi *et al.* employed a modified synthesis approach using electrodeposition in order to increase the PEC performance; they electrodeposited Fe_2O_3 precursor on polystyrene colloids, and by this way, they could tune thickness by electrodeposition time and geometry (Shi *et al.*, 2013). Zhang *et al.* reported that, it was also possible to control the amount of nanoparticle decoration or doping percent by adjusting the electrodeposition time (Zhang *et al.*, 2018).

In the synthesis of semiconductors, the ratio of precursors and additional agents are important in the result. For example, it was possible to obtain and tune the structure of biphasic Cd_2SnO_4 by addition of urea in different amounts; the junction of crystal structures improved PEC performance by reducing surface recombination of electron and holes (Deshpande *et al.*, 2014). Similarly, changing the acid type, optimizing the concentration and reaction time resulted in performance improvement of BiVO_4 photo anode (Iwase and Kudo, 2010).

In Figure 4.17, the effect of synthesis method on the crystal structure and band gap energies are given for the most frequent bottom materials used in the dataset. It is seen that BiVO_4 was used mostly in the monoclinic form and only sol-gel method resulted in different crystal structure, tetragonal. The lowest band gap value was observed for tetragonal structure of BiVO_4 synthesized by sol-gel method. Overall, BiVO_4 had the average band gap value of 2.42 eV, but sol-gel method resulted in average of 2.19 eV. Fe_2O_3 was used in hexagonal and rhombohedral crystal structures in the dataset and lowest band gap value (1.85 eV) was

achieved with hexagonal Fe_2O_3 prepared by hydrothermal method while the overall band gap for Fe_2O_3 was 2.19 eV. For TiO_2 , mostly anatase phase was utilized (around 60%). It is seen that rutile phase had generally lower band gap than anatase. For WO_3 , lowest band gaps were observed in monoclinic crystal structure prepared by sol-gel synthesis and chemical solution route. The average band gap for all WO_3 bottoms in the dataset was 2.57 eV. ZnO was mostly in hexagonal structure in the dataset; Its average band gap value of ZnO is 2.94 eV and electrodeposition method gave the lowest average band gap of 2.63 eV with rhombohedral crystal structure. Figure 4.19 gives more general perspective to the discussion above. The colours of bubble represent a certain class (range) of bandgap while their sizes show the fraction of data corresponds to that class. For instance, when chemical deposition is used in the synthesis of bottom layer, it resulted in medium class band gap (between 2.4-3 eV). However, the electrodeposition gave mostly lower band gap. Some synthesis methods gave only one class of band gap because they were utilized for specific materials. For example, the etching is only used for Si material, and it only gives low class band gap because Si has the band gap around 1.5 eV.

In addition to the synthesis of semiconductors, the methods used in the electrode fabrication are also important. It is possible to fabricate electrodes in-situ; the conductive substrate will be coated as the semiconductor is synthesized. In our dataset 70% of the experiments (412 instance) used in-situ fabrication using methods like hydrothermal, sol-gel, chemical deposition, vapour deposition and electrodeposition. When in-situ is not preferred, spin coating (Hsiao *et al.*, 2011; Kochuveedu *et al.*, 2013; Liu *et al.*, 2013; Luo *et al.*, 2013; Pilli *et al.*, 2012, 2011; Saito *et al.*, 2012) (50 instance), dip coating (Lin *et al.*, 2013; Pinheiro *et al.*, 2014; Y. Wang *et al.*, 2017; J. Wei *et al.*, 2018) (15 instance), doctor blade (Deshpande *et al.*, 2014; Kelkar *et al.*, 2012; Sfaelou *et al.*, 2016) (14 instance) and electrodeposition (Fang *et al.*, 2015; Guo *et al.*, 2017; L. Zhang *et al.*, 2014) (14 instance) are used to prepare electrodes from previously synthesized semiconductors. In spin-coating and dip coating, the thickness of the coating layer can be controlled via cycle number whereas the deposition time plays an important role in methods like electrodeposition as explained above.

4.2.3. Light Properties

Photo electrodes absorb sun light and use the photo-excited electrons for redox reactions in water splitting (Bosserez *et al.*, 2015). The spectrum of sunlight mainly consists of visible light (Cho *et al.*, 2011), hence it is important to utilize visible light portion of spectrum for high efficiency. In the dataset, different wavelengths and intensities of light are used. Increasing the light intensity resulted in photocurrent increase since more photons were involved in the excitement of electrons (Hu *et al.*, 2008; Shen *et al.*, 2013; Zhang *et al.*, 2016). Cut-off filters were used to investigate the performance of photo electrodes under different wavelengths, and the improvements in the visible light region were reported using such filters (Kitano *et al.*, 2007; Reddy *et al.*, 2019; Y. Sun *et al.*, 2008; Zhao *et al.*, 2007).

The widely used type of light is Xe lamp with 487 instances (82%) in the dataset. The intensity varies between 0.06 mW/cm^2 and 1000 mW/cm^2 , and mostly 100 mW/cm^2 is used (404 instance, 69%). The use of AM 1.5G filter, which is a standard used to simulate solar light on the earth's surface (Gueymard *et al.*, 2002), was reported in 376 instances (64%).

4.2.4. Electrolyte Properties

For a good photoelectrochemical performance, the choice of electrolyte is important since reaction on the catalyst surface is affected by the concentration, pH and type of ions (Ding *et al.*, 2017). Electrolytes are the media for ion transfer that can also prevent electron-hole recombination (Abe, 2010). Ion types in the electrolyte is important for the water splitting performance of electrodes. For TiO_2 , the photoelectrochemical performance was highest for Li^+ containing alkaline electrolytes compared to the that of Na^+ and K^+ containing alternatives (C. Ding *et al.*, 2015); however, when Li^+ was used instead of Na^+ , the injection of Li^+ ions to mesostructure was easier and resulted in electrochromism on WO_3 electrodes (Sfaelou *et al.*, 2016). Oxidation/reduction potentials of ions are important to achieve effective water splitting. For example, anions like iodine and bromide are oxidized easier than water, therefore they should not be used in electrolytes to have water oxidation reaction; however, the chlorine ions can be used because they have higher oxidation potential than water (Crawford *et al.*, 2009). For ZnO , it was reported that using Na_2SO_3 electrolyte instead of NaSO_4 resulted in 6 times higher water splitting performance due to consumption of holes by SO_3^{2-} ions, which promotes electron-hole separation (Guo *et al.*, 2018). In Figure 4.18, it is clearly seen that, in order to maximize the performance of the photo electrodes,

different electrolytes must be used depending to the electrolyte material type; WO_3 gave better results in acidic conditions whereas for TiO_2 similar results were obtained in basic solutions. As well as concentration, pH of the electrolyte is important since the gas evolution reactions and the rate determining steps depend on pH (Jozwiak *et al.*, 2020). Yourey *et al.*, used electrolytes with different pH values (7, 3, and 1) for CuWO_4 , and observed that, as the pH of buffered solution decreases, current decreases (Yourey and Bartlett, 2011).

Additives can also be used in electrolytes to promote photoelectrochemical performance of the electrodes. These additives work as electron donors (hole scavenger) or acceptors and prevent electron-hole recombination. For example, methanol is widely used as electron donor in both photocatalytic and photoelectrochemical water splitting (Pinheiro *et al.*, 2014). Yourey and Bartlett used methanol as additive since it is a strong electron donor, and doubled the photocurrent achieved by CuWO_4 photoelectrode (Yourey and Bartlett, 2011). Similarly, ethanol is employed as additive for photoelectrochemical system with WO_3 photoelectrode and photocurrent was increased (Sfaelou *et al.*, 2016). H_2O_2 is another type of additive that can be used to increase the photocurrent (Bohra and Smith, 2015; Bu *et al.*, 2019; Hill and Choi, 2013; L. Zhang *et al.*, 2014), but not preferred due to its corrosive nature (Kuo and Klabunde, 2012). Other additives used in the dataset are ethylene glycol (Momeni *et al.*, 2020; Chen Xu *et al.*, 2013) and EDTA (J. Cai *et al.*, 2016). Na_2S and Na_2SO_3 additives were mostly used as hole scavenger to promote sulphide oxidation on the working electrode (Berglund *et al.*, 2012; Cho *et al.*, 2011; Hong *et al.*, 2016; Park *et al.*, 2014; Zhang *et al.*, 2014; Zhang *et al.*, 2016).

In our dataset, NaOH is the most dominant electrolyte with 169 instance (29%), followed by Na_2SO_4 with 147 instance (25%) and then KOH with 111 instance (19%); other electrolyte types are less than 5% of the overall dataset. For additives, 481 instances did not use any additive in the electrolyte and 38 instances out of 102 (37%) used Na_2SO_3 as additional electrolyte/hole scavenger.

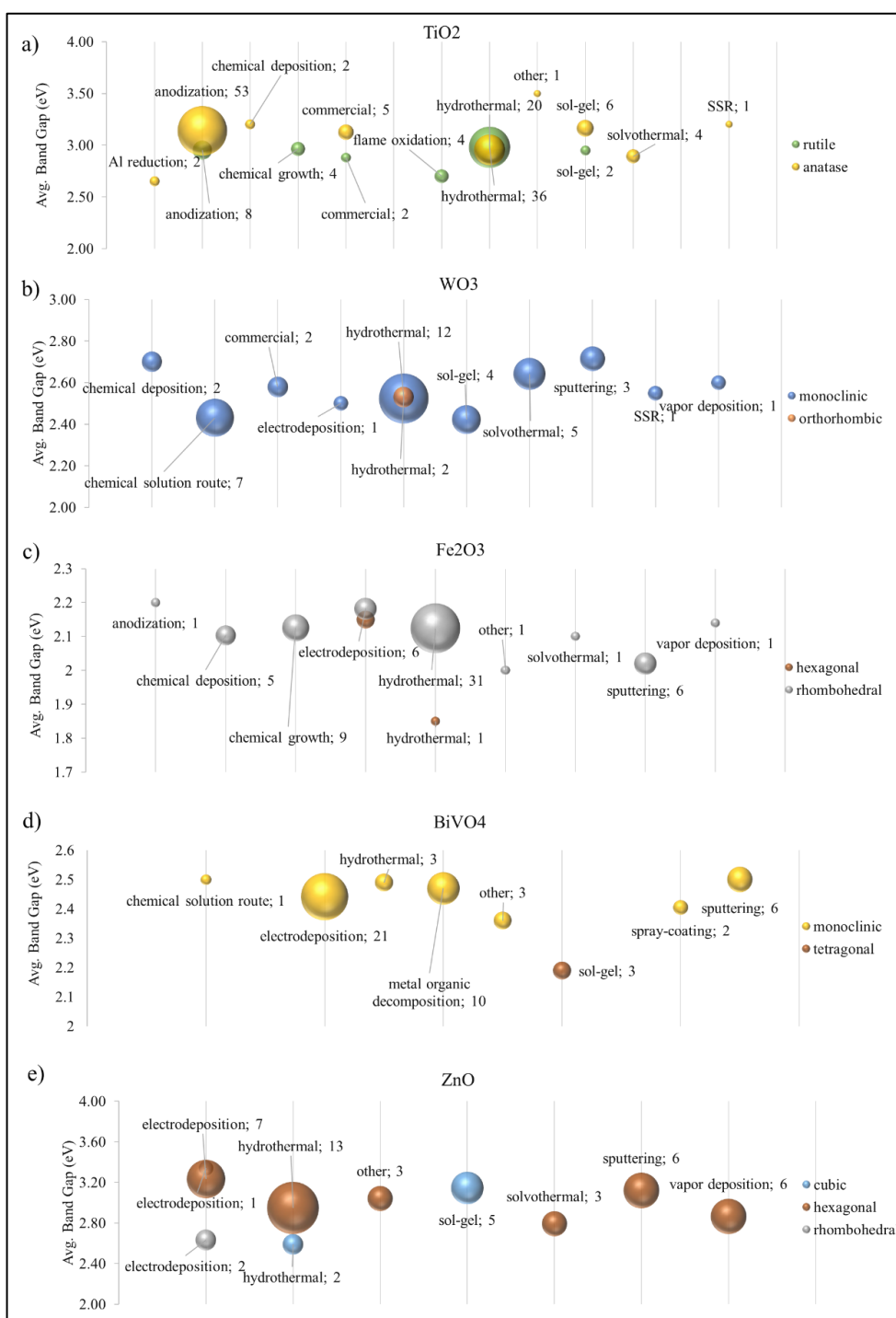


Figure 4.17. The effect of synthesis method on the crystal structure and band gap are given for the most frequent bottom materials, a) TiO₂, b) WO₃, c) Fe₂O₃, d) BiVO₄, and e) ZnO.

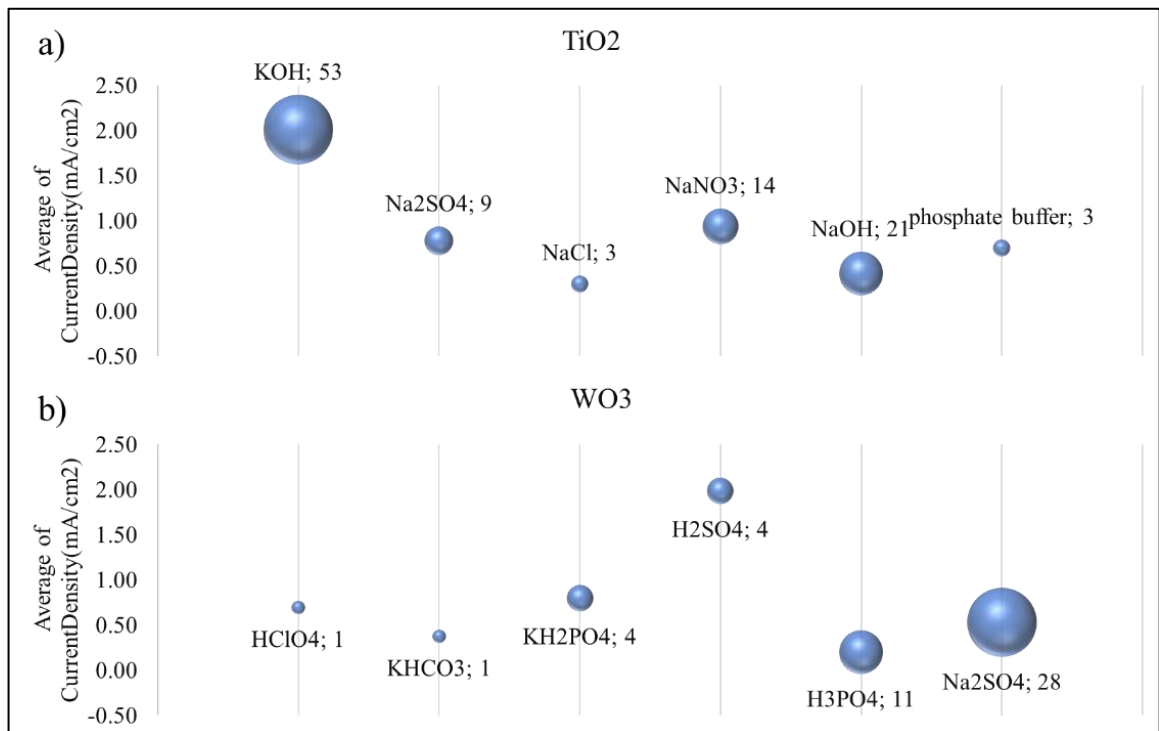


Figure 4.18. The effect of electrolytes for different semiconductors, a) TiO₂ and b) WO₃.

4.2.5. Association Rule Mining

Association rule mining (ARM) helps to reveal the previously unknown or frequently occurring if-then relations between input and output variables within a large dataset. ARM algorithm is appropriate for only categorical dataset and if there are continuous variables in the dataset they should be discretized in a meaningful way. For this study, the details of discretization were explained in Section 3.1.2. The terms related with ARM should be clarified to interpret the rules created. At the beginning two variables (possibly relatable) should be determined as antecedent (left hand side – LHS) and consequent (right hand side – RHS); antecedent and consequent can be considered as “if” and “then” part of the statement, respectively. In the rule below, the preparation method of anodization for the first layer of the semiconductor was selected as LHS and the band gap class of “*3high*” was chosen as RHS. By that rule, it was aimed to learn if the high band gap value of semiconductor is related with the preparation method of anodization or not; if there is a relation, the probability of seeing together in the population compared to the probability of seeing them individually can be calculated as

$$\{Method.I = anodization\} \rightarrow \{BandGapClass = 3high\}$$

support: 0.095 confidence: 0.72 lift: 2.25.

For qualitative interpretation, the terms of *support*, *confidence*, and *lift* should be clarified. The *support* is the probability of instances that meet both conditions in RHS and LHS in the population. For example, the number of instances in which anodization was selected as the bottom layer preparation method and the band gap value of synthesized semiconductor is higher than 3 eV is 38. Hence, the support value of 0.095 can be found by dividing 38 to 401 (total number of instances). The *confidence* is the ratio of the number of instances that meet the conditions defined in RHS and LHS to the number of instances that meet the condition in only LHS. In the given example below the first number is 38 as explained before, and the number of instances used anodization is 53 in total; so, the confidence 0.72 equals to 38 over 53. In ARM, the minimum support (0.0001) and confidence (0.0001) values were specified to eliminate the rare and irrelevant association rules. The algorithm gives also *lift* for each rule developed to make a comparison. The lift equals to the division of confidence with the fraction of instances meet the restriction in RHS. In the previous example the lift value of 2.25 can be calculated by dividing 0.72 (the confidence) to the fraction of instances have low band gap value in population (128/401) which is 0.32. In general, lift of one indicates no relation between the conditions defined in RHS and LHS while the lift larger (smaller) than one implies a positive (negative) relation between the conditions in RHS and LHS. Therefore, by considering the example above, it can be concluded that if anodization is selected as the preparation method for the bottom layer of the semiconductor, it is more likely (2.25 times more likely than being in population) to get a semiconductor having band gap value greater than 3 eV. In the Figure 4.19, the x axis represents bottom layer preparation method, y axis stands for the band gap classes, and the bubble size is correlated with the lift value of the association rule in which x-coordinate represents LHS and y-coordinate represents RHS.

The red, yellow, and purple colors refer to the band gap classes of *1low*, *2medium*, and *3high*, respectively. The effect of different bottom layer preparation methods on band gap value of semiconductor synthesized can be observed and compared by investigating Figure 4.19. For example, if the first layer is prepared by chemGro, that semiconductor tend to have a band gap value smaller than 2.40 eV (*1low* class), since the radius of red bubble is bigger than the radius of yellow and purple bubbles.

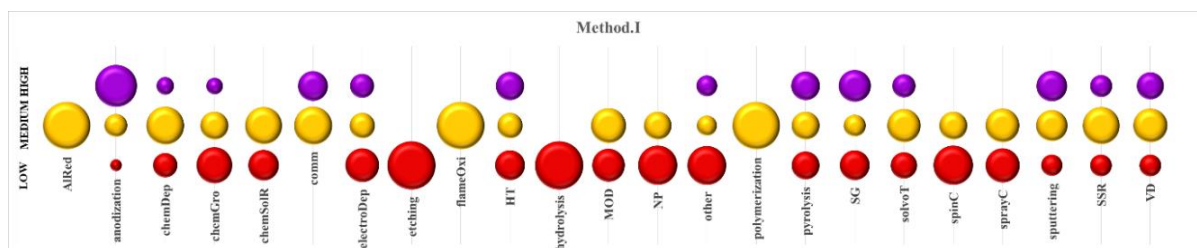


Figure 4.19. The effect of bottom layer preparation method on the band gap value of that semiconductor.

4.2.6. Band Gap Prediction by Random Forest Algorithm

Band gap can be assumed as one of the key factors in photo active reactions. However, band gap values of working electrodes were missing for some of the articles (78 out of 479 unique experiments) in the dataset. Those missing values were predicted by using random forest algorithm successfully. The input variables were selected as electrode materials, preparation methods, treatment conditions and calcination specifications, which are known as effective on band gap of a semiconductor. As explained in the Computational Details section, several *ntree* and *nodesize* values were tested to find the optimum RF model with the highest accuracy and lowest complexity. The 5-fold cross validation procedure was applied on 401 instances with known band gap values; 100 instances (out of 401) were defined as testing subset while 301 instances (out of 401) were divided into five subsets to be used in 5-fold CV. By evaluating the results of 5-fold CV procedure, the *ntree* and *nodesize* were determined as 20 and 5, respectively. The average validation (average of 5 folds) error, RMSE, and R^2 were calculated as 0.20, 0.31 and 0.61; while the testing error, RMSE, and R^2 were evaluated as 0.12, 0.17, and 0.87, respectively. In Figure 4.20 the predicted versus real band gap values were given for validation and testing.

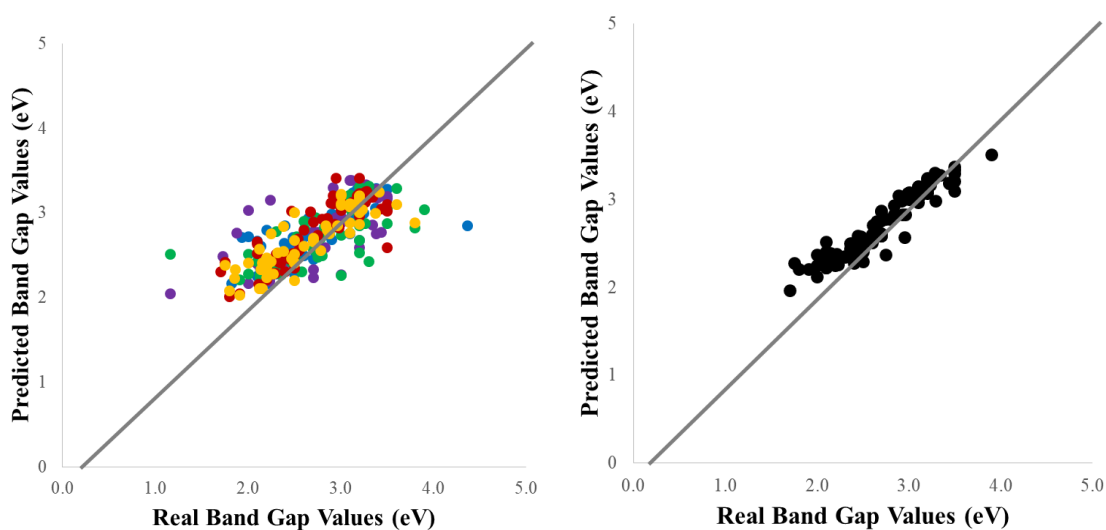


Figure 4.20. Predicted versus real band gap values by using random forest model for a) validation set, b) testing set.

In Figure 4.21a and Figure 4.21b the input variables vs the reduction in sum of squared errors (IncNodePurity), and the percentage increase in mean square error (%IncMSE) were given, respectively. Those IncNodePurity and %IncMSE values can be considered as their relative importance or contribution to the model improvement (see Appendix B for details). As can be seen from both Figure 4.21a and Figure 4.21b, the material used, the doping element and the preparation method for the bottom layer is very effective on the band gap prediction of the semiconductor used in PEC studies.

4.2.7. Band Gap Classification by Decision Tree Algorithm

As explained in Section 3.1.2, several DT models were created by changing the *cp* (complexity parameter) in a way that the minimum validation error with the simplest tree could be obtained. In Figure 4.22 the optimum decision tree for band gap classification with a *cp* value of 0.1 was given. It should be noted that the given tree developed by using the instances except those in testing subset.

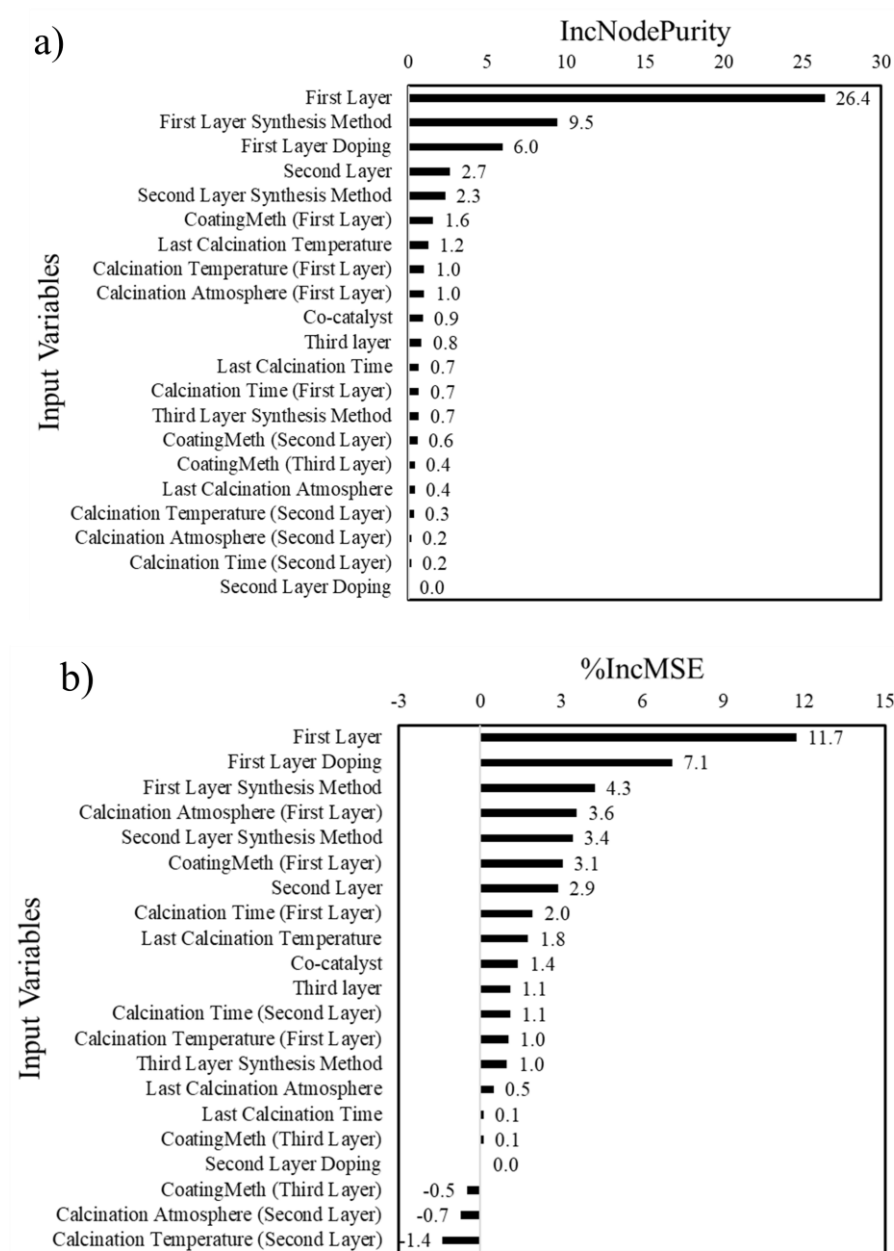


Figure 4.21. The input variables vs a) the reduction in sum of squared errors (IncNodePurity) and b) the percentage increase in mean square error (%IncMSE).

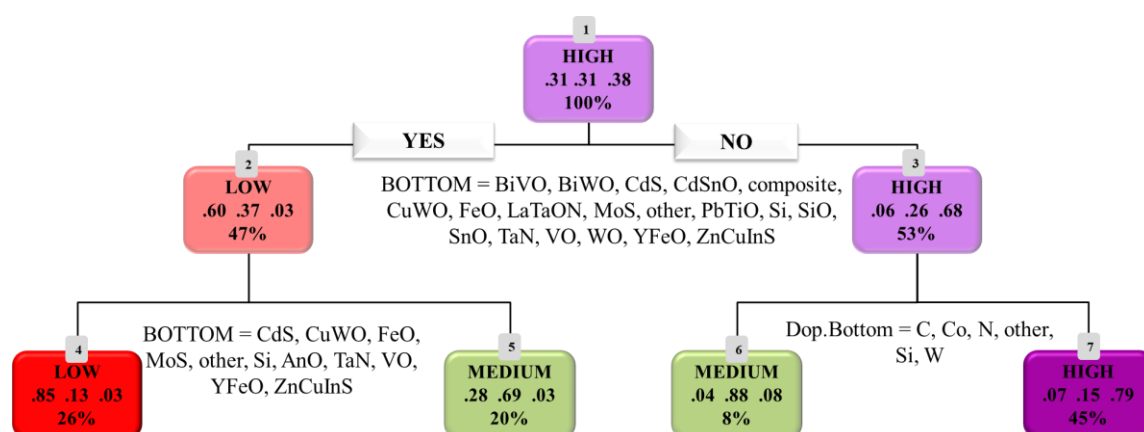


Figure 4.22. The decision tree for band gap classification.

The regular testing accuracy of the model is 72%, and the balanced testing accuracy (explained in Appendix B) is computed as 75%; the small difference between those values implies the distribution of the instances through classes are well balanced. The model predicted the instances in *class 1low*, *2medium*, and *3high* with the accuracies of 60%, 70%, and 95%, respectively (the confusion matrix given in Appendix Table B.3 with details).

In Figure 4.22, the nodes containing instances with mostly low band gap values (*class 1low*) are colored as red while those having mostly high band gap values (*class 3high*) are purple; the nodes contain instances with medium band gap values are indicated as green. The colors of the nodes become darker from root to the leaf nodes, i.e., with increasing size of the tree. The first line in a node indicates the band gap class of that node which is determined with majority voting. The second line gives the real fractions of each band gap class in that node and the last line is the percentage of the instances existed in that node among total instances. For example, in Figure 4.22, the band gap class of first node is decided as *3high* since it includes 38% instances belonging *3high* class. Then the branching starts with the first rule of whether the bottom layer includes some definite semiconductors or not; and if those specified semiconductors are in the bottom layer, then the tree continues to grow leftward, if not then it goes to the rightward. The decision tree should be followed like this from roof to the leaf (terminal) nodes. As explained before in the terminal node the first line represents the band gap class. For example, in node 4 the tree indicates that if the previous statements are answered as “yes” then the band gap of semiconductor obtained will be in *1low* class. The accuracy of that decision equals to the real fraction of the instances in *1low* class in node 4, which is 0.85. The reliability of that rule can be also evaluated with the

number of instances meeting that condition; in node 4 it is declared that 26% (78 out of 301) of instances meet those conditions (rules).

4.2.8. Photocurrent Density at Bias of 1 V Classification by Decision Tree Algorithm

Due to photocurrent density results being a large dataset, a reliable regression model for photocurrent density prediction could not be developed. Therefore, the dataset was simplified by only taking one bias value (1 V vs RHE since it is the most repeated one); 438 instances correspond to this filtering. In the development of random forest regression model, the same steps explained in the computational details were followed. The optimum parameters were found to be 68 for number of trees and 18 for *nodesize* with validation error, RMSE and R^2 of 0.60, 1.07 and 0.70 and testing error, RMSE and R^2 of 1.3, 1.67 and 0.50. When Figure 4.23 is investigated, it is seen that in training data model tends to predict lower photocurrent density values.

4.2.9. Photocurrent Density Classification

In classification of photocurrent density values, both random forest and decision tree algorithms were utilized. Details about RF classification is given in Appendix Table B.4. The model has high accuracy in validation set and good predictive power for high photocurrent density class (class A). Lower prediction power of class B is expected because middle classes tend to have leaks to other classes. From variable importance graph it is seen that in classification of photocurrent density, bias value, pH and molarity of the electrolyte and band gap are important decision variables.

Since RF algorithm is a black box model, it only gives insights about important variables but fails to return heuristics for achieving high photocurrent density values. Therefore, a decision tree model is also developed. In Table 4.5, the confusion matrix for decision tree model of the photocurrent density is given. When recall and precision values are analyzed, it is seen that model fails to predict medium class (B) correctly, which is usually the case due to the leaks from both sides, but the prediction of high and low classes (A and C respectively) are reasonably well.

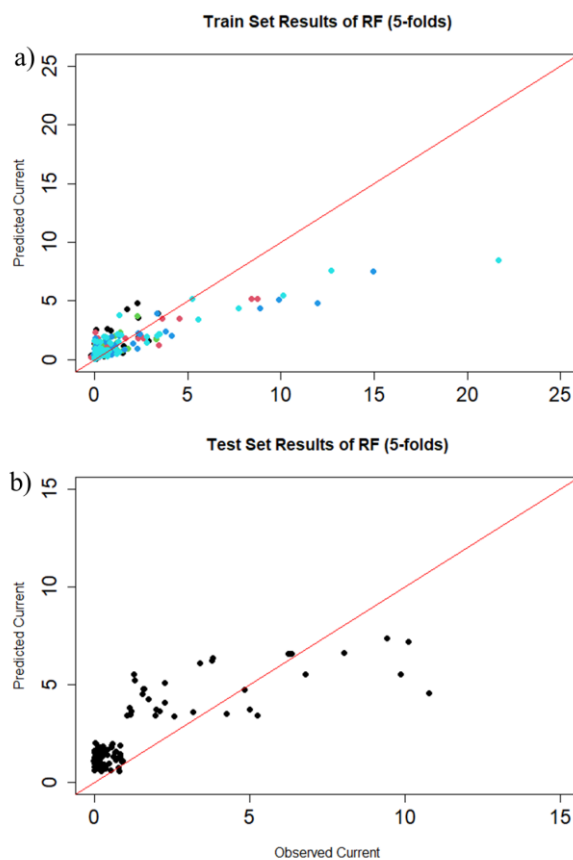


Figure 4.23. Predicted versus actual photocurrent values by using random forest model for a) validation sets, b) testing set.

Table 4.5. The confusion matrix for decision tree model for photocurrent density prediction.

	Overall Accuracy	Real Class	Total number	Predicted Class			Class Accuracy (Recall)
				A	B	C	
Validation	0.61	A	2671	197 7	438	256	0.74
		B	2623	908	106 3	652	0.40
		C	2622	326	469	182 7	0.70
			Precision	0.62	0.54	0.67	
Test	0.54	A	865	543	187	135	0.63
		B	909	276	348	285	0.38
		C	869	129	222	518	0.60
			Precision	0.57	0.46	0.55	

In Figure 4.24, the optimum decision tree model is given ($cp=0.01$). When the rules are read from top node to the bottom, some general trends can be observed for the high performance (if the number of data and the purity of the node are high). For example, it can be said that if the applied bias is less than 0.99 V and the molarity of the electrolyte is greater than 4 M, the tree results in class A (representing high class) with 0.92 purity suggesting that these combinations can be considered as a reliable rule to follow. It can be inferred that the bottom material of the photo anode, its synthesis method and electrolyte choice are important decision variables in producing high photocurrent densities.

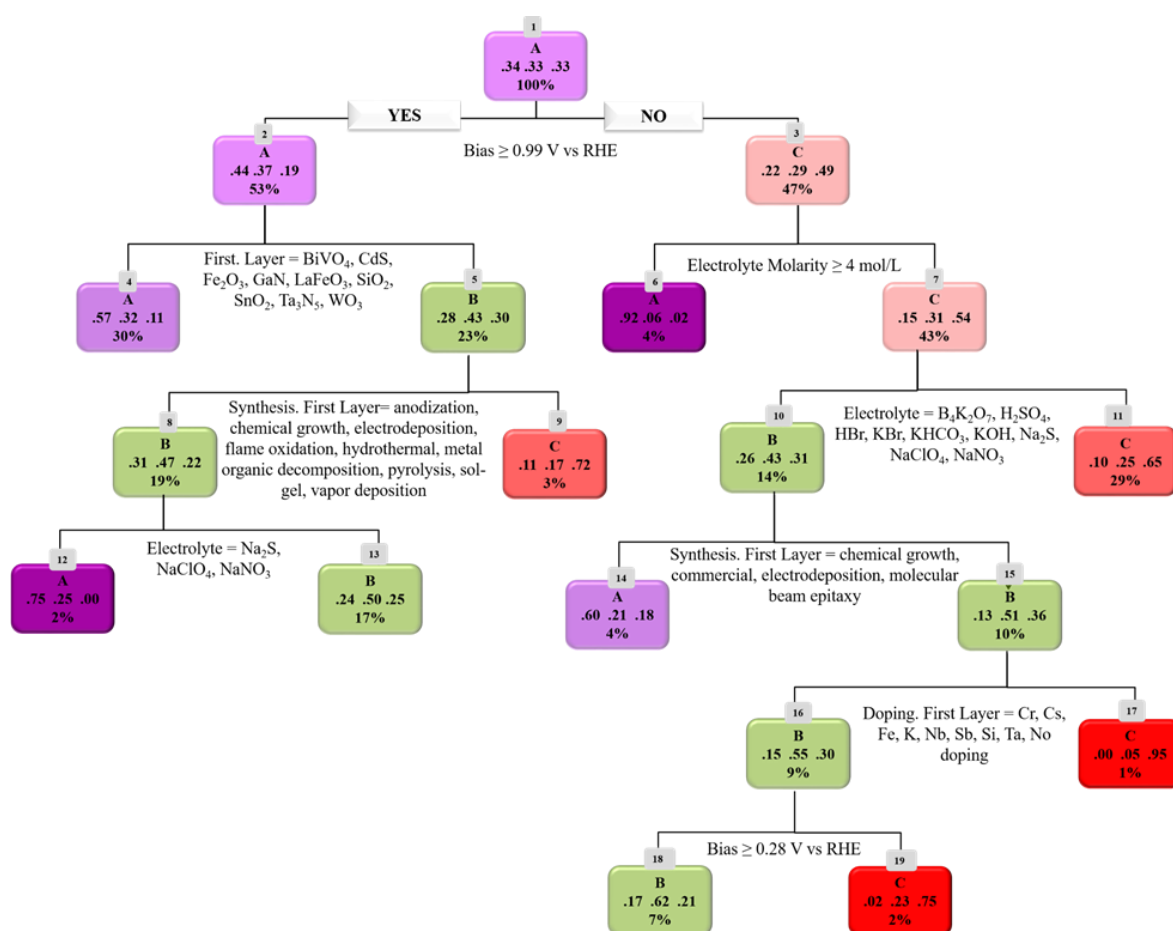


Figure 4.24. The decision tree for current density classification.

4.3. Machine Learning for Predicting Water Solubility in Ionic Liquids

The dataset for water solubility in ionic liquids was built computationally by Ahsan Jalal (PhD in Chemical Engineering, Koc University, Istanbul); the water solubility in ILs was computed by COSMO-RS while the molecular descriptors for the individual cations and anions were determined by semi-empirical PM3 method. Then the dataset was analyzed with machine learning tools, and the information extracted from ML models was compared with the literature to verify the models developed (BU undergraduate student İrem Gülçin Zırhlioğlu also contributed to the development of deep learning models). The results given in Section 4.3 has been published as a scientific article (Can et al., 2021) and revised for this thesis.

4.3.1. Pre-analysis of Data

Before implementing the ML algorithms mentioned above, some simple descriptive statistics were used to understand the data structure better and identify some basic patterns if there were any. For this purpose, the data were divided into certain groups using the range of descriptors, and the average solubility in each group were monitored against *the bin average* of descriptors so that the effects of each specific descriptor could be seen clearly (see Table C.3 in Appendix for the bin ranges). The change of average solubility is presented as the function of bin average of cation descriptors in Figure 4.25; the ball colors represent individual cation groups as given at the bottom of the Figure 4.25 while the corresponding sizes of the balls represent the number of ILs in that group with the given bin average. Figure 4.25 indicates that the water solubility in phosphonium group ILs is generally higher than the ILs in other groups while the lower LUMO energy, higher dipole, higher HBD and higher HBA count seem to promote lower water solubility. We also analyzed the dataset to see whether the values of some cations descriptors significantly differ in cations groups in a way that it may explain the patterns observed in Figure 4.25. Indeed, the range of seven descriptors (MW, CPK ovality, Polarizability, HBA count, HBD count, CPK Area and ZPE) are clearly distinguishable (lower or higher) for phosphonium group ILs, which dissolve water more compared to others. However, high LUMO energy level, low dipole, low HBA count and HDB count, which seem to be good for lower water solubility, are not clearly assignable to any specific cation groups (see Figure C.1 in Appendix). Hence, it may be concluded from the pre-analysis that the descriptors seem to account the differences in the

water solubility in ILs with different cation groups even though they cannot be directly related to any functional cations groups (such as alky groups, halogens etc.) as mentioned in Introduction. This result is probably because the ILs under the same cation groups have significant differences in their behavior depending on the properties of specific cation and anion as well as their combined effects; the ARM and decision tree analysis in the following section are more suitable to determine such specific effects.

4.3.2. Analysis of Descriptor Effects by Association Rule Mining

ARM was used to analyze the relations between the IL descriptors and the corresponding water solubility in the entire dataset as well as the subsets for the individual cation groups (i.e., IL families). The water solubility values (i.e., output) were divided into three classes as described in Computational details (*class A* for instances with 1.5 mol/mol and lower solubility, *class C* for instances with 15 mol/mol and higher solubility and *class B* for remaining data). The numerical cation and anion descriptors were also discretized in 10 categories as it was already discussed above (and presented in the Table C.3 in Appendix). The results obtained for the dataset consisting of the imidazolium family is given in Table 4.6 together with the values of the corresponding *support*, *confidence*, *lift* and *count*, which are the major parameters in ARM. Only the descriptor values having a *lift* value higher than one is given here because this level is a critical indicator as it will be apparent below; the complete list of factors for all IL families were given in supplementary material file of our article published recently (Can et al., 2021). Here, *support* shows the fraction of instances that meet the conditions in the second (descriptors) and the third (performance class) columns in the dataset. For example, there are 205 ILs meeting the conditions that anion LUMO is greater than or equal to 1 eV and less than 2 eV, and solubility is in low class (less than 1.5 mol/mol); this result leads to the *support* values of $205/4059=0.0505$ (4059 is the total number of imidazolium ILs as given in Table 3.8) as shown in first row in Table 4.6. The *confidence* is the fraction of instances that meet the conditions in the second column ($1 \text{ eV} \leq \text{anion LUMO} < 2 \text{ eV}$) and the third column (low solubility) to total number of data having $1 \text{ eV} \leq \text{anion LUMO} < 2 \text{ eV}$ in entire imidazolium dataset. Then, the confidence is $205 / 205 = 1.00$ indicating that water solubility is low in all ILs with this LUMO

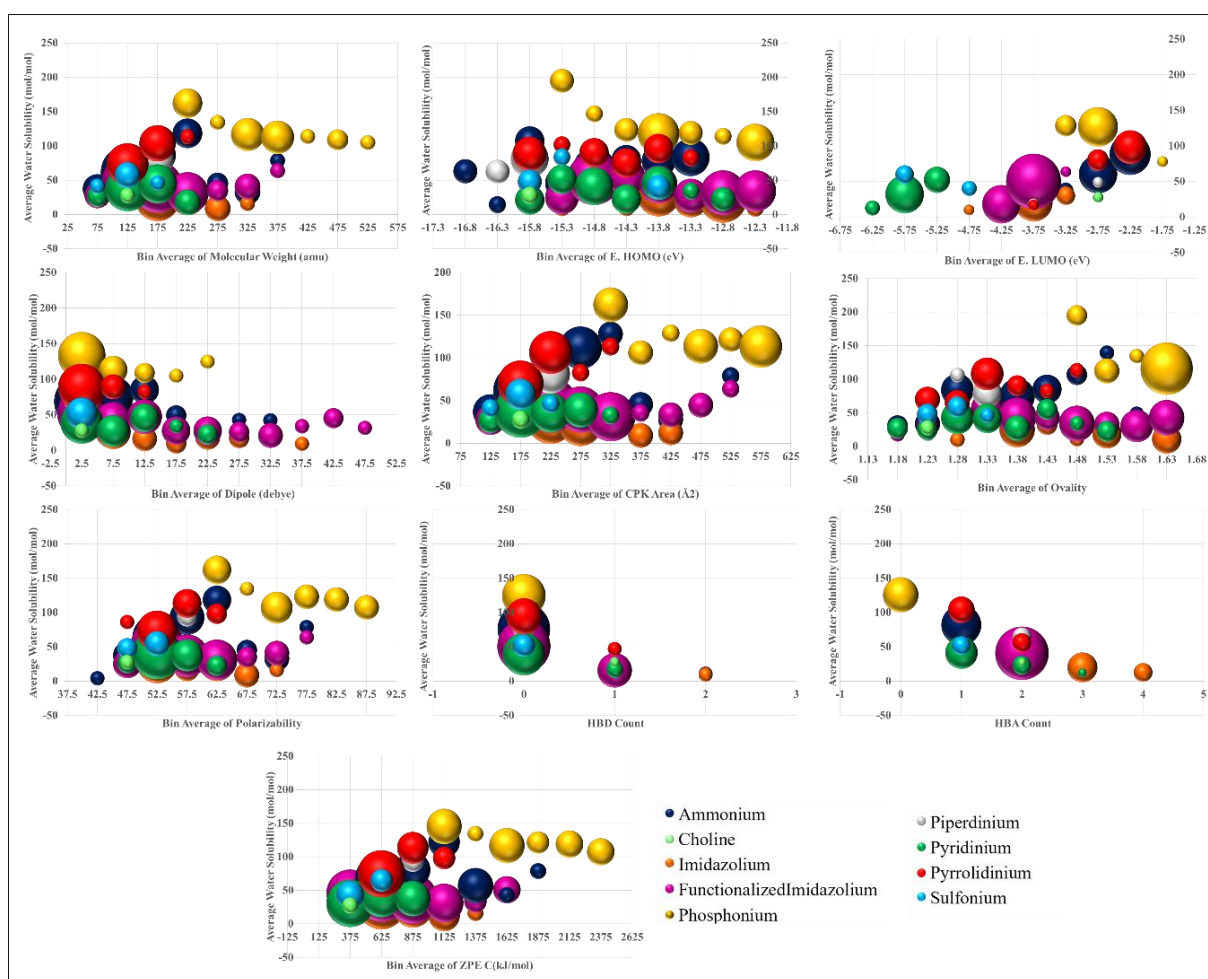


Figure 4.25. Effects of cation descriptors on average water solubility for various IL groups (See Table C.3 in Appendix for the range of bins used for the average values of descriptors).

level. Finally, the *lift* is calculated by dividing the confidence with the fraction of instances meeting the performance in column 3 (i.e., fraction of low solubility ILs in entire imidazolium dataset as given in Table 3.8 as $1495/4059=0.368$). Hence, the *lift* can be calculated for $1\text{ eV} \leq \text{anion LUMO} < 2\text{ eV}$ as $1/0.368 = 2.72$ ($205/205$)/($1495/4059$). This result means that *the probability of having low water solubility is 2.72 times higher in ILs with $2\text{ eV} > \text{anion LUMO} \geq 1\text{ eV}$ than all ILs in the imidazolium dataset*; hence, having an anion with this LUMO range is favorable for low water solubility. As can be clearly seen from this example, the factors having the positive impact on the desired outcome will have the *lift* value of higher than one, and the higher values are better. Similar discussion can be carried out for all entries in Table 4.6.

Table 4.6 clearly indicates that the anion properties are much more influential for water solubility in imidazolium ILs; this inference was also found to be the case for the other cation groups as well (see Supplementary Material file of the related article (Can et al., 2021)). Hence, we also tested whether anion effects can be generalized to the entire dataset without dividing into cation groups, and obtained the results given in Table 4.7. Although the exact order is not the same (for example, instead of anion LUMO, anion HOMO has the highest lift here); the dominance of the anion characteristics is also quite clear in the entire dataset.

Then, we inspected the results further, and noticed that, regardless of the cation used, about one-third of the anions result in ILs with mostly low (and rarely medium but never high) solubility while approximately one-third gave ILs with mostly high (and rarely medium but never low) solubility. The remaining one third form mostly moderate water solubility (and rarely low or high but never both). The results are presented in Figure 4.26 for all solubility levels; the ball sizes denote the number of ILs synthesized with that anion so that the relative frequency of these anions in the dataset can be also seen (not all anions formed IL with all cations; hence, they are not represented in equal numbers in the database).

As explained in Computational details, the low, moderate, and high solubility subsets, which were obtained by dividing into three equal parts based on solubility; this class structure seems to also divide the anions approximately the same way as the further indicator for the deterministic role of anions for solubility. The presence of small fraction of ILs in the neighbors of the dominant class for an anion in Figure 4.26 merely indicates that the low, medium, and high solubility ranges for individual anions are not exactly the same as the entire set (but they are quite close; hence only small fractions of data may be placed to a neighboring classes).

The anions that form ILs with low water solubility have relatively higher molecular weights, lower HOMO, lower LUMO, and lower dipole values than the other ILs in the dataset. This inference can be explained using Koopman's theorem stating that the negative HOMO and the negative LUMO energies are directly related to ionization potential and electron affinity respectively (Zhan *et al.*, 2003). According to Shriver and Atkins, if a molecule has a very high electron affinity and/or very high ionization energy, that molecule is also highly electronegative (Atkins and Paula, 2006). Consequently, it can be deduced that

negative HOMO and negative LUMO energies are positively correlated with the electronegativity. It should be also considered that the polarity of a molecule could be derived and IUPAC labels all polar compounds as hydrophilic while nonpolar compounds are classified as hydrophobic (Galvert, 1990). Hence, we posit that negative HOMO and negative LUMO energies can be also directly linked to hydrophilicity. Similarly, the dipole moment is also a measure of the polarity and therefore hydrophilicity (Blaber, 2014) while the electronegativity decreases with the increasing size of the atoms as can be also clearly observed in the periodic table; consequently, the molecular weight should also affect the hydrophilicity positively. Finally, it is also known that the anion part is effective in the hydrophobic or hydrophilic behavior of the IL (Ranke *et al.*, 2009), and the ILs with a hydrophilic anion has relatively higher water solubility than those with hydrophobic anions (Boruń, 2019; Isosaari *et al.*, 2019; Lawal *et al.*, 2019; Marcinkowska *et al.*, 2019; Nawala *et al.*, 2018; Sajid, 2019; Verma *et al.*, 2019). Indeed, most of the anions in Figure 4.26 upper panel are hydrophobic while the anions in Figure 4.26 lower panel have hydrophilic character explaining the corresponding capacity of resulting IL toward water. As the result, our finding related to the desired anion properties for low water solubility (or high solubility as well) are in close agreement with the theoretical and experimental evidence reported in literature.

4.3.3. Heuristics for Cation/Anion Pairing by Decision Tree

As explained in the Section 3.1.3, the datasets were divided into three classes in water solubility: 0-1.5 mol/mol (including 1.5 mol/mol) as *class A*, 1.5-15 mol/mol (including 15 mol/mol) as *class B*, and higher than 15 mol/mol as *class C*. According to this division, the imidazolium data set contained 1499 instances in *class A*, 1288 instances in *class B* and 1272 instances in *class C* (see Table 3.8). Since the number of instances in classes are quite close to each other, we did not have any precaution (like random sampling) for class imbalance problem (Bramer, 2008) during tree construction; however, we calculated the balanced accuracy for the optimum model to verify that the class imbalance problem was indeed not serious (given below). Before starting the analysis, 25% of the data was randomly selected as the *test* set while the remaining 75% was used for model construction; the 5-fold cross validation procedure (i.e., 60% of data was used for *training* and 15% as *validation*), as explained in Section 3.1.3. After all the model parameters were decided through 5-fold cross

validation procedure (i.e., after the model was fully constructed), the classification accuracy of the model was tested using the 25% testing data, which was not seen by the model during construction.

Table 4.6. Association rules leading class A (low solubility) for imidazolium dataset.

<i>LHS^a</i>	<i>RHS^b</i>	<i>Support^c</i>	<i>Confidence^d</i>	<i>Lift^e</i>	<i>Count^f</i>
{2 eV>Anion.E.LUMO ≥1 eV}	Class A (solubility < 1.5 mol/mol)	0.05	1.00	2.72	205
{-9 eV>Anion.E.HOMO}		0.05	1.00	2.72	205
{-7eV>Anion.E.HOMO≥-8 eV}		0.09	1.00	2.72	369
{3 debye>Anion.Dipole≥1 debye}		0.10	0.91	2.48	412
{Anion.CPK.Area>260 Å ² }		0.09	0.85	2.30	348
{300 amu>Anion.MW≥250 amu}		0.06	0.78	2.13	225
{Anion.HBA.Count=0}		0.16	0.77	2.09	630
{1 debye>Anion.Dipole}		0.20	0.74	2.02	794
{3 eV>Anion.E.LUMO≥2 eV}		0.07	0.70	1.90	287
{50 kJ/mol>Anion.ZPE}		0.12	0.64	1.73	469
{140 Å ² >Anion.CPK.Area≥120 Å ² }		0.08	0.62	1.69	332
{-6 eV>Anion.E.HOMO≥-7 eV}		0.12	0.60	1.63	469
{1 eV>Anion.E.LUMO}		0.05	0.60	1.62	220
{120 Å ² >Anion.CPK.Area≥100 Å ² }		0.07	0.55	1.48	291
{49>Anion.Polarizability≥46}		0.12	0.53	1.44	501
{5 eV>Anion.E.LUMO≥4 eV}		0.07	0.49	1.33	281
{220 Å ² >Anion.CPK.Area≥180 Å ² }		0.05	0.47	1.29	214
{55>Anion.Polarizability≥52}		0.05	0.47	1.27	211
{Cation.CPK.Ovality≥1.6}		0.06	0.44	1.20	262
{Anion.HBD.Count=0}		0.36	0.44	1.19	1475
{Cation.E.HOMO≥-12.5 eV}		0.06	0.43	1.17	255

Table 4.6. Association rules leading class A (low solubility) for imidazolium dataset.

(cont.)

<i>LHS^a</i>	<i>RHS^b</i>	<i>Support^c</i>	<i>Confidence^d</i>	<i>Lift^e</i>	<i>Count^f</i>
{100 kJ/mol>Anion.ZPE $\geq$ 50 kJ/mol}		0.06	0.43	1.16	245
{Anion.HBA.Count=4}		0.11	0.42	1.14	432
{200 amu>Anion.MW $\geq$ 150 amu}		0.09	0.39	1.07	371
{1.24>Anion.CPK.Ovality $\geq$ 1.18}		0.08	0.39	1.05	333
{1250 kJ/mol>Cation.ZPE $\geq$ 1000 kJ/mol}		0.06	0.39	1.05	229
{-12.5 eV>Cation.E.HOMO $\geq$ -13 eV}		0.07	0.39	1.05	267
{65>Cation.Polarizability $\geq$ 60}		0.07	0.38	1.03	300
{350 Å ² >Cation.CPK.Area $\geq$ 300 Å ² }		0.07	0.38	1.03	300
{250 amu>Cation.MW $\geq$ 200 amu}		0.07	0.38	1.03	300
{Cation.HBD.Count=0}		0.27	0.38	1.02	1077
{15 debye>Cation.Dipole $\geq$ 10 debye}		0.05	0.37	1.01	222
{-4.75 eV>Cation.E.LUMO $\geq$ -5 eV}		0.25	0.37	1.01	998

^a The features of anion or cation which gives high solubility in related dataset^b Solubility level^c This is the fraction of all data belonging to class in RHS and has feature in LHS^d This fraction of data in RHS has feature in LHS^e This fraction of data in RHS has a feature in LHS / (this fraction of data in all data has a feature in LHS)^f The # of instances having the factors in LHS and belonging to the efficiency class in RHS

Table 4.7. Association rules leading class A (low solubility) for entire dataset.

<i>LHS</i>	<i>RHS</i>	<i>Support</i>	<i>Confidence</i>	<i>Lift</i>	<i>Count</i>
{-7 eV>Anion.E.HOMO $\geq$ -8 eV}	Class A (solubility < 1.5 mol/mol)	0.09	0.97	2.73	1421
{3 debye>Anion.Dipole $\geq$ 1 debye}		0.10	0.88	2.48	1581
{Anion.CPK.Area>260 Å ² }		0.08	0.83	2.32	1345
{300 amu>Anion.MW $\geq$ 250 amu}		0.05	0.76	2.12	862
{Anion.HBA.Count=0}		0.15	0.74	2.09	2414
{1 debye>Anion.Dipole}		0.19	0.72	2.02	3035
{3 eV>Anion.E.LUMO $\geq$ 2 eV}		0.07	0.68	1.92	1114
{50 kJ/mol>Anion.ZPE}		0.11	0.61	1.72	1788
{140 Å ² >Anion.CPK.Area $\geq$ 120 Å ² }		0.08	0.60	1.70	1280
{-6 eV>Anion.E.HOMO $\geq$ -7 eV}		0.11	0.58	1.63	1793
{1 eV>Anion.E.LUMO}		0.05	0.58	1.62	846
{120 Å ² >Anion.CPK.Area $\geq$ 100 Å ² }		0.07	0.53	1.49	1119
{49>Anion.Polarizability $\geq$ 46}		0.12	0.52	1.46	1939
{5 eV>Anion.E.LUMO $\geq$ 4 eV}		0.07	0.47	1.31	1065
{220 Å ² >Anion.CPK.Area $\geq$ 180 Å ² }		0.05	0.45	1.27	808
{Cation.CPK.Ovality $\geq$ 1.6}		0.07	0.42	1.20	1177
{Anion.HBD.Count=0}		0.35	0.42	1.19	5667
{100 kJ/mol>Anion.ZPE $\geq$ 50 kJ/mol}		0.06	0.42	1.17	948
{Anion.HBA.Count=4}		0.10	0.41	1.15	1658
{-4.25 eV>Cation.E.LUMO $\geq$ 4.5 eV}		0.08	0.38	1.08	1252
{200 amu>Anion.MW $\geq$ 150 amu}		0.09	0.38	1.07	1424
{1.24>Anion.CPK.Ovality $\geq$ 1.18}		0.08	0.38	1.06	1289
{-13.5 eV>Cation.E.HOMO $\geq$ -14 eV}		0.07	0.37	1.04	1135
{Cation.HBD.Count=0}		0.29	0.37	1.03	4646
{65>Cation.Polarizability $\geq$ 60}		0.06	0.36	1.02	932
{-4.75 eV>Cation.E.LUMO $\geq$ -5 eV}		0.10	0.36	1.01	1562

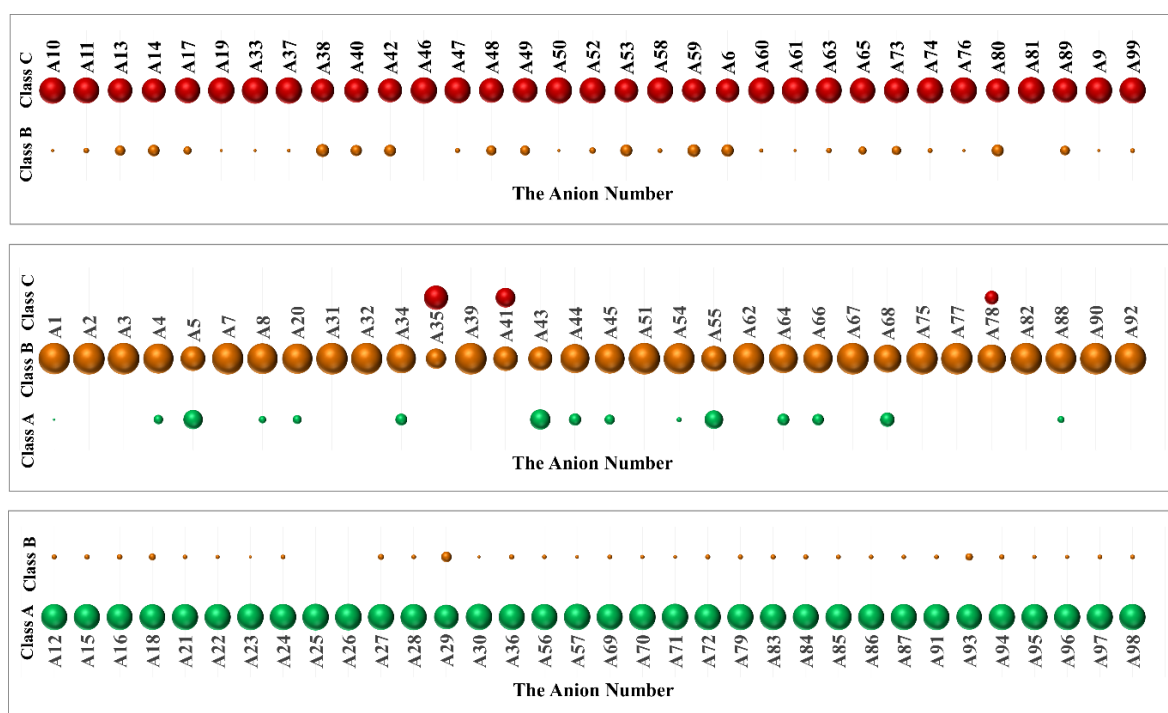


Figure 4.26. Distribution of anions leading various solubility classes. Upper panel shows high water solubility class (*class C*) while middle and lower panels indicate intermediate (*class B*) and low water solubility (*class A*) classes respectively (A# represents the individual anions; the list of anions corresponding the index number # is given in Appendix C).

Various DT models were created by changing cp (the complexity parameters) in a way that the minimum validation error could be obtained. The DT model given in Figure 4.27, with a cp value of 0.011, was found to be optimum for the imidazolium dataset. The testing accuracy of the model is 92.2%, which is incredibly high indicating the high predictive power of the model. The balanced accuracy (see Appendix B for details) was computed as 92.0%, which is quite close to testing accuracy indicating that we do not have a serious class imbalance problem. The model predicted the instances in *class A* with the accuracy of 94.3% while the B and C classes were predicted with an accuracy of 88.3% and 93.6% respectively; it is expected to have a lower accuracy for *class B* because of the misplaced data from both A and C sides. The confusion matrix, which contains the class recalls (fraction of data in class X that was correctly classified as class X) and class precisions (fraction of actual class X cases in the data that were classified as class X) are presented in Table C.4-C.11 in Appendix. The higher prediction accuracy for the *class A* is

more desirable because we would like to have heuristic rules for this class (if one is interested in high solubility, the accuracy of *class C* is also remarkably good). The purity of a terminal node in *class A* (i.e., high fraction of *class A* data in that specific node) is also important. Since the tree split the data using majority voting, the high purity of a terminal node in desired class indicates that the criteria (rules) used for splitting the data through the branch are identified correctly to promote that class performance (meanwhile the lower fraction of other classes will be incorrectly placed to that node). Consequently, if a node has higher purity in *class A*, the rules that will be deduced from that branch will be more reliable. As can be seen from Figure 4.27, and as will be more apparent below, the purity of the terminal nodes containing *class A* ILs is very high.

In Figure 4.27, the nodes containing mostly *class A* ILs are colored as green while those having mostly *class C* instances are red; the nodes with *class B* majority are marked as yellow. The colors of the nodes become darker with increasing *depth* (of the tree because purity increases). The first line in a node indicates the index for the majority class while the second line shows the fraction of *class A*, *B*, and *C* in that node respectively; the percentage in the last line represents the fraction of data obeying the rules described in that branch from top to down up to that node. The statement just below the node refers to the conditions used for splitting after that node; the branch on the left-hand side contains the data obeying the rule, while the data on the right do just the opposite.

The first node has the total imidazolium dataset, which is labelled as *A* because this class has the highest number of ILs (37%, 1499/4095) in the entire dataset. The branching starts with the value of anion HOMO (eV) and continues to the left if anion HOMO is lower than -6.3 eV followed by the anion ZPE reaching to the left most terminal node (node 4) if the anion ZPE is smaller than 509 kJ/mol. This node contains 25% of data, which are all *class A* (100% pure); thus, we can deduce that *if the anion HOMO is lower than -6.3 eV and ZPE is lower than 509 kJ/mol, the water solubility in the resulting IL will be low (lower than 1.5 mol/mol)*. This inference can be considered as a reliable heuristic rule for low water solubility because large fraction of the *class A* data obeys this rule without any false classification from the other classes. It should be also noted that the model reaches to 96% accuracy in the first division (node 2) based on anion HOMO value; hence, a fairly correct selection can be made by just checking the anion HOMO energy without even considering the anion ZPE value. Such rules can constitute a guideline for a researcher to synthesize ILs

with low water solubility or to avoid producing ILs with high water solubility by just computing HOMO energy for the potential anions. In this respect, calculating the related descriptors at a semi empirical PM3 level offers a significant flexibility as these calculations can even be performed on freely available software. Indeed, when the dataset is closely inspected, it can be found that there are 25 anions (see Appendix C for details) out of 99 leading to node 4 and these anions usually offer a low (and rarely medium) water solubility as it was also detected in ARM analysis in previous section 4.3.2.

A similar analysis can be also performed for the other branches. For example, the branches leading to the terminal nodes of 12 and 16 may be also used to deduce some other heuristics for low water solubility; although the numbers of instances are low compared to node 4, the purity of the nodes are still 100% indicating the reliability of these results. The terminal node 14 is also rich in *class A*; however, the rules described by the branches that lead to this node is not as strong as the first three discussed above, because both the number of instances (61 ILs) and purity of the node (76 %) are relatively lower (there are 61 ILs obeying these rules and 46 is *class A*). It should be noted that, however, even such accuracy and purity may be considered as high in DT applications in many other fields.

As clearly seen from Figure 4.27, the tree mostly created branches according to the properties of anion with minor contributions from cation properties verifying the results obtained with ARM. The same is also true for the other IL groups as presented in Figure C.2- C.8 in Appendix; the low water solubility always depends on the anion properties for all cation groups. Considering that mostly anions determine the water solubility in ILs, and some anions are good candidate for most of the cation groups, we also tested the possibility of developing a single DT model from the entire dataset to deduce common rules that are applicable to all IL families. The same class structure and the 5-fold cross-validation procedure that were employed for the individual cation groups were also used for the entire dataset. Figure 4.28 represents the optimum decision tree with the cp of 0.011; the testing accuracy was 92.2%, which is quite high as it was the case for the models for the individual cation groups while the balanced accuracy was 92.1%. The accuracy rates for the individual classes are also given as the confusion matrix in Appendix C. Such high accuracy rates indicate that even a single model constructed without considering the presence of different cation groups is sufficient to represent the data; this is also evident from the tree structure showing that the anion properties are also dominant in this model. For example, the leftmost

branch that led to the terminal group (node 4) with the low solubility instances (*class A*); again, the division is based on anion HOMO and then ZPE values as in the case of imidazolium dataset. Such similarities also exist in all branches in tree indicating the sufficiency of the single model; the conditions from top down to the terminal nodes with the significant number of low solubility data and high purity can be treated again, as heuristics to obtain ILs that dissolves the water in low levels.

4.3.4. Predictive Models for Water Solubility by Deep Learning

A multilayer fully connected neural network was used to develop predictive models (see Section 3.1.3 for details of deep learning model). As in the decision tree analysis, 25% of data was randomly selected for testing while the remaining 75% was used for model building in a 5-fold cross validation (i.e., 60% for training and 15% for validation in rotation) for each cation groups. For each dataset, a detailed grid search was employed to optimize the model hyperparameters for maximum validation accuracy (or minimum validation loss) with smallest possible model. The optimum values of these parameters were found to be the same for the individual cation groups, which have sufficiently large number of data and entire dataset (DL models were not built for the small IL families in the dataset).

The plot of actual versus predicted water solubility for testing (for the data not seen by the model) is given in Figure 4.29a while the relative importance of the descriptors is presented in Figure 4.29b. The model predictions are quite high with the testing accuracy rate of 89%; however, it is also clearly seen from the Figure 4.29a that the model predictions for higher solubility are less accurate. In fact, the model failed to predict the water solubility values higher than 250 mol/mol for 78 ILs (out of 4059, see Table 3.8) accurately. These cases were not included to Figure 4.29a even though they were not removed from the training data to prevent information leak (meaning that information obtained after testing would be used back in training to correct the model and diminish the independence of testing step). It was turned out that the cause of the problem was inaccurate COSMO-RS predictions because the experimental solubility values for some of these 78 ILs were actually low (see Section 4.3.5 for model verification). We could not identify the reasons for such unexpectedly high COSMO-RS results for these ILs and decided that such high predictions by the model should be used with cautions because any problems associated with the nature of COSMO-RS calculations also affect the model. However, the model predictions for the remaining ILs

(low to moderate solubility cases) seem to be quite reliable as the corresponding *RMSE* and R^2 for the predicted versus computed solubility plot after removal of data for these 78 ILs (Figure 4.29a) are 5.0 and 0.99, respectively (as also verified by the experimental results in Section 4.3.5). In addition to more accurate COSMO-RS predictions in low solubility, the higher frequency of the data in low and moderate solubility range (1/3 and 2/3 of ILs have lower solubility than 1.5 mol/mol and 15 mol/mol, respectively) also influenced the model hyperparameters more causing that the model represents the low and moderate solubility cases better.

The high prediction accuracy (hence, low error) also indicates that the model identified the significant relations between the descriptors and output variable (water solubility); this result may be used to determine the relative importance of descriptors, and this information may be utilized in the selection of anions and cations for a low water solubility. We used the *permutation feature importance method*, which randomly shuffles the values of one descriptor (so that its relationship with output is broken), and then measure of the changes in the validation loss. By repeating this analysis for each descriptor, the relative importance of descriptors, which indicates percent changes in response with the change of the value of a descriptor, can be established as presented in Figure 4.29b for imidazolium dataset. The anion descriptors, starting from Anion ZPE and Anion HOMO are also dominant here as in the case of decision tree analysis.

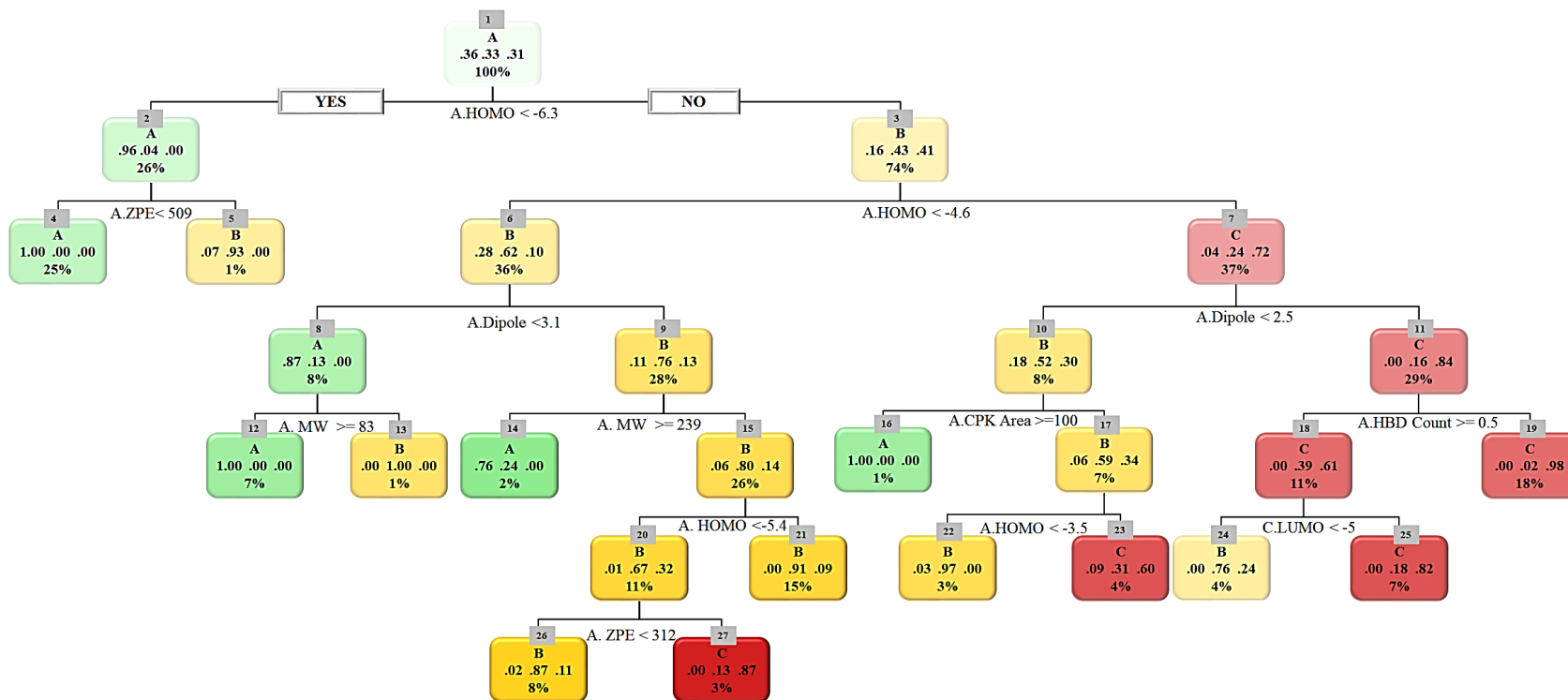


Figure 4.27. Optimum decision tree structure for imidazolium dataset.

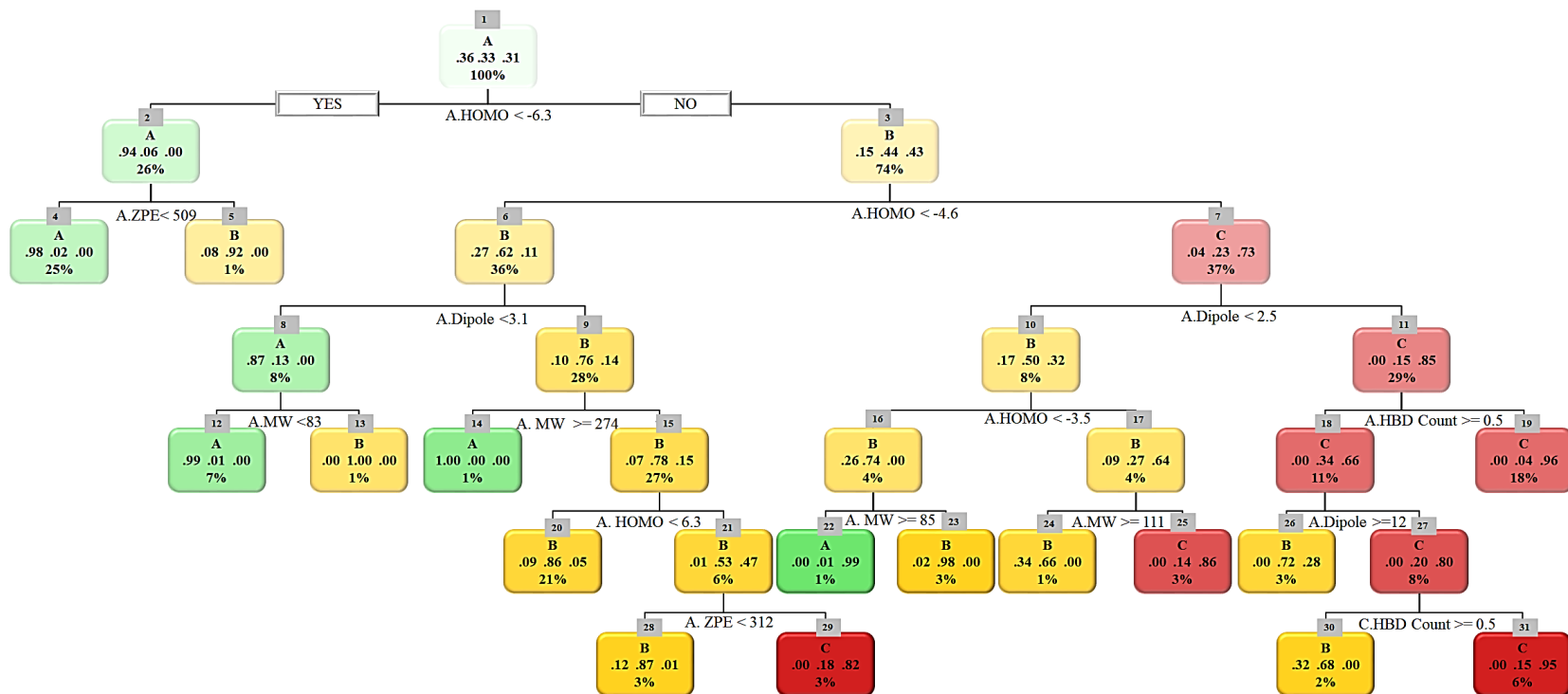


Figure 4.28. Optimum decision tree structure for entire dataset.

However, the cation descriptors are also placed at the top of the list here. Apparently, the anion descriptors are not sufficient for themselves for *point prediction*, which is more difficult than decision tree classification, which can be considered as *interval prediction* for water solubility. Considering the high accuracy of the DL model discussed above and the success obtained in developing a decision tree model for entire data set, a single DL model was also developed from the entire data set to predict the water solubility of ILs from any cation groups. The same 5-fold cross validation procedure, with the same ratio of data for training, validation, and testing; the same hyperparameters were also used in the model development. The model predictions for high solubility cases were, again, not as accurate as the low solubility ILs probably for the same reasons discussed above. The model was quite successful with an accuracy level of 89 %, which is quite similar to those for the imidazolium dataset. Again, the model predictions (higher than 400 mol/mol solubility) for 313 ILs (out of 16137 in total) were not included in the analysis even though they used in the training (to prevent information leak). After removal of these cases, the DL predicted versus the COSMO-RS computed solubility plot in Figure 4.30a resulted in the RMSE and R^2 values of 23.4 and 0.84, respectively; the RMSE is clearly larger than the imidazolium case, but it is still low enough to consider the model as successful.

The list of the significant descriptors (Figure 4.30b) is also close to those obtained for imidazolium data. However, the contributions of cation descriptors are more apparent here as expected; although the anion properties are more significant for water solubility in general, the cation properties are also needed to show the difference between water solubility in ILs belonging to different cation groups.

As the last set of analysis, we developed DL models using smaller fractions of data (for training and validation) to test the remaining larger portions to see whether the model is going to be still significant as the indicator of extendibility. We randomly separated 50% of the data for testing and constructed a model with the remaining 50% data (for training and validation), and the accuracy decreased only to 88% (error increased only 1%). Then, we proceeded further, and constructed a model using 4000 data (25% of total) for training and validation and tested the model with the remaining approximately 12000 data; the accuracy was still about 82%. This result indicates that the model is likely to predict the new data as long as it is one of the ILs families used in the dataset.

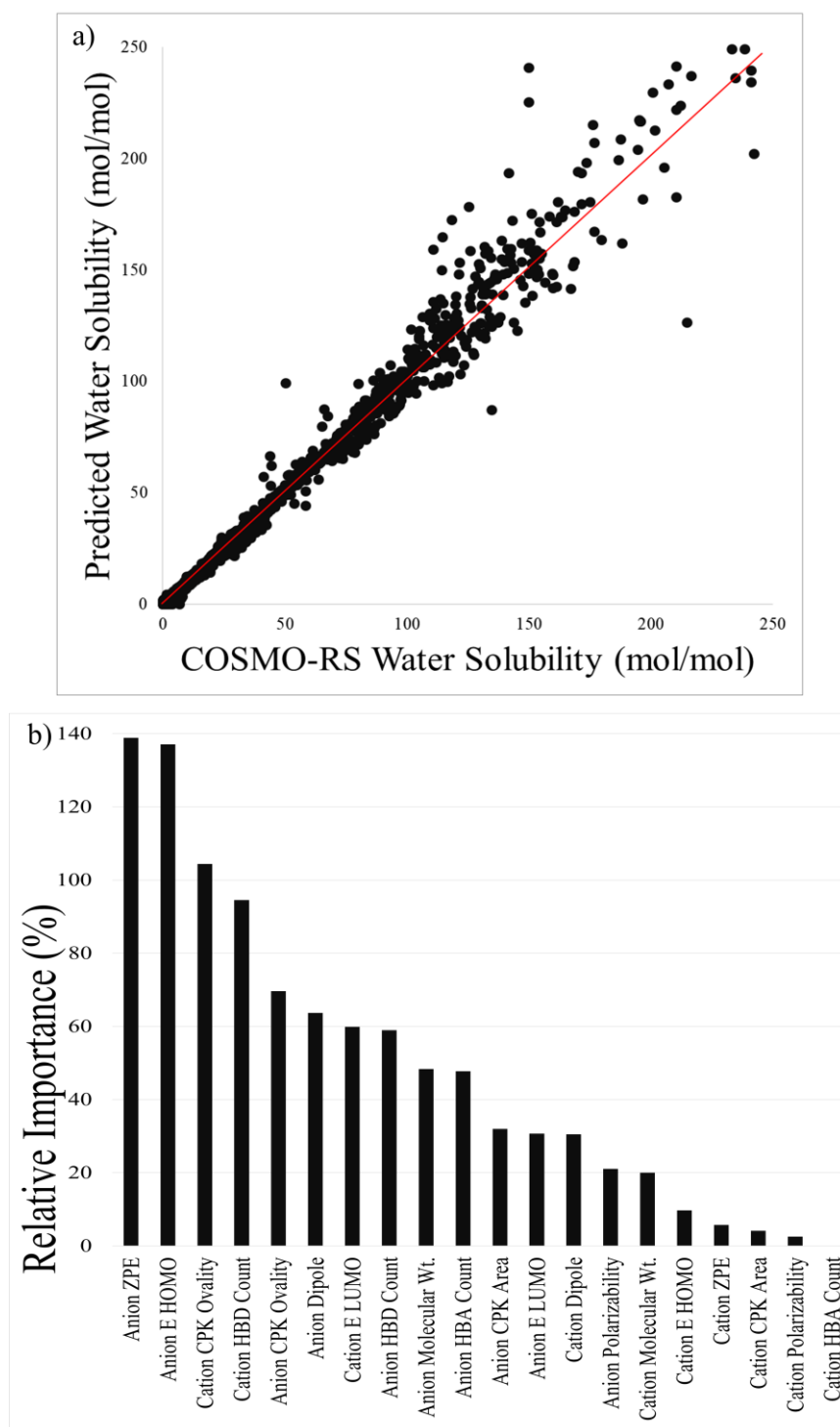


Figure 4.29. Testing results of optimum deep learning model developed for imidazolium dataset. (a) Predicted versus computed solubility for testing, (b) relative importance of descriptors (see Section 3.1.3 for model hyperparameters).

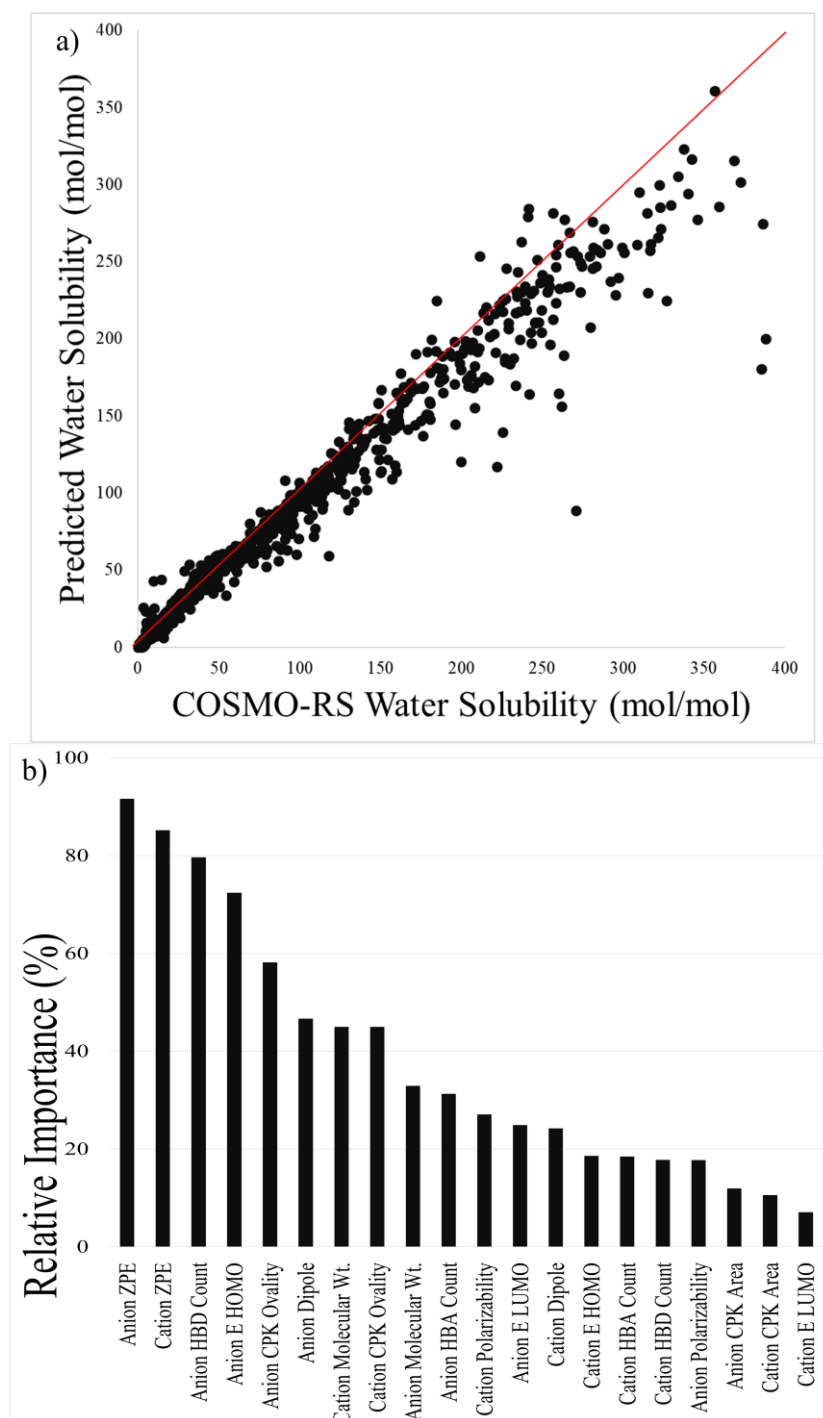


Figure 4.30. Testing results of optimum deep learning model developed for entire dataset. (a) Predicted versus computed solubility for testing, (b) relative importance of descriptors (see Section 3.1.3 for model hyperparameters).

4.3.5. Verification of Models Using Experimental Data from Literature

It is known that COSMO-RS can predict the solubility with an order of magnitude accuracy and can provide qualitative comparison among the alternative materials if their solubility is also computed in the same way; the low-cost semi-empirical methods used for the calculation of descriptors should also make significant contribution to the error in the dataset. Although the error associated with the ML techniques should be also considered, they seem to be not high in the present case considering that all the models seem to represent the experimental data quite accurately.

In order to verify the decision tree and deep learning models, we extracted 49 experimental solubility data from 13 articles from the literature (detailed information about those 49 ILs were given in Appendix C); although we cannot claim that this is an exhausted list, it seems unlikely that we can find sufficient number of experimental data for training the models or modifying the computational dataset. However, this number should be sufficient to verify the models already developed by the computational data. First, we tested the decision tree model from entire data set and found that the model correctly classified 46 of 49 data points corresponding to 93% accuracy, which is remarkably high. Then, we tested the deep learning model with the same experimental data; however, since the experimental data are mostly accumulated in the low solubility region (mostly less than 1 mol/mol) and the COSMO-RS predictions are generally reliable within an order of magnitude, the direct comparison would not be realistic. Hence, we used logarithmic scale to compare the experimental data with corresponding COSMO-RS calculations and deep learning predictions using the model developed for entire data set. As it is clearly seen in Figure 4.31, the COSMO-RS computed solubility (and consequently DL predictions) values are generally in agreement with most of the experimental solubility data (the same 46 of 49 IL came out in decision tree tests) in logarithmic scale. Some DL predications could not be presented in the plot (due to the logarithmic scale) because the model assigned zero or very small negative values (within the error range) for the solubility of ILs whose COSMO-RS computed, or experimental solubility values were very close to zero. However, the predictions for few ILs are much higher compared to the experimental results (the points that are close to zero in x-axis, and higher than 10 in y-axis). These data points happen to be among those over-predicted by COSMO-RS and already discussed in previous section. From these results, it

can be concluded that, except for a small fraction of ILs (with no apparent similarity) for which the COSMO-RS computes unrealistically high solubility, the ML predication for the overwhelming majority of ILs are sufficiently accurate.

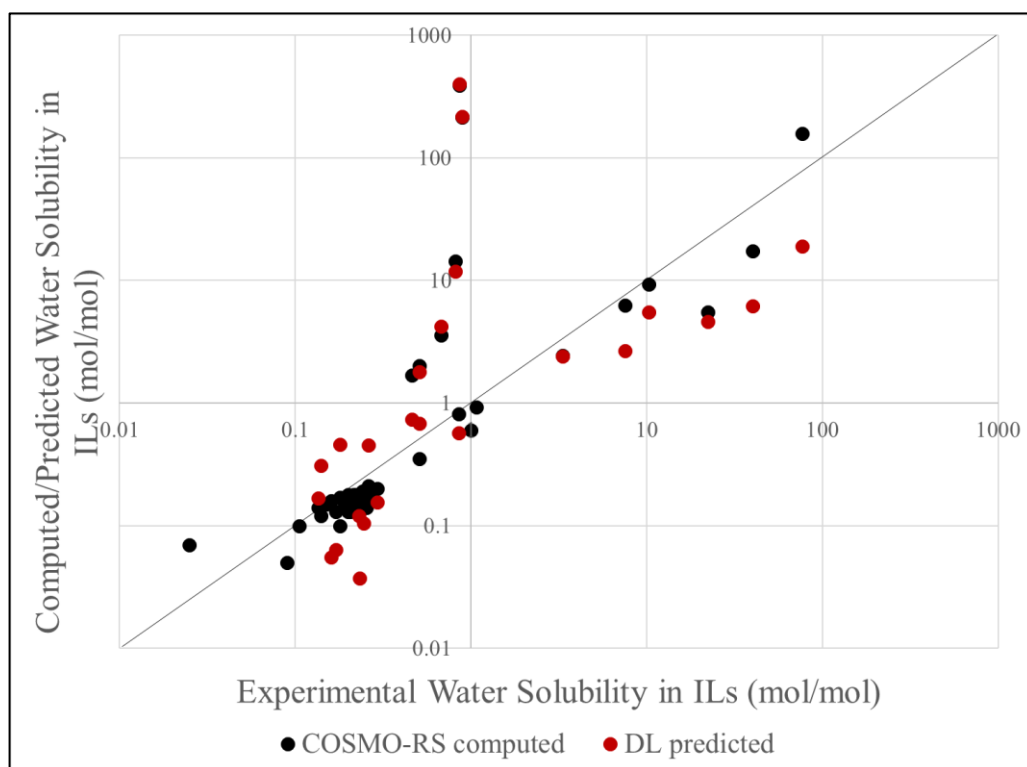


Figure 4.31. Experimental versus computed (by COSMO) and predicted (by multilayer fully connected network) solubility for various ILs. (see Appendix C for the experimental dataset).

4.4. Experimental Studies Using SrTiO_3 as Photocatalyst in PWS Reaction

The initial aim of our experimental work was to improve the performance of perovskites in PWS reaction with dye sensitization and then to increase the stability and further improve the activity of dye sensitized perovskite with ionic liquid coating. As the first attempt, the perovskite of SrTiO_3 was preferred because it is one of the most common perovskites used for PWS reaction; additionally, the physical and optical characteristics of SrTiO_3 can be easily tuned (see Section 3.1.1). The series of experiments were carried out by changing precursors and precursor compositions of SrTiO_3 , doping element, promoter, preparation method, and post-treatment temperature to determine the most appropriate strontium titanate form for dye sensitization and ionic liquid coating.

The commercial SrTiO_3 was tried in PWS reaction, but no hydrogen production was observed during the reaction. As the beginning, SrTiO_3 was prepared with the simplest and well-known method, which is solid state reaction (SSR). The reactants were kept at 1000°C to produce SrTiO_3 and 1%Pt (Pt nitrate salt as precursor) was impregnated onto the semiconductor and ionic liquid was deposited as explained in Section 3.1. At the first stage, no calcination was applied to the semiconductor between the steps of promoter impregnation and IL deposition. In those experiments, one ml of ionic liquid dissolved in five ml of dichloromethane or just one ml of ionic liquid (without DCM solvent) was impregnated onto five grams of semiconductor. The experiments were conducted in 20% (by volume) aqueous methanol reaction solution under solar light simulator. The results given in Figure 4.32 indicates that all ionic liquids ($[\text{BMIM}][\text{BF}_4]$, $[\text{BMIM}][\text{PF}_6]$, and $[\text{EMIM}][\text{N}(\text{Tf})_2]$) improved the photoactivity of semiconductor and the direct impregnation of IL (without the use of solvent DCM) appeared as more effective than impregnation of IL dissolved in DCM. Also, as can be seen from Figure 4.32, the activity of Pt/ SrTiO_3 photocatalyst (without any IL) starts to decrease after 3rd hour and it became zero at 5th or 6th hours of experiment; at the same time, dark spots were observed on the surface of reactor (shown in Figure 4.33). Since IL coated particles did not suffer from activity loss during PWS reaction, the ratio between volume of ionic liquid and weight of semiconductor was changed (ie. 0.75ml/5g, 1ml/5g, and 1.5ml/5g) and their photocatalytic performances were compared to observe whether the amount of IL is influential on that loss or not. In those experiments only $[\text{BMIM}][\text{BF}_4]$ was used as the representative of ionic liquids due to its better performance than the other ILs (see Figure 4.32). As can be seen from Figure 4.34, the best ratio between IL and SrTiO_3 is determined as 1ml over 5g to be used in the rest of SrTiO_3 experiments. The favorable impact of IL coating on Pt/ SrTiO_3 is understood from these experiments but the problem of sudden decrease in the photocatalytic performance of plain Pt/ SrTiO_3 could not be explained without additional experiments.

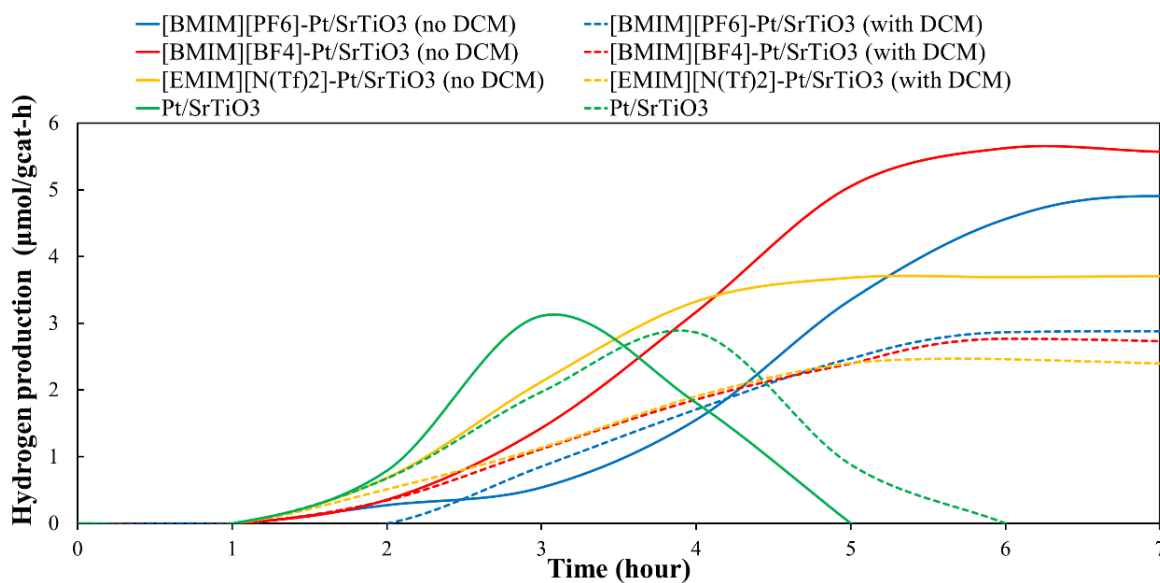


Figure 4.32. The hydrogen production vs time for ionic liquid coated and plain 1%Pt/SrTiO₃ (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).

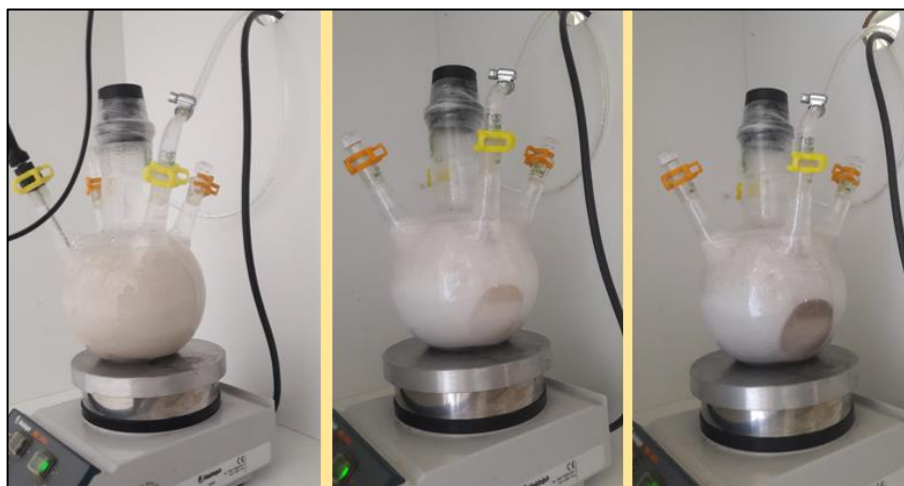


Figure 4.33. The images of darkening spot on the reactor through reaction.

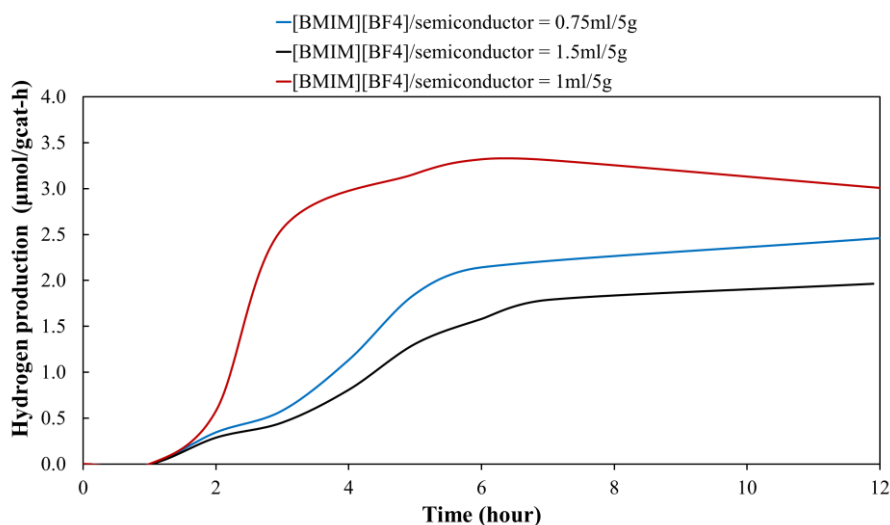


Figure 4.34. The effect of ratio between ionic liquid and semiconductor on hydrogen production (catalyst: Pt/SrTiO₃, ionic liquid; [BMIM][BF₄], preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).

4.4.1. Structural Modifications of SrTiO₃

One of the most common modifications is to use A and B precursors in an unbalanced molar ratio while producing A_xB_{2-x}O_{3±y} type perovskites, which can have different properties and photocatalytic performance than ABO₃. In the literature, it was seen that those modifications might be effective on visible light harvesting and photocatalytic activity of related semiconductor (Lee *et al.*, 2013; Wang *et al.*, 2012). So, the molar ratio between precursors (SrCO₃ and TiO₂) of SrTiO₃ was changed from 1:1 to 2:3 leading Sr₄Ti₆O_y semiconductor. The procedure for 1% Pt impregnation and reaction conditions were exactly the same with the previous experiments. The performances of Pt/Sr₄Ti₆O_y and Pt/SrTiO₃ are compared in Figure 4.35. Although there is a slight increment in the hydrogen production of Pt/Sr₄Ti₆O_y with respect to Pt/SrTiO₃, the problem of activity loss after 3rd hour continued.

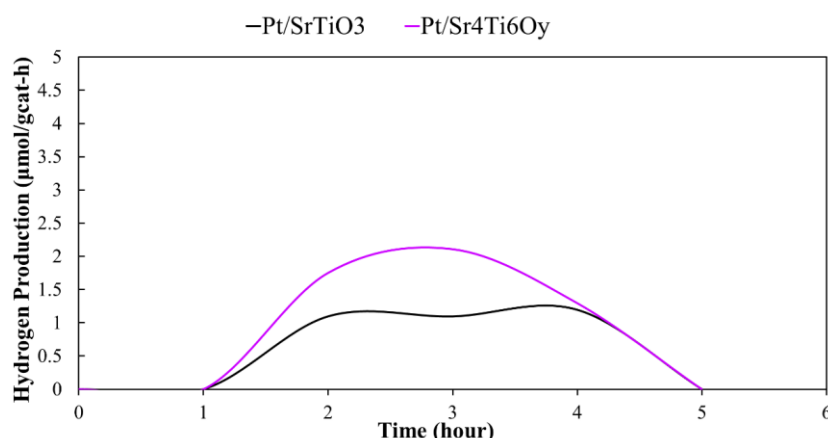


Figure 4.35. The comparison of Pt/Sr₄Ti₆O_y and Pt/SrTiO₃ semiconductors in PWS reaction (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, no calcination after impregnation).

Before continuing with ionic liquid coating and/or dye sensitization, the issue of unexpected activity loss for uncoated (in the absence of IL) strontium titanate particles was focused. To understand the reason of the problem several process variables such as quality and condition of precursors (i.e., purity or humidity and the brand of precursors), production method, and light source were reviewed, and several experiments were conducted by changing those variables. In the end, the problem was solved by applying post calcination to Pt/SrTiO₃ after Pt impregnation at 400°C for four hours; the results are shown in Figure 4.36. Since the problem was solved, the experiments were continued with another structure modification, which is element doping into the lattice structure of perovskite. The detailed review and statistical analysis of literature over perovskites in PWS has been already discussed. By considering the knowledge extracted from literature (see Section 3.1.2), the elements of La and Al, which are two of most common doping elements for A and B sites respectively, were determined as doping agents to SrTiO₃. Al or La doped strontium titanate particles were prepared by solid state reaction; required amount of Al₂O₃ or La₂O₃ (so as molar ratio between Al:Sr or La:Ti is 0.01:0.99) were mixed with the precursors of SrTiO₃ and heated in a muffle furnace. The performances of Pt/Sr₄Ti₆O_y, Pt/Al:Sr₄Ti₆O_y, and Pt/La:Sr₄Ti₆O_y photocatalysts are compared in Figure 4.37; Al or La doping failed to improve hydrogen production.

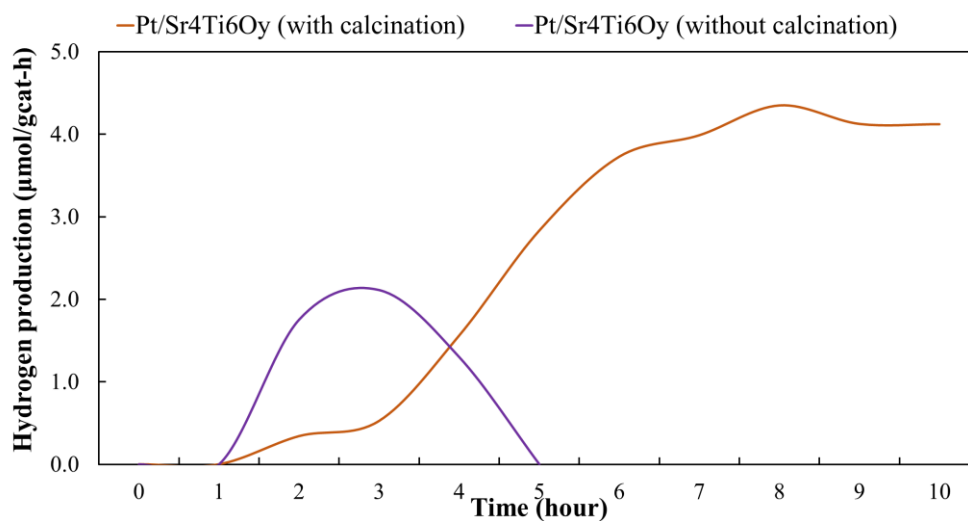


Figure 4.36. The performance of calcined and not calcined Pt/Sr₄Ti₆O_y semiconductors in PWS reaction (catalyst: Pt/SrTiO₃, preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt).

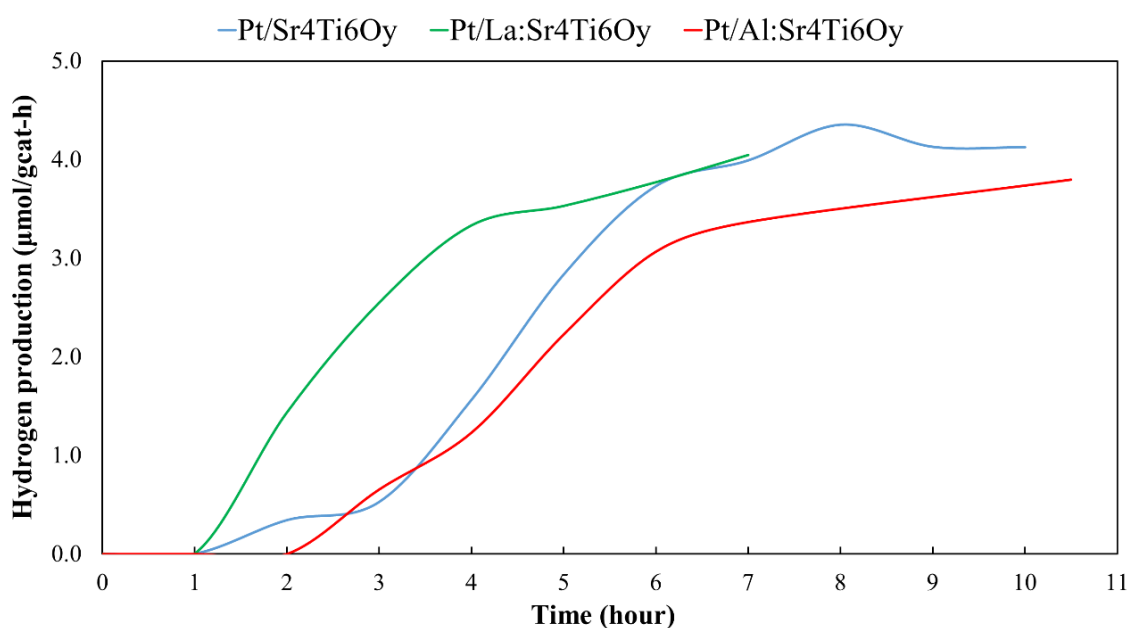


Figure 4.37. The performance of Pt/Sr₄Ti₆O_y, Pt/Al:Sr₄Ti₆O_y, and Pt/La:Sr₄Ti₆O_y semiconductors in PWS reaction (preparation method: SSR, preparation temperature: 1000°C, Pt precursor: Pt nitrate salt, calcination: 400°C and 4 hours).

It was decided to move on with undoped photocatalyst since doping has no positive effect on hydrogen production. On the other hand, the little improvement that was achieved

by changing precursor ratio of SrTiO_3 fell short of the expectations. As a result, since doping or changing precursor ratio are not as effective as expected, it was not preferred to take the risk of complications due to those modifications, so plain SrTiO_3 was determined as photocatalyst for following experiments.

4.4.2. The Effect of Synthesis Temperature for SrTiO_3

As the first step, the synthesis temperature of SrTiO_3 by solid state reaction was determined as 1000 C, which is the most common temperature in literature (Can and Yildirim, 2019b), but a higher synthesis temperature such as 1100°C was also tried to see how it effects hydrogen production (Ham *et al.*, 2016; Liu *et al.*, 2008). Figure 4.38. demonstrates the comparison of hydrogen production of semiconductors synthesized at two different solid state reaction temperatures. The hydrogen production is only slightly higher for semiconductor synthesized at 1100°C than that synthesized at 1000°C. Since 1100°C is too close to the maximum applicable temperature of muffle furnace in our laboratory. It was determined to continue with 1000°C as synthesis temperature of SrTiO_3 in the case of solid state reaction.

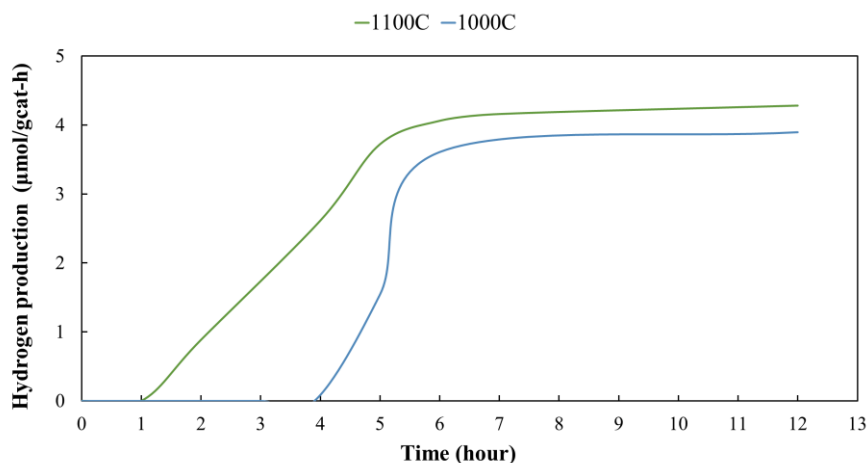


Figure 4.38. The performance of Pt/SrTiO_3 synthesized at two different temperatures (1000°C and 1100°C) in PWS reaction (catalyst: Pt/SrTiO_3 , preparation method: SSR, Pt precursor: Pt nitrate salt, calcination: 400°C and 4 hours).

4.4.3. The Effect of Calcination Temperature After Pt Impregnation onto SrTiO₃

As it was mentioned above, the problem of activity loss for plain Pt/SrTiO₃ has been solved by calcining Pt impregnated SrTiO₃ particles at 400°C for 4 hours; thus, the impact of calcination on hydrogen production is confirmed. By considering this, the effect of calcination temperature on the hydrogen production was also studied and several experiments were conducted in the presence or absence of ionic liquid coating. Figure 4.39. shows the hydrogen productions of ionic liquid coated and uncoated Pt/SrTiO₃ semiconductors calcined at four different conditions (no calcination, 200°C, 300°C, and 400°C). The optimum calcination temperature was decided as 400°C for Pt/SrTiO₃ semiconductor after those experiments. However, in contrast to the previous tests, it was observed that the ionic liquid photocatalysts performed worse than uncoated photocatalysts.

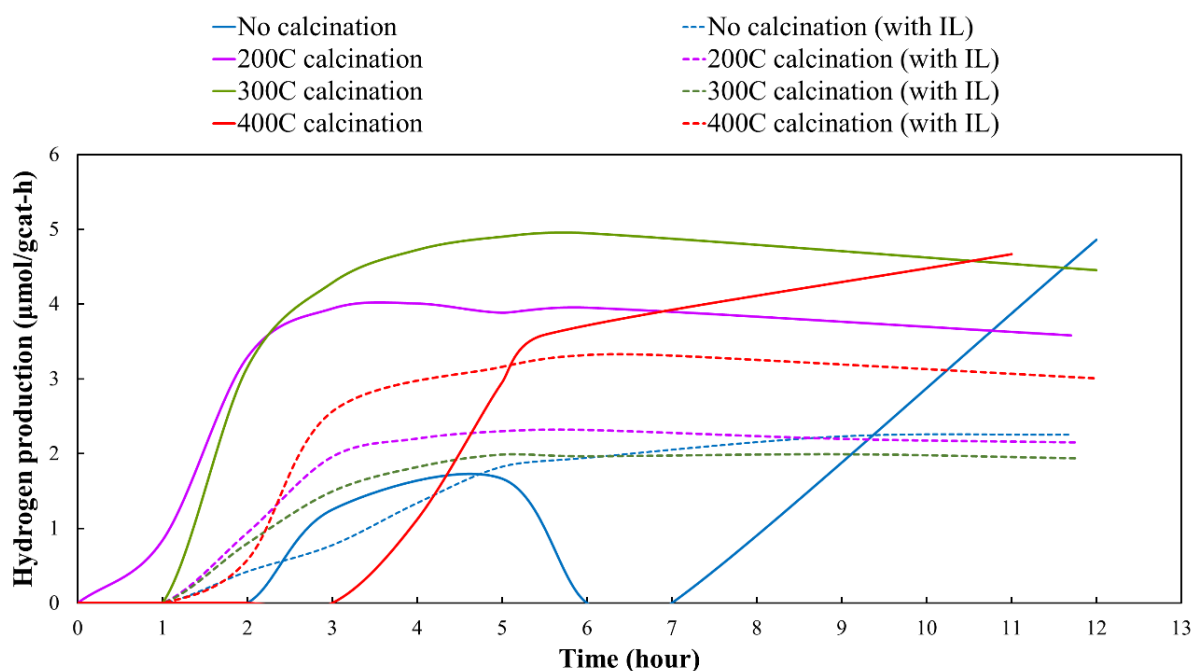


Figure 4.39. The performance of Pt/SrTiO₃ calcined at different conditions (catalyst: Pt/SrTiO₃, preparation method: SSR, synthesis temperature: 1000°C, Pt precursor: Pt nitrate salt, IL: [BMIM][BF₄]).

The experiments above were repeated by using Pt hydrate salt instead of Pt nitrate salt and the results are given in Figure 4.40. Since there is almost no difference between

performances of semiconductors impregnated with different Pt salts and Pt hydrate salt is more sensitive to ambient conditions (i.e., humidity), the Pt nitrate salt was used for impregnation for the rest of experiments.

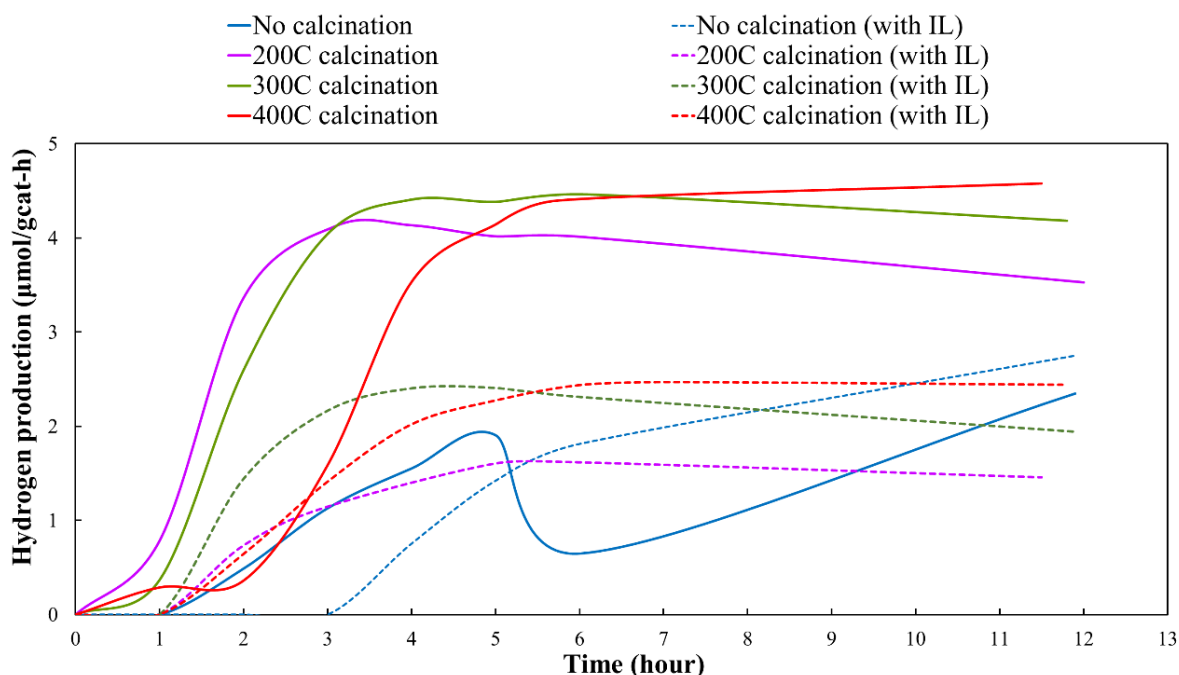


Figure 4.40. The performance of Pt/SrTiO₃ calcined at different conditions (catalyst: Pt/SrTiO₃, preparation method: SSR, synthesis temperature: 1000°C, Pt precursor: Pt hydrate salt, IL: [BMIM][BF₄]).

4.4.4. The PWS Experiments with Au as Promoter

The Au/SrTiO₃ semiconductor was already used in photocatalytic hydrogen production reaction and the details can be found in the study of Saadetnejad and Yildirim (Saadetnejad and Yildirim, 2018). As the last step of experiments related with promoter, the best performed IL, which was [BMIM][BF₄] coated onto 1%Au/SrTiO₃ and then the performance of semiconductor produced was tested in PWS reaction. The Au deposition procedure was imitated from the work of Saadetnejad and Yildirim. The results shown in Figure 4.41. indicates that Pt favors hydrogen production more than Au in the presence of IL.

After all these tests, Pt was selected as promoter, and the nitrate salt of Pt was found as more practical than hydrate due to its stability and resistances to environmental

conditions. Calcination after impregnation was proved as an obligatory step for Pt/SrTiO₃ photocatalyst and 400°C and 4 hours are determined as the best calcination temperature and time for Pt/SrTiO₃, respectively.

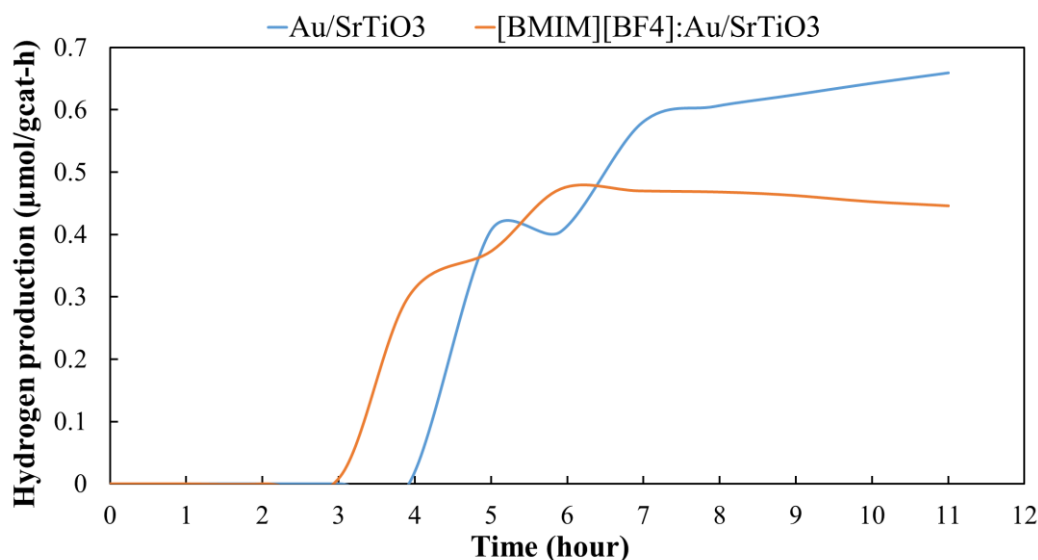


Figure 4.41. The performance of Au/SrTiO₃ and IL:Au/SrTiO₃ in PWS reaction (catalyst: Au/SrTiO₃, preparation method: SSR, synthesis temperature: 1000°C, Au precursor: Au chloride salt, IL: [BMIM][BF₄]).

4.4.5. The Effect of Production Method for SrTiO₃

Zwara *et al.* recently synthesized SrTiO₃ particles using hydrothermal method, and they investigated the effect of imidazolium ILs on the morphology of semiconductor (Zwara *et al.*, 2019). That procedure was repeated to produce Pt/SrTiO₃, and the performances of Pt/SrTiO₃ photocatalysts produced by SSR and this method are compared in Figure 4.42. Since a significant improvement was obtained with the use of new method, the effect of coating with ILs on PWS reaction was also examined and presented in Figure 4.43. It was observed that two ionic liquids of [BMIM][BF₄] and [EMIM][N(Tf)₂] favored hydrogen production when they were coated on Pt/SrTiO₃ (produced with hydrothermal method). Although the samples (with or without IL) synthesized by using hydrothermal method performed much better than those synthesized with solid state reaction, two important problems were encountered. One of them is the decrease of hydrogen production in the case of ionic liquid coated semiconductors (see Figure 4.43). The other important issue is that the

results (shown in Figure 4.43) were unrepeatable. Those experiments were repeated for several times under exactly same conditions, and it was expected to see the hydrogen production would be higher when [BMIM][BF₄] and [EMIM][N(Tf)₂] coated Pt/SrTiO₃ were used with respect to plain Pt/SrTiO₃. However, the ionic liquid coated particles, not only [BMIM][PF₆] (as in Figure 4.43) but also [BMIM][BF₄] and [EMIM][N(Tf)₂] coated particles, performed worse than plain Pt/SrTiO₃, which means the results shown in Figure 4.43 are unrepeatable.

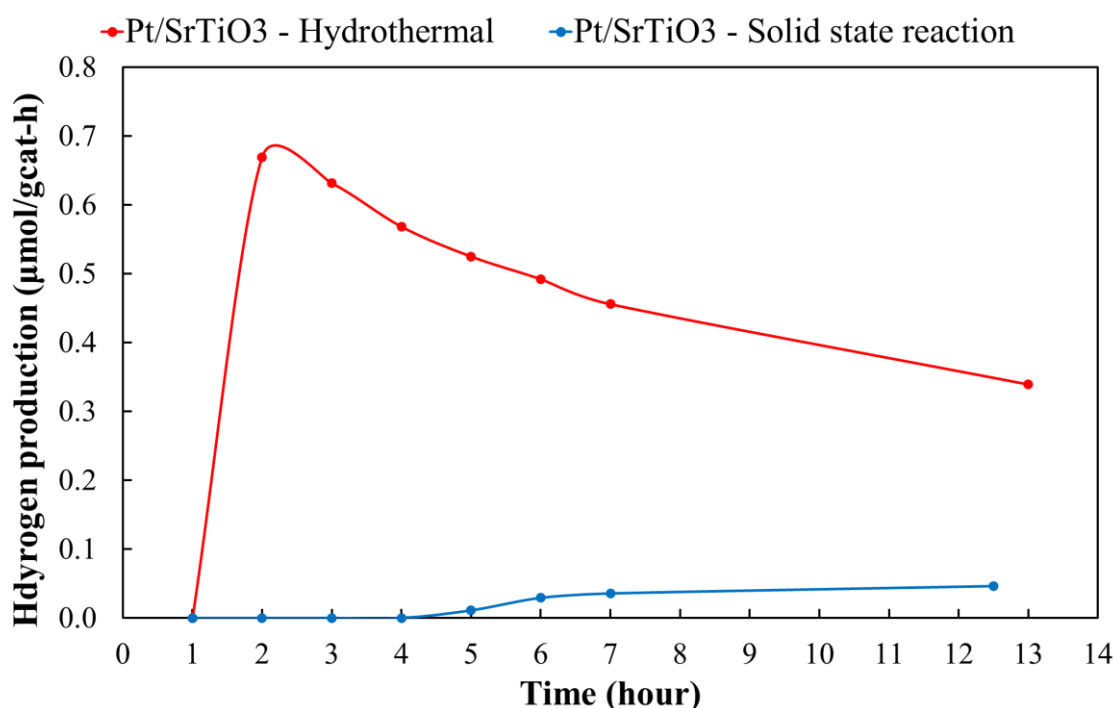


Figure 4.42. The performance of Pt/SrTiO₃ produced with two different methods (catalyst: Pt/SrTiO₃, SSR temperature: 1000°C, Pt precursor: Pt nitrate salt).

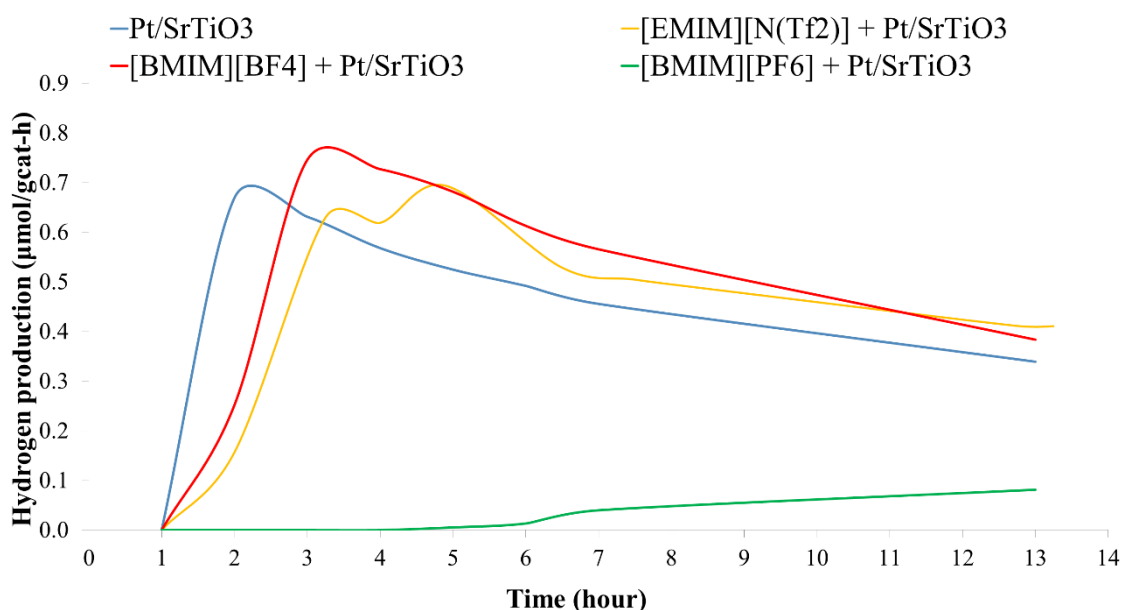


Figure 4.43. The performance of plain and IL coated Pt/SrTiO₃ in PWS reaction (catalyst: Pt/SrTiO₃, synthesis method: hydrothermal, Pt precursor: Pt nitrate salt).

In conclusion, the hydrogen production rate over SrTiO₃ photocatalysts could not be repeated at different times, and the causes of this problem could not be identified. Hence, the experiments involving dye sensitization and ionic liquid coating were continued with TiO₂, which is a well-known semiconductor in the field of photocatalysis.

4.4.6. The Characterization of SrTiO₃ Prepared by Solid State Reaction

The SrTiO₃ photocatalysts prepared by solid state reaction are characterized with SEM, UV-Vis, XRD, and FTIR methods. Since the photocatalysts prepared with hydrothermal method has a reproducibility problem, only samples (IL coated and uncoated) prepared by solid state reaction, whose performances are given in Figure 4.32 were characterized and discussed below.

Figure 4.44 shows SEM images of Pt/SrTiO₃ and [BMIM][BF₄]+Pt/SrTiO₃. In Figure 4.44a, the cubic particles may represent the unreacted precursor SrCO₃, and the spherical ones may come from the unreacted precursor TiO₂. The brighter and smaller particles in both images (Figure 4.44a and Figure 4.44b) represents Pt since it has higher atomic number than Sr and Ti metals; Pt region looks brighter in those SEM images.

In Figure 4.44b, which indicates SEM result of [BMIM][BF₄]+Pt/SrTiO₃, the images become blurred, and this may be attributed to the presence of IL layer on those particles; however, no information about IL layer thickness or clear evidence indicating presence of IL could be extracted from SEM results.

The band gap energies of Pt/SrTiO₃ and [BMIM][BF₄]+Pt/SrTiO₃ semiconductors are determined as 3.43 eV and 3.40 eV by UV-Vis spectrophotometry measurements. The [BMIM][PF₆]+Pt/SrTiO₃ and [EMIM][N(Tf)₂]+Pt/SrTiO₃ particles also have the same band gap with [BMIM][BF₄]+Pt/SrTiO₃ as shown in Figure 4.45. Hence, it can be concluded that ILs are not very influential on the band gap energies of SrTiO₃ semiconductor.

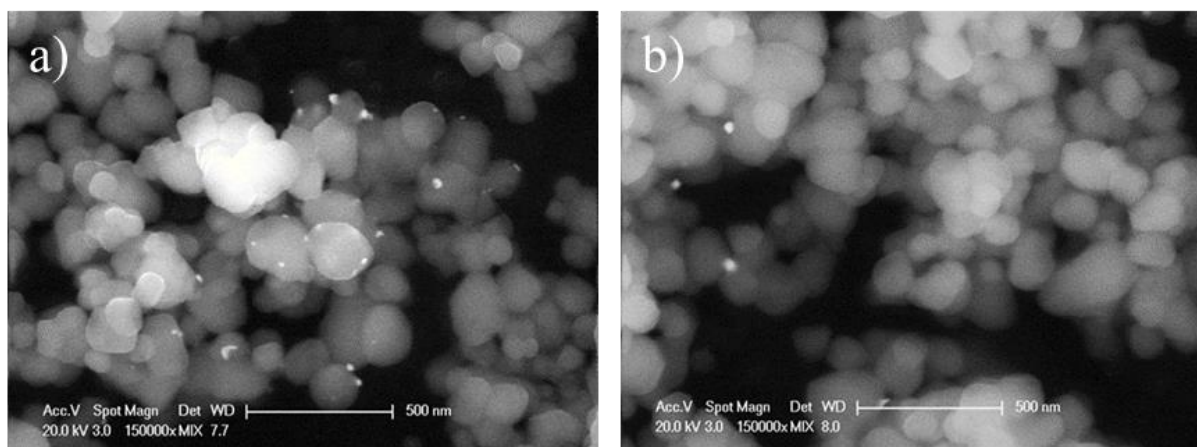


Figure 4.44. SEM images of SrTiO₃ photocatalysts produced with solid state reaction a) Pt/SrTiO₃ b) [BMIM][BF₄]+Pt/ SrTiO₃.

In FTIR analysis, Pt/SrTiO₃, [BMIM][BF₄]+Pt/SrTiO₃, and [EMIM][N(Tf)₂]+Pt/SrTiO₃ photocatalysts were investigated since [BMIM][BF₄] and [EMIM][N(Tf)₂] were the best performed ILs in PWS experiments. The FTIR analysis of empty sample containers were also performed, and their images were subtracted from the FTIR images of actual samples. As explained before, the particles studied in FTIR analysis are the photocatalysts used in the first experiments (for results see Figure 4.32) having stability problem; The primary aim of this analysis was to identify the activity loss of uncoated Pt/SrTiO₃ catalyst when compared to IL coated ones. Figure 4.46 demonstrates FTIR results of Pt/SrTiO₃, [EMIM][N(Tf)₂]+Pt/SrTiO₃, and [BMIM][BF₄]+Pt/ SrTiO₃ before and after reaction. In these figures, the specific peaks belong to SrTiO₃ or ILs are identifiable, but FTIR images

of IL coated particles (see Figure 4.46) are different before and after reaction indicating that IL could not hold on to the surface of semiconductor through the reaction. This was another reason (in addition to reproducibility problem) to switch the tests to TiO_2 as the semiconductor.

Figure 4.47a and 4.47b demonstrates XRD analyses of Pt/SrTiO_3 before and after the reaction. It can be deduced that the crystallography of Pt/SrTiO_3 did not change during reaction and the crystal structure of Pt/SrTiO_3 calculated by XRD is consistent with the literature (H. Yu *et al.*, 2011). There are peaks belong to TiO_2 and SrCO_3 because there may be unreacted precursors as observed by SEM analysis. In addition to the peaks of SrTiO_3 , TiO_2 , and SrCO_3 , there are also peaks that correspond to TiO_2 (rutile phase), SrCO_3 , $\text{Sr}(\text{OH})_2 \cdot \text{H}_2\text{O}$, Sr_2TiO_4 , and $\text{Sr}_4\text{Ti}_3\text{O}_{10}$. Pt impregnation did not cause any change in the crystal structure of SrTiO_3 as it was expected. By looking at Figure 4.47c it can be deduced that crystal structure of ILs coated Pt/SrTiO_3 catalysts are very similar to that of uncoated Pt/SrTiO_3 . Moreover, there is no obvious differences between the crystal structures of semiconductors before and after reaction. However, in Figure 4.47c there is a wider peak around 2θ equals to $10-20^\circ$, and that peak disappears in Figure 4.47d indicating the possibility for degradation of IL coating on the surfaces of semiconductor. That indication coincides with the results of FTIR analysis.

4.5. Experimental Studies Using TiO_2 as Photocatalyst in PWS Reaction

Up to this point, are two main challenges: the reproducibility problem in PWS reaction when SrTiO_3 is used as photocatalyst and uncertainty about the stability of IL coating on the SrTiO_3 surface. Those problems directed us to find a new photocatalyst material and methods for IL coating onto surface of a photocatalyst. Hence, a well-known semiconductor of TiO_2 was decided as photocatalyst to be used in following reactions to eliminate problems related with complicated perovskite structure.

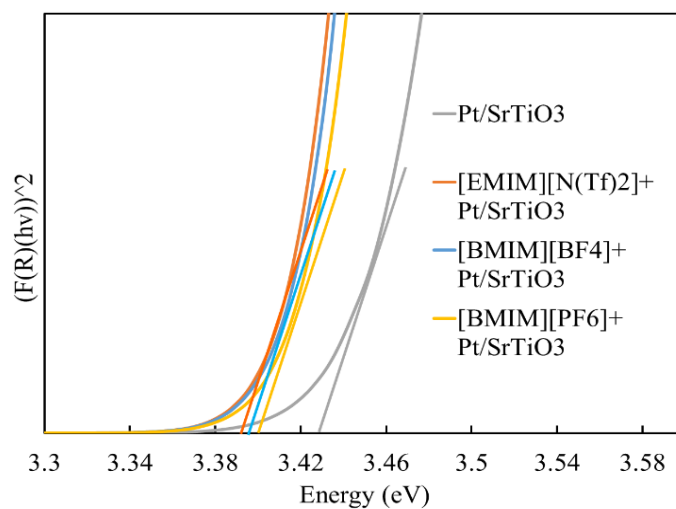


Figure 4.45. Band gaps of Pt/SrTiO₃, [EMIM][N(Tf)₂]+Pt/SrTiO₃, [BMIM][BF₄]+Pt/SrTiO₃, and [BMIM][PF₆]+Pt/SrTiO₃ (converted to absorbance using Kubelka-Munk function)

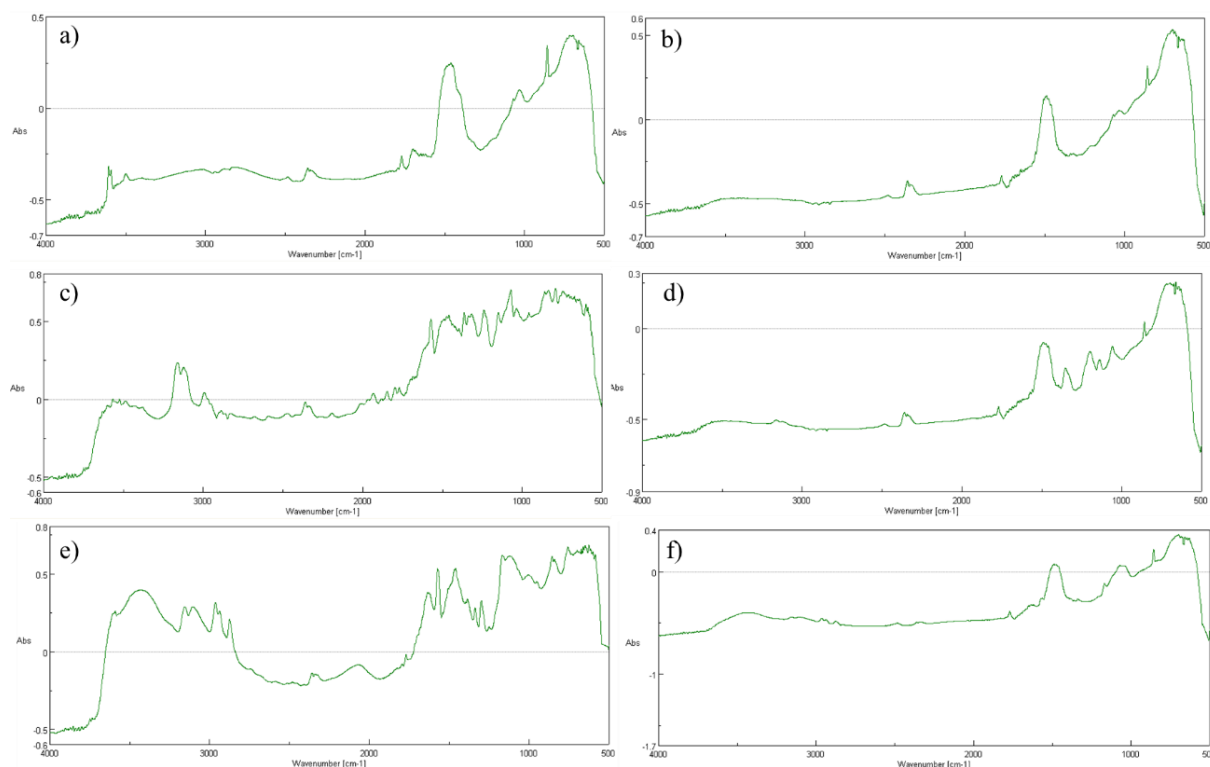


Figure 4.46. FTIR analysis of semiconductors before and after reaction for a-b) Pt/SrTiO₃, c-d) [EMIM][N(Tf)₂]+Pt/SrTiO₃, and e-f) [BMIM][BF₄]+Pt/SrTiO₃, respectively.

4.5.1. Physical and Chemical Characterization

The morphologies of Pt/TiO₂, Dye:Pt/TiO₂, and IL-Dye:Pt/TiO₂ particles were determined by using SEM-EDAX and Contrast Transmission Electron Microscopy (CTEM), which is more suitable for materials containing organic compounds as our last two samples. As can be seen from Figure 4.48a, which shows CTEM image of IL-Dye:Pt/TiO₂, the photocatalyst particles are almost spherical, and their size varies between 20 and 40 nm indicating that the shape of TiO₂ nanopowder was preserved through the catalyst preparation procedure involving Pt, dye loading and IL coating steps. Although it was not possible to detect the Pt particles conclusively due to the resolution limitation of CTEM, the small (less than 5 nm) and darker particles marked with purple circles, likely represent Pt considering its atomic density difference from the other ingredients and the contrast between the color of these particles and the others. The effects of the dye and ionic liquid could not be observed in the CTEM images. However, the SEM-EDAX image, presented in Figure 4.48b clearly indicates that both dye loading and IL coating on Pt/TiO₂ particles were successful; homogeneously dispersed F and Ru atoms are attributable to [BMIM][BF₄] ionic liquids and N719 dye respectively since they are the only sources for these elements in the medium. Pt seems to be dispersed well while its fraction found by EDAX analysis is also consistent with its loading.

The XRD patterns of TiO₂, Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, and IL-Dye:Pt/TiO₂ presented in Figure 4.49 indicate that the phase structure of TiO₂ did not change during the photocatalyst preparation as the patterns for all samples mentioned above are essentially the same. The peaks at 25.3°, 37.9°, 49.0°, 54.0°, and 55.1° correspond to the anatase phase (JCPDS Card no.99-0008) while the peaks at 27.4°, 54.4°, and 62.8° refer to the rutile phase (JCPDS Card no.87-0710) indicating Pt/TiO₂ involves both anatase and rutile phase as expected because the nanopowder TiO₂ was in P25 form. This is an expected result because the calcination temperature of 400° C (which was applied after Pt impregnation) was not sufficiently high to change the phase structure of TiO₂ (Mohamed *et al.*, 2014; Wetchakun *et al.*, 2012) while dye sensitization and IL coating were not expected to have any impact as well.

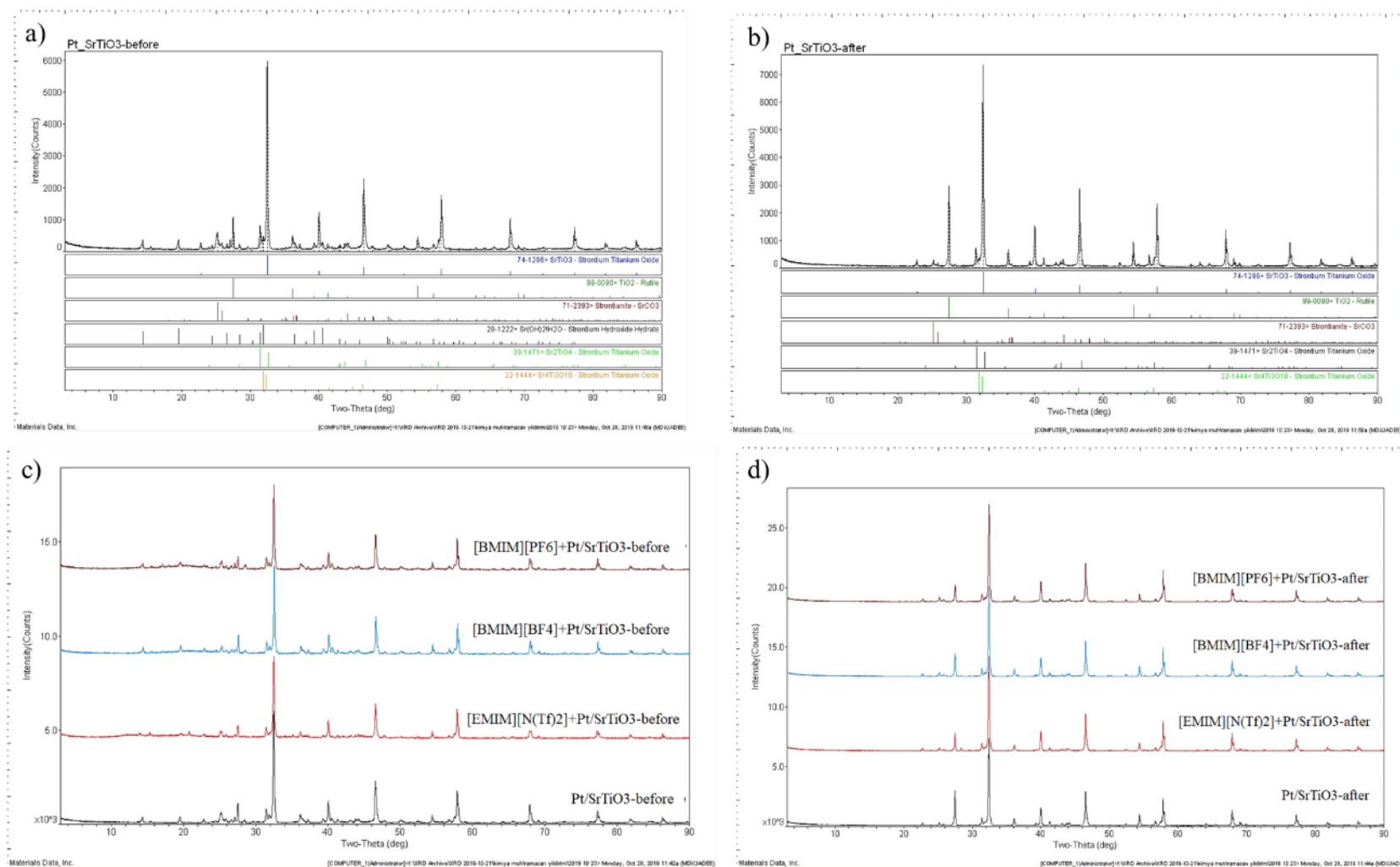


Figure 4.47. XRD analysis of semiconductors a) Pt/SrTiO₃ before reaction, b) Pt/SrTiO₃ after reaction, c) IL coated semiconductors before reaction, and d) IL coated semiconductors after reaction

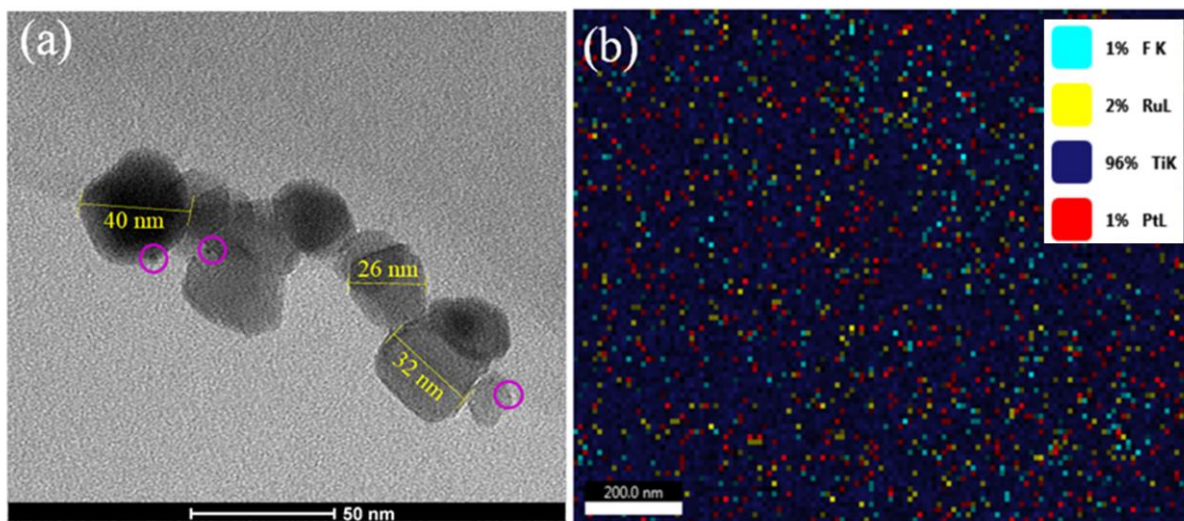


Figure 4.48. a) CTEM of IL-Dye:Pt/TiO₂, b) SEM-EDAX analysis of IL-Dye:Pt/TiO₂.

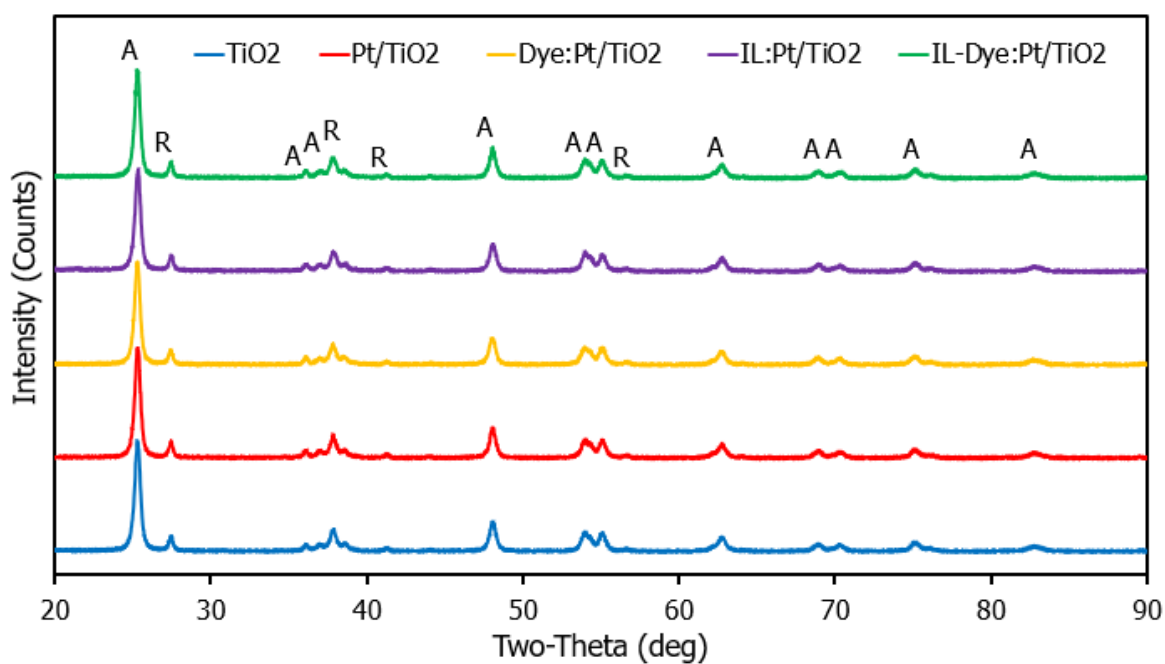


Figure 4.49. XRD of TiO₂, Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂ and IL-Dye:Pt/TiO₂.

The UV-visible absorbance spectrum of the dye (N719) is presented in Figure 4.50a while UV-Vis absorbance spectra expressed as Kubelka-Munk unit of TiO₂, Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, and IL-Dye:Pt/TiO₂ are given in Figure 4.50b; the band gap energies of all four photocatalysts are determined by using Kubelka-Munk function as shown

in Figure 4.50c. Figure 4.50a shows that the maximum absorption wavelengths of N719 dye are 320, 485 and 540 nm, which are in visible region consistent with the literature (Toor *et al.*, 2016; Xiaohu *et al.*, 2012). As can be seen from Figure 4.50b, the absorbance by the photocatalysts increases with the use of dye even though the patterns in pure N719 spectrum are not clearly observable probably due to the low dye concentration. Only a small peak at 500-600 nm is noticeable for the dye containing photocatalysts (Dye:Pt/TiO₂, and IL-Dye:Pt/TiO₂). The band gap value of 3.56 eV for pure TiO₂ (without any treatment) decreased to 3.45 eV after Pt impregnation and measured as 3.49 eV after dye sensitization; this slight increase after dye loading may be due to experimental error considering that Pt/TiO₂, IL:Pt/TiO₂ and IL-Dye:Pt/TiO₂ have the same band gap value of 3.45 eV (Al-Shomar, 2020). We concluded from these measurements that dye sensitization and IL coating does not change the band gap of Pt/TiO₂ (Mora-Seró *et al.*, 2010), while Pt impregnation lowers it slightly. However, Figure 4.50b depicts that, as expected, the absorbance was higher for Dye:Pt/TiO₂ while IL coating decreased it slightly.

FTIR spectra of the synthesized photocatalysts are given in Figure 4.51. The FTIR spectra of TiO₂ and Pt/TiO₂ particles are nearly the same indicating that Pt impregnation did not affect the physical state of TiO₂ as it was expected (Ahmed *et al.*, 2014; Wang *et al.*, 2020). The broad peak between 500 and 1000 cm⁻¹ is due to the Ti-O and Ti-O-Ti stretching modes (Kuvarega *et al.*, 2014; Yu *et al.*, 2006). Peaks at 1630 cm⁻¹ and the wider peak around 3300 cm⁻¹ were attributed to the H-OH bending and the O-H stretching of the hydroxyl group, respectively. Those hydroxyl groups on the surface of P25, and water adsorbed by the surface enhance the photocatalytic water splitting since they prevent recombination reaction and promote charge transfer by acting as molecular adsorption and hole scavenging sites (Ao *et al.*, 2009; Du *et al.*, 2008; Kuvarega *et al.*, 2014; Yu *et al.*, 2007; Zhou *et al.*, 2012); the peaks at 1630 cm⁻¹ and 3300 cm⁻¹ are also apparent in other photocatalysts.

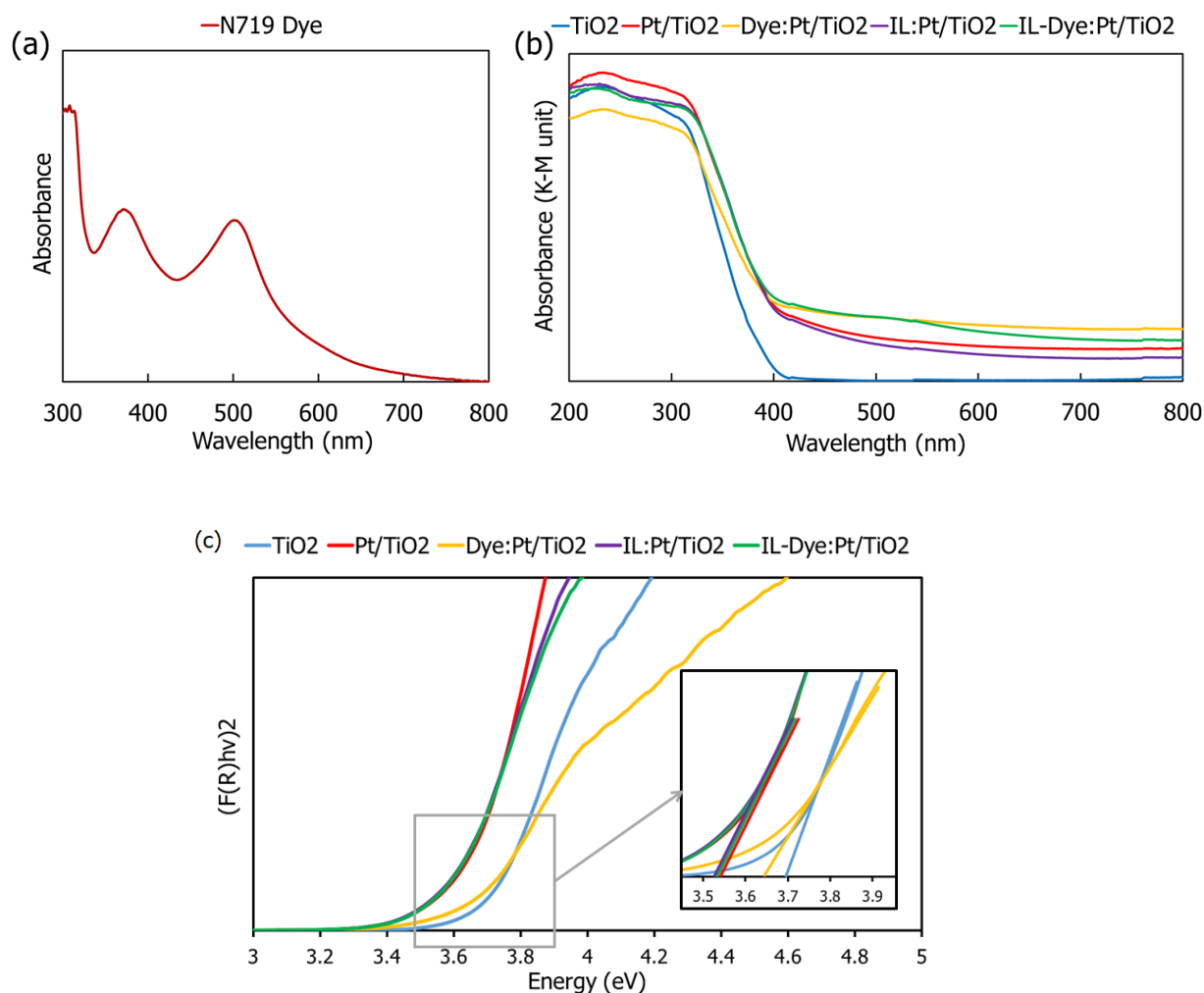


Figure 4.50. a) UV-Vis absorbance spectra of N719 dye, b) Kubelka-Munk diffuse reflectance absorption spectra of TiO₂, Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, IL-Dye:Pt/TiO₂, c) Band gap energies TiO₂, Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, IL-Dye:Pt/TiO₂.

The peaks appear between 1400 and 1700 cm⁻¹ in the spectrum of Dye:Pt/TiO₂ (but not in the spectrum of plain P25) correspond to the stretching C=O groups of carboxylic acid in dye molecule (Arifin *et al.*, 2017; Wei *et al.*, 2015). The presence of tetrabutylammonium (TBA) counter-ions and carboxylate groups (C-O) in N719 dye also reveal themselves with the peaks around 1550 and 1650 cm⁻¹. The characteristic peak at 2104 cm⁻¹ representing thiocyanato (NCS group) and several peaks between 1250 and 1750 cm⁻¹ belonging to carboxylic acid group in the spectrum of N719 did not appear in the spectra of dye containing photocatalysts indicating that these groups enabled the chemisorption of N719 to the surface of TiO₂ (Cai *et al.*, 2015; L. Wei *et al.*, 2015). Similarly, all characteristic sharp peaks of N719 between 1400 and 1700 cm⁻¹ become broader when dye anchored to TiO₂ surface (Cai

et al., 2015). In the spectrum of IL coated samples, the broad peak in the range of 3130 – 3200 cm^{-1} is due to quaternary amine salt formation with tetrafluoroborate, which is the anion part ($[\text{BF}_4]^-$) of ionic liquid used for coating (Karakaya and Uzun, 2016); the peaks at 1450, 2890, and 2970 cm^{-1} belong to the aliphatic C-H stretching in the cation part ($[\text{BMIM}]^+$) of IL while peaks at 1570 and 1170 cm^{-1} are for C-N stretching and C-H bending vibration (W. Sun and Xu, 2016). Finally, the wider peak between 970 and 1000 cm^{-1} corresponds to the methyl groups in the cation part of $[\text{BMIM}][\text{BF}_4]$. The IL coating seems to dominate the spectrum of IL-Dye:Pt/TiO₂ photocatalyst and make the signs of dye unobservable as the spectrum of IL:Pt/TiO₂ and IL-Dye:Pt/TiO₂ are nearly the same.

4.5.2. Photocatalytic Activity of Pt/TiO₂, Dye:Pt/TiO₂, IL-Dye:Pt/TiO₂

In Figure 4.52, the hydrogen production rates over Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, and IL-Dye:Pt/TiO₂ are given as a function of time. The catalyst weight to reaction solution ratio was kept constant at 2 grams of catalyst (gcat) per liter of solution in all experiments; 10%

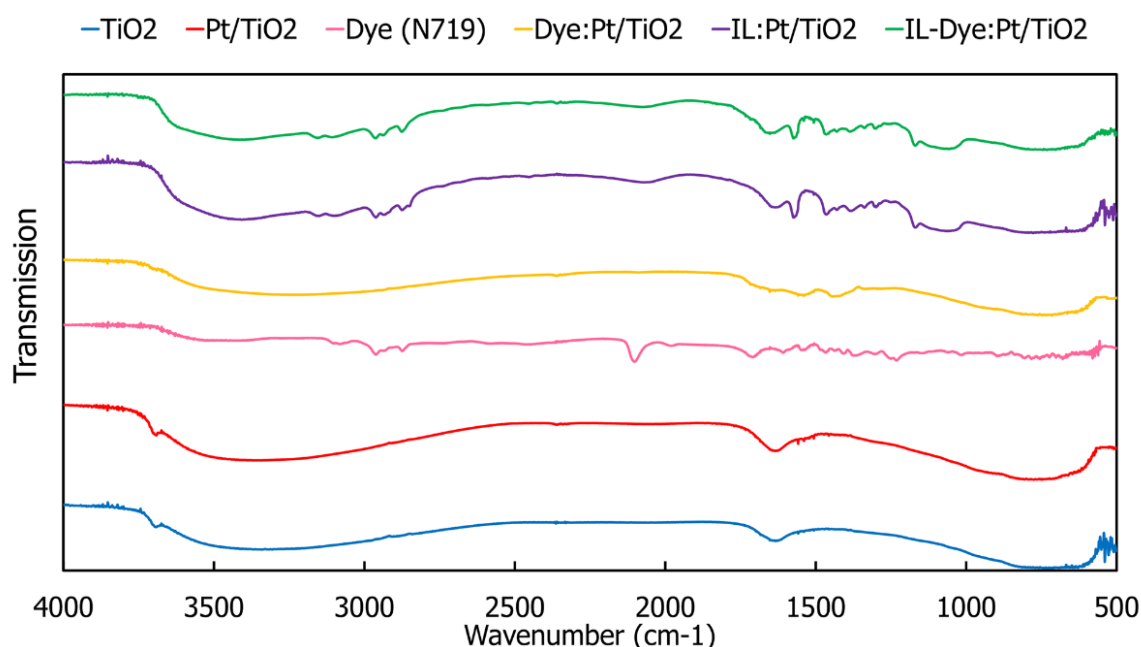


Figure 4.51. FTIR of (blue) TiO₂, (red) Pt/TiO₂, (pink) Dye, (yellow) Dye:Pt/TiO₂, (purple) IL:Pt/TiO₂, and (green) IL-Dye:Pt/TiO₂.

TEOA was also present in the water as the sacrificial donor. The weight contribution of the dye loaded was negligible; hence, it was not counted during reaction. On the other hand, the weight ratio of ionic liquid to TiO_2 was 1:5 (corresponding to 0.4 g/L), which is not negligible if IL is contributing to the function of TiO_2 . Hence, the photocatalytic activity of IL:Pt/ Al_2O_3 was also tested under the same conditions; since Al_2O_3 is not an active semiconductor, any activity observed would be attributed to IL ([BMIM][BF₄]). However, no hydrogen production was observed; therefore, the weight of IL was not taken into account in calculating the amount of catalyst. Only the weight of Pt/ TiO_2 was considered and kept constant at 2 g/L in all experiments regardless of the presence or absence of the dye or ionic liquid (i.e., the hydrogen production rates were computed as $\mu\text{mol}/\text{hour per gram } 1\% \text{ Pt}/\text{TiO}_2$) so that the results obtained under different formulations would be comparable.

As seen in Figure 4.52, hydrogen production first increases with time and subsequently stabilizes after the four hours. In the stabilized rate region, the lowest (about 20-22 $\mu\text{mol}/\text{h.gcat}$) hydrogen production was observed over Pt/ TiO_2 ; this value is low compared to the results reported in literature under UV because we used visible light solar simulator to be able to see the use of ionic liquid in more realistic conditions. The rate slightly increased (about 20%) with dye sensitization (Dye:Pt/ TiO_2) while the IL coating without dye (IL:Pt/ TiO_2 ,) resulted in much higher enhancement (about three times) reaching to the production rate of about 60 $\mu\text{mol}/\text{h.gcat}$. The highest performance was observed with the IL coating of the dye-sensitized photocatalyst (IL-Dye:Pt/ TiO_2) with a further 20% increase (exceeding 70 $\mu\text{mol}/\text{h.gcat}$). Generally speaking, the N719 dye sensitization favors hydrogen production because it has a wider absorption spectrum in visible range than TiO_2 leading more photon absorption (this is also evident from Figure 4.50b). Additionally, the LUMO energy of N719 is more negative than that of TiO_2 while its HOMO level is lower (Toyao *et al.*, 2013); this enables easy transfer of photogenerated electrons from N719 dye to the TiO_2 enhancing the photocatalytic activity. This explains the small increase in hydrogen production in Figure 4.52 even though the enhancement is much less than to be significant (or expected) (Chen *et al.*, 2021; Latorre-Sánchez *et al.*, 2012; Peng *et al.*, 2008). It is likely that the dye molecules were removed from the surface as the result of stirring, and they could not contribute as much as they normally would as it was evident from the tests involving IL coated catalyst.

Figure 4.52 also shows that IL coating of the Pt/TiO₂ catalyst resulted in much higher enhancement in hydrogen production. As discussed in Section 3.1, neither the crystal structure nor the band gap of Pt/TiO₂ were affected by the IL coating (Figure 4.49 and 4.50); hence, IL coating could not contribute to the photocatalytic activity; the experiments with IL:Pt/Al₂O₃ mentioned above indicate that IL could not contribute to the role of semiconductor either. Additionally, we tested the performance of Pt/TiO₂ with the same amount of IL added to the reaction solution instead of coating over the photocatalyst to see whether the IL acts as an electron donor as well. No improvement was observed in the performance of Pt/TiO₂, ruling out this possibility. Hence, we concluded that the performance enhancement with IL coating should be the result of its ability to facilitate the charge transfer between photocatalyst and aqueous solution, and from its ion pairing structure, which inhibits the recombination of photogenerated electron-hole pairs (Gusain *et al.*, 2016; Jing *et al.*, 2015; Opallo and Lesniewski, 2011; Safavi *et al.*, 2006; Zwara *et al.*, 2019). The maximum enhancement was obtained when the photocatalyst was coated with IL after dye sensitization. This indicates that the IL keeps the dye molecules on the TiO₂ surface enabling them to sensitize the photocatalyst more effectively while making its own contribution as discussed above. The improvement provided by the dye sensitization can be still considered low (compared to the effect of IL). However, our primary focus was to demonstrate the effect of IL encapsulation; we did not attempt to optimize the sensitization conditions (such as the use of other dyes, testing different dye concentrations or testing different sacrificial agents) that may result in better photocatalytic performances.

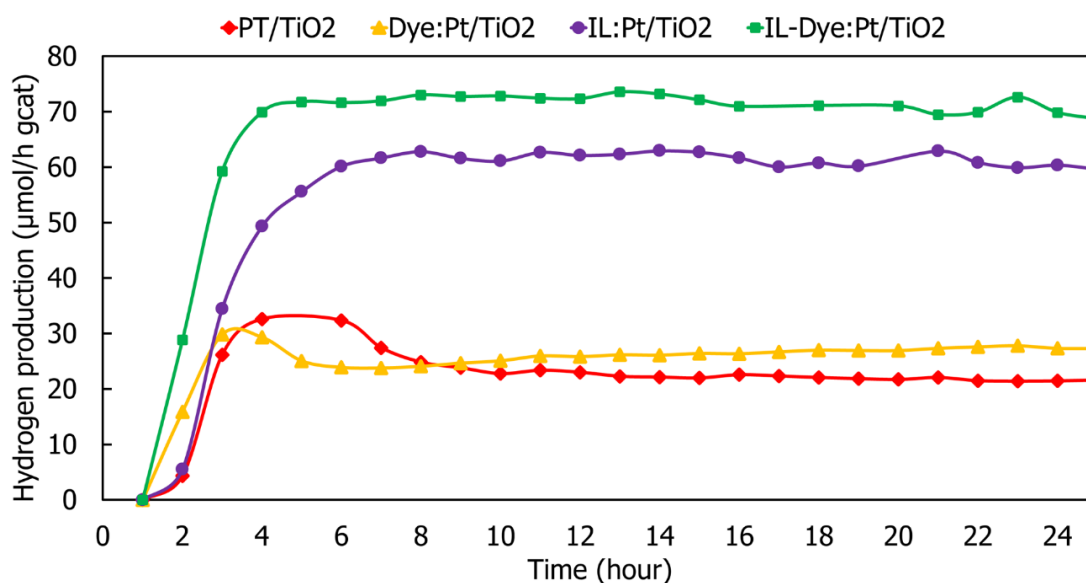


Figure 4.52. Hydrogen production of Pt/TiO₂, Dye:Pt/TiO₂, IL:Pt/TiO₂, and IL-Dye:Pt/TiO₂ under 300W solar light illumination in TEOA solution 10% (600 ml reaction solution, 2g active catalyst/L).

As it is also seen from Figure 4.52, the hydrogen production rate increased sharply at the beginning when Pt/TiO₂ or IL:Pt/TiO₂ were used as catalysts while it started to increase with a concave upward trend when Dye:Pt/TiO₂ or IL-Dye:Pt/TiO₂ were used. This is attributable to the fact that Pt nanoparticles absorb oxygen from air easily after impregnation process delaying the activation of catalyst to produce hydrogen about one hour. Such delay was not observed in dye containing photocatalysts because the hydrophilic N719 dye provides better interaction of the photocatalyst with the aqueous reaction solution (Manfredi *et al.*, 2016).

The performance of IL-Dye:Pt/TiO₂ photocatalyst was also monitored for 48 hours as an initial indicator of its stability even though neither this period was sufficient nor the other conditions were suitable for a realistic stability test. As demonstrated in Figure 4.53, the activity of that photocatalyst slightly decreased (especially after 35 hours) indicating that the viscous hydrophobic ionic liquid coating of dye-sensitized Pt/TiO₂ has a potential to improve the photocatalytic activity in a stable manner.

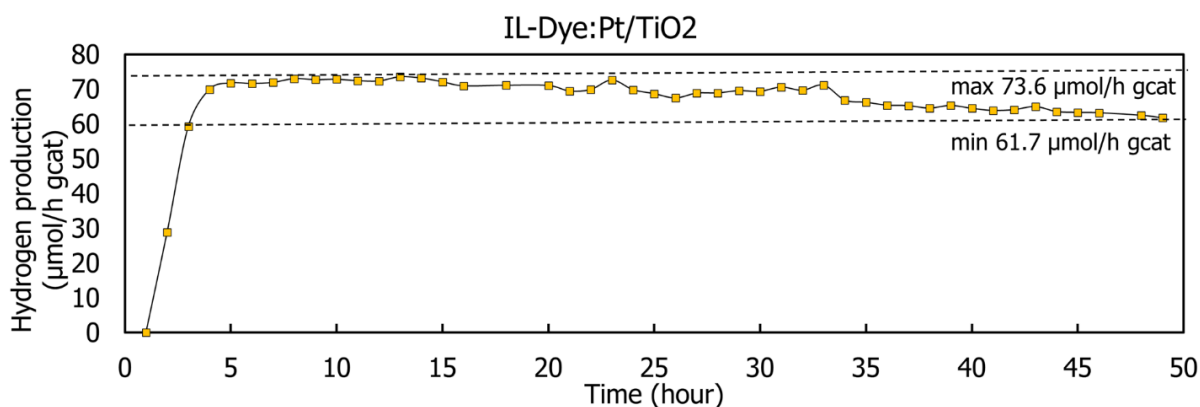


Figure 4.53. Stability test for hydrogen production of IL-Dye:Pt/TiO₂ under 300W solar light illumination in TEOA solution 10% (600 ml reaction solution, 2g active catalyst/L).

4.5.3. Electrochemical Analysis of TiO₂, IL:TiO₂ and IL-Dye:TiO₂ Photoelectrodes

The performance of photoelectrodes can be evaluated by measuring the photocurrent density and photoconversion efficiency as the photocurrent is proportional to the hydrogen/oxygen evolution rate (Jiang *et al.*, 2018; Liu *et al.*, 2013; Shi *et al.*, 2015). Photoelectrochemical measurements were performed on TiO₂, IL:TiO₂ and IL-Dye:TiO₂ photoelectrodes to further elucidate the role of the IL on photocatalytic activity. TiO₂ and IL:TiO₂ photoelectrodes were stable, and no degradation of the electrodes was observed in long times (three sets of five CV cycles with ~30 min waiting time between the sets). Meanwhile, the IL-Dye:TiO₂ photoelectrode exhibited some degradation after each test. Figure 4.54a shows the cyclic voltammetry profiles of the photoelectrodes in the third CV cycle of the third set. Although some degradation occurred in IL-Dye:TiO₂ photoelectrode, it exhibited the highest anodic vertex current density at 1 V (3.3 mA/cm²).

Figure 4.54b depicts the linear sweep voltammograms of the photoelectrodes in the range -0.95 V to 0.6 V. The current density increased significantly in the presence of the ionic liquid coating; and was further enhanced when the TiO₂ photoelectrode was coated both by the dye sensitizer and the ionic liquid. We could not properly measure the performance of the TiO₂ photoelectrode coated only with dye due to the fast deterioration of the dye layer in the absence of the IL. IL:TiO₂ exhibited a photocurrent density of 0.99 mA/cm² at 0.5 V while IL-Dye:TiO₂ was at 1.14 mA/cm² compared to 0.33 mA/cm² obtained with TiO₂ electrode at the same conditions. We also calculated the photoconversion efficiencies of the photoelectrodes using

$$\eta\% = \left[\frac{J_P(E_{rev}^0 - E_{Bias})}{I_0} \right] \times 100. \quad (4.1)$$

where J_P is the photocurrent density (mA/cm^2), I_0 is the incident light ($100 \text{ mW}/\text{cm}^2$), E_{rev}^0 is the standard reversible potential for water splitting (1.23 V) and E_{Bias} is the applied potential given by $E_{meas} - E_{oc}$ where E_{meas} and E_{oc} are the electrode potential (vs SCE) and the open circuit potential (vs SCE), respectively. As given in Figure 4.54c, the maximum photoconversion efficiency of 0.74% at -0.72V (vs SCE) was reached with the dye-sensitized and IL coated TiO_2 photoanode (IL-Dye: TiO_2). The maximum photoconversion efficiencies of IL- TiO_2 and TiO_2 photoanodes were measured as 0.61% and 0.35% at -0.7 V (vs SCE) and -0.58 V (vs SCE), respectively. These results are in good agreement with the results and discussions presented in Section 3.2 regarding the performance improvements provided by the ionic liquid ([BMIM][BF₄]) coating, which enhanced the charge transfer between photoelectrode and electrolyte, and inhibited the recombination of photogenerated electron-hole pairs. The current density and efficiency of the IL-Dye: TiO_2 photoelectrode was further enhanced due to the light harvesting capability of the dye-sensitized photoelectrode.

4.6. Study Of Critical Factors Affecting the Hydrogen Evolution Via Mass Transfer Limitations

In photocatalytic reactions, determining the amount of energy adsorbed and utilized by the photocatalyst and calculating the efficiency or any other performance measures in comparable manner are quite difficult for various reasons. To begin with, most of the time the description of the light sources such as *300 W xenon arc lamp* is not sufficient because the fraction of UV photons emitted from such lamp changes even among the models produced by the same company. A certified solar simulator may solve the problem related the source; however, the differences in distance and angle of light as well as the reflection, transmission and absorption characteristics of reactor materials and reaction medium will affect the amount of energy reaching to the catalyst surface.

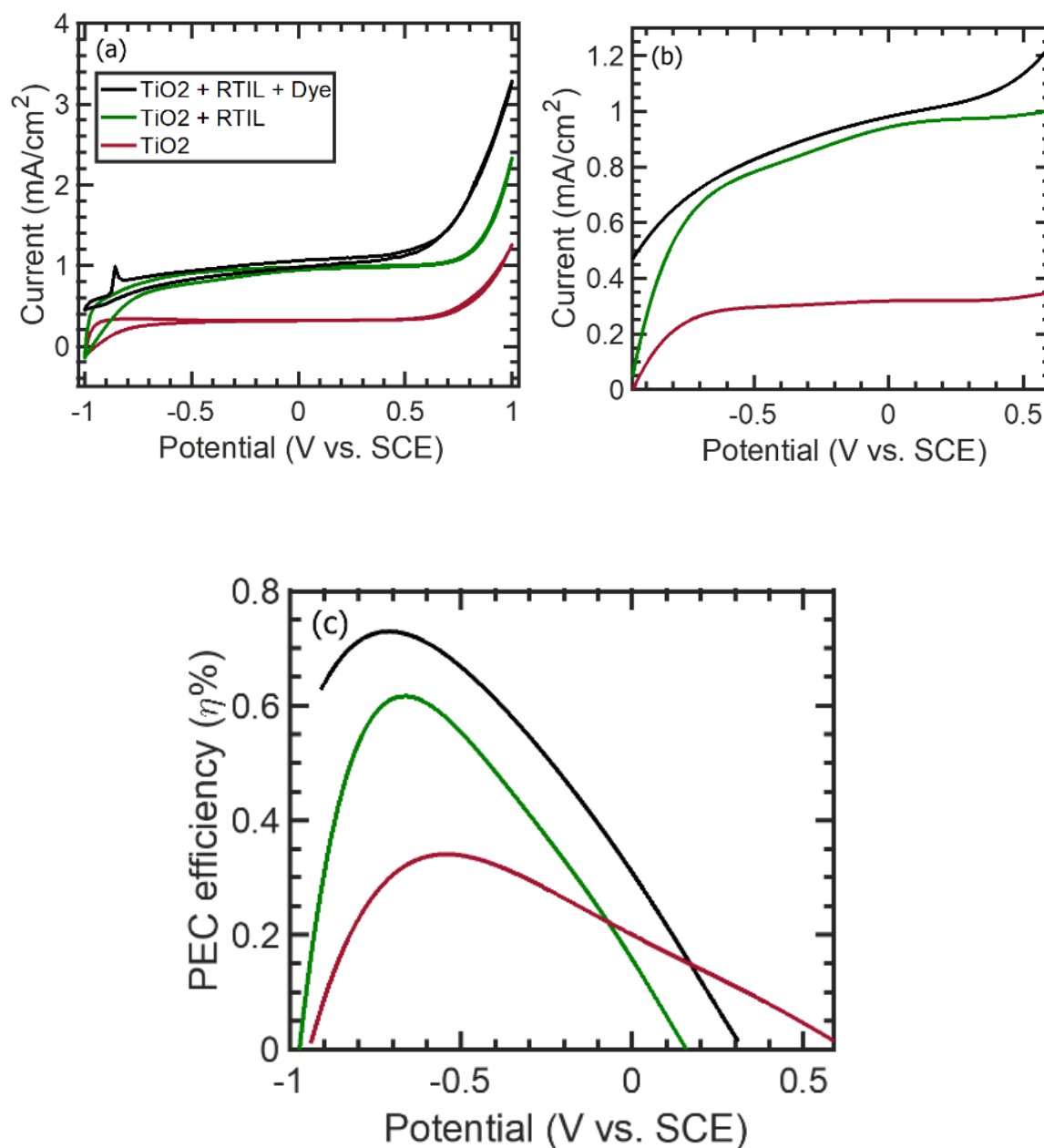


Figure 4.54. a) Photoelectrochemical profile of the TiO₂, IL:TiO₂ and IL-Dye:TiO₂ photoelectrodes (cyclic voltammetry profiles are recorded in a TEOA/Na₂SO₃ aqueous solution), b) variation of the photocurrent density as a function of measured potential, c) photoconversion efficiency as a function of measured potential.

Researchers generally agree that the quantum yield, which is an indicator of the number of photons absorbed by the catalyst and/or the photonic yield, which is related to the number of photons reached to the internal surface of photocatalyst, should be measured to evaluate the true performance of photocatalytic reactions (Braslavsky *et al.*, 2011; Buriak *et*

et al., 2014; Kisch and Bahnemann, 2015; Melchionna and Fornasiero, 2020; Qureshi and Takanabe, 2017; Rajeshwar *et al.*, 2015; Serpone *et al.*, 2016). However, it is not easy to calculate these measures accurately for several reasons. For example, the scattering of the light is a major problem because it has highly negative impact on the number of photons reached on the catalyst surface, therefore on the photonic yield (Braslavsky *et al.*, 2011; Kisch and Bahnemann, 2015; Kočí *et al.*, 2008; Rajeshwar *et al.*, 2015); consequently, the reactor geometry (Serpone *et al.*, 1996), catalyst loading (Buriak *et al.*, 2014; Kočí *et al.*, 2008; Melchionna and Fornasiero, 2020; Serpone *et al.*, 1996), catalyst surface area (Braslavsky *et al.*, 2011; Buriak *et al.*, 2014; Dong *et al.*, 2009; He and Que, 2013; He *et al.*, 2011; Ho *et al.*, 2006; Peng *et al.*, 2010), particle radius or thickness of catalyst film (Ballari *et al.*, 2010, 2008; Buriak *et al.*, 2014; Chen *et al.*, 2000; Dingwang *et al.*, 2001), reaction media (Kisch and Bahnemann, 2015; Melchionna and Fornasiero, 2020), temperature (Melchionna and Fornasiero, 2020; Serpone *et al.*, 1996), and pH of solution (Serpone *et al.*, 1996), which all have strong impact on light scattering, inevitably affect the quantum or photonic efficiency as well. Since it is nearly impossible to control all those factors, the quantum or photonic yield measured are (or should be) reported as apparent or incident yield rather than true yield (Kisch and Bahnemann, 2015; Melchionna and Fornasiero, 2020), and it should not be used for comparison. Another option would be the use of excess light (sun is free anyway) and just focus on the other problematic issues in the process as we discussed below; however, the excess energy may increase the temperature and inversely effect the photocatalytic yield indirectly while it may also influence the mass transfer limitations (Yang and Liu, 2007).

These problems involving the source and system properties affecting the light utilization efficiency have been well known and discussed in various publications for many years. However, there are other factors, especially those have high impact on mass transfer limitations, are usually overlooked.

As the last step of this thesis, we investigated the influence of the dead volume, in which the gas products bubbled from the reaction medium are collected before the analysis and the interfacial area between the reaction solution and dead volume on the hydrogen evolution rate. The gas phase concentration in a static process depends on the total volume in the gas phase. In a dynamic process in which gas phase is swept by an inert carrier gas, on the other hand, the residence time of the carrier gas (the flow rate of gas/ dead volume)

becomes important in determining the gas phase concentration. In addition to the possible effect of gas phase concentration to the mass transfer limitation from liquid to gas phase, the long residence times may cause erroneous measurements especially in the early stage of process in which the hydrogen production rate changes with time; the detection of the change in the hydrogen production by the measurement device may be delayed. Although the effects of these factors on the mass transfer from liquid to gas and therefore the measurement of hydrogen production are well-known, they are not usually treated (even reported) as a parameter; they are usually imposed by the geometry of the reactor making erroneous comparison of results obtained in two different works. We also studied the other factors, which may be influential, such as the sweep gas flow rate, stirring speed, and catalyst concentration.

4.6.1. Effect of Dead Volume

Clearly, the size of the dead volume should be important in the batch reactors because it may affect the reaction yield by affecting the mass transfer limitations in gas-liquid interface; the increasing hydrogen partial pressure and total pressure may affect the equilibrium and mass transfer resistances in the gas-liquid interface. Although this problem should be in lesser extend in semi-batch systems, a certain dead volume is still required to collect and carry the product gases to the GC (He and Que, 2013; Schwarze *et al.*, 2013; Takata *et al.*, 2015; Teixeira *et al.*, 2020), and it may affect the measurements through residence time especially in early stage of the reaction, in which the hydrogen production rate changes with time. It may also have some other effects such as changing the interfacial area and distance between the light source and the liquid gas interface.

In this part of the work, we repeated the experiments under five different dead volumes (15, 45, 75, 190, and 425 ml) under the constant liquid volume (310 ± 10 ml) and interfacial area (30 cm^2) using the reactor geometry given in Table 3.14 (R1, R2, R3, R4, and R5). The hydrogen production rates as a function of time are given in Figure 4.55 while the amount of hydrogen produced (cumulatively), which was calculated from the rates given in Figure 4.55a, are presented in Figure 4.55b. As Fig 4.55a indicates, the measured hydrogen production rates decrease with increasing dead volume (or decreasing residence time) even though the other factors such as light source, liquid to gas mass transfer area, catalyst type, etc. were exactly same. More interestingly, the hydrogen production rates

obtained with different dead volumes are still different from each other even after they became constant at the 4th -5th hours, in which some researchers stop measurements. Consequently, the cumulative hydrogen production also increases with decreasing dead volume (or residence time). These differences observed in the rate of hydrogen production or total amount of hydrogen up to a certain time may be, at least, indicating the possibility of different level of error in hydrogen measurements at different dead volumes. The residence time is much higher for higher dead volumes; hence, the change in the hydrogen production between two measurements may not be captured correctly.

4.6.2. Effect of Liquid to Gas Mass Transfer Area

Figure 4.56a exhibits the effect of gas-liquid interfacial area on hydrogen production rate while Figure 4.56b exhibit total hydrogen production; the three different interfacial areas as 42 cm², 72 cm², and 90 cm² using reactors R6, R7, and R8 respectively, were tested with 2 g/L catalyst concentration. Considering that the interfacial area may influence the mass transfer limitation and gas-liquid hydrogen equilibrium, which also depend on hydrogen concentration, we also repeated the experiments with 3 g/L catalyst concentrations, which should normally result in higher amount of hydrogen production. Different catalyst concentrations may also influence the internal mass transfer limitation between solid particle and liquid reaction media, which is also highly important for photocatalytic reactions as emphasized by Ipek and Uner (Ipek and Uner, 2020).

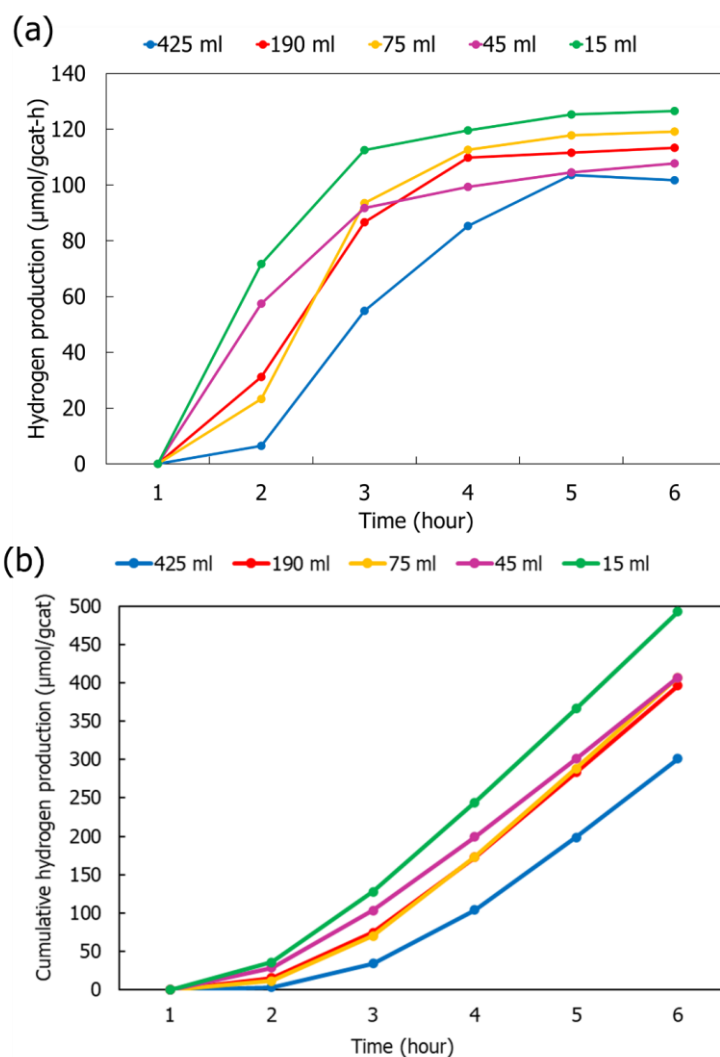


Figure 4.55. a) The effect of dead volume in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and dead volumes (mass transfer area: 72 cm^2 , catalyst concentration: 2 g/L , stirring speed: 420 rpm , N_2 flow rate: 30 ml/min , under Science-Tech solar simulator).

The lowest hydrogen production rate (per gram catalyst) was obtained at 42 cm^2 , and the rate at 2 g/L was significantly higher than that at 3 g/L . Hydrogen production was almost doubled when liquid to gas interfacial area increased from 42 cm^2 to 72 cm^2 ; in contrast to the case at 42 cm^2 , the hydrogen production rate at 2 g/L and 3 g/L was very close to each other (the rate at 3 g/L was even slightly higher) when the interfacial area was 72 cm^2 . Further increase of interfacial area to 90 cm^2 , on the other hand, resulted in decreasing hydrogen production rate slightly suggesting that there is an optimum interfacial area;

additionally, the difference between the rates at 2 g/L and 3 g/L catalyst became more apparent again.

The results summarized above clearly indicate that the change of gas-liquid interfacial area affects the hydrogen production rate; the difference continues even after the 5th hour, at which the rates become constant (as many researchers stop measurements). Although the change in mass transfer limitation may be a plausible explanation, some other factors associated with the reactor shape may be also influential. For example, the shape of R6 (erlenmeyer) is significantly different than the other two (balloon), and the mixing in R6 (especially at lower edges) may not be as good as that in R7 and R8. Even the conditions for R7 and R8 are not the same; in addition to the difference in mixing patterns due to difference in diameters, even the light absorbance (due to the difference in light exposure area and the distance to the light source) in two reactors may be different. It is not possible to draw the definitive conclusions about the dominant cause at this stage; however, it is clear that the effect of interfacial area (and other factor related to reactor shape) affects the hydrogen production rate making the comparison of the results from different works impossible.

4.6.3. Effect of Sweep Gas Flow Rate

In most of the photocatalytic water splitting reactions studied in the literature, the product gases are analyzed by using gas chromatography as in our reaction system. The flow rate of the sweep gas should be sufficient to bubble reaction solution and push product gases dissolved in reaction solution to GC. Qureshi and Takanabe emphasized the importance of reporting the flow properties for sweep gas in heterogeneous photocatalytic reactions because different flow rates caused to different partial pressures of product gases (Qureshi and Takanabe, 2017). A similar study with different results was published by Hernandez *et al.* in 2013; they demonstrated that higher flow rates of sweep gas increase the oxygen yield because the O₂ (the product for that system) saturation in the reaction media becomes lower and reaction is no more limited by mass transfer (Hernández *et al.*, 2014). We only saw the positive effects of flowrate at the initial stage of reaction (see 4.57a), at which the hydrogen production rates are slow and increases with time; it is likely that the increasing hydrogen production (and increasing H₂ concentration in liquid phase) is felt in the gas phase faster if the sweep gas flowrate is higher. Indeed, the trend disappeared after three hours; the

hydrogen production at the sweep flowrate of 15 ml/min was slightly higher. As the result, it can be said that the change in gas flow rate was not as influential on the hydrogen production as the dead volume and the interfacial area discussed above. The differences become less noticeable when the total hydrogen production was plotted against time as given in Figure 4.57b.

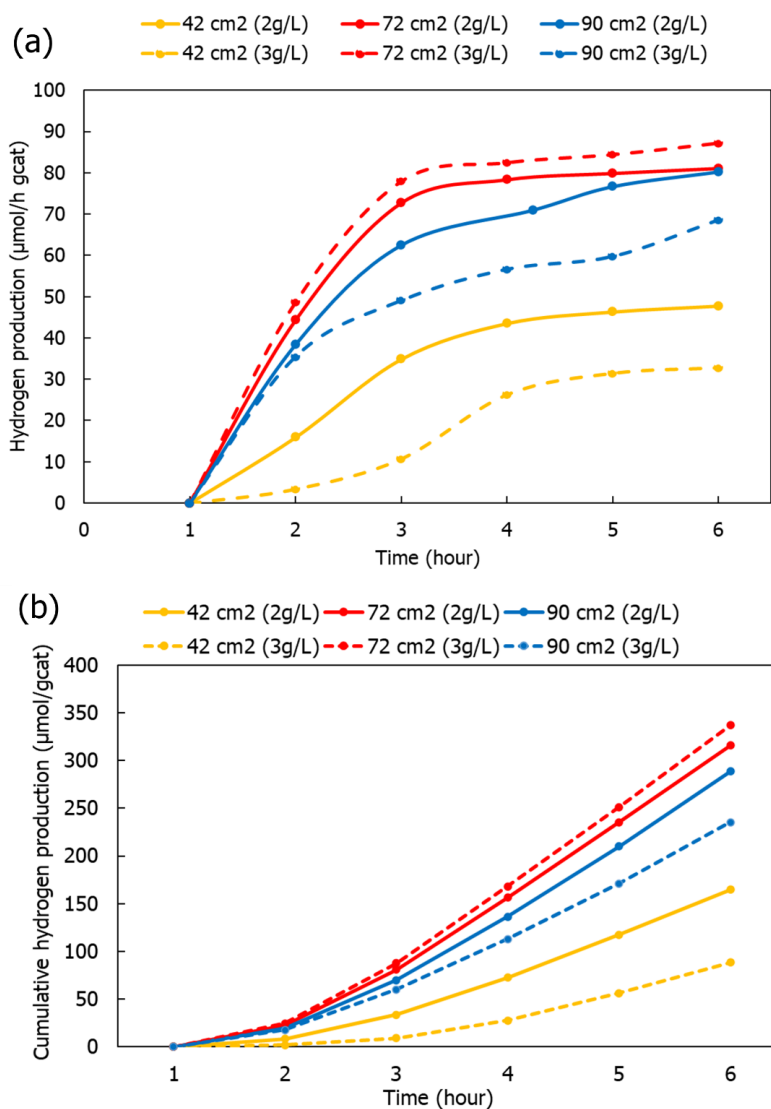


Figure 4.56. a) The effect of liquid to gas mass transfer area at two different catalyst concentration in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and liquid to gas mass transfer area (dead volume: 150 ml, stirring speed: 420 rpm, N₂ flow rate: 30 ml/min, under Science-Tech solar simulator).

4.6.4. Effect of Stirring Speed

The catalyst suspension was mixed during the reaction at constant stirring speed using magnetic stirrer. The situation is similar for most of the photocatalytic hydrogen production systems reported in literature (Schwarze *et al.*, 2013; Takata *et al.*, 2015; Teixeira *et al.*, 2020); researchers decide on a stirring rate to provide sufficient mixing for physical interaction between catalyst, light, and reaction solution enabling photocatalytic reaction. The necessity of good mixing in heterogeneous reactions is stated in several studies (Braslavsky *et al.*, 2011; Ipek and Uner, 2020; Kočí *et al.*, 2008) due to its strong impact on the elimination of mass transfer limitations. However, in most of the scientific papers, the stirring speed is still not reported neglecting its importance on photocatalytic efficiency. In Figure 4.58a and Figure 4.58b the effect of stirring speed on hydrogen production was demonstrated; some stirring rates are repeated to show that the differences in hydrogen production rate are not merely experimental error. It is clear that 200 rpm is not sufficient to mix the catalyst suspension and reaction solution; the hydrogen production rates are significantly lower than those obtained at high stirring rates indicating that the higher stirring rates decrease the mass transfer limitations at the solid-liquid interface. However, the hydrogen production rate does not always increase after certain speed was reached. Most likely, the vortex formed at high speeds brings up new complications to the reaction and makes it hard to see the relation between speed and hydrogen production rate clearly (it may even affect the reproducibility). These results support Ballari *et al.* stating that the perfect mixing is one of the simple and most essential factors to eliminate mass transfer limitations in slurry photocatalytic reactors (Ballari *et al.*, 2010). Consequently, the stirring rate should be at least reported and preferably similar to make comparison easier; however, it may not be easy to achieve this because the optimum stirring rate may be different for different works due to the differences in reactor structure.

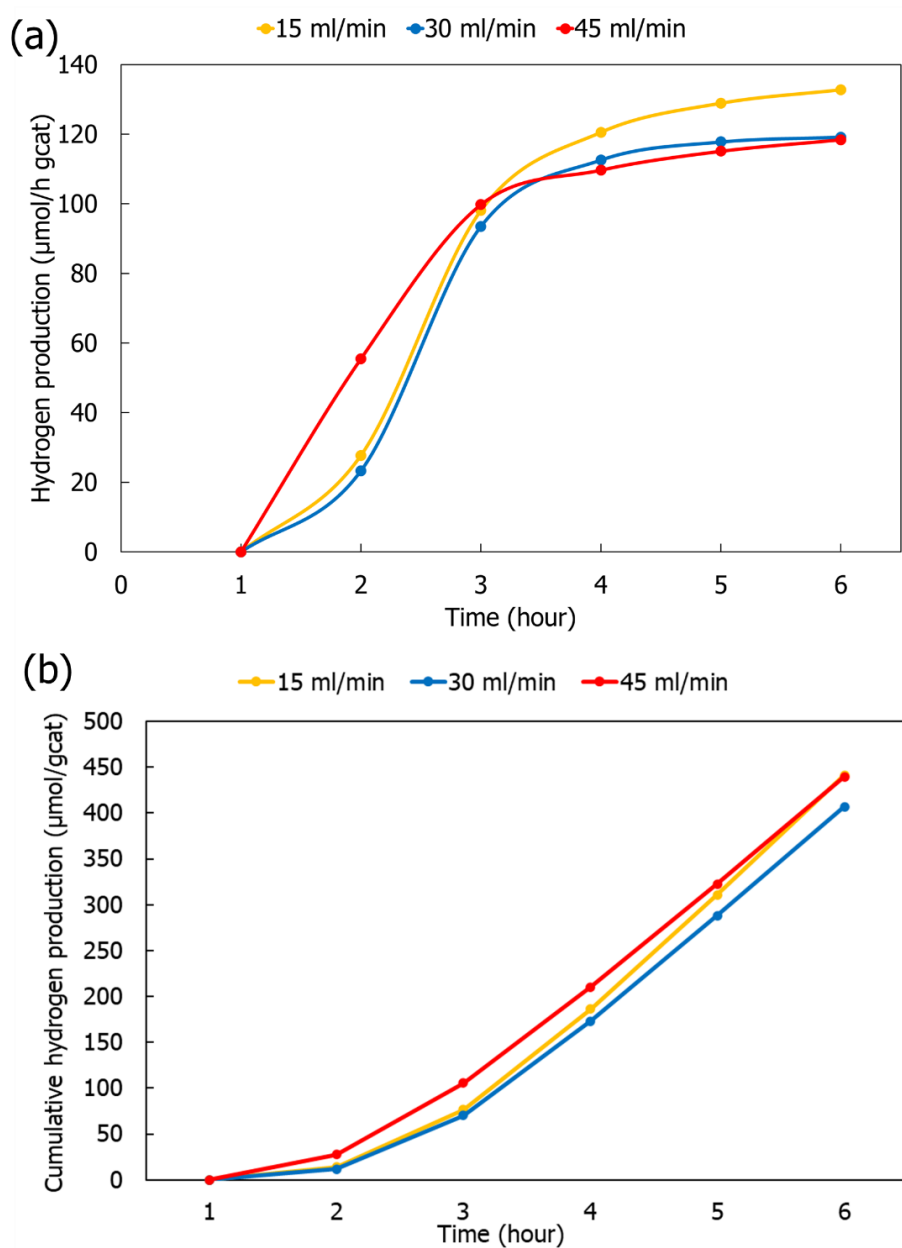


Figure 4.57. a) The effect of sweep gas flow rate in semi-batch reactors on hydrogen production, b) The relationship between cumulative hydrogen production and sweep gas flow rate (dead volume: 75 ml, mass transfer area: 30 cm^2 , catalyst concentration: 2g/L, N_2 flow rate: 30 ml/min, stirring speed: 420 rpm, under Science-Tech solar simulator).

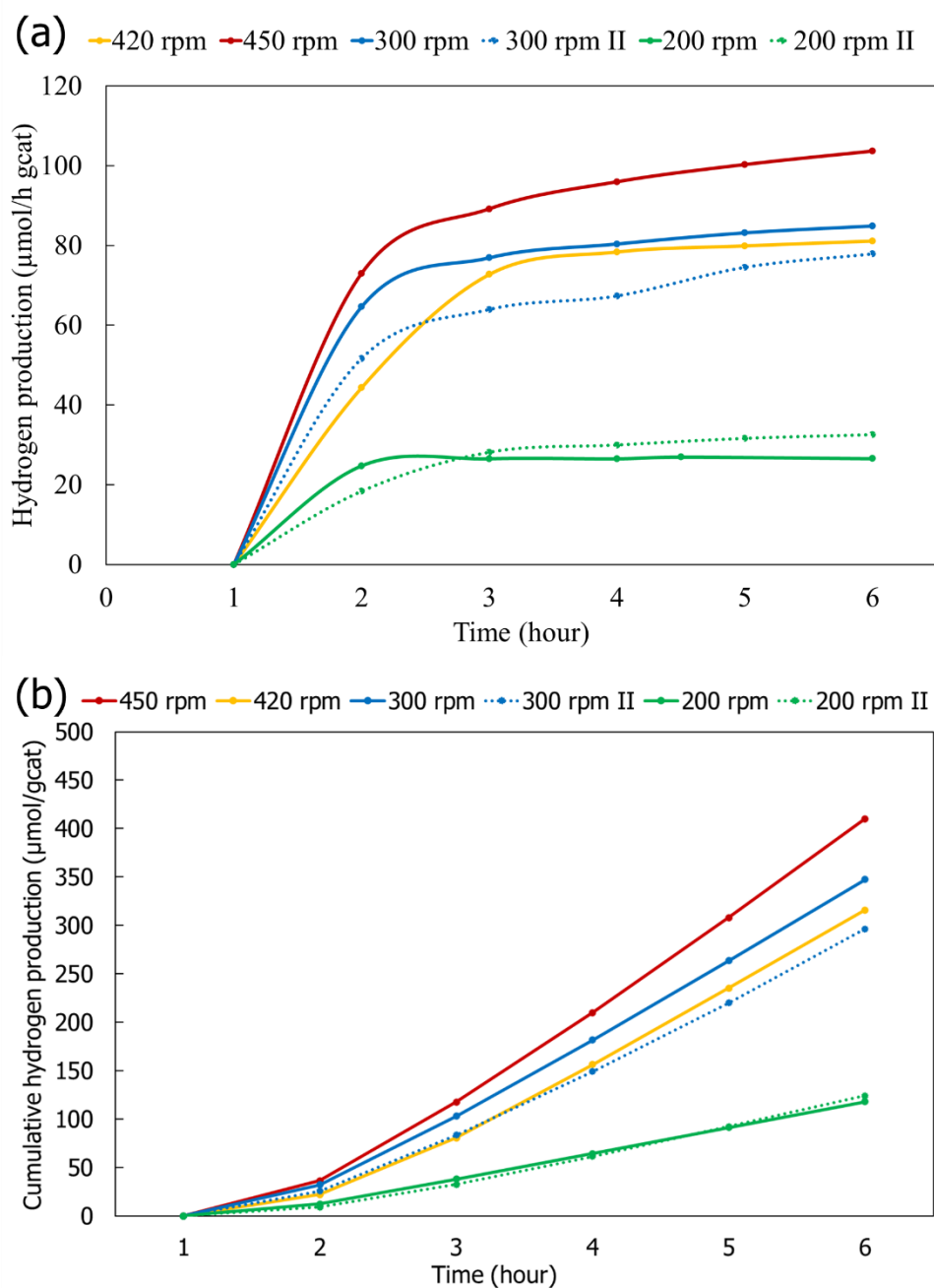


Figure 4.58. a) The comparison of hydrogen production at different stirring rates with repeated experiments, b) The relationship between cumulative hydrogen production and stirring speed (dead volume: 150 ml, mass transfer area: 72 cm², catalyst concentration: 2g/L, N₂ flow rate: 30 ml/min, under Science-Tech solar simulator).

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

The major objective of this research is to develop active and stable dye sensitized semiconductor encapsulated by an ionic liquid (IL) to be used in photocatalytic hydrogen production and to clarify the roles of each constituent of the photocatalyst. The conclusions drawn from this work will be presented in five parts. The first three parts involve implementation of machine learning (ML) tools such as association rule mining (ARM), decision tree (DT), random forest (RF), and deep learning (DL) for knowledge extraction from databases established on the topic of photocatalytic water splitting (PWS), photoelectrochemical water splitting (PECWS), and water solubility in ionic liquids. The motivation of the fourth part is to design an effective N719 dye sensitized 1%Pt/SrTiO₃ and 1%Pt/TiO₂ semiconductors coated by three acidic imidazolium based ionic liquids ([EMIM][N(Tf)₂], [BMIM][BF₄], and [BMIM][PF₆]). In the last part, the critical factors affecting the hydrogen evolution via mass transfer limitations over Pt/TiO₂ was studied.

At first, the past 151 publications on PWS over perovskites using ML algorithms were analyzed to assess the present status, identify the major trends and patterns in past works; the conclusions drawn can be summarized as follows:

- The most common perovskites in PWS are tantalates followed by titanates, niobates and indium-based perovskites. NaTaOs are the most frequently studied tantalates, and they are especially effective under UV light. The most common titanates are SrTiOs, which are also better under UV-light, but they are mostly tested under visible light.
- About half (48%) of perovskites in database are doped to A-site (31%) and/or B-site (27%). The effect of doping on band gap is observable; however, only some portion of doped catalysts had better PWS activity than plain perovskites. The doping also improves stability in some cases.

- Solid-state reaction, hydrothermal synthesis, polymerized-complex and sol-gel are the most common methods for perovskite preparation; their average performances are similar for UV tests despite the reports claiming the limitations of solid-state reaction. However, hydrothermal synthesis clearly performs better in average under the visible light. Effects of preparation method on surface area, band gap and crystal structure are also observable; the last two partially explain the hydrogen production rates as well.
- Approximately 80% of cases contain co-catalyst (Pt is the most common). There are various successful applications; however, in overall, only some fraction of co-catalyst results in higher performance than the bare perovskites.
- The methanol (20 % or less) is used as sacrificial agent in about 80% of the tests (both UV and visible light) and found to be beneficial; some other organic sacrificial agents are also used. The inorganics additives, on the other hand, are usually used in visible light tests.
- Association rule mining clearly identifies the most influential factors for high hydrogen production, and verifies the conclusions obtained from the review of literature. However, no single factor dominates the result; large number of factors affect the performance.
- Decision tree analysis provides some simple and easy to follow selection rules and heuristics for both hydrogen production and band gap. The selection rules seem to be clearer for the band gap; the doping (especially in B site) is the most determining factor for the band gap as intended.
- Random forest models for hydrogen production (especially for visible light) are quite successful in predicting the data not seen before; hence, they can be used for initial estimates even though they are not sufficient for the high accuracy predictions. The same is also true for band gap prediction. The random forest model for the band gap prediction is also very successful; however, a model directly correlating the band gap with the hydrogen production rate could not be established.

Considering the clear trends in analyses and the high fitness of the models constructed using the data from 151 different sources, it can be concluded that PWS over perovskites is predictable. However, non-standard testing procedure makes the generalization harder (especially for the stability photocatalysts).

It is hard to identify few significant factors that may be fine-tuned and improve PWS; many factors seem to be effective, and the improvement obtained is always limited. The sound solutions like ion doping to change band gap, use of co-catalyst for charge separation or use of additives as sacrificial agents help to improve the results but not as much as it is desired; there seems to be a need for new approaches.

In the second computational study, a comprehensive database for PECWS is constructed with 180 scientific articles and then analyzed by using ARM, RF, and DT. The main results are given as follows:

- The association rules revealed the effect of material, preparation method and post-treatment conditions on the band gap of the electrode produced.
- The band gap was predicted successfully with the testing error, RMSE, and R^2 values of 0.12, 0.17, and 0.87, respectively.
- The RF algorithm failed to predict the photocurrent density; in best, a model could be developed for single bias value 1 V vs RHE with the testing error, RMSE, and R^2 values of 1.3, 1.67 and 0.50, respectively.
- Both random forest and decision tree models were developed for photocurrent density classification; the random forest model had poor predictive power for high photocurrent density class. Even though overall performance of the decision tree model was not very good, it could be used in deducing some rules that leads to high photocurrent density.
- The results of band gap prediction can be a guide to researchers to foresee the possible band gap of photo anode to be produced. Similarly, ARM and DT algorithms may be helpful to determine optimum photo anode specifications and reaction conditions to reach the band gap value (in intended interval) and high photocurrent density.

As the last step of machine learning applications, the water solubility data (prepared by Ahsan Jalal, Ph.D in Koc University) of 16137 ILs from most commonly used 163 cations (in nine groups) and 99 anions were analyzed using ML tools of ARM, DT and DL algorithms. The major conclusions obtained from the analysis are summarized below.

- The accuracy of all models was remarkably high. The association rule mining analysis was clearly identified the descriptors leading low water capacity of IL while the decision tree analysis provided information to develop rules for cation and anion pairing to form ILs with low water capacity. The prediction accuracy of fully connected neural network models were also remarkable high.
- The decision tree and neural network models constructed from the entire data set had almost the same accuracy level with the models formed for each individual cation group indicating that the cationic and anionic descriptors represent the water solubility in these groups of ILs sufficiently well. This result was also supported by the fact that even a small fraction of data (about 4000 out of 16137) was sufficient to develop a model that can predict the water solubility in other ILs in the database.
- Preliminary analysis using simple descriptive statistics indicated that the water capacity for the phosphonium group ILs were generally higher than that for others while the value of cationic descriptors for this group was also distinguishably different. No other distinct pattern was observed in the dataset without the detailed ML analysis.
- Analysis using all three machine learning tools clearly showed that the anionic descriptors were much more influential than the cationic counterparts in determining the water capacity of ILs. For example, all decision tree models start to split the data based on anion HOMO energy; the anion properties such as ZPE, dipole and LUMO also used through the branches as the most influential descriptors; the association rule mining and deep learning models also have similar results.
- It was observed that if an anion forms ILs with low water capacity, it is most likely that the same anion does the same even if the cation was changed; the same is also true for the anions leading to high water capacity ILs.
- All ILs that are resulted from anions A25 (bis[1,2-benzenedolato]borate) and A26 (bisbiphenyldiolatoborate) have low water while anions A46 (fluoride) and A81(sulfate) produced ILs having high water solubility only with all the cations tested.
- The predictive powers of decision tree and deep learning models were also verified by the experimental water solubility of 49 ILs from 13 published articles in literature.

The prediction of experimental solubility data was accurate within an order of magnitude, consistent with known accuracy level of COSMO-RS, and usually sufficient for the rough estimate of water solubility in ILs.

In the fourth part of this study, the photocatalytic hydrogen production over 1 wt% Pt/TiO₂ sensitized with N719 dye (di-tetrabutylammonium cis-bis(isothiocyanato)bis(2,2'-bipyridyl-4,4'-dicarboxylato) ruthenium(II)) and coated by a thin layer of ionic liquid [BMIM][BF₄] (1-butyl-3-methylimidazolium tetrafluoroborate) was investigated. The conclusions are summarized below.

- SEM-EDAX analysis demonstrated that the ionic liquid coating after dye sensitization was successful as indicated by the homogenous distribution of Ru and F atoms representing the N719 dye and [BMIM][BF₄] ionic liquids, respectively.
- The crystal structure of TiO₂ did not change with Pt, dye or IL loading as expected.
- The band gap of TiO₂ slightly decreased upon Pt loading apparently because of calcination after Pt impregnation; no further change was observed after dye loading or IL coating.
- The hydrogen production rate increased about 15-20 % (from ~20-22 to 27 $\mu\text{mol/h.gcat}$) with dye sensitization while [BMIM][BF₄] coating alone improved the hydrogen production three times (to approximately 60 $\mu\text{mol/h.gcat}$). The ionic liquid coating after the dye, on the other hand, further improved the performance to approximately 70 $\mu\text{mol/h.gcat}$.
- Electrochemical analysis of the respective photoelectrodes further verified the performance enhancement by dye and IL coatings as demonstrated by current density and photoconversion efficiency profiles.

These findings demonstrate that coating with ionic liquid increased the performance of the catalyst by providing better charge transfer between the photocatalyst and the aqueous reaction conditions and preventing the recombination of the photogenerated electron-hole pairs.

It should be also noted that as the initial attempt of fourth part, Pt/SrTiO₃ photocatalysts produced with various methods were also coated with three ILs mentioned. However, the minor differences in the components of that photocatalytic process had a major

effect on the efficiency creating reproducibility problems. Unfortunately, those critical variables, which are so effective on the reaction, could not be determined exactly.

The fifth and final part focused on the effects of factors related to reactor structure (liquid to gas mass transfer area and dead volume) and operational conditions (flow rate of sweep gas and stirring speed of reaction media) on photocatalytic hydrogen production efficiency by using Pt/TiO₂ as catalyst in aqueous methanol (20% by volume) solution. The conclusions can be drawn as follows:

- The dead volume may affect the measured hydrogen production rate especially at the initial period of reaction; the rate is higher at lower dead volumes at constant sweep gas flowrate and decreases with the increasing dead volume. One possible cause of this is that the long residence times at high dead volumes causes measurement errors because the measured rate may not represent the actual rate on the solution due to lag time. There may be also changes in mass transfer limitations.
- The same is also true for the sweep gas flowrate because it also changes the residence time (in opposite way with dead volume); however, the effects of this variable on total hydrogen production seems to be insignificant.
- The gas liquid interfacial area is also influential on the hydrogen production rate probably through liquid-gas mass transfer limitation.
- The stirring rate seems to affect the rate increasing the solid-liquid interactions.
- These variables seem to be also correlated directly (like dead volume and sweep gas flowrate through residence time) or indirectly due to the limitations in reactor geometry (like interfacial area and dead volume).
- Any changes in these variables change the reactor structure and may change the light absorption characteristics due to the change in the distance and angle between the light source and reaction medium.

Due to all these uncertainties and well-known (and well discussed) uncertainty in determining the amount and characteristics of light absorbed by the reaction solution, it is almost impossible to compare the results obtained from different works unless some standard measurement and testing protocols (like in solar cells) are developed.

5.2. Recommendations

By considering the challenges and limitations encountered in both computational and experimental parts of this study, the following works can be proposed as recommendations for the future studies.

- Structural, thermodynamic, electronic, optical, and mechanical characteristics of energy materials are obtained and stored in some databases such as the Cambridge Structure Database (CSD) (Kennard, 1965) and the Inorganic Crystal Structure Database (ICSD) (Bergerhoff and Brown, 1978). Those properties of semiconductors can be included to the database developed from experimental articles in the context of PWS and PECWS aiming to describe the material better and to enhance the learning ability of the machine learning algorithms in related topics.
- The inconsistencies and incomparable results from different articles on PWS and PECWS create a very noisy dataset, and the simplest way to overcome this problem is to expand the dataset by scanning a larger portion of the literature.
- Gradient Boost algorithm and AdaBoost algorithm are the most promising tree-based algorithms developed recently. They can be used for both classification and regression problems instead of standard decision tree and random forest models when DT and RF fail to identify patterns in data or to predict the output successfully as in the case of photocurrent density prediction by using PECWS dataset.
- The rate slightly enhanced (about 20%) with dye sensitization (Dye:Pt/TiO₂) and also the influence of dye sensitization on the UV-Vis absorption spectra of photocatalyst could not be observed. Although the amount of dye loading was decided by literature research, to investigate the impact of dye loading on hydrogen yield could be beneficial.

- N719 is the most popular dye in solar cell applications but it may not be the best sensitizer for PWS or PECWS photocatalysis. Other ruthenium based dyes such as Z907 ([cis-Bis(isothiocyanato)(2,2'-bipyridyl-4,4'-dicarboxylato) (4,4'-di-nonyl-2'-bipyridyl) ruthenium(II)]) can be also tested for future research.
- The surface charge density of photocatalysts can be evaluated by measuring point of zero charge with changing pH. That information helps to understand the mechanisms of IL functionalization or dye sensitization on photocatalyst surface better.
- The electrochemical impedance spectroscopy can be carried out on photoanodes fabricated to reveal their electron transport and recombination characteristics exhaustively.
- Photoluminescence emission spectroscopy and Mott Schottky analysis on electrodes can provide an insight about the photocatalytic activity by considering recombination rate of photo excited charge carriers.
- Onset potential, over potential, open circuit potential, and Tafel slope of photoanode (if it is stable) can be reported to express its photoelectrochemical specifications thoroughly.
- Besides all those issues, the standardized reporting in photocatalytic hydrogen production is an absolute necessity to be able to compare the results with those published in the literature; a consensus to develop some standard measurement, testing and reporting protocols should be sought for this purpose.

REFERENCES

- Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis and Dean, J., 2016, "Tensorflow: A System for Large-Scale Machine Learning Martín", *12th USENIX Symposium on Operating Systems Design and Implementation*, pp. 265–283.
- Abe, R., 2010, "Recent Progress on Photocatalytic and Photoelectrochemical Water Splitting Under Visible Light Irradiation", *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, Vol. 11, pp. 179–209,.
- Abe, R., M. Higashi, A. Sayama, Y. Abe, and H. Sugihara, 2006, "Photocatalytic Activity of R_3MO_7 and $R_2Ti_2O_7$ ($R = Y, Gd, La$; $M = Nb, Ta$) for Water Splitting into H_2 and O_2 ", *The Journal of Physical Chemistry B*, Vol. 110, pp. 2219–2226.
- Acar, C. and I. Dincer, 2015, "Impact Assessment and Efficiency Evaluation of Hydrogen Production Methods", *Special Issue on Global Warming and Energetic Solutions*, Vol. 39, pp. 1757–1768.
- Ager, J. W., M. R. Shaner, K. A. Walczak, I. D. Sharp and S. Ardo, 2015, "Experimental Demonstrations of Spontaneous, Solar-Driven Photoelectrochemical Water Splitting", *Energy and Environmental Science*, Vol. 8, pp. 2811–2824.
- Ahmed, L. M., I. Ivanova, F. H. Hussein and D. W. Bahnemann, 2014, "Role of Platinum Deposited on TiO_2 in Photocatalytic Methanol Oxidation and Dehydrogenation Reactions", *International Journal of Photoenergy*, Vol. 2014, pp. 1-9.
- Ahn, C. W., P. H. Borse, J. H. Kim, J. Y. Kim, J. S. Jang, C. R. Cho and J. S. Lee, 2018, "Effective Charge Separation in Site-Isolated Pt-Nanodot Deposited $PbTiO_3$ Nanotube Arrays for Enhanced Photoelectrochemical Water Splitting", *Applied Catalysis B: Environmental*, Vol. 224, pp. 804–809.
- Al-Shomar, S. M., 2020, "Investigation The Effect of Doping Concentration in Ruthenium-Doped TiO_2 Thin Films for Solar Cells and Sensors Applications", *Materials Research Express*, Vol. 7, p. 036409.
- Alotaibi, B., M. Harati, S. Fan, S. Zhao, H. P. T. Nguyen, M. G. Kibria and Z. Mi, 2013,

- “High Efficiency Photoelectrochemical Water Splitting and Hydrogen Generation Using Gan Nanowire Photoelectrode”, *Nanotechnology*, Vol. 24, p. 175401.
- Altintas, C., O. F. Altundal, S. Keskin and R. Yildirim, 2021, “Machine Learning Meets with Metal Organic Frameworks for Gas Storage and Separation”, *Journal of Chemical information and Modeling*, Vol. 61, pp. 2131-2146.
- Altomare, M., K. Lee, M. S. Killian, E. Selli, and p. Schmuki, 2013, “Ta-Doped TiO₂ Nanotubes for Enhanced Solar-Light Photoelectrochemical Water Splitting”, *Chemistry - A European Journal*, Vol. 19, pp. 5841–5844.
- Amarasekara, A. S., 2016, “Acidic Ionic Liquids”, *Chemical Reviews*, Vol. 116, pp. 6133–6183.
- Amarasekara, A. S., P. Yu, Y. Lin, L. Xiang, L. Su, J. J. Zhang and B. J. Lowe, 2019, “Oxygen Reduction in Nanoporous Metal-Ionic Liquid Composite Electrocatalysts”, *Trac - Trends in Analytical Chemistry*, Vol. 19, pp. 11835–11850.
- Amedi, H. R., A. Baghban and M. A. Ahmadi, 2016, “Evolving Machine Learning Models To Predict Hydrogen Sulfide Solubility in The Presence of Various Ionic Liquids”, *Journal of Molecular Liquids*, Vol. 216, pp. 411–422.
- Ao, Y., J. Xu and D. Fu, 2009, “Study on The Effect of Different Acids on The Structure and Photocatalytic Activity of Mesoporous Titania”, *Applied Surface Science*, Vol. 256, pp. 239–245.
- Arai, N., N. Saito, H. Nishiyama, Y. Shimodaira, H. Kobayashi, Y. Inoue and K. Sato, 2008, “Photocatalytic Activity for Overall Water Splitting of RuO₂-Loaded Y_xIn_{2-x}O₃ (x = 0.9–1.5)”, *The Journal of Physical Chemistry C*, Vol. 112, pp. 5000–5005.
- Arifin, Z., S. Soeparman, D. Widhiyanuriyawan and S. Suyitno, 2017, “Performance Enhancement of Dye-Sensitized Solar Cells Using A Natural Sensitizer”, *International Journal of Photoenergy*, Vol. 1788, p. 030123.
- Armenante, P. M. and D. J. Kirwan, 1989, “Mass Transfer To Microparticles in Agitated Systems”, *Chemical Engineering Science*, Vol. 44, pp. 2781–2796.

- Arney, D., L. Fuoco, J. Boltersdorf and p. A. Maggard, 2013, "Flux Synthesis of $\text{Na}_2\text{Ca}_2\text{Nb}_4\text{O}_{13}$: The influence of Particle Shapes, Surface Features, and Surface Areas on Photocatalytic Hydrogen Production", *Journal of The American Ceramic Society*, Vol. 96, pp. 1158–1162.
- Asai, R., H. Nemoto, Q. Jia, K. Saito, A. Iwase and A. Kudo, 2014, "A Visible Light Responsive Rhodium and Antimony-Codoped SrTiO_3 Powdered Photocatalyst Loaded with An IrO_2 Cocatalyst for Solar Water Splitting", *Chemical Communications*, Vol. 54, pp. 2543–2546.
- Atabaev, T. S. and S. Atabaev, 2016, "Titania Coated Hematite Nanostructures for Solar Water Splitting Applications", *Nano LIFE*, Vol. 06, p. 1650008.
- Atkins, P. W. and J. De Paula, 2006, "*Atkins' Physical Chemistry*", Oxford University Press, Oxford.
- Babucci, M., A. Akçay, V. Balci and A. Uzun, 2015, "Thermal Stability Limits of Imidazolium Ionic Liquids Immobilized on Metal-Oxides", *Langmuir*, Vol. 31, pp. 9163–9176.
- Ballari, M. De Los M., O. M. Alfano and A. E. Cassano, 2010, "Mass Transfer Limitations in Slurry Photocatalytic Reactors: Experimental Validation", *Chemical Engineering Science*, Vol. 65, pp. 4931–4942.
- Ballari, M. De Los M., R. Brandi, O. Alfano and A. Cassano, 2008, "Mass Transfer Limitations in Photocatalytic Reactors Employing Titanium Dioxide Suspensions: External and internal Particle Constrains for The Reaction", *Chemical Engineering Journal*, Vol. 136, pp. 242–255.
- Baysal, M., M. E. Günay and R. Yıldırım, 2017, "Decision Tree Analysis of Past Publications on Catalytic Steam Reforming To Develop Heuristics for High Performance: A Statistical Review", *International Journal of Hydrogen Energy*, Vol. 42, pp. 243–254.
- Bemana, H. and S. Rashid-Nadimi, 2019, "Incorporation of NiO Electrocatalyst with $\text{A-Fe}_2\text{O}_3$ Photocatalyst for Enhanced and Stable Photoelectrochemical Water Splitting",

Surfaces and interfaces, Vol. 14, pp. 184–191.

- Benimam, H., C. Si-Moussa, M. Laidi and S. Hanini, 2020, “Modeling The Activity Coefficient At infinite Dilution of Water in Ionic Liquids Using Artificial Neural Networks and Support Vector Machines”, *Neural Computing and Applications*, Vol. 32, pp. 8635–8653.
- Bergerhoff, G. and I. D. Brown, 1978, “*Inorganic Crystal Structure Database (ICSD)*”, <https://icsd.products.fiz-karlsruhe.de/>, accessed in August 2021.
- Berglund, S. P., A. J. E. Rettie, S. Hoang and C. B. Mullins, 2012, “Incorporation of Mo and W into Nanostructured BiVO₄ Films for Efficient Photoelectrochemical Water Oxidation”, *Physical Chemistry Chemical Physics*, Vol. 14, pp. 7065–7075.
- Bini, R., C. Chiappe, C. Duce, A. Micheli, R. Solaro, A. Starita and M. R. Tiné, 2008, “Ionic Liquids: Prediction of Their Melting Points By A Recursive Neural Network Model”, *Green Chemistry*, Vol. 10, pp. 306–330.
- Birol, F., 2020a, “*Hydrogen*”, <https://www.iea.org/fuels-and-technologies/hydrogen>, accessed in August 2021.
- Birol, F., 2020b, “*Key World Energy Statistics 2020*”, <https://www.iea.org/reports/key-world-energy-statistics-2020>, accessed in August 2021.
- Blaber, M., 2014, “*Bond Polarity and Dipole Moments*”, <https://chem.libretexts.org/@go/page/15100>, accessed in August 2021.
- Bohra, D. and W. A. Smith, 2015, “Improved Charge Separation Via Fe-Doping of Copper Tungstate Photoanodes”, *Physical Chemistry Chemical Physics*, Vol. 17, pp. 9857–9866.
- Boruñ, A., 2019, “Conductance and Ionic Association of Selected Imidazolium Ionic Liquids in Various Solvents: A Review”, *Journal of Molecular Liquids*, Vol. 276, pp. 214–224.
- Bosserez, T., J. Rongé, J. Van Humbeeck, S. Haussener and J. Martens, 2015, “Conception De Cellules Photoélectrochimiques Compactes Pour La Décomposition De L’eau”, *Oil and Gas Science and Technology*, Vol. 70, pp. 877–889.

- Bosso, P., A. Milella, G. Barucca, P. Mengucci, V. Armenise, F. Fanelli and F. Fracassi, 2021, "Plasma-Assisted Deposition of Iron Oxide Thin Films for Photoelectrochemical Water Splitting", *Plasma Processes and Polymers*, Vol. 18, pp. 1–12.
- Bramer, M., 2018, "Artificial intelligence in Theory and Practice II", *IFIP 20th World Computer Congress*, International Federation for Information Processing, Milano, Italy.
- Braslavsky, S. E., A. M. Braun, A. E. Cassano, A. V. Emeline, M. I. Litter, L. Palmisano and N. Serpone, 2011, "Glossary of Terms Used in Photocatalysis and Radiation Catalysis (IUPAC Recommendations 2011)", *Pure and Applied Chemistry*, Vol. 83, pp. 931–1014.
- Bu, X., Y. Gao, S. Zhang and Y. Tian, Y., 2019, "Amorphous Cerium Phosphate on P-Doped Fe_2O_3 Nanosheets for Efficient Photoelectrochemical Water Oxidation", *Chemical Engineering Journal*, Vol. 355, pp. 910–919.
- Buriak, J. M., P. V. Kamat and K. S. Schanze, 2014, "Best Practices for Reporting on Heterogeneous Photocatalysis", *ACS Applied Materials and interfaces*, Vol. 6, pp. 11815–11816.
- Cai, C. Y., S. K. Tseng, M. Kuo, K. Y. Andrew Lin, H. Yang and R. H. Lee, 2015, "Photovoltaic Performance of A N719 Dye Based Dye-Sensitized Solar Cell with Transparent Macroporous Anti-Ultraviolet Photonic Crystal Coatings", *RSC Advances*, Vol. 5, pp. 102803–102810.
- Cai, F., Y. Meng, B. Hu, Y. Tang and W. Shi, 2015, "Microwave-Assisted Synthesis of La–Cr Co-Doped SrTiO_3 Nano-Particles and Their Use in Photocatalytic Hydrogen Evolution Under Visible Light", *RSC Advances*, Vol. 5, pp. 57354–57360.
- Cai, J., Y. Liu, S. Li, M. Gao and G. Qin, 2016, "Photoelectrochemical Behavior of Sn-Doped A- Fe_2O_3 Photoanode with Different Reducer", *Chinese Journal of Chemistry*, Vol. 34, pp. 778–782.
- Can, E. and R. Yildirim, 2019, "Data Mining in Photocatalytic Water Splitting Over Perovskites Literature for Higher Hydrogen Production", *Applied Catalysis B:*

Environmental, Vol. 242, pp. 267-283.

- Can, E., A. Jalal, I. G. Zirhlioglu, A. Uzun and R. Yildirim, 2021, "Predicting Water Solubility in Ionic Liquids Using Machine Learning Towards Design of Hydrophilic/Phobic Ionic Liquids", *Journal of Molecular Liquids*, Vol. 332, p. 115848.
- Cao, Y., Y. Chen, X. Sun, Z. Zhang and T. Mu, 2012, "Water Sorption in Ionic Liquids: Kinetics, Mechanisms and Hydrophilicity", *Physical Chemistry Chemical Physics*, Vol. 14, pp. 12252–12262.
- Carraro, G., C. Maccato, A. Gasparotto, M. E. A. Warwick, C. Sada, S. Turner and D. Barreca, 2017, "Hematite-Based Nanocomposites for Light-Activated Applications: Synergistic Role of TiO₂ and Au introduction", *Solar Energy Materials and Solar Cells*, Vol. 159, pp. 456–466.
- Chahrour, K. M., F. K. Yam and A. M. Eid, 2020, "Water-Splitting Properties of Bi-Phased TiO₂ Nanotube Arrays Subjected To High-Temperature Annealing", *Ceramics International*, Vol. 46, pp. 21471–21481.
- Chen, D., F. Li and A. K. Ray, 2000, "Effect of Mass Transfer and Catalyst Layer Thickness on Photocatalytic Reaction", *Aiche Journal*, Vol. 46, pp. 1034–1045.
- Chen, D., F. Li and A. K. Ray, 2001, "External and Internal Mass Transfer Effect on Photocatalytic Degradation", *Catalysis Today*, Vol. 66, pp. 475–485.
- Chen, D., Z. Liu, J. Li and L. Han, 2018, "Fabrication and Photoelectrochemical Properties of A Promising Flaky-Structured NaInS₂ Photoelectrode", *Journal of Photochemistry and Photobiology A: Chemistry*, Vol. 356, pp. 627–632.
- Chen, W., X. Chen, Y. Yang, J. Yuan and W. Shangguan, 2014, "Synthesis and Performance of Layered Perovskite-Type H-ABi₂Ta₂O₉ (A= Ca, Sr, Ba, K_{0.5}La_{0.5}) for Photocatalytic Water Splitting", *International Journal of Hydrogen Energy*, Vol. 39, pp. 13468–13473.
- Chen, W., C. Li, H. Gao, J. Yuan, W. Shangguan, J. Su and Y. Sun, 2012, "Photocatalytic Water Splitting on Protonated form of Layered Perovskites K_{0.5}La_{0.5}Bi₂M₂O₉ (M = Ta; Nb) by Ion-Exchange", *International Journal of Hydrogen Energy*, Vol. 37, pp. 12846–

12851.

- Chen, X., S. Shen, L. Guo and S. S. Mao, 2010, "Semiconductor-Based Photocatalytic Hydrogen Generation", *Chemical Reviews*, Vol. 110, pp. 6503–6570.
- Chen, Y., Y. Liu and Z. Ma, 2021, "Graphitic C₃N₄ Modified By Ru(II)-Based Dyes for Photocatalytic H₂ Evolution", *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, Vol. 614, p. 126119.
- Chen, Z., T. F. Jaramillo, T. G. Deutsch, A. Kleiman-Shwarscstein, A. J. Forman, N. Gaillard and H. N. Dinh, 2010, "Accelerating Materials Development for Photoelectrochemical Hydrogen Production: Standards for Methods, Definitions, and Reporting Protocols", *Journal of Materials Research*, Vol. 25, pp. 3–16.
- Chesterfield, D. and A. A. Adesina, 2009, "Evidence-Based Design and Optimisation of Titania Photocatalysts Via Artificial Neural Network Analysis", *Journal of Chemical Engineering of Japan*, Vol. 42, pp. 185–198.
- Chiou, Y. C., U. Kumar and J. C. S. Wu, 2009, "Photocatalytic Splitting of Water on NiO/InTaO₄ Catalysts Prepared By An innovative Sol-Gel Method", *Applied Catalysis A: General*, Vol. 357, pp. 73–78.
- Cho, S., J. W. Jang, S. H. Lim, H. J. Kang, S. W. Rhee, J. S. Lee and K. H. Lee, 2011, "Solution-Based Fabrication of ZnO/ZnSe Heterostructure Nanowire Arrays for Solar Energy Conversion", *Journal of Materials Chemistry*, Vol. 21, pp. 17816–17822.
- Chou, J. C., S. A. Lin, C. Y. Lee and J. Y. Gan, 2013, "Effect of Bulk Doping and Surface-Trapped States on Water Splitting with Hematite Photoanodes", *Journal of Materials Chemistry A*, Vol. 1, pp. 5908–5914.
- Coşgun, A., M. E. Günay and R. Yıldırım, 2021, "Exploring The Critical Factors of Algal Biomass and Lipid Production for Renewable Fuel Production By Machine Learning", *Renewable Energy*, Vol. 163, pp. 1299–1317.
- Crawford, S., E. Thimsen and P. Biswas, 2019, "Impact of Different Electrolytes on Photocatalytic Water Splitting", *Journal of The Electrochemical Society*, Vol. 156, p. H346.

- Cui, H., W. Zhao, C. Yang, H. Yin, T. Lin, Y. Shan and F. Huang, 2014, "Black TiO₂ Nanotube Arrays for High-Efficiency Photoelectrochemical Water-Splitting", *Journal of Materials Chemistry A*, Vol. 2, pp. 8612–8616.
- Das, C., P. Roy, M. Yang, H. Jha and p. Schmuki, 2011, "Nb Doped TiO₂ Nanotubes for Enhanced Photoelectrochemical Water-Splitting", *Nanoscale*, Vol. 3, pp. 3094–3096.
- Davran-Candan, T., M. E. Günay and R. Yildirim, 2010, "Structure and Activity Relationship for CO and O₂ Adsorption Over Gold Nanoparticles Using Density Functional Theory and Artificial Neural Networks", *Journal of Chemical Physics*, Vol. 132, p. 174113.
- De Souza, R. F., J. C. Padilha, R. S. Gonçalves and J. Rault-Berthelot, 2006, "Dialkylimidazolium Ionic Liquids As Electrolytes for Hydrogen Production from Water Electrolysis", *Electrochemistry Communications*, Vol. 8, pp. 211–216.
- Deshpande, A., S. Kelkar, S. Rayalu and S. Ogale, 2014, "Orthorhombic/Cubic Cd₂SnO₄ Nanojunctions: Enhancing Solar Water Splitting Efficiency by The Suppression of Charge Recombination", *Journal of Materials Chemistry A*, Vol. 2, pp. 492–499.
- Diedenhofen, M., F. Eckert, A. Klamt, B. Strasse and D. Leverkusen, 2003, "Compounds in Ionic Liquids Using COSMO-RS", *Engineering*, Vol. 1, pp. 475–479.
- Diedenhofen, M. and A. Klamt, 2010, "COSMO-RS As A Tool for Property Prediction of IL Mixtures - A Review", *Fluid Phase Equilibria*, Vol. 294, pp. 31–38.
- Dincer, I. and C. Acar, 2014, "Review and Evaluation of Hydrogen Production Methods for Better Sustainability", *International Journal of Hydrogen Energy*, Vol. 40, pp. 11094–11111.
- Ding, C., J. Shi, Z. Wang and C. Li, 2017, "Photoelectrocatalytic Water Splitting: Significance of Cocatalysts, Electrolyte, and interfaces", *ACS Catalysis*, Vol. 7, pp. 675–688.
- Ding, C., X. Zhou, J. Shi, p. Yan, Z. Wang, G. Liu and C. Li, 2015, "Abnormal Effects of Cations (Li⁺, Na⁺, and K⁺) on Photoelectrochemical and Electrocatalytic Water Splitting", *Journal of Physical Chemistry B*, Vol. 119, pp. 3560–3566.

- Ding, J. R. and K. S. Kim, 2016, "Flame Synthesized Single Crystal Nanocolumn-Structured WO_3 Thin Films for Photoelectrochemical Water Splitting", *Journal of Nanoscience and Nanotechnology*, Vol. 16, pp. 1578–1582.
- Dong, F., H. Wang and Z. Wu, 2009, "One-Step "Green" Synthetic Approach for Mesoporous C-Doped Titanium Dioxide with Efficient Visible Light Photocatalytic Activity", *Journal of Physical Chemistry C*, Vol. 113, pp. 16717–16723.
- Dong, Z., D. Ding, T. Li and C. Ning, 2018, "Facile Fabrication of Si-Doped TiO_2 Nanotubes Photoanode for Enhanced Photoelectrochemical Hydrogen Generation", *Applied Surface Science*, Vol. 436, pp. 125–133.
- Drummond, C., R. C. Holte, N. V. Chawla, V. S. Sheng, B. Gu, W. Fang and J. Wu, 2003, "Exploiting The Cost (In)Sensitivity of Decision Tree Splitting Criteria", *International Conference on Machine Learning*, Vol. 66, pp. 239–246.
- Du, p., A. Bueno-López, M. Verbaas, A. R. Almeida, M. Makkee, J. A. Moulijn and G. Mul, 2008, "The Effect of Surface OH-Population on The Photocatalytic Activity of Rare Earth-Doped P25- TiO_2 in Methylene Blue Degradation", *Journal of Catalysis*, Vol. 260, pp. 75–80.
- Eckert, F. and A. Klamt, 2002, "Fast Solvent Screening Via Quantum Chemistry: COSMO-RS Approach", *Aiche Journal*, Vol. 48, pp. 369–385.
- Eckert, F. and A. Klamt, 2004, "Cosmotherm (Version Version C1.2, Release 01.04)", *Cosmologic GmbH & Co. KG*, Leverkusen, Germany.
- Emery, A. A., J. E. Saal, S. Kirklin, V. I. Hegde and C. Wolverton, 2016, "High-Throughput Computational Screening of Perovskites for Thermochemical Water Splitting Applications", *Chemistry of Materials*, Vol. 28, pp. 5621–5634.
- Endres, F., 2010, "Physical Chemistry of Ionic Liquids", *Physical Chemistry Chemical Physics*, Vol. 12, pp. 1648–1648.
- Escudero, J. C., R. Simarro, S. Cervera-March and J. Giménez, 1989, "Rate-Controlling Steps in A Three-Phase (Solid-Liquid-Gas) Photoreactor: A Phenomenological Approach Applied To Hydrogen Photoproduction Using Pt/TiO_2 Aqueous Suspensions",

Chemical Engineering Science, Vol. 44, pp. 583–593.

- Everett, D. H., 1965, "Effect of Gas Imperfection on G.L.C. Measurements: A Refined Method for Determining Activity Coefficients and Second Virial Coefficients", *Transactions of The Faraday Society*, Vol. 61, pp. 1637–1645.
- Fallanza, M., M. González-Miquel, E. Ruiz, A. Ortiz, D. Gorri, J. Palomar and I. Ortiz, 2013, "Screening of RTILs for Propane/Propylene Separation Using COSMO-RS Methodology", *Chemical Engineering Journal*, 220, pp. 284–293.
- Fan, W. J., Z. F. Zhou, W. B. Xu, Z. F. Shi, F. M. Ren, H. H. Ma and S. W. Huang, 2010, "Preparation of ZnIn_2S_4 /Fluoropolymer Fiber Composites and Its Photocatalytic H_2 Evolution from Splitting of Water Using Xe Lamp Irradiation", *International Journal of Hydrogen Energy*, Vol. 35, pp. 6525–6530.
- Fan, W. Q., X. Q. Yu, S. Y. Song, H. Y. Bai, C. Zhang, D. Yan and W. D. Shi, 2014, "Fabrication of TiO_2 - BiOCl Double-Layer Nanostructure Arrays for Photoelectrochemical Water Splitting", *Crystengcomm*, Vol. 16, pp. 820–825.
- Fang, T., Y. Guo, S. Cai, N. Zhang, Y. Hu, S. Zhang and Z. Zou, 2017, "Improved Charge Separation Efficiency of Hematite Photoanodes By Coating An Ultrathin P-Type LaFeO_3 Overlayer", *Nanotechnology*, Vol. 28, p. 394003.
- Fang, Y., W. C. Lee, G. E. Canciani, T. C. Draper, Z. F. Al-Bawi, J. S. Bedi and Q. Chen, 2015, "Thickness Control in Electrophoretic Deposition of WO_3 Nanofiber Thin Films for Solar Water Splitting", *Materials Science and Engineering B: Solid-State Materials for Advanced Technology*, Vol. 202, pp. 39–45.
- Fei, H., Y. Yang, D. L. Rogow, X. Fan and S. R. J. Oliver, 2010, "Polymer-Templated Nanospider TiO_2 Thin Films for Efficient Photoelectrochemical Water Splitting" *ACS Applied Materials and interfaces*, Vol. 2, pp. 974–979.
- Frasco, M. M., A. Ben, M. Frasco and E. Ledell, 2018, "Package 'Metrics' ", <https://github.com/mfrasco/metrics>, accessed in August 2021.
- Freire, M. G., C. Neves, S. P. M. Ventura, M. J. Pratas, I. M. Marrucho, J. Oliveira and A. M. Fernandes, 2010, "Solubility of Non-Aromatic Ionic Liquids in Water and

- Correlation Using A QSPR Approach", *Fluid Phase Equilibria*, Vol. 294, pp. 234–240.
- Freire, M. G., L. M. N. B. F. Santos, A. M. Fernandes, J. A. P. Coutinho and I. M. Marrucho, 2007, "An Overview of The Mutual Solubilities of Water-Imidazolium-Based Ionic Liquids Systems", *Fluid Phase Equilibria*, Vol. 261, pp. 449–454.
- Fu, Z., T. Jiang, Z. Liu, D. Wang, L. Wang and T. Xie, 2014, "Highly Photoactive Ti-Doped A-Fe₂O₃ Nanorod Arrays Photoanode Prepared By A Hydrothermal Method for Photoelectrochemical Water Splitting", *Electrochimica Acta*, Vol. 129, pp. 358–363.
- Fu, Z., T. Jiang, L. Zhang, B. Liu, D. Wang, L. Wang and T. Xie, 2014, "Surface Treatment with Al³⁺ on A Ti-Doped A-Fe₂O₃ Nanorod Array Photoanode for Efficient Photoelectrochemical Water Splitting", *Journal of Materials Chemistry A*, Vol. 2, pp. 13705–13712.
- Fujishima, A. and K. Honda, 1972, "Electrochemical Photolysis of Water At A Semiconductor Electrode", *Nature*, Vol. 238, pp. 37–38.
- Fukuzumi, S., D. Hong and Y. Yamada, 2013, "Bioinspired Photocatalytic Water Reduction and Oxidation with Earth-Abundant Metal Catalysts", *Journal of Physical Chemistry Letters*, Vol. 4, pp. 3458–3467.
- Galvert, J. G., 1990, "Glossary of Atmospheric Chemistry Terms", *Pure and Applied Chemistry*, Vol. 62, pp. 2167–2219.
- Gao, S., X. Chen, R. Abro, A. A. Abdeltawab, S. S. Al-Deyab and G. Yu, 2015, "Desulfurization of Fuel Oil: Conductor-Like Screening Model for Real Solvents Study on Capacity of Ionic Liquids for Thiophene and Dibenzothiophene", *Industrial and Engineering Chemistry Research*, Vol. 54, pp. 9421–9430.
- Ge, M., J. Cai, J. Iocozzia, C. Cao, J. Huang, X. Zhang and Z. Lin, 2017, "A Review of TiO₂ Nanostructured Catalysts for Sustainable H₂ Generation", *International Journal of Hydrogen Energy*, Vol. 42, pp. 8418–8449.
- Gillham, R., S. Hovington, B. Hunter, C. Clarizio, M. Gallant, R. Hardy and S. Flatteau-Taylor, 2006, "Book Reviews", *Healthcare Counselling & Psychotherapy Journal*, Vol. 6, pp. 33–35.

- Gołabiewska, A., M. Paszkiewicz-Gawron, A. Sadzińska, W. Lisowski, E. Grabowska, A. Zaleska-Medynska and J. Luczak, 2018, "Fabrication and Photoactivity of Ionic Liquid-TiO₂ Structures for Efficient Visible-Light-Induced Photocatalytic Decomposition of Organic Pollutants in Aqueous Phase", *Beilstein Journal of Nanotechnology*, Vol. 9, pp. 580–590.
- Gómez-Solís, C., M. A. Ruiz-Gómez, L. M. Torres-Martínez, I. Juárez-Ramírez and D. Sánchez-Martínez, 2014, "Facile Solvo-Combustion Synthesis of Crystalline NaTaO₃ and Its Photocatalytic Performance for Hydrogen Production", *Fuel*, 130, 221–227.
- Gorman, J., 2001, "Faster, Better, Cleaner?", *Science News*, Vol. 160, p. 156.
- Graser, J., S. K. Kauwe and T. D. Sparks, 2018, "Machine Learning and Energy Minimization Approaches for Crystal Structure Predictions: A Review and New Horizons", *Chemistry of Materials*, Vol. 30 pp. 3601-3612.
- Grätzel, M., 2005, "Solar Energy Conversion By Dye-Sensitized Photovoltaic Cells", *Inorganic Chemistry*, Vol. 44, pp. 6841–6851.
- Gueymard, C. A., D. Myers and K. Emery, 2002, "Proposed Reference Irradiance Spectra for Solar Energy Systems Testing", *Solar Energy*, Vol. 73, pp. 443–467.
- Gülsoy, Z., K. B. Sezginel, A. Uzun, S. Keskin and R. Yıldırım, 2019, "Analysis of CH₄ Uptake Over Metal-Organic Frameworks Using Data-Mining Tools", *ACS Combinatorial Science*, Vol. 21, pp. 257–268.
- Günay, M. E. and R. Yıldırım, 2010, "Analysis of Selective CO Oxidation Over Promoted Pt/Al₂O₃ Catalysts Using Modular Neural Networks: Combining Preparation and Operational Variables", *Applied Catalysis A: General*, Vol. 377, pp. 174–180.
- Günay, M. E. and R. Yıldırım, 2021, "Recent Advances in Knowledge Discovery for Heterogeneous Catalysis Using Machine Learning", *Catalysis Reviews - Science and Engineering*, Vol. 63, pp. 120–164.
- Guo, S., X. Zhao, W. Zhang and W. Wang, 2018, "Optimization of Electrolyte To Significantly Improve Photoelectrochemical Water Splitting Performance of ZnO Nanowire Arrays", *Materials Science and Engineering B: Solid-State Materials for*

- Advanced Technology*, Vol. 227, pp. 129–135.
- Guo, Y., N. Zhang, H. Huang, Z. Li and Z. Zou, 2017, “A Novel Wide-Spectrum Response Hexagonal YFeO_3 Photoanode for Solar Water Splitting”, *RSC Advances*, Vol.7, pp. 18418–18420.
- Gusain, R., N. Singhal, R. Singh, U. Kumar and O. P. Khatri, 2016, “Ionic-Liquid-Functionalized Copper Oxide Nanorods for Photocatalytic Splitting of Water”, *Chempluschem*, Vol. 81, pp. 489–495.
- Hagfeldt, A., G. Boschloo, L. Sun, L. Kloo and H. Pettersson, 2010, “Dye-Sensitized Solar Cells”, *Chemical Reviews*, Vol. 11, pp. 6595–6663.
- Hahsler, M., T. Giallanza and S. Chelluboina, 2021, “Package Arulesviz ”, <https://doi.org/10.32614/rj-2017-047>, accessed in August 2021.
- Ham, Y., T. Hisatomi, Y. Goto, Y. Moriya, Y. Sakata, A. Yamakata and K. Domen, 2016, “Flux-Mediated Doping of SrTiO_3 Photocatalysts for Efficient Overall Water Splitting”, *J. Mater. Chem. A*, Vol. 4, pp. 3027–3033.
- Han, J., Z. Liu, K. Guo, X. Zhang, T. Hong and B. Wang, 2015, “ AgSbS_2 Modified ZnO Nanotube Arrays for Photoelectrochemical Water Splitting”, *Applied Catalysis B: Environmental*, Vol. 179, pp. 61–68.
- Hassan, N. K., M. R. Hashim and N. K. Allam, 2012, “ZnO Nano-Tetrapod Photoanodes for Enhanced Solar-Driven Water Splitting”, *Chemical Physics Letters*, Vol. 549, pp. 62–66.
- Hayes, R., G. G. Warr and R. Atkin, 2015, “Structure and Nanostructure in Ionic Liquids”, *Chemical Reviews*, Vol. 115, pp. 6357–6426.
- He, W., W. Wu, Q. Li, K. Chen and X. Lu, 2020, “Facile Fabrication of Ga_2O_3 Nanorods for Photoelectrochemical Water Splitting”, *ChemNanoMat*, Vol. 6, pp. 208–211.
- He, Z. and W. Que, 2013, “Surface Scattering and Reflecting: The Effect on Light Absorption or Photocatalytic Activity of TiO_2 Scattering Microspheres”, *Physical Chemistry Chemical Physics*, Vol. 15, pp. 16768–16773.

- He, Z., Z. Zhu, J. Li, J. Zhou and N. Wei, 2011, "Characterization and Activity of Mesoporous Titanium Dioxide Beads with High Surface Areas and Controllable Pore Sizes", *Journal of Hazardous Materials*, Vol. 190, pp. 133–139.
- Hernández, S., S. Bensaid, M. Armandi, A. Sacco, A. Chiodoni, B. Bonelli and G. Saracco, 2014, "A New Method for Studying Activity and Reaction Kinetics of Photocatalytic Water Oxidation Systems Using A Bubbling Reactor", *Chemical Engineering Journal*, Vol. 238, pp. 17–26.
- Hill, J. C. and K. S. Choi, 2013, "Synthesis and Characterization of High Surface Area CuWO_4 and Bi_2WO_6 Electrodes for Use As Photoanodes for Solar Water Oxidation", *Journal of Materials Chemistry A*, Vol. 1, pp. 5006–5014.
- Hill, J. C., Y. Ping, G. A. Galli and K. S. Choi, 2013, "Synthesis, Photoelectrochemical Properties, and First Principles Study of n-Type $\text{CuW}_{1-x}\text{Mo}_x\text{O}_4$ Electrodes Showing Enhanced Visible Light Absorption", *Energy and Environmental Science*, Vol. 6, pp. 2440–2446.
- Ho-Kimura, S., S. J. A. Moniz, A. D. Handoko and J. Tang, 2014, "Enhanced Photoelectrochemical Water Splitting by Nanostructured $\text{BiVO}_4\text{-TiO}_2$ Composite Electrodes", *Journal of Materials Chemistry A*, Vol. 2, pp. 3948–3953.
- Ho, W., J. C. Yu and S. Lee, 2006, "Synthesis of Hierarchical Nanoporous F-Doped TiO_2 Spheres with Visible Light Photocatalytic Activity", *Chemical Communications*, Vol. 2016, pp. 1115–1117.
- Hong, S. J., S. Lee, J. S. Jang and J. S. Lee, 2011, "Heterojunction $\text{BiVO}_4/\text{WO}_3$ Electrodes for Enhanced Photoactivity of Water Oxidation", *Energy and Environmental Science*, Vol. 4, pp. 1781–1787.
- Hong, T., Z. Liu, J. Zhang, G. Li, J. Liu, X. Zhang and S. Lin, 2016, "Flower-Like $\text{Cu}_2\text{In}_2\text{ZnS}_5$ Nanosheets: A Novel Promising Photoelectrode for Water Splitting", *ChemCatChem*, Vol. 8, pp. 1288–1292.
- Hou, Y., C. Zheng, Z. Zhu and X. Wang, 2016, "Microwave-Assisted Fabrication of Porous Hematite Photoanodes for Efficient Solar Water Splitting", *Chemical Communications*,

Vol. 52, pp. 6888–6891.

- Hsiao, p. T., L. C. Chen, T. L. Li and H. Teng, 2011, “Vapor Treatment of Nanocrystalline WO_3 Photoanodes for Enhanced Photoelectrochemical Performance in The Decomposition of Water”, *Journal of Materials Chemistry*, Vol. 21, pp. 19402–19409.
- Hu, C. C. and H. Teng, 2007, “Influence of Structural Features on The Photocatalytic Activity of NaTaO_3 Powders from Different Synthesis Methods”, *Applied Catalysis A: General*, Vol. 331, pp. 44–50.
- Hu, C. C. and H. Teng, 2010, “Structural Features of p-Type Semiconducting NiO As A Co-Catalyst for Photocatalytic Water Splitting”, *Journal of Catalysis*, Vol. 272, pp. 1–8.
- Hu, C. C., T. F. Yeh and H. Teng, 2013, “Pyrochlore-Like $\text{K}_2\text{Ta}_2\text{O}_6$ Synthesized from Different Methods As Efficient Photocatalysts for Water Splitting”, *Catalysis Science and Technology*, Vol. 3, pp. 1798–1804.
- Hu, C., W. Y. Teoh, S. Ji, C. Ye, C. and A. Iwase, 2016, “In Situ Metal Doping During Modified Anodization Synthesis of Nb_2O_5 with Enhanced Photoelectrochemical Water Splitting”, *Aiche Journal*, Vol. 62, pp. 352–358.
- Hu, S., A. Wang, X. Li, Y. Wang and H. Löwe, 2010, “Hydrothermal Synthesis of Ionic Liquid [BMIM]OH-Modified TiO_2 Nanoparticles with Enhanced Photocatalytic Activity Under Visible Light”, *Chemistry - An Asian Journal*, Vol. 5, pp. 1171–1177.
- Hu, Yaping, X. Yan, Y. Gu, X. Chen, Z. Bai, Z. Kang and Y. Zhang, 2015, “Large-Scale Patterned ZnO Nanorod Arrays for Efficient Photoelectrochemical Water Splitting”, *Applied Surface Science*, Vol. 339, pp. 122–127.
- Hu, Y-S., A. Kleiman-Shwarsstein, A. J. forman, D. Hazen, J. Park, E. W. Mcfarland and S. Barbara, 2008, “Pt-Doped $\text{R-Fe}_2\text{O}_3$ Thin Films Active for Photoelectrochemical Water Splitting”, *Chemistry of Materials*, Vol. 20, pp. 3803–3805.
- Huang, J. F., Y. Lei, T. Luo and J. M. Liu, 2020, “Photocatalytic H_2 Production from Water By Metal-Free Dye-Sensitized TiO_2 Semiconductors: The Role and Development Process of Organic Sensitizers”, *ChemSusChem*, Vol. 13, pp. 5863–5895.

- Huang, X., Y. Li, X. Gao, Q. Xue, R. Zhang, Y. Gao and M. Shao, 2020, "The Effect of The Photochemical Environment on Photoanodes for Photoelectrochemical Water Splitting", *Dalton Transactions*, Vol. 49, pp. 12338–12344.
- Huang, Y., Y. Li, Y. Wei, M. Huang and J. Wu, 2011, "Photocatalytic Property of Partially Substituted Pt-Intercalated Layered Perovskite, $\text{ASr}_2\text{Ta}_x\text{Nb}_{3-x}\text{O}_{10}$ (A=K, H; x=0, 1, 1.5, 2 and 3)", *Solar Energy Materials and Solar Cells*, Vol. 95, pp. 1019–1027.
- Huddleston, J. G., A. E. Visser, W. M. Reichert, H. D. Willauer, G. A. Broker and R. D. Rogers, 2011, "Characterization and Comparison of Hydrophilic and Hydrophobic Room Temperature Ionic Liquids Incorporating The Imidazolium Cation", *Green Chemistry*, Vol. 3, pp. 156–164.
- Huerta-Flores, A. M., L. M. Torres-Martínez and E. Moctezuma, 2017, "Overall Photocatalytic Water Splitting on $\text{Na}_2\text{Zr}_x\text{Ti}_{6-x}\text{O}_3$ (x= 0, 1) Nanobelts Modified with Metal Oxide Nanoparticles As Cocatalysts", *International Journal of Hydrogen Energy*, Vol. 42, pp. 14547–14559.
- Husin, H., S. Y. Adisalamun, T. M. Asnawi and F. Hasfita, 2017, "Pt Nanoparticle on $\text{La}_{0.02}\text{Na}_{0.98}\text{TaO}_3$ Catalyst for Hydrogen Evolution from Glycerol Aqueous Solution", *AIP Conference Proceedings*, Vol. 1788, pp. 0–6.
- Husin, H., H. M. Chen, W. N. Su, C. J. Pan, W. T. Chuang, H. S. Sheu and B. J. Hwang, 2011, "Green Fabrication of La-Doped NaTaO_3 Via H_2O_2 Assisted Sol-Gel Route for Photocatalytic Hydrogen Production", *Applied Catalysis B: Environmental*, Vol. 102, pp. 343–351.
- Ipek, B. and D. Uner, 2020, "On The Limits of Photocatalytic Water Splitting", *Water Chemistry*, <https://doi.org/10.5772/intechopen.89235>, accessed in August 2021.
- Irfan, S., Y. Shen, S. Rizwan, H. C. Wang, S. B. Khan and C. W. Nan, 2017, "Band-Gap Engineering and Enhanced Photocatalytic Activity of Sm and Mn Doped BiFeO_3 Nanoparticles", *Journal of The American Ceramic Society*, Vol. 100, pp. 31–40.
- Ishihara, T., H. Nishiguchi, K. Fukamachi and Y. Takita, 1999, "Effects of Acceptor Doping To KTaO_3 on Photocatalytic Decomposition of Pure H_2O ", *The Journal of Physical*

Chemistry B, Vol. 103, pp. 1–3.

Ishii, T., H. Kato and A. Kudo, 2004, “H₂ Evolution from An Aqueous Methanol Solution on SrTiO₃ Photocatalysts Codoped with Chromium and Tantalum Ions Under Visible Light Irradiation”, *Journal of Photochemistry and Photobiology A: Chemistry*, Vol. 163, pp. 181–186.

Ismail, A. A. and D. W. Bahnemann, 2014, “Photochemical Splitting of Water for Hydrogen Production By Photocatalysis: A Review”, *Solar Energy Materials and Solar Cells*, Vol. 128, pp. 85–101.

Isosaari, P., V. Srivastava and M. Sillanpää, 2019, “Ionic Liquid-Based Water Treatment Technologies for Organic Pollutants: Current Status and Future Prospects of Ionic Liquid Mediated Technologies”, *Science of The Total Environment*, Vol. 690, pp. 604–619.

Iwakura, H., H. Einaga and Y. Teraoka, 2010, “Relationship Between Cation Arrangement and Photocatalytic Activity for Sr-Al-Nb-O Double Perovskite”, *Inorganic Chemistry*, Vol. 49, pp. 11362–11369.

Iwase, A., H. Kato and A. Kudo, 2009, “The Effect of Alkaline Earth Metal Ion Dopants on Photocatalytic Water Splitting By NaTaO₃ Powder”, *ChemSusChem*, Vol. 2, pp. 873–877.

Iwase, A. and A. Kudo, 2010, “Photoelectrochemical Water Splitting Using Visible-Light-Responsive BiVO₄ Fine Particles Prepared in An Aqueous Acetic Acid Solution”, *Journal of Materials Chemistry*, Vol. 20, pp. 7536–7542, 2010.

Jalal, A., E. Can, S. Keskin, R. Yildirim and A. Uzun, 2019, “Selection Rules for Estimating The Solubility of C₄-Hydrocarbons in Imidazolium Ionic Liquids Determined By Machine-Learning Tools”, *Journal of Molecular Liquids*, Vol. 284, pp. 511–521.

Jana, p., C. Mata Montero, p. Pizarro, J. M. Coronado, D. P. Serrano and V. A. De La Peña O’Shea, 2014a, “Photocatalytic Hydrogen Production in The Water/Methanol System Using Pt/RE:NaTaO₃ (RE = Y, La, Ce, Yb) Catalysts”, *International Journal of Hydrogen Energy*, Vol. 39, pp. 5283–5290.

- Jeong, H., T. Kim, D. Kim and K. Kim, 2006, "Hydrogen Production By The Photocatalytic Overall Water Splitting on NiO/Sr₃Ti₂O₇: Effect of Preparation Method", *International Journal of Hydrogen Energy*, Vol. 31, pp. 1142–1146.
- Ji, I. A., M. J. Park, J. Y. Jung, M. J. Choi, Y. W. Lee, J. H. Lee and J. H. Bang, 2012, "One-Dimensional Core/Shell Structured TiO₂/ZnO Heterojunction for Improved Photoelectrochemical Performance", *Bulletin of The Korean Chemical Society*, Vol. 33, pp. 2200–2206.
- Jia, Y., Z. Wang, Y. Ma, J. Liu, W. Shi, Y. Lin and K. Zhang, 2019, "Boosting interfacial Charge Migration of TiO₂/BiVO₄ Photoanode By W Doping for Photoelectrochemical Water Splitting", *Electrochimica Acta*, Vol. 300, pp. 138–144.
- Jiang, H., Y. Chen, L. Li, H. Liu, C. Ren, X. Liu and G. Tian, 2020, "Hierarchical ZnO Nanorod/ZnFe₂O₄ Nanosheet Core/Shell Nanoarray Decorated with Pbs Quantum Dots for Efficient Photoelectrochemical Water Splitting", *Journal of Alloys and Compounds*, Vol. 828, p. 154449.
- Jiang, L., H. Li, J. Mu and Z. Ji, 2018, "Manipulation of Surface Plasmon Resonance of Sputtered Gold-Nanoparticles on TiO₂ Nanostructured Films for Enhanced Photoelectrochemical Water Splitting Efficiency", *Thin Solid Films*, Vol. 661, pp. 32–39.
- Jiang, S., Y. Li, X. Zhang and Y. Li, 2015, "Enhancing The Photoelectrochemical Water Splitting Activity of Rutile Nanorods by Removal of Surface Hydroxyl Groups", *Catalysis Today*, Vol. 259, pp. 360–367.
- Jing, L., M. Wang, X. Li, R. Xiao, Y. Zhao, Y. Zhang and K. Sun, 2015, "Covalently Functionalized TiO₂ with Ionic Liquid: A High-Performance Catalyst for Photoelectrochemical Water Oxidation", *Applied Catalysis B: Environmental*, Vol. 166, pp. 270–276.
- Jozwiak, L., J. Balcerzak and J. Tyczkowski, 2020, "Plasma-Deposited Ru-Based Thin Films for Photoelectrochemical Water Splitting", *Catalysts*, Vol. 10, p. 278.
- Jun, H., B. Im, J. Y. Kim, Y. O. Im, J. W. Jang, E. S. Kim and J. S. Lee, 2012,

- “Photoelectrochemical Water Splitting Over Ordered Honeycomb Hematite Electrodes Stabilized By Alumina Shielding”, *Energy and Environmental Science*, Vol. 5, pp. 6375–6382.
- Kado, Y., C. Y. Lee, K. Lee, J. Müller, M. Moll, E. Spiecker and p. Schmuki, 2012, “Enhanced Water Splitting Activity of M-Doped Ta₃N₅ (M = Na, K, Rb, Cs)”, *Chemical Communications*, Vol. 48, pp. 8685–8687.
- Kalanur, S. S., Y. J. Hwang, S. Y. Chae and O. S. Joo, 2013, “Facile Growth of Aligned WO₃ Nanorods on FTO Substrate for Enhanced Photoanodic Water Oxidation Activity”, *Journal of Materials Chemistry A*, Vol. 1, pp. 3479–3488.
- Kang, H. W., S. N. Lim and S. B. Park, 2012, “Co-Doping Schemes To Enhance H₂ Evolution Under Visible Light Irradiation Over SrTiO₃:Ni/M (M = La Or Ta) Prepared by Spray Pyrolysis”, *International Journal of Hydrogen Energy*, Vol. 37, pp. 5540–5549.
- Kang, H. W. and S. B. Park, 2011, “H₂ Evolution Under Visible Light Irradiation from Aqueous Methanol Solution on SrTiO₃:Cr/Ta Prepared by Spray Pyrolysis from Polymeric Precursor”, *International Journal of Hydrogen Energy*, Vol. 36, pp. 9496–9504.
- Kang, M. J. and Y. S. Kang, 2015, “Ultrathin insulating Under-Layer with A Hematite Thin Film for Enhanced Photoelectrochemical (PEC) Water Splitting Activity”, *Journal of Materials Chemistry A*, Vol. 3, pp. 15723–15728.
- Kang, Q., J. Cao, Y. Zhang, L. Liu, H. Xu and J. Ye, 2013, “Reduced TiO₂ Nanotube Arrays for Photoelectrochemical Water Splitting”, *Journal of Materials Chemistry A*, Vol.1, pp. 5766–5774.
- Kanhere, p., J. Zheng and Z. Chen, 2012, “Visible Light Driven Photocatalytic Hydrogen Evolution and Photophysical Properties of Bi³⁺ Doped NaTaO₃”, *International Journal of Hydrogen Energy*, Vol. 37, pp. 4889–4896.
- Karakaya, M. and F. Uzun, 2012, “H₂-Anion interactions and Energy Calculations for Imidazolium-Based Ionic Liquids As Hydrogen Storage Materials”, *International*

Journal of Engineering Technologies, IJET, Vol. 2, p. 1.

- Kato, H., K. Asakura and A. Kudo, 2003, "Highly Efficient Water Splitting into H₂ and O₂ Over Lanthanum-Doped NaTaO₃ Photocatalysts with High Crystallinity and Surface Nanostructure", *Journal of The American Chemical Society*, Vol. 125, pp. 3082–3089.
- Kato, H., H. Kobayashi and A. Kudo, 2002, "Role of Ag⁺ in The Band Structures and Photocatalytic Properties of AgMO₃ (M: Ta and Nb) with The Perovskite Structure", *Journal of Physical Chemistry B*, Vol. 106, pp. 12441–12447.
- Kato, H. and A. Kudo, 2001, "Water Splitting into H₂ and O₂ on Alkali Tantalate Photocatalysts ATaO₃ (A = Li, Na, and K)", *Journal of Physical Chemistry B*, Vol. 105, pp. 4285–4292.
- Kato, H. and A. Kudo, 2002a, "Photocatalytic Reduction of Nitrate Ions Over Tantalate Photocatalysts", *Physical Chemistry Chemical Physics*, Vol.4, pp. 2833–2838.
- Kato, H. and A. Kudo, 2002b, "Visible-Light-Response and Photocatalytic Activities of TiO₂ and SrTiO₃ Photocatalysts Codoped with Antimony and Chromium", *Journal of Physical Chemistry B*, Vol. 106, pp. 5029–5034.
- Kelkar, S. A., P. A. Shaikh, P. Pachfule and S. B. Ogale, 2012, "Nanostructured Cd₂SnO₄ As An Energy Harvesting Photoanode for Solar Water Splitting", *Energy and Environmental Science*, Vol. 5, pp. 5681–5685.
- Kennard, O., 1965, "Cambridge Structure Database (CSD)", <https://www.ccdc.cam.ac.uk/solutions/csd-core/components/csd/>, accessed in August 2021.
- Kernchen, U., B. Etzold, W. Korth and A. Jess, 2007, "Solid Catalyst Ionic Liquid Layer (SCILL) - A New Concept To Improve Selectivity Illustrated By Hydrogenation of Cyclooctadiene", *Chemical Engineering and Technology*, Vol. 30, pp. 985–994.
- Khan, A., X. Lu, L. Aldous and C. Zhao, 2013, "Oxygen Reduction Reaction in Room Temperature Protic Ionic Liquids", *Journal of Physical Chemistry C*, Vol. 117, pp. 18334–18342.

- Khan, H. R., M. Aamir, M. A. Malik, A. A. Tahir, B. Akram, G. Murtaza and J. Akhtar, 2019, "Chemically Vaporized Cobalt Incorporated Wurtzite As Photoanodes for Efficient Photoelectrochemical Water Splitting", *Materials Science in Semiconductor Processing*, Vol. 101, pp. 223–229.
- Khenkin, M. V., E A. Katz, A. Abate, G. Bardizza, J. J. Berry, C. Brabec and M. Lira-Cantu, 2020, "Consensus Statement for Stability Assessment and Reporting for Perovskite Photovoltaics Based on ISOS Procedures", *Nature Energy*, Vol. 5, pp. 35–49.
- Kilic, A., R. Yildirim and D. Eroglu, 2021, "Machine Learning Analysis of Ni/SiC Electrodeposition Using Association Rule Mining and Artificial Neural Network", *Journal of Electrochemical Society*, Vol. 168, p. 062514.
- Kim, J. H., J. W. Jang, Y. H. Jo, F. F. Abdi, Y. H. Lee, R. Van De Krol and J. S. Lee, 2016, "Hetero-Type Dual Photoanodes for Unbiased Solar Water Splitting with Extended Light Harvesting", *Nature Communications*, Vol. 7, pp. 1–9.
- Kim, J. H. and J. S. Lee, 2019, "Elaborately Modified BiVO₄ Photoanodes for Solar Water Splitting", *Advanced Materials*, Vol. 31, pp. 1–35.
- Kim, J. K., K. Shin, S. M. Cho, T. W. Lee, T and J. H. Park, 2011, "Synthesis of Transparent Mesoporous Tungsten Trioxide Films with Enhanced Photoelectrochemical Response: Application To Unassisted Solar Water Splitting", *Energy and Environmental Science*, Vol. 4, pp. 1465–1470.
- Kim, S. H., S. Park, C. W. Lee, B. S. Han, S. W. Seo, J. S. Kim and K. S. Hong, 2012, "Photophysical and Photocatalytic Water Splitting Performance of Stibiotantalite Type-Structure Compounds, SbMO₄ (M = Nb, Ta)", *International Journal of Hydrogen Energy*, Vol. 37, pp. 16895–16902.
- Kingma, D. P. and J. L. Ba, 2015, "Adam: A Method for Stochastic Optimization", *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, San Diego.
- Kisch, H., 2013, "Semiconductor Photocatalysis - Mechanistic and Synthetic Aspects", *Angewandte Chemie - International Edition*, Vol. 52, pp. 812–847.

- Kisch, H. and D. Bahnemann, 2015, "Best Practice in Photocatalysis: Comparing Rates or Apparent Quantum Yields?", *Journal of Physical Chemistry Letters*, Vol. 6, pp. 1907–1910.
- Kitano, M., M. Takeuchi, M. Matsuoka, J. M. Thomas and M. Anpo, 2007, "Photocatalytic Water Splitting Using Pt-Loaded Visible Light-Responsive TiO₂ Thin Film Photocatalysts", *Catalysis Today*, Vol. 12, pp. 133–138.
- Klamt, A., 2003, "Prediction of The Mutual Solubilities of Hydrocarbons and Water with COSMO-RS", *Fluid Phase Equilibria*, Vol. 206, pp. 223–235.
- Klamt, A. and F. Eckert, F., 2000, "COSMO-RS: A Novel and Efficient Method for The A Priori Prediction of Thermophysical Data of Liquids", *Fluid Phase Equilibria*, Vol. 172, pp. 43–72.
- Kment, S., F. Riboni, S. Pausova, L. Wang, L. Wang, H. Han and R. Zboril, 2017, "Photoanodes Based on TiO₂ and α -Fe₂O₃ for Solar Water Splitting-Superior Role of 1D Nanoarchitectures and of Combined Heterostructures", *Chemical Society Reviews*, Vol. 46, pp. 3716–3769.
- Kochuveedu, S. T., Y. H. Jang, Y. J. Jang and D. H. Kim, 2013, "Visible Light Active Photocatalysis on Block Copolymer induced Strings of ZnO Nanoparticles Doped with Carbon", *Journal of Materials Chemistry A*, Vol. 1, pp. 898–905.
- Kočí, K., L. Obalová, D. Plachá and Z. Lacný, 2008, "Effect of Temperature, Pressure and Volume of Reacting Phase on Photocatalytic CO₂ Reduction on Suspended Nanocrystalline TiO₂", *Collection of Czechoslovak Chemical Communications*, Vol. 73, pp. 1192–1204.
- Kölle, P. and R. Dronskowski, 2004, "Hydrogen Bonding in The Crystal Structures of The Ionic Liquid Compounds Butyldimethylimidazolium Hydrogen Sulfate, Chloride, and Chloroferrate(II,III)", *Inorganic Chemistry*, Vol. 43, pp. 2803–2809.
- Kondratenko, E. V., M. Schlüter, M. Baerns, D. Linke and M. Holena, 2015, "Developing Catalytic Materials for The Oxidative Coupling of Methane Through Statistical Analysis of Literature Data", *Catalysis Science and Technology*, Vol. 5, pp. 1668–1677.

- Kothari, R., D. Buddhi and R. L. Sawhney, 2008, "Comparison of Environmental and Economic Aspects of Various Hydrogen Production Methods", *Renewable and Sustainable Energy Reviews*, Vol. 12, pp. 553–563.
- Kudo, A. and H. Kato, 2000, "Effect of Lanthanide-Doping into NaTaO₃ Photocatalysts for Efficient Water Splitting", *Chemical Physics Letters*, Vol. 331, pp. 373–377.
- Kudo, A. and Y. Miseki, 2009, "Heterogeneous Photocatalyst Materials for Water Splitting", *Chemical Society Reviews*, Vol. 38, pp. 253–278.
- Kumar, D., S. Singh and N. Khare, 2018, "Plasmonic Ag Nanoparticles Decorated NaNbO₃ Nanorods for Efficient Photoelectrochemical Water Splitting", *International Journal of Hydrogen Energy*, Vol. 43, pp. 8198–8205.
- Kuo, Y. and K. J. Klabunde, 2012, "Hydrogen Generation from Water/Methanol under Visible Light Using Aerogel Prepared Strontium Titanate (SrTiO₃) Nanomaterials Doped with Ruthenium and Rhodium Metals", *Nanotechnology*, Vol.23, p. 294001.
- Kuvarega, A. T., R. W. M. Krause and B. B. Mamba, 2014, "Comparison Between Base Metals and Platinum Group Metals in Nitrogen, M Codoped TiO₂ (M = Fe, Cu, Pd, Os) for Photocatalytic Removal of An Organic Dye in Water", *Journal of Nanomaterials*, Vol. 2014, p. 962102.
- Kwon, J., H. Cho, J. Jung, H. Lee, S. Hong, J. Yeo and S. H. Ko, 2018, "ZnO/CuO/M (M = Ag, Au) Hierarchical Nanostructure By Successive Photoreduction Process for Solar Hydrogen Generation", *Nanomaterials*, Vol. 8, p. 323.
- Latorre-Sánchez, M., C. Lavorato, M. Puche, V. Fornés, R. Molinari and H. Garcia, 2012, "Visible-Light Photocatalytic Hydrogen Generation by Using Dye-Sensitized Graphene Oxide As A Photocatalyst", *Chemistry - A European Journal*, Vol. 18, pp. 16774–16783.
- Lawal, I. A., M. Klink, P. Ndungu and B. Moodley, 2019, "Brief Bibliometric Analysis of "Ionic Liquid" Applications and Its Review As A Substitute for Common Adsorbent Modifier for The Adsorption of Organic Pollutants", *Environmental Research*, Vol. 175, pp. 34–51.

- Lawson, B., L. Leemis and V. Kudlay, 2021, "Package SimEd", <https://cran.r-project.org/package=simed>, accessed in August 2021.
- Lee, B. S. and S. T. Lin, 2015, "Screening of Ionic Liquids for CO₂ Capture Using The COSMO-SAC Model", *Chemical Engineering Science*, Vol. 121, pp. 157–168.
- Lee, C. W., D. W. Kim, I. S. Cho, S. Park, S. S. Shin, S. W. Seo and K. S. Hong, 2012, "Simple Synthesis and Characterization of SrSnO₃ Nanoparticles with Enhanced Photocatalytic Activity", *International Journal of Hydrogen Energy*, Vol. 37, pp. 10557–10563.
- Lee, H. J., S. H. Shin, K. T. Nam, J. Nah and M. H. Lee, 2016, "Spontaneously Polarized Lithium-Doped Zinc Oxide Nanowires As Photoanodes for Electrical Water Splitting", *Journal of Materials Chemistry A*, Vol. 4, pp. 3223–3227.
- Lee, H., Y. Park and M. Kang, 2013, "Synthesis and Characterization of Zn_xTi_yS and Its Photocatalytic Activity for Hydrogen Production from Methanol/Water Photo-Splitting", *Journal of industrial and Engineering Chemistry*, Vol. 19, pp. 1162–1168.
- Lee, J. G., D. Y. Kim, J. J. Park, Y. H. Cha, J. Y. Yoon, H. S. Jeon and S. S. Yoon, 2014, "Graphene-Titania Hybrid Photoanodes by Supersonic Kinetic Spraying for Solar Water Splitting", *Journal of The American Ceramic Society*, Vol. 97, pp. 3660–3668.
- Letcher, T. M. and W. C. Moollan, 1995, "The Determination of Activity Coefficients at Infinite Dilution Using G.L.C. with A Moderately Volatile Solvent (Dodecane) at The Temperatures 280.15 K and 298.15 K", *The Journal of Chemical Thermodynamics*, Vol. 27, pp. 1025–1032.
- Li, F., J. Luo, G. Chen, Y. Fan, Q. Huang, Y. Luo and Q. Meng, Q., 2014, "Hydrothermal Synthesis of Zinc Indium Sulfide Microspheres with Ag⁺ Doping for Enhanced H₂ Production by Photocatalytic Water Splitting under Visible Light", *Catal. Sci. Technol.*, Vol. 4, pp. 1144–1150.
- Li, G., T. Kako, D. Wang, Z. Zou and J. Ye, 2008, "Synthesis and Enhanced Photocatalytic Activity of NaNbO₃ Prepared by Hydrothermal and Polymerized Complex Methods", *Journal of Physics and Chemistry of Solids*, Vol. 69, pp. 2487–2491.

- Li, H., F. Zhao, J. Zhang, L. Luo, X. Xiao, Y. Huang and Y. Tong, 2017, "Ag-C₃N₄/WO₃ Photoanode with Exceptional Ability for Photoelectrochemical Water Splitting", *Materials Chemistry Frontiers*, Vol.1, pp. 338–342.
- Li, Jiangtian and N. Wu, 2015, "Semiconductor-Based Photocatalysts and Photoelectrochemical Cells for Solar Fuel Generation: A Review", *Catalysis Science and Technology*, Vol. 5, pp. 1360–1384.
- Li, Juanjuan, L. Jia, W. Fang and J. Zeng, 2010, "Enhancement of Activity of LaNi_{0.7}Cu_{0.3}O₃ for Photocatalytic Water Splitting by Reduction Treatment at Moderate Temperature", *International Journal of Hydrogen Energy*, Vol. 35, pp. 5270–5275.
- Li, Juanjuan, J. Zeng, L. Jia and W. Fang, 2010, "Investigations on The Effect of Cu²⁺/Cu¹⁺ Redox Couples and Oxygen Vacancies on Photocatalytic Activity of Treated LaNi_{1-x}Cu_xO₃ (X=0.1, 0.4, 0.5)", *International Journal of Hydrogen Energy*, Vol. 35 pp. 12733–12740.
- Li, Qian, J. Bian, N. Zhang and D. H. L. Ng, 2015, "Loading Ni(OH)₂ on The Ti-Doped Hematite Photoanode for Photoelectrochemical Water Splitting", *Electrochimica Acta*, Vol. 155, pp. 383–390.
- Li, Qiuye, T. Kako and J. Ye, 2011, "Facile Ion-Exchanged Synthesis of Sn²⁺ incorporated Potassium Titanate Nanoribbons and Their Visible-Light-Responded Photocatalytic Activity", *International Journal of Hydrogen Energy*, Vol. 36, pp. 4716–4723.
- Li, Yibin, Y. Huang, J. Wu, M. Huang and J. Lin, 2010, "Photocatalytic Activities for Hydrogen Evolution of New Layered Compound Series HLaTa_{x/3}Nb_{2-x/3}O₇/Pt (x=0, 2, 3, 4, and 6)", *Journal of Hazardous Materials*, Vol. 177, pp. 458–464.
- Li, Yibin, J. Wu, Y. Huang, M. Huang and J. Lin, J., 2009, "Photocatalytic Water Splitting on New Layered Perovskite A_{2.33}Sr_{0.67}Nb₅O_{14.335} (A = K, H)", *International Journal of Hydrogen Energy*, 34, 7927–7933.
- Li, Yingxuan, G. Chen, H. Zhang and Z. Li, 2009a, "Electronic Structure and Photocatalytic Water Splitting of Lanthanum-Doped Bi₂AlNbO₇", *Materials Research Bulletin*, Vol. 44, pp. 741–746.

- Li, Yingxuan, G. Chen, H. Zhang and Z. Li, 2009b, "Photocatalytic Water Splitting of $\text{La}_2\text{AlTaO}_7$ and The Effect of Aluminum on The Electronic Structure. *Journal of Physics and Chemistry of Solids*, Vol. 70, pp. 536–540.
- Li, Yingxuan, G. Chen, H. Zhang, Z. Li and J. Sun, 2008, "Electronic Structure and Photocatalytic Properties of $\text{ABi}_2\text{Ta}_2\text{O}_9$ ($\text{A} = \text{Ca}, \text{Sr}, \text{Ba}$)", *Journal of Solid State Chemistry*, Vol. 181, pp. 2653–2659.
- Li, Yingxuan, G. Chen, H. Zhang and Z. Lv, 2010, "Band Structure and Photocatalytic Activities for H_2 Production of $\text{ABi}_2\text{Nb}_2\text{O}_9$ ($\text{A} = \text{Ca}, \text{Sr}, \text{Ba}$)", *International Journal of Hydrogen Energy*, Vol. 35, pp. 2652–2656.
- Li, Yiyang, Y. K. Peng, L. Hu, J. Zheng, D. Prabhakaran, S. Wu and S. C. E. Tsang, 2019, "Photocatalytic Water Splitting by N-TiO₂ on MgO (111) with Exceptional Quantum Efficiencies at Elevated Temperatures", *Nature Communications*, Vol. 10, pp. 1–9.
- Lin, J., P. Hu, Y. Zhang, M. Fan, Z. He, C. K. , Ngaw and T. T. Y. Tan, 2013, "Understanding The Photoelectrochemical Properties of A Reduced Graphene Oxide-WO₃ Heterojunction Photoanode for Efficient Solar-Light-Driven Overall Water Splitting", *RSC Advances*, Vol. 3, pp. 9330–9336.
- Ling, Y., G. Wang, D. A. Wheeler, J. Z. Zhang and Y. Li, 2011, "Sn-Doped Hematite Nanostructures for Photoelectrochemical Water Splitting", *Nano Letters*, Vol.11, pp. 2119–2125.
- Liu, C., F. Wang, J. Zhang, K. Wang, Y. Qiu, Q. Liang and Z. Chen, 2018, "Efficient Photoelectrochemical Water Splitting by g-C₃N₄/TiO₂ Nanotube Array Heterostructures", *Nano-Micro Letters*, Vol.10, pp. 1–13.
- Liu, G., L. C. Yin, J. Wang, P. Niu, C. Zhen, Y. Xie and H. M. Cheng, 2012, "A Red Anatase TiO₂ Photocatalyst for Solar Energy Conversion", *Energy and Environmental Science*, Vol. 5, pp. 9603–9610.
- Liu, J., M. Shahid, Y. S. Ko, E. Kim, T. K. Ahn, J. H. Park and Y. U. Kwon, 2013, "Investigation of Porosity and Heterojunction Effects of A Mesoporous Hematite Electrode on Photoelectrochemical Water Splitting", *Physical Chemistry Chemical*

Physics, Vol. 15, pp. 9775–9782.

- Liu, J. W., G. Chen, Z. H. Li and Z. G. Zhang, 2007, “Hydrothermal Synthesis and Photocatalytic Properties of ATaO_3 and ANbO_3 ($A = \text{Na}$ and K)”, *International Journal of Hydrogen Energy*, Vol. 32, pp. 2269–2272.
- Liu, K. I., Y. C. Hsueh, C. Y. Su and T. P. Perng, 2013, “Photoelectrochemical Application of Mesoporous TiO_2/WO_3 Nanohoneycomb Prepared by Sol-Gel Method”, *International Journal of Hydrogen Energy*, Vol. 38, pp. 7750–7755.
- Liu, K., G. Wang, M. Meng, S. Chen, J. Li, X. Sun and N. Qin, 2017, “ TiO_2 Nanotube Photonic Crystal Fabricated by Two-Step Anodization Method for Enhanced Photoelectrochemical Water Splitting”, *Materials Letters*, Vol. 207, pp. 96–99.
- Liu, Yang, L. Xie, Y. Li, R. Yang, J. Qu, Y. Li and X. Li, 2008, “Synthesis and High Photocatalytic Hydrogen Production of SrTiO_3 Nanoparticles from Water Splitting Under UV Irradiation. *Journal of Power Sources*, Vol. 183, pp. 701–707.
- Liu, Yichong, Z. Kang, H. Si, P. Li, S. Cao, S. Liu and Y. Zhang, 2017, “Cactus-Like Hierarchical Nanorod-Nanosheet Mixed Dimensional Photoanode for Efficient and Stable Water Splitting”, *Nano Energy*, Vol. 35, pp. 189–198.
- Liu, Yichong, B. Xiao, H. Chen, Y. Li and S. Chang, 2019, “Decreased Surface Photovoltage of ZnO Photoanode Films Via Optimal Annealing Temperature for Enhanced Photoelectrochemical Performance”, *Journal of Nanomaterials*, Vol. 2019, p. 9367573.
- Lu, L., S. Ni, G. Liu and X. Xu, 2017, “Structural Dependence of Photocatalytic Hydrogen Production Over La/Cr Co-Doped Perovskite Compound ATiO_3 ($A = \text{Ca}$, Sr and Ba)”, *International Journal of Hydrogen Energy*, Vol. 42, pp. 23539–23547.
- Lucas-Granados, B., R. Sánchez-Tovar, R. M. Fernández-Domene and J. García-Antón, 2017, “Controlled Hydrodynamic Conditions on The formation of Iron Oxide Nanostructures Synthesized by Electrochemical Anodization: Effect of The Electrode Rotation Speed”, *Applied Surface Science*, Vol. 392, pp. 503–513.
- Luo, J. and P. A. Maggard, 2006, “Hydrothermal Synthesis and Photocatalytic Activities of SrTiO_3 -Coated Fe_2O_3 and BiFeO_3 ”, *Advanced Materials*, Vol. 18, pp. 514–517.

- Luo, W., Z. Li, X. Jiang, T. Yu, L. Liu, X. Chen and Z. Zou, 2008, "Correlation Between The Band Positions of $(\text{SrTiO}_3)_{1-x}(\text{LaTiO}_2\text{N})_x$ Solid Solutions and Photocatalytic Properties under Visible Light Irradiation", *Physical Chemistry Chemical Physics*, Vol. 10, pp. 6717–6723.
- Luo, W., J. Wang, X. Zhao, Z. Zhao, Z. Li and Z. Zou, 2013, "Formation Energy and Photoelectrochemical Properties of BiVO_4 After Doping At Bi^{3+} or V^{5+} Sites with Higher Valence Metal Ions", *Physical Chemistry Chemical Physics*, Vol. 15, pp. 1006–1013.
- Lv, J., T. Kako, Z. Li, Z. Zou and J. Ye, 2010, "Synthesis and Photocatalytic Activities of NaNbO_3 Rods Modified By In_2O_3 Nanoparticles", *Journal of Physical Chemistry C*, Vol. 114, pp. 6157–6162.
- Ma, S. S. K., K. Maeda, R. Abe and K. Domen, 2012, "Visible-Light-Driven Nonsacrificial Water Oxidation over Tungsten Trioxide Powder Modified with Two Different Cocatalysts", *Energy and Environmental Science*, Vol. 5, pp. 8390–8397.
- Ma, Y., X. Wang, Y. Jia, X. Chen, H. Han and C. Li, 2014, "Titanium Dioxide-Based Nanomaterials for Photocatalytic Fuel Generations", *Chemical Reviews*, Vol. 114, pp. 9987–10043.
- Macfarlane, D. R., N. Tachikawa, M. Forsyth, J. M. Pringle, P. C. Howlett, G. D. Elliott and C. A. Angell, 2014, "Energy Applications of Ionic Liquids", *Energy and Environmental Science*, Vol. 7, pp. 232–250.
- Mahajan, V. K., S. K. Mohapatra and M. Misra, 2008, "Stability of TiO_2 Nanotube Arrays in Photoelectrochemical Studies", *International Journal of Hydrogen Energy*, Vol. 33, pp. 5369–5374,.
- Manfredi, N., B. Cecconi, V. Calabrese, A. Minotti, F. Peri, R. Ruffo and A. Abboto, 2016, "Dye-Sensitized Photocatalytic Hydrogen Production: Distinct Activity in A Glucose Derivative of A Phenothiazine Dye", *Chemical Communications*, Vol. 52, pp. 6977–6980.
- Mansilla, C., C. Bourasseau, C. Cany, B. Guinot, A. Duigou, P. Leand Lucchese, 2018,

- “Hydrogen Applications: Overview of The Key Economic Issues and Perspectives”, Elsevier Ltd. Academic Press, Cambridge, United States.
- Mao, A., G. Y. Han and J. H. Park, 2010, “Synthesis and Photoelectrochemical Cell Properties of Vertically Grown A-Fe₂O₃ Nanorod Arrays on A Gold Nanorod Substrate”, *Journal of Materials Chemistry*, Vol. 20, pp. 2247–2250.
- Marciniak, A., 2010, “Influence of Cation and Anion Structure of The Ionic Liquid on Extraction Processes Based on Activity Coefficients at infinite Dilution: A Review”, *Fluid Phase Equilibria*, Vol. 294, pp. 213–233.
- Marcinkowska, R., K. Konieczna, Ł. Marcinkowski, J. Namieśnik and A. Kloskowski, A., 2019, “Application of Ionic Liquids in Microextraction Techniques: Current Trends and Future Perspectives”, *Trac - Trends in Analytical Chemistry*, Vol. 119, p. 115614.
- Marsen, B., E. L. Miller, D. Paluselli and R. E. Rocheleau, 2007, “Progress in Sputtered Tungsten Trioxide for Photoelectrode Applications”, *International Journal of Hydrogen Energy*, Vol. 32, pp. 1310–1315.
- Martins, M. A. R., J. A. P. Coutinho, S. P. Pinho and U. Domańska, 2015, “Measurements of Activity Coefficients At infinite Dilution of Organic Solutes and Water on Polar Imidazolium-Based Ionic Liquids”, *Journal of Chemical Thermodynamics*, Vol. 91, pp. 194–203.
- Matheswaran, P., C. D. Wilfred, K. A. Kurnia and A. Ramli, 2016, “Overview of Activity Coefficient of Thiophene at infinite Dilution in Ionic Liquids and Their Modeling Using COSMO-RS”, *Industrial and Engineering Chemistry Research*, Vol. 55, pp. 788–797.
- Melchionna, M. and P. Fornasiero, 2020, “Updates on The Roadmap for Photocatalysis”, *ACS Catalysis*, Vol. 10, pp. 5493–5501.
- Mele, B. and G. Altarelli, 1993, “Lepton Spectra As A Measure of B Quark Polarization At LEP”, *Physics Letters B*, Vol. 299, pp. 345–350.
- Mesbah, M., S. Shahsavari, E. Soroush, N. Rahaei and M. Rezakazemi, 2018, “Accurate Prediction of Miscibility of CO₂ and Supercritical CO₂ in Ionic Liquids Using Machine Learning”, *Journal of CO₂ Utilization*, Vol. 25, pp. 99–107.

- Miao, J., H. B. Yang, S. Y. Khoo and B. Liu, 2013, "Electrochemical Fabrication of ZnO-CdSe Core-Shell Nanorod Arrays for Efficient Photoelectrochemical Water Splitting", *Nanoscale*, Vol. 5, pp. 11118–11124.
- Minggu, L. J., W. R. Wan Daud and M. B. Kassim, 2010, "An Overview of Photocells and Photoreactors for Photoelectrochemical Water Splitting", *International Journal of Hydrogen Energy*, Vol. 35, pp. 5233–5244.
- Modak, B. and S. K. Ghosh, 2016, "An Efficient Strategy for Controlled Band Gap Engineering of Ktao_3 ", *Journal of Physical Chemistry C*, Vol. 120, pp. 6920–6929.
- Mohajernia, S., S. Hejazi, A. Mazare, N. T. Nguyen and P. Schmuki, 2017, "Photoelectrochemical H_2 Generation from Suboxide TiO_2 Nanotubes: Visible-Light Absorption Versus Conductivity", *Chemistry - A European Journal*, Vol. 23, pp. 12406–12411.
- Mohamed, M. A., W. N. Wan Salleh, J. Jaafar and N. Yusof, 2014, "Preparation and Photocatalytic Activity of Mixed Phase Anatase/Rutile TiO_2 Nanoparticles for Phenol Degradation", *Jurnal Teknologi (Sciences and Engineering)*, Vol. 70, pp. 65–70.
- Mohamed, N. A., J. Safaei, A. F. Ismail, M. N. Khalid, M. F. A. Mohd Jailani, M. F. M. Noh and M. A. M. Teridi, 2020, "Boosting Photocatalytic Activities of BiVO_4 By Creation of $\text{g-C}_3\text{N}_4/\text{ZnO}@\text{BiVO}_4$ Heterojunction", *Materials Research Bulletin*, Vol. 125, p. 110779.
- Molak, A. and M. Pilch, 2016, "Visible Light Absorbance Enhanced By Nitrogen Embedded in The Surface Layer of Mn-Doped Sodium Niobate Crystals, Detected by Ultra Violet - Visible Spectroscopy, X-Ray Photoelectron Spectroscopy, and Electric Conductivity Tests", *Journal of Applied Physics*, Vol. 119, p. 204901.
- Momeni, M. M. and Y. Ghayeb, 2016, "Fabrication, Characterization and Photoelectrochemical Performance of Chromium-Sensitized Titania Nanotubes As Efficient Photoanodes for Solar Water Splitting", *Journal of Solid State Electrochemistry*, Vol. 20, pp. 683–689.
- Momeni, M. M., Y. Ghayeb and F. Ezati, 2020, "Investigation of The Morphology,

- Structural, Optical, and Photoelectrochemical Properties of $\text{WO}_3\text{-Fe}_2\text{O}_3/\text{CrTiO}_2$ Thin-Film Photoanodes for Water Splitting", *Applied Physics A: Materials Science and Processing*, Vol. 126, pp. 1–9.
- Mora-Seró, I., V. Likodimos, S. Giménez, E. Martínez-Ferrero, J. Albero, E. Palomares and J. Bisquert, 2010, "Fast Regeneration of CdSe Quantum Dots By Ru Dye in Sensitized TiO_2 Electrodes", *Journal of Physical Chemistry C*, Vol. 114, pp. 6755–6761.
- Müller, W., D. Pallmer, D. Mühlenberg, I. Loumiotis, K. Remoundou, P. Kosmides and K. Demestichas, 2020, "Machine Learning for Discovery Analytics To Support Criminal Investigations", *Big Data II: Learning, Analytics, and Applications*, p. 1139504.
- Naik, B., S. Martha and K. M. Parida, 2011, "Facile Fabrication of $\text{Bi}_2\text{O}_3/\text{TiO}_{2-x}\text{N}_x$ Nanocomposites for Excellent Visible Light Driven Photocatalytic Hydrogen Evolution", *International Journal of Hydrogen Energy*, Vol. 36, pp. 2794–2802.
- Nain, A., S. Paul, and M. Maynar-Reid, 2021, "Keras", <https://keras.io/>, accessed in August 2021.
- Natarajan, K., P. Bhatt, P. Yadav, K. Pandey, B. Tripathi and M. Kumar, 2017, "Investigating The Role of Substrate Tin Diffusion on Hematite Based Photoelectrochemical Water Splitting System", *Journal of Nanoscience and Nanotechnology*, Vol. 18, pp. 1856–1863.
- Nawała, J., B. Dawidziuk, D. Dziedzic, D. Gordon and S. Popiel, 2018, "Applications of Ionic Liquids in Analytical Chemistry with A Particular Emphasis on Their Use in Solid-Phase Microextraction", *Trac - Trends in Analytical Chemistry*, Vol. 105, pp. 18–36.
- Ni, M., M. K. H. Leung, D. Y. C. Leung and K. Sumathy, 2007, "A Review and Recent Developments in Photocatalytic Water-Splitting Using TiO_2 for Hydrogen Production", *Renewable and Sustainable Energy Reviews*, Vol. 11, pp. 401–425.
- Nikolaïdis, P. and A. Poullikkas, 2017, "A Comparative Overview of Hydrogen Production Processes", *Renewable and Sustainable Energy Reviews*, Vol. 67, pp. 597–611.
- NREL, 2020, "Solar Spectra", <http://rredc.nrel.gov/solar/spectra/am1.5/>, accessed in August

2021.

- Nwankpa, C., W. Ijomah, A. Gachagan and S. Marshall, 2018, "Activation Functions: Comparison of Trends in Practice and Research for Deep Learning", *Computer Science*, pp. 1–20.
- O'Mahony, A. M., D. S. Silvester, L. Aldous, C. Hardacre and R. G. Compton, 2008, "Effect of Water on The Electrochemical Window and Potential Limits of Room-Temperature Ionic Liquids", *Journal of Chemical and Engineering Data*, Vol. 53, pp. 2884–2891.
- Obee, T. N. and R. T. Brown, 1995, "TiO₂ Photocatalysis for indoor Air Applications: Effects of Humidity and Trace Contaminant Levels on The Oxidation Rates of formaldehyde, Toluene, and 1,3-Butadiene", *Environmental Science and Technology*, Vol. 29, pp. 1223–1231.
- Odabaşı, Ç., M .E. Günay and R. Yildirim, 2014, "Knowledge Extraction for Water Gas Shift Reaction Over Noble Metal Catalysts from Publications in The Literature Between 2002 and 2012", *International Journal of Hydrogen Energy*, Vol. 39, pp. 5733–5746.
- Odabaşı, Ç. and R. Yıldırım, 2020, "Assessment of Reproducibility, Hysteresis, and Stability Relations in Perovskite Solar Cells Using Machine Learning", *Energy Technology*, Vol. 8, pp. 1–12.
- Olivier-Bourbigou, H., L. Magna and D. Morvan, 2010, "Ionic Liquids and Catalysis: Recent Progress from Knowledge To Applications", *Applied Catalysis A: General*, Vol. 373, pp. 1–56.
- Opallo, M. and A. Lesniewski, 2011, "A Review on Electrodes Modified with Ionic Liquid", *Journal of Electroanalytical Chemistry*, Vol. 656, pp. 2–16.
- Ouaer, H., A. H. Hosseini, M. N. Amar, M. E. A. Seghier, M. A. Ben, Ghriga, N. Nabipour and S. Shamshirband, 2020, "Rigorous Connectionist Models To Predict Carbon Dioxide Solubility in Various Ionic Liquids", *Applied Sciences*, Vol. 10, p. 304.
- Paduszyński, K., 2017, "An Overview of The Performance of The COSMO-RS Approach in Predicting The Activity Coefficients of Molecular Solutes in Ionic Liquids and

- Derived Properties at Infinite Dilution", *Physical Chemistry Chemical Physics*, Vol. 19, pp. 11835–11850.
- Park, J. J., D. Y. Kim, J. G. Lee, Y. H. Cha, M. T. Swihart and S. S. Yoon, 2014, "Supersonic Aerosol-Deposited TiO_2 Photoelectrodes for Photoelectrochemical Solar Water Splitting", *RSC Advances*, Vol. 4, pp. 8661–8670.
- Park, Y., D. Kang and K. S. Choi, 2014, "Marked Enhancement in Electron-Hole Separation Achieved in The Low Bias Region Using Electrochemically Prepared Mo-Doped BiVO_4 Photoanodes", *Physical Chemistry Chemical Physics*, Vol. 16, pp. 1238–1246.
- Peng, L., T. Xie, Y. Lu, H. Fan and D. Wang, 2010, "Synthesis, Photoelectric Properties and Photocatalytic Activity of The $\text{Fe}_2\text{O}_3/\text{TiO}_2$ Heterogeneous Photocatalysts", *Physical Chemistry Chemical Physics*, Vol. 12, pp. 8033–8041.
- Peng, T., D. Ke, P. Cai, K. Dai, L. Ma and L. Zan, 2008, "Influence of Different Ruthenium(II) Bipyridyl Complex on The Photocatalytic H_2 Evolution Over TiO_2 Nanoparticles with Messtructures", *Journal of Power Sources*, Vol. 180, pp. 498–505.
- Pihosh, Y., I. Turkevych, K. Mawatari, J. Uemura, Y. Kazoe, S. Kosar and T. Kitamori, 2015, "Photocatalytic Generation of Hydrogen By Core-Shell $\text{WO}_3/\text{BiVO}_4$ Nanorods with Ultimate Water Splitting Efficiency", *Scientific Reports*, Vol. 5, pp. 1–2.
- Pilli, S. K., T. G. Deutsch, T. E. Furtak, J. A. Turner, L. D. Brown and A. M. Herring, 2012, "Light induced Water Oxidation on Cobalt-Phosphate (Co-Pi) Catalyst Modified Semi-Transparent, Porous $\text{SiO}_2\text{-BiVO}_4$ Electrodes", *Physical Chemistry Chemical Physics*, Vol. 14, pp. 7032–7039.
- Pilli, S. K., T. E. Furtak, L. D. Brown, T. G. Deutsch, J. A. Turner and A. M. Herring, 2011, "Cobalt-Phosphate (Co-Pi) Catalyst Modified Mo-Doped BiVO_4 Photoelectrodes for Solar Water Oxidation", *Energy and Environmental Science*, Vol. 4, pp. 5028–5034.
- Pinheiro, A. N., E. G. S. Firmiano, A. C. Rabelo, C. J. Dalmaschio and E. R. Leite, 2014, "Revisiting SrTiO_3 As A Photoanode for Water Splitting: Development of Thin Films with Enhanced Charge Separation Under Standard Solar Irradiation", *RSC Advances*, Vol. 4, pp. 2029–2036.

- Puangpetch, T., T. Sreethawong and S. Chavadej, 2010a, "Hydrogen Production Over Metal-Loaded Mesoporous-Assembled SrTiO_3 Nanocrystal Photocatalysts: Effects of Metal Type and Loading", *International Journal of Hydrogen Energy*, Vol. 35, pp. 6531–6540.
- Qamar, M., M. Abdalwadoud, M. I. Ahmed, A. M. Azad, B. Merzougui, S. Bukola and M. N. Siddiqui, 2015, "Single-Pot Synthesis of <001> Faceted N-Doped Nb_2O_5 Reduced Graphene Oxide Nanocomposite for Efficient Photoelectrochemical Water Splitting", *ACS Applied Materials and interfaces*, Vol. 7, pp. 17954–17962.
- Qi, L., J. Yu and M. Jaroniec, 2013, "Enhanced and Suppressed Effects of Ionic Liquid on The Photocatalytic Activity of TiO_2 ", *Adsorption*, Vol. 19, pp. 557–561.
- Qin, D. D., C. L. Tao, S. A. Friesen, T. H. Wang, O. K. Varghese, N. Z. Bao and C. A. Grimes, 2012, "Dense Layers of Vertically Oriented WO_3 Crystals As Anodes for Photoelectrochemical Water Oxidation", *Chemical Communications*, Vol. 48, pp. 729–731.
- Qiu, P., F. Li, H. Zhang, S. Wang, Z. Jiang and Y. Chen, 2020, "Photoelectrochemical Performance of $\text{A-Fe}_2\text{O}_3/\text{NiOOH}$ Fabricated with Facile Photo-Assisted Electrodeposition Method", *Electrochimica Acta*, Vol. 358, p. 136847.
- Qureshi, M. and Takanabe, K., 2017, "Insights on Measuring and Reporting Heterogeneous Photocatalysis: Efficiency Definitions and Setup Examples", *Chemistry of Materials*, 29, 158–167.
- Rajeshwar, K., A. Thomas and C. Janaky, 2015, "Photocatalytic Activity of Inorganic Semiconductor Surfaces: Myths, Hype, and Reality", *Journal of Physical Chemistry Letters*, Vol. 6, pp. 139–147.
- Ramachandran, K., M. Geerthana, P. Maadeswaran, B. Liang and R. Ramesh, 2020, "Enhanced Photoelectrochemical Water Splitting Performance of Hematite Photoanodes by Hybrid Microwave Annealing Process", *Optik*, Vol. 212, pp. 164658.
- Ran, J., J. Zhang, J. Yu, M. Jaroniec and S. Z. Qiao, 2014, "Earth-Abundant Cocatalysts for Semiconductor-Based Photocatalytic Water Splitting", *Chemical Society Reviews*, Vol.

43, pp. 7787–7812.

- Ranke, J., A. Othman, P. Fan and A. Müller, 2009, “Explaining Ionic Liquid Water Solubility in Terms of Cation and Anion Hydrophobicity”, *International Journal of Molecular Sciences*, Vol. 10, pp. 1271–1289.
- Ravishankar, T. N., M. De, T. Ramakrishnappa, S. R. Teixeira and J. Dupont, 2019, “Ionic Liquid–Assisted Hydrothermal Synthesis of Nb/TiO₂ Nanocomposites for Efficient Photocatalytic Hydrogen Production and Photodecolorization of Rhodamine B Under UV-Visible and Visible Light Illumination”, *Materials Today Chemistry*, Vol. 12, pp. 373–385.
- Ravishankar, T. N., M. D. O. Vaz, T. Ramakrishnappa, S. R. Teixeira and J. Dupont, 2017, “Ionic Liquid Assisted Hydrothermal Syntheses of Au Doped TiO₂ NPs for Efficient Visible-Light Photocatalytic Hydrogen Production from Water, Electrochemical Detection and Photochemical Detoxification of Hexavalent Chromium (Cr⁶⁺)”, *RSC Advances*, Vol. 7, pp. 43233–43244.
- Reddy, I. N., A. Sreedhar, J. Shim and J. S. Gwag, 2019, “Multifunctional Monoclinic VO₂ Nanorod Thin Films for Enhanced Energy Applications: Photoelectrochemical Water Splitting and Supercapacitor”, *Journal of Electroanalytical Chemistry*, Vol. 835, pp. 40–47.
- Reese, M. O., S. A. Gevorgyan, M. Jørgensen, E. Bundgaard, S. R. Kurtz, D. S. Ginley and F. C. Krebs, 2011, “Consensus Stability Testing Protocols for Organic Photovoltaic Materials and Devices”, *Solar Energy Materials and Solar Cells*, Vol. 95, pp. 1253–1267.
- Rieland, J. M. and B. J. Lowe, 2020, “Ionic Liquids: A Milestone on The Pathway To Greener Recycling of Cellulose from Biomass”, *Resources, Conservation and Recycling*, Vol. 155, p. 104678.
- Roy, P., C. Das, K. Lee, R. Hahn, T. Ruff, M. Moll and P. Schmuki, 2011, “Oxide Nanotubes on Ti-Ru Alloys: Strongly Enhanced and Stable Photoelectrochemical Activity for Water Splitting”, *Journal of The American Chemical Society*, Vol. 133, pp. 5629–5631.

- Saadetnejad, D. and R. Yıldırım, 2018, "Photocatalytic Hydrogen Production by Water Splitting Over Au/Al-SrTiO₃", *International Journal of Hydrogen Energy*, Vol. 43, pp. 1116–1122.
- Safavi, A., Z. Movahedi, D. Mohajer and N. Maleki, 2006, "Dramatic Effects of Ionic Liquid on Platinum Electrode Surface and Electron-Transfer Rates of Meso-Tetraphenylporphyrins", *Electroanalysis*, Vol. 18, pp. 1227–1229.
- Saito, R., Y. Miseki and K. Sayama, 2012, "Highly Efficient Photoelectrochemical Water Splitting Using A Thin Film Photoanode of BiVO₄/SnO₂/WO₃ Multi-Composite in A Carbonate Electrolyte", *Chemical Communications*, Vol. 48, pp. 3833–3835.
- Sajid, M., 2019, "Magnetic Ionic Liquids in Analytical Sample Preparation: A Literature Review", *Trac - Trends in Analytical Chemistry*, Vol. 113, pp. 210–223.
- Sánchez-Tovar, R., E. Blasco-Tamarit, R. M. Fernández-Domene, B. Lucas-Granados and J. García-Antón, 2017, "Should TiO₂ Nanostructures Doped with Li⁺ Be Used As Photoanodes for Photoelectrochemical Water Splitting Applications?", *Journal of Catalysis*, Vol. 349, pp. 41–52.
- Sánchez-Tovar, R., E. Blasco-Tamarit, R. M. Fernández-Domene, M. Villanueva-Pascual and J. García-Antón, 2020, "Electrochemical formation of Novel TiO₂-ZnO Hybrid Nanostructures for Photoelectrochemical Water Splitting Applications", *Surface and Coatings Technology*, Vol. 388, p. 125605.
- Schwarze, M., D. Stellmach, M. Schröder, K. Kailasam, R. Reske, A. Thomas and R. Schomäcker, 2013, "Quantification of Photocatalytic Hydrogen Evolution", *Physical Chemistry Chemical Physics*, Vol. 15, pp. 3466–3472.
- Şener, A. N., M. E. Günay, A. Leba and R. Yıldırım, 2018, "Statistical Review of Dry Reforming of Methane Literature Using Decision Tree and Artificial Neural Network Analysis", *Catalysis Today*, Vol. 299, pp. 289–302.
- Senthil, V., T. Badapanda, A. Chithambararaj, A. Chandra Bose and S. Panigrahi, 2016, "Impedance Spectroscopy and Photocatalysis Water Splitting for Hydrogen Production with Cerium Modified SrBi₂Ta₂O₉ Ferroelectrics", *International Journal of Hydrogen*

Energy, Vol. 41, pp. 22856–22865.

- Serpone, N., A. V. Emeline, V. K. Ryabchuk, V. N. Kuznetsov, Y. M. Artem'Ev and S. Horikoshi, 2016, "Why Do Hydrogen and Oxygen Yields from Semiconductor-Based Photocatalyzed Water Splitting Remain Disappointingly Low? intrinsic and Extrinsic Factors Impacting Surface Redox Reactions", *ACS Energy Letters*, Vol. 1, pp. 931–948.
- Serpone, N., G. Sauvé, R. Koch, H. Tahiri, P. Pichat, P. Piccinini and H. Hidaka, 1996, "Standardization Protocol of Process Efficiencies and Activation Parameters in Heterogeneous Photocatalysis: Relative Photonic Efficiencies", *Journal of Photochemistry and Photobiology A: Chemistry*, Vol. 94, pp. 191–203.
- Sfaelou, S., L. C. Pop, O. Monfort, V. Dracopoulos and P. Lianos, 2016, "Mesoporous WO₃ Photoanodes for Hydrogen Production By Water Splitting and Photofuelcell Operation", *International Journal of Hydrogen Energy*, Vol. 41, pp. 5902–5907.
- Shaban, Y. A. and S. U. M. Khan, 2007, "Surface Grooved Visible Light Active Carbon Modified (CM)-N-TiO₂ Thin Films for Efficient Photoelectrochemical Splitting of Water", *Chemical Physics*, Vol. 339, pp. 73–85.
- Shaban, Y. A. and S. U. M. Khan, 2008, "Visible Light Active Carbon Modified N-TiO₂ for Efficient Hydrogen Production by Photoelectrochemical Splitting of Water", *International Journal of Hydrogen Energy*, Vol. 33, pp. 1118–1126.
- Shafiei, A., M. A. Ahmadi, S. H. Zaheri, A. Baghban, A. Amirfakhrian and R. Soleimani, 2014, "Estimating Hydrogen Sulfide Solubility in Ionic Liquids Using A Machine Learning Approach", *Journal of Supercritical Fluids*, Vol. 95, pp. 525–534.
- Shah, R., 2021, "Introduction To K-Nearest Neighbors (KNN) Algorithm", <https://ai.plainenglish.io/introduction-to-k-nearest-neighbors-knn-algorithm-e8617a448fa8>, accessed in August 2021.
- Shaheen, B. S., A. M. Hafez, B. Murali, A. R. Kirmani, A. Amassian, O. F. Mohammed and N. K. Allam, 2016, "10-Fold Enhancement in Light-Driven Water Splitting Using Niobium Oxynitride Microcone Array Films", *Solar Energy Materials and Solar Cells*,

Vol. 151, pp. 149–153.

- Sharma, K., V. Sharma and S. S. Sharma, 2018, “Dye-Sensitized Solar Cells: Fundamentals and Current Status”, *Nanoscale Research Letters*, Vol. 13, p. 381.
- Sharma, V., M. Dixit, V. R. Satsangi, S. Dass, S. Pal and R. Shrivastav, R., 2014, “Photoelectrochemical Splitting of Water with Nanocrystalline $\text{Zn}_{1-x}\text{Mn}_x\text{O}$ Thin Films: First-Principle DFT Computations Supporting The Systematic Experimental Endeavor”, *International Journal of Hydrogen Energy*, Vol. 39, pp. 3637–3648.
- Shen, S., J. Jiang, P. Guo and L. Guo, 2013, “Facile Growth of Porous Hematite Films for Photoelectrochemical Water Splitting”, *International Journal of Photoenergy*, Vol. 2013, pp. 1–9.
- Shi, J. and L. Guo, 2012, “ ABO_3 -Based Photocatalysts for Water Splitting”, *Progress in Natural Science: Materials International*, Vol. 22, pp. 592–615.
- Shi, J., J. Ye, Q. Li, Z. Zhou, H. Tong, G. Xi and L. Guo, 2012, “Single-Crystal Nanosheet-Based Hierarchical AgSbO_3 with Exposed {001} Facets: Topotactic Synthesis and Enhanced Photocatalytic Activity”, *Chemistry - A European Journal*, Vol. 18, pp. 3157–3162.
- Shi, J., J. Ye, Z. Zhou, M. Li and L. Guo, 2011, “Hydrothermal Synthesis of $\text{Na}_{0.5}\text{La}_{0.5}\text{TiO}_3$ - LaCrO_3 Solid-Solution Single-Crystal Nanocubes for Visible-Light-Driven Photocatalytic H_2 Evolution”, *Chemistry - A European Journal*, Vol. 17, pp. 7858–7867.
- Shi, W., X. Lv and Y. Shen, 2018, “ BiOI/WO_3 Photoanode with Enhanced Photoelectrochemical Water Splitting Activity”, *Frontiers of Optoelectronics*, Vol. 11, pp. 367–374.
- Shi, X., L. Cai, M. Ma, X. Zheng and J. H. Park, 2015, “General Characterization Methods for Photoelectrochemical Cells for Solar Water Splitting”, *Chemsuschem*, Vol. 8, pp. 3192–3203.
- Shi, X., K. Zhang, K. Shin, J. H. Moon, T. W. Lee and J. H. Park, 2013, “Constructing Inverse Opal Structured Hematite Photoanodes Via Electrochemical Process and Their

- Application To Photoelectrochemical Water Splitting", *Physical Chemistry Chemical Physics*, Vol. 15, pp. 11717–11722.
- Shibli, S. M. A., P. S. Arun and A. V. Raj, 2015, "Exploration of Octahedrally Shaped MnCo_2O_4 Catalyst Particles for Visible Light Driven Photocatalytic Water Splitting Reaction", *RSC Advances*, Vol. 5, pp. 19393–19399.
- Shimizu, K. I., S. Itoh, T. Hatamachi, T. Kodama, M. Sato K. and Toda, 2005, "Photocatalytic Water Splitting on Ni-Intercalated Ruddlesden-Popper Tantalate $\text{H}_2\text{La}_{2/3}\text{Ta}_2\text{O}_7$ ", *Chemistry of Materials*, Vol. 17, pp. 5161–5166.
- Singh, A. P., S. Kumari, R. Shrivastav, S. Dass and V. R. Satsangi, 2008, "Iron Doped Nanostructured TiO_2 for Photoelectrochemical Generation of Hydrogen", *International Journal of Hydrogen Energy*, Vol. 33 pp. 5363–5368.
- Skoda-Földes, R., 2014, "The Use of Supported Acidic Ionic Liquids in Organic Synthesis", *Molecules*, Vol. 19, pp. 8840–8884.
- Snyder, J., T. Fujita, M. W. Chen and J. Erlebacher, 2010, "Oxygen Reduction in Nanoporous Metal-Ionic Liquid Composite Electrocatalysts", *Nature Materials*, Vol. 9, pp. 904–907.
- Soares, E. T., M. A. Lansarin and C. C. Moro, 2007, "A Study of Process Variables for The Photocatalytic Degradation of Rhodamine B", *Brazilia Journal of Chemical Engineering*, Vol. 24, pp. 29–36.
- Soares, R. D. P., 2011, "The Combinatorial Term for COSMO-Based Activity Coefficient Models", *industrial and Engineering Chemistry Research*, Vol. 50, pp. 3060–3063.
- Song, K., R. Zhu, F. Tian, G. Cao and F. Ouyang, 2015, "Effects of indium Contents on Photocatalytic Performance of ZnIn_2S_4 for Hydrogen Evolution Under Visible Light", *Journal of Solid State Chemistry*, Vol. 232, pp. 138–143.
- Stanescu, D., M. Piriyeu, V. Villard, C. Mocuta, A. Besson, D. Ihiwakrim and S. Stanescu, 2020, "Characterizing Surface States in Hematite Nanorod Photoanodes, Both Beneficial and Detrimental To Solar Water Splitting Efficiency", *Journal of Materials Chemistry A*, Vol. 8, pp. 20513–20530.

- Su, F., J. Lu, Y. Tian, X. Ma and J. Gong, 2013, "Branched TiO₂ Nanoarrays Sensitized with Cds Quantum Dots for Highly Efficient Photoelectrochemical Water Splitting", *Physical Chemistry Chemical Physics*, Vol. 15, pp. 12026–12032.
- Su, F., T. Wang, R. Lv, J. Zhang, P. Zhang, J. Lu and J. Gong, 2013, "Dendritic Au/TiO₂ Nanorod Arrays for Visible-Light Driven Photoelectrochemical Water Splitting", *Nanoscale*, Vol. 5, pp. 9001–9009.
- Subramanian, A., Z. Pan, H. Li, L. Zhou, W. Li, Y. Qiu and Y. Zhang, 2017, "Synergistic Promotion of Photoelectrochemical Water Splitting Efficiency of TiO₂ Nanorods Using Metal-Semiconducting Nanoparticles", *Applied Surface Science*, Vol. 420, pp. 631–637.
- Subramanyam, P., T. Khan, G. Neeraja Sinha, D. Suryakala and C. Subrahmanyam, 2020, "Plasmonic Bi Nanoparticle Decorated BiVO₄/RGO As An Efficient Photoanode for Photoelectrochemical Water Splitting", *International Journal of Hydrogen Energy*, Vol. 45, pp. 7779–7787.
- Sun, Q., T. Cheng, Z. Liu and L. Qi, 2020, "A Cobalt Silicate Modified BiVO₄ Photoanode for Efficient Solar Water Oxidation", *Applied Catalysis B: Environmental*, Vol. 277, p. 119189.
- Sun, W. and H. Xu, H., 2016, "Enhanced Photovoltaic Performances of Novel All-Ionic Liquid integrated Poly(Acrylic Acid/Polyethylene Glycol) Gel Electrolytes in Quasi-Solid-State Dye-Sensitized Solar Cells", *RSC Advances*, Vol. 6, pp. 68193–68198.
- Sun, Y., C. J. Murphy, K. R. Reyes-Gil, E. A. Reyes-Garcia, J. P. Lilly and D. Raftery, 2008, "Carbon-Doped In₂O₃ Films for Photoelectrochemical Hydrogen Production", *International Journal of Hydrogen Energy*, Vol. 33, pp. 5967–5974.
- Sun, Z., G. Fang, J. Li, J. Mo, X. He, X. Wang and Z. Yu, 2020, "Preparation of (Ti, Zr) Co-Doped Hematite Photoanode for Enhanced Photoelectrochemical Water Splitting", *Chemical Physics Letters*, Vol. 754, p. 137736.
- Takata, T. and K. Domen, 2009, "Defect Engineering of Photocatalysts By Doping of Aliovalent Metal Cations for Efficient Water Splitting", *Journal of Physical Chemistry*

- C, Vol. 113, pp. 19386–19388.
- Takata, T. and K. Domen, 2017, “Development of Non-Oxide Semiconductors As Light Harvesting Materials in Photocatalytic and Photoelectrochemical Water Splitting”, *Dalton Transactions*, Vol. 46, pp. 10529–10544.
- Takata, T., J. Jiang, Y. Sakata, M. Nakabayashi, N. Shibata, V. Nandal and K. Domen, 2020, “Photocatalytic Water Splitting with A Quantum Efficiency of Almost Unity”, *Nature*, Vol. 581, pp. 411–414.
- Takata, T., C. Pan and K. Domen, 2015, “Recent Progress in Oxynitride Photocatalysts for Visible-Light-Driven Water Splitting”, *Science and Technology of Advanced Materials*, Vol. 16, p. 33506.
- Takayama, T., K. Tanabe, K. Saito, A. Iwase and A. Kudo, 2014, “The $\text{KCaSrTa}_5\text{O}_{15}$ Photocatalyst with Tungsten Bronze Structure for Water Splitting and CO_2 Reduction”, *Physical Chemistry Chemical Physics*, Vol. 16, pp. 24417–24422.
- Tan, X., J. Zhang, D. Tan, J. Shi, X. Cheng, F. Zhang and B. Han, 2019, “Ionic Liquids Produce Heteroatom-Doped Pt/TiO_2 Nanocrystals for Efficient Photocatalytic Hydrogen Production”, *Nano Research*, Vol. 12, pp.1967–1972.
- Tapan, N. A., M. E. Günay and R. Yildirim, 2016, “Constructing Global Models from Past Publications To Improve Design and Operating Conditions for Direct Alcohol Fuel Cells”, *Chemical Engineering Research and Design*, Vol. 105, pp. 162–170.
- Teixeira, I. F., J. Quiroz, M. S. Homsí and P. H. C. Camargo, 2020, “An Overview of The Photocatalytic H_2 Evolution By Semiconductor-Based Materials for Nonspecialists”, *Journal of The Brazilian Chemical Society*, Vol. 31, pp. 211–229.
- Therneau, T., B. Atkinson and B. Ripley, 2018, “*Package Rpart*”, <https://cran.r-project.org/web/packages/rpart/index.html>, accessed in August 2021.
- Toor, R. A., M. H. Sayyad, N. Nasr, S. Sajjad, S. A. A. Shah and T. Manzoor, 2016, “Efficiency Enhancement of Dye Sensitized Solar Cells with A Low Cost Co-Adsorbant in N719 Dye”, *International Journal of Sustainable Energy and Environmental Research*, Vol. 5, pp. 46–50.

- Torres-Martínez, L. M., R. Gómez, O. Vázquez-Cuchillo, I. Juárez-Ramírez, A. Cruz-López and F. J. Alejandre-Sandoval, 2010, "Enhanced Photocatalytic Water Splitting Hydrogen Production on $\text{RuO}_2/\text{La:NaTaO}_3$ Prepared by Sol-Gel Method", *Catalysis Communications*, Vol. 12, pp. 268–272.
- Toyao, T., M. Minakata, K. Iyatani, A. Ebrahimi, P. C. Chen, C. B. Tsai and M. Anpo, M., 2013, "Development of Dye-Sensitized Solar Cells Based on Visible-Light-Responsive TiO_2 Thin Films with A Unique Columnar Structure", *Research on Chemical intermediates*, Vol. 39, pp. 415–424.
- Tsai, C. C. and H. Teng, 2008, "Chromium-Doped Titanium Dioxide Thin-Film Photoanodes in Visible-Light-Induced Water Cleavage", *Applied Surface Science*, Vol. 254, pp. 4912–4918.
- Ullah, R., M. Pei, J. Wu, Y. Tian, Z. Gu, Q. Zhang, J. Zhang, 2020, "Bifunctional Photoelectrode Driven by Charged Domain Walls in Ferroelectric Bi_2WO_6 ", *ACS Applied Energy Materials*, Vol. 3, pp. 4149–4154.
- Van Doorslaer, C., J. Wahlen, P. Mertens, K. Binnemans and D. De Vos, 2010, "Immobilization of Molecular Catalysts in Supported Ionic Liquid Phases", *Dalton Transactions*, Vol. 39, pp. 8377–8390.
- Venkatraman, V. and B. K. Alsberg, 2017, "Predicting CO_2 Capture of Ionic Liquids Using Machine Learning", *Journal of CO2 Utilization*, Vol. 21, pp. 162–168.
- Verma, C., A. Mishra, S. Chauhan, P. Verma, V. Srivastava, M. A. Quraishi and E. E. Ebenso, 2019, "Dissolution of Cellulose in Ionic Liquids and Their Mixed Cosolvents: A Review", *Sustainable Chemistry and Pharmacy*, Vol. 13, p. 100162.
- Wadhawan, J. D., U. Schröder, A. Neudeck, S. J. Wilkins, R. G. Compton, F. Marken and J. Dupont, 2000, "Ionic Liquid Modified Electrodes. Unusual Partitioning and Diffusion Effects of $\text{Fe}(\text{CN})_6^{4-/3-}$ in Droplet and Thin Layer Deposits of 1-Methyl-3-(2,6-(s)-Dimethylocten-2-yl)-Imidazolium Tetrafluoroborate", *Journal of Electroanalytical Chemistry*, Vol. 493, pp. 75–83.
- Wang, D. H., L. Wang and A. W. Xu, 2012, "Room-Temperature Synthesis of $\text{Zn}_{0.80}\text{Cd}_{0.20}\text{S}$

- Solid Solution with A High Visible-Light Photocatalytic Activity for Hydrogen Evolution", *Nanoscale*, Vol. 4, pp. 2046–2053.
- Wang, D., J. Ye, T. Kako and T. Kimura, T., 2010, "Photophysical and Photocatalytic Properties of SrTiO₃ Doped with Cr Cations on Different Sites", *Journal of Physical Chemistry B*, Vol. 110, pp. 15824–15830.
- Wang, D., J. Huang, F. Liu, X. Xu, X. Fang, J. Liu and X. Wang, 2020, "Rutile RuO₂ Dispersion on Rutile and Anatase TiO₂ Supports: The Effects of Support Crystalline Phase Structure on The Dispersion Behaviors of The Supported Metal Oxides", *Catalysis Today*, Vol. 339, pp. 220–232.
- Wang, F., Z. Zheng and F. Jia, 2012, "Enhanced Photoelectrochemical Water Splitting on Pt-Loaded TiO₂ Nanorods Array Thin Film", *Materials Letters*, Vol. 71, pp. 141–144.
- Wang, G., Y. Ling, H. Wang, X. Yang, C. Wang, J. Z. Zhang and Y. Li, 2012, "Hydrogen-Treated WO₃ Nanoflakes Show Enhanced Photostability", *Energy and Environmental Science*, Vol. 5, pp. 6180–6187.
- Wang, Lei, C. Y. Lee and P. Schmuki, 2013, "Solar Water Splitting: Preserving The Beneficial Small Feature Size in Porous A-Fe₂O₃ Photoelectrodes During Annealing", *Journal of Materials Chemistry A*, Vol.1, pp. 212–215.
- Wang, Luand Wang, W., 2012, "Photocatalytic Hydrogen Production from Aqueous Solutions over Novel Bi_{0.5}Na_{0.5}TiO₃ Microspheres", *International Journal of Hydrogen Energy*, Vol. 37, pp. 3041–3047.
- Wang, M., M. Fang, X. Min, Z. Huang, C. Tang, Y. Liu and X. Wu, 2017, "Molten Salt Synthesis of NaNb_xTa_{1-x}O₃ Perovskites with Enhanced Photocatalytic Activity", *Chemical Physics Letters*, Vol. 686, pp. 18–25.
- Wang, Q. and K. Domen, 2020, "Particulate Photocatalysts for Light-Driven Water Splitting: Mechanisms, Challenges, and Design Strategies", *Chemical Reviews*, Vol. 120, pp. 919–985.
- Wang, X, Z. Liu and Z. Liu, 2017, "A Dumbbell CaBi₂O₄ Photoelectrode for Photoelectrochemical Water Splitting", *Chemcatchem*, Vol. 9, pp. 4029–4034.

- Wang, Xuewen, G. Liu, L. Wang, J. Pan, G. Q. Lu and H. M. Cheng, 2011, "TiO₂ Films with Oriented Anatase {001} Facets and Their Photoelectrochemical Behavior As CdS Nanoparticle Sensitized Photoanodes", *Journal of Materials Chemistry*, Vol. 21, pp. 869–873.
- Wang, Y., L. Jiang, Y. Liu, D. Tang, F. Liu and Y. Lai, 2017, "Facile Synthesis and Photoelectrochemical Characterization of Sb₂O₃ Nanoprism Arrays", *Journal of Alloys and Compounds*, Vol. 727, pp. 469–474.
- Watanabe, M., 2017, "Dye-Sensitized Photocatalyst for Effective Water Splitting Catalyst", *Science and Technology of Advanced Materials*, Vol. 18, pp. 705–723.
- Wei, J., C. Zhou, Y. Xin, X. Li, L. Zhao and Z. Liu, 2018, "Cooperation Effect of Heterojunction and Co-Catalyst in BiVO₄/Bi₂S₃/NiOOH Photoanode for Improving Photoelectrochemical Performances", *New Journal of Chemistry*, Vol. 42, pp. 19415–19422.
- Wei, L., Y. Yang, R. Fan, Y. Na, P. Wang and Y. Dong, 2015, "Effects of Rubrene Co-Sensitized TiO₂ Photoanode on The Performance of Ruthenium Dye N719 Sensitized Solar Cells", *Thin Solid Films*, Vol. 592, pp. 14–23.
- Wei, T., V. Simko, M. Levy, Y. Xie, Y. Jin, J. Zemla and T. Protivinsky, 2021, "Package Corrplot ", <https://github.com/taiyun/corrplot>, accessed in August 2021.
- Wei, Y., J. Li, Y. Huang, M. Huang, J. Lin and J. Wu, 2009, "Photocatalytic Water Splitting with In-Doped H₂LaNb₂O₇ Composite Oxide Semiconductors", *Solar Energy Materials and Solar Cells*, Vol. 93, pp. 1176–1181.
- Weingärtner, H., 2008, "Understanding Ionic Liquids at The Molecular Level: Facts, Problems, and Controversies. *Angewandte Chemie - International Edition*, Vol. 47, pp. 654–670.
- Wender, H., A. F. Feil, L. B. Diaz, C. S. Ribeiro, G. J. Machado, P. Migowski and S. R. Teixeira, 2011, "Self-Organized TiO₂ Nanotube Arrays: Synthesis By Anodization in An Ionic Liquid and Assessment of Photocatalytic Properties", *ACS Applied Materials and interfaces*, Vol. 3, pp. 1359–1365.

- Weng, B., F. Xu and J. Xu, 2014, "Core-Shell Photoanode Developed By Atomic Layer Deposition of Bi_2O_3 on Si Nanowires for Enhanced Photoelectrochemical Water Splitting", *Nanotechnology*, Vol. 25, p. 455402.
- Wetchakun, N., B. Incessungvorn, K. Wetchakun and S. Phanichphant, 2012, "Influence of Calcination Temperature on Anatase To Rutile Phase Transformation in TiO_2 Nanoparticles Synthesized by The Modified Sol-Gel Method", *Materials Letters*, Vol. 82, pp. 195–198.
- Williams, G., M. V. Culp, E. Cox, A. Nolan, D. White, D. Medri and C. Chisholm, 2020, "Package Rattle ", <https://cran.r-project.org/web/packages/rattle/index>, accessed in August 2021.
- Wu, P., J. Shi, Z. Zhou, W. Tang and L. Guo, 2012, "CaTaO₂N-CaZrO₃ Solid Solution: Band-Structure Engineering and Visible-Light-Driven Photocatalytic Hydrogen Production", *International Journal of Hydrogen Energy*, Vol. 37, pp. 13704–13710.
- Wu, Z., G. Li, F. Zhang and W. Zhang, 2014, "Photocatalytic Activity of NaTaO₃:La Prepared Under Different Atmospheres", *Applied Surface Science*, Vol. 319, pp. 372–375.
- Xiao, J., X. Hou, L. Zhao and Y. Li, 2017, "A Carbon-Quantum-Dot-Sensitized ZnO:Ga/ZnO Multijunction Composite Photoanode for Photoelectrochemical Water Splitting Under Visible Light Irradiation", *Journal of Catalysis*, Vol. 346, pp. 70–77.
- Xu, Chen, Y. Song, L. Lu, C. Cheng, D. Liu, X. Fang and D. Li, 2013, "Electrochemically Hydrogenated TiO_2 Nanotubes with Improved Photoelectrochemical Water Splitting Performance. *Nanoscale Research Letters*, Vol. 8, pp. 1–7.
- Xu, C., Y. A. Shaban, W. B. Ingler and S. U. M. Khan, 2007, "Nanotube Enhanced Photoresponse of Carbon Modified (CM)-N-TiO₂ for Efficient Water Splitting", *Solar Energy Materials and Solar Cells*, Vol. 91, pp. 938–943.
- Xu, D., S. Yang, Y. Jin, M. Chen, W. Fan, B. Luo and W. Shi, 2015, "Ag-Decorated ATaO₃ (A = K, Na) Nanocube Plasmonic Photocatalysts with Enhanced Photocatalytic Water-Splitting Properties", *Langmuir*, Vol. 31, pp. 9694–9699.

- Yang, D., Q. Zhu and B. Han, 2020, "Electroreduction of CO₂ in Ionic Liquid-Based Electrolytes", *Innovation (United States)*, Vol.1, p. 100016.
- Yang, L. and Z. Liu, 2007, "Study on Light intensity in The Process of Photocatalytic Degradation of indoor Gaseous formaldehyde for Saving Energy", *Energy Conversion and Management*, Vol. 48, pp. 882–889
- Ye, K. H., J. Y. Wang, N. Li, Z. Q. Liu, S. H. Guo, Y. P. Guo and Y. Z. Su, 2014, "A Facile Way To Synthesize Er₂O₃@ZnO Core-Shell Nanorods for Photoelectrochemical Water Splitting", *inorganic Chemistry Communications*, Vol. 45, pp. 116–119.
- Yin, J., Z. Zou and J. Ye, 2003a, "A Novel Series of The New Visible-Light-Driven Photocatalysts MCo_{1/3}Nb_{2/3}O₃ (M = Ca, Sr, and Ba) with Special Electronic Structures", *Journal of Physical Chemistry B*, Vol. 107, pp. 4936–4941.
- Yin, J., Z. Zou and J. Ye, 2003b, "Photophysical and Photocatalytic Properties of MIn_{0.5}Nb_{0.5}O₃ (M = Ca, Sr, and Ba)", *The Journal of Physical Chemistry B*, Vol. 107, pp. 61–65.
- Yılmaz, B. and R. Yıldırım, 2021, "Critical Review of Machine Learning Applications in Perovskite Solar Research", *Nano Energy*, Vol. 80, p. 105546.
- Yoshida, H., M. Takeuchi, M. Sato, L. Zhang, T. Teshima and M. G. Chaskar, 2014, "Potassium Hexatitanate Photocatalysts Prepared by A Flux Method for Water Splitting", *Catalysis Today*, Vol. 232, pp. 158–164.
- Yourey, J. E. and B. M. Bartlett, 2011, "Electrochemical Deposition and Photoelectrochemistry of CuWO₄, A Promising Photoanode for Water Oxidation", *Journal of Materials Chemistry*, Vol. 21, pp. 7651–7660.
- Yu, H., S. Ouyang, S. Yan, Z. Li, T. Yu and Z. Zou, 2011, "Sol-Gel Hydrothermal Synthesis of Visible-Light-Driven Cr-Doped SrTiO₃ for Efficient Hydrogen Production", *Journal of Materials Chemistry*, Vol. 21, pp. 11347–11351.
- Yu, J. M., J. Lee, Y. S. Kim, J. Song, J. Oh, S. M. Lee and J. W. Jang, 2020, "High-Performance and Stable Photoelectrochemical Water Splitting Cell with Organic-Photoactive-Layer-Based Photoanode", *Nature Communications*, Vol.11, pp. 1–9.

- Yu, J., Y. Su, B. Cheng and M. Zhou, 2006, "Effects of pH on The Microstructures and Photocatalytic Activity of Mesoporous Nanocrystalline Titania Powders Prepared via Hydrothermal Method", *Journal of Molecular Catalysis A: Chemical*, Vol. 258, pp. 104–112.
- Yu, J., G. Wang, B. Cheng and M. Zhou, 2007, "Effects of Hydrothermal Temperature and Time on The Photocatalytic Activity and Microstructures of Bimodal Mesoporous TiO₂ Powders", *Applied Catalysis B: Environmental*, Vol. 69, pp. 171–180.
- Yu, P., Y. Lin, L. Xiang, L. Su, J. Zhang and L. Mao, 2005, "Molecular Films of Water-Miscible Ionic Liquids formed on Glassy Carbon Electrodes: Characterization and Electrochemical Applications", *Langmuir*, Vol.21, pp. 9000–9006.
- Yu, X., X. An, A. Shavel, M. Ibáñez and A. Cabot, 2014, "The Effect of The Ga Content on The Photocatalytic Hydrogen Evolution of CuIn_{1-x}Ga_xS₂ Nanocrystals", *Journal of Materials Chemistry A*, Vol. 2, pp. 12317–12322.
- Yu, Z., F. Li and L. Sun, 2015, "Recent Advances in Dye-Sensitized Photoelectrochemical Cells for Solar Hydrogen Production Based on Molecular Components", *Energy and Environmental Science*, Vol. 8, pp. 760–775.
- Yuan, Y., J. Gu, K. H. Ye, Z. Chai, X. Yu, X. Chen and W. Mai, 2016, "Combining Bulk/Surface Engineering of Hematite To Synergistically Improve Its Photoelectrochemical Water Splitting Performance", *ACS Applied Materials and interfaces*, Vol.8, pp. 16071–16077.
- Zambrano-Bigiarini, M., 2020, "Package Hydrogof", <https://github.com/hzambran/hydrogof>, accessed in August 2021.
- Zeng, C., T. Hu, N. Hou, S. Liu, W. Gao, R. Cong and T. Yang, 2015, "Photocatalytic Pure Water Splitting Activities for ZnGa₂O₄ Synthesized by Various Methods", *Materials Research Bulletin*, Vol. 61, pp. 481–485.
- Zhan, C. G., J. A. Nichols and D. A. Dixon, 2003, "Ionization Potential, Electron Affinity, Electronegativity, Hardness, and Electron Excitation Energy: Molecular Properties from Density Functional Theory Orbital Energies", *Journal of Physical Chemistry A*,

Vol. 107, pp. 4184–4195.

- Zhang, G. R. and B. J. M. Etzold, 2016, “Ionic Liquids in Electrocatalysis”, *Journal of Energy Chemistry*, Vol. 25, pp. 199–207.
- Zhang, H., G. Chen, X. Li and Q. Wang, 2009, “Electronic Structure and Water Splitting Under Visible Light Irradiation of $\text{BiTa}_{1-x}\text{Cu}_x\text{O}_4$ ($x = 0.00-0.04$) Photocatalysts”, *International Journal of Hydrogen Energy*, Vol. 34, pp. 3631–3638.
- Zhang, K., X. J. Shi, J. K. Kim and J. H. Park, 2012, “Photoelectrochemical Cells with Tungsten Trioxide/Mo-Doped BiVO_4 Bilayers”, *Physical Chemistry Chemical Physics*, Vol. 14, pp. 11119–11124.
- Zhang, L., Y. Song, J. Feng, T. Fang, Y. Zhong, Z. Li and Z. Zou, 2014, “Photoelectrochemical Water Oxidation of LaTaON_2 Under Visible-Light Irradiation”, *International Journal of Hydrogen Energy*, Vol. 39, pp. 7697–7704.
- Zhang, Q., Z. Li, S. Wang, R. Li, X. Zhang, Z. Liang and C. Li, 2016, “Effect of Redox Cocatalysts Location on Photocatalytic Overall Water Splitting over Cubic NaTaO_3 Semiconductor Crystals Exposed with Equivalent Facets”, *ACS Catalysis*, Vol. 6, pp. 2182–2191.
- Zhang, S., M. Feng, Y. Liu and D. Wang, 2020, “ Ta_2O_5 NTs- TiO_2 Nanodots Heterostructure Photocatalyst Material for Enhanced Photodegradation and Photoelectrochemical Performance Under Simulated Solar Light”, *Journal of Nanoparticle Research*, Vol. 22, p. 375.
- Zhang, S, G. She, S. Li, L. Mu and W. Shi, 2018, “Si-H induced Synthesis of $\text{Si/Cu}_2\text{O}$ Nanowire Arrays for Photoelectrochemical Water Splitting”, *Nanotechnology*, Vol. 29, p. 035601.
- Zhang, T., K. Zhao, J. Yu, J. Jin, Y. Qi, H. Li and G. Liu, 2013, “Photocatalytic Water Splitting for Hydrogen Generation on Cubic, Orthorhombic, and Tetragonal KNbO_3 Microcubes”, *Nanoscale*, Vol. 5, pp. 8375–8383.
- Zhang, W., W. Wang, H. Shi, Y. Liang, J. Fu and M. Zhu, 2018, “Surface Plasmon-Driven Photoelectrochemical Water Splitting of Aligned ZnO Nanorod Arrays Decorated with

- Loading-Controllable Au Nanoparticles", *Solar Energy Materials and Solar Cells*, Vol. 180, pp. 25–33.
- Zhang, X, T. Peng and S. Song, 2016, "Recent Advances in Dye-Sensitized Semiconductor Systems for Photocatalytic Hydrogen Production", *Journal of Materials Chemistry A*, Vol. 4, pp. 2365–2402.
- Zhang, X., U. Veikko, J. Mao, P. Cai and T. Peng, 2012, "Visible-Light-Induced Photocatalytic Hydrogen Production Over Binuclear Ru II-Bipyridyl Dye-Sensitized TiO₂ Without Noble Metal Loading", *Chemistry - A European Journal*, Vol. 18, pp. 12103–12111.
- Zhang, X, Y. Liu, S. T. Lee, S. Yang and Z. Kang, 2014, "Coupling Surface Plasmon Resonance of Gold Nanoparticles with Slow-Photon-Effect of TiO₂ Photonic Crystals for Synergistically Enhanced Photoelectrochemical Water Splitting", *Energy and Environmental Science*, Vol. 7, pp. 1409–1419.
- Zhang, Z., C. Gao, Z. Wu, W. Han, Y. Wang, W. Fu and E. Xie, 2016, "Toward Efficient Photoelectrochemical Water-Splitting by Using Screw-Like SnO₂ Nanostructures As Photoanode After Being Decorated with Cds Quantum Dots", *Nano Energy*, Vol. 19, pp. 318–327.
- Zhao, W., K. Maeda, F. Zhang, T. Hisatomi and K. Domen, 2014, "Effect of Post-Treatments on The Photocatalytic Activity of Sm₂Ti₂S₂O₅ for The Hydrogen Evolution Reaction", *Physical Chemistry Chemical Physics*, Vol. 16, p.12051.
- Zhao, X., Y. Wu, W. Yao and Y. Zhu, 2007, "Photoelectrochemical Properties of Thin Bi₂WO₆ Films", *Thin Solid Films*, Vol. 515, pp 4753–4757.
- Zhen, C., L. Wang, L. Liu, G. Lu, G. Q. Lu and H. M. Cheng, 2013, "Nonstoichiometric Rutile TiO₂ Photoelectrodes for Improved Photoelectrochemical Water Splitting", *Chemical Communications*, Vol. 49, pp. 6191–6193.
- Zhou, C., G. Chen, Y. Li, H. Zhang and J. Pei, 2009, "Photocatalytic Activities of Sr₂Ta₂O₇ Nanosheets Synthesized by A Hydrothermal Method", *International Journal of Hydrogen Energy*, Vol.34, pp. 2113–2120.

- Zhou, T., L. Chen, Y. Ye, L. Chen, Z. Qi, H. Freund and K. Sundmacher, 2012, "An Overview of Mutual Solubility of Ionic Liquids and Water Predicted by COSMO-RS", *industrial and Engineering Chemistry Research*, Vol. 51, pp. 6256–6264.
- Zhou, X. T., H. B. Ji and X. J. Huang, 2012, "Photocatalytic Degradation of Methyl Orange Over Metalloporphyrins Supported on TiO₂ Degussa P25", *Molecules*, Vol. 17, pp. 1149–1158.
- Zielińska, B., E. Borowiak-Palen and R. J. Kalenczuk, 2008, "Photocatalytic Hydrogen Generation Over Alkaline-Earth Titanates in The Presence of Electron Donors", *International Journal of Hydrogen Energy*, Vol. 33, pp. 1797–1802.
- Zielińska, B., E. Mijowska and R. J. Kalenczuk, 2012, "Synthesis, Characterization and Photocatalytic Properties of Lithium Tantalate", *Materials Characterization*, Vol. 68, pp. 71–76.
- Zwara, J., M. Paszkiewicz-Gawron, J. Łuczak, A. Pancielejko, W. Lisowski, G. Trykowski and E. Grabowska, 2019, "The Effect of Imidazolium Ionic Liquid on The Morphology of Pt Nanoparticles Deposited on The Surface of SrTiO₃ and Photoactivity of Pt–SrTiO₃ Composite in The H₂ Generation Reaction", *International Journal of Hydrogen Energy*, Vol. 44, pp. 26308–26321.

APPENDIX A: ADDITIONAL INFORMATION FOR ANALYSIS OF PWS DATABASE

Table A.1. Distribution of data points among the performance classes for significant variables (visible light dataset).

LHS	RHS	Support	Confidence	Lift
{Undoped_Pervoskite=NaTaO}	{ProdRate=HIGH}	0.079	0.630	1.889
{Undoped_Pervoskite=NaTaO}	{ProdRate=MEDIUM}	0.009	0.074	0.225
{Undoped_Pervoskite=NaTaO}	{ProdRate=LOW}	0.037	0.296	0.877
LHS	RHS	Support	Confidence	Lift
{Undoped_Pervoskite=SrTiO}	{ProdRate=HIGH}	0.04	0.16	0.47
{Undoped_Pervoskite=SrTiO}	{ProdRate=MEDIUM}	0.13	0.55	1.67
{Undoped_Pervoskite=SrTiO}	{ProdRate=LOW}	0.07	0.29	0.87
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=SSR}	{ProdRate=HIGH}	0.03	0.15	0.44
{Prep_Meth=SSR}	{ProdRate=MEDIUM}	0.07	0.33	1.01
{Prep_Meth=SSR}	{ProdRate=LOW}	0.12	0.52	1.54
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=HT}	{ProdRate=HIGH}	0.12	0.34	1.03
{Prep_Meth=HT}	{ProdRate=MEDIUM}	0.15	0.42	1.28
{Prep_Meth=HT}	{ProdRate=LOW}	0.08	0.24	0.70
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=PC}	{ProdRate=HIGH}	0.00	0.00	0.00
{Prep_Meth=PC}	{ProdRate=MEDIUM}	0.01	0.20	0.61
{Prep_Meth=PC}	{ProdRate=LOW}	0.03	0.70	2.07
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=SG}	{ProdRate=HIGH}	0.01	0.30	0.90
{Prep_Meth=SG}	{ProdRate=MEDIUM}	0.00	0.00	0.00
{Prep_Meth=SG}	{ProdRate=LOW}	0.03	0.60	1.78
LHS	RHS	Support	Confidence	Lift
{Promoter=Pt}	{ProdRate=HIGH}	0.09	0.21	0.63
{Promoter=Pt}	{ProdRate=MEDIUM}	0.18	0.41	1.25
{Promoter=Pt}	{ProdRate=LOW}	0.17	0.38	1.12

Table A.1. Distribution of data points among the performance classes for significant variables (visible light dataset). (cont.)

LHS	RHS	Support	Confidence	Lift
{Promoter=NAN}	{ProdRate=HIGH}	0.17	0.58	1.73
{Promoter=NAN}	{ProdRate=MEDIUM}	0.06	0.22	0.67
{Promoter=NAN}	{ProdRate=LOW}	0.06	0.20	0.60
LHS	RHS	Support	Confidence	Lift
{Crystal=cubic}	{ProdRate=HIGH}	0.12	0.30	0.91
{Crystal=cubic}	{ProdRate=MEDIUM}	0.12	0.30	0.92
{Crystal=cubic}	{ProdRate=LOW}	0.16	0.40	1.17
LHS	RHS	Support	Confidence	Lift
{Crystal=orthorombic}	{ProdRate=HIGH}	0.10	0.48	1.43
{Crystal=orthorombic}	{ProdRate=MEDIUM}	0.09	0.43	1.32
{Crystal=orthorombic}	{ProdRate=LOW}	0.02	0.09	0.26
LHS	RHS	Support	Confidence	Lift
{Alcohol%=NAN}	{ProdRate=HIGH}	0.18	0.42	1.26
{Alcohol%=NAN}	{ProdRate=MEDIUM}	0.11	0.25	0.75
{Alcohol%=NAN}	{ProdRate=LOW}	0.14	0.33	0.99
LHS	RHS	Support	Confidence	Lift
{Alcohol%=20}	{ProdRate=HIGH}	0.08	0.33	1.00
{Alcohol%=20}	{ProdRate=MEDIUM}	0.06	0.24	0.73
{Alcohol%=20}	{ProdRate=LOW}	0.11	0.43	1.26
LHS	RHS	Support	Confidence	Lift
{SacAgent1=Na ₂ S}	{ProdRate=HIGH}	0.13	0.76	2.27
{SacAgent1=Na ₂ S}	{ProdRate=MEDIUM}	0.04	0.22	0.66
{SacAgent1=Na ₂ S}	{ProdRate=LOW}	0.00	0.00	0.00

Table A.2. Distribution of data points among the performance classes for significant variables (UV dataset).

LHS	RHS	Support	Confidence	Lift
{Undoped_Pervoskite=NaTaO}	{ProdRate=HIGH}	0.10	0.48	1.45
{Undoped_Pervoskite=NaTaO}	{ProdRate=MEDIUM}	0.07	0.35	1.08
{Undoped_Pervoskite=NaTaO}	{ProdRate=LOW}	0.03	0.17	0.49

Table A.2. Distribution of data points among the performance classes for significant variables (UV dataset). (cont.)

LHS	RHS	Support	Confidence	Lift
{Undoped_Perovskite=SrTiO}	{ProdRate=HIGH}	0.03	0.38	1.13
{Undoped_Perovskite=SrTiO}	{ProdRate=MEDIUM}	0.02	0.33	1.03
{Undoped_Perovskite=SrTiO}	{ProdRate=LOW}	0.02	0.29	0.85
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=SSR}	{ProdRate=HIGH}	0.13	0.30	0.90
{Prep_Meth=SSR}	{ProdRate=MEDIUM}	0.11	0.26	0.79
{Prep_Meth=SSR}	{ProdRate=LOW}	0.19	0.44	1.29
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=HT}	{ProdRate=HIGH}	0.07	0.31	0.92
{Prep_Meth=HT}	{ProdRate=MEDIUM}	0.08	0.35	1.07
{Prep_Meth=HT}	{ProdRate=LOW}	0.08	0.35	1.01
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=PC}	{ProdRate=HIGH}	0.03	0.27	0.80
{Prep_Meth=PC}	{ProdRate=MEDIUM}	0.06	0.49	1.50
{Prep_Meth=PC}	{ProdRate=LOW}	0.03	0.24	0.71
LHS	RHS	Support	Confidence	Lift
{Prep_Meth=SG}	{ProdRate=HIGH}	0.05	0.57	1.71
{Prep_Meth=SG}	{ProdRate=MEDIUM}	0.03	0.32	0.99
{Prep_Meth=SG}	{ProdRate=LOW}	0.01	0.11	0.31
LHS	RHS	Support	Confidence	Lift
{Promoter=Pt}	{ProdRate=HIGH}	0.08	0.33	0.98
{Promoter=Pt}	{ProdRate=MEDIUM}	0.11	0.42	1.30
{Promoter=Pt}	{ProdRate=LOW}	0.06	0.25	0.74
LHS	RHS	Support	Confidence	Lift
{Promoter=NAN}	{ProdRate=HIGH}	0.15	0.35	1.06
{Promoter=NAN}	{ProdRate=MEDIUM}	0.13	0.29	0.90
{Promoter=NAN}	{ProdRate=LOW}	0.15	0.35	1.04

Table A.2. Distribution of data points among the performance classes for significant variables (UV dataset). (cont.)

LHS	RHS	Support	Confidence	Lift
{Crystal=cubic}	{ProdRate=HIGH}	0.07	0.26	0.77
{Crystal=cubic}	{ProdRate=MEDIUM}	0.09	0.32	1.00
{Crystal=cubic}	{ProdRate=LOW}	0.12	0.42	1.22
LHS	RHS	Support	Confidence	Lift
{Crystal=orthorombic}	{ProdRate=HIGH}	0.08	0.33	0.98
{Crystal=orthorombic}	{ProdRate=MEDIUM}	0.05	0.21	0.66
{Crystal=orthorombic}	{ProdRate=LOW}	0.11	0.46	1.35
LHS	RHS	Support	Confidence	Lift
{Alcohol%=NAN}	{ProdRate=HIGH}	0.12	0.22	0.67
{Alcohol%=NAN}	{ProdRate=MEDIUM}	0.19	0.37	1.14
{Alcohol%=NAN}	{ProdRate=LOW}	0.21	0.41	1.20
LHS	RHS	Support	Confidence	Lift
{Alcohol%=20}	{ProdRate=HIGH}	0.14	0.61	1.83
{Alcohol%=20}	{ProdRate=MEDIUM}	0.04	0.17	0.52
{Alcohol%=20}	{ProdRate=LOW}	0.05	0.22	0.64
LHS	RHS	Support	Confidence	Lift
{SacAgent1=Na ₂ S}	{ProdRate=HIGH}	0.00	0.00	0.00
{SacAgent1=Na ₂ S}	{ProdRate=MEDIUM}	0.02	0.22	0.69
{SacAgent1=Na ₂ S}	{ProdRate=LOW}	0.06	0.78	2.33

Table A.3. Results of association rule mining for visible light data for production rate class *high*.

LHS		RHS	Support	Confidence	Lift
HIGH	350 ($\mu\text{mol/gcat.h}$) and above	{Undoped_Pervoskite=ZnCdS}	0.07	0.21	3.00
		{Prep_Meth=novel}	0.13	0.39	2.33
		{SacAgent1=Na ₂ S}	0.13	0.39	2.27
		{SacAgent2=Na ₂ SO ₃ }	0.13	0.39	2.27
		{SurfaceArea=9}	0.07	0.22	2.09
		{Calctime=2}	0.05	0.15	1.94
		{Undoped_Pervoskite=NaTaO}	0.08	0.24	1.89
		{Promoter=NAN}	0.17	0.51	1.73
		{Calctime=NAN}	0.19	0.57	1.60
		{Crystal=hexagonal}	0.06	0.18	1.56
		{CalcT=NAN}	0.14	0.43	1.48
		{Crystal=orthorombic}	0.10	0.31	1.43
		{Prom_Meth=NAN}	0.19	0.58	1.38
		{CalcT=5}	0.07	0.22	1.33
		{Alcohol%=NAN}	0.18	0.54	1.26
		{SurfaceArea=4}	0.16	0.49	1.15
		{B_1stdoped=NAN}	0.21	0.63	1.05
		{B_2nddoped=NAN}	0.33	1.00	1.03
		{Prep_Meth=HT}	0.12	0.36	1.03
		{Alcohol%=2}	0.08	0.25	1.00
		{A_2nddoped=NAN}	0.33	1.00	1.00

Table A.4. Results of Association rule mining for UV light data for production rate class *high*.

LHS		RHS	Support	Confidence	Lift
HIGH	350 ($\mu\text{mol/gcat.h}$) and above	{SacAgent1=NaNO ₂ }	0.06	0.19	2.33
		{Alcohol%=2}	0.14	0.43	1.83
		{SurfaceArea=6}	0.05	0.16	1.46
		{Undoped_Perovskite=NaTaO}	0.10	0.29	1.45
		{CalcT=8}	0.08	0.25	1.42
		{Calctime=NAN}	0.07	0.22	1.36
		{CalcT=NAN}	0.07	0.22	1.36
		{Crystal=octahedral}	0.07	0.20	1.22
		{Calctime=7}	0.06	0.17	1.21
		{Calctime=2}	0.07	0.20	1.20
		{Prom_Meth=PD}	0.09	0.28	1.16
		{Promoter=NAN}	0.15	0.46	1.06
		{Prom_Meth=NAN}	0.15	0.46	1.06
		{A_2nddoped=NAN}	0.32	0.96	1.04
		{SacAgent2=NAN}	0.33	1.00	1.00
		{Promoter=NiO}	0.06	0.18	1.00
		{B_2nddoped=NAN}	0.33	1.00	1.00

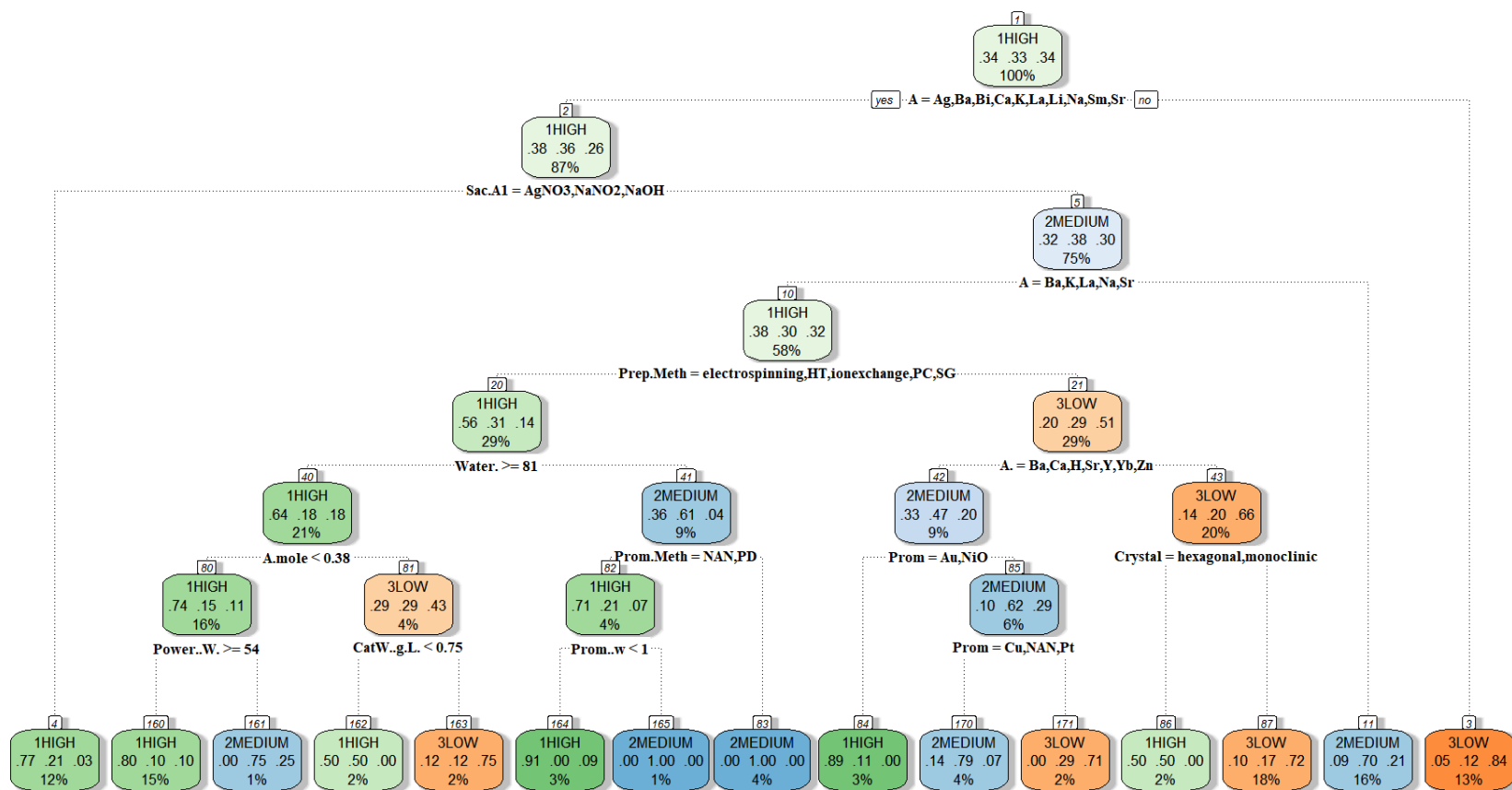


Figure A.1. Optimal decision tree for UV light data.

Table A.5. Confusion matrix for visible light data*.

Actual Data		Predictions			Prediction
Class	# of data	LOW	MEDIUM	HIGH	Accuracy %
LOW	73	64	8	1	87.67
MEDIUM	71	6	58	7	81.69
HIGH	72	2	15	55	76.39

*First column on the right shows the actual number of data in each classes. The columns labelled as “prediction” indicate the classes in which the data placed. For example, there are 73 data points in high classes and tree accurately placed 64 of them while it placed 8 into medium and 1 into low classes. Consequently the prediction accuracy for high classes is $(64/73)*100 = 87.67$

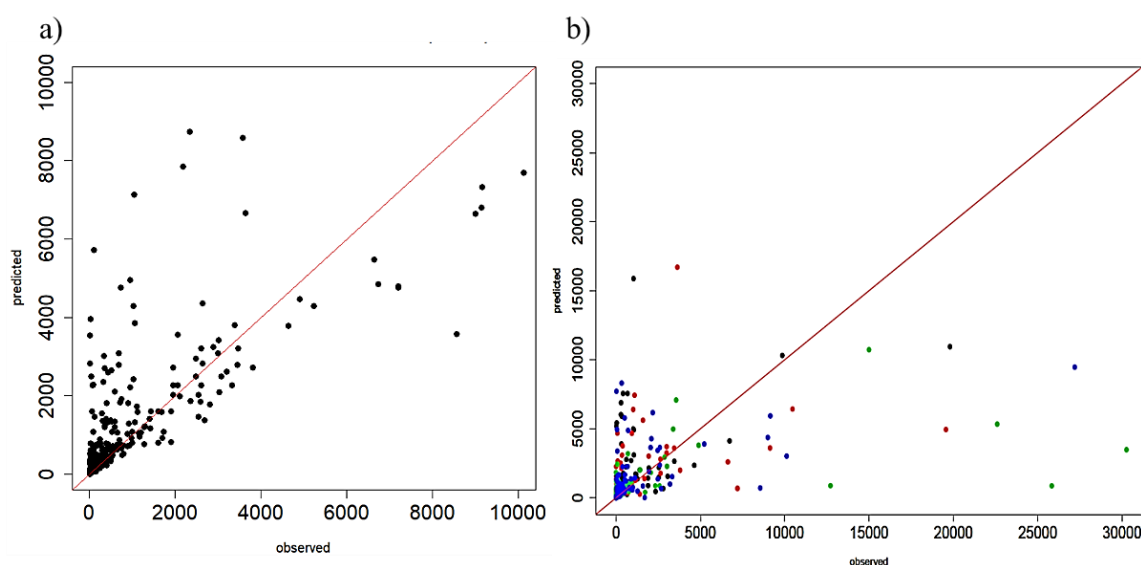


Figure A.2. Predictions by random forest model for UV a) testing set and b) training set.

APPENDIX B: ADDITIONAL INFORMATION FOR ANALYSIS OF PECWS DATABASE

Table B.1. Discretization of the numeric input in ten levels.

CalcT.I		CalcTime.I		CalcT.II	
intervals	given name	intervals	given name	intervals	given name
0-100	0	0-1	0	0-100	0
100-400	1	1-2	1	100-400	1
400-500	2	2-3	2	400-500	2
500-600	3	3-4	3	500-600	3
600-700	4	4-5	4	600-700	4
700-800	5	5-6	5	700-800	5
800-900	6	6-7	6	800-900	6
900-1000	7	7-8	7	900-1000	7
1000+	8	8+	8	1000+	8
CalcTime.II		LastCalcT		LastCalcTime	
intervals	given name	intervals	given name	intervals	given name
0-1	0	0-100	0	0-1	0
1-2	1	100-400	1	1-2	1
2-3	2	400-500	2	2-3	2
3-4	3	500-600	3	3-4	3
4-5	4	600-700	4	4-5	4
5-6	5	700-800	5	5-6	5
6-7	6	800-900	6	6-7	6
7-8	7	900-1000	7	7-8	7
8+	8	1000+	8	8+	8
Power.W		Wavelength.nm		Intensity.mW.cm2	
intervals	given name	intervals	given name	intervals	given name
00-100	0	0-300	0	0-10	0
100-150	1	300-400	1	10-50	1
150-200	2	400-410	2	50-75	2
200-250	3	410-420	3	75-100	3
250-300	4	420-450	4	100	4
300-350	5	450-500	5	100+	5
350-400	6	500+	6		
400-450	7				
450-500	8				
500+	9				

Table B.1. Discretization of the numeric input in ten levels. (cont.)

E.Molarity.M	A.Molarity.M	pH
--------------	--------------	----

intervals	given name	intervals	given name	intervals	given name
0-0.1	0	0-0.1	0	0-1	1
0.1-0.25	1	0.1-0.25	1	1-2	2
0.25-0.5	2	0.25-0.5	2	2-3	3
0.5-1	3	0.5-1	3	3-4	4
1-1.25	4	1-1.25	4	4-5	5
1.25-1.5	5	1.25-1.5	5	5-6	6
1.5-2	6	1.5-2	6	6-7	7
2-3	7	2-3	7	7-8	8
3+	8	3+	8	8-9	9
				9-10	10
				10-11	11
				11-12	12
				12-13	13
				13-14	14

Evaluation criteria for Variable Importance

IncNodePurity: This term is the abbreviation for the phrase of “increase in node purity”. In the process of random forest application, several trees were developed by using subset of the input variables and subset of the instances at each time. Therefore, the algorithm can calculate the change in purity of the nodes by evaluating the models formed in the absence and presence of a definite input variable. That change can be considered as the contribution of that input variable to the accuracy of random forest model.

%IncMSE: This term is the abbreviation for the phrase of “increase in mean square error”. Similar to previous one, for each tree, the prediction error on the out-of-bag subset through the dataset is recorded during RF application. Then the error value represents the importance of that input variable in the prediction of output variable.

The difference between regular accuracy and balanced accuracy

The regular accuracy and balanced accuracy are the same thing if the number of instances in each class are well balanced. However, in a classification task if there is a class imbalance problem and it cannot be solved, then balanced accuracy should be reported. Here are the equations for regular and balanced accuracy.

Table B.2. The confusion matrix for sample subset.

		Predicted values	
		A	B
Real values	A	3	5
	B	4	8

$$\text{Regular accuracy} = \frac{3 + 8}{3 + 5 + 4 + 8} = 0.55$$

$$\text{Balanced accuracy} = (\frac{3}{3 + 5} + \frac{8}{8 + 4})/2 = 0.38$$

Table B.3. The confusion matrix for bandgap testing subset (PECWS).

Optimum parameters: nodesize=5 ntree=20

			Predicted Class			Recall
			1low	2medium	3high	
Overall Accuracy 0.72	Real Class	1low	21	9	5	0.60
		2medium	3	31	10	0.70
		3high	1	0	20	0.95
	Precision		0.84	0.78	0.57	

Table B.4. The confusion matrix for photocurrent density testing and validation subsets (PECWS).

Optimum parameters: nodesize=6 ntree=166

<i>Confusion Matrix</i>	Overall Accuracy	Real Class	Total number	Predicted Class		
				A	B	C
Validation	0.89	A	2671	2489	176	6
		B	2623	303	2222	98
		C	2622	76	196	2350
			Precision	0.87	0.86	0.96
Test	0.60	A	865	734	87	44
		B	909	357	432	120
		C	869	121	262	486
			Precision	0.60	0.55	0.75

APPENDIX C: ADDITIONAL INFORMATION FOR ANALYSIS OF WATER SOLUBILITY IN ILs DATABASE

Explanation for the descriptors included in the datasets

In this study, we selected 10 descriptors for each cation and anion. These descriptors are explained briefly as follows;

Molecular weight (MW): the summation of atomic masses forming anion/cation in grams.

HOMO and LUMO energies (E_{HOMO} , E_{LUMO}): the energy of highest occupied and lowest unoccupied molecular orbital for each specie in eV.

Dipole (μ): the charge separation within the specie in Debye.

CPK-Area (A): the area as determined by space filling model where, CPK are second name initials of chemists who developed this model in Å².

CPK Ovality (O): the measure of sphericity of the specie as determined by space filling model.

Polarizability (P): the ability of cation/anion for inducing dipole in m³.

Hydrogen bond donor count (HBD) and hydrogen bond acceptor count (HBA): the number of hydrogen bond donor and acceptor atoms, respectively.

ZPE: the vibrational energy of nuclei at 0 K in kJ/mol.

Table C.1. Full names of cations.

Cation Full Name	Abbreviation
butyl-diethanolammonium_cation	CAmm1
hexyltrimethylammonium_cation	CAmm10
methyl-trioctyl-ammonium_cation	CAmm11
n-hexyl-n,n,n-triethylammonium_cation	CAmm12
octyltrimethylammonium_cation	CAmm13
tetra-ethylammonium_cation	CAmm14
tetra-methylammonium_cation	CAmm15

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
tetra-n-butylammonium_cation	CAmm16
tetradecyltrimethylammonium_cation	CAmm17
tetrapropylammonium_cation	CAmm18
tributylmethylammonium_cation	CAmm19
butyltrimethylammonium_cation	CAmm2
triethylheptylammonium_cation	CAmm20
triethyloctylammonium_cation	CAmm21
triethylpentylammonium_cation	CAmm22
trimethylethylammonium_cation	CAmm23
dimethyl-tetradecyl-benzylammonium_cation	CAmm24
cyanomethyl-dimethyl-ethylammonium_cation	CAmm25
dodecyl-dimethyl-3-sulfopropylammonium_cation	CAmm26
dodecyl-dimethyl-benzylammonium_cation	CAmm27
ethyl-(3-hydroxypropyl)-dimethylammonium_cation	CAmm28
ethyl-(3-methoxypropyl)-dimethylammonium_cation	CAmm29
decyl-dimethyl-benzylammonium_cation	CAmm3
ethyl-dimethyl-2-methoxyethylammonium_cation	CAmm30
(2-ethoxyethyl)-ethyl-dimethylammonium_cation	CAmm31
2-hydroxyethyl-ammonium_cation	CAmm32
bis(2-methoxyethyl)ammonium_cation	CAmm33
diethanolammonium_cation	CAmm4
diethyl-methyl-propylammonium_cation	CAmm5
dimethylethanolammonium_cation	CAmm6
ethyl-dimethyl-butylammonium_cation	CAmm7
ethyl-dimethyl-propylammonium_cation	CAmm8
heptyltrimethylammonium_cation	CAmm9
Choline_cation	Cch1
1,3-didecyl-2-methyl-imidazolium_cation	CIm41
1,3-diethylimidazolium_cation	CIm14
1,3-dihexyl-imidazolium_cation	CIm17

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
1,3-dipentyl-imidazolium_cation	CIm16
1,3-methyl-imidazolium_cation_	CIm13
1-butyl-2-3-methyl-imidazolium_cation	CIm34
1-butyl-3-ethylimidazolium_cation	CIm38
1-butyl-3-methyl-imidazolium_cation	CIm20
1-butyl-imidazolium_cation	CIm4
1-decyl-3-ethyl-imidazolium_cation	CIm40
1-decyl-3-methyl-imidazolium_cation	CIm26
1-dodecyl-3-methyl-imidazolium_cation	CIm27
1-ethyl-2-3-methyl-imidazolium_cation	CIm31
1-ethyl-3,4-dimethylimidazolium_cation	CIm32
1-ethyl-3-hexyl-imidazolium_cation	CIm39
1-ethyl-3-methyl-imidazolium_cation	CIm18
1-ethyl-3-propylimidazolium_cation	CIm37
1-heptyl-3-methyl-imidazolium_cation	CIm23
1-hexadecyl-2-3-methyl-imidazolium_cation	CIm36
1-hexadecyl-3-methyl-imidazolium_cation	CIm29
1-hexyl-2-3-methyl-imidazolium_cation	CIm35
1-hexyl-3-methyl-imidazolium_cation	CIm22
1-methyl-3-nonylimidazolium_cation	CIm25
1-octadecyl-3-methyl-imidazolium_cation	CIm30
1-octyl-3-methyl-imidazolium_cation	CIm24
1-pentyl-3-methyl-imidazolium_cation	CIm21
1-propyl-2-3-methyl-imidazolium_cation	CIm33
1-propyl-3-methyl-imidazolium_cation	CIm19
1-tetradecyl-3-methyl-imidazolium_cation	CIm28
3-decyl-imidazolium_cation	CIm10
3-dodecyl-imidazolium_cation	CIm12
3-ethyl-imidazolium_cation	CIm2
3-heptyl-imidazolium_cation	CIm7

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
3-hexyl-imidazolium_cation	CIm6
3-methyl-imidazolium_cation	CIm1
3-nonyl-imidazolium_cation	CIm9
3-octyl-imidazolium_cation	CIm8
3-pentyl-imidazolium_cation	CIm5
3-propyl-imidazolium_cation	CIm3
3-undecyl-imidazolium_cation	CIm11
dibutylimidazolium_cation	CIm15
1,3-bis[(heptyloxy)methyl]-imidazolium_cation	CIm42
1-benzyl-3-methyl-imidazolium_cation	CIm43
1-butoxymethyl-3-propoxymethylimidazolium_cation	CIm44
1-isobutenyl-3-methylimidazolium_cation	CIm45
1-methyl-3-(3-oxobutyl)-imidazolium_cation	CIm46
1-methyl-3-heptoxymethylimidazolium_cation	CIm47
1-methyl-3-nonoxymethylimidazolium_cation	CIm48
1-propenyl-boronic-acid-3-decyl-imidazolium_cation	CIm49
1-propenyl-boronic-acid-3-methyl-imidazolium_cation	CIm50
1-propenyl-boronic-acid-3-octyl-imidazolium_cation	CIm51
2-methyl-1-(phenylmethyl)-3-propylimidazolium_cation	CIm52
3-(3-carboxypropyl)-1-methyl-imidazolium_cation	CIm53
3-(3-methoxypropyl)-1-methyl-imidazolium_cation	CIm54
3-(4-hydroxybutyl)-1-methyl-imidazolium_cation	CIm55
3-(7-carboxyheptyl)-1-methyl-imidazolium_cation	CIm56
3-(butoxymethyl)-imidazolium_cation	CIm57
3-(cyanomethyl)-1-methyl-imidazolium_cation	CIm62
3-[(dodecyloxy)methyl]-imidazolium_cation	CIm61
3-[(heptyloxy)methyl]-imidazolium_cation	CIm59
3-[(hexyloxy)methyl]-imidazolium_cation	CIm58
3-[(octyloxy)methyl]-imidazolium_cation	CIm60
benzyl-triphenyl-phosphonium_cation	CPhos1

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
butyl-trihexyl-phosphonium_cation	CPhos2
decyl-trihexyl-phosphonium_cation	CPhos3
dodecyl-trihexyl-phosphonium_cation	CPhos19
ethyl-trihexyl-phosphonium_cation	CPhos16
heptyl-trihexyl-phosphonium_cation	CPhos17
hexadecyl-trihexyl-phosphonium_cation	CPhos18
methyltris(2-methylpropyl)phosphonium_cation	CPhos15
tetrabutyl-phosphonium_cation	CPhos14
tetrahexyl-phosphonium_cation	CPhos13
tributyl-ethyl-phosphonium_cation	CPhos10
tributylmethylphosphonium_cation	CPhos12
tributyl-tetradecylphosphonium_cation	CPhos11
trihexyl(methoxymethyl)phosphonium_cation	CPhos4
trihexyl-2-propen-1-ylphosphonium_cation	CPhos5
trihexyl-octyl-phosphonium_cation	CPhos6
trihexyl-propyl-phosphonium_cation	CPhos7
trihexyl-tetradecyl-phosphonium_cation	CPhos8
triisobutyl-methyl-phosphonium_cation	CPhos9
1-(2-methoxyethyl)-1-methylpiperidinium_cation	Cpip1
1-(3-hydroxypropyl)-1-methylpiperidinium_cation	Cpip2
1-(3-methoxypropyl)-1-methylpiperidinium_cation	Cpip3
1-(cyanomethyl)-1-methylpiperidinium_cation	Cpip4
1-(ethoxymethyl)-1-methylpiperidinium_cation	Cpip5
1-butyl-1-methylpiperidinium_cation_	Cpip6
1-methyl-1-propylpiperidinium_cation_	Cpip7
1-(2-ethoxyethyl)pyridinium_cation	CPyr11
1-(2-hydroxyethyl)pyridinium_cation	CPyr10
1-(cyanomethyl)pyridinium_cation	CPyr12
1-(ethoxymethyl)pyridinium_cation	CPyr13
1-[(heptyloxy)methyl]-3-hydroxypyridinium_cation	CPyr21

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
1-butyl-pyridinium_cation	CPyr14
1-decylpyridinium_cation	CPyr15
1-ethyl-3-methylpyridinium_cation_	CPyr20
1-ethyl-pyridinium_cation	CPyr16
1-hexyl-3-methyl-pyridinium_cation	CPyr17
1-hexyl-4-methyl-pyridinium_cation	CPyr18
1-hexyl-pyridinium_cation	CPyr19
1-methyl-pyridinium_cation_	CPyr9
1-octyl-pyridinium_cation	CPyr1
1-pentylpyridinium_cation	CPyr2
1-propylpyridinium_cation	CPyr3
1-tert-butyl-pyridinium_cation_	CPyr4
4-methyl-n-butylpyridinium_cation	CPyr5
n-(3-hydroxypropyl)pyridinium_cation	CPyr6
n-(3-sulfopropyl)pyridinium_cation	CPyr7
pyridinium_cation_	CPyr8
1-(2-hydroxyethyl)-1-methylpyrrolidinium_cation	CPyrr12
1-(2-methoxyethyl)-1-methylpyrrolidinium_cation	CPyrr11
1-(3-hydroxypropyl)-1-methylpyrrolidinium_cation	CPyrr13
1-(3-methoxypropyl)-1-methylpyrrolidinium_cation	CPyrr14
1-(cyanomethyl)-1-methylpyrrolidinium_cation	CPyrr15
1,1-dihexylpyrrolidinium_cation	CPyrr3
1,1-dimethyl-pyrrolidinium_cation_	CPyrr1
1,1-dipropyl-pyrrolidinium_cation	CPyrr2
1-butyl-1-ethyl-pyrrolidinium_cation	CPyrr6
1-butyl-1-methyl-pyrrolidinium_cation	CPyrr5
1-ethyl-1-methyl-pyrrolidinium_cation	CPyrr4
1-hexyl-1-methyl-pyrrolidinium_cation	CPyrr9
1-methyl-1-pentylpyrrolidinium_cation	CPyrr8
1-methyl-1-propylpyrrolidinium_cation	CPyrr7

Table C.1. Full names of cations. (cont.)

Cation Full Name	Abbreviation
1-octyl-1-methyl-pyrrolidinium_cation	CPyrr10
diethyl-methylsulfonium_cation	Csulf1
dimethyl-phenylsulfonium_cation	Csulf4
ethyl-dimethylsulfonium_cation	Csulf2
ethyl-phenyl-methylsulfonium_cation	Csulf5
triethylsulfonium_cation	Csulf3

Table C.2. Full names of anions.

Anion Full Name	Abbreviation
(2r,3r)-2,3-dihydroxybutanedioate_anion	A1
1-butanesulfonate_anion	A2
1-hexanesulfonate_anion	A3
2-(2-methoxyethoxy)ethylsulfate_anion	A4
2-hydroxy-1,2,3-propanetricarboxylate_anion	A5
2-hydroxyacetate_anion	A6
2-hydroxybutanedioate_anion	A7
acesulfamate_anion	A8
acetate_anion	A9
alaninate_anion	A10
argininate_anion	A11
hexafluoroarsenate_anion	A12
asparaginate_anion	A13
aspartate_anion	A14
tetrachloroborate_anion	A15
chlorotrifluoroborate_anion	A16
benzoate_anion	A17
tetrafluoroborate_anion	A18
bis(2,4,4-trimethylpentyl)phosphinate_anion	A19
bis(2-methylactato)borate_anion	A20

Table C.2. Full names of anions. (cont.)

Anion Full Name	Abbreviation
bis(fluorosulfonyl)imide_anion	A21
bis(pentafluoroethylsulfonyl)amide_anion	A22
bis(trifluoromethyl)imide_anion	A23
bis(trifluoromethylsulfonyl)methane_anion	A24
bis[1,2-benzenediolato]borate_anion	A25
bisbiphenyldiolatoborate_anion	A26
bismalonatoborate_anion	A27
bisoxalatoborate_anion	A28
bis-pentafluoroethyl-phosphinate_anion	A29
bissalicylatoborate_anion	A30
bromide_anion	A31
butanedioate_anion	A32
butanoate_anion	A33
butylsulfate_anion	A34
chloride_anion	A35
chlorate_anion	A36
decanoate_anion	A37
dibutylphosphate_anion	A38
dicyanamide_anion	A39
diethylphosphate_anion	A40
dihydrogen-phosphate_anion	A41
dimethylphosphate_anion	A42
dodecylsulfate_anion	A43
ethoxyethylsulfate_anion	A44
ethylsulfate_anion	A45
fluoride_anion	A46
formate_anion	A47
glutamate_anion	A48
glutamate_anion	A49
glycinate_anion	A50

Table C.2. Full names of anions. (cont.)

Anion Full Name	Abbreviation
heptafluorobutanoate_anion	A51
histidinate_anion	A52
hydrogencarbonate_anion	A53
hydrogensulfate_anion	A54
iodide_anion	A55
triiodide_anion	A56
imidodiphosphorylfluoride_anion	A57
isoleucinate_anion	A58
lactate_anion	A59
leucinate_anion	A60
lysinate_anion	A61
methanesulfonate_anion	A62
methioninate_anion	A63
methoxyethylsulfate_anion	A64
methylcarbonate_anion	A65
methylsulfate_anion	A66
nitrate_anion	A67
octylsulfate_anion	A68
pentafluoroethyltrifluoroborate_anion	A69
perfluorobutanesulfonate_anion	A70
perfluorooctanesulfonate_anion	A71
hexafluorophosphate_anion	A72
phosphonate_anion	A73
prolinate_anion	A74
propanedioate_anion	A75
propanoate_anion	A76
saccharinate_anion	A77
salicylate_anion	A78
hexafluoroantimonate_anion	A79
serinate_anion	A80

Table C.2. Full names of anions. (cont.)

Anion Full Name	Abbreviation
sulfate_anion	A81
sulfamate_anion	A82
tetrachloroaluminate_anion	A83
tetrachlorogallate_anion	A84
tetrachloroindium_anion	A85
tetracyanoborate_anion	A86
tf2n_anion	A87
thiocyanate_anion	A88
threoninate_anion	A89
toluene-4-sulfonate_anion	A90
tricyanomethane_anion	A91
trifluoroacetate_anion	A92
trifluoromethane-sulfonate_anion	A93
trifluoromethyltrifluoroborate_anion	A94
trifluorotris(perfluoropropyl)phosphate_anion	A95
tris(nonafluorobutyl)trifluorophosphate_anion	A96
tris(pentafluoroethyl)trifluorophosphate_anion	A97
tris(trifluoromethylsulfonyl)methide_anion	A98
valinate_anion	A99

Table C.3. Discretization of the numeric input in ten levels.

Cation molecular weight (amu)			Cation E.HOMO (eV)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<100	990	1	<-16.5	198
2	100≤<150	4950	2	-16.5≤<-16	99
3	150≤<200	4554	3	-16≤<-15.5	1485
4	200≤<250	2574	4	-15.5≤<-15	1584
5	250≤<300	792	5	-15≤<-14.5	2772
6	300≤<350	1188	6	-14.5≤<-14	1782
7	350≤<400	693	7	-14≤<-13.5	3069
8	400≤<450	99	8	-13.5≤<-13	1683
9	450≤<500	198	9	-13≤<-12.5	1881
10	500≤	99	10	-12.5≤	1584
Cation E.LUMO (eV)			Cation dipole (debye)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<-6	198	1	<5	7425
2	-6≤<-5.75	1584	2	5≤<10	3168
3	-5.75≤<-5.5	594	3	10≤<15	1980
4	-5.5≤<-5.25	297	4	15≤<20	990
5	-5.25≤<-5	1782	5	20≤<25	1188
6	-5≤<-4.75	4356	6	25≤<30	495
7	-4.75≤<-4.5	990	7	30≤<35	396
8	-4.5≤<-4.25	3267	8	35≤<40	198
9	-4.25≤<-4	2970	9	40≤<45	198
10	-4≤	99	10	≤45	99
Cation CPK area (Å²)			Cation CPK ovality		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<150	990	1	<1.2	495
2	150≤<200	4059	2	1.2≤<1.25	1386
3	200≤<250	3861	3	1.25≤<1.3	1980
4	250≤<300	2376	4	1.3≤<1.35	2475
5	300≤<350	1782	5	1.35≤<1.4	2475
6	350≤<400	693	6	1.4≤<1.45	1089
7	400≤<450	693	7	1.45≤<1.5	1287
8	450≤<500	693	8	1.5≤<1.55	1386
9	500≤<550	396	9	1.55≤<1.6	792
10	≤550	594	10	≤1.6	2772
Cation polarizability			Cation ZPE (kJ/mol)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<45	99	1	<45	0
2	45≤<50	1386	2	45≤<50	2871
3	50≤<55	6138	3	50≤<55	5346
4	55≤<60	2871	4	55≤<60	3168
5	60≤<65	2574	5	60≤<65	2079
6	65≤<70	792	6	65≤<70	891
7	70≤<75	1188	7	70≤<75	891
8	75≤<80	495	8	75≤<80	297
9	80≤<85	297	9	80≤<85	297
10	≤85	297	10	≤85	297

Table C.3. Discretization of the numeric input in ten levels. (cont.)

Anion molecular weight (amu)			Anion E.HOMO (eV)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<50	489	1	<-9	815
2	50≤<100	3423	2	-9≤<-8	326
3	100≤<150	4727	3	-8≤<-7	1467
4	150≤<200	3749	4	-7≤<-6	3097
5	200≤<250	1141	5	-6≤<-5	3749
6	250≤<300	1141	6	-5≤<-4.5	1141
7	300≤<350	163	7	-4.5≤<-4	3912
8	350≤<400	489	8	-4≤<-3	1141
9	400≤<450	326	9	-3≤<-2	163
10	≤450	489	10	≤-2	326
Anion E.LUMO (eV)			Anion dipole (debye)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<-1	1467	1	<1	4238
2	1≤<2	815	2	1≤<3	1793
3	2≤<3	1630	3	3≤<6	1956
4	3≤<4	2608	4	6≤<9	2445
5	4≤<5	2282	5	9≤<12	2119
6	5≤<6	1956	6	12≤<15	1630
7	6≤<7	2934	7	15≤<18	815
8	7≤<8	1793	8	18≤<21	326
9	8≤<9	489	9	21≤<24	326
10	≤9	163	10	≤24	489
Anion CPK area (Å ²)			Anion CPK ovality		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<60	652	1	<1.06	815
2	60≤<80	815	2	1.06≤<1.12	652
3	80≤<100	2119	3	1.12≤<1.18	1956
4	100≤<120	2119	4	1.18≤<1.24	3423
5	120≤<140	2119	5	1.24≤<1.3	2934
6	140≤<160	1793	6	1.3≤<1.36	2608
7	160≤<180	2282	7	1.36≤<1.42	1304
8	180≤<220	1793	8	1.42≤<1.48	978
9	220≤<260	815	9	1.48≤<1.54	652
10	≤260	1630	10	≤1.54	815
Anion polarizability			Anion ZPE (kJ/mol)		
Category #	Intervals	Number of Instances	Category #	Intervals	Number of Instances
1	<43	1304	1	<50	2934
2	43≤<46	3749	2	50≤<100	2282
3	46≤<49	3749	3	100≤<150	1467
4	49≤<52	3423	4	150≤<200	978
5	52≤<55	1793	5	200≤<300	2445
6	55≤<58	652	6	300≤<400	3260
7	58≤<61	489	7	400≤<500	1304
8	61≤<64	326	8	500≤<600	489
9	64≤<67	326	9	600≤<700	326
10	≤67	326	10	≤700	652

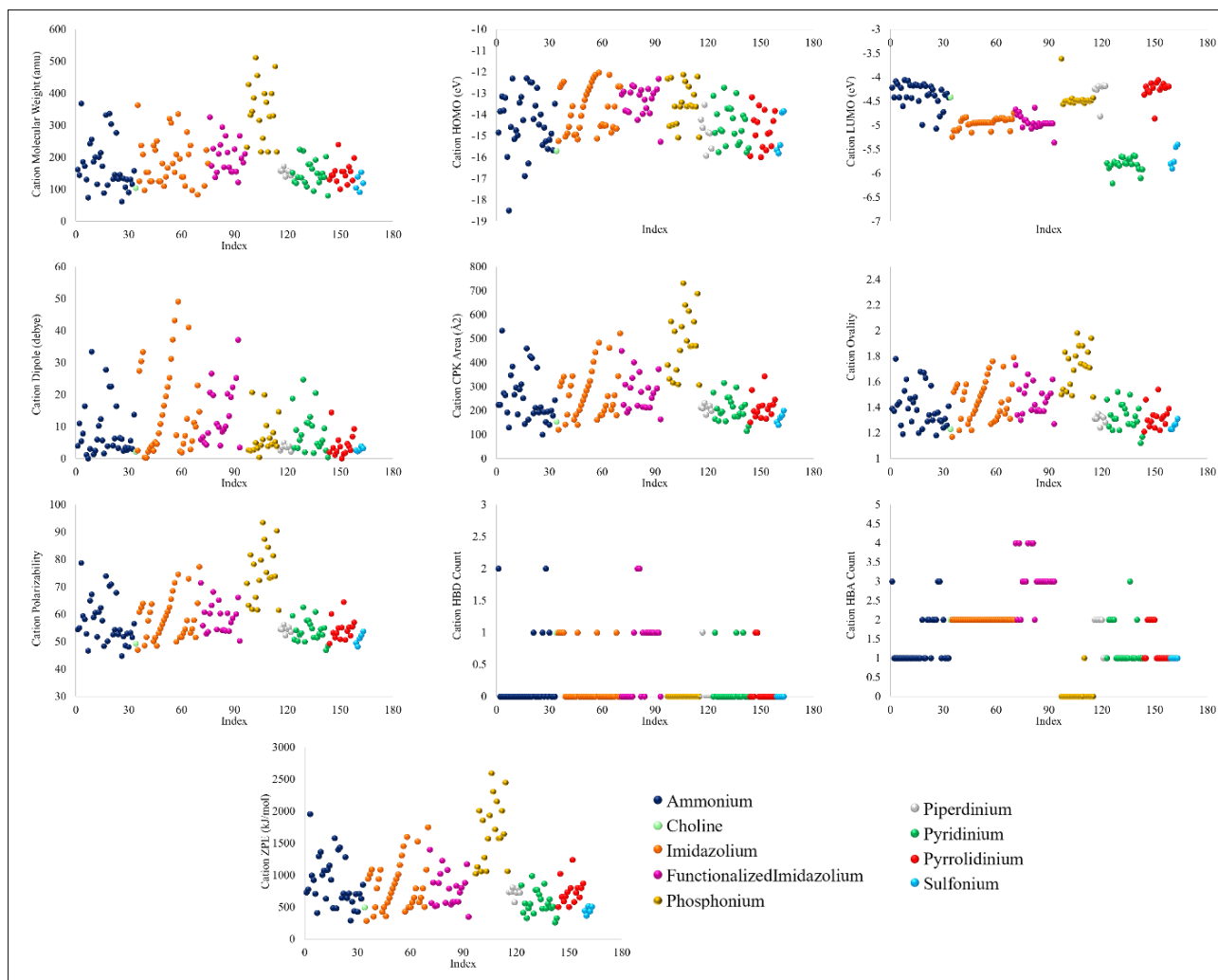


Figure C.1. The distribution of input variables for cations in the entire dataset.

Table C.4. Confusion matrix for imidazolium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	324	32	4	0.90
	B	22	279	16	0.88
	C	0	5	333	0.98
	recall	0.94	0.88	0.94	

$$accuracy = \frac{324 + 279 + 333}{324 + 32 + 4 + 22 + 279 + 16 + 0 + 5 + 333} = 0.922$$

$$accuracy = 0.922$$

$$balanced\ accuracy = \left(\frac{324}{324 + 22 + 0} + \frac{279}{32 + 279 + 5} + \frac{333}{4 + 16 + 333} \right) / 3$$

$$balanced\ accuracy = 0.920$$

Table C.5. Confusion matrix for ammonium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	237	41	5	0.84
	B	19	248	24	0.85
	C	0	10	233	0.96
	recall	0.92	0.83	0.89	

Table C.6. Confusion matrix for functionalized imidazolium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	90	12	0	0.88
	B	41	203	10	0.80
	C	0	1	163	0.99
	recall	0.69	0.94	0.94	

Table C.7. Confusion matrix for phosphonium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	176	2	0	0.99
	B	2	104	3	0.95
	C	0	1	182	0.99
	recall	0.99	0.97	0.98	

Table C.8. Confusion matrix for piperidinium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	51	7	0	0.88
	B	8	46	6	0.77
	C	0	1	54	0.98
	recall	0.86	0.85	0.90	

Table C.9. Confusion matrix for pyridinium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	159	31	1	0.83
	B	0	137	4	0.97
	C	0	9	179	0.95
	recall	1.00	0.77	0.97	

Table C.10. Confusion matrix for pyrrolidinium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	141	16	0	0.90
	B	0	90	2	0.98
	C	0	0	122	1.00
	recall	1.00	0.85	0.98	

Table C.11. Confusion matrix for sulfonium dataset.

		ACTUAL			
		A	B	C	precision
PRED	A	41	8	0	0.84
	B	4	28	2	0.82
	C	0	1	40	0.98
recall		0.91	0.76	0.95	

The list of 25 anion out of 99 leading to node 4 in the decision tree (Figure 4.27) for imidazolium dataset

A12-hexafluoroarsenate	A79- hexafluoroantimonate
A18- tetrafluoroborate	A83-tetrachloroaluminate
A21- bis(fluorosulfonyl)imide	A84- tetrachlorogallate
A22-bis(pentafluoroethylsulfonyl)amide	A85- tetrachloroindium
A24- bis(trifluoromethylsulfonyl)methane	A86- tetracyanoborate
A27-bismalonatoborate	A87- tf2n
A28- bisoxalatoborate	A93- trifluoromethane-sulfonate
A29- bis-pentafluoroethyl-phosphinate	A94- trifluoromethyltrifluoroborate
A36-chlorate	A95-trifluorotris(perfluoropropyl)phosphate
A69- pentafluoroethyltrifluoroborate	A96- tris(nonafluorobutyl)trifluorophosphate
A70- perfluorobutanesulfonate	A97-tris(pentafluoroethyl)trifluorophosphate
A71-perfluorooctanesulfonate	A98- tris(trifluoromethylsulfonyl)methide)
A72- hexafluorophosphate	

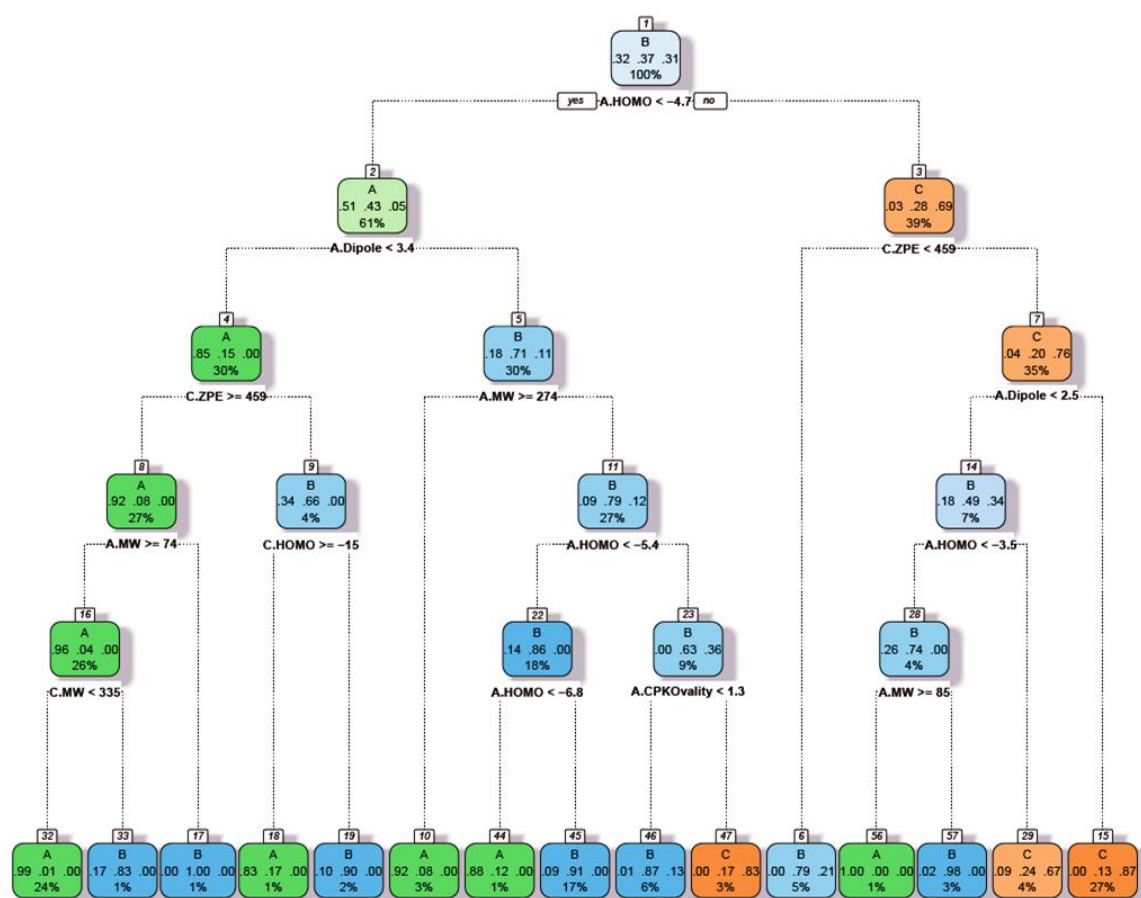


Figure C.2. Optimum decision tree structure for ammonium dataset.

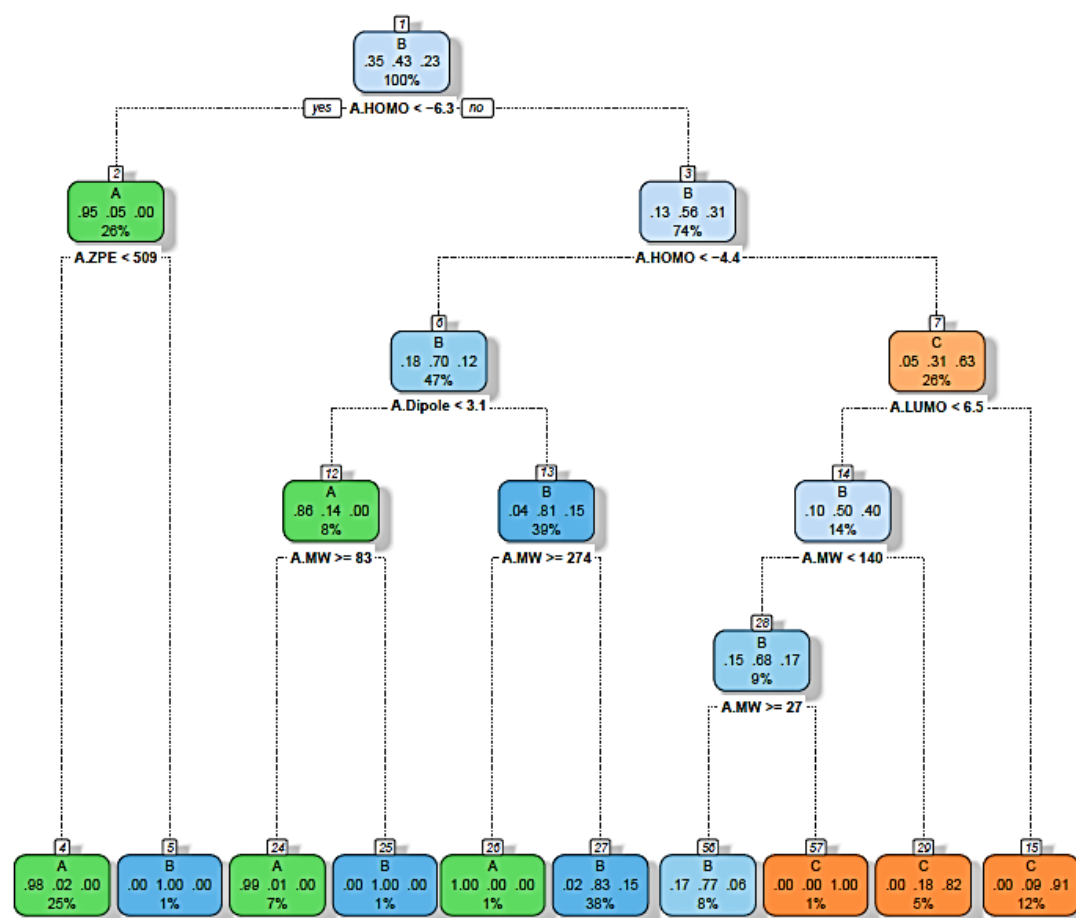


Figure C.3. Optimum decision tree structure for functionalized imidazolium dataset.

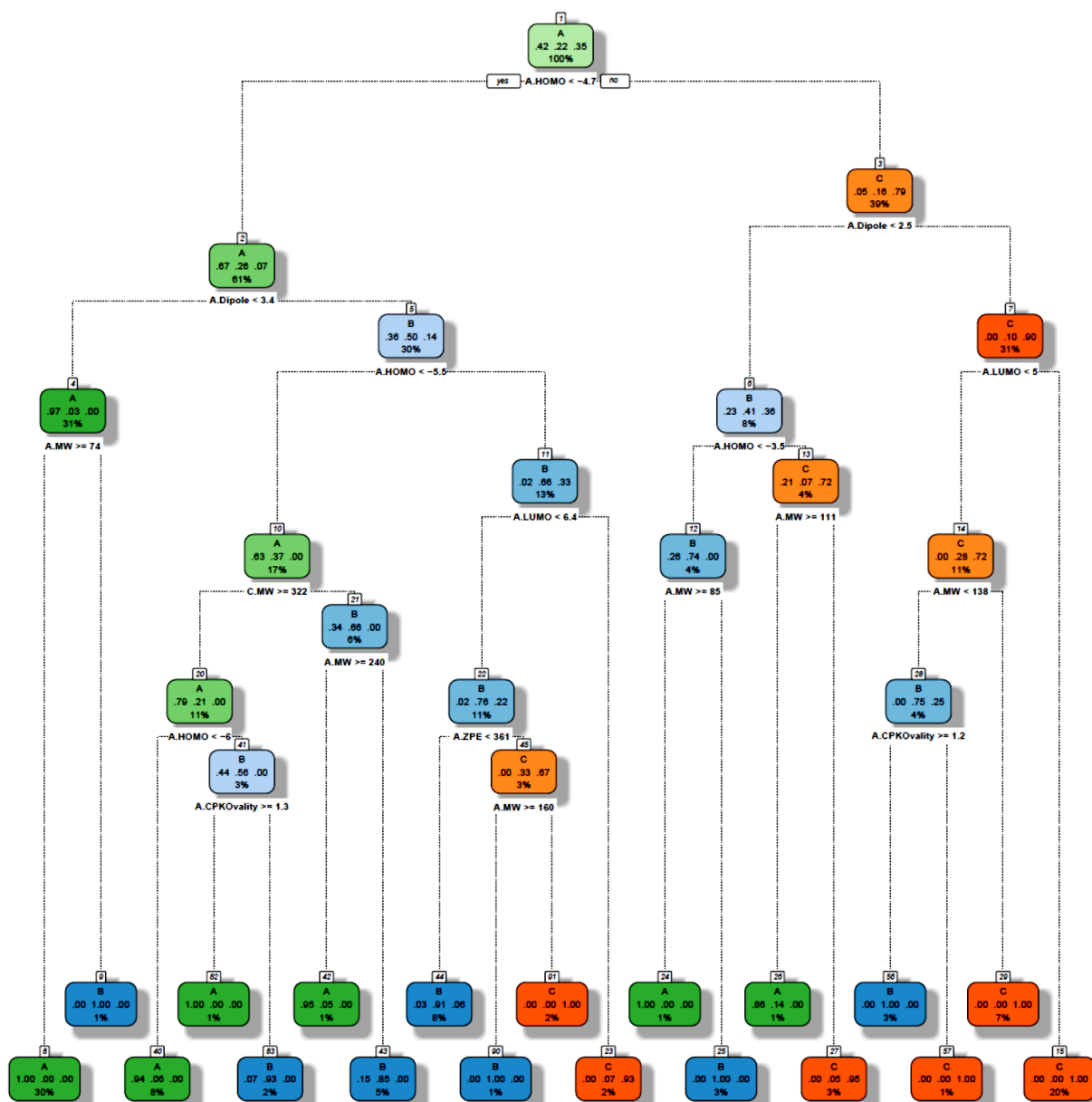


Figure C.4. Optimum decision tree structure for phosphonium dataset.

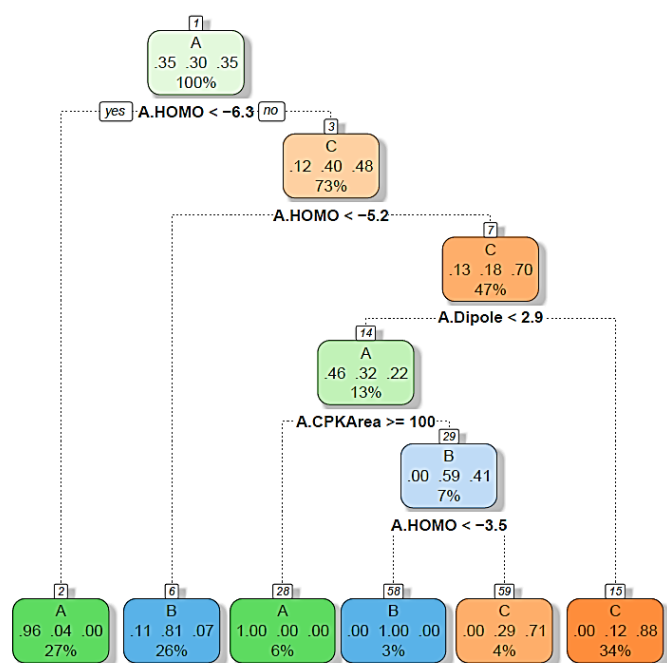


Figure C.5. Optimum decision tree structure for piperidinium dataset.

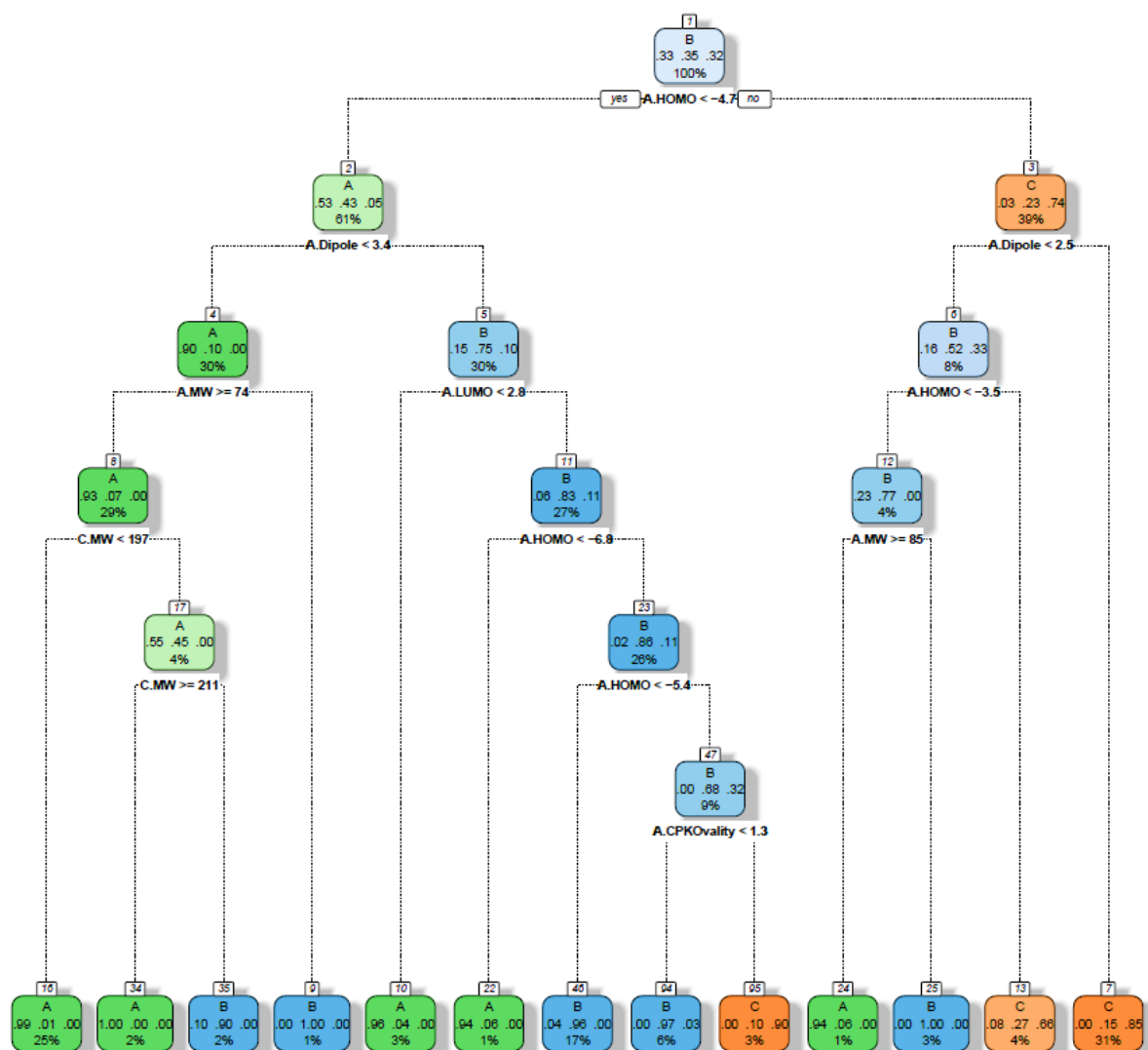


Figure C.6. Optimum decision tree structure for pyridinium dataset.

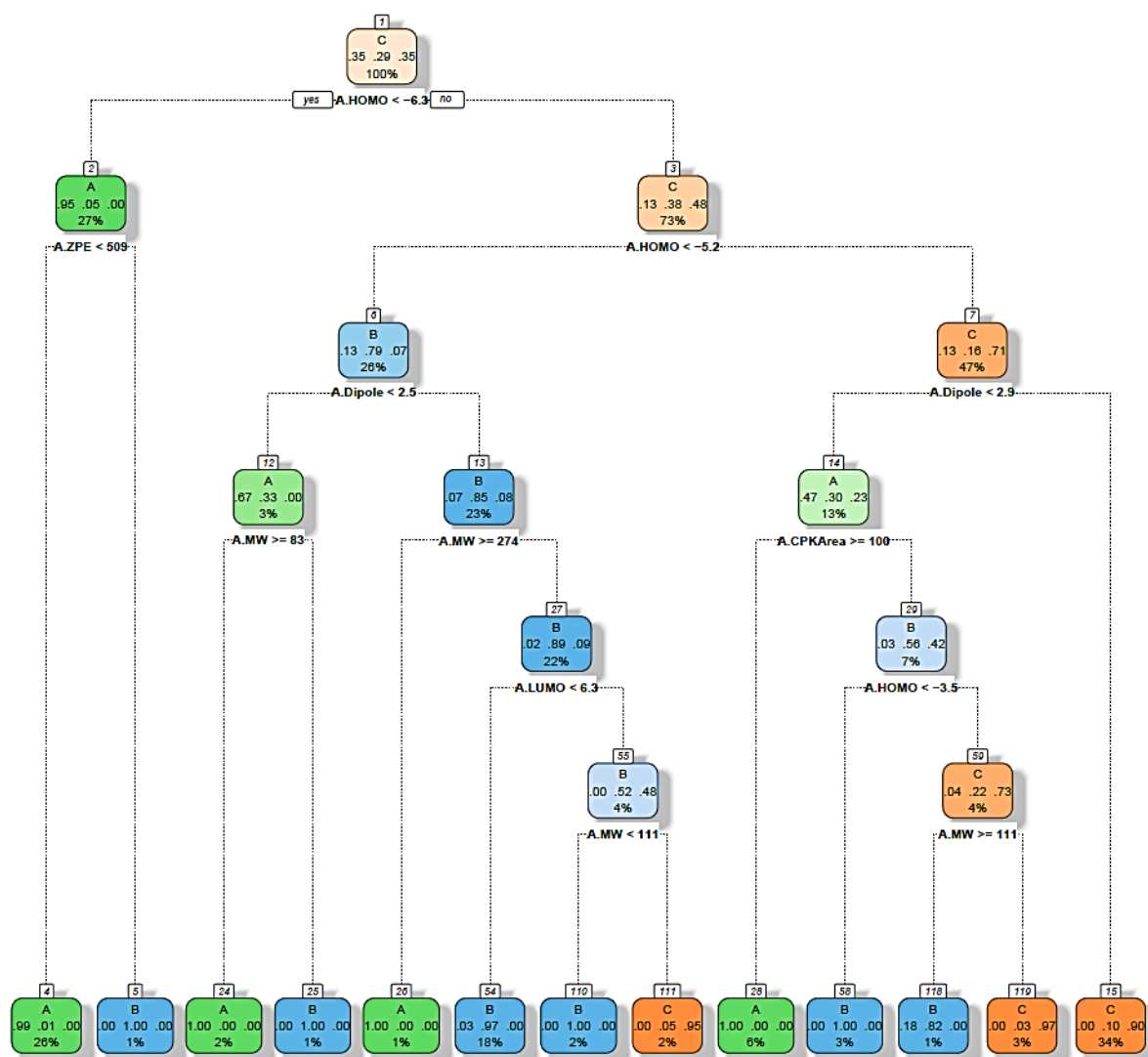


Figure C.7. Optimum decision tree structure for pyrrolidinium dataset.

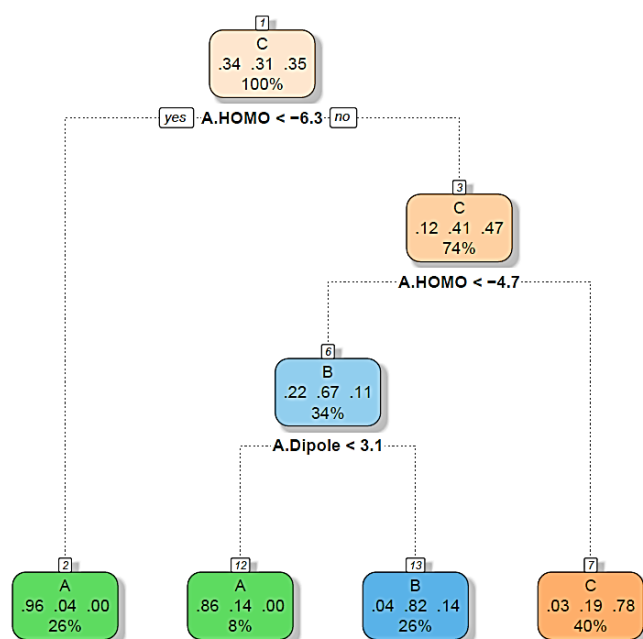


Figure C.8. Optimum decision tree structure for sulfonium dataset.

Table C.12. Detailed information about the 49 ILs.

Cation	Anion	Cation Code	Anion Code	Experimental solubility values found in literature	REFs
methyl-trioctyl-ammonium_cation	tf2n_anion	CAmm11	A87	0.03	S1
tributylmethylammonium_cation	tf2n_anion	CAmm19	A87	0.11	S1
butyltrimethylammonium_cation	tf2n_anion	CAmm2	A87	0.23	S1
1,3-diethylimidazolium_cation	tf2n_anion	CIIm14	A87	0.29	S2
1,3-dihexyl-imidazolium_cation	tf2n_anion	CIIm17	A87	0.14	S3
1-butyl-2-3-methyl-imidazolium_cation	hexafluorophosphate_anion	CIIm34	A72	0.18	S4
1-butyl-2-3-methyl-imidazolium_cation	tf2n_anion	CIIm34	A87	0.21	S2
1-butyl-3-ethylimidazolium_cation	tf2n_anion	CIIm38	A87	0.22	S5
1-butyl-3-methyl-imidazolium_cation	acetate_anion	CIIm20	A9	76.92	S6
1-butyl-3-methyl-imidazolium_cation	tetrafluoroborate_anion	CIIm20	A18	1.00	S5
1-butyl-3-methyl-imidazolium_cation	bromide_anion	CIIm20	A31	22.22	S6
1-butyl-3-methyl-imidazolium_cation	chloride_anion	CIIm20	A35	40.00	S6
1-butyl-3-methyl-imidazolium_cation	iodide_anion	CIIm20	A55	0.46	S7
1-butyl-3-methyl-imidazolium_cation	methanesulfonate_anion	CIIm20	A62	10.31	S6
1-butyl-3-methyl-imidazolium_cation	hexafluorophosphate_anion	CIIm20	A72	0.25	S4
1-butyl-3-methyl-imidazolium_cation	tf2n_anion	CIIm20	A87	0.24	S4
1-butyl-3-methyl-imidazolium_cation	thiocyanate_anion	CIIm20	A88	3.31	S6
1-butyl-3-methyl-imidazolium_cation	tricyanomethane_anion	CIIm20	A91	0.85	S4
1-butyl-3-methyl-imidazolium_cation	trifluoroacetate_anion	CIIm20	A92	7.52	S6
1-butyl-3-methyl-imidazolium_cation	trifluoromethane-sulfonate_anion	CIIm20	A93	1.07	S6
1-butyl-3-methyl-imidazolium_cation	tris(trifluoromethylsulfonyl)methide_anion	CIIm20	A98	0.14	S5
1-decyl-3-methyl-imidazolium_cation	tf2n_anion	CIIm26	A87	0.15	S3
1-ethyl-3-methyl-imidazolium_cation	tf2n_anion	CIIm18	A87	0.26	S4
1-heptyl-3-methyl-imidazolium_cation	tf2n_anion	CIIm23	A87	0.18	S4
1-hexyl-2-3-methyl-imidazolium_cation	tf2n_anion	CIIm35	A87	0.21	S8

Table C.12. Detailed information about the 49 ILs. (cont.)

Cation	Anion	Cation Code	Anion Code	Experimental solubility values found in literature	REFs
1-hexyl-3-methyl-imidazolium_cation	hexafluorophosphate_anion	CIm22	A72	0.20	S4
1-hexyl-3-methyl-imidazolium_cation	tf2n_anion	CIm22	A87	0.20	S4
1-methyl-3-nonylimidazolium_cation	tf2n_anion	CIm25	A87	0.16	S3
1-octyl-3-methyl-imidazolium_cation	hexafluorophosphate_anion	CIm24	A72	0.20	S5
1-octyl-3-methyl-imidazolium_cation	tf2n_anion	CIm24	A87	0.19	S9
1-pentyl-3-methyl-imidazolium_cation	tf2n_anion	CIm21	A87	0.22	S10
1-propyl-3-methyl-imidazolium_cation	tf2n_anion	CIm19	A87	0.26	S10
trihexyl-tetradecyl-phosphonium_cation	bis(2,4,4-trimethylpentyl)phosphinate_anion	CPhos8	A19	0.89	S11
trihexyl-tetradecyl-phosphonium_cation	bromide_anion	CPhos8	A31	0.68	S11
trihexyl-tetradecyl-phosphonium_cation	chloride_anion	CPhos8	A35	0.82	S11
trihexyl-tetradecyl-phosphonium_cation	decanoate_anion	CPhos8	A37	0.86	S11
trihexyl-tetradecyl-phosphonium_cation	dicyanamide_anion	CPhos8	A39	0.51	S11
trihexyl-tetradecyl-phosphonium_cation	tf2n_anion	CPhos8	A87	0.09	S5
trihexyl-tetradecyl-phosphonium_cation	tricyanomethane_anion	CPhos8	A91	0.51	S5
1-butyl-1-methylpiperidinium_cation_c0.cosmo	tf2n_anion	Cpip6	A87	0.17	S12
1-methyl-1-propylpiperidinium_cation_c0.cosmo	tf2n_anion	Cpip7	A87	0.23	S2
1-butyl-pyridinium_cation	tf2n_anion	CPyr14	A87	0.27	S2
1-hexyl-pyridinium_cation	tf2n_anion	CPyr19	A87	0.25	S2
1-octyl-pyridinium_cation	tf2n_anion	CPyr1	A87	0.22	S2
4-methyl-n-butylpyridinium_cation	tf2n_anion	CPyr5	A87	0.24	S13
1-butyl-1-methyl-pyrrolidinium_cation	tf2n_anion	CPyrr5	A87	0.23	S2
1-methyl-1-propylpyrrolidinium_cation	tf2n_anion	CPyrr7	A87	0.25	S2
diethyl-methylsulfonium_cation	tf2n_anion	Csulf1	A87	0.25	S1
triethylsulfonium_cation	tf2n_anion	Csulf3	A87	0.20	S1

- (S1) Kurnia, K. A., Lima, F., Cláudio, A. F. M., Coutinho, J. A. P. and Freire, M. G. (2015). Hydrogen-bond acidity of ionic liquids: an extended scale. *Physical Chemistry Chemical Physics*, 17, 18980–18990.
- (S2) Freire, M. G., Santos, L. M. N. B. F., Fernandes, A. M., Coutinho, J. A. P. and Marrucho, I. M. (2007). An overview of the mutual solubilities of water-imidazolium-based ionic liquids systems. *Fluid Phase Equilibria*, 261, 449–454.
- (S3) Kurnia, K. A., Neves, C. M. S. S., Freire, M. G., Santos, L. M. N. B. F. and Coutinho, J. A. P. (2015). Comprehensive study on the impact of the cation alkyl side chain length on the solubility of water in ionic liquids. *Journal of Molecular Liquids*, 210, 264–271.
- (S4) Freire, M. G., Neves, C. M. S. S., Ventura, S. P. M., Pratas, M. J., Marrucho, I. M., Oliveira, J., ... Fernandes, A. M. (2010). Solubility of non-aromatic ionic liquids in water and correlation using a QSPR approach. *Fluid Phase Equilibria*, 294, 234–240.
- (S5) Zhou, T., Chen, L., Ye, Y., Chen, L., Qi, Z., Freund, H. and Sundmacher, K. (2012). An overview of mutual solubility of ionic liquids and water predicted by COSMO-RS. *Industrial and Engineering Chemistry Research*, 51, 6256–6264.
- (S6) Passos, H., Khan, I., Mutelet, F., Oliveira, M. B., Carvalho, P. J., Santos, L. M. N. B. F., ... Coutinho, J. A. P. (2014). Vapor-liquid equilibria of water + alkylimidazolium-based ionic liquids: Measurements and perturbed-chain statistical associating fluid theory modeling. *Industrial and Engineering Chemistry Research*, 53, 3737–3748.
- (S7) Yokozeki, A. and Shiflett, M. B. (2010). Water solubility in ionic liquids and application to absorption cycles. *Industrial and Engineering Chemistry Research*, 49, 9496–9503.
- (S8) Chapeaux, A., Simoni, L. D., Stadtherr, M. A. and Brennecke, J. F. (2007). Liquid phase behavior of ionic liquids with water and 1-octanol and modeling of 1-octanol/water partition coefficients. *Journal of Chemical and Engineering Data*, 52, 2462–2467.
- (S9) Kakiuchi, T. (2008). Mutual solubility of hydrophobic ionic liquids and water in liquid-liquid two-phase systems for analytical chemistry. *Analytical Sciences*, 24, 1221–1230.

- (S10) Zhai, N., Wang, Y. H., Zhu, L. M., Wang, J. H., Sun, X. H., Hu, X. Bin, ... Luo, Y. H. (2015). Sensitivity and specificity of diffusion-weighted magnetic resonance imaging in diagnosis of bladder cancers. *Clinical and Investigative Medicine*, 38, E173–E184.
- (S11) Freire, M. G., Carvalho, P. J., Gardas, R. L., Santos, L. M. N. B. F., Marrucho, I. M. and Coutinho, J. A. P. (2008). Solubility of water in tetradecyltriethylphosphonium-based ionic liquids. *Journal of Chemical and Engineering Data*, 53, 2378–2382.
- (S12) Kurnia, K. A., Sintra, T. E., Neves, C. M. S. S., Shimizu, K., Canongia Lopes, J. N., Gonçalves, F., ... Coutinho, J. A. P. (2014). The effect of the cation alkyl chain branching on mutual solubilities with water and toxicities. *Physical Chemistry Chemical Physics*, 16, 19952–19963.
- (S13) Freire, M. G., Neves, C. M. S. S., Shimizu, K., Bernardes, C. E. S., Marrucho, I. M., Coutinho, J. A. P., ... Rebelo, L. P. N. (2010). Mutual solubility of water and structural/positional isomers of N-alkylpyridinium-based ionic liquids. *Journal of Physical Chemistry B*, 114, 15925–15934.

PERMISSIONS

About Figure 2.2

Order Date	30-Aug-2021	Type of Use	Republish in a thesis/dissertation
Order License ID	1144258-1	Publisher	ROYAL SOCIETY OF CHEMISTRY
ISSN	1463-9076	Portion	Image/photo/illustration

LICENSED CONTENT

Publication Title	Physical chemistry chemical physics : PCCP	Rightholder	Royal Society of Chemistry
Article Title	Quantification of photocatalytic hydrogen evolution.	Publication Type	Journal
Author/Editor	Royal Society of Chemistry (Great Britain), Deutsche Bunsen-Gesellschaft für Physikalische Chemie., Koninklijke Nederlandse Chemische Vereniging., Società chimica italiana.	Start Page	3466
Date	01/01/1999	Issue	10
Language	English	Volume	15
Country	United Kingdom of Great Britain and Northern Ireland		

REQUEST DETAILS

Portion Type	Image/photo/illustration	Distribution	Worldwide
Number of images / photos / illustrations	1	Translation	Original language of publication
Format (select all that apply)	Print, Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	No
Duration of Use	Life of current edition	Incidental promotional use?	Yes
Lifetime Unit Quantity	Up to 499	Currency	EUR
Rights Requested	Main product		

NEW WORK DETAILS

Title	PHOTOCATALYTIC HYDROGEN PRODUCTION OVER IONIC LIQUID COATED SEMICONDUCTORS	Institution name	Bogazici University
Instructor name	Ramazan Yildirim	Expected presentation date	2021-09-07

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Elif Can
------------------------	-----	---	----------

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 1	Title of the article/chapter the portion is from	Quantification of photocatalytic hydrogen evolution.
Editor of portion(s)	Schwarze, Michael; Stellmach, Diana; Schröder, Marc; Kailasam, Kamalakannan; Reske, Rulle; Thomas, Arne; Schomäcker, Reinhard	Author of portion(s)	Schwarze, Michael; Stellmach, Diana; Schröder, Marc; Kailasam, Kamalakannan; Reske, Rulle; Thomas, Arne; Schomäcker, Reinhard
Volume of serial or monograph	15	Issue, if republishing an article from a serial	10
Page or page range of portion	3466	Publication date of portion	2013-02-13

About Figure 2.5 and Figure 2.6

ISSN:	0360-3199	Publisher:	PERGAMON
Type of Use:	Republish in a thesis/dissertation	Portion:	Image/photo/illustration

LICENSED CONTENT

Publication Title	International journal of hydrogen energy	Rightholder	Elsevier Science & Technology Journals
Article Title	An overview of photocells and photore...	Publication Type	Journal
Author/Editor	INTERNATIONAL ASSOCIATION FOR HY...	Start Page	5233
Date	01/01/1976	Issue	11
Country	United Kingdom of Great Britain and N...	Volume	35

REQUEST DETAILS

Number of images / photos / illustrations	2	Distribution	Worldwide
Format (select all that apply)	Print, Electronic	Translation	Original language of publication
Who will republish the content?	Academic institution	Copies for the disabled?	No
Duration of Use	Life of current edition	Minor editing privileges?	No
Lifetime Unit Quantity	Up to 499	Incidental promotional use?	Yes
Rights Requested	Main product	Currency	EUR

NEW WORK DETAILS

Title	PHOTOCATALYTIC HYDROGEN PRODU...	Institution name	Bogazici University
Instructor name	Ramazan Yildirim	Expected presentation date	2021-09-07

ADDITIONAL DETAILS

The requesting person / organization to appear on the license	Elif Can
---	----------

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Figure 2 and Figure 3	Title of the article/chapter the portion is from	An overview of photocells and photore...
Editor of portion(s)	Minggu, Lorna Jeffery; Wan Daud, Wan...	Author of portion(s)	Minggu, Lorna Jeffery; Wan Daud, Wan...
Volume of serial or monograph	35	Issue, if republishing an article from a serial	11
Page or page range of portion	5233-5244	Publication date of portion	2010-06-01

About Figure 2.9

Order Date	30-Aug-2021	Type of Use	Republish in a thesis/dissertation
Order License ID	1144265-1	Publisher	ROYAL SOCIETY OF CHEMISTRY
ISSN	1477-9234	Portion	Image/photo/illustration

LICENSED CONTENT

Publication Title	Dalton transactions	Rightsholder	Royal Society of Chemistry
Article Title	Immobilization of molecular catalysts in supported ionic liquid phases.	Publication Type	e-Journal
Author/Editor	Royal Society of Chemistry (Great Britain)	Start Page	8377
Date	01/01/2003	Issue	36
Language	English	Volume	39
Country	United Kingdom of Great Britain and Northern Ireland		

REQUEST DETAILS

Portion Type	Image/photo/illustration	Distribution	Worldwide
Number of images / photos / illustrations	1	Translation	Original language of publication
Format (select all that apply)	Print, Electronic	Copies for the disabled?	No
Who will republish the content?	Academic institution	Minor editing privileges?	No
Duration of Use	Life of current edition	Incidental promotional use?	Yes
Lifetime Unit Quantity	Up to 499	Currency	EUR
Rights Requested	Main product		

NEW WORK DETAILS

Title	PHOTOCATALYTIC HYDROGEN PRODUCTION OVER IONIC LIQUID COATED SEMICONDUCTORS	Institution name	Bogazici University
Instructor name	Ramazan Yildirim	Expected presentation date	2021-09-07

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Elif Can
------------------------	-----	---	----------

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Scheme 1	Title of the article/chapter the portion is from	Immobilization of molecular catalysts in supported ionic liquid phases.
Editor of portion(s)	Van Doorslaer, Charlie; Wahlen, Joos; Mertens, Pascal; Binnemans, Koen; De Vos, Dirk	Author of portion(s)	Van Doorslaer, Charlie; Wahlen, Joos; Mertens, Pascal; Binnemans, Koen; De Vos, Dirk
Volume of serial or monograph	39	Issue, if republishing an article from a serial	36
Page or page range of portion	8377	Publication date of portion	2010-08-31