##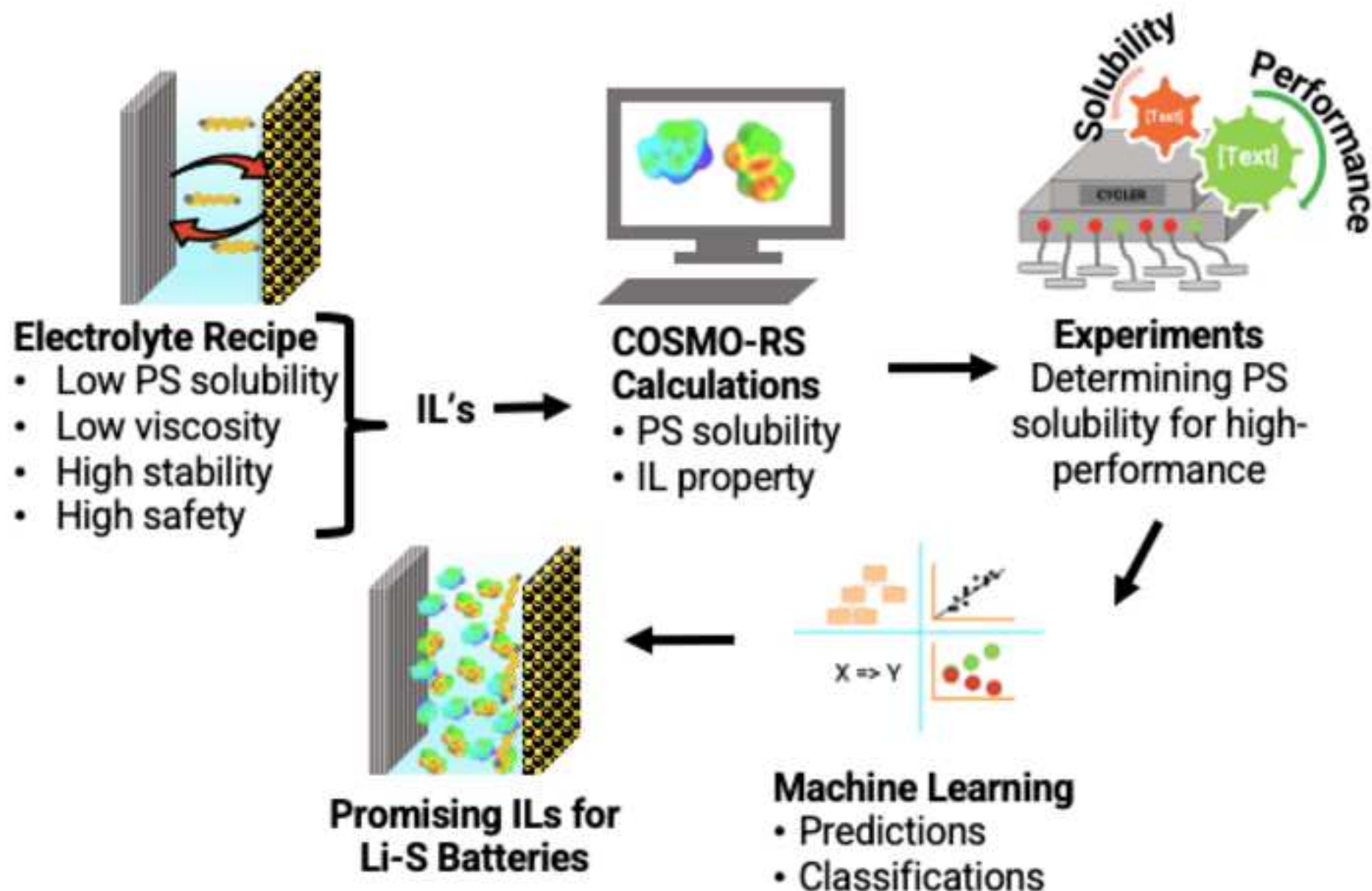 Selection of Ionic Liquid Electrolytes for High-Performing Lithium-Sulfur Batteries: An Experiment-Guided High-Throughput Machine Learning Analysis
--Manuscript Draft--

Manuscript Number:	CEJ-D-24-05279
Article Type:	Research Paper
Section/Category:	Novel Materials for Energy and Advanced Applications
Keywords:	Li-S battery; electrolyte design; ionic liquid; polysulfide solubility; electrolyte viscosity; ionic liquid descriptor
Abstract:	The polysulfide (PS) shuttle mechanism (PSM) is one of the most significant challenges of lithium-sulfur (Li-S) batteries in achieving high capacity and cyclability. One way to minimize the shuttle effect is to limit the PS solubilities in the battery electrolyte. Ionic liquids (IL) are particularly suited as electrolyte solvents because of their tunable physical and chemical properties. In this work, thousands of ILs are screened to narrow down potentially viable candidates to be used as electrolytes in Li-S batteries. To that end, the COnductor-like Screening Model for Realistic Solvents (COSMO-RS) calculations are performed over more than 36000 ILs. An extensive database containing PS solubilities and other relevant properties is constructed at 25 °C. First, the effectiveness of the COSMO-RS calculations is experimentally tested with six different ILs having a wide range of solubility and viscosity values; a strong correlation between the PS solubility and battery performance is obtained. After specifying the target limits for promising ILs using the experimental battery performance data, machine learning (ML) tools are used to predict and identify the relationship between IL properties and PS solubilities and structural and molecular descriptors of ILs. The extreme gradient boosting (XGBoost) method successfully predicts the solubility and property values. Association rule mining (ARM) and the feature importance analysis show that anion descriptors are more dominant, whereas cations have less impact on the solubilities and properties of ILs. Finally, the imidazolium and pyridinium ILs with bisimide and borate anion groups are identified as the most promising ones.

Graphical Abstract

- COSMO-RS calculations performed over more than 36000 ILs
- Experiments show a strong correlation between PS solubility and battery performance
- Machine learning models developed using experiment-guided target solubility limits
- XGBoost models successfully predict the PS solubility and IL properties
- ARM and feature importance analysis show that anion descriptors are more dominant

1 **Selection of Ionic Liquid Electrolytes for High-Performing Lithium-Sulfur Batteries: An**
2 **Experiment-Guided High-Throughput Machine Learning Analysis**
3
4
5 *Aysegul Kilic[a], Omar Abdelaty[a], Muhammad Zeeshan[b,e], Alper Uzun[b,c,d], Ramazan Yildirim[a*],*

6 *Damla Eroglu[a*]*

7 A. Kilic, O. Abdelaty, M. Zeeshan, Prof. A. Uzun, Prof. R. Yildirim, Prof. D. Eroglu

8 [a] Bogazici University, Department of Chemical Engineering, 34342 Istanbul, Turkey
9 [b] Koç University, Department of Chemical and Biological Engineering, 34450 Istanbul, Turkey
10 [c] Koç University, Koç University TÜPRAŞ Energy Center (KUTEM), 34450 Istanbul, Turkey
11 [d] Koç University, Koç University Surface Science and Technology Center (KUYTAM), 34450
12 Istanbul, Turkey
13 [e] Case Western Reserve University, Department of Chemical and Biomolecular Engineering,
14 44106, Cleveland, OH, USA
15 [*]Corresponding authors

16 E-mail:

17 *R. Yildirim: yildirra@boun.edu.tr

18 *D. Eroglu: eroglud@boun.edu.tr

19
20
21 **Abstract**

22 The polysulfide (PS) shuttle mechanism (PSM) is one of the most significant challenges of lithium-

23 sulfur (Li-S) batteries in achieving high capacity and cyclability. One way to minimize the shuttle

24 effect is to limit the PS solubilities in the battery electrolyte. Ionic liquids (IL) are particularly

25 suited as electrolyte solvents because of their tunable physical and chemical properties. In this

26 work, thousands of ILs are screened to narrow down potentially viable candidates to be used as

27 electrolytes in Li-S batteries. To that end, the COnductor-like Screening Model for Realistic

28 Solvents (COSMO-RS) calculations are performed over more than 36000 ILs. An extensive

29 database containing PS solubilities and other relevant properties is constructed at 25 °C. First, the

30 effectiveness of the COSMO-RS calculations is experimentally tested with six different ILs having

31 a wide range of solubility and viscosity values; a strong correlation between the PS solubility and

32 battery performance is obtained. After specifying the target limits for promising ILs using the

33 experimental battery performance data, machine learning (ML) tools are used to predict and

34 identify the relationship between IL properties and PS solubilities and structural and molecular

35 descriptors of ILs. The extreme gradient boosting (XGBoost) method successfully predicts the

36 solubility and property values. Association rule mining (ARM) and the feature importance analysis

37 show that anion descriptors are more dominant, whereas cations have less impact on the solubilities

38 and properties of ILs. Finally, the imidazolium and pyridinium ILs with bisimide and borate anion

39 groups are identified as the most promising ones.

40

41 Keywords: Li-S battery, ionic liquid, polysulfide solubility, electrolyte viscosity, ionic liquid
42 descriptor
43

44

45 **List of Abbreviation**

Abbreviation	Long name
ARM	Association rule mining
BETA	Bis(pentafluoroethanesulfonyl)amide
BF_4	Tetrafluoroborate ($BF4-$)
BMIM	1-butyl-3-methyl-imidazolium
C4dmim	1-butyl-2,3-dimethyli-imidazolium
CF_3SO_3	Trifluoromethane-sulfonate
COSMO-RS	COnductor-like Screening Model for Realistic Solvents
CPK	Corey, Pauling, and Koltun space filling model
DEME	N,N-diethyl-N-methyl-N-(2-methoxyethyl)-ammonium
DFT	Density functional theory
DME	1,2-Dimethoxyethane
DOL	1,3-Dioxolane
FSI	Bis(fluorosulfonyl)imide
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor

HOMO	Highest occupied molecular orbital
IL	Ionic liquid
Li-S	Lithium-sulfur
LUMO	Lowest unoccupied molecular orbital
$MeSO_4$	Methyl-sulfate
ML	Machine learning
MW	Molecular weight
OTF	Trifluoromethane-sulfonate
P13	1-methyl-1-propyl-pyrrolidinium
P14	1-butyl-1-methyl-pyrrolidinium
P2225	Triethyl-pentyl-phosphonium
PF_6	Hexafluorophosphate
phosh./phosphin.	Phosphate/phosphininate
PP13	1-methyl-1-propyl-piperidinium
PP14	1-butyl-1-methylpiperidinium
PS	Polysulfide
PSM	Polysulfide shuttle mechanism
RMSE	Root-mean-square-error
SM	Supplementary Material
TBMA	Tributylmethylammonium
TFSI	Bis(trifluoromethane)sulfonimide
XGBoost	Extreme gradient boosting
ZPE	Zero point energy

46

1. Introduction

48 Lithium-ion (Li-ion) batteries have so far successfully met the requirements of many applications

49 and have been widely accepted as state-of-the-art owing to their high voltage range (2.5–4 V),

50 negligible memory effect, long life (>1000 cycles), relatively high energy density (240 Wh/kg;

51 640 Wh/L), and high energy conversion efficiency for storage purposes. Despite these advantages,

52 Li-ion batteries have a theoretical energy density of 800 Wh/kg, which may be insufficient to meet

53 the targets for future transport and grid storage applications. Furthermore, limited availability and

54 the high cost of some active materials manufacturing the Li-ion batteries prevent their wider

55   adaptation. Therefore, beyond Li-ion battery technologies have been explored to reach the set

56   targets regarding energy density, cost, and other essential properties of future applications [1–4].

57   In this respect, lithium-sulfur (Li-S) batteries are one of the best alternatives to Li-ion batteries.

58   They have high theoretical gravimetric and volumetric energy densities of 2600 Wh/kg and 2800

59   Wh/L, respectively. Sulfur, with a theoretical specific capacity of 1675 mAh/g, is an abundant,

60   non-toxic, and environmentally friendly cathode active material. However, Li-S batteries suffer

61   from major drawbacks, namely the shuttling of the polysulfide (PS) intermediates, which is called

62   the polysulfide shuttle mechanism (PSM), the insulating nature of sulfur and Li_2S, and lithium

63   anode degradation, that must be resolved before their practical applications and broader

64   commercial adaptations.

65   The PSM is a fatal problem in the Li-S batteries [3–5], typically addressed using specialized

66   cathodes, separators, and/or electrolytes. In general, electrolytes of any battery chemistry are

67   optimized to have high ionic transport (attributed to low viscosity), good chemical, thermal, and

68   electrochemical stability (i.e., wide electrochemical window), low cost, and high safety. The

69   shuttle effect puts an additional constraint on standard electrolyte properties by requiring that the

70   electrolyte has low but finite PS solubility and mobility to minimize the loss of active material and

71   cathode structure degradation while retaining sufficient kinetics in Li-S batteries, especially at high

72   sulfur loadings [6]. This constraint is further complicated by the fact that the PSs form a highly

73   dynamic system that undergoes several rapid disproportionation reactions whereby S-S bonds are

74   continuously being broken and recreated [7,8]. Hence, it is difficult to measure or predict the

75   solubility of PSs in the electrolyte.

76   The typically used liquid electrolytes are ether- or aprotic solvent-based electrolytes that offer

77   excellent ionic transport and stability of PS ions. Specifically, the mixture of 1,3-Dioxolane (DOL)

78   and 1,2-Dimethoxyethane (DME) with the lithium bis(trifluoromethane)sulfonimide (LiTFSI) salt

79   is the most common electrolyte used in Li-S batteries. Unfortunately, such solvents suffer from

80   very high PS solubility and, hence, excessive PS shuttle rate. In addition, the low viscosity of the

81   DOL:DME electrolyte also facilitates the fast diffusion of PSs inside the cell. Ionic liquids (ILs)

82   are proposed as alternative electrolyte solvents in the battery literature [1,9,10]. IL electrolytes

83   have relatively low PS solubilities and excellent chemical, electrochemical, and thermal stabilities.

84   Many ILs have sufficiently low melting points as well. However, their high viscosity and, thus,

85   moderate ionic conductivity challenge pose a downside. In the literature, typically

86   bis(trifluoromethane)sulfonimide ([TFSI]$^-$) anion is explored with cations from different groups,

87   including pyrrolidinium, piperidinium, sulfonium, and imidazolium as Li-S battery electrolytes

88   [11–13]. Recently, ILs have been used in hybrid electrolyte systems to combine the benefits and

89   mitigate the drawbacks of both ether-based and IL solvents [12,13]. The most common ethers

90   include DOL, DME, and tetraethylene glycol dimethyl ether (TEGDME).

91   To reap the benefits of IL electrolytes, screening for ILs with optimum electrolyte properties,

92   including low melting point, viscosity, and PS solubility, is necessary. Among them, the most

93   difficult property to measure and predict is the solubility of PSs. Such difficulty arises from the

94   dynamic nature of these ions, in addition to the inherent difficulty of measuring solubility values.

95   To simplify the problem, Ueno et al. [14] measured the total solubility of atomic sulfur, which

96   included all PS species in the system, in the solvate IL electrolytes of equimolar mixtures of lithium

97   salts and triglyme and tetraglyme (Li(Glyme)X) combined with various anions. In the end, they

98   concluded that the stronger the donor ability of the IL, the higher their ability to solvate PSs. More

99   specifically, they found that [TFSI]$^-$ and bis(pentafluoroethanesulfonyl)amide ([BETA]$^-$) anions

100   effectively suppressed the dissolution of Li$_2$S$_x$ species while maintaining reasonable conductivity,

101   viscosity, and stability. Park et al. [15] from the same group conducted a similar study to measure

102   the solubility of PSs in various common ILs. For the [P14]-[OTF](trifluoromethane-sulfonate) IL,

103   the solubility of PSs was relatively high, which causes poor cycling performance when used as the

104   electrolyte in a Li-S battery. The results indicated a close connection between the PS solubility in

105   the electrolyte and battery performance. In addition, it was also discussed that when the PS

106   solubilities are similar, the viscosities of the electrolytes become important for good ionic transport

107   in the Li-S cells. Similarly, low PS solubility and high Li^+ ion diffusion were found to be the key

108   criteria, especially at low electrolyte-to-sulfur ratios, to attain high energy density Li-S batteries

109   [16].

110   The presence of such complex and diverse constraints makes the search for optimum ILs a

111   complex, multi-dimensional, and non-linear problem. Extensive experimental testing of a wide set

112   of ILs is not feasible considering that the number of available cations, anions, and their

113   combinations in ILs is staggeringly massive, and experimental testing of even a handful of ILs is

114   heavily time-consuming. Instead, experts can use chemical intuition and educated guessing to find

115   and test potentially suitable candidates as a common practice in materials science. This chemical

116   intuition is learned by experience over a long period of time, and learning from previous

117   experiences is the basis for the machine learning (ML) [17]. On the other hand, fast screening

118   calculations are helpful in narrowing down potential targets for further investigation, which only

119   requires crude and quick models. One such model is COSMO-RS, which stands for Conductor-

120   Like Screening Model for Real Solvents [18]. COSMO-RS is a well-established quantum

121   chemistry-based computational model to estimate semi-quantitatively the thermodynamic

122   properties of chemical mixtures [19]. The basic idea behind the COSMO-RS model is to treat the

123 interactions in a molecular system as if they are between an ensemble of discrete surface screening

124 charge densities, denoted as sigma (σ), for each molecule or ion in the system [20–22].

125 Although the COSMO-RS model calculations are sufficiently fast, they depend on the surface

126 charge densities of the solvent and solute ions, which are not readily available. Therefore, analysis

127 of a sufficiently large representative set of ILs must be done and interpreted to generalize these

128 results on all ILs. With the assistance of ML algorithms, it is possible to find the hidden

129 correlations between the structural, physical, and chemical properties of the IL and the estimated

130 solubility or other predicted properties. ML is essential for materials search and is becoming more

131 feasible with the continuous growth of large materials databases and simulation tools [17,23].

132 Similar to our objective here, some studies in the literature utilize both COSMO-RS solubility

133 predictions and ML techniques to determine suitable ILs for their objectives [24–27]. In fact, water

134 and C_4 hydrocarbon solubilities in ILs were determined in our previous studies [28,29].

135 Meanwhile, only two studies in the Li-S battery field have used the COSMO-RS method for

136 solubility calculations. The first study focused on predicting cyclic sulfur solubility in ten solvents,

137 and they correctly captured experimental values [30]. The same group used ML classification

138 based on PS solubilities by using the electrolyte descriptors obtained from sigma profiles of

139 electrolytes calculated by the COSMO-RS method for a limited number of solvate-ILs, salt-in-

140 solvent and solvent-in-salt electrolytes [31]. These studies indicate that combining COSMO-RS

141 solubility predictions with ML methods is highly promising. Yet, the number of electrolytes

142 investigated in the previous studies is minimal to develop heuristic rules or predictive models. In

143 contrast, this research focuses on the solubility of Li_2S_8 in around 36,000 ILs for fast screening,

144 building ML models for predictions, and showing structure-property relationships, which have not

145 been addressed before.

146   In this work, around 36,000 IL solvents are screened for their potential as Li-S battery electrolytes

147   based on their PS (Li_2S_8) solubilities, viscosities, and melting points obtained by COSMO-RS

148   calculations. Moreover, the cycling performance of Li-S cells with six commercially available ILs

149   having low and high viscosities and PS solubilities is examined experimentally to determine the

150   selection criteria used for ML models; a critical connection between the COSMO-RS solubilities

151   and viscosities and the battery performance is identified. The results are further used to train ML

152   prediction models, specifically XGBoost, and the structure-property relationship in ILs is

153   identified using association rule mining (ARM) using ten descriptors for each anion and cation

154   determined by the density functional theory (DFT) calculations.

155   **2.   Computational Methods**

156   *2.1. Solubility and Property Calculations:*

157   COSMOThermX software was used for the estimation of Li_2S_8 solubilities in ILs at 25 ºC by

158   performing COSMO-RS calculations [32]. The IL database, COSMObaseIL, was used to form the

159   dataset for this study. COSMOThermX uses the σ-profiles calculated using the DFT functional BP

160   and def2-TZVP level. The dataset comprised 98 anions and 370 cations, which adds up to 36,260

161   pairs of ILs. The long chain Li_2S_8 was used to model the PSs in the system. In accordance with

162   previous reports about Li_2S_8 conformation in the solvents [8,33], the linear conformation of Li_2S_8

163   was first optimized using DFT by employing the B3LYP/def2-TZVP level theory with the ORCA

164   quantum chemistry software package [34]. The optimized structure conformation was then used

165   as an input file for TMOLEX (v.4.5.3)[35] to generate the standard σ-profile file (*.cosmo* file). The

166   sigma profile and the corresponding surface of the Li_2S_8 molecule are given in Figure S1 in the

167   Supplementary Material (SM).

168   The *IL Screening Module* was used to calculate the PS solubility in the ILs at 25 °C via equation

169   1, where C_i^∞ and γ_i^∞ are the capacity (mol/mol solubility) and activity coefficient of the PS at

170   infinite dilution, respectively.

171

$$C_i^\infty = \frac{1}{\gamma_i^\infty} \qquad\qquad (1)$$

172

173   The melting point, viscosity, and electrical conductivity were also calculated at 25 °C using the IL

174   Properties Module, which also uses the same σ-profile files for the ions. The densities of ILs were

175   calculated using the molecular volume obtained by the COSMO-RS volume calculations, whereas

176   COSMO-RS enthalpies were used for the melting point calculations. Finally, the same formula

177   utilizing ionic radius and dielectric energies was used for the viscosity and electronic conductivity

178   calculations. The implemented correlation coefficients were naturally different for each

179   calculation**.**

180   *2.2. Structural Descriptors for ILs*

181   To calculate structural descriptors for ILs, molecular geometries of cations and anions were

182   individually optimized using Spartan'14 by employing the PM3 semi-empirical method with the

183   default convergence criteria. Similar to our previous studies [28,29], ten descriptors were

184   calculated for each anion and cation separately: Molecular weight (MW in amu), highest occupied

185   molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies (E_{HOMO},

186   E_{LUMO} in eV), CPK-area (in Å^2) and CPK ovality (O) obtained from the space-filling model by

187   Corey, Pauling, and Koltun (CPK), dipole (μ in D), polarizability (m^3), vibrational zero point

188   energy (ZPE in kJ/mol), hydrogen bond donor count (HBD), and hydrogen bond acceptor count

189 (HBA). Consequently, there are 20 descriptors for each IL pair. The distribution of ions in terms

190 of these structural descriptors is given in Figure 1. The complete dataset, including all the

191 properties and the descriptors, can be found in SM as a Microsoft Office Excel^{TM} file.

192 *2.3. ML Modeling*

193 Distinct prediction models were trained for each IL property of interest: the solubility of Li_2S_8,

194 melting point, viscosity, and electrical conductivity. Random sampling was employed to negate

195 the bias towards the majority class. Afterward, the dataset was partitioned into training and test

196 sets using a 75%-25% split. Given that PS solubility highly depends on the anion type, random

197 assignment of the data entries into each subset was avoided as it overfitted the ions in the set and

198 performed poorly on new ions (due to the strong effects of the anions). Therefore, the splitting was

199 done either according to the anion or the cation groups. In practice, such splitting means that a

200 random 25% fraction of anions/cations in each anion/cation group was included in the test set,

201 while the remaining ions were used in the training set. This way, the test set contains new

202 anions/cations not encountered during training or validation. The same strategy was used for 5-

203 fold cross-validation, where 20% of the training ions were selected in the validation sets, and the

204 rest were used for training. Even with this additional constraint, successful and robust models were

205 developed. Next, the partitioned dataset, along with the respective values of properties calculated

206 by COSMOThermX, were used to train ML models. XGBoost algorithm was used in the

207 predictions. The hyperparameters were optimized with a grid search according to the performance

208 criteria of root mean square error (RMSE), which shows the average difference between the

209 predicted and the COSMO-RS calculated values; hence, the lower it is, the better the predictions.

210 In addition, the R-squared (R^2) value, the proportion of the difference between the two, ranging

211 between 0 and 1, was also reported [36]. The optimized hyperparameters for XGBoost are the

212   maximum depth, number of trees, and learning rate (η) values, which were found to be 3, 225, and

213   0.1, respectively, for the solubility predictions. The descriptor importance was calculated with

214   these hyperparameters over the training set.

215   After the prediction models were built, the ARM method was utilized to classify the promising

216   anion and cations together with their descriptors. The classification models were trained as binary

217   classifiers: class A and class B are the favorable and non-favorable levels of related properties and

218   solubilities, respectively.

219   The thresholds of these classes were determined according to the experimental results and the

220   desired characteristics of a battery electrolyte. In the case of solubility, a threshold value of -0.7 to

221   0.1 mol/mol in the log scale was selected as it is comparable to ILs tested in the literature and the

222   performance data obtained in our research [15]. Similarly, limits for the viscosity, conductivity,

223   and melting point classifications were determined as 100 mPa.s, 2 mS/cm, and 0 □C, respectively.

224   In the ARM analysis, all these four targets were held to see the most promising ILs as electrolytes

225   of Li-S batteries.

226   The relationships between the anion/cation groups and their descriptors with class A in all four

227   criteria were independently analyzed by a single-factor ARM algorithm. Since ARM only works

228   for categorical data, HBA and HBD counts were turned into factors, and the remaining eight

229   numeric descriptors were categorized into ten multiple intervals with a similar number of data

230   points. The performance metrics, such as support, confidence, and lift, were used together to

231   determine the reliability of a rule. The frequency of a certain descriptor coinciding with the target

232   class in the entire dataset, $P(D \cap A)$, is termed support. Confidence is the probability that the IL

233   pair belongs to class A given that a certain descriptor is within a particular interval, $P(A|D)$. Lift

234   calculates the correlation by comparing the frequency or the support as compared to random

235   independent coincidence $P(D \cap A)/P(A) * P(D)$. Hence, the lift values larger and smaller than

236   1.0 correspond to positive and negative correlation, respectively. The more significant the lift

237   value, the stronger the correlation is between the descriptor and that class [37]. Thus, the lift value

238   was considered the ultimate criterion to assess the association between a descriptor and the

239   solubility, as long as the rule clears the minimum support and confidence thresholds, taken as 0.1%

240   and 3%, respectively. All the modeling and figure creation were performed in the R Studio

241   environment [38].

242   **3. Results and Discussion**

243   *3.1. Pre-analysis of the Dataset*

244   In this screening part, we aimed to use a dataset representative of most ILs commonly used and

245   studied, as well as the rare ones in the battery literature. Hence, our dataset spans several cation

246   groups, including imidazolium, pyridinium, and ammonium, in addition to anions of different

247   types, including fluorinated, chlorinated, carboxylates, oxyanions, amino acids, etc., and more

248   commonly investigated anions in IL electrolytes such as tetrafluoroborate ($[BF_4]^-$),

249   hexafluorophosphate ($[PF_6]^-$) and $[TFSI]^-$. The complete lists of the cations and anions groups are

250   given in Table 1, whereas members of each group are presented in the Excel file provided in SM.

251   The table shows that the most crowded groups are imidazolium and pyridinium for cations, while

252   amino acids and carboxylates for the anions.

253   The COSMO-RS solubility screening results, including solubility, viscosity, conductivity, and

254   melting points, are shown in Figure 2 depending on anion groups, whereas the distributions on the

255   cation groups are presented in Figure S3. When these distributions are compared, solubilities and

256   properties show similar distributions and ranges regardless of the cation groups. On the other hand,

11

257 noticeable differences are observed when anion groups are considered. This clearly indicates

258 anionic effect dominance on the properties of the ILs.

259 Figure 2a and Figure S2 for the whole dataset show that the solubility values for each anion group

260 obtained from the calculations span a wide scale of many orders of magnitude from 10^{-9} to almost

261 10^{20} in mol/mol units. This is expected primarily due to the large difference between the activity

262 coefficients of the solute (Li_2S_8) and the solvents (ILs). Still, the calculations can compare small

263 and larger values, but the extreme results should be treated cautiously [39,40]. When ILs with

264 extremely low and high PS solubilities are analyzed, it is seen that some anions give these

265 abnormal solubility values regardless of the cation types. This may indicate COSMO-RS's failure

266 to predict the properties of these anions. According to the graph, most amino acids, carboxylates,

267 halogens, non-metal oxides, and phosphines fall into the extreme solubility ranges, below 10^{-3} and

268 above 10^3 mol/mol, and should be treated cautiously. On the other hand, bis_imide, borate, halo-

269 elemental, and others (the rare ones that do not belong to any of the anion groups listed) anion

270 groups show reasonable solubility values.

271 The PS solubility is not the only criterion when selecting electrolytes for Li-S batteries. Other

272 properties are also essential in the screening process of ILs in search of suitable electrolytes for

273 Li-S batteries; the melting point, viscosity, and conductivity are all significant criteria in Li-S cell

274 electrolyte selection to ensure smooth cell operation. This way, the utilization of highly viscous IL

275 electrolytes must be avoided to allow for appreciable Li^+ diffusion. 1 M Li salt containing

276 DOL:DME electrolyte has a viscosity of 1.6 mPa.s and shows sufficient ionic conductivity [16].

277 In the dataset, the lowest viscosity value calculated using the COSMO-RS is 9.1 mPa.s, which is

278 almost 5-fold higher than that of DOL:DME electrolytes and only 25% of 36,260 ILs have

12

279 viscosity values below 134 mPa.s. On the other hand, around 15% and 7% of the dataset do not

280 have melting points below 0 °C and ln(electronic conductivity) values below 2, respectively.

281 The anion and cation descriptors have been shown to strongly correlate with target properties in

282 previous studies, including the solubility of water and C_4 hydrocarbons in various ILs and the

283 physicochemical properties of ILs [28,29]. The descriptors are simple yet essential structural,

284 electronic, and energetic factors. CPK area and ovality are related to the geometry of ions

285 calculated based on the space-filling model, which is a crucial indicator of the area of potential

286 interactions with other ions or the PSs. Naturally, the higher the area is, the stronger the interactions

287 in the solution are. However, the ovality effect also plays an important role in the space-filling of

288 ions surrounding the PSs and each other. Dipole and polarizability represent ions' charge

289 distribution and susceptibility of that distribution to deformation when interacting with other

290 molecules. These are highly important in studying solvation energy and solubility because they

291 directly relate to the interactions during solvation. The potential for hydrogen bond formation is

292 given by the hydrogen bond donor (HBD) and acceptor (HBA) counts [41]. PSs are all known to

293 be weak bases, in other words, hydrogen bond acceptors [42]. However, this also indicates the

294 potential for the cation and anion to be attracted to each other, which will also affect the solubility.

295 Other important descriptors include electronic and energetic ones, which may affect the solubility

296 less directly. The values of HOMO and LUMO energies calculated by simple DFT calculations

297 contain important information related to ion stability, its potential for electronic interactions, and

298 bond strength. The vibrational zero-point energy is the lowest vibrational energy level and

299 determines the flexibility or stiffness of the bonds in the molecule to stretching and bending. This

300 property describes the flexibility of ions to structural deformation during the solvation process

301 [43,44]. Finally, the molecular weight of ions is also useful as it gives information about ion

302 diffusion coefficients, density, and viscosity, which indirectly affect solubility.

303 *3.2. IL Selection Criteria for High-Performance Li-S Batteries*

304 Experiments were carried out to see if there is a correlation between PS solubilities and IL

305 properties calculated using the COSMO-RS method and Li-S battery performance. Li-S cells with

306 standard cathodes, carbon black mixed with sulfur using the melt-diffusion strategy, and mixed

307 electrolytes containing organic and IL solvents (75:25 vol %) were prepared. As the organic

308 electrolyte, 1 M LiTFSI and 0.1 M lithium nitrate ($LiNO_3$) containing DOL:DME (1:1 vol.%) was

309 used. Meanwhile, six commercially available ILs listed in Table 2 were used in the mixed

310 electrolytes, and cycling experiments were done for 50 cycles at 0.1 C (the experimental details

311 are provided in SM). As seen in the table, three different cations, 1-butyl-1-methylpiperidinium

312 ($[PP14]^+$),   N,N-diethyl-N-methyl-N-(2-methoxyethyl)-ammonium   ($[DEME]^+$),   and

313 tributylmethylammonium   ($[TBMA]^+$)   with $[TFSI]^-$ and   three   different   anions,

314 hexafluorophosphate ($[PF_6]^+$), trifluoromethane-sulfonate ($[CF_3SO_3]^+$), and methylsulfate

315 ($[MeSO_4]^+$) with 1-butyl-3-methyl-imidazolium ($[BMIM]^+$) are used in the experiments. These ILs

316 are commercially available, hence easily accessible, and have low/high solubility and viscosity

317 values (melting point and electrical conductivity are only used to confirm the suitable liquid phase

318 of the electrolyte and to ensure no electron flow through the electrolyte). Hence, four of these ILs

319 are projected to have low solubilities ($[PP14]$-$[TFSI]$, $[DEME]$-$[TFSI]$, $[TBMA]$-$[TFSI]$,

320 $[BMIM]$-$[PF_6]$), whereas the other two ($[BMIM]$-$[CF_3SO_3]$ and $[BMIM]$-$[MeSO_4]$) have

321 extremely high values. Moreover, ILs with low and high viscosities also have both high and low

322 solubility cases. Hence, it is possible to identify the effect of COSMO-RS predicted viscosity and

323 solubility values on the performance of Li-S batteries.

14

324   The results presented in Figure 3 supported the discussion on the importance of low solubility but

325   also highlighted the effect of viscosity on the Li-S cell performance. Although high viscosity may

326   suppress the PSM by restricting the PS movement, it also prevents the diffusion of Li^+ ions. In this

327   respect, the ILs of [PP14]-[TFSI] and [DEME]-[TFSI] with both low solubility and viscosity show

328   the best cycling performance. On the other hand, Li-S cells with ILs ([TBMA]-[TFSI], [BMIM]-

329   [PF_6]) with low PS solubility but higher viscosity performed moderately. Finally, [BMIM]-

330   [CF_3SO_3] and [BMIM]-[$MeSO_4$] showed almost zero capacity over cycling. These ILs are

331   predicted to have high solubility, indicating that even though low viscosity is required for high

332   performance, low PS solubility is a more critical property. Consistent with this inference, [BMIM]-

333   [CF_3SO_3] performs poorly, even though it has a low viscosity, proving that mostly PS solubility

334   determines the performance.

335   The experimental results show that it is possible to use the COSMO-RS predictions for assessing

336   Li-S battery performance. However, these results are only limited to six ILs, and further validation

337   of COSMO-RS results is still needed. Unfortunately, a sufficiently comprehensive experimental

338   study to evaluate the predictions made by the COSMO-RS model is excessively difficult,

339   necessitating a comparison with the available data in the literature. The few studies regarding

340   solubility had limited scope due to difficulties discussed previously. Among them, Park et al.

341   reported PS solubilities for various ILs. The experimentally measured PS solubilities account for

342   a mixture of PSs in the solution due to *disproportionation reactions* and the subsequent cascade

343   of PS reactions. This difference is partially responsible for the departure of predictions of Li_2S_8

344   solubility from experimental measurements of all PSs. Indeed, the measured solubilities are the

345   sum of solubilities of all PS species with various chain lengths and are provided in terms of total

346   atomic sulfur concentration. Despite this and other simplifying assumptions, the COSMO-RS

347 model results showed an excellent correlation with the experimental measurements as shown in

348 Figure 4. Although there is a significant numerical difference between the COSMO-RS calculated

349 absolute solubility values and experimental ones, the resulting linear correlation has an R^2 score

350 of 0.97, showing the success of our method.

351 Since the actual properties of ILs are not known, determining the limits for high performance

352 should also be done according to the COSMO-RS calculated values. Because of the

353 aforementioned reasons, Park et al. performance data were also included in our analysis. When the

354 50th cycle capacities of this experimental set are considered, a clear trend for high specific capacity

355 can be determined, as shown in Figure 5. The PS solubility should neither be too low or too high

356 due to the sluggish reaction kinetics and sulfur loss from the cathode in extremely low and high

357 PS solubilities, respectively [45]. Therefore, determining the solubility limits is very critical to

358 assess whether an IL will perform well or not as an electrolyte in Li-S batteries. According to

359 Figure 5, the solubility should be between -0.7 to 0.1 mol/mol in log scale.

360 All these discussions show the importance of solubilities and properties in selecting suitable ILs

361 for Li-S battery applications and the success of the COSMO-RS calculations. The COSMO-RS

362 calculations make it possible to calculate the properties of thousands of ILs using ".cosmo" files

363 using special packages like COSMOthermX. On the other hand, once the dataset is constructed,

364 it is valuable to have ML methods that can predict IL properties using the IL descriptors calculated

365 from more conveniently found methods compared to generating ".cosmo" files and using special

366 packages. In addition, hidden relations between ILs and their properties can be identified. In this

367 respect, ML models for predicting PS solubility and the IL properties are developed with a dataset

368 consisting of 36260 ILs and 20 IL descriptors, computed from PM3 semi-empirical calculations,

369 in the search for promising ILs for Li-S battery electrolytes. In addition, using the limit for

370   solubility (-0.7 to 0.1 mol/mol in log scale) for high performance obtained from experimental

371   results, the factors leading to desired properties are obtained.

372   *3.3. Solubility and Property Predictions*

373   The ML models are trained to further understand the trends and heuristics of the solubility

374   correlation with the specified descriptors and to make predictions for new ILs without using the

375   COSMO-RS software. The training of the XGBoost model on randomly chosen combinations of

376   anions and cations performs nearly perfectly with a 5-fold cross-validation R-squared (R^2) score

377   of 0.99. It is essential to note the excessive dependence of the predictions on the anion descriptors

378   and the fact that randomly splitting the data generally means that the same anions are present in

379   both training and validation sets. To avoid such overlap, which may lead to the model

380   "memorizing" the anions instead of learning descriptor correlations, the data was split such that

381   randomly chosen anions from each anion group are only present in the validation set. For example,

382   there are four bis_imide anions including [TFSI]$^-$ in the dataset. Knowing that anions are more

383   dominant over solubility, once the solubility of [TFSI]$^-$ is seen by the model, it will automatically

384   determine the solubility of IL containing [TFSI]$^-$ without paying attention to the cation types in

385   the validation set. However, restricting [TFSI]$^-$ to only train set while including the

386   bis(fluorosulfonyl)imide ([FSI]$^-$) in the validation set, makes the model more robust. This way, the

387   training model learns from similar but non-identical anions. With this restriction, the XGBoost

388   model performance R^2 score dropped to 0.98 with a root-mean-square-error (RMSE) of 1.4, which

389   still indicates an excellent performance on the available data and eliminates the risk of bias due to

390   data overlap. Meanwhile, the RMSE of the test set increases to 3.05, which shows that the model

391   can capture the PS solubilities in an order of magnitude scale.

392   These results are somewhat similar to water solubility prediction results [28], in the sense that the

393   anion properties are more dominant over the solubilities. However, in this work, the effect of cation

394   properties was also found to be important. This inference is observable in Figure 6, showing the

395   descriptor importance obtained from the analysis of XGBoost results. Some cation descriptors,

396   such as the CPK area and dipole, have noticeable importance. While the anions can still roughly

397   estimate the solubilities, the cations refine these results. The top three descriptors having the

398   dominant effect on PS solubility were electronic HUMO-LUMO levels and dipoles of the anions.

399   The HUMO-LUMO levels were also found to have a significant contribution to ML models on the

400   CO_2 solubility [46].

401   Literature review indicates that ML techniques have been repeatedly used to predict different IL

402   properties, including melting point, density, viscosity, ionic conductivity, and even surface tension

403   [47,48]. Nevertheless, the presented property prediction models based on our descriptors were also

404   trained to a high degree of success. Herein, the sampling in the training of these predictions was

405   also anion-based rather than cation-based, as in the solubility classifier. This is because the results

406   vary significantly according to the anion groups in these predictions, and the cation group-based

407   distributions are similar (Figure S3). The anion-based sampling for melting point, viscosity, and

408   conductivity performed well with RMSE score values of 19.4, 0.48, and 1.80, respectively. The

409   descriptor importance plots are given for each property prediction in SM, along with each model's

410   performance metrics. In the prediction of IL properties of the interest, the cationic properties have

411   found more significance in comparison to solubility predictions. The most significant cation

412   descriptors were found to be MW, dipole, and HOMO level. In parallel with our findings, Zheng

413   et al. also found that the alkyl chain lengths of ILs having different imidazolium cations with

18

414   [TFSI]⁻ anion, therefore, MW and dipole, considerably affect the viscosities [49]. Another study

415   found that the anion and cation properties both affect the final IL properties [50].

416   *3.4. Identifying Promising ILs for Li-S Batteries using ARM*

417   Lastly, the descriptor-property and solubility correlations were analyzed using ARM. The ARM

418   method needs both categorical descriptors and outcomes; hence, most descriptors were divided

419   into ten intervals to see the characteristics of the desired ILs, while the HBA and HBD counts were

420   defined as factors. The limits for both solubility and viscosity values were set according to the

421   experimental findings reported in Section 3.2. In this respect, the solubility was categorized into a

422   binary class, and class A includes log(solubility) values between -0.7 and 0.1 mol/mol as

423   determined from experimental results reported in Figure 5. On the other hand, viscosities lower

424   than 100 mPa.s were decided to be class A. Given the viscosity of the standard DOL:DME

425   electrolyte is 1.6 mPa.s [16], 100 mPa.s is a relevant limit for IL viscosity, set to avoid severe Li⁺

426   ion resistance problems. In addition, since the potential applications of Li-S batteries are prone to

427   various climate conditions, the melting point of ILs should be low enough so that the cell remains

428   functional in cold climates. Towards that end, ILs with melting points below 0□C are desired, and

429   fortunately, 86.1% of the dataset satisfies this condition. Finally, the electronic conductivity limit

430   is taken as 2 mS/cm. This rough estimation allows us to exclude ILs with considerable electronic

431   conductivity. With these four criteria in place, the number of potential IL candidates drops to only

432   650 from 36260 data points. The selection process is illustrated in Figure 7.

433   First, the associations between the anion or cation groups or cation-anion pairs and the desired

434   properties were investigated. After refining the generated rules using the support and confidence

435   thresholds, the lift value was used to extract rules with the highest correlation, and the results were

436   sorted according to the lift values. Although the definitions of support, confidence, and lift are

437   already provided in the previous section, an example is provided here to understand the results

438   better. The dataset used in the ARM analysis contained 36260 total data, with only 650 of them

439   having class A solubility and property (viscosity, melting point, and conductivity). In the dataset,

440   there are 1480 ILs with the bis_imide anion group, and only 194 (count value in Table S4) of them

441   are in class A for the four criteria. Hence, the support, confidence, and lift values are calculated as

442   194/36260 = 0.005, 194/650 = 0.298, and (194/650)/(1480/36260) = 7.31, respectively. As seen in

443   Table S4, rules satisfying the confidence and support thresholds with the highest three lift values

444   correlate with the anion group, indicating the trends in anion descriptors are more reliable and

445   determinative than the ones in cation descriptors. These trends can be seen in Figure 8a for cation

446   and anion groups and Figure 8b for cation-anion pairs. The bis_imide group is found to be the

447   most promising one with the highest lift, but borates and "others" groups are also good candidates

448   for both cases. However, although imidazolium and pyridinium groups have lower lifts than

449   piperidinium and pyrrolidinium, their synergistic effects are stronger with the anion groups of

450   borates and bis_imides. Specifically, imidazolium_borate and imidazolium_bis_imide ILs are

451   around ten- and five-times more favorable than other ILs for Li-S batteries, respectively.

452   Imidazolium and [TFSI]⁻ are the most common cation and anion in the Li-S literature, respectively

453   [9]. In addition, it has been reported that combining them improves the Li metal ion morphology

454   and, therefore, increases the cycle life [51]. Given that pyridinium and imidazolium are similar in

455   the completely delocalized aromatic rings [51], pyridinium is the second promising cation group

456   paired with borates. On the other hand, the borate anion group has 12 anions in total, including

457   [BF$_4$]⁻, which is reported to be reactive towards polysulfides [15]. Fortunately, this anion is not

458   present in the promising IL list. Although no additional articles use ILs with borate anions, in some

459 studies, borate anions, specifically bisoxalatoborate [52–54], are used in Li salts, which positively

460 affect the capacities.

461 Now that the promising groups are identified, we discuss the ARM results on a descriptor basis to

462 identify the rules ending up with favorable ILs. To better extract the trends, Figure 9 summarizes

463 the ARM results for the anions more concisely. Upon examining the results, the most pronounced

464 rule concerns the anion HOMO energy value. In the database, low-lying anion HOMO energies in

465 the range of (-7.5):(-6.7) eV result in a nearly five-times more chance of having low PS solubility

466 and viscosity values. This drops significantly to slightly over 2.5-times if the HOMO value dips

467 lower than -7.5 going to -9.9 eV. However, the correlation stands that low HOMO values correlate

468 strongly with low solubility. A similar yet less strong correlation can be seen for the other

469 properties. The other rules indicate that desirable anions for low PS solubility and viscosity are

470 more likely to have moderately low LUMO energy, no HBD sites, relatively high MW but

471 moderate CPK area, and moderate polarizability. Figure 10 shows that cation descriptors have

472 clear trends for low solubility and viscosity; lift values above one are obtained for low values of

473 each descriptor, except HBA, HBD, and LUMO. It is important to emphasize that these trends

474 imply an increased likelihood rather than a confident prediction.

475 Can et al. [28] reported similar results about the dominance of anion descriptors in the ARM

476 analysis when examining the association between water solubility in ILs and molecular

477 descriptors. However, the importance of each descriptor differs significantly from the results

478 reported here. That is to be expected because water and Li_2S_8 have different solvation processes.

479 On the contrary, the solubility of hydrocarbons in imidazolium-based ILs showed a strong

480 correlation with both anion and cation descriptors [29].

481   **4. Conclusions**

482   In this study, a large-scale screening of ILs was conducted in search of potential Li-S battery

483   electrolytes not only with low PS solubility to limit the shuttle effect but also with low viscosity,

484   electronic conductivity, and melting point. The screening was performed using the COSMO-RS

485   model on the COSMObaseIL dataset of over 36000 ion pairs to produce rough estimations of the

486   solubility of Li_2S_8 as the model PS and IL properties. The predictive model developed by the

487   XGBoost algorithm was quite successful, indicating that the solubility can be easily predicted for

488   new ILs. Ten descriptors were analyzed for each anion and cation along with the solubility

489   estimations using association rule mining to find valuable correlations between solubility,

490   property, and IL structural descriptors; solubility limits, which were used as the selection criterion

491   for high performance, were determined experimentally. The results showed that the anion

492   descriptors correlate more strongly with the PS solubility than the cation descriptors. The ML

493   models showed significant overfitting when exposed to the same anions in both training and test

494   sets. This further provides evidence of the strong correlation between the COSMO-RS-predicted

495   PS solubility/property and anions descriptors. The feature importance analysis also showed that

496   anion descriptors are leading in solubility prediction, with anion LUMO and HUMO energies,

497   dipole, and CPK area being the most important. Although calculating the descriptors is comparable

498   in terms of computational cost to the COSMO-RS solubility calculations, the COSMO-RS

499   solubility calculations require the generation of .cosmo files and the use of COSMO

500   thermodynamic software. In contrast, these descriptors are more readily available and easier to set

501   up on a larger scale. Consequently, the ML models developed here are significantly valuable for

502   future studies. According to the results, imidazolium and pyridinium were the most suitable

503   cations, whereas borates and bis_imides were determined to be the anion choices for high-

504   performing Li-S batteries. To conclude, this study not only offers machine-learning models that

22

505 can be easily utilized to identify promising IL electrolytes for Li-S batteries by an original

506 integrated methodology coupling high-throughput COSMO-RS and DFT calculations and

507 experimental characterization but also leads to a better understanding of the structure-solubility-

508 performance relation for IL electrolytes.

509

513 **References**

514 [1] N. Angulakshmi, A.M. Stephan, Efficient Electrolytes for Lithium-Sulfur Batteries, Front.
515   Energy Res. 3 (2015). https://doi.org/10.3389/fenrg.2015.00017.
516 [2] T.B. Reddy, D. Linden, eds., Linden's handbook of batteries, 4th ed, McGraw-Hill, New
517   York, 2011.
518 [3] X. Shen, H. Liu, X.-B. Cheng, C. Yan, J.-Q. Huang, Beyond lithium ion batteries: Higher
519   energy density battery systems based on lithium metal anodes, Energy Storage Materials 12
520   (2018) 161−175. https://doi.org/10.1016/j.ensm.2017.12.002.
521 [4] H. Zhao, N. Deng, J. Yan, W. Kang, J. Ju, Y. Ruan, X. Wang, X. Zhuang, Q. Li, B. Cheng,
522   A review on anode for lithium-sulfur batteries: Progress and prospects, Chemical
523   Engineering Journal 347 (2018) 343−365. https://doi.org/10.1016/j.cej.2018.04.112.
524 [5] S. Evers, L.F. Nazar, New Approaches for High Energy Density Lithium−Sulfur Battery
525   Cathodes, Acc. Chem. Res. 46 (2013) 1135−1143. https://doi.org/10.1021/ar3001348.
526 [6] L. Wang, Y. Ye, N. Chen, Y. Huang, L. Li, F. Wu, R. Chen, Development and Challenges of
527   Functional Electrolytes for High-Performance Lithium-Sulfur Batteries, Adv. Funct. Mater.
528   28 (2018) 1800919. https://doi.org/10.1002/adfm.201800919.
529 [7] R. Steudel, T. Chivers, The role of polysulfide dianions and radical anions in the chemical,
530   physical and biological sciences, including sulfur-based batteries, Chem. Soc. Rev. 48 (2019)
531   3279−3319. https://doi.org/10.1039/C8CS00826D.
532 [8] M. Vijayakumar, N. Govind, E. Walter, S.D. Burton, A. Shukla, A. Devaraj, J. Xiao, J. Liu,
533   C. Wang, A. Karim, S. Thevuthasan, Molecular structure and stability of dissolved lithium
534   polysulfide species, Phys. Chem. Chem. Phys. 16 (2014) 10923−10932.
535   https://doi.org/10.1039/C4CP00889H.
536 [9] N. Angulakshmi, R.B. Dhanalakshmi, S. Sathya, J. Ahn, A.M. Stephan, Understanding the
537   Electrolytes of Lithium−Sulfur Batteries, Batteries & Supercaps 4 (2021) 1064−1095.
538   https://doi.org/10.1002/batt.202000273.
539 [10] J. Liao, Z. Ye, Nontrivial Effects of "Trivial" Parameters on the Performance of Lithium−
540   Sulfur Batteries, Batteries 4 (2018) 22. https://doi.org/10.3390/batteries4020022.

541 [11] H. Lu, Z. Chen, H. Du, K. Zhang, J. Wang, Z. Hou, J. Fang, The enhanced performance of
542   lithium sulfur battery with ionic liquid-based electrolyte mixed with fluorinated ether, Ionics
543   25 (2019) 2685–2691. https://doi.org/10.1007/s11581-018-2814-x.
544 [12] H. Lu, Y. Zhu, B. Zheng, H. Du, X. Zheng, C. Liu, Y. Yuan, J. Fang, K. Zhang, A hybrid
545   ionic liquid-based electrolyte for high-performance lithium–sulfur batteries, New J. Chem.
546   44 (2020) 361–368. https://doi.org/10.1039/C9NJ03790J.
547 [13] S. Suriyakumar, M. Kathiresan, A.M. Stephan, Charge–Discharge and Interfacial Properties
548   of Ionic Liquid-Added Hybrid Electrolytes for Lithium–Sulfur Batteries, ACS Omega 4
549   (2019) 3894–3903. https://doi.org/10.1021/acsomega.8b03544.
550 [14] K. Ueno, J.-W. Park, A. Yamazaki, T. Mandai, N. Tachikawa, K. Dokko, M. Watanabe,
551   Anionic Effects on Solvate Ionic Liquid Electrolytes in Rechargeable Lithium–Sulfur
552   Batteries, J. Phys. Chem. C 117 (2013) 20509–20516. https://doi.org/10.1021/jp407158y.
553 [15] J.-W. Park, K. Ueno, N. Tachikawa, K. Dokko, M. Watanabe, Ionic Liquid Electrolytes for
554   Lithium–Sulfur Batteries, J. Phys. Chem. C 117 (2013) 20531–20541.
555   https://doi.org/10.1021/jp408037e.
556 [16] M. Yanagi, K. Ueno, A. Ando, S. Li, Y. Matsumae, J. Liu, K. Dokko, M. Watanabe, Effects
557   of Polysulfide Solubility and Li Ion Transport on Performance of Li–S Batteries Using
558   Sparingly Solvating Electrolytes, J. Electrochem. Soc. 167 (2020) 070531.
559   https://doi.org/10.1149/1945-7111/ab7a81.
560 [17] A. Chen, X. Zhang, Z. Zhou, Machine learning: Accelerating materials development for
561   energy storage and conversion, InfoMat 2 (2020) 553–576.
562   https://doi.org/10.1002/inf2.12094.
563 [18] A. Klamt, Conductor-like Screening Model for Real Solvents: A New Approach to the
564   Quantitative Calculation of Solvation Phenomena, J. Phys. Chem. 99 (1995) 2224–2235.
565   https://doi.org/10.1021/j100007a062.
566 [19] A. Klamt, F. Eckert, W. Arlt, COSMO-RS: An Alternative to Simulation for Calculating
567   Thermodynamic Properties of Liquid Mixtures, Annu. Rev. Chem. Biomol. Eng. 1 (2010)
568   101–122. https://doi.org/10.1146/annurev-chembioeng-073009-100903.
569 [20] S. Boobier, D.R.J. Hose, A.J. Blacker, B.N. Nguyen, Machine learning with physicochemical
570   relationships: solubility prediction in organic solvents and water, Nat Commun 11 (2020)
571   5753. https://doi.org/10.1038/s41467-020-19594-z.
572 [21] R. Franke, B. Hannebauer, On the influence of basis sets and quantum chemical methods on
573   the prediction accuracy of COSMO-RS, Phys. Chem. Chem. Phys. 13 (2011) 21344.
574   https://doi.org/10.1039/c1cp22317h.
575 [22] J. Reinisch, M. Diedenhofen, R. Wilcken, A. Udvarhelyi, A. Glöß, Benchmarking Different
576   QM Levels for Usage with COSMO-RS, J. Chem. Inf. Model. 59 (2019) 4806–4813.
577   https://doi.org/10.1021/acs.jcim.9b00659.
578 [23] D.H. Barrett, A. Haruna, Artificial intelligence and machine learning for targeted energy
579   storage solutions, Current Opinion in Electrochemistry 21 (2020) 160–166.
580   https://doi.org/10.1016/j.coelec.2020.02.002.
581 [24] S.C. Balchandani, R. Singh, B. Mandal, Experimental and COSMO-RS analysis of CO2
582   solubility in novel aqueous blends of 1-butyl-3-methyl-imidazolium tetrafluoroborate
583   activated by 2-aminoethyl piperazine and bis(3-aminopropyl) amine for post combustion
584   carbon capture, Journal of Environmental Chemical Engineering 11 (2023) 109099.
585   https://doi.org/10.1016/j.jece.2022.109099.

586 [25] Y. Cao, Z. Wu, Y. Zhang, Y. Liu, H. Wang, Screening of alternative solvent ionic liquids for
587   artemisinin: COSMO-RS prediction and experimental verification, Journal of Molecular
588   Liquids 338 (2021) 116778. https://doi.org/10.1016/j.molliq.2021.116778.

589 [26] R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Bostani, A. Hemmati-
590   Sarapardeh, A. Mohaddespour, Solubility of gaseous hydrocarbons in ionic liquids using
591   equations of state and machine learning approaches, Sci Rep 12 (2022) 14276.
592   https://doi.org/10.1038/s41598-022-17983-6.

593 [27] V. Venkatraman, B.K. Alsberg, Predicting CO_2 capture of ionic liquids using machine
594   learning, Journal of CO_2 Utilization 21 (2017) 162–168.
595   https://doi.org/10.1016/j.jcou.2017.06.012.

596 [28] E. Can, A. Jalal, I.G. Zirhlioglu, A. Uzun, R. Yildirim, Predicting water solubility in ionic
597   liquids using machine learning towards design of hydro-philic/phobic ionic liquids, Journal
598   of Molecular Liquids 332 (2021) 115848. https://doi.org/10.1016/j.molliq.2021.115848.

599 [29] A. Jalal, E. Can, S. Keskin, R. Yildirim, A. Uzun, Selection rules for estimating the solubility
600   of C4-hydrocarbons in imidazolium ionic liquids determined by machine-learning tools,
601   Journal of Molecular Liquids 284 (2019) 511–521.
602   https://doi.org/10.1016/j.molliq.2019.03.182.

603 [30] S. Jeschke, P. Johansson, Predicting the Solubility of Sulfur: A COSMO- RS- Based
604   Approach to Investigate Electrolytes for Li–S Batteries, Chemistry A European J 23 (2017)
605   9130–9136. https://doi.org/10.1002/chem.201701011.

606 [31] S. Jeschke, P. Johansson, Supervised Machine Learning- Based Classification of Li−S
607   Battery Electrolytes, Batteries & Supercaps 4 (2021) 1156–1162.
608   https://doi.org/10.1002/batt.202100031.

609 [32] COSMOThermX Version C30_1601, COSMOlogic GmbH & Co. kG (2005).
610   http://www.solventextract.org/Organization/COSMOlogic-GmbH-Co-kG (accessed
611   November 17, 2023).

612 [33] M.F. Sgroi, D. Pullini, A.I. Pruna, Lithium Polysulfide Interaction with Group III Atoms-
613   Doped Graphene: A Computational Insight, Batteries 6 (2020) 46.
614   https://doi.org/10.3390/batteries6030046.

615 [34] F. Neese, F. Wennmohs, U. Becker, C. Riplinger, The ORCA quantum chemistry program
616   package, The Journal of Chemical Physics 152 (2020) 224108.
617   https://doi.org/10.1063/5.0004608.

618 [35] C. Steffen, K. Thomas, U. Huniar, A. Hellweg, O. Rubner, A. Schroer, TmoleX—A graphical
619   user interface for TURBOMOLE, J Comput Chem 31 (2010) 2967–2970.
620   https://doi.org/10.1002/jcc.21576.

621 [36] D.L.J. Alexander, A. Tropsha, D.A. Winkler, Beware of R^2: Simple, Unambiguous
622   Assessment of the Prediction Accuracy of QSAR and QSPR Models, J. Chem. Inf. Model.
623   55 (2015) 1316–1322. https://doi.org/10.1021/acs.jcim.5b00206.

624 [37] P. Yu, D.J. Wild, Fast rule-based bioactivity prediction using associative classification
625   mining, J Cheminform 4 (2012) 29. https://doi.org/10.1186/1758-2946-4-29.

626 [38] Posit team, RStudio: Integrated Development Environment for R, (2023).
627   http://www.posit.co/.

628 [39] T. Mu, J. Rarey, J. Gmehling, Performance of COSMO-RS with Sigma Profiles from
629   Different Model Chemistries, Ind. Eng. Chem. Res. 46 (2007) 6612–6629.
630   https://doi.org/10.1021/ie0702126.

25

631 [40] K. Paduszyński, An overview of the performance of the COSMO-RS approach in predicting
632   the activity coefficients of molecular solutes in ionic liquids and derived properties at infinite
633   dilution, Phys. Chem. Chem. Phys. 19 (2017) 11835–11850.
634   https://doi.org/10.1039/C7CP00226B.
635 [41] K. Dong, S. Zhang, J. Wang, Understanding the hydrogen bonds in ionic liquids and their
636   roles in properties and reactions, Chem. Commun. 52 (2016) 6744–6764.
637   https://doi.org/10.1039/C5CC10120D.
638 [42] J. Tan, J. Matz, P. Dong, M. Ye, J. Shen, Appreciating the role of polysulfides in lithium-
639   sulfur batteries and regulation strategies by electrolytes engineering, Energy Storage
640   Materials 42 (2021) 645–678. https://doi.org/10.1016/j.ensm.2021.08.012.
641 [43] M. Karelson, V.S. Lobanov, A.R. Katritzky, Quantum-Chemical Descriptors in
642   QSAR/QSPR Studies, Chem. Rev. 96 (1996) 1027–1044. https://doi.org/10.1021/cr950202r.
643 [44] L. Wang, J. Ding, L. Pan, D. Cao, H. Jiang, X. Ding, Quantum chemical descriptors in
644   quantitative structure–activity relationship models and their applications, Chemometrics and
645   Intelligent Laboratory Systems 217 (2021) 104384.
646   https://doi.org/10.1016/j.chemolab.2021.104384.
647 [45] Z. Wang, Y. Li, H. Ji, J. Zhou, T. Qian, C. Yan, Unity of Opposites between Soluble and
648   Insoluble Lithium Polysulfides in Lithium–Sulfur Batteries, Advanced Materials 34 (2022)
649   2203699. https://doi.org/10.1002/adma.202203699.
650 [46] M. Aghaie, S. Zendehboudi, Estimation of CO2 solubility in ionic liquids using connectionist
651   tools based on thermodynamic and structural characteristics, Fuel 279 (2020) 117984.
652   https://doi.org/10.1016/j.fuel.2020.117984.
653 [47] P. Dhakal, J.K. Shah, Developing machine learning models for ionic conductivity of
654   imidazolium-based ionic liquids, Fluid Phase Equilibria 549 (2021) 113208.
655   https://doi.org/10.1016/j.fluid.2021.113208.
656 [48] S. Koutsoukos, F. Philippi, F. Malaret, T. Welton, A review on machine learning algorithms
657   for the ionic liquid chemical space, Chem. Sci. 12 (2021) 6820–6843.
658   https://doi.org/10.1039/D1SC01000J.
659 [49] W. Zheng, A. Mohammed, L.G. Hines, D. Xiao, O.J. Martinez, R.A. Bartsch, S.L. Simon, O.
660   Russina, A. Triolo, E.L. Quitevis, Effect of Cation Symmetry on the Morphology and
661   Physicochemical Properties of Imidazolium Ionic Liquids, J. Phys. Chem. B 115 (2011)
662   6572–6584. https://doi.org/10.1021/jp1115614.
663 [50] A.B. Pereiro, J.M.M. Araújo, S. Martinho, F. Alves, S. Nunes, A. Matias, C.M.M. Duarte,
664   L.P.N. Rebelo, I.M. Marrucho, Fluorinated Ionic Liquids: Properties and Applications, ACS
665   Sustainable Chem. Eng. 1 (2013) 427–439. https://doi.org/10.1021/sc300163n.
666 [51] N. Nakamura, S. Ahn, T. Momma, T. Osaka, Future potential for lithium-sulfur batteries,
667   Journal of Power Sources 558 (2023) 232566.
668   https://doi.org/10.1016/j.jpowsour.2022.232566.
669 [52] N. Azimi, Z. Xue, L. Hu, C. Takoudis, S. Zhang, Z. Zhang, Additive Effect on the
670   Electrochemical Performance of Lithium–Sulfur Battery, Electrochimica Acta 154 (2015)
671   205–210. https://doi.org/10.1016/j.electacta.2014.12.041.
672 [53] S. Xiong, X. Kai, X. Hong, Y. Diao, Effect of LiBOB as additive on electrochemical
673   properties of lithium–sulfur batteries, Ionics 18 (2012) 249–254.
674   https://doi.org/10.1007/s11581-011-0628-1.
675 [54] H. Xiang, P. Shi, P. Bhattacharya, X. Chen, D. Mei, M.E. Bowden, J. Zheng, J.-G. Zhang,
676   W. Xu, Enhanced charging capability of lithium metal batteries based on lithium

677 bis(trifluoromethanesulfonyl)imide-lithium bis(oxalato)borate dual-salt electrolytes, Journal
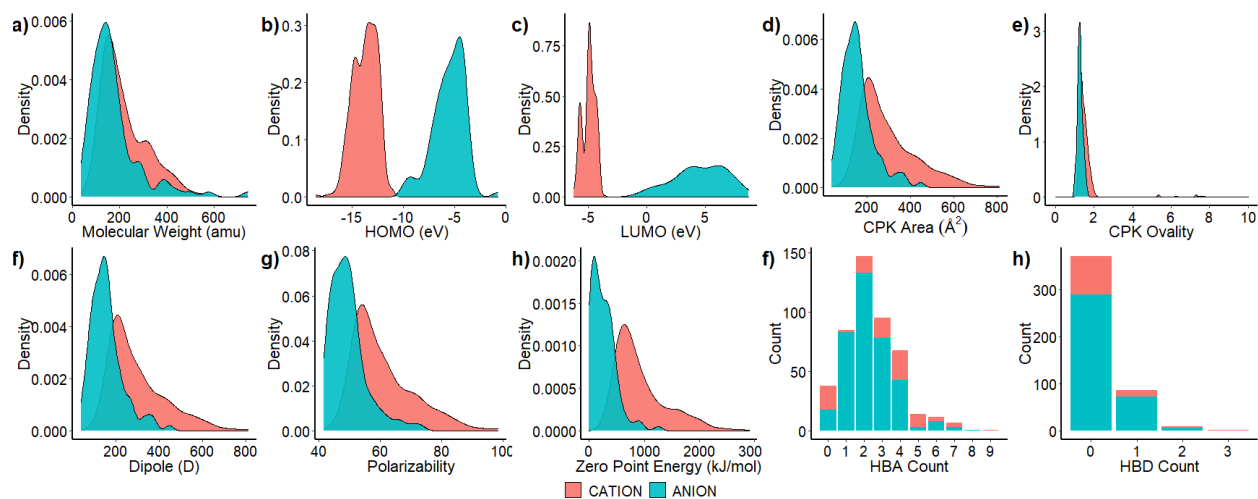678 of Power Sources 318 (2016) 170–177. https://doi.org/10.1016/j.jpowsour.2016.04.017.
679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694
695

696
697 **Figure 1.** The distribution of descriptors for cations and anions in the dataset
698
699
700
701
702
703
704
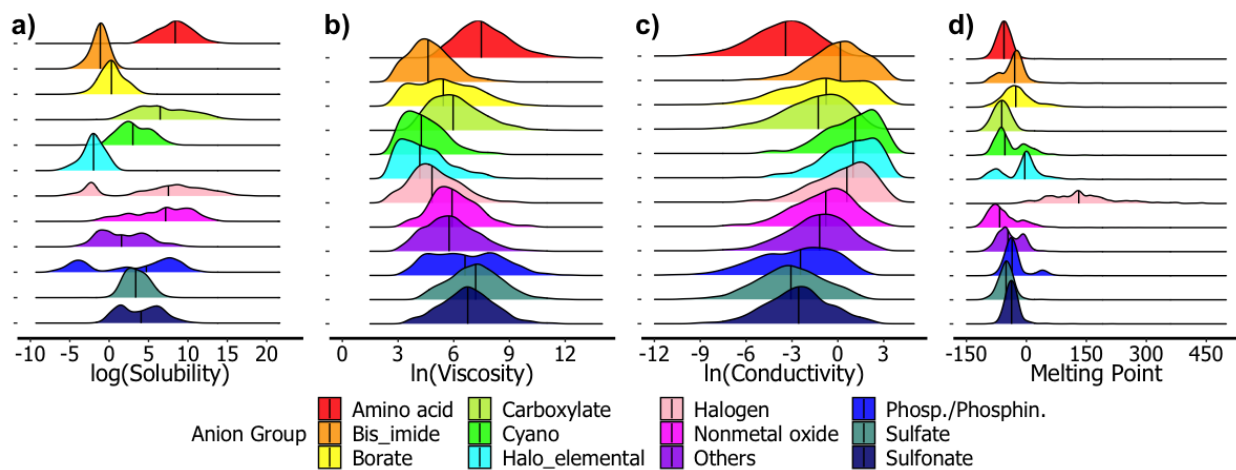
705
706 **Figure 2.** The distribution of solubility (a), ln(viscosity(mPa.s)) (b), ln(conductivity(mS/cm))
707 (c), and melting point(°C) (d), depending on the anion group.
708
709

28
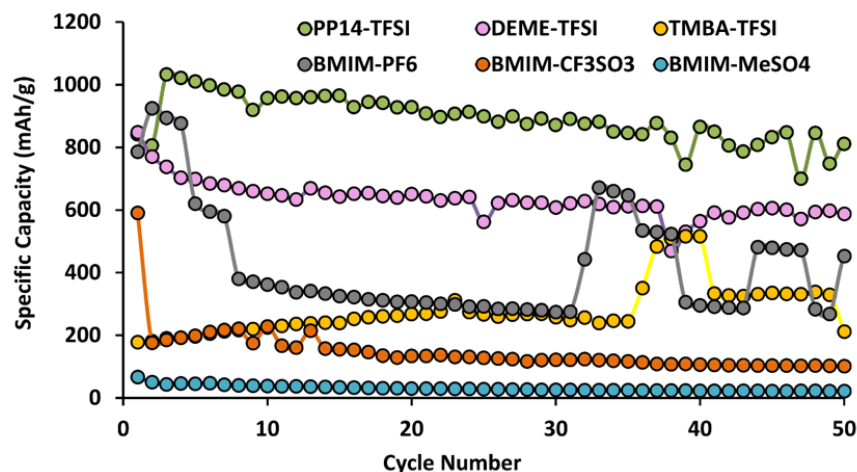
710
711 **Figure 3.** The cycling performance of six ILs tested at 0.1C

712
713
714
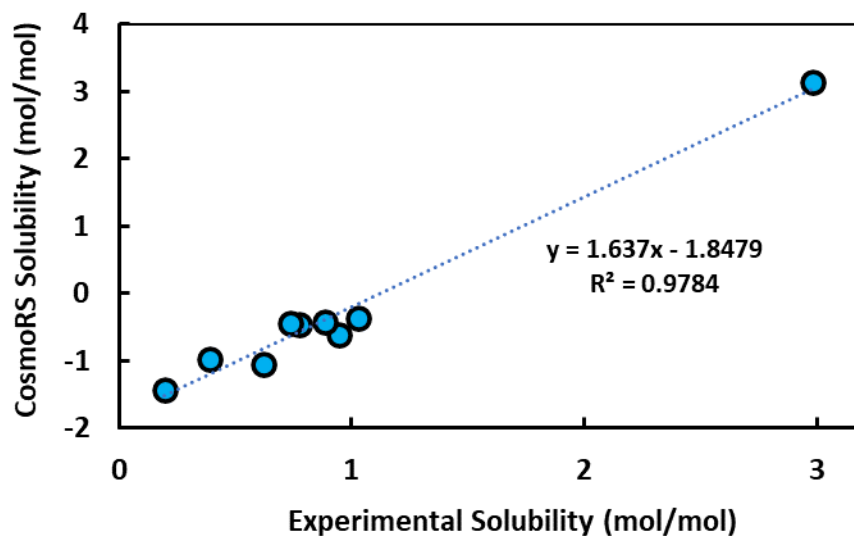
715
716
717 **Figure 4.** Experimental polysulfide solubility plotted against the predicted solubility values by
718 COSMO-RS *in log scale*. The dotted line is the best-fit line. The experimental data was obtained
719 from [15].

720
721
722
723
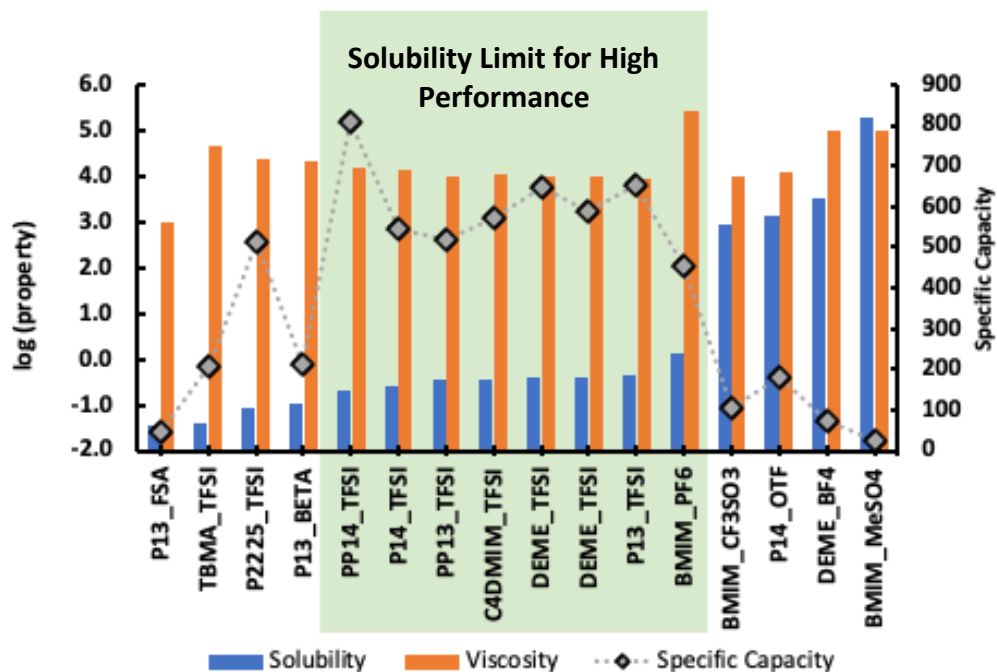
Figure 5. Experimental 50th cycle capacities (mAh/g) matched with the predicted COSMO-RS solubility and viscosity results. The experimental data was obtained from our results in Figure 3 and from [15].
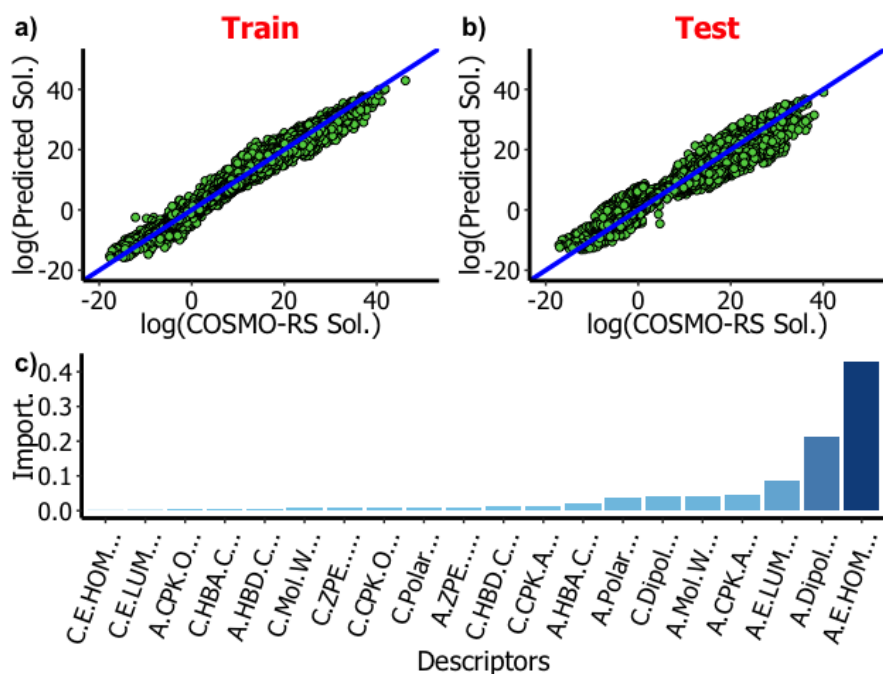
Metrics	Train	Test
RMSE	1.40	3.05
R-squared	0.98	0.92

Figure 6. The XGBoost prediction results for ln(COSMO-RS Solublity (mol/mol)) for the train (a), test (b) sets and relative importance (c) of the descriptors in the determinative power of the model
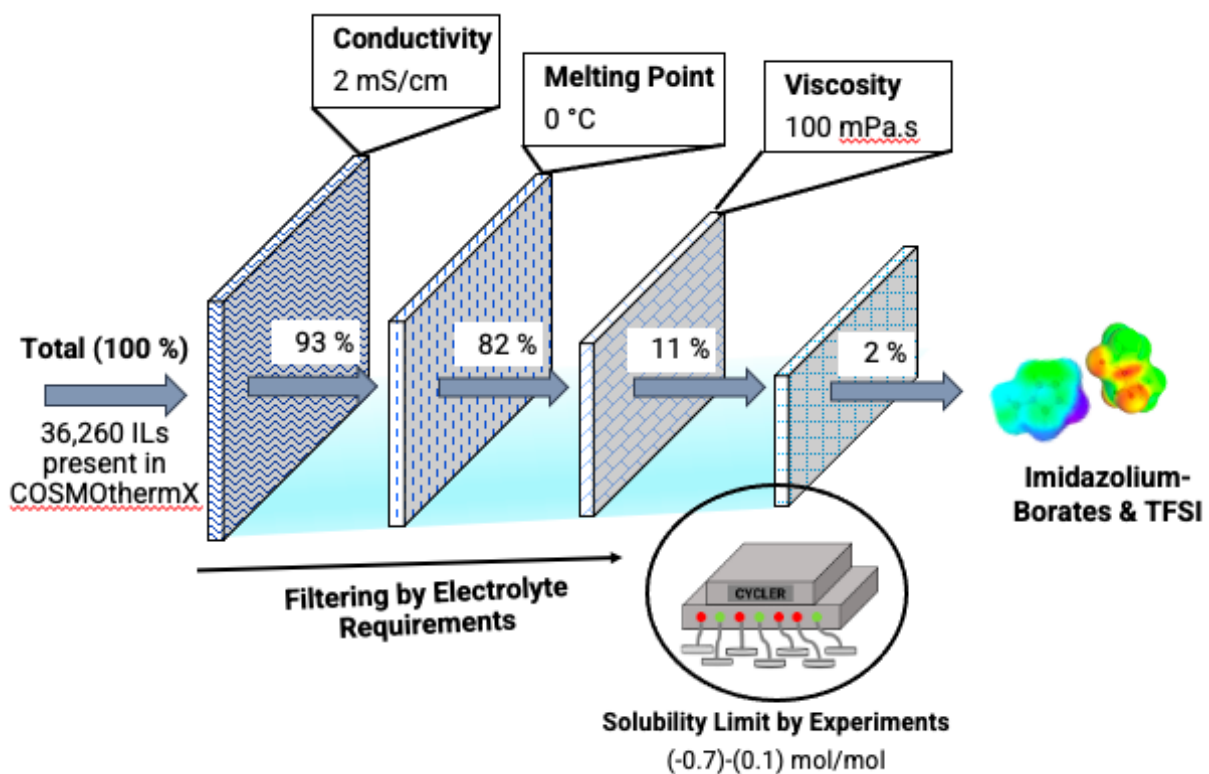
744
745 **Figure 7.** The determination of suitable ILs with the help of experimental results
746
747
748
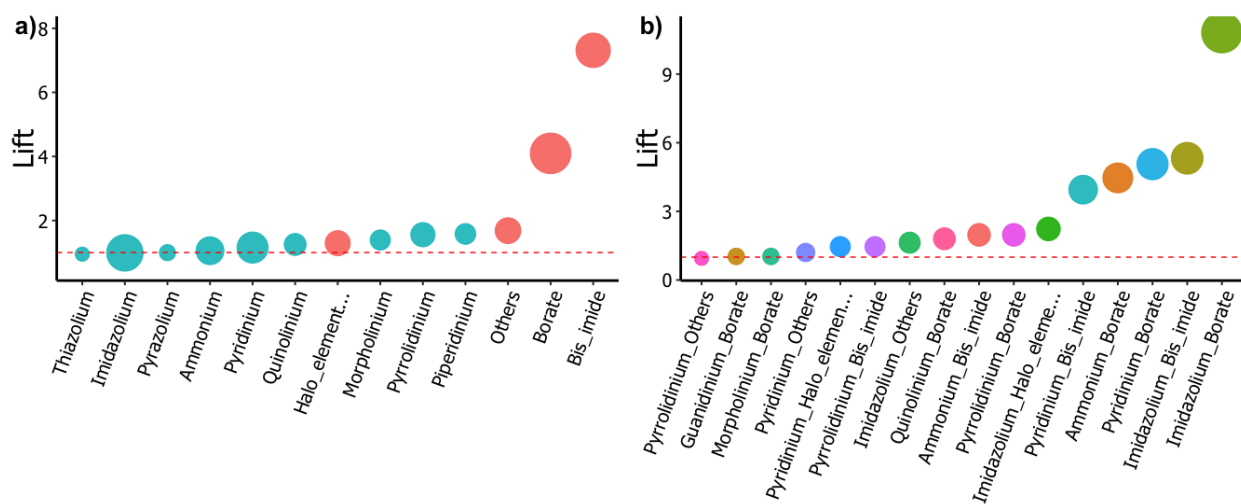
749
750 **Figure 8.** ARM results for anion and cation groups independently (a), and anion-cation pairs (b)
751 for low solubility and viscosity
752
753
754
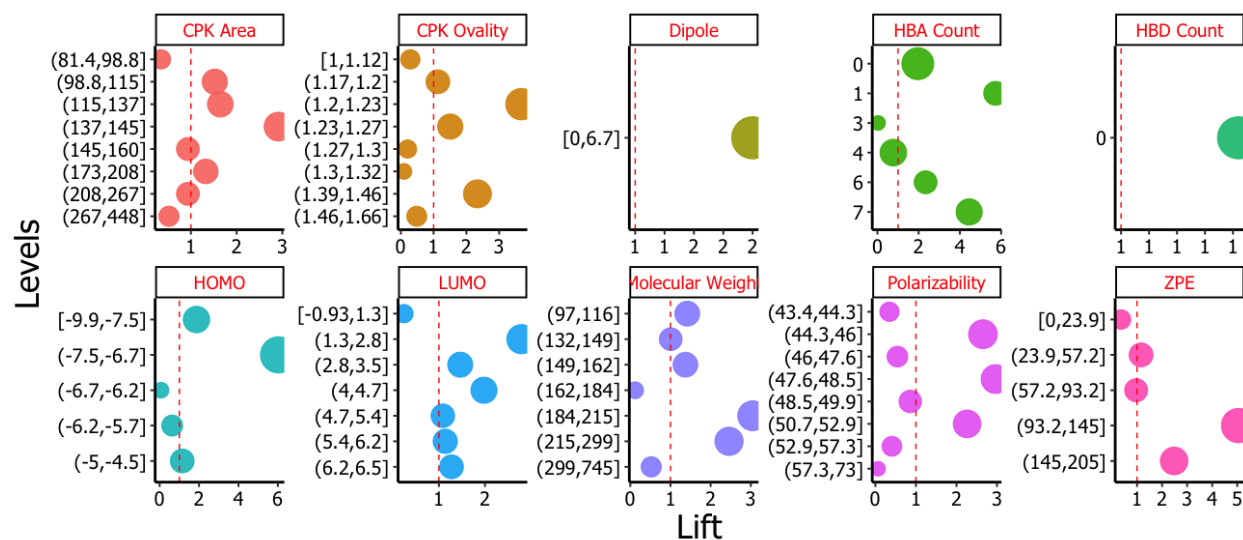
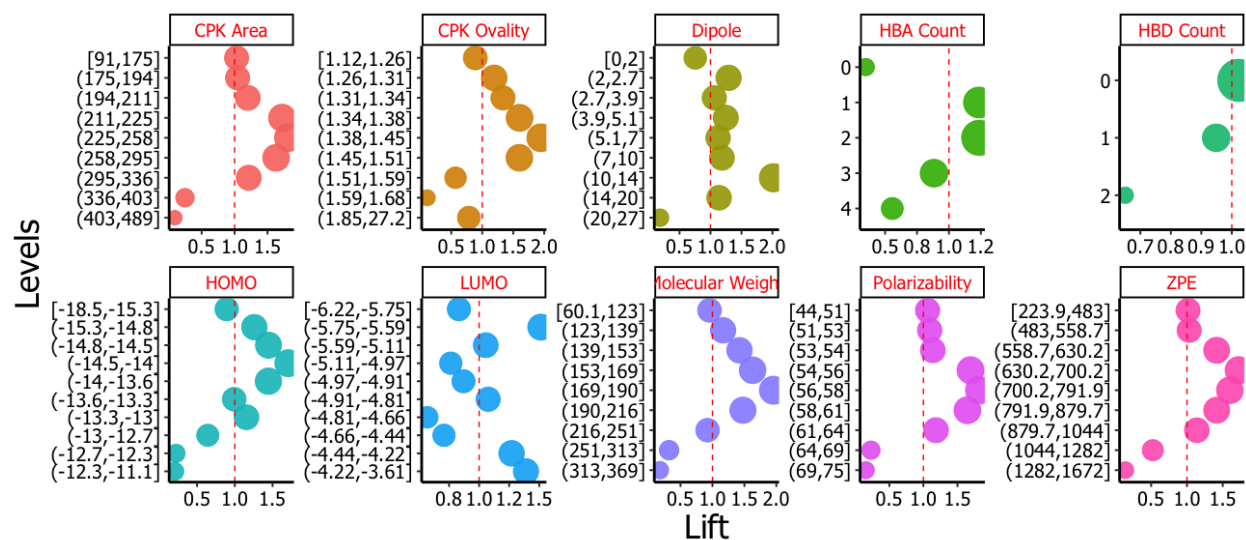Figure 9. ARM results for anion descriptors; each point size correlates with support.

Figure 10. ARM results for cation descriptors; each point size correlates with support.

774 **Table 1.** The list of cation and anion groups present in the dataset

Cation Group	Cation Count	Anion Group	Anion Count
Ammonium	49	Amino acid	16
Choline	1	Bis_imide	4
Guanidinium	20	Borate	12
Imidazolium	134	Carboxylate	16
Morpholinium	9	Cyano	3
Others	4	Halo_elemental_complexes	7
Phosphonium	28	Halogen	4
Piperidinium	9	Nonmetal oxide	5
Pyrazolium	4	Others	6
Pyridinium	67	Phosphate/ Phosphinate*	9
Pyrrolidinium	19	Sulfate	9
Quinolinium	15	Sulfonate	7
Sulfonium	5		
Thiazolium	3		
Uronium	3		
Total	370	Total	98
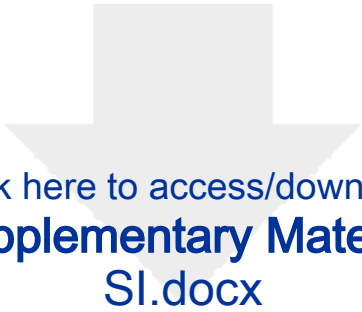
775 *Abbreviated as Phosp./Phosphin.

776
777
778 **Table 2.** Experimentally tested six ionic liquids*
779
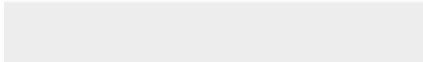
Cation	Anion	COSMO-RS PS Solubility (mol/mol)	COSMO-RS Viscosity (mPa.s)
PP14	TFSI	0.20	65
DEME	TFSI	0.38	54
TBMA	TFSI	0.04	107
BMIM	PF6	1.29	223
BMIM	CF3SO3	847	56
BMIM	MeSO4	182635	146

780
781 *BMIM: 1-butyl-3-methyl-imidazolium, CF3SO3: trifluoromethane-sulfonate, DEME: N,N-diethyl-N-methyl-N-(2-
782 methoxyethyl)ammonium, MeSO4: methylsulfate, PF6: hexafluorophosphate, PP14: 1-butyl-1-methylpiperidinium,
783 TBMA: tributylmethylammonium, TFSI: bis(trifluoromethane)sulfonimide
784 **These cells include 25 vol. % of the selected IL in the standard electrolyte of 1 M LiTFSI and 0.1 M LiNO$_3$ in
785 DOL:DME (1:1 vol/vol)
786
787

Click here to access/download
Supplementary Material
SI.docx

Supplementary Material

Click here to access/download
Supplementary Material
complete_dataset.xlsx

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: