



**HAL**  
open science

## Chemoinformatics-Driven Design of New Physical Solvents for Selective CO<sub>2</sub> Absorption

Alexey Orlov, Daryna Yu. Demenko, Charles Bignaud, Alain Valtz, Gilles Marcou, Dragos Horvath, Christophe Coquelet, Alexandre Varnek, Frédérick de Meyer

► **To cite this version:**

Alexey Orlov, Daryna Yu. Demenko, Charles Bignaud, Alain Valtz, Gilles Marcou, et al.. Chemoinformatics-Driven Design of New Physical Solvents for Selective CO<sub>2</sub> Absorption. Environmental Science and Technology, 2021, 55 (22), pp.15542-15553. 10.1021/acs.est.1c04092. hal-03435228

**HAL Id: hal-03435228**

**<https://hal.science/hal-03435228>**

Submitted on 24 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Chemoinformatics-driven design of new physical solvents for selective CO<sub>2</sub> absorption

*Alexey A. Orlov<sup>§</sup>, Daryna Yu. Demenko<sup>§</sup>, Charles Bignaud<sup>†</sup>, Alain Valtz<sup>‡</sup>, Gilles Marcou<sup>§</sup>, Dragos Horvath<sup>§</sup>, Christophe Coquelet<sup>‡</sup>, Alexandre Varnek<sup>§\*</sup>, Frédérick de Meyer<sup>†‡\*</sup>.*

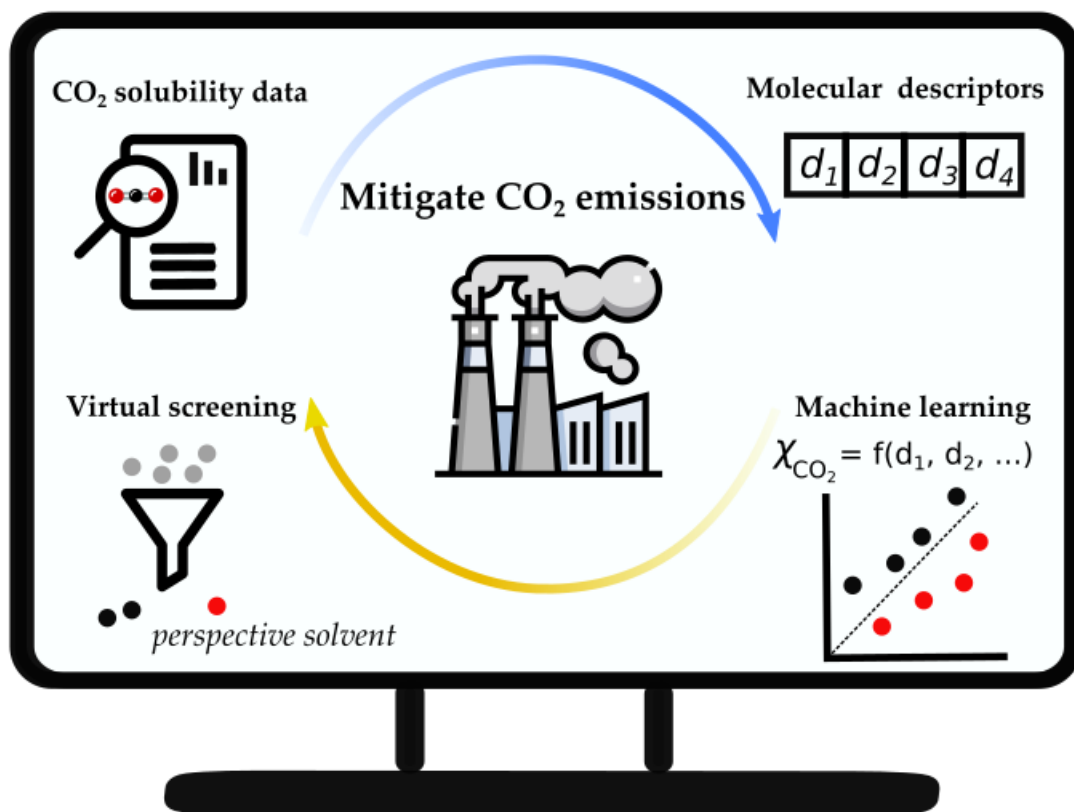
<sup>§</sup> Laboratory of Chemoinformatics, Faculty of Chemistry, University of Strasbourg, Strasbourg, 67081 France

<sup>†</sup> TotalEnergies S.E., Exploration Production, Development and Support to Operations, Liquefied Natural Gas – Acid Gas Entity, CCUS R&D Program, Paris, 92078 France

<sup>‡</sup> MINES ParisTech, PSL University, Centre de thermodynamique des procédés (CTP), 35 rue St Honoré, 77300 Fontainebleau, France

*Keywords: gas solubility, industrial gases, carbon dioxide, methane, nitrogen, hydrogen, carbon monoxide, chemoinformatics, machine learning*

14 TOC GRAPHIC



15

## 16 **ABSTRACT**

17 The removal of CO<sub>2</sub> from gases is an important industrial process in the transition to a low-  
18 carbon economy. The use of selective physical (co-)solvents is especially perspective in cases  
19 when the amount of CO<sub>2</sub> is large as it enables one to lower the energy requirements for solvent  
20 regeneration. However, only a few physical solvents have found industrial application and the  
21 design of new ones can pave the way to more efficient gas treatment techniques. Experimental  
22 screening of gas solubility is a labor-intensive process, and solubility modeling is a viable  
23 strategy to reduce the number of solvents subject to experimental measurements. In this paper, a  
24 chemoinformatics-based modeling workflow was applied to build a predictive model for the  
25 solubility of CO<sub>2</sub> and four other industrially important gases (CO, CH<sub>4</sub>, H<sub>2</sub>, N<sub>2</sub>). A dataset  
26 containing solubilities of gases in 280 solvents was collected from literature sources and  
27 supplemented with the new data for six solvents measured in the present study. A modeling  
28 workflow based on the usage of several state-of-the-art machine learning algorithms was applied  
29 to establish quantitative structure-solubility relationships. The best models were used to perform  
30 virtual screening of the industrially produced chemicals. It enabled the identification of  
31 compounds with high predicted CO<sub>2</sub> solubility and selectivity towards the other gases. The  
32 prediction for one of the compounds – 4-Methylmorpholine was confirmed experimentally.

## 33 **SYNOPSIS STATEMENT**

34 Developing better solvents for selective CO<sub>2</sub> capture is crucial for reaching net-zero emissions  
35 targets.

## 36 INTRODUCTION

37 Global warming due to increasing levels of greenhouse gases CO<sub>2</sub> and CH<sub>4</sub> in the atmosphere  
38 has become a major public issue. Several companies and countries have announced ambitious  
39 plans to reach net zero CO<sub>2</sub> emissions by 2050. According to the International Energy Agency,  
40 Carbon Capture, Utilization or Storage (CCUS) will likely play an important role in achieving  
41 this goal.<sup>1</sup> Numerous materials for CO<sub>2</sub> capturing from gases were suggested, including chemical  
42 and physical solvents, zeolites, metal oxides, metal-organic frameworks, and membranes.<sup>2,3</sup> The  
43 applicability of a certain technology in each case depends on many factors, including the  
44 concentrations of CO<sub>2</sub> and of other components in the gas, the pressure of the gas feed, the  
45 temperature, etc. In cases, wherein the partial pressure of CO<sub>2</sub> in a gas mixture is sufficiently  
46 large, physical solvents represent a perspective alternative to the conventionally used aqueous  
47 amines mixtures because of the lower energy requirement. Indeed, a large part of a physical  
48 solvent can be regenerated by pressure swing and air stripping, while the regeneration of  
49 chemical solvents requires heating and steam stripping<sup>3,4</sup> Pre-combustion CO<sub>2</sub> capture is a key  
50 example of a case with a CO<sub>2</sub> partial pressure sufficiently high to use physical solvents. In a pre-  
51 combustion process the feed (e.g. coal, natural gas, biomass, etc.) is converted into syngas (H<sub>2</sub>  
52 and CO) via gasification, steam reforming, auto thermal reforming or partial oxidation and  
53 subsequently the CO is further converted into CO<sub>2</sub> and H<sub>2</sub> via the water gas shift reaction.  
54 Typical CO<sub>2</sub> concentrations are in the range of 15 to 60 mol% for a total pressure of 2 to 7 MPa,  
55 thus, the CO<sub>2</sub> can be captured with a physical solvent<sup>5</sup>. The other components are mainly H<sub>2</sub>, but  
56 also CO, N<sub>2</sub>, CH<sub>4</sub>, H<sub>2</sub>O (saturation). The composition strongly depends on the feedstock and on  
57 the process. At the moment, only a limited number of physical solvents such as methanol  
58 (Rectisol® process), propylene carbonate (Fluor® process), N-acetyl and N-formyl morpholines

59 (Morphysorb® process), 1-methylpyrrolidin-2-one (Purisol® process), polyethylene glycol  
60 ethers (Selexol® process), found application in the industrial CO<sub>2</sub> capture processes. Physical  
61 solvents are also often added to chemical solvents (so-called hybrid solvents), for example, to  
62 increase the selectivity of absorption towards a specific gas component, to lower the regeneration  
63 energy, etc. Examples are sulfolane (Sulfinol® process), thiodiglycol (Hysweet ® process), etc.  
64 The search for new physical (co-)solvents is thus an important task.<sup>6</sup>

65 New suitable solvents should satisfy many criteria, among which are a decent capacity to  
66 absorb CO<sub>2</sub>, a competitive price, a low volatility (to avoid solvent losses), a low viscosity, etc.  
67 Another very important criterium is the selectivity towards CO<sub>2</sub> which should be high enough to  
68 obtain a CO<sub>2</sub> stream of acceptable purity for re-utilization or storage, and to avoid losses of  
69 valuable chemicals like CH<sub>4</sub>. Depending on the source of the CO<sub>2</sub>, the selectivity criteria are  
70 different. For the removal of CO<sub>2</sub> from natural gas the co-absorption of mainly CH<sub>4</sub>, but also of  
71 N<sub>2</sub> and H<sub>2</sub>O and eventually H<sub>2</sub>S should be limited. For a steam methane reformer (SMR), which  
72 in a near future is likely to play a key role in the massive production of blue hydrogen from  
73 natural gas, the absorption of CO<sub>2</sub> with a high selectivity towards H<sub>2</sub> is important<sup>7,8</sup>, but also  
74 towards CO, H<sub>2</sub>O, N<sub>2</sub> and CH<sub>4</sub><sup>7,8</sup>. The CO<sub>2</sub> present in flue gas from boilers should be removed  
75 with a high selectivity towards N<sub>2</sub>, but the co-absorption of water and SO<sub>x</sub>, NO<sub>x</sub> should be low  
76 too.

77 Experimental screening of gas solubility is a time and labor-intensive process, and solubility  
78 modeling is a viable strategy to reduce the costs of the required experiments. There were  
79 numerous approaches suggested for modeling gas solubility in pure physical solvents. In the  
80 work by Pirig et al.<sup>9</sup> a five-parameter linear equation based on the experimentally measured  
81 properties was used to model the solubility in 58 solvents. In the work by Li et al.<sup>10</sup>, artificial

82 neural networks were used to model CO<sub>2</sub> mole fraction solubility in 11 solvents (alcohols, ethers,  
83 ketones) at different temperatures and pressures. Several structural features (number of C-H, O-  
84 H, C-O, C=O bonds, number of rotatable bonds, etc.) and physical properties of the compounds  
85 (density, dipole moment, etc.) were used as descriptors. It enabled achieving high precision of  
86 predictions for certain types of physical solvents under varying experimental conditions.  
87 Nonetheless, the major disadvantage of the modeling approaches based on experimentally  
88 measured parameters – the limited number of compounds for which the parameters are available,  
89 complicates their usage for large-scale virtual screening of solvent candidates.

90 Alternative strategies for modeling CO<sub>2</sub> solubility, that require less preliminary knowledge of  
91 the experimental properties, were also suggested. In the works of Li et al.<sup>11</sup> and Shi et al.<sup>12</sup>  
92 molecular simulations were used to predict CO<sub>2</sub> solubility in nine and twenty-seven physical  
93 solvents respectively. Although high predictive accuracy was achieved in both works, molecular  
94 simulations are also not very convenient for the large-scale virtual screening, as they are time-  
95 consuming and require significant computational resources. Alternatively, the conductor-like  
96 screening model for real solvents (COSMO-RS),<sup>13</sup> a method combining quantum chemical  
97 calculations with statistical thermodynamics, was suggested to rapidly screen large sets of  
98 various materials. In the work by Kim et al.<sup>14</sup> CO<sub>2</sub> and CH<sub>4</sub> Henry's law coefficients were  
99 predicted by COSMO-RS for 63 common liquid solvents and 10 ionic liquids at 300 K.  
100 Unfortunately, there was no comparison of the predicted solubilities with the available  
101 experimental data.

102 A pool of gas solubility data accumulated in scientific literature supports applying machine  
103 learning for quantitative structure-solubility relationships (QSPR) modeling. In this approach  
104 chemical structures of compounds are encoded as vectors of molecular descriptors and machine

105 learning algorithms are then applied for modeling the property of interest. As compared to the  
106 modeling based on the usage of experimentally determined parameters or resource-intensive  
107 molecular simulations, this method allows efficient screening of large numbers of compounds.  
108 To our knowledge, there was only one work related to QSPR modeling of CO<sub>2</sub> solubility in  
109 physical solvents published.<sup>15</sup> In the paper of Gorji et al.,<sup>15</sup> Henry coefficients for 22 solvents  
110 composed only of carbon, oxygen, and hydrogen elements at different temperatures were used to  
111 build a multiple linear regression model with Dragon<sup>16</sup> descriptors. Although good predictive  
112 performance was achieved, the applicability domain (AD) of this model is limited to the specific  
113 classes of compounds used for model building. Except for recent publications on H<sub>2</sub>S solubility  
114 modeling<sup>17,18</sup>, to our knowledge there were no papers describing the application of the  
115 chemoinformatics-driven methods for physical solubility modeling of the major components  
116 encountered in natural gas treatment or in the pre-combustion CO<sub>2</sub> capture process : carbon  
117 dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), carbon monoxide (CO), hydrogen (H<sub>2</sub>), and nitrogen (N<sub>2</sub>). Hence,  
118 the investigation of the perceptiveness of using chemoinformatics for the rational design of new  
119 solvents for the absorption of CO<sub>2</sub> and other industrial gases is an important task.

120

## 121 **MATERIALS AND METHODS**

### 122 *Data collection and preprocessing*

123 A dataset containing mole fraction solubility values ( $\chi$ ) for 280 liquid solvents at 298.15 K and  
124 1 atm was collected (Table S1, Supporting Information) from IUPAC reports<sup>19-23</sup>, scientific  
125 literature<sup>24-75</sup>, and patents<sup>76,77</sup>. The mole fraction solubility for a binary (gas-liquid) system is  
126 defined<sup>19</sup> as:



127 
$$\chi = \frac{n(g)}{n(g)+n(l)} \quad (1)$$

128  $n(g)$  – an amount of substance in a gas phase,  $n(l)$  – an amount of substance in a liquid phase.

129 Median values were then taken for the solvents associated to several reliable measurements of  
130  $\chi$  at the particular temperature. The final dataset used for modeling is present in Table S1  
131 (Supporting Information). Extrapolation and interpolation of the data to 298.15 K were  
132 performed assuming a linear variation of Henry's coefficients with temperature or by the  
133 equations suggested in the IUPAC report or corresponding papers. The compounds that are  
134 structural outliers with respect to the training set majority, comprising water, carbon disulfide,  
135 octamethylcyclotetrasiloxane and hydrazines were not included to the dataset. These compounds  
136 contain rare or unique fragments significantly affecting their gas-absorbing properties, and,  
137 hence, confident predictions cannot be obtained for them by statistical modeling.

138 The collected mole fraction solubilities were converted to the Kuenen coefficients  $S$  using the  
139 following formula<sup>78</sup>:

140 
$$S = \frac{R \times T \times P}{M_w} \times \frac{\chi}{1 - \chi} \quad (2)$$

141  $S$  – Kuenen coefficient ( $\text{m}^3\text{kg}^{-1}$ ),  $R$  – ideal gas constant ( $8.314 \text{ m}^3\text{PaK}^{-1}\text{mol}^{-1}$ ),  $T$  and  $P$  –  
142 standard temperature and pressure (273.15 K and 101.325 kPa),  $M_w$  – molecular weight of  
143 compound ( $\text{kgmol}^{-1}$ ),  $\chi$  – mole fraction solubility value.

144 The Kuenen coefficient is the volume of saturated gas reduced at 273.15 K and 1 atm pressure,  
145 which is dissolved by unit mass of pure solvent at the temperature of measurement and partial  
146 pressure of 1 atm. This parameter is widely used in industrial applications, as it enables one to

147 directly estimate the efficiency of the particular solvent related to its cost and dimensions of the  
148 required industrial unit (design-capital expenses cost CAPEX). Here, Kuenen coefficients were  
149 used for the data analysis and models interpretation.

150 The selectivity index SI was calculated using the following formula:

$$151 \quad SI = \frac{\chi_{CO_2}}{\chi_{gas}} \quad (3)$$

152 SI – selectivity index, experimental or predicted  $\chi_{CO_2}$  and  $\chi_{gas}$  – mole fraction solubilities of  
153 CO<sub>2</sub> and other gases respectively.

154 All  $\chi$  values were also transformed to a logarithmic scale, i.e. the negative value of the decimal  
155 logarithm was taken (Figure S1).

### 156 *Modeling*

### 157 *Standardization*

158 All compound structures were standardized using in-house standardization procedures based  
159 on KNIME,<sup>79</sup> which included aromatization, stereochemistry depletion, etc.

### 160 *Descriptors*

161 193 different ISIDA fragment descriptor sets were generated using the Fragmentor17  
162 software.<sup>80,81</sup> ISIDA fragments represent either sequences (the shortest topological paths with an  
163 explicit representation of all atoms and bonds), atom-centered fragments (all connected atoms to  
164 a certain topological distance), or triplets (all the possible combinations of 3 atoms in a graph  
165 with the topological distance between each pair indicated). The number of fragments in each set

166 varied from 30-40 (for short sequences of atoms/bonds) to 400-1200 (for long sequences up to 6  
167 atoms) for different gases.

168 Quantum chemical descriptors resulted from DFT calculations in the gas phase, with model  
169 wB97X-D 6-31G\* performed with the Spartan 18.0 program<sup>82</sup>. Default QSAR descriptors  
170 available in Spartan including energy, dipole moment,  $E_{\text{HOMO}}$ , and  $E_{\text{LUMO}}$  were calculated.

### 171 *Machine learning algorithms*

172 Random forest (RF): RF algorithm<sup>83</sup> implemented in sci-kit learn library (v. 0.22.1)<sup>84,85</sup> was  
173 used. The following hyperparameters were tuning during optimization (grid search): number of  
174 trees (100, 300, 1000), number of features (all features, one-third of all features,  $\log_2$  of the  
175 number of features), the maximum depth of the tree (5, 10, full tree), bootstrapping (with and  
176 without the usage of bootstrap samples for building the tree).

177 XGBoost (XGB): XGBoost algorithm<sup>86</sup> as implemented in XGBoost python module (v.1.2.0)<sup>87</sup>  
178 was used. The following hyperparameters were tuning during optimization (grid search): number  
179 of trees (50, 100, 300, 500), number of features (all features, 70% of all features), number of  
180 samples (all samples, 70% of all samples), the maximum depth of the tree (3, 5, 10), learning rate  
181 (0.3, 0.1, 0.5, 0.05), the minimum sum of instance weight needed in a node (1, 5, 10). All other  
182 parameters were left as default.

183 Support vector regression (SVR): SVR algorithm<sup>88</sup> implemented in sci-kit learn library (v.  
184 0.22.1), was used. The descriptors were scaled to the [0,1] range before applying the algorithm.  
185 The following hyperparameters were tuning during optimization (grid search): kernel (linear, rbf,

186 poly, sigmoid), kernel coefficient (1, 0.1, 0.01, 0.001, 0.0001), regularization parameter (0.1, 1,  
187 10, 100, 1000).

### 188 *Model validation workflow*

189 The modeling workflow was implemented using sci-kit learn library (v. 0.22.1) in python 3.7  
190 scripting language. Identical modeling workflows were used for solubility modeling (expressed  
191 as  $-\lg \chi$ ) of all gases. At the first stage of the modeling, a machine learning algorithm: RF, SVR  
192 and XGB were tested in 5-fold cross-validation, which was repeated 5 times (Figure S2). For  
193 each descriptor set, the model's measures of performance were calculated and several models  
194 with a coefficient of determination  $Q^2_{CV} \geq 0.7$  were selected for consensus modeling.

195 The following equations were used to calculate the measures of the model's performance in  
196 cross-validation:

$$197 \quad Q^2_{CV} = \frac{\sum_{j=1}^5 \left(1 - \frac{\sum_{i=1}^n (y_{i,exp} - y_{i,pred})^2}{\sum_{i=1}^n (y_{i,exp} - \bar{y})^2}\right)}{5} \quad (4)$$

$$198 \quad RMSE_{CV} = \frac{\sum_{j=1}^5 \sqrt{\frac{\sum_{i=1}^n (y_{i,exp} - y_{i,pred})^2}{n}}}{5} \quad (5)$$

$$199 \quad MAE_{CV} = \frac{\sum_{j=1}^5 \sum_{i=1}^n \frac{|y_{i,exp} - y_{i,pred}|}{n}}{5} \quad (6)$$

200 Above, n is the number of compounds in the entire learning set,  $y_{i,exp}$ ,  $y_{i,pred}$  experimental and  
201 values predicted in 5-fold cross-validation for compound i from the learning set, j is the index of  
202 the repetition of the 5-fold cross-validation procedure. For each measure of the model's  
203 performance, the standard deviation over 5 repetitions was calculated.

204 Each of the selected models was then associated with an Applicability Domain (AD), defined  
205 as a bounding box.<sup>89</sup> Hence, the pool of selected models extracted from the given data set was  
206 used as a consensus predictor, returning for each input solvent candidate a mean value of  
207 solubility estimates and its standard deviation, taken over the predictions returned by each model  
208 in the pool, if the compounds appeared outside AD of all the models, or, alternatively, over the  
209 predictions returned by only those models having the candidate within their AD.

210 Outlying data points were defined as the data points for which absolute errors ( $|\chi_{\text{exp}} - \chi_{\text{pred}}|$ )  
211 from cross-validation were larger than  $2 \times \text{RMSE}_{\text{CV}}$  threshold.

#### 212 *Y-randomization test*

213 The absence of chance correlation was checked through the Y-randomization procedure. Y-  
214 randomization test was performed in the following way:  $-\lg\chi$  values (y values) were shuffled,  
215 surrogate models from the cross-validation were built using shuffled values and the values from  
216 the corresponding cross-validation test set were calculated. This procedure was repeated 100  
217 times for each fold and the maximum values of the coefficient of determination were compared  
218 with the coefficient of determination obtained for the original  $-\lg\chi$  values.

#### 219 *Virtual screening*

220 An in-house dataset comprising 4,082 industrially produced compounds and their  
221 structural analogs was screened in the following way. Only structures containing the same atoms  
222 (C, H, N, O, S, P, halogens) as in the learning set were kept. All structures were standardized and  
223 ISIDA descriptors were calculated for them as described above. Individual ISIDA models  
224 refitted to the entire dataset with the hyperparameters selected in the cross-validation were used

225 to compose the final consensus model as described above. Then, predictions were made using the  
226 ISIDA consensus model. Only compounds that were inside the applicability domain defined as  
227 bounding box for at least three ISIDA fragment types were considered.

### 228 *Software implementation*

229 The developed model was implemented into the ISIDA-Predictor software.<sup>81</sup>

### 230 *Experimental measurement of CO<sub>2</sub> solubility*

231 A “static-synthetic” technique based on a closed-circuit method<sup>90–92</sup> was used for the  
232 determination of CO<sub>2</sub> solubility in the solvents. In this method, which is explained in detail in the  
233 supplementary information (Text S1, Figures S3, S4), the system pressure is measured at  
234 constant temperature for different overall compositions. To determine the global compositions,  
235 the quantities of pure substances charged into the stirred equilibrium cell, which is evacuated and  
236 placed in a thermostatic liquid bath, need to be known precisely. The purified and degassed  
237 solvents are charged into the cell as compressed liquids using thermostatted piston injectors.  
238 Then, the gas is added stepwise as a liquefied gas using the same injection pumps or as a gaseous  
239 component using a thermo-regulated gas bomb. Knowing the pressure, temperature, and volume  
240 of the gas bomb, the amount of gas inside the bomb can be calculated using correlated PvT data  
241 of the gas. Thus, the injected amount of gas can be obtained from the pressure difference in the  
242 bomb before and after each injection.

243 Since only temperature, pressure, total loadings, gas-liquid interface level and total volumes  
244 are measured, the compositions of the coexisting phases need to be determined by the evaluation  
245 of the raw data. From the known amount of solvent, the liquid phase volume is determined using

246 precise information about the density of the liquid solution inside the equilibrium chamber. From  
247 the total volume of the cell, the remaining gas phase volume can be calculated precisely (see  
248 supplementary information). At given equilibrium conditions (temperature, gas phase volume,  
249 and gas pressure) the amounts of gas in the gas phase and thus, also in the liquid phase are  
250 obtained. In this approach, several effects influence the resulting liquid phase compositions.  
251 These effects are the small amounts of solvents in the gas phase, the compressibility of the  
252 solvent under the gas pressure, the partial molar volume of the dissolved gas and the solvent  
253 activity coefficient. All effects are considered in an isothermal and isochoric algorithm by  
254 solving the mass and volume balances.

255 The partial pressure is obtained during the iterative procedure:

$$256 \quad P_{\text{gas}} = P_{\text{sys}} - P_{\text{solvent}} \quad (6)$$

257 where  $P_{\text{gas}}$  – partial pressure of the acid gas in the system,  $P_{\text{sys}}$  – total pressure in the system,  
258  $P_{\text{solvent}}$  – partial pressure of a solvent vapour. This equation is valid at low pressure and in the  
259 absence of chemical reactions in the gas phase. The method to calculate the uncertainty of the  
260 measured experimental data is explained in the supplementary information. The uncertainty of  
261 the measured  $\text{CO}_2$  solubility at 1 atm. is equal or lower than 1%.

## 262 **RESULTS AND DISCUSSION**

### 263 *Data collection, preprocessing and analysis*

264 IUPAC reports on gas solubilities in non-aqueous solvents contain to our knowledge the most  
265 complete and carefully analyzed publicly available data on gas solubility. The data from these  
266 reports were used to compose the “cores” of our datasets. As the largest number of data points

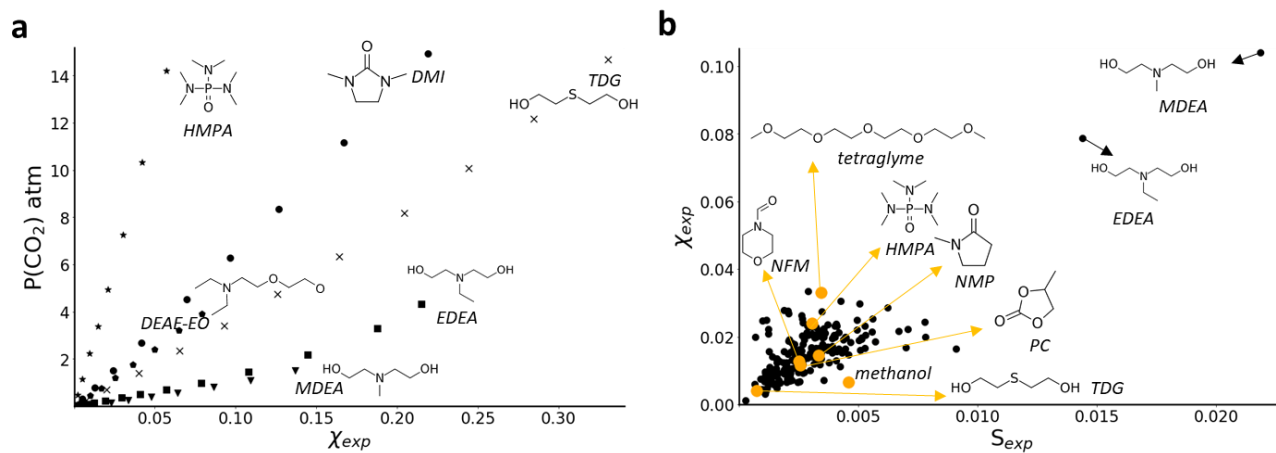
267 for various solvents was available at 298.15 K and 1 atm, the data at this temperature and  
268 pressure was chosen for modeling. Since the mole fraction values can vary significantly  
269 depending on the experimental methods being used, we chose only the data points which were  
270 considered as the most reliable by IUPAC's or Total's experts. Data from recent publications  
271 either at 298.15 K or obtained by extrapolation or interpolation of the data measured at close  
272 temperatures were also added to the dataset.

273 Besides the data collected from IUPAC reports and literature, data points for six compounds,  
274 hexametapol (HMPA), 1,3-Dimethylimidazolidin-2-one (DMI), thiodiglycol (TDG), and three  
275 tertiary amines 2-[2-hydroxyethyl(methyl)amino]ethanol (MDEA), 2-[ethyl(2-  
276 hydroxyethyl)amino]ethanol (EDEA), and 2-[2-(diethylamino)ethoxy]ethanol (DEAE-EO) were  
277 measured experimentally and added to the dataset (Figure 1a; Table S1). The choice of the  
278 solvents was motivated by their wide application in industrial processes and the absence of  
279 consistent data at 298.15 K for them in the literature. TDG is employed in a commercial mixed  
280 chemical/physical solvent formulation for sour gas treating (HySWEET technology) developed  
281 by TotalEnergies S.E.<sup>93</sup> HMPA and DMI are being used as solvents for gases, polymers, and in  
282 organic synthesis. Aqueous amines are used as chemical solvents, and only little is known about  
283 the physical solubility of gases in pure amines. One of the few examples is MDEA, which, in an  
284 aqueous solution, is commonly used for industrial gas treatment, and for which the CO<sub>2</sub> mole  
285 fraction solubility can be estimated from Skylogianni.<sup>94</sup> The value is extremely high (~0.04 at 1  
286 atm and 313K) as compared to other physical solvents (see below). Hence, considering the  
287 growing interest in water-lean solvents<sup>95</sup>, including the ones based on pure amines<sup>96</sup>, we have  
288 chosen three industrial amines (MDEA, EDEA, DEAE-EO) for the experimental assessment of  
289 CO<sub>2</sub> physical solubility.



290 All the solvents showed close to linear variation of mole fraction solubility vs partial pressure  
291 of the gas in the pressure range 0-2 atm (Figure 1a). Estimated mole fraction solubilities for TDG  
292 and DMI (0.0041; 0.0150) are in good agreement with the data that can be obtained by  
293 extrapolation from recent publications.<sup>27,97</sup> On the contrary, CO<sub>2</sub> mole fraction solubility in  
294 HMPA (0.024) is lower, than the one suggested by IUPAC's expert<sup>19</sup> (0.031), but is close to the  
295 one obtained by Schay et al.<sup>19</sup> (0.028). The mole fractions values in MDEA (0.1) and EDEA  
296 (0.08) are remarkably high and in agreement with the aforementioned data for MDEA from the  
297 Skylogianni et al. obtained at higher temperatures.<sup>94</sup> At the same time, the CO<sub>2</sub> solubility in  
298 DEAE-EO is much lower (0.02) and is the same as in two other tertiary amines present in the  
299 dataset: triethylamine (0.02) and perfluorotributylamine (0.02).

300 The high CO<sub>2</sub> solubility in pure MDEA and EDEA triggers the question whether the  
301 absorption is purely physical. It is commonly reported in the literature that, contrary to primary  
302 and secondary amines, tertiary amines cannot chemically absorb CO<sub>2</sub> in the absence of water<sup>98,99</sup>  
303 (we have verified that no water was present in the solvent in the experiments performed in the  
304 present paper). This view has been challenged by Maddox<sup>100</sup> and more recently by Heldebrandt  
305 et al.<sup>101,102</sup> who studied the reaction of CO<sub>2</sub> with anhydrous tertiary amines. Both conclude that  
306 reversible Lewis acid-base adducts are formed at high pressure (note that in this work we  
307 compare solubilities at low pressure, 1 atm.). Heldebrandt<sup>101</sup> suggests that the difference in  
308 absorption capacity between pure amines can be explained by solvent polarity effects. Either  
309 way, anhydrous tertiary amines do absorb less CO<sub>2</sub> than aqueous tertiary amines, but some still  
310 absorb significant amounts of CO<sub>2</sub>.



311  
 312 Figure 1. (a) Variation of mole fraction with partial pressure for CO<sub>2</sub> in TDG (×), DMI (★), HMPA (●), MDEA  
 313 (▼), EDEA (■), and DEAE-EO (□) at 298.15 K experimentally measured for this paper. (b) Plot of experimental  
 314 molecular fraction values ( $\chi_{exp}$ ) vs Kuenen coefficients ( $S_{exp}$ ) at 298.15 K and 1 atm for CO<sub>2</sub>.

315 There were 211 mole fraction solubility values collected for CO<sub>2</sub>. The largest mole fraction  
 316 CO<sub>2</sub> solubility was for tertiary amines MDEA and EDEA (Figure 1b). Among other classes of  
 317 compounds with large CO<sub>2</sub> solubility were phosphoric acid esters, long chain ethers, and esters.  
 318 To estimate the efficiency of a solvent related to its cost and dimensions of the required  
 319 industrial unit, mole fractions were converted to Kuenen coefficients. MDEA and EDEA also  
 320 have the largest Kuenen coefficients. By contrast to the trend observed for mole fractions, the  
 321 largest Kuenen coefficient values in other compound classes were for small polar compounds:  
 322 nitriles, ketones (acetone, butan-2-one), tetrahydrofuran (THF). Notably, the solvents, which are  
 323 used in industrial gas treatment processes are not among the best ones in terms of CO<sub>2</sub> solubility  
 324 (Figure 1b).

325 Other gases are less studied as compared to CO<sub>2</sub>. There were less than 105 mole fraction  
 326 solubility values collected for each of other gases (N<sub>2</sub>, H<sub>2</sub>, CO, CH<sub>4</sub>). It is worth noting, that  
 327 polar CO and non-polar CH<sub>4</sub>, N<sub>2</sub>, H<sub>2</sub> showed similar solubility trends. The largest values were

328 for non-polar compounds, including perfluorated alkanes, and long-chain n-alkanes (Figure S5).  
329 The minimal  $\chi$  values were for polar solvents, such as methanol, *N,N*-dimethylformamide  
330 (DMF).

331 The collected experimental data were used to analyze the trends in the selectivity of CO<sub>2</sub>  
332 absorption towards other gases (Figure S6). The only industrial solvent for which the data were  
333 available for all the solvents is methanol, which is not selective at 298.15 K and 1 atm. The  
334 industrially used solvents were among the best ones in terms of the CO<sub>2</sub>/CH<sub>4</sub> selectivity. The  
335 highest selectivity index (SI<sub>exp</sub>) was for dimethyl sulfoxide (DMSO, SI<sub>exp</sub>=24) and *N*-formyl  
336 morpholine (NFM, SI<sub>exp</sub>=21). Among other most selective solvents were industrially used  
337 propylene carbonate (PC) and *N*-Methyl-2-pyrrolidone (NMP). The large selectivity stems from  
338 the extremely low solubility of CH<sub>4</sub> in these solvents. The same observation was made for all  
339 other gases: the most selective are the polar solvents, such as DMF, 1,4-dioxane, DMSO, etc.  
340 (Figure 1b, Figures S3-S4).

#### 341 *Quantitative structure solubility relationships*

342 Application of the machine learning allowed one to establish quantitative structure-  
343 solubility relationships. Reasonable predictive accuracy was achieved in the repeated cross-  
344 validation procedure for all the gases (Table 1). None of the models has shown chance  
345 correlation in the y-scrambling procedure.

346 Table 1. Performance estimation for modeling of mole fraction solubility expressed as  $-\lg\chi$ .

| Gas             | ISIDA consensus models | $Q^2_{CV}$ | RMSE <sub>CV</sub> | MAE <sub>CV</sub> |
|-----------------|------------------------|------------|--------------------|-------------------|
| CO <sub>2</sub> | 20 RF, 17 XGBoost      | 0.71±0.01  | 0.12±0.01          | 0.08±0.01         |

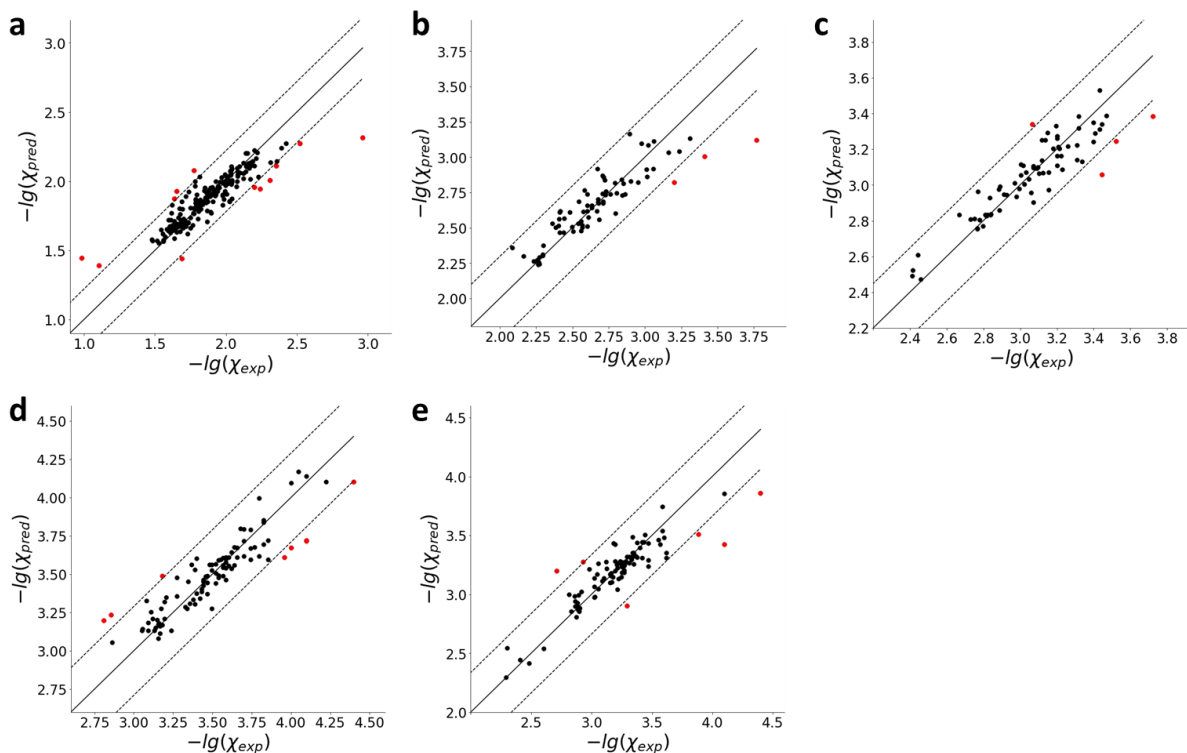
|                 |                         |           |           |           |
|-----------------|-------------------------|-----------|-----------|-----------|
| CH <sub>4</sub> | 9 SVR, 15 XGBoost       | 0.77±0.02 | 0.15±0.01 | 0.10±0.01 |
| CO              | 5 SVR, 8 XGBoost        | 0.78±0.03 | 0.12±0.01 | 0.09±0.01 |
| H <sub>2</sub>  | 2 RF, 3 SVR, 23 XGBoost | 0.77±0.04 | 0.15±0.01 | 0.10±0.01 |
| N <sub>2</sub>  | 9 SVR, 12 XGBoost       | 0.75±0.06 | 0.17±0.01 | 0.11±0.01 |

347

348 Since the collected datasets are small, the presence of compounds containing rare fragments, or  
 349 compounds with noise in the experimental data lead to unstable modeling results. Several  
 350 compounds, which were systematically mispredicted (the absolute error >0.7 log units) in the  
 351 cross-validation procedure were removed: dodecanal and dodecene for the H<sub>2</sub> model, and  
 352 dimethyl ether and dodecanal for the CO model. The values for all these compounds are  
 353 significantly different from their close structural analogs. They were obtained by interpolation  
 354 and additional experimental confirmation is required to assess whether the values are reliable.  
 355 After the removal of outliers, the models with reasonable figures of merit were obtained for each  
 356 gas (Table 1). The lowest mean absolute error (MAE<sub>CV</sub>) was for the CO<sub>2</sub> model, which is based  
 357 on the largest pool of data. Note, that MAE<sub>CV</sub> is close to the variance in the experimental data.  
 358 For example, the standard deviation for propylene carbonate based on IUPAC's data<sup>19</sup> and the  
 359 recently published data<sup>36</sup> can be estimated as 0.05 log units.

360 To check if some other descriptor types can lead to significantly better results, we calculated  
 361 quantum chemical descriptors using Spartan software. The results of modeling were on average  
 362 comparable to those obtained by the usage of ISIDA fragments (Table S2). Considering the  
 363 advantages of ISIDA fragments, i.e. speed of calculation and intuitive interpretation of structure-  
 364 property relationships, we further focused on this descriptor type.

365 For each model, the compounds for which absolute errors were larger than  $2 \times \text{RMSE}_{\text{CV}}$   
366 threshold were analyzed (Figure 2, Table S3). These compounds either contain rare fragments or  
367 can be considered as “solubility cliffs”: small changes in structure (e.g. replacement of hydrogen  
368 atom by methyl group) lead to large changes in solubility (see Figure S7 and discussion below).  
369 For example, hexafluorobenzene – the only polyhalogenated aromatic compound and MDEA –  
370 one of few representatives of alkanolamines in the dataset were among the compounds with the  
371 largest errors for the  $\text{CO}_2$  model. The datasets for the gases are rather chemically diverse. Many  
372 compounds containing rare fragments appear outside AD of the models in the cross-validation.  
373 There were 15% of compounds appearing to be outside AD in the cross-validation for  $\text{CO}_2$ ,  
374 while about 20-25% of compounds were outside AD for other gases. The presence of compounds  
375 with rare fragments leads to high variance of predictions in the cross-validation. From the  
376 learning curve (Figure S8), one can see that adding data improves the performance on the  
377 validation sets, and, therefore, decreases the gap between prediction accuracy on validation and  
378 training sets. Hence, further accumulation of the experimental data on gas solubility organic  
379 solvents is required for building more robust models with enlarged applicability domains and the  
380 extended range of temperature and pressure values.



381

382 Figure 2. Plot of predicted ( $-\lg\chi_{\text{pred}}$ ) vs experimental ( $-\lg\chi_{\text{exp}}$ ) values for ISIDA consensus model in cross-  
 383 validation procedure for  $\text{CO}_2$  (a),  $\text{CH}_4$  (b),  $\text{CO}$  (c),  $\text{H}_2$  (d),  $\text{N}_2$  (e). The predicted values are calculated as an average  
 384 of 5 folds. Compounds for which absolute errors were larger than  $2 \times \text{RMSE}_{\text{CV}}$  are shown in red. Dash lines indicate  
 385  $\pm 2 \times \text{RMSE}_{\text{CV}}$  threshold.

386

387 In contrast to the above examples, solubility of  $\text{CO}_2$  in alcohols, glycols and ethers was  
 388 systematically studied. Yet, one of the largest absolute errors were for glycols (glycerol, ethane-  
 389 1,2-diol), which is related to a sharp change in solubility with the replacement of  $-\text{OH}$  group to  $-$   
 390  $\text{OCH}_3$  (Figure S7). For example, the mole fraction solubility in glycerol is more than three-time  
 391 smaller, than the solubility in its closest structural analog – propylene glycol (PG). At the same  
 392 time solubility in another structural analog – diethylene glycol (DEG), containing the same  
 393 number of carbon and oxygen atoms as glycerol, is six time higher. This phenomenon can be

394 explained by considering forces driving the process of gas dissolution. The mechanistic  
395 interpretation of this process assumes the formation of a cavity capable of accommodating a gas  
396 molecule by breaking solvent-solvent bonds and introduction and fixation of a gas molecule in  
397 this cavity due to gas-solvent interactions. Hence, solubility of gases in liquids depends upon two  
398 types of interactions: gas-solvent and solvent-solvent.<sup>103,104</sup> Strong gas-solvent and weak solvent-  
399 solvent interactions lead to greater solubility. In line with that, CO<sub>2</sub> solubility in glycols and their  
400 ethers follows the cohesive energy density values trend: the solubility is increasing from glycerol  
401 to DEG with the decreasing cohesive energy density (glycerol: 1142 MPa; DEG: 615 MPa).<sup>105</sup>  
402 However, the cohesive energy density is not the only factor affecting the solubility of CO<sub>2</sub>.  
403 Although carbon dioxide is nonpolar, its appreciable polarizability and ability to accept hydrogen  
404 bonds from suitable donor solvents<sup>106</sup> makes structure solubility landscape more complex. For  
405 example, CO<sub>2</sub> solubility is lower in hexane than in dimethyl ether of ethylene glycol (DMEG),  
406 while the cohesive energy density of DMEG is higher than that of hexane (DMEG: 317 MPa;  
407 hexane: 222 MPa)<sup>105,107</sup>. Thus, the interplay between cohesive energy density and solvent-CO<sub>2</sub>  
408 interactions should be taken into account in the process of the design of new solvents.

#### 409 *Virtual screening*

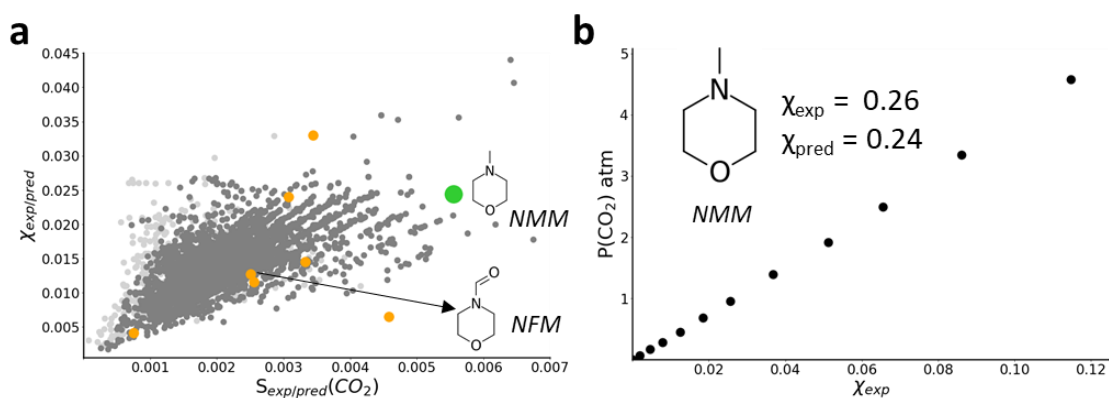
410 To find new solvents with high CO<sub>2</sub> solubility and high selectivity towards other gases, we  
411 performed the virtual screening of the in-house library of industrially produced chemicals and  
412 their close structural analogs comprising more than 4,000 chemicals (Figure S9). It is worth  
413 noting that the experimentally measured physico-chemical properties such as melting and boiling  
414 points, density, flash points, etc. were available only for a small fraction of the dataset and thus,  
415 we did not check if the compounds possess plausible values of properties at 298.15 K and 1 atm.  
416 Most of the screened compounds (87%) appeared to be inside AD of the CO<sub>2</sub> model. There were

417 numerous compounds found with high predicted CO<sub>2</sub> mole fractions values and Kuenen  
418 coefficients, several of which were superior to the existing industrially used solvents (Figure 3a).  
419 Among the best CO<sub>2</sub> solvents according to mole fraction solubilities were tertiary amines and  
420 long-chain esters (e.g., dioctyl adipate,  $\chi_{\text{pred}} = 0.27$ ), while the largest Kuenen coefficients were  
421 for tertiary amines and the close structural analogs of the compounds with the largest Kuenen  
422 coefficients from the learning set: ethers (e.g. ethyl methyl ether,  $S_{\text{pred}} = 0.0068$ ), ketones (e.g.  
423 methoxyacetone,  $S_{\text{pred}} = 0.0053$ ) and nitriles (e.g. butyronitrile,  $S_{\text{pred}} = 0.0052$ ). One of the  
424 tertiary amines with the largest Kuenen coefficient ( $S_{\text{pred}} = 0.0054$ ) – 4-Methylmorpholine  
425 (NMM) was selected for the experimental measurement of solubility. NMM showed linear  
426 variation of mole fraction solubility vs partial pressure of the gas in the pressure range indicating  
427 pure physical solubility (Figure 3b). The experimental mole fraction solubility (0.26) matched  
428 the predicted one (0.24) well. The CO<sub>2</sub> solubility in NMM is appreciably higher than in  
429 industrially used NFM, which instead of a tertiary amine group contains an amide group. We  
430 have measured the dynamic viscosity of NMM: 0,92 cP at 20 °C. The dynamic viscosity of water  
431 at 20 °C is 1 cP. The viscosity of NMM is thus comparable to water. NMM is much less viscous  
432 than e.g. pure MDEA (100 cP at 20 °C) or pure EDEA (90 cP at 20 °C), which is a significant  
433 advantage. On the other hand, the boiling point of NMM is 116 °C (MDEA 243 °C). NMM is  
434 thus more volatile than MDEA. The NMM solvent should thus be used at a lower temperature, to  
435 minimize the solvent losses. This is e.g. also done in the Rectisol process which uses methanol  
436 (boiling point 65 °C).

437 The environment, health and safety (EHS) of solvents for CO<sub>2</sub> capture is a potential issue. In  
438 principle, amine emissions should not be an obstacle because the causes are well known and  
439 counter-measures can be put in place (operating temperature and pressure, water wash, Brownian



440 demister, reclaiming units, etc.). One clear advantage of using physical solvents is that there is  
441 much less thermal and oxidative solvent degradation because the regeneration is not thermal and  
442 the high operating pressure avoids oxygen ingress. According to the safety datasheet<sup>108</sup> NMM is  
443 flammable, corrosive and harmful, but the substance contains no components considered to be  
444 either persistent, bioaccumulative and toxic.



445  
446 Figure 3. (a) A plot of CO<sub>2</sub> molecular fraction values vs Kuenen coefficients. Experimental values – black and  
447 orange (solvents used in the industry). (b) Variation of mole fraction with partial pressure for CO<sub>2</sub> in NMM at  
448 298.15 K experimentally measured in this paper;  $\chi_{exp}$  – experimental mole fraction value at 1 atm and 298.15 K,  $\chi_{pred}$   
449 – predicted value. Predicted values for compounds inside AD – grey, outside AD – light grey, for NMM – green.  
450 The position of the NFM on the plot is shown for comparison.

451  
452 The selectivity of the NMM calculated from predicted mole fraction values is comparable to  
453 the one of industrially used solvents (Figure S10). Other tertiary amines were also among the  
454 most selective solvents. For example, the 3-(Dimethylamino)-1,2-propanediol was among the  
455 most selective solvents for all gases. Among other classes of solvents with the highest selectivity

456 indexes were cyclic amides (e.g. 5-(hydroxymethyl)-1-methylpyrrolidin-2-one) and ketones (e.g.  
457 methoxyacetone).

458 To conclude, the rational approach to the design of new physical solvents based on the usage  
459 of machine learning for modeling of structure-solubility relationships was suggested in this  
460 paper. The collected data on solubility of gases were used to build QSPR models, which were  
461 then applied to identify compounds potentially superior to the existing ones via virtual screening  
462 of industrially produced chemicals. We have identified pure tertiary amines with a remarkable  
463 CO<sub>2</sub> absorption capacity. Previously, the team of Heldebrandt et al.<sup>101</sup> has investigated the use of  
464 pure, anhydrous amines for high pressure CO<sub>2</sub> absorption. They compared the performance of  
465 anhydrous EDEA to the Fluor solvent (propylene carbonate), to Selexol, and to aqueous MDEA  
466 for a representative absorber. Despite the attractiveness due to lower energy consumptions, the  
467 use of anhydrous or water-lean amines faces numerous challenges, for example, their lower  
468 absorption capacity and their higher viscosity.<sup>6</sup> In this work we have focused on the gas  
469 solubility. A further extension of the chemoinformatics workflow for the prediction of other  
470 industrial important solvent properties might be very useful in the identification of the most  
471 suitable physical (co-)solvent (optimal absorption properties, selectivity, viscosity,<sup>109</sup> EHS  
472 impact,<sup>110</sup> etc.) for a given application.

473

474 ASSOCIATED CONTENT

475 **Supporting Information.**

476

477 Table S1. Mole fraction solubilities of gases in organic liquid solvents at 298.15 K and 1 atm.

478 Figure S1. Negative logarithm of mole fraction solubility values distribution.

479 Figure S2. A scheme for building the ISIDA consensus model.

480 Text S1. Description of the experimental protocols.

481 Figure S3a. Schematic diagram of apparatus used to determine gas solubility.

482 Figure S3b. Flow diagram of the synthetic apparatus used to determine gas solubility.

483 Figure S4. A pictorial view of CO<sub>2</sub> loading arrangement.

484 Figure S5. Plots of experimental molecular fraction values vs Kuenen coefficients for (CH<sub>4</sub>, H<sub>2</sub>,  
485 CO, N<sub>2</sub>).

486 Figure S6. Plots of selectivity indexes based on experimental molecular fraction values vs CO<sub>2</sub>  
487 Kuenen coefficients.

488 Table S2. Outliers from ISIDA consensus models.

489 Table S3. Performance estimation for modeling using quantum-chemical descriptors.

490 Figure S7. Variation of CO<sub>2</sub> solubility in several series of structurally similar solvents.

491 Figure S8. Learning curves for machine learning models. Figure S9. Distribution of molecular  
492 descriptors for the compounds from the screening library and the learning set.

493 Figure S10. Plots of predicted selectivity indexes vs Kuenen coefficient of CO<sub>2</sub>.

494

## 495 AUTHOR INFORMATION

### 496 **Corresponding Author**

497 \*Professor Alexandre Varnek. Laboratory of Chemoinformatics, Faculty of Chemistry,

498 University of Strasbourg, 4, Blaise Pascal Str., 67081, Strasbourg, France. email:

499 varnek@unistra.

500 \*Doctor Frédéric de Meyer. TotalEnergies S.E., Exploration Production, Development and

501 Support to Operations, Liquefied Natural Gas – Acid Gas Entity, CCUS R&D Program, Paris,

502 92078 France. email: [frederick.de-meyer@totalenergies.com](mailto:frederick.de-meyer@totalenergies.com).

### 503 **Author Contributions**

504 The manuscript was written through contributions of all authors. All authors have given approval

505 to the final version of the manuscript.

### 506 **Funding Sources**

507 This work was supported by the Carbon Capture Utilization and Storage (CCUS) transverse

508 R&D program from TotalEnergies S.E.

## 509 ACKNOWLEDGMENT

510 The authors are grateful to Dr. Fanny Bonachera for her help with the implementation of

511 models to the Predictor software. The authors sincerely acknowledge the Laboratory for

512 Thermophysical Properties in Oldenburg, Germany, for performing a part of the experimental

513 work.

## 514 ABBREVIATIONS

- 515 Caroxin D – 1,1,2,2,3,3,4,4-octafluoro-1,4-bis(1,1,1,2,3,3,3-heptafluoropropan-2-yloxy)butane
- 516 Caroxin F – 1,1,1,2,2,3,3,4,4,5,5,6,6-tridecafluoro-6-(1,1,1,2,3,3,3-heptafluoropropan-2-yloxy)hexane
- 517
- 518 DEG – 2-(2-hydroxyethoxy)ethanol (diethylene glycol)
- 519 DEGM – 2-(2-Methoxyethoxy)ethan-1-ol (diethylene glycol monomethyl ether)
- 520 diglyme – 1-methoxy-2-(2-methoxyethoxy)ethane
- 521 DMF – N,N-dimethylformamide
- 522 DMI – 1,3-Dimethylimidazolidin-2-one
- 523 DMSO – methylsulfinylmethane (dimethyl sulfoxide)
- 524 EG – ethane-1,2-diol (ethylene glycol)
- 525 glycerol – propane-1,2,3-triol
- 526 HMPA – N-[bis(dimethylamino)phosphoryl]-N-methylmethanamine (hexametapol)
- 527 M2CA – methyl 2-cyanoacetate
- 528 MDEA – 2-[2-hydroxyethyl(methyl)amino]ethanol 2-[ethyl(2-hydroxyethyl)amino]ethanol
- 529 EDEA – 2-[ethyl(2-hydroxyethyl)amino]ethanol
- 530 DEAE-EO – 2-[2-(diethylamino)ethoxy]ethanol
- 531 methoxyacetone – 1-methoxypropan-2-one
- 532 NMM – 4-methylmorpholine

- 533 NFM – morpholine-4-carbaldehyde (N-formylmorpholine)
- 534 NMP – 1-methylpyrrolidin-2-one
- 535 PC – 4-methyl-1,3-dioxolan-2-one (propylene carbonate)
- 536 TDG – 2-(2-hydroxyethylsulfanyl)ethanol (thiodiglycol)
- 537 pentaglyme – 1-methoxy-2-[2-[2-[2-(2-methoxyethoxy)ethoxy]ethoxy]ethoxy]ethane
- 538 perflubron – 1-bromo-1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-heptafluorooctane
- 539 perfluoroheptane – 1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,7-hexadecafluoroheptane
- 540 perfluoro(methylcyclohexane) – 1,1,2,2,3,3,4,4,5,5,6-undecafluoro-6-
- 541 (trifluoromethyl)cyclohexane
- 542 perfluorooctane – 1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-octafluorooctane
- 543 perfluorotributylamine – 1,1,2,2,3,3,4,4,4-nonafluoro-N,N-bis(1,1,2,2,3,3,4,4,4-
- 544 nonafluorobutyl)butan-1-amine
- 545 THF – oxolane (tetrahydrofuran)
- 546 TPrP – tripropyl phosphate
- 547  $\chi$  – mole fraction solubility
- 548 S – Kuenen coefficient
- 549 SI – Kuenen coefficients selectivity index
- 550 squalane – 2,6,10,15,19,23-hexamethyltetracosane

551

552 REFERENCES

- 553 (1) Net Zero by 2050, A Roadmap for the Global Energy Sector, IEA Report, May 2021.  
554 <https://www.iea.org/reports/Net-Zero-by-2050> (Accessed 17.06.2021).
- 555 (2) Sifat, N. S.; Haseli, Y. A Critical Review of CO<sub>2</sub> Capture Technologies and Prospects for  
556 Clean Power Generation. *Energies* 2019, 12 (21), 4143.  
557 <https://doi.org/10.3390/en12214143>.
- 558 (3) Abdulsalam, J.; Mulopo, J.; Amosa, M. K.; Bada, S.; Falcon, R.; Oboirien, B. O. Towards a  
559 Cleaner Natural Gas Production: Recent Developments on Purification Technologies. *Sep.*  
560 *Sci. Technol.* 2019, 54 (15), 2461–2497. <https://doi.org/10.1080/01496395.2018.1547761>.
- 561 (4) N.Borhani, T.; Wang, M. Role of Solvents in CO<sub>2</sub> Capture Processes: The Review of  
562 Selection and Design Methods. *Renew. Sustain. Energy Rev.* 2019, 114, 109299.  
563 <https://doi.org/10.1016/j.rser.2019.109299>.
- 564 (5) Wang, X.; Song, C. Carbon Capture From Flue Gas and the Atmosphere: A Perspective.  
565 *Front. Energy Res.* 2020, 8, 560849. <https://doi.org/10.3389/fenrg.2020.560849>.
- 566 (6) Wanderley, R. R.; Pinto, D. D. D.; Knuutila, H. K. From Hybrid Solvents to Water-Lean  
567 Solvents – A Critical and Historical Review. *Sep. Purif. Technol.* 2021, 260, 118193.  
568 <https://doi.org/10.1016/j.seppur.2020.118193>.
- 569 (7) Collodi, G.; Azzaro, G.; Ferrari, N.; Santos, S. Techno-Economic Evaluation of Deploying  
570 CCS in SMR Based Merchant H<sub>2</sub> Production with NG as Feedstock and Fuel. *Energy*  
571 *Procedia* 2017, 114, 2690–2712. <https://doi.org/10.1016/j.egypro.2017.03.1533>.
- 572 (8) Yan, Y.; Thanganadar, D.; Clough, P. T.; Mukherjee, S.; Patchigolla, K.; Manovic, V.;  
573 Anthony, E. J. Process Simulations of Blue Hydrogen Production by Upgraded Sorption

- 574 Enhanced Steam Methane Reforming (SE-SMR) Processes. *Energy Convers. Manag.* 2020,  
575 222, 113144. <https://doi.org/10.1016/j.enconman.2020.113144>.
- 576 (9) Pirig N. Ya.; Polyuzhin I. V.; Makitra R. G. Carbon Dioxide Solubility. *Russ. J. Appl.*  
577 *Chem.* 1993, 4 (66), 691–695.
- 578 (10) Li, H.; Yan, D.; Zhang, Z.; Lichtfouse, E. Prediction of CO<sub>2</sub> Absorption by Physical  
579 Solvents Using a Chemoinformatics-Based Machine Learning Model. *Environ. Chem. Lett.*  
580 2019, 17 (3), 1397–1404. <https://doi.org/10.1007/s10311-019-00874-0>.
- 581 (11) Li, H.; Tang, Z.; He, Z.; Gui, X.; Cui, L.; Mao, X. Structure-Activity Relationship for CO<sub>2</sub>  
582 Absorbent. *Energy* 2020, 197, 117166. <https://doi.org/10.1016/j.energy.2020.117166>.
- 583 (12) Shi, W.; Thompson, R. L.; Macala, M. K.; Resnik, K.; Steckel, J. A.; Siefert, N. S.;  
584 Hopkinson, D. P. Molecular Simulations of CO<sub>2</sub> and H<sub>2</sub> Solubility, CO<sub>2</sub> Diffusivity, and  
585 Solvent Viscosity at 298 K for 27 Commercially Available Physical Solvents. *J. Chem.*  
586 *Eng. Data* 2019, 64 (9), 3682–3692. <https://doi.org/10.1021/acs.jced.8b01228>.
- 587 (13) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the  
588 Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* 1995, 99 (7), 2224–2235.  
589 <https://doi.org/10.1021/j100007a062>.
- 590 (14) Kim, J.; Maiti, A.; Lin, L.-C.; Stolaroff, J. K.; Smit, B.; Aines, R. D. New Materials for  
591 Methane Capture from Dilute and Medium-Concentration Sources. *Nat. Commun.* 2013, 4  
592 (1), 1694. <https://doi.org/10.1038/ncomms2697>.
- 593 (15) Gorji, A. E.; Gorji, Z. E.; Riahi, S. Quantitative Structure-Property Relationship (QSPR) for  
594 Prediction of CO<sub>2</sub> Henry's Law Constant in Some Physical Solvents with Consideration of  
595 Temperature Effects. *Korean J. Chem. Eng.* 2017, 34 (5), 1405–1415.  
596 <https://doi.org/10.1007/s11814-017-0018-0>.



- 597 (16) Kode Srl, *Dragon (Software for Molecular Descriptor Calculation) Version 7.0.8, 2017,*  
598 *https://chm.kode-solutions.net.*
- 599 (17) Orlov, A. A.; Marcou, G.; Horvath, D.; Cabodevilla, A. E.; Varnek, A.; Meyer, F. de.  
600 Computer-Aided Design of New Physical Solvents for Hydrogen Sulfide Absorption. *Ind.*  
601 *Eng. Chem. Res.* 2021, 60 (23), 8588–8596. <https://doi.org/10.1021/acs.iecr.0c05923>.
- 602 (18) H. Rostami; Riahi, S. Quantitative Structure–Property Relationship Study on Solubility of  
603 Hydrogen Sulfide in Organic Solvent; Kish, Iran, 2014.
- 604 (19) Carbon Dioxide in Non-Aqueous Solvents at Pressures Less than 200 KPa. In IUPAC  
605 Solubility Data Series (Volume 50); Fogg, P. G. T.; Ed.; Pergamon: Amsterdam, 61988; Pp.  
606 1-483, 1992.
- 607 (20) Methane. In IUPAC Solubility Data Series (Volume 27/28); Clever, H. L.; Young, C. L.;  
608 Eds.; Pergamon: Amsterdam, 61988; Pp 1–783, 1987.
- 609 (21) Carbon Monoxide. In IUPAC Solubility Data Series (Volume 43); Cargill, R.W.; Ed; Eds.;  
610 Pergamon: Amsterdam, 61988; Pp 1–783, 1990.
- 611 (22) Hydrogen and Deuterium. In IUPAC Solubility Data Series (Volume 5/6); Young, C. L.;  
612 Ed; Pergamon: Amsterdam, 61988; Pp 1–646, 1981.
- 613 (23) Nitrogen and Air. In IUPAC Solubility Data Series (Volume 10); Battino, R; Ed;  
614 Pergamon: Amsterdam, 61988; Pp 1–570, 1982.
- 615 (24) Décultot, M.; Ledoux, A.; Fournier-Salaün, M.-C.; Estel, L. Solubility of CO<sub>2</sub> in Methanol,  
616 Ethanol, 1,2-Propanediol and Glycerol from 283.15 K to 373.15 K and up to 6.0 MPa. *J.*  
617 *Chem. Thermodyn.* 2019, 138, 67–77. <https://doi.org/10.1016/j.jct.2019.05.003>.

- 618 (25) Yamamoto, H.; Kamei, H.; Tokunaga, J. Solubilities of Argon, Oxygen and Nitrogen in  
619 1,2-Propanediol + Water Mixed Solvent at 298.15 K and 101.33 kPa. *J. Chem. Eng. Japan.*  
620 1994, 27 (4), 455-459.
- 621 (26) Li, Y.; Liu, Q.; Huang, W.; Yang, J. Solubilities of CO<sub>2</sub> Capture Absorbents Methyl  
622 Benzoate, Ethyl Hexanoate and Methyl Heptanoate. *J. Chem. Thermodyn.* 2018, 127, 25–  
623 32. <https://doi.org/10.1016/j.jct.2018.07.010>.
- 624 (27) Li, X.; Jiang, Y.; Han, G.; Deng, D. Investigation of the Solubilities of Carbon Dioxide in  
625 Some Low Volatile Solvents and Their Thermodynamic Properties. *J. Chem. Eng. Data*  
626 2016, 61 (3), 1254–1261. <https://doi.org/10.1021/acs.jced.5b00893>.
- 627 (28) Henni, A.; Tontiwachwuthikul, P.; Chakma, A. Solubilities of Carbon Dioxide in  
628 Polyethylene Glycol Ethers. *Can. J. Chem. Eng.* 2008, 83 (2), 358–361.  
629 <https://doi.org/10.1002/cjce.5450830224>.
- 630 (29) Zhao, Z.; Xing, X.; Tang, Z.; Zhao, Y.; Fei, W.; Liang, X.; He, Z.; Zhang, S.; Guo, D.  
631 Solubility of CO<sub>2</sub> and H<sub>2</sub>S in Carbonates Solvent: Experiment and Quantum Chemistry  
632 Calculation. *Int. J. Greenh. Gas Control* 2017, 59, 123–135.  
633 <https://doi.org/10.1016/j.ijggc.2017.02.011>.
- 634 (30) Deng, D.; Han, G.; Jiang, Y.; Ai, N. Solubilities of Carbon Dioxide in Five Biobased  
635 Solvents. *J. Chem. Eng. Data* 2015, 60 (1), 104–111.
- 636 (31) Yogish, K. Solubility of CO<sub>2</sub> in Some Physical Solvents. *J. Chem. Eng. Jpn.* 1991, 24 (1),  
637 135–137. <https://doi.org/10.1252/jcej.24.135>.
- 638 (32) Miller, M. B.; Chen, D.-L.; Luebke, D. R.; Johnson, J. K.; Enick, R. M. Critical Assessment  
639 of CO<sub>2</sub> Solubility in Volatile Solvents at 298.15 K. *J. Chem. Eng. Data* 2011, 56 (4), 1565–  
640 1572. <https://doi.org/10.1021/je101161d>.

- 641 (33) Hansen, C. M. *Hansen Solubility Parameters: A User's Handbook*, 2nd ed.; CRC Press:  
642 Boca Raton, 2007.
- 643 (34) Gennaro, A.; Isse, A. A.; Vianello, E. Solubility and Electrochemical Determination of  
644 CO<sub>2</sub> in Some Dipolar Aprotic Solvents. *J. Electroanal. Chem. Interfacial Electrochem.*  
645 1990, 289 (1–2), 203–215. [https://doi.org/10.1016/0022-0728\(90\)87217-8](https://doi.org/10.1016/0022-0728(90)87217-8).
- 646 (35) Anouti, M.; Dougassa, Y. R.; Tessier, C.; El Ouatani, L.; Jacquemin, J. Low Pressure  
647 Carbon Dioxide Solubility in Pure Electrolyte Solvents for Lithium-Ion Batteries as a  
648 Function of Temperature. Measurement and Prediction. *J. Chem. Thermodyn.* 2012, 50, 71–  
649 79. <https://doi.org/10.1016/j.jct.2012.01.027>.
- 650 (36) Li, Y.; You, Y.; Huang, W.; Yang, J. Solubility Measurement and Thermodynamic  
651 Properties Calculation for Several CO<sub>2</sub> + Ether Absorbent Systems. *J. Chem. Eng. Data*  
652 2019, 64 (3), 1020–1028. <https://doi.org/10.1021/acs.jced.8b00936>.
- 653 (37) Li, Y.; Zheng, D.; Dong, L.; Xiong, B. Solubilities of Carbon Dioxide in 2-Methoxyethyl  
654 Acetate, 1-Methoxy-2-Propyl Acetate and 3-Methoxybutyl Acetate. *J. Chem. Thermodyn.*  
655 2014, 74, 126–132. <https://doi.org/10.1016/j.jct.2014.01.019>.
- 656 (38) Li, Y.; Liu, Q.; Huang, W.; Yang, J. Below the Room Temperature Measurements of  
657 Solubilities in Ester Absorbents for CO<sub>2</sub> Capture. *J. Chem. Thermodyn.* 2018, 127, 71–79.  
658 <https://doi.org/10.1016/j.jct.2018.07.021>.
- 659 (39) Flowers, B. S.; Mittenthal, M. S.; Jenkins, A. H.; Wallace, D. A.; Whitley, J. W.; Dennis,  
660 G. P.; Wang, M.; Turner, C. H.; Emel'yanenko, V. N.; Verevkin, S. P.; Bara, J. E. 1,2,3-  
661 Trimethoxypropane: A Glycerol-Derived Physical Solvent for CO<sub>2</sub> Absorption. *ACS*  
662 *Sustain. Chem. Eng.* 2017, 5 (1), 911–921.  
663 <https://doi.org/10.1021/acssuschemeng.6b02231>.

- 664 (40) Li, Y.; Huang, W.; Zheng, D.; Mi, Y.; Dong, L. Solubilities of CO<sub>2</sub> Capture Absorbents 2-  
665 Ethoxyethyl Ether, 2-Butoxyethyl Acetate and 2-(2-Ethoxyethoxy)Ethyl Acetate. *Fluid*  
666 *Phase Equilibria* 2014, 370, 1–7. <https://doi.org/10.1016/j.fluid.2014.02.029>.
- 667 (41) Schappals, M.; Breug-Nissen, T.; Langenbach, K.; Burger, J.; Hasse, H. Solubility of  
668 Carbon Dioxide in Poly(Oxymethylene) Dimethyl Ethers. *J Chem Eng Data* 2017, 5.
- 669 (42) Gui, X.; Tang, Z.; Fei, W. Solubility of CO<sub>2</sub> in Alcohols, Glycols, Ethers, and Ketones at  
670 High Pressures from (288.15 to 318.15) K. *J. Chem. Eng. Data* 2011, 56 (5), 2420–2429.  
671 <https://doi.org/10.1021/je101344v>.
- 672 (43) Gui, X.; Wang, W.; Wang, C.; Zhang, L.; Yun, Z.; Tang, Z. Vapor–Liquid Phase  
673 Equilibrium Data of CO<sub>2</sub> in Some Physical Solvents from 285.19 K to 313.26 K. *J. Chem.*  
674 *Eng. Data* 2014, 59 (3), 844–849. <https://doi.org/10.1021/je400985u>.
- 675 (44) Jou, F.-Y.; Otto, F. D.; Mather, A. E. Solubility of H<sub>2</sub>S and CO<sub>2</sub> in Diethylene Glycol at  
676 Elevated Pressures. *Fluid Phase Equilibria* 2000, 175 (1), 53–61.  
677 [https://doi.org/10.1016/S0378-3812\(00\)00440-4](https://doi.org/10.1016/S0378-3812(00)00440-4).
- 678 (45) F. Blanchard; B. Carre; F. Bonhomme; P. Biensan; D. Lemordat. Solubility of Carbon  
679 Dioxide in Alkylcarbonates and Lactones. *Can. J. Chem.* No. 81, 385–391.
- 680 (46) Wu, F.; Zhao, Q.; Tao, L.; Danaci, D.; Xiao, P.; Hasan, F. A.; Webley, P. A. Solubility of  
681 Carbon Monoxide and Hydrogen in Methanol and Methyl Formate: 298–373 K and 0.3–3.3  
682 MPa. *J. Chem. Eng. Data* 2019, 64 (12), 5609–5621.  
683 <https://doi.org/10.1021/acs.jced.9b00676>.
- 684 (47) Qureshi, M. S.; Le Nedelec, T.; Guerrero-Amaya, H.; Uusi-Kyyny, P.; Richon, D.;  
685 Alopaeus, V. Solubility of Carbon Monoxide in Bio-Oil Compounds. *J. Chem. Thermodyn.*  
686 2017, 105, 296–311. <https://doi.org/10.1016/j.jct.2016.10.030>.

- 687 (48) Brunner, E. Solubility of Hydrogen in 10 Organic Solvents at 298.15, 323.15, and 373.15  
688 K. *J. Chem. Eng. Data* 1985, 30 (3), 269–273. <https://doi.org/10.1021/je00041a010>.
- 689 (49) Brunner, E. Solubility of Hydrogen in Diols and Their Ethers. *J. Chem. Thermodyn.* 1980,  
690 12 (10), 993–1002. [https://doi.org/10.1016/0021-9614\(80\)90140-8](https://doi.org/10.1016/0021-9614(80)90140-8).
- 691 (50) Purwanto; Deshpande, R. M.; Chaudhari, R. V.; Delmas, H. Solubility of Hydrogen,  
692 Carbon Monoxide, and 1-Octene in Various Solvents and Solvent Mixtures. *J. Chem. Eng.*  
693 *Data* 1996, 41 (6), 1414–1417. <https://doi.org/10.1021/je960024e>.
- 694 (51) Krüger, M. B.; Selle, C.; Heller, D.; Baumann, W. Determination of Gas Concentrations in  
695 Liquids by Nuclear Magnetic Resonance: Hydrogen in Organic Solvents. *J. Chem. Eng.*  
696 *Data* 2012, 57 (6), 1737–1744. <https://doi.org/10.1021/je2013582>.
- 697 (52) Henni, A.; Tontiwachwuthikul, P.; Chakma, A. Solubility Study of Methane and Ethane in  
698 Promising Physical Solvents for Natural Gas Sweetening Operations. *J. Chem. Eng. Data*  
699 2006, 51 (1), 64–67. <https://doi.org/10.1021/je050172h>.
- 700 (53) Hesse, P. J.; Battino, R.; Scharlin, P.; Wilhelm, E. Solubility of Gases in Liquids. 21.  
701 Solubility of He, Ne, Ar, Kr, N<sub>2</sub>, O<sub>2</sub>, CH<sub>4</sub>, CF<sub>4</sub>, and SF<sub>6</sub> in 2,2,4-Trimethylpentane At T=  
702 298.15 K. *J. Chem. Thermodyn.* 1999, 31 (9), 1175–1181.  
703 <https://doi.org/10.1006/jcht.1999.0529>.
- 704 (54) Battino, R.; Rettich, T. R.; Tominaga, T. The Solubility of Nitrogen and Air in Liquids. *J.*  
705 *Phys. Chem. Ref. Data* 1984, 13 (2), 563–600. <https://doi.org/10.1063/1.555713>.
- 706 (55) Bo, S.; Battino, R.; Wilhelm, E. Solubility of Gases in Liquids. 19. Solubility of He, Ne,  
707 Ar, Kr, Xe, N<sub>2</sub>, O<sub>2</sub>, CH<sub>4</sub>, CF<sub>4</sub>, and SF<sub>6</sub> in Normal 1-Alkanols n-C<sub>1</sub>H<sub>2</sub>l+1OH (1 .Lto req. 1  
708 .Lto req. 11) at 298.15 K. *J. Chem. Eng. Data* 1993, 38 (4), 611–616.  
709 <https://doi.org/10.1021/je00012a035>.

- 710 (56) Hesse, P. J.; Battino, R.; Scharlin, P.; Wilhelm, E. Solubility of Gases in Liquids. 20.  
711 Solubility of He, Ne, Ar, Kr, N<sub>2</sub>, O<sub>2</sub>, CH<sub>4</sub>, CF<sub>4</sub>, and SF<sub>6</sub> in n-Alkanes n-C<sub>1</sub>H<sub>2</sub>l+2 (6 ≤ l ≤  
712 16) at 298.15 K. *J. Chem. Eng. Data* 1996, 41 (2), 195–201.  
713 <https://doi.org/10.1021/je9502455>.
- 714 (57) Pardo, J.; López, M. C.; Santafé, J.; Royo, F. M.; Urieta, J. S. Solubility of Gases in  
715 Butanols. I. Solubilities of Nonpolar Gases in 1-Butanol from 263.15 to 303.15 K at 101.33  
716 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1995, 109 (1), 29–37.  
717 [https://doi.org/10.1016/0378-3812\(95\)02712-N](https://doi.org/10.1016/0378-3812(95)02712-N).
- 718 (58) Pardo, J.; López, M. C.; Mayoral, J. A.; Royo, F. M.; Urieta, J. S. Solubility of Gases in  
719 Butanols. III. Solubilities of Non-Polar Gases in 2-Butanol from 263.15 to 303.15 K at  
720 101.33 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1997, 134 (1–2), 133–140.  
721 [https://doi.org/10.1016/S0378-3812\(97\)00064-2](https://doi.org/10.1016/S0378-3812(97)00064-2).
- 722 (59) Pardo, J.; López, M. C.; Santafé, J.; Royo, F. M.; Urieta, J. S. Solubility of Gases in  
723 Butanols II. Solubilities of Nonpolar Gases in 2-Methyl-1-Propanol from 263.15 to 303.15  
724 K at 101.33 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1996, 119 (1), 165–173.  
725 [https://doi.org/10.1016/0378-3812\(95\)02984-2](https://doi.org/10.1016/0378-3812(95)02984-2).
- 726 (60) Pardo, J.; Mainar, A. M.; Lopez, M. C.; Royo, F.; Urieta, J. S. Solubility of Gases in  
727 Butanols IV. Solubilities of Nonpolar Gases in 2-Methyl-2-Propanol at 303.15 K and  
728 101.33 KPa Partial Pressure of Gas. *Fluid Ph. Equilibria*. 1999, 155 (1), 127-137.
- 729 (61) Weng, W.-L.; Chen, J.-T.; Chang, J.-S.; Chang, S.-L. Vapor–Liquid Equilibria for Nitrogen  
730 with 2-Hexanol, 2-Heptanol, or 2-Octanol Binary Systems. *Fluid Phase Equilibria* 2006,  
731 248 (2), 168–173. <https://doi.org/10.1016/j.fluid.2006.08.005>.

- 732 (62) Gallardo, M. A.; Melendo, J. M.; Urieta, J. S.; Losa, C. G. Solubility of Non-Polar Gases in  
733 Cyclohexanone between 273.15 and 303.15 K at 101.32 KPa Partial Pressure of Gas. *Can.*  
734 *J. Chem.* 1987, 65 (9), 2198–2202. <https://doi.org/10.1139/v87-368>.
- 735 (63) Gallardo, M. A.; López, M. C.; Urieta, J. S.; Losa, C. G. Solubility of He, Ne, Ar, Kr, Xe,  
736 H<sub>2</sub>, D<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, CF<sub>4</sub>, SF<sub>6</sub> and CO<sub>2</sub> in Cyclopentanone from 273.15 K  
737 to 303.15 K and Gas Partial Pressure of 101.33 KPa. *Fluid Phase Equilibria* 1989, 50 (1),  
738 223–233. [https://doi.org/10.1016/0378-3812\(89\)80292-4](https://doi.org/10.1016/0378-3812(89)80292-4).
- 739 (64) Gallardo, M.A.; López, M.C.; Urieta, J.S.; Gutierrez Losa, C. Solubility of 15 non-polar  
740 gases (He, Ne, Ar, Kr, Xe, H<sub>2</sub>, D<sub>2</sub>, N<sub>2</sub>, O<sub>2</sub>, CH<sub>4</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, CF<sub>4</sub>, SF<sub>6</sub> and CO<sub>2</sub>) in  
741 cycloheptanone. *Fluid Ph. Equilibria*. 1990, 58 (1–2), 1990, 159-172.
- 742 (65) Arai, C; Yoshitama, T.; Nishihara, K.; Sano, Y. Gas Solubilities in Esters of Oleic Acid.  
743 *Kagaku Kougaku Ronbunshu* 1989, 15 (6), 1193–1195.
- 744 (66) Lizano, L. P.; López, M. C.; Royo, F. M.; Urieta, J. S. Solubility of Non Polar Gases in  
745 Formaldehyde Diethyl Acetal Between -10 and 30°C, and 1 Atm Partial Pressure of Gas. *J.*  
746 *Solut. Chem.* 1990, 19 (7), 721–728. <https://doi.org/10.1007/BF00647390>.
- 747 (67) Urieta, J. S.; Gibanel, F.; Martínez-López, J. F.; Pardo, J. I.; Mainar, A. M. Solubilities of  
748 Gases in Cycloethers. The Solubility of 13 Nonpolar Gases in 2,5-Dimethyltetrahydrofuran  
749 at 273.15 to 303.15 K and 101.32 KPa. *J. Chem. Thermodyn.* 2019, 132, 306–315.  
750 <https://doi.org/10.1016/j.jct.2018.12.037>.
- 751 (68) Gibanel, F.; López, M. C.; Royo, F. M.; Rodríguez, V.; Urieta, J. S. Solubility of Nonpolar  
752 Gases in Tetrahydropyran at 0 to 30°C and 101.33 KPa Partial Pressure of Gas. *J. Solut.*  
753 *Chem.* 1994, 23 (11), 1247–1256. <https://doi.org/10.1007/BF00974033>.

- 754 (69) Gibanel, F.; López, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutiérrez Losa, C. Solubility of  
755 Nonpolar Gases in Hexamethylenoxide. *Fluid Phase Equilibria* 1988, 42, 261–268.  
756 [https://doi.org/10.1016/0378-3812\(88\)80063-3](https://doi.org/10.1016/0378-3812(88)80063-3).
- 757 (70) Mainar, A.M.; Pardo, J.; Royo, F.M.; Lopez, M.C.; Urieta, J.S. Solubility of nonpolar gases  
758 in 2,2,2-trifluoroethanol at 25C and 101.33 kPa partial pressure of gas. *J. Solution. Chem.*  
759 1996, 25 (6) , 589-595.
- 760 (71) Lopez, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutierrez Losa, C. Solubility of Nonpolar  
761 Gases in Halogenated Compounds. 1. Solubility of Hydrogen, Deuterium, Nitrogen,  
762 Oxygen, Methane, Ethylene, Ethane, Carbon Tetrafluoride, Sulfur Hexafluoride and Carbon  
763 Dioxide in Chlorocyclohexane at 263.15-303.15 K and 101.32 KPa of Partial Pressure of  
764 Gas. *J. Chem. Eng. Data* 1987, 32 (4), 472–474. <https://doi.org/10.1021/je00050a027>.
- 765 (72) Lopez, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutierrez Losa, C. Solubility of Nonpolar  
766 Gases in Halogenated Compounds. 2. Solubility of Hydrogen, Deuterium, Nitrogen,  
767 Oxygen, Methane, Ethylene, Ethane, Carbon, Tetrafluoride, Sulfur Hexafluoride and  
768 Carbon Dioxide in Bromocyclohexane at 263.15- to 303.15 K and 101.32 KPa Partial  
769 Pressure of Gas. *J. Chem. Eng. Data* 1989, 34 (2), 198–200.  
770 <https://doi.org/10.1021/je00056a015>.
- 771 (73) Nitta, T.; Nakamura, Y.; Ariyasu, H.; Katayama, T. Solubilities of Nitrogen in Binary  
772 Solutions of Acetone with Cyclohexane, Benzene, Chloroform and 2-Propanol. *J. Chem.*  
773 *Eng. Jpn.* 1980, 13 (2), 97–103. <https://doi.org/10.1252/jcej.13.97>.
- 774 (74) Akimoto, T.; Nitta, T.; Katayama, T. Nitrogen Solubility and Vapor Pressure of Binary  
775 Mixed Solvents Containing Benzene, Carbon Tetrachloride, Cyclohexane and 1-Hexane. *J.*  
776 *Chem. Eng. Jpn.* 1984, 17 (6), 637–641. <https://doi.org/10.1252/jcej.17.637>.



- 777 (75) Mainar, A. M.; Pardo, J.; García, J. I.; Royo, F. M.; Urieta, J. S. Solubility of Gases in  
778 Fluoroorganic Alcohols. *J. Chem. Soc. Faraday Trans.* 1998, 94 (24), 3595–3599.  
779 <https://doi.org/10.1039/A807488G>.
- 780 (76) Atlani, M.; Loutaty, R.; Wakselman, C.; Yacono, C. Method of Purifying a Gas Mixture  
781 Containing Undesirable Gas Compounds. 4504287, March 12, 1985.
- 782 (77) Barber, R.F.G; Ritter, T.J.; Sweeney, C.W. Removing Sulfur Compounds from Gases.  
783 2245889A, January 15, 1992.
- 784 (78) Gamsjäger, H.; Lorimer, J. W.; Salomon, M.; Shaw, D. G.; Tomkins, R. P. T. The IUPAC-  
785 NIST Solubility Data Series: A Guide to Preparation and Use of Compilations and  
786 Evaluations (IUPAC Technical Report). *Pure Appl. Chem.* 2010, 82 (5), 1137–1159.  
787 <https://doi.org/10.1351/PAC-REP-09-10-33>.
- 788 (79) KNIME. <https://www.knime.com/open-for-innovation-0> (accessed 2021-03 -03).
- 789 (80) Varnek, A.; Fourches, D.; Hoonakker, F.; Solov'ev, V. P. Substructural Fragments: An  
790 Universal Language to Encode Reactions, Molecular and Supramolecular Structures. *J.*  
791 *Comput. Aided Mol. Des.* 2005, 19 (9–10), 693–703. [https://doi.org/10.1007/s10822-005-](https://doi.org/10.1007/s10822-005-9008-0)  
792 9008-0.
- 793 (81) Varnek, A.; Fourches, D.; Horvath, D.; Klimchuk, O.; Gaudin, C.; Vayer, P.; Solov'ev, V.;  
794 Hoonakker, F.; Tetko, I.; Marcou, G. ISIDA - Platform for Virtual Screening Based on  
795 Fragment and Pharmacophoric Descriptors. *Curr. Comput. Aided-Drug Des.* 2008, 4 (3),  
796 191–198. <https://doi.org/10.2174/157340908785747465>.
- 797 (82) Spartan 18.0; Wavefunction, Inc.:[www.wavefun.com](http://www.wavefun.com).
- 798 (83) Breiman, L. Random Forests. *Mach. Learn.* 2001, 45 (1), 5–32.  
799 <https://doi.org/10.1023/A:1010933404324>.

- 800 (84) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel,  
801 M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.;  
802 Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *J.*  
803 *Mach. Learn. Res.* 2011, *12* (85), 2825–2830.
- 804 (85) <https://scikit-learn.org/stable/> (accessed 23.10.20).
- 805 (86) Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the*  
806 *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*;  
807 ACM: San Francisco California USA, 2016; pp 785–794.  
808 <https://doi.org/10.1145/2939672.2939785>.
- 809 (87) XGBoost, [https://xgboost.readthedocs.io/en/latest/python/python\\_intro.html](https://xgboost.readthedocs.io/en/latest/python/python_intro.html) (accessed  
810 31.05.21).
- 811 (88) Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* 1995, *20* (3), 273–297.  
812 <https://doi.org/10.1007/BF00994018>.
- 813 (89) Sahigara, F.; Mansouri, K.; Ballabio, D.; Mauri, A.; Consonni, V.; Todeschini, R.  
814 Comparison of Different Approaches to Define the Applicability Domain of QSAR Models.  
815 *Molecules* 2012, *17* (5), 4791–4810. <https://doi.org/10.3390/molecules17054791>.
- 816 (90) Descamps, C.; Coquelet, C.; Bouallou, C.; Richon, D. Solubility of Hydrogen in Methanol  
817 at Temperatures from 248.41 to 308.20K. *Thermochim. Acta* 2005, *430* (1), 1–7.  
818 <https://doi.org/10.1016/j.tca.2004.12.001>.
- 819 (91) Soubeyran, A.; Rouabhi, A.; Coquelet, C. Thermodynamic Analysis of Carbon Dioxide  
820 Storage in Salt Caverns to Improve the Power-to-Gas Process. *Appl. Energy* 2019, *242*,  
821 1090–1107. <https://doi.org/10.1016/j.apenergy.2019.03.102>.

- 822 (92) Dicko, M.; Coquelet, C.; Theveneau, P.; Mougin, P. Phase Equilibria of H<sub>2</sub>S-Hydrocarbons  
823 (Propane, n-Butane, and n-Pentane) Binary Systems at Low Temperatures *J. Chem. Eng.*  
824 *Data* 2012, *57*, 5, 1534–1543.
- 825 (93) Cadours, R.; Shah, V.; Weiss, C.; Roquet, D.; Lallemand, F. Industrial Operation of  
826 HySWEET®, a New Hybrid Solvent for Improved Mercaptan Removal. In *Proceedings of*  
827 *the 2nd Annual Gas Processing Symposium*; Elsevier, 2010; pp 221–228.  
828 [https://doi.org/10.1016/S1876-0147\(10\)02024-0](https://doi.org/10.1016/S1876-0147(10)02024-0).
- 829 (94) Skylogianni, E.; Wanderley, R. R.; Austad, S. S.; Knuutila, H. K. Density and Viscosity of  
830 the Nonaqueous and Aqueous Mixtures of Methyldiethanolamine and Monoethylene  
831 Glycol at Temperatures from 283.15 to 353.15 K. *J. Chem. Eng. Data* 2019, *64* (12), 5415–  
832 5431. <https://doi.org/10.1021/acs.jced.9b00607>.
- 833 (95) Heldebrant, D. J.; Koech, P. K.; Glezakou, V.-A.; Rousseau, R.; Malhotra, D.; Cantu, D. C.  
834 Water-Lean Solvents for Post-Combustion CO<sub>2</sub> Capture: Fundamentals, Uncertainties,  
835 Opportunities, and Outlook. *Chem. Rev.* 2017, *117* (14), 9594–9624.  
836 <https://doi.org/10.1021/acs.chemrev.6b00768>.
- 837 (96) Zheng, R. F.; Barpaga, D.; Mathias, P. M.; Malhotra, D.; Koech, P. K.; Jiang, Y.; Bhakta,  
838 M.; Lail, M.; V. Rayer, A.; Whyatt, G. A.; Freeman, C. J.; Zwoster, A. J.; Weitz, K. K.;  
839 Heldebrant, D. J. A Single-Component Water-Lean Post-Combustion CO<sub>2</sub> Capture Solvent  
840 with Exceptionally Low Operational Heat and Total Costs of Capture – Comprehensive  
841 Experimental and Theoretical Evaluation. *Energy Environ. Sci.* 2020, *13* (11), 4106–4113.  
842 <https://doi.org/10.1039/D0EE02585B>.

- 843 (97) Vahidi, M.; Shokouhi, M. Experimental Solubility of Carbon Dioxide and Hydrogen  
844 Sulfide in 2,2'-Thiodiglycol. *J. Chem. Thermodyn.* 2019, 133, 202–207.  
845 <https://doi.org/10.1016/j.jct.2019.02.024>.
- 846 (98) Versteeg, G. F.; van Swaaij, W. P. M. On the Kinetics between CO<sub>2</sub> and Alkanolamines  
847 Both in Aqueous and Non-Aqueous Solutions—II. Tertiary Amines. *Chem. Eng. Sci.* 1988,  
848 43 (3), 587–591. [https://doi.org/10.1016/0009-2509\(88\)87018-0](https://doi.org/10.1016/0009-2509(88)87018-0).
- 849 (99) Versteeg, G. F.; van Swaaij, W. P. M. On the Kinetics between CO<sub>2</sub> and Alkanolamines  
850 Both in Aqueous and Non-Aqueous Solutions—I. Primary and Secondary Amines. *Chem.*  
851 *Eng. Sci.* 1988, 43 (3), 573–585. [https://doi.org/10.1016/0009-2509\(88\)87017-9](https://doi.org/10.1016/0009-2509(88)87017-9).
- 852 (100) Maddox, R. N.; Mains, G. J.; Rahman, M. A. Reactions of Carbon Dioxide and Hydrogen  
853 Sulfide with Some Alkanolamines. *Ind. Eng. Chem. Res.* 1987, 26 (1), 27–31.  
854 <https://doi.org/10.1021/ie00061a006>.
- 855 (101) Rainbolt, J. E.; Koech, P. K.; Yonker, C. R.; Zheng, F.; Main, D.; Weaver, M. L.;  
856 Linehan, J. C.; Heldebrant, D. J. Anhydrous Tertiary Alkanolamines as Hybrid Chemical  
857 and Physical CO<sub>2</sub> Capture Reagents with Pressure-Swing Regeneration. *Energy Environ.*  
858 *Sci.* 2011, 4 (2), 480–484. <https://doi.org/10.1039/C0EE00506A>.
- 859 (102) Mathias, P. M.; Jasperson, L. V.; VonNiederhausern, D.; Bearden, M. D.; Koech, P. K.;  
860 Freeman, C. J.; Heldebrant, D. J. Assessing Anhydrous Tertiary Alkanolamines for High-  
861 Pressure Gas Purifications. *Ind. Eng. Chem. Res.* 2013, 52 (49), 17562–17572.  
862 <https://doi.org/10.1021/ie4020974>.
- 863 (103) Battino, R.; Clever, H. L. The Solubility of Gases in Liquids. *Chem. Rev.* 1966, 66 (4),  
864 395–463. <https://doi.org/10.1021/cr60242a003>.

- 865 (104) Pierotti, R. A. A Scaled Particle Theory of Aqueous and Nonaqueous Solutions. *Chem.*  
866 *Rev.* 1976, 76 (6), 717–726. <https://doi.org/10.1021/cr60304a002>.
- 867 (105) Abboud, J.-L. M.; Notari, R. Critical Compilation of Scales of Solvent Parameters. Part I.  
868 Pure, Non-Hydrogen Bond Donor Solvents. *Pure Appl. Chem.* 1999, 71 (4), 645–718.  
869 <https://doi.org/10.1351/pac199971040645>.
- 870 (106) Lewis, M.; Wu, Z.; Glaser, R. Polarizabilities of Carbon Dioxide and Carbodiimide.  
871 Assessment of Theoretical Model Dependencies on Dipole Polarizabilities and Dipole  
872 Polarizability Anisotropies. *J. Phys. Chem. A* 2000, 104 (48), 11355–11361.  
873 <https://doi.org/10.1021/jp002927r>.
- 874 (107) Zeng, W.; Du, Y.; Xue, Y.; Frisch, H. L. Solubility Parameters. In *Physical Properties of*  
875 *Polymers Handbook*; Mark, J. E., Ed.; Springer New York: New York, NY, 2007; pp 289–  
876 303. [https://doi.org/10.1007/978-0-387-69002-5\\_16](https://doi.org/10.1007/978-0-387-69002-5_16).
- 877 (108) Sigma-Aldrich Catalog, <https://www.sigmaaldrich.com/FR/En/Sds/Aldrich/M56557>  
878 (Accessed 11.08.21).
- 879 (109) Kauffman, G. W.; Jurs, P. C. Prediction of Surface Tension, Viscosity, and Thermal  
880 Conductivity for Common Organic Solvents Using Quantitative Structure–Property  
881 Relationships. *J. Chem. Inf. Comput. Sci.* 2001, 41 (2), 408–418.  
882 <https://doi.org/10.1021/ci000139t>.
- 883 (110) Singh, A. K.; Bilal, M.; Iqbal, H. M. N.; Raj, A. Trends in Predictive Biodegradation for  
884 Sustainable Mitigation of Environmental Pollutants: Recent Progress and Future Outlook.  
885 *Sci. Total Environ.* 2021, 770, 144561. <https://doi.org/10.1016/j.scitotenv.2020.144561>.  
886