The Effect of Solvent Polarity on Autocatalytic Furfural Production Confirmed by Multivariate Statistical Analysis


Autocatalytic dehydration of xylose to furfural was studied in pure aqueous and monophasic organic/water mixtures to determine the effect reaction media and conditions have on conversion and yield. This study identified that the severity ($R_s$) of the reaction and polarity, as determined by the Hansen Solubility Parameter, $\delta_p$, strongly correlate with xylose conversion and furfural yield. Increasing the $R_s$ and $\delta_p$ increased both conversion and yield in pure aqueous and organic/water mixtures of sulfolane, $\gamma$-butyrolactone, $\gamma$-valerolactone, $\gamma$-hexalactone, and tetrahydrofuran. Additionally, it was found that at a specified $R_s$ and $\delta_p$ similar conversions and yields were achieved using different combinations of time, temperature, and solvent mixture. Using principal component analysis and projection to latent structures, a semi-empirical model was developed that provided estimates of xylose conversion and furfural yield over a range of experimental $R_s$ and $\delta_p$ values.

Furfural has been recognized as an important building block for fuels, solvents, and other value-added chemicals,[11] and numerous recent publications have focused on optimizing furfural production from model sugars and lignocellulosic biomass. The use of organic solvents for xylose dehydration reactions has recently gained interest due to faster reaction rates[2] and the higher furfural yields obtained, which have been reported as high as 80% in biphasic systems.[3] Solvent selection has been identified as an important reaction variable with mentions of solvent polarity,[4,5] and the number of oxygen containing groups[6] being potential explanations as to why solvents improve furan yields, but currently no systematic method for solvent selection exists. Researchers have previously used Kamlet-Taft theory[6] or Hansen Solubility Parameters (HSPs)[7] to explain differences in selectivities, conversions, and yields; however, the HSP parameters (dispersion ($\delta_d$), polarity ($\delta_p$), and hydrogen bonding ($\delta_h$)) are functions of both solvent mixture and temperature,[8] which often times are not considered.

This work aims to bridge the gap between xylose conversion, furfural yield, and solvent properties to guide solvent selection in monophasic, autocatalytic systems. Previous research has explored acid-catalyzed reactions in organic solvent mixtures for both monophasic and biphasic systems.[9] In biphasic systems using aluminium sulfate as the catalyst, Yang et al. hypothesized that as the polarity of a solvent increases, furfural yields increase in $\gamma$-valerolactone (GVL), methyl isobutyl ketone, tetrahydrofuran (THF), and 2-methyltetrahydrofuran;[10] however, no direct measure of solvent polarity was provided.

Although acid catalysts increase the rate of reaction, furans can be produced through an autocatalytic process (Scheme 1), with the mechanism initiated by the formation of hydronium ions in high temperature water.[11] Additionally, formic acid, a degradation product, has been shown to contribute to the reaction activity.[12] Previous publications on autocatalytic furfural production have focused on pure aqueous systems,[12a,13] and although organic solvent mixtures have been explored for autocatalytic 5-hydroxymethylfurfural production,[14] to our knowledge, the present work is the first that studied monophasic, organic solvent mixtures for autocatalytic furfural production.

The dehydration of xylose to furfural is a function of both time and temperature, which can be combined into a single variable, severity ($R_s$), defined as [Eq. (1)]:

$$R_s = \log_{10} \left[ \frac{t_{\exp} \exp \left( \frac{T_{\exp} - 100}{14.75} \right)}{100} \right]$$

(1)

where $T_{\exp}$ is the reaction temperature [°C] and $t_{\exp}$ is the reaction time [min].[15] Herein, the effect $R_s$ and $\delta_p$ have on xylose conversion and furfural yield is reported for pure aqueous and organic/water solvent mixtures of either sulfolane,
\( \gamma \)-butyrolactone (GBL), GVL, \( \gamma \)-hexalactone (GHL), or THF in autocatalytic reactions. Using principal components analysis (PCA) and projection to latent structures (PLS), the strength of the correlations between \( R_0 \) and \( \delta_P \) to xylose conversion and furfural yields were determined and a semi-empirical model was developed that estimates conversion and yield within a range of \( R_0 \) and \( \delta_P \).

At a temperature of 463 K, increasing the reaction time, which increases \( R_0 \) increased both xylose conversion (Figure 1A) and furfural yield (Figure 1B) for most solvent mixtures. The maximum furfural yields of the 50/50 w/w solvent mixtures ranged from 51 % for THF (12 h; \( R_0 = 5.5 \)) to 61 % for sulfolane (5.5 h; \( R_0 = 5.2 \)) at 96 % and 93 % xylose conversion, respectively. These results were comparable to the pure aqueous system (59 % furfural yield at 95 % xylose conversion; Figure S1), which was a higher yield than previously reported.[9a,16] Unsurprisingly, in the pure aqueous system at severities near 100 % xylose conversion (\( R_0 > 5.45 \); Figure S1), furfural yields decreased due to the rate of furfural degradation exceeding the rate of furfural formation, which was also seen for sulfolane at 12 h (\( R_0 = 5.5 \); Figure 1B).

Since \( R_0 \) is a function of time and temperature, as time increases, \( R_0 \) increases at a given temperature as shown in Figure 1. Xylose conversion and furfural yield typically increased as \( R_0 \) increased, with sulfolane and THF achieving the highest and lowest values, respectively, at most severities. Correlation plots of kinetic parameters (temperature, time, and \( R_0 \)) and solvent properties (\( \delta_P \), \( \delta_H \), and \( \delta_D \)) showed that both xylose conversion and furfural yield positively correlated with \( R_0 \) at the 5 % significance level (Figure S2). This result affirms the validity of \( R_0 \) to represent the reaction conditions (time and temperature). Additional affirmation was found at an \( R_0 \) of 5.2, where similar conversion and yields were achieved for different solvent systems at varying combinations of time (21 h and 5.5 h) and temperature (443 K and 463 K; Figure 2, blue and orange points). Likewise, the pure aqueous system showed similar conversions and yields at an \( R_0 \) of 5.0 using different combinations of time and temperature (Table S1). Interestingly, a relationship between \( \delta_P \) and xylose conversion or furfural yield can be seen on Figure 1, where there is indeed a clear

Figure 1. Effect of increased reaction time on A) conversion and B) yield in 50/50 w/w solvent systems (463 K; 2 wt% xylose) with sulfolane (△), GBL (○), GVL (●), GHL (□), and THF (○).

Figure 2. Effect of \( R_0 \) and \( \delta_P \) on A) conversion and B) yield at 443 K (orange) and 463 K (all others) in water (■) and 50/50 w/w solvent systems of sulfolane (△), GBL (○), GVL (●), GHL (□), and THF (○). Best fit lines are shown for severities of 5.5 (red), 5.2 (gray), and 4.9 (black).
trend that as the $\delta_R$ increased (Table S2), the conversion and yield generally increased at a given $R_o$ (sulfolane > GBL > GVL \approx GHL > THF). Regardless of solvent type, as $\delta_R$ increased, xylose conversion and furfural yield trended upward for all severities (Figure 2); however, at an $R_o$ of 5.5, the dependence on $\delta_R$ decreased, likely due to the nearly complete conversion of xylose (Figure 2A). The correlation between $\delta_R$ and furfural yield may be explained by the proposed mechanism for xylose to furfural in which a carbocation is formed and could be stabilized by a polar solvent through H-bond formation.\[17\]

The presence of formic acid as a degradation product (Scheme 1) was observed in the pure aqueous and most monophasic organic/water systems at xylose conversions above 30%. However, the formic acid yields tended to be less than 10% and there was no significant correlation between formic acid and xylose conversion or furfural yield at the 5% significance level. This indicates the autocatalytic reaction mechanism has little contribution from the formic acid.

Since the $\delta_R$ for mixtures is a function of volume percent, it can be tuned within the range of the $\delta_R$ of water and the $\delta_R$ of the pure solvent (Table S2). Observing all of the solvents collectively (Figure S3), increasing the $\delta_R$ of the organic/water mixture by changing the solvent ratio resulted in a nonlinear trend in both xylose conversion and furfural yield. For mixtures with a pure solvent $\delta_R$ greater than that of water, such as GBL and sulfolane, increasing the weight percent of organic solvent increases the $\delta_R$ of the system, which led to increased conversions and yields. In the sulfolane/water system, increasing $\delta_R$ with a 70/30 w/w mixture led to a maximum furfural yield of 67% at 97% xylose conversion and an $R_o$ of 4.9 (Table S3). Further increases in $R_o$ led to decreased furfural yields since the rate of furfural degradation exceeded production near 100% xylose conversion. Interestingly, GVL (14.8) has a pure solvent $\delta_R$ near that of water (14.3), limiting the possible $\delta_R$ range to 14.37–14.63 at 463 K for the mixtures used in this study. As expected, xylose conversions and furfural yields were limited (48–66% and 29–34%, respectively, $R_o=4.9$) for mixtures ranging from 20/80 to 70/30 w/w GBL/water.

In contrast, for organic solvents that have a pure solvent $\delta_R$ less than the $\delta_R$ of water, as the percent organic solvent in the mixture increases, $\delta_R$ decreases, resulting in lower conversions and yields. For example, in THF/water mixtures increasing the solvent from 20/80 to 70/30 w/w decreased xylose conversion from 52 to 42% and furfural yield from 30 to 4%, respectively. Similar trends were seen for furfural yields in GVL/water mixtures (Table S3), but this decrease in yield with increasing solvent percent differs from what has been reported previously for monophasic systems using an acid catalyst.\[18\] This opposing trend may be explained by the large effect pH has on severity,\[19\] which could minimize the effect of $\delta_R$ or the effect the acid has on $\delta_R$, which is not easily accounted for.

To determine the contribution that $R_o$ and $\delta_R$ each have on xylose conversion and furfural yield, multivariate statistical modeling (MSM) methods, PCA and PLS,\[19\] were used to account for collinearity and correlation among the variables. To maximize the usefulness of the MSM methods, the data were limited to 11.89 \leq \delta_R \leq 15.11 and 4.9 \leq R_o \leq 5.33 to avoid the highly nonlinear region of the data (Figure S3). Additionally, to generate a more robust model, a data set referred to as the calibration data was created that included the experimental data supplemented with random variability from normal distributions defined by the known experimental errors (Supplementary 1.2). It is noteworthy that the coefficients of the model used to estimate the dependent variables were identified simultaneously, which demonstrates the power of the PLS method to fit the correlated and collinear data.

Although multiple combinations of variables were considered (Table S5), the PLS analysis using $R_o$ and $\delta_R$ as the regressor variables demonstrated the greatest correlation to the dependent variables (xylose conversion and furfural yield), as evidenced by the higher amount of variance accounted for (85.47%; Table S6). Almost all of the estimates of the model data (Equation S8) fell within the 95% prediction interval (Figure 3, black points); however, some of the estimated points were outside the experimental error (±6% for xylose conversion and ±4% for furfural yield). Near the boundaries of the model range, the estimates of the model data were less accurate, likely...
due to nonlinear correlations unaccounted for by the PLS method (Figure S3). Importantly, the multivariate statistical analyses confirmed that over a range of $R_o$ and $\delta_o$, xylose conversion and furfural yield can be estimated within experimental error in most cases.

Given the regressor variables, $R_o$ (Equation 1) and $\delta_o$, the model is an implicit function of time, temperature, and solvent type in its estimates of xylose conversion and furfural yield. A key finding supported by the analyses, is that given a specified $R_o$ and $\delta_o$, the same xylose conversion and furfural yield can be achieved within experimental error regardless of the organic/solvent mixture (Table 1). This important discovery must be interpreted with caution, however, as the experimental furfural yield in the 20/80 w/w THF/water mixture was approximately 10% higher than all other solvent mixtures at a similar $\delta_o$ (Figure S3; $\delta_o = 11.93$; $R_o = 4.9$) and the xylose conversion of GHL (Table 1; $\delta_o = 12.61$; $R_o = 5.2$) was lower than model estimates. Since these experimental data were confirmed with multiple experiments, discrepancies may be the result of an inexact pure solvent $\delta_o$ value given by the HSPIP program (see the Supporting Information 1.1), repeated experimental errors, or an unaccounted for difference in the reaction mechanism.

The prediction ability of the regression model (Equation S8) was tested using validation data not included in the calibration data (Figure 3, pink points; Table S7). The model’s estimates of xylose conversion and furfural yield for the validation data set were estimated within the 95% prediction interval; however, validation points containing 50 wt% or more organic solvent were underestimated for conversion. The underestimation may be explained in part due to increased rate of reactions as Mellmer et al. reported when using GVL.[1] Thus, the accuracy of the model estimates for higher weight percent solvent mixtures may be improved if reaction kinetics were taken into account.

Even with the limited number of solvents explored (sulfolane, GBL, water, GVL, GHL, and THF), the regression model (Equation S8) successfully estimated xylose conversion and furfural yield for acetone (Figure 3 and Table S7), which was not included in the calibration data. Remarkably, even when considering another new solvent, 1-butanol (BuOH), the multivariate statistical analyses confirmed experimental findings that for values of $R_o$ and $\delta_o$ that are within the range of the model’s calibration data, similar values of xylose conversions and furfural yield (within experimental error) will result (Table 1). The success of the regression model in confirming the significance of $R_o$ and $\delta_o$ should be tempered by the obvious limitation that the range of data used to develop the model exhibited a high degree of linearity. For example, at the lower bounds of $\delta_o$, although the model accurately estimates the conversion and yield of the 50/50 w/w GHL/water mixture, it does not accurately estimate the 70/30 w/w GVL/water mixture (Table S7). Additionally, the model is not valid above an $R_o$ of 5.3, where most solvents achieve 100% xylose conversion and furfural degradation may occur.

In summary, the effect of $R_o$ and solvent parameters was studied to determine the relationship they have on xylose conversion and furfural yield. Increasing $R_o$ and $\delta_o$, regardless of the time and temperature combination or organic solvent, increased xylose conversion and furfural yield. The availability of HSP data allowed for easy solvent selection based on $\delta_o$ and reaction conditions can be at tailored for temperature and time restrictions using $R_o$. Correlations and collinearity among these variables were confirmed by multivariate statistical analyses enabling a more systematic way for choosing experimental conditions to achieve high xylose conversions and furfural yields in xylose to furfural autocatalytic reactions.

### Table 1. Experimental and estimated values of conversion and yield for different solvent blends with similar HSP polarity values. Reaction conditions: $R_o = 5.2$, 2 wt% xylose.

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Organic solvent percent [wt%]</th>
<th>HSP polarity [(\delta)]</th>
<th>Xylose conversion [%]</th>
<th>Furfural yield [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>THF</td>
<td>17</td>
<td>12.26</td>
<td>76</td>
<td>72</td>
</tr>
<tr>
<td>GVL</td>
<td>50</td>
<td>12.61</td>
<td>74</td>
<td>73</td>
</tr>
<tr>
<td>GHL</td>
<td>35</td>
<td>12.61</td>
<td>57</td>
<td>73</td>
</tr>
<tr>
<td>BuOH</td>
<td>15</td>
<td>12.65</td>
<td>67</td>
<td>73</td>
</tr>
</tbody>
</table>

### Acknowledgements
This work was supported in part by National Science Foundation – CBET – Energy for Sustainability award number 1705490 and we thank Prof. Steven Abbott for his valuable discussions.

### Conflict of Interest
The authors declare no conflict of interest.

### Keywords:
Autocatalysis · Furfural · Hansen Solubility Parameters · Multivariate statistical modelling · Sustainable chemistry


