Comparison of ANFIS and statistical modeling for estimation of chemical oxygen demand parameter in textile effluent

S. Akilandeswari\textsuperscript{1}\ast and B. Kavitha\textsuperscript{2}

\textsuperscript{1}Department of Physics, Annamalai University, Annamalai Nagar, Tamil Nadu, India
\textsuperscript{2}Department of Engineering Physics, Annamalai University, Annamalai Nagar, Tamil Nadu, India

ABSTRACT

The large volumes of wastewater generated at different stages of textile processing. The wastewater derived from the textile industry can caused serious environmental impact in the neighboring water bodies, because of the presence of toxic chemical residues and enhances the value of chemical oxygen demand (COD) of water. COD is one of the major parameter used to find the quality of wastewater. In this paper prediction of COD from textile effluent according to their physicochemical parameters such as pH, Total Suspended Solids (TSS), Sulphate (SO\textsubscript{4}), Chloride (Cl\textsubscript{2}) and Total Dissolved Solids (TDS) have been determined by Adaptive Neuro Fuzzy Inference System (ANFIS) modeling and Statistical modeling. The results compared by calculating Average percentage error, Chi-Squared test and Worst Case Error.

Key words: Textile Industry, environmental impact, COD.

INTRODUCTION

The rapid industrialization is accompanied by both direct and adverse effect on environment. The rate of contamination of natural water bodies increases with increased industrialization [1, 2]. The textile industry is one of the important industrial sectors of India based on earnings foreign exchange and labour employment. The chemical reagents used in textile industries are very diverse in chemical composition, ranging from in-organic compounds to polymers and organic products [3]. The most common textile-processing sector consists of de-sizing, scouring, bleaching, mercerizing and dyeing processes. The dyeing is the process of adding colours to fibers, which normally requires large volume of water not only in the dye path, but also during the rinsing sector. Depending on the dyeing process, many chemicals like metals, salts, surfactants, organic processing assistants, sulphide and formaldehyde may be added to improve dye absorption on the fibers. Waste water generated by different production steps of a textile mill have high pH, suspended and dissolved solids, dispersants, leveling agents, toxic and non biodegradable matter, colour and alkalinity. Wastewater from fabric and yarn printing and dyeing produces a serious environmental problem [4-7]. Because of these characteristics, wastewater from textile industry must be treated before discharged into natural water system [8-10]. Chemical oxygen demand parameter used to determine the quality of effluent. Many methods are available to predict the value of COD for effluent with five independent parameters such as pH, TSS, SO\textsubscript{4}, Cl\textsubscript{2} and TDS. Statistical method like linear are multiple regression analysis are usual procedures adopted to calculate the COD with some determining parameters [11]. However, these methods fail to calculate the accurate value of COD. The literature survey showed that for water and waste water treatment process, most of the artificial intelligence (AI) based prediction model were introduce to estimate the value of COD.
COD. Among these AI-based prediction methods adaptive neuro fuzzy inference system (ANFIS) have recently become a popular universal approximator that represent high non linear function. ANFIS is an adaptive network, which permits the usage of neural network topology together with the fuzzy logic. In fuzzy section, only zero or first order Sugeno inference system or Tsukamoto inference system can be used [12,13]. Even if the targets are not given, ANFIS may reach their accurate result rapidly. Models performance evaluated by sufficiently fitted training and testing data. Moreover model performance evaluate error values such as Training Root Mean Square Error (Trn RMSE) which are in term minimized by back propagation and hybrid learning algorithm allowed by ANFIS. The main objective of the study is to compare the predictive ability of ANFIS modeling and statistic modeling for the estimation of COD in the textile effluent.

MATERIALS AND METHODS

Wastewater samples from textile industry collected with the help of clean plastic container for physico chemical analysis over a period of one year. The collected samples were brought to the laboratory and stored at 4° C temperature. Selected parameters such as pH, TSS, SO₄, Cl, TDS, BOD and COD in the wastewater analyzed as per standard procedure [14].

ANFIS modeling: ANFIS combines both neural network and fuzzy logic; it is capable of handling complex and non-linear problems. Operation of ANFIS looks like feed-forward back-propagation network. Consequent parameters calculated forward while premise parameters are backward. Even if the targets are not given, ANFIS may reach the optimum result rapidly. There are two learning methods in neural section of the system: Hybrid learning method and Back Propagation learning method. The architecture of ANFIS consists of five layers and the number of neurons in each layer equals the number of rules. In ANFIS interpretation, the Sugeno inference system be used [15, 16]. In addition there is no vagueness in ANFIS is opposed to neural network [17]. In this study, the program has written in command window performed with the help of Matlab version (5.3) [18]. Out of 40 data obtained from textile effluent, it has split into 30 training sets and 10 checking sets. The training data sets presented in table.1. The training process has completed when the Training Root Mean Square Error (Trn RMSE), Check Root Mean Square Error (Chk RMSE) is minimum and the radius, epochs values assigned. The prediction of COD compared with the observed value by evaluating the Average Percentage Error (APE) using the relation,

\[ APE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{COD_{\text{obs}} - COD_{\text{pred}}}{COD_{\text{obs}}} \right| \times 100\% \]

Where n is the number of data pairs, COD(obs) represents observed values of COD, COD(pred) represents predicted values of COD, and i=1,2,3,…….n.

Statistical modeling: The data obtained from the textile industry subjected to statistical analysis. The present work determines the COD, using five independent parameters such as pH, TSS, SO₄, Cl, and TDS. The multiple regression analyses used to calculate the value of COD by statistical modeling, because it has more input. The multiple regression equation of the type is \( y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \) used. Where y is independent parameter, \( x_1, x_2 \) are independent parameters and \( \beta_0, \beta_1, \beta_2 \) etc, are regression coefficients. The prediction of COD in the textile effluent by ANFIS and statistical modeling, compared with observed value using Chi-Squared value (\( \chi^2 \)). The formula for calculating Chi-Square value (\( \chi^2 \)) is

\[ \chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \]

Where \( O_i \) is observed value of COD and \( E_i \) is the predicted value of COD. The statistical analyses are performed with the help of SPSS 13.0. The checked data are used in ANFIS, the COD are predicted by statistical modeling using the same data are represented in Table.2, and their corresponding APE(%), WE and Chi Square Value (\( \chi^2 \)) are given in table2.
Table 1: Effluent characteristic parameters used as ‘Training Data Set’ in the present work:

<table>
<thead>
<tr>
<th>Sample No</th>
<th>pH  (mg/l)</th>
<th>TSS  (mg/l)</th>
<th>SO$_4$  (mg/l)</th>
<th>Cl$_2$  (mg/l)</th>
<th>TDS  (mg/l)</th>
<th>COD  (mg/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.69</td>
<td>109</td>
<td>85</td>
<td>380</td>
<td>1424</td>
<td>140</td>
</tr>
<tr>
<td>2</td>
<td>8.59</td>
<td>157</td>
<td>168</td>
<td>1525</td>
<td>4244</td>
<td>140</td>
</tr>
<tr>
<td>3</td>
<td>8.2</td>
<td>128</td>
<td>120</td>
<td>1100</td>
<td>3252</td>
<td>139</td>
</tr>
<tr>
<td>4</td>
<td>7.4</td>
<td>200</td>
<td>114</td>
<td>1050</td>
<td>3060</td>
<td>140</td>
</tr>
<tr>
<td>5</td>
<td>6.52</td>
<td>60</td>
<td>347</td>
<td>285</td>
<td>1416</td>
<td>56</td>
</tr>
<tr>
<td>6</td>
<td>7.5</td>
<td>99</td>
<td>190</td>
<td>433</td>
<td>1800</td>
<td>130</td>
</tr>
<tr>
<td>7</td>
<td>7.11</td>
<td>160</td>
<td>299</td>
<td>1799</td>
<td>3776</td>
<td>137</td>
</tr>
<tr>
<td>8</td>
<td>6.92</td>
<td>100</td>
<td>145</td>
<td>890</td>
<td>2432</td>
<td>96</td>
</tr>
<tr>
<td>9</td>
<td>7.03</td>
<td>84</td>
<td>137</td>
<td>1924</td>
<td>3536</td>
<td>185</td>
</tr>
<tr>
<td>10</td>
<td>6.42</td>
<td>168</td>
<td>249</td>
<td>1799</td>
<td>4256</td>
<td>131</td>
</tr>
<tr>
<td>11</td>
<td>7.55</td>
<td>85</td>
<td>30</td>
<td>1650</td>
<td>1650</td>
<td>145</td>
</tr>
<tr>
<td>12</td>
<td>6.77</td>
<td>98</td>
<td>120</td>
<td>1175</td>
<td>2196</td>
<td>140</td>
</tr>
<tr>
<td>13</td>
<td>8.83</td>
<td>92</td>
<td>224</td>
<td>1350</td>
<td>2324</td>
<td>56</td>
</tr>
<tr>
<td>14</td>
<td>6.77</td>
<td>68</td>
<td>240</td>
<td>120</td>
<td>1952</td>
<td>89</td>
</tr>
<tr>
<td>15</td>
<td>7.72</td>
<td>208</td>
<td>156</td>
<td>1749</td>
<td>4244</td>
<td>151</td>
</tr>
<tr>
<td>16</td>
<td>6.75</td>
<td>100</td>
<td>250</td>
<td>200</td>
<td>2700</td>
<td>108</td>
</tr>
<tr>
<td>17</td>
<td>7.1</td>
<td>76</td>
<td>276</td>
<td>835</td>
<td>1840</td>
<td>64</td>
</tr>
<tr>
<td>18</td>
<td>6.41</td>
<td>12</td>
<td>134</td>
<td>145</td>
<td>568</td>
<td>82</td>
</tr>
<tr>
<td>19</td>
<td>7.31</td>
<td>40</td>
<td>149</td>
<td>145</td>
<td>640</td>
<td>51</td>
</tr>
<tr>
<td>20</td>
<td>6.84</td>
<td>52</td>
<td>127</td>
<td>690</td>
<td>1624</td>
<td>94</td>
</tr>
<tr>
<td>21</td>
<td>7.32</td>
<td>104</td>
<td>86</td>
<td>800</td>
<td>1972</td>
<td>152</td>
</tr>
<tr>
<td>22</td>
<td>7.18</td>
<td>48</td>
<td>108</td>
<td>565</td>
<td>1760</td>
<td>160</td>
</tr>
<tr>
<td>23</td>
<td>7.11</td>
<td>40</td>
<td>60</td>
<td>450</td>
<td>1184</td>
<td>152</td>
</tr>
<tr>
<td>24</td>
<td>6.84</td>
<td>48</td>
<td>121</td>
<td>690</td>
<td>1624</td>
<td>94</td>
</tr>
<tr>
<td>25</td>
<td>7.58</td>
<td>272</td>
<td>218</td>
<td>124</td>
<td>3196</td>
<td>86</td>
</tr>
<tr>
<td>26</td>
<td>7.18</td>
<td>16</td>
<td>154</td>
<td>739</td>
<td>1344</td>
<td>58</td>
</tr>
<tr>
<td>27</td>
<td>7.12</td>
<td>60</td>
<td>113</td>
<td>256</td>
<td>1620</td>
<td>104</td>
</tr>
<tr>
<td>28</td>
<td>7.14</td>
<td>40</td>
<td>223</td>
<td>916</td>
<td>1556</td>
<td>168</td>
</tr>
<tr>
<td>29</td>
<td>7.64</td>
<td>12</td>
<td>134</td>
<td>532</td>
<td>1364</td>
<td>108</td>
</tr>
<tr>
<td>30</td>
<td>6.82</td>
<td>70</td>
<td>329</td>
<td>530</td>
<td>1752</td>
<td>152</td>
</tr>
</tbody>
</table>

Table 2: Effluent characteristic parameters used as ‘Check Data’, the observed and predicted values of COD:

<table>
<thead>
<tr>
<th>Sample No</th>
<th>pH  (mg/l)</th>
<th>TSS  (mg/l)</th>
<th>SO$_4$  (mg/l)</th>
<th>Cl$_2$  (mg/l)</th>
<th>TDS  (mg/l)</th>
<th>COD  (mg/l)</th>
<th>Observed</th>
<th>ANFIS predicted</th>
<th>Statistical model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8.22</td>
<td>100</td>
<td>95</td>
<td>313</td>
<td>3958</td>
<td>156</td>
<td>155.6352</td>
<td>161.78</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7.51</td>
<td>88</td>
<td>186</td>
<td>437</td>
<td>1802</td>
<td>131</td>
<td>130.1929</td>
<td>138.27</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7.52</td>
<td>82</td>
<td>72</td>
<td>140</td>
<td>1638</td>
<td>144</td>
<td>143.9770</td>
<td>123.47</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7.01</td>
<td>121</td>
<td>278</td>
<td>960</td>
<td>3545</td>
<td>141</td>
<td>140.1285</td>
<td>155.02</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.45</td>
<td>108</td>
<td>306</td>
<td>150</td>
<td>2222</td>
<td>101</td>
<td>101.296</td>
<td>122.49</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6.86</td>
<td>114</td>
<td>420</td>
<td>448</td>
<td>1920</td>
<td>164</td>
<td>163.8026</td>
<td>145.95</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7.34</td>
<td>106</td>
<td>82</td>
<td>895</td>
<td>3695</td>
<td>157</td>
<td>156.3159</td>
<td>146.05</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>6.29</td>
<td>108</td>
<td>259</td>
<td>630</td>
<td>3510</td>
<td>136</td>
<td>135.0848</td>
<td>127.90</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6.85</td>
<td>50</td>
<td>349</td>
<td>520</td>
<td>3592</td>
<td>145</td>
<td>144.6686</td>
<td>134.55</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>7.38</td>
<td>24</td>
<td>135</td>
<td>410</td>
<td>382</td>
<td>102</td>
<td>101.8079</td>
<td>113.41</td>
<td></td>
</tr>
</tbody>
</table>

RESULTS AND DISCUSSION

In this paper, adaptive neuro fuzzy inference system and statistical modeling for the estimation of chemical oxygen demand of textile effluents compared. From table 2, it can be evident that the predictive ability of COD by ANFIS modeling agreed well with observed value because Chi-Square value ($\chi^2$) and Average Percentage Error (APE) value are low for ANFIS method. Fig.1 shows that the plot between observed values of COD and predicted values of COD from ANFIS modeling and statistical modeling. The fig.1 illustrates that predicted values from ANFIS is closer to the actual values than that from statistical modeling. It is believed that with more training data sets, best results could be achieved by ANFIS modeling.
CONCLUSION

The overall results indicated that ANFIS provided higher accuracy for the prediction of COD than statistical modeling. ANFIS is a valuable method for the determination of COD. Because it combines the advantages of both neural network and fuzzy logic which offers good results.

REFERENCES