ESTIMATE OF THE HEAD PRODUCED BY ELECTRICAL SUBMERSIBLE PUMPS ON GASEOUS PETROLEUM FLUIDS, A RADIAL BASIS FUNCTION NETWORK APPROACH

Morteza Mohammadzaheri\textsuperscript{1}, Mojataba Ghodsi\textsuperscript{1} and Abdullah AlQallaf\textsuperscript{2}

\textsuperscript{1}Department of Mechanical and Industrial Engineering, Sultan Qaboos University, Muscat, Oman
\textsuperscript{2}Department of Electrical Engineering, Kuwait University, Kuwait City, Kuwait

\textbf{ABSTRACT}

This paper reports successful development of an exact and an efficient radial basis function network (RBFN) model to estimate the head of gaseous petroleum fluids (GPFs) in electrical submersible pumps (ESPs). Head of GPFs in ESPs is now often estimated using empirical models. Overfitting and its consequent lack of model generality data is a potentially serious issue. In addition, available data series is fairly small, including the results of 110 experiments. All these limits were considered in RBFN design process, and highly accurate RBFNs were developed and cross validated.

\textbf{KEYWORDS}

Electrical Submersible Pump(ESP), Radial Basis Function Network (RBFN), Model, Petroleum, Gaseous, Head Estimation

1. \textbf{INTRODUCTION}

ESPs are widely used to lift large volume of fluid from downhole at different well conditions \cite{1,2}. As a critical stage at design of offshore petroleum systems, size of ESPs should be chosen correctly as over- or under-sizing leads to premature equipment failure or low petroleum fluid recovery. In order to facilitate size selection, manufacturers normally provide curves depicting generated head versus liquid volumetric flow rate for each ESP size. The correct choice is the smallest ESP which can generate sufficient head to lift petroleum fluids within the designed range of flow rates.

However, the aforementioned curves are not valid for gaseous fluids; while, ESPs are frequently utilised to pump petroleum fluids with high gas content \cite{1}. The solution is to develop models to estimate the generated head by ESPs on GPFs. This paper focuses on head estimating models and exclude other types of models developed for GPFs in ESPs, e.g. the ones which estimate surging or stability border \cite{3}, gas bubble size \cite{4} or in-situ gas volume fraction \cite{5}.

Tens of year research led to head-estimating models based on analytical, numerical and empirical approaches \cite{6}. Available analytical models, derived on the basis of mass and momentum balances \cite{7,8}, use unrealistic assumptions and oversimplified physics of two-phase fluids. This has harmed their reliability. Numerical models are normally formulated based on one-
dimensional two-fluid conservations of mass and momentum along streamlines; also they demand prediction of surging initiation in ESPs [9]. As a result, they are not assumed to be trustworthy. On the contrary, empirical models of GFPs in ESPs are widely trusted and used in practice [10, 11].

2. CURRENTLY PREVALENT MODELS

In this section, homogenous model and a number of empirical head-estimating models of GFPs in ESPs are briefly introduced. The values presented for empirical models’ parameters have been identified using the data collected from experiments on diesel fuel/carbon dioxide mixtures, reported in [12]. Advantageously, aforementioned mixtures are similar to petroleum fluids [13], dissimilar to air/water mixtures, which have also been used in experiments to develop ESP empirical models [14-16]. Latter models have been excluded from this paper due to significant difference of their tested fluids and GFPs.

2.1. Model 1

The oldest ESP model for GFPs is the homogenous model. This model is in fact an analytical model, based on oversimplification of two-phase physics of GFPs. Homogenous model receives an input from the curve provided by the manufacturer: the generated head by ESP if pure liquid was pumped instead of GFP ($H_l$) [17]. This head is modified with assumption that the fluid motion is homogenous i.e. liquid and gas have equal speeds:

$$\hat{H}_m = (1 - \alpha)\rho_l + \alpha\rho_g)H_l,$$

where $\rho$, $H$ and indices $l$, $g$ and $m$ stand for density, head, liquid, gas, and mixture, respectively. $\alpha$ is gas void fraction. $\hat{ }$ shows the head is estimated (not experimentally measured).

2.2. Model 2

The second model was developed by Turpin et al in 1986 [18]. Similar to (1), this model was focused to modify the output of manufacturer curve, $H_l$:

$$\hat{H}_m = H_l \exp\left[346430 \left(\frac{q_l}{p_m q_l}\right)^2 - 410 \left(\frac{q_g}{p_m q_l}\right)\right].$$

where $q_l$ and $q_g$ are liquid and gas volumetric flow rates in gallons per minutes (gpm), $p_m$ is intake pressure in psi.

2.3. Model 3

Sachdeva et al, in 1992 [19] proposed a model which did not need any inputs from manufacturer’s catalogue:

$$\hat{H}_m = \frac{K_1}{\rho_m g} (p_m^{E_1} \alpha^{E_2} q_l^{E_3}).$$

where $g$ stands for gravity acceleration. The values of $E_1$, $E_2$ and $E_3$ and $K_1$ are listed in [10] for multiple stages of electrical submersible pumps. As an example, for 8 stages of I-42B radial ESP, $K_1=1.1545620$, $E_1=0.943308$, $E_2=-1.175596$ and $E_3=-1.300093$. Similar to Model 2, (3) is
convertible to a linear equation through taking algorithm. Parameters of a linear equation can be evidently identified with use of least square of error algorithm [20]

2.4. Model 4

This model was presented by Zhou and Sachdeva in 2010 [10]:

\[ \hat{H}_m = H_{\text{max}} K_2 (C_{p_m} \alpha)^{E_4} (1 - \alpha)^{E_5} \left(1 - \frac{q_m}{q_{\text{max}}}\right)^{E_6}, \]

where \( K_2 \) is a unit-less coefficient, \( C \) is pressure unit factor, e.g. 1, 1000 or 0.145 for psi, ksi or kPa. \( H_{\text{max}} \) and \( q_{\text{max}} \) are nominal maximum head and flow rate of the ESP; \( q_m \) is mixture or GPF flow rate where \( q_m = q_f + q_g = q_f / \alpha \).

Mathematical structure of this model seems more meaningful than Model 3; as if gas void fraction and flow rate equal zero, estimated head is definitely \( H_{\text{max}} \). According to [10], for 8 stages of I-42B radial ESP, \( K_2 = 1.971988 \), \( E_4 = 1.987838 \), \( E_5 = 9.659664 \) and \( E_6 = 0.905908 \).

2.5. Summary and Limits of Empirical Models

All presented models have three input variables amongst \( p_{\text{in}}, \rho_{l}, q_{l}, q_{g}, q_m, \) or \( \alpha \). Two other potential input variables, pump rotational speed and temperature appear in neither empirical models of GPFs in ESPs nor in the homogenous model. As a matter of fact, the parameters of the presented empirical models have been identified based on the data collected at a fixed rotational speed of 3500 rpm; thus, the models are valid at this speed only. The estimated head can be adapted for other rotational speeds using ‘affinity laws’ detailed in [2, 10].

3. MODEL DEVELOPMENT

In this section, development of a radial basis function network (RBFN) is reported which is designed to estimate the head of mixtures of carbon dioxide/ diesel fuel pumped by eight stages of an I-42B radial ESP. The reason to choose an RBFN as the head-estimating model is the fact that RBFNs are mathematically proven universal approximators [21]. Inspired by existing empirical models, a single output of \( H_m \) and triple inputs of \( p_{\text{in}}, q_m \) and \( \alpha \) were opted for the RBFN model. The same experimental data, used to identify the parameters of empirical models 2-4, were employed to develop and test the RBFN. Hence, the proposed model and presented empirical models are comparable. The data collected from 110 experiments cover a wide range of gas void fractions [0-0.5], intake pressures [50 to 400] psi and heads [1-55] ft. 93 data sets were used for modelling and the rest to test the RBFN.

An RBFN has two layers, the first layer receives inputs array, \( U \), and produces the ‘layer output’, \( O \). The second layer receives \( O \) and produces the ‘network output’, \( Y \). In this problem, there are three inputs and one output; thus, the input and output arrays are \( U_{3xQ} \) and \( Y_{1xQ} \). \( Q \) is the number of data sets which are fed to the RBFN at once. For instance, if the inputs of the modelling data are fed into the model altogether, then \( Q=93 \).

The first layer has an array of weights (\( W_{3x3} \)) and a scalar namely Spread (\( S \)). The components of layer output, \( O_{3xQ} \), are calculated as presented in (5):

\[ O_{j} = \sum_{i=1}^{3} W_{ij} U_{i} + S, \quad j = 1, 2, \ldots, Q, \]

where \( j \) is the number of data sets and \( i \) is the number of input variables.
\[ O_d = \exp \left( -S \sum_{j=1}^{3} (W_{ij} - U_{jk})^2 \right). \]  

The second layer has an array of weights \((X_{1\times R})\) and an array of biases \((B_{1\times Q})\). The output array is calculated according to (6):

\[ Y_{1\times Q} = X_{1\times R} \times O_{R\times Q} + B_{1\times Q}, \]  

Combination of (5) and (6) is the structure of the RBFN. The major task is to identify the model’s unknown parameters \(R, W, S, X\) and \(B\) using the modelling data presented as input and output vectors of \(U_{3\times 93}\) and \(Y_{1\times 93}\).

From (5), it is clear that the range of \(O\) components is [0 1]; also, if \(i^{th}\) row of \(W\) and \(k^{th}\) column of \(U\) are identical, \(O_{ik}\) will be at its maximum, 1; or simply, maximum values of \(O\) components happen if the rows of \(W\) are same as the columns of \(U\). From (6), it can be seen larger components of \(O\) are more influential on the network output. As a result, in order to maximise the effect of the modelling data on parameter selection, it has been suggested to set \(W = U^T\), consequently \(R = Q\). Then, \(O\) can be calculated with mere use of \(U_{3\times 93}\) and \(S\).

Solution of (7), a rearranged version of (6), determines \(B\) and \(X\) components.

\[ Y_{93\times 93} = [X \ B]_{4\times 86} \begin{bmatrix} O \\ I \end{bmatrix}, \]  

where \(I\) is a unique matrix with size of 93x93.

By this point, it has been shown how to find all unknowns of (5) and (6) except for \(S\). The developed model is called an ‘exact’ RBFN. Such a model evidently provides exact estimation for the modelling data; however, a serious concern about exact RBFNs is model generality or the accuracy of estimation outside the operating points where the modelling data have been collected from or the modelling points. \(S\) is chosen by the designer as a tool to balance accuracy inside and outside the modelling points. A large spread \((S)\) \((S >> 1\) in (5)) can smoothen the model output and generalise the network [22].

Here is a pseudo-algorithm of exact RBFN modelling (to find \(R, W, X, B\) using the input and output arrays of the modelling data, \(U_{3\times 93}\) and \(Y_{1\times 93}\), and a well-selected \(S\))

1. Set \(W_{93\times 3} = U^T_{93\times 3}\)
2. Choose a large \(S\) to generalise the developed RBFN
3. Calculate \(O_{93\times 93}\) from (5) with \(U_{3\times 93}\) (from the modelling data), \(W_{93\times 3}\) and \(S\) defined at steps 1 and 2, respectively.
4. Form (7) with \(Y_{1\times 93}\) (from the modelling data) and \(O\) calculated at step 3.
5. Solve (7) to find \(X_{1\times 93}\) and \(B_{1\times 93}\)

A straightforward non-iterative parameter identification algorithm is an advantage of exact RBFNs; however, this method creates models with a lot of parameters: a disadvantage for working on small data series. In this research, the exact RBFN has 466 parameters; whereas, only 93 modelling data sets, in total 372 pieces of input/output data, are available for modelling.
Excessive number of parameters and focus of the algorithm on exact fitting of the model to the modelling data increases the risk of ‘overfitting’ or lack of generality (see [23, 24]). The only tool to enhance generality in exact RBFNs is spread with a limited influence [23].

Efficient RBFNs, alternatively, may have much fewer parameters than exact RBFNs, which means a lower risk of overfitting. In efficient RBFNs, some columns of \( U \) (not all of them) are selected and transposed to form \( W \). Therefore, \( W \) array is smaller. In order to select \( U \) columns to be used as \( W \) rows, first, every single column is transposed and tried as a single-row \( W \). The column of \( U \) which leads to the smallest modelling error (see the second appendix of [25] about the modelling error) is selected, transposed and used as the first row of \( W \). At the next iteration, another column of \( U \) in which the merger of its transpose to \( W \) leads to the largest drop in the modelling error is chosen and added to \( W \). This continues till the number of \( W \) rows (\( R \)) reaches its pre-defined maximum (\( R_{\text{max}} \)) or the modelling error reaches its predefined target (\( E_t \)). It should be noted that if a too small modelling error (e.g. 0) is targeted, overfitting is more likely to happen.

Here is a pseudo-algorithm of efficient RBFN modelling:

1. \( W=\text{null}, U_{\text{rem}}=U, U_{\text{opt}}=\text{null}, E=1000 \) (a large number)
2. Choose a large \( S \) to generalise the developed RBFN
3. Choose \( R_{\text{max}} \) and target modelling error, \( E_t \)
4. Set \( R=1 \)
5. Set \( k=1 \)
6. Add transpose of \( k^{\text{th}} \) column of \( U_{\text{rem}} \) as the \( R^{\text{th}} \) row of \( W \)
7. Calculate \( O \) from (5) with \( U_{3\times93} \) (from the modelling data), \( W_{R\times3} \) and \( S \) defined at steps 6 and 2.
8. Solve \( Y_{1\times93} = \left[ X B \right]_{d\times(R+93)} \begin{bmatrix} O \\ I \end{bmatrix} \) (from \( Y_{1\times93} \) and \( O \) are available from the modelling data and step 7)
9. Find the Modelling Error (\( ME \)) from comparison of \( Y_{\text{model}} \) (calculated from (5) and (6)) and \( Y \)
10. if \( ME<E_t \), then \( E=ME \) and \( U_{\text{opt}}=U_k \)
11. if \( k \leq 93-R \) then go to 6
12. Remove \( U_{\text{opt}} \) from \( U_{\text{rem}} \)
13. if \( R=R_{\text{max}} \) and \( E>E_t \) then go to 5

4. RESULTS AND DISCUSSION

Both exact and efficient RBFN modelling methods were employed to develop models for GPFs pumped by eight stages of an I-42B radial ESP, using 93 sets of experimental data as detailed in section 3. The models were tested with 17 data sets not used for modelling, ‘test data’, including an input and outputs arrays of \(^1U_{3\times17} \) and \(^1Y_{17\times1} \). Upper left index of \( T \) refers to ‘test’. The vector
of estimated outputs with inputs of $^T U_{3 \times 17}$ is named $^T \hat{Y}_{17 \times 1}$. The ‘test error’ or $TE$, introduced in (8), was employed as the accuracy criterion for models.

$$TE = \frac{1}{17} \sum_{i=1}^{17} \left| Y_i - {^T \hat{Y}}_i \right|,$$

(8)

A model should have a reasonably low $TE$ to be cross-validated [25, 26]. The developed exact RBFN, with 466 parameters, has a rather low $TE$, 2.7683 ft, at a very high value of spread, $S=125$. At lower values of $S$, the modelling error may be misleadingly small. For example, at $S=1$, with use of the modelling data in (8), the resultant error is 0.0645 ft; whereas, the test error is 44.8621 ft, an evident example of overfitting to the modelling data.

The efficient RBFN, however, provides better results with fewer parameters and a lower spread. Development of an efficient RBFN with a spread of 20 and target modelling error ($E_t$) of 1.2 ft result in $R=74$, 390 parameters in total and a very low test error of 1.8648 ft equal to 3.45% of head range. With such a low, this model is definitely considered as cross-validated.

Table 1 compares the test error ($TE$) and number of parameters in empirical models 1-4 (M1-M4) and developed exact and efficient RBFNs:

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>Exact RBFN</th>
<th>Efficient RBFN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Error</td>
<td>8.85</td>
<td>7.36</td>
<td>12.13</td>
<td>5.16</td>
<td>2.77</td>
<td>1.87</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>466</td>
<td>390</td>
</tr>
</tbody>
</table>

Figures 1-3 compare estimated heads of different models with the real head at three different operating areas. In this paper, an operating area is the collection of operation points with same intake pressure and gas void ratio, e.g. $P_{in}=100 \ psi$ and $\alpha=0.2$. Table 2 shows the mean of absolute estimation error for different operating areas. Table 2 and Figs.1-3 show the results for the entire available experimental data in operating areas, not only the test data.

![Figure 1](image-url)
Figure 2. Real and estimated head (by six models) for a mixture of carbon dioxide and diesel fuel pumped by eight stages of an I-42B radial ESP; intake pressure is 100 psi and gas void fraction is 0.1.

Figure 3. Real and estimated head (by six models) for a mixture of carbon dioxide and diesel fuel pumped by eight stages of an I-42B radial ESP; intake pressure is 400 psi and gas void fraction is 0.5.
Table 2. Mean of absolute head estimation error in ft for different models at various operating areas.

<table>
<thead>
<tr>
<th>$P_{in}$</th>
<th>$\alpha$</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>Exact RBFN</th>
<th>Efficient RBFN</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.10</td>
<td>4.66</td>
<td>14.7</td>
<td>24.1</td>
<td>5.00</td>
<td>1.42</td>
<td>0.33</td>
</tr>
<tr>
<td>50</td>
<td>0.15</td>
<td>12.1</td>
<td>12.0</td>
<td>11.5</td>
<td>7.32</td>
<td>4.62</td>
<td>1.85</td>
</tr>
<tr>
<td>50</td>
<td>0.20</td>
<td>15.8</td>
<td>8.63</td>
<td>6.85</td>
<td>8.63</td>
<td>1.32</td>
<td>0.71</td>
</tr>
<tr>
<td>50</td>
<td>0.30</td>
<td>16.4</td>
<td>6.63</td>
<td>5.06</td>
<td>4.34</td>
<td>4.97</td>
<td>1.68</td>
</tr>
<tr>
<td>50</td>
<td>0.40</td>
<td>17.1</td>
<td>3.00</td>
<td>1.39</td>
<td>2.44</td>
<td>2.31</td>
<td>1.81</td>
</tr>
<tr>
<td>100</td>
<td>0.10</td>
<td>5.66</td>
<td>3.21</td>
<td>22.5</td>
<td>4.24</td>
<td>2.29</td>
<td>0.90</td>
</tr>
<tr>
<td>100</td>
<td>0.15</td>
<td>6.28</td>
<td>4.61</td>
<td>13.3</td>
<td>5.94</td>
<td>1.98</td>
<td>1.06</td>
</tr>
<tr>
<td>100</td>
<td>0.20</td>
<td>8.25</td>
<td>4.96</td>
<td>10.4</td>
<td>6.68</td>
<td>1.29</td>
<td>1.30</td>
</tr>
<tr>
<td>100</td>
<td>0.30</td>
<td>10.1</td>
<td>10.3</td>
<td>7.76</td>
<td>4.65</td>
<td>3.00</td>
<td>1.00</td>
</tr>
<tr>
<td>100</td>
<td>0.40</td>
<td>11.7</td>
<td>6.40</td>
<td>3.81</td>
<td>2.89</td>
<td>1.33</td>
<td>0.43</td>
</tr>
<tr>
<td>400</td>
<td>0.30</td>
<td>5.47</td>
<td>3.73</td>
<td>9.92</td>
<td>5.84</td>
<td>0.80</td>
<td>1.09</td>
</tr>
<tr>
<td>400</td>
<td>0.40</td>
<td>4.45</td>
<td>2.95</td>
<td>8.27</td>
<td>4.30</td>
<td>0.57</td>
<td>0.52</td>
</tr>
<tr>
<td>400</td>
<td>0.50</td>
<td>5.69</td>
<td>9.04</td>
<td>7.51</td>
<td>5.79</td>
<td>0.39</td>
<td>1.18</td>
</tr>
</tbody>
</table>

According to Table 2, the efficient and exact RBFN models outperform all comparable empirical models, M1-M4, in 12 and 11 operating areas out of 13 operating areas covered by experiments, respectively. As an exception, Model 3 presents a lower error than RBFNs, at pressure of 50 psi and gas void ratio of 0.4. As a justification, the results of only 2 experiments are available in this operating area, too few to assess modelling approaches. As to Table 2 and Figs. 1-3, different empirical models may perform better in specific operating areas; so-called critical models [6] have been proposed to define these high-performance or validity areas of a number of empirical models. However, RBFN models are advantageously valid for the whole operating areas where the modelling and test data have been collected from.

5. CONCLUSION

This paper initially presented existing empirical models, which estimate the head of gaseous petroleum fluids in ESPs as well as a simple analytical model (the homogenous model). These models are called head estimating models.

Then, exact and efficient RBFN models were developed to serve the same function as the existing head estimating models using the same data used to identify the parameters of aforementioned empirical models. The developed models outperformed existing models, and the efficient RBFN particularly estimated head highly accurately with a test error equivalent to 3.45% of head range. It was also shown that if some popular values were opted for RBFN design factors, e.g., spread of 1 and target modelling error of 0, the developed model would fail to fulfill cross-validation requirements due to overfitting. In the exact RBFN, a very large spread, 125, was shown to effectively reduce overfitting. Overfitting was diminished with use of a large spread, 20, and a fairly large target modelling error for the efficient RBFN.
ACKNOWLEDGEMENT

This work was supported by Kuwait University Research Grant EE02/16.

REFERENCES


AUTHORS

Morteza Mohammadzaheri received his PhD from School of Mechanical Engineering, University of Adelaide, Australia in 2011. He has published/presented more than 90 peer-reviewed articles in technical journals and conferences. He is now an Assistant Professor of Dynamic Systems and Control at the Department of Mechanical and Industrial Engineering of Sultan Qaboos University, Oman.

Mojtaba Ghodsi received his B.Sc. degree in Mechanical Engineering from Isfahan University of Technology in 1999, the M.Sc. degree in Applied Mechanics from Tehran Polytechnic in 2001 and continued his research as Ph.D. (2007) and JSPS postdoctoral fellow (2009) in Precision Engineering Department of the University of Tokyo, Japan. Currently, he is pursuing his career at Sultan Qaboos University, Oman in department of Mechanical and Industrial Engineering. His main research interests include Smart Materials for Actuators, Sensors and Energy Harvesting, NDT and development of Mechatronic systems and devices. Dr. Ghodsi is a Member of IEEE and International Society of Optics and Photonics (SPIE).

Abdullah AlQallaf is an Assistant Professor with the Department of Electrical Engineering, Kuwait University. He received his Ph.D. degree in Electrical Engineering from the University of Minnesota—twin cities, St. Paul, MN, in 2009. Alqallaf’s research interests are Microwave Imaging Techniques, Multimedia Signal Processing, Communication, Bioinformatics and Medical Image Analysis. Alqallaf is an IEEE Senior Member and an IEEE Board Member-Educational and Professional Activities-Kuwait section.