Determination of distance between DC traction power centers in a 1500-V DC subway line with artificial intelligence methods

Mehmet Taciddin AKÇAY¹,*, İlhan KOCAARSLAN²

¹Directorate of Rail Systems, Istanbul Metropolitan Municipality, İstanbul, Turkey
²Department of Electrical-Electronics Engineering, Faculty of Engineering, Istanbul University, Istanbul, Turkey

Abstract: The electrification system in rail systems is designed with regard to the operating data and design parameters. While the electrification system is formed, the minimum voltage rating that the traction force requires during the operation needs to be provided. The highest value of the voltage drop occurring on the line is determined by the distance between power centers. This value needs to be kept within certain limits for the continuity of operation. In this study, the determination of the distance between DC traction power centers for a 1500-V DC-fed rail system is done by means of the adaptive neuro-fuzzy inference system (ANFIS), support vector machines (SVMs), and artificial neural networks (ANNs). The distance occurring on the line is calculated with regard to the operating parameters by means of the ANFIS, SVMs, and ANNs. The ANFIS, SVMs, and ANNs are explained and a comparison is made. The data created regarding one-way and two-way supply conditions are examined for simulation. The main contribution of this paper is the determination of the distance between railway traction power centers with artificial intelligence methods.

Key words: ANFIS, artificial neural network, DC traction power, subway electrification, support vector machine

1. Introduction

In railways, there are various operation parameters affecting the system performance. Artificial intelligence is expedient for numerous operation conditions and various railway traffic. In the literature, there are some prediction works and research about these issues. One such topic is the use of neural networks (NNs) to predict voltage drop during the starting of medium-voltage induction motors. Other is prediction of the voltage drop due to the diode commutation process in the excitation system of salient-pole synchronous generators. DC railway system emulator design for stray current, touch voltage prediction, and a fast scheme for fault detection in DC microgrids based on voltage prediction are other works [1–4]. In this study, determination of the distance between the power centers is done with artificial intelligence methods.

A 1500-V DC supply voltage is used for the traction force system in DC-supplied railways. The supply voltage that the traction force uses is acquired through an interconnected network, which has 34.5-kV phase-to-phase voltage. Two transformers of 34.5 kV/1.2 kV are present in the substations and the transformers can operate as back-up [5]. The equivalent circuit model of a DC railway is presented in Figure 1. The equation regarding the supplying status from a single substation is given with Eq. (1). The resistance values of the feeder cables were also added to R₁ and R₃. R₁ and R₃ values change in accordance with the distance depending on the location of the vehicle. V₁ is the voltage of the vehicle, Vₙ₁ indicates the nominal supply voltage, and

*Correspondence: taciddin.akcay@ibb.gov.tr

This work is licensed under a Creative Commons Attribution 4.0 International License.
I_{vehicle} indicates the vehicle current. The maximum traction force of the vehicles in the railway vehicles with high power consumption can increase to 20 MVA [5–7].

\[ V_1 = V_{n1} - I_{vehicle}\times R_1 - I_{vehicle}\times R_3 \]  

(1)

Since the voltage drop occurring on the line and the currents drawn do not reach high values under normal operating conditions, the distances between the supply stations may be longer. As the number of traction supply stations and the efficiency of the traction force system increase, the voltage drop on the line and the losses decrease [8–12]. The single-line scheme of a DC-fed line is presented in Figure 2.

\[ F_{motion} = A + B \times V + C \times V^2 \]  

(2)

\[ F_{slope} = m \times g \times \sin(\gamma) \]  

(3)
\[ F_{\text{curve}} = (m \times g \div 1000) \times (C_1 - C_2 \times R) \div (R - C_3) \] (4)

\[ F_{\text{traction}} = F_{\text{motion}} + F_{\text{slope}} + F_{\text{curve}} + ma \] (5)

\[ P_{\text{vehicle}} = F_{\text{traction}} \times V \] (6)

The vehicle power increases with traction force and vehicle speed. The equivalent circuit given in Figure 1 was simulated with different operating parameters and 1000 data arrays were obtained regarding different operating conditions. The parameters used in the simulation are the number of vehicles, acceleration-mass value of the vehicle, vehicle motion resistance, curve radius, slope, the length of the supply line, internal consumption current of the vehicle, electric resistance, and inductance of the line; the calculated value is the highest voltage drop value occurring on the line. Random values were assigned to all the input parameters used in the simulation. For the simulation, the number of vehicles varying between 0 and 10 was used and vehicle placement was performed by taking the maximum voltage drop into consideration.

2. Materials and methods

The diversity of the parameters and the variability in operating conditions in the simulation render the solution of this problem complex. Artificial intelligence is a science that deals with enabling machines to produce solutions to complex problems as humans. This is generally performed by taking the characteristics of human intelligence and applying them to a computer as an algorithm. In accordance with the demanded or desired needs, which mental attitude will be presented to which effect with less or more flexible or effective approaches can be displayed. Artificial intelligence was preferred in this study due to the stated advantages. In this study, the adaptive fuzzy inference system (ANFIS), support vector machine (SVM), and artificial neural network (ANN), among the artificial intelligence applications, were used for the simulation. The ANFIS is a hybrid artificial intelligence method that uses the parallel computing and learning capability of ANNs and the inferential characteristic of fuzzy logic. The SVM is one of the quite effective and simple methods used in classification. For classification, it is possible to divide two groups by drawing a line between two groups on a plane. The location on which this line will be drawn should be the farthest place from the members of both groups. The SVM determines how this line will be drawn. The ANN is a method that functions by imitating the way of work of a simple biological nervous system. MATLAB and Weka programs were used for the simulation. Weka was used for the SVM simulation for a better performance and is in widespread use in academic articles.

One thousand data arrays different from each other were used for the MATLAB simulation. The simulations were run for 1000 different operation conditions. The MATLAB simulation screen is given in Figure 3.

2.1. Adaptive neuro fuzzy inference system (ANFIS)

The ANFIS is a class of adaptive networks functionally equivalent to a fuzzy inference system. The ANFIS can be given more integrated with some characteristics of controllers, learning ability, parallel processing, structured knowledge representation, and other supervision and design methods. Fuzzy logic and NNs are supplementary means used together in developing smart systems [13–15]. The ANFIS consists of 6 layers. This system is displayed in Figure 4. The node functions of every layer in the ANFIS structure and the operation of the layers are respectively as follows [15]. The first layer is the input layer. \( X_1 \) and \( X_2 \) are the input signals. The
Figure 3. MATLAB simulation screen.
input signals obtained from every node in this layer are transmitted to other layers. The second layer is the fuzzification layer. In separating the input values into fuzzy sets, Jang’s ANFIS model uses the Bell activation function generalized as a membership function [15]. Here, the output of each node consists of membership degrees based on the input values and the membership function used and the membership values obtained from the 2nd layer are presented as $\mu_{A_j}(x)$ and $\mu_{B_j}(y)$. The third layer is the layer of rules. Each node in this layer expresses the rules established in accordance with the Sugeno fuzzy logic inference system and their numbers [15]. The output of each rule node $\mu_i$ turns out to be the multiplication of membership degrees that arrive from the 2nd layer. The acquisition of $\mu_i$ values, on the condition that $(j = 1, 2)$ and $(i = 1, \ldots, n)$ is as follows:

$$y_3^i = \mu_{A_j}(x) \times \mu_{B_j}(y) = \mu_i \quad (7)$$

Here, $y_3^i$ represents the output values of the 3rd layer; $n$ represents the number of nodes in this layer. The fourth layer is the normalization layer. Each node in this layer regards all the nodes coming from the rule layer as input values and computes the normalized ignition level of each rule. The computing of the normalized ignition level $\bar{\mu}_i$ is performed in accordance with the following formula:

$$y_4^i = N_i = \frac{\mu_i}{\sum_{i=1}^{n} \mu_i} = \bar{\mu}_i \quad (i = 1, n) \quad (7)$$

The fifth layer is the purification layer. The resulting weighted values of a given rule in each node in the purification layer are calculated. The output value of the ith node in the 5th layer is as follows.

$$y_5^i = \bar{\mu}_i [p_i x_1 + q_i x_2 + r_i], \quad (i = 1, n) \quad (8)$$

The $(p_i, q_i, r_i)$ variables here are the outcome parameter set of the ith rule. The sixth is the sum layer. There is only one node in this layer and it is labeled as $\Sigma$. The output value of each node in the 5th layer is summed here so that the actual value of the ANFIS system is obtained. The computing of $y$, which is the output value of the system, is performed in accordance with the equation below [15].

$$y = \sum_{i=1}^{n} \bar{\mu}_i [p_i x_1 + q_i x_2 + r_i] \quad (9)$$

### 2.2. Support vector machine (SVM)

SVMs can be employed in classification and regression problems. The basic idea in the SVM regression method is finding the linear separator function that reflects the characteristic of the educational data available in a way as close to reality as possible and suits the statistical learning theory. Similarly to the classification, in the regression, the core functions are used for the nonlinear situations to be processed. The most significant advantage of SVMs is to solve the classification problem by converting it to a squared optimization problem.
This way, in the learning stage regarding the solution of the problem, the number of operations decreases and the solution is reached more rapidly when compared to other techniques/algorithms. The technique, due to this characteristic of it, provides a great advantage, especially in bulky datasets. Furthermore, since it is optimization-based, it is more successful in terms of the classification performance, computational complexity, and practicality when compared to other techniques [16–19].

A SVM constitutes an n-dimensional hyperplane, which optimally divides the data into two categories. SVM models are closely related to ANNs, and the SVM, which uses a sigmoid kernel function, has a two-layer, feedforward ANN. The interesting characteristic of the SVM is that it functions with the quality of structural risk minimization in statistical learning theory rather than the empirical risk minimization principle derived by minimizing the mean squared error on the dataset. One of the basic assumptions of the SVM is the independent and similar distribution of all samples in the education set. The SVM can be employed in classification and regression problems. The basic idea in the SVM regression method is finding the linear separator function, which reflects the characteristic of the educational data available in a way as closes to reality as possible and suits the statistical learning theory. Similarly to the classification, in the regression, the core functions are used for nonlinear situations to be processed. Two situations that can be encountered with SVMs are the data being of a structure that can be linearly separated or cannot be linearly separated. The SVM network structure is given in Figure 5. X indicates the input vector and b is the bias value. In the system, support vectors are shown with K, whereas a is the Lagrange coefficient. Y is the output of the system.

![Figure 5. Support vector machine structure.](image_url)

The SVM is a controlled classification algorithm based on statistical learning theory. The mathematical algorithms that the SVM has were initially designed for the classification problem of two-class linear data but later they were generalized for the classification of multiclass and nonlinear data.

SVM regression uses a set of core functions for simulations. In this study, the normalized polynomial kernel was selected and is given by Eqs. (10) and (11).

\[ K(x, y) = \langle x, y \rangle = \sqrt{\langle x, \langle x, y \rangle \rangle} \]  \hspace{1cm} (10)

\[ \langle x, y \rangle = \text{PolyKernel}(x, y) \]  \hspace{1cm} (11)

2.3. Artificial neural network (ANN)

294
ANNs emerged as a mathematical method from the latest outputs of endeavors to study and imitate human nature. They take computing and data processing power from their parallel distributed structure and their capability to learn and generalize. Generalization is defined as an ANN producing proper reactions to inputs that have not been experienced in the course of education or learning. These characteristics indicate the problem-solving capability of ANNs [20–25].

A biological neuron consists of a nucleus, body, and two extensions. The structure of the ANN is given in Figure 6. The 1st layer is the input layer. Data are received here and entered into the system. The 2nd layer is the hidden layer. Its use depends on the simulation. The 3rd layer is the output layer. Inputs are processed and received here. Each sphere (nerve) has a function and a threshold value [26–30].

The output of a neuron is given with Eq. (12) as a function formed by adding a bias \( b \) value to the sum of the input data in specific weights. I indicates input, while \( W \) represents the coefficients that the input values take.

\[
\text{Output} = f(i_1W_1 + i_2W_2 + i_3W_3 + b)
\]  

3. Results and discussion

One thousand data arrays different from each other were used for the calculation of distances. A portion of the data used is displayed in Table 1. These data were used for simulation with the ANFIS, SVM, and ANN. Ten input data and 1 output datum are used for the design. These data are the operational data obtained from the train traffic. Input 1 is the number of the supply direction system. In DC rail systems this value may be one or two. Input 2 is the number of the vehicles, which depends on the operational traffic. Input 3 is the product of mass and acceleration. Input 4 is the vehicle motion resistance. Input 5 is curve radius of the route. Input 6 is the slope of the location. Input 7 is the voltage drop at the line. Input 8 is the vehicle current of the internal consumption. Input 9 and input 10 are resistances of catenary and rail. The output is the length of the supply line, which obtains the traction power distances.

3.1. Simulation results with the ANFIS

The structure of the system created for the ANFIS and the simulation results are given below. A structure with 10 inputs and 2 membership functions created for the ANFIS after much testing is given in Figure 7. The input data are given in Table 1 for the ANFIS system. The length of the supply line constitutes the output data. A triangular-shaped membership function was used for the simulation. This triangular-shaped membership function was chosen because of the performance of the better results. The triangular curve is a function of a vector that depends on scalar parameters. A total of \( 2^{10} = 1024 \) rules were established for the ANFIS design.
The ANFIS architecture is shown in Figure 8. The system consists of the input, input MF, rule, output MF, and output modules. The proposed ANFIS detector is a first-order Sugeno type fuzzy inference system with 10 inputs and 1 output. Each input has 2 triangular-type membership functions.

The realized values and calculated values of all data are shown with the ANFIS simulation in Figure 9. The regression value for all data is 0.77. The other performance calculations for the ANFIS and other methods are given in Table 2.

### 3.2. Simulation results with the SVM

The 10 input data and 1 output datum used in the ANFIS simulation were simulated with the SVM method. By trying different variations to obtain better results in the simulation, the SVM parameters were eventually selected as follows. The complexity parameter “c = 1” was selected. The performance of the classifier depends on this parameter. The normalized polynomial kernel function was selected as the core function and the exponent value was taken as “e = 3”. Test mode 10-fold cross-validation was selected in Weka for a better performance. Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model and a test set to evaluate it.
Table 1. A portion of the data set that used.

<table>
<thead>
<tr>
<th>Input 1 - Number of supply directions</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input 2 - Number of vehicles</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Input 3 - ma value (kN)</td>
<td>241</td>
<td>227</td>
<td>290</td>
<td>280</td>
<td>226</td>
</tr>
<tr>
<td>Input 4 - Vehicle motion resistance (kN)</td>
<td>74</td>
<td>74</td>
<td>73</td>
<td>75</td>
<td>66</td>
</tr>
<tr>
<td>Input 5 - Curve radius (m)</td>
<td>851</td>
<td>802</td>
<td>997</td>
<td>856</td>
<td>859</td>
</tr>
<tr>
<td>Input 6 - Slope</td>
<td>21</td>
<td>281</td>
<td>134</td>
<td>133</td>
<td>383</td>
</tr>
<tr>
<td>Input 7 - Voltage drop (V)</td>
<td>257</td>
<td>66</td>
<td>219</td>
<td>378</td>
<td>794</td>
</tr>
<tr>
<td>Input 8 - Internal consumption current of the vehicle (A)</td>
<td>241</td>
<td>218</td>
<td>176</td>
<td>160</td>
<td>168</td>
</tr>
<tr>
<td>Input 9 - Catenary resistance (Ω)</td>
<td>0.0271</td>
<td>0.0112</td>
<td>0.0275</td>
<td>0.0142</td>
<td>0.0182</td>
</tr>
<tr>
<td>Input 10 - Rail resistance (Ω)</td>
<td>0.0312</td>
<td>0.0275</td>
<td>0.0176</td>
<td>0.0206</td>
<td>0.0293</td>
</tr>
<tr>
<td>Output - Length of the supply line (km)</td>
<td>37</td>
<td>26</td>
<td>29</td>
<td>16</td>
<td>32</td>
</tr>
</tbody>
</table>

Figure 9. ANFIS regression graph.

Table 2. The simulation results of all methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>ANFIS</th>
<th>SVM</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean absolute error (MAE)</td>
<td>2.46</td>
<td>2.74</td>
<td>0.57</td>
</tr>
<tr>
<td>Root mean squared error (RMSE)</td>
<td>6.84</td>
<td>3.83</td>
<td>0.75</td>
</tr>
<tr>
<td>Relative absolute error (RAE)</td>
<td>0.3261</td>
<td>0.3627</td>
<td>0.0763</td>
</tr>
<tr>
<td>Root relative squared error (RRSE)</td>
<td>0.7813</td>
<td>0.4378</td>
<td>0.0864</td>
</tr>
<tr>
<td>Total number of instances</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
</tbody>
</table>
The realized values and calculated values of all the data are observed in Figure 10. The regression value is shown with R and, as seen in the figure, this value is 0.90.

\[ \text{Output} = 0.82 \times \text{Target} + 4.6 \]

\[ R = 0.89917 \]

\begin{figure}
\centering
\includegraphics[width=\textwidth]{SVM_regression_graph}
\caption{SVM regression graph.}
\end{figure}

3.3. Simulation results with the ANN

Ten input data, 10 hidden neurons, 1 output neuron, and 1 output datum were used for the ANN architecture used in the design. The network was trained with the Levenberg-Marquardt backpropagation algorithm. The selection of the number of hidden neurons is important to inhibit overfitting. With increasing hidden neuron number, ANN mapping accuracy increases given the training.

Input and output data are given in Table I. The ANN architecture used is given in Figure 11. Seventy percent of the data used for simulation were used for education, 15% for validation, and 15% for the testing. As seen in Figure 12, the best validation value was reached at the 124th iteration by inhibiting overfitting in the simulation. The lowest mean squared error value is 0.57113. The training, validation, and test data produced by the system displayed similar characteristics. Since the validation error value increased in the course of 6 iterations, the simulation was stopped at the end of 130 iterations.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{ANN_architecture}
\caption{ANN architecture designed [MATLAB R2015b].}
\end{figure}

The backpropagation gradient value is given on a logarithmic scale for each iteration in Figure 13. The difference between the test values and validation values is predicted. Validation checks and MATLAB stop the...
simulation with the increase in the MSE value of the validation values in order to inhibit overfitting at the end of 6 iterations. The MSE performance is given with the training state graph. Gradient = 3.5011 at epoch 130, \( \mu = 0.001 \) at epoch 130, and validation checks = 6 at epoch 130.

The error histogram is shown in Figure 14. The differences between the realized values and calculated
values are seen with this graph. The distribution of the errors of the training data is shown with blue, validation data with green, and test data with red. The errors are mostly concentrated between –1.829 and 1.784.

![Error Histogram with 20 Bins](image)

Figure 14. Error histogram.

The realized and calculated values of the training, validation, and test data are seen in Figure 15. The regression value is shown with R, and as seen in Figure 15, these values are 0.99673 for training, 0.99651 for validation, and 0.99366 for the test data. The R value is 0.99627 for all data. As this value approaches 1, the accuracy of the data calculated by the system increases.

3.4. The comparison of the ANFIS, SVM, and ANN results

The ANN regression result is 0.99627. The results show fair results with maximum percent error 0.37%. In the SVM method, the regression value is 0.90. Maximum percent error is 1%. The ANFIS regression value is 0.77 where the maximum percent error is 23%. The accuracy for the realized values and calculated values of the ANN and SVM methods is better than that of the ANFIS results. When the ANFIS, SVM, and ANN results are compared, the ANN results are observed to be better. The performance calculation results of all methods are given in Table 2.

4. Conclusions

In this study, the prediction of the distance between DC traction power centers on a DC-supplied railway with regard to the operating data was performed. One thousand random input data arrays and the calculated output data were used for the simulation. In the analyses carried out, the ANFIS, SVM, and ANN techniques were used. The distance value was predicted. The RRSE value for the data obtained for the ANFIS in the calculations carried out is 78%, while this value is 44% for the SVM and 9% for the ANN. The RMSE values are 7 V for the ANFIS simulation, 4 V for the SVM, and 1 V for the ANN. The MAE value acquired for the ANFIS is 2 V and for the SVM is 3 V, while this value is 1 V for the ANN. The RAE value for the ANFIS is 33% and for the SVM is 36%, while this value is 8% for the ANN. When the data obtained from the simulations are compared, the prediction values produced with the ANN are observed to be better. When the prediction data produced for all techniques are compared with the real data, it is observed that errors are at an acceptable rate and that the prediction data produced are usable.
Figure 15. ANN regression graph.

References


