**ORIGINAL PAPER**

CrossMark

# Estimating power factor of induction motors at any loading conditions using support vector regression (SVR)

**M. Khodapanah**[1] · **A. F. Zobaa**[1] · **M. Abbod**[2]

## Abstract

Three-phase induction motors are industrial work horses known as inductive loads. Inductive loads always create low power factor due to consuming more reactive power. Low power factor not only makes a penalty charge for costumers, but also produces energy losses in electrical systems. To prevent these problems, the power factor which is the ratio of active and reactive power must be maintained toward unity. The power factor can be controlled by manipulating either of active or reactive power. In induction motors, active power is proportional to the motor load that variation of motor load results in increase or decrease in power factor. However, adding reactive power by capacitors would be a substantial solution to improve and control the power factor in unity. Many researches expressed that injecting improper reactive power to improve the power factor creates under- and over-correction. To prevent such problems, equation of power factor correction can be employed in order to obtain the optimal value of reactive power. In this equation, the presence of power factor at every single loading point is required. Estimation techniques can be a significant key to determine the power factor at every loading point. In this paper, several statistical methods including, kriging, regression, artificial neural network and support vector regression, are tested in three induction motors. A comparison is presented to verify the great performance of support vector regression method.

**Keywords** Induction motors · PF · MCMD · Kriging · Regression · ANN · SVR

## 1 Introduction

Three-phase induction motors (IMs) are the main industrial work horses that consume both active and reactive power [1, 2]. IMs are known inductive loads that produce a high power quality problem in electrical systems [3]. Inductive loads consume more reactive power. Consuming more reactive power in loads such as IMs creates energy loss and voltage drop in the electrical systems [4]. To reduce the energy loss and voltage drop, PF of IMs which is the ratio of active and reactive power must be maintained toward unity [5, 6]. In IMs,

✉ M. Khodapanah
Mohammadali.khodapanah@brunel.ac.k

A. F. Zobaa
azobaa@ieee.org

M. Abbod
Maysam.abbod@brunel.ac.uk

[1] Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UK

[2] Department of Electrical Engineering, College of Engineering, Brunel University London, Uxbridge, UK

the active and reactive power varies while the motor load changes from no-load to full-load that consequently PF will be changed [7]. Bimbahara [8] described the reason that at no-load condition there is no mechanical resistance, the only magnetization reactance and motor resistance losses are presented. The stator current will be divided in two components active and reactive current in order to supply both mechanical resistance and magnetization reactance. Since the resistance losses (friction and windage loss) are quit small, only a few active current pass through the resistance losses, but the majority of stator current as a reactive current flow in magnetization reactance and therefore the no-load current creates high-angle legs the stator voltage in the range of 75°–85° [9]. The stator PF at no-load will be approximately between 0.1 and 0.3. However, as mechanical resistance increases, active current or power increases gradually and flow in rotor side to supply mechanical resistance. Then, it decreases the angle and improves the power factor in stator side about 0.8–0.9 [9].

In industrial factories, many motor loads are changing or even sometimes are working at light-load that causes low PF. The low PF in majority of IMs consuming more current is

Springer

stored in the windage magnetic field and regenerated back to the grid line at each AC cycle [10, 11]. This exchange is known reactive current which can be a cost factor. To reduce this cost, generating reactive power is required in order to improve low PF [12]. Capacitors bank is one of the significant solutions to generate reactive power and correct the low PF. Obtaining the optimum value of required reactive power still is an unresolved challenge because in many cases the improper capacitors bank creates under- and over-correction in which under-correction produces low PF and over-correction causes self-excitation in IMs [13]. Presence of PF at any loading condition can obtain the proper size of capacitors. The PF can be determined by the equivalent circuit of IM that presents rotor and stator parameters. However, the equivalent circuit parameters need no-load test and lock rotor test that create a difficulty [2].

Many papers presented methods to determine the parameters using measured data and other available information from manufacturer data. For instance, Pedra [14], Haque [15] and Marcondes and Guimaraes [16] presented determination of IM parameters from manufacturer data. Estimation of IM parameters with genetic algorithm is reported by Phumiphak and Chat-uthai [17]. However, by using the equivalent circuit method the value of PF can be obtained only at no-load and full-load. In this method, to determine the PF at any loads, the slip or rotor speed requires that providing these parameters can create a difficulty in measurement [18].

Ukil presented a method using measured current and manufacturer data (MCMD) to determine the PF of a small IM [9]. It also used the voltage and motor current measurements with zero crossing method and instantaneous power method to obtain the PF at any loads. In this method, the indicated results showed poor results particularly at large IMs due to variation of reactive current. A normal meter device creates a difficulty to measure the PF at every single loading points due to numerical fluctuation. Power analyzer can be used to measure and record the PF at every loading points to resolve the reading issues. However, power analyzer not only is an expensive device, but also is required the motor being switch off for cable connection [19].

Therefore, estimation techniques would be an economical solution to predict the PF at every loading point. Hence, these techniques provide an online monitoring in order to enhance reliability and security of power quality in electrical systems. In this research, several statistical methods with two categories are used. Kriging and polynomial regression as numerical techniques, artificial neural network (ANN) and support vector regression (SVR) as intelligent techniques are used to estimate the PF of IM 100HP from no-load to full/over-load conditions. These methods required input data. The input values can be taken either from motor datasheet or by a few measurement points of voltage, current and input power from no-load to full-load conditions. In this paper, the
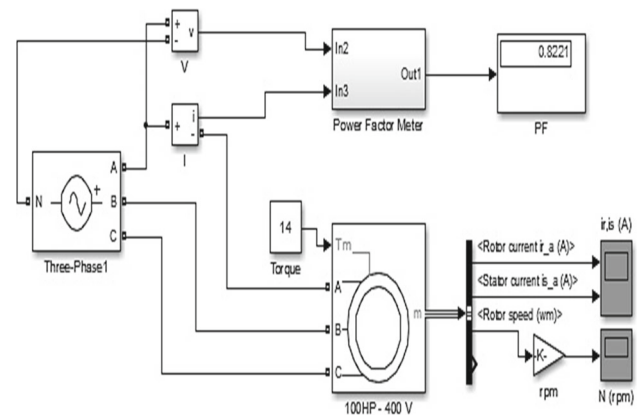


**Fig. 1** Simulated IM by MATLAB/Simulink

estimated PF is compared with measured PF by simulation and practical work as well.

## 2 Case studies

In this study, a three-phase IM with range of 100HP is considered from a stone cutting machine. The measurement procedure took place when the operator moves the blade for cutting the stone by variable volume gradually. A power analyzer named Uni power (UP-2210) is used to measure and record all components of three phases including voltage, current, active and reactive power, PF and harmonics. The power analyzer stored all components in 6-min interval from no-load to full/over-load conditions. In addition, a three-phase IM in size of 100 HP with the same specification of IM from industrial is modeled by MATLAB/Simulink. A torque meter is used to increase motor load step by step. Then, a simulated PF meter measures the PF from no-load to full-load. The simulation diagram is shown in Fig. 1. The measured PF by power analyzer and simulation is illustrated in Fig. 2 in which the measured PF by simulation indicated a result close to the measured PF by power analyzer.

## 3 Kriging technique

Kriging is a geostatistical method, which is known as an interpolation technique. The kriging estimates unknown values based on nearby observed values at surrounding location and weights them in order to minimize the error of a predicted value. The kriging is more applicable in cases that the distance between each observed points and an unknown point is known. Therefore, the general kriging equation can be expressed in Eq. (1).

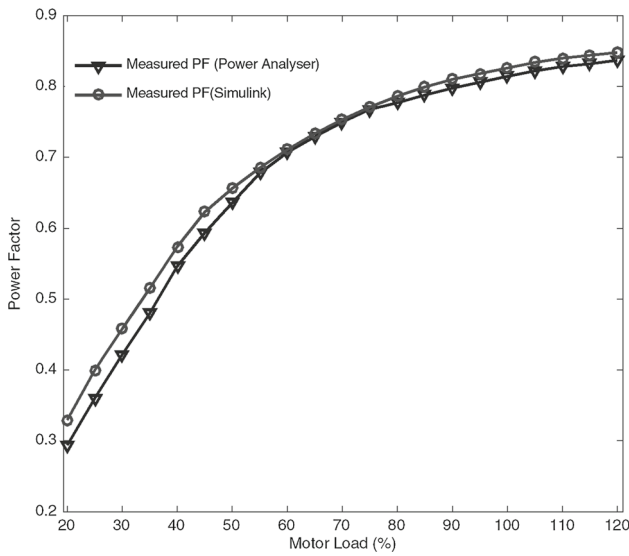$$\hat{Z}_{(S_0)} = \sum_{i=1}^{N} W_i Z_{(S_i)} \qquad (1)$$

**Fig. 2** Measured PF by MATALB/Simulink

where $Z_{(S_i)}$ is the observed value at the $i$th location. $W_i$ is an unknown weight for the observed value at the $i$th location. $S_0$ is the prediction location. $N$ is the number of observed values. To apply this equation, obtaining of weighs $W_i$ are important in which $W_i$ can be computed by a semivariogram. The semivariogram is a function that relates semivariance of data points. It also describes the spatial autocorrelation of the observed values. There are many semivariogram models such as exponential, gaussian and spherical models. Exponential model, which can be a suitable for estimating PF, is applied in this study. Therefore, the exponential function will be expressed in Eq. (2)

$$\gamma(h) = c\left(1 - \exp\left(\frac{-3h}{a}\right)\right) \tag{2}$$

where $c$ is sill that semivariance at which levelling takes place, $h$ is a distance between variables and $a$ is a range that represents the maximum distance in $x$-axis of semivariagram model. The important key point of this method is applying a suitable semivariogram model to perform high output accuracy. Selecting the exponential model is more applicable since it is similar to the PF curve. Therefore, in Eq. (2), $c$ can be replaced as a rated PF at maximum load ($m_{PF}$), $h$ is a distance between all load points and $a$ is replaced as a maximum load ($m_L$). $\gamma(i)$ is semivariagram of exponential model. Lagrange matrix will be applied to obtain the weights of observed values. In the matrix, two main vectors are needed. One is obtained values of semivariogram model $\gamma(i)$ and another is the distance between observed value and the point that will be estimated. Then, the Lagrange multiplayer matrix can be expressed in Eq. (3).

$$\begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mn} & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \vdots \\ \gamma_{n0} \\ 1 \end{bmatrix} \tag{3}$$

In Lagrange multiplication, $W_i$ is ($m \times 1$ matrix) the weight of actual and estimated points which is unknown, $\gamma_i$ is ($m \times n$ matrix) output of the semivariagram function and $\gamma_{no}$ is a vector ($m \times 1$) between the unknown loading points and observed loading points. Thus, from the obtaining values of $W_1, W_2, \ldots, W_n$, the unknown PF can be estimated by Eq. (4).

$$PF = W_1 F_1 + W_2 F_2 + W F_3 + \cdots W_n F_n \tag{4}$$

where $W_i$ is a weight between an estimated point and observed points and $F_i$ is the observed PF. Then, multiplying observed points to obtained weights, PF at a desired point is going to be estimated. In the kriging algorithm, a loop function has been applied in order to estimate the PF at any desired loading condition [20, 21].

## 4 Regression technique

Regression analysis is a kind of statistical modeling that uses to describe the relationships between the independent variable $x$ and the dependent variable $y$. Regression analysis is also applied to predict values named coefficients ($\beta_i$) between a dependent variables ($y_i$) and independent variables ($x_i$).

Predicted coefficients ($\beta_i$) with independent variables ($x_i$) create a new dependent variables ($\hat{y}_i$) with a significant model. Least squares method which is more used of regression analysis estimates the coefficients. There are many functions to use in least squares method including polynomial, exponential, logarithmic and power. Among these functions, polynomial function is more suitable in least square methods because of providing a model by nth degrees [15]. The polynomial function can be expressed by Eqs. (5) and (6).

$$y_i = \sum_{i=1}^{n} f(x_i, \beta) + \varepsilon_i \tag{5}$$

$$\hat{y}_i = f(x_i, \beta) = \beta_0 + \beta_1 x_1 + \beta_2 x_2^2 + \cdots + \beta_m x_n^m \tag{6}$$

where in Eq. (5), $y_i$ is observed value. $f(x_i, \beta)$ is a polynomial function $\varepsilon_i$ is error between observed and estimated values. In Eq. (6), ($\beta_0, \beta_1, \beta_2, \ldots \beta_m$) are the coefficients of polynomial where $m$ indicates the coefficients. $\left(x_1 + x_2^2 + \cdots x_n^m\right)$ are independent variables where $m$ and $n$

are number of polynomial degree and number of variables, respectively. In Eq. (7), $\beta$ multiplied by $x_i$, provide an estimated value as $\hat{y}_i$. The difference between $y_i$ and $\hat{y}_i$ is $\varepsilon_i$. Now, by obtaining the value of $\beta$ and $\varepsilon_i$ with new set of $x_i$, the values of $y_i$ will be determined. Therefore, Eqs. (5) and (6) can be described in Eq. (7).

$$[Y] = [X][\beta] + [\varepsilon] \tag{7}$$

where $[Y]$ is $n$-by-1 vector of dependent variables, $[X]$ is $n$-by-$m$ matrix of estimators with one column for each estimator and one row for each observation. $[\beta]$ is a $m$-by-1 vector of unknown parameters that can be obtained by Eq. (8). $[\varepsilon]$ is an $n$-by-1 vector. In order to minimize the errors $\varepsilon$, the least square procedure is applied in Eq. (9) in which Vondermonde matrix will be used to solve this equation.

$$\beta = \left(X^T X\right)^{-1} X^T Y \tag{8}$$

$$\text{SSE} = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{9}$$

In polynomial regression, the number of polynomial degrees is described as following. For instance, first order creates a linear model. Second and third orders are named as quadratic and cubic make nonlinear models. Therefore, the presence of the real model is important for determination of polynomial degree in order to obtain the best fitting [18, 20].

Polyfit and Polyval are substantial functions in statistical MATLAB tools. Polyfit ($p$) function is used to obtain the coefficients for polynomial of degree ($n$). It can be also described as $p = \text{polyfit}(x, y, n)$ in which $x$ is observed point and determined as an independent value. $y$ is observed value and is determined as a dependent value. $n$ is a degree of polynomial that specifies the polynomial power of the coefficient in $p$. $p$ obtains the polynomial coefficients by using least squares procedure with selecting the number of degree (the length of $p$ is $n+1$). In the procedure of polyfit, independent axis ($x$) requires forming a Vondermonde matrix with $n+1$ columns. Polyfit solves the polynomial coefficients with $p = V/y$ that expressed in Eq. (10).

$$\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} x_1^{n+1} & x_1^n & \cdots & 1 \\ x_2^{n-1} & x_2^n & \cdots & 1 \\ \vdots & \vdots & \ddots & 1 \\ x_n^{n+1} & x_n^n & \vdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{10}$$

Polyval is a function that evaluates $p$ at query points. The function can be describes as $y = \text{polyval}(p, x)$ in which $y$ output is the polynomial coefficients of degree $n$ evaluated at query points $x$. Therefore, combination of both functions with required number of degree can predict values at unknown points with significant accuracy.

## 5 Artificial neural network

In artificial neural network (ANN), the back-propagation is a multilayer feed-forward, and it is one of the most applied neural network models. The back-propagation utilizes the methods of mean square error and gradient descent to realize the modification to the connection weight of the minimum error sum of squares. In this algorithm, some measured values are given to the network as a training sample. Then, the initial values are assigned for the connection weights [22]. For updating weights, the error between the estimated and measured values is back-propagated via the network.

Then, reducing the error between estimated and measured values will be done after procedure of supervised learning. In the nonlinear model, the network is structured in three layers feed-forward back-propagation.

From Fig. 2, the structure of this network contains an input layer, a hidden layer of neurons (nonlinear transfer function) and an output layer of neurons (linear transfer functions). $x_j (j = 1, 2, \ldots, n)$ indicates the input variables, $z_i (i = 1, 2, \ldots, m)$ describes the output of neurons in the hidden layers, and $y_t (t = 1, 2, \ldots, l)$ states the output of the neural network [23].

Neural network enables to create any kind of patterns by given sufficient input values. Training the network by suitable method such as Levenberg–Marquardt back-propagation will determine the excellent weight in order to fit the inputs and targets. The training process of updating the weights values can be done with two important steps (Fig. 3).

The first step is hidden layer that the below function in Eqs. (11) and (12) explains the calculation of hidden layer for outputs of all neurons. $\text{net}_i$ is the activation value of the $i$th node, $z_i$ is the output of the hidden layer, and $f_H$ in Eq. (13) is a activation function that in this case sigmoid function is determined.

$$\text{net}_i = \sum_{j=0}^{n} w_{ji} x_j v_i = 1, 2, \ldots m \tag{11}$$

$$z_i = f_H(\text{net}_i) i = 1, 2, \ldots, m \tag{12}$$

$$f_H(x) = \frac{1}{1 + \exp(-x)} \tag{13}$$

The second step is the output in which the below function in Eq. (14) shows the output of all neurons in the output layer.

$$y_t = f_t \left(\sum_{i=0}^{m} w_{it} z_i \right) t = 1, 2, \ldots, l \tag{14}$$
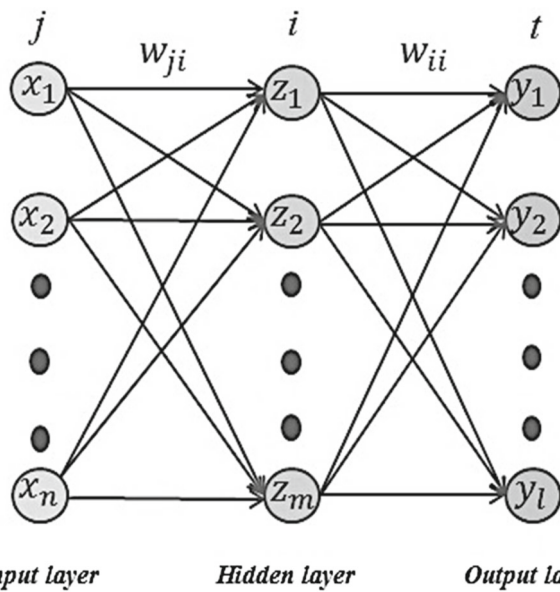
**Fig. 3** Structure of NNBP layers. (Reproduced with permission from [23])



**Fig. 4** Error $\varepsilon$ and limits $\xi$ in the $\varepsilon$-insensitive function. (Reproduced with permission from [25])

where $(f_t (t = 1, 2, \ldots, l)$ is a line function. The weights set with observed values and are minimized by the delta rule according to the learning samples. The topology in this study determines by a set of observed values and errors in order to select the suitable number of neurons.

## 6 Support vector regression

The theory of SVR was developed by Vapnik in 1997 and is known as one of the significant technique in terms of solving regression problem. The SVR method constructs a hyperplane in high-dimensional space in order to minimize the generalization error between defined upper and lower bound [24]. SVR can only act in linear way but by mapping the main space into the higher-dimensional space, it would construct a set of hyperplanes close to the all data points to solve a nonlinear model. The data point is $D = \left\{ X_i, t_i \right\}_i^n$ where $x_i$ is the input vector, $t_i$ is the target output and $n$ is the number of data sample. Therefore, the regression function can be expressed in Eq. (15).

$$Y = f(x) = w\phi(x) + b \tag{15}$$

where $\phi(x)$ is the hyperplane in high-dimensional space. $X$ is a $m$-dimensional feature space. $w$ and $b$ are coefficients of SVR that solve the regression problem. $w$ and $b$ are required to be found by minimizing the regularized empirical risk function in Eq. (16) and a loss function in Eq. (17).
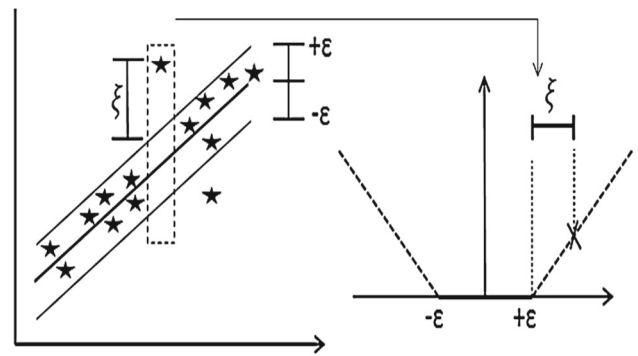
$$R_{\text{emp}} = C\frac{1}{n} \sum_{i=1}^{n} L_\varepsilon(t_i, y_i) + \frac{1}{2} w^{\mathrm{T}} w \tag{16}$$

$$L_\varepsilon(t_i - y_i) = \begin{cases} 0 & |t_i - y_i| \leq \varepsilon \\ |t_i - y_i| - \varepsilon & \text{otherwise} \end{cases} \tag{17}$$

In the risk function, $C\frac{1}{n} \sum_{i=1}^{n} L_s(t_i, y_i)$ is empirical risk error and $\frac{1}{2} w^{\mathrm{T}} w$ is a regularization term or the flatness of the function that needs to be minimized for simplification of the model. $L_\varepsilon(t_i - y_i)$ is an intensive loss function. Parameter $C$ is named as a capacity of the SVR that decides the trade between the regularization term and the empirical risk. E is named as size of the hyper-dimensional cylinder that covers the function with the training data points. SVR performs linear regression in high-dimensional feature space using $\varepsilon$-insensitive loss, and at the same time tries to reduce model complicity by minimizing $w^{\mathrm{T}} w$. The minimization would be determined by introducing slack variables $\xi_i^-, \xi_i^+ \quad i = 1, \ldots n$ since $\varepsilon$-insensitive loss is equal to slack variables. Figure 4 indicates $\varepsilon$ and limits $\xi$ in the $\varepsilon$-insensitive function.

The parameters $C$ and $\varepsilon$ will be set by designer during training step for optimizing slack variables [25]. To calculate the parameters of $w$ and $b$,, Eq. (16) changes to Eq. (17). Slack variables $\xi_i^-$ and $\xi_i^+$ represent upper and lower limits in the output and minimized by Eq. (18).

Minimize $\quad \frac{1}{2} w^{\mathrm{T}} w + C \sum_{i=1}^{n} \left( \xi_i^- + \xi_i^+ \right)$

Subject to the constraints:

$$\begin{cases} \alpha_i^+ \ -t_i + y_i + \varepsilon + \xi_i^+ \geq 0 \\ \alpha_i^- \ t_i + y_i + \varepsilon + \xi_i^- \geq 0 \\ \mu_i^+ \ \xi_i^+ \geq 0 \\ \mu_i^- \ \xi_i^- \geq 0 \end{cases} \tag{18}$$

where $\alpha_i^+$, $\alpha_i^-$ and $\mu_i^+$, $\mu_i^-$ are the coefficients of Lagrange multipliers. The $w$ will be obtained by applying partial derivative in Eq. (19).

$$\frac{\partial R_f}{\partial w} = 0 w = \sum_i (\alpha_i^+ - \alpha_i^-) x_i$$

$$\frac{\partial R_f}{\partial b} = 0 \sum_i (\alpha_i^+ - \alpha_i^-) = 0$$

$$\frac{\partial R_f}{\partial \xi_i^+} = 0 \alpha_i^+ + \mu_i^+ = C$$

$$\frac{\partial R_f}{\partial \xi_i^-} = 0 \alpha_i^- + \mu_i^- = C \qquad (19)$$

Moreover, for obtaining the value of $b$, two main parameters are required. One is $w$ which is calculated by Eq. (19) and another is $S$ (Support vector) that can be considered from Eq. (20). Therefore, considering both Eqs. (19) and (20), $b$ would be determined in Eq. (21). Therefore, regression function in Eq. (22) solves the nonlinear problem.

$$S = \left\{ i \mid \ 0 < \alpha_i^+ + \alpha_i^- < C \right\} \qquad (20)$$

$$b = \frac{1}{|S|} \sum_{i \in S} \left[ t_i - w^T - x_i - \text{sign}(\alpha_i^+ - \alpha_i^-)\varepsilon \right] \qquad (21)$$

$$y = \sum_i (\alpha_i^+ - \alpha_i^-) K(X_i, X_j) + b \qquad (22)$$

where $(\alpha_i^+ - \alpha_i^-)$ is support vector coefficients and $K(X_i, X_j)$ is the kernel function. There are several kernel functions to solve the minimization problem. In this study, radial basis function (RBF) using Gaussian is used by Eq. (23) where $\sigma$ is the dispersion coefficient of the Gaussian.

$$K(X_i, X_j) = \exp\left( -\frac{1}{2\sigma^2} \| X_i - X_j \| \right) \qquad (23)$$

In Lagrange multipliers, the following Karush–Kuhn—Tucker (KKT) and the quadratic programming will consider nonzero values to the $\alpha_i^+$, $\alpha_i^-$ which are defined support vectors. By multiplying the support vectors to the kernel $K(X_i, X)$, the output provide errors equal, less or greater than $\varepsilon$. Kernel function is equal to vectors $X_i$ and $X_j$ in the feature space as $\phi(X_i)$ and $\phi(X_j)$ where $K(X_i, X_j) = \phi(X_i) * \phi(X_j)$. Therefore, the training of the SVR can solve a quadratic and convex optimization problem [25].

## 7 Results and discussions

Input power measurement method is applied for motor load calculation in order to indicate PF against motor load. Determining the PF at any load points leads to select the proper size of capacitors in order to prevent under- or over-correction.
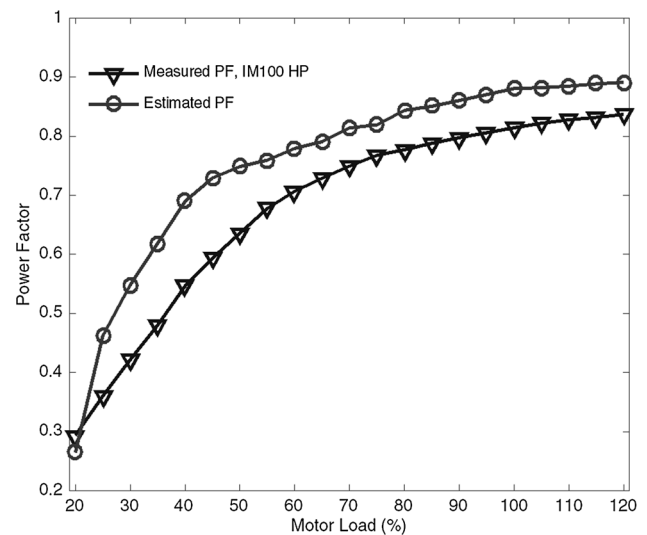


**Fig. 5** Results of estimated PF by MCMD method

Under-correction indicates low PF that produces a penalty of charge. Over-correction generates more reactive power or current than the motor needed. In such cases, self-excitation takes place due to higher reactive current than magnetizing current.

Hence, aforementioned reasons can prove the importance of the PF from no-load to full/over-load conditions. Ukil published a method using measured current and manufacturer data (MCMD) to estimate the PF from no-load to full/over-load condition, and the result of IM 100HP in Fig. 5 shows that the MCMD method produced hug errors from no-load to full/over-load conditions in large IMs. This method only provided satisfied performance in small IMs because in the small IM the reactive current is almost constant from no-load to full/over-load.

However, the method was not able to estimate the PF in the large IM because reactive current cannot be constant due to high air gap that variation of motor loads from no-load to full/over-load causes reactive current to be changed highly. In this paper, to resolve these issues, several estimation techniques have been implemented in order to minimize the errors. Kriging and regression as numerical techniques are used to estimate the PF. In Kriging, the distance between target points and observed load points is considered. Then, by using exponential function and Lagrange matrix, the weight of observed points is computed. Multiplying of observed points in their weights computed the PF at a desired point. In this method, a loop function is applied to predict the PF at every loading from no-load to full-load conditions. The obtained weights times the observed PF provided the PF at target load points.

Figure 6 indicated that Kriging generated results very close to the measured points from no-load to full-load. However, it can be seen huge errors between estimated and
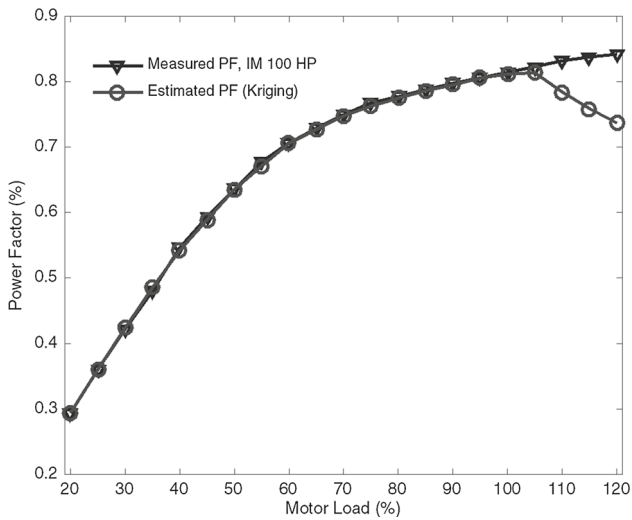
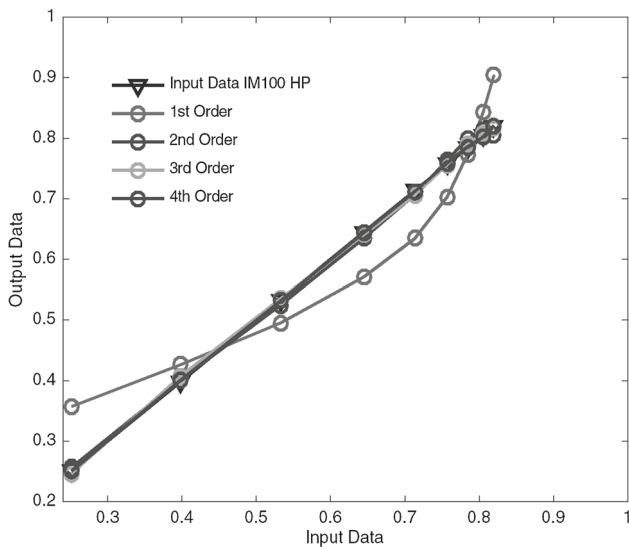Fig. 6 Results of estimated target PF by Kriging method



Fig. 7 Results of fitness in polynomial regression

**Table 1** Predicted coefficients of polynomial degrees

| Orders | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ | $\beta_5$ |
|---|---|---|---|---|---|
| First order | 0.0063 | 0.2498 | – | – | |
| Second order | $-9.1E-05$ | 0.0173 | $-0.0121$ | – | |
| Third order | $4.8E-07$ | $-1.8E-04$ | 0.0220 | $-0.079$ | – |
| Fourth order | $1.3E-08$ | $-2.7E-06$ | $8.7E-05$ | 0.0133 | 0.0108 |
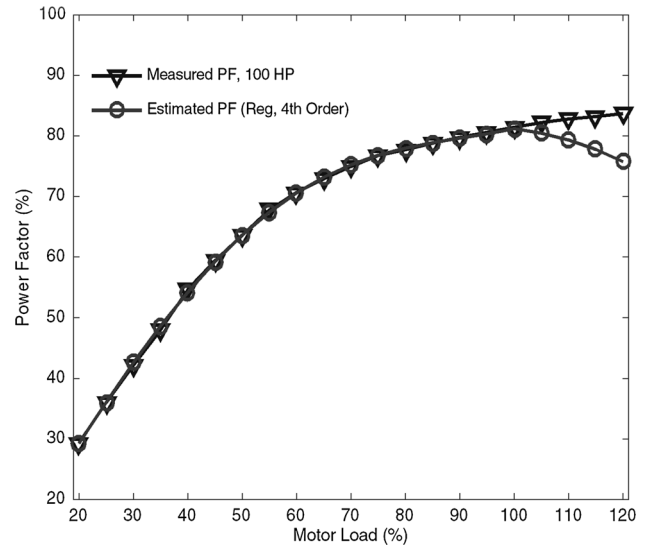


Fig. 8 Results of estimated PF in polynomial regression

measured values from full-load to over-load. Since Kriging is an interpolation technique, it is not able to make an extrapolation at over-load condition. However, in regression, polynomial function is applied in which polynomial degrees provided significant roles that each number of polynomial degree creates different models. Polyval and Polyfit in MATLAB can be significant functions to determine the polynomial coefficients and then create a model fit to the observed PF curve.

Figure 7 and Table 1 illustrated the trained data and predicted polynomial coefficients in first, second, third and fourth orders. The evidence confirmed that polynomial regression with fourth order produced best fitting to the observed PF curve. Therefore, based on the existed model, Fig. 8 indicates the estimated unknown PF from no-load to

full/over-load conditions. However, from full-load to overload there is a huge gap between estimated and measured PF. Although both methods produced results from no-load to full-load very close to the measured points, kriging and regression methods could not fit the model in the over-load condition that results indicated extreme errors at over-load condition.

The reason is that both methods are not able to extrapolate the data at unseen points. Figures 6 and 8 indicated that these methods obtained unacceptable results with huge errors at over-load conditions. To optimize these issues, the study found intelligent techniques including ANN and SVR in order to estimate the PF not only between the known observation, but also to estimate the PF at over-load conditions with high performance. In ANN method, feed-forward backpropagation algorithm is used in which five and three of input data are selected as training and testing, respectively. Considering three hidden layer and a Levenberg–Marquardt to train the algorithm indicated a significant generalization.
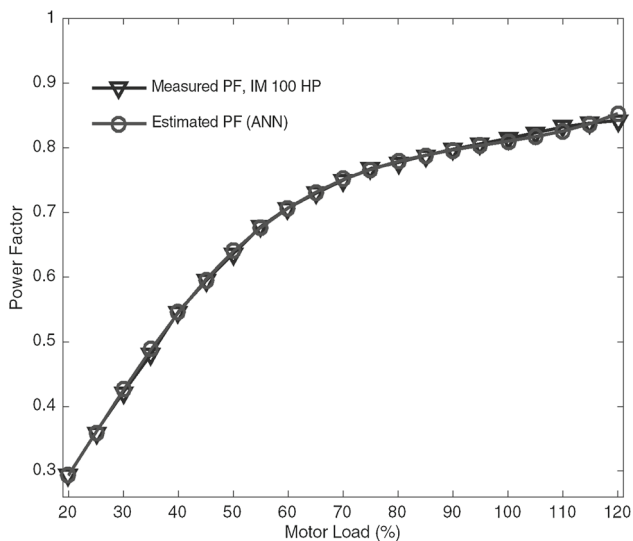
Figure 9 shows the fitness between observed points and output points where the output values are fitted to the input data. Figure 10 illustrates the estimated PF from no-load to
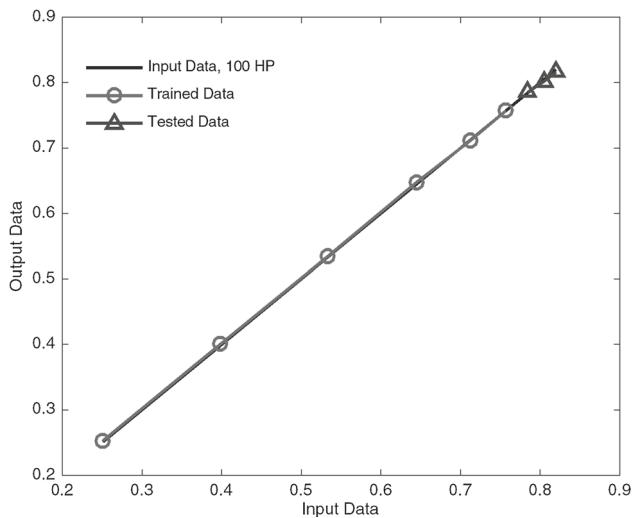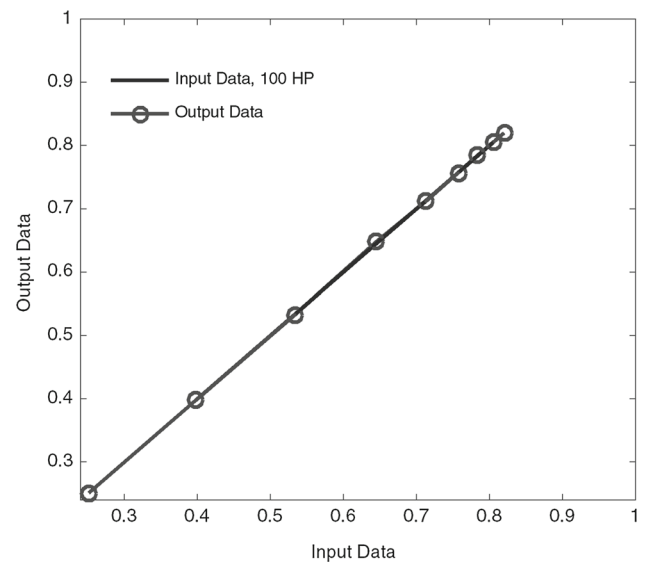
**Fig. 9** Results of fitness in NNBP
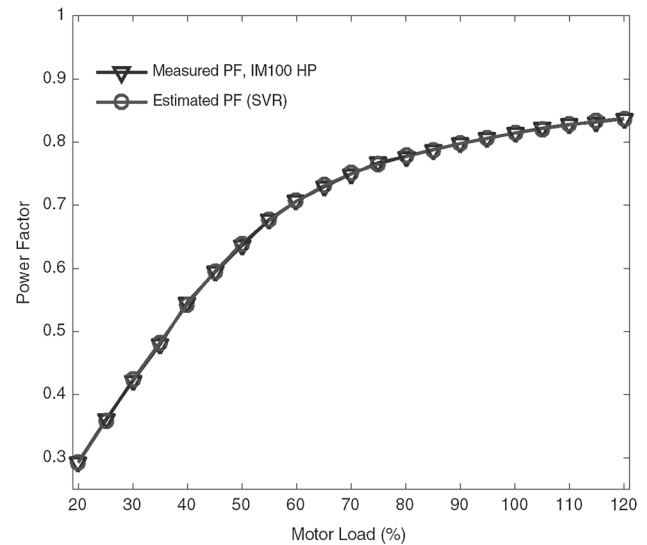


**Fig. 11** Results of estimated PF in SVR



**Fig. 10** Results of estimated PF in NNBP



**Fig. 12** Results of fitness in SVR

full-load and over-load. The results illustrated that NNBP performed a great fitting from no-load to over-load. In spite of the fact that NNBP produced the results very close to measured points with small error, several times are applied to run the algorithm that each running generated different results. Therefore, in this method, obtaining the best result creates a difficulty due to running the algorithm more than once. This can be a main disadvantage of ANN method. To optimize the issue of NNBP, the SVR is used to provide a fixed model and estimate the PF at any loading point. The strategy of SVR is constructed a set of hyperplanes close to the all data points with a lower and upper bound. The RBF kernel function is used to obtain the support vectors. The parameters of SVR have a significant role in terms of creating a suitable model. In this case, the proper design of parameters indi-

cated a model very close to the observed points. The excited model leads to estimate the unknown points from no-load to full-load and over-load conditions properly. The estimated PF from no-load to over-load points is shown in Fig. 11. As a result, the comparison between implemented methods expressed that the SVR method obtained satisfactory results in small, medium and large IMs. The possibility of adjusting the parameters is one of the main advantages of this method that is able to provide desired models (Fig. 12).

The designed parameters of SVR are shown in Table 2.

As a result, Table 3 illustrates the fitness and accuracy of existed models where the fitness is computed by R-squared and the accuracy obtained by minimum value divided by maximum value at each points, and then the average value

**Table 2** Parameters of SVR

| Parameters | Values |
| --- | --- |
| $C$ | 23 |
| $\xi$ | 25 |
| $\varepsilon$ | 0.00015 |

**Table 3** Validity and accuracy of proposed methods

| Methods | Fitness $R_2$ | Error (%) | Accuracy (%) | Computation time (s) |
| --- | --- | --- | --- | --- |
| Simulation | – | 2.946 | 96.864 | 0.032 |
| MCMD | – | 14.448 | 85.552 | 0.005 |
| Kriging | – | 1.681 | 98.319 | 0.015 |
| PR first order | 0.8855 | 9.749 | 91.454 | 0.047 |
| PR second order | 0.9975 | 2.036 | 97.971 | 0.043 |
| PR third order | 0.9993 | 0.860 | 99.150 | 0.046 |
| PR fourth order | 0.9999 | 0.932 | 99.104 | 0.048 |
| ANN | 0.9997 | 0.364 | 99.443 | 0.370 |
| SVR | 0.9999 | 0.348 | 99.653 | 0.272 |

provided accuracy. Error of estimated points is obtained by mean absolute percentage error (MAPE). In Kriging and MCMD, the input data are not trained as like as other methods due to different strategies. The error results observed that MCMD produced a lower accuracy in 85.5%. However, the SVR provided a high accuracy in 99.6% compared with others. The error is calculated by mean absolute percentage error (MAPE) which is shown in. The computation time in proposed methods are illustrated in second.

## 8 Conclusions

The power factor of induction motors is one of the significant elements that must be maintained toward unity. The power factor is variable while the motor load changes from no-load to full/over-load. This variation caused monitoring and determining the low power factor at any loading condition becomes important due to finding the optimal reactive power for power factor compensation. In this paper, several estimation techniques are applied to estimate the power factor at any loading conditions. Kriging and regression which are numerical methods estimated the power factor with reasonable results from no-load to full-load. However, both methods produced very poor results from full-load to over-load. Neural network and support vector regression which are intelligent techniques created greater results from no-load to full/over-load conditions in which the support vector

regression method indicated a satisfactory performance with accurate results greater than ANN and numerical methods.

## References

1. Fuchs EF (2008) Power quality in power systems and electrical machines. Academic Press, Cambridge
2. Chapman S (2004) Electric machinery fundamentals. McGraw-Hill Education, New York
3. Orsag P (2014) Impact of mains power quality on operation characteristics of induction motor. In: 14th international conference on environment and electrical engineering (EEEIC), 10–12 May 2014, Ostrava, Czech Republic
4. Zahir J (2009) Estimation of power factor by the analysis of power quality data for voltage unbalance. In: ICEE, Melborne
5. Kumar CP, Sabberwal SP, Mukharji AK (1994) Power factor measurement and correction techniques. Electric Power Syst Res 32:141–143
6. Guo L, Cheng Y, Zhang L, Huang H (2008) Research on power-factor regulating tariff standard. In: Proc. 2008 IEEE electricity distribution conf., pp 1–5
7. Lalotra J (2016) Examination of the change in the power factor due to loading effect. Int Res J Adv Eng Sci 1(1):25–28
8. Bimbhra P (1997) Electrical machinery. Khanna, Delhi
9. Ukil A, Bloch R, Andenna A (2011) Estimation of induction motor operating power factor from measured current and manufacturer data. IEEE Trans Energy Convers 26(2):699–706
10. Khanchi S (2013) Power factor improvement of induction motor by using capacitors. Int J Eng Trends Technol 4(7):2967–2971
11. Ermis M (2003) Self-excitation of induction motors compensated by permanently connected capacitors and recommendations for IEEE Std 141-1993. IEEE Trans Ind Appl 39(2):313–324
12. Adisa A (2006) A study of improving the power factor of a three-phase induction motor using static switched capacitor. In: Power electronics and motion control conference
13. Jimoh AA (2006) A study of improving the power factor of a three-phase induction motor using a static switched capacitor. In: 12th International power electronics and motion control conference, Pretoria North, South Africa 30 Aug–1 Sep 2006, pp 1088–1093
14. Pedra J (2008) On the determination of induction motor parameters from manufacturer data for electromagnetic transient programs. IEEE Trans Power Syst 23(4):1709–1718
15. Haque MH (2008) Determination of NEMA design induction motor parameters from manufacturer data. IEEE Trans Energy Convers 23(4):997–1004
16. Marcondes J, Guimaraes C (2014) Parameter determination of asynchronous machines from manufacturer data sheet. IEEE Trans Energy Convers 29(3):689–697
17. Phumiphak T, Chat-uthai C (2002) Estimation of induction motor parameters based on field test coupled with genetic algorithm. Int Conf Power Syst 2:1199–1203
18. Orman M (2011) Slip estimation of a large induction machine based on MCSA. In: Diagnostics for electric machines, power electronics and drives (SDEMPED), 5–8 Sept. 2011
19. Khodapanah M, Zobaa AF, Abbod M (2016) Estimating power factor of induction motors using regression techniques. In: 17th

International conference on harmonics and quality of power (ICHQP), Belohorizento, Brezil, Oct 16–19, 2016, pp 502–507

20. Khodapanah M, Zobaa AF, Abbod M (2015) Monitoring of power factor for induction machine using estimation technique. In 2015 50th international universities power engineering conference, UPEC 2015, Stoke on Trent, United Kingdom, September 1–4, 2015, pp 1–5

21. Guntaka R (2014) Regression and Kriging analysis for grid power factor estimation. J Electr Syst Inf Technol 1:223–233

22. Sagiroglu S (2006) Power factor correction technique based on artificial neural networks. Energy Convers Manag 47:3204–3215

23. Al-hnaity B (2015) Predicting FTSE 100 close price using hybrid model. In: SAI intelligent systems conference (IntelliSys), 10–11 Nov. 2015, London, UK, pp 49–54

24. Smola A (2004) A tutorial on support vector regression. Stat Comput 14(4):199–222

25. Villazana S (2006) A novel method to estimate the rotor resistance of the induction motor using support vector machines. In: IEEE industrial electronics