Power Electronics and Drives
Volume 8(43), 2023 DOI: 10.2478/pead-2023-0003

Power Electronics and Drives

Analysis of the Stator Current Prediction Capabilities in Induction Motor Drive Using the LSTM Network

Research paper

Krystian Teler*, Teresa Orłowska-Kowalska

Department of Electrical Machines, Drives and Measurements, Wrocław University of Science and Technology, Wrocław, Poland

Received: 06, November 2022; Accepted: 13, December 2022

Abstract: In modern areas of knowledge related to electric drive automation, there is often a need to predict the state variables of the drive system state variables, such as phase current and voltage, electromagnetic torque, stator and rotor flux, and others. This need arises mainly from the use of predictive control algorithms but also from the need to monitor the state of the drive to diagnose possible faults that have not yet occurred but may occur in the future. This paper presents a method for predicting stator phase current signals using a network composed of long-short-term memory units, allowing the simultaneous prediction of two signals. The developed network was trained on a set of current signals generated by software. Its operation was verified by simulation tests in a direct rotor flux-oriented control (DRFOC) structure for an induction motor drive in the Matlab/Simulink environment. An important property of this method is the possibility of obtaining a filtering action on the output of the network, whose intensity can be controlled by varying the sampling frequency of the training signals.

Keywords: *current sensors • faults • signal prediction • LSTM • induction motor drive*

1. Introduction

Today, the use of electric drives with induction motors (IMs) or permanent magnet synchronous motors (PMSMs) has become very common in industry. This is due to their very high efficiency and very good control characteristics (Orlowska-Kowalska and Dybkowski, 2016). Unfortunately, the use of advanced control methods requires the construction of an appropriate power electronics system consisting of many components. Any of these can fail, which is why it is important to design fault-tolerant control systems. In these systems, the priority is to ensure that the system operates correctly at the earliest possible point after a fault has occurred (Blanke et al., 2006). Tools for predicting future system states, and in particular, the values of the state variables, appear to be invaluable. Vector control structures such as direct rotor flux-oriented control (DRFOC) or direct torque control require information about the present values of state variables, including stator phase currents, in order to function properly. Therefore, it is important to develop methods to detect and compensate for faults in current sensors (CSs), for which signal prediction methods may be useful.

The need for phase current prediction of electric motors is mainly due to the use of prediction control, which has gained popularity in recent years. Li et al. (2016) present a method for predicting the currents of an IM in a rectangular *x*–*y* synchronous coordinate system, based on its mathematical model. The disadvantage of this method, as with all methods based on mathematical models, is the problem of accurately determining the motor parameters, which may vary under real operating conditions. Similar methods based on a mathematical model of the machine can be found for PMSMs (Ha, 2010; Im and Kim, 2018; Peralta-Sanchez et al., 2009; Wang et al., 2022). To make the predictor immune to changes in IM parameters, a Luenberger observer applied to the feedback loop

* *Email: krystian.teler@pwr.edu.pl*

is presented by Yan et al. (2020). By using the Lyapunov stability theory, the problem of the time-varying values of the stator resistance and the main inductance of the machine is solved.

Solutions that do not use the mathematical model of the motor can also be found in the literature. A method for predicting IM stator currents using the difference between the present and historical values of current samples is presented by Wang et al. (2020). The method is robust to changing machine parameters, which is important in real-world conditions. In a previous paper (Li et al., 2019), a method for the prediction of stator currents in the FOC control structure of a PMSM is described. This uses the delay information introduced by the converter and the pulse-width modulator (PWM), and then the equations describing the future value of the current sample are determined. The equations do not use a mathematical model; however, the exact determination of the mentioned delay values is required.

In the literature presented above, current prediction algorithms are usually designed to be as computationally efficient as possible and to best represent future current measurements. This approach is necessary due to the use of prediction signals in control. However, to the best of our knowledge, there is a lack of information in the literature on methods for predicting currents in IMs or PMSMs using artificial neural networks (NNs), including long short-term memory (LSTM) networks, which deal very well with time series. However, there is little literature in the field of electric drive automation on the prediction of signals using these networks. An example is the method described by Li and Akilan (2022), in which the LSTM network was used to predict the winding, tooth, and stator yoke temperatures, as well as the temperature of permanent magnets in a PMSM.

Although the presented literature is rich in model-based and model-free solutions in the area of predictive control, none of them has been used for prediction of stator current due to different fault types of CS. Nevertheless, these types of methods are commonly used for CS fault detection and compensation in IM and PMSM drives, and a review of the recent proposals can be found in the Introduction of the article by Adamczyk and Orlowska-Kowalska (2022). However, there is a lack of fault classification methods based on NNs that use current prediction. For example, a previous work (Skowron et al., 2022) indeed uses an NN for classification, but as mentioned, the developed method does not use current prediction values. Therefore, the main reason for the research presented in this article was to explore the possibility of predicting stator currents in the IM drive using LSTM networks, including selected CS fault types such as gain change, offset, and saturation. During the research, it was discovered that changing the sampling frequency of the training and testing signals significantly affects the output values of the network. The simulations carried out can help to understand how artificial NNs should be designed and trained so that the results obtained meet the designer's expectations.

After this Introduction, Section 2 describes the methodology for simulating the phase current signals used to train the network. Section 3 presents a description of the basic LSTM cell and its training method. Section 4 provides information on the adopted network structure, its operation, and the training parameters used. The results of the simulation studies, including the DRFOC system, are presented in Section 5. The article ends with the conclusions of the presented studies.

2. Description of the Methodology for Simulating Current Signals

The research carried out in this paper was based on the mathematical model of an IM, which is usually developed based on well-known simplifying assumptions (Leonhard, 1996). Taking these into account, this model can be represented in a stationary coordinate system (α–β) using spatial vectors, in relative units [p.u.] as follows:

$$\frac{d}{dt}\boldsymbol{\Psi}_s = \frac{1}{T_N}(\boldsymbol{u}_s - r_s\boldsymbol{i}_s)\,, \tag{1}$$

$$\frac{d}{dt}\boldsymbol{\Psi}_r = \frac{1}{T_N}\left(\frac{r_r}{l_r}\left(l_m\boldsymbol{i}_s - \boldsymbol{\Psi}_r\right) + j\omega_m\boldsymbol{\Psi}_r\right), \tag{2}$$

$$\boldsymbol{\Psi}_s = l_s\boldsymbol{i}_s + l_m\boldsymbol{i}_r\,, \tag{3}$$

$$\boldsymbol{\Psi}_r = l_r\boldsymbol{i}_r + l_m\boldsymbol{i}_s\,, \tag{4}$$

$$\frac{d}{dt}\omega_m = \frac{1}{T_m}\left(t_{em} - t_L\right), \tag{5}$$

$$t_{em} = \mathrm{Im}\left(\boldsymbol{\Psi}_s^* \boldsymbol{i}_s\right), \tag{6}$$

where $\boldsymbol{u}_s$, $\boldsymbol{i}_s$ – stator voltage and current vector; $\boldsymbol{i}_r$ – rotor current vector; Ψ_s, Ψ_r – stator and rotor fluxes; r_s, r_r – stator and rotor windings' resistance, respectively; l_m, l_s, l_r – main, stator, and rotor inductances, respectively; ω_m – rotor angular speed; $T_N = 1/(2\pi f_{sN})$ – electromechanical constant; T_m – mechanical constant; and t_L – load torque.

Parameters of the tested motor are given in Table A1 of the Appendix.

In this study, the time waveforms of the phase currents measured by the CSs in the DRFOC structure of the IM and the signals from the stator current estimator, i.e., the virtual CS (VCS) (Adamczyk and Orlowska-Kowalska, 2019) were used. Signals from the VCS estimator were converted from the α–β coordinate system to the three-phase *A*–*B*–*C* system according to the following formula:

$$\begin{bmatrix} i_{sA} \\ i_{sB} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & 0 \\ -1 & \sqrt{3} \end{bmatrix}\begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \end{bmatrix}, \tag{7}$$

where i_{sA}, i_{sB} are the currents measured in the stator winding phases, respectively, and $i_{s\alpha}$, $i_{s\beta}$ are the currents converted to the α–β system, respectively.

During the research, artificial NNs were used, the performance of which largely depends on the quantity and quality of the training data used in their learning process. To omit the time-consuming step of simulating the operation of the drive under various operating conditions and stator CS faults, it was decided that the current signals required for training would be generated by software.

2.1. Phase currents of the resistor–inductor–electromotive force (RLE) receiver

The steady-state IM current signal is a sinusoidal signal, whose frequency depends on the rotor angular velocity and amplitude depends on the load torque. Furthermore, there are harmonics in the signal related to the operation of the PWM and voltage inverter of the order $k = 1 + 6g$, where $g = 0, \pm1, \pm2, \ldots$. Thus, the phase current of an IM, which for a frequency converter is an RLE-type load (Figure 1), can be described by the following formula (Tunia et al., 1982):

$$i_{sA}(t) = I_m\left(\sin(\omega t - \varphi_1) - \frac{E_{om}}{U_m}\sin(\omega t - \varphi_1 + \varphi_E) + \sum_{k=5}^{\infty}\frac{1}{k}\sqrt{\frac{1+\left(\omega\frac{L_o}{R_o}\right)^2}{1+\left(k\omega\frac{L_o}{R_o}\right)^2}}\sin(k\omega t - \varphi_1)\right), \tag{8}$$

where: $I_m = \dfrac{U_m}{\sqrt{R_o^2 + (\omega L_o)^2}}$ - amplitude of the first harmonic of the current; R_o – resistance of the load; L_o – inductance of the load; $U_m = 2U_d/\pi$ – amplitude of the first harmonic of the phase voltage of the star connected receiver;

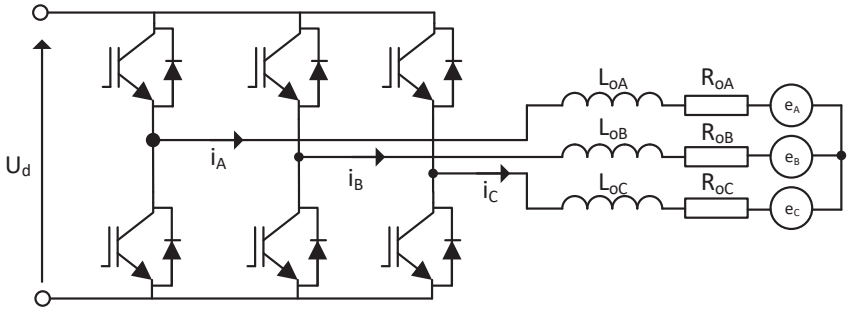

Fig. 1. Scheme of the voltage inverter with RLE-type load.
RLE, resistor–inductor–electromotive force.

E_{om} – amplitude of sinusoidal voltage source e_o; φ_E – phase shift angle between the voltage e_o and the first harmonic of inverter output voltage; $\varphi_1 = \text{atan}(L_o/R_o)$; $\varphi_k = \text{atan}(k\omega L_o/R_o)$; ω – pulsation of the first harmonic of voltage and current.

As the harmonic order k increases, its amplitude decreases, so the influence of higher harmonics on the form of the current signal is much smaller than in the case of harmonics of lower orders. In the case of the inverter phase current, the phase shift of successive harmonics φ_k, which depends on the ratio of inductance and load resistance of the inverter, as well as on the harmonic order, must also be taken into account.

Given the above considerations, it was assumed that harmonics of orders k = 5, 7, 11, 13, 17, 19, 23, 25, 29, 31 would be added to the generated current signals. The amplitudes of harmonics of higher orders are relatively small and do not significantly affect the shape of the current. To check the generalisability of the proposed method using the LSTM network, it was assumed that the parameters of the virtual receiver would take the values $E_{om} = 0$ V, R_o = 1 Ω, and L_o = 0.1 H. Then, the amplitudes of subsequent harmonics were determined in relation to the amplitude of the first harmonic according to the following formula:

$$h_{no\min al} = \frac{1}{k}\sqrt{\frac{1+\left(\omega\frac{L_o}{R_o}\right)^2}{1+\left(k\omega\frac{L_o}{R_o}\right)^2}}\cdot 100\% \,. \tag{9}$$

To diversify the results obtained, the amplitudes of the successive harmonics were randomly drawn from the range $[h_{min}, h_{max}]$, where $h_{min} = h_{nominal}/2$ and $h_{max} = 3h_{nominal}/2$. Noise was also added to each signal sample, with a random value described by a uniform distribution with a range of ±0.02 p.u.

To represent the different operating conditions of the drive, sinusoidal signals (one full period) were generated representing signals estimated by the VCS estimator with different frequencies f = 5, 20, 35, and 50 Hz and current amplitudes expressed in relative units I_m = 0.4, 0.55, 0.7, 0.85, and 1.0 p.u. The signal was sampled at a frequency of f_p = 8 kHz. The successive stages of signal generation are shown in Figure 2.

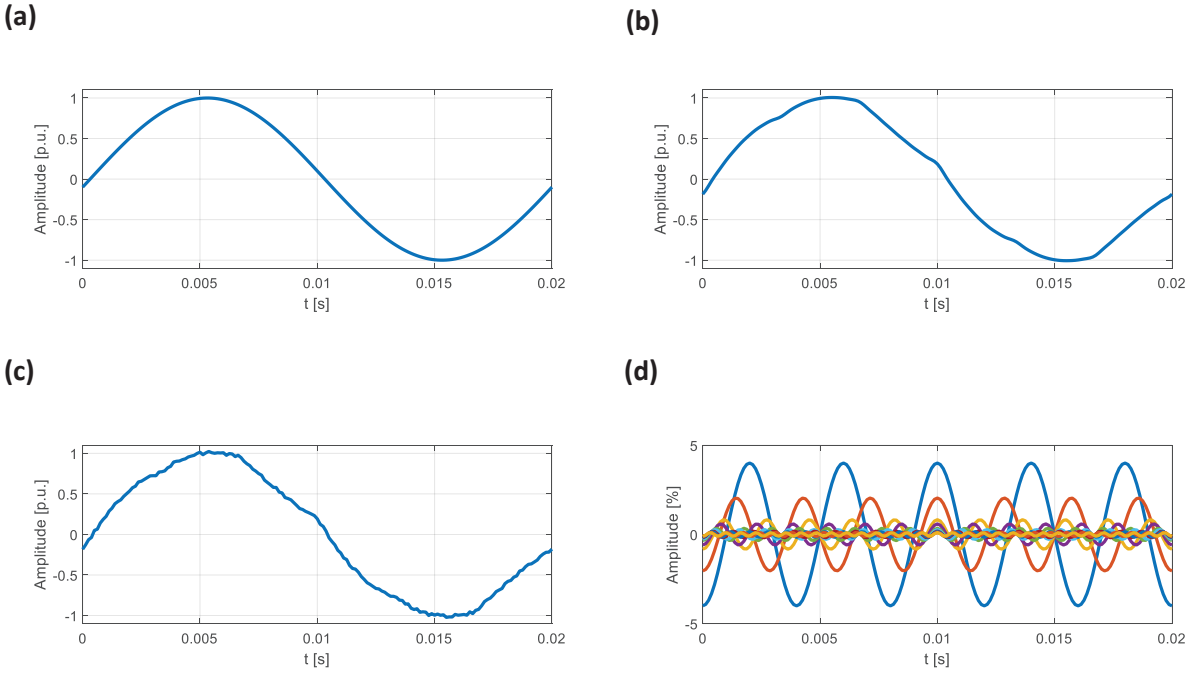

Fig. 2. Data generation process: basic sinusoidal signal with frequency f = 50 Hz and amplitude I_m = 1 p.u. (a), signal containing higher harmonics (b), signal containing higher harmonics and noise (c), and harmonic waveforms that have been added to the basic signal (d).

2.2. Phase current signals measured by faulty CSs

In addition to sinusoidal signals representing the estimates of phase currents, waveforms were also generated that could be recorded by a faulty CS. The principle of operation of typical CSs used in industry is to measure the magnetic field generated by the current flowing in the primary circuit. By design, a distinction is made between open-loop and closed-loop CSs. The second type of sensor has properties that compensate for the effects of temperature drift (Ziegler et al., 2009). The internal structure of typical closed-loop current transducers used in industry consists of a magnetic core with a wound coil, a Hall sensor located in the air gap of the core, an operational amplifier, and a measurement resistor (Figure 3).

The magnetic field generated by the current i_c flowing in the primary circuit closes in the core and acts on a Hall sensor located in the air gap of the core. An operational amplifier then amplifies the Hall voltage, forcing the current to flow through the coil wound on the core. The flowing current produces a magnetic flux that compensates for the flux produced by the primary circuit. By measuring the voltage drop across a resistor R_s connected in series with the coil circuit, the value of the current that flows in the primary circuit can be determined. The way in which the sensors operate and their design allows six different types of faults to be distinguished (Lee and Ryu, 2003), as shown in Table 1.

The waveforms of the faulty signals that were generated by the software included the following faults: change in sensor gain (gain change) caused by a change in the resistance of the measurement resistor, e.g., as a result of temperature changes; addition of a DC component to the measurements (offset) caused by an asymmetry in the

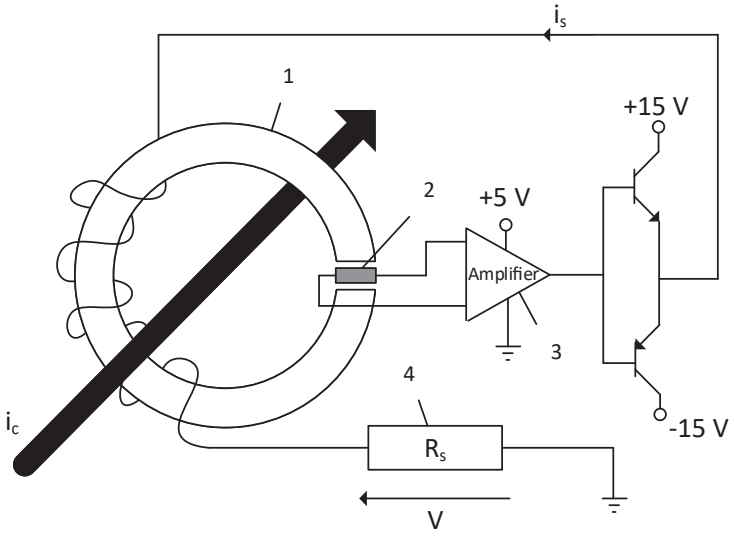

Fig. 3. Operating diagram of a typical current feedback converter. Magnetic core (1), Hall sensor (2), operational amplifier (3), and measurement resistor (4).

Fault type	Mathematical model
Open circuit	0
Gain change	$\varepsilon I_m \sin(\omega t)$
Offset	$I_m \sin(\omega t) + I_{offset}$
Saturation	$\begin{cases} I_{sat} & \text{for} \quad I_m \sin(\omega t) \geq I_{sat} \\ I_m \sin(\omega t) & \text{for} \quad \lvert I_m \sin(\omega t) \rvert < I_{sat} \\ -I_{sat} & \text{for} \quad I_m \sin(\omega t) \leq -I_{sat} \end{cases}$
Disconnections	$\begin{bmatrix} 0 & I_m \sin(\omega t) \end{bmatrix}$
Noise	$I_m \sin(\omega t) + n(t)$

Table 1. Types of CS faults.

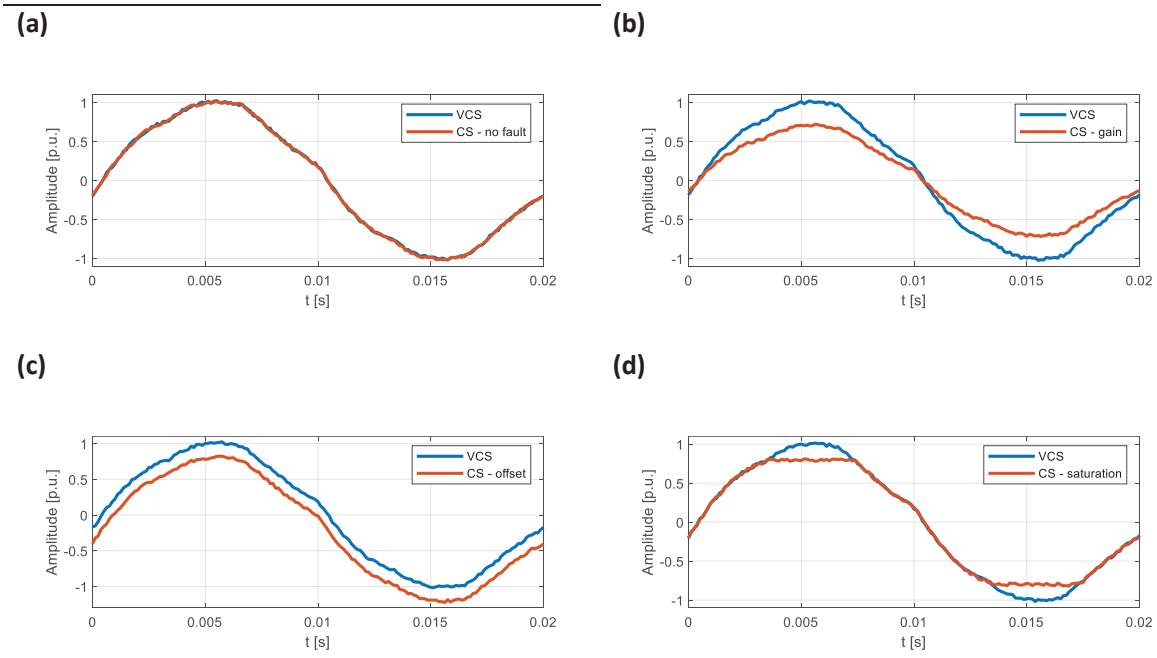

Fig. 4. Software-generated signals due to stator CS fault: no fault (a), gain change with ε = 0.7 (b), offset with I_{offset} = –0.3 p.u. (c), and saturation with I_{sat} = 0.8 p.u. (d). The frequency of the fundamental harmonic was f = 50 Hz, and its amplitude was I_m = 1 p.u.
CS, current sensor; p.u., per unit; VCS, virtual CS.

Hall sensor voltage, e.g., also as a result of temperature changes; and saturation of the sensor core (saturation) caused by demagnetisation of the core, e.g., as a result of current flow in the primary circuit with a value exceeding the limits allowed by the sensor manufacturer. Example fault signals generated for the frequency f = 50 Hz and amplitude I_m = 1 p.u. are shown in Figure 4.

3. Basic Information on LSTM Networks

3.1. Structure of the LSTM network

LSTM networks are a specially modified version of recurrent NNs (RNNs). Their advantage over RNNs lies in the ability to determine the effect of successive input signal samples on the network response. This means that during the processing of successive input vectors, information that is irrelevant or of little relevance to the task being performed may appear. The LSTM network is then able to reduce the impact of such information on the network output, which will be the conclusion of all the input vectors that appear. This property is particularly useful in signal processing.

An LSTM network is made up of cells whereby each cell decides which information is relevant and should be further processed and which information should be ignored. Every cell is composed of gates, and a distinction is made between an input gate, an output gate, and a forget gate, as well as a cell state (Figure 5). Data processing in the basic version of the LSTM cell can be divided into four stages, as described below.

The first step is to decide whether the output h_{j-1} returned in the previous calculation step is relevant for further information processing. The forgetting gate is responsible for this. The new input vector x_j and the previous output vector h_{j-1} are multiplied by the respective weight matrices W_{fh} and W_{fx}, and then the resulting values are summed together with the bias vector b_f. The value obtained is given as the argument of the sigmoidal activation function $y = \sigma(x)$, whose values are in the interval (0,1), allowing the calculation of the value f_t of the forgetting gate according to the following formula:

$$f_j = \sigma(W_{fx}x_j + W_{fh}h_{j-1} + b_f)\,. \tag{10}$$

In the next step, a decision is made as to which information to keep in the present cell state. This is done by calculating the value of the input gate i_j and the vector associated with the cell state update according to the following relationships:

$$i_j = \sigma(W_{ix}x_j + W_{ih}h_{j-1} + b_i)\,, \tag{11}$$

$$\tilde{c}_j = \tanh(W_{cx}x_j + W_{ch}h_{j-1} + b_j)\,. \tag{12}$$

Once the values f_j, i_j, and $\tilde{c}_j$ have been determined, the new cell state c_j can be calculated by multiplication of the corresponding elements of the vectors, which is the Hadamard product denoted by $\odot$:

$$c_j = f_j \odot c_{j-1} + i_j \odot \tilde{c}_j\,. \tag{13}$$

The first multiplication in Eq. (13) corresponds to the operation of forgetting irrelevant information for the cell. The second multiplication takes into account the relevance of the new information by which the state of the cell needs to be updated.

In the last step, the output vector h_j of the cell is determined. Its value is influenced by the present state of cell c_j, the input vector x_j, and the previous output of cell h_{j-1}. To determine the output vector h_j, it is necessary first to calculate the value of the output gate o_j and the value of the hyperbolic tangent function with the argument being the present state of the cell c_j:

$$o_j = \sigma(W_{ox}x_j + W_{oh}h_{j-1} + b_o)\,, \tag{14}$$

$$h_j = o_j \odot \tanh(c_j)\,. \tag{15}$$

3.2. Training methods for LSTM networks

LSTM cells are constructed from gates whose activation functions are differentiable functions. Therefore, each cell can be treated as a certain differentiable approximating function that can be trained using classical optimisation methods. These methods require the definition of an objective function $E(\theta)$, and one of the simplest methods for training the network is the gradient descent method. In the mentioned method, in each iteration of the training process, the vector of network parameters is updated towards the negative values of the gradient of the objective function according to the following formula:

$$\theta_{j+1} = \theta_j - \alpha \nabla E\left(\theta_j\right), \tag{16}$$

where θ – parameter vector, j – iteration number, $\alpha > 0$ – learning rate, and $E(\theta)$ – objective function.

In LSTM networks, it is very difficult to obtain satisfactory results when training them using the gradient descent method due to the existence of recursion in the network, whereby the gradient of the objective function can quickly reach very large (gradient explosion) or very small values (vanishing of the gradient). This in turn leads to parameter vector values that are far from optimal values (Goodfellow et al., 2016). Another disadvantage is the constant learning rate α. When the gradient of the objective function takes on small values, further training of the network is very slow. To speed up the training process, optimisation algorithms that have a learning rate adaptation mechanism can be used. One such algorithm is ADAM (a name derived from the term “adaptive moment estimation”) (Kingma and Ba, 2015), whose successive steps to update the parameters θ are described by Eqs. (17)–(21).

$$m_j = \beta_1 m_{j-1} + \left(1-\beta_1\right)\nabla E\left(\theta_j\right), \tag{17}$$

$$v_j = \beta_2 v_{j-1} + \left(1-\beta_2\right)\left[\nabla E\left(\theta_j\right)\right]^2, \tag{18}$$

$$\hat{m}_j = \frac{m_j}{\sqrt{1-\beta_1^j}}\,, \tag{19}$$

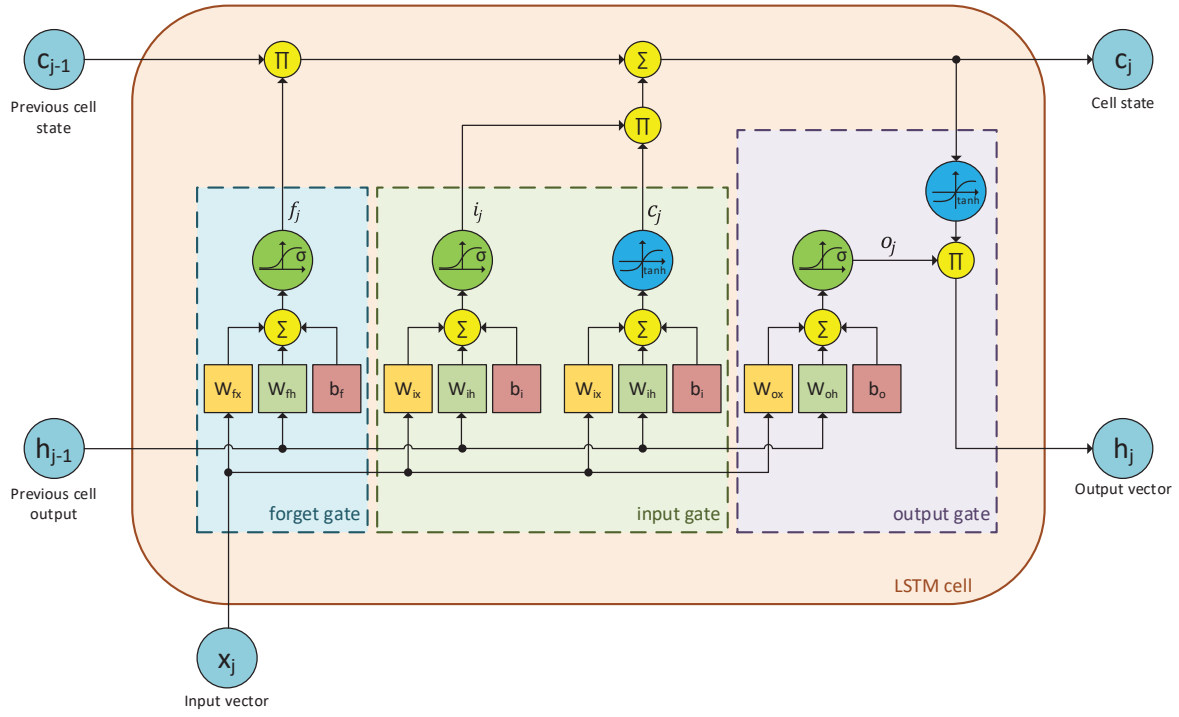

Fig. 5. Scheme of the LSTM cell.
LSTM, long short-term memory.

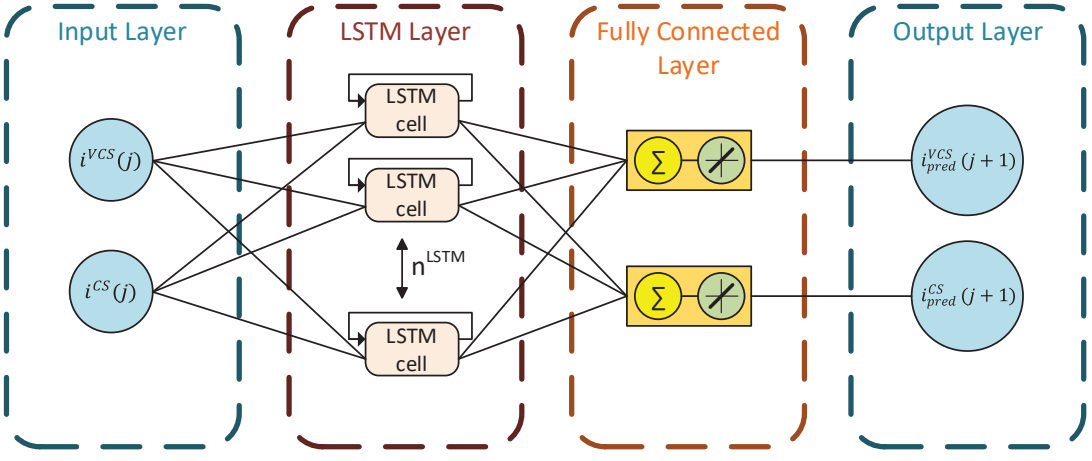

Fig. 6. Structure of the stator current predictor using the LSTM network.
LSTM, long short-term memory.

$$\hat{v}_j = \frac{v_j}{\sqrt{1-\beta_2^j}}, \tag{20}$$

$$\theta_j = \theta_{j-1} - \frac{\alpha \hat{m}_j}{\sqrt{\hat{v}_j} + \varepsilon}. \tag{21}$$

The algorithm updates the exponential moving average of the gradient m_j and the square of the gradient v_j at each step. The parameters $\beta_1, \beta_2 \in [0,1)$ control how quickly the effect of historical values on the present value of the moving average decreases. The vectors m_0 and v_0 are initialised with zeroes. Therefore, the moving average

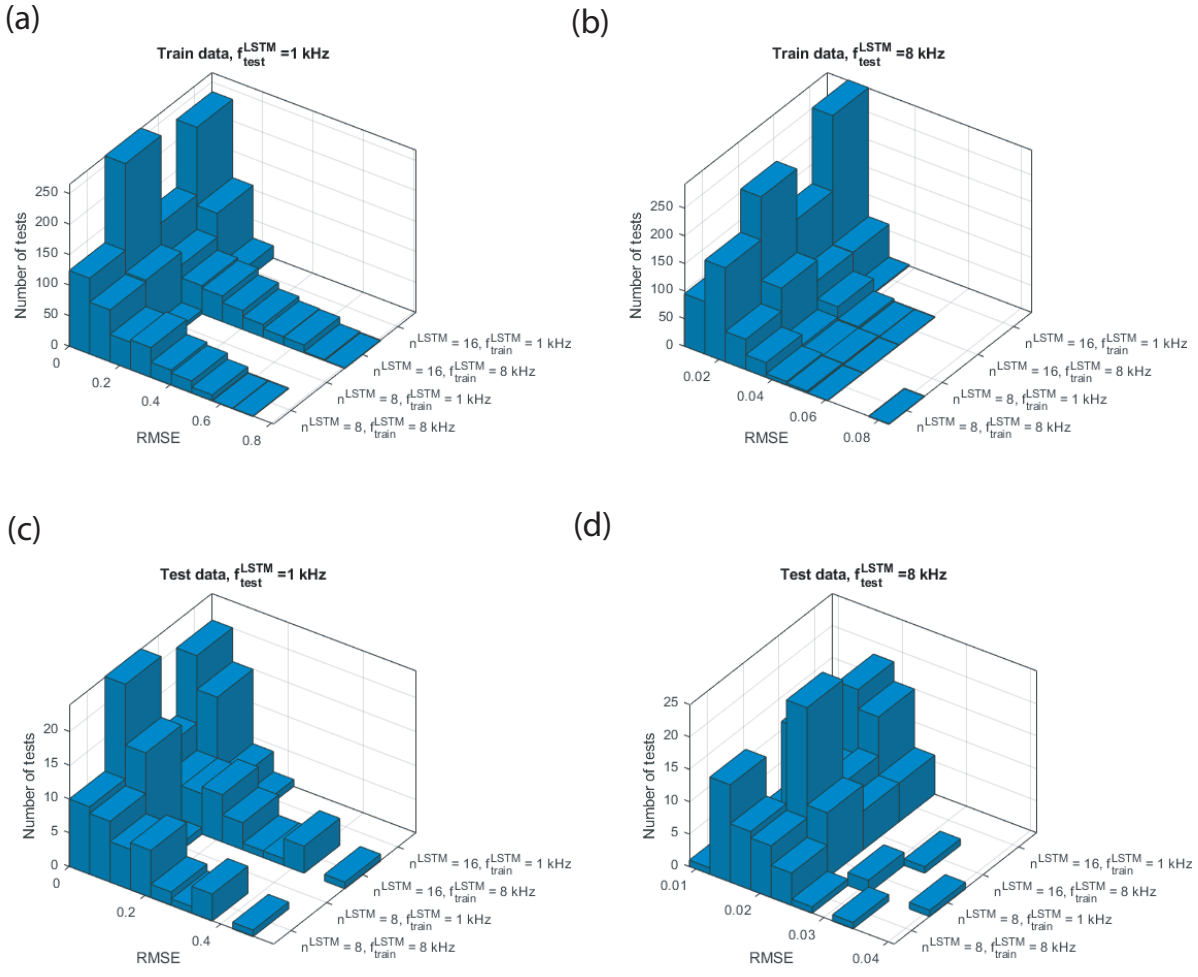

Fig. 7. Histograms showing the distribution of RMSE values on Y_{train} (a, b) and Y_{test} (c, d) training data for different structures containing $n^{LSTM} = 8$ and $n^{LSTM} = 16$ LSTM cells and trained for two different training signal sampling frequencies ($f_{train}^{LSTM} = 1$ kHz and $f_{train}^{LSTM} = 8$ kHz) and tested for signals sampled at $f_{test}^{LSTM} = 1$ kHz (a, c) and $f_{test}^{LSTM} = 8$ kHz (b, d).
LSTM, long short-term memory; RMSE, root mean square error.

is shifted towards zero, and the values $\hat{m}_j, \hat{v}_j$ are calculated for correction. The values of β_1^j and β_2^j denote an exponentiation, where the exponent is *j*. The factor ε in Eq. (21) takes on very small values and is added to avoid division by zero. In many cases, $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$ are good values with which to begin an attempt to train the network; all the networks described in the following sections were also trained with these parameters.

4. LSTM Network Training and Testing Methodology Applied to IM Stator Current Prediction

The structure of the developed network included an input layer of dimension 2, an LSTM layer of dimension n^{LSTM} (which was modified in the study), a fully connected layer, and an output layer of dimension 2 (Figure 6).

The input vector *X* consisted of two values, which were samples of the signal at time instant *j*:

$$X(j) = \begin{bmatrix} i^{VCS}(j) \\ i^{CS}(j) \end{bmatrix}, \tag{22}$$

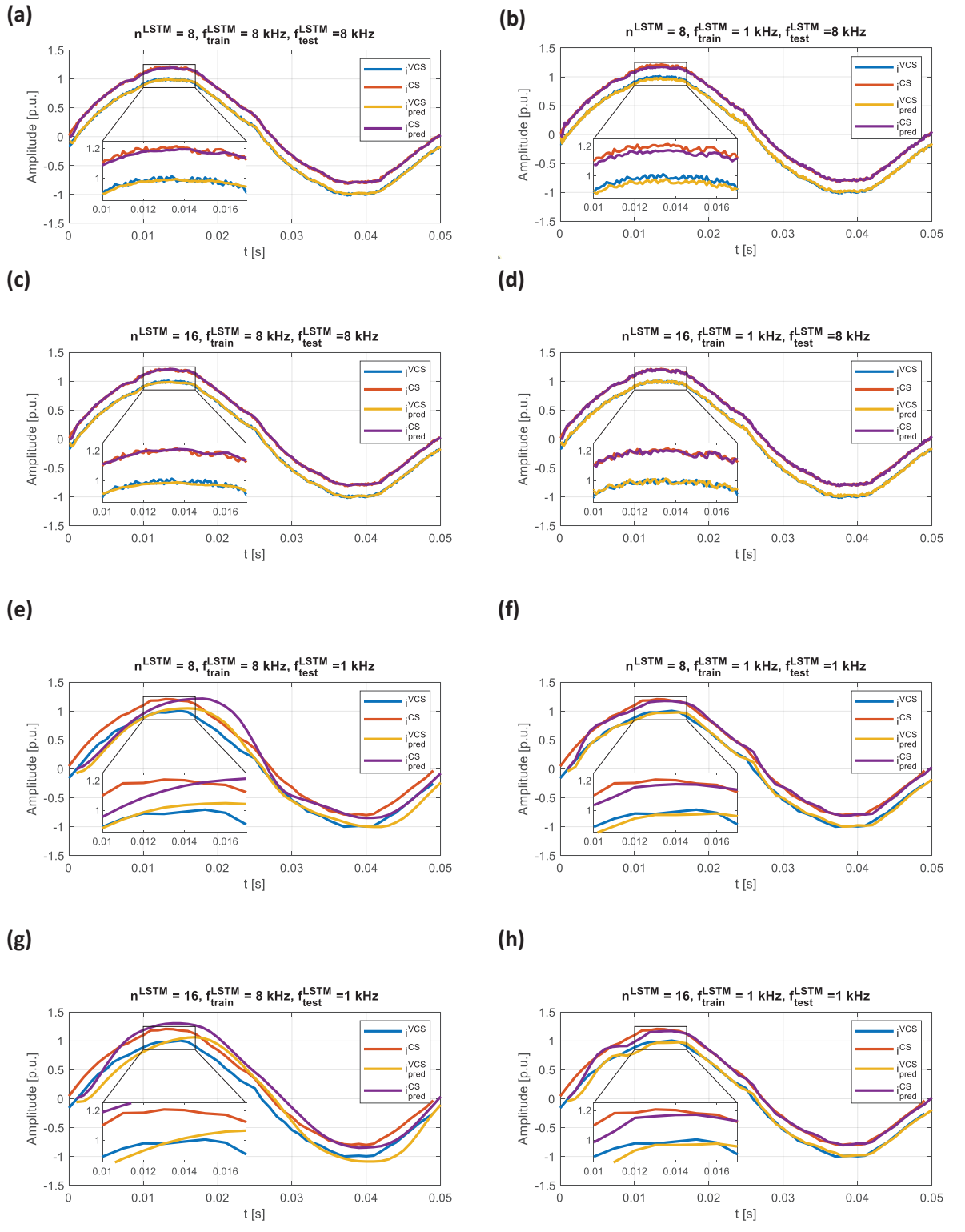

Fig. 8. Current signal prediction waveforms by networks containing different numbers of LSTM cells and trained for different training signal decimation values. Current signal parameters: f = 35 Hz, I_m = 0.4 p.u., and offset damage with I_{offset} = 0.3 p.u. Network operating frequency on test data: f_{test}^{LSTM} = 8 kHz (a, b, c, d) and f_{test}^{LSTM} = 1 kHz (e, f, g, h).
CS, current sensor; LSTM, long short-term memory; VCS, virtual CS.

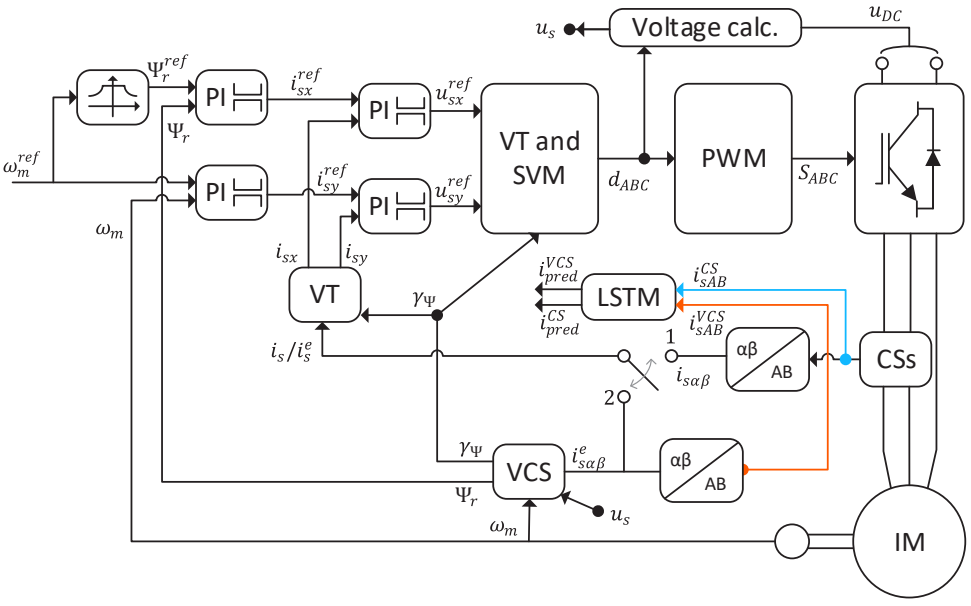

Fig. 9. DRFOC structure used in the study.

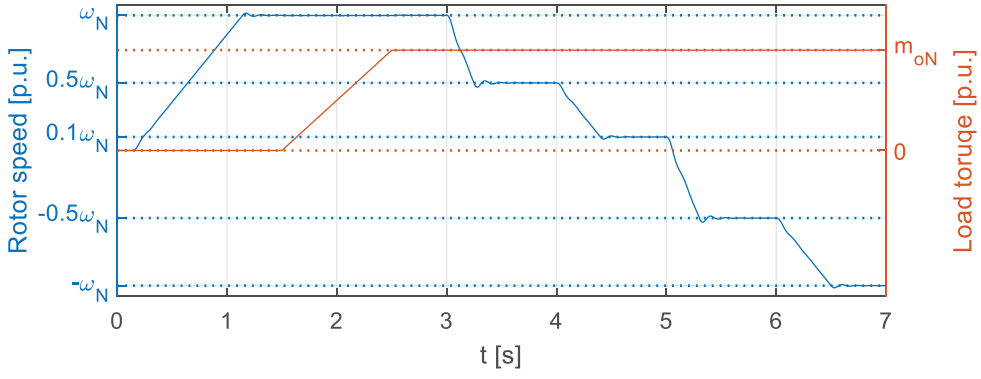

Fig. 10. Simulated rotor speed waveform (blue) and set load torque (red).

where: i^{VCS} – the value of the sinusoidal current signal estimated with VCS, and i^{CS} – the value of the measured current signal (e.g., in the case of a saturation-type fault). The output vector Y contained the prediction values of the signals fed into the network input, which were subscripted "pred":

$$Y(j)=\begin{bmatrix} i_{pred}^{VCS}(j+1) \\ i_{pred}^{CS}(j+1) \end{bmatrix}. \tag{23}$$

The current signals used in the training process were generated by software, as described in Section 2. Four different frequencies (f = 5, 20, 35, and 50 Hz) and five different amplitudes (I_m = 0.4, 0.55, 0.7, 0.85, and 1.0 p.u.) of the first harmonic were considered. For each amplitude and frequency, in addition to the correct signals, faulty signals were generated considering six different values of the gain ε at the gain fault (ε = 0.5, 0.7, 0.9, 1.1, 1.3, and 1.5), six different values of the constant I_{offset} in offset damage (I_{offset} = -0.3, -0.2, -0.1, 0.1, 0.2, and 0.3 p.u.), and six different values of the core saturation I_{sat} in saturation damage (I_{sat} = 0.70, 0.75,0.80, 0.85, 0.90, and 0.95 p.u.). The signal generator sampling frequency was f_p = 8 kHz. The length of the generated signals was one period of the first

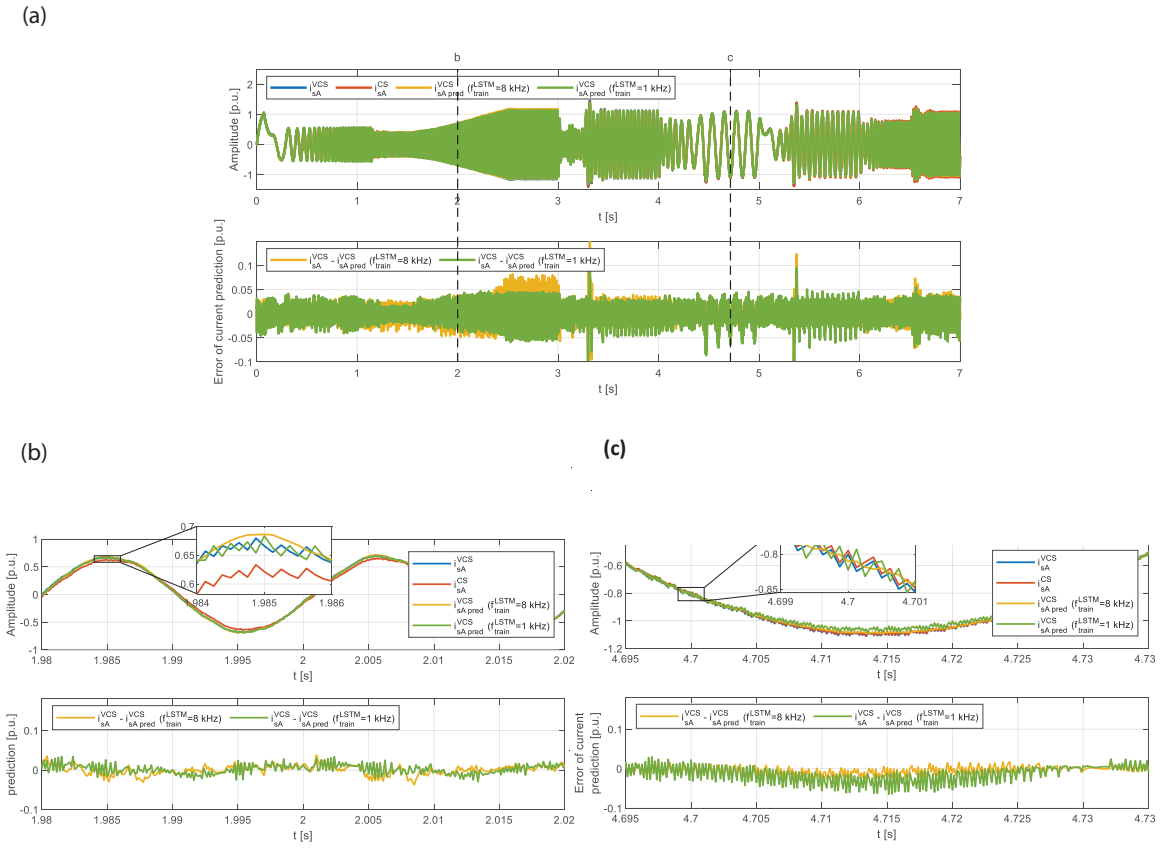

Fig. 11. Simulations of the LSTM network in the case of an undamaged CS: n^{LSTM} = 16, f_{train}^{LSTM} = 8 kHz, f_{test}^{LSTM} = 8 kHz (yellow); and n^{LSTM} = 16, f_{train}^{LSTM} = 1 kHz, f_{test}^{LSTM} = 8 kHz (green) in a DRFOC induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a). For the yellow waveforms, signal-smoothing properties are visible.
CS, current sensor; DRFOC, direct rotor flux-oriented control; LSTM, long short-term memory; VCS, virtual current sensor.

harmonic. The data set prepared this way was randomly divided into training data, further denoted in the text as X_{train} (342 signals: 90%) and test data X_{test} (38 signals: 10%). The corresponding training output data Y_{train} and Y_{test} were generated by shifting the input vector sequence by one sample.

During the training process, the input and output data were normalised by subtracting their mean value and dividing by the standard deviation according to the following relationships:

$$\tilde{X}_{train} = \frac{X_{train} - \mu_X}{\sigma_X}, \tag{24}$$

$$\tilde{Y}_{train} = \frac{Y_{train} - \mu_Y}{\sigma_Y}, \tag{25}$$

where $\tilde{X}_{train}$, $\tilde{Y}_{train}$ – input and output training vectors, respectively, after normalisation; μ_X, μ_Y – mean values of the input and output training set, respectively; σ_X, σ_Y – standard deviation from the input and output training set, respectively. When testing the network, the prediction results obtained are normalised values; so, in order to obtain the real values, the inverse operation to Eq. (25) must be applied on the output data according to the following formula:

$$Y = \tilde{Y} \cdot \sigma_Y + \mu_Y \cdot \tag{26}$$

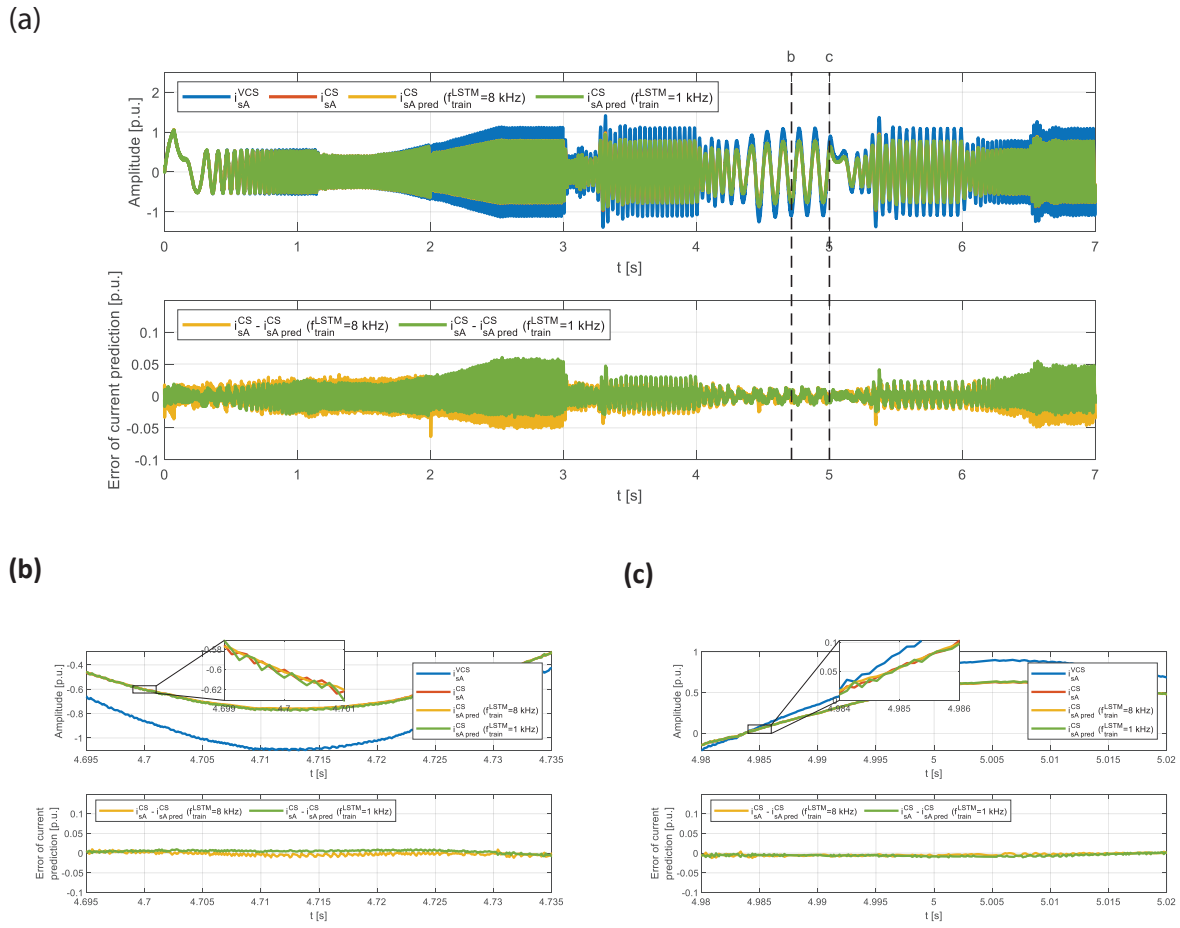

Fig. 12. Simulations of the LSTM network in the case of a CS gain-type fault: $n^{LSTM} = 16$, $f_{train}^{LSTM} = 8$ kHz, $f_{test}^{LSTM} = 8$ kHz (yellow), and $n^{LSTM} = 16$, $f_{train}^{LSTM} = 1$ kHz, $f_{test}^{LSTM} = 8$ kHz (green) in a DRFOC field-controlled induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a). For the yellow waveforms, signal-smoothing properties are visible.
CS, current sensor; DRFOC, direct rotor flux-oriented control; LSTM, long short-term memory; VCS, virtual current sensor.

Two networks containing $n^{LSTM} = 8$ and $n^{LSTM} = 16$ LSTM cells were included in the study. Each of these networks was trained for two different values of the input signal sampling frequency. In the first case, it was assumed that current signals from the generator with a sampling frequency of $f_p = 8$ kHz were fed to the input of the network sample by sample; so, the LSTM network was trained with X_{train} and Y_{train} signals with a sampling frequency of $f_{train}^{LSTM} = 8$ kHz. In the second case, the current signals were downsampled, and the sampling frequency of the X_{train} and Y_{train} network training signals was reduced to $f_{train}^{LSTM} = 1$ kHz, which corresponds to feeding the network input with every eighth sample of the generator signal.

LSTM networks were trained using the ADAM optimisation algorithm. The parameters of the optimisation algorithm, described in Subsection 3.2, were $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$. The number of epochs was $n_{epoch} = 1{,}000$ in each case, and in each epoch, all learning data were reshuffled.

The training process used $p = 342$ signal sequences. This was done using a minibatch data set. In each epoch, $m_b = 114$ sequences were randomly selected from all p sequences, based on which the values of the network parameters were updated. The length of each sequence, denoted by S_b, varied and depended on the frequency of the first harmonic of the current. The objective function for the minibatch set was equal to the arithmetic mean of the objective functions calculated for each of the sequences in the minibatch set according to the following formula:

$$E(\theta) = \frac{1}{m_b}\sum_{b=1}^{m_b}\left\{\frac{1}{2S_b}\sum_{j=1}^{S_b}\left\{\left[i^{VCS}(b,j) - i_{pred}^{VCS}(b,j)\right]^2 + \left[i^{CS}(b,j) - i_{pred}^{CS}(b,j)\right]^2\right\}\right\}. \quad (27)$$

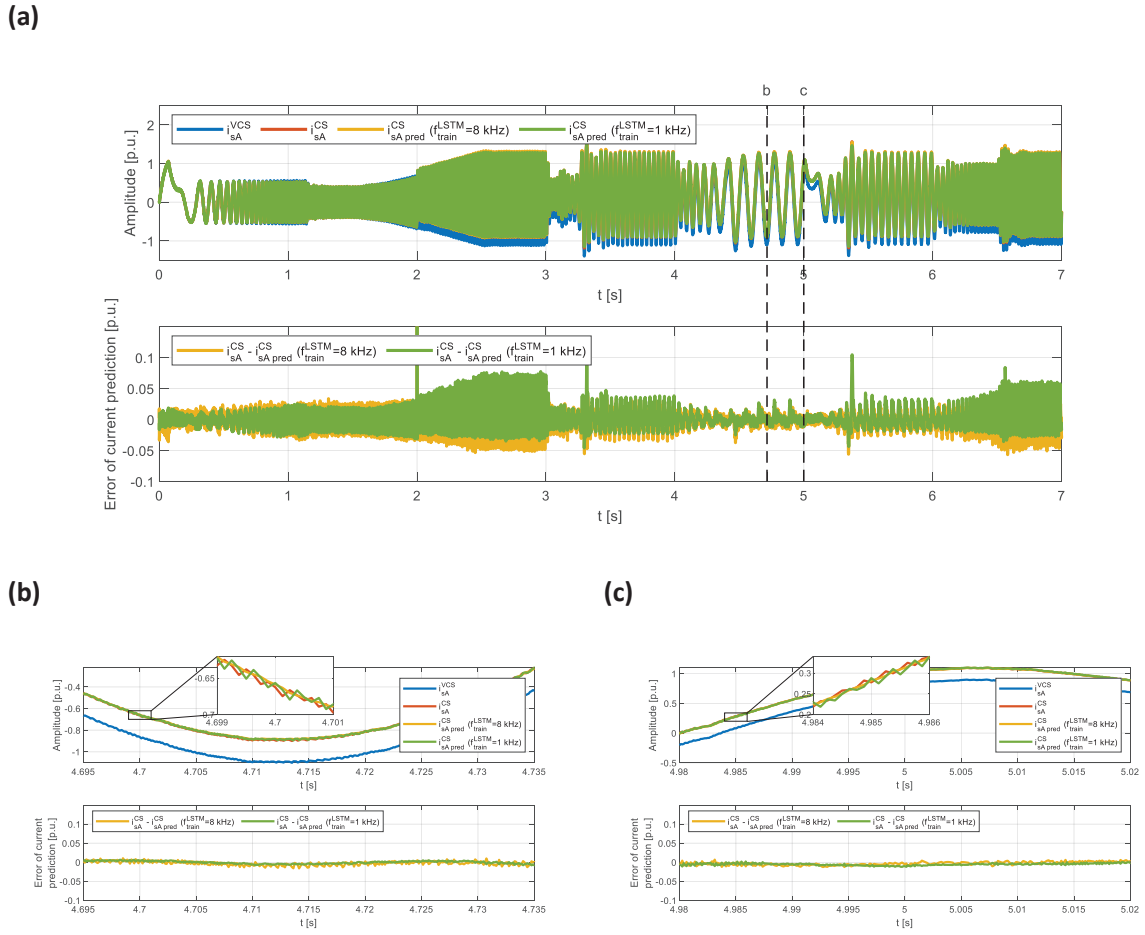

Fig. 13. Simulations of the LSTM network in the case of a CS offset-type fault: n^{LSTM} = 16, f_{train}^{LSTM} = 8 kHz, f_{test}^{LSTM} = 8 kHz (yellow); and n^{LSTM} = 16, f_{train}^{LSTM} = 1 kHz, f_{test}^{LSTM} = 8 kHz (green) in a DRFOC field-controlled induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a). For the yellow waveforms, signal-smoothing properties are visible.
CS, current sensor; DRFOC, direct rotor flux-oriented control; LSTM, long short-term memory; VCS, virtual current sensor.

In the next steps, a set of other sequences was selected, and the parameters were updated again until all *p* training sequences had been used for the epoch.

5. Analysis of Current Prediction Results

5.1. Assessing the quality of current prediction using LSTM networks

The comparative indicator of network performance quality and prediction precision was the root mean square error (RMSE) calculated between the prediction values of *Y* and the expected values of the Y_{test} for each test case *r* according to the following formula:

$$RMSE_r = \sqrt{\frac{1}{S_r}\sum_{j=1}^{S_r}\left[Y(r,j) - Y_{test}(r,j)\right]^2}\ , \tag{28}$$

where S_r is the length of the test sequence numbered *r*. Results showing histograms of the RMSE values for different network structures and the sampling frequencies of the training and testing signals are shown in Figure 7.

Networks with the number of cells n^{LSTM} = 16 obtained significantly better results in terms of the distribution of RMSE values in test studies than networks containing a smaller number of cells, i.e., n^{LSTM} = 8. However, increasing

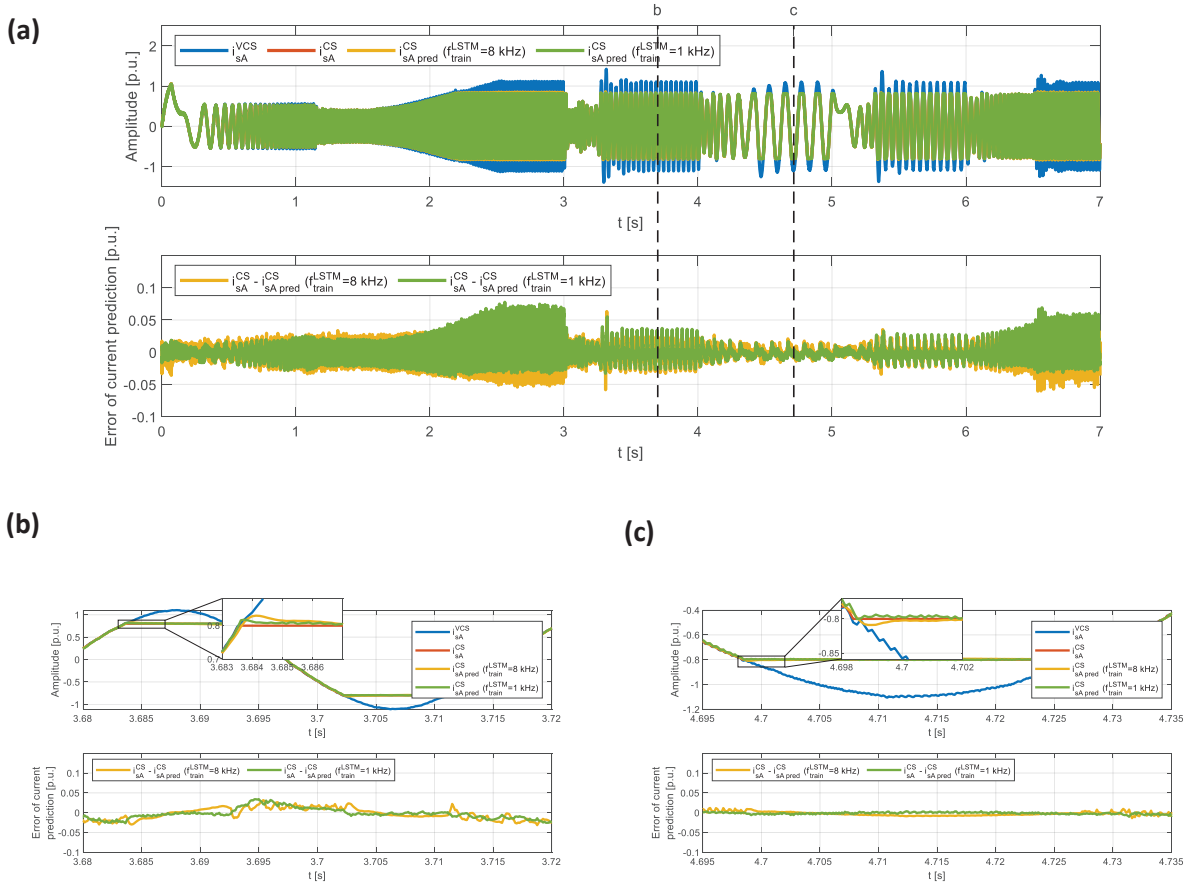

Fig. 14. Simulations of the LSTM network in the case of a CS saturation-type fault: $n^{LSTM} = 16$, $f_{train}^{LSTM} = 8$ kHz, $f_{test}^{LSTM} = 8$ kHz (yellow) and $n^{LSTM} = 16$, $f_{train}^{LSTM} = 1$ kHz, $f_{test}^{LSTM} = 8$ kHz (green) in a DRFOC field-controlled induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a). For the yellow waveforms, signal-smoothing properties are visible.
CS, current sensor; DRFOC, direct rotor flux-oriented control; LSTM, long short-term memory; VCS, virtual current sensor.

the number of cells in a network does not have the greatest impact on the quality of signal prediction. The most significant is the sampling rate of the training signals f_{train}^{LSTM} and the sampling rate of the test signals f_{test}^{LSTM}. The histograms clearly indicate significantly better performance of networks trained with signals with a lower sampling frequency f_{train}^{LSTM}, regardless of the sampling frequency of the test signals f_{test}^{LSTM}. Considering the results of the distribution of RMSE values shown in Figure 7, the best results were obtained for a network containing $n^{LSTM} = 16$ LSTM cells, trained with signals with sampling frequency $f_{train}^{LSTM} = 1$ kHz and tested with signals with sampling frequency $f_{test}^{LSTM} = 8$ kHz. Example network prediction plots for the test data are shown in Figure 8.

As with the analysis of the RMSE index, a visual comparison of the waveforms in Figure 8 reveals a marginally significant improvement in the quality of current prediction with an increase in the number of LSTM cells from $n^{LSTM} = 8$ to $n^{LSTM} = 16$ (Figures 8a, c; Figures 8b, d; Figures 8e, g; and Figures 8f, h). The sampling frequency of the training and test signal has a significant effect. Training the network at a training signal sampling rate $f_{train}^{LSTM} = 8$ kHz and then testing it on data with sampling rate $f_{test}^{LSTM} = 1$ kHz delays the performance of the prediction network (Figures 8e, g). This way of using the network disqualifies it from further use for signal analysis in an electric drive. The network trained for frequency $f_{train}^{LSTM} = 1$ kHz and tested for frequency $f_{test}^{LSTM} = 1$ kHz (Figures 8f, h) works correctly, but the prediction results are not ideal in the initial stage of operation. The network trained for frequency $f_{train}^{LSTM} = 8$ kHz and tested for frequency $f_{test}^{LSTM} = 8$ kHz also works correctly, but the effect of filtering noise from the signal is visible (Figures 8a, c). An interesting effect can be observed for a network that was trained on data with sampling frequency $f_{train}^{LSTM} = 1$ kHz and tested on data with sampling frequency $f_{test}^{LSTM} = 8$ kHz (Figures 8b, d).

Prediction error				Fault type			
		f_{train}^{LSTM}	f_{test}^{LSTM}	No fault (NF)	Gain (G)	Offset (OFF)	Saturation (SAT)
$ERR^{VCS} = i_{sA}^{VCS} - i_{sA\ pred}^{VCS}$	$\sigma(ERR^{VCS})$	8 kHz	8 kHz	0.0156	0.0138	0.0156	0.0144
		1 kHz		0.0171	0.0157	0.0152	0.0157
$ERR^{CS} = i_{sA}^{CS} - i_{sA\ pred}^{CS}$	$\sigma(ERR^{CS})$	8 kHz		0.0139	0.0134	0.0132	0.0143
		1 kHz		0.0154	0.0117	0.0152	0.0130

CS, current sensor; ERR, error; LSTM, long short-term memory; VCS, virtual current sensor.

Table 2. Standard deviation values for different CS faults and for different LSTM networks containing $n^{LSTM} = 16$.

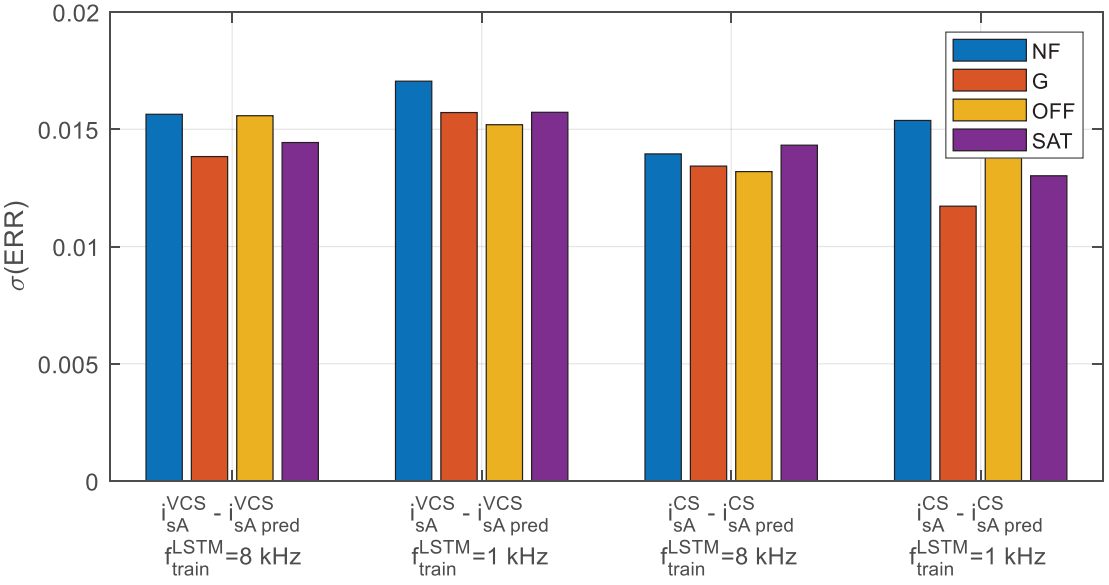

Fig. 15. Standard deviation of the prediction error for different CS faults and for different LSTM networks.
CS, current sensor; ERR, error; G, gain; LSTM, long short-term memory; NF, no fault; OFF, offset; SAT, saturation; VCS, virtual current sensor.

In this case, the network does not have the properties of filtering the signal but tries to reproduce it in such a way that it is as consistent as possible with the actual input.

5.2. Simulation tests of the LSTM predictor in the DRFOC structure

The predictor developed based on the LSTM network was tested in a DRFOC structure for an IM. This structure, together with the VCS current estimator and the LSTM network used, is shown in Figure 9.

Simulations were performed in the Matlab/Simulink environment with a computational step of $\Delta t_s = 6.25 \times 10^{-6}$ s, and the Euler method (ODE, ordinary differential equation; ode1) was used to solve the differential equations describing the simulation model. Simulations were carried out for different motor speeds ($-\omega_N$, $-0.5\omega_N$, $0.1\omega_N$, $0.5\omega_N$, and ω_N – rated speed of the IM). The load torque varied from 0 p.u. to the rated m_N during the simulation. The measured speed and load torque waveforms are shown in Figure 10.

The tests were carried out for two trained networks containing $n^{LSTM} = 16$ cells. One was trained with signals with sampling frequency $f_{train}^{LSTM} = 1$ kHz and the other with frequency $f_{train}^{LSTM} = 8$ kHz. Both networks were used to predict the signals from the CS and the VCS estimator in Phase A only. Their operating frequency was $f_{test}^{LSTM} = 8$ kHz. The selected networks were characterised by the best results obtained in the tests performed on the test data, Y_{test}.

Figure 11 shows the results of the simulations carried out in the DRFOC structure of the electric drive when the CS has not suffered any damage. The waveforms of the output signals from the LSTM network have been shifted by one sample on the graph, so that the actual measured or estimated CS signal and the prediction for that time instant can be shown. The prediction error signal, calculated as the difference between the measured current and the predicted current, was used to compare the performance quality of the two networks. As with the

Prediction error				Parameter change					
		f_{train}^{LSTM}	f_{test}^{LSTM}	0.75r_s	1.25r_s	0.75r_r	1.25r_r	0.75l_m	1.25l_m
$ERR^{VCS} = i_{sA}^{VCS} - i_{sA\ pred}^{VCS}$	$\sigma(ERR^{VCS})$	8 kHz	8 kHz	0.0112	0.0112	0.0181	0.0172	0.0246	0.0190
		1 kHz		0.0140	0.0147	0.0155	0.0186	0.0186	0.0172
$ERR^{CS} = i_{sA}^{CS} - i_{sA\ pred}^{CS}$	$\sigma(ERR^{CS})$	8 kHz		0.0135	0.0147	0.0144	0.0214	0.0199	0.0170
		1 kHz		0.0130	0.0137	0.0141	0.0156	0.0157	0.0145

CS, current sensor; ERR, error; l_m, the main inductance; LSTM, long short-term memory; r_s, stator winding resistance; r_r, the rotor resistance; VCS, virtual current sensor.

Table 3. Standard deviation values of prediction errors for different variations of induction motor parameters with undamaged CS for LSTM network containing $n^{LSTM} = 16$.

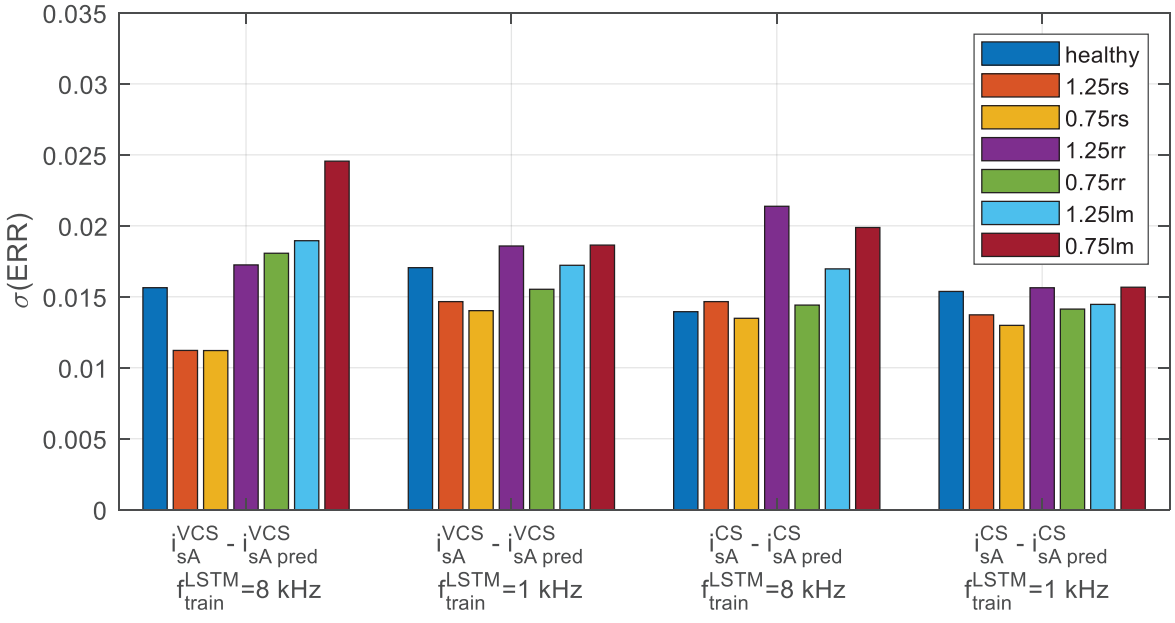

Fig. 16. Standard deviation of the prediction error when the IM parameters are changed for different LSTM networks. The number of LSTM cells was $n^{LSTM} = 16$ and the sampling frequency of the test signals was $f_{test}^{LSTM} = 8$ kHz.
CS, current sensor; ERR, error; IM, induction motor; l_m, the main inductance; LSTM, long short-term memory; r_s, stator winding resistance; r_r, the rotor resistance; VCS, virtual current sensor.

test results shown in Figure 8, the filtering properties of the network trained on a signal sampled at a frequency $f_{train}^{LSTM} = 8$ kHz can be seen in Figure 11 (yellow waveforms). In each of the time windows shown, this network performs prediction and simultaneous filtering of the signal, in contrast to the network that was trained on a signal sampled at $f_{train}^{LSTM} = 1$ kHz (green waveforms). The second network ($f_{train}^{LSTM} = 1$ kHz), on the other hand, makes a prediction of the signal while retaining the noise present.

The next simulations were performed in a system with a faulty CS in Phase A. The simulated faults included gain change (Figure 12), addition of a fixed value to the measurements (offset) (Figure 13), and core saturation (Figure 14). The faults were implemented at time $t = 2$ s and, from this point on, the operation of the speed control system used the currents estimated by the VCS estimator in its structure, and the signals from the faulty CSs did not affect the operation of the system. The rotor speed and load torque waveforms were as before (Figure 10). In each case, the same observations can be made as previously, i.e., the manifestation of the filtering properties of the network trained on data with sampling frequency $f_{train}^{LSTM} = 8$ kHz (yellow waveforms) and the absence of these properties for the network trained on data with sampling frequency $f_{train}^{LSTM} = 1$ kHz (green waveforms).

To assess the quality of the prediction, the standard deviation of the prediction error was calculated as the difference between the actual signal and the prediction signal. Networks were tested on a signal sampled at a

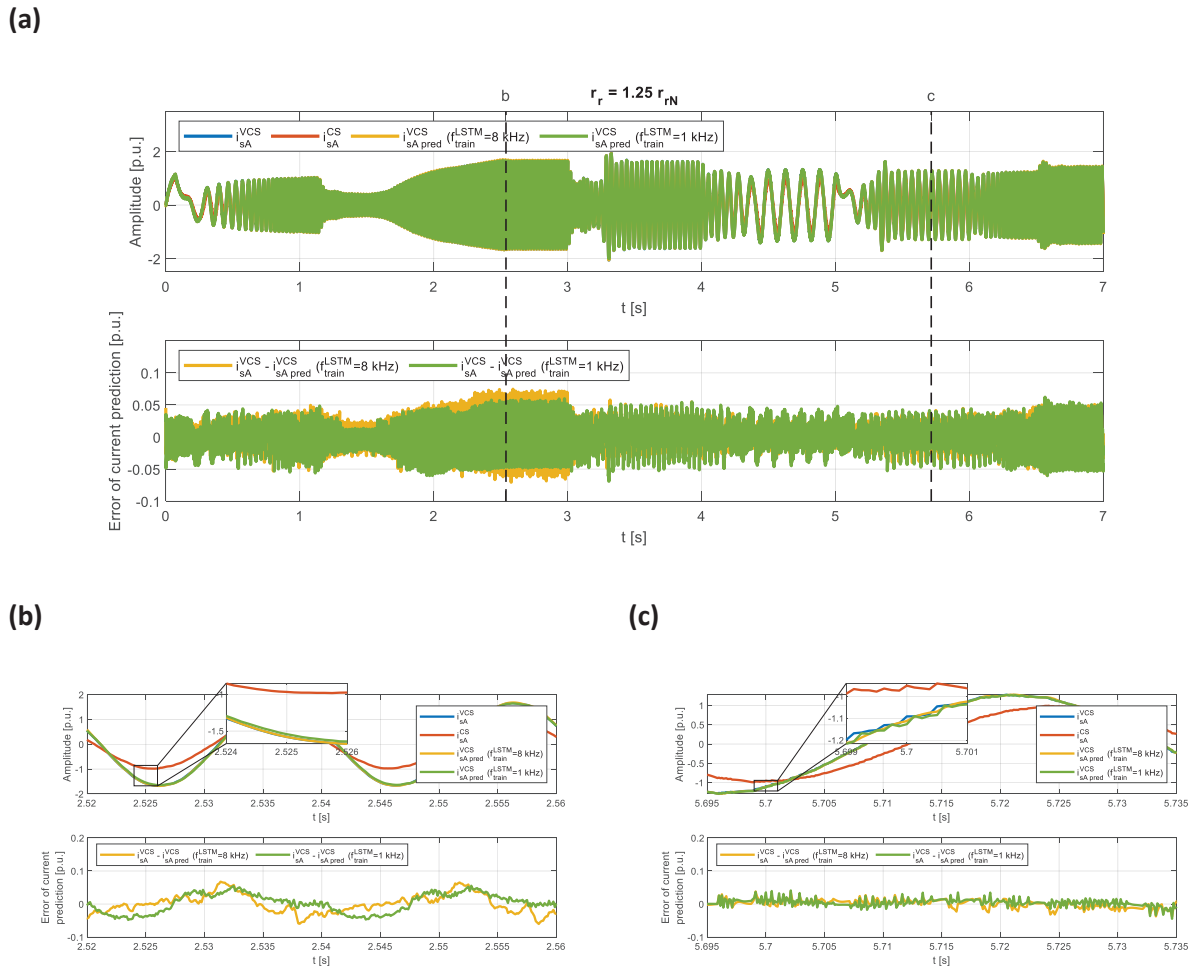

Fig. 17. Simulations showing the prediction of VCS currents using the LSTM network when rotor resistance r_r changes: $r_r = 1.25 r_{rN}$, $n^{LSTM} = 16$, $f_{train}^{LSTM} = 8$ kHz, $f_{test}^{LSTM} = 8$ kHz (yellow); and $n^{LSTM} = 16$, $f_{train}^{LSTM} = 1$ kHz, $f_{test}^{LSTM} = 8$ kHz (green) in a DRFOC induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a).
CS, current sensor; DRFOC, direct rotor flux-oriented control; ERR, error; l_m, the main inductance; LSTM, long short-term memory; r_s, stator winding resistance; r_r, the rotor resistance; VCS, virtual current sensor.

frequency $f_{test}^{LSTM} = 8$ kHz. The statistics were determined for the entire run, i.e., from 0 s to 7 s of simulation, and are summarised in Table 2 and plotted in Figure 15.

By analysing the data in Figure 15, it can be seen that the standard deviation values calculated for the different CS faults and for the different LSTM networks are comparable. This means that the networks trained both for $f_{train}^{LSTM} = 1$ kHz and for $f_{train}^{LSTM} = 8$ kHz can be considered as good predictive models of stator currents, regardless of the type of damage that occurred to the CS.

5.3. Testing the robustness of the LSTM predictor to changes in IM parameters

The described current prediction method was tested for robustness to changing machine parameters. The parameters that were changed were the stator winding resistance r_s, the rotor resistance r_r, and the main inductance l_m. Changing the values of these parameters has the greatest impact on the performance of the VCS estimator (Adamczyk and Orlowska-Kowalska, 2022). Simulation tests were carried out in which the CSs measured the correct value and were not damaged in any way. The parameter variations were 75% or 125% of the rated values (given in Table A1 of the Appendix). The standard deviation results for a network containing $n^{LSTM} = 16$ cells, tested for an input signal sampling frequency $f_{test}^{LSTM} = 8$ kHz and trained for two different sampling frequencies are shown in Table 3 and Figure 16.

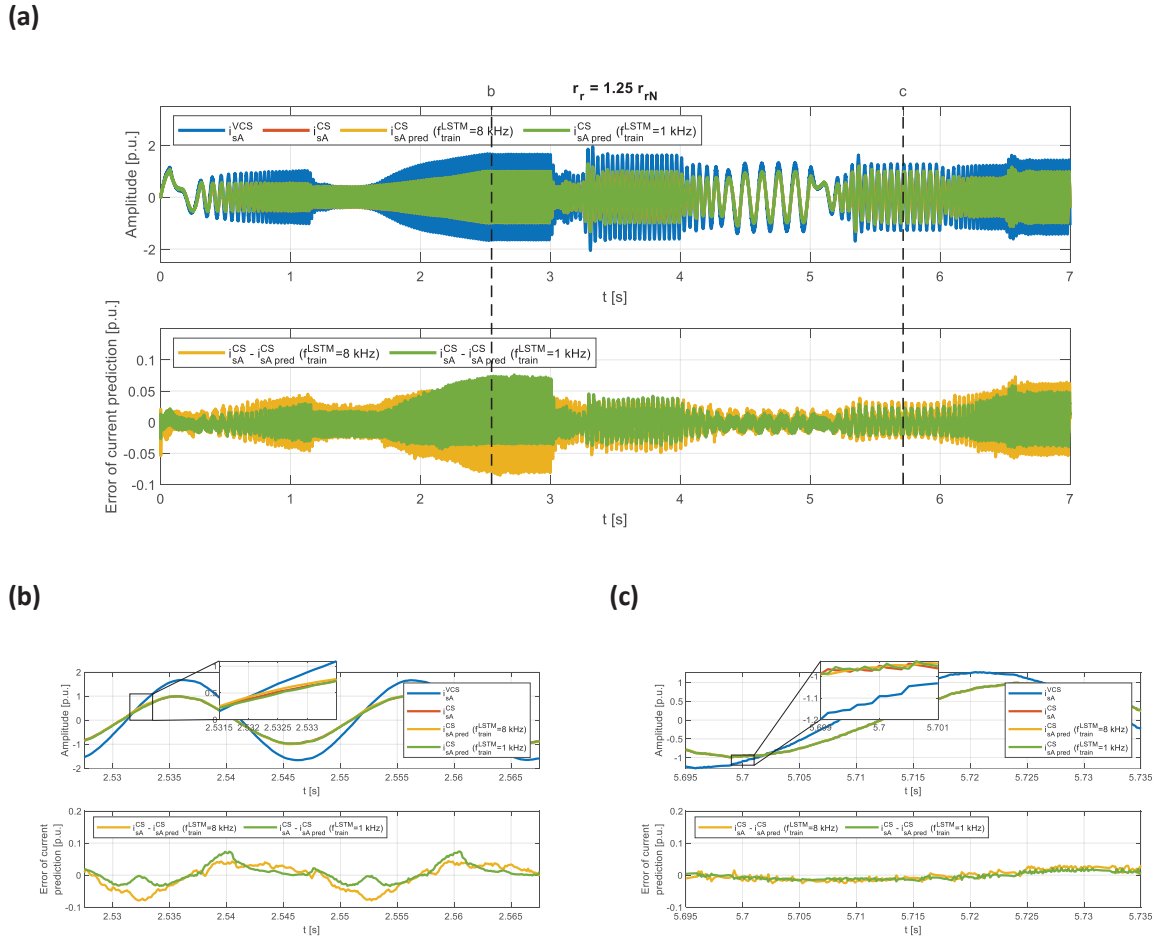

Fig. 18. Simulations showing the prediction of CS currents using the LSTM network when rotor resistance r_r changes: $r_r = 1.25r_{rN}$, $n^{LSTM} = 16$, $f_{train}^{LSTM} = 8$ kHz, $f_{test}^{LSTM} = 8$ kHz (yellow) and $n^{LSTM} = 16$, $f_{train}^{LSTM} = 1$ kHz, $f_{test}^{LSTM} = 8$ kHz (green) in a DRFOC induction motor drive. Current waveforms in Phase A from the VCS estimator, CS, and LSTM predictor and the corresponding prediction error value (a), example approximations of the time instants (b, c) marked on the waveforms from (a).
CS, current sensor; DRFOC, direct rotor flux-oriented control; ERR, error; l_m, the main inductance; LSTM, long short-term memory; r_s, stator winding resistance; r_r, the rotor resistance; VCS, virtual current sensor.

The results shown above clearly demonstrate the robustness of the developed predictor to changing motor parameters. The results of the standard deviation of the prediction errors for the system operating at nominal parameters and for the system with changed parameters are comparable. Example waveforms for $r_r = 1.25r_{rN}$ are shown in Figures 17 and 18.

The significant differences between the currents estimated with VCS and those measured with CS (Figures 17 and 18) are due to changes in machine parameters and are normal for this estimator, when parameter updates are not applied (Adamczyk and Orlowska-Kowalska, 2019, 2022). However, they do not affect the prediction values of the currents and the developed LSTM predictor correctly determines the subsequent values of the signals. Thus, it can be concluded that the proposed current prediction method is robust to changes in motor parameters.

6. Conclusions

The results of tests carried out on a network composed of LSTM cells, whose task was to predict the currents in the stator phases, are presented. The prediction was based on the currently measured and estimated values. A comparison was made between different structures, which were trained and tested with software-generated signals at different sampling frequencies. Particularly interesting properties were discovered regarding the noise filtering

at the output of the network, as well as how to train and use the network to obtain the output filter. In the future, it may be important to use a filter on the output of the network to detect stator CS faults. The same effect can be achieved by training the network with appropriate signals, which will reduce the amount of computation required. On the other hand, in predictive control, it is important to obtain the most accurate current values possible. Based on the results, it can be concluded that testing the network with a signal sampled at a lower frequency than that at which the network was trained significantly delays signal prediction (Figures 8e, g). Testing the network with a signal sampled at the same frequency as the training signal allows filtering out of the input noise (Figures 8a, c). Testing the network with a signal sampled at a higher frequency than that at which the network was trained allows the signal to be reconstructed while retaining the input noise (Figures 8b, d). Particularly noteworthy is the fact that the LSTM network trained on software-generated current signals (one period of the signal with different frequencies and amplitudes) correctly reproduced the IM stator currents in the DRFOC structure, in the case of various CS faults. The tests carried out can be used in the future to respond to emerging CS faults in advance, thus maintaining the continuity of drive control.

Acknowledgment

This research was funded by the National Science Center, Poland, under Project No. 2021/41/B/ST7/02971.

References

Adamczyk, M. and Orlowska-Kowalska, T. (2019). Virtual Current Sensor in the Fault-Tolerant Field-Oriented Control Structure of An Induction Motor Drive. *Sensors (Switzerland)*, 19(22), art. 4979.

Adamczyk, M. and Orlowska-Kowalska, T. (2022). Postfault Direct Field-Oriented Control of Induction Motor Drive Using Adaptive Virtual Current Sensor. *IEEE Transactions on Industrial Electronics*, 69(4), pp. 3418–3427.

Blanke, M., Kinnaert, M., Lunze, J. and Staroswiecki, M. (2006). *Diagnosis and Fault-Tolerant Control*, 2nd ed. Heidelberg: Springer Berlin.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. The MIT Press.

Ha, J. I. (2010). Current Prediction in Vector-Controlled PWM Inverters Using Single DC-Link Current Sensor. In: *IEEE Transactions on Industrial Electronics*. February 2010, pp. 716–726.

Im, J. H. and Kim, R. Y. (2018). Improved Saliency-Based Position Sensorless Control of Interior Permanent-Magnet Synchronous Machines with Single DC-Link Current Sensor Using Current Prediction Method. *IEEE Transactions on Industrial Electronics*, 65(7), pp. 5335–5343.

Kingma, D. P. and Ba, J. (2015). Adam: A Method for Stochastic Optimization. In: *Proceedings of the 3rd International Conference for Learning Representations,* San Diego, December 2015.

Lee, K. S. and Ryu, J. S. (2003). Instrument Fault Detection and Compensation Scheme for Direct Torque Controlled Induction Motor Drives. *IEE Proceedings: Control Theory and Applications,* 150, pp. 376–382.

Leonhard, W. (1996). *Control of Electrical Drives,* 2nd ed. Berlin, Heidelberg: Springer Berlin Heidelberg.

Li, J. and Akilan, T. (2022). Global Attention-based Encoder-Decoder LSTM Model for Temperature Prediction of Permanent Magnet Synchronous Motors. arXiv preprint http://arxiv.org/abs/2208.00293.

Li, Z., Gao, Q. and Zhang, W. (2016). An improved current sampling scheme using three resistors for induction motor drives based on current prediction. In: *2016 IEEE 8th International Power Electronics and Motion Control Conference, IPEMC-ECCE Asia 2016*. 13 July 2016, Hefei, China, pp. 104–108.

Li, P., Liao, Y., Lin, H. and Yan, L. (2019). An Improved Three-Phase Current Reconstruction Strategy Using Single Current Sensor with Current Prediction. In: *2019 22nd International Conference on Electrical Machines and Systems (ICEMS).* 11–14 August 2019, Harbin, China.

Orlowska-Kowalska, T. and Dybkowski, M. (2016). Industrial Drive Systems. Current State and Development Trends. *Power Electronics and Drives*, 1(36), pp. 5–25.

Peralta-Sanchez, E., Al-Rifai, F. and Schofield, N. (2009). Direct torque control of permanent magnet motors using a single current sensor. In: *2009 IEEE International Electric Machines and Drives Conference, IEMDC '09,* Miami, FL, USA, 2009, pp. 89–94.

Skowron, M., Teler, K., Adamczyk, M. and Orlowska-Kowalska T. (2022). Classification of Single Current Sensor Failures in Fault-Tolerant Induction Motor Drive Using Neural Network Approach. *Energies*, 15(18), art. 6646.

Tunia, H., Smirnow, A., Nowak, M. and Barlik, R. (1982). *Power Electronics Systems – Calculation, Modeling, Design*. Scientific-Technical Editorial Office (In Polish).

Wang, X., Zhang, Y., Yang, H., Zhang, B., Rodriguez, J. and Garcia, C. (2020). A Model-Free Predictive Current Control of Induction Motor Based on Current Difference. In: *2020 IEEE 9th International Power Electronics and Motion Control Conference, IPEMC 2020 ECCE Asia*. 29 November 2020, Nanjing, China, pp. 1038–1042.

Wang, W., Yan, H., Xu, Y., Zou, J. and Buticchi, G. (2022). Improved Three-Phase Current Reconstruction Technique for PMSM Drive with Current Prediction. *IEEE Transactions on Industrial Electronics*, 69(4), pp. 3449–3459.

Yan, L., Wang, F. and Rodriguez, J. (2020). Luenberger prediction model-based robust predictive current control of induction machine drives. In: *2020 IEEE 9th International Power Electronics and Motion Control Conference, IPEMC 2020 ECCE Asia*. 29 November 2020, Nanjing, China, pp. 1006–1010.

Ziegler, S., Woodward, R. C., Iu, H. H. C. and Borle, L. J. (2009). Current Sensing Techniques: A Review. *IEEE Sensors Journal*, 9(4), pp. 354–376.

Appendix

Symbol	[ph.u.]	[p.u.]
Rated phase voltage, U_N	230 V	0.707
Rated phase current, I_N	2.5 A	0.707
Rated power, P_N	1.1 kW	0.638
Rated speed, n_N	1390 rpm	0.927
Rated torque, M_{eN}	7.56 Nm	0.688
Number of pole pairs, p_b	2	-
Rotor winding resistance, r_r	4.968 Ω	0.0540
Stator winding resistance, r_s	5.114 Ω	0.0556
Rotor leakage inductance, $l_{\sigma r}$	31.6 mH	0.1079
Stator leakage inductance, $l_{\sigma s}$	31.6 mH	0.1079
Main inductance, l_m	541.7 mH	1.8498
Rated rotor flux, Ψ_{rN}	0.7441 Wb	0.7187
Mechanical time constant, T_M	0.25 s	-

Table A1. Induction motor parameters.