Deep Learning Based Predictive Compensation of Flicker, Voltage Dips, Harmonics and Interharmonics in Electric Arc Furnaces

Ebrahim Balouji 1

1Chalmers University of Technology

October 30, 2023

Abstract

In this research work, deep machine learning based methods together with a novel data augmentation are developed for predicting flicker, voltage dip, harmonics and interharmonics originating from highly time-varying electric arc furnace (EAF) currents and voltage. The aim with the prediction is to counteract both the response and reaction time delays of active power filters (APFs) specifically designed for electric arc furnaces (EAF). Multiple synchronous Reference frame (MSRF) analysis is used to decompose the frequency components of the EAF current and voltage waveforms into dqo components. Then using low-pass filters and prediction of the future values of these dqo components, reference signals for APFs are generated. Three different methods have been developed. In two of them, a low-pass Butterworth filter is used together with a linear FIR based prediction or long short-term memory network (LSTM) for prediction. In the third method, a deep convolutional neural network (CNN) combined with a LSTM network is used to filter and predict at the same time. For a 40 ms prediction horizon, the proposed methods provide 2.06%, 0.31%, 0.99% prediction errors of the dqo components for the Butterworth and linear prediction, Butterworth and LSTM and CNN with LSTM, respectively. The error of the predicted reconstructed waveforms of flicker, harmonics, and interharmonics resulted in 8.5%, 1.90%, and 3.2% reconstruction errors for the above-mentioned methods. Finally, a Simulink and GPU based implementation of predictive APF using Butterworth filter + LSTM and a trivial APF resulted 96% and 60% efficiency on compensation of EAF current interharmonics.
Deep Learning Based Predictive Compensation of Flicker, Voltage Dips, Harmonics and Interharmonics in Electric Arc Furnaces

Ebrahim Balouji1, Özgül Salor2,3, Senior Member, IEEE, Tomas McKelvey1, Senior Member, IEEE
1Chalmers University of Technology, Dept. of Electrical Engineering, Gothenburg, Sweden
2Gazi University, Dept. of Electrical and Electronics Engineering, Ankara, Turkey
3Kyrgyz-Turkish Manas University, Dept. of Electrical and Electronics Engineering, Bishkek, Kyrgyzstan

Abstract—In this research work, deep machine learning based methods together with a novel data augmentation are developed for predicting flicker, voltage dip, harmonics and interharmonics originating from highly time-varying electric arc furnace (EAF) currents and voltage. The aim with the prediction is to counteract both the response and reaction time delays of active power filters (APFs) specifically designed for electric arc furnaces (EAF). Multiple synchronous Reference frame (MSRF) analysis is used to decompose the frequency components of the EAF current and voltage waveforms into $dqo$ components. Then using low-pass filters and prediction of the future values of these $dqo$ components, reference signals for APFs are generated. Three different methods have been developed. In two of them, a low-pass Butterworth filter is used together with a linear FIR based prediction or long short-term memory network (LSTM) for prediction. In the third method, a deep convolutional neural network (CNN) combined with a LSTM network is used to filter and predict at the same time. For a 40 ms prediction horizon, the proposed methods provide $2.06\%$, $0.31\%$, $0.99\%$ prediction errors of the $dqo$ components for the Butterworth and linear prediction, Butterworth and LSTM and CNN with LSTM, respectively. The error of the predicted reconstructed waveforms of flicker, harmonics, and interharmonics resulted in $8.5\%$, $1.90\%$, and $3.2\%$ reconstruction errors for the above-mentioned methods. Finally, a Simulink and GPU based implementation of predictive APF using Butterworth filter + LSTM and a trivial APF resulted $96\%$ and $60\%$ efficiency on compensation of EAF current interharmonics.

Index Terms—active power filter, Butterworth filter, convolutional neural networks, deep learning, long short-term memory, linear prediction, multiple synchronous reference frame, electric arc furnace, flicker, harmonics, interharmonics, voltage dip.

I. INTRODUCTION

Electric arc furnaces (EAF) in an electricity network are often the largest consumers of electric power and hence the most power-quality (PQ)-threatening customers of the utility. For example, a single medium-sized EAF can induce a sudden demand for 100 MVA electric power. Such sudden peaks of energy demands repeated frequently at random intervals of a standard tap-to-tap time of an EAF, causes EAFs to be categorized as non-stationary loads, with undesired PQ events and variations such as flicker and voltage dips. Due to their operation principles, EAFs are also highly non-linear loads, which draw currents from the power system rich in interharmonics and non-characteristic harmonics. These current harmonics often contribute significantly to the increase of voltage harmonics and interharmonics, in the transmission part of the power systems. Moreover, during resonance conditions of electric power systems, the amplitudes of these undesired frequency components can increase and damage or cause malfunction of some electronic devices. One of the most common solutions to such problematic issues is to employ active power filters (APF). These devices continuously make estimations of amplitudes and phases of the undesired frequency components and mitigated those components by supplying currents with the same amplitudes but half-cycle-shifted phases. Hence, each regenerated current component should ideally be out-of-phase and exactly at the same amplitude so that perfect compensation is achieved at each harmonic or interharmonic frequency. Therefore, one of the most important key requirements in these types of methods is a fast and accurate estimation of the amplitudes and phases of the undesired frequency components, since inaccurate or delayed estimations could even cause amplification of the undesired current components. Having estimated the components to be mitigated, the second key requirement is a fast and accurate reaction of APFs hardware to supply the required mitigating currents to the system. Therefore improving the performances of APFs can be investigated in two main areas: algorithmic quality and hardware aspects. To improve the algorithmic quality, studies have been focused on reducing time delays of filters [1][3], reducing the computation times of the control algorithms [1][4][7], and reducing the time spent on other algorithms [8][11]. The time delay due to the algorithms is called response time delay [12][17], whereas time delay caused by the hardware limitations is called reaction time delay [17][21]. For hardware improvement, researchers have tried to adapt faster data acquisition units [22][24], more accurate and faster measurement sensors [25][27], to increase
the switching frequency of power electronics switches \cite{28,30} and to optimize LC filter parameters \cite{31,35}.

All these aforementioned studies provide valuable contributions; however, the contributions to algorithms to reduce response time delays remain limited. Moreover, solutions to the reaction time delays, such as generating the switches with the possibility of higher switching frequencies are quite expensive and limited. Using larger inductance and capacitance in LC filter design will help to mitigate the high-frequency harmonics and being less sensitive to impedance changes of the grid but this will in-turn add more delay to the overall reaction time of the APF. On the other hand, fast and more accurate data acquisition units and sensors are also expensive and still, the most expensive ones have delays. Very limited studies had been proposed to tackle both algorithmic and hardware efficiency of APFs.

In this paper three different approaches have been employed to address the above-mentioned problems of EAF-current compensation by reducing the response time delay. A multiple synchronous reference frame (MSRF) analysis is used to decompose the frequency components of the EAF current and voltage waveforms into \textit{dqo} components. Then using low-pass filters and prediction of the future values of these \textit{dqo} components, reference signals for APFs are generated. Three different methods have been developed with the aim to reduce the response time delay. The first approach employs a low-pass Butterworth filter together with linear prediction. In the second approach, long short-term memory (LSTM) deep learning (DL) method together with a low-pass Butterworth filter is used. With this method enough time can be gained to mitigate reaction and response time delays. In the final approach, a convolutional neural network (CNN) together with LSTM is proposed which combines the filter and prediction functions in one network. All of these approaches differ from the previous work of the authors in \cite{36} such that the work in \cite{36} mostly contributes on the mitigation of algorithmic delays only. The work presented in this paper proposes three different approaches to mitigate both algorithmic and hardware delays by increasing the prediction accuracy and horizon. Furthermore, besides predicting EAF current harmonics and interharmonics, solutions for efficient compensation of voltage fluctuations (flicker) and voltage dips originating from EAFs are also proposed. Finally, we extend the conference version of this paper \cite{37} with implementation of the above mentioned methods on embedded computing system and an APF model in the Simulink environment. The results of employing the predictive methods in an APF are compared with a traditional non-predictive APF. The contributions and some important key-points of the three predictive methods presented in this research work can be listed as follows:

- With a novel proposed data augmentation method we can significantly increase the prediction accuracy
- Algorithmic and hardware delays can be mitigated thanks to the successful prediction of the future waveforms
- The methods are designed to compensate for several PQ problems at the same time, namely, harmonics, interharmonics, flicker, and voltage dips, based on the fact that they all appear to be interharmonics in the spectrum
- It is possible to do soft switching by predicting the dynamic voltage and current waveforms and this reduces the switching losses
- Since the filtering in the proposed method where CNN and LSTM are combined, there is no need to design an explicit low-pass filter
- Inductance and capacitance values of the LCL filter of the APF can be optimized and the problems coming from the grid side can be reduced. This can make it possible for APF/inverters to have less conduction losses and be less sensitive to network impedance changes and more important than that, to the frequency deviations of the electric power systems.
- A Virtual design of the ML based APF prototype is successfully implemented using the Simulink environment and embedded AI computing system Jetson TX2.

II. EVERYTHING IS INTERHARMONICS

A. Flicker and voltage dip modeled as interharmonics

The flicker phenomenon is an objectionable consequence of the random or periodic fluctuations on the voltage waveform envelope. These fluctuations have lower frequencies compared to the fundamental frequency and it can be shown that flicker appears as interharmonics around the fundamental and the harmonics in the frequency spectrum of the waveform \cite{38,41}. According to the standards IEEE 1159 and IEC 61000-4-15, flicker can be expressed as an amplitude modulation (AM) signal with a carrier frequency which corresponds to the fundamental frequency of the power system \cite{42,44}.

\[
y(t) = (A + m(t))c(t) = (A + M \cos (w_m t + \phi)) \sin (w_c t)
\]

where \(M\) and \(w_m\) are the amplitude and frequency of the flicker fluctuation and \(A\) and \(w_c\) are the amplitude and frequency of the power system, respectively. Furthermore, \( y(t) \) in (1) can be reformulated as,

\[
y(t) = A \sin (w_c t) + \frac{M}{2} \left[ \sin \left( (w_c + w_m) t + \phi \right) + \sin \left( (w_c - w_m) t - \phi \right) \right].
\]

(2)

As it can be seen in (2) the flickers which is the lower frequency fluctuation at \(w_m\) of the fundamental amplitude, appears as interharmonics at frequencies \(w_c - w_m\) and \(w_c + w_m\), around the main component. Note that, flicker or any other fluctuation, does not need to be exactly sinusoidal shaped. However, it is also known that any envelope shape can also be modeled as the summation of various low frequency components. Then more number of frequency components will appear around the fundamental frequency, \(w_c\). It can also be stated that the flicker frequency components will appear
as interharmonics around the existing harmonics as well. For example in the case of a second harmonic in the system equation (2) can be formulated as,

\[ y(t) = (A + M \cos(w_m t + \phi)) [\sin(w_c t) + M_2 \sin(2w_c t)] \]

and \( y(t) \) can be expressed as

\[
y(t) = A \sin (w_c t) + AM_2 \sin (2w_c t) \\
+ \frac{M}{2} \sin [(w_c + w_m) t + \phi] + \sin [(w_c - w_m) t + \phi] \\
+ \frac{M}{2} \sin [(2w_c + w_m) t + \phi] + \sin [(2w_c - w_m) t + \phi] 
\]

(4)

where \( AM_2 \) corresponds to the amplitude of the second harmonic. Here, for the sake of simplicity it is assumed that the fundamental component and second harmonic have the same phase difference \( \phi \) with respect to the flicker frequency component.

Furthermore, it can be stated that voltage dips can be modelled as changes of the waveform envelope in a random manner and that envelope can also be modelled as the modulation given in (1) with several modulating frequency components, whose summation forms the envelope. Therefore, it is possible to say that, a voltage dip will also appear as interharmonics around the fundamental and the existing harmonics. Hence, any method which quickly and accurately mitigates interharmonics around the fundamental frequency, will also compensate flicker and voltage dips. In order to be able compensate voltage dips, they should be predicted as soon as they start to appear in a few cycles of the fundamental and the APF should start compensating it by mitigating the related interharmonics immediately.

**B. A Solution and its Drawbacks**

In the past, many studies proposed solutions to estimate and compensate interharmonics [16], [45]. Only limited solutions have been presented for flicker compensation. To extract the interharmonics and harmonics components the proposed MSRF-based method in [11], [46] is used in this paper and is illustrated in Fig. 1. For example, to extract the 45 Hz frequency components, \( w \) in Fig. 1 needs to be defined as \( w = 2\pi f_0 \). In the extraction process of the phase and amplitude of the 45 Hz frequency component in the dqo frame, the low-pass filter block is a key component in the analysis chain and is needed due to time variation together with the effect of other frequency components in the signal, e.g. the blue signal shown in Fig. 2. After low-pass filtering of the dqo frame component, the reverse operation dqo\(^{-1}\) can be used to obtain the pure 45 Hz component.

The choice of LP-filter in the MSRF analysis chain is important to obtain high performance. One of the well-known choices in the industry is the Butterworth filter for low-pass filtering of dqo components. In this research work, a 4th order Butterworth filter with 1.5 Hz cutoff frequency is used for low-pass filtering and an example of using this filter on filtering of \( d^+ \) component is shown in Fig. 2. The drawback of a this causal Butterworth filter is that it has a settling time delay of almost 0.8 cycle of the fundamental frequency (16 ms for 50.0 Hz) in following the trend of the \( d^+ \) components. This delay is measured by comparing the results obtained using a zero-phase digital filter (in MATLAB called `filtfilt`) which implements the non-causal filter and is only applicable for off-line analysis [47]. Later these values will be used as the desired true values for training ML methods to predict the future values of dqo components. The settling time delay of the Butterworth filter is one part of the total algorithmic delay. Since hardware delays also exist it is desirable to be able to provide prediction which are even further ahead than the delays incurred by the casual Butterworth filter. In the next section, three possible solutions for mitigating these delays are addressed.

**III. Deep Machine Learning for Predictive Filtering**

In this section, the three proposed approaches are explained for predictive low-pass filtering of dqo components illustrated in Fig. 1. The overall schemata of the proposed solutions are illustrated in Fig. 3. In \( PS_1 \), we use low-pass Butterworth filter to filter out the unwanted frequency components. Following this step, linear prediction is used to estimate the future samples of the zero-phase filtered dqo. Similarly, in \( PS_2 \) we use LSTM to predict the future values of the zero-phase filtered dqo values. The aim of using linear prediction and LSTM is to predict far ahead to both mitigate the lag of the Butterworth filter and contributes to cancelling the rest of algorithmic delays and reaction time delay of APFs. Note that the same Butterworth filter defined in II-B is used as low-pass filter in \( PS_1 \) and \( PS_2 \). In \( PS_3 \), a combination of
be the same for voltage waveform if
of Matlab).

filtering and zero-phase Butterworth filter (using
ponent with its low-pass filtered version using Butterworth
and PLF stand for
Proposed Solution
Fig. 3: Proposed solutions developed in this research work. PS
Fig. 2: The example of d+ component for f_{IH} = 45 Hz com-
component with its low-pass filtered version using Butterworth
filtering and zero-phase Butterworth filter (using filtfilt
of Matlab).

Fig. 3: Proposed solutions developed in this research work. PS
and PLF stand for Proposed Solution and Predictive Low-pass
Filter respectively.

CNN and LSTM networks are employed to do the filtering
and prediction task at the same time. The theory of these
methods will be explained in detail in the next subsections.
Please note that all notations for proposed method carried out
on current waveform of three-phase EAF (i_{a},i_{b},i_{c}) and it will
be the same for voltage waveform if (v_{a},v_{b},v_{c}) are considered
instead.

A. Low-Pass Predictor for Butterworth Linear Prediction

The linear prediction method aims to forecast the amplitude
of the zero-phase low-pass filtered d+o components at time t
by a FIR filter if length M which provide a weighted sum of
the M past samples from the output of the casual Butterworth
filter [48]. For example, predicting the N^{th} future sample of
the Butterworth low-pass filtered d+BW component can be
defined as:

\[ d^{+}_{BW}(t + N) = \sum_{k=1}^{M} h_{k}d^{+}_{BW}(t - k) \]  

where \( h_{k} \) are the coefficients of the linear predictor and BW is
notation of Butterworth filter. These coefficients are obtained
from training data by solving the least-squares problem:

\[ \min_{\{h_{k}\}_{k=1}^{M}} \sum_{t} \left[ \left( d^{+}_{BW}(t+N) - \sum_{k=1}^{M} h_{k}d^{+}_{BW}(t-k) \right) \right]^{2} \]  

where \( d^{+}_{BW}(t) \) is the signal obtained from the anti-causal
zero-phase filtering using filtfilt. Note that the linear
prediction coefficients are obtained by training on the output
of the Butterworth filter shown in Fig. 3 and the desired output
are future samples of the zero-phase filter output. Later these
coefficients will be used in low-pass filtering and prediction
in the MSRF method, which can be seen in pseudo-code in
Algorithm 1.

Algorithm 1: The algorithm of MSRF with PS1:
Inference in time-domain at time t to decompose three-
phase set to its desired frequency components

1. M : length of sequence used for prediction
2. N : prediction horizon
3. BW : Butterworth filter
4. f_{c}:frequently components
5. f_{p} = \{5, 10, \ldots, 200\} Hz (interharmonics)
6. f_{c} = \{50, 100, \ldots, 5000\} Hz (fundamental component and harmonics)
7. Read \( (i_{a}(t),i_{b}(t),i_{c}(t)) \)
8. do in parallel
9. \[ \tilde{i}_{DQO_{f_{k}}}^{+}(t) = d_{qo}(\tilde{i}_{a_{h}/i_{h}}(t),\tilde{i}_{b_{h}/i_{h}}(t),\tilde{i}_{c_{h}/i_{h}}(t)) \]
10. Read \( i_{DQO} = [i_{DQO_{f_{k}}}^{+}(t-M:t),i_{DQO_{f_{k}}}^{-}(t-M:t)] \) to feed to linear predictor
11. end
B. Butterworth LSTM based low-pass predictor

In this approach, different from the previous solution, structure of LSTM allows to use a novel data augmentation step. Simply, the selected window of data is flipped from buffer and concatenated with itself:

\[ d_{\text{Augmented}}^{BW} = \begin{bmatrix} d_1^{BW} & d_1^{BW} & d_2^{BW} & \ldots & d_M^{BW} \end{bmatrix} \]

This augmentation will act as regularizer and can be helpful to increase the prediction robustness and accuracy. Another difference from linear prediction is that LSTM will provide a nonlinear prediction function. In the training phase, at each time step, we select a window of data with the length of \( M \) together with the future value of \( d_{\text{dqo}} \) from the zero-phase digital filter, i.e. \( d_{\text{dqo}}^{BW}(t + N) \). We flip the data and feed it to the LSTM networks by augmentation. By using chain derivative and Adam optimizer for gradient upgrade the weights on each layer of LSTM cells will be trained. With this style, it can create a model of patterns by trained weights and can predict the future values of \( d_{\text{dqo}} \) component. Fig. [4] illustrates the basic schema of LSTM Cells. The stateful operator of each cell is shown by \( h \) that is obtained using some operations on \( d_{\text{dqo}}^{BW} \) at index \( n \), and \( h_{n-1} \) and information of previous cells \( C_n \). The cell has three gates to control the information flow. The forget gate \( f_n \) controls if the previous state of the cell should be erased or kept. The input gate \( i_n \) controls how much information from other cells should be added to the current cell. The output gate \( o_n \) controls the information that goes on to the next cells. Note that the computed states \( h \) are fed to the LSTM cells in the next LSTM layer. The obtained states \( h_{1 \ldots n} \) from the while loop of Algorithm [2] in the last layer of the LSTM are fed to fully connected layers (FC). In the fully connected layers these extracted states \((h_1, h_2, \ldots, h_n)\) are used to obtain the future value of \( d_{\text{dqo}} \) components. The process of using the \( PS_2 \) in the MSRF method is addressed in Algorithm [2].

C. CNNLSTM based low-pass predictor

In this case of study, we use the same data augmentation suggested in the previous subsection as a regularizer to increase the prediction accuracy and result robustness. Also, different from \( PS_1 \) and \( PS_2 \) in \( PS_3 \) we use a combination of CNN and LSTM to not only low-pass filter the \( d_{\text{dqo}} \) components but also predict the future value of them. The idea is to use the CNN to filter selected window of the \( d_{\text{dqo}} \) components data with zero delays and represent them with a smaller number of features. This will help the LSTM and deep neural network in FC layers to see the highlighted features of the selected sequence. The detail of this solution together with MSRF is given in Algorithm [2]. We select several filters and a window of data from \( d_{\text{dqo}} \) components after MSRF inversion (see Fig. [1]) and convolve the initialized weights of the filters through the selected sequence. Then the averagePool will obtain the average of multiplication of weights to data on each filter and pass them to the next layer. Note that since we are striding the weights in the convolution we simply lowering the dimension of the input sequence at the same time. In the second step, the extracted features by CNN are fed to the deep neural network to create a new sequence with the desired length (preferably considerably smaller than the selected window of \( d_{\text{dqo}} \) component). The extracted features which are smoother version of \( d_{\text{dqo}} \) components and matched to downsampled and smoother version of the timeline in the raw \( d_{\text{dqo}} \) component.

Algorithm 2: MSRF with \( PS_2 \) : Inference in time domain at time \( t \) to decompose, low-pass filter and predict three phase set to its all frequency components

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( M ) : length of sequence used for prediction</td>
</tr>
<tr>
<td>2</td>
<td>( P ) : Predicted ( N ) : prediction horizon</td>
</tr>
<tr>
<td>3</td>
<td>( BW ) : Butterworth filter</td>
</tr>
<tr>
<td>4</td>
<td>( f_k ) : frequently components</td>
</tr>
<tr>
<td>5</td>
<td>( f_k = 1, 5, 10, \ldots, 200 ) Hz (interharmonics)</td>
</tr>
<tr>
<td>6</td>
<td>( f_k = 50, 100, \ldots, 5000 ) Hz (fundamental component and harmonics)</td>
</tr>
<tr>
<td>7</td>
<td>Read ((i_n(t), i_n(t), i_n(t)))</td>
</tr>
<tr>
<td>8</td>
<td>do in parallel</td>
</tr>
<tr>
<td>9</td>
<td>( i_{DQf_k}^{q_n}(t) = d_{\text{dqo}}g_k(i_{n+1/2}(t), i_{n+1/2}(t), i_{n+1/2}(t)) )</td>
</tr>
<tr>
<td>10</td>
<td>( i_{DQf_k}^{c_n}(t) = BW i_{DQf_k}^{g_n}(t) )</td>
</tr>
<tr>
<td>11</td>
<td>( i_{DQf_k}^{h_n}(t) = d_{\text{dqo}}g_k(i_{n+1/2}(t), i_{n+1/2}(t), i_{n+1/2}(t)) )</td>
</tr>
<tr>
<td>12</td>
<td>( i_{DQf_k}^{h_n}(t) = BW i_{DQf_k}^{h_n}(t) )</td>
</tr>
<tr>
<td>13</td>
<td>Read ((i_{DQf_k}^{q_n}(t - M : t), i_{DQf_k}^{c_n}(t - M : t))) to feed LSTM predictor</td>
</tr>
<tr>
<td>14</td>
<td>Data augmentation: ( i_{DQf_k} = [i_{DQf_k}; \text{flip}(i_{DQf_k})] )</td>
</tr>
<tr>
<td>15</td>
<td>LSTM Networks:</td>
</tr>
<tr>
<td>16</td>
<td>while ( i &lt;= n ) (length of ( i_{DQf_k} )) do</td>
</tr>
<tr>
<td>17</td>
<td>First LSTM layer:</td>
</tr>
<tr>
<td>18</td>
<td>( n \leftarrow n + 1 )</td>
</tr>
<tr>
<td>19</td>
<td>( i_{q_n} = \sigma_g(W_i i_{DQf_k}^{q_n} + U_i h_{n-1} + b) ) ( \triangleright \text{input gate} )</td>
</tr>
<tr>
<td>20</td>
<td>( f_n = \sigma_g(W_f i_{DQf_k}^{q_n} + U_f h_{n-1} + b) ) ( \triangleright \text{forget gate} )</td>
</tr>
<tr>
<td>21</td>
<td>( o_n = \sigma_g(W_o i_{DQf_k}^{q_n} + U_o h_{n-1} + b) ) ( \triangleright \text{output gate} )</td>
</tr>
<tr>
<td>22</td>
<td>( c_n = \tanh(W_c i_{DQf_k}^{q_n} + U_c h_{n-1} + b) ) ( \triangleright \text{cell information} )</td>
</tr>
<tr>
<td>23</td>
<td>( h_n = o_n \odot \tanh(c_n) ) ( \triangleright \text{state goes to next layer} )</td>
</tr>
<tr>
<td>24</td>
<td>First FC layer:</td>
</tr>
<tr>
<td>25</td>
<td>( FC^1 = W_p h_{n-1}^{\text{hid}} )</td>
</tr>
<tr>
<td>26</td>
<td>( i_{DQf_k}^{q_n} = i_{DQf_k}^{q_n} )</td>
</tr>
<tr>
<td>27</td>
<td>( i_{DQf_k}^{c_n} = i_{DQf_k}^{c_n} )</td>
</tr>
<tr>
<td>28</td>
<td>( i_{DQf_k}^{h_n} = i_{DQf_k}^{h_n} )</td>
</tr>
<tr>
<td>29</td>
<td>end</td>
</tr>
</tbody>
</table>

\[ P \] : Predicted \( N \) : prediction horizon
Finally, LSTM sequentially analyzes the extracted features to do the prediction task.

IV. APPLICATION OF PROPOSED SOLUTIONS

The proposed methods are implemented in a Python software environment using a workstation with an Intel i7 3.40 GHz CPU, 48GB RAM, and an NVIDIA Titan XP 12GB GPU. The proposed methods are trained and tested on 35 GB of data collected from four transformer substation feeding EAFs. The recordings were carried out at a 25.6 kHz sampling rate from 154 kV transmission lines between two to three hours duration from each substation. We have done two different kinds of training and testing of the proposed algorithms. In the first scenario, 80% of the data is used for training and 20% for testing. To test the robustness of the proposed algorithms we also the train the networks on data from three separate EAFs and test on the data set from the last EAF in a cross-validation format. The three methods are evaluated by averaging the prediction error results over the resulting four sets of models. The overall results of the proposed solutions in the previous section are presented in Table I. The results are based on the \( d^{+} \) component for the 45 Hz interharmonic extracted for both voltage and current waveforms of the EAF and the prediction horizon is set to 40 ms. It can be seen that in training with cross-validation we have a slight increase in prediction error rate. It can also be seen that PS\(_2\) perform best of the three approaches. Compared with PS\(_2\), the prediction error increases slightly for the PS\(_3\) approach where CNN and LSTM are integrated. In PS\(_3\), Linear prediction of low-pass filtered \( d^{+} \) component combined with low-pass filtering has the highest prediction error. An investigation of the relation between the prediction horizon with respect to the prediction error is shown in Fig 5. The prediction error increases approximately exponentially for all three methods. In Table I, the decrease in percentage points of the prediction error for the proposed data augmentation methods is shown. The augmentation reduces the prediction error significantly for PS\(_2\) and PS\(_3\). This can be because of the way the weights are trained towards modeling the data to give the lowest prediction error. In another word, in the prediction tasks, when data is nonlinear and complex, the weights cannot model the trend of the desired signal and can end up be trained to just pass along the last value of the selected data point. But flipping data and feeding with the original window of selected data will act as a regularizer which makes the weights to train in a way that they are focused on learning the pattern of the window of data and predict future value. Finally, to see the prediction accuracy and delays mentioned earlier, for different proposed solutions we show results from the prediction of \( d^{+} \) component for the second harmonic in Fig. 6. The prediction is carried out on one second of data (25600 samples) recorded from the boring phase of EAF. Note that the amplitude of the \( d^{+} \) component is normalized with the max-min method to keep the values between zero and one. The sequence is the same zoomed sequence illustrated in Fig. 2. To avoid clutter in the figures we refrained from plotting the non-filtered \( d^{+} \) components. From Fig. 6 it is clear that the PS\(_1\) method has a higher prediction error compared to PS\(_2\) and PS\(_3\). The method PS\(_1\) can follow the trend of the desired signal but with a slight delay. As it can be seen PS\(_2\) follow the pattern of the desired signal of the zero-phase filtered \( d^{+} \) component very well. Recall from subsection II-B, the desired signal are the future values of zero-phase Butterworth filter obtained by \texttt{filtfilt} command of Matlab. For the case of PS\(_3\) the prediction error and delay is lower than the PS\(_1\) but higher than PS\(_2\).

V. DISCUSSION ON COMPENSATION OF FLICKER, INTERHARMONICS, HARMONICS AND VOLTAGE DIPS

A. Flicker and Voltage Dip

As it is addressed in Section II it is observed that a voltage fluctuation with 5-Hz frequency can generate flicker and it can be expressed as interharmonics with 45 and 55 Hz frequency. Therefore for flicker compensation, it is needed to compensate...
Algorithm 3: MSRF with $PS_3$ Algorithm: Inference in time domain at time $t$ to decompose, low-pass filter and predict three phase set for all frequency components using $PS_3$ method

1. $M$ : length of sequence used for prediction
2. $P$ : Predicted $N$ : prediction horizons
3. $BW$ : Butterworth filter
4. $f_k$ : frequently components
5. $f_k = 1, 5, 10, ..., 200$ Hz (interharmonics)
6. $f_k = 50, 100, ..., 5000$ Hz (fundamental component and harmonics)
7. Read $(x_i(t), x_{i+1}(t), x_{i+2}(t), x_{i+3}(t), x_{i+4}(t))$
8. do in parallel
9. $i_{DQ}^k(t) = dqqf_k \left( i_{s,k}/i_n(t), i_{p,k}/i_n(t), i_{c,k}/i_n(t) \right)$
10. $i_{DQ}^k(t) = dqqf_k \left( i_{s,k}/i_n(t), i_{p,k}/i_n(t), i_{c,k}/i_n(t) \right)$
11. Read $i_{DQ}^k = [i_{DQ}^r(t - M : t), i_{DQ}^f(t - M : t)]$
12. Data augmentation:
13. $i_{DQO}^k = [dqqf(p(i_{DQ}^k))$
14. CNN Networks:
15. for $j = 1$ : number of filter do
16. for $k = 1$: stride : length (i.e: $DQ$) do
17. First CNN layer:
18. $x_{[j,k]} = $ AveragePool$(\text{Relu}(W_{\text{CNN}}^j i_{DQO}(k : \text{filter size})))$
19. end
20. end
21. First FC layer:
22. $FC^1 = W_F^i (x_{[j,k]})$
23. LSTM Networks:
24. while $i < n$ (number of neurons in $FC^1$) do
25. First LSTM layer:
26. $n \leftarrow n + 1, i_n = \sigma_g (W_n FC^1 + U_n h_{n-1} + b_i) \triangleright$ input gate
27. $f_n = \sigma_g (W_n FC^1 + U_n h_{n-1} + b_f) \triangleright$ forget gate
28. $c_n = \sigma_g (W_n FC^1 + U_n h_{n-1} + b_o) \triangleright$ output gate
29. $\tilde{c}_n = \text{tanh}(W_n FC^1 + U_n h_{n-1} + b_c)$
30. $c_n = f_n \cdot c_{n-1} + \tilde{c}_n \cdot c_n \triangleright$ cell information
31. $h_n = o_n \cdot \text{tanh}(c_n) \triangleright$ state goes to next layer
32. $\cdot$
33. $\cdot$
34. $\cdot$
35. $3^{\text{rd}}$ LSTM layer:
36. $i_n = \sigma_g (U_n h_{n-1} + b_i) \triangleright$ input gate
37. $f_n = \sigma_g (U_n h_{n-1} + b_f) \triangleright$ forget gate
38. $o_n = \sigma_g (U_n h_{n-1} + b_o) \triangleright$ output gate
39. $\tilde{c}_n = \text{tanh}(U_n h_{n-1} + b_c)$
40. $c_n = f_n \cdot c_{n-1} + \tilde{c}_n \cdot c_n \triangleright$ cell information
41. $h_n = o_n \cdot \text{tanh}(c_n) \triangleright$ state goes to FC layer second FC layer:
42. $FC^2 = W_F^i (h_n^1, h_n^2, ..., h_n^k)$
43. Third FC layer:
44. $i_{DQ}^k(t + N), i_{DQ}^k(t + N) = W_F^i(FC^2)$
45. end
46. $i_{DQ}^k(t), i_{DQ}^k(t), i_{DQ}^k(t)$ $\leftarrow$ dqqf $\left( i_{DQ}^k(t + N), i_{DQ}^k(t + N) \right)$
47. end

TABLE I: Investigation of proposed solutions and their results for 40 ms prediction horizon of $d^+_p$ component of 45 Hz interharmonics of EAF current and voltage. Abbreviations used on this table: prediction (Pred.), error (Err.), validation (Val.)

<table>
<thead>
<tr>
<th>Scenario \ Methods</th>
<th>$d^+_p$</th>
<th>$PS_1$</th>
<th>$PS_2$</th>
<th>$PS_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 ms Pred. Err. (%)</td>
<td>$d^+_p$</td>
<td>2.06</td>
<td>0.32</td>
<td>0.99</td>
</tr>
<tr>
<td>80 (%) Training 20 (%) Val.</td>
<td>$d^+_p$</td>
<td>1.91</td>
<td>0.30</td>
<td>0.82</td>
</tr>
<tr>
<td>Data augmentation</td>
<td>$d^+_p$</td>
<td>0.15</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>effect on Pred. Err. (%) points</td>
<td>$d^+_p$</td>
<td>-</td>
<td>0.41</td>
<td>0.83</td>
</tr>
<tr>
<td>40 ms Pred. Err. (%)</td>
<td>$d^+_p$</td>
<td>2.32</td>
<td>0.42</td>
<td>1.22</td>
</tr>
<tr>
<td>Cross Val.</td>
<td>$d^+_p$</td>
<td>2.01</td>
<td>0.37</td>
<td>0.89</td>
</tr>
<tr>
<td>Waveform Pred. Err. (%)</td>
<td>$d^+_p$</td>
<td>7.59</td>
<td>1.92</td>
<td>3.5</td>
</tr>
<tr>
<td>with 40 ms prediction</td>
<td>$d^+_p$</td>
<td>8.59</td>
<td>1.5</td>
<td>3.01</td>
</tr>
</tbody>
</table>

Fig. 5: Investigation of prediction error for $d^+_p$ as a function of the prediction horizon for proposed solutions. Prediction horizon increased every cycle/10

be accomplished it shows that it is possible to filter out the fluctuations by injecting the reference signal with out of phase for flicker compensation. For the case of a voltage dip, the injection will be the same phase as the fundamental component. To illustrate the prediction accuracy, the envelope of both $V_a$ and the predicted version of it are also shown in the figure. It can be seen that the predicted signal can follow the trend of the original $V_a$ two cycles ahead and the reconstruction error is 1.35 % for the selected segment of data in the figure.

B. Harmonics and Interharmonics in current

Similar to the case of voltage flicker, in this subsection a reconstruction of the current waveform of an EAF using MSRF based on $PS_2$ is presented. The reconstruction of current waveform based on predicted frequency components resulted with 1.79% error rate on selected window of data shown in Fig. 3. Since the current waveform of the EAF has more stochastic behavior compared to the voltage, all of the proposed solutions resulted slightly less accurate pre-
dictability. The reconstruction error rates of current waveform using $PS_1$ and $PS_3$ are 7.22% and 2.99% respectively. The accuracy of reconstruction using $PS_2$ still has highest accuracy for a prediction horizon of 40 ms. The error on the reconstructed current waveform is slightly higher than the average prediction error of the 45 Hz frequency component only. This can be reasoned around possible overlap in estimation of different frequency components and slight increase in prediction error of high frequency components such as the 9th harmonic. To avoid to keep the presentation clear, only the reconstruction results of $i_a$ using $PS_2$ is shown in Fig. 8 where the reconstructed waveform is generated by summing up the fundamental component and the components from the 2nd harmonic, the odd harmonics up to the 10th, and interharmonics with a 5 Hz resolution up to 200 Hz. It can be seen that the predicted and reconstructed $i_a$ follow the pattern of the original $i_a$ with a good accuracy before (40 ms) it appears and the delay effect of the Butterworth filter is mitigated. Thus this extra available time contributes to the mitigation of the rest of reaction time delay and the response time delay of the APF.

C. Implementation of APF in Simulink and methods embedded system

In this subsection, a virtual prototype of ML based and a trivial APF is implemented. The aim is to evaluate the performance and effectiveness of proposed predictive algorithms mentioned in previous sections. The $PS_2$ based approach
is used for the prediction and low-pass filtering task in Simulink since, as shown in the previous sections, it gave the best performance. The block schema and specification of the developed shunt APF in Simulink environment is illustrated in Fig. 9.

The Simulink model consists of three main power electronics-based blocks including AC grid, EAF, and Shunt Active filter. To make sure the model is close to the real environment we used data that is recorded from transformers feeding an EAF and give it as the reference signal to a six gate IGBT bridge with a PI controller. This block will act as a current source that regenerates the EAF current. We also replicated the AC grid block diagram using a similar six gates IGBT bridge. The voltage in the AC grid block has to have a constant level (400 KV) and be in phase with EAF to avoid the effect phase overlap resulting in the extra frequency deviation problems. Hence, the reference current waveform is obtained based on the difference of the reference voltage (400 KV) and measured voltage from EAF. Note that this is an example model of compensating interharmonics and harmonics including variation in the fundamental component and models specifications for the purpose of flicker and voltage dip compensation. With a successful compensation, \( R_L \) in Fig. 9 will have a pure sinusoidal current flow. Two scenarios are defined to show the effect of the predictive based approaches versus a trivial APF solution. The first scenario is an implementation of a trivial APF where the reference signal is estimated with only MSRF analysis + Butterworth filter. In the second scenario the reference signal is predicted using MSRF + \( PS_2 \). In both scenarios, at first, the voltages and currents of EAF together with voltage and current of AC grid are measured and sent to GPU accelerated processor (Jetson TX2) via TCP link using Ethernet cable. In the second step, we estimate the \( dq \) components using the MSRF box. Then in predictive APF, \( PS_2 \) is used for low-pass filtering and 40 ms prediction of the future value of the \( dq \) components. Note that this 40 ms delay originates from the lag of the Butterworth filter mentioned in 11-E (16 ms), PI controller response time (21 ms), the data communication delay between Simulink and Jetson TX2 (2.5 ms), actuation of APF (0.8 ms) and processing time of data on MSRF blocks and \( PS_2 \) (0.2 ms). Furthermore, the reference signal is generated using by MSRF\(^{-1}\) on predicted and low-pass filtered values of the \( dq \) components. Similar approach is carried out for the trivial APF scenario but the only difference is that there is no LSTM for prediction of the filtered \( dq \) components. As an example, one minute of EAF data is generated by the EAF block with the APF generated signal for the two scenarios. The MSRF decomposition includes frequency components from 45 Hz to 55 Hz with 1 Hz resolution. The EAF current and the compensation results on 0.9 s of data between 1.1 s to 1.8 s using the trivial APF and predictive APF with \( PS_2 \) are illustrated in Fig. 10 (a) and Fig. 10 (b) respectively.

The results shown in Figure 10 (a) and (b) clearly illustrate the importance and value of a predictive-based approach. The performance evaluation on the selected window of data where we measure the remaining undesired frequency components after compensation resulted in 60% and 96% efficiency for the trivial and \( PS_2 \) based APF respectively. The trivial APF reduce the undesired frequency components at a lower compensation efficiency. However, during transients it can cause an increase in the interharmonics which is evident from Figure 10 (a) at time\( \geq 1.7 \) s. The reason is that in the case of highly time-varying EAF current and voltage, the amplitude and the phase of frequency components are varying almost every cycle. Thus, having a 40 ms delay means that the APF compensates the interharmonics based on measurements two cycle ago which are no longer valid at present time.

VI. CONCLUSIONS

In this research work, it is shown that it is possible to predict highly time-varying voltage and current waveforms of EAFs up to two cycles (40 ms) of the fundamental frequency using a deep-learning based method. This is to provide reference signals to active power filters (APFs) so that harmonics, interharmonics, flicker and voltage dips can be mitigated. The key point is choosing the right prediction method and pre-processing steps which is decomposing the waveform into enough number sinusoidal components. These sinusoidal signals are represented with \( dq \) frames using the MSRF method. It has been shown that \( dq \) frame analysis after low-pass filtering of \( dq \) components became smoother and less time-varying over time. Hence, it is thus easier to predict the \( dq \) components rather then the current or voltage waveforms directly. Three prediction methods are investigated. \( PS_1 \) is a combination of low-pass filtering and linear prediction. \( PS_2 \) is a combination of low-pass filtering and non-linear prediction using an LSTM model. \( PS_3 \) is a combination of CNN and LSTM architectures which together perform the prediction without an explicit low-pass filter. For the \( PS_2 \) and \( PS_3 \) solutions a particular data augmentation is employed, which is flipping the selected window of data and feed it together with the original window, can act as regularization and increase the prediction robustness and decrease the prediction error. It has been shown that the flicker and voltage dip phenomenons can be represented as interharmonic problems and they can be compensated and predicted along with compensating harmonics and interharmonics of both voltage and current waveform of EAFs. Finally, it can be concluded that each of the proposed methods have their own pros and cons. Implementation of \( PS_1 \) is easy and less complicated but its average prediction error is almost seven times and two times larger compared to \( PS_2 \) and \( PS_3 \), respectively. This can be an essential factor when it comes to predicting highly time-varying signals such as EAF current and voltage waveform. The \( PS_3 \) made it possible to develop a DL-based low-pass filtering which mitigates the need for tuning of Butterworth filter and automatically extract the future values of low-pass filtered \( dq \) components.
Fig. 9: GPU accelerated Simulink based implementation of a trivial and ML based shunt APF. The elements used in Simulink can be listed as: $R_G = 1\, \Omega$, $R_L = 10\, k\Omega$ and $L_E = L_G = L_C = 3\, mH$.

Finally, we observed that the implementation of predictive based methods on a replicated electric power system using Simulink improve the performance compared to a trivial APFs by more than 36%.

ACKNOWLEDGEMENTS

The authors would like to thank the Chalmers University of Technology for financially supporting this research work. The authors would also like to thank the Public Research Institutions and Development Projects Support Group (KAMAG) of the Scientific and Technological Research Council of Turkey (TUBITAK) for providing EAF field data.

REFERENCES


M. A. M. Radzi and N. A. Rahim, “Neural network and bandless
M. El-Habrouk, M. Darwish, and P. Mehta, “Active power filters: A
E. Balouji, “On machine learning based signal processing for power
S. A. O. da Silva, P. F. Donoso-Garcia, P. C. Cortizo, and P. F. Seixas,
M. C. Cirrincione, M. Pucci, and G. Vitale, “A single-phase DG gener-
J. D. Bull, N. A. F. Jaeger, and F. Rahmatian, “A new hybrid current
E. Balouji, K. Bäckström, T. McKelvey, and Ö. Salor, “Deep-learning-
J. Tyster et al., “Power semiconductor nonlinearities in active du/dt
M. A. M. Radzi and N. A. Rahim, “Real-time detection of interharmonics and harmonics of AC electric arc furnaces on GPU
time of reactive power based on different definitions and algorithms,”
M. A. M. Radzi and N. A. Rahim, “Neural network and bandless
M. A. A. Mohd Zainuri, M. A. Mohd Radzi, A. Che Soh, N. Mariun,
M. A. A. Mohd Zainuri, M. A. Mohd Radzi, A. Che Soh, N. Mariun,
M. El-Habrouk, M. Darwish, and P. Mehta, “Active power filters: A
s.a. ABB n.v. Power Quality Products, “Active filters - PQFI-PQFM-
J. Yuster et al., “Power semiconductor nonlinearities in active du/dt
E. Balouji, K. Bäckström, T. McKeilvey, and Ö. Salor, “Deep-learning-
E. Balouji, O. Salor, and T. McKeilvey, “Predictive compensation of EAF
O. Salor, B. Gultekin, S. Buhan, B. Boyrazoglu, T. Inan, T. Atalık,
O. Salor, B. Gultekin, S. Buhan, B. Boyrazoglu, T. Inan, T. Atalık,
E. Balouji, K. Bäckström, T. McKeilvey, and Ö. Salor, “Deep-learning-
(a)
(b)
Fig. 10: Comparison of the result of the Simulink based
Fig. 10: Comparison of the result of the Simulink based
(a) Compensation of the frequency components between 45:1:55 Hz without prediction. (b) Compensation of the frequency components between 45:1:55 Hz with PSO of dq components.


Ebrahim Balouji (S’13) received M.Sc degree in electrical and electronics engineering from Gazi University, 2015. From 2015-2017 he worked as a researcher of Turkey’s National Power Quality Project. Since 2017, he has researched analyzing power system data using machine learning and signal processing tools towards his Ph.D. studies in the electrical engineering department of Chalmers University of Technology, Gothenburg, Sweden. He received Gunnar Engström honer reward from ABB foundation for his research in 2021.

Özgül Salor (S’98–M’05–SM’18) received the B.Sc., M.Sc., and Ph.D. degrees from Middle East Technical University, Ankara, Turkey, in 1997, 1999, and 2005, respectively, all in electrical and electronics engineering. From 2001 to 2003, she was a Professional Researcher with the Center for Spoken Language Research, University of Colorado, Boulder, CO, USA. From 2006 to 2012, she was with the Power Electronics Department, Space Technologies Research Institute, Scientific and Technological Research Council of Turkey. Since 2012, she has been a Professor with the Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey. Her current research interests include machine learning techniques and signal processing for power quality analysis. Currently, she is with the Kyrgyz-Turkish Manas University in Kyrgyzstan to start the new department of Electrical and Electronics Engineering since August 2019. Dr. Salor was the recipient of the Outstanding Paper Award and Certificate of Recognition Award from the Metal Industry Committee of the IEEE Industry Applications Society in 2009 and 2015 and the Meritorious Paper Award of MIC in 2013 and 2019.

Tomas McKelvey received the M.Sc. degree in Electrical Engineering from Lund University, Lund, Sweden, in 1991, and the Ph.D. degree in Automatic Control from Linköping University, Linköping, Sweden, in 1995. He held research and teaching positions with Linköping University from 1995 to 1999, where he became a Docent in 1999. From 1999 and 2000, he was a Visiting Researcher with the University of Newcastle, Newcastle, NSW, Australia. Since 2000, he has been with the Chalmers University of Technology, Gothenburg, where he has held a Full Professor position and has been the head of the Signal Processing Group since 2006 and 2011, respectively. His current research interests include model-based and statistical signal processing, system identification, and optimal control with applications to radar systems, electrical power systems, and vehicle propulsion systems. systems.