Harmonic Current Estimation of Unmonitored Harmonic Sources with a Novel Oversampling Technique for Limited Datasets

Ahmadreza Eslami, Student Member, IEEE, Michael Negnevitsky, Senior Member, IEEE, Evan Franklin, Member, IEEE Sarah Lyden, Member, IEEE

Centre for Renewable Energy and Power Systems, School of Engineering, University of Tasmania (UTAS), Hobart, TAS 7005, Australia

Corresponding author: Michael Negnevitsky (e-mail: michael.negnevitsky@utas.edu.au).

ABSTRACT In modern power systems, harmonics are amongst the significant issues attributed to renewable energy sources and nonlinear loads. Direct harmonic monitoring of the entire power system may be too costly or impractical, and measured data could be limited. In this paper, a new methodology is proposed to estimate harmonic current rms values of unmonitored harmonic sources, based on harmonic voltage rms magnitudes only, measured at a limited number of monitored buses. A new technique of output curve-normalization is employed in per-processing. Subsequently, a method is proposed to refine the architecture of the Artificial Neural Networks’ (ANNs), after which ANN-based harmonic current estimators are developed for each harmonic order and each harmonic source. Furthermore, a novel Neural Oversampling Consensus Algorithm for Regression (NOCAR) is proposed to improve estimation accuracy. K-Nearest Neighbor (KNN) and ANN are combined in developing NOCAR. A comparison is made with state-of-the-art techniques by using synthetic data, which demonstrates both the proposed method’s robustness and its capability to perform when minimal information is available. The implementation for real data demonstrates the efficiency of the ANN-based harmonic current estimators with oversampling. The influence of the number of harmonic meters is investigated, revealing the ability of this data-driven technique to reduce the number of harmonic meters, and hence monitoring costs. Moreover, the correlation between different harmonic orders is studied, with results suggesting that, unlike the widely accepted notion, this correlation should not be ignored in harmonic analysis. This study highlights the advantages of integrating intelligent techniques into harmonic monitoring systems.

INDEX TERMS artificial neural networks, harmonic current estimation, harmonic voltage monitors, oversampling

Nomenclature

| Symbol | Description |
|--------|-------------|
| $b_{j}^{k+1}$ | bias of node $j$ in Layer $k+1$ |
| $D$ | dataset for ANN training and testing |
| $d_{j}^{k+1}$ | output of node $j$ in Layer $k+1$ |
| $D_{l,new}$ | final dataset after lower-oversampling |
| $D_{u,new}$ | final dataset after upper-oversampling |
| $e$ | measurement error including noise |
| $F$ | nonlinear mapping function from rms voltages to rms currents |
| $f$ | activation function |
| $G$ | Inverse of $F$ |
| $H$ | measurement matrix |
| $h$ | harmonic order |
| $I_{r}$ | resultant current vector by consensus |
| $p$ | number of ANN layers |
| $P_{l}$ | lower-oversampling percentile |
| $P_{u}$ | upper-oversampling percentile |
| $Q$ | number of points for each layer in $SP_{L}$ |
| $Reg_{MSE}$ | regularized Mean Squared Error |
| $S_{l}$ | size of layer $l$ of the overfitted ANN |
| $SP_{L}$ | discrete search space for size of ANN layers |
| $V_{D}$ | input voltages in $D$ |
| $V_{D_{l,new}}$ | subset of final input vectors after lower-oversampling |
| $V_{D_{u,new}}$ | subset of final input vectors after upper-oversampling |
| $(V_{D_{l,new}}, I_{D_{l,new}})$ | subset for lower-oversampling |
| $(V_{D_{u,new}}, I_{D_{u,new}})$ | subset for upper-oversampling |
Harmonics and waveform distortion are of major concern for operators of modern power systems, due to the ever-increasing penetration of Renewable Energy Sources (RES) and nonlinear loads. RES, primarily wind and solar photovoltaics, generally connect to the grid via inverters, while nonlinear loads typically incorporate power electronics switching devices and other nonlinear elements. In addition, the emerging tendency towards use of DC microgrids, solid-state transformers and high-voltage dc transmission [1] also increases the degree of nonlinearity in power systems. A high degree of nonlinearity will increase the flow of harmonic currents, which in turn will result in harmonic voltages and currents throughout. Power quality (PQ) standards, such as the IEEE Std 519-2014 [2] and IEC 61000-3 [3], impose limits on the levels of harmonic voltages and currents. Reliable and effective harmonic analysis and mitigation techniques are vital in order to ensure that these PQ standards are met [4].

A. Inspiration and Motivation
Harmonic analysis is generally quite complex, particularly given the typically high correlation between harmonic variables within a power network. The authors in [5], for example, reported a high correlation between harmonic orders 5 and 7 in low-voltage networks, and also high correlation between the 5th harmonic in both the low-voltage and medium-voltage sections of a network. The intermittent and uncertain nature of wind and solar generation, which represent the main sources of harmonics on the generation side, further adds to complexity. On the other hand, direct monitoring of the entire power system may be too costly or
impractical, with deregulated power systems now comprising many prosumers and various ownership models [6]. Therefore, the concept of Harmonic State Estimation (HSE) has been addressed by researchers [7]-[10]. HSE involves solving the reverse of the harmonic power flow problem, to provide estimation of harmonic voltages and currents of unmonitored harmonic sources, based on measurements at other buses and lines. HSE is thus a problem which requires analysis in the frequency domain. While conventional signal processing techniques such as Fourier Transform (FT), S-Transform (ST) and Wavelet Transform (WT) can be employed to convert data from time domain to frequency domain, they are not able to be applied to the HSE problem itself. HSE requires more sophisticated methods. This is illustrated in Fig. 1, which depicts analysis methods typically used in power systems, both in time domain and frequency domain. FT, ST and WT transform data from time/frequency domain to frequency/time domain, while HSE is conducted in frequency domain. In essence, conventional techniques can only be employed to determine harmonic content at buses where time domain measurements have been made, while HSE is able to be used to determine harmonic content at other buses which have not been observed at all. A good methodology for HSE can yield full harmonic observability for partially harmonic-monitored power systems [11]. In addition to monitoring occurring at a limited number of locations, measured data itself may also be limited, with for example harmonic phase angles being unavailable. However, even a limited and incomplete dataset still can contain useful information and provide the network operator with immediate insight into harmonic levels in a power system. Hence, HSE methods which are able to address limited information could be valuable for network operators. The main motivation of this paper is to develop a harmonic monitoring technique that works when the number of direct measurements in a network are limited and/or the measured data contains incomplete information.

“See Fig. 1 at the end of the manuscript”

B. Assessment of the State-of-the-Art

Estimation of harmonic voltages and currents of unmonitored harmonic sources has been studied both by using conventional analytical models and by using data-driven models. Least squares (LS) methods and their variants, such as weighted LS and Genetic Algorithm LS (GALS), have been used to address HSE [9], [12]-[14]. In spite of the interesting technique for locating harmonic sources, the GALS method in [13] is only applicable when the number of simultaneously operating harmonic sources is lower than the number of measurements. In [12], weighted LS was used to solve the HSE problem, by considering uncertainty in network parameters. The number of measurements was around 1.35 times the number of state variables. Kalman filter (KF) methods have also been applied in HSE [7], [15] and were reported to be fast. An adaptive KF was applied in [7], which did not require explicit knowledge of the noise covariance matrix. Other numerical methods, such as sparsity maximization assuming sparsity in state variables (harmonic source injections) [8] and combined GA and sensitivity analysis [16] have also been applied to solve the HSE problem. In [8], [13] and [16], small numbers of harmonic sources were considered to operate simultaneously, compared to the number of measurements. However, the number of sources is likely to be much higher in power systems with high penetration of RES.

When a limited number of harmonic meters are available, conventional HSE methods may reveal some errors due to the lack of harmonic observability [8]. Hence, non-conventional methods have been proposed to minimize the number of harmonic meters required, while preserving the accuracy of HSE and harmonic observability. For example, a method based on elimination of the row with minimum condition in the measurement matrix, and GA combined with grid topology analysis were proposed for optimal harmonic meter placement in [17] and [18], respectively. In spite the merits of such techniques, they did not address the variation of network parameters.

Researchers have moved toward Artificial Intelligence (AI) methods for HSE because of their learning capabilities. Traditional HSE methods require pre-knowledge of power system parameters (impedance/admittance model) and quite a large number of harmonic monitors. These requirements are reduced by applying AI. AI techniques for HSE encompass Artificial Neural Networks (ANNs) and its variants, Bayesian Learning (BL), Fuzzy Inference System (FIS), and hybrid techniques. In early works of [19] and [20], ANN was combined with traditional HSE and constrained estimation, respectively, to estimate current injection of harmonic sources. The estimates by ANNs were treated as pseudo-measurements and modified in the complementary stages of either HSE or constrained estimation. In [21], ANN was employed to estimate voltage and Total Harmonic Distortion (THD) of an unmonitored bus, based on the voltage/current measurements at Distributed Generation (DG) buses. The algorithm was tested for a small power system and under slightly varying load condition. Authors in [22] proposed a method based on Nonlinear AutoRegressive models with eXogenous inputs (NARX), which estimated the voltage of an unmonitored sensitive bus as a function of voltage and current at a monitored bus. However, no test set was used to assess the generality of the trained model. BL and Sparse BL were utilized to solve HSE in [23] and [24], respectively. The former employed regression analysis and recurrent NN for power flow calculation and demand prediction, respectively. In [25] and [26], FIS and ANN were combined to locate harmonic sources based on the signs of active and reactive harmonic powers while their harmonic levels were estimated as qualitative variables.
Conventional techniques such as LS and statistical methods such as BL usually require an accurate network model and many harmonic meters, to ensure accurate estimation. Their accuracy is generally low for underdetermined systems. In addition, techniques such as [8], [13] and [16] assume sparsity of simultaneous harmonic sources, which is likely not the case for systems with high penetration of RES. The iterative nature of methods of [13] and [24] make them inappropriate for real-time applications. Last but not least, ANN techniques for HSE generally require large datasets. To the best of the authors’ knowledge, application of HSE techniques in systems where harmonic phase angles and network models are both not available and where the number of datapoints is limited, has not been addressed in the literature.

C. Contribution and Paper Organization

In this paper, harmonic current estimation of unmonitored harmonic sources based on harmonic voltage measurements at monitored buses of a power system has been addressed. ANNs have been used to develop the harmonic monitoring system, with input features extracted using FT. After preprocessing, the feature vectors are fed to the ANNs, to map harmonic voltage rms values at monitored buses to harmonic current rms values of the harmonic sources. The methodology is demonstrated first on synthetic data using the IEEE 14-bus network, and then real data measured in the Tasmanian transmission network, which features wind farms as the major harmonic sources. It is notable that the test sets, which have not been used in any stage of training the harmonic estimators, are employed to evaluate the performance and to compare the proposed technique with others reported in the literature including GALS [13], Singular Value Decomposition (SVD) [17] and BL [24]. The contributions of the paper are as follows.

1) Compared to other HSE techniques in the literature, the proposed technique requires minimum information. The proposed method does not require any knowledge of the network model or of harmonic voltage phase angles. It only needs harmonic voltage rms values at monitored buses to estimate rms currents injected by unmonitored harmonic sources.

2) Regarding the development of AI harmonic estimators, input normalization and a newly proposed output curve-normalization are performed for the data, in the preprocessing stage. In addition, the refined structure of the ANN-based monitoring system has been obtained by solving an optimization problem defined on the training sets.

3) A new Neural Oversampling Consensus Algorithm for Regression (NOCAR) has been developed to address the overfitting problem and increase the generality of the estimators. NOCAR aims to increase the number of datapoints in limited datasets and consequently, to improve the performance of the harmonic current estimators on the test sets. NOCAR employing different soft computing technique including FT, K-Nearest Neighbor (KNN) algorithm and ANN, can be useful in any regression problem for any application.

4) For real data measured in the Tasmanian transmission network, cases of different numbers of harmonic voltage monitors have been studied. In addition, harmonic orders in the input feature vectors have been considered separately and simultaneously. In the former, harmonic voltages of each order are only considered in estimating the harmonic currents of the same order while in the latter, harmonic voltages of all orders under study, are considered in estimating the harmonic currents of each order. The analysis of the real data suggests that unlike the widely accepted convention in the literature, the correlation between different harmonic orders is important for harmonic analysis.

The rest of the paper is organized as follows. Harmonic current estimation is presented in Section II. Section III presents the proposed methods and the mathematical model for developing the AI system for HSE. Comparison of the proposed technique with the state-of-the-art techniques is made in Section IV. Implementation for a real transmission network and the discussion of the results are presented in Section V. Finally, the main findings of the paper and suggestions for further research are provided in Section VI.

II. HARMONIC CURRENT ESTIMATION

The basic formulation of conventional HSE is a mapping of monitored harmonic voltage and current phasors to unmonitored harmonic voltage and current phasors as (1) [11]. For a limited dataset, where only harmonic rms values (or magnitudes) are available, but not phase angles, the linear matrix equation of (1) does not hold and traditional HSE methods are not applicable. However, a mapping of monitored harmonic voltage/current rms values to unmonitored harmonic voltages/currents can be defined as (2), where $F$ and $G$ are nonlinear functions. While it may be analytically very hard or even impossible to derive these nonlinear functions, ANNs can efficiently address such problems by learning the nonlinear mapping from data of input/output pairs.

$$Z_{(M \times l)} = H_{(M \times N)}X'_{(N \times l)} + e$$

$$|Z_{(M \times l)}| = F(|X'_{(N \times l)}|) \rightarrow |X'_{(N \times l)}| = G(|Z_{(M \times l)}|)$$

The relationship between measurements and current injected by harmonic sources can be defined as (3) and (4). Two approaches can then be considered: 1) unknown harmonic currents of each harmonic order are obtained only as a function of measurements of the same harmonic order, the HSE problem for each harmonic order being independent of other harmonic orders, as in (3); and, 2) simultaneous consideration of all relevant harmonic orders in estimating currents of each harmonic order, as in (4). The independent harmonic approach is commonly employed for conventional harmonic analysis [27]. In (3) and (4), $H_{(M \times N)}$ is the vector of $h$-order rms harmonic currents injected by $N$ harmonic sources and $V_{h}^{i}_{(M \times l)}$ is the vector of $h$-order rms harmonic voltages measured at $M$.
metering locations. ANNs will be employed to estimate either the function $G$ or the functions $G_h$ for each harmonic order $h$.

$$I_{HS}^{h}(N_{x1}) = G_h(|V_{(M_{x1})}|), h = h_1, h_2, h_3, \ldots \quad (3)$$

$$\begin{bmatrix}
I_{HS}^{h_1}(N_{x1}) \\
I_{HS}^{h_2}(N_{x1}) \\
\vdots \\
I_{HS}^{h_i}(N_{x1}) \\
\vdots \\
I_{HS}^{h_n}(N_{x1})
\end{bmatrix} = G
\begin{bmatrix}
|V_{(M_{x1})}| \\
|V_{(M_{x1})}| \\
\vdots \\
|V_{(M_{x1})}| \\
\vdots \\
|V_{(M_{x1})}|
\end{bmatrix} \quad (4)$$

III. THE PROPOSED AI SYSTEM FOR HSE

To develop the AI system for harmonic current estimation, three stages are employed: data refinement and feature extraction (pre-processing stage), ANN estimation, and, post-processing. FT, conventional scaling normalization of the ANN inputs and a proposed normalization technique for the ANN outputs, called curve-normalization, are employed in the pre-processing stage. In developing ANNs, the training input/output pairs have been first used to refine the ANN structure by solving an optimization problem. ANNs, with this refined structure, are then trained using training sets. In the final stage, denormalization of the outputs is conducted. The test sets are only used to evaluate the performance of the AI system and are not used in the development stages. Fig. 2 depicts the flowchart of the AI harmonic estimators, each stage of which is detailed in the following subsections.

"See Fig. 2 at the end of the manuscript"

A. PRE-PROCESSING STAGE: FOURIER TRANSFORM AND DATA NORMALIZATION

Frequency domain analysis is a common practice in HSE and harmonic estimation. FT and its variants such as Fast FT (FFT) and Short-Time FT (STFT) can be used to transform from time domain to frequency domain [28]. The intelligent harmonic estimators, in this paper, are also developed in the frequency domain.

Data normalization is used to improve the generality of AI systems. One of the prevalent data normalization techniques, known as feature scaling, is based on the minimum and the maximum of each variable in the dataset. This normalization is usually applied to input features of the AI system so that each feature is transformed to a range between 0 to 1. Consider $Z$ as the $i$th feature vector in the dataset, feature scaling results in (5) [29].

$$Z_{i,\text{Norm}} = \frac{Z_i - Z_{i,\text{min}}}{Z_{i,\text{max}} - Z_{i,\text{min}}} \quad (5)$$

Sometimes, different output variables of an AI system may also have high variance and wide ranges. The same normalization approach can be followed to normalize the training set outputs. However, this approach cannot be applied to the test set, since the maximum and minimum parameters of the test sets are assumed to be unknown in the training stage. Accordingly, an output normalization method is proposed which operates just based on the minimum and the maximum values of the target outputs in training sets. This approach is also applicable for other normalization methods such as standardization based on mean and standard deviation parameters. Consider the output vector $X$ which is divided into two portions, $X_{\text{train}}$ and $X_{\text{test}}$. The normalized output vectors will be

$$X_{\text{train \ Norm}} = \frac{X_{\text{train}} - X_{\text{train \ min}}}{X_{\text{train \ max}} - X_{\text{train \ min}}} \quad (6a)$$

$$X_{\text{test \ Norm}} = \frac{X_{\text{test}} - X_{\text{test \ min}}}{X_{\text{test \ max}} - X_{\text{test \ min}}} \quad (6b)$$

As can be seen, the original value of any arbitrary data point in the dataset can be recovered.

In spite of the benefit of the output normalization, there still may be deficiencies in the process of learning for datasets containing outliers. This is the case for our study which uses data obtained from a real power system. There are spikes in the injected harmonic currents of the windfarms due to wind speed change, particularly when an interval of no wind or low wind is followed by an interval of high wind, or vice versa. The data outliers and significant variations within the data can be problematic when using normalization, since the vast majority of normalized data points become crowded near to zero. Hence it can be advantageous to further pre-process the data in such a way as to stretch out data points that are close to 0 and compress data points close to 1. In the domain of [0,1] and for datapoints which are crowded near to zero with outliers close to one, $n$th root functions operate to reduce the variance of the dataset and the effect of spikes. For example, cube roots of 0.1 and 0.9 are 0.464 and 0.965, respectively. The cube root function is a continuous and invertible function on $\mathbb{R}$, and hence the original data can be easily recovered. The process of curve-normalization (C-Norm), via application of the cube root function, is depicted in (7). As a result, the ANNs will subsequently be trained to estimate the curve-normalized value of the outputs, from which the original outputs can be obtained, accordingly.

$$X_{\text{train \ C-Norm}} = \sqrt[3]{X_{\text{train \ Norm}}} \quad (7a)$$

$$X_{\text{test \ C-Norm}} = \sqrt[3]{X_{\text{test \ Norm}}} \quad (7b)$$

B. PROCESSING STAGE: ANN HARMONIC CURRENT ESTIMATOR
A Feed Forward Neural Network (FFNN) is used to estimate the function $G$ and hence the harmonic current rms values. An FFNN is comprised of an input layer, several hidden layers, and an output layer. Each layer consists of a number of neurons (nodes) with activation functions such as sigmoid and hyperbolic tangent. The neurons in Layer $k$ are connected to neuron in Layer $k+1$ through adjustable synaptic weights [30]. In the feed forward calculation, the output of node $j$ in Layer $k+1$ is derived from the nodes in Layer $k$ as follows.

$$d_{j}^{k+1} = f(b_{j}^{k+1} + \sum_{i=1}^{n} w_{ji}^{k}d_{i}^{k})$$

(8)

where $f$ is an activation function such as sigmoid, tanh or a linear function.

In a supervised manner, NN Learning is an iterative process where the ANN outputs are calculated in the feed forward path. The ANN estimated outputs are compared with the target output to form a loss function. The gradient of the loss function with respect to different weights of each layer are then back propagated and the weights are updated. This process will continue until the loss function is reduced to a desirable threshold or the maximum number of epochs is reached. Other stop criteria such as validation failure can also be used. Different training algorithms such as back-propagation gradient descent, stochastic gradient descent and Levenberg–Marquardt (LM) have been applied in the state-of-the-art literature [31].

Fig. 3 shows an eight-layer FFNN initially used in this paper which will later be replaced by a more appropriate architecture obtained from the optimization. The inputs to this ANN are the normalized harmonic voltage rms values of the monitored buses, and its output will be the curve-normalized $h$-order rms harmonic currents of the $n$th harmonic source. Hence, an ANN-based harmonic current estimator will be trained and developed for each harmonic order injected by each harmonic source. This procedure outperforms the approach where only one ANN is considered to estimate all variables, mainly because with one separate ANN trained for each harmonic, the task of each ANN will be more specific. This approach has also been recommended in [19].

“See Fig. 3 at the end of the manuscript”

The input harmonic voltage vector contains harmonic voltage rms values at monitored buses. For example, for two monitors $(m_1, m_2)$ and two harmonic orders $(h_1, h_2)$, the input vector for simultaneous consideration of harmonic orders is $[|V_{m1}h_1| \ [V_{m1}h_2]| [V_{m2}h_1]| [V_{m2}h_2]|]^T$, while the input vectors for separate consideration of harmonic orders are $[|V_{m1}h_1|^2 \ |V_{m1}h_2|^2 |V_{m2}h_1|^2 |V_{m2}h_2|^2]^T$ where $T$ denotes the vector transpose operator.

1) REFINEMENT OF THE ANN ARCHITECTURE

When an AI system overfits a dataset, it performs almost perfectly on the training set while the same performance is not achieved on the test set. The gap between the training error and test error is generally observed to be too large [32]. In other words, the generality of the model is low in the case of overfitting. A validation failure check is a process which checks the performance on a validation set as part of the training process. The learning process will be stopped if the performance on the validation set is not improving for several consecutive iterations [32]. It is worth noting that validation is part of the training itself and thus contributes to the learning process, whereas testing is a process conducted after training and is used to evaluate the generality of the model.

Without any validation stopping criterion, if the initial architecture of the ANN leads to overfitting, there is no need to make the ANN more complex. Instead, lighter structures may lead to better generality. Consider a neural network with $p$ layers as $L_1, l = 1, 2, 3, ..., p$. The overfitted network has the size of $S_l = S_1, S_2, S_3, ..., S_p$. A discrete search space (SP) is defined on $\mathbb{N}^P$ as $SP_l = \{0, [S_1/Q], 2\times[S_1/Q], 3\times[S_1/Q], ..., S_l \}$ $l=1,2,...,P$ where square brackets, $[$, denotes the floor function and $Q$ is the number of discrete points in the search space of each layer. Each member in $SP_l$ corresponds to the size of a layer and zeros is correspondent to removing the layer. To illustrate this with an example, consider the ANN of Fig. 3 with $Q=2$. The search space for layers $L1-L6$ will be $L1 \in \{0, 16, 32\}, L2 \in \{0, 5, 10\}, L3 \in \{0, 5, 10\}, L4 \in \{0, 16, 32\}, L5 \in \{0, 16, 32\}$ which implies $3^6=729$ combinations. Each combination can be modelled as a ternary number with six digits. For example, 201112 corresponds the ANN architecture of $\{32, 10, 16, 16\}$ where layer $B$ is removed and so forth. This way of defining the search space can be suitable for evolutionary optimization algorithms like GA.

In this paper, the regularized Mean Squared Error (Reg_MSE) and conventional MSE have been used for training and testing as depicted in (9a) and (9b), respectively. Reg_MSE can be considered as the objective function used to refine the ANN structure, and the sizes of layers (number of neurons in each layer) are the control variables. In Section V, a minimization problem has been solved to obtain the refined ANN architecture by using a brute force algorithm [33] and by considering the training sets with validation check.

$$Re \ g_{-MSE} = 1 \frac{M}{M} \sum_{m=1}^{M} \sum_{s} (x_{s}^{\text{Target}} - y_{s}^{\text{Estimate}})^2 + \frac{\lambda}{2} \frac{M}{M} \sum_{l=1}^{L} w_{l}^{2}$$

(9a)

$$MSE = 1 \frac{M}{M} \sum_{m=1}^{M} (x_{s}^{\text{Target}} - y_{s}^{\text{Estimate}})^2$$

(9b)

2) OVERSAMPLING FOR OVERFITTED ANNS

An oversampling technique which is applicable to regression problem could be useful in dealing with limited and skewed datasets. However, conventional oversampling techniques are not applicable to regression problems while they are efficient.
tools to deal with unbalanced datasets and minor classes in classification tasks by AI. In this subsection, the new Neural Oversampling technique, NOCAR, is proposed which can efficiently address the task of oversampling for any regression problem. As the datasets used in this paper are limited, NOCAR can lead to improvements in accuracy of the harmonic current estimators.

Traditional AI classifiers are vulnerable in learning highly skewed data as they are designed to expect different classes to contribute equally to minimization of the classifiers’ loss functions [34]. One of the well-known oversampling techniques called Synthetic Minority Over-sampling Technique (SMOTE) was proposed in [35] which formed the basis of many other oversampling techniques for classification such as hybrid K-means SMOTE [36] and SMOTE combined with self-organizing maps [34]. The main reason which hinders the application of oversampling for regression problems, is the determination of the target output of the oversampled datapoints. This issue does not happen in classification problems, as the generated datapoints will inherit its output class from the original datapoint. A few methods have been proposed on oversampling for regression. In [37], a SMOTE for Regression (SMOTER) was proposed which oversampled the ANN input data based on the K-means method in the original SMOTE algorithm [35] and estimated the target output of the oversampled data by interpolation between two datapoints contributing to the generation of the new point. SMOTER may be insufficient as the interpolation between two original datapoints does not necessarily follow the innate pattern in the dataset which ANNs aims to find. It can lead to distortion of the dataset from the original pattern in dimensionally large problems.

In NOCAR, the approach for generating ANN input datapoints is similar to the traditional SMOTE algorithm. However, the process of generating the final datasets (combination of newly generated datasets and original datasets) and the process of calculating the target outputs of the generated datapoints are completely novel. In addition, a consensus stage is also proposed for any regression task performed by the AI system. Two oversampling procedures can be conducted: 1) Oversampling on the upper portion of the output data, and 2) Oversampling on the lower portion of the output data. The target outputs of the newly generated datapoints are obtained by training additional ANNs for the initial upper and lower portions of the dataset. Then, a consensus between the upper-oversampling, the lower-oversampling and no-oversampling will be formed.

Let us consider a dataset \( D: (V_d, I_d) \) which is comprised of \( N \) datapoints (input-output pairs). First, \( I_d \) should be sorted in the descending order, and two data percentiles \( p_u \% \) and \( p_l \% \) should be selected for upper-oversampling and lower-oversampling, respectively. The corresponding inputs to the percentiles are \( V_d(p_u \%) \) and \( V_d(p_l \%) \). Based on the definition of percentile, 100-\( p_u \% \) of the data have output values greater than \( I_d(p_u \%) \) and \( p_l \% \) of the data have output values lower than \( I_d(p_l \%) \). Hence, correspondent to \( p_u \% \) and \( p_l \%, \) two subsets for upper-oversampling \((V_d,u, I_d,u)\) and lower-oversampling \((V_d,l, I_d,l)\) are extracted. Two oversampling numbers are chosen as \( N_u \) and \( N_l \) which are the number of generated datapoints per each original datapoint. Initially, for each oversampling subset, an ANN with the similar structure to the main ANN will be trained to map the input-output pairs e.g. \((V_{d,u}, I_{d,u})\) which is denoted by \( ANN_u \) (or \( ANN_l \)). Afterward, for all candidate samples in oversampling subsets \((V_{d,u}, I_{d,u})\) and \((V_{d,l}, I_{d,l})\), the \( N_u \) and \( N_l \) nearest neighbors, respectively, will be obtained by KNN algorithm. The difference between input features of each obtained neighbor and the initial sample is calculated. A uniformly random array (Rand) with the same size of the input feature vector is also generated. The generated sample’s input feature vector, \( OV_u \), is depicted in (10) (a similar equation can be used to calculate \( OV_l \)). Then, the target output vectors corresponding to \( OV_u \) and \( OV_l \) will be obtained using the trained ANNs as \( O_{U} \) and \( O_{L} \). By using ANNs (\( ANN_u \) or \( ANN_l \)) to calculate the target output of the generated datapoints, the features of the oversampled subsets will be better projected in the generated data. However, still the new datapoints are randomly generated and they should not dominate the datasets. Hence, for each generated sample, the original sample will be replicated in the dataset. For example, if two new datapoints are generated for each candidate in \( V_{d,u} \) \((N_u=2)\), the final dataset will be \( V_{d,u,new} = \{ V_{d,u}, OV_{a1}, V_{d,u}, OV_{a2}, V_{d,u} \} \) and \( I_{d,u,new} = \{ I_{d,u}, I_{d,u}, I_{d,u}, I_{d,u} \} \). By replicating the datapoints in the upper/lower subsets plus oversampling, the contribution of the upper/lower portions of the datasets will increase in minimization of the loss function which could be a good practice in dealing with outliers. Choice of upper-oversampling, lower-oversampling or both depends on the application and on the problem which is supposed to be solved. Algorithm 1 depicts the step-by-step implementation of oversampling part of NOCAR.

\[
OV_u = V_{d,u}(initial) + rand \times (V_{d,u}(neighbor) - V_{d,u}(initial))
\]

(10)

NOCAR also contains a consensus stage. After the oversampling stage, three categories of datasets are obtained: \( D: (V_d, I_d), D_{u,new}(V_{d,new}, I_{d,new}), D_{l,new}(V_{d,new}, I_{d,new}) \). The number of generated datasets can be more using more oversampling runs. For each dataset, a neural network will be trained to estimate the outputs (\( \text{ANN}_u \), \( \text{ANN}_l \), and \( \text{ANN}_{u,new} \)). The consensus between the trained ANNs can be achieved by averaging as (11). For the case study of this paper, averaging consensus revealed accurate performance. Other statistical indices can also be employed based on application. Fig. 4 depicts the flowchart of the proposed consensus algorithm.

\[
Output_{final} = \frac{Output_{\text{ANN}} + Output_{\text{ANN}_{u,new}} + Output_{\text{ANN}_{l,new}}}{M_{\text{cons}}}
\]

(11)
Algorithm 1- Proposed Oversampling Algorithm

1. Sort \((V_{p0}, I_{D})\) in the descending order based on \(I_{D}\)
2. Pre-set percentiles \((p_{n}\% \text{ and } p_{l}\%)\) and the oversampling rate \((N_{u}, N_{l})\)
3. Extract subsets \(D_{u}\) and \(D_{l}\)
   3.1) \((V_{p0}, I_{D})_{D_{u}}=\{(V_{p0}, I_{D}) | I_{D} \geq I_{D}(p_{u}\%\)}\)
   3.2) \((V_{p0}, I_{D})_{D_{l}}=\{(V_{p0}, I_{D}) | I_{D} \leq I_{D}(p_{l}\%\)}\)
4. Train ANN\(_{u}\) to map \(V_{D,u}\) to \(I_{D,u}\)
5. Train ANN\(_{l}\) to map \(V_{D,l}\) to \(I_{D,l}\)
6. \(OV_{u}=\text{OversampleInput}(V_{D,u})\)
7. \(OV_{l}=\text{OversampleInput}(V_{D,l})\)
8. \(OI_{u}=\text{ANN}_{u}(OV_{u})\)
9. \(OI_{l}=\text{ANN}_{l}(OV_{l})\)
10. \({\{V_{D_{u,\text{new}}, I_{D_{u,\text{new}}}}\}}=\{V_{D,I_{D}}\}+\{O_{V_{u},O_{I_{u}}}\}+\{V_{D,I}_{D,u}\}\)
    // plus (+) is concatenate operator
11. \({\{V_{D_{l,\text{new}}, I_{D_{l,\text{new}}}}\}}=\{V_{D,I_{D}}\}+\{O_{V_{l},O_{I_{l}}}\}+\{V_{D,I}_{D,l}\}\)
    // plus (+) is concatenate operator

OversampleInput Function

Function \(OV_{u}=\text{OversampleInput}(V_{D,u}, I_{D,u}, N_{u})\)
// For lower oversampling, \(u\) is replaced by \(l\)
// Size is the number of samples in \(D_{u}\) or \(D_{l}\)

1. \(I_{\text{sample}}=V_{D,u}(i)\) // Initial Sample
2. \(V_{\text{sample}}=V_{D,u}(i)\)
3. \(I_{\text{sample}}=I_{D,u}(i)\)
4. \(I_{\text{neighbor}}=K\)-nearest Neighbors \((I_{\text{sample}})\)
// K nearest neighbors to \(I_{\text{sample}}\) in \(I_{D,u}\)

for \(j=1\) to \(K\)
5. \(I_{d}=\text{find} (I_{D,u} = I_{\text{neighbor}}(j))\)
6. \(V_{\text{neighbor}}(j) = V_{D,u}(I_{d})\)

for \(i=1\) to \(\text{Size}\)
7. \(OV_{u} = V_{D,u}(\text{initial}) + \text{rand} \times (V_{D,u}(\text{neighbor}) - V_{D,u}(\text{initial}))\)
8. \(\text{end for}\)
9. \(\text{end for}\)

“See Fig. 4 at the end of the manuscript”

IV. COMPARISON WITH STATE-OF-THE-ART TECHNIQUES

To show the competence of the proposed technique, it has been compared with GALS [13], SVD [17] and BL [24]. The methods have been compared for a synthetic dataset generated on the IEEE 14-bus test system as shown in Fig. 5. The network information can be found in [38]. The uncertainty in harmonic voltage measurements and network parameters has been modelled similar to [24]. An undetermined harmonic state estimation has been performed for four harmonic sources at Buses 5, 10, 12 and 13 and harmonic orders 3, 5, 7 and 11. Three harmonic voltage meters are assumed at Buses 2, 6 and 11, based on the optimal meter placement technique of [17]. As the location of harmonic sources are assumed to be known, the harmonic location algorithm of [13] was skipped. However, the method of [13] requires information of at least one Zero Injection Bus (ZIB) for each harmonic order. Hence, an extra equation regarding zero harmonic current injection at Bus 3 is included to implement the method of [13]. The proposed technique was simulated with the oversampling parameters of \(p_{n}\% \in \{80,70\}\) and \(p_{l}\% \in \{20,30\}\), \(N_{u}=N_{l}=2\) and consensus of five members. 1000 datapoints were simulated and randomly divided into the training sets, validation sets and the test set with the ratio of 65%, 15% and 20%. The ANN structure of Fig. 3 is utilized in this section. Hyperbolic tangent and linear functions have been used as activation functions in the hidden layers and output layers, respectively, which revealed better performance than the sigmoid function for the case study. \(\text{Reg} \_\text{MSE}\) is used for training and \(\text{MSE}\) for performance evaluation and testing. The error threshold is set to 0.0001; regularization factor and learning rate are chosen 0.05 and 0.1, respectively; maximum number of epochs is 100. The proposed methodology was implemented in MATLAB 2020b.

The results are presented for the test sets, which were not used in the training stage. Fig. 6 shows the comparison of \(\text{MSE}\) on per unit harmonic currents for different HSE techniques. It can be seen that the proposed method is performing as well as GALS and SVD, even though it does not employ harmonic phase angle measurements. The identical behavior of GALS and SVD methods is because the two techniques are mathematically analogous when the location of harmonics sources are known. For harmonic sources at Buses 6 and 10, the estimation by all four methods is accurate. BL revealed more error in some cases, owing to the undetermined nature of the problem. Although the proposed technique may have high computational burden during the training stage, it only requires one calculation of the forward path for each ANN at the estimation stage, in contrast to the iterative nature of BL [24] and GALS [13]. Accordingly, the simulations show that only SVD has a faster response than the proposed technique. Table I summarizes the measurements and information required for each technique, highlighting how the proposed technique has the minimum requirement. Although currently not used, harmonic phase angles could also be embedded into the proposed method, by using real and imaginary values of harmonic voltage measurements instead of rms values, potentially leading to further accuracy improvements.

“See Fig. 5 at the end of the manuscript”

“See Fig. 6 at the end of the manuscript”

Table I. Required measurements and information for different techniques

| Method | GALS [13] | SVD [17] | BL [24] | proposed |
|--------|-----------|---------|---------|----------|
| harmonic voltage magnitude | × | × | × | × |
V. IMPLEMENTATION FOR A REAL TRANSMISSION NETWORK

In this section, a real transmission network with high penetration of wind energy (four operating wind farms: wf1-wf4) has been studied. The simplified single line diagram of the Tasmanian transmission network is shown in Fig. 7. On the generation side, wind farms are one of the main harmonic sources. There are eight harmonic voltage meters to measure harmonics of different orders. Four of these meters (M1, M5, M7 and M8) are electrically closer to the wind farms than the other four. The recorded data are 10-min average harmonic rms values in per unit. A limited dataset of around 1000 data points measured from March 29th, 2021, to April 4th, 2021, has been used. The data only contains harmonic voltage rms values and harmonic current rms values. Wind farms on average produced 20% of the total energy generation of the network during this period, with maximum output from the wind farms reaching 500 MW (about 60% of total system generation).

“See Fig. 7 at the end of the manuscript”

The proposed harmonic estimation method uses harmonic voltage rms values of monitored buses to determine harmonic current rms values injected by each of the wind farms. The problem is made challenging because the ANNs’ inputs, harmonic voltages, have low variance while the target outputs, harmonic currents, have high variance. In addition, only 10-minute average harmonic data is available, and so wind farm output and thus injected harmonic currents can change considerably between two adjacent samples in the dataset. In this case study, harmonic current estimation has been evaluated for harmonic orders \( h=3, 5, 7, 11 \). \( P_{sh} \) should be replaced by \( P_{sh} \) in (3) and (4) where \( P_{sh} \) is the vector of \( h \)-order harmonic current rms values injected by the wind farms.

For each harmonic order and each wind farm, an intelligent harmonic monitor comprised of several ANNs are trained. Hence, 16 intelligent monitors have been trained separately. Consideration of different number of harmonic voltage meters has been assessed in developing the intelligent harmonic monitors. In addition, simultaneous and separate consideration of harmonic orders in the input feature vectors has been compared. This evaluation addresses the question whether variables of a harmonic order can be analyzed and estimated independently of other harmonic orders.

The ANNs have been trained with the same activation functions and training parameters mentioned in Section IV. The data is again randomly divided into 65% training, 15% validation and 20% test sets. For all the methods and procedures elaborated in Section III, only training sets have been used applying LM training algorithm. It is worth noting that all error curves and error values presented in this section are calculated using curve-normalized outputs while the estimated/real value graphs are in electrical per unit. Similar to section IV, MATLAB 2020b has been used to implement AI harmonic current estimators.

A. OVERFITTING IN THE HARMONIC ESTIMATORS

Initially, the ANN structure depicted in Fig. 3 has been used to develop the harmonic current estimators of the wind farms, with monitors at M2, M3, M4 and M6. Validation and oversampling are not used in this instance, revealing that the ANNs may overfit the data. The estimation is almost exact on the training sets for all ANNs although the performance on the test sets is poor. Fig. 8 shows the estimated and real values of 3rd harmonic current of wf4 and 7th harmonic current of wf2 for 100 datapoints in the training sets and in the test sets. As can be seen from the figure, ANNs are suffering from overfitting because of the limited datasets. Remedial actions are required to mitigate this problem.

“See Fig. 8 at the end of the manuscript”

B. REFINING THE ANN ARCHITECTURE

The initial ANN architecture leads to overfitting, but can be simplified and lightened to improve accuracy. Optimizing the ANN structure is not the focus of this paper; however, the process proposed and described earlier can be used to obtain a more suitable structure. First, ANNs with the initial architecture of Fig. 3 are trained using 15% of the training sets randomly selected for validation in each run. Fig. 9 shows the MSE values for the test sets for each of the 16 harmonic current estimators. Since the outputs of all estimators are curve-normalized, their MSEs are comparable with each other. As can be seen in Fig. 9, the error for harmonic 7 of wf2 is significantly higher than the errors for other estimated harmonics. Hence, this estimator \((h=7, \text{wf2})\) has been used to find a unique refined structure of all ANNs. This methodology could be applied to each ANN separately, but this is not the focus of this paper.

“See Fig. 9 at the end of the manuscript”

The refined structure of ANNs is obtained by using training dataset of 7th harmonic of wf2. Based on the coding method introduced in section III, the initial ANN structure is coded as (22222). Every candidate ANN structure has a 6-digit ternary code. To illustrate the search space and the objective function in 3-D, the first three digits are considered on X-axis and the second three digits are considered on Y-axis. For example...
will be divided into $X=(212)_3=23$ and $Y=(022)_2=8$. Accordingly, the discrete range of $X$-axis and $Y$-axis is 0 to 26. The search space and the best solution are shown in Fig. 10. The optimal solution is $X=26=(222)_3$ and $Y=16=(121)_2$, which implies ANN layer sizes of $\{32, 10, 10, 16, 32, 8\}$. This refined structure is used for all ANNs in the upcoming subsections of this study. Fig. 11a and Fig. 11b show the real values and estimated values of 3rd harmonic current of $wf_4$ and of 7th harmonic current of $wf_2$ for the test sets, respectively, by using the refined ANN structure. In spite of improvement, it can be seen that the accuracy is not high at some points, yet. In Subsection E, oversampling has been employed to improve the performance further.

“See Fig. 10 at the end of the manuscript”

“See Fig. 11 at the end of the manuscript”

C. IMPACT OF NUMBER OF HARMONIC VOLTAGE MONITORS ON HSE

Reducing the number of required harmonic monitors leads to a reduction in cost of a harmonic monitoring system [17, 39]. There are eight harmonic voltage meters in the transmission network ($M1-M8$) of our case study. There is a trade-off between the cost of the harmonic monitoring system and harmonic observability. By reducing the number of monitors below the number of unknown harmonic variable, the system of equations will be mathematically underdetermined which is more challenging to solve.

Different cases of harmonic voltage monitors have been studied to estimate the harmonic currents of the wind farms, as shown in Table II. We study the impact of the number of voltage meters used, by creating eight separate cases, each with successively fewer meters being used to train and test ANNs, as per Table II. Case 1 employs the maximum number of harmonic voltage monitors available; however, the wind farms still remain unmonitored. Case 5 is based on four harmonic voltage meters which are further from the wind farms than other harmonic meters. Hence, for choice of four harmonic monitors out of eight, Case 5 is the most challenging. Fig. 12 shows the box plot of MSE for the test sets for 16 trained ANNs. The estimation errors increase with decreasing the number of monitors used, which is generally as expected. The red pluses (+) considered as outliers in the box plot, correspond to the estimation of 7th harmonic current of $wf_2$, in each case. Figs. 13a-13d depict the 3rd harmonic current of $wf_1$, 5th harmonic current of $wf_2$, 7th harmonic current of $wf_3$ and 11th harmonic current of $wf_4$, respectively. The ANN estimation and real value on 50 datapoints of the test sets for Cases 1, 5 and 7 are compared. It can be seen that Case 1 and Case 5 with a greater number of monitors than in Case 7, are performing better than Case 7 for majority of the samples. The accuracy in Case 1 is mostly better than the accuracy in Case 5. The network is not thriving in estimating the high harmonic values as there are not enough of such points in the datasets for ANNs to learn their features which will be addressed by the proposed oversampling technique in Subsection E.

| Case # | Harmonic Voltage Monitors |
|--------|---------------------------|
| Case 1 | $\{M1, M2, M3, M4, M5, M6, M7, M8\}$ |
| Case 2 | $\{M2, M3, M4, M5, M6, M7, M8\}$ |
| Case 3 | $\{M2, M3, M4, M5, M6\}$ |
| Case 4 | $\{M2, M3, M4, M5, M6\}$ |
| Case 5 | $\{M2, M3, M4, M6\}$ |
| Case 6 | $\{M3, M4, M6\}$ |
| Case 7 | $\{M3, M6\}$ |
| Case 8 | $\{M3\}$ |

“See Fig. 12 at the end of the manuscript”

“See Fig. 13 at the end of the manuscript”

D. SIMULTANEOUS AND SEPARATE CONSIDERATION OF HARMONIC ORDERS

It is a common approach to analyze each harmonic order independently of other harmonic orders [23], [40], [41]. However, there is correlation between different harmonic orders which may not be negligible, particularly in power systems with high penetration of RES. In the dataset of the transmission network, there is high correlation between different harmonic orders of the same wind farm as they are generated by the identical harmonic source. In this subsection, harmonic currents have been estimated by considering harmonic voltage orders, simultaneously and separately. Cases of eight harmonic monitors and four harmonic monitors (Case 1 and Case 5 in Table II) have been evaluated. Fig. 14 depicts MSE of the test sets for Case 1 and Case 5 considering harmonic orders simultaneously and separately. For all harmonic current estimators, separate consideration of harmonic voltages of each order leads to larger error than its corresponding case of simultaneous consideration. Some of the ANN estimators are even unable to learn the features on the training sets when harmonic orders are analyzed separately. For example, ANNs developed based on separate consideration of harmonic orders could not appropriately learn to estimate 5th harmonic current of $wf_3$ with 8 monitors and 11th harmonic current of $wf_4$ with 4 monitors throughout training, as shown in Figs. 15a and 15b, respectively. This suggests that the correlation between different harmonic orders should not be ignored.

“See Fig. 14 at the end of the manuscript”

“See Fig. 15 at the end of the manuscript”

E. NOCAR TO IMPROVE HARMONIC CURRENT ESTIMATION

In this subsection, the dataset will be oversampled for cases of eight and four harmonic voltage monitors (Case 9 and Case 10

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
in Table III). The locations of harmonic voltage meters are similar to Case 1 and Case 5 of Table II. Two upper percentiles of \( p=\{90,80\} \) and two lower percentiles \( p=\{30,40\} \) have been utilized. As there are many zeros in the output data, 30% and 40% have been chosen for lower oversampling. Lower values of \( p \) undesirably would have led to majority of oversampled datapoints being zeros. For the upper oversampling, values of 80% and 90% were chosen to make the generated datasets inclined toward the upper portion of data which contains the outliers. The number of generated datapoints for each oversampled datapoint is \( N_t=5 \) , \( N_t=2 \) for upper oversampling and lower oversampling. It should be noted that for each generated datapoint, the original datapoint is replicated so that the newly generated data does not dominate the dataset. Three consensus indices are defined for both cases as in (12) with parameters shown in Table III.

Oversampling is performed on both lower percentile and upper percentile of the data to keep the balance and hence, conserve the original features of the data. \( I_{c31} \) and \( I_{c32} \) represent consensus of five participants for cases of four and eight harmonic voltage monitors, respectively.

\[
I_{c3i} = \frac{I_{1i} + I_{2i} + I_{5i}}{3} \\
I_{c2i} = \frac{I_{1i} + I_{3i} + I_{4i}}{3} \\
I_{c3i} = \frac{I_{1i} + I_{2i} + I_{3i} + I_{4i} + I_{5i}}{5}
\]

\( i = 1,2 \)

**Table III. Output current parameters of oversampling cases for the real dataset**

| Case # | Number of Harmonic Monitors | Original Data | Oversampling parameters |
|--------|---------------------------|---------------|-------------------------|
| 9      | 4                         | \( I_{11} \)   | \( p=90 \) %        |
| 10     | 8                         | \( I_{12} \) | \( p=80 \) %        |
|        |                            | \( I_{22} \)   | \( p=30 \) %        |
|        |                            | \( I_{31} \)   | \( p=40 \) %        |
|        |                            | \( I_{42} \)   |                         |
|        |                            | \( I_{51} \)   |                         |
|        |                            | \( I_{52} \)   |                         |

Moreover, hardware-based solutions such as installing more harmonic monitors can be much more costly than data analysis and intelligent techniques.

“See Fig. 16 at the end of the manuscript”

Harmonic current estimations by various ANNs are presented in Fig. 17a-17h to prove the effectiveness of NOCAR along with other techniques proposed in this paper. Harmonic current outputs of consensus \( I_{c31} \) and \( I_{c32} \) are compared with real values for 100 data points. A couple of points are noticeable in Fig. 17. First, the ANN-based harmonic current estimations are accurate for the majority of datapoints and the performance has been improved by the integration of ANN and oversampling as NOCAR. However, the network performance on 7th harmonic is not as good as other harmonic orders. This can also be understood from the outliers of box plots in Fig. 16.

Error Reduction Rate (ERR) as a measure to show the improvement introduced by NOCAR is defined as (13). Fig. 18 shows ERR of 16 harmonic current estimators for consensus #3 in cases of four monitors and eight harmonic voltage monitors (\( I_{c31} \) and \( I_{c32} \)). As revealed in the figure, NOCAR has led to error reduction and accuracy improvement for both cases of four and eight harmonic voltage monitors. The improvement for some cases of eight monitors is less than that of four monitors which is because the initial error for cases of eight monitors has been less.

\[
ERR(\%) = \frac{\text{MSE}_{\text{no-oversample}} - \text{MSE}_{\text{NOCAR}}}{\text{MSE}_{\text{no-oversample}}} \times 100
\]

“See Fig. 17 at the end of the manuscript”

“See Fig. 18 at the end of the manuscript”

In spite of the improvement provided by NOCAR, there are still singular points in the test sets with relatively high magnitude for which ANNs are not performing as well as other datapoints. These singular points are accounted as outliers and there are not enough samples of them for the ANNs to learn their features, properly. However, these spiky outliers could be unimportant as long as the PQ indices introduced by the standards are estimated accurately. It is recommended in IEEE std 519 that 95th and 99th percentiles of weekly harmonic currents should not violate the provided harmonic current limits. Figs. 19a and 19b depict the real values and the estimated values of 95th and 99th percentiles of the harmonic currents of the wind farms, respectively. The estimations are based on consensus with five members. As can be seen from Fig. 19a, the estimation of 95th percentiles is quite accurate at all instances for case of eight harmonic monitors. High accuracy in 95th percentiles of harmonic currents is also achieved for majority of instances of four monitors. While still good accuracy is obtained in estimating the 99th percentiles for case of eight monitors, the 99th percentiles for case of four monitors is slightly different from...
real values. Larger datasets can improve the results of cases with four harmonic monitors.

“See Fig. 19 at the end of the manuscript”

VI. CONCLUSION

In this paper, an intelligent monitoring system based on ANN was developed for harmonic current estimation of unmonitored harmonic sources, based on harmonic voltages at monitored buses. Synthetic and real datasets were used, both of which were limited in terms of the number of meters, the number of datapoints and the measured variables available. The developed system maps rms harmonic voltages measured in some locations in a power system to injected rms harmonic currents at other locations. A new curve-normalization of the target outputs was introduced to reduce the skewness of the outputs, and a novel neural oversampling technique, NOCAR, was proposed to address overfitting. Estimation performance was improved when these proposed methodologies were employed. Comparison of performance with GALS, SVD and BL, state-of-the-art techniques in the literature, demonstrated the effectiveness of the proposed technique. Applied to data collected from a real power system, the performance of the harmonic estimators was studied with limited number of harmonic voltage meters. Moreover, both simultaneous or separate consideration of harmonic voltage orders in training ANN-based current estimators for each harmonic order were investigated and analyzed.

A. MAJOR FINDINGS AND FUTURE RESEARCH

The major findings of this paper are as follows.

1) Simultaneous analysis of different harmonic orders yields better harmonic estimation accuracy than when each harmonic order is estimated independently of other orders.

2) It was demonstrated that by employing NOCAR and ANNs, the performance of harmonic current estimators with only four voltage monitors got close to the performance achieved with eight harmonic monitors. This suggests that software-based data analysis techniques are worthy of investigation, as opposed to only relying on hardware-based remedies such as installation of new monitors at harmonic source injection points.

3) NOCAR, as an oversampling technique for the regression problem, introduced some improvement in the performance of the proposed harmonic monitoring system. In addition, it can be a beneficial tool in the development of AI systems for other applications with regression tasks.

4) Compared to other techniques, ANN-based harmonic estimators can provide accurate results in real-time with minimum required information.

The following research directions could be pursued in future.

1) The advantage of curve-normalization could be revealed for estimation of all harmonic current sources using only one ANN.

2) By combining data techniques and direct harmonic monitoring, develop methods for determination of the minimum number of harmonic monitors and their optimal locations.

3) Integration of harmonic phase angles as input features of the ANNs for more accurate harmonic estimation.

4) Accurate assessment of harmonic measurement outliers, and analysis of their importance from an electrical system outlook.

ACKNOWLEDGMENT

We would like to acknowledge the assistance of TasNetworks and Mr. Pablo Oriol for providing data used in this research.

REFERENCES

[1] N. Ertugrul and D. Abbott, “DC is the Future [Point of View],” Proc. IEEE, vol. 108, no. 5, pp. 615-624, May 2020, DOI: 10.1109/JPROC.2020.2982707.

[2] IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems,” IEEE Std 519-2014 (Revision of IEEE Std 519-1992), 2014.

[3] Electromagnetic compatibility (EMC) - Part 3: Limit - ALL PARTS, IEC 61000-3, 2022.

[4] A. Esram, M. Negnevitsky, E. Franklin, and S. Lyden, “Review of AI applications in harmonic analysis in power systems,” Renew. Sust. Energy, Rev., vol. 154, p. 111897, Feb. 2022.

[5] N. Nakhodchi, T. Busatto, and M. Boilén, “Measurements of Harmonic Voltages at Multiple Locations in LV and MV Networks,” in 19th Int. Conf. on Harm. Quality Power (ICHQP), Dubai, UAE, 2020, pp. 1-5.

[6] L. Whei-Min, L. Chia-Hung, T. Keng-Pang, and W. Chien-Hsien, "Multiple harmonic source detection and equipment identification with cascade correlation network," IEEE Trans. Power Del., vol. 20, no. 3, pp. 2166-2173, June 2005, DOI: 10.1109/TPWRD.2004.843462.

[7] K. K. C. Yu, N. R. Watson, and J. Arrillaga, “An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking,” IEEE Trans. Power Del., vol. 20, no. 2, pp. 1577-1584, April 2005, DOI: 10.1109/TPWRD.2004.838643.

[8] H. Liao, "Power System Harmonic State Estimation and Observability Analysis via Sparsity Maximization,” IEEE Trans. Power Syst., vol. 22, no. 1, pp. 15-23, Feb. 2007, DOI: 10.1109/TPWRS.2006.887957.

[9] E. Moghanlu, M. Göld, and S. Ö. "State-Estimation-Based Determination of Harmonic Current Contributions of Iron and Steel Plants Supplied from PCC," IEEE Trans. on Ind. Appl., vol. 52, no. 3, pp. 2654-2663, May 2016, DOI: 10.1109/TIA.2016.2521598.

[10] G. T. Heydt, "Identification of harmonic sources by a state estimation technique,” IEEE Trans. Power Del., vol. 4, no. 1, pp. 569-576, Jan. 1989, DOI: 10.1109/61.19248.

[11] N. Kano, M. Yamashita, H. Yanagida, M. Mizukami, Y. Hayashi, and J. Matsuki, "Power system harmonic analysis using state-estimation method for Japanese field data," IEEE Trans. Power Del., vol. 20, no. 2, pp. 970-977, April 2005, DOI: 10.1109/TPWRD.2004.838632.

[12] C. Rakpenthai, S. Uatrongjit, N. R. Watson, and S. Premrudefeechacharn, "On Harmonic State Estimation of Power System with Uncertain Network Parameters,” IEEE Trans. Power Syst., vol. 28, no. 4, pp. 4829-4838, Nov. 2013, DOI: 10.1109/TPWRS.2013.227943.

[13] F. Xu, C. Wang, K. Guo, Q. Shu, Z. Ma, and H. Zheng, “Harmonic Sources’ Location and Emission Estimation in Underdetermined Measurement System,” IEEE Trans. Instrum. Meas., vol. 70, pp. 1-11, May 2021, DOI: 10.1109/TIM.2021.3077658.

[14] M. Bettayeb and Q. Uvais, "A hybrid least squares-GA-based algorithm for harmonic estimation," IEEE Trans. Power Del., vol. 18, no. 2, pp. 377-382, April 2003, DOI: 10.1109/TPWRD.2002.807458.

[15] I. Molina-Moreno, A. Medina, R. Cisneros-Magaña, O. Anaya-Lara, and J. A. Salazar-Torres, "Enhanced harmonic state estimation in unbalanced three-phase electrical grids based on the Kalman filter and
physical scale-down implementation,” *Int. J. Electr. Power Energy Syst.*, vol. 123, p. 106243, Dec. 2020.

[16] V. Bčeirović, I. Pavić, and B. Filipović-Grčić, “Sensitivity analysis of method for harmonic state estimation in the power system,” *Elect. Power Syst. Res.*, vol. 154, pp. 515-527, Jan. 2018.

[17] C. Madhurad, S. Premrudeepreechacharn, N. R. Watson, and S.-U. Rachwai, “An optimal measurement placement method for power system harmonic state estimation,” *IEEE Trans. Power Del.*, vol. 20, no. 2, pp. 1514-1521, April 2005, DOI: 10.1109/TPWRD.2004.841309.

[18] C. F. M. Almeida and N. Kagan, “Harmonic state estimation through optimal monitoring systems,” *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 467-478, March 2013, DOI: 10.1109/TSG.2012.2235472.

[19] R. K. Hartana and G. G. Richards, “Harmonic source monitoring and identification using neural networks,” *IEEE Trans. Power Syst.*, vol. 5, no. 4, pp. 1098-1104, Nov. 1990, DOI: 10.1109/5.999358.

[20] R. K. Hartana and G. G. Richards, “Constrained neural network-based identification of harmonic sources,” *IEEE Trans. Ind. Appl.*, vol. 29, no. 1, pp. 202-208, Jan. 1993, DOI: 10.1109/28.195908.

[21] B. Adineh, M. R. Habibi, A. N. Askolat, and F. Blaabjerg, “Sensorless Voltage Estimation for Total Harmonic Distortion Calculation Using Artificial Neural Networks in Microgrids,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 7, pp. 2583-2587, Feb. 2021, DOI: 10.1109/TCSII.2021.3095410.

[22] A. Rahmani and A. Dehimi, “Reduction of harmonic monitors and estimation of voltage harmonics in distribution networks using wavelet analysis and NARX,” *Elect. Power Syst. Res.*, vol. 178, p. 100464, Jan. 2020.

[23] W. Zhou, O. Ardakian, H. T. Zhang, and Y. Yuan, “Bayesian Learning-Based Harmonic State Estimation in Distribution Systems with Smart Meter and DPMU Data,” *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 832-845, Sep. 2020, DOI: 10.1109/TSG.2019.2938733.

[24] G. D. Antona, C. Muscas, P. A. Pegoraro, and S. Sulis, “Harmonic Source Estimation in Distribution Systems,” *IEEE Trans. Instrum. Meas.*, vol. 60, no. 10, pp. 3351-3359, Apr. 2011, DOI: 10.1109/TIM.2011.2126910.

[25] D. Saxena, S. Bhaumik, and S. N. Singh, “Identification of Multiple Harmonic Sources in Power System Using Optimally Placed Voltage Measurement Devices,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2483-2492, July 2014, DOI: 10.1109/TIE.2013.2270218.

[26] A. Moradifar, A. Akbari Foroud, and K. Gorgani Firouzjah, "Comprehensive identification of multiple harmonic sources using fuzzy logic and adjusted probabilistic neural network," *Neural Comput. Appl.*, vol. 31, no. 1, pp. 543-556, Jan. 2019.

[27] A. Alavi, N. Alavi, A. R. Kaliar, Z. Saleem, and N. Khan, "Review of harmonic analysis, modeling and mitigation techniques," *Renew. Sust. Energy Rev.*, vol. 78, pp. 1152-1187, Oct. 2017.

[28] S. K. Jain and S. N. Singh, "Harmonics estimation in emerging power system: Key issues and challenges," *Elect. Power Syst. Res.*, vol. 81, no. 9, pp. 1754-1766, Sep. 2011.

[29] P. Trebula, J. Halčinová, M. Filo, and J. Markovič, "The importance of normalization and standardization in the process of clustering," in *IEEE 12th Int. Symp. Appl. Mach. Intell. Inform. (SAMII)*, Herľany, Slovakia, 2014, pp. 23-25.

[30] M. Negnevitsky, “Artificial Neural Networks” in *Artificial intelligence: a guide to intelligent systems.*, 2nd ed., Essex, England: Pearson Ed, 2005, ch. 6, pp. 165-216.

[31] M. Tkáč and R. Vermer, "Artificial neural networks in business: Two decades of research," *Applied Soft Comput.*, vol. 38, pp. 788-804, Jan. 2016.

[32] I. Goodfellow, Y. Bengio, and A. Courville, “Machine Learning Basics” in *Deep learning*. 2nd ed., Cambridge, USA: MIT press, 2017, ch. 5, sec. 5.2, pp. 125-135.

[33] G. Grothendieck, “Non-linear regression with brute force,” *R package version 0.2*, 2013. Available: https://cran.r-project.org/web/packages/rpart/index.html

[34] G. Douzas, R. Rauch, and F. Bacao, "G-SOMO: An oversampling approach based on self-organized maps and geometric SMOTE," *Expert Syst. Appl.*, vol. 183, p. 115230, Nov. 2021.

[35] N. V. Chawlà, K. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *J. Artif. intel. res.*, vol. 16, pp. 321-357, June 2002.

[36] Z. Xu, D. Shen, T. Nie, Y. Kou, N. Yin, and X. Han, “A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data,” *Inform. Sci.*, vol. 572, pp. 574-589, Sep. 2021.

[37] L. Torgo, R. P. Ribeiro, B. Pfahringer, and P. Branco, “SMOTE for Regression,” in *Progress in Artificial Intelligence*, Berlin, Germany: Springer, 2013, pp. 378-389.

[38] R. Christie, J. Dabbagchi. "Power Systems Test Case Archive." University of Washington. Available: https://labs.ece.uw.edu/psta/pf14/pg_tca14bus.htm (accessed 25.02.2022).

[39] I. D. Melo, J. L. R. Pereira, P. F. Ribeiro, A. M. Variz, and B. C. Oliveira, "Harmonic state estimation for distribution systems based on optimization models considering daily load profiles," *Elect. Power Syst. Res.*, vol. 170, pp. 303-316, May 2019.

[40] L. Whei-Min, Z. Tung-Sheng, and T. Ming-Tong, "Multiple-frequency three-phase load flow for harmonic analysis," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 897-904, May 2004, DOI: 10.1109/TPWRS.2004.825906.

[41] M. Amini, A. Jalilian, and M. R. Pour Behbahani, "Fast network reconfiguration in harmonic polluted distribution network based on developed backward/forward sweep harmonic load flow," *Elect. Power Syst. Res.*, vol. 168, pp. 295-304, Jan. 2019.

Ahmadreza Eslami received B.Sc. and M.Sc. degrees in electrical engineering from Shiraz University in 2011 and M.Sc. degree in power systems from Isfahan University of Technology, Iran, in 2014. He is currently pursuing his Ph.D. degree with the college of science and engineering, University of Tasmania, Hobart, TAS, Australia. His current research interests include modern power system analysis, power quality, integration of renewable energy sources and artificial intelligence applications in power systems.

Michael Negnevitsky (Senior Member, IEEE) received the B.S.E.E. (Hons.) and Ph.D. degrees from the Belycrousian University of Technology, Minsk, Belarus, in 1978 and 1983, respectively. He is currently the Chair Professor in power engineering and computational intelligence, University of Tasmania, Hobart, TAS, Australia. His research interests include power system analysis, power economics, microgrids and smart grids, and intelligent systems applications in power systems. He is a Chartered Professional Engineer and Fellow of Engineers Australia.

Evan Franklin (Member, IEEE) received the BE (Hons.) degree from the University of Tasmania in 1996 and the Ph.D. degree from the Australian National University in 2006. He has authored numerous journal papers, conference papers and patents, and has a strong record of industry collaboration. His research interests include the integration of renewable energy generation into distribution networks and power systems, and the role of energy storage in future energy systems.
Sarah Lyden (S’12) received the Graduate degree with a B.Sc.-B.E.(Hons.) and the Ph.D. degree in engineering from the University of Tasmania, Hobart, Tas., Australia, in 2011 and 2015, respectively. She is currently a Lecturer at the School of Engineering and ICT, University of Tasmania. Her research interests include control and grid integration of renewable energy sources, microgrid systems, and engineering education.
Fig. 1. Data analysis in power systems: time domain (left) vs frequency domain (right)

Fig. 2. Flowchart of the AI harmonic Estimator
Fig. 3. Initial Architecture of the ANNs for harmonic current estimation

![Diagram](image1)

**Data** 

- Upper-Oversampling
- **Data** 
  - **Data** 
    - **ANNU**
  - **ANNU**

- Lower-Oversampling
- **Data** 
  - **Data** 
    - **ANNL**
  - **ANNL**

- **Average**
- **Final Output**

Fig. 4. Flowchart of the proposed NOCAR

![Diagram](image2)

Fig. 5. IEEE 14-bus power system
Fig. 6. MES on per unit currents for different HSE techniques.

Fig. 7. Simplified single line diagram of Tasmanian transmission network with wind farms and voltage meters.
Fig. 8. Harmonic Current Estimation with No oversampling or Validation: 1) 3rd harmonic current of wf4 a) training set, b) test set 2) 7th harmonic current of wf2 c) training set d) test set

Fig. 9. MSE of test sets for initial ANN architecture

Fig. 10. Search space and the best solution for refining the structure of ANNs
Fig. 11. Harmonic current estimation of the test sets using the refined ANN structure without oversampling: a) 3rd harmonic current of wf4, b) 7th harmonic current of wf2

Fig. 12. Plot Box of ANN errors on the test sets for different number of harmonic voltage monitors
Fig. 13. Harmonic current estimation considering different number of monitors: a) 3rd harmonic current of $w_f$1 b) 5th harmonic of $w_f$2 c) 7th harmonic current of $w_f$3 d) 11th harmonic current of $w_f$4.

Fig. 14. MSE for ANN estimators on the testing set and harmonic voltage orders considered separately and simultaneously.
Fig. 15. Estimation of harmonic currents on the training sets considering harmonic orders simultaneously and separately: a) 8 harmonic monitors b) 4 harmonic monitors

Fig. 16. MSE Boxplot for original outputs and NOCAR cases as presented in table II
Fig. 17. Harmonic current estimation of different wind farms by consensus #3 for 4 and 8 monitor cases
Fig. 18. Error Reduction Rate for NOCAR consensus 3: $I_{31}$-four harmonic monitors, $I_{32}$-eight harmonic monitors

Fig. 19. Comparision of Estimation and real values of a) 95\textsuperscript{th} percentiles of harmonic current rms values, b) 99\textsuperscript{th} percentiles of harmonic current rms values for cases of four monitors and eight monitors