



DATA-DRIVEN VALIDATION AND CORRECTION OF THE DISTORTION FACTOR ESTIMATED FROM CURRENT TOTAL HARMONIC DISTORTION USING INDUSTRIAL FIELD MEASUREMENTS

Faruk KÜRKER^{1*}, Ramazan TAŞALTIN²

¹Adıyaman University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 02040, Adıyaman, Türkiye

²Kahramanmaraş Sütçü İmam University, Faculty of Engineering and Architecture, Department of Electrical and Electronics Engineering, 46040, Kahramanmaraş, Türkiye

Abstract: In this study, the validity of the theoretical distortion factor based on total harmonic distortion of current (THDI) under industrial field conditions was evaluated, and low-complexity, data-driven correction models were developed to reduce formula error. The dataset was compiled from three-phase field measurements at 12 industrial measurement points. The measurement-based distortion factor was calculated from the PF/dPF ratio, while the theoretical distortion factor was calculated from the THDI-based equation. Since the phases at the same measurement point were not considered independent, the models were evaluated using a 12-cycle out-of-sample cross-validation method. The theoretical formula produced a mean absolute error of 0.015662, a root mean square error of 0.019681, and a mean bias of -0.009357. Among the compared models, restricted second-order regression produced the lowest error values, providing improvements of 31.83% and 15.70% in MAE and RMSE, respectively. Ridge regression, on the other hand, reduced the MAE to 0.012263 and brought the mean bias close to zero. However, the data-driven models did not reduce the maximum absolute error. At the measurement-point level, the Wilcoxon signed-rank test indicated lower paired errors after Ridge correction ($P=0.0342$), whereas the 95% bootstrap confidence interval for the mean paired MAE difference included zero. Taken together, these findings provide positive but limited evidence regarding the error-reduction potential of the Ridge correction. The results show that the THDI-based formula provides an interpretable initial estimate but cannot, on its own, account for field conditions such as phase asymmetry, load level, power factor, and voltage distortion. The proposed correction approach should be considered not as a universal formula but as an empirical calibration approach specific to the available field data.

Keywords: Distortion factor, Total harmonic distortion of current, Power factor, Industrial field measurements, Ridge regression, Out-of-measurement-point cross-validation

*Corresponding author: Adıyaman University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 02040, Adıyaman, Türkiye

E mail: fkurker@adiyaman.edu.tr (F. KÜRKER)

Faruk KÜRKER  <https://orcid.org/0000-0003-1544-9743>

Ramazan TAŞALTIN  <https://orcid.org/0000-0003-2026-2430>

Received: July 18, 2026

Accepted: August 20, 2026

Published: September 15, 2026

Cite as: Kürker, F., & Taşaltın, R. (2026). Data-driven validation and correction of the distortion factor estimated from current total harmonic distortion using industrial field measurements. *Black Sea Journal of Engineering and Science*, 9(5), 2639-2651.

1. Introduction

The reliable and efficient use of electrical energy depends on voltage and current waveforms having acceptable power quality characteristics. However, variable-speed drives, rectifiers, uninterruptible power supplies, switched-mode power supplies, arc furnaces, and other power electronics-based equipment generate harmonic components due to their nonlinear current-voltage behavior. When harmonic currents flow through the grid impedance, they can cause distortion of the voltage waveform and affect other loads connected to the same electrical system. The level of harmonic distortion depends not only on the type of load but also on the load factor, the topology of the power electronics converter, existing distortion in the supply voltage, system impedance, and the simultaneous operating conditions of

loads connected to the same busbar (Kürker, 2017; Senra et al., 2017; Michalec et al., 2021; Hu et al., 2023).

Harmonic distortion can lead to technical consequences such as a decrease in power factor, increased conductor and transformer losses, thermal stress, an increase in neutral current, the risk of resonance, and impaired performance of measurement or protection systems. Therefore, the evaluation of harmonics is not merely a waveform analysis but also an engineering problem aimed at assessing system capacity and energy usage efficiency. IEEE Std 519-2022, published by the Institute of Electrical and Electronics Engineers, defines limits and design objectives for controlling steady-state voltage and current distortions at the common connection point (IEEE, 2022). IEC 61000-4-7, published by the International Electrotechnical Commission, explains the



principles of measurement categorization using harmonic and interharmonic measurement devices (IEC, 2009), IEC 61000-4-30 describes the methods and measurement classes for conducting field power quality measurements in a repeatable manner (IEC, 2025). Studies conducted on industrial systems also demonstrate that harmonic currents can interact with system impedance to cause voltage distortion and that harmonic evaluation depends on the measurement point (Obulesu et al., 2014; Ugwuagbo et al., 2021).

The interpretation of the power factor under harmonic conditions is more comprehensive than the classical $\cos\phi$ approach used in sinusoidal systems. While the displacement power factor represents the phase difference between the voltage and current at the fundamental frequency, the total power factor also accounts for the effect of harmonic components on apparent power. IEEE Std 1459-2025 provides an expanded framework for defining power quantities in non-sinusoidal, unbalanced, and polyphase systems (IEEE, 2025). The physical meaning and definition of power components in systems with harmonics have long been debated; it has been shown that a single definition of reactive or distortion power may be insufficient to account for all non-sinusoidal conditions (Czarnecki, 1987; Ferrero and Superti-Furga, 1991; Emanuel, 2004; Willems, 2004). Therefore, the total power factor, displacement power factor, and harmonic distortion indices should be considered complementary quantities representing different electrical effects, rather than interchangeable ones.

In a simplified scenario where the supply voltage is ideally sinusoidal and non-fundamental components are primarily present in the load current, the current distortion factor can be expressed using the $\frac{1}{\sqrt{1+THDI^2}}$ relationship. This approach allows for a quick and easily interpretable estimate of the distortion factor based on the THDI value. However, in real industrial systems, voltage distortion, phase asymmetry, light-load conditions, time-dependent operating points of loads, and voltage-current interactions outside the fundamental frequency can undermine the assumptions of this simple relationship (IEEE, 2025; Emanuel, 2004; Willems, 2004; Senra et al., 2017). In particular, it has been experimentally demonstrated that the harmonic currents drawn by a nonlinear load under a distorted supply voltage can vary depending on the amplitude and phase of the voltage harmonics (Senra et al., 2017). Therefore, the extent to which the theoretical distortion factor calculated using only THDI represents the PF/dPF ratio measured under field conditions must be verified with direct measurement data.

The presence of multiple nonlinear loads in the same system further complicates the formula-measurement relationship. Harmonic components may not simply add up arithmetically across loads; depending on the phase angles of the components, partial addition or cancellation

may occur. Measurements taken with different load combinations in a low-voltage grid have shown that the THDI value of the total load current may be lower than the THDI values of individual devices (Michalec et al., 2021). Similarly, the harmonic voltage observed at industrial common connection points may consist of the phasor sum of grid-background harmonics and consumer-generated harmonics (Ugwuagbo et al., 2021; Liu et al., 2023). These findings demonstrate that the relationship between THDI and power factor depends not only on the magnitude of the harmonics but also on the electrical context of the system.

Although THDI, THDU, total power factor, and displacement power factor are widely used in power quality studies, these metrics can yield different or seemingly contradictory assessments for certain loads. For this reason, new metrics that more comprehensively describe the effect of harmonics on power and losses have been proposed (Coelho et al., 2024). In the analysis of measurement data, artificial neural networks, Random Forest, and other data-driven methods have been used for monitoring harmonic distortion and for prediction problems related to electrical energy (Nascimento et al., 2013; Nur et al., 2025). However, high-capacity models may lead to overly optimistic performance estimates and the risk of overfitting in field studies with a limited number of independent measurement units. In such datasets, models with lower complexity that are interpretable and evaluated using an appropriate cross-validation scheme offer a more defensible approach (Hoerl and Kennard, 1970; Roberts et al., 2017; Vabalas et al., 2019).

In particular, when repeated observations from the same facility, measurement point, or piece of equipment are randomly assigned to training and test sets as if they were independent, the model may detect characteristics specific to that group during both the training and evaluation phases. This situation can lead to an underestimation of the actual generalization error. For this reason, it is important to construct cross-validation sets based on independent units in hierarchical or grouped data (Roberts et al., 2017). Furthermore, the relationship between a new computational approach and reference values should be examined not only using correlation or aggregated error measures but also in conjunction with mean bias and the 95% limits of agreement (Bland and Altman, 1986). Small-sample uncertainty can be assessed using the bootstrap resampling method (Efron, 1979), while differences in error at the matched field level can be evaluated using the Wilcoxon signed-rank test, which does not require the assumption of a parametric distribution (Wilcoxon, 1945).

The current literature contains comprehensive studies on the modeling of harmonic sources, the separation of harmonic contributions, filter design, the development of new power quality indicators, and data-driven estimation of harmonics (Nascimento et al., 2013; Obulesu et al.,

2014; Ugwuagbo et al., 2021; Hu et al., 2023; Liu et al., 2023; Coelho et al., 2024). However, within the scope of the reviewed studies, there are few that directly compare the measured distortion factor—based on the PF/dPF ratio from field measurements obtained at different industrial loads and connection points—with the theoretical factor calculated solely from THDI; correct the residual error of the theoretical formula using low-complexity models; and simultaneously address group structure, limits of agreement, and small-sample uncertainty. This gap necessitates an investigation not only of the average performance of the theoretical relationship under field conditions but also of the measurement points and operating conditions where it produces significant deviations.

The objective of this study is to evaluate the validity of the THDI-based theoretical distortion factor using three-phase field measurements from 12 industrial measurement points and to investigate whether the observed formula deviation can be reduced using low-complexity, data-driven models. The measurements used were obtained as part of a previously completed doctoral dissertation, reorganized for the present study, and analyzed after verifying the integrity of the calculations (Kürker, 2017). The main contributions of the study are as follows: (i) verification of the consistency between reported and recalculated PF/dPF ratios and THDI-based theoretical distortion factors; (ii) comparison of linear regression, robust linear regression, Ridge regression, restricted second-order regression, and small regression trees under out-of-sample cross-validation; (iii) evaluation of model performance using a combination of mean and maximum error, bias, Bland–Altman 95% limits of agreement, the Wilcoxon test at the measurement point level, and bootstrap confidence intervals; and (iv) examination of formula errors in the context of phase, load level, total harmonic distortion of voltage (THDU), and phase asymmetry. Within this framework, the study aims not to propose a universal corrected THDI equation, but rather to highlight the strengths and limitations of the theoretical approach in industrial field data and to cautiously evaluate the potential benefits that data-driven correction could provide.

2. Materials and Methods

2.1. Research Design and Data Source

This study is based on a retrospective and secondary analysis of three-phase power quality measurements previously conducted at various industrial loads and connection points. The field data used were obtained as part of the doctoral dissertation titled “Harmonic Analysis and Harmonic Mitigation Techniques in Power Systems,” completed in 2017 at the Institute of Natural Sciences at Harran University (Kürker, 2017). No new field measurements were conducted in the present study; the phase-level electrical quantities reported in the source measurements were reorganized, checked for

computational consistency, and reanalyzed to evaluate the THDI-based distortion factor.

The dataset consists of 12 three-phase measurement points representing different processes and electrical connection conditions. Since the phases at the same measurement point share common electrical and operational conditions, they were not considered independent observations; the measurement point was defined as the independent unit of analysis for model validation and statistical comparisons. For this reason, low-complexity and interpretable correction models were used instead of high-capacity models.

2.2. Measurement System and Characteristics of Source Measurements

Field measurements were performed using the Amprobe Energy Test 2020 power and harmonic analyzer. The thesis states that the device can measure current and voltage harmonics, active power, reactive power, apparent power, total power factor, and displacement power factor in three-phase systems. The device also has the capability to record the measured electrical quantities and transfer them to a computer.

In this study, the root-mean-square (RMS) value of phase current (Irms), total power factor (PF), displacement power factor (dPF), total current harmonic distortion (THDI), and total voltage harmonic distortion (THDU) were used. In the source thesis, dPF is defined as the power factor related to the phase difference between the current and voltage at the fundamental frequency, while PF represents the total power factor under harmonic operating conditions.

In the source thesis, the device’s calibration date, measurement uncertainty, IEC measurement class, and harmonic binning window were not explicitly reported. Therefore, no additional assumptions regarding the device’s measurement class or absolute metrological accuracy were made in the present study.

2.3. Data Preparation and Calculation of Distortion Factors

The analysis data table consists of measurement point ID, measurement location, phase, Irms, PF, dPF, THDI, THDU, reported distortion factors, and recalculated control variables. The data columns and variable names used in the code are clearly defined at the beginning of the analysis.

The measurement-based distortion factor was calculated as the ratio of the total power factor (PF) to the displacement power factor (dPF) (equation 1):

$$DF_{measured} = \frac{PF}{dPF} \quad (1)$$

The THDI-based theoretical distortion factor is determined by the following equation 2:

$$DF_{THDI} = \frac{1}{\sqrt{1 + \left(\frac{THDI}{100}\right)^2}} \quad (2)$$

The deviation from the theoretical formula at the phase level is defined as the difference between the measured and theoretical distortion factors (equation 3):

$$\Delta DF = DF_{measured} - DF_{THDI} \quad (3)$$

The absolute formula error is defined in the text as $AE = |\Delta DF|$. The relative error, on the other hand, is calculated by expressing the ratio of the ΔDF value to the measured distortion factor as a percentage.

The PF/dPF and theoretical distortion factor values

reported in the source tables were compared with the values recalculated using equations 1 and 2. In subsequent modeling and error analyses, the recalculated values were used instead of the reported rounded values. The basic and derived variables used in the analysis are summarized in Table 1.

Table 1. Basic and derived variables used in the analysis

Variable	Description	Unit or Calculation	Role in the Analysis
I_{rms}	RMS phase current	A	Load-level indicator
PF	Total power factor	Dimensionless	Measured DF calculation
dPF	Displacement power factor	Dimensionless	Measured DF calculation
THDI	Current total harmonic distortion	%	Theoretical DF calculation
THDU	Voltage total harmonic distortion	%	Correction model input
$DF_{measured}$	Measurement-based distortion factor	PF/dPF	Reference output
DF_{THDI}	THDI-based theoretical distortion factor	Equation (2)	Baseline model
ΔDF	Formula deviation	$DF_{measured} - DF_{THDI}$	Regression target
AE	Absolute formula error	$ \Delta DF $	Absolute error and exploratory classification metric
$\log(I_{rms})$	Logarithmic current variable	$\ln(1 + I_{rms})$	Correction model input

2.4. Exploratory Data Analysis and Outlier Screening

Prior to modeling, the minimum, maximum, mean, median, standard deviation, variance, range, interquartile range, coefficient of variation, skewness, and kurtosis values were calculated for I_{rms} , PF, dPF, THDI, THDU, the measured distortion factor, the theoretical distortion factor, and ΔDF .

Variable distributions were explored using histograms, box plots, and violin plots; bivariate relationships between variables were examined using a correlation matrix and scatter plots. Theoretical normal distribution curves added to the histograms were used solely for visual identification purposes, and no formal conclusion regarding the normality of the distributions was drawn from these plots.

Potential outliers were screened using Tukey's 1.5×IQR method. The lower limit was set at $Q1 - 1.5IQR$, and the upper limit at $Q3 + 1.5IQR$. Observations falling outside these limits were flagged but were not removed from the dataset solely because they were identified as outliers. The flagged values were evaluated in conjunction with the measurement point, phase, load level, and phase asymmetry.

2.5. Correction Objective and Selection of Explanatory Variables

The objective of data-driven modeling was defined as estimating the deviation from the THDI-based theoretical formula, rather than directly re-estimating the measured distortion factor. The corrected distortion factor was calculated as follows (equation 4):

$$DF_{corrected} = DF_{THDI} + \widehat{\Delta DF} \quad (4)$$

Here, $\widehat{\Delta DF}$ represents the formula correction estimated

by the data-driven model.

Because current values vary over a wide range, I_{rms} has been subjected to the $\ln(1+I_{rms})$ transformation. THDU and $\ln(1+I_{rms})$ were used as explanatory variables in the main correction models. Since THDI is used directly in the calculation of the theoretical distortion factor, it was not included in the main model inputs to limit the reweighting of the same information in the correction layer.

Multicollinearity among the candidate variables was assessed using the correlation matrix and the variance inflation factor (VIF). VIF values were calculated using the R^2 values obtained from the linear models established for each variable with the other candidate variables.

2.6. Data-Driven Correction Models

Since the independent analysis unit was limited to 12 measurement points, high-capacity machine learning and deep learning models were not used. Instead, five low-complexity models that were more suitable for the sample structure and more interpretable were compared:

- . Ordinary least squares-based linear regression,
- . robust linear regression,
- . L2-regularized Ridge regression,
- . restricted second-order regression including the $THDU^2$ term,
- . a small regression tree consisting of leaves with at most three splits and at least five observations.

In the linear, robust, and Ridge models, the explanatory variables used were THDU and $\ln(1+I_{rms})$. In the restricted second-order regression, the $THDU^2$ term was added to the predictors; no interaction terms or

additional higher-order terms were included. In the regression tree, the settings MaxNumSplits=3 and MinLeafSize=5 were applied.

Model complexities were pre-constrained, and no separate hyperparameter optimization was performed. In Ridge regression, the regularization type was set to L2; however, the regularization coefficient was not explicitly fixed in the current code.

2.7. Out-of-Measurement-Point Cross-Validation

The out-of-measurement-point cross-validation method was used to prevent phases belonging to the same measurement point from being distributed together across the training and test sets. In each validation cycle, three phases belonging to a single measurement point were assigned to the test set, while the remaining 33 phases from the other 11 measurement points were used as the training set. The process was repeated 12 times, ensuring that each measurement point appeared in the test set exactly once.

In each cycle, the model was trained using only the ΔDF values from the training measurement points, and the predicted error for each phase at the test measurement point was added to the corresponding DF_{THDI} value. Thus, each corrected estimate was obtained from a validation cycle in which the relevant measurement point was not included in the model training.

2.8. Evaluation of Model Performance

The theoretical baseline formula and data-driven models were compared based on LOSO prediction errors. The prediction error for a given observation was defined as $e_i = DF_{measured,i} - \widehat{DF}_i$.

The mean absolute error was calculated as follows (equation 5):

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (5)$$

Root mean square error (equation 6):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (6)$$

and mean bias (equation 7):

$$Bias = \frac{1}{n} \sum_{i=1}^n e_i \quad (7)$$

have been determined.

In addition, the maximum absolute error and the median absolute error have been reported. The improvement percentages for MAE and RMSE were calculated by dividing the reduction in error between the model and the theoretical baseline formula by the error of the theoretical baseline formula. Positive improvement values indicate a lower error compared to the theoretical formula.

2.9. Statistical Comparison and Uncertainty Analysis

Although the restricted second-order regression produced the lowest aggregate MAE and RMSE values, Ridge regression was selected as the focal model for the subsequent agreement, statistical uncertainty, and coefficient-stability analyses. This choice was based on its additive linear structure, L2-regularized formulation,

and lower functional complexity compared with the restricted second-order regression, which additionally includes a squared THDU term. Given the limited number of independent measurement points, Ridge regression was therefore treated as a parsimonious and cautious reference model for evaluating the behavior of a regularized correction. This selection does not imply that Ridge regression was the best-performing model according to the aggregate prediction-error metrics.

Since the phase records are not independent within the same measurement point, the statistical comparison between the theoretical baseline formula and the Ridge model was performed at the measurement point level. The mean absolute error (MAE) for the three phases was calculated for each measurement point, and the 12 paired measurement-point-level MAE values of the theoretical formula and the Ridge model were compared using the Wilcoxon signed-rank test.

The uncertainty in the mean measurement-point-level MAE difference was evaluated using a nonparametric bootstrap analysis with 5,000 resamples. First, the MAE values of the theoretical formula and the Ridge model were calculated separately for each of the 12 measurement points using the three corresponding phase-level errors. A paired MAE difference was then defined for each measurement point as $MAE_{theoretical} - MAE_{Ridge}$, such that a positive value indicated a lower error for the Ridge model. The 12 paired measurement-point-level MAE differences were resampled with replacement. In each bootstrap iteration, 12 paired differences were drawn with replacement and their arithmetic mean was calculated. The 95% percentile confidence interval was obtained from the 2.5th and 97.5th percentiles of the bootstrap distribution. For reproducibility, the random number seed was fixed at 42. The stability of the Ridge coefficients was examined based on the mean, standard deviation, and sign consistency of the THDU and logarithmic current coefficients obtained over 12 LOSO cycles.

2.10. Bland-Altman and Residual Analyses

The agreement between the theoretical formula and the Ridge-corrected values was evaluated using Bland-Altman analysis. The mean difference was taken as the bias, and the approximate 95% limits of agreement (LoA) were calculated as follows (equation 8):

$$LoA = Bias \pm 1.96SD(e) \quad (8)$$

Bland-Altman plots were used to assess descriptive agreement at the phase level; statistical inferences were based on Wilcoxon and bootstrap analyses at the measurement point level.

The structure of the Ridge residuals was visually assessed using residual-THDU, residual- $\ln(1+I_{rms})$, and normal probability plots. No formal hypothesis test was conducted to assess residual normality or constant variance.

2.11. Exploratory Agreement Classes

To describe the behavior of the THDI-based theoretical

formula at different error levels, phase records were divided into three exploratory classes based on absolute error (equation 9):

$$\text{Agreement class} = \begin{cases} \text{High agreement,} & AE < 0.01, \\ \text{Moderate agreement,} & 0.01 \leq AE < 0.03, \\ \text{Low agreement,} & AE \geq 0.03. \end{cases} \quad (9)$$

These thresholds were not derived from any standard but were determined solely for the purpose of exploratory evaluation of the existing dataset. To examine threshold sensitivity, the baseline thresholds were tightened and relaxed by 50%, and the resulting class frequencies were compared.

The relationship between the exploratory agreement classes and THDI, THDU, and logarithmic current was explored using a small classification tree with a maximum of three splits and leaves containing at least five observations. This model was not evaluated as a generalizable classifier.

2.12. Anomaly and Measurement Context Indicators

In data quality control, the conditions $PF/dPF > 1$, missing or zero THDU, and $AE \geq 0.03$ were used as rule-based anomaly indicators. The phase-level anomaly score was calculated as the number of conditions met.

Phase current imbalance at the measurement point level was determined as follows (equation 10):

$$I_{\text{imbalance}} = \frac{\max(I_{rms}) - \min(I_{rms})}{I_{rms}} \quad (10)$$

In addition, the power factor variation between phases was calculated as $PF_{\text{range}} = \max(PF) - \min(PF)$. Records with phase current below 10 A were flagged as low load, while records with a PF range above 0.30 within the measurement point were flagged for phase asymmetry. These thresholds are not standard compliance criteria but rather screening thresholds specific to the current dataset. No observations were excluded from the analysis due to these flags.

2.13. PCA and Hierarchical Clustering

Rather than classifying the general load types of the measurement points, complementary principal component analysis and hierarchical clustering were applied solely to examine their similarities in terms of the error behavior of the THDI-based formula.

For each measurement point, the mean absolute error, mean THDI, mean THDU, and standard deviation of ΔDF were calculated for the three phases. This resulted in a field-level data matrix consisting of 12 measurement points and four features. To minimize the effect of variables being on different scales on the PCA and distance calculations, the features were standardized using z-score transformation.

PCA was performed on the standardized 12×4 data matrix, and the measurement points were plotted in the score plane of the first two principal components. Hierarchical clustering was performed using the same standardized feature matrix, Euclidean distance, and Ward's linkage method. No predefined number of clusters or dendrogram cutoff distance was applied.

Therefore, the results were interpreted as exploratory findings illustrating relative similarities among measurement points; no fixed or generalizable industrial load classes were derived.

2.14. Analysis Workflow and Software

Figure 1 summarizes the complete analysis workflow, including the data source and anonymization, recalculation of the distortion factors, data-integrity checks, exploratory analysis, LOSO validation, model comparison, statistical uncertainty analysis, agreement and residual diagnostics, exploratory agreement-class analysis, and the PCA and hierarchical clustering stages.

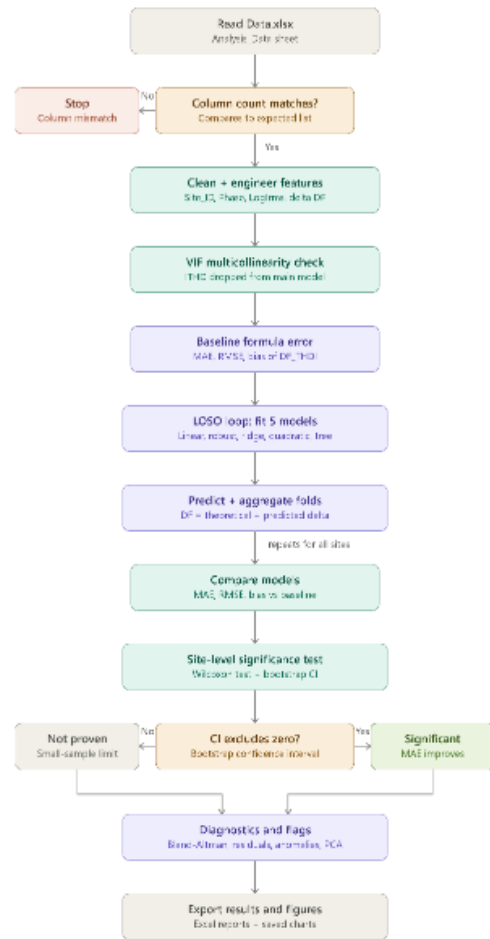


Figure 1. Analysis workflow for the validation of the THDI-based distortion factor using industrial field measurements and its data-driven correction.

3. Results

3.1. Data integrity and the fundamental performance of the THDI-based theoretical distortion factor

The dataset for this study consists of three-phase field measurements from 12 industrial measurement points. In the dataset, the total harmonic distortion of current (THDI) ranged from 11.946% to 64.618%, the total harmonic distortion of voltage (THDU) ranged from 0.953% to 8.870%, and the phase currents ranged from 0.886 to 290.760 A. This variability allowed for the evaluation of the THDI-based theoretical distortion factor

under different load levels and field operating conditions. First, the reported distortion factors were compared with the recalculated values. The highest absolute difference between the values calculated from the PF/dPF ratio for the measured distortion factor and the reported values was 0.000479; for the THDI-based theoretical distortion factor, it was determined to be 0.000459. Therefore, it was concluded that the minor differences between the reported and recalculated values were primarily due to rounding. Specifically, for the third phase S12, the measured distortion factor calculated from the values $PF = 0.179$ and $dPF = 0.1811$ is 0.988404. Therefore, the high relative error value observed earlier stems not from a calculation error but from the actual difference between the measured and THDI-based theoretical distortion factors.

The mean absolute error, root mean square error, and mean bias of the THDI-based theoretical formula across the entire dataset were calculated as 0.015662, 0.019681, and -0.009357 , respectively. For the majority of phase-level measurements, the measured distortion factor remained below the theoretical value. The negative mean bias indicates that the theoretical formula tends to overestimate the distortion factor on average in the analyzed dataset. However, the error distribution is not entirely one-sided; error values ranged from -0.038720 to $+0.055927$. This result shows that, despite a systematic tendency to overestimate the distortion factor on average, the theoretical formula can also

underestimate the measured distortion factor under certain load and phase conditions.

In the multicollinearity analysis conducted for variables that could be used in the correction model, the variance inflation factors for the THDI, THDU, and $\log(I_{rms})$ variables were determined to be 1.15, 1.09, and 1.23, respectively. These values indicate that there is no significant multicollinearity issue among the candidate variables. However, since THDI is already used directly in the calculation of the theoretical distortion factor, it was not included in the set of independent variables to prevent the same information from being reweighted in the correction model and to ensure that the model focuses on the contribution of additional explanatory variables.

3.2. LOSO Performance of Data-Driven Correction Models

The generalization performance of the models was evaluated using a measurement-point-excluded cross-validation method, which prevents the three phases associated with the same measurement point from being distributed across the training and test sets. In this approach, the three phases associated with a single measurement point were assigned to the test set together in each iteration, and a total of 12 validation cycles were performed. The performance of the theoretical baseline formula and the linear, robust, Ridge, restricted second-order regression, and decision tree models is compared in Table 2.

Table 2. Comparison of LOSO performance between the THDI-based theoretical formula and data-driven correction models

Model	MAE	RMSE	Bias	Maximum Absolute Error	Median Absolute Error	MAE Improvement (%)	RMSE Improvement (%)
Theoretical Baseline Formula	0.01566	0.019681	-0.00936	0.055927	0.011937	0.00	0.00
Linear	0.01229	0.018678	0.000037	0.066484	0.009878	21.50	5.09
Robust	0.01131	0.018435	0.002029	0.068745	0.006487	27.79	6.33
Ridge	0.01226	0.018660	0.000057	0.066482	0.009772	21.70	5.19
Restricted Second-Order Regression	0.01068	0.016590	-0.00046	0.068687	0.007089	31.83	15.70
Decision Tree	0.01413	0.018824	-0.00067	0.067629	0.010357	9.79	4.35

The restricted second-order regression yielded the best results in terms of average performance metrics, with an MAE of 0.010677 and an RMSE of 0.016590. Relative to the theoretical baseline formula, these values correspond to improvements of 31.83% and 15.70%, respectively. The robust model produced the lowest median absolute error, at 0.006487. However, a convergence warning indicating that the iteration limit had been reached was generated during robust regression fitting. Therefore, the robust regression results were retained only as a complementary comparison and should be interpreted

with caution. No claim of model superiority is made on the basis of these results, because convergence was not fully established under the applied fitting settings.

Ridge regression, considered the main regularization model in this study, reduced the MAE value of the theoretical formula from 0.015662 to 0.012263, achieving a 21.70% improvement. The improvement in RMSE was more limited, remaining at 5.19%. The most pronounced effect of the Ridge model was observed on the mean bias. The bias, which was -0.009357 in the theoretical formula, approached 0.000057 after Ridge

correction. This result indicates that the model largely eliminated the systematic tendency toward high overestimation present in the theoretical formula.

Although the restricted second-order regression remained the best-performing model in terms of aggregate MAE and RMSE, Ridge regression was retained for the detailed diagnostic analyses because it provides a regularized linear correction without introducing the additional nonlinear THDU² term. Thus, the subsequent focus on Ridge regression reflects a preference for a more parsimonious and interpretable reference model under the limited independent-sample structure, rather than a claim that it achieved the lowest overall prediction error.

However, none of the data-driven models were able to reduce the maximum absolute error. The maximum absolute error, which was 0.055927 in the theoretical formula, increased to 0.066482 in the Ridge model and to 0.068687 in the restricted second-order regression. Thus, data-driven corrections reduced the mean error but produced larger estimation errors in some outlier observations. This finding indicates that model evaluation should not be based solely on aggregated metrics such as MAE or RMSE.

The relationship between the measured distortion factors and the estimates produced by the theoretical baseline formula and the Ridge model is shown in Figure 2. The dashed 1:1 line represents equality between the measured and predicted distortion factors. The Ridge correction moved a substantial proportion of the observations closer to the 1:1 line; however, noticeable deviations persisted, particularly for the extreme phase-level observations.

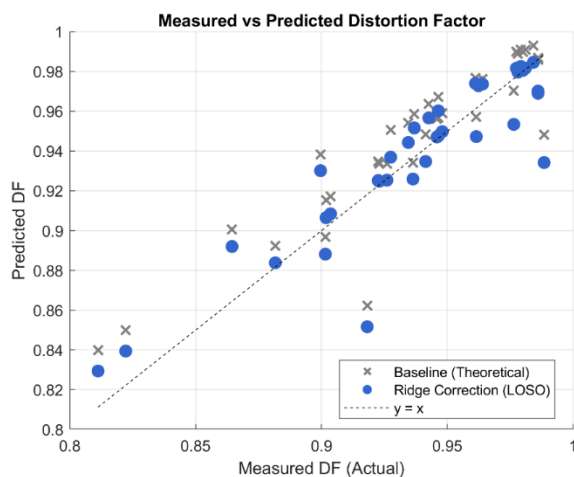


Figure 2. Comparison of the measured distortion factor with the estimates from the THDI-based theoretical baseline formula and the LOSO Ridge correction. The dashed line represents the 1:1 line.

3.3. Bias, 95% Limits of Agreement, and Statistical Uncertainty

The agreement between the theoretical baseline formula and the Ridge-corrected values was further examined

using the Bland-Altman plots shown in Figure 3. For the theoretical formula, the mean difference between the measured and theoretical distortion factors was -0.009357 , and the approximate 95% limits of agreement ranged from -0.043774 to 0.025059 . For the Ridge model, the mean difference was 0.000057 , while the approximate 95% limits of agreement ranged from -0.037035 to 0.037148 .

These results indicate that Ridge regularization nearly eliminated the mean bias but did not reduce the dispersion of the differences. The total width of the 95% limits of agreement was approximately 0.06883 for the theoretical formula and 0.07418 for the Ridge model. Therefore, the primary benefit of the Ridge correction was to center the differences closer to zero rather than to narrow their variability at the phase or measurement-point level.

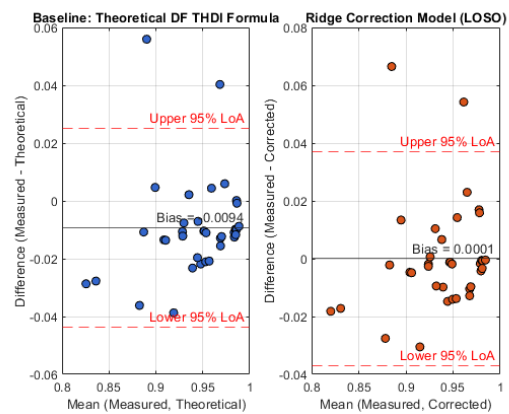


Figure 3. Bland-Altman plots showing the bias and approximate 95% limits of agreement for the THDI-based theoretical formula and the LOSO Ridge correction.

In the Wilcoxon signed-rank test based on the 12 paired measurement-point-level MAE values, the difference between the theoretical formula and the Ridge model yielded a p-value of 0.0342. This rank-based paired result indicates a systematic tendency toward lower measurement-point-level errors after Ridge correction. In the bootstrap analysis based on 5,000 resamples of the 12 paired measurement-point-level MAE differences, the estimated mean difference, defined as $MAE_{theoretical} - MAE_{Ridge}$, was 0.00339, with a 95% percentile confidence interval ranging from -0.00093 to 0.00657 . Because this interval included zero, a positive improvement in the arithmetic mean could not be established with sufficient precision using the current 12 measurement points. These findings are not directly contradictory because the two analyses evaluate different statistical features of the paired differences. The Wilcoxon signed-rank test evaluates a rank-based paired location shift, whereas the bootstrap confidence interval quantifies uncertainty in the arithmetic mean of the paired MAE differences. Taken together, the results provide positive but limited evidence regarding the error-reduction potential of the Ridge correction rather

than definitive statistical confirmation of improvement.

3.4. Error Behavior at the Measurement Point and Phase Levels

In the measurement-point-based analysis, it was determined that the Ridge model reduced the MAE value at 11 out of 12 measurement points. The most significant absolute reductions were observed at S01 (from 0.010188 to 0.000977), S06 (from 0.010643 to 0.001411), S04 (from 0.012582 to 0.003802), and S07 (from 0.011003 to 0.002675). In contrast, while the field-based MAE value of the theoretical formula in S10 was only 0.002295, it rose to 0.018605 in the Ridge model. In S12, the MAE decreased from 0.023535 to 0.023417, showing only a limited change.

The loss of performance in S10 demonstrates that a

general correction model can produce unnecessary corrections at measurement points where the theoretical formula already provides high accuracy. Therefore, rather than applying the correction model unconditionally to every site, it may be more appropriate to use an application or reliability criterion that evaluates the model's expected benefit based on the theoretical formula.

Observations with an absolute distortion factor error above 0.03 are presented in Table 3. It is noteworthy that all four high-error observations occurred in the third phase. However, this pattern, based on only four observations, should not be interpreted as indicating a general systematic error for the third phase.

Table 3. Phase observations with an absolute distortion factor error greater than 0.03

Measurement Point	Phase	THDI (%)	PF	dPF	Measured DF	Theoretical DF	Error	Absolute Relative Error (%)
S05 – Steam Fixation Process	3	58.75	0.83	0.904	0.918142	0.862215	0.055927	6.091
S08 – Water System, Inverter Input	3	48.26	0.338	0.391	0.86445	0.900597	-0.03615	4.181
S09 – Water System, Grid Input	3	36.85	0.233	0.259	0.899614	0.938334	-0.03872	4.304
S12 – Laser Supply 3	3	33.53	0.179	0.1811	0.988404	0.94812	0.040284	4.076

The signs of the high-error observations are not the same. While the distortion factor measured in the third phases of S05 and S12 was higher than the theoretical value, the value measured in the third phases of S08 and S09 was below the theoretical value. This bidirectional error behavior indicates that a constant or unidirectional correction applied to all observations would not be sufficient to explain the extreme cases.

The result for the third phase of S12 also demonstrates the need to interpret the distortion factor and the total power factor together. In this observation, although the PF/dPF ratio is 0.988404, both PF and dPF are approximately 0.18. Therefore, a distortion factor close to 1 does not necessarily imply a high total power factor. In this case, the dominant component of the low total power factor is related to phase shift, and the ratio of PF to dPF alone does not represent the load's overall power factor performance.

Significant PF differences between phases were identified at some of the S08–S12 measurement points. The PF range is 0.541 at S08, 0.693 at S09, 0.792 at S10, 0.565 at S11, and 0.704 at S12. The S03 and S07 observations, meanwhile, have been flagged separately for reliability purposes due to low current levels. These flags should not be interpreted as direct data errors; however, they indicate that the variability of the PF/dPF ratio may increase under single-phase-dominant loads, phase asymmetry, and low-load conditions.

3.5. Residual Analysis and Model Stability

Figure 4 shows the distribution of residuals from the

Ridge model with respect to THDU and logarithmic current level, together with a normal probability plot. Although the model reduced the mean bias, several relatively large residuals remained. Because no formal statistical tests of residual normality or heteroscedasticity were conducted, the plots were interpreted solely as visual diagnostic tools.

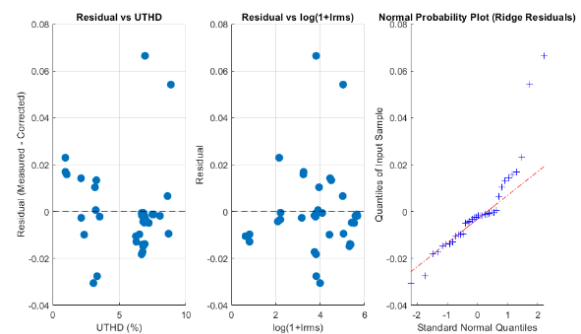


Figure 4. Residual diagnostics and normal probability plot for the LOSO Ridge correction model based on THDU and $\log(I_{rms})$.

When examining the stability of the Ridge coefficients over 12 LOSO cycles, the mean of the THDU coefficient was found to be 0.00007 and the standard deviation 0.00047; the coefficient took on positive values in eight of the 12 cycles. The mean of the $\log(I_{rms})$ coefficient was -0.00020 , with a standard deviation of 0.00057, and the coefficient remained positive in only five cycles. The variation in the signs of the coefficients across validation

cycles indicates that the physical effects of the explanatory variables cannot be reliably determined with the current dataset. This alone does not prove that the model is invalid or overfitted; however, it limits the physical interpretation of the coefficients and points to high model uncertainty. Therefore, the Ridge model should be presented not as a physical equation explaining the causal effect of THDI or the current distortion factor, but as an empirical correction model aimed at reducing the average bias of the theoretical formula based on field data.

4. Discussion

4.1. Exploratory Agreement Classes and Threshold Sensitivity

Using the exploratory absolute-error thresholds of 0.01 and 0.03, 11 observations were assigned to the high-agreement class, 21 to the moderate-agreement class, and four to the low-agreement class. The distribution of these exploratory agreement classes in the THDI-THDU plane is shown in Figure 5. The observations in the low-agreement class were not concentrated within a specific THDI or THDU region but occurred at different distortion levels.

When the thresholds were tightened by 50%, the distribution changed to five high-agreement, 19 moderate-agreement, and 12 low-agreement observations. When the thresholds were relaxed by 50%, the distribution changed to 24 high-agreement, 11 moderate-agreement, and one low-agreement observation. These changes demonstrate that the classification is sensitive to the selected absolute-error thresholds. Therefore, the thresholds of 0.01 and 0.03 should not be interpreted as standard, universal, or compliance limits; they are working thresholds used solely for the exploratory evaluation of the current dataset.

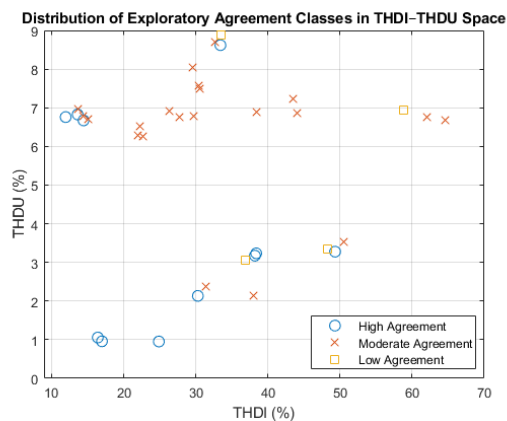


Figure 5. Distribution of the exploratory agreement classes, defined according to the absolute distortion-factor error, in the THDI-THDU plane.

The high-, moderate-, and low-agreement classes were defined using the thresholds $|e| < 0.01$, $0.01 \leq |e| < 0.03$, and $|e| \geq 0.03$, respectively. These thresholds were not

derived from a standard and were used solely for the exploratory evaluation of the current dataset.

4.2. Exploratory Multivariate Analysis of Measurement Points

The similarities among the measurement points in terms of their error profiles relative to the theoretical formula were explored using principal component analysis (PCA) and hierarchical clustering. The distribution of the measurement points in the PCA score plot is shown in Figure 6, and the hierarchical clustering dendrogram is presented in Figure 7.

The first principal component of PCA explained 64.5% of the total variance, while the second principal component explained 27.0%. Thus, the first two principal components together accounted for 91.5% of the total variance. In Figure 6, it can be seen that the measurement points S01, S02, S03, S04, S06, and S07 are located in the upper-left quadrant of the PCA plane and are relatively close to one another. S08, S09, and S11, on the other hand, are located in the lower region where the second principal component is negative. S05 exhibits a distribution that is more distinct from the other measurement points in the region where the first principal component has a high positive value; S10 is in the region where both the first and second principal components are negative; and S12 is in the region where both principal components are positive. These findings indicate that the error behavior of the theoretical formula is not entirely homogeneous across the measurement points and that some points exhibit distinct multivariate error profiles.

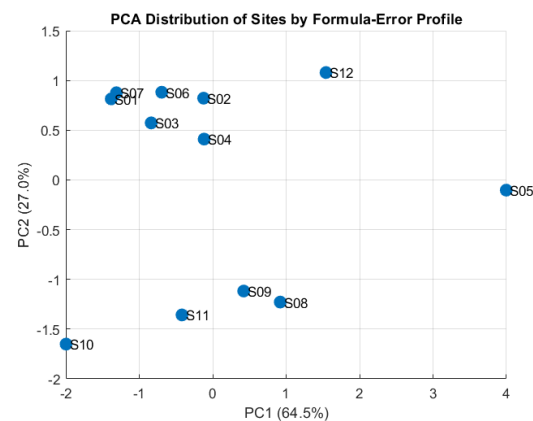


Figure 6. Distribution of principal component analysis scores based on the error profiles of the theoretical formula across the measurement points. The first and second principal components explain 64.5% and 27.0% of the total variance, respectively, and together account for 91.5% of the total variance.

The results of hierarchical clustering showed that there were varying degrees of similarity among the measurement points. In Figure 7, S01 and S07 are the measurement points that matched at the lowest clustering distance. The pairs S04 and S06, and S08 and S09, also clustered together at relatively low clustering

distances. The S01–S07 pair subsequently merged with S03, while the S04–S06 pair formed a group with S02 at a higher hierarchical level. Similarly, the S08–S09 pair first merged with S11 and then with S10. S05 and S12, on the other hand, formed a branch among themselves but joined the main structure composed of the other measurement points at a higher clustering distance. However, since the analysis was based on only 12 measurement points and no predefined cluster cutoff height was used, a fixed number of generalizable load classes was not derived from the dendrogram. Therefore, the PCA and hierarchical clustering results should be evaluated not as independent evidence of classification, but as complementary exploratory analyses that visualize the similarities and differences in error profiles among the measurement points. The resulting similarities should not be interpreted as universal clusters or causal classes representing specific industrial load types.

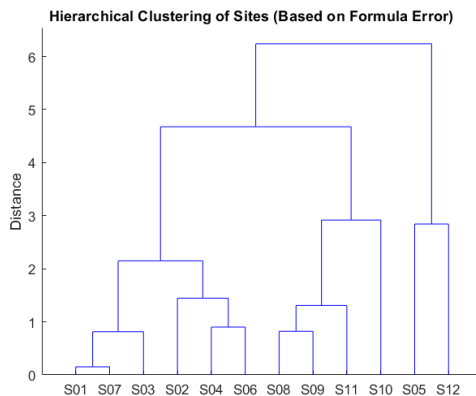


Figure 7. A hierarchical clustering dendrogram based on the theoretical formula error profiles of the measurement points. The merging distances in the dendrogram indicate the relative differences among the multivariate formula-error profiles of the measurement points. Since no predefined cluster cutoff height was used, a fixed number of clusters were not defined.

4.3. General Evaluation

The results show that the theoretical distortion factor calculated using THDI alone provides a useful initial approximation for industrial field data, but does not represent the measured PF/dPF ratio with the same accuracy under all load and phase conditions. The theoretical formula overestimated the distortion factor by an average of 0.009357 and produced higher deviations, particularly under conditions of phase asymmetry, low power factor, or unusual load conditions. Data-driven models reduced the aggregate MAE and RMSE values. While the restricted second-order regression produced the lowest average error, the Ridge model brought the systematic bias of the theoretical formula close to zero. At the measurement-point level, the Wilcoxon signed-rank test indicated a systematic tendency toward lower paired errors, whereas the bootstrap confidence interval for the mean paired MAE

difference included zero. These results are not directly contradictory because the Wilcoxon test evaluates a rank-based paired location shift, while the bootstrap interval quantifies uncertainty in the arithmetic mean difference. Their combined interpretation provides positive but limited evidence regarding the error-reduction potential of the Ridge correction. Together with the failure to reduce the maximum absolute error, the absence of narrower 95% limits of agreement, and the limited coefficient stability, these findings warrant caution regarding the generalizability of the correction. Therefore, the resulting Ridge model should not be defined as a universal “corrected THDI formula,” but rather as a field-calibrated correction approach that reduces the average formula deviation observed at the 12 measurement points. For a stronger generalization, external validation using independent data obtained from different facilities, different loading levels, and repeated operating conditions is required.

5. Conclusion

In this study, the performance of the THDI-based theoretical distortion factor under field conditions was evaluated using three-phase field measurements from 12 industrial measurement points, and low-complexity, data-driven models aimed at reducing the error of the theoretical formula were compared. Since the phases at the same measurement point were not considered independent observations, the models were evaluated using a 12-cycle cross-validation method, in which one measurement point was set aside for testing in each cycle.

The theoretical formula produced a mean absolute error (MAE) of 0.015662, a root mean square error (RMSE) of 0.019681, and an average bias of -0.009357 . The negative bias indicated that the formula tended to overestimate the distortion factor on average in the analyzed dataset. Among the compared models, the restricted second-order regression provided the lowest MAE and RMSE values; the Ridge model reduced the MAE to 0.012263, bringing the mean bias in the theoretical formula close to zero. However, the fact that data-driven models could not reduce the maximum absolute error showed that the improvement in average performance was not valid for all phases and measurement points.

At the measurement-point level, the Wilcoxon signed-rank test indicated a systematic tendency toward lower paired MAE values after Ridge correction ($P=0.0342$). However, the 95% percentile bootstrap confidence interval for the mean paired MAE difference included zero. Because the Wilcoxon test evaluates a rank-based paired location shift, whereas the bootstrap interval quantifies uncertainty in the arithmetic mean difference, these findings are not directly contradictory. Taken together, they provide positive but limited evidence regarding the error-reduction potential of the Ridge correction rather than definitive statistical confirmation of improvement. Therefore, the Ridge model should be

regarded as an empirical calibration approach for the available field data rather than as a universally validated corrected THDI formula.

The fact that high absolute errors are not concentrated solely at a specific THDI or THDU level, and that the error signs point in different directions, indicates that the formula deviation cannot be explained by a single constant correction. In particular, wide PF differences between phases, low power factor, and low-load conditions are measurement contexts that must be considered when interpreting the theoretical formula.

In conclusion, the THDI-based theoretical distortion factor provides a quick and interpretable initial estimate in field applications; however, it does not represent all load and phase conditions on its own. For a more reliable assessment, the theoretical factor must be interpreted in conjunction with PF, dPF, THDU, phase currents, and phase asymmetry. Future studies should include external validation using independent measurements obtained from different facilities, load levels, and time intervals.

Author Contributions

The percentages of the authors' contributions are presented below. Both authors reviewed and approved the final version of the manuscript.

	F.K.	R.T.
C	60	40
D	60	40
S	20	80
DCP	100	0
DAI	60	40
L	100	0
W	100	0
CR	40	60
SR	80	20
PM	70	30
FA	100	0

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The authors declare that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because the research did not involve human participants, human biological materials, animals, or animal experiments. The study was based on the analysis of previously collected industrial electrical measurement data.

Acknowledgements

The authors would like to thank Aydemir Kalebek and Kayhan Kalebek for their technical support and contributions during the course of this study.

References

- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 1(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Coelho, R. A., Araújo, B. V. S., Xavier, G. V. R., Rodrigues, G. A., Ferreira, T. V., & Brito, N. S. D. (2024). Novel indices for power quality assessment of non-linear loads. *Electric Power Systems Research*, 236, Article 110952. <https://doi.org/10.1016/j.epsr.2024.110952>
- Czarnecki, L. S. (1987). What is wrong with the Budeanu concept of reactive and distortion power and why it should be abandoned. *IEEE Transactions on Instrumentation and Measurement*, 36(3), 834–837. <https://doi.org/10.1109/TIM.1987.6312797>
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26. <https://doi.org/10.1214/aos/1176344552>
- Emanuel, A. E. (2004). Summary of IEEE Standard 1459: Definitions for the measurement of electric power quantities under sinusoidal, nonsinusoidal, balanced, or unbalanced conditions. *IEEE Transactions on Industry Applications*, 40(3), 869–876. <https://doi.org/10.1109/TIA.2004.827452>
- Ferrero, A., & Superti-Furga, G. (1991). A new approach to the definition of power components in three-phase systems under nonsinusoidal conditions. *IEEE Transactions on Instrumentation and Measurement*, 40(3), 568–577. <https://doi.org/10.1109/19.87021>
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
- Hu, Z., Han, Y., Zalhaf, A. S., Zhou, S., Zhao, E., & Yang, P. (2023). Harmonic sources modeling and characterization in modern power systems: A comprehensive overview. *Electric Power Systems Research*, 218, Article 109234. <https://doi.org/10.1016/j.epsr.2023.109234>
- IEEE. (2022). *IEEE standard for harmonic control in electric power systems* (IEEE Std 519-2022). Institute of Electrical and Electronics Engineers.
- IEEE. (2025). *IEEE standard definitions for the measurement of electric power quantities under sinusoidal, nonsinusoidal, balanced, or unbalanced conditions* (IEEE Std 1459-2025). Institute of Electrical and Electronics Engineers.
- IEC. (2009). *Electromagnetic compatibility (EMC)—Part 4-7: Testing and measurement techniques—General guide on harmonics and interharmonics measurements and instrumentation, for power supply systems and equipment connected thereto* (IEC 61000-4-7:2002+AMD1:2008 CSV). International Electrotechnical Commission.
- IEC. (2025). *Electromagnetic compatibility (EMC)—Part 4-30: Testing and measurement techniques—Power quality measurement methods* (IEC 61000-4-30:2025). International Electrotechnical Commission.
- Kürker, F. (2017). *Enerji sistemlerinde harmonik analizi ve harmonik azaltma teknikleri* [Harmonic analysis and harmonic mitigation techniques in energy systems] [Publication No. 468305] [Doctoral dissertation, Harran University]. YÖK Tez Merkezi.
- Liu, Q., Liu, F., Li, Y., & Wang, S. (2023). Harmonic modeling and

- harmonic contribution determination: A case study of industrial power supply system. *Electric Power Systems Research*, 217, Article 109122. <https://doi.org/10.1016/j.epsr.2023.109122>
- Michalec, Ł., Jasiński, M., Sikorski, T., Leonowicz, Z., Jasiński, Ł., & Suresh, V. (2021). Impact of harmonic currents of nonlinear loads on power quality of a low voltage network—Review and case study. *Energies*, 14(12), Article 3665. <https://doi.org/10.3390/en14123665>
- Nascimento, C. F., de Oliveira, A. A., Jr., Goedel, A., & Dietrich, A. B. (2013). Harmonic distortion monitoring for nonlinear loads using neural-network-method. *Applied Soft Computing*, 13(1), 475–482. <https://doi.org/10.1016/j.asoc.2012.08.043>
- Nur, A., Güre, B., Rüstemli, S., & Bezek Güre, Ö. (2025). Wind speed prediction in microgrid energy management using the random forest and artificial neural networks methods. *Middle East Journal of Science*, 11(2), 276–289. <https://doi.org/10.51477/mejs.1738993>
- Obulesu, Y. P., Venkateswara Reddy, M., & Kusumalatha, Y. (2014). A %THD analysis of industrial power distribution systems with active power filter—Case studies. *International Journal of Electrical Power & Energy Systems*, 60, 107–120. <https://doi.org/10.1016/j.ijepes.2014.02.037>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Senra, R., Boaventura, W. C., & Mendes, E. M. A. M. (2017). Assessment of the harmonic currents generated by single-phase nonlinear loads. *Electric Power Systems Research*, 147, 272–279. <https://doi.org/10.1016/j.epsr.2017.02.028>
- Ugwuagbo, E., Balogun, A., Olajube, A., Omeje, O., Awelewa, A., & Abba-Aliyu, S. (2021). Experimental data on power quality assessment at point of common coupling of a steel mill to an electric power grid. *Data in Brief*, 39, Article 107681. <https://doi.org/10.1016/j.dib.2021.107681>
- Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLOS ONE*, 14(11), Article e0224365. <https://doi.org/10.1371/journal.pone.0224365>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83. <https://doi.org/10.2307/3001968>
- Willems, J. L. (2004). Reflections on apparent power and power factor in nonsinusoidal and polyphase situations. *IEEE Transactions on Power Delivery*, 19(2), 835–840. <https://doi.org/10.1109/TPWRD.2003.823182>