Optimised Time for Travelling Wave Fault Locators in the Presence of Different Disturbances Based on Real-World Fault Data

MAHMOOD PARSI1 (Member, IEEE), AND PETER A. CROSSLEY2 (Senior Member, IEEE)

1Fundamentals Ltd., Swindon SN5 5FZ, U.K.
2College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter EX4 4PY, U.K.
CORRESPONDING AUTHOR: M. PARSI (mmd.parsi@gmail.com)

The work of Mahmood Parsi was supported by the Engineering and Physical Sciences Research Council (EPSRC) through the Centre for Doctoral Training (CDT) in Power Networks under Grant EP/L016141/1.

ABSTRACT The real-world travelling wave fault data investigated in this paper indicate disturbances generate unpredictable, non-stationary and random waveforms which may cause maloperation of protection and control elements in a power system including travelling wave fault locators (TWFL). This type of fault locator is directly dependent on the detection of an accurate time of arrival (ToA) of travelling waves (TW) generated by a fault. This detection becomes complicated in the presence of disturbances when their ToAs are detected earlier than the fault TWs. Since travelling waves occur in the high-frequency bands (e.g. >50 kHz), in this paper a capacitor voltage transformer is employed to measure the TW voltage signals; this involves acquiring the current flowing to the ground and removing the low-frequency components (50/60 Hz). Disturbances create high magnitude pulses in the pre-fault section of a TW fault signal that last for a short time. Therefore, the time when a TWFL starts its computations requires to be optimised so that the effect of the disturbances is eliminated. The analysis techniques mentioned in this paper are based on real-world travelling wave fault data, and the solution uses statistical tools, such as cost function, mean and standard deviation, alongside Digital Signal Processing algorithms.

INDEX TERMS Fault location, optimised time, power network faults, power system disturbance, real-world data, short-circuit fault, travelling waves, travelling wave fault location.

I. INTRODUCTION

A C CURATE data analysis is important in ensuring a power system can deliver a stable and secure supply of electrical power to consumers. A recent development is the use of the travelling waves instigated by a short-circuit fault for determining the fault location. With the applications of communications and intelligent devices, it is feasible to access high-quality and wide bandwidth recorded data from transmission lines. Hence, the recorded data is analysed in intelligent electronic devices (IED) such as travelling wave fault locators (TWFL). This can help prevent failures by accurately locating points of weaknesses in a power system.

Disturbances often appear in recorded faults that occur on transmission lines due to various phenomena such as cloud discharges and vegetation. Disturbances may cause maloperation in TWFLs especially when they appear in the pre-fault section of a fault signal, and thus, the detection of the first ToA of the fault initiated surges becomes complex. The data in this study is obtained from various transmission lines with different geometry, but all based on a double-ended TWFL with a TWFL at the two ends of each line. A Double-ended (Type-D) fault location normally uses GPS for time synchronization. GPS has a high level of precision, 1 µs, and as travelling waves (TWs) propagate approximately at the speed of light, the time settings have to be reasonably accurate. As a result, 1 µs inaccurate detection of a fault surge’s ToA results in 150 m error in fault location estimation, so that a disturbance with a few microseconds of inaccuracy adversely impacts the functionality of the fault location algorithm.

Reliable detection of ToAs provides accurate fault location, and minimises the maintenance time and operation costs since minimum resources are utilised to find the fault and restore the lines. The functionality and accuracy of a TWFL are based on a correct detection of TW ToAs at the ends of the fault.
A disturbance is introduced here as a rising magnitude in the pre-fault section of a recorded fault signal and is detected as the first ToA which is far away from the first ToA, which is not related to the fault inception time or fault location. A disturbance and a fault have several characteristics in common, e.g. both have a non-stationary waveform. As an instance, if a disturbance is 20 micro-second away from a correct TW ToA, a maloperation of a TWFL creates 3 km inaccuracy in fault location estimation. Thus, time optimisation of TWFLs has become an essential part of a real-world power system.

The main contribution of the paper is using real-world TW fault signals to demonstrate a new algorithm for fault location, capable of distinguishing a true fault from possible disturbances. The solution begins by examining and assimilating how the human brain is remarkably capable of recognizing which part of a signal is a disturbance. The solution applies statistical tools to analyse the data so that the fault and the disturbance initiated waves are distinguishable from each other. Therefore, an optimised time between these two points is selected as the start point for a TWFL algorithm to process the rest of the signal, avoiding the disturbance.

This paper is organised as follows. Section II briefly explains the double-ended fault location technique, with a summary of the background research. Section III explores a measurement method used by the industrial collaborator of the research and the characteristics of the recorded signals. Section IV explains what type of disturbances a TWFL may encounter and the different segments of the signals that need to be examined. The methodology of the proposed solution is represented in section V followed by demonstrations of results in section VI.

II. BACKGROUND

A. HISTORY OF TW FAULT LOCATORS

TW Fault location methods commonly apply current transformers (CTs) rather than potential transformers (PTs) since CTs not only reproduce the transient surges on the secondary side, but also as the number of lines increases, the summated current signals become double amplitude [1]. A setup of double-ended fault locators is depicted in Fig. 1 which shows the propagation of the TWs from the fault location to fault locators at the two ends of the line; they are synchronised via a global timing system, e.g. GPS. The detection of the first ToAs has a crucial role in accurate fault location estimations [1]. TWFLs are designed to detect the ToAs at each end which are the basis for an accurate fault location. This detection is often affected by the presence of disturbances, e.g. cloud discharge or a tree contact with the lines, which result in large inaccuracies when the ToAs are recorded very precisely. Hence, inaccurate fault location decisions are likely if disturbances with high magnitudes occur at a time close to the actual fault initiated travelling wave arrival times.

The source of a disturbance may vary in different areas of the world as climate and environment are different, and the cause of most faults is natural phenomena (54% [2]). For example, a lightning flashover in a cloud may initially introduce a partial discharge into transmission line without causing a fault but after a time delay, it causes an insulator flashover and a fault. The problem is that a fault location device recognises the discharge as a fault and consequently delivers a wrong fault location estimation.

B. ABRUPT CHANGING POINT DETECTION

In the electrical engineering field, there have been limited studies based on change detection in a dataset; most of the relevant developments have been in mechanical engineering and numerous investigations have relied on the proposed solution described in a paper by Huang [3]. They mainly examine the behaviour of rotating machinery which generated non-stationary and non-linear waves, similar to recorded travelling wave fault signals from transmission lines.

The research in [4] uses a linear regression model for optimal segmentation and also the study in [5] has a broad discussion about segmentation and abrupt change detection of a signal, including a few algorithms for different scenarios, e.g. change detection algorithm. To categorise a non-stationary signal, the research [6] introduced different segments of a signal as: Abnormal Magnitudes, Rolling, Noise, Dependency Faults, Significant Rise/Fall, Spikes, Flat Intervals, Oscillations. It presents automatic methods for segmenting signals and to extract their important features, with the introduction of a quasi-optimal algorithm to select a subset of features.

The study in [7] proposes a signal processing method for clustering-based segmentation for fault detection in a mechanical bearing by applying the Fast Fourier Transform (FFT) and the Hilbert transform. The FFT is used to analyse the signal in the frequency domain and the Hilbert transform filters the signal and extracts the envelope of the band-passed signal. Furthermore, a method for signal segmentation of fault records based on Empirical Mode Decomposition is represented in [8]. It applies a cubic spline function to interpolate between the minimum and the maximum costs to detect abrupt changes in currents or voltages. However, the study has a deficiency in investigating fault signals with different characteristics. Moreover, the sampling frequency.

FIGURE 1. Travelling waves propagation towards to ends of transmission lines.
applied for the research (2.5 kHz) is not useful since a TWFL requires a high sampling frequency, e.g. up to 1.25 MHz in our recorded fault signals.

Modern tools like wavelet transform were significantly explored during the past few decades, for example, the research in [9] elucidates an abrupt change detection technique using a wavelet decomposition method. However, using Daubechies 1 and 4 do not provide excellent precision and accuracy in our real-world TW fault records. This is because of the broad frequency bandwidth (625 kHz) in the recorded signals and thus, it is not feasible to define one scaling level which can accurately extract the disturbance and fault characteristics. Also, implementing a wavelet solution is considerably more difficult than statistical-based solutions, especially when the optimum time for a TWFL can be an approximation.

The development in [9] discusses how significant changes in the distributional properties of a dataset have been studied in numerous different fields, ranging from finance and economics to genomics. The most commonly used method is based on a cost function and was first introduced by Scott and Knott [10]. Also, the study in [11] proposes a dynamic programming algorithm for change point detection. This was used for further developments, the pruned exact linear time (PELT) method, by Killick et al. [8]. This applies a pruning step within the dynamic program to observe the change points based on statistical criteria such as penalised likelihood, quasi-likelihood [12], and the cumulative sum of squares [13], [14]. However, the study is based on simulated models and uses a cost function for change detection, e.g.:

$$\sum_{i=1}^{m+1} \left[ C \left( \gamma(\tau_{i-1}+1), \tau_{i} \right) \right] + \beta f(m)$$  \hspace{1cm} (1)

where C is a cost function for a segment and $\beta f(m)$ is a penalty to guard against overfitting. The paper explains different penalty choices, but they will not be applied in our algorithm since our estimation is an approximation and will not require precise computations. To solve the problem in TWFLs caused by a disturbance, an excellent outcome will approximately set a starting time for a TWFL to operate after the disturbance using statistical and signal processing techniques.

III. MEASUREMENT METHOD

Recorded fault data is obtained by applying CTs with split-core linear couplers that connect a double-ended travelling wave system to a transmission network. The linear coupler has an air-gap so that power-line frequency (50/60 Hz) is filtered leaving the high-frequency travelling waves. Fault TWs have high frequencies and the aim is to detect the first abrupt increase rising to a peak, and thus, low-frequency components are removed. Another advantage of using the linear coupler is the non-intrusive installation which can often be done without a line outage. Fig. 2 depicts a typical voltage measurement device on a transmission line.

The linear coupler is capable to monitor the line signals, with a sampling rate of 1 to 20 MHz. However, it is 1.25 MHz in all the records described in this paper. This provides a high level of time precision equal to 0.8 $\mu$s timestep. Higher sampling frequencies deliver higher levels of precision in data acquisition but they increase computation load and are not essential for introducing an optimum time for a TWFL.

Instead of applying CTs directly to the transmission lines, their voltage can be measured using the capacitors within the Capacitive Voltage Transformer (CVT). A CT fitted in the earth path measures high-frequency currents that flow to the earth and the split-core CT alternates the low frequencies. Thus, pre-fault part of signals has a limited magnitude and the focus is on the detection of high-frequency components that carry ToAs.

The recorded fault signals studied in this paper come from various transmission lines, and so, their geometries differ. As TWs are used in this study, the required parameters are line length, the velocity of propagation and velocity factor. Velocity factor is introduced because the line length and propagation characteristics of the line are not as the speed of light due to heat, age of materials, weather, etc [16].

IV. DISTURBANCES

The main focus of this article is on disturbances that appear in the immediate pre-fault section of a recorded fault current. The aim is to ensure the fault location algorithm starts at an optimal point in time, avoiding the disturbance and therefore, prevents maloperation of the TWFL.

The study in [6] elucidates a technique to segment signals using signal processing tools by categorising different parts of each signal. However, the difficulty with disturbances in travelling wave fault data is to find only one optimum time for use in the fault location algorithm. For this goal and fault location automation, waveforms of recorded fault signals are initially analysed. Generally, they consist of two main parts: pre- and post-fault sections. The first part comprises the fault disturbance, which is recognisable by a human brain but difficult for a machine. In Fig. 3, the absolute values of a fault signal show clearly how it is segmented into three parts and are described as follows:
Parsi and Crossley: Optimised Time for TWFL in the Presence of Different Disturbances Based on Real-World Fault Data

FIGURE 3. Three Different sections of a fault signal.

FIGURE 4. Examples of major rises with/out a disturbance: (a) typical fault rise (b) fault rise taking longer to reach a peak (c) TWs with a high frequency (d) intense wave with disturbances in the pre-fault area (e) intense wave without a disturbance in the pre-fault area (f) rising waves with spike disturbances.

• Flat Interval: This section of a signal is almost flat and close to zero; it corresponds to the steady-state part. Note, the power frequency signals are severely alternated as explained previously. Power frequency signals without a disturbance should result in a flat interval not only in the pre-disturbance section of a fault signal but also after fault clearance in the steady-state area of the post-fault section.

• Disturbance: The second section is where the disturbance occurs, the magnitude increases for a short time in the pre-fault section of the fault signal, i.e. before fault occurrence.

• Abrupt Rise: The third part is associated with the fault which has a larger magnitude than the disturbance. It also lasts longer with a larger mean value.

The waveforms in Fig. 4 indicate various shapes in the above three sections. The graph in (a) shows an abrupt rising magnitude without disturbance in the pre-fault area whereas (b) and (c) have small disturbances with wide and short time windows respectively before reaching a peak at the fault inception times. The waveforms in (d), (e) and (f) show a general view of the whole fault signals. The signal in (d) has significant disturbances in the pre-fault area while (e) and (f) indicate flat intervals with minor disturbances in their pre-fault sections. These examples show the difficulty to find an optimum time for a TWFL to start its algorithm from. A well-designed method to detect disturbances reduces the operation costs of the power system and improves the accuracy of TWFLs.

V. METHODOLOGY

A short circuit fault on a transmission line normally creates an extreme increase in the magnitude of the measured current. This abrupt change is where the statistical parameters change significantly. Assume the order of the dataset sequence is:

$$y = (y_1, y_2, y_3, \ldots, y_k)$$

and significant changes occur at $$y_m$$ and $$y_n$$ where the former is associated with the ToAs of the disturbance and the fault-initiated wave. The optimum time is a time between these two times. The mean value of the signal in the post-fault area has a large magnitude and is considerably greater than a similar value from the pre-fault area of the fault signal. This observation is the foundation of differentiating an abrupt change point from the disturbance data as:

$$\text{Mean of Post} - \text{Fault} > \text{Mean of Pre} - \text{Fault}$$

This paper introduces a solution by choosing the maximum value of the fault signal as the reference point, see Fig. 5. At the reference point, the signal is split into two areas called Production (left side) and Cost (right side). The aim is to calculate the mean of each area followed by a cost function. The cost function will generate two peaks associated with the disturbance and the fault-initiated wave.
However, in calculating the mean values and when the fault is cleared, the mean value for the cost area decreases significantly again. Hence, to keep the mean of the post-fault area very large, signal processing tools are applied to flip the signal, so that the signal is reversed from the end of the data set to the maximum peak. Fig. 6 indicates the flipping of the data in Fig. 5. The first peak above 50% of the maximum peak is chosen as the last data in the cost function. The selection allows several peaks with high magnitudes to be considered for our mean calculations. This is considered as the cost section and the left side of the reference point is the production part.

To confirm if the mean values are correctly detecting an abrupt change in a signal, there will be a double-checking by applying the Standard Deviation (SD). Thus, a second measure is used to examine the effectiveness of the initial measure (mean). This will improve our confidence in using the proposed method. SD is defined as a measure to quantify the variation or dispersion of a set of data values. Consequently, when the cost function, based on mean values, is used to find a data point, its trueness is checked by SD. The mean values ($\bar{x}$) and the SD are obtained as follows:

\[
\bar{x} = \frac{1}{n} \left( \sum_{i=1}^{n} x(i) \right) \\
SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x(i) - \bar{x})^2}
\]

In an iterative process, the mean and SD values are stored for further calculations, i.e., until a disturbance is detected. In the first iteration, the cost data is between the maximum peak and its 50%. For the second iteration, this area is removed and the maximum peak is detected for the remaining dataset which then divides the dataset into cost and production areas. The mean and SD values are calculated for these areas. The algorithm continues to process until reaching an endpoint. The endpoint is a small peak at the beginning of the fault signal, e.g., 0.1% of the maximum. For example, Fig. 7 and 8 depict subsequent segmentations by moving from the reference point to the fault inception time and the disturbance by calculating the mean values per segment. Having saved values of each segment, the cost function is introduced as it records the difference between each subsequent mean and SD values respectively as follows:

\[
CT_{Mean}(i) = |\bar{x}(i + 1) - \bar{x}(i)|
\]

\[
CT_{SD}(i) = |SD(i + 1) - SD(i)|
\]

By Calculating the cost functions, the difference between two mean or SD values is substantial at a certain point. That is where the mean or SD of the cost decreases significantly and will be where a fault occurs. The rest of the solution shows a second large change associated with the disturbance, e.g., the disturbance peak in Fig. 9. Therefore, approximately half the distance between these two points is the optimal time for starting a TW fault location algorithm.

Once the abrupt change associated with the fault-initiated wave is detected by the method mentioned above, a disturbance in the pre-fault area of a signal is detected by applying the same method to the remaining data in the last production
section after finding an abrupt change in the signal. This process creates a second peak in the cost functions and is associated with the disturbance. The time between these two points is the optimum time to start a travelling fault location. The algorithm is as represented in Table 1.

### TABLE 1. Optimised time algorithm.

| TWFL Start Time Optimisation Algorithm |
|---------------------------------------|
| **Inputs**: fault signal with time and magnitude related to each data |
| **Output**: Optimum time for a TWFL to start a computation |
| The sequence of the algorithm: |
| 1. obtain the absolute values of the input signal coefficients |
| 2. flip the signal and find the first peak above 50% of the maximum peak of the whole signal |
| 3. find time associated with the peak in step 2 from the non-flipped original data |
| 4. find the length of the signal in step 1 |
| 5. find a small peak from the start of the signal that can be considerably smaller than any possible disturbance |
| 6. find the time associated with step 5 |
| 7. the first segmented signal is between the time in step 3 and the time in step 5 |
| 8. process the following steps until reaching the end pint: |
| a. find the time associated with the maximum of the segmented signal |
| b. compute the Mean and SD on the left (Production) and Right (Cost) sides of a partitioned dataset |
| c. store the left (Production) side of segmented signal for the next iteration |
| d. if (time in 8.a) < (time in 6) then exit the process, otherwise continue segmenting until this condition is met |
| e. have the left side (Production) of signal in 8.a as the new segmented signal for the next iteration |
| f. end the process |
| 9. find the difference between per two subsequent Mean and SD values |
| 10. acquire the max of calculations in step 9 as the fault inception time |
| 11. find the second max of step 9 as the disturbance time |
| 12. the optimum time for a fault location is between the times obtained in steps 10 and 11 |

The red and blue patterns in Fig. 10 are related to the mean values of the cost and production sides respectively. The red pattern illustrates differences among individual points.
VI. RESULTS AND DISCUSSION

The investigation demonstrates that results from the cost area deliver more reasonable outcomes compared to the production side. Fig. 12 shows where the maximum difference between two mean and SD values are when they reach peaks relevant to the fault and the disturbance times. Considering the non-stationery and frequency-varying signal at the beginning, this graph displays smooth and clear outcomes.

The proposed method calculated the correct time for all eighty-five real-world travelling wave fault signals. The plot in Fig. 13 demonstrates the achieved performance of the optimised time algorithm; the blue-square and green-triangle data are the fault inception and disturbance times respectively for each dataset. The optimised time for a fault location must be between these two points as indicated by the red circles. The results indicate an excellent time optimization is achieved for the entire examined cases.

To demonstrate the results more accurately, Fig. 14 shows $t_{T/2}$ as half the time gap (T) between the disturbance and the fault inception times. It introduces a range of acceptable tolerances, e.g. 1%, 5% and 50% of $t_{T/2}$. It displays examples of acceptable tolerance ranges where the optimised times could be detected.

Table 2 represents how many of the results are within the acceptable tolerance range. It can be observed, not only are

| Range of Acceptable Tolerance | Number of Results within the Tolerance Range |
|------------------------------|---------------------------------------------|
| $t_{\text{optimised}}$ ± 1% $t_{\text{optimised}}$ | 29                                           |
| $t_{\text{optimised}}$ ± 5% $t_{\text{optimised}}$ | 43                                           |
| $t_{\text{optimised}}$ ± 10% $t_{\text{optimised}}$ | 52                                           |
| $t_{\text{optimised}}$ ± 20% $t_{\text{optimised}}$ | 59                                           |
| $t_{\text{optimised}}$ ± 30% $t_{\text{optimised}}$ | 67                                           |
| $t_{\text{optimised}}$ ± 50% $t_{\text{optimised}}$ | 81                                           |
| within $t_{\text{dis}}$ and $t_{\text{fault}}$ | 85                                           |

Table 3 represents the coverage quantity of the tolerance intervals.

| Range of Acceptable Tolerance | Tolerance Coverage |
|------------------------------|--------------------|
| $t_{\text{optimised}}$ ± 1% $t_{\text{optimised}}$ | 34.12%             |
| $t_{\text{optimised}}$ ± 5% $t_{\text{optimised}}$ | 50.59%             |
| $t_{\text{optimised}}$ ± 10% $t_{\text{optimised}}$ | 61.18%             |
| $t_{\text{optimised}}$ ± 20% $t_{\text{optimised}}$ | 69.41%             |
| $t_{\text{optimised}}$ ± 30% $t_{\text{optimised}}$ | 78.82%             |
| $t_{\text{optimised}}$ ± 50% $t_{\text{optimised}}$ | 95.29%             |
| within $t_{\text{dis}}$ and $t_{\text{fault}}$ | 100%               |
TABLE 4. The exact binomial confidence interval for each accepted tolerance range.

| Range of Acceptable Tolerance | CI = 90% | CI = 95% | CI = 99% |
|--------------------------------|----------|----------|----------|
|                                |          |          |          |
| $t_{\text{optimised}} \pm 1\%$ | 0.25599 $\leq p \leq 0.43491$ | 0.24175 $\leq p \leq 0.45202$ | 0.21507 $\leq p \leq 0.48559$ |
| $t_{\text{optimised}} \pm 5\%$ | 0.41176 $\leq p \leq 0.59969$ | 0.39520 $\leq p \leq 0.61614$ | 0.36338 $\leq p \leq 0.64371$ |
| $t_{\text{optimised}} \pm 10\%$ | 0.52894 $\leq p \leq 0.71148$ | 0.51811 $\leq p \leq 0.72637$ | 0.47838 $\leq p \leq 0.75445$ |
| $t_{\text{optimised}} \pm 20\%$ | 0.60174 $\leq p \leq 0.77602$ | 0.58473 $\leq p \leq 0.78951$ | 0.55120 $\leq p \leq 0.81426$ |
| $t_{\text{optimised}} \pm 30\%$ | 0.70232 $\leq p \leq 0.85843$ | 0.68613 $\leq p \leq 0.86940$ | 0.65368 $\leq p \leq 0.88931$ |
| $t_{\text{optimised}} \pm 50\%$ | 0.89554 $\leq p \leq 0.98377$ | 0.88387 $\leq p \leq 0.98703$ | 0.85916 $\leq p \leq 0.99198$ |

all 85 fault cases timely optimised to be after the disturbance, but more than fifty per cent (43 cases) were optimised whilst the tolerance range was $\pm 5\%$ of the optimised time.

To evaluate the level of confidence based on the tolerance ranges in Table 2, tolerance interval is applied which is defined in two quantities. The first quantity is coverage that is equal to the proportion of the population ($p$) covered by the interval or the percentage of successes in a population; calculated as [17]:

$$\text{coverage} = \frac{\text{No. of results within the tolerance range}}{\text{total population}} \quad (6)$$

The second quantity is the probabilistic confident of a population proportion covered in the tolerance interval. It is a binary indication of being within a tolerance interval (success) or not (failure). This is assessed using exact Binomial Confidence Interval (BCI) defined respectively for lower and upper bounds of a tolerance at a desired confidence ($\alpha$) as:

$$\sum_{k=\text{int}}^{n} \binom{n}{k} p_{ub}^k (1 - p)^{(n-k)} = \frac{\alpha}{2} \quad (7)$$

$$\sum_{k=0}^{x} \binom{n}{k} p_{lb}^k (1 - p)^{(n-k)} = \frac{\alpha}{2} \quad (8)$$

where $p_{ub}$ and $p_{lb}$ are the upper and lower bounds of the tolerance range respectively, $n$ is the number of trials and $k$ is the number of successes in $n$ trials [17]. Therefore, the coverage of the tolerance ranges mentioned in Table 2 are as shown in Table 3. The tolerance coverage means, for example for 50% tolerance, the correct optimised times are 95.29% covered within a 50% tolerance range from the middle of the time spectrum between the disturbance and the fault inception times.

Table 4 shows the results for the exact binomial confident interval for each tolerance range. They are cited within boundaries according to the tolerance ranges and the boundaries determine the confidence intervals. For instance, if the accepted tolerance rate is set to 50%, we would be 95% confident that the proportion of the correct optimised times falls between 88.387% and 98.703%. As discussed before, the eighty-five fault signals are correctly optimised, which according to Table 3 means that the proposed algorithm gains 99% confidence that 93.957% to 100.00% of the population are correctly optimised.

Extracting data from Tables 3 and 4 manifests the coverage levels in each tolerance intervals and the binomial confidence intervals. They are depicted in Fig. 15. In each tolerance range, each bar chart cites the coverage level with three margins of the BCIs. They are shown in three different colours when the confidence interval is equal to 90%, 95% and 99%. Note, the margins must not be more than 100% in the last bar, but between a lower percentage and 100%.

VII. CONCLUSION

Disturbances in transmission lines cause maloperation in fault locators which have an adverse impact on the stability of electrical supply. The sources of the disturbances vary, e.g. vegetation or cloud discharge, and can occur with various time range and magnitude. The fundamentals of travelling wave fault location methods were elucidated, including a description of a CVT measurement method that our industrial collaborator uses.

A time optimisation technique was proposed in this paper for use with travelling wave-based fault locators, this avoids disturbances, and as a result, prevents inaccurate decisions and possible maloperation. Statistical tools and digital signal processing techniques were employed to investigate where in the fault inception and disturbances occur. Furthermore, a cost-production function is defined to distinguish the significant changes in the recorded signals, that are associated with the disturbance and fault inception times.
The paper demonstrated mean and standard deviation were appropriate tools to find the disturbance and fault inception times. The research indicated segmentation of a signal into production and cost areas simplifies the process of computation. The cost side of the segmented signal demonstrated differences across each individual data more explicitly than the production side. Since these quantities reach peaks where the fault inception and disturbance times are, the proposed algorithm finds the optimum time as the half of their time gap. The optimised time is where a fault location starts its computation to avoid a disturbance and reduce the risk of maloperation.

All eighty-five fault signals were timely optimised by detecting the disturbance and fault inception times. The statistical analysis showed high levels of confidence intervals when an acceptable tolerance range is larger than five percent. Since the research applied a variety of signals and all were successfully optimised, the interval between a disturbance and a fault inception time appeared to have a negligible effect on the accuracy of the proposed method, it however will increase the processing time to compute segmenting such signals. Further research into the proposed methodology and access to more data may increase the accuracy levels within each tolerance interval. However, access to a reasonable number of fault data in transmission systems is exceptionally challenging while they do not occur frequently in many areas of the world.

ACKNOWLEDGMENT
The authors were with the Department of Electrical and Electronic Engineering, School of Engineering, The University of Manchester, Manchester M13 9PL, U.K.

REFERENCES
[1] IEEE Guide for Determining Fault Location on AC Transmission and Distribution Lines, I. S. Assoc., New York, NY, USA, 2014.
[2] T. Shibata, T. Fujita, K. Yoshimura, and H. Nagashima, “Improvements to fault location analysis based on voltage sag data—Verifying the improvements to the fault location accuracy of the prototype tool,” presented at the IEEE 2nd Int. Power and Energy Conf., Johor Bahru, Malaysia, 2008.
[3] N. E. Huang et al., “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” Proc. Roy. Soc. London, Ser. A, Math., Phys. Eng. Sci., vol. 454, no. 1971, pp. 903–995, Mar. 1998. [Online]. Available: https://www-jstor-organ.manchester.idm.oclc.org/stable/53161?seq=1#metadata_info_tab_contents
[4] T. D. Popescu and D. Aiorodachioae, “Fault detection of rolling element bearings using optimal segmentation of vibrating signals,” Mech. Syst. Signal Process., vol. 116, pp. 370–391, Feb. 2019. doi: 10.1016/j.ymssp.2018.06.033.
[5] M. Basseville and I. V. Nikiforov, Detection of Abrupt Changes: Theory and Application. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
[6] J. A. Crossman, H. Guo, Y. L. Murphy, and J. Cardillo, “Automotive signal fault diagnostics. I. Signal fault analysis, signal segmentation, feature extraction and quasi-optimal feature selection,” IEEE Trans. Veh. Technol., vol. 52, no. 4, pp. 1063–1075, Jul. 2003. doi: 10.1109/TVT.2002.807635.
[7] H. Hou, M. Liang, Y. Zhang, and C. Li, “Vibration signal demodulation and bearing fault detection: A clustering-based segmentation method,” Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci., vol. 228, no. 11, pp. 1888–1899, Aug. 2014, doi: 10.1177/0954406213514960.
[8] R. Killick, P. Fearnhead, and I. A. Eckley, “Optimal detection of changepoints with a linear computational cost,” J. Amer. Stat. Assoc., vol. 107, no. 500, pp. 1590–1598, Dec. 2012, doi: 10.1080/01621459.2012.737745.
[9] A. Ukil and R. Ivanovic, “Application of abrupt change detection in power systems disturbance analysis and relay performance monitoring,” IEEE Trans. Power Del., vol. 22, no. 1, pp. 59–66, Jan. 2007, doi: 10.1109/TPWRD.2006.887096.
[10] A. J. Scott and M. Knott, “A cluster analysis method for grouping means in the analysis of variance,” Biometrics, vol. 30, no. 3, pp. 507–512, 1974, doi: 10.2307/2529204.
[11] B. Jackson et al., “An algorithm for optimal partitioning of data on an interval,” IEEE Signal Process. Lett., vol. 12, no. 2, pp. 105–108, Feb. 2005, doi: 10.1109/LSP.2001.838216.
[12] J. Braun, R. K. Braun, and H. G. Müller, “Multiple changepoint fitting via quasi-likelihood, with application to DNA sequence segmentation,” Biometrika, vol. 87, no. 2, pp. 301–314, Jun. 2000, doi: 10.1093/biomet/87.2.301.
[13] F. Picard, E. Lebarbier, M. Hoebeke, G. Rigail, B. Thiam, and S. Robin, “Joint segmentation, calling, and normalization of multiple CGH profiles,” Biostatistics, vol. 12, no. 3, pp. 413–428, Jul. 2011, doi: 10.1093/biostatistics/kxq076.
[14] C. Inclán and G. C. Tiao, “Use of cumulative sums of squares for retrospective detection of changes of variance,” J. Amer. Stat. Assoc., vol. 89, no. 427, pp. 913–923, Sep. 1994, doi: 10.1080/01621459.1994.10476824.
[15] F. Ghassemi, P. F. Gale, B. Clegg, T. Cumming, and C. Coutts, “Method to measure CVT transfer function,” IEEE Trans. Power Del., vol. 17, no. 4, pp. 915–920, Oct. 2002, doi: 10.1109/TPWRD.2002.803739.
[16] M. Gilany, D. K. Ibrahim, and E. S. T. Eldin, “Traveling-wave-based fault-location scheme for multiend-aged underground cable system,” IEEE Trans. Power Del., vol. 22, no. 1, pp. 82–89, Jan. 2007.
[17] P. Mayfield. Exact Confidence Interval. SigmaZone. Accessed: May 2020. [Online]. Available:https://sigmazone.com/binomial-confidence-intervals

MAHMOOD PARI (Member, IEEE) received the M.Sc. degree in embedded systems engineering from the University of Leeds, U.K., in 2015. He was a Ph.D. Researcher with the Power Networks Group, The University of Manchester, U.K. He worked as an Electrical Engineer and a Teaching Assistant. In 2020, he joined the Research and Development Team with Fundamentals Ltd., U.K., as a Power Systems and Algorithms Engineer. His study was funded by EPSRC Centre for Doctoral Training (CDT). His research interests include power system protection, fault locators, and data analysis in power systems. He is a member of IET and CIGRE-UK.

PETER A. CROSSLEY (Senior Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from The University of Manchester Institute of Science and Technology, Manchester, U.K., in 1977, and the Ph.D. degree in power system engineering from the University of Cambridge, U.K., in 1983. He was a Professor of electrical power systems engineering with The University of Manchester, and the Director of the Doctoral Training Centre, Manchester, U.K. In 2021, he joined Exeter University as a Professor with the Smart Grids Research Group. He is also the author or coauthor of more than 250 technical articles. He is also an active member of various CIGRE, IEEE, and IET committees.