Supplementary Material: A Method for Fault Classification in Transmission Lines with Sampling Rate Flexibility and Generalization Capacity

Leandro Augusto Ensina[a,b], Luiz Eduardo Soares de Oliveira[a], Eduardo Cunha de Almeida[a], Signie Laureano França Santos[c], Leandro Silva Bernardino[d]

[a] *Department of Informatics (DInf), Federal University of Paraná (UFPR), Curitiba, PR, Brazil*
[b] *Computer Engineering Coordination (COENC), Federal University of Technology - Paraná (UTFPR), Toledo, PR, Brazil*
[c] *Institute of Technology for Development (LACTEC), Curitiba, PR, Brazil*
[d] *Energy Company of Paraná (COPEL), Curitiba, PR, Brazil*

Abstract

This Supplementary Material companion the paper entitled "A Method for Fault Classification in Transmission Lines with Sampling Rate Flexibility and Generalization Capacity", published in Electric Power Systems Research, vol. XX, no. XX, pp. XX–XX, year 202X, doi: XXXXXXXXX. It contains complementary information regarding the proposed method for fault type classification. This document is not intended to be self-contained and should be read jointly with the original paper.

1. Additional information

Table 1 shows the p-values for the statistical comparisons performed among our method and the related works, as reported in Section 5.2 (page 17) of the paper.

Table 1: P-values for the pairwise comparison using Wilcoxon signed test. The mark • represents the statistically significant difference for a confidence interval of 95%.

Sampling rate	**Proposed method *vs* Ferreira *et al.* [1]**		**Proposed method *vs* Coban & Tezan [2]**	
	TL1	**TL2**	**TL1**	**TL2**
256 Hz	1.7344E-6 •	1.0246E-5 •	0.0035 •	1.7344E-6 •
1 kHz	2.5614E-6 •	2.0603E-5 •	7.6195E-6 •	1.9209E-6 •
5 kHz	3.7896E-6 •	2.8786E-6 •	7.8294E-6 •	1.7344E-6 •
10 kHz	3.7869E-6 •	3.1652E-6 •	5.4086E-6 •	1.7322E-6 •

Another analysis conducted is related to the importance of the features for the predictions, which enables us to understand their relevance for the algorithm used in the proposed method (i.e., Random Forest - RF). Table 2 shows the feature's importance for the fault type classification for the RF algorithm, in which the higher the value, the more relevant the attribute. The scores presented in this table represent the average and the standard deviation for each feature considering three repetitions (once per training set). The relevance of a feature is computed based on the Gini importance with a normalized range of scores (the values sum to 1) [3].

Table 2: The features' importance for the proposed method considering the Random Forest algorithm.

Feature	**256 Hz**	**1 kHz**	**5 kHz**	**10 kHz**
maxPhaseA	0.14 ± 0.02	0.13 ± 0.02	0.13 ± 0.01	0.12 ± 0.01
maxPhaseB	0.13 ± 0.02	0.14 ± 0.01	0.15 ± 0.00	0.14 ± 0.02
maxPhaseC	0.12 ± 0.01	0.12 ± 0.00	0.13 ± 0.02	0.13 ± 0.01
minPhaseA	0.13 ± 0.02	0.13 ± 0.00	0.14 ± 0.01	0.13 ± 0.01
minPhaseB	0.15 ± 0.01	0.14 ± 0.02	0.13 ± 0.02	0.15 ± 0.01
minPhaseC	0.11 ± 0.00	0.13 ± 0.01	0.12 ± 0.00	0.13 ± 0.02
groundDetection	0.22 ± 0.01	0.21 ± 0.01	0.20 ± 0.01	0.20 ± 0.01

As can be identified, groundDetection is the feature that manifested the higher contribution to the predictions. This aspect can be explained by the analysis conducted in Section 5.2 of the paper, which revealed that the groundDetection is crucial to the RF algorithm to distinguish among the fault types regarding the involvement or not of the ground in the failure, e.g., discerning between AB and ABG.

In contrast, the rest of the features presented similar scores. This behavior was expected since each pair of attributes (max and min) aimed to contribute only to its respective phase, identifying variations individually for each one of them. In other words, these features aim to enable the algorithm (RF) to recognize if a fault affected a particular phase.

Also, Random Forest outperforms the other algorithms because it is an ensemble technique composed of multiple specialists, ensuring diversity among them by restricting classifiers to work on different random subsets of the entire feature space. Therefore, the combination of a group of estimators with uncorrelated errors can identify distinct patterns in the feature set, which can result in a more accurate model for a particular problem [4], such as fault classification.

Table 3 presents all experimental data for the several scenarios evaluated in this paper in order to enable a complete and detailed analysis. The performances correspond to the results for both terminals together.

Table 3: Average accuracy and standard deviation for fault type classification for the experiments.

		TL1 (validation set)		TL1 (testing set)	TL2		
		Without normalization and groundDetection feature	**Without groundDetection feature**	**Proposed Method**	**Without normalization and groundDetection feature**	**Without groundDetection feature**	**Proposed Method**
	AG	99.92% ± 0.01%	99.92% ± 0.02%	**99.97% ± 0.02%**	52.18% ± 27.83%	98.39% ± 0.43%	**97.82% ± 0.46%**
	BG	99.96% ± 0.02%	99.97% ± 0.03%	**99.92% ± 0.04%**	51.41% ± 21.63%	95.06% ± 3.11%	**94.60% ± 1.04%**
	CG	99.94% ± 0.03%	99.97% ± 0.02%	**99.97% ± 0.03%**	71.31% ± 21.49%	97.60% ± 0.51%	**98.39% ± 1.05%**
	AB	99.80% ± 0.14%	99.86% ± 0.07%	**100.00% ± 0.00%**	0.01% ± 0.02%	32.08% ± 6.63%	**100.00% ± 0.00%**
256 Hz	**AC**	99.84% ± 0.06%	99.87% ± 0.06%	**100.00% ± 0.00%**	0.00% ± 0.00%	44.03% ± 28.45%	**100.00% ± 0.00%**
	BC	99.68% ± 0.21%	99.71% ± 0.23%	**100.00% ± 0.00%**	0.00% ± 0.00%	17.06% ± 7.34%	**100.00% ± 0.00%**
	ABG	99.09% ± 0.58%	99.60% ± 0.16%	**99.98% ± 0.01%**	83.47% ± 18.04%	98.06% ± 0.62%	**100.00% ± 0.00%**
	ACG	99.07% ± 0.39%	99.56% ± 0.10%	**99.97% ± 0.03%**	71.58% ± 27.27%	98.34% ± 0.34%	**99.98% ± 0.01%**
	BCG	99.21% ± 0.52%	99.40% ± 0.41%	**99.98% ± 0.02%**	80.82% ± 19.80%	97.47% ± 2.06%	**99.82% ± 0.14%**
	ABC	100.00% ± 0.00%	99.99% ± 0.01%	**100.00% ± 0.00%**	86.39% ± 14.85%	99.99% ± 0.01%	**100.00% ± 0.00%**
Total		99.65% ± 0.44%	99.79% ± 0.24%	**99.98% ± 0.03%**	49.72% ± 38.58%	77.81% ± 32.70%	**99.06% ± 1.73%**
	AG	99.96% ± 0.03%	99.93% ± 0.04%	**99.98% ± 0.02%**	61.10% ± 20.44%	98.07% ± 1.21%	**98.05% ± 0.65%**
	BG	99.97% ± 0.01%	99.96% ± 0.02%	**99.99% ± 0.01%**	67.84% ± 23.04%	97.60% ± 2.33%	**98.44% ± 0.96%**
	CG	99.97% ± 0.02%	99.97% ± 0.03%	**99.98% ± 0.01%**	64.08% ± 20.84%	98.53% ± 1.49%	**99.54% ± 0.44%**
	AB	99.85% ± 0.02%	99.78% ± 0.21%	**100.00% ± 0.00%**	0.00% ± 0.00%	85.65% ± 4.24%	**100.00% ± 0.00%**
1 kHz	**AC**	99.84% ± 0.08%	99.90% ± 0.09%	**100.00% ± 0.00%**	0.00% ± 0.00%	65.79% ± 2.90%	**100.00% ± 0.00%**
	BC	99.89% ± 0.06%	99.90% ± 0.06%	**100.00% ± 0.00%**	0.00% ± 0.00%	63.01% ± 10.76%	**100.00% ± 0.00%**
	ABG	99.46% ± 0.45%	99.38% ± 0.54%	**99.99% ± 0.01%**	80.08% ± 21.42%	98.90% ± 0.27%	**100.00% ± 0.00%**
	ACG	99.50% ± 0.35%	99.45% ± 0.48%	**99.98% ± 0.01%**	77.80% ± 23.12%	98.67% ± 1.19%	**100.00% ± 0.00%**
	BCG	99.61% ± 0.18%	99.54% ± 0.26%	**99.94% ± 0.04%**	80.80% ± 20.34%	98.04% ± 0.50%	**98.87% ± 0.14%**
	ABC	99.99% ± 0.01%	99.99% ± 0.01%	**100.00% ± 0.00%**	91.87% ± 8.90%	100.00% ± 0.00%	**100.00% ± 0.00%**
Total		99.80% ± 0.26%	99.78% ± 0.32%	**99.98% ± 0.02%**	52.36% ± 38.87%	90.42% ± 14.17%	**99.59% ± 0.78%**
	AG	99.98% ± 0.01%	99.94% ± 0.06%	**99.99% ± 0.01%**	58.10% ± 14.26%	99.23% ± 0.43%	**99.20% ± 0.56%**
	BG	99.98% ± 0.02%	99.98% ± 0.01%	**99.99% ± 0.01%**	66.51% ± 23.40%	98.90% ± 1.14%	**99.52% ± 0.29%**
	CG	99.97% ± 0.02%	99.96% ± 0.04%	**99.99% ± 0.01%**	56.11% ± 24.55%	98.88% ± 1.14%	**99.73% ± 0.51%**
	AB	99.85% ± 0.11%	99.90% ± 0.11%	**100.00% ± 0.00%**	0.00% ± 0.00%	82.34%± 8.51%	**99.99% ± 0.01%**
5 kHz	**AC**	99.94% ± 0.03%	99.99% ± 0.01%	**100.00% ± 0.00%**	0.00% ± 0.00%	85.26% ± 9.04%	**100.00% ± 0.00%**
	BC	99.83% ± 0.13%	99.85% ± 0.16%	**100.00% ± 0.00%**	0.00% ± 0.00%	74.06% ± 3.37%	**99.98% ± 0.02%**
	ABG	99.66% ± 0.13%	99.60% ± 0.19%	**99.99% ± 0.01%**	79.66% ± 21.46%	95.83% ± 3.76%	**99.95% ± 0.07%**
	ACG	99.69% ± 0.16%	99.69% ± 0.09%	**99.99% ± 0.01%**	77.86% ± 22.34%	95.46% ± 6.57%	**100.00% ± 0.00%**
	BCG	99.62% ± 0.16%	99.69% ± 0.13%	**99.95% ± 0.01%**	80.14% ± 20.20%	85.57% ± 16.32%	**99.93% ± 0.09%**
	ABC	99.99%± 0.01%	99.99% ± 0.01%	**100.00% ± 0.00%**	92.61% ± 8.09%	100.00% ± 0.00%	**100.00% ± 0.00%**
Total		99.85% ± 0.16%	99.86% ± 0.17%	**99.99% ± 0.01%**	51.10% ± 38.49%	91.55% ± 10.80%	**99.83% ± 0.35%**
	AG	99.98% ± 0.02%	99.94% ± 0.05%	**99.99% ± 0.01%**	55.13% ± 20.24%	99.00% ± 0.60%	**99.36% ± 0.53%**
	BG	99.98% ± 0.02%	99.97% ± 0.03%	**99.99% ± 0.01%**	66.77% ± 20.78%	99.07% ± 0.92%	**99.81% ± 0.17%**
	CG	99.98% ± 0.02%	99.99% ± 0.01%	**100.00% ± 0.00%**	54.72% ± 25.61%	98.89% ± 0.99%	**99.84% ± 0.20%**
	AB	99.88% ± 0.10%	99.95% ± 0.05%	**100.00% ± 0.00%**	0.00% ± 0.00%	82.74% ± 8.39%	**100.00% ± 0.00%**
10 kHz	**AC**	99.94% ± 0.04%	99.98% ± 0.02%	**100.00% ± 0.00%**	0.00% ± 0.00%	86.03% ± 8.67%	**100.00% ± 0.00%**
	BC	99.83% ± 0.15%	99.87% ± 0.15%	**100.00% ± 0.00%**	0.00% ± 0.00%	76.14% ± 2.02%	**99.98% ± 0.03%**
	ABG	99.70% ± 0.10%	99.63% ± 0.18%	**99.98% ± 0.01%**	80.91% ± 20.49%	91.52% ± 9.70%	**100.00% ± 0.00%**
	ACG	99.67% ± 0.08%	99.68% ± 0.12%	**99.98% ± 0.01%**	78.52% ± 22.38%	97.76% ± 0.90%	**100.00% ± 0.00%**
	BCG	99.61% ± 0.13%	99.70% ± 0.11%	**99.96% ± 0.01%**	80.06% ± 20.40%	88.99% ± 11.09%	**99.91% ± 0.16%**
	ABC	99.99% ± 0.01%	100.00% ± 0.00%	**100.00% ± 0.00%**	92.67% ± 8.03%	100.00% ± 0.00%	**100.00% ± 0.00%**
Total		99.85% ± 0.16%	99.87% ± 0.16%	**99.99% ± 0.01%**	50.88% ± 38.70%	92.01% ± 9.74%	**99.89% ± 0.26%**

References

[1] V. H. Ferreira, R. Zanghi, M. Z. Fortes, S. G. Júnior, A. P. Alves da Silva, Probabilistic transmission line fault diagnosis using autonomous neural models, Electr. Power Syst. Res. 185 (2020) 1–10. doi:https://doi.org/10.1016/j.epsr.2020.106360.

[2] M. Coban, S. S. Tezcan, Detection and classification of short-circuit faults on a transmission line using current signal, Bull. Pol. Acad. Sci.: Tech. Sci. 69 (4) (2021). doi:https://doi.org/10.24425/bpasts.2021.137630.

[3] H. Ishwaran, The effect of splitting on random forests, Mach. Learn. 99 (1) (2015) 75–118. doi:10.1007/s10994-014-5451-2.

[4] T. K. Ho, The random subspace method for constructing decision forests, IEEE Trans. Pattern Anal. Mach. Intell. 20 (8) (1998) 832–844. doi:https://doi.org/10.1109/34.709601.