Compare and contrasting IEEE C57.104-2019 on utility transformer fleets.

Zachary H. Draper, Ph.D.
Research & Development
Delta-X Research, Inc.
Victoria, BC, Canada
zhdraper@deltaxresearch.com

James J. Dukarm, Ph.D.
Research & Development
Delta-X Research, Inc.
Victoria, BC, Canada
jim@deltaxresearch.com

***Abstract*—Dissolved gas analysis is an important tool for utilities to watch over a fleet of transformers for abnormalities. It does not provide direct condition monitoring, but can be a cost-effective way to find assets in trouble, while keeping them in service. Recently the IEEE Transformers committee approved a new version of the C57.104 which offers guidelines on how to perform DGA screening of transformers. In this paper we compare and contrast the previous and current version of the C57.104, as well as the IEC 60599-2015, in order to understand the effects on a utility company by adopting the new IEEE standard. A large data set from 4 different utilities are used in combination with the 3 different methods to determine the overall fraction of "abnormal" transformers being flagged, as well as the level of disagreement between the methods. The veracity of the IEEE v2019 will require significant verification steps from utilities because there is a strong shift in which transformers are being flagged. Furthermore, the false-negative rate is estimated based upon a population of failed-in-service transformers and found to be the highest in IEEE v2019 method. Inconsistencies within the IEEE v2019 method are also discussed. Overall, it is suggested that DGA methods should put stronger focus on understanding DGA results from failure cases, rather than finding statistical outliers from a large population.**

***Index Terms*—Transformers, Dissolved Gas Analysis, DGA**

I. INTRODUCTION

Recently the IEEE Transformers committee approved the IEEE Std C57.104-2019 [2], a new guide for the interpretation of dissolved gas analysis in oil-immersed transformers. Adopting a new DGA standard can lead to a dramatic shift in the number and classification of transformers that engineers were previously concerned with. Utility companies will have to adapt to large fractions of their fleet being reclassified as abnormal and therefore requiring attention. The 2019 version is not simply an update to the percentile concentration limits used in the 2008 version. In the 2019 version, there is now a more significant focus on the rates of change to monitor for active gassing events rather than high gas concentrations. Since the IEC 60599-2015 also provides guidance on transformer DGA and is focused more on rates of change in gases, this method is also used to compare against the IEEE methods. There are also different limits to apply based on further qualifications. For example, different limits apply based on the O_2/N_2 ratio to try to compensate for potential gas loss. Furthermore, the age of the transformer will change the concentration limits to compensate for a build-up of gases over time.

This paper seeks to provide context and comparison points on the effects of using either of the two methods on real-world fleets at utility companies, rather than trying to prove the veracity of the methods. Ultimately the degradation of transformer insulating material is a slow process and only through the application of the new standard will we be able to learn how effective this method is at flagging transformers which require more physical testing and/or maintenance.

II. DESCRIPTION OF DGA GUIDELINES

A. IEEE C57.104 v.2008

The broad approach of the IEEE v2008 guidelines is to apply gas concentration limits where progressively higher level limits lead to higher level "DGA Condition Codes", representing the severity of the problem [3]. There is also the inclusion of monitoring the rates of change in the TDCG, or total dissolved combustible gas, to gauge severity and recommendations as well. Limits are established by looking at a large data set of DGA samples and measuring the 90th, 95th and 99th percentiles in order to flag statistical outliers. The implied assumption being that higher gas values represent worse condition.

B. IEEE C57.104 v.2019

In IEEE v2019, the methodology is again identifying statistical outliers with 90th and 95th percentiles in a large database of DGA samples [2]. One of the main changes is a recognition of the fact that DGA analysis is a symptom-based approach and not a condition-based analysis. Therefore the 4 "condition codes" are now 3 levels of "status codes". Furthermore, greater emphasis is applied towards deltas or rates of change on individual gases than in the previous version. This represents an improvement from v2008 because transformers with low gas values would not have been flagged as abnormal, even if they had a high rate of change. By applying rates of change on individual gases v2019 can remove the TDCG-based rates of change. TDCG measurements are sometimes flawed because it is more sensitive to certain fault types. For example, C_2H_2 seen in electrical arcing can go unnoticed at low levels compared to other thermal fault gases, even though it would be alarming to see. Furthermore, TDCG simply sums together gases from

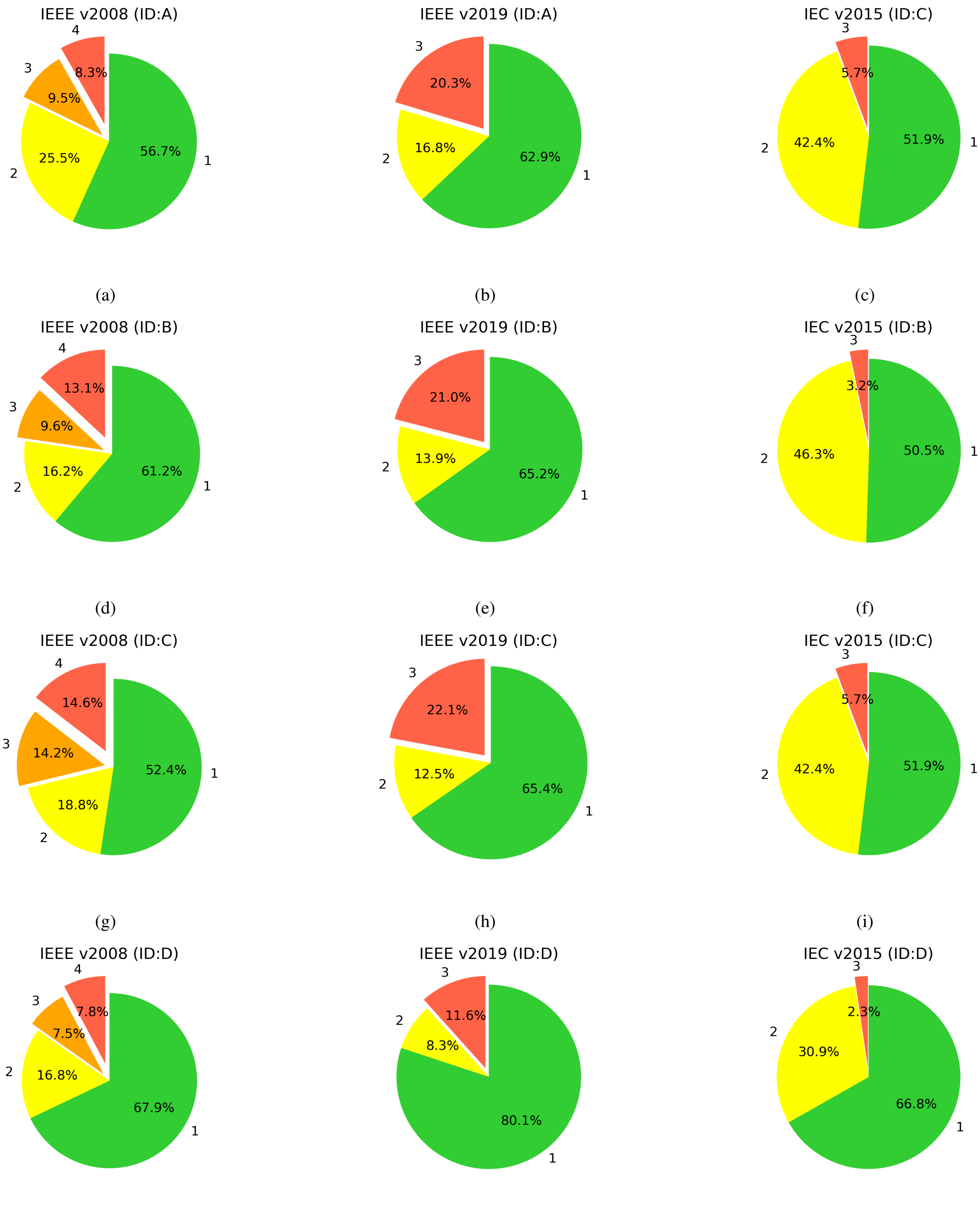


Fig. 1: *Pie charts illustrating what fractions of utility transformer fleets that will be flagged as abnormal by each of the three methods applied to the 4 utility fleets. "Abnormal" is defined in this paper as anything other than a status code 1.*

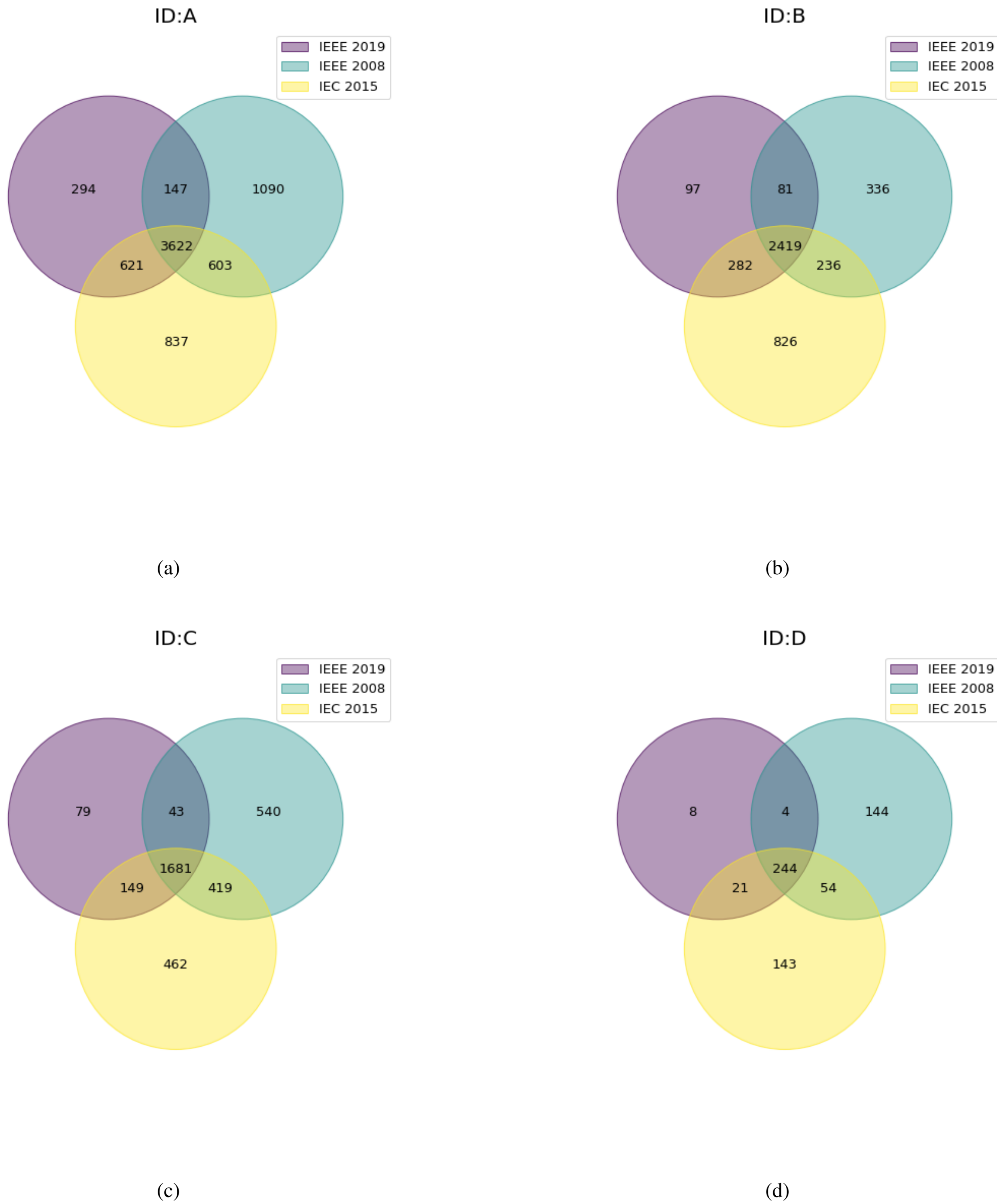


Fig. 2: *Venn diagrams showing the level of agreement and disagreement between various methods on transformers that were flagged as abnormal with a status code greater than 1. When* O_2/N_2 *and age are considered, IEEE v2019 tends to flag the same or fewer units overall, but there is a large discrepancy in which units are actually flagged as abnormal. The veracity of each method will have to be tested by identifying the transformers uniquely flagged by a given method and conducting follow-up investigations to verify abnormalities.*

different insulation systems, carbon monoxide from paper and hydrocarbons from the oil.

C. *IEC 60599 v.2015*

IEC v2015 does not use status codes but has three general states being "Normal", "Alert" and "Alarm". For comparison purposes to IEEE standards we will equate those as status codes 1, 2, and 3, respectively. The primary methodology is more similar to IEEE v2019 than v2008. First you have to exceed an alert level from individual gas concentrations, then only through a rate of change increase can you achieve an alarm state. One exception is in the case of a D2 fault diagnosis where you have significant amounts of C_2H_2. All of the gas concentration and rate of change limits are established from 90th percentiles from a large data set. IEEE v2019 will still flag an asset status code 3 if the concentration levels themselves are above the 95th percentile. In general IEC will flag fewer status code 3 as a result because it is not focusing on higher gas concentrations.

III. Application of the Standards

A. Data used in this study

In this study, DGA data from 4 large North American utility companies was used. The companies operate diverse fleets across North America with subdivisions working in distribution, generation and transmission. The data sets typically have yearly annual sampling over a decade or two, but only the most recent samples will play into their classification by IEEE and IEC guidelines. Utility Source A has 13,034 transformers, Utility B has 14,518 transformers, Utility C has 5,992 transformers and Utility D has 1,449 transformers.

In order to attempt to measure the relative effectiveness of each method at flagging abnormal transformers, a selection of transformers that failed in service at the utility companies involved in this study is used. The total in this set was 127 units. Transformers which failed due to external phenomenon, not predictable by DGA were excluded. The last sample used in the analysis was the sample just prior to failure and not samples taken in a postmortem analysis.

B. Fraction of fleets being flagged

Overall, the 3 methods across the 4 fleets flag about a third of transformers as abnormal. The IEEE v2019 tends to flag fewer assets with the largest discrepancy being Utility D. The overall distribution of status codes within a utility fleet is consistent with Figure A.9 in C57.104 v2019 were the method was employed on real cases [2]. What also becomes clear is the difference in prioritization. IEC has the smallest fraction of status code 3, which means it predominately focusing on units with high rates of change, but also putting many more units in status code 2 with lower on-average concentration limits. What also comes across is that in IEEE v2019, by getting rid of status code 4, it effectively merged them down into status code 3. Therefore prioritizing the worst units becomes harder without establishing higher limits. The IEEE v2019 guidelines also describes what it calls "Extreme" DGA events that will require higher priority, but no concrete way to identify them. It should also be noted that using the correct age and O_2/N_2 ratio is crucial for setting the limits in IEEE v2019. By conservatively using limits based on young, high O_2/N_2 limits, the results will drastically increase the number of assets that are flagged. Therefore the typical breathing type and age of a utility fleet will become extremely important to use the IEEE v2019 correctly.

C. Disagreement between methods

In Fig. 2, it can be seen that even if the overall levels are roughly similar, a large sea change occurs in which units are actually being flagged as abnormal. For example in Utility A, IEEE v2008 and v2019 methodology agree that 3,769 (3622 + 147) of the transformers in the fleet are abnormal in some way based on their DGA data. However, the methods also disagree on 2,608 of the transformers (294 + 621 + 1090 + 603) as being normal or abnormal. In each of the 4 utilities the number of disagreements is comparable to the number of agreements. This will represent a significant shift in the assets each utility company would be keeping track of. Since IEC is agnostic of air exposure or age, it similarly flags different assets despite being rooted in rates of change on individual gases like the new IEEE v2019 guidelines. Depending on the age or O_2/N_2 ratio the IEC method will either be more or less conservative than the IEEE v2019. Greater consideration of the equipment properties will be needed to understand the differences in the DGA status between those methods.

For transformers that are uniquely identified by IEEE v2019 the most common cases are the high O_2/N_2 ratio, which represent the lower bar on concentration and rate of change limits. For units uniquely flagged by IEC the most common cases are in the low O_2/N_2 set. This is consistent with the idea that low oxygen limits in IEEE v2019 tend to be higher than the IEC limits, which do not distinguish base on the oxygen content and presumed gas loss.

D. Estimated False-Negative Rates

One way to test the veracity of any given testing method is to estimate its false-negative rate. The false-negative rate is the chance that an applied test will return a negative result for something that we know in fact was positive. One way to approach this with transformers is to look at failure cases, apply the DGA methods, and see how many of the failed transformers were not flagged as "abnormal" before failure. To do this we have to limit the transformer failures to cases where the failure could, in principle, be flagged by DGA as a result of some internal fault. Transformer failures from vandalism, wildlife, etc. have to be ignored.

In Figure 3, a set of transformer failures from the utilities used in this study was run through each of the 3 methods. The results show that IEEE v2019 method has the highest false-negative rate at $\sim$45% compared to the lowest being the IEC method which fails to flag the failures $\sim$34% of the time. However, IEC flags a higher proportion of the overall fleet (See Figure 1), so the odds that it will flag the failures by chance are higher. Therefore understanding the false-positive rates for each method is also very important to consider when comparing the methods. It is much harder to estimate a false-positive rate since the outcome of a transformer flagged as abnormal can either fail or not fail in the near future. A balance between minimizing false-positives and false-negatives will be important for any DGA method to be efficient. False-positives will represent more effort for the utility to investigate and false-negatives will represent a failure that was not on the utilities radar as an issue.

It can also be seen that the IEEE v2008 method puts more failures in status code 2, than status code 3. This suggests that prioritization based on gas concentration levels is not always appropriate. For example, gas loss either on purpose or by accident may underestimate the severity of the faults. If your DGA method is based primarily around gas concentration limits on a sample, rather than rates of change (like IEEE v2019 or IEC), you may not get an accurate prioritization of the “abnormal” transformers that require attention.

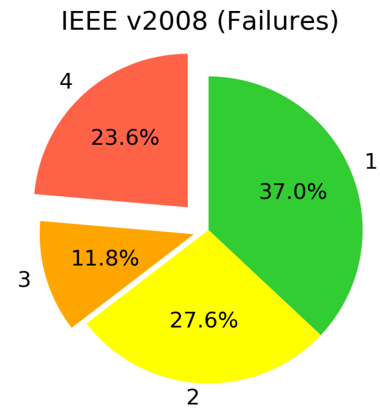
In Figure 4, the agreement and disagreement between the methods can also be determined. In the cases where transformers were flagged as abnormal by one of the 3 methods, each method uniquely identifies a subset of the failed transformers as abnormal. This would represent unique saves by using this method. The sacrifice is in the set of transformers that are not flagged as abnormal by choosing that method over the others. This represents the number of unique losses. If we subtract unique saves from the unique losses we get a relative number of transformers that were potentially lost by adopting one method over another. This is just another way to represent the relative outcome of the false-negative rate in terms of absolute number of failed transformers by choosing one only of the methods. IEEE v2019 has the highest at 26, IEEE v2008 at 10, and IEC at 5. Ideally, the best method would have the fewest relative, unique losses to saves.

IV. Conclusions

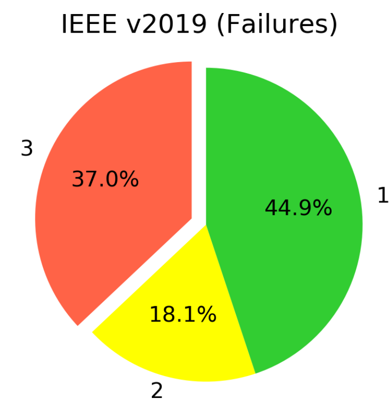
A. Inconsistencies with the IEEE v2019 guidelines.

The IEEE C57.104 v.2019 has a few inconsistencies, which users should be cautious about before blindly applying limits taken from the tables and not reading the guidelines in detail. In general, the change in limits based on different age ranges allows for lower limits on younger transformers and higher gas concentrations for older transformers. An explanation might be that older transformers have built up the gases over time through some minor heating or aging effects and not necessarily warrant a higher status code for the transformer. However, in Table 2 under the $O_2/N_2 > 0.2$ category, the methane concentration limit for younger transformers is higher than it is for older transformers. This would create a greater allowance for low temperature faults in younger transformers than older ones, which is inconsistent with the general idea that the gases will gradually increase over time. It may in fact be the case that transformers in the IEEE data set with high oxygen content at older ages have less methane, but that may be a quirk of the data through some cross-correlated factor, rather than being rooted in a causal mechanism and logic.

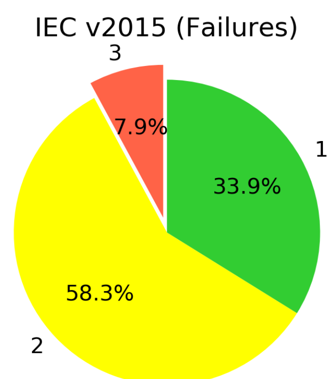
An even more pressing inconsistency is the fact that the concentration limits in Table 2 for C_2H_2, where $O_2/N_2 > 0.2$, actually increases from 1-2 ppm up to 4-7 ppm. To quote the C57.104-2019 in section 5.4, ”The O_2/N_2 ratio was proposed for evaluation as a proxy for distinguishing sealed



(a) Failures: IEEE v2008



(b) Failures: IEEE v2019



(c) Failures: IEC v2015

Fig. 3: *The fraction of transformers flagged as abnormal by each method in the population of transformers that failed-in-service. Units that failed but are given a status code 1 for “normal” would represent the rate of false-negatives. IEC appears to have the lowest false-negative rate, but may also have a higher false-positive rate since it flags more units overall. For IEEE v2008 the, higher number of status code 2 to status code 3, suggests that the prioritization due to higher gas concentrations is not an optimal way of prioritizing compared to gassing rates of change or deltas. Ideally units flagged with a higher status code should be more likely to fail, not less.*

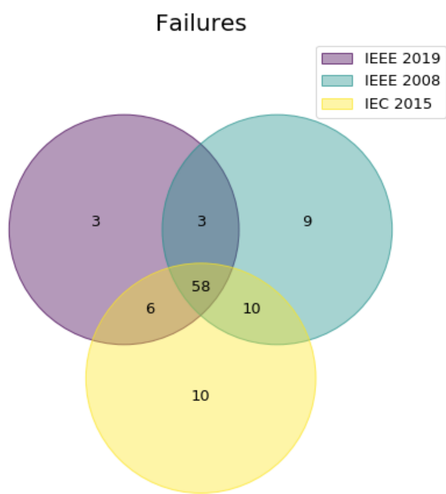


Fig. 4: *Venn diagram showing the relative number for transformers flagged as "abnormal" by each method. The number of unique saves vs unique losses by adopting only one of theses DGA strategies for flagging transformers can be compared. IEEE v2019 for example will uniquely save 3 transformers, while uniquely missing 29 (9+10+10), by not employing the other DGA screening methods. This means at the end of the day you would have missed flagging about 26 transformer failures relative to adopting the other methods. The other methods would have lost by comparison 5 transformers for IEC and 10 transformers for IEEE v2008.*

units from free-breathing ones". While not an exact 1:1 with breathing type, the reason to have lower gas concentration limits for free-breathing transformers more broadly is because some of the gases created by the fault will be lost to the atmosphere [4, 1]. Therefore an air-breathing transformer can have just as serious a fault as a sealed transformer, while having lower gas concentrations in the oil. The tables in IEEE v2019 compensate for this by decreasing the concentration limits to apply. The one notable exception is C_2H_2, despite how concerning it is to see in transformer DGA because it can be indicative of electrical arcing. Again this may be true based on the data set and methods employed by IEEE, but does not fit with a consistent logic on air exposure. This could be dangerous by allowing for a higher tolerance of electrical arcing in free-breathing transformers when the opposite would be true.

Another thing to be aware of is the fact that the rates of change measurements are based on 3-6 data points in a 4-month to 2-year window. Therefore if you are only doing routine yearly sampling then you would not apply the rate of change limits. You would apply the rates of change only in cases where you have investigative sampling of at least 3 points in a 4- to 24-month period, after having fit a linear regression to all of the data. If the investigative sampling goes on for a long period of time, then the limits on the rates of change will actually decrease after 9 months. This might reduce false positives from measurement noise by requiring more samples for more stringent limits, but may cause the status code to flip as you continue sampling, but the gassing rate remains static.

B. *Overall changes.*

DGA is only a screening method to broadly identify and prioritize transformers that are "abnormal" through gas production generated by fault in the transformer. The gases themselves generally do not fail the transformer, but the faults generating the gas may. Only though physical testing, contemporaneous data sets, and inspections can the condition of the transformer be determined. The veracity of any DGA method will require more testing to be done in order to verify that the method is accurately flagging transformers as abnormal. IEEE v2019 may or may not prove to be successful in time. What is clear is that a significant shift and undertaking will be required to actually validate one method over the other given the sea change in the underlying methodology. It is crucially important that any DGA method try to optimize the number of potential saves to losses. This requires DGA interpretation methods to be more strongly rooted in understanding actual failure cases, rather than assuming high gas levels represent a worse status and finding statistical outliers in a population. To that end, measuring and comparing the false-negative rates for DGA methods could be a good way to optimize a DGA strategy.

References

[1] "DGA in Non-Mineral Oils and Load Tap Changers and Improved DGA Diagnosis Criteria". In: *CIGRE Technical Brochures* (2010).

[2] "IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers". In: *IEEE Std C57.104-2019 (Revision of IEEE Std C57.104-2008)* (Nov. 2019), pp. 1–98. ISSN: null. DOI: 10.1109/{IEEESTD}.2019.8890040.

[3] "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers". In: *IEEE Std C57.104-2008 (Revision of IEEE Std C57.104-1991)* (Feb. 2009), pp. 1–36. DOI: 10.1109/IEEESTD.2009.4776518.

[4] *Mineral oil-filled electrical equipment in service – Guidance on the interpretation of dissolved and free gases analysis*. 3.0. IEC 60599-2015-09. International Electrotechnical Commission, Sept. 2015.