Fault detection and prediction in Smart Grids

Christian Andre Andresen
Energy Systems
SINTEF Energy Research
Trondheim, Norway
christian.andresen@sintef.no

Bendik Nybakk Torsæter
Energy Systems
SINTEF Energy Research
Trondheim, Norway
bendik.torsater@sintef.no

Hallvar Haugdal
Department of Electric
Power Engineering
NTNU
Trondheim, Norway
hallvar.haugdal@ntnu.no

Kjetil Uhlen
Department of Electric Power
Engineering
NTNU
Trondheim, Norway
kjetil.uhlen@ntnu.no

***Abstract*—Modern society is to a larger and larger extent dependant on electric energy, and hence the reliance on and utilization of the electric grid is increasing steadily. At the same time the production and consumption patterns are changing from large centralized generation of electric power and pure consumers to distributed generation (DG) and more complex consumers. This transition causes higher stress on an aging infrastructure and major investments are required over the coming years to maintain a reliable supply of electric energy. Better monitoring solutions and predictive methods can increase the possible utilization of the existing grid and reduce the fault frequency. This paper presents some current challenges in the grid and a possible monitoring solution and fault prediction method. This is exemplified with statistics and field-measurements from the Norwegian power grid.**

Keywords—PMU, PQA, fault prediction, WAMS, statistical learning

I. INTRODUCTION

The use of electric power has been increasing over the last decades and is expected to further increase globally in both the short and long-term future [1]. At the same time the grid infrastructure in Europe is aging and changing towards a Smart Grid architecture [2], where there is more flexibility in the grid to meet varying demand. On the production side a substantial amount of intermittent energy sources such as PV and wind has been introduced into the energy mix, to some degree replacing controllable generation such as coal and nuclear-powered plants. On the consumption side the consumers have become more complex in their behaviour. Many modern electric appliances, such as induction heaters and chargers for electric vehicles (EV), have a more challenging consumption pattern than traditional appliances, which causes higher stress on the distribution grid. In addition, some end-users, previously considered pure consumers, have installed local generation capability and are at times net producers of electric energy termed *prosumers*.

As a consequence of these trends traditional generation methods that included the use of large rotating mass, such as large turbine generators, are increasingly being replaced by power electronics that digitally control the frequency and voltage levels. This development gives less robustness in the overall electric system towards faults and disturbances. Without the stabilizing nature of rotating masses cascading faults on a system wide level may become more prevalent in the future. In the transition towards a more flexible and adaptable Smart Grid, it is essential that such challenges are handled. Enabling the grid to meet the resulting change in load profile and the increased variability in demand is a massive challenge on a European level.

However, there is a parallel development that is helping to provide the means to meet this challenge. The development in instrumentation, communication and data analysis has been significant, and novel solutions for monitoring and managing the grid is becoming reality [2, 3]. Of special interest to this paper is the development in monitoring instrumentation, such as Power Quality Analysers (PQA) and Phasor Measurement Units (PMU). In combination with the development of machine learning (ML) techniques, it will be possible to extract and efficiently analyse the vast amounts of data generated from these devices. This article argues that modern machine learning algorithms may utilize the gathered data to predict or give early warnings about faults and instabilities in the grid both on a system-wide level, such as frequency oscillations between generators, and on the local level, such as impeding component and line failure. To do so, the faults and instabilities needs a preceding signature to manifest itself in the measurements before the operational ability of the grid is affected. This paper seeks to demonstrate how a sufficiently large dataset may be collected for such signatures to be detected and used to give effective decision support for the grid operators.

It is beyond the scope of this paper to give a comprehensive overview of challenges in the smart grid field. This is an active area of research and this paper seeks to give some insight into a possible route with special focus on the Norwegian grid as an example case.

II. FAULT DETECTION AND MITIGATION

The last few years the amount of measuring instruments and sensors in the power grid has increased significantly, through PQAs, PMUs, smart meters (AMI – Advanced Metering Infrastructure) and other types of measurements (temperature, wind, humidity, etc.) and sensors (line angle, vibration, etc.). The increased amount of measurement data from the power grid, in combination with new methods within machine learning, gives new possibilities in terms of fault detection and mitigation. By a system-wide deployment of PQA and PMU devices the grid may be monitored in real-time via an efficient communication network. This enables the continuous evaluation of the current state of the grid, indicating congestions, frequency oscillations and overall load distribution. It also enables the real-time application of predictive algorithms for fault probability assessments that have the potential of reducing the fault frequency if appropriate mitigating actions can be found.

A. Fault prediction and system requirements

The power grid is exposed to many possible types of disturbances and faults, and all of them have different signatures and propagations. In order for a system to be able to detect and predict a future fault, the fault needs to have a signature in advance of the decisive fault event. It is assumed that many of the most frequent fault events in the power grid develop over time and have a fault development signature ahead of the actual malfunction such as investigated in [4]. As an example, an insulator failure leading to an earth fault or short circuit might have discharges in advance that could be detected using measuring instruments.

There is some uncertainty related to whether a prewarning ahead of a fault makes it possible to mitigate and avoid a crucial fault event. The uncertainty is related to how much time that is needed to initiate preventive measures, and what kind of action that is needed in order to avoid the potential fault event and customer outage. According to Norwegian grid companies, a prewarning of at least a few minutes is needed for a grid operator to manually do a preventive network configuration, e.g. disconnecting a faulty line. If the fault can't be mitigated from the control room, a prewarning of minimum a few hours is needed for a technician to solve the problem. In the above paradigm it is thought that a predictive algorithm would give decision support to a human operator, and that the operator would have to make a decision and initiate an action manually. The application of automatic corrective actions initiated by the system itself is emerging. In such a paradigm, where a machine makes alterations to the grid to mitigate faults and customers outages, a warning horizon of seconds or milliseconds may suffice.

The fault prewarning that is received in the operation centre must be reliable in order for the warning system to be trusted by the grid operator. In addition, the system should contain detailed information on the type of fault, whether the fault is imminent and how critical the fault potentially could be for grid customers.

B. Predictable faults and disturbances

Traditionally, monitoring based on measuring instruments in the grid has been sparse. Due to this, existing research on the power quality signature of fault events and disturbances is limited. However, some types of faults and disturbances are more likely to have a signature in advance of the decisive fault event than others. Problems related to component failure, caused by humidity, salt, ageing etc., generally have a development over time. It is therefore assumed that it should be possible to detect such faults prior to their occurrence using power quality measurements. In the Norwegian grid, faults and disturbances related to extreme weather conditions are common. Such faults can be earth faults and short-circuits caused by wind (vegetation) and icing. Even though such faults normally don't have a signature in advance of the decisive fault event, it could be possible to anticipate such faults and quantify the probability of their occurrence by analysing power quality measurements in combination with weather data.

While it is assumed that some faults and disturbances can be prewarned, other fault events are unfeasible or prohibitively hard to include in a system for fault prediction and mitigation. Such faults are related to lightning strikes, human errors, animals etc.

Transformer faults have been pointed out by operators as a motivation for deploying an early warning system. Severe transformer faults could cause long-lasting outages due to lengthy and complicated repairs, delivery and transport, and occurs frequently enough to represent a significant cost [5]. Reference [6] considers a transformer subject to mechanical degradation and quantifies the resulting introduction of harmonics in the current and voltage waveforms. This is a developing signature that could potentially be detected by a ML-based early warning system.

Earth faults in the medium-voltage grid can cause dangerous situations and disturbances if they are not handled efficiently. As an example, intermittent earth faults caused by e.g. vegetation and weather (snow, ice, wind etc.) can be difficult to detect and avoid. In Figure 1, an intermittent earth fault in the medium-voltage grid is shown. The figure shows that the earth fault is unstable for a period of more than 12 hours before it is cleared. Such an intermittent earth fault could be detected and handled at an earlier stage using measurement data, which in turn could make fault detection and location more reliable and reduce the danger for grid customers.

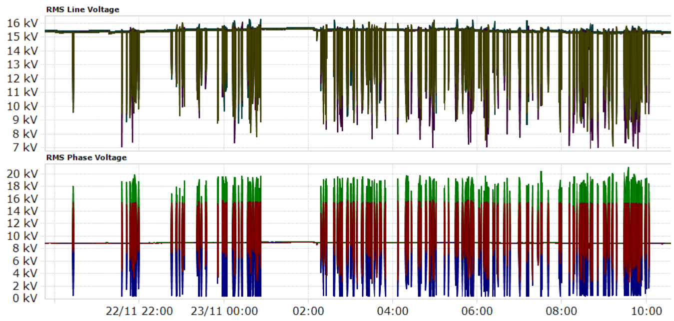

Figure 1: RMS voltage in grid with intermittent earth fault

III. MEASUREMENT INFRASTRUCTURE

Real-time monitoring schemes requires high-resolution measurements that are reported with a low time delay (latency) to a centralized computing unit. Two relatively extensively deployed measurement infrastructures in the Norwegian grid are based on PQAs and PMUs, which both answer relatively well to the demands of real-time monitoring. The characteristics of these are outlined in the following, along with examples on what information can be retrieved.

A. Power Quality Analysers

Power Quality Analysers, also known as PQAs, are commonly used to analyse the power quality in the grid.

Traditionally, PQAs have been installed in important junctions in the grid, such as the bus bar on the secondary side of a HV/MV transformer. PQAs can give valuable information on all voltage quality parameters, e.g. voltage variation, transients and harmonic distortion. Compared to PMUs, PQAs are superior when it comes to sampling rate. Some PQA devices have a bandwidth of up to 25 kHz and higher and are able to efficiently reproduce and store continuous waveforms for large periods of time (multiple years without generating extreme data volumes). This is achieved by compression/decompression of signals when storing/reading. The extent to which this has a negative impact on time delay must be kept in mind when designing a real-time application, since fast-acting systems might be crucial to achieve rapid alert and remedial action for certain types of faults.

In Figure 2, the RMS and waveform of a line voltage with high harmonic distortion is depicted in the same pane. When analysing the RMS voltage, it appears that the voltage is stable and without significant noise. However, when the waveform of the voltage is analysed in detail, it is apparent that the voltage waveform is highly distorted. Thus, if analysing only the RMS voltage in this case, valuable information would be lost. This clearly shows the advantages of a PQA compared to a PMU, as the higher bandwidth makes it possible to analyse voltage quality on a much higher resolution than 50 Hz RMS.

Considering the degradation of a transformer mentioned in Section II.B, monitoring the RMS would not yield sufficient information to be able to detect a change in harmonic distortion. For this type of signature, a higher resolution on the level of that provided by a PQA would be required.

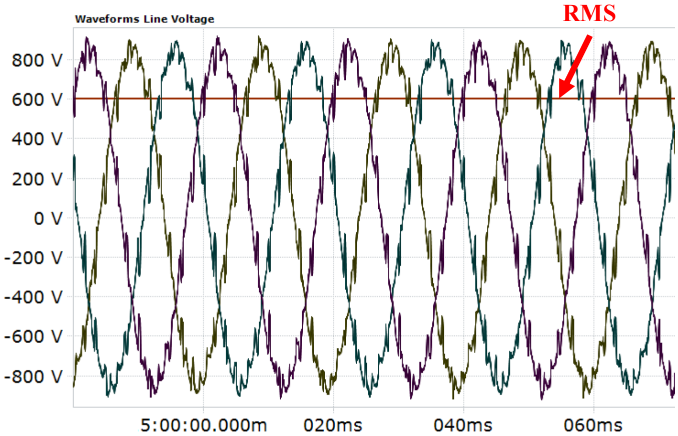

Figure 2: RMS and waveform of line voltage with high harmonic distortion

An increasing instrumentation in the power grid also provides a better fundament for wide area monitoring systems (WAMS) [7]. Both DSOs and TSOs invest in a higher number of high-quality measuring instruments, which can be used to anticipate and detect faults and disturbances, and to find their location in the grid. In Figure 3, an example of wide area monitoring using PQAs is presented. The figure displays the measured line voltage in three different locations, using three different PQAs. The PQAs are placed in three different MV grids, but they are all in the same geographic region.

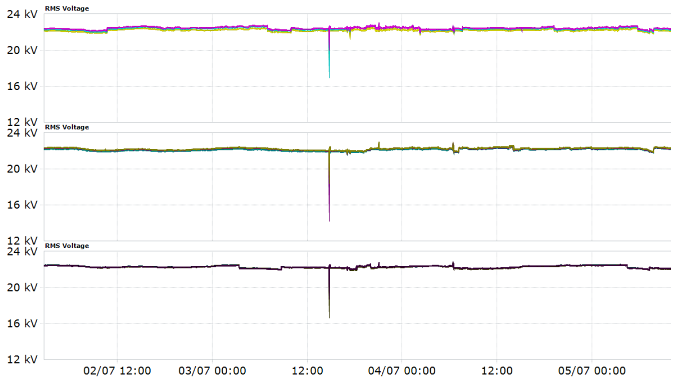

Figure 3: WAMS using PQAs in three different locations in the same geographic region

During the measurement period displayed in the figure, a voltage dip occurs in the grid. As can be seen from the figure, the voltage dip has different amplitude in the three locations. It is apparent that the voltage dip influences the voltage level in the middle pane more than in the two others. Using this information, it is reasonable to assume that the PQA in the middle pane has a higher electrical proximity to the root cause of the voltage dip than the two other PQAs.

B. Phasor Measurement Units

PMUs (as standardized in [8] and [9]) provide time-synchronized voltage and current phasors at a reporting rate in the range of 10-60 Hz, usually 50 Hz in the Norwegian grid. Each reported frame contains phasors constituted by an angle and a magnitude, which corresponds to the phase and amplitude of the measured quantity, usually three phase voltages and currents. Due to the highly accurate time-synchronization, which is often achieved by using GPS-signals, providing an accuracy in the range of milliseconds, the angles of phasors situated at widely separated geographical locations can be compared with high accuracy. PMU measurements therefore gives an advantage in the wide area perspective, making this infrastructure suitable for monitoring the operational state and stability of the power system at the transmission level.

PMUs report measurements to Phasor Data Concentrators, which transmit the data further to other units, for instance a computer in a control centre. The transfer protocol is very efficient, resulting in a low latency in the range of milliseconds [9]. PMUs are therefore suitable for real-time applications, and monitoring applications based on PMU measurements are already used extensively in power systems around the world [10].

Figure 4 and Figure 5 indicate how information and knowledge of a disturbance can be retrieved from PMU measurements recorded at three widely separated locations. The figures show frequency and power flow, respectively, in the different locations during a loss of load.

The increasing frequency in all locations indicates surplus production of power, i.e. increased generation or loss of load. Looking at the power flows in Figure 5, which flow from Location 2 to Location 1, it can be deduced that a loss of load probably occurred near Location 1. Further, looking closer at the frequencies in the three locations, it is observed that the frequency in Location 1 rises quickly before the other frequencies. This also indicates that the loss of load occurs close to Location 1.

The trained expert would easily be able to perform the above reasoning by studying measurements retrospectively. However, if such knowledge could instead be retrieved by an online ML-based algorithm, the operators could potentially be notified almost immediately about the root cause of the recorded deviations.

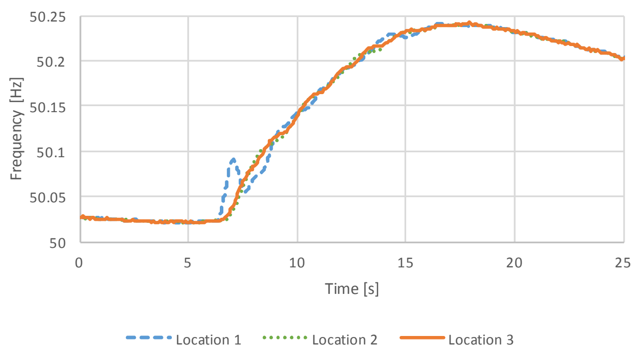

Figure 4: Frequency excursion measured by PMUs at three locations approximately 500 km apart. A loss of load occurs at Location 1, causing the frequency to rise rapidly in this area, before Locations 2 and 3 starts following.

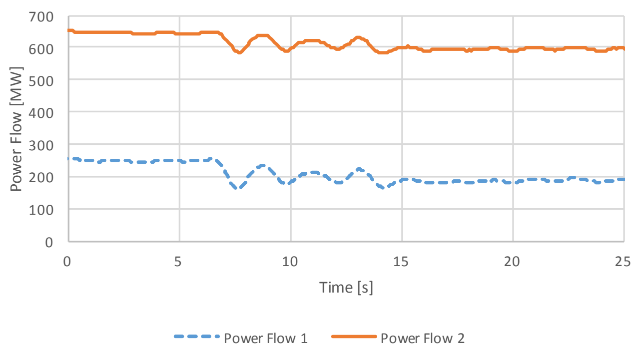

Figure 5: Power flows (derived from PMU measurements) from Location 2 in the direction of Location 1, where a loss of load occurred.

A benefit of wide area information is that it provides for improved situational awareness, helping operators to quickly understand the criticality of disturbances and take correct actions. Looking further, it is envisioned that WAMS will be used in automatic control systems to perform very fast and targeted remedial actions. Actions including fast ramping, or shedding of generation or loads, will then be performed without the operators' intervention, thus greatly reducing the risk of blackouts. Solutions for system protection or control schemes that utilize synchronized phasor measurements are commonly referred to as wide area protection (WAPS) or wide area control systems (WACS) or combined WAMPAC [11].

IV. Measurement Resolution Aspects

The requirements for fault logging with respect to PQA and PMU data varies dramatically. It is not always immediately clear what is the required data quality needed for event identification. However, modern power quality analysers enable logging at least as rapidly as 1024 samples per cycle using standard commercially available products. Logging strategies also differ between event-based logging and continuous logging. Event-based logging enables high-resolution sampling of parameters such as voltage and current in the immediate time interval surrounding some event, e.g. a voltage transient. Such logging often does not allow the determination of the development of the fault, and root causes can be harder to identify lacking such information. In addition, continuous time series allows for a much richer analysis of the power system enabling a broader range of model-building and statistical learning.

In the following, an earth fault is identified in a 22 kV line in the Norwegian power grid. Figure 6 illustrates what information would be available for analysis of such a fault with different logging strategies. In the top pane cycle-by-cycle RMS values of the phase voltages are shown. The pane shows that the magnitude of the first event (at about 4800 seconds) is taken to be of approximately the same magnitude as the main earth fault (at about 7000 seconds). In the middle pane, 1-second averages are shown. In this case, the preceding event is displayed to be of significant lower magnitude than the later earth fault. In the bottom pane, 1-minute averaged RMS values are shown. In this case, it is impossible to see that there was an earth fault in the grid at all. It should be noted that if max/min values for each sample period for each resolution were plotted, the information would be more complete even for the larger time-period averages. It is assumed that higher time resolution is more beneficial for the application of algorithms for fault prediction.

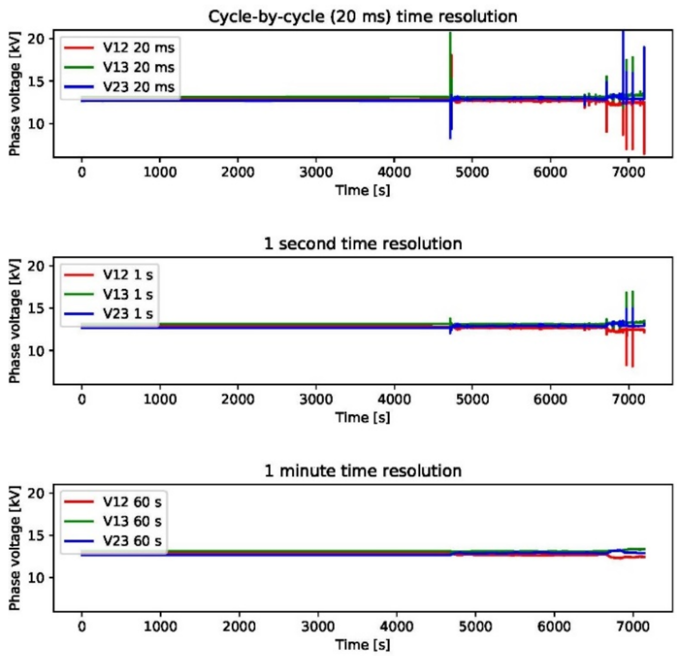

Figure 6: An earth fault in the Norwegian 22 kV grid displayed with three levels of resolution. In the top pane cycle-by-cycle RMS values are shown, in

the middle pane 1-second RMS averages are shown and in the bottom pane 1-minute RMS values are shown.

V. GRID FAULT STATISTICS

Faults and disturbances in the power grid are recorded by several actors, such as end-users, DSOs and TSOs. In the Norwegian power grid there is a mandatory responsibility given to the DSOs and the TSO to report faults and disturbances via a national reporting structure called FASIT [12]. The report must specify the type of incident, the time and duration, the number of affected phases, the voltage level and plausible root cause and contributing causes. This reporting gives a statistical basis for national reports on the frequency of faults and disturbances in the Norwegian grid both regarding voltage levels and causes. Every year the Norwegian TSO Statnett publishes statistics regarding faults in the 1-22 kV grid and in the 33-420 kV grid [5], and the Norwegian governmental directorate NVE publishes statistics of end-user power quality and disruptions [13]. In Figure 7 and Figure 8 respectively, the number of events reported and the resulting energy not delivered is shown for the 1-22 kV grid and for the 33-420 kV grid. Data is publicly available and collected from Statnett's web pages [9]. As can be seen, the number of events and consequences is far larger in the 1-22 kV grid. Vegetation and wind is two of the major root causes [5]. Both these categories lead to faults and disturbances where a significant portion is expected to be of a developing nature where there are indications of such events before the critical fault occurs. By analysing historical PQA and PMU data in combination with the recorded faults that have both time, location and cause information, it may be possible to apply modern machine learning algorithms to enable prediction of the occurrence of such faults as discussed below.

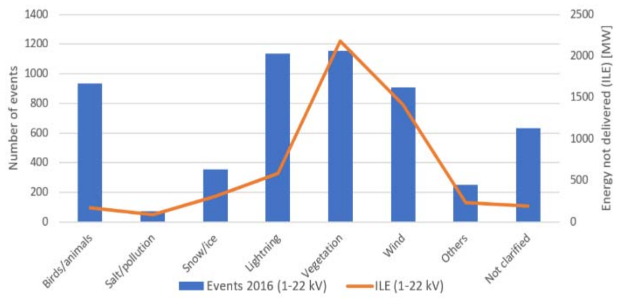

Figure 7: Number of events reported in the Norwegian 1-22 kV grid in 2016 and the resulting energy not delivered (ILE) as function of root cause.

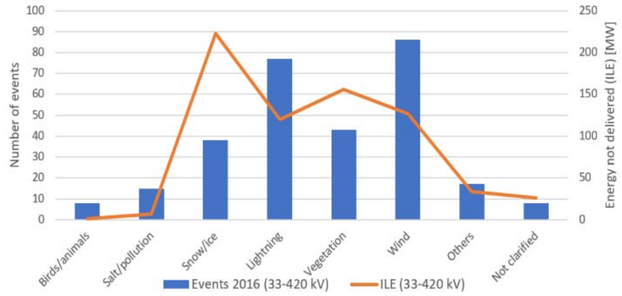

Figure 8: Number of events reported in the Norwegian 33-420 kV grid in 2016 and the resulting energy not delivered (ILE) as function of root cause.

VI. DISCUSSION

As has been demonstrated in the sections above, there is a significant development in the available grid monitoring data alongside an expectation that the loading profile of the grid will become more challenging for the operators in the future. This development goes alongside an unprecedented evolution in the fields of big-data analysis and machine learning, such as deep neural networks [14] [15]. These developments are made possible by better computational resources (such as HPC and GPU utilization), better machine learning algorithms and more accessible libraries making the application of the methods considerably easier and more efficient for researchers in fields outside computer science and mathematics.

There are several initiatives seeking to exploit modern machine learning methods to both detect and predict faults in the grid, such as GRIP [16], EarlyWarn [17] and others. One of the central challenges in the machine learning domain is the large amount of training data that is required to train the algorithms. In traditional grid fault analysis, experts study each fault occurrence separately and seek the root cause and/or mitigating actions to be implemented. There is a limit to the number of cases that can be studied by a person or team in this manner. In addition, manual analysis is not suitable for real-time applications. The training of machine learning algorithms typically applies thousands or millions of examples before robust functionality is achieved. Each example needs to be labelled (annotated) with the desired outcome of the algorithm for that particular example. As an example, algorithms designed to do picture recognition have utilized millions of pictures that have been manually labelled with their class ("cat", "dog", "horse" etc.) for the algorithms to be trained to the level where they can automatically distinguish between different classes of images.

For the utilization of machine learning towards fault classification and prediction, a similar dataset of labelled input will be needed. The above fault reporting structure combined with the continuous logging of the PQA and PMU units described enables the automatic generation of such training material. Since high-resolution time series are becoming available for periods over years for a large number of locations, training datasets can be generated with the required time resolution as needed by the algorithm. Manually extracting fault statistics for thousands or millions of fault occurrences would require enormous resources and raise data-right issues. There are however a number of methods to automatically detect and classify faults in PQA and PMU data series [3]. By running such classification algorithms over continuously collected time-series data for an extensive grid area (such as the Norwegian grid) over a prolonged period of time, a sufficient number of events can be collected to facilitate the application of machine learning methods such as deep neural networks. It is important to note that time series of "non-events" (time series where there are no faults) also needs to be collected for the algorithms to learn to accept normal variations in the field data. The above-reported statistics indicate the number of events that can be expected to be found in time series from the Norwegian grid. The data suggests that there are a few thousand events per year, and continuous monitoring has been conducted for more than a decade at some locations. Keeping in mind that a balanced dataset of time series with and without faults should be collected, a training set of tens of thousands of samples should be feasible. It remains to be established what time resolution and duration that is needed for robust fault detection and prediction.

A distinction should be drawn between the task of fault *identification* and fault *prediction*. In the former case of identification, the task is to quickly determine that a fault has occurred, where it has occurred and what type of fault that has occurred. This information then needs to be communicated to an operator or system that takes appropriate action. Time is of significant importance in the execution of this task. Traditionally some instrumentation has had batching of data before communicating it to a central where any fault detecting activities would be carried out. This batching could give a latency of 5 minutes or more in commonly found industrial applications. Recent improvements in communication could potentially bring this latency down to milliseconds. Employing rapid algorithms for fault identification could enable the location and classification in near real-time. In the case of prediction, the time scale of importance varies according to the fault to be predicted, as discussed above. However, the same argument regarding low latency data communication and rapid algorithms holds for this case. Machine learning algorithms have demonstrated very rapid processing times once they have been trained properly, and it is likely that they will contribute to bringing rapid identification and prediction to the power grid as well.

VII. Conclusion

This paper presents the foreseen challenges arising the in the operation of the grid, and some instrumentation and monitoring solutions that may help in the operation of the grid. The increasing availability of long time series of high quality monitoring data with low latency combined with advances in big-data analysis and machine learning gives rise to efficient fault identification and prediction. Developments in this research area is expected to enable a more robust and efficient utilization of the grid in the transition towards a smarter grid.

Acknowledgment

The authors would like to thank the Research Council of Norway and industry partners for the support in writing this paper under project 268193/E20 EarlyWarn.

VIII. References

[1] IEA, “World Energy Outlook 2017,” IEA, 2017.

[2] N. IqtiyaniIlham, M. Hasanuzzaman and M. Hosenuzzamana, “European smart grid prospects, policies, and challenges,” *Renewable and Sustainable Energy Reviews,* vol. 67, pp. 776-790, 2016.

[3] S. M. Muyeen and S. Rahman, Communication, Control and Security Challenges for the Smart Grid, Institution of Engineering and Technology, 2017.

[4] B. D. Russell, R. M. Cheney, T. L. Anthony, C. L. Brenner, C. F. Wallis and W. E. Muston, “Reliability improvement of distribution feeders through real-time, intelligent monitoring,” in *Power & Energy Society General Meeting, 2009. PES '09. IEEE*, Calgary, 2009.

[5] Statnett, “Statnett feilstatistikk,” Statnett, [Online]. Available: http://www.statnett.no/Kraftsystemet/Nedlastingssenter/Feilstatistikk/. [Accessed 14 May 2018].

[6] A. S. Masoum, N. Hashemnia, A. Abu-Siada, M. A. S. Masoum and S. M. Islam, “Online transformer internal fault detection based on instantaneous voltage and current meassurements considering impact of harmonics,” *IEEE Trans. Power Deliv.,* vol. 32, no. 2, pp. 587-598, 2017.

[7] A. Chakrabortty and P. P. Khargonekar, “Introduction to Wide-Area Control of Power Systems,” in *2013 AMERICAN CONTROL CONFERENCE (ACC)*, Washington, DC, 2013.

[8] IEEE, *IEEE standard for synchrophasor meassurements for power systems, IEEE Standard C37.118.1-2011,* 2011.

[9] IEEE, *IEEE standard for synchrophasor data transfer for power systems, IEEE Standard C37.118.2-2011,* 2011.

[10] U. ANNAKAGE, A. RAJAPAKSE, B. BHARGAVA, N. R. CHAUDHURI, S. RIBEIRO CAMPOS ANDRADE, V. PATHIRANA, K. V. KATSAROS, A. CHAKRABORTTY and V. TERZIJA, “Application of Phasor Measurement Units for Monitoring Power System Dynamic Performance,” in *Cigré*, Paris, 2017.

[11] V. Terzija, G. Valverde, D. Y. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic and A. Phadke, “Wide-Area Monitoring, Protection, and Control of Future Electric Power Networks,” *PROCEEDINGS OF THE IEEE,* vol. 99, no. 1, pp. 80-93, 2011.

[12] FASIT, “FASIT - Reporting structure for faults and disturbances in the Norwegian power grid,” [Online]. Available: www.fasit.no. [Accessed 2018 2018 2018].

[13] NVE, “NVE - Avbruddsstatistikk,” NVE, [Online]. Available: https://www.nve.no/reguleringsmyndigheten-for-energi-rme-marked-og-monopol/nettjenester/leveringskvalitet/leveringspalitelighet/avbruddsstatistikk/. [Accessed 14 May 2018].

[14] Y. LeCunn, Y. Bengio and G. Hinton, “Deep Learning,” *Nature,* vol. 521, pp. 436-44, 2015.

[15] F. Chollet, Deep learning with Python, New York: Manning Publications Co., 2018.

[16] SLAC National Accelerator Laborotory, [Online]. Available: https://www6.slac.stanford.edu/news/2017-09-14-slac-led-project-will-use-artificial-intelligence-prevent-or-minimize-electric-grid. [Accessed 14 May 2018].

[17] S. E. AS, “EarlyWarn,” [Online]. Available: https://www.sintef.no/en/projects/earlywarn/. [Accessed 14 May 2018].