24th International Conference & Exhibition on Electricity Distribution (CIRED)

12-15 June 2017

Session 3: Operation, control and protection

Risk prediction in distribution networks based on the relation between weather and (underground) component failure

ISSN 2515-0855
doi: 10.1049/oap-cired.2017.0610
www.ietdl.org

Tongyou Gu ✉, *Jerom Janssen, Edwin Tazelaar, Geert Popma*

Alliander N.V., Duiven, Gelderland, The Netherlands
✉ *E-mail: tongyou.gu@alliander.com*

Abstract: Weather affects the reliability of distribution networks. Extreme weather conditions can initiate outages, but already normal weather conditions affect the failure of the components in the system, such as cables and cable joints. This study analyses the historical weather and failure records of Alliander, a Dutch distribution system operator. The study discusses the correlation between failure rates and different weather factors. It presents a predictive model using basket analysis. This predictive model is verified using a data set from recent component failures.

1 Introduction

Weather is a dominant factor affecting the reliability of distribution systems [1]. Especially, the failures of components, such as cables and joints, are heavily influenced by weather conditions. Studies [2, 3] demonstrate the influences of extreme weather conditions such as hurricanes and ice storms on power failures. However, even normal weather conditions already correlate with the reliability of components such as cables and joints. For example, historical failure data of Alliander, a distribution system operator (DSO), shows that in summer the System Average Interruption Duration Index of the Amsterdam area is significantly higher than other time of the year.

By analysing the impact of weather factors on (underground) component failures, the expected risk in different regions of the distribution networks can be predicted. Based on a daily or hourly prediction, the grid operator can make a more efficient planning of the number of staffs available for fault clearing service; more crew members can be assigned to the region with higher failure risk. Thus, in case of a power outage due to component failure, there are sufficient crew members who can react immediately for the power restoration. This helps to decrease the failure handling time and reduce the outage duration.

This paper investigates the impact of different weather factors on the failure of (underground) components in medium voltage (MV) network. It cross-examines failure records from Alliander and weather records from the Royal Dutch Meteorological Institute (KNMI). A number of key weather factors over a day are selected for big-data analysis, including maximum/minimum/average temperature, the difference between maximum and minimum temperature of 1 day, precipitation, vaporisation, and solar radiation.

Fig. 1 shows the approach of the study. To obtain the local weather information for each failure, the weather data from 35 weather stations in the Netherlands is interpolated to the location of the failure. Weather data of up to 5 days before the failure is included to investigate any delay effects. The (cor)relation between weather and failures is statistically examined. The weather factors with the most dominant impact on failures are further used to create a predictive model which predicts the risk in different regions on the next day. The (cor)relation between weather and the failures of different types of components is also examined.

The predictive model is tested by the weather and failure data of the summer of 2016. The test results show that it is possible to give a moderately accurate prediction of power failure risk based on the forecast weather information. This model will be further implemented within Alliander to assist the scheduling of crew members for power restoration.

2 Interpolation of weather data

The KNMI has 35 weather stations over the country. Fig. 2 shows their locations. To obtain local weather information for component failures, we use the KNMI data.

For each rayon, an operational region, we use the data from the stations inside this rayon. This provides regional average weather data. We investigated several methods for interpolation of weather data. Other methods provide more granulated weather data, which slightly improves the correlation between component failure and weather data. However, the method with a regional average per rayon has the fastest calculating time. In some instances, there is missing weather data, if this is the case we use weather from another weather station nearby. In the rayon of Amsterdam, there is no weather station. Here we use the nearest station.

3 Normalisation of failure data

For each of the key weather factors selected for this study, the (expected) failure frequency is normalised to avoid the effect that certain weather conditions are more common in the country and occur in more days, which results in more failures occurring under these weather conditions. For example, in the Netherlands an over-average amount of failures occur at temperatures around 10 or 20°C simply because there are a lot of days with these temperatures.

This normalisation is achieved by the following equation:

$$(\text{expected})\ \text{failure frequency} = \frac{\text{total number of failures under these weather conditions}}{\text{number of days with these weather conditions}}$$

Here, the total number of failures and the number of days are counted based on the interpolated weather information.

The failure data used in this analysis is the historical failure records of Alliander from January 2007 to June 2016. The failure records contain the type of component, date and time, and the location.

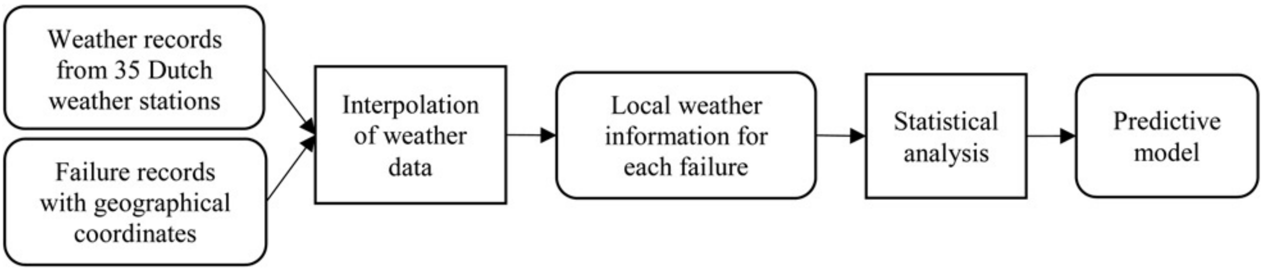


Fig. 1 *Approach of investigating influence of weather conditions on failure rates*

Fig. 2 *Weather stations in the Netherlands*

4 Statistical analysis

The interpolated weather data and the normalised failure frequency are further used in the statistical analysis.

4.1 Correlation

The correlation between weather and the normalised (expected) failure frequency is analysed. The impact of weather conditions on component failure may be cumulative. To study the time delay effect, the correlation between the failure rate and the weather data over five preceding days is calculated. For example, Fig. 3 shows the correlation coefficient between the failure rate and different weather factors over 5 days in Amsterdam area. Temperature difference (3 days average) and global radiation (2 days average) give the highest correlation. This result verifies that weather conditions have an impact on component failures. Moreover, the weather impact could be cumulative over the preceding days.

The time delay effect for each weather factor is also obtained. For each weather factor, there is a highest correlation with the *x*-day average, e.g. for the temperature difference the 3 days' average value is the highest, thus the representative for this weather factor. The representative values (*x*-day average) of all the weather factors are further used to generate predictive rules.

4.2 Predictive model

A predictive model is made to predict the risk in different regions based on the interpolated weather forecast information.

Basket analysis, an a priori method [4], is applied to generate prediction rules. Basket analysis is a modelling technique based upon the theory that if you buy a certain group of items, you are more (or less) likely to buy another group of items. In our case, we would like to generate rules that in certain weather conditions,

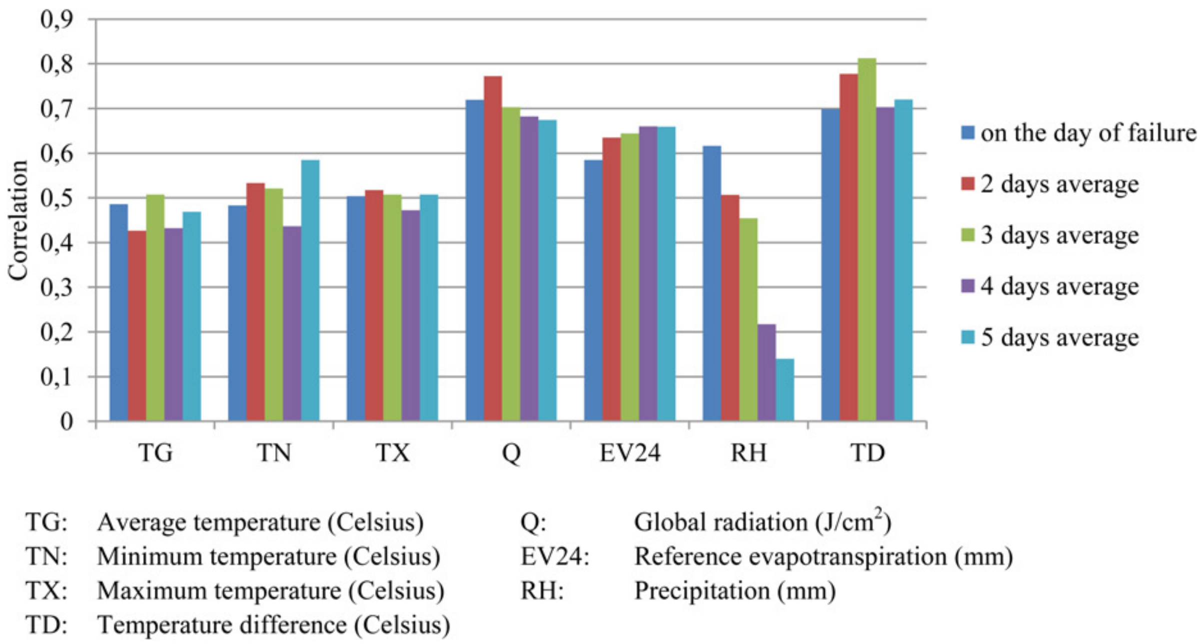


Fig. 3 *Correlation between failure rate and weather conditions in Amsterdam area*

Table 1 Definition of high, medium, and low for basket analysis

	TG	TN	TX	TD	Q	EV24	RH	Number of failures
high	(25, +∞)	(18, +∞)	(30, +∞)	(15, +∞)	(2500, +∞)	(5, +∞)	(20, +∞)	(0, +∞)
medium	(−10, 25]	(−10, 18]	(−5, 30]	—	—	—	—	—
low	(−∞, −10]	(−∞, −10]	(−∞, −5]	(−∞, 15]	(−∞, 2500]	(−∞, 5]	(−∞, 20]	0

Table 2 Predictive rules for expected failure rate in Amsterdam area

If	Then	Support (>0.001)	Confidence (>0.65)
{TX_2 = high, Q_2 = high}	num_fault = high	0.001142531	0.8000000
{TX_2 = high}	num_fault = high	0.002285061	0.7272727
{EV24_4 = high}	num_fault = high	0.001142531	0.6666667

Table 3 Predictive rules for expected failure rate in Amsterdam area

If	Then	Support (>0.001)	Confidence (>0.2)
{EV24_1 = high}	num_fault = high	0.002158114	0.2054381
{TX_2 = high, Q_2 = high}	num_fault = high	0.002189851	0.2005814

it is more likely to have higher failure frequency. The basket analysis targets high failure risk and seeks relevant conditions with a certain confidence level.

The input items of the basket analysis need to be discrete quantities. Therefore, the historical data of the weather factors are categorised as 'high', 'medium' or 'low'. The boundaries are defined based on the statistical distribution of the data. Table 1 shows the category boundaries. Since the number of failures in the MV network of a certain region is relatively low from a statistical perspective, it is defined as 'high' when there is one (or more) failure.

Table 2 illustrates the predictive rules for Amsterdam area. The rules state that when the maximum temperature (2 days average) or evapotranspiration (4 days average) is high, more failures are expected in Amsterdam area with a confidence level >0.65. The support value is larger than 0.001, which indicates that >4 days (0.001×3501 days in January 2007–June 2016) in the dataset have the weather condition as in each rule. Grid operators can use the predictive rules as an indication about which regions are with higher risk.

5 Analysis for different types of components

Predictive rules are also generated for different types of components in the MV network: cables, cable joints, transformers, and secondary installations.

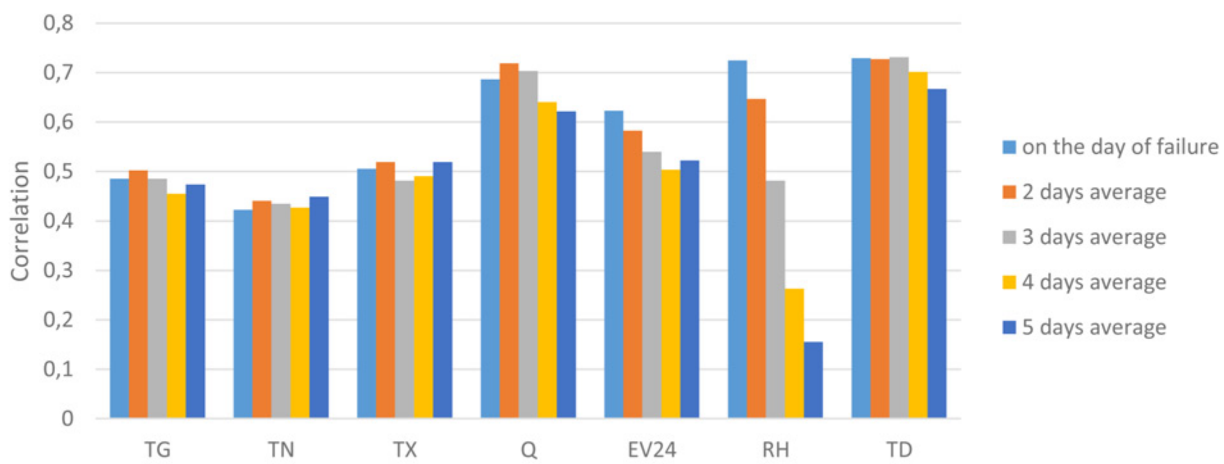
As an example, Fig. 4 shows the correlation between failure rate of cable joints and different weather factors over 5 days. Precipitation (on the day of failure), temperature difference (3 days average), and global radiation (2 days average) give the highest correlation.

Again, the representatives for each weather factor are used to generate the predictive rules. Table 3 illustrates the resulted rules, in which the reference evapotranspiration, maximum temperature and global radiation are the most predictive factors. The confidence level in Table 3 is lower than in Table 2, because the total number of failures of a certain component type is much smaller than the number of all the failures, which leads to lower confidence in statistical analysis. It is discovered that for different components the dominant weather factor varies.

6 Test results

The predictive model is tested using the weather and failure data of the summer of 2016. Within the 60 days of July and August, there were 159 power failures in 10 Alliander operational regions. For every region on each day (in total 620 cases), the model is applied to predict the risk of power failures.

Fig. 5 illustrates the summary of the test results. When it is predicted as 'low risk', in 72% of the cases there was indeed no failure. When it is predicted as 'high risk', in 64% of the cases, there was power failure(s). These results show that based on the



TG: Average temperature (Celsius)
TN: Minimum temperature (Celsius)
TX: Maximum temperature (Celsius)
TD: Temperature difference (Celsius)
Q: Global radiation (J/cm^2)
EV24: Reference evapotranspiration (mm)
RH: Precipitation (mm)

Fig. 4 *Correlation between failure rate of cable joints and weather conditions*

CIRED, Open Access Proc. J., 2017, Vol. 2017, Iss. 1, pp. 1442–1445

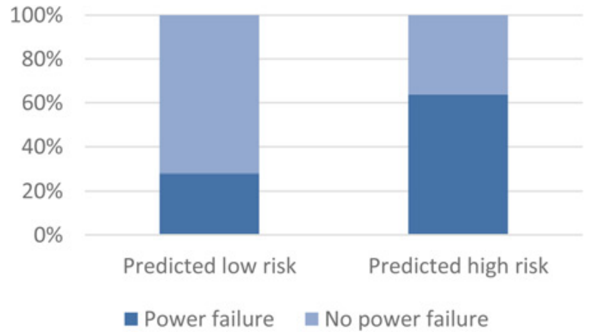


Fig. 5 *Test results of predictive model*

forecast weather information, it is possible to give a moderately accurate prediction.

7 Conclusion

This paper presents an analysis of historical weather and failure records. It first discusses the correlation between the failures and different weather factors, and then constructs a predictive model using basket analysis.

The results show that some weather factors are clearly correlated with the failure rates of certain component type. The dominant weather factors (highest correlation coefficient) vary between different regions. A possible explanation is that the type and age of the cables in each region are different due to the diversification of regional grid development over the past decades. It is also discovered that the dominant weather factor varies for different components.

Using the x-day average representative for each weather factor, a basket analysis is performed to build the predictive model. The basket analysis targets high failure risk and seeks relevant conditions with sufficient confidence level. The predictive model can be used to give a high-risk alarm when certain weather conditions occur. The model is tested by the recent failures and weather records in the summer of 2016. The test results show moderate accuracy of the prediction.

Weather impact is one of the (indirect) factors that could cause a component failure. The study suggests to take the root causes into account to make a more complete predictive model of failure risk, e.g. weather influence on the physical aging process of components and/or digging damages of underground components because of improper weather conditions for digging work.

Alliander is implementing the model in the optimisation of the scheduling of outage service crew members. This helps to decrease the failure handling time and reduce the outage duration.

8 References

1 Zhou, Y., Pahwa, A., Yang, S.S.: ‘Modeling weather-related failures of overhead distribution lines’, *IEEE Trans. Power Syst.*, 2006, **21**, (4), pp. 1683–1690

2 Billinton, R., Acharya, J.: ‘Consideration of multi-state weather models in reliability evaluation of transmission and distribution systems’. Canadian Conf. Electrical and Computer Eng., 2005, Saskatoon, Sask., 2005, pp. 916–922

3 Brostrom, E., Ahlberg, J., Soder, L.: ‘Modelling of ICE storms and their impact applied to a part of the Swedish transmission network’. 2007 IEEE Lausanne Power Tech, Lausanne, 2007, pp. 1593–1598

4 Trnka, A.: ‘Market basket analysis with data mining methods’. 2010 Int. Conf. Networking and Information Technology, Manila, 2010, pp. 446–450