MÁSTER UNIVERSITARIO EN SMART GRIDS

TRABAJO FIN DE MÁSTER

Detection of Hotspot Anomalies in LV Networks via Data Analysis

Autor: Carlos Prieto Rodríguez de Vera

Director: Dr. Miguel Ángel Sanz Bobi

Co-Director: Jesús Gutiérrez Serrano

Co-Directora: Itziar Lumbreras Basagoiti

Madrid
Agosto de 2025

Carlos Prieto Rodríguez de Vera, declara bajo su responsabilidad, que el Proyecto con título **Detection of Hotspot Anomalies in LV Networks via Data Analysis** presentado en la ETS de Ingeniería (ICAI) de la Universidad Pontificia Comillas en el curso académico 2024/25 es de su autoría, original e inédito y no ha sido presentado con anterioridad a otros efectos. El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido tomada de otros documentos está debidamente referenciada.

Fdo.: Carlos Prieto .... Fecha: 20 ...... / 08 ...... / 2025 ....

Autoriza la entrega:

EL DIRECTOR DEL PROYECTO

Dr. Miguel Ángel Sanz Bobi

Fdo.: .... Fecha: ...... / ...... / ....

Detection of Hotspot Anomalies in LV Networks via Data Analysis

Carlos Prieto Rodríguez de Vera
Master in Smart Grids (MSG)
Universidad Pontificia Comillas - ICAI
Madrid, Spain
carlosprietorodriguezdevera@gmail.com

Dr. Miguel Ángel Sanz Bobi
Director
Universidad Pontificia Comillas - ICAI
Madrid, Spain
masanz@comillas.edu

Jesús Gutiérrez Serrano
Co-director: Industrial Supervisor
i-DE redes inteligentes, Iberdrola
Madrid, Spain
jmgutierrez@iberdrola.es

Itziar Lumbreras Basagoiti
Co-director: Industrial Supervisor
i-DE redes inteligentes, Iberdrola
Madrid, Spain
ilumbreras@iberdrola.es

***Abstract*—Low-voltage (LV) distribution networks are undergoing rapid complexity growth driven by distributed energy resources (DER), electric vehicles and bidirectional power flows, which challenges conventional protection and maintenance strategies. This paper develops and validates a predictive maintenance framework to anticipate hotspot anomalies, localized thermal degradations, in secondary substations (SS) and underground cable pits (UCPs) by leveraging the advanced LV supervision SABT and AMI smart-meter ecosystem. For SS, we propose a hybrid scheme comprising (i) a physically based analytical threshold model, (ii) a general multilayer perceptron (MLP) regressor trained on healthy behaviour and (iii) a substation-specific MLP; anomaly flags arise from persistent prediction errors above a 3σ error threshold. For UCPs, where direct measurements are absent, we construct probabilistic models on smart-meter event logs using Gaussian Mixture Models (GMMs) with BIC-guided selection and late-window change detection. On real OMS-labelled incidents, the SS combined policy (AND of the three models) achieves TPR 60.71% with FPR 1.33% and delivers the best annualized net economic benefit (~€944k/yr), while the analytical model alone achieves the highest TPR/FPR ratio (7.71) and the general ML model had the highest TPR at 92.86%. The UCP GMM attains TPR 46% with effective FPR < 1%, demonstrating feasibility but not yet break-even under direct-cost assumptions; however, safety and reputational externalities argue for continued development. The results indicate that data-driven predictive maintenance in LV networks is technically viable at scale and economically attractive for secondary substations, with a clear upgrade path for underground cable pits. All metrics and content are sourced from this project and will be presented in later sections.**

***Index Terms*—Smart grids, Low-voltage distribution networks, Predictive maintenance, Anomaly detection, Multilayer Perceptron (MLP), Gaussian Mixture Model (GMM), Advanced Metering Infrastructure (AMI), SABT, Economic analysis, secondary substations, underground cable pits, hotspot events**

I. Introduction

A. Background

The traditional unidirectional electricity supply chain is dissolving as DER, battery systems, and EV integration introduce reverse power flows and time-varying operational states, complicating LV operation [1] [2], forecasting and pressuring quality indices such as SAIFI/SAIDI. This constant changes demand for a modernisation in operations, necessitating a change in the way data is used for everyday operations and maintenance schemes [3]. Digitalization via SABT and AMI increases LV observability, enabling continuous collection of measurement profiles and asynchronous event logs, which are essential to shift from corrective to predictive maintenance and reduce non-served energy and reduce OPEX and CAPEX [4] [5].

B. Problem Statement

Hotspot anomalies are localized thermal degradations that mainly originate from loose or aged contacts, insulation deterioration, moisture ingress, phase unbalance and overloading [6] [7] [8]. They can escalate to service-affecting incidents and safety risks (SS and UCPs [9] alike). While thermography and maintenance are used reactively, a validated predictive solution leveraging SABT/AMI data and aligned with DSO economics [10] [11] has not been publicly established for LV hotspot-driven anomalies [12]. Figure: 1, 2

(a) Damage caused by a hot spot event detected in a routine maintenance

(b) Thermal imaging of the hot spot event

Fig. 1: Images depicting hotspot events and consequences in a secondary substation [source: internal maintenance reports from i-DE]

Fig. 2: Underground cable pit thermal event in Gastiez, Spain [*Source*: Newspaper Gastiez-hoy: "La Calle Diputación registra varias explosiones en una arqueta eléctrica"]

C. Objectives and Scope

This work targets industrial-grade predictive prevention of hotspots in (i) secondary substations and in (ii) underground cable pits. The objectives include OMS-driven historic incident detection and labelling, predictive maintenance algorithm development for (i) with both analytical methods and ML techniques, validation of results with real incident and field data and finally, extend the methodology to underground cable pits and develop a predictive maintenance algorithm for hotspot events prediction in such grid elements, this final part introduces additional complexity due to the indirect nature of the data.

D. Contributions

A hybrid predictive framework for SS that fuses a physically interpretable thresholds with general/specific MLP regressors, yielding TPR 60.71% with FPR 1.33% under an AND combination and >€900k/yr net economic benefit; individually, the analytical model achieves TPR/FPR = 7.71 and the general MLP achieves the highest TPR (92.86%). A GMM-based UCP hotspot predictor on event logs that captures late-window probability inflation and emergent Gaussian components, achieving TPR 46% with effective FPR < 1%. An economic break-even model that ties TPR/FPR to crew visits, repairs and incident avoidance, together with compute-cost accounting, revealing that SS are ROI-positive today while UCPs require either additional telemetry or inclusion of economic externalities like safety and press. A potentially scalable data pipeline within SABT/AMI databases, with OMS text filtering for robust labelling.

E. Paper Organization

Section II reviews maintenance strategies, LV anomaly-detection approaches and ML techniques; Section III outlines data sources, measurements and labelling; Section IV details the SS and UCP methodologies; Section V presents statistical, economical and model results; Section VI discusses trade-offs and deployment; Section VII states limitations of the project; Section VIII concludes and adds future research options to further develop the project.

II. Related Work and Technical Context

A. Maintenance Strategies in Power Distribution

Maintenance strategies encompass corrective, preventive, and predictive paradigms, with complementary philosophies such as RCM, CBM, risk-based, TPM, and lean maintenance guiding prioritization and workflow efficiency [13]. Corrective maintenance minimizes planning overhead but increases unplanned downtime and risk exposure; preventive maintenance schedules inspections at fixed intervals, reducing surprise failures but incurring in unnecessary interventions; predictive maintenance leverages continuous monitoring and analytics to act on condition-derived risks, reducing OPEX and extending asset life [5]. In LV grids, preventive schemes are the most common practice [14] but, predictive strategies are emerging as SABT/AMI expand observability [15].

B. LV Anomaly Detection

Within LV networks, instantaneous anomaly detection is practised for operational awareness (e.g., PQ alarms, thermography), but predictive maintenance, anticipating temperature-driven hotspot anomalies, is not commonly established. Industrial initiatives emphasize thermography [6] [7] and dielectric hotspot inspection [8]; academic efforts focus on fault localization, cybersecurity anomaly detection [16] and non-technical loss analytics rather than pre-failure thermal anomalies in LV assets. The gap motivating this work is predictive, hotspot detection for SS and UCPs via SABT/AMI analytics, validated against OMS-labelled incidents.

C. Analytical and ML Approaches for Anomaly Detection

Techniques include physical/analytical models, statistical thresholds, supervised/unsupervised ML, and image-based CNNs [17]. Analytical thresholds are interpretable and suited for MVPs/auditability but are limited for multi-factor complex relationships; supervised models and deep learners (MLP, CNN, RNN/LSTM) capture complex patterns but some are more computationally expensive [18]; unsupervised models (One-Class SVM, Isolation Forest, LOF, Autoencoders) detect anomalies without labels [17], [19], [20]; clustering (k-means, mean-shift, DBSCAN, spectral, GMM) offers complementary structure discovery [21]. In contrast with other deep learners, MLPs reduce model complexity and training time [22]. GMM provides a probabilistic, interpretable model with flexible covariance structures and BIC-based selection, advantageous when features are few and event statistics dominate [23].

III. Data and System Overview

A. System Architecture and Data Sources

Three main databases are used as data sources. ICDS (OMS) is integrated for information related to incident logs and affected elements, it is essential as labelling ground truth. GENESIS (GIS) is used for inventory characteristics and nominal parameters required for data normalization. STG is the source of SABT line-supervisor profiles and smart-meter event logs. SABT profiles (currents, voltages, power, ambient temperature, and computed thermal image) arrive at 5-min interval resolution; events arrive asynchronously from LS cards (secondary substation) and RTU endpoints (customer smart meters).

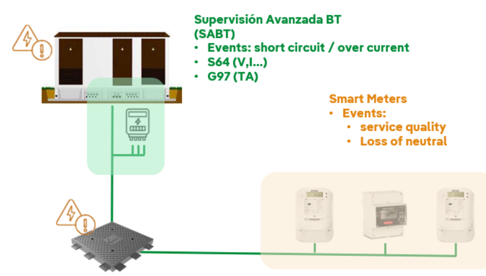


Fig. 3: Data sources (SABT and Smart Meters) and visual representation of a secondary substation and an underground cable pit

B. Measurement Set and Derived Variables

The secondary substation predictive algorithms use three-phase and neutral currents, phase voltages, ambient temperature (measured by sensors placed near each conductor, therefore having a relationship with conductors temperature), geographical ambient temperature (from open-source weather API: Open-meteo) and a computed thermal image per IEC/EN 60255 [24]. With constants (TPTMax, TPTlimit, TPInom, TPThermalC, TPCurrentMeasMax) and measured ambient temperature TPTamb and current, the conductor temperature is [25]:

$$F_a = \frac{\mathrm{TPTMax} - \mathrm{TPTlimit}}{\mathrm{TPTMax} - \mathrm{TPTamb}} \tag{1}$$

$$H(t) = \left(\frac{I(t)}{\mathrm{TPInom}}\right)^2 \cdot \frac{15}{\mathrm{TPThermalC} + 15} \cdot F_a + \frac{\mathrm{TPThermalC}}{\mathrm{TPThermalC} + 15} \cdot H(t-1) \tag{2}$$

$$\mathrm{Temp}(t) = \mathrm{TPTamb} + H(t) \cdot [\mathrm{TPTMax} - \mathrm{TPTamb}] \tag{3}$$

For UCP analysis, we mainly exploit power-quality event groups/types from RTU (customer smart meters) and LS cards (secondary substation supervisors). We focus on voltage deviations (low/high limits), long-duration outages, overloads, short-circuit detection and neutral loss detection events [25].

C. Data labelling

OMS extraction, text normalization/keyword filtering, manual vetting, and linkage to the most-affected element by downtime duration produce labelled data. For SS, we build paired windows: months pre-incident (damaged) and post-repair (healthy), plus preventive-maintenance cases and a control population of similar SS, equal loadings, characteristics, geographical location and date times (at a 1:10 ratio). For UCPs, we aggregate 120 days of downstream events per line, normalised for total number of meters and include seasonally overloaded but healthy lines as hard negatives, as the objective is to detect only those lines that contain hotspot events that can lead to a downtime incident with loss of power.

D. Data Volumes and Practical Limits

At a 5-min frequency and system scale (~100k secondary substations, ~550k lines, 11M meters), naive annual ingestion approaches ~9.3×10^{11} points (~900 GB) for decided features; this necessitates sampling and modular model design. Nevertheless, due to the critical nature of the study and safety risks involved, all models presented have been trained with the data of all detected hotspot anomalies in the past year with an additional control population. Further studies also demonstrated that compute costs are negligible compared to field OPEX, these will be developed in Section V.

IV. Methodology

A. Secondary Substations (SS) Method

1) Historic Incidents and Statistical Analysis

Historic incidents are detected by applying the required filters to the OMS results and then filtering the descriptions first by key word inclusion, unwanted key word exclusion and then a by a hand filter (which can eventually be automated via a LLM). Around 4.6M time stamp instances were then recorded of both healthy and damaged instances for the model to train, test and validate on. This pre-extraction filters alongside the word filter reduced the size of the incident list from hundreds of thousands to about one hundred instances. It is possible some hotspot events were lost to some filter but several filtering techniques and different word combinations were tested to loose as few hotspots as possible whilst keeping the filter strict enough to avoid poor labelling, as this would greatly affect the performance of the models.

This data times were then analysed statistically for each of the relevant features (60M data points approximately). Mainly temperature and current distributions and their relationships via pair-plots were analysed for pattern detection that proved prediction algorithms had a statistical bases.

2) Economic Analysis

We link model TPR/FPR to annual economics: visit cost €120 per flagged SS; repair cost €1,200 at 30% of TPs (as some hotspot repairs require a smaller cost); incident-avoidance benefit €120,000 at 35% of incidents (as not in all incidents does the full SS get damaged). The values were agreed on by field experts on SS maintenance. Model execution costs were assumed negligible when compared with these costs, results will be provided in the next section. There is additionally a significant and complex cost related to the safety hazard and social impact caused by such thermal incidents, that is not included in the scope of this study but should be further analysed.

3) Analytical Model

A physically grounded, criteria-based model was developed to distinguish healthy from damaged lines in secondary substations. The model relies on expert knowledge and physical phenomena, requiring minimal training and enabling iterative refinement. Over 20 hotspot-related criteria were analysed, where possible basing thresholds on statistical deviations rather than fixed values. Also, the optimal frequency of flagging for each criteria was analysed and criteria combinations were explored to capture more complex anomaly patterns over time. An agile workflow was followed, starting with evident criteria and incrementally expanding the model as new relationships emerged.

4) General and specific ML Model

The general ML approach focused on developing a predictive anomaly detection system for secondary substations using Multilayer Perceptron (MLP) regressors. MLPs were chosen for their simplicity, effectiveness in modelling non-linear relationships and suitability for detecting abrupt changes rather than long-term patterns [26] [27]. Two models were built: a general model trained on data from all secondary substations for scalability and computation efficiency and a specific model tailored to each secondary substation for higher accuracy. Both models predict ambient temperature based on multivariate time series profiles from the SABT - STG database: date-time, phase currents, phase temperatures, zone code, outside temperature, ventilation type and secondary substation type.

Both models are trained with healthy line data and anomalies are flagged when prediction errors exceed a 3-sigma threshold for three consecutive time steps, following control charts theory and using stricter conditions [28]. The methodology involved careful feature selection, exhaustive ML architecture optimisation, iterative training, and statistical assumptions about error distribution, all within an agile development framework. Several assumptions like statistical stationarity over training for line deterioration, prediction error normal distribution and proper prior labelling.

B. Underground Cable Pits (UCP) Method

1) Historic Incidents and Statistical Analysis

Historic hotspot incidents in underground cable pits were identified using a process analogous to that for secondary substations, with adaptations for data sources, features, and filters. All incidents from the previous year were retrieved from the OMS database and filtered for categories likely to include UCP hotspots (e.g., unplanned outages, Other Attentions on supply nodes, LV lines, electrical boxes). Descriptions were normalized (ASCII, lowercase, no accents ...) to mitigate operator entry inconsistencies.

Incidents were screened for hotspot-related terms (e.g., smoke, heat, sulphated) and location indicators (underground cable pit, LV underground line), while excluding MV-specific descriptors. Root-based matching reduced misspell sensitivity. This reduced the dataset from hundreds of thousands to ~700 candidates, which were then manually validated to ensure labelling quality—critical for model performance. For each validated incident, the most affected element (longest downtime) was identified, typically the line requiring corrective maintenance. Line meters were then extracted, including SABT supervisors (LVS) and downstream smart meters (CN, T4, T4MI).

For each line, event logs from all associated meters were collected for the 120 days preceding the incident (excluding day 0 to preserve preventive action feasibility). Events were aggregated daily and normalized by meter count to form a per-line event-rate series. To challenge the model, a control set of healthy but seasonally overloaded lines (e.g., coastal feeders during summer) was included, as these exhibit voltage deviations similar to damaged lines. Lines with zero events were excluded to avoid trivial separability.Faulty meters generating >1000 daily events were removed. Final features represent normalized daily event counts per line over 120 days.

The initial analysis of smart-meter event logs involved plotting daily event counts per line; however, this visualization failed to reveal actionable patterns, prompting a shift toward statistical characterization. Using Python statistical libraries, the study generated kernel density estimates (KDE) for individual lines, an aggregated KDE, histograms, and per-line probability density plots, applying refined filters for relevant event groups and types across SABT (LVS) and downstream meters. These analyses aimed to uncover latent distributional differences between damaged and healthy lines, guiding the subsequent selection of probabilistic modelling techniques.

2) Economic Analysis

An economic feasibility study was conducted to determine the precision threshold required for the UCP anomaly detection model to achieve break-even. The analysis considered direct costs: crew dispatch for flagged lines (€120 per visit), corrective maintenance for confirmed hotspots (€800 at 50% as some hotspot repairs require a smaller cost) against the benefit of avoiding hotspot failures, approximated at €1,000 per incident (discounted to 90%). Model execution costs were assumed negligible when compared with these costs, results will be provided in the next section. Using estimated annual incident rates and the population of LV lines, the study revealed that the high cost of field visits relative to replacement costs imposes a stringent TPR/FPR ratio for economic viability. While indirect benefits such as improved public safety and reputational risk mitigation were excluded from this calculation, their inclusion could significantly enhance the business case and should be addressed in future assessments.

3) Gaussian Mixture Model

Model selection for hotspot anomaly detection in underground cable pits was driven by the prior statistical analysis, which indicated that a probabilistic approach could effectively capture the observed late-window event escalation. Among various clustering techniques [29], the Gaussian Mixture Model (GMM) was chosen for the available data characteristics, its ability to handle unknown and variable cluster densities, computational efficiency and interpretability, critical for large-scale deployment across ~550,000 lines. The model operates on aggregated daily event counts from SABT supervisors and downstream smart meters, normalized by meter count, over a 120-day window preceding each incident. Its objective is to identify anomalous behaviour characterized by an increasing probability of event occurrence, particularly when this rise accelerates in the final 30 days.

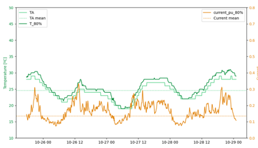
The implementation involves sequential steps: data extraction and cleaning (removing outliers and faulty meters), event-type filtering, daily aggregation and normalization, and iterative GMM fitting across covariance types and up to 40 components, selecting the configuration minimizing BIC. Anomalies are flagged based on probability thresholds and structural changes, such as the emergence of new Gaussian components after the 90-day mark. While this approach balances scalability and interpretability, it assumes event frequency correlates with degradation, a limitation in cases of maintenance activity. Fixed windowing and daily aggregation, chosen as a result of the statistical analysis and for computational efficiency, may overlook finer temporal dynamics. But the main limitation is the absence of direct thermal measurements constrains benchmarking against alternative models. Future improvements include incorporating maintenance logs, adaptive windowing and instalment of new direct data sources to enhance predictive accuracy without compromising scalability.

V. Results

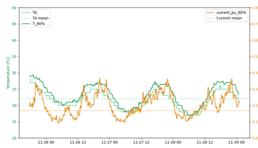
A. Secondary Substations

1) Statistical and Economical Analysis

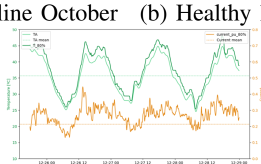
Ambient-temperature distributions shift with hotspots, showing a high-temperature tail; phase-current KDEs reveal stronger unbalances in damaged lines; pairplots show tighter temperature-current coupling and elevated temperatures at lower currents; temporal evolution indicates gradual degradation with late acceleration; inter-line comparisons in the same SS reveal a temperature offset on the hotspot line at similar loading when compared to the other lines. Some of the main analysis result graphs are presented below:



(a) Healthy line October



(b) Healthy line November



(c) Damaged line December

Fig. 4: Cable degradation with passing time

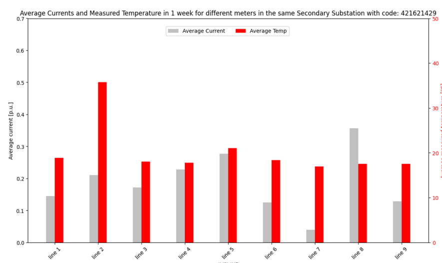


Fig. 5: Gradient temperature between lines in a secondary substation

In terms of the economic analysis, the assumptions stated in the methodology were used to develop a model to compute the minimum KPI (TPR/FPR) required for economic break-even. With the limitation that this model does not include the social and safety risk involved costs which are significant. There are around 100 incidents per year, of which 40% <have SABT measurements. With that the outcome is a required minimum TPR of 8.43%, FPR of 2.92%, **KPI Limit (TPR/FPR) of 2.89**, to obtain a final economic value of -0.01€. Therefore any model that surpasses 2.89 will be economically feasible.

2) Model Performance: Analytical Model

Of the +20 criteria developed with physical and statistical thresholds 7 main ones and 2 combinations were selected to create the final model which is an and combination of criteria 2 & 14 & A & B. These criteria are presented below alongside the optimal threshold (the minimum required instances of the toal data recorded for the given study period that these criteria must be flagged before the line is to be considered damaged by the model)

TABLE I: Performance Metrics for Toggle Criteria

Criteria	**Description**	**Precision**	**Recall**	**Optimal Threshold**
1	Temperature exceeds a limit	100.0%	1.3%	0%
2	Temperature gradient inside a secondary substation	98.0%	26.5%	3.0%
3	Relationship between temperature and current squared	42.3%	1.13%	–
4	Deviation from mean temperature in that secondary substation	66.9%	9.6%	37.0%
10	High current standard deviation in that secondary substation	74.6%	58.3%	–
12	Phase unbalances	63.4%	87.3%	–
14	Close relationship between ambient temperature and thermal image	21.4%	8.42%	22%
A	Combination: Crit3 & Crit4 & Crit10 & Crit12	58.1%	0.12%	1%
B	Combination: Crit4 & Crit10	51.0%	16.0%	26%

		Prediction outcome		
		p	**n**	**total**
Actual value	**p′**	24	4	28
	n′	25	200	225
	total	49	204	253

TABLE II: Confusion Matrix

In terms of computations: 0.47 processing minutes per secondary substation, 313 total processing hours, for an Azure Standard_D96ls_v5 VCPU the

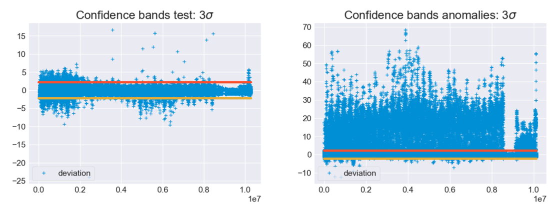
KPI	Result
True Positive Rate	85.71%
False Positive Rate	11.11%
TPR / FPR	7.71
Economic result	€891,000 / yr

TABLE III: Economic results of the final analytical model

cost is 7.7571€/hr and has 96 vCPUs and 192 GiB (enough for the required task). For a total processing cost of 25.32€ per analysis (assumed negligible compared to the other costs and benefits)

3) Model Performance: Global ML Model

Global MLP predicts ambient temperature; alarms occur when the 3σ confidence bands error limit is exceeded for 3 consecutive points. Optimal architecture after iterative search (20,20) with learning rate 0.01;



(a) General ML model test error compared with 3 sigma limits

(b) General ML model hotspot containing substation error compared with 3 sigma limits

Fig. 6: Model error for training with healthy lines and validating with damaged lines

		Prediction outcome		
		p	**n**	**total**
Actual value	**p′**	26	2	28
	n′	64	161	106
	total	63	71	253

TABLE IV: Confusion Matrix

KPI	Result
True Positive Rate	92.86%
False Positive Rate	28.44%
TPR / FPR	3.26
Economic result	€178,000 / yr

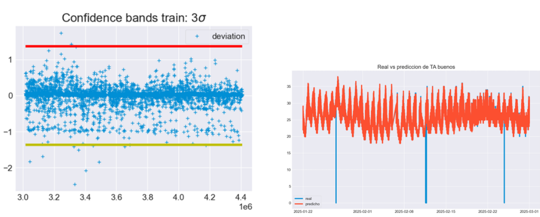
TABLE V: Economic results of the general ML model

General ML model is better at predicting, has a higher TPR at the expense of an increased FPR making the model less economically viable.

In terms of computation costs: 0.02 processing minutes per secondary substation, 12 total processing hours, for an Azure Standard_D4as_v5 VCPU the cost is 0.1375€/hr and has 4 vCPUs and 16 GiB (enough for the required task). For a total processing cost of 1.02€ per training (assumed negligible compared to the other costs and benefits).

4) Model Performance: Specific ML Model

Specific ML model follows the same steps as the general model but a different model is generated per secondary substation and is trained only on such secondary substation data. As it is specific, less false positives are flagged and there fore is economically more viable. But, as the 3 sigma - 3 consecutive instances rule is kept, now that the model is more specific this rule is more strict and therefore less positives are detected in total terms.

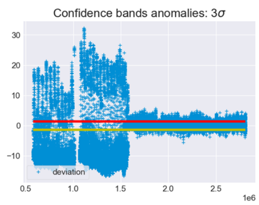


(a) Specific ML model train error compared with 3 sigma limits

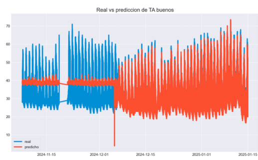
(b) Specific ML model ambient temperature prediction vs real measured value

Fig. 7: Specific ML model predicts accurately for a healthy line

In terms of computational costs: 0.176 processing minutes per secondary substation, 117 total processing hours, for an Azure Standard_D96ls_v5



(a) Specific ML model test error compared with 3 sigma limits for a repair case



(b) Specific ML model ambient temperature prediction for a repair case

Fig. 8: Model performance for a complex case where a damaged line gets repaired due to a preventive maintenance and then the model starts predicting accurately

		Prediction outcome		
		p	**n**	**total**
Actual value	**p′**	18	10	28
	n′	31	194	225
	total	49	204	253

TABLE VI: Confusion Matrix

KPI	Result
True Positive Rate	64.29%
False Positive Rate	13.78%
TPR / FPR	4.7
Economic result	€407,000 / yr

TABLE VII: Economic results of the specific ML model

VCPU the cost is 7.7571€/hr and has 96 vCPUs and 192 GiB (enough for the required task). For a total processing cost of 9.48€ per round of training, meaning training required for all secondary substations (assumed negligible compared to the other costs and benefits)

5) Model Performance: Combination Model

The economically optimum model is a logical AND combination across (Analytical Rule-Set) $\wedge$ (General MLP) $\wedge$ (Specific MLP). With the limitations and assumptions explained for the economical model, if safety risks costs were included each model would be significantly more economically viable but this study was outside the scope of the project. The summary of the combination model and the comparison between all models is presented below:

KPI	Result
True Positive Rate	60.71%
False Positive Rate	1.33%
TPR / FPR	45.5
Economic result	€944,000 / yr

TABLE VIII: Economic results of the combined model

In terms of computational costs: 0.8 processing minutes per secondary substation, 533 total processing hours, for an Azure Standard_D96ls_v5 VCPU the cost is 7.7571€/hr and has 96 vCPUs and 192 GiB (enough for the required task). For a total processing cost of 43.10€ for training (the highest of the prior but still assumed negligible compared to the other costs and benefits)

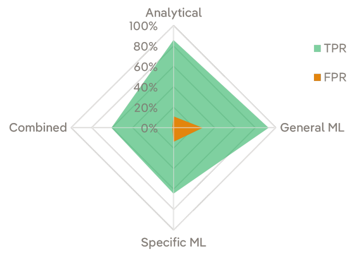
B. Underground Cable Pits

1) Statistical and Economical Analysis

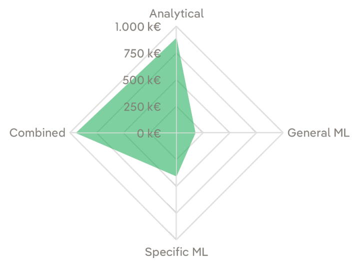
The objective of this section of the project is to create a predictive maintenance model capable of pre-emptively flagging unhealthy underground cable pits that require maintenance due to a hotspot event. The challenge is that there is no direct measurement available or purpose built sensor to detect this. Meaning data analytics must be used with the current sources to develop the model. The measurements deemed most connected with the problem were SABT and Smart Meter asynchronous event logs, specifically the power quality events of over/under voltage, neutral wire disconnection and short-circuit.

Data was aggregated at a daily bases, normalised for the number of meters connected to the studied line and faulty meters removed form the study. To ensure prediction capability a prior statistical analysis was performed as the visual time bases analysis yielded no results.

Event-density rises after the 90-day mark across lines, motivating late-window probabilistic logic and backing up model selection: GMM.



(a) Main KPIs summary



(b) Economic summary

Fig. 9: KPI and economic summary for: analytical model, general ML model, specific ML model, combination of models

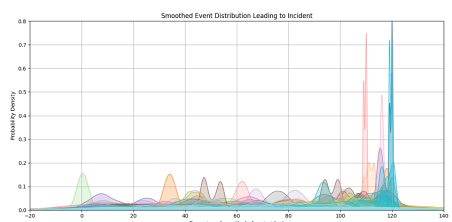


Fig. 10: Probability distribution of number of events per day

An economic analysis was then performed to analyse the minimum KPI (TPR/FPR) required for economic break-even. Nevertheless, given the low CAPEX characteristics of underground cable pits and the high cost associated with visiting false positives relative to the modest corrective maintenance expenses, economic viability depends on a high accuracy of the predictive model, accuracy which is hindered by the limited availability of direct data for this specific case. With the limitation that this model does not include the social and safety risk

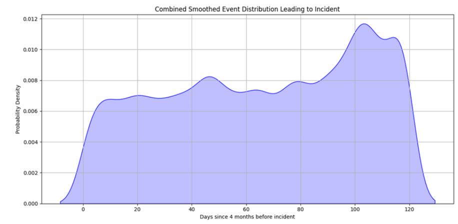


Fig. 11: Combined probability distribution shows a constant value up to the 90 day mark prior to the incident where probability increases

involved costs which are significant.

There are around 500 incidents per year of the 550,000 LV lines. With that the outcome is a required minimum TPR of 65.70%, FPR of 0.21%, **KPI Limit (TPR/FPR) of 315.47**, to obtain a final economic value of -0.01€. Therefore any model that surpasses 315.47 will be economically feasible.

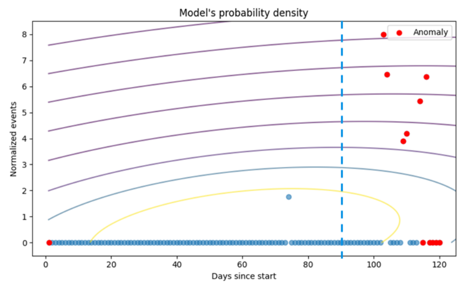
2) GMM Performance

The process of finding the optimal model was iterative and involved preprocessing data differently and setting certain criteria. The optimal final model does a per line fit of GMMs, finding iteratively the optimum: up to 40 components with four covariance types and selected by minimum BIC. Characterize anomalies via least-probable points considered anomalous point, and doing an anomalous rate comparison between before and after the 90-day split.

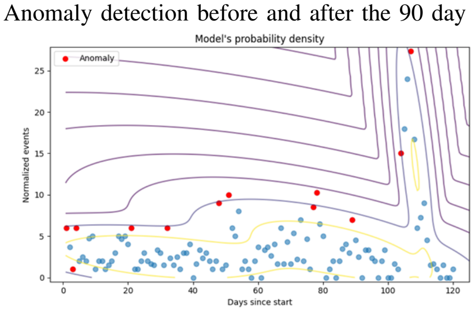
For extreme cases with vast amounts of anomalous points where new Gaussians were created for those anomalous points, detect new-Gaussian emergence by comparing minimum number of Gaussians when training for 90 first days and when training for the full 120 days. If such new Gaussians had an average number of events twice the mean number of events of the pre 90 day Gaussians then flag the line.

The results of the best several models tested and the combination that encompasses the most detection rate are presented below:

All FPR presented are calculated over a population of overloaded lines with significant meter events related to other healthy lines, this is to increase the complexity of the model as normal



(a) Anomaly detection before and after the 90 day mark



(b) Example of a case where a new Gaussian is needed to explain the last days of a damaged line and such gaussian has more than twice the average nº of events

Fig. 12: Models 13 and 68 visual representation

TABLE IX: Performance metrics of hotspot anomaly prediction GMM models for underground cable pits

Model Nº	Model Description	TPR	FPR
13	Rate of anomaly after > 2 × rate before (using p90 probability for 'before')	31%	4%
55	Number of normal instances after > number of normal instances before	73%	88%
64	Mean of normal instances after > 2 × mean of normal instances before	39%	4%
68	Combination of models 55 and 64	36%	4%
100	Logical OR between models 13 and 68	46%	8%

healthy lines basically have no events and can easily be separated from these 2 groups (damaged hotspot containing lines and seasonally overloaded lines with several daily events).

Model 100 achieves TPR 46% and FPR 8% on overloaded lines; effective FPR ~0.83% (as overloaded lines acount only for 43,000 out of the 550,000 lines and using a 0.2% FPR for the healthy non event containing lines. Achieving a KPI of 55.73, break-even KPI for UCPs (315.47) is not met under direct-cost accounting (€350k deficit), but safety/reputational externalities exposed by internal experts suggest continued development.

Finally, in erms of computational costs: 0.24 processing minutes per secondary substation, 1800 total processing hours, for an Azure Standard_D96ls_v5 VCPU the cost is 7.7571€/hr and has 96 vCPUs and 192 GiB (enough for the required task). For a total processing cost of 145.45€ for training, assumed negligible compared to the other costs

VI. Discussion

Analytical rules maximize interpretability and TPR/FPR; general MLP maximizes recall; specific MLP reduces FPR for a relatively more strict flagging; AND combination optimizes economics by minimising FPR while retaining useful TPR. For UCPs, GMM is appropriate given data constraints but requires added telemetry or a deeper external costs study for ROI. All whilst computational costs remain negligible, nevertheless more efficient model can be studied to reduce model OPEX.

VII. Limitations

Absence of direct UCP measurements enforces use of event proxies; OMS labelling requires manual vetting until LLM is available internally; SS MLPs assume stationary healthy regimes and normal residuals; study cohorts, while representative, may not cover all regions; economic model excludes costs related to safety and reputational externalities, which are significant but complex.

VIII. Conclusion and Future Work

This work demonstrates a scalable predictive maintenance framework for LV hotspot anomalies using SABT/AMI data analytics. For secondary substations, the hybrid approach (analytical + general MLP + specific MLP) results in economic optimality (TPR 60.71%, FPR 1.33%, €944k/yr).

Whilst the analytical model has the highest individual KPI and one of the lowest computational costs, also being the most physically explainable. For underground cable pits, GMM-based event modelling attains TPR 46% at effective FPR 0.83% but is not yet break-even without cost externalities.

Future studies could include: enrich UCP telemetry (temperature/current, humidity/gases); integrate maintenance/work-order logs to reset counters for UCP modelling; explore more efficient learners (autoencoders, contrastive, lightweight RNN/LSTM), dynamic windows, and seasonality-aware thresholds for secondary substations; automate labelling with OMS text filter via LLM to obtain more training instances; expand economics to include safety risks and reputation costs.

ACKNOWLEDGMENT

The author thanks i-DE (Iberdrola) for data access and operational insights, ICAI - Universidad Pontificia Comillas and the University of Strathclyde for academic support. Academic and industrial supervisors that helped with model analysis and development with their technical and expertise insights: Dr. Miguel Ángel Sanz Bobi, Jesús Gutiérrez Serrano and Itziar Lumbreras Basagoiti.

REFERENCES

[1] Damiano Bracci and Simone Paglia. Advanced LV Management for Electrical Utilities. White paper, Gridspertise, 2023.

[2] M. Arnold, W. Friede, and J.M.A. Myrzik. Challenges in future distribution grids - a review. *Proc. ICREPQ*, 20:1–6, 03 2013.

[3] International Energy Agency. Unlocking Smart Grid Opportunities in Emerging Markets and Developing Economies. Technical report, International Energy Agency (IEA), February 2024.

[4] Enric R. Bartlett Castellà. Impulsar la digitalización del sistema eléctrico y el papel activo de los consumidores, February 2021.

[5] Dave Anny. Impact on Predictive Maintenance and Reliability Engineering. *ResearchGate*, 2023.

[6] François Mirallès, Luc Cauchon, Marc-André Magnan, François Grégoire, Mouhamadou Makhtar Dione, and Arnaud Zinflou. Towards reliable detection of dielectric hotspots in thermal images of the underground distribution network. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, page 3566–3574, New York, NY, USA, 2022. Association for Computing Machinery.

[7] Roberto Poyato. TERMOGRAFÍA EN SISTEMAS DE DISTRIBUCIÓN ELÉCTRICA. Nota técnica, Fluke Ibérica.

[8] Ana Carolina Martinazzo Kanitz, Euler Ribeiro, Marcos Aurélio Izumida Martins, Kleber Duarte Tomaz, and Silvia de Francisci. Hot spot analysis in asset inspections in the electricity distribution area. In *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pages 1–5, 2021.

[9] Insulated Conductors Committee. Ieee guide for early detection, mitigation, preventative measures, and response to smoke, fire, and explosions in underground electrical structures. *IEEE Std 2417-2022*, pages 1–69, 2022.

[10] Iberdrola. Buscamos alianzas para potenciar la detección de incidencias en nuestras redes de distribución, 2022.

[11] Philipp zur Heiden, Jennifer Priefer, and Daniel Beverungen. Predictive maintenance on the energy distribution grid–design and evaluation of a digital industrial platform in the context of a smart service system. *IEEE Transactions on Engineering Management*, 71:3641–3655, 2024.

[12] Muhammed Fatih Pekşen, Ulaş Yurtsever, and Yılmaz Uyaroğlu. Enhancing electrical panel anomaly detection for predictive maintenance with machine learning and iot. *Alexandria Engineering Journal*, 96:112–123, 2024.

[13] ABB. 4 types of maintenance strategy, which one to choose?, 2024.

[14] Endesa. Maintaining the power distribution grid, 2025.

[15] Amaia Amantegui Escribano. Anomalías en CT relacionadas con el riesgo incendios, May 2025. Article published in the 'Gestión del Cambio' at i-DE (Iberdrola).

[16] Akash Sharma and Rajive Tiwari. Anomaly detection in smart grid using optimized extreme gradient boosting with scada system. *Electric Power Systems Research*, 235:110876, 2024.

[17] Martin Neumayer, Dominik Stecher, Sebastian Grimm, Andreas Maier, Dominikus Bücker, and Jochen Schmidt. Fault and anomaly detection in district heating substations: A survey on methodology and data sets. *Energy*, 276:127569, 2023.

[18] Rawal Keerti and Ahmad Aijaz. *5 Big Data Analytical Techniques for Electrical Energy Forecasting in Smart Grid Paradigm*, pages 101–126. 2023.

[19] Malin Bülund. Exploring integration of predictive maintenance using anomaly detection: Enhancing productivity in manufacturing. Master's thesis, KTH Royal Institute of Technology, 2024.

[20] Ahish Shylendra, Priyesh Shukla, Saibal Mukhopadhyay, Swarup Bhunia, and Amit Ranjan Trivedi. Low power unsupervised anomaly detection by non-parametric modeling of sensor statistics. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2020.

[21] Practicus AI. Medium: The 5 clustering algorithms data scientists need to know, February 2018.

[22] Vatsal Raval. When to use MLP, CNN or RNN?, January 2019.

[23] Joaquín Amat Rodrigo. Detección de anomalías con Gaussian Mixture Model (GMM) y Python, December 2020.

[24] IEC-CENELEC. *Measuring Relays and Protection Equipment - Part 149: Functional Requirements for Thermal*

Electrical Relays (IEC 60255-149:2013). AENOR, Avenue Marnix 17, B-1000 Brussels, 2013.

[25] Iberdrola Group, STAR project i-DE Project Team. Et funcionalidad supervision avanzada bt – edition 9.6. Technical report, Iberdrola, March 2022. Internal technical report, Edition 9.6 dated 09/03/2022.

[26] Ming Zhou. Anomaly detection in smart distribution grids with deep neural network. Master's thesis, University of Alberta, 2022.

[27] Ming Zhou and Petr Musilek. Real-time anomaly detection in distribution grids using long short term memory network. In *2021 IEEE Electrical Power and Energy Conference (EPEC)*, pages 208–213, 2021.

[28] Quality Gurus. Nelson Rules (and Western Electric Rules) for Control Charts, 2020.

[29] Jiawei Han, Micheline Kamber, and Jian Pei. 11 - advanced cluster analysis. In Jiawei Han, Micheline Kamber, and Jian Pei, editors, *Data Mining: Concepts and Techniques (Third Edition)*, The Morgan Kaufmann Series in Data Management Systems, pages 497–541. Morgan Kaufmann, Boston, third edition edition, 2012.