Hybrid Transformer Prognostics Framework for Enhanced Probabilistic Predictions in Renewable Energy Applications

Jose Ignacio Aizpurua, *Senior Member, IEEE*, Ibai Ramirez, Iker Lasa, Luis del Rio, Alvaro Ortiz, and Brian G. Stewart *Member, IEEE*

***Abstract*—The intermittent nature of renewable energy sources (RESs) hamper their integration to the grid. The stochastic and rapid-changing operation of RES technologies impact on power equipment reliability. Transformers are key integrative assets of the power grid and it is crucial to monitor their health for the reliable integration of RESs. Existing models to transformer lifetime estimation are based on point forecasts or steady-state models. In this context, this paper presents a novel hybrid transformer prognostics framework for enhanced probabilistic predictions in RES applications. To this end, physics-based transient thermal models and probabilistic forecasting models are integrated using an error-correction configuration. The thermal prediction model is then embedded within a probabilistic prognostics framework to integrate forecasting estimates within the lifetime model, propagate associated uncertainties and predict the transformer remaining useful life with prediction intervals. Prediction intervals vary for each prediction according to the propagated uncertainty and they inform about the confidence of the model in the predictions. The proposed approach is tested and validated with a floating solar power plant case study. Results show that, from the insulation degradation perspective, there may be room to extend the transformer useful life beyond initial lifetime assumptions.**

***Index Terms*—Distribution transformers, prognostics, hybrid model, probabilistic forecasting, uncertainty.**

I. Introduction

THE reliable operation and integration of renewable energy sources (RESs) to the power grid is crucial [1]. However, complex and intermittent RES dependencies on weather conditions, operation dynamics, such as fast-switching transient events, and ever-dominant power-electronics reliability complicates the efficient and reliable operation of RESs and associated power equipment [2].

Prognostics and health management (PHM) is at the hearth of condition monitoring technology, where datasets and engineering knowledge cooperate to develop anomaly detection, diagnostics and prognostics solutions [3]. Failure prognostics aims to predict the remaining useful life (RUL) of assets to assist asset-management teams in condition-based monitoring and maintenance activities and is the focus of this work [3].

J. I. Aizpurua is with the Electronics & Computing Department, Mondragon University, Arrasate, Spain and also with IKERBASQUE, Basque Foundation For Science, Bilbao, Spain (e-mail: jiaizpurua@mondragon.edu);

I. Ramirez is with Electronics & Computing Department, Mondragon University, Arrasate, Spain (e-mail: iramirezg@modragon.edu);

I. Lasa, L. Rio and A. Ortiz are with Ormazabal, Boroa, Spain (e-mail: {ilo, lre, aog}@ormazabal.com);

B. G. Stewart is with the University of Strathclyde, Glasgow, UK (e-mail: brian.stewart.100@strath.ac.uk);

Transformers are key components for the efficient operation of power and energy systems, and it is crucial to monitor their health for the reliable integration of RESs to the grid [4]. The degradation mechanisms of transformer subsystems are complex, surrounded by uncertainty, and in some cases, not directly measurable [5]. This is the case of the insulation degradation, which is one of the major transformer failure cause [6]. The main factor that affects insulation degradation is the hottest-spot temperature (HST) [7]. The transformer insulation heat is distributed over different surface areas. This complicates the temperature estimation process, which is often estimated from indirect measurements. The HST may be combined with other parameters which can accelerate the transformer insulation ageing rate. For example, the presence of moisture under high loading can lead to bubble formation and potential catastrophic failure [8].

HST estimation models have been integrated with transformer lifetime models to predict the transformer RUL. These models, specified as deterministic and steady-state models, have been used for transformer prognostics predictions in distribution networks [9], [10], electric vehicles [11], photovoltaic (PV) generation [12], [13], [14], [15], wind energy [16], nuclear power plants [17], [18] and smart grid infrastructure [19]. However, these formulations do not integrate uncertainties present in the lifetime prediction [9], [10], [11], [12], [13], [14], [15], [16], [19], and result in point forecasts, *i.e.* RUL forecasts assume a deterministic relation between measurements, models and predictions. Uncertainty modelling is crucial for transformer RUL estimation and the uncertainty will only increase with the increasing amount of RESs.

In order to properly propagate uncertainty, it is necessary to model and propagate predictions as full probability density function (PDF) of plausible values of the estimated parameter consistent with the underlying modelled process. In this context, probabilistic forecasting approaches are gaining traction with predictive monitoring solutions [20], [21]. Bracale *et al.* presented a dynamic thermal rating (DTR) approach, which uses probabilistic forecasting estimates to predict ambient temperature and load, and calculate the maximum hourly transformer load that does not pass a given HST threshold [20]. This approach is continued in [21] with a probabilistic transformer risk management tool, which based on DTR threshold criteria determines if an alarm will be activated. These methods implement probabilistic forecasting models for transformer DTR applications, but they do not focus on

uncertainty modelling, propagation and RUL estimation.

Furthermore, it can be observed that health-state estimation methods focus on steady-state assessments with an hour-based temporal resolution, assuming constant health-state degradation between samples [11], [13], [14], [16], [17], [18], [19], [20], [21]. However, with the intermittency of RESs, this assumption may not adhere to real operation and degradation conditions of RESs and it may be beneficial to adopt thermal models which depart from steady-state operation [22]. Given the costs of power interruptions that may cause transformer outages, asset management teams should rely upon informative and accurate predictive distributions.

Recent prognostics developments focus on data-driven methods, which can generate accurate RUL predictions [23]. However, their ability to account for uncertainties is limited [3]. In this context, novel hybrid configurations have emerged combining physics and data-driven models [24], [25], which obtain accurate and explainable prediction results. Accordingly, this research aims to contribute in the development of novel hybrid transformer prognostics solutions.

To the best of the authors' knowledge, probabilistic forecasting models along with physics-based thermal models have not been integrated with prognostics models, and accordingly, the main contribution of this paper is the development of a novel hybrid transformer prognostics model.

The model integrates physics-based and probabilistic forecasting models in a error-correction configuration adapted for transformers operated in RES applications. The approach is embedded in a probabilistic RUL estimation framework to integrate the hybrid forecasting estimates in the lifetime model and propagate associated uncertainties. The integration of uncertainty enables designing a prognostics approach which provides uncertainty bounds according to the model confidence. The approach is tested and validated with distribution transformer data operated on a floating solar power plant.

The proposed approach departs from previous work [17], [18], focused on steady-state modelling in nuclear power plants, through the integration of probabilistic forecasting and transient thermal models in the probabilistic prognostics framework. The approach informs asset-management teams on how they can improve transformer lifetime prediction practice by accurately considering uncertainty and fast-switching events coming from RESs, and hints for evaluating the influence of sampling rates on forecasting estimates.

The rest of this paper is organized as follows. Section II defines the transformer thermal and lifetime modelling. Section III presents the methodology, Section IV applies the approach to case study, and Section VII draws conclusions.

II. Empirical Background

The insulation is the most critical part of the transformer which can cause its failure [6]. The winding hottest-spot temperature (HST) determines the insulation ageing, and accordingly, insulation health can be monitored through the HST. Moving beyond steady-state estimations, this work focuses on transient thermal modelling and lifetime estimation [6].

A. Transient thermal model

Differential equations are used to estimate the transformer transient HST defined as follows:

$$\Theta_{\mathrm{H}}(t) = \Theta_{\mathrm{TO}}(t) + \Delta\Theta_{\mathrm{H}}(t) \tag{1}$$

where $\Delta\Theta_{\mathrm{H}}(t)$ is the HST rise over top-oil temperature (TOT) and $\Theta_{\mathrm{TO}}(t)$ is the TOT. The TOT is defined as follows:

$$\Delta\Theta_{\mathrm{TO,R}}\left(\frac{1+K(t)^2R}{1+R}\right)^x = k_{11}\tau_{\mathrm{TO}}\frac{d\Theta_{TO}(t)}{dt} + \Theta_{\mathrm{TO}}(t) - \Theta_{\mathrm{A}}(t) \tag{2}$$

where $K(t) = i(t)/i_{\mathrm{r}}$, $i(t)$ being the load at instant t and i_{r} the rated load, R is the ratio of load losses to no-load losses, x is the oil exponent constant, which models the exponential power of total losses with respect to TOT heating, τ_{TO} is the oil time constant, $\Delta\Theta_{\mathrm{TO,R}}$ is the TOT rise at rated load, k_{11} is a thermal constant determined through experimentation, and $\Theta_{\mathrm{A}}(t)$ is the ambient temperature.

The HST rise is calculated as:

$$\Delta\Theta_H(t) = \Delta\Theta_{H_1}(t) - \Delta\Theta_{H_2}(t) \tag{3}$$

where $\Delta\Theta_{H_1}$(t) models the oil heating without HST variations:

$$k_{21}\Delta\Theta_{H,R}K(t)^y = k_{22}\tau_w\frac{\Delta\Theta_{H_1}(t)}{dt} + \Delta\Theta_{H_1}(t) \tag{4}$$

$\Delta\Theta_{H_2}(t)$ models the HST variations:

$$(k_{21}-1)\Delta\Theta_{H,R}K(t)^y = (\tau_{\mathrm{TO}}/k_{22})\frac{\Delta\Theta_{H_2}(t)}{dt} + \Delta\Theta_{H_2}(t) \tag{5}$$

where τ_{w} is the winding time constant, $\Delta\Theta_{\mathrm{H,R}}$ is the HST rise at rated load, y is the winding exponent constant, which models the loading exponential power with the heating of the windings, k_{21} and k_{22} are the transformer thermal constants.

Fig. 1 synthesizes the transformer thermal model, where s is the Laplace variable. Constants $k_{11}, k_{21}, k_{22}, \tau_O, \tau_w$, are determined through a heat-run test, and x, y are determined through extended thermal experiments.

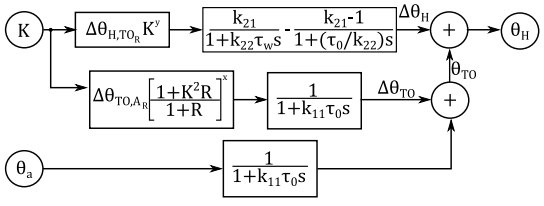


Fig. 1: Transient thermal model — adapted from [6].

Assuming d/dt is a small time difference, differential equations can be turned into difference equations, so that it is feasible to estimate the transient thermal values. The TOT in (2) is calculated follows [6]:

$$\begin{aligned}\Theta_{\mathrm{TO}}(t) &= D\Theta_{\mathrm{TO}}(t) + \Theta_{\mathrm{TO}}(t-1)\\ D\Theta_{\mathrm{TO}}(t) &= \frac{\Delta t}{k_{11}\tau_{\mathrm{TO}}}\left[\Delta\Theta_{H,R}\left(\frac{1+K(t)^2R}{1+R}\right)^x + \Theta_A(t) - \Theta_{\mathrm{TO}}(t-1)\right]\end{aligned} \tag{6}$$

where D denotes a difference operation of the associated variable within the period Δt. Similarly, (4) and (5) can be rewritten as follows:

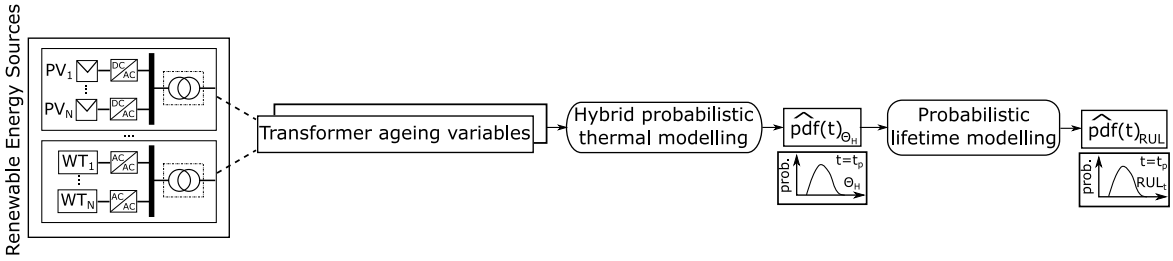


Fig. 2: Conceptual diagram of the hybrid probabilistic prognostics estimation framework.

$$\Delta\Theta_{H_i}(t) = D\Delta\Theta_{H_i}(t) + \Delta\Theta_{H_i}(t-1) \tag{7}$$

for i={1,2}, where

$$\begin{aligned} D\Delta\Theta_{H_1}(t) &= \frac{\Delta t}{k_{22}\tau_w}\left[k_{21}\Delta\Theta_{\mathrm{H,R}}K(t)^y - \Delta\Theta_{H_1}(t-1)\right] \\ D\Delta\Theta_{H_2}(t) &= \frac{k_{22}\Delta t}{\tau_{\mathrm{TO}}}\left[(k_{21}-1)\Delta\Theta_{\mathrm{H,R}}K(t)^y - \Delta\Theta_{H_2}(t-1)\right] \end{aligned} \tag{8}$$

Finally, assuming steady-state initial conditions, time-derivatives in (2), (4), and (5) are null, and the initial conditions are defined as follows:

$$\begin{aligned} \Theta_{\mathrm{H}}(0) &= \Theta_{\mathrm{TO}}(0) + k_{21}\Delta\Theta_{\mathrm{H,R}}K(0)^y - (k_{21}-1)\Delta\Theta_{\mathrm{H,R}}K(0)^y \\ \Theta_{\mathrm{TO}}(0) &= \Theta_{\mathrm{A}}(0) + \Delta\Theta_{\mathrm{TO,R}}\left(\frac{1+K(0)^2R}{1+R}\right)^x \end{aligned} \tag{9}$$

Δt should be as small as possible, and never greater than the half of the smaller time constant.

B. Lifetime model

The IEC 60076-7 standard defines insulation paper aging acceleration factor at time t, $F_{\mathrm{AA}}(t)$, as [6]:

$$F_{\mathrm{AA}}(t) = 2^{(98-\Theta_H(t))/6} \tag{10}$$

The RUL at t, $RUL(t)$, can be obtained by turning (10) into a Markovian relation model:

$$RUL(t) = RUL(t-\Delta t) - F_{\mathrm{AA}}(t) = RUL(t-\Delta t) - 2^{(98-\Theta_H(t))/6} \tag{11}$$

The lifetime model starts from the initial RUL estimate, RUL_0, which is iteratively re-calculated and updated to model the latest RUL estimate.

III. Hybrid Probabilistic Prognostics Framework

Fig. 2 shows the conceptual block-diagram of the developed hybrid probabilistic transformer prognostics framework, based on the thermal and lifetime modelling processes described in Section II with the integration of probabilistic forecasting and lifetime estimation stages. The focus of the proposed method is on the distribution transformers that are used to connect RESs and the grid.

The hybrid probabilistic thermal modelling block receives and processes transformer ageing variables and integrates physics-based analytic models and data-driven error-correction models, including different sources of uncertainty. At the prediction time instant, t_p, this block estimates the PDF of the HST, $\hat{pdf}(t)_{\Theta_H}$, and it will input the probabilistic lifetime modelling phase along with different sources of uncertainty, including the initial health state and the process degradation uncertainty. The final outcome of the hybrid probabilistic framework will be the PDF of the transformer RUL, $\hat{pdf}(t)_{RUL}$.

A. Hybrid Probabilistic Thermal Modelling

Fig. 3 shows the hybrid thermal modelling approach.

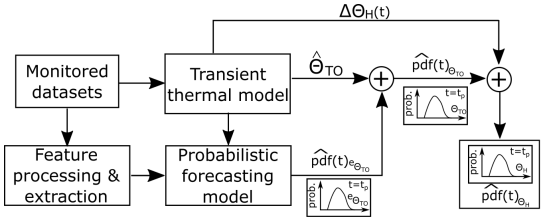


Fig. 3: Hybrid-probabilistic transient thermal modelling.

The transient thermal model estimates the deterministic HST using physics-based difference equations (cf. Section II). This estimate is corrected with a data-driven probabilistic error-correction phase through the probabilistic estimation of the physics-based model error:

$$\hat{pdf}_{\Theta_{\mathrm{H}}}(t) = \hat{\Theta}_{\mathrm{TO}}(t) + \hat{pdf}_{e_{\Theta\mathrm{TO}}}(t) + \Delta\Theta_H(t) \tag{12}$$

where, $\Delta\Theta_H(t)$ is the HST rise over TOT [cf. (8)], $\hat{\Theta}_{\mathrm{TO}}(t)$ is the physics-based TOT estimate, and $\hat{pdf}_{e_{\Theta\mathrm{TO}}}(t)$ is the PDF of the residual error, which is estimated using probabilistic forecasting methods through the function $f_{ML}(.)$:

$$\hat{pdf}_{e_{\Theta\mathrm{TO}}}(t) = f_{ML}(f_1(t-k), \ldots, f_N(t-k)) \tag{13}$$

where $\langle f_1(t-k), \ldots, f_N(t-k)\rangle$ denotes the set of features obtained at the time instant $t-k$, that best represent the error estimate at instant t.

The probabilistic forecasting stage focuses on one-step ahead predictions. Owing to the lack of monitored HST data, the data-driven error correction phase focuses on the TOT estimate and for the HST estimation physics-based equations are employed. In order to train the probabilistic forecasting model, the historic TOT time-series must be first estimated through the physics-based difference equations and compute the error (cf. Section II).

1) Probabilistic Forecasting Models: The probabilistic thermal modelling phase aims to predict the HST under uncertainty. This is performed in two connected steps (i) TOT prediction via probabilistic forecasting models and (ii) HST estimation via experimental models. The uncertainty associated with the TOT prediction is modelled with $\hat{pdf}(t)_{\Theta_{TO}}$, which is propagated to estimate the PDF of the HST $\hat{pdf}(t)_{\Theta_H}$ through (12). Among forecasting models that integrate uncertainties and produce a PDF for each prediction estimate, this paper focuses on Quantile Regression Forest (QRF), Quantile Gradient Boosting (QGB) and Quantile Regression (QR) [26].

QRF are based on Random Forests (RF) with a powerful predictive performance [27]. A RF grows an ensemble of regression trees using N independent observations $\{y, x_i\}_{i=1}^{N}$, where y is the TOT prediction error, $e_{\Theta_{TO}}(t)$, $x_i \in X$ are the error prediction variables, $\langle f_1(t-k), \ldots, f_N(t-k)\rangle$, and N is the length of training data. QRFs draw prediction intervals from RF predictions. The prediction becomes a conditional distribution function $F(y|X = x_i)$, *i.e.* probability of the predicted TOT error value y, given the error prediction variables X, which is approximated by the weighted mean of x_i over the observations:

$$\hat{F}(y|X = x_i) = \sum_{i=1}^{N} w_i(x_i) 1_{\{y<y_i\}} \tag{14}$$

where $w_i(x_i) = K^{-1} \sum_{k=1}^{K} w_i(x_i\theta_k)$ is the weighted vector and $1_{\{.\}}$ is an indicator function.

The ϕ-quantile, $q_\phi(x_i) = \phi$, is defined such that the probability of $y < q_\phi(x_i) = \phi$. That is, a quantile of order ϕ is a value where the distribution function crosses ϕ. Quantiles provide a complete distribution information of y as a function of explanatory features X. For example, for a feature set, X, and target variable y, 80% prediction intervals (PI) are estimated as:

$$PI(X) = [q_{0.1}(y|x = X), q_{0.9}(y|x = X)] \tag{15}$$

Fig. 4 shows a distribution function example, including quantiles, where the interval [a, c] covers 80% PI.

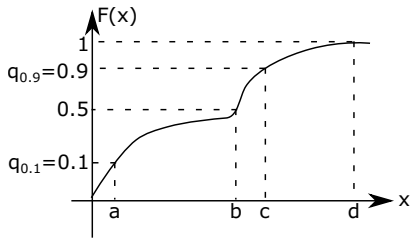


Fig. 4: Cumulative distribution function (CDF) and quantiles.

Using quantiles and interpolation methods it is possible to build a CDF. In this case the monotonic piecewise interpolation method was used [28] — see example in Fig. 11.

QRF is implemented in the `skgarden` Python package. Model tuning has been done from a predefined grid of parameters including the minimum number of samples required to split an internal node (min_samples_split), minimum number of samples required to be at a leaf node (min_samples_leaf) and number of trees (n_estimators). These parameters aim at preventing overfitting and were considered as follows: min_samples_split=[5, 10, 20], min_samples_leaf=[1, 2, 5] and n_estimators=[10, 100, 1000, 2000].

QGB are based on boosting methods that sequentially combine an ensemble of weak learners as a weighted sum of base-learner models to reduce the ensemble error [29]:

$$\hat{y}_t = F_N(x_t) + \varepsilon_t = \sum_{n=1}^{N} f_n(x_t) + \varepsilon_t \tag{16}$$

where $F_N(x_t)$ is the ensemble of N regression trees, each $f_n(x_t)$ is a regression tree and ε_t is an error term. The new regression tree $f_{n+1}(x_t)$ for the quantile loss function $L(.)$ is estimated as follows:

$$\underset{f_{n+1}}{argmin} \sum_{t} L(y_t, F_N(x_t) + f_{n+1}(x_t)) \tag{17}$$

This optimization is solved through the steepest descent algorithm [29], where each $f_n(x_t)$ is designed to be maximally correlated $F_N(x_t)$. The implementation of the quantile loss function enables the probabilistic prediction [30].

QGB is implemented in the `sklearn` Python package. Model tuning has been done by adjusting the learning rate, which controls the contribution of each regression tree (learning_rate), the number of boosting iterations which defines the number of base learners in the final model (n_estimators) and maximum depth of the individual regression estimators (max_dept). These parameters have been evaluated from a predefined grid of parameters: learning_rate=[0.01, 0.05, 0.1], n_estimators=[100, 250, 500] and max_depth=[2, 5, 8].

QR is an extension of linear regression (LR) models, which assume a linear model as follows [31]:

$$\hat{y}_t = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_K X_K + \varepsilon \tag{18}$$

where $\hat{y}$ is the response variable and $X = \{X_1, \ldots, X_K\}$ is the set of explanatory variables and ε is the error often represented as zero-mean Gaussian variable.

Prediction errors in the LR model are denoted residuals, and the ordinary least square minimization procedure is used to find the optimal parameters β of the LR model that minimize the squared sum of residuals. Instead of focusing on the mean value as in LR estimate, QR generates quantiles of the target value [32]. The QR model of the quantile ϕ is:

$$Q_\phi(\hat{y}_t) = \beta_0(\phi) + \beta_1(\phi) X_1 + \ldots + \beta_K(\phi) X_K + \varepsilon \tag{19}$$

where it can be observed that β coefficients are functions of a quantile value, which are determined through minimization of the mean absolute deviation (MAD):

$$MAD = \frac{1}{N} \sum_{i=1}^{N} \rho_\phi(y_t - \beta_0(\phi) + \beta_1(\phi) X_1 + \ldots + \beta_K(\phi) X_K + \varepsilon) \tag{20}$$

where ρ_ϕ is a function which weights the loss function according to the weight of the quantile.

QR is implemented in the `statsmodel` Python package, solved through iterative weighted least squares, Epanechnikov kernel, robust standard errors and Hall-Sheather bandwidth selection, maximum iterations=1e5 and tolerance error=1e-3.

2) Preprocessing & Feature Extraction: Data preprocessing steps are comprised of filtering and normalization. Outliers and invalid sensor readings caused by unknown phenomena, *i.e.* missing-at-random, are pre-processed with classical imputation approaches using a rolling window mean value.

Feature selection has been driven by the analytics relation between the TOT and explanatory variables, including lagged signals due to the oil constant propagation τ_{TO}. From (2) it is observed that TOT is related with the ambient temperature and input load. As the transformer is operated in a RES plant, the influence of the natural resource on the generated energy is also relevant, and potentially helpful for TOT forecasting. Assuming that the transformer is operated in a PV plant, the power generated by a PV module can be defined as [33]:

$$P(t) = \eta A R(t) sin(\alpha) \tag{21}$$

where η is the efficiency, α is the PV module angle, A is the module area $[m^2]$ and $R(t)$ is the solar irradiance $[W/m^2]$.

Accordingly, the solar irradiance is included as an explanatory variable, which is obtained from the numerical weather predictions (NWPs) produced by the European Centre for Medium-Range Weather Forecasts [34]. NWPs are based on a set of mathematical equations, which describe the physical state and dynamic motion of the atmosphere [34]. The most up-to-date solar irradiance estimate at instant t is used, $\hat{R}(t)$.

The remainder of features are one step lagged signals, $\langle \theta_A(t-1), i(t-1) \rangle$, oil constant lagged load and ambient temperature signals $\langle i(t-\tau_{TO}), \theta_A(t-\tau_{TO}) \rangle$, and finally, the cross-correlation between load and TOT, and ambient temperature and TOT has been analysed. The cross-correlation finds the empirical lag between signals, τ_X, in which the correlation between the analysed signals is maximum $\langle i(t-\tau_X), \theta_A(t-\tau_X) \rangle$. The set of extracted features are finally post-processed to create different predictive models and select the optimal model structure (see Sec. V-A).

3) Performance Metrics: The main criteria to evaluate the probabilistic forecasts has been the continuously ranked probability score (CRPS). From the probabilistic forecast PDF, $f(z)$, with its cumulative distribution function (CDF), $F(z)$, and observation y, the $CRPS(F, y)$ is defined as [35]:

$$CRPS(F, y) = \int_{\mathbb{R}} \left(F(z) - \mathbb{1}\{y \leq z\} \right)^2 dz \tag{22}$$

where $\mathbb{1}\{y \leq z\}$ is the indicator function, which is one if $y \leq z$ and zero otherwise.

The CRPS calculates the discrepancy between the forecast CDF and the empirical CDF of the observation $\mathbb{1}\{y \leq z\}$ which is considered as a step function because the observations y are deterministic values. The CRPS is a better suited metric for categorizing the predictive power of probabilistic models that generate a PDF as a prediction because it quantifies the error of each probabilistic value with respect to the observation [35]. Other metrics, such as the mean average error (MAE), are based on deterministic point estimate error values and, thus, they have to quantify the MAE value for every quantile to have an overall idea of the predictive power.

B. Probabilistic Lifetime Modelling

1) Sources of Uncertainty: The transformer hottest-spot temperature is estimated from indirect measurements, which may include measurement errors. This is captured through the probabilistic thermal model in (12). Furthermore, the insulation degradation model may be inaccurate due to the empiric-process based specification [cf. (11)], along with the initial insulation estimate. Integrating these sources of uncertainty into the lifetime model results in:

$$RUL(t) = RUL(t-\Delta t) + \omega_{\mathrm{RUL}_{t-\Delta t}} - 2^{\left(98 - pdf_{\Theta_{\mathrm{H}}}(t)\right)/6 + \omega_{\mathrm{t}}} \tag{23}$$

where $\omega_{\mathrm{RUL}_{t-\Delta t}}$ denotes the uncertainty of the RUL estimate at $t-\Delta t$, ω_{t} is the variability of the degradation trajectory and $pdf_{\Theta_{\mathrm{H}}}(t)$ is defined in (12). The initial error estimate, ω_{RUL_0}, will be iteratively propagated via (23).

By checking deterministic lifetime formulation (11) with (23), one can infer that HST and RUL estimates are influenced by different sources of uncertainty.

2) Probabilistic Lifetime Framework: The goal of this phase is the transformer RUL estimation under uncertainty. To this end, hybrid probabilistic forecasting model results from the thermal modelling phase are connected with experimental degradation equations. This is achieved through a Particle Filtering (PF), which is a tracking approach for state-estimation under uncertainty using Bayesian inference [36].

Let k be a discrete time-step, $k \in \mathbb{N}$, PF diagnoses the transformer health state x_k through a iterative combination of the health degradation model $f(\cdot)$ and its influencing measurements $h(\cdot)$:

$$\begin{aligned} x_{\mathrm{k}} &= f(x_{\mathrm{k\text{-}1}}, \upsilon_{\mathrm{k\text{-}1}}) \\ z_{\mathrm{k}} &= h(x_{\mathrm{k}}, \varphi_{\mathrm{k}}) \end{aligned} \tag{24}$$

where x_{k} is the transformer RUL at the discrete time instant k, $f(\cdot)$ is the insulation degradation [cf. (23)], υ_{k} is the degradation uncertainty vector $\upsilon_{\mathrm{k}} = \langle \omega_{\mathrm{k}}, \omega_{\mathrm{RUL_k}} \rangle$, z_k is the HST at time instant k, $h(\cdot)$ is the HST estimation function that correlates transformer health measurements with the RUL estimate, and φ_{k} is the measurement uncertainty vector.

The HST model integrates load measurement errors, probabilistic distribution of the TOT forecasting estimate including uncertainties along with transformer design parameters, and it estimates the PDF of the HST, $\hat{pdf}(t)_{\Theta_{H_k}}$.

The insulation degradation function in (23) integrates the process noise ω_{k} and calculates the insulation RUL from the HST and initial health state, which is then iteratively updated. Fig. 5 shows the probabilistic lifetime modelling approach.

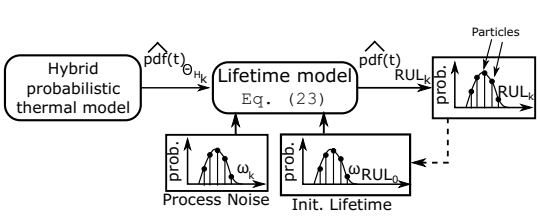


Fig. 5: Probabilistic lifetime modelling.

At a time $k-1$, assuming that the *prior* PDF, $p(x_{\mathrm{k\text{-}1}}|z_{0:\mathrm{k\text{-}1}})$, is available, the one-step ahead prediction is estimated from:

$$p(x_{\mathrm{k}}|z_{0:\mathrm{k}\text{-}1}) = \int p(x_{\mathrm{k}}|x_{\mathrm{k}\text{-}1})p(x_{\mathrm{k}\text{-}1}|z_{0:\mathrm{k}\text{-}1})dx_{\mathrm{k}\text{-}1} \tag{25}$$

where $p(x_{\mathrm{k}}|x_{\mathrm{k}\text{-}1})$ is the state-evolution defined in (24). Using the Bayes rule, when a new observation is collected at k, z_{k}, an updated state-estimate can be obtained [36]:

$$p(x_{\mathrm{k}}|z_{0:\mathrm{k}}) = p(x_{\mathrm{k}}|z_{0:\mathrm{k}\text{-}1})p(z_{\mathrm{k}}|x_{\mathrm{k}})/p(z_{\mathrm{k}}|z_{0:\mathrm{k}\text{-}1}) \tag{26}$$

The PDF $p(x_{\mathrm{k}}|z_{0:\mathrm{k}})$ models the solid insulation state x_{k} given measurements z_{k} up to k, which is synthesized in the recursive filtering expression:

$$p(x_{\mathrm{k}}|z_{0:\mathrm{k}}) = \underbrace{p(z_{\mathrm{k}}|x_{\mathrm{k}})}_{\text{Likelihood}} \int \underbrace{p(x_{\mathrm{k}}|x_{\mathrm{k}\text{-}1})}_{\text{Transition}} \underbrace{p(x_{\mathrm{k}\text{-}1}|z_{0:\mathrm{k}\text{-}1})}_{\text{Filtering}} dx_{\mathrm{k}\text{-}1} \tag{27}$$

The resolution of (27) is obtained through the PF approach via prediction, update and resampling steps [36], which has been adapted to include probabilistic forecasting estimates of the measurement function.

In order to model the initial transformer health state at k=0, N_p random samples, known as particles, are drawn $\{x_{k=0}^i\}_{i=1}^{N_p}$ from the initial transformer insulation health-state conditions. Without loss of generality, throughout this work N_p=5000 particles have been used achieving an acceptable trade-off between computational cost and variance [37].

Prediction. Transformer degradation-state samples are estimated from $p(x_{\mathrm{k}}|z_{0:\mathrm{k}\text{-}1})$. This is achieved through the insulation degradation model (24) integrating sources of uncertainty drawn from the corresponding probability density function. The result of the prediction step are a set of health-state samples x_k^i which are realizations of the predicted distribution.

In order to draw particles from the predicted HST, $\hat{pdf}(t)_{\Theta_{H_k}}$, the inverse sampling method is applied. Let $\hat{F(t)}_{\Theta_{H_k}}$ be the cumulative distribution function inferred from the probabilistic forecasting model, r_{Θ_H} the random variable drawn from the uniform distribution $r_{\Theta_H} \sim U([0,1])$, then the inverse sampling method applies the relation $\hat{F}_{\Theta_H}^{-1}(r_{\Theta_H}) = \hat{\Theta}_H$.

Update. For new observations at instant k, z_{k}, particle weights are assigned estimating their probability:

$$w_k^i = p(z_{\mathrm{k}}|x_k^i)/\sum_{j=1}^{N_{\mathrm{p}}} p(z_{\mathrm{k}}|x_k^j) \tag{28}$$

Resampling. In order to avoid the concentration of weights on one particle, known as weight degeneracy [36], a resampling threshold is defined: $N_{\mathrm{e}} = 1/\sum_{i=1}^{N_{\mathrm{p}}} w_k^i$. Particles are resampled if $N_{\mathrm{e}} < N_{\mathrm{p}}/2$ [36].

The PDF, $p(x_{\mathrm{k}}|z_{0:\mathrm{k}})$, analytically defined in (27), is now inferred from the weighted particles $\{x_k^i, w_k^i\}_{i=1}^{N_p}$.

IV. Case Study

Sierra Brava is a grid-connected floating PV plant operated by Acciona with the objective of evaluating the performance and operation and maintenance (O&M) costs. Floating PV plants shows an improved performance due to the decreased ambient temperature, high solar irradiation and less shading [38]. These factors, along with the reuse of existing grid connections, may compensate the initial installation costs.

The installation is located close to the southern shore of the Sierra Brava reservoir (Cáceres, Spain). Fig. 6 shows the layout configuration. Designed to cover around 12000 m^2, the installation consists of 5 floating systems with different configuration (orientation, inclination). Each system has 600 PV panels with an estimated capacity of 1.125 MW peak.

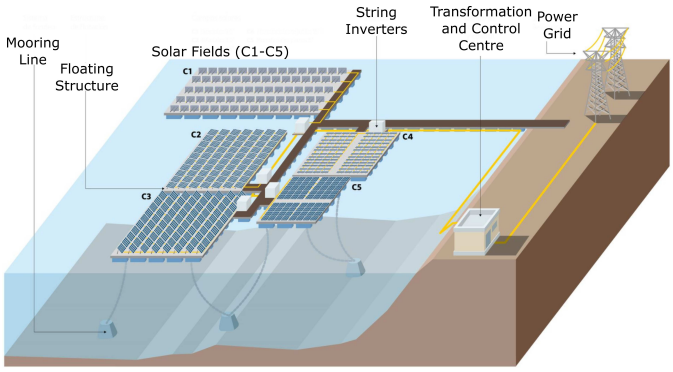


Fig. 6: Sierra Brava floating solar PV configuration.

Monitored data includes minutely sampled 63 days from 30/05/2021 to 31/07/2021. After cleaning and preprocessing, results in 60 days (86400 samples per variable). In addition to transformer-specific monitored variables, solar irradiance data has been extracted from meteorological databases [34].

The focus of the case study is on the O&M evaluation of the distribution transformer located in the transformation centre. Table I displays the transformer nameplate rating values.

TABLE I: Transformer parameter values.

Parameter	Value
Rating [kVA], V_1/V_2	1100, 22000/400
R=Load losses/No load losses [W]	9800/842
$\Delta\Theta_{H,R}, \Delta\Theta_{TO,R}$ [°C]	15.1, 54.26
k_{11}, k_{21}, k_{22}	0.75, 2.32, 2.05
τ_0, τ_w [min.]	266.8, 9.75

A. Transformer heat-run tests

A short-circuit heat-run test was carried out to determine the transformer top-oil and winding temperature rise [39]. A voltage test is performed so that measured active power is equal to the total losses and held constant until reaching the steady-state. Next, the load is reduced to the rated current and the winding resistance is measured as the transformer cools down to determine the average winding temperature. After this process, thermal parameters k_{11}, τ_0, k_{21}, k_{22} and τ_w are determined through the following functions [6]:

$$f_1(t) = 1 - e^{-\frac{t}{k_{11}\tau_0}} \tag{29a}$$

$$f_2(t) = k_{21}\left(1 - e^{-\frac{t}{(k_{22}\cdot\tau_w)}}\right) - (k_{21}-1)\left(1 - e^{-\frac{t}{(\tau_0\,/\,k_{22})}}\right) \tag{29b}$$

with the robust optimization problem defined as:

$$\underset{k_{11},\tau_0}{\text{argmin}}\{\sum_{i=0}^{n-1}\rho\left(\left(f_1(t_i,k_{11},\tau_0)-\frac{\Theta_{\text{TO}i}-\Theta_{\text{A}i}}{\Delta\Theta_{TO,R}}\right)^2\right)\}, \quad (30a)$$

$$\underset{k_{21},k_{22},\tau_w}{\text{argmin}}\{\sum_{i=0}^{n-1}\tilde{\rho}\left(\left(f_2(t_i,k_{21},k_{22},\tau_w,\tau_0^*)-\frac{\Theta_{\text{H}i}-\Theta_{\text{TO}i}}{\Delta\Theta_{H,R}}\right)^2\right)\} \quad (30b)$$

where the loss functions ρ and $\tilde{\rho}$, are used to smooth the effect of outliers via least square error.

The trust-region-reflective algorithm is used for robust non-linear optimization [40], with initial values and parameter bounds in agreement with [6] and a tolerance of 10^{-8} for convergence. Using the resulting optimal oil time constant, τ_0^*, thermal parameters k_{21}, k_{22}, τ_w are fitted to the winding temperature rise according to f_2 in (29b). Results are given in Table I and the fitting is shown in Fig. 7.

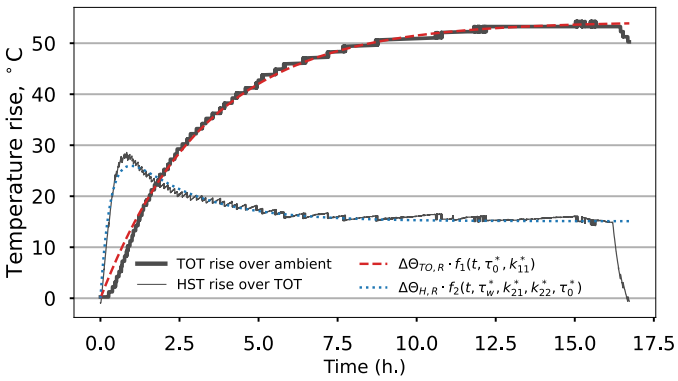


Fig. 7: Heat-run test based thermal parameter fitting.

Thermal parameters of the HST have been collected through experimental heat-run tests. However, note that the monitored HST was not available.

V. Numerical Applications

A. Hybrid Probabilistic Thermal Modelling

The accuracy of the transient thermal model is enhanced via probabilistic forecasting and dynamic error-correction. The dataset has been divided into training and testing sets, with a proportion of 70% and 30%, respectively. The training set has been used for feature selection and hyper-parameters tuning via 3-fold cross validation, and the testing set has been reserved to examine the predictive capacity on unseen data. Fig. 8 shows the monitored variables and the solar irradiation obtained from ERA5 [34]. Vertical dashed line divides the series data into train and test sets.

For the feature selection process, different models have been designed including different signals as shown in Table II. Monitored signals, $i(t-1), \theta_A(t-1)$, NWP solar irradiation, $\hat{R}(t)$ and oil constant lagged load, $i(t-\tau_{TO})$ signals have been considered across all the models, and combinations of oil constant lagged ambient temperature and cross-correlation lagged signals, $i(t-\tau_X), \theta_A(t-\tau_X)$, have been analysed.

The predictive power of the models have been analysed for different folds of the training set and evaluated through the mean CRPS, as shown in Fig. 9. For each model configuration

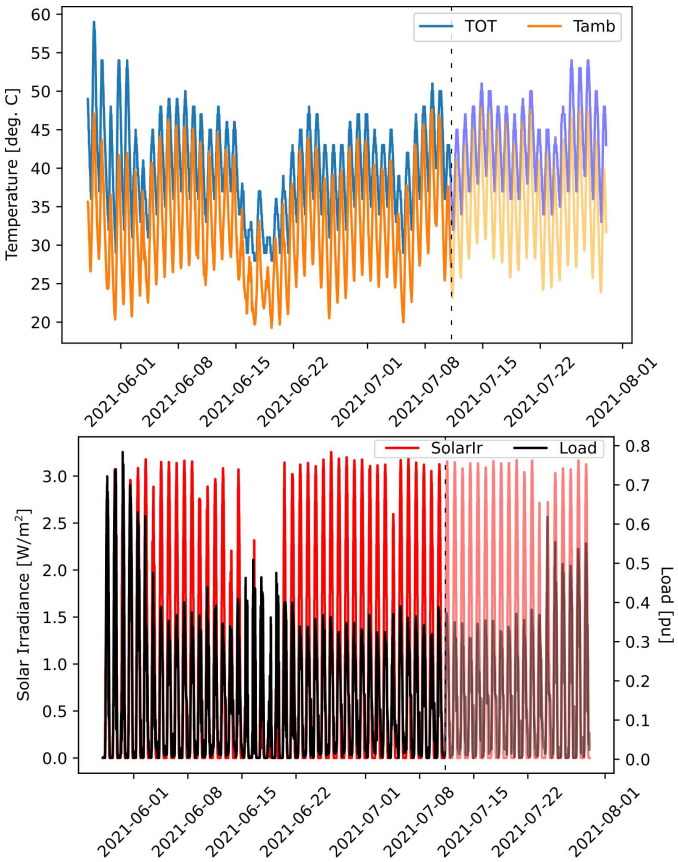


Fig. 8: Sierra Brava monitored time-series

.

TABLE II: Analysed models in the feature selection process.

Feature	Model Number							
	1	2	3	4	5	6	7	8
$i(t-1)$	✓	✓	✓	✓	✓	✓	✓	✓
$\theta_A(t-1)$	✓	✓	✓	✓	✓	✓	✓	✓
$\hat{R}(t)$	✓	✓	✓	✓	✓	✓	✓	✓
$i(t-\tau_{TO})$	✓	✓	✓	✓	✓	✓	✓	✓
$\theta_A(t-\tau_{TO})$	×	×	×	×	✓	✓	✓	✓
$i(t-\tau_X)$	×	✓	×	✓	×	✓	×	✓
$\theta_A(t-\tau_X)$	×	×	✓	✓	×	×	✓	✓

in Table II, two model types have been tested: (i) hybrid configurations (H) using QRF, QGB and QR probabilistic forecasting models as error-correction models and (ii) data-driven (DD) configurations without error-correction directly using QRF, QGB and QR as probabilistic forecasting models.

From Fig. 9 it can be observed that the performance of the hybrid models is superior to data-driven predictive models. Among the tested models, features in the sixth model configuration obtain the best mean result for the hybrid approaches, and the fifth and seventh model configurations for data-driven approaches, which are selected for subsequent modelling stages. Finally, it can be observed that the performance trends of QRF and QGB models are similar.

The cross-validation process implements iterative training and validation procedure on the set of parameters defined in the grid-search strategy (cf. Section III-A1), and reports the best model based on the loss function. The parameter-tuning

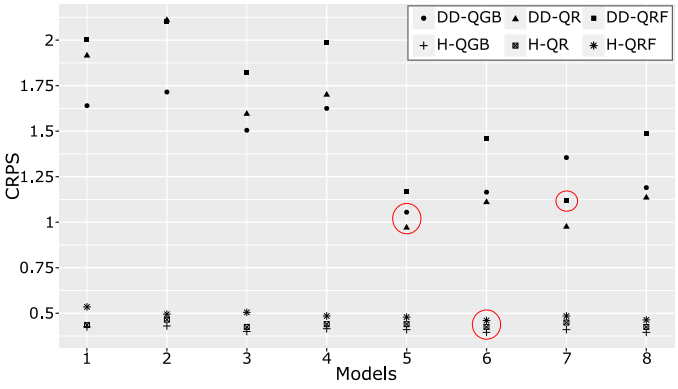


Fig. 9: Feature selection performance results

.

process has been implemented individually for each quantile prediction model including 99%, 95%, 90%, 50%, 10%, 5% and 1% prediction quantiles.

Table III displays probabilistic forecasting error results for probabilistic forecasting models including hybrid configurations, data-driven configurations and the analytic IEC model. Probabilistic prediction models have been evaluated with CRPS and MAE metrics, using the median probabilistic prediction values to infer the MAE, *i.e.* 50% prediction quantile. As for the IEC model, which provides deterministic point estimates, it is not possible to estimate CRPS values, and testing set prediction results have been directly used to estimate the top-oil temperature and the associated MAE.

TABLE III: TOT forecasting error results for various methods.

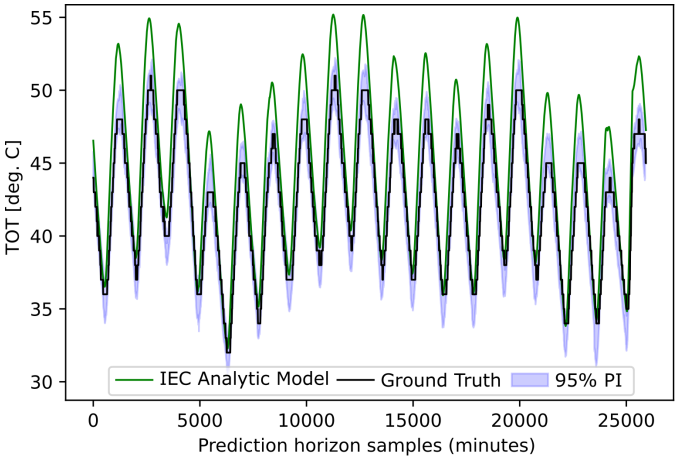
Method	CRPS	MAE
H-QGB	**0.19±0.13**	0.41±0.3
H-QR	0.25±0.16	0.53±0.38
H-QRF	0.26±0.22	0.51±0.39
DD-QGB	0.59±0.37	1.06±0.81
DD-QRF	0.83±0.95	1.43±1.77
DD-QR	0.97±0.39	1.5±1.13
IEC	2.56	

From Table III it can be observed that the best probabilistic forecasting configuration is the hybrid configuration using the QGB as the probabilistic error-correction model. The performance evaluation metric for probabilistic forecasting models is the CRPS. However, even using the MAE with median probabilistic prediction values, results are still coherent. The performance of data-driven configurations drops compared with error-correction models. Finally, the analytic IEC model shows deterministic TOT estimates, which performs worse due to the lack of statistical learning and adaptation steps.

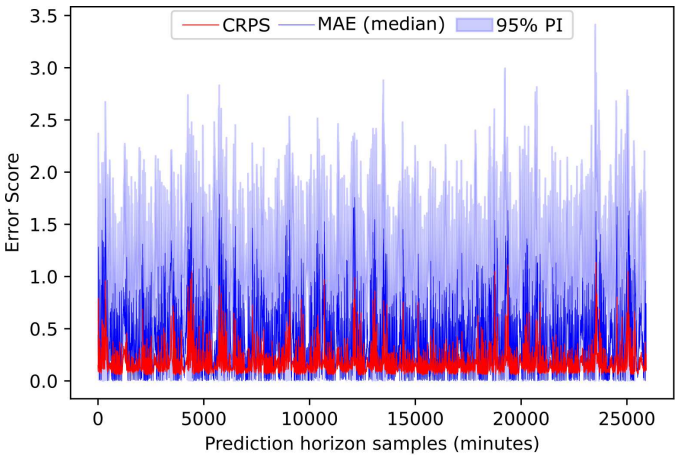
After selecting the best method, additional quantile models have been tuned and inferred for the H-QGB model to obtain a more accurate representation of the underlying distribution.

Fig. 10a shows the probabilistic forecasting estimates with 95% prediction intervals for the best hybrid-configuration, *i.e.* 18 days of testing (25920 minutes). Fig. 10b shows the corresponding CRPS and MAE results, calculated with the median and 95% prediction interval.

Fig. 10 shows that the hybrid prediction model confidence bounds are within the measured TOT values, which confirms



(a) Probabilistic TOT forecasting results.



(b) Error scores: CRPS & MAE (median ± 95% PI).

Fig. 10: One-step ahead probabilistic prediction results.

that effectively corrects the analytic model predictions.

Probabilistic predictions take the form shown in Fig. 11, with a probability value assigned to each error prediction value, with the median probabilistic prediction error of -5°C, and 95% prediction interval limits within -5.76°C to -4.3°C.

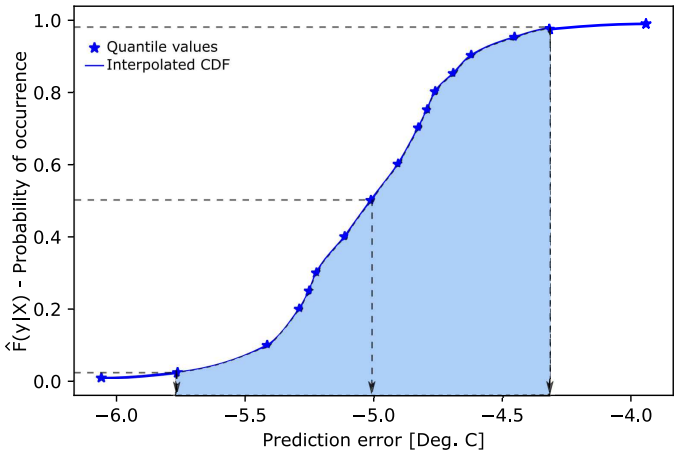


Fig. 11: Probabilistic error prediction results (1000-th sample).

Probabilistic error distributions vary for each prediction

according to the propagated uncertainty. Predicted error distribution informs about the model confidence on the predictions. This information is connected with probabilistic lifetime estimation models to predict the RUL (cf. Section V-B).

B. Prognostics modelling

Taking the H-QGB prediction model, two transformer degradation scenarios have been analysed:

(#1) Deterministic RUL model in (11), using measured TOT and load data and HST calculated via (1);

(#2) Application of the framework in Fig. 2, with probabilistic RUL model in (23), one-step ahead transient probabilistic TOT forecasting results and HST calculated via (12)

In both scenarios, an initial lifetime of 262800 hours is assumed (30 years) with a variance of 1 hour, $RUL_0 = N(262800, 1)$, and process degradation uncertainty of $w_k = N(0, 1)$ [6]. So as to evaluate the likely future degradation under these scenarios, the monitored operation profile of one week has been repeated over 6 months (cf. Fig. 12).

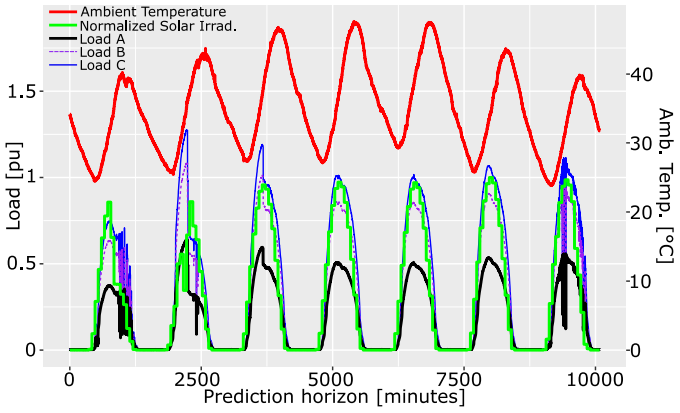


Fig. 12: Operational load profiles for one week.

Accordingly, Fig. 13 shows the obtained prognostics results using the operational data and reference load A in Fig. 12.

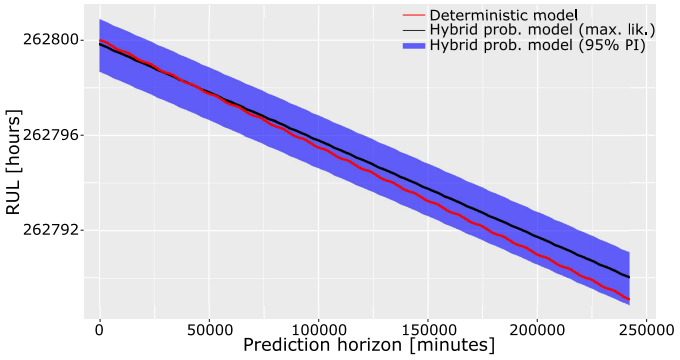


Fig. 13: Prognostics predictions for different configurations.

The degradation after 6 months of operation (4032 hours) is mild with a maximum likelihood value of 262790 hours and 95% prediction intervals of [262788.9, 262791.1] hours in #2 and 262789.1 hours in #1. The difference between both configurations is on the uncertainty information propagated by #2, which informs about the likelihood and confidence of the predicted RUL values (cf. load A in Fig. 14).

So as to test the model with other operation contexts, two different loading profiles have been analysed (cf. load B and C, Fig. 12). Fig. 14 shows the obtained PDF results.

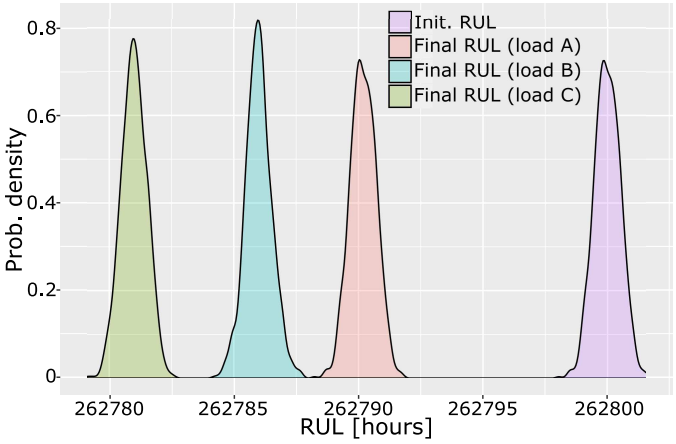


Fig. 14: PDF of RUL estimates for different operation profiles

.

The maximum likelihood RUL after 6 months is 262786 hours and 262781 hours for load B and C respectively, and 95% prediction intervals within [262787, 262784.9] and [262781.9, 262779.8] hours, respectively, which suggests that the dependency on the solar energy generation leads to an intermittent and mild insulation degradation and there may be room for increasing the ageing limit of transformers.

C. Analysis of Sampling Rates

In order to evaluate the influence of signals with different sampling rates, the H-QGB model has been tested with changing sampling rates of the monitored signals. Table IV shows the probabilistic forecasting error results.

TABLE IV: TOT forecasting error for various sampling rates.

Sampling Rate	CRPS	MAE
1 min.	**0.19±0.13**	0.41±0.3
5 min.	0.2±0.13	0.4±0.3
10 min.	0.2±0.14	0.41±0.32
30 min.	0.21±0.17	0.43±0.31
60 min.	0.25±0.19	0.48±0.38

It can be observed that changing the sampling rate of the monitored signals affects the precision of the probabilistic model. That is, the higher the sampling rate, the lower the error score. This is in line with the thermal difference equations and parameters, which lose precision at lower sampling rates, and better capture the changing dynamics with higher sampling rate signals.

VI. Discussion

A comprehensive transformer prognostics framework is shown in this paper. However, before drawing definitive conclusions about the influence of RES on transformer RUL, further work is necessary testing the approach for other periods, and possibly monitoring additional degradation factors.

The proposed analysis suggests worst-case operation and degradation context which drive transformer O&M strategies. Transformer degradation has been evaluated for the warmest period of the year and load profiles which can emerge from extreme meteorological conditions (cf. Fig. 12) have been also analysed to evaluate the effect on transformer RUL estimation. However, in order to generalize the influence of RES on the transformer degradation, a larger dataset covering different seasons and a few years of operation may be needed.

With a larger dataset, it may be possible to infer seasonality patterns so as to build error-correction approaches adapted to each period including summer, winter and shoulder seasons of spring and fall. The same model design procedure would be applied with appropriate feature-selection adjusted to each season, e.g. the influence of the solar irradiance and PV conversion efficiency may change between different seasons, as well as the empirical cross-correlation lags between the analysed signals.

Additionally, the degradation rate and RUL estimations have been obtained from the influence of the HST on the insulation ageing. However, there are other factors, such as moisture, which can also accelerate the insulation ageing rate [8], [15]. Accordingly, obtained results must be interpreted with care, recognising that synergistic degradation effects may also play a part.

Finally, probabilistic forecasting models have been used due to their ability to convey uncertainty information in the predicted estimate. There are black-box deterministic estimates [3], mostly based on deep-learning modelling strategies, e.g. [23]. However, although they may accurately estimate the predicted value, the uncertainty value is lost. Accordingly, this may limit the applicability and potential of the proposed approach, and precisely, this is the main novelty of the proposed framework. Without probabilistic estimates, it is possible to rely on already existing prognostics frameworks [17].

The dropout mechanism of deep-learning models, which was originally introduced to avoid overfitting, has been extended to capture uncertainty through Monte Carlo simulations and variational inference [41]. Recent probabilistic extensions of deep-learning models open the way to enhance the robustness of deep-learning predictions including uncertainties [42], [43]. The introduction of additional parameters may result in deep-learning architectures that require more time to converge [41]. The integration and comparison of the proposed hybrid approach with probabilistic deep-learning models is part of the future work of the authors.

VII. Conclusions

The operation and integration of RES with traditional power equipment creates novel operation and degradation scenarios, including fast-switching events and strong weather-dependency, which impact on the reliability of power assets. Focusing on transformers, key assets for RES integration, their reliability should be modelled with appropriate health-monitoring methodologies suited for their operation with RES.

In this context, this paper has proposed a hybrid prognostics framework for the health assessment of transformers operated with RES. The approach combines physics-based modelling with probabilistic forecasting in a error-correcting configuration, and this is embedded in a probabilistic lifetime framework to predict the transformer RUL with prediction intervals. Prediction intervals vary for each prediction according to the propagated uncertainty and they inform about the confidence of the model in the predictions.

The approach is applicable to oil-immersed distribution transformers that are used to connect RESs and the grid. However, it may be necessary to adapt the probabilistic forecasting stage, with an application-specific feature selection. Results obtained from the application to a floating photovoltaic power plant, demonstrate that the proposed approach adapts physics-based methodologies to the RES context through a data-driven probabilistic forecasting strategy.

Prognostics results have shown that the transformer insulation degradation is mild due to the intermittency of the generated energy. This suggests that there may be room to extend the operation condition beyond the initially assumed lifetime and flexibility to increase power production for future integration. However, obtained degradation results should be complemented with other degradation-influencing variables, such as moisture, so as to obtain a more complete picture of the insulation health.

Analysis of the sampling rate of monitored signals indicate that a decrease in the sampling rate affects the prediction performance, due to the loss-of-information in-between samples, and therefore, it suggests that it is important to accurately capture transformer ageing variables and model degradation dynamics.

The proposed approach assists operation management decisions by better understanding likely transformer degradation in RESs from the insulation degradation perspective.

References

[1] S. R. Sinsel, R. L. Riemke, and V. H. Hoffmann, "Challenges and solution technologies for the integration of variable renewable energy sources—a review," *renewable energy*, vol. 145, pp. 2271–2285, 2020.

[2] H. Wang and F. Blaabjerg, "Power electronics reliability: State of the art and outlook," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 6, pp. 6476–6493, 2021.

[3] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani, "Failure prognosis and applications—a survey of recent literature," *IEEE Transactions on Reliability*, vol. 70, no. 2, pp. 728–748, 2021.

[4] M. J. Heathcote, *J & P Transformer Book*, Thirteenth ed. Newnes, 2007.

[5] "Ieee guide for evaluation and reconditioning of liquid immersed power transformers," *IEEE Std C57.140-2017 (Revision of IEEE Std C57.140-2006)*, pp. 1–88, 2017.

[6] IEC, "Loading guide for oil-immersed power transformers," *IEC 60076-7*, 2018.

[7] X. Zhang, Z. Wang, and Q. Liu, "Interpretation of hot spot factor for transformers in od cooling modes," *IEEE Transactions on Power Delivery*, vol. 33, no. 3, pp. 1071–1080, 2018.

[8] R. D. Medina, A. A. Romero, E. E. Mombello, and G. Rattá, "Assessing degradation of power transformer solid insulation considering thermal stress and moisture variation," *Electric Power Systems Research*, vol. 151, pp. 1–11, 2017.

[9] M. Ariannik, A. A. Razi-Kazemi, and M. Lehtonen, "An approach on lifetime estimation of distribution transformers based on degree of polymerization," *Reliability Eng. Sys. Safety*, vol. 198, p. 106881, 2020.

[10] W. Huang, C. Shao, M. Dong, B. Hu, W. Zhang, Y. Sun, K. Xie, and W. Li, "Modeling the aging-dependent reliability of transformers considering the individualized aging threshold and lifetime," *IEEE Transactions on Power Delivery*, pp. 1–1, 2022.

[11] S. F. Abdelsamad, W. G. Morsi, and T. S. Sidhu, "Probabilistic impact of transportation electrification on the loss-of-life of distribution transformers in the presence of rooftop solar photovoltaic," *IEEE Trans. Sustainable Energy*, vol. 6, no. 4, pp. 1565–1573, 2015.

[12] A. A. Taheri, A. Abdali, and A. Rabiee, "A novel model for thermal behavior prediction of oil-immersed distribution transformers with consideration of solar radiation," *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1634–1646, 2019.

[13] S. A. El-Bataway and W. G. Morsi, "Distribution transformer's loss of life considering residential prosumers owning solar shingles, high-power fast chargers and second-generation battery energy storage," *IEEE Trans. Industrial Informatics*, vol. 15, no. 3, pp. 1287–1297, 2019.

[14] Y. Li, Y. Wang, and Q. Chen, "Optimal dispatch with transformer dynamic thermal rating in adns incorporating high pv penetration," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 1989–1999, 2021.

[15] X. Zhong, C. Ekanayake, H. Ma, and T. K. Saha, "Ageing analysis of solar farm inverter transformers," *IEEE Transactions on Power Delivery*, vol. 36, no. 6, pp. 3815–3824, 2021.

[16] A. Molina, K. Morozovska, T. Laneryd, and P. Hilber, "Optimal sizing of the wind farm and wind farm transformer using milp and dynamic transformer rating," *International Journal of Electrical Power & Energy Systems*, vol. 136, p. 107645, 2022.

[17] J. Aizpurua, S. McArthur, B. Stewart, B. Lambert, J. Cross, and V. Catterson, "Adaptive power transformer lifetime predictions through machine learning & uncertainty modeling in nuclear power plants," *IEEE Trans. Industrial Electronics*, vol. 66, no. 6, pp. 4726–4737, 2019.

[18] J. I. Aizpurua, U. Garro, E. Muxika, M. Mendicute, I. P. Gilbert, B. G. Stewart, S. D. J. McArthur, and B. Lambert, "Probabilistic power transformer condition monitoring in smart grids," in *2019 6th Int. Advanced Research Workshop on Transformers*, 2019, pp. 42–47.

[19] C. d. M. Affonso and M. Kezunovic, "Technical and economic impact of pv-bess charging station on transformer life: A case study," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4683–4692, 2019.

[20] A. Bracale, G. Carpinelli, M. Pagano, and P. De Falco, "A probabilistic approach for forecasting the allowable current of oil-immersed transformers," *IEEE Transactions on Power Delivery*, vol. 33, no. 4, pp. 1825–1834, 2018.

[21] A. Bracale, P. Caramia, G. Carpinelli, and P. De Falco, "Smartrafo: A probabilistic predictive tool for dynamic transformer rating," *IEEE Transactions on Power Delivery*, vol. 36, no. 3, pp. 1619–1630, 2021.

[22] C.-M. Lai and J. Teh, "Comprehensive review of the dynamic thermal rating system for sustainable electrical power systems," *Energy Reports*, vol. 8, pp. 3263–3288, 2022.

[23] F. Wang, J. Du, Y. Zhao, T. Tang, and J. Shi, "A deep learning based data fusion method for degradation modeling and prognostics," *IEEE Transactions on Reliability*, vol. 70, no. 2, pp. 775–789, 2021.

[24] L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, and P. L. McMahon, "Deep physical neural networks trained with backpropagation," *Nature*, vol. 601, no. 7894, pp. 549–555, 2022.

[25] S. Cuomo, V. S. D. Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, "Scientific machine learning through physics-informed neural networks: Where we are and what's next," *ArXiv*, vol. abs/2201.05624, 2022.

[26] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, vol. 1, no. 1, pp. 125–151, 2014.

[27] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896 – 913, 2016.

[28] F. N. Fritsch and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 238–246, 1980.

[29] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232, 2001.

[30] H. Verbois, A. Rusydi, and A. Thiery, "Probabilistic forecasting of day-ahead solar irradiance using quantile gradient boosting," *Solar Energy*, vol. 173, pp. 313 – 327, 2018.

[31] Z. Xiao and R. Koenker, "Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models," *J. of American Statistical Association*, vol. 104, no. 488, pp. 1696–1712, 2009.

[32] R. Koenker and K. F. Hallock, "Quantile regression," *Journal of Economic Perspectives*, vol. 15, no. 4, pp. 143–156, December 2001.

[33] F. Chiacchio, D. D'Urso, F. Famoso, S. Brusca, J. I. Aizpurua, and V. M. Catterson, "On the use of dynamic reliability for an accurate modelling of renewable power plants," *Energy*, vol. 151, pp. 605–621, 2018.

[34] Hersbach, Hans and Bell, et al., "The era5 global reanalysis," *Quarterly J. Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020.

[35] H. Hersbach, "Decomposition of the continuous ranked probability score for ensemble prediction systems," *Weather and Forecasting*, vol. 15, no. 5, pp. 559–570, 2000.

[36] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb 2002.

[37] R. Karlsson, T. Schon, and F. Gustafsson, "Complexity analysis of the marginalized particle filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4408–4411, 2005.

[38] S. Oliveira-Pinto and J. Stokkermans, "Assessment of the potential of different floating solar technologies – overview and analysis of different case studies," *Energy Conv. and Management*, vol. 211, p. 112747, 2020.

[39] IEC, "Temperature rise for oil-immersed power transformers," *IEC 60076-2*, 2011.

[40] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, vol. 6, no. 2, pp. 418–445, 1996.

[41] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*. PMLR, 2016, pp. 1050–1059.

[42] C. Chen, N. Lu, B. Jiang, Y. Xing, and Z. H. Zhu, "Prediction interval estimation of aeroengine remaining useful life based on bidirectional long short-term memory network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021.

[43] H. Pei, X.-S. Si, C. Hu, T. Li, C. He, and Z. Pang, "Bayesian deep-learning-based prognostic model for equipment without label data related to lifetime," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–14, 2022.