A Physics-Inspired and Data-Driven Approach for Temperature-Based Condition Monitoring

Giacomo Garegnani[1], Kai Hencken[1], and Frank Kassubek[1]

[1] *ABB Switzerland Ltd., Corporate Research*
giacomo.garegnani@ch.abb.com
kai.hencken@ch.abb.com
frank.kassubek@ch.abb.com

ABSTRACT

System overheating is a common problem in electric equipment, as degradation of contacts lead to an increase in Ohmic resistance and increased thermal losses. Temperature measurements are widely employed to monitor a device's health status, to estimate its remaining useful life, and to inform maintenance strategies. An issue that is special to electrical distribution networks is the varying heating power, which is in turn due to changes in the current. This results in varying temperatures, which in addition can often be delayed compared to the currents. Simple threshold-based diagnostics approaches are therefore not reliable for detecting faults in contacts. It is common to analyze the thermal behavior of electric devices using thermal networks, for both design and diagnostic purposes. Unfortunately, identifying the parameters of thermal networks from measured temperature data is a challenging problem, mainly due to identifiability issues and to numerical instabilities in parameter estimation. We propose an alternative data-driven strategy to compute the state-of-health of electrical devices which does not resort to thermal networks. Our approach consists in informing physics-based reduced models of the thermal response with sensor data. We show that our method is linked to the thermal network approach but avoids the full identification of the system, leading to better stability as well as less computational effort in the determination of its parameters. Rigorous testing with synthetic and experimental data confirms the efficacy of our methodology.

1. INTRODUCTION

The effective monitoring of the operational health of electric devices is of utmost importance to guarantee the secure and steady functioning of industrial facilities (Hoffmann et al., 2020). Among the vulnerabilities encountered by these devices, the issue of overheating due to Joule's effect stands out prominently. A significant portion of the heat generated within these devices comes from electric contacts. The deterioration of contacts results in an increase in their electrical contact resistance (ECR), which, in turn, triggers pronounced overheating. Such overheating not only disrupts operational stability but also exposes the devices to the imminent risk of irreparable harm.

With the rise in connectivity of industrial devices – the industrial internet of things – the potential of monitoring algorithms for predictive maintenance has grown considerably. In the present context, temperature data can be leveraged to prevent excessive overheating and monitor the health state of devices. While simple algorithms monitor the temperature and raise alerts based on critical levels, the dynamic nature of the thermal response to time-dependent current loads yields more insight into the root cause of the problem.

The method we propose in this report is computationally light and memory-efficient (in contrast to numerical solvers of partial differential equations), and is robust when confronted with real data (unlike thermal networks). Despite its simplicity, we believe this method can be effectively used for making thermal predictions and detect anomalous behavior for a wide range of electric devices.

Thermal networks can be cumbersome to set up and train (O. M. Brastein et al., 2019; O. Brastein et al., 2020; Boodi et al., 2022), but they are nevertheless a flexible tool to model the temperature response of the device. Conversely, the method which we propose here does not allow to predict temperatures away from the sensor and requires some dedicated training of the response of each device using a specific current profile. Memory-wise, the method also requires some limited storage of past current values. Finding a mathematical equivalence of thermal networks and the proposed method may therefore lead to the development of a method that combines the two approaches and retains the advantages of both.

The remainder of this article is as follows. In Section 2 we introduce the thermal kernel method, focusing on how to cope with noise in the data, and on how to infer variations in the ECR values. The equivalence of the kernel method and the thermal network approach is then discussed in Section 3. In Section 4 we test the method against synthetic data, checking the accuracy of the method as well as its robustness against model misspecification and noise in temperature data. Finally, in Section 5 we draw our conclusions and propose an outlook for future work.

2. THE THERMAL KERNEL METHOD

We consider an electric device which is equipped for simplicity with a single temperature sensor (see Section 2.4 below for the case of multiple sensors). We model the over-temperature T at the sensor location for a new and healthy device as

$$T(t) = \int_0^t k_0(t-s)\, I^2(s)\, \mathrm{d}s, \tag{1}$$

where k_0 is an unknown kernel function, and where I denotes the electric current, which we assume to satisfy $I(t) = 0$ for all $t \leq 0$. Note that k_0 can be seen as a Green function of the thermal problem, and I^2 is proportional to the thermal input due to Joule's law of heating. The kernel k_0 captures all linear thermal influences on the temperature measured by the sensor, and in particular the heat generated both by bulk conductors (e.g., busbars in electric devices) and the heat generated at imperfect contacts, which are both proportional to I^2. The existence of such a kernel function is guaranteed if we assume that the system is linear both with respect to the heat flow and the dependency on I^2.

Let N be the number of electrical contacts in the device, and denote by ΔR_i the variation (typically an increase) of the ECR in the i-th contact for $i = 1, \ldots, N$. Then, we assume that there exist kernels k_i for all $i = 0, 1, \ldots, N$ such that the temperature at the sensor location reads

$$\begin{aligned} T(t) = &\int_0^t k_0(t-s) I^2(s)\, \mathrm{d}s \\ &+ \sum_{i=1}^{N} \Delta R_i \int_0^t k_i(t-s) I^2(s)\, \mathrm{d}s. \end{aligned} \tag{2}$$

With the formula in Eq. (2), we assume that the thermal response at the sensor location of the device after contact degradation is encoded by the kernel

$$k_\Delta(t) = k_0(t) + \sum_{i=1}^{N} \Delta R_i k_i(t), \tag{3}$$

where the kernels $k_i(t)$ model the thermal response due to a change of resistance at contact i. The kernel functions k_i for $i = 0, 1, \ldots, N$ are unknown and depend on the device geometry, on how heat is exchanged with the surrounding environment, and on the thermal interconnections of the device components.

2.1. Determination of the kernel functions

In order to determine the kernel functions we use the response of the system to a step excitation, i.e., by imposing a constant current $I(t) = I_0$ for all $t \geq 0$. First, we fix $\Delta R_i = 0$ for all i and derive both sides of Eq. (1) with respect to t to obtain

$$k_0(t) = \frac{\dot{T}(t)}{I_0^2}. \tag{4}$$

Note that the temperature derivative $\dot{T}$ may not be available from sensor data, but can easily be reconstructed in practice by means of a finite difference formula from the measured temperature T. Note that numerical differentiation may amplify noise on the signal. We tackle this issue in Section 2.2 below. Given k_0 we can then determine the remaining N kernels by increasing the ECRs by a known quantity one by one. Indeed, if it holds $I(t) = I_0$ and $\Delta R_j = 0$ for all $j \neq i$ for a fixed index i, we have from Eq. (2)

$$\dot{T}(t) = \left(k_0(t) + \Delta R_i k_i(t)\right) I_0^2.$$

If the ECR increase ΔR_i is known and we measure the corresponding temperature T, then the kernel k_i is given by

$$k_i(t) = \frac{1}{I_0^2 \Delta R_i} \left(\dot{T}(t) - I_0^2 k_0(t)\right). \tag{5}$$

It might be unpractical or impossible in some scenarios to increase the ECR by a known quantity. Determining kernel functions may then involve data generation through a high fidelity simulation.

2.2. Noisy or short temperature data: Exponential fit

Let us assume that the temperature T is observed for a finite time interval $0 \leq t \leq t_{\mathrm{end}}$ and that observations are subject to measurement noise. In this case, the kernel k_0 given by Eq. (4) (and similarly the kernels k_i, $i = 1, \ldots, N$) should be post-processed to obtain a smooth kernel that can also be evaluated for times $t > t_{\mathrm{end}}$. For this purpose, we can introduce the natural assumption that k_0 is given by an infinite sum of negative exponential functions, as in

$$k_0(t) = \sum_{j=1}^{\infty} a_j \exp(-\lambda_j t),$$

where $a_j \in \mathbb{R}$, $\lambda_j \in \mathbb{R}^+$ for all $j = 1, 2, \ldots$. We then truncate the sum to an integer number N_{exp} of exponential functions and write

$$\widetilde{k}_0(t) = \sum_{j=1}^{N_{\mathrm{exp}}} a_j \exp(-\lambda_j t).$$

A suitable value $N_{\exp}$ can be chosen with a model selection algorithm. Finally, we determine a_j and λ_j by maximizing the likelihood of the noisy kernel k_0 given by Eq. (4) applied with the data sequence T. This ansatz can be motivated by the equivalence between thermal networks and kernels (or more in general by any finite-dimensional approximation of the full heat problem). More details can be found in Section 3.

The approach of fitting exponential functions to k_0 could be problematic in case temperature data are corrupted by noise. Indeed, noise is amplified when computing the time derivative $\dot{T}$ of the temperature. In this case, it is more robust to fit directly the temperature data T, which under the assumption above is approximated by

$$\widetilde{T}(t) = I_0^2 \sum_{j=1}^{N_{\exp}} \frac{a_j}{\lambda_j} \left(1 - \exp(-\lambda_j t)\right).$$

We can therefore fit the curve above directly to the temperature data and determine the values of a_j and λ_j which fully define the kernel function k_0.

We can repeat the same reasoning for the kernels $\{k_i\}_{i=1}^N$ modeling the thermal response at the sensor location due to (additional) heat generated at the contacts. We make the guess that for all $i = 1, \dots, N$ it holds

$$k_i(t) = \sum_{j=1}^{N_{\exp}} a_{ij} \exp(-\lambda_{ij} t).$$

Manipulating Eq. (5) with similar calculations as above we obtain

$$T(t) - I_0^2 \int_0^t k_0(t-s)\,\mathrm{d}s = I_0^2 \Delta R_i \sum_{j=1}^{N_{\exp}} \frac{a_{ij}}{\lambda_{ij}} \left(1 - \exp(-\lambda_{ij} t)\right).$$

The left-hand side of this equation is known. Fitting the coefficients a_{ij} and λ_{ij} to data then defines the kernel k_i. Note that this approach assumes that the coefficients λ are independent of each other for k_0 and each k_i. Since the thermal time scales should be the same for the nominal value of the resistance and increased resistances by linearity, the values λ_{ij} should be shared by the fit to k_0. A more robust approach, which we do not investigate here, would therefore consist in fitting the kernel functions simultaneously.

2.3. Inference of the resistance variations

In this section, we describe how knowledge of the kernel functions can be combined with temperature data to infer online a variation of the ECR of the N contacts, and consequently deduce their health status. Assume that all the kernels k_i have been determined and denote for $i = 0, \dots, N$ by K_i the integrated quantity

$$K_i(t) = \int_0^t k_i(t-s) I^2(s)\,\mathrm{d}s,$$

where I is the measured current. Then, we can rewrite Eq. (2) as

$$T(t) = K_0(t) + \sum_{i=1}^{N} \Delta R_i K_i(t). \tag{6}$$

Assume that the current and the temperature at the sensor have been measured on a set of times $\mathbf{t} = (t_0, t_1, \dots, t_M)$, where $t_j = t_s \cdot j$ and t_s is the sampling time. We can then assemble M-dimensional vectors

$$\mathbf{T} = T(\mathbf{t}), \qquad \mathbf{K}_i = K_i(\mathbf{t}),$$

where $T(\mathbf{t}) = (T(t_0), T(t_1), \dots, T(t_M))^\top$. Using the vectorial notation, the discrete version of Eq. (6) is

$$\mathbf{T} = \mathbf{K}_0 + \sum_{i=1}^{N} \Delta R_i \mathbf{K}_i.$$

An estimator $\widehat{\Delta R} \in \mathbb{R}^N$ of the vector of ECRs can be defined as the least square estimator

$$\widehat{\Delta R} = \arg\min_{\Delta R} \left\| \mathbf{K}\Delta R - (\mathbf{T} - \mathbf{K}_0) \right\|, \tag{7}$$

where $\mathbf{K}$ is the $M \times N$ matrix with columns $\mathbf{K}_i$ for $i = 1, \dots, N$. The minimization problem is overdetermined whenever $M \geq N$, i.e., the number of time instants for the measurements exceeds the number of contacts in the system, which is most likely verified. Hence, the estimator in Eq. (7) should be determined as the solution of the $N \times N$ linear system

$$\mathbf{K}^\top \mathbf{K} \widehat{\Delta R} = \mathbf{K}^\top (\mathbf{T} - \mathbf{K}_0).$$

Note that in real applications we expect the values of ΔR_i to increase rather than decrease due to contact degradation. A physically meaningful solution could therefore be enforced by using the constrained minimizer

$$\widehat{\Delta R} = \arg\min_{\Delta R \geq 0} \left\| \mathbf{K}\Delta R - (\mathbf{T} - \mathbf{K}_0) \right\|,$$

where the symbol $\geq$ is meant component-wise.

2.4. Multiple temperature sensors

We now consider a device which is equipped with multiple temperature sensors, and explain how more information can be leveraged to obtain a possibly more precise estimation of variations in the ECRs.

Assume that we have J temperature sensors. The temperature

of each sensor $j = 1, \dots, J$ can be expressed as

$$T^j(t) = \int_0^t k_0^j(t-s)I^2(s)\,\mathrm{d}s + \sum_{i=1}^{N} \Delta R_i \int_0^t k_i^j(t-s)I^2(s)\,\mathrm{d}s.$$

Note that the resistance increase ΔR_i is common for all sensors, as contacts are the same. Conversely, the temperature response is different across sensors, hence typically $k_i^{j_1} \neq k_i^{j_2}$ for $j_1 \neq j_2$. Kernels k_i^j can be determined as outlined in Section 2.1 for each $i = 0, \dots, N$ and $j = 1, \dots, J$. Similarly to Section 2.3, we then write $\mathbf{T}^j = T^j(\mathbf{t})$ and $\mathbf{K}_i^j = K_i^j(\mathbf{t})$ where

$$K_i^j(t) = \int_0^t k_i^j(t-s)I^2(s)\,\mathrm{d}s.$$

Calling $\mathbf{K}^j$ the $N \times M$ matrix whose columns are the vectors $\mathbf{K}_i^j$ for $i = 1, \dots, N$, we have J linear equations for ΔR

$$\mathbf{K}^j \Delta R = \mathbf{T}^j - \mathbf{K}_0^j, \qquad j = 1, \dots, J.$$

In order to compute the least square solution $\widehat{\Delta R}$ we assemble a $NJ \times M$ matrix $\mathbf{K}$ and NJ vectors $\mathbf{T}$ and $\mathbf{K}_0$ by stacking vertically the J equations as

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}^1 \\ \mathbf{K}^2 \\ \vdots \\ \mathbf{K}^J \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} \mathbf{T}^1 \\ \mathbf{T}^2 \\ \vdots \\ \mathbf{T}^J \end{pmatrix}, \quad \mathbf{K}_0 = \begin{pmatrix} \mathbf{K}_0^1 \\ \mathbf{K}_0^2 \\ \vdots \\ \mathbf{K}_0^J \end{pmatrix}.$$

The least square estimate $\widehat{\Delta R}$ is then the solution of the $N \times N$ linear system

$$\mathbf{K}^\top \mathbf{K} \widehat{\Delta R} = \mathbf{K}^\top (\mathbf{T} - \mathbf{K}_0),$$

and similarly to the single-sensor case a non-negative constraint can be imposed to the least-square solution. We note that in this case we expect an improvement by enforcing the time scale parameters to be the same across resistances when performing an exponential fit as in Section 2.2.

3. Equivalence with thermal networks

Thermal networks have been used to model the temperature of electric devices, and to infer health status given temperature measurements (Stosur et al., 2016). In this section, we describe how our approach simplifies thermal networks, whose parameters are notoriously difficult to estimate from data (O. M. Brastein et al., 2019; O. Brastein et al., 2020; Boodi et al., 2022). For a general discussion on identifiability of linear models, we refer the reader to (Raue et al., 2014).

We call thermal network a model which splits the device into an integer number N_c of compartments, whose temperature is assumed to be sufficiently homogeneous to be described by a single over-temperature T_i, for $i = 1, \dots, N_c$. We assume that the i-th compartment has a heat capacity C_i for $i = 1, \dots, N_c$. The compartments are thermally interconnected so that the heat flowing between the compartments indexed by i and j is proportional to their temperature difference with a constant h_{ij}. If two compartments are not directly connected thermally, we trivially set $h_{ij} = 0$. Moreover, we assume that the heat flowing towards the environment is proportional to the over-temperature T_i with a constant α_i. Finally, we assume that all elements in the network represent parts of the device which are subject to an electrical current $I = I(t)$, so that the thermal input to the i-th element is given by $u_i(t) = R_i I^2(t)$ by Ohmic heating. Under these assumptions, the over-temperature T_i of the i-th compartment of the network, for $i = 1, \dots, N_c$, satisfies the ordinary differential equation (ODE)

$$C_i \dot{T}_i(t) = \sum_{j=1, j\neq i}^{N_c} h_{ij}(T_j - T_i) - \alpha_i T_i + u_i(t). \tag{8}$$

In this section, we show how the temperature evolution of each compartment in a thermal network satisfies Eq. (1), i.e., there exist kernels k_0^i such that

$$T_i(t) = \int_0^t k_0^i(t-s)I^2(s)\,\mathrm{d}s, \tag{9}$$

for each $i = 1, \dots, N_c$, and that the kernel function can be written as a sum of exponential functions as in Section 2.2. Hence, a system whose thermal response can be described accurately by a thermal network can also be described by thermal kernels, with the advantage that in the kernel approach less parameters need to be determined from temperature measurements.

To start the derivation, we notice that the ODE system Eq. (8) can be written in matrix form as

$$\mathbf{C}\dot{\mathbf{T}}(t) = \mathbf{H}\mathbf{T}(t) + \mathbf{R}I^2(t), \tag{10}$$

where $\mathbf{T}$ is a vector with the temperatures of all compartments, where $\mathbf{R}$ is a N_c-dimensional vector containing the values of the resistances, and where $\mathbf{C}$ and $\mathbf{H}$ are appropriate matrices containing the values of the coefficients h, α and C. Let us rewrite Eq. (10) as the generic linear system

$$\dot{\mathbf{T}}(t) = -\mathbf{A}\mathbf{T}(t) + \mathbf{F}(t), \tag{11}$$

where $\mathbf{A} = -\mathbf{C}^{-1}\mathbf{H}$ and $\mathbf{F}(t) = \mathbf{C}^{-1}\mathbf{R}I^2(t)$. Let $\mathbf{T}(0) = \mathbf{T}_0$ be a known initial condition. It is simple to verify by differentiation that the solution of Eq. (11) is given by

$$\mathbf{T}(t) = \mathbf{e}^{-\mathbf{A}t}\mathbf{T}_0 + \int_0^t \mathbf{e}^{-\mathbf{A}(t-s)}\mathbf{F}(s)\,\mathrm{d}s, \tag{12}$$

where we denote by $\mathbf{e}^{-\mathbf{A}t}$ the matrix exponential to distin-

guish it from the scalar exponential (e.g., e^t).

The matrix $\mathbf{A}$ is not symmetric but it is diagonalizable with real eigenpairs.[1] Recall that for any diagonalizable matrix $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$, where $\mathbf{V}$ is the matrix with the eigenvectors $\{\mathbf{v}_j\}_{j=1}^{N_c}$ of $\mathbf{A}$ as columns, and $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_{N_c})$ is the matrix of the eigenvalues, it holds

$$\mathbf{e}^{-\mathbf{A}t} = \mathbf{V}\mathbf{e}^{-\mathbf{\Lambda}t}\mathbf{V}^{-1} = \mathbf{V}\mathrm{diag}\left(e^{-\lambda_1 t}, \dots, e^{-\lambda_{N_c} t}\right)\mathbf{V}^{-1}.$$

This implies that if (λ, v) is an eigenpair of $\mathbf{A}$, then $(e^{-\lambda t}, v)$ is an eigenpair of $\mathbf{e}^{-\mathbf{A}t}$. Let $\mathbf{w}$ be an arbitrary vector in $\mathbb{R}^{N_c}$ and let $\{c_j = (\mathbf{V}^{-1}\mathbf{w})_j\}_{j=1}^{N_c}$ be the components[2] of $\mathbf{w}$ in the basis formed by the eigenvectors of $\mathbf{A}$, i.e., the scalars such that

$$\mathbf{w} = \sum_{j=1}^{N_c} c_j \mathbf{v}_j.$$

Hence, applying $\mathbf{e}^{-\mathbf{A}t}$ to $\mathbf{w}$ gives

$$\mathbf{e}^{-\mathbf{A}t}\mathbf{w} = \sum_{j=1}^{N_c} c_j \mathbf{e}^{-\mathbf{A}t}\mathbf{v}_j = \sum_{j=1}^{N_c} c_j e^{-\lambda_j t}\mathbf{v}_j.$$

Assume for simplicity and without loss of generality that $\mathbf{T}_0 = 0$. Replacing the decomposition above into Eq. (12) with $\mathbf{w} = \mathbf{C}^{-1}\mathbf{R}I^2(s)$ shows that

$$\mathbf{T}(t) = \int_0^t \sum_{j=1}^{N_c} c_j \mathbf{v}_j e^{-\lambda_j (t-s)} I^2(s)\,\mathrm{d}s,$$

where $c_j = (\mathbf{V}^{-1}\mathbf{C}^{-1}\mathbf{R})_j$. Hence, the temperature of the i-th compartment satisfies

$$T_i(t) = \int_0^t \sum_{j=1}^{N_c} \alpha_{ij} e^{-\lambda_j (t-s)} I^2(s)\,\mathrm{d}s,$$

where $\alpha_{ij} = \mathbf{V}_{ij} c_j$. This shows that the temperature of the i-th compartment can be indeed written as in Eq. (9) for

$$k_0^i(t) = \sum_{j=1}^{N_c} \alpha_{ij} e^{-\lambda_j t},$$

which is a sum of exponential functions as the approximations we employ in Section 2.2. Consider now that for each $k = 1, \dots, N_c$ the resistance of the k-th compartment increases by a quantity ΔR_k. We can write the overall kernel defining the temperature of the i-th compartment as

$$k^i(t) = \sum_{j,k=1}^{N_c} \mathbf{V}_{ij}\left(\mathbf{V}^{-1}\mathbf{C}^{-1}\right)_{jk} R_k e^{-\lambda_j t} + \sum_{j,k=1}^{N_c} \mathbf{V}_{ij}\left(\mathbf{V}^{-1}\mathbf{C}^{-1}\right)_{jk} \Delta R_k e^{-\lambda_j t}.$$

We see that k^i has the form of the kernel of Eq. (3) with

$$k_k^i(t) \coloneqq \sum_{j=1}^{N_c} \mathbf{V}_{ij}\left(\mathbf{V}^{-1}\right)_{jk} e^{-\lambda_j t},$$

which is the kernel associated to an increase in the k-th resistance as seen by the i-th element of the thermal network. Note that since the resistances do not appear in the expression of the system matrix $\mathbf{A} = -\mathbf{C}^{-1}\mathbf{H}$, the time scales λ_j in the kernels k_k^i are the same as the ones of the original kernel.

3.1. Generalization: kernel structure of thermal problems

The considerations above for thermal networks and the kernel structure of their solution applies more widely. In a linear approximation, heat transfer can be described by

$$\mathbf{C}\dot{\mathbf{T}} = \mathbf{L}\mathbf{T} + u, \tag{13}$$

where $\mathbf{L}$ is an operator describing both heat conduction $\mathbf{H}$ and coupling to the ambient α, and u is the heat injected in the system. In the specific case of a thermal network, the temperatures are vectors and the operators (finite dimensional) matrices. However, this equation may also describe a temperature field with a partial differential operator describing heat conduction on a physical domain Ω. For $x \in \Omega$, the local operator $\mathbf{C} = \mathbf{C}(x)$ is the specific heat capacity and the differential operator $\mathbf{L}(x) = -\nabla k(x) \cdot \nabla - k(x)\Delta$ describes heat conduction with a space-dependent heat conductivity k defined on Ω.

Note that Eq. (13) is linear in temperature, the operator $\mathbf{L}$ is self adjoint due to the symmetric nature of heat diffusion, and the field $\mathbf{C}$ is a (local) positive scalar. Normalizing the temperature $\widetilde{\mathbf{T}} = \mathbf{C}^{1/2}\mathbf{T}$ and multiplying Eq. (13) by $\mathbf{C}^{-1/2}$, we see that the operator occurring on the right hand side of the equation for $\widetilde{\mathbf{T}}$ ($\mathbf{C}^{-1/2}\mathbf{L}\mathbf{C}^{-1/2}$) is also self-adjoint. The spectral theorem then guarantees that this operator has real eigenvalues and orthogonal eigenfunctions that span the full space. Formally, the solution can be expressed in terms of the exponential operator $\mathbf{e}^{\mathbf{L}t}$ as

$$\widetilde{\mathbf{T}} = \int_0^t \mathbf{e}^{\mathbf{L}(t-s)}\widetilde{u}(s)\,\mathrm{d}s, \tag{14}$$

[1] Since $\mathbf{A} = -\mathbf{C}^{-1}\mathbf{H}$, with $\mathbf{H}$ symmetric and $\mathbf{C}$ diagonal and positive definite, we can write

$$\mathbf{A} = \mathbf{C}^{-1/2}\widetilde{\mathbf{A}}\mathbf{C}^{1/2},$$

where $\widetilde{\mathbf{A}} = -\mathbf{C}^{-1/2}\mathbf{H}\mathbf{C}^{-1/2}$. The matrix $\widetilde{\mathbf{A}}$ is real and symmetric, and hence can be diagonalized with real eigenpairs, which in turn implies that $\mathbf{A}$ is diagonalizable with real eigenpairs.

[2] Since the matrix $\mathbf{A}$ is in general not symmetric, the vectors $\mathbf{V}$ do not form a orthonormal basis of $\mathbb{R}^{N_c}$. If $\mathbf{A}$ is symmetric, it holds $\mathbf{V}^{-1} = \mathbf{V}^\top$ and

$$c_j = (\mathbf{V}^{-1}\mathbf{w})_j = \sum_{i=1}^{N_c} (\mathbf{V}^\top)_{ji}\mathbf{w}_i = \sum_{i=1}^{N_c} \mathbf{V}_{ij}\mathbf{w}_i = \langle \mathbf{v}_j, \mathbf{w}\rangle,$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean scalar product, which gives the more recognizable decomposition on a basis of orthonormal eigenvectors.

where $\widetilde{u} = \mathbf{C}^{-1/2}u$. For practical calculations, one has to expand in the eigenvectors as shown in the explicit example above. The general solution (14) has the same structure as the thermal kernels (1), which is hence a generic form for this type of linear heat diffusion problems. Therefore the exponential form of the kernel function can be derived independent of the assumption of an underlying thermal network as an approximation taking the dominant eigenmodes of $\mathbf{L}$ into account.

4. Numerical Experiments

In this section, we present a series of numerical experiments demonstrating the usefulness, accuracy, and robustness of our approach.

4.1. Scenario 1: Simple network

The first test setup we employ in experiments is represented schematically in Fig. 1(a). We consider an electrical device, e.g., a power protection device such as a breaker or a switch, which protects an electrical installation. The device connects the installation to a power source (e.g., the grid) with two electric contacts between busbars, one per side of the device. We assume that the device is equipped with a temperature sensor. We consider the problem of monitoring the ECR of the two contacts using the temperature sensor of the device.

In order to simulate this scenario, we use a three-compartments thermal network as shown in Fig. 1(b). In the network, the center element represents the device, and the lateral elements the two contacts. We suppose that the three compartments are exposed to the same ambient temperature T_{amb}, which we assume without loss of generality to be equal to zero.

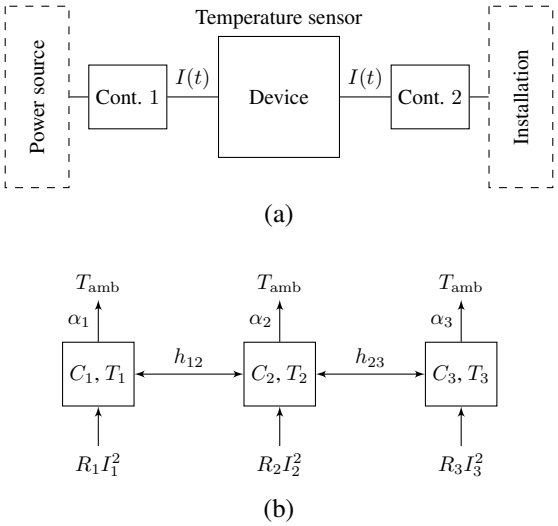


Figure 1. Setup for numerical experiments. (a) Schematic representation of a power protection device connecting an installation to a power source with two electric contacts. (b) Thermal network used to simulate the scenario.

The values of the coefficients C_i, α_i, R_i for $i = 1, \ldots, 3$, as well as of the h_{ij} for $(i, j) \in \{(1, 2), (2, 3)\}$, given in Table 1, are fixed to values which are realistic for a typical electric device. We determine the base kernel k_0 associated with the temperature sensor placed on the device fixing $I = 1\,\mathrm{kA}$ and simulating the network temperatures for $0 \leq t \leq 5\,\mathrm{h}$. Simulated data are obtained with an implicit numerical discretization of Eq. (10) on a time grid with time step equal to $1\,\mathrm{min}$. We then extract the device temperature T_2 and compute k_0 using Eq. (4), where $\dot{T}_2$ is computed by finite differences. We determine the kernels k_1 and k_3 associated to an increase of R_1 and R_3 following the procedure outlined in Section 2.1 with $\Delta R_i = R_i$, i.e., we double the ECR value to determine the kernel associated to a fault in the i-th contact.

Table 1. Coefficients of the thermal network in Fig. 1.

	$\alpha\,[\mathrm{W\,K^{-1}}]$	$R\,[\mu\Omega]$	$C\,[\mathrm{J\,K^{-1}}]$	$h\,[\mathrm{W\,K^{-1}}]$
1	1.0	100	3500	–
2	2.0	50	3500	–
3	3.0	100	3500	–
12	–	–	–	0.75
23	–	–	–	0.55

We measure the error on the i-th resistance as

$$\mathrm{err}_i = \frac{\left|\widehat{\Delta R}_i - \Delta R_i\right|}{R_i + \Delta R_i}, \tag{15}$$

where $\widehat{\Delta R}_i$ is the inferred increase in resistance and R_i is the nominal value of the i-th resistance (i.e., before increase). Note that the numerator in the right-hand side of Eq. (15) is equal to $|R_i + \widehat{\Delta R}_i - (R_i + \Delta R_i)|$, i.e., the absolute difference between the increased resistance and its inferred value. Hence, the error metric above is a relative error between the inferred and the true values of the increased *resistance*, rather than the resistance increase.

We generate 200 values of resistance increases $(\Delta R_1, \Delta R_3)$ randomly as $\Delta R_i \sim \mathcal{U}(0, R_i)$, independently for $i = 1, 3$. This means that the ECR degrades in all experiments, with values up to twice the original. For each pair of increases in the resistances, we generate 12 hours of temperature T_2 with sampling time $1\,\mathrm{min}$, always with the same current I defined by

$$I(t) = \begin{cases} 1\,\mathrm{kA}, & t \leq 1\,\mathrm{h}, \\ 0\,\mathrm{kA}, & 2\,\mathrm{h} < t \leq 5\,\mathrm{h}, \\ 0.7\,\mathrm{kA}, & 5\,\mathrm{h} < t \leq 9\,\mathrm{h}, \\ 0.3\,\mathrm{kA}, & t > 9\,\mathrm{h}. \end{cases} \tag{16}$$

The error in the estimation procedure, computed using Eq. (15), is summarized with boxplots in Fig. 2(a). We see that both resistances are estimated very accurately over the whole dataset of 200 experiments. Specifically, the error on R_1 never exceeds 0.1%, and the error on R_3 never exceeds 1%.

We repeat the same experiment but increase either R_1 or R_3 while keeping the other resistance to its nominal value. We repeat the inference 200 times per resistance with random increments as above. This experiment is relevant for applications, as the ECR of one contact only could undergo a rapid degradation, while all others could stay constant. Results, given in Figs. 2(b) and 2(c), demonstrates that also in this case the inference procedure is very accurate in determining the increased resistance values.

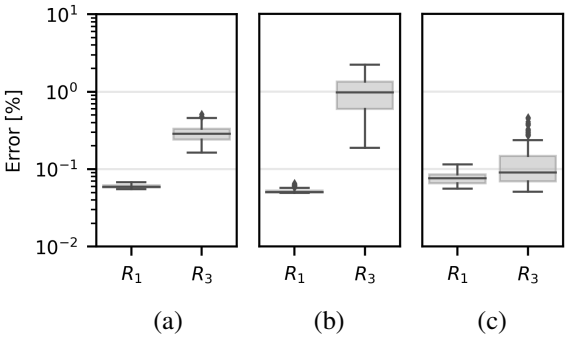


Figure 2. Percentage relative error in inference of two increased resistances given one temperature sensor. (a) Simultaneous increase of R_1 and R_3. (b) Only R_1 is increased. (c) Only R_3 is increased. Model configurations given in Fig. 1, test setup given in Section 4.1. Box-plot whiskers indicate 1.5 times the interquartile range, dots indicate outliers.

4.2. Scenario 2: The impact of adding a sensor

We consider a more complex configuration consisting of a thermal network with 6 compartments connected on a line, i.e., such that $h_{ij} = 0$ if $j \notin \{i-1, i+1\}$ for $i, j = 1, \dots, 6$. We assume that all parameters appearing in Eq. (8) are known, including nominal resistance values. Nominal parameter values are of the same magnitude as those of Table 1. Similarly to Section 4.1, we then increase randomly the resistances up to double their value and infer the increase increase with the procedure described in Section 2.3. The current used to excite the network with increased resistances is given in Eq. (16). We compare results obtained observing one temperature of the network only, T_2, and with two temperatures, T_2 and T_6. Note that when we observe one temperature we have one kernel k_0 and 6 additional kernels for the increase of R_i, $i = 1, \dots, 6$. When we observe two temperatures, we have one base kernel per sensor, and 6 additional kernels corresponding to an increase in resistance per sensor, for a total of 14 kernel functions. We recall that the method to infer the resistance increase with multiple sensors is described in Section 2.4.

Results, given in Fig. 3, demonstrate that errors can be as high as 60% on the fifth and sixth resistance (using the metric of Eq. (15)) when only the temperature R_2 is measured. This is because the thermal impact of the sixth compartment on the second is weak, and diluted by heat diffusion through the network. If we observe both T_2 and T_6, the error on all resistances is extremely low in most cases (below 0.01%), except of some outliers for which the error is above 50% error on R_6. This experiment nevertheless shows the benefit of equipping an electric device with an additional temperature sensor, especially if the device consists of many components that are thermally interconnected.

4.3. Scenario 3: Model misspecification

The method we present in this report to determine contact resistances relies on accurate determination of the kernel functions k_0 and k_i for $i = 1, \dots, N$. In a realistic setting, the kernel k_0 can be simply determined by applying a step current and measuring the temperature increase, or with any other system identification approach using data measured on the real device. For the kernels k_i, instead, we would need to increase artificially each resistance by a known quantity before applying a step current. It could be difficult, or unfeasible, to obtain such a controlled increase in practice, especially in a device-specific fashion. We could instead determine kernels that fit an entire fleet of devices, modulo the variability due to different installations. Specifically, we could use an experimental or simulated setup to determine universal resistance kernels $\widetilde{k}_i$ that are common to a whole fleet of devices, maintaining a base kernel k_0 that is specific to an individual installed device. The inferred resistances are then obtained as the solution to the linear system

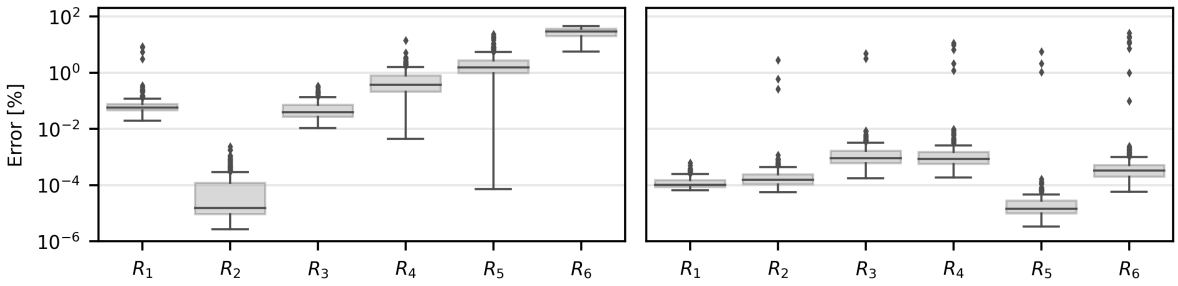
$$\widetilde{\mathbf{K}}^\top \widetilde{\mathbf{K}} \widehat{\Delta R} = \widetilde{\mathbf{K}}^\top (\mathbf{T} - \mathbf{K}_0), \tag{17}$$

where $\widetilde{\mathbf{K}}$ is built as in Section 2.3 using the fixed, universal kernels $\widetilde{k}_i$. The major concern with this approach is the misspecification between the real kernel function k_0 and the universal ones k_i, especially in terms of incompatible time scales.

Summarizing, the procedure that we propose to deal with installation specificity would consist of the following steps:

- Determine a universal base kernel $\widetilde{k}_0$ in an experimental or simulated setup;
- Use $\widetilde{k}_0$ to determine universal kernels $\widetilde{k}_i$ for each resistance that needs to be monitored;
- For each installation of the device, redetermine device-specific base kernel k_0 applying constant current load;
- When needed, infer an increase in resistances using Eq. (17).

We test the procedure above using the three-compartment network of Fig. 1, with coefficients given in Table 1. In order to simulate installation-specific conditions, we modify multiple times the value of the coefficients α_i as $\widetilde{\alpha}_i \sim \log \mathcal{N}(\alpha_i, \sigma)$, for $i = 1, \dots, 3$, where a large value of σ mimics devices that are very sensitive to different installations. We consider



(a) Compartment 2 sensed

(b) Compartments 2 and 5 sensed

Figure 3. Inference of five resistances varied simultaneously in a six-compartment network. (a) One temperature sensor. (b) Two temperature sensors. Test setup given in Section 4.2. Box-plot whiskers indicate 1.5 times the interquartile range, and dots indicate outliers.

$\sigma = 0.4, 0.2, 0.1, 0.05$, and for each of these values we generate 200 values at random of the coefficients α to simulate 200 installations of the same device. Then, we infer the resistances R_1 and R_3 using Eq. (17). Note that we do not apply a resistance increase in this case, and just attempt to infer how impactful is a change of the nominal conditions onto the kernels.

Results, given in Fig. 4, demonstrate that re-calibrating only the base kernel k_0 for each installation is sufficient for keeping good accuracy in the inference of the resistances. Moreover, we see that a good inference result (below 1% except some outliers) can be achieved even for devices that are subject to high variability when installed (see the spread in temperature development in case $\sigma = 0.4$). We note that lower installation specificity results in smaller variability in the inferred resistances (see the width of the box-plots in case $\sigma = 0.05$).

4.4. Scenario 4: Noisy data

In all experiments above, we employed noiseless data for determining the kernel functions and for estimating the resistances in the model. In this section, we assess the impact of these two sources of noise on the estimation variability. We consider the simple three-element network of Fig. 1 with parameters as in the experiments above. We compute the base kernel k_0 and the kernels k_i associated to resistances $i = 1, 3$ by perturbing the temperature response to a step current with a Gaussian source of noise $\eta_k \sim \mathcal{N}(0, \sigma_k^2)$, where $\sigma_k > 0$. Then, we excite the system with the current profile of Eq. (16) and perturb the temperature response with a Gaussian source of noise $\eta_d \sim \mathcal{N}(0, \sigma_d^2)$, where $\sigma_d > 0$. We then infer the resistance increase without changing its value in the model, i.e., data is generated by imposing $\Delta R_i = 0$. We repeat the experiment for noise scales σ_d and σ_k ranging between 10^{-4} and 10^{-1}, and for each combination of σ_k and σ_d we repeat the experiment $M = 50$ times. At each j-th repetition, we record the estimated resistance variations $\widehat{\Delta R}_1^{(j)}$ and $\widehat{\Delta R}_3^{(j)}$. We measure variability in the estimation as the sum of the population standard deviations of the two estimated resistance increases, i.e.,

$$\text{variability} = \text{std}\left(\left\{\widehat{\Delta R}_1^{(j)}\right\}_{j=1}^{M}\right) + \text{std}\left(\left\{\widehat{\Delta R}_3^{(j)}\right\}_{j=1}^{M}\right),$$

where $\text{std}(\cdot)$ denotes population standard deviation. We repeat the estimation twice: once with raw kernel functions computed with Eqs. (4) and (5), and once by fitting exponential functions as explained in Section 4.4

Results, given in Fig. 5, demonstrate that the method we develop here is robust with respect to random sources of noise. As expected, the variability is a growing function of both σ_k and σ_d. We remark that fitting exponential functions to the kernels has a beneficial effect on the inference accuracy. Indeed, it can be noticed in Fig. 5 that the variability is slightly lower when thermal kernels are fitted with exponential functions.

5. Conclusion

We introduced a novel method based on thermal kernels to monitor the condition of an electric device given temperature measurements. This method allows the calculation of temperatures at specific locations for general linear heat diffusion problems including thermal networks, for which we demonstrated an equivalence analytically.

Thermal kernels are simple to fit to data due to their non-parametric nature, which prevents issues of poor identifiability. Indeed, the parameters of thermal networks as simple as the one of Fig. 1 can be cumbersome or even impossible to determine if only one of the compartments equips a temperature sensor, unless good priors on the parameters are available due to physical considerations or from the results of high-fidelity and high-cost simulations. This issue is completely circumvented by thermal networks, which absorb the effects of all the parameters of an equivalent network approach into

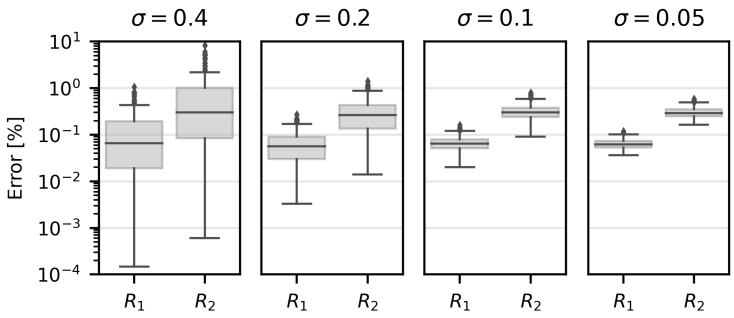
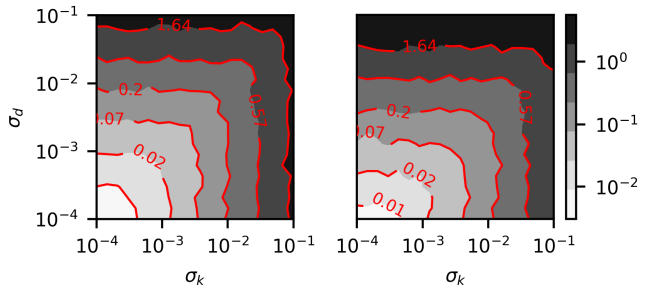


Figure 4. Results for four scales of model misspecification σ. Experiment setup in Section 4.3.



(a) Fitted kernels (b) Raw kernels

Figure 5. Estimation variability as a function of noise in the determination of the kernel functions (σ_k, horizontal axis), and in the data used for estimating the resistance values (σ_d, vertical axis). The contour values are in μΩ (a): Exponential fit for the kernel functions. (b): Raw kernel functions. Experiment setup in Section 4.4.

a simple data-driven linear transfer function.

We believe that thermal kernels should be preferred to thermal networks to monitor the linear heat sources (e.g., electrical resistances) of devices that are not equipped with a multitude of sensors, which would be required to fit the parameters of the network.

We suggest that future investigation may exploit the equivalence of thermal kernels and networks, e.g., to study whether knowing the former can be beneficial to improve the identifiability of the latter.

REFERENCES

Boodi, A., Beddiar, K., Amirat, Y., & Benbouzid, M. (2022). Building thermal-network models: a comparative analysis, recommendations, and perspectives. *Energies*, *15*(4), 1328.

Brastein, O., Ghaderi, A., Pfeiffer, C., & Skeie, N.-O. (2020). Analysing uncertainty in parameter estimation and prediction for grey-box building thermal behaviour models. *Energy and Buildings*, *224*, 110236.

Brastein, O. M., Lie, B., Sharma, R., & Skeie, N.-O. (2019). Parameter estimation for externally simulated thermal network models. *Energy and Buildings*, *191*, 200–210.

Hoffmann, M. W., Wildermuth, S., Gitzel, R., Boyaci, A., Gebhardt, J., Kaul, H., ... Tornede, T. (2020). Integration of novel sensors and machine learning for predictive maintenance in medium voltage switchgear to enable the energy and mobility revolutions. *Sensors*, *20*(7), 2099.

Raue, A., Karlsson, J., Saccomani, M. P., Jirstrand, M., & Timmer, J. (2014). Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics*, *30*(10), 1440–1448.

Stosur, M., Szewczyk, M., Sowa, K., Dawidowski, P., & Balcerek, P. (2016). Thermal behaviour analyses of gas-insulated switchgear compartment using thermal network method. *IET Generation, Transmission & Distribution*, *10*(12), 2833–2841.

BIOGRAPHIES

Giacomo Garegnani is a scientist at ABB corporate research. He obtained a PhD in Mathematics from EPFL in 2021, with a thesis on inverse problems involving partial and stochastic differential equations, and on probabilistic numerical methods. His research interests include uncertainty quantification of numerical solvers, model identifiability, and statistical inference for condition monitoring.

Kai Hencken is a corporate research fellow at ABB corporate research. He obtained a PhD in Theoretical Physics from the University of Basel in 1994. He was a post-doc at the University of Washington from 1995 to 1997 and at the University of Basel from 1997 to 2005, where he received his Habilitation in 2000 and is a lecturer since. In 2005 he joined the theoretical Physics group at ABB corpo-

rate research. His research interests include the combination of physical modeling with statistical methods to solve problems related to industrial devices, as well as developing diagnostics and prognostics approaches.

Frank Kassubek obtained a PhD in Physics from the University of Freiburg in 2000 (“Electrical and Mechanical Properties of Metallic Nanowires”). At ABB corporate research, he works on a wide range of topics including modeling of electrical systems and sensors, plasma and arc physics, and PHM topics.