Design and Experimental Validation of a Multiphysics Twin of a High-Voltage EV Motor

R. Torchio, F. Conte, A. Martin, N. Bianchi, *Fellow, IEEE*, M. De Soricellis, F. Toso, F. Pase, M. Scarpa, M. Filippini, M. Lurtz, and D. Szepanski

***Abstract*— The integration of electric motors into various industrial and automotive applications emphasizes the critical necessity for reliable performance and operational efficiency. The advent of advanced digital technologies offers opportunities for predictive maintenance strategies. Digital twins (DTs), mathematical models simulating a system's physical behavior in real-time, present a transformative approach to enhance real-time monitoring of critical quantities, which is imperative to improve operational efficiency and minimize downtime. In this article, we explore the feasibility and efficacy of deploying real-time physics-based DTs for condition monitoring in electric motor applications. Particularly, we focus on employing on-the-edge DTs, implemented on low-power on-board microprocessors, ensuring continuous communication with the physical asset for reliable real-time monitoring. The study applies DT technology to a high-voltage high-density electric vehicle (EV) motor, assessing its predictive capabilities in a real-world scenario. Results showcase the potential of DTs in revolutionizing condition monitoring, thereby meeting the evolving operational and maintenance requirements of contemporary electric motor systems.**

***Index Terms*— Automotive, digital twin (DT), electric motors, electric vehicle (EV), fault detection, predictive maintenance, real time.**

I. Introduction

THE increasing integration of electric motors in various industrial and automotive applications underscores the critical need for reliable performance and operational efficiency. As electric motors become pivotal components in modern systems, ensuring their continuous operation while minimizing downtime due to faults has emerged as a paramount concern [1], [2], [3], [4].

In traditional approaches, fault detection and diagnosis in electric motors have predominantly relied on postfault analysis or rudimentary monitoring techniques, often leading to reactive maintenance practices and unexpected downtime. However, with the advent of advanced digital technologies and the growing demand for predictive maintenance strategies, there arises an opportunity to revolutionize the way we monitor and manage electric motors [5], [6].

The evolution of electric motor applications demands a paradigm shift in operational and maintenance strategies to meet contemporary requirements efficiently. Conventional approaches in material research and hardware design, although effective to a certain extent, are no longer adequate in addressing the evolving operation and maintenance (O&M) needs of electric motor systems. The prevailing industrial practice often involves overengineering components and collecting extensive data, resulting in unsustainable costs and diminished profit margins [7].

Amidst these challenges, the emergence of digital twins (DTs), characterized as mathematical models capable of real-time simulation of a system's physical behavior, presents a promising solution to revolutionize condition monitoring and fault detection in electric motor systems [8], [9]. Analogous to their application in power electronics [10], [11], DTs offer a transformative approach to enhancing operational efficiency and minimizing downtime by enabling proactive maintenance strategies [12], [13]. The concept involves creating a virtual model of the electric motor physical attributes and operational characteristics within the motor control unit (MCU). This DT emulates the motor behavior under various operating conditions and load scenarios in real time [2].

The significance of such a digital replica lies not only in its ability to accurately simulate the motor performance but also in its integration with sophisticated condition monitoring and fault detection algorithms. By placing virtual sensors at critical points within the motor, including rotor magnets, windings, and bearings, the DT enables continuous monitoring of key parameters indicative of impending faults or performance degradation [1], [14].

One of the primary areas where DT technology can revolutionize electric motor applications lies in thermal management [15], [16], [17]. Temperature, particularly within critical components, such as rotor magnets, windings, and bearings, significantly influences motor performance, safety, and lifespan. However, traditional thermal management techniques

Manuscript received 19 March 2024; revised 27 May 2024; accepted 30 July 2024. Date of publication 2 August 2024; date of current version 3 February 2025. *(Corresponding author: R. Torchio.)*

R. Torchio is with the Department of Industrial Engineering and the Department of Information Engineering, University of Padova, 35131 Padua, Italy (e-mail: riccardo.torchio@unipd.it).

F. Conte and N. Bianchi are with the Department of Industrial Engineering, University of Padova, 35131 Padua, Italy.

A. Martin, M. De Soricellis, F. Toso, F. Pase, and M. Filippini are with Newtwen, 35127 Padua, Italy.

M. Scarpa is with the Department of Information Engineering, University of Padova, 35131 Padua, Italy.

M. Lurtz and D. Szepanski are with SEG Automotive Germany GmbH, 70499 Stuttgart, Germany.

Digital Object Identifier 10.1109/TTE.2024.3437475

often rely on sensors placed away from the crucial components, leading to latency and imprecise observations of temperature dynamics [18].

Moreover, the state of the art, which is characterized by oversensing, thermal network models with poor resolution for real-time execution, and oversized control safety margins, cannot meet the current market demand for high power density applications, particularly prevalent in the automotive sector [19], [20]. The reliance on these outdated methodologies hampers the ability to achieve optimal performance and efficiency, limiting the competitiveness of electric motor systems in the rapidly evolving automotive market.

In this article, we aim to analyze the feasibility and efficacy of deploying real-time physics-based digital replicas for condition monitoring. In contrast to the prevailing approach found in most existing literature, which utilizes in-cloud DTs, this article focuses on employing on-the-edge DTs [1], [9]. In this article, DTs are implemented on the low-power MCU already available on-board. This approach ensures the continuous communication of the DT with its corresponding physical asset, a crucial aspect when monitoring critical quantities in real time. This shift in methodology holds significant importance as it guarantees reliable and uninterrupted data exchange between the DT and its associated physical asset, enhancing the efficacy of real-time monitoring processes.

The remainder of this article is organized as follows. In Section II, the problem is discussed, and in Section III, the high power density EV motor is described, along with the test bench adopted for the experiments. Section IV provides a thorough description of the DT generation workflow, beginning with the construction of the finite element method (FEM)-based high fidelity model and its augmentation with state observers. Then, in Section V, emphasis is placed on avoiding the drift of the DT w.r.t. the manufactured physical asset by augmenting the physics-based model with data-driven artificial intelligence (AI)-based approaches. Finally, Section VI presents and discusses the results of the developed on-the-edge DT in terms of accuracy and computational cost, while conclusion is drawn in Section VII.

II. Problem Statement

The relentless pursuit of higher power densities and cost reductions in electric motor design to meet the ever-growing demand presents a formidable challenge: the complexity of thermal management. Overheating emerges as a critical issue, often leading to electric motor failures. Such overheating can stem from various factors, including insufficient cooling systems, elevated ambient temperatures, or prolonged operation under heavy loads. This thermal stress not only compromises the immediate performance of rare earth elements, resulting in significant efficiency loss over time but also accelerates the degradation of crucial component materials like the insulation film, potentially leading to short circuits or other electrical faults within the motor [21].

In addition, age and wear are inevitable factors in electromechanical devices, with high-voltage electric motors experiencing gradual performance decline over time [22].

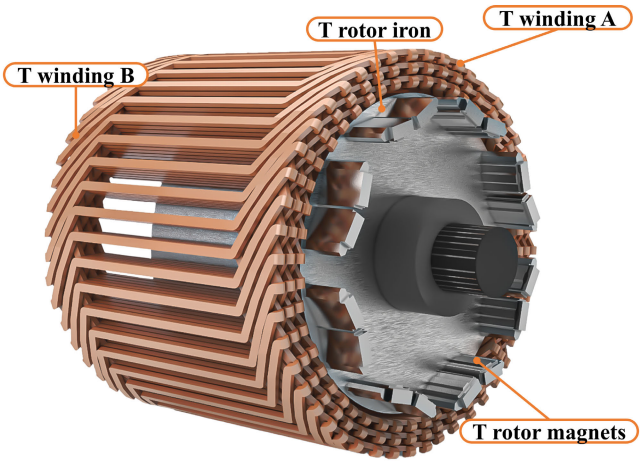

Fig. 1. EV motor rendering and sensor locations.

Addressing these challenges necessitates advancements in control systems and temperature monitoring techniques. However, current state-of-the-art approaches rely on physical sensors that cannot be feasibly implemented in critical areas, such as rotor magnets and bearings. Moreover, existing temperature estimation methods, typically based on thermal network models, lack the spatial resolution necessary for components with complicated thermal behavior. Incorporating nonlinearities, uncertainties in boundary conditions, time-varying parameters, and thermal exchange with fluids into thermal networks presents significant challenges. Indeed, although it is possible to add equivalent parameters in thermal networks, doing so while maintaining a coherent physical interpretation without resorting to a multiphysics model is particularly complex, especially in contexts such as liquid cooling systems for automotive motors [23].

The implementation of DT technology emerges as a promising solution to tackle these complex challenges. In this study, it has been applied this cutting-edge technology to a high-voltage electric vehicle (EV) motor. Utilizing a specialized prototype sample and test bench setup, we have rigorously assessed the performance of the DT. The aim is to demonstrate its predictive capabilities in a real-world scenario.

III. Test Bench and EV Motor Prototype Sample

To objectively and experimentally validate the methodology applied to the electric motor case, it was necessary to develop a special prototype of a sensorized motor with thermocouples integrated into the rotor at various depths. Specifically, the rotor was axially and radially drilled, and the thermocouples were located on the internal magnets, in the rotor iron core, and on the shaft. A wireless transducer was designed and attached to the shaft to acquire the measurements. It collects and transfers the real-time data required for model validation. Other thermocouples were attached to the stator iron, the casing, the windings, and the two stator end windings. In Fig. 1, the positioning of four specific thermocouples can be observed: two are in the end-windings, one is in the magnets, and one is in the rotor iron. The test bench is composed in this way to validate the model over a wide operating range, verifying

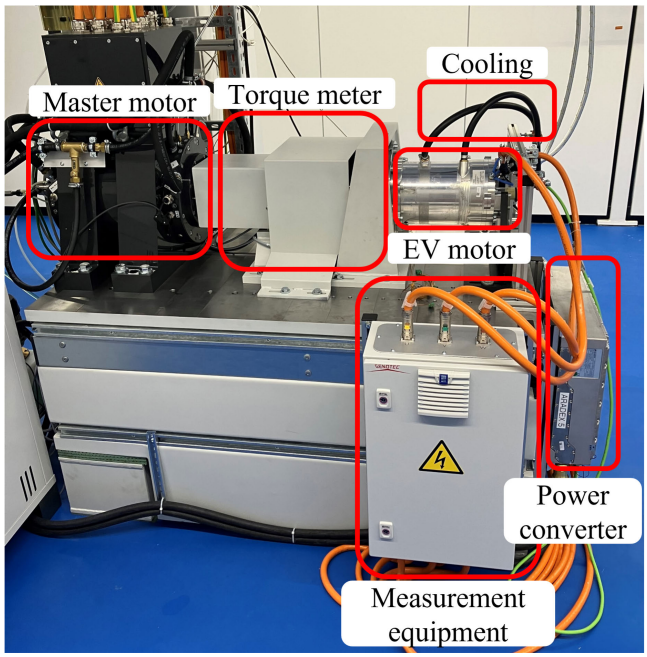

Fig. 2. Test bench provided for the study.

the performance at various speeds, load points in max torque per ampere (MTPA) and flux weakening (FW), and different flow rates and coolant temperatures. It consists of the motor under test, a master motor that can be controlled in torque (in which case it acts as a brake) or imposes a speed (thus acting as a motor), a torque meter that measures the torque at the mechanical coupling between the two motors, two inverters with separate control, and a cooling circuit regulated by a chiller that controls the flow rates and temperature of the coolant (water–glycol mixture) leaving it. In Fig. 2, the test bench just described can be observed, which was used for all the tests analyzed subsequently in this article.

IV. Twin Generation Workflow

A. High Fidelity Model

The model of the EV motor has an intrinsically multiphysics nature since electromagnetic (EM), thermal, and fluid dynamic effects must be considered to define the overall behavior of the device. With the final objective of generating an on-the-edge DT of the device for the real-time monitoring of critical temperatures, dedicated modeling strategies have been used to consider these three physics. In the following, these three models are described.

1) Power Loss Model: As in [24] and [25], an established approach for analyzing power losses in electrical machines employs the EM finite element analysis (FEA). This technique involves creating a detailed 2-D/3-D model that captures the geometry of the machine stator and rotor pole pair, alongside an accurate representation of EM material properties. By solving Maxwell's equations numerically, it is possible to carry out multiple EM simulations across the electrical period at different rated frequencies, thereby obtaining the field distribution. Subsequent postprocessing of these EM solutions allows for the estimation of fundamental losses in the copper windings, stator iron, rotor iron, and magnets using various numerical

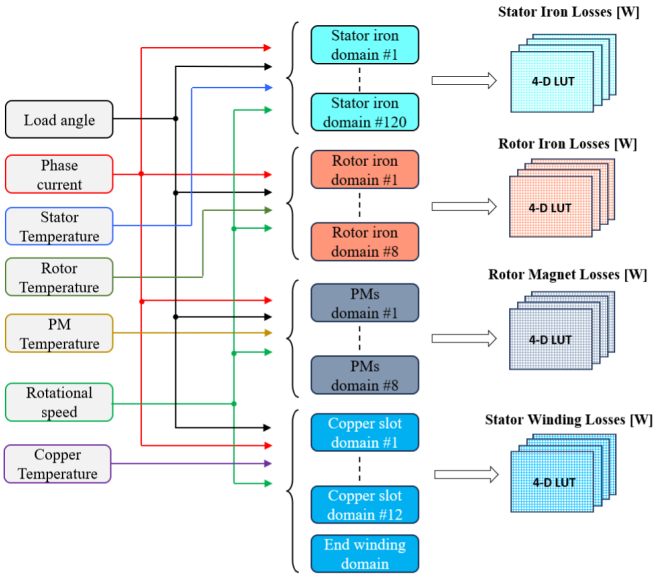

Fig. 3. Schematic losses structure of the EM model.

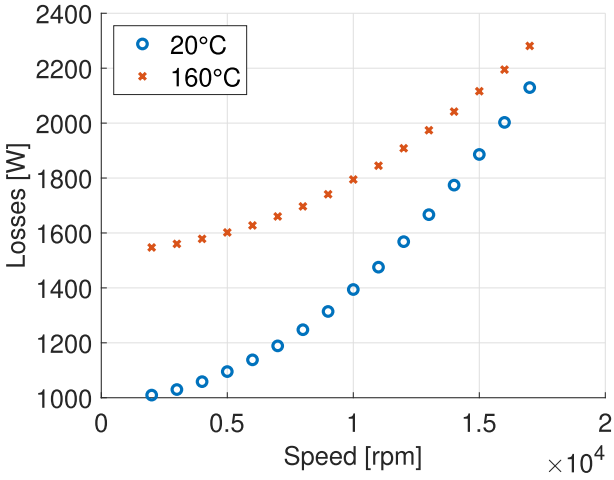

Fig. 4. Winding losses versus speed at 20 °C and 160 °C.

methods. These losses can be cataloged across a spectrum of operating conditions to generate look-up tables (LUTs). The tables account for variables, such as the amplitude and phase of current, rotational speed, and temperature across the full operational range of the electrical machines. Fig. 3 shows the loss structure developed for this study, while Figs. 4–6 show the losses related to different speed, temperatures of the copper windings and permanent magnets, and torque values.

As depicted in Fig. 3, the power loss model consists of a set of 4-D LUTs that depend on the load angle, the phase current, the rotational speed, and the average temperature of the stator, rotor, magnets, and windings. The load angle, phase current, and rotational speed will serve as inputs for the final DT model. Meanwhile, the temperatures will be generated by the DT itself and utilized as feedback for the power loss model.

Concerning the computational cost to extract the LUTs from EM simulations, they required about 1 h for a single point and about 2 days of overall computation time.

2) Thermal Model: The thermal model of the EV motor must be capable of providing the dynamic evolution of the temperature in the critical points shown in Fig. 1.

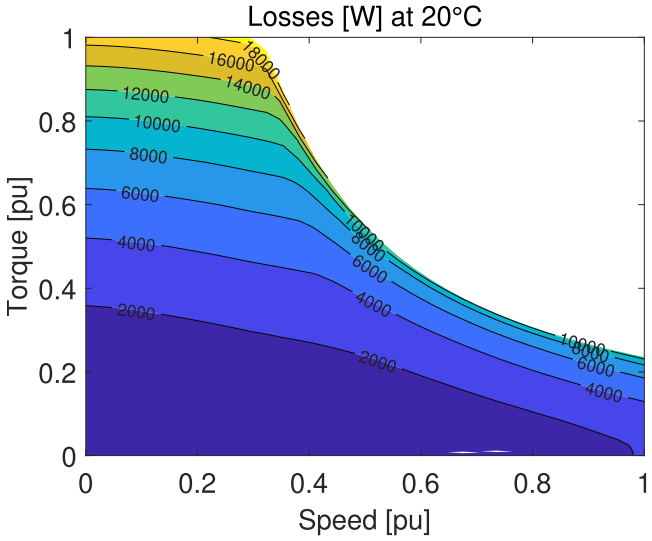

Fig. 5. Total losses computed at 20 °C.

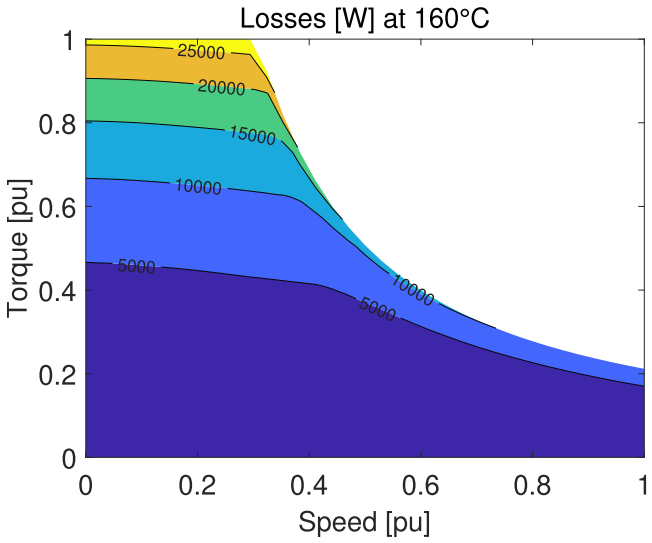

Fig. 6. Total losses computed at 160 °C.

The inputs of the thermal model are the power losses obtained from the EM model described in Section IV-A1 and the velocity field of the coolant provided by the fluid dynamic model described in Section IV-A3. The thermal model is described by the following well-known equation, i.e.,

$$\rho c_p \frac{\partial T}{\partial t} + \rho c_p \mathbf{v} \cdot \nabla T - \nabla \cdot k \nabla T = q \tag{1}$$

where ρ is the density, c_p is the heat capacity at constant pressure, T is the temperature, k is the thermal conductivity, q is the power density, and $\mathbf{v}$ is the velocity field (which is not zero only in the fluid region). The dependence w.r.t. the position has been neglected for simplicity. Equation (1) is then complemented by a convective boundary condition which is valid on the border of the model ($\partial\Omega$), i.e.,

$$\mathbf{n} \cdot k \nabla T = h(T_{\text{ext}} - T) \tag{2}$$

where $\mathbf{n}$ is the normal versor of the boundary of the motor, h is the convective coefficient, and T_{ext} is the external temperature.

TABLE I
MATERIAL THERMAL PROPERTIES

Material	Thermal Conductivity W/(m K)	Specific Heat J/(kg K)	Density kg/m^3
Stator Iron	40	434	7854
Rotor Iron	40	434	7854
Copper	401	385	8933
Permanent Magnet	10	500	7500
Aluminum	140	900	2680
Insulation	0.08	1100	1380
Shaft Iron	34	460	7850

The cooling conditions (i.e., the flow rate ϕ and the inlet temperature T_{inlet} of the coolant fluid) may change during the operation of the e-motor. Such dependence should be considered in the model. However, this would require the fluid velocity map, $\mathbf{v}$, to be time-varying. Although theoretically possible, as thoroughly discussed in Section IV-A3, it may result in a complex model, challenging to integrate on the edge for real-time execution. Therefore, the advection term associated with the heat exchange between the EV motor and the coolant fluid has been removed and replaced with a time-varying equivalent convective condition, i.e.,

$$\mathbf{n} \cdot k \nabla T = h_{\text{fluid}}(\phi)(T_{\text{fluid}}(\phi, T_{\text{inlet}}) - T) \tag{3}$$

where h_{fluid} is the equivalent heat transfer coefficient which depends on ϕ, and T_{fluid} is the temperature map of the fluid which depends on ϕ and T_{inlet}.

To generate a dynamic numeric model of (1)–(3), FEM is applied for the discretization. Thus, a computational model of the EV motor is generated, and a tetrahedral mesh is constructed (see Fig. 7).

Since the FEM model must be manipulated to apply model order reduction (MOR) techniques (as will be described in Section IV-B), a proprietary FEM code has been used. The model consists of about $N = 1.2 \times 10^6$ mesh elements. Thus, the final discretized model can be written as follows:

$$\mathbf{M}\frac{d\mathbf{x}}{dt} + (\mathbf{K} + \mathbf{H} + \mathbf{H_f})\mathbf{x} = \mathbf{Q_p}\mathbf{p} + \mathbf{Q_c}T_{\text{ext}} + \mathbf{Q_f}T_{\text{fluid}} \tag{4}$$

where $\mathbf{x}$ is the array of nodal temperature of dimension N, $\mathbf{M}$ is the mass matrix, $\mathbf{K}$ is the stiffness matrix, $\mathbf{H}$ and $\mathbf{H_f}$ are the stiffness matrices related to the convective boundary conditions, $\mathbf{p}$ is the power loss array of dimension N_p storing the losses (in [W]) for each domain (see Section IV-A1), $\mathbf{Q_p}$ is the $N \times N_p$ matrix which maps $\mathbf{p}$ into the rhs of the thermal model, and $\mathbf{Q_c}$ and $\mathbf{Q_f}$ are the array mapping T_{ext} and T_{fluid}, respectively, into the right-hand side (rhs) of the thermal model related to the convective boundary conditions.

The values of the material properties (i.e., density, heat capacity, and thermal conductivity) have been taken from data sheets and data reported in the literature. Table I shows such values.

Concerning the thermal model of the air gap, to account for the effect of the rotation speed on the heat exchange between stator and rotor through the air gap, as proposed by several works in the literature [26], an equivalent heat

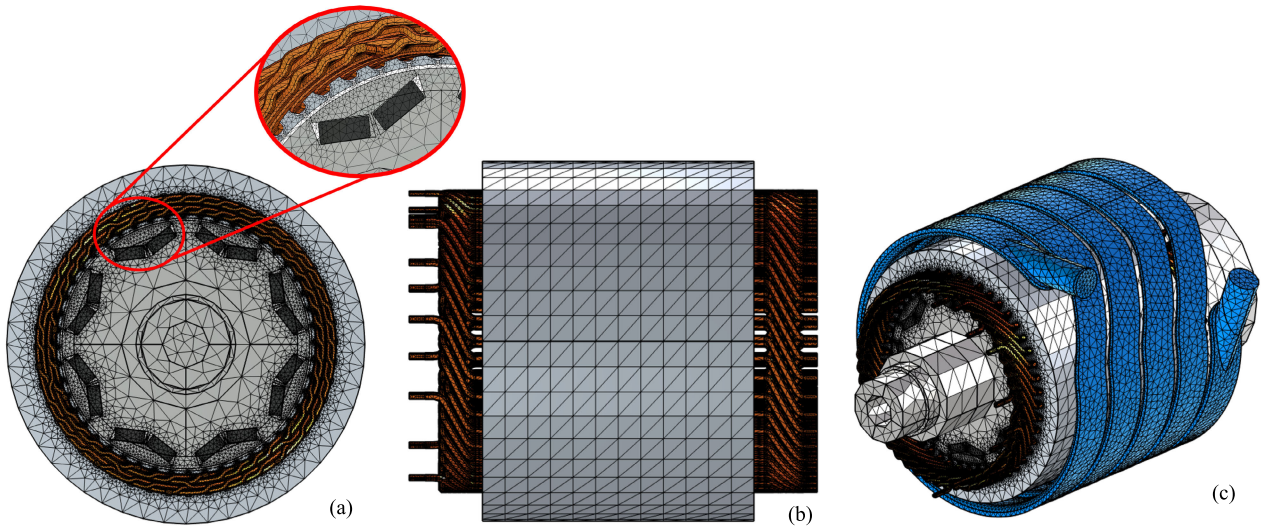

Fig. 7. Mesh of the 3-D FEM model. (a) Frontal view and (b) axial view of the motor. (c) 3-D view including the cooling system.

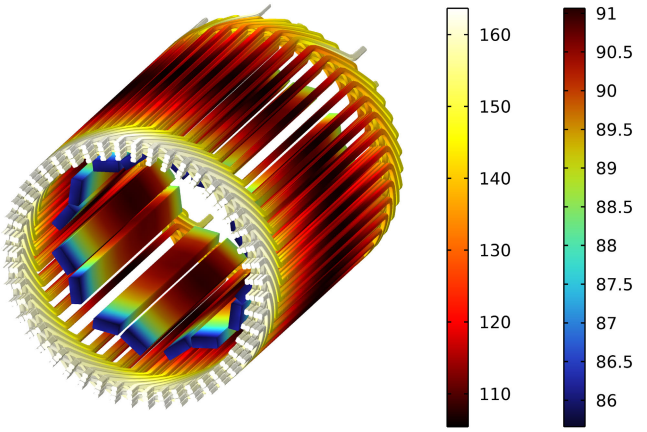

Fig. 8. Temperature distribution on windings and magnets. Values are in degree celsius.

transfer coefficient (which depends on the rotor speed) can be introduced, i.e.,

$$h_{\text{air gap}} = \frac{\text{Nu} \cdot k}{D_h} \tag{5}$$

where Nu is the Nusselt number, D_h is the hydraulic diameter [26], and k is the thermal conductivity of the air. Then, since the air gap is considered as a meshed domain in the thermal model, an equivalent, speed-dependent, thermal conductivity can be obtained from (5). However, this approach leads to a parametric model that, although compatible with the MOR strategies discussed in Section IV-B, would be too computationally intensive for real-time on-the-edge execution since one should handle a model with time-varying parameters. Therefore, an average value of the equivalent thermal air-gap conductivity has been selected, and its dependency on rotational speed is accounted for as discussed in Section V.

Concerning the computational cost, the computation time to solve (11) at each time step is approximately 5 min by using an algebraic multigrid solver with tolerance 10^{-6}.

Fig. 8 shows a snapshot (steady-state solution) of the temperature distribution on windings and magnets for the case of Calibration Set 2 (48 000:end) (see Table II).

3) Fluid Dynamics: The exchange of heat with the coolant fluid strongly affects the thermal behavior of the system. When it comes to modeling this phenomenon within the framework of electric machines, at least two strategies emerge as viable options, each with its advantages and limitations.

1) Using computational fluid dynamics (CFD) simulations to obtain the velocity map $\mathbf{v}$ which is then inserted in (1), resulting in the advection term. This approach allows for high accuracy and physics realism. However, solving CFD simulations is computationally complex, and the insertion of the advection term in (1) make its solution particularly challenging from the numerical point of view: even fine meshes lead to Peclet number Pe $>$ 1, which results in large node-to-node oscillations. To remove such oscillations, standard stabilization techniques (e.g., based on streamline upwind Petrov–Galerkin (SUPG) [27]) can be adopted.
2) Replacing the fluid with an equivalent convective boundary condition, i.e., (3). At the cost of sacrificing some physics realism, this solution significantly reduces the computational complexity of the model.

In this work, since the final model has to be compatible with on-the-edge implementation and real-time execution, the second approach was adopted in the final model but a simulation campaign with CFD simulations based on the first approach has been conducted to infer $h_{\text{fluid}}(\phi)$ and $T_{\text{fluid}}(\phi)$ to be used in (3). For the CFD simulations, the $k-\omega$ Reynolds-averaged (RANS) turbulence model (where k is the kinetic energy and ω is the specific dissipation rate) has been used

$$\begin{cases} \rho\dfrac{\partial k}{\partial t} + \rho(\mathbf{v}\cdot\nabla k) = P_k - \rho\beta^* k\omega + \nabla\cdot(\mu\sigma^*\mu_T)\nabla k \\ \rho\dfrac{\partial \omega}{\partial t} + \rho(\mathbf{v}\cdot\nabla \omega) = \alpha\dfrac{\omega}{k}P_k - \rho\beta\omega^2 + \nabla\cdot(\mu\sigma\mu_T)\nabla\omega. \end{cases} \tag{6}$$

The full definition of symbols in (6) is given in [28]. By solving (6) for several flow rate conditions (i.e., by varying ϕ), the velocity map as a function of the flow rate was obtained i.e., $\mathbf{v} = \mathbf{v}(\phi)$.

From the simulation campaign, the following models for h_{fluid} and T_{inlet} have been synthesized as a trade-off between accuracy and computational complexity

$$h_{\text{fluid}} \approx 9725\left(\frac{\phi}{\phi_{\text{nom}}}\right)^{0.65} \text{W/(m}^2 \cdot \text{K)} \tag{7}$$

$$T_{\text{fluid}} \approx T_{\text{inlet}} \ ^\circ\text{C} \tag{8}$$

where ϕ_{nom} is the nominal flow rate. To avoid dealing with a model with time-varying parameters, (7) has been replaced with an average value, i.e., $h_{\text{fluid}} = 9725$ W/(m^{2}K), this introduces an approximation that is treated in Section IV-C.

Concerning the computational cost, the number of DoFs was about 10^6 and the computation time required to solve the CFD problem was about 3 h.

B. Model Order Reduction

The primary feature distinguishing a DT from a high-fidelity model lies in its ability to be deployed either in the cloud or on-edge hardware, facilitating real-time or even faster than real-time execution. This enables a seamless data exchange between the physical asset and its corresponding DT. In this article, the DT is specifically designed for real-time monitoring of critical parameters. Relying solely on in-cloud implementations may prove unreliable due to inevitable communication delays. However, recent advancements in microprocessor technology offer a solution through on-chip DTs, where digital replicas are directly integrated into the available hardware on board.

Obviously, due to its large dimension, the high-fidelity thermal model described in Section IV-A is not directly compatible with the on-chip implementation. To solve this problem, MOR techniques can be used [29], [30].

For the real-time monitoring of critical temperatures, the FEM thermal model, i.e., (4), must be solved in real-time. Thus, its dimension must be significantly reduced to allow on-chip implementation. To do that, MOR strategies based on, e.g., balanced truncation, moment matching, or proper orthogonal decomposition can be used. The interested reader is referred to, see [31] for more details about different MOR strategies, which can be applied to both continuous and discrete models. Regardless of the adopted technique, MOR allows for projecting the original full order model (FOM) (4) into a reduced order space, i.e.,

$$\begin{aligned} \hat{\mathbf{E}}\frac{d\hat{\mathbf{x}}}{dt} &= \hat{\mathbf{A}}\hat{\mathbf{x}} + \hat{\mathbf{B}}\mathbf{u} \\ \mathbf{y} &= \hat{\mathbf{C}}\hat{\mathbf{x}} \end{aligned} \tag{9}$$

where $\hat{\mathbf{E}}$, $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, and $\hat{\mathbf{C}}$ have been obtained by writing (4) in (descriptor) state space form and then projecting the corresponding FOM matrices into the reduced order space. In (9), $\hat{\mathbf{x}}$ is the reduced order state, i.e., $\mathbf{x} \approx \mathbf{V}\hat{\mathbf{x}}$, where $\mathbf{V}$ is the projection basis function constructed by the adopted MOR strategy. The projection matrix $\mathbf{V}$ has dimension $N \times N_r$, where N_r is the dimension of the reduced order space and $N_r \ll N$. $\mathbf{y}$ is instead the vector that stores the temperature of interest. Since $\mathbf{V}$ has in general a limited number of columns (i.e., a small reduced order space is sufficient to accurately represent the dynamic of the quantity of interest stored in $\mathbf{y}$), the computational cost of solving the reduced order model (ROM) (9) is much smaller than the one required to solve the FOM, making it compatible with on-the-edge implementation. The ROM of the EV motor has been constructed by using moment matching techniques [32], [33], leading to a ROM of dimension $N_r = 22$.

Finally, (9) is discretized in time by applying a backward Euler scheme and it is written in (descriptor) state space form, i.e.,

$$\begin{aligned} \hat{\mathbf{E}}_{\mathbf{d}}\mathbf{x}_k &= \hat{\mathbf{A}}_{\mathbf{d}}\mathbf{x}_{k-1} + \hat{\mathbf{B}}_{\mathbf{d}}\mathbf{u}_{k-1} \\ \mathbf{y}_k &= \hat{\mathbf{C}}_{\mathbf{d}}\hat{\mathbf{x}}_k \end{aligned} \tag{10}$$

where $\mathbf{d}$ indicates that the matrices are the ones of the model discretized in time and $\mathbf{k}$ indicates the time instant $t = k\Delta t$, with $k = 0, \ldots, N_T$, and $\Delta_t = 100$ ms.

It is worth noting that more advanced time-stepping techniques may be applied to discretize (9). However, advanced time-stepping techniques may not be compatible with the final on-chip implementation of the DT. The backward Euler scheme is instead simple enough to be implemented in a standard microprocessor and, by choosing a small enough value of Δt, a good level of accuracy can be guaranteed.

C. State Observer

One of the key features of the DT is its real-time implementation. This allows the model to interact bidirectionally with the actual device through measurements from real sensors T_{meas} implemented in the system (i.e., T winding B in Fig. 1) and control actions from the electric drive implemented in the MCU. This enables the implementation of state-space observers such as the Kalman filter, moving horizon estimator, and particle filter, which are tasked with mitigating model uncertainties based on a priori information about the confidence given to both the model itself and the measurement system [2], [34].

In the specific case of the EV motor under test, an augmented Kalman filter (AKF) was applied to the system in the form of discrete reduced state space. Specifically, the reduced state was augmented with one new state variable: the inlet temperature T_{inlet} of the coolant fluid. The reason for this choice is primarily due to the prior information that this variable is time-varying in the real system, and its measurement is generally imprecise or sometimes absent. However, this is an extremely important input variable for modeling the heat exchange of the EV motor with the coolant fluid, as presented in the previous sections. Furthermore, by acting on the value of T_{inlet} it is even possible to compensate for uncertainties on the value of flow rate ϕ that is also a time-varying quantity.

Therefore, the final augmented state system, which also integrates the observer to be implemented in real-time, can

still be rewritten in discrete-time state-space form as follows:

$$\begin{aligned}\mathbf{X}_k &= \mathbf{A}_{\text{aug}}\mathbf{X}_{k-1} + \mathbf{B}_{\text{aug}}\mathbf{U}_k \\ \mathbf{y}_k &= \mathbf{C}_{\text{aug}}\mathbf{X}_k \end{aligned} \tag{11}$$

where$\mathbf{X} = [\mathbf{x}; T_{\text{inlet}}]$, $\mathbf{U} = [\mathbf{u}; T_{\text{meas}}]$, and $\mathbf{A}_{\text{aug}}$, $\mathbf{B}_{\text{aug}}$, and $\mathbf{C}_{\text{aug}}$ see [35] for the algebraic manipulations details.

V. From Model-as-Designed to Model-as-Manufactured

A. Limitations of a Pure Physic-Based as Designed Model

This section serves as one of the cornerstones regarding the novelty of this research work because the concept of transforming from an as-designed model to an as-manufactured model is key to completing the definition of a DT [36]. A model, no matter how complex and comprehensive, remains a simplification and an approximate representation of reality for the following main reasons in the specific case of the EV motor under test.

1) Approximated geometry due to the bottleneck of mesh size.
2) The partial differential equations of heat equation and Navier–Stokes for the fluid require numerical methods such as finite elements and finite volumes analysis to be solved with a certain degree of approximation.
3) The parameters of these equations, i.e., material properties, are inherently affected by uncertainty.
4) MOR techniques cause approximation and are effective only for linear or mildly nonlinear problems; hence, some physical phenomena such as radiation are not considered in the initial equations.
5) The uncertainty and variability of boundary conditions make them difficult to model, leading to a significant source of approximation.
6) Manufacturing errors, e.g., welding and stacking of stator and rotor stacks, leave the motor unique fingerprint, introducing a degree of uniqueness such that it is unthinkable to have a single model based solely on physics capable of accurately representing all samples of motor production.
7) Aging, wear, and degradation make material property parameters time-varying, but the temporal evolution functions to model these effects are unknown.

These inherent complexities highlight the need for a comprehensive approach that bridges the gap between the idealized as-designed model and the reality of the as-manufactured system, thereby embodying the essence of the DT concept in addressing the intricacies of real-world electric motor applications.

The proposed methodology for transforming an as-designed model into an as-manufactured model is based on the utilization of data, coupling a physical model with a data-driven model capable of mitigating the uncertainties and approximations just mentioned. This hybrid architecture, combining physics-based and data-driven approaches, aims to address the inherent complexities of real-world electric motor systems. By integrating such models within the MCU, this hybrid approach becomes the standard for implementing real-time models, providing enhanced accuracy and robustness in monitoring and controlling electric motor systems.

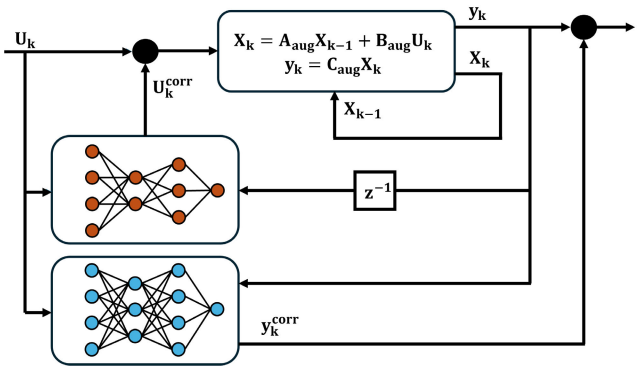

Fig. 9. Hybrid model architecture (physics-based and AI data-driven-based).

B. Model Architecture

The topology proposed in Fig. 9 to mitigate various sources of uncertainty and approximation involves the final reduced model obtained through the procedures described above and two feed forward neural networks (FFNNs) [37]. The first FFNN is used to correct potential uncertainties in the model inputs, such as incorrect distribution of losses in various domains of the stator, windings, rotor, and magnets, as well as potential uncertainties in boundary conditions such as external ambient temperature. Meanwhile, the second FFNN directly corrects the model output of interest, thus mitigating all those intrinsic errors of the physics-based model listed from items 1 to 4 in Section V-A. The FFNNs take as input the power losses generated by the power loss models and the temperatures estimated by the state space model. Then, they were trained from the mean-squared error loss computed against the measured temperatures and the estimated ones provided by the DT. It is important to note that the temperature inputs of the FFNN that acts on the inputs of the state space model are time-delayed. In essence, this FFNN adjusts the state space model input by using the DT output from the previous time step, which is known before the execution of the state space model. The final model architecture is calibrated using a substantially reduced experimental dataset, as much of the system dynamics are already present as intrinsic information from the physics-based model and do not need to be inferred. In addition, thanks to the presence of the physics-based model, the number of layers and the dimensionality of the FFNNs can be substantially reduced compared with a fully data-driven approach, thus maintaining a structure and computational complexity suitable for real-time integration on a microcontroller. The choice of the specific FFNNs is derived after a series of studies on different architectures, including recurrent neural networks (RNNs). These models were compared against the results obtained with unseen data during the validation phase. Other more complex architectures did not yield significant performance improvements to justify the increased complexity introduced by those models, in terms of both memory usage and computation time. Therefore, we decided to use FFNNs, with the final architecture parameterized as outlined

in Section VI, achieving a favorable balance between accuracy and complexity.

C. Calibration

Training FFNNs involves optimizing model parameters and selecting methods to enhance predictive performance. Common techniques include gradient-based optimization such as stochastic gradient descent (SGD) and Adam algorithm [38], along with backpropagation for efficient gradient computation. Regularization methods like $\mathcal{L}_1$- and $\mathcal{L}_2$-regularization, and dropout are employed to prevent overfitting. PyTorch and TensorFlow are prominent libraries for FFNN development. PyTorch, known for its flexibility and dynamic computation graph, offers extensive support for model building [39]. TensorFlow, backed by Google Brain, provides a high-level API and tools for deployment. Keras, integrated into TensorFlow, remains a popular choice for its user-friendly interface. Training and calibrating FFNNs require careful selection and application of optimization algorithms, regularization techniques, and appropriate libraries. PyTorch and TensorFlow stand out as leading frameworks due to their rich features, active communities, and widespread adoption in both research and industry.

Concerning the training of the FFNNs, the experimental dataset has been obtained from the test bench described in Section III. The dataset comprises a sequence of experiments performed under several operating conditions. We then split those experiments into two subsets. The former, with approximately 80% of the available data collected, was used to train the model. The remaining data were used after the training to ensure the optimized model could well generalize the correction needed under different scenarios.

D. Validation

Generalization of the final model architecture is crucial for ensuring robust and reliable performance in real-world applications. This is particularly significant when integrating physics-based models with FFNNs, as it allows the model to capture underlying physical principles while leveraging the flexibility of neural networks for complex pattern recognition. Integrating physics-based models into FFNN architectures enhances the interpretability and physical plausibility of the model predictions. By incorporating domain knowledge and fundamental principles, these models provide constraints that guide the learning process, promoting better generalization to unseen data. Maintaining a certain degree of physicality in the model architecture is essential to prevent overfitting and increase reliability. Overfitting occurs when the model learns to memorize training data rather than capturing underlying patterns, leading to poor generalization. By incorporating physics-based constraints, the model is less likely to extrapolate erroneously and more capable of making accurate predictions in diverse operating scenarios. To evaluate the generalization capability of the final model architecture, it is crucial to test its performance in operating scenarios never seen during the training phase. This ensures that the model can effectively extrapolate beyond the training data and provides confidence in its reliability for real-world applications.

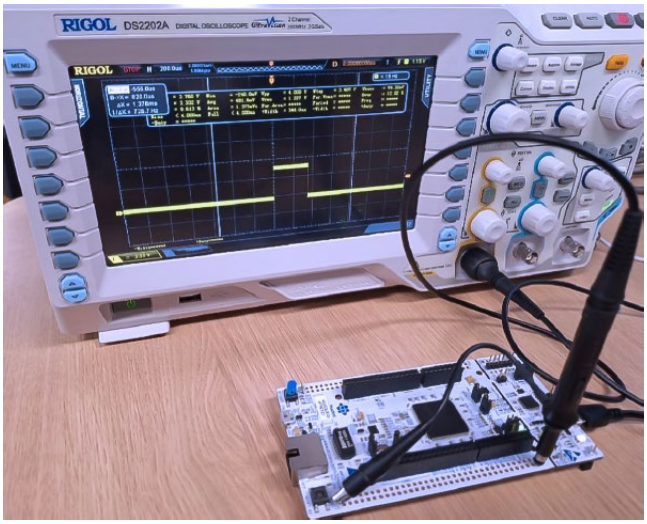

Fig. 10. Execution time measurement.

TABLE II
CONTINUOUS OPERATING CONDITIONS

Test	Current I_d [A]	Current I_q [A]	Speed ω [rpm]
Calibration Set 1 (0:48000)	-151	66	10000
Calibration Set 1 (48000:end)	-133	84	8000
Calibration Set 2 (0:48000)	-80	129	4000
Calibration Set 2 (48000:end)	-89	138	2000
Calibration Set 3 (0:48000)	-76	114	6000
Calibration Set 3 (48000:96000)	-82	46	10000
Calibration Set 3 (96000:end)	-45	55	8000
Validation Set (0:48000)	-37	83	4000
Validation Set (48000:end)	-37	83	2000

VI. RESULTS, DISCUSSION, AND ADDED VALUE

In this section, the results obtained from the calibration of the DT described earlier and the validation tests to demonstrate its reliability and accuracy are presented. The final DT consists of the power loss model (i.e., LUTs), the thermal model (a state space model), the AKF, and the FFNNs. Thus, the inputs of the power loss model are also the inputs of the final DT, i.e., the phase current (amplitude and load angle) and the rotational speed. Note that the temperatures needed to extract the losses from the LUTs are generated by the DT itself, thus they are internal states and outputs of the DT. The thermal sensor (T_{meas}) exploited by the AKF to infer T_{inlet} is also considered as input of the DT.

The final state-space model has 22 degrees of freedom (DoFs), resulting in a state matrix A of size 22×22. The two FFNNs each have 419 parameters, with three intermediate hidden layers consisting of 8, 12, and 8 neurons, respectively, and their weight optimization took approximately 2 h considering both the calibration and validation phase. The entire model architecture is executed on an STM32-based evaluation board with an execution time of 256.4 μs, as observed from the real-time feasibility test depicted in Fig. 10. In Table II, the different operating points of the various tests are shown. As can be observed, various conditions were tested in terms of

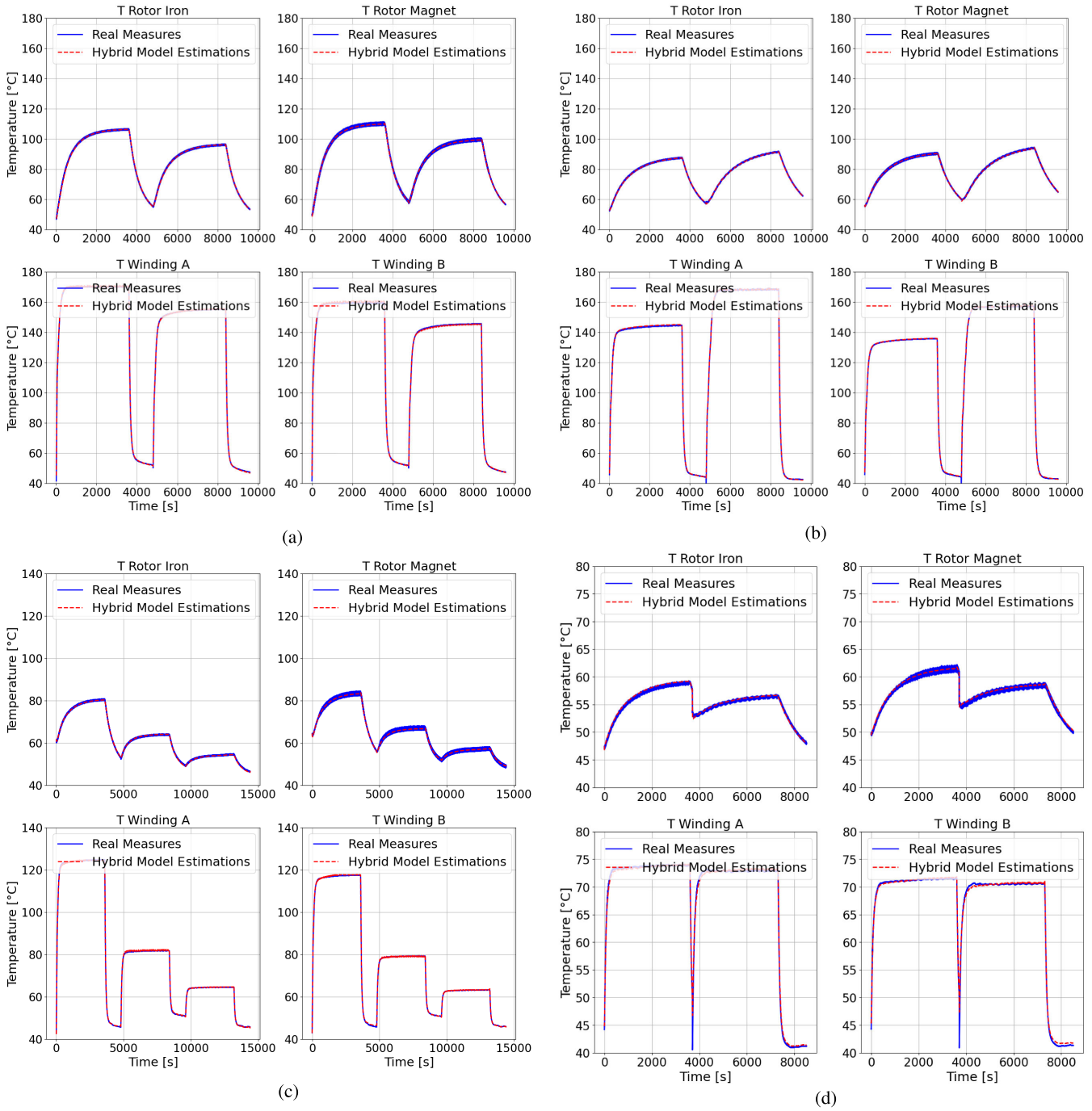

Fig. 11. Calibration and validation sets. (a) Calibration Set 1. (b) Calibration Set 2. (c) Calibration Set 3. (d) Validation set.

torque and speed to observe different thermal loads in terms of absolute value and distribution of losses in the various components, e.g., stator, rotor, and windings. The first three sets of data recorded with the special prototype sample and used in the calibration phase for the hybrid physics-based-data-driven model architecture are shown in Fig. 11(a)–(c), respectively. The post-calibration error is less than 1.5 °C across the entire explored dynamics. A notable difference in dynamics can be observed between the stator (liquid-cooled with a water–glycol cooling circuit with a flow rate of 10 L/min and a temperature of 50 °C) and the rotor insulated from air at the air gap, which has a significantly larger thermal time constant compared to the stator. The validation test performed and shown in Fig. 11(d) yields very accurate results, with a maximum error of 2.5 °C across the entire range of tests conducted at a working point never analyzed during training. This makes the model extremely reliable throughout the machine's operational range, allowing it to be used for control purposes, especially for managing power derating while considering magnet temperature faithfully, thereby avoiding local demagnetization and minimizing safety margins to maximize machine functionality.

Another very important case study, especially for automotive applications, is the peak torque tests, crucial for one

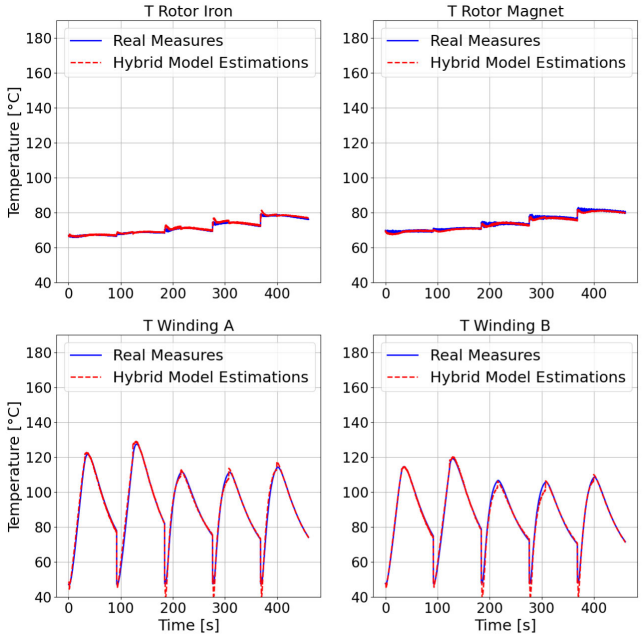

Fig. 12. Peak torque calibration set. Note that, for the sake of conciseness, five different tests are reported in this figure.

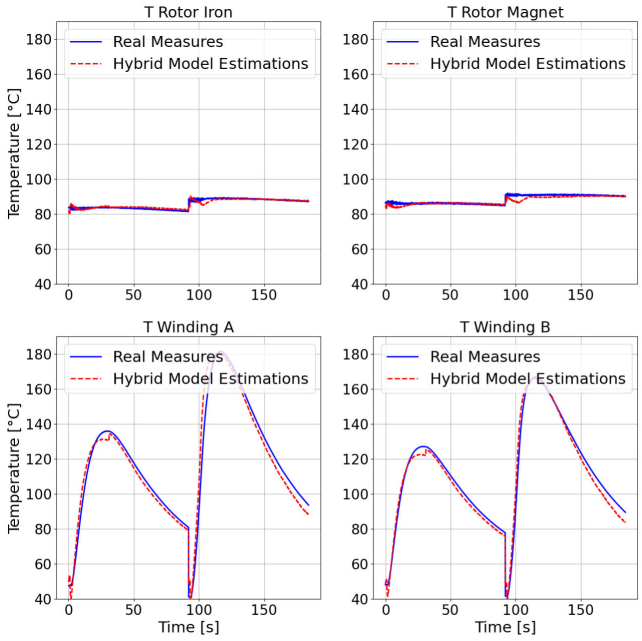

Fig. 13. Peak torque validation set. Note that, for the sake of conciseness, only two different tests are reported in this figure.

of the most important key performance indicators (KPIs) in automotive, namely, acceleration from 0 to 100 km/h. These stress tests require a high current value on the stator, operating at predefined I_d and I_q working points depending on the set rotor speed. They are short-duration tests, less than 100 s, to verify the maximum duration of the peak torque that the motor can achieve and the consequent thermal loads it can withstand. Fig. 12 shows the tests used by the model training algorithm to integrate extreme operating points and learn their dynamics. In particular, five tests at different currents and speeds were used during calibration, the details of the operating conditions of which are reported in Table III.

TABLE III
PEAK TORQUE OPERATING CONDITIONS

Test	Current I_d [A]	Current I_q [A]	Speed ω [rpm]
Calibration Set (0:92)	-150	200	4000
Calibration Set (92:184)	-150	200	2000
Calibration Set (184:276)	-300	100	10000
Calibration Set (276:368)	-350	100	8000
Calibration Set (368:end)	-350	150	6000
Validation Set (0:92)	-350	350	4000
Validation Set (92:end)	-400	400	2000

Regarding the validation of the model, the operating points of maximum current were used precisely to verify the reliability and degree of generalization obtained from the final model architecture. In Fig. 13, it is possible to observe the real-time estimation performance obtained, where once again the metric on the maximum acceptable error (±2.5 °C) has been met. We can especially notice in these tests the difference in stator and rotor dynamics, suggesting that, especially for random ON/OFF applications like automotive, without a reliable model estimating the rotor temperature, the control is completely blind. This is the main reason why this technology can enhance the control performance of electric machines, significantly reducing the overly cautious safety margins previously used in the state-of-the-art due to concerns about overheating the rotor and irreversibly compromising the machine's efficiency.

Finally, it is worth remarking that the online monitoring of temperature in critical points of the EV motor also unlocks the possibility of monitoring the remaining useful life (RUL) of the device. For instance, as described in [22] and [40], the RUL of the insulation directly depends on the temperature. Thus, an accurate and continuous knowledge of the temperature can be used to predict the RUL of the insulation, allowing for predictive maintenance and preventing faults.

VII. CONCLUSION

This article demonstrates the feasibility and effectiveness of employing physics-based DTs to monitor critical temperatures in high-power density EV motors in real time. We extensively discuss the workflow for generating the physics-based model, emphasizing efforts to reduce computational complexity without compromising fidelity to enable real-time execution and on-the-edge implementation. This facilitates seamless data exchange between the physical asset and its corresponding DT. We leverage state observers and AI data-driven augmentation to mitigate DT drifting, ensuring high accuracy during real-time operation.

Experimental results validate the model accuracy under realistic conditions, including peak torque tests, where errors smaller than 2.5 °C are achieved. Future research will explore leveraging DTs to implement advanced and informed control strategies. Future works will explore different approaches to construct the DT, such as different MOR strategies, different schemes to join the physics-based and data-driven models, and different AI techniques. Moreover, comparison in terms of accuracy, computational effort, and automatization of the whole approach will also be carried out in dedicated studies.

REFERENCES

[1] S. Venkatesan, K. Manickavasagam, N. Tengenkai, and N. Vijayalakshmi, "Health monitoring and prognosis of electric vehicle motor using intelligent-digital twin," *IET Electr. Power Appl.*, vol. 13, no. 9, pp. 1328–1335, Sep. 2019, doi: 10.1049/iet-epa.2018.5732.

[2] O. Wallscheid, "Thermal monitoring of electric motors: State-of-the-art review and future challenges," *IEEE Open J. Ind. Appl.*, vol. 2, pp. 204–223, 2021, doi: 10.1109/OJIA.2021.3091870.

[3] J. Medina-García, T. Sánchez-Rodríguez, J. Galán, A. Delgado, F. Gómez-Bravo, and R. Jiménez, "A wireless sensor system for real-time monitoring and fault detection of motor arrays," *Sensors*, vol. 17, no. 3, p. 469, Feb. 2017, doi: 10.3390/s17030469.

[4] K. Mykoniatis, "A real-time condition monitoring and maintenance management system for low voltage industrial motors using Internet-of-Things," *Proc. Manuf.*, vol. 42, pp. 450–456, Jan. 2020, doi: 10.1016/j.promfg.2020.02.050.

[5] S. K. Gundewar and P. V. Kane, "Condition monitoring and fault diagnosis of induction motor," *J. Vib. Eng. Technol.*, vol. 9, no. 4, pp. 643–674, Jun. 2021, doi: 10.1007/s42417-020-00253-y.

[6] L. Magadán, F. J. Suárez, J. C. Granda, and D. F. García, "Low-cost real-time monitoring of electric motors for the industry 4.0," *Proc. Manuf.*, vol. 42, pp. 393–398, Jan. 2020, doi: 10.1016/j.promfg.2020.02.057.

[7] J. Antonino-Daviu, "Electrical monitoring under transient conditions: A new paradigm in electric motors predictive maintenance," *Appl. Sci.*, vol. 10, no. 17, p. 6137, Sep. 2020, doi: 10.3390/app10176137.

[8] G. Bhatti, H. Mohan, and R. R. Singh, "Towards the future of smart electric vehicles: Digital twin technology," *Renew. Sustain. Energy Rev.*, vol. 141, May 2021, Art. no. 110801, doi: 10.1016/j.rser.2021.110801.

[9] Y. Jiang, S. Yin, K. Li, H. Luo, and O. Kaynak, "Industrial applications of digital twins," *Phil. Trans. Roy. Soc. A: Math., Phys. Eng. Sci.*, vol. 379, no. 2207, Oct. 2021, Art. no. 20200360, doi: 10.1098/rsta.2020.0360.

[10] H. Chen, Z. Zhang, P. Karamanakos, and J. Rodriguez, "Digital twin techniques for power electronics-based energy conversion systems: A survey of concepts, application scenarios, future challenges, and trends," *IEEE Ind. Electron. Mag.*, vol. 17, no. 2, pp. 20–36, Nov. 2022, doi: 10.1109/MIE.2022.3216719.

[11] B. Rodríguez, E. Sanjurjo, M. Tranchero, C. Romano, and F. González, "Thermal parameter and state estimation for digital twins of E-powertrain components," *IEEE Access*, vol. 9, pp. 97384–97400, 2021, doi: 10.1109/ACCESS.2021.3094312.

[12] S. V. N. Sreenivasu, T. S. Kumar, O. B. Hussain, A. R. Yeruva, S. R. Kabat, and A. Chaturvedi, "Cloud based electric vehicle's temperature monitoring system using IotT," *Cybern. Syst.*, pp. 1–16, 2023, doi: 10.1080/01969722.2023.2176649.

[13] J. F. D. Santos et al., "Digital twin-based monitoring system of induction motors using IoT sensors and thermo-magnetic finite element analysis," *IEEE Access*, vol. 11, pp. 1682–1693, 2023, doi: 10.1109/ACCESS.2022.3232063.

[14] F. Toso, R. Torchio, A. Favato, P. G. Carlet, S. Bolognani, and P. Alotto, "Digital twins as electric motor soft-sensors in the automotive industry," in *Proc. IEEE Int. Workshop Metrol. Automot. (MetroAutomotive)*, Jul. 2021, pp. 13–18, doi: 10.1109/MetroAutomotive50197.2021.9502885.

[15] M. Al-Gabalawy, A. H. Elmetwaly, R. A. Younis, and A. I. Omar, "Temperature prediction for electric vehicles of permanent magnet synchronous motor using robust machine learning tools," *J. Ambient Intell. Humanized Comput.*, vol. 15, no. 1, pp. 243–260, Jan. 2024.

[16] J. Ma, Y. Sun, S. Zhang, J. Li, and S. Li, "Experimental study on the performance of vehicle integrated thermal management system for pure electric vehicles," *Energy Convers. Manage.*, vol. 253, Feb. 2022, Art. no. 115183, doi: 10.1016/j.enconman.2021.115183.

[17] S. Schaut, E. Arnold, and O. Sawodny, "Predictive thermal management for an electric vehicle powertrain," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 2, pp. 1957–1970, Feb. 2023, doi: 10.1109/TIV.2021.3131944.

[18] X. Wang et al., "A critical review on thermal management technologies for motors in electric cars," *Appl. Thermal Eng.*, vol. 201, Jan. 2022, Art. no. 117758, doi: 10.1016/j.applthermaleng.2021.117758.

[19] R. Burke, A. Giedymin, Z. Wu, H. Chuan, N. Bourne, and J. G. Hawley, "A lumped parameter thermal model for single-sided AFPM machines with experimental validation," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 3, pp. 1065–1083, Sep. 2020, doi: 10.1109/TTE.2020.2998110.

[20] W. Kirchgässner, O. Wallscheid, and J. Böcker, "Estimating electric motor temperatures with deep residual machine learning," *IEEE Trans. Power Electron.*, vol. 36, no. 7, pp. 7480–7488, Jul. 2021, doi: 10.1109/TPEL.2020.3045596.

[21] S. Ramarathnam, A. K. Mohammed, B. Bilgin, A. Sathyan, H. Dadkhah, and A. Emadi, "A review of structural and thermal analysis of traction motors," *IEEE Trans. Transport. Electrific.*, vol. 1, no. 3, pp. 255–265, Oct. 2015, doi: 10.1109/TTE.2015.2476478.

[22] P. Giangrande, V. Madonna, S. Nuzzo, and M. Galea, "Moving toward a reliability-oriented design approach of low-voltage electrical machines by including insulation thermal aging considerations," *IEEE Trans. Transport. Electrific.*, vol. 6, no. 1, pp. 16–27, Mar. 2020, doi: 10.1109/TTE.2020.2971191.

[23] J. Huang et al., "A hybrid electric vehicle motor cooling system—Design, model, and control," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4467–4478, May 2019, doi: 10.1109/TVT.2019.2902135.

[24] A. Tameemi et al., "Power loss and performance analysis of a permanent magnet synchronous motor for actuator applications," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, Sep. 2022, pp. 807–813, doi: 10.1109/ICEM51905.2022.9910911.

[25] E. Campara and A. Solakovic, "Iron loss modelling for a permanent magnet synchronous motor," M.S thesis, Competence Center Electr. Power Eng., Permanent Magnet Drives (PMD) Res. Group, Chalmers Univ. Technol., Göthenburg, Sweden, 2020.

[26] D. A. Howey, P. R. N. Childs, and A. S. Holmes, "Air-gap convection in rotating electrical machines," *IEEE Trans. Ind. Electron.*, vol. 59, no. 3, pp. 1367–1375, Mar. 2012, doi: 10.1109/TIE.2010.2100337.

[27] R. Codina, "On stabilized finite element methods for linear systems of convection–diffusion-reaction equations," *Comput. Methods Appl. Mech. Eng.*, vol. 188, nos. 1–3, pp. 61–82, Jul. 2000, doi: 10.1016/s0045-7825(00)00177-8.

[28] D. C. Wilcox, "Formulation of the k-w turbulence model revisited," *AIAA J.*, vol. 46, no. 11, pp. 2823–2838, Nov. 2008, doi: 10.2514/1.36541.

[29] D. Hartmann, M. Herz, and U. Wever, "Model order reduction a key technology for digital twins," in *Reduced-Order Modeling (ROM) for Simulation and Optimization*. Cham, Switzerland: Springer, 2018, pp. 167–179.

[30] M. G. Kapteyn, D. J. Knezevic, and K. Willcox, "Toward predictive digital twins via component-based reduced-order models and interpretable machine learning," in *Proc. AIAA Scitech Forum*, Jan. 2020, p. 0418, doi: 10.2514/6.2020-0418.

[31] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. M. Silveira, Eds., *Model Order Reduction: Volume 1: System- and Data-Driven Methods and Algorithms*. Berlin, Germany: De Gruyter, 2021.

[32] X. Peng, Q. Bai, X. Xia, Z. Huang, K. Saenko, and B. Wang, "Moment matching for multi-source domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1406–1415.

[33] L. Codecasa, D. D'Amore, and P. Maffezzoni, "Multipoint moment matching reduction from port responses of dynamic thermal networks," *IEEE Trans. Compon. Packag. Technol.*, vol. 28, no. 4, pp. 605–614, Dec. 2005, doi: 10.1109/TCAPT.2005.859741.

[34] P. N. Phuc, D. Bozalakov, H. Vansompel, K. Stockman, and G. Crevecoeur, "Rotor temperature virtual sensing for induction machines using a lumped-parameter thermal network and dual Kalman filtering," *IEEE Trans. Energy Convers.*, vol. 36, no. 3, pp. 1688–1699, Sep. 2021, doi: 10.1109/TEC.2021.3060478.

[35] R. G. Brown and P. Y. C. Hwang, Eds., *Introduction to Random Signals and Applied Kalman Filtering*. New Jersey, NJ, USA: Logo, 2012.

[36] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, "A probabilistic graphical model foundation for enabling predictive digital twins at scale," *Nature Comput. Sci.*, vol. 1, no. 5, pp. 337–347, May 2021, doi: 10.1038/s43588-021-00069-0.

[37] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.*, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.

[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[39] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32. Red Hook, NY, USA: Curran Associates, 2019, pp. 1–12.

[40] H. Oraee, "A quantative approach to estimate the life expectancy of motor insulation systems," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 7, no. 6, pp. 790–796, Dec. 2000, doi: 10.1109/94.891990.

Open Access funding provided by 'Università degli Studi di Padova' within the CRUI CARE Agreement