PAPER • OPEN ACCESS

Development and benchmarking of a thermal lumped model for the magnetron injection gun of the MW-power ITER European gyrotron

To cite this article: A. Antonione *et al* 2025 *Nucl. Fusion* **65** 036034

OPEN ACCESS
IOP Publishing | International Atomic Energy Agency
Nuclear Fusion
Nucl. Fusion **65** (2025) 036034 (16pp)
https://doi.org/10.1088/1741-4326/ada8be

Development and benchmarking of a thermal lumped model for the magnetron injection gun of the MW-power ITER European gyrotron

A. Antonione[1,2,3,*], R. Bertazzoni[4], G. Carannante[5], K. Cindric[3], M. Ferrari[5], A. Leggieri[6], F. Sanchez[3] and G. De Tommasi[1,2]

[1] Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università degli Studi di Napoli Federico II, via Claudio 21, 80125 Napoli, Italy
[2] Consorzio CREATE, via Claudio 21, 80125 Napoli, Italy
[3] Fusion for Energy,Josep Pla 2, Barcelona 08019, Spain
[4] Institute for Pulsed Power and Microwave Technology (IHM), 76131 Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[5] ITER Organization, Route de Vinon sur Verdon, 13067 Saint Paul Lez Durance Cedex, France
[6] Microwave & Imaging Sub-Systems, THALES Vélizy-Villacoublay, F-78141 Vélizy-Villacoublay, France

E-mail: andrea.antonione@unina.it

Received 9 October 2024, revised 17 December 2024
Accepted for publication 10 January 2025
Published 25 February 2025

Abstract

With ITER's re-baseline, the Electron Cyclotron Heating system has become of even higher importance with the multiple purposes of plasma heating, current drive, and instability suppression. ITER's Electron Cyclotron Heating system is based on several gyrotrons, such as radio-frequency sources supposed to provide *on-demand* power in the order of megawatts in the millimeter-wave domain. Due to the non-linear dynamics involved in the gyrotron operation and the structure of the machine itself, controlling a gyrotron and preparing it in the shortest time possible to provide radiofrequency power represents a challenging task for the tokamak operations. Reliable models of the power source are essential to design effective control strategies. To this aim, we further developed a semi-physical thermal lumped model of the ITER European Gyrotron's Magnetron Injection Gun available in literature. The model has been re-implemented in C++ for better computational performances and therefore making it particularly suitable to be exploited by data-driven model identification algorithms and real-time control codes. In this paper the model parameters are identified by using a Particle Swarm Optimization algorithm on experimental data and the model validation is finally carried out on a dedicated dataset, aiming at providing a tool to design and to be integrated into the control system of the machine.

* Author to whom any correspondence should be addressed.

Keywords: gyrotron, magnetron, ITER, control system, modeling, simulation, heuristic

(Some figures may appear in colour only in the online journal)

1. Introduction

In the magnetic confinement fusion research field, additional plasma heating represents a crucial feature in performing long continuous pulses and maintaining the plasma stability for the fusion reaction, by heating plasma up to the temperature to have enough alpha particles to maintain a burning plasma. Besides the naturally occurring Ohmic heating, three main additional heating sources are generally implemented in tokamaks: the Neutral Beam Injection, providing energy to plasma by collision with neutral particles, the Ion Cyclotron Resonance Heating, exploiting the cyclotron resonance of the heavy ions, and the Electron Cyclotron Resonance Heating (ECRH), exploiting the cyclotron resonance of the plasma electrons [1]. The core of the latter lies in the employment of gyrotrons.

Gyrotrons are oscillators based on the resonance of weakly relativistic electrons to produce a coherent electromagnetic beam in the radiofrequency millimeter-wave domain [2, 3]. In the ITER baseline, each gyrotron is expected to produce 1 MW of radiofrequency (RF) power at 170 GHz on demand, and to release it in specific regions of the plasma. Foreseeing to operate up to one hour, the Plasma Control System [4, 5] expects to have a close-to-100% reliability on the ECRH system and gyrotrons available in the millisecond time scale for the different plasma operations. Due to its flexibility and localized absorption, the main use of the ECRH consists in start-up assist, heating, profile tailoring, impurity control, and magnetohydrodynamics control [6].

During gyrotron operation, the emitting component of the device's Magnetron Injection Gun (MIG) is heated up by a filament, which represents the input power provider of the system. To generate the electromagnetic beam, electrons are extracted from the emitter via thermionic emission and are accelerated via an applied potential towards the resonator of the gyrotron, called the cavity. Here, the particles are coupled with a magnetic field to produce the electromagnetic wave of the desired frequency. The mechanical design of the gyrotron and its operational parameters determine the desired electromagnetic radiation mode. Afterward, the RF beam is extracted from the device through a series of mirrors and injected into the transmission line that leads to the tokamak launchers. The remaining electrons are dumped in the so-called collector, where their energy is dissipated via heat exchange with the cooling water. A simplified schematic representation of the above-mentioned gyrotron components is shown in figure 1.

A gyrotron's operation can be divided in two phases:

- The **on-phase**, during which a high voltage (∼70 kV) is applied to the MIG and the heating filament provides power to the system. The electron current is emitted into the cavity, and the RF beam is extracted from the tube. In this phase, both thermal (MIG heating and cooling, collector heating) and electromagnetic (RF wave generation, arcing) phenomena occurs in the gyrotron.
- The **off-phase**, during which no voltage is applied to the MIG, therefore no electron current is emitted and the RF beam is not generated. The filament still provides power to the MIG. This phase is dominated by thermal phenomena, since no potential is applied to the system.

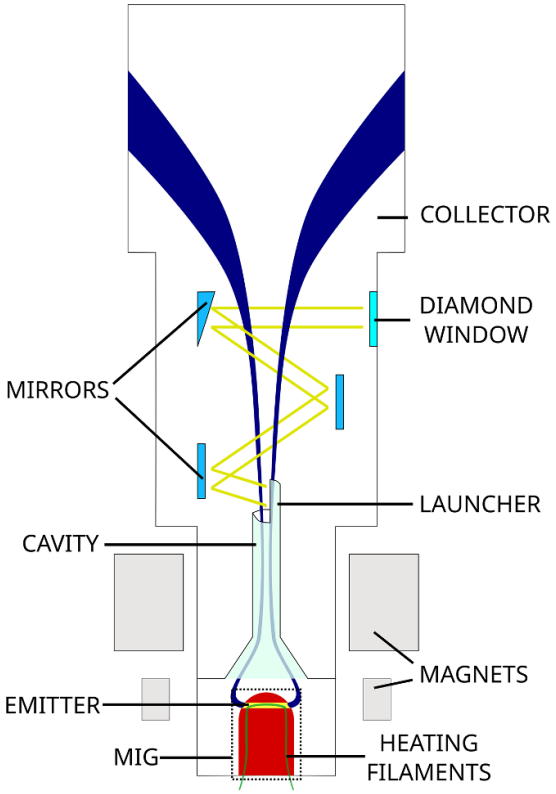


Figure 1. Schematic representation of a gyrotron.

Each phase has different dynamics and must be controlled accordingly. Moreover, by switching from one phase to another, the system experiences a transient, during which its state rapidly changes due to the different dynamics involved.

The on-phase is conventionally controlled via an experimentally tuned feedforward current waveform applied to the filament, meant to mitigate the transient effects at the beginning of the emission and compensate the additional cooling given by the electron beam current, called Nottingham Effect [7]. In recent developments, a beam current feedback with an integral controller has been successfully superimposed to the feedforward, to reach the electron beam target in long pulses [8].

No active control has been proposed for the off-phase so far, also due to the lack of reliable models describing the MIG behavior during this stage of operation. With the

common gyrotron experiments, the tube activation is scheduled in advance and the expected electron beam emission is tuned via microsecond pulses before the actual pulse, which lead to negligible additional cooling and provide an insight of the MIG state, that is then adjusted accordingly by modulating the input power. After the RF emission, the MIG experiences a transient where its temperature rises due to the lack of Nottingham Effect and to the delayed action of the on-phase control (e.g. figure 3). The filament input current is set again to the nominal value for this phase and the system is subject to a slow return (of the order of magnitude of about 10 min) to its steady state. For the following emission, the microsecond-pulse tuning is repeated.

In the ITER baseline, gyrotrons need to be available on demand at any time, therefore, a gyrotron might be called to action right after an RF emission during its transient time and might not be usable to inject power again into the plasma. A controller for the off-phase not based on microsecond pulses and able to drastically reduce the preparation time is of crucial importance. This paper contributes toward reliable and reproducible operations of tokamak's ECRH systems, by providing an enhanced version of the semi-physical MIG thermal lumped model originally proposed in Badodi *et al* [9]. Our contribution is manifold: the model has been further developed by redefining its parameters and their range of optimization and by tuning the model geometrical characteristics; the implementation of the code itself has been changed and the enhanced version is now in C++, resulting in a computationally more efficient and open source software, available to be integrated in external codes. The increased computational performances allowed us to employ a data-driven model parameters identification, based on the Particle Swarm Optimization algorithm [10, 11], on a much larger training data-set than the original implementation. The validation of the trained model on a dedicated set of experimental data has been also carried out.

All the data used for identification and validation has been provided by the FALCON facility at Ecole Polytechnique Fédérale de Lausanne. FALCON is a joint laboratory setup by Fusion For Energy and the Swiss Plasma Centre, where the Thales European gyrotron prototypes are tested [12].

The purpose of the proposed model is to provide a reliable tool that can adequately simulate different gyrotrons and adapt to each gyrotron during different stages of its lifetime, provided that sufficient data for model parameters tuning are available. Indeed, the tool enables model-based design and validation of controllers for both on- and off- phases, before the actual test on the real plant. In particular, for the on-phase, the model permits to carry out accurate study on the filament control effort and the controller's operational space. Moreover, the provided model can play an active role in the off-phase control since it provides a real-time estimation of the MIG temperature and expected electron beam emission, which can be used as feedback for the controller, overcoming the limitation of the microsecond-pulse tuning approach, which is not applicable for all the operation regimes and is in general slow.

The remaining of the paper is structured as follows: in section 2 the general MIG geometry, the physics behind the electron beam emission, the original thermal model, and a brief description of the FALCON test facility are presented. Section 3 describes in detail the developments of the thermal lumped model, while in section 4 introduces the adopted optimization method. Section 5 is devoted to both the model parameter identification and the model validation; in particular the obtained performance is shown. Finally, in section 6 the conclusions are drawn as well as possible further improvements.

Table 1. MIG components physical properties - Where ρ is the density, k is the thermal conductivity, c_p is the specific heat capacity at constant pressure, and ϵ is the emissivity.

Component	$\rho\,[\frac{g}{cm^3}]$	$k\,[\frac{W}{mK}]$	$c_p\,[\frac{J}{kgK}]$	ϵ
Filaments	19.30	180	134.4	0.50
Filament insulator coatings	3.975	36	756.0	0.14
Shields, Nosecone, Prolongator	10.22	130	251.0	0.40
Emitter	15.44	125	142.0	0.60

2. Background

This section gives an overview of the MIG geometry, the physical basis behind the electron beam emission, the model starting point, and the FALCON dataset exploited for model identification and validation.

2.1. Magnetron injection gun

The MIG is a complex multi-body subsystem of the gyrotron. To produce a 1 MW radiofrequency wave with an efficiency of 50% for ITER baseline, the MIG generates a 2 MW electron beam via temperature-limited thermionic emission from a ring-shaped emitter made of barium-calcium-aluminate dioxide nominally operating at temperatures of the order of 10^3 K.

The geometry of the ITER 1 MW European Gyrotron's MIG and the energy exchanges happening in the system are extensively described in Badodi *et al* [9], therefore only a summary of magnetron's components is provided here to understand the model implementation that will follow:

- heating filaments covered by an insulator coating wrapped inside the cathode structure, representing the input power providers of the system.
- 2 thermal shielding structures, meant to focus the power emitted by the filament towards the emitter.
- 1 emitter made of barium-calcium-aluminate oxide-impregnated tungsten, the core component from which the electrons are emitted.
- 1 nosecone and 1 prolongator, belonging to the structure supporting the system.

Materials are not provided here due to confidentiality reasons. However, in table 1, the described components physical properties can be found.

2.2. Thermionic emission, Nottingham effect, and delayed input power effect

The thermionic emission in the temperature-limited approximation is described by the so-called Richardson-Dushman formula

$$j(T,E) = \eta \frac{4\pi e k_B^2}{h^3} T^2 \exp\left(-\frac{\phi(T,E)}{k_B T}\right) \left[\frac{A}{m^2}\right],$$

where T and ϕ are respectively the surface temperature and working function of the emitting body [13]. The scaling factor η is an experimental correction to the current density often introduced in order to take into account possible uncertainties in the working function, or when it is assumed to be independent of temperature. Many values can be found in literature, ranging from 0.25 to ∼1 (see [14, 15]).

The working function $\phi(T,E)$ depends on both the temperature of the body and the applied electric field, i.e.

$$\phi(T,E) = \phi_0 + f(T) - \Delta\phi(E), \tag{1}$$

where the term $\Delta\phi$ represents the field-dependent Schottky lowering barrier effect in vacuum, such as the effective decrease of the average energy needed to extract an electron from the material, that can be expressed as

$$\Delta\phi(E) = \sqrt{\frac{e^3 E}{4\pi\epsilon_0}}. \tag{2}$$

The term $f(T)$ in (1) depends on the Fermi energy, the thermal expansion coefficient, the dimensionality, and different temperature powers of the emitting material [16, 17]. With temperatures of the order of 10^3 K, the higher order corrections in temperature are negligible and the dependence can be considered linear [18].

According to Nottingham [19], each electron leaving the emitter surface carries an average energy given by

$$q_e = 2k_B T + \phi, \tag{3}$$

implying that the thermionic emission must be considered in the energy balance of the body.

The heat lost by the emitting body during the electron beam emission can be easily calculated by multiplying (3) by the total current density leaving the body and by its surface area, which gives the additional negative term in the right-hand side of

$$\frac{\mathrm{d}E_{\mathrm{on}}}{\mathrm{d}t} - \frac{\mathrm{d}E_{\mathrm{off}}}{\mathrm{d}t} = -q_e S j + \Delta P_{\mathrm{in}}, \tag{4}$$

where ΔP_{in} is the variation of external input power and no differences in power leaks between phases are assumed for the sake of simplicity. The power balance difference between the two operational stages of the gyrotron highlights the additional dynamics brought by the electron emission. The term $q_e S j$ in (4) represents the Nottingham Effect, which, with a negative sign, leads to an additional cooling of the system. If the additional input power does not balance this term, the net effect on the system is an electron beam current-dependent cooling.

During the on-phase, the Nottingham Effect adds up to the power exchanges the MIG is subject to. This behavior can be clearly seen in the operations when the heating filament input current provided to the system does not change between phases, as reported in figure 2.

After the end of the emission, the additional cooling term is removed from the balance equation and the system is subject to a transient phase in which the MIG temperatures vary rapidly. If the system had additional input power to mitigate the current drop, right after the end of the pulse, even with the filament input current back to the initial off-phase value, a remaining delayed energy transfer will continue to heat-up the emitting surface. This effect can be easily understood by analysing multiple long pulses in row, as the ones reported in figure 3.

2.3. Model overview

The MIG semi-physical thermal lumped model initially developed by Badodi *et al* [9] simulates the thermal dynamics of the bodies explained in section 2.1. Here, a generic description of the original implementation is provided.

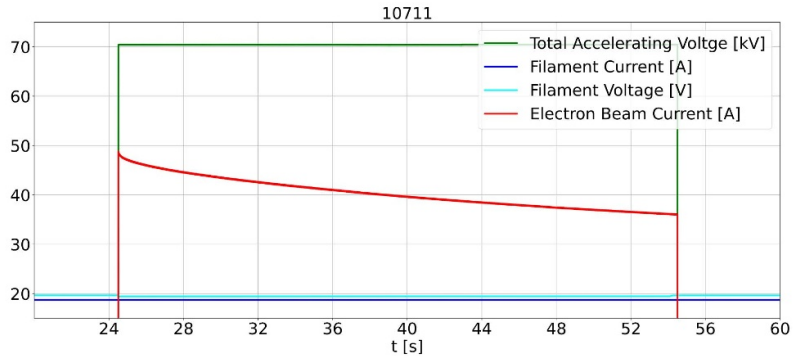
The emitter is irradiated by the filaments on one side and, during RF production, is cooled on the opposite side by the Nottingham Effect. To adequately model the temperature gradient persisting in the object during operation while maintaining the lumped approach (each body with a homogeneous temperature), the body is divided into multiple parts, implying the introduction of a number of parameters to be optimized (sizes of the sub-components and thermal resistances between them).

One more hyper-parameters is related to the emitter: the **correction to the beam heat** (**$\mathbf{c_2 b}$**). In order to match the Nottingham Effect in the simulations and tune the electron beam current decrease, the **$\mathbf{c_2 b}$** parameter is applied to the thermionic power loss. Note that this factor is unrelated to the correction to the current density explained in section 2.2, which is assumed unitary in this case.

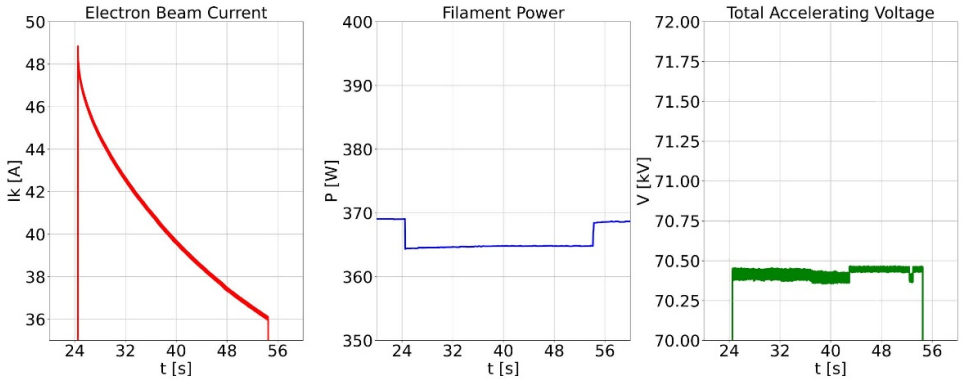
All the emitter-related parameters contribute to the power balance of the body and thus to the thermionic cooling, therefore they have a high mutual dependence and cannot be studied individually.

The last two hyper-parameters are the energy sink of the system, such as the bodies with approximately constant temperatures during operations. In the original implementation, these components were identified in the nosecone and prolongator, therefore their temperatures were kept constant.

As explained in section 2.2, the thermionic emission is dependent on the electric field applied to the emitting body with a net effect of lowering the working function of the material and increasing the emitted current. In FALCON experiments, the accelerating voltage applied to the emitter is



(a) Electron beam current (red time trace), total accelerating voltage (green time trace), filament current (blue time trace) and filament voltage (cyan time trace)



(b) Electron beam current (red time trace), filament power (blue time trace), total accelerating voltage (green time trace)

Figure 2. FALCON experimental pulse #10711 - The electron beam current has a first peak of ∼49 A. During the pulse, no additional power is provided to the system with respect to the off-phase (the difference in the filament power trend in is due to the 70 kV total accelerating voltage influence on the filament voltage measurement). This leads to a time-varying cooling of the system that is reflected on the measured current.

measured and the model needs to directly make use of this proxy for the simulations, since the conversion to the electric field applied to the emitter surface is not trivial. To do so, a voltage-electric field conversion factor K_{EV} was evaluated via Finite Element Method (FEM) simulations computing the Poisson equation on the surface of the emitter. As a consequence, the Schottky lowering barrier effect implemented in the thermal model is taken equal to

$$\Delta\phi\,(K_{\mathrm{EV}}, V) = \sqrt{\frac{eK_{\mathrm{VE}}V}{4\pi\,\epsilon_0}},$$

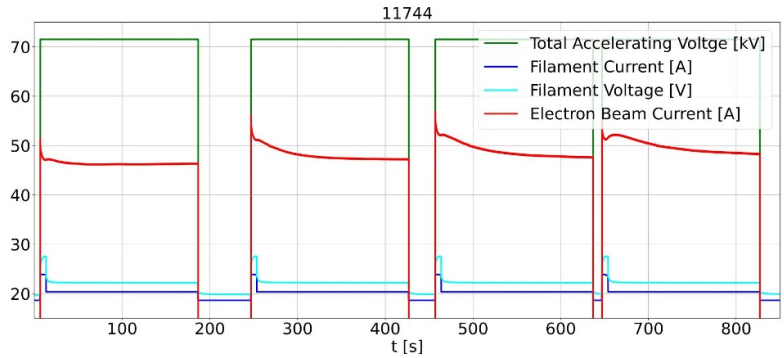
i.e. it provides a formula directly dependent on the experimental accelerating voltage.

In Introini *et al* [20], a sensitivity analysis (SA) on the parameter set of a chosen model implementation is shown, describing the possible influences of the parameters on the electron beam simulation. Even if they cannot be considered uncorrelated to each other and the importance of each one of them depends on the specific implementation employed, the SA provides us with a general understanding of the role of the parameters in the model: the boundaries will affect the steady state the most, as they determine the overall temperatures of the system at the equilibrium, while the parameters related to the emitter will mainly influence the simulation during the emission, since they rule both the gradient in the emitter and the emitter surface cooling, both crucially influencing the dynamics of the electron beam current.

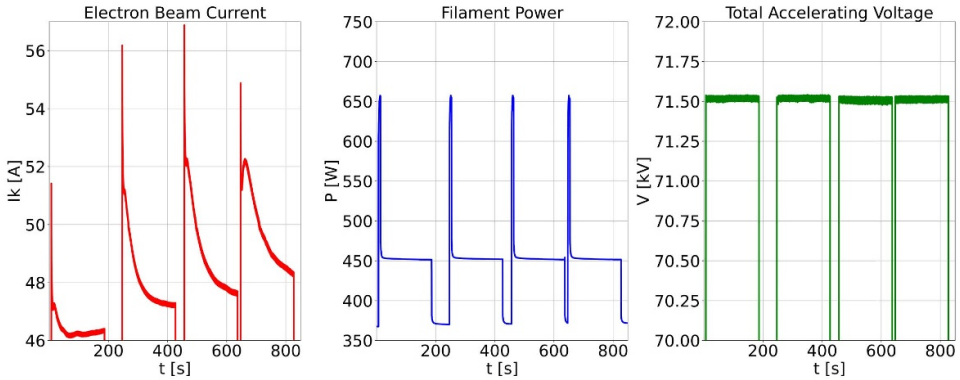
2.4. FALCON data

The FALCON test facility is located at the Swiss Plasma Centre in the Ecole Polytechnique Fédérale de Lausanne in Switzerland and hosts the Fusion for Energy gyrotron programme. Here, the 1 MW-class European Thales gyrotrons operating at 170 GHz are conditioned and tested. The facility allows long pulses experiments, with the longest registered so far of 1000 s, with a demonstrated power of 950 kW [21].

During operations, more than 300 signals collected from the plant and from ancillary systems (temperatures, flows, voltages, interlocks, etc) are acquired at 20 Hz, 1 kHz,



(a) Electron beam current (red time trace), total accelerating voltage (green time trace), filament current (blue time trace) and filament voltage (cyan time trace)



(b) Electron beam current (red time trace), filament power (blue time trace), total accelerating voltage (green time trace)

Figure 3. FALCON experimental pulse #11744 - Four pulses are emitted in row, with 60 s, 30 s, and 10 s intervals. As it can be seen from the filament data, the system input power is brought to the nominal initial off-phase value between emissions, but it does not prevent the system to heat up and have a first peak for the following pulse higher than the expected one as for the first pulse, which was performed at the steady state of the machine.

or 2 MHz, depending on the expected timescale of the measured variable, and are stored in a dataset for each experiment. Moreover, data of the states (mainly temperatures) of and around the gyrotron are continuously acquired each day, independently on the operation.

The facility allows us to access the experimental data needed to train, validate, and test the proposed MIG model. For the electron beam emission, we are mainly interested in 3 physical quantities, provided by the combination of 5 signals:

- the total accelerating voltage the MIG is subject to, given by the sum of two voltages applied into the system;
- the system input power, given by the product of the heating filament current and voltage;
- the emitted cathode current.

The model uses the first four signals as inputs, and the reconstruction of the cathode current is employed as a benchmark of its performances. A possible correlation between the beam reconstruction accuracy and the rest of the dataset is explored in section 5.

3. Model developments

The MIG semi-physical thermal lumped model firstly introduced in [9] has been further developed, optimized, and validated against experimental data. In what follows, the improvements to the model introduced in the previous section are described in detail.

3.1. Parameters

The emitter has been divided into three parts, therefore introducing 4 hyper-parameters:

- The **Bulk**, heated up by the filament.
- The **Surface**, cooled down by the Nottingham Effect during operation.
- The **Intermediate**, meant to emulate the temperature gradient in the lumped system.
- The **Bulk-intermediate contact thermal resistance**.
- The **Intermediate-surface contact thermal resistance**.

Table 2. Model parameters - Note that the three Emitter components are linked by the total size of the body, therefore just two of them are optimized independently.

Parameter	Meaning
Bulk	Emitter Bulk size
Intermediate	Emitter Intermediate size
Surface	Emitter Surface size
R_{BI}	Bulk-Intermediate thermal contact resistance
R_{IS}	Intermediate-Surface thermal contact resistance
Gyrotron Internal Temperature	Gyrotron internal temperature facing the Surface
c_2b	Correction to the Nottingham Effect
Environment Temperature	Effective temperature in contact with the MIG
R_{NE}	Nosecone-Environment thermal resistance
R_{PE}	Prolongator-Environment thermal resistance

The total mass of the body remains consistent with the real emitter, therefore just two out of the three emitter's components are optimized: the third is evaluated as consequence.

An additional interaction has been considered for the emitter: the grey body emission towards the inside of the gyrotron, consisting mainly in the anode from the point of view of the **Surface**. The inside of the gyrotron will reflect a certain amount of the incident energy and it will have a grey body emission itself due to its (unknown and unmeasured) **temperature**. This forces us to introduce the latter as a parameter in the energy exchange computation. Considering also the $\mathbf{c_2 b}$, this makes a total of 6 parameters concerning the emitter.

The nosecone and the prolongator have been implemented with evolving temperatures, therefore, to simulate the energy sink of the system, three parameters are introduced:

- The **Environment Temperature**, an effective constant temperature for all the components interacting with the MIG, such as the cooling oil system, etc.
- The **Nosecone-Environment contact thermal resistance**.
- The **Prolongator-Environment contact thermal resistance**.

With the boundary conditions, the model has 9 parameters to be optimized together, which are reported in table 2. In figure 4, a schematic of the MIG transverse section with the emitter divided in three parts can be seen.

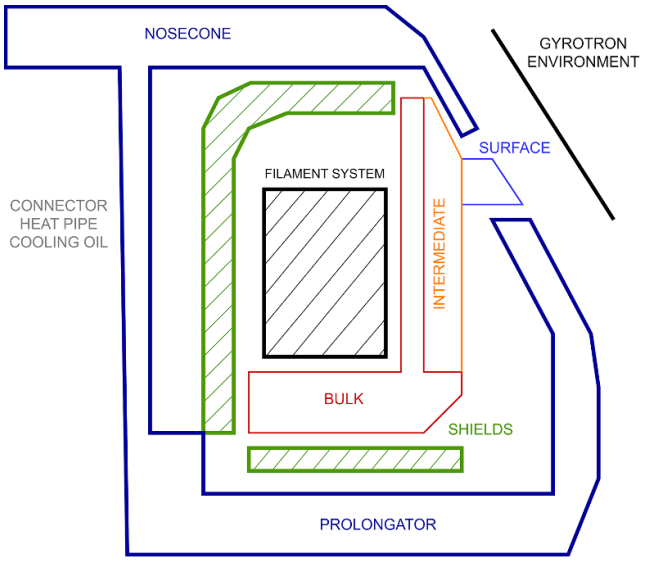


Figure 4. Schematic of MIG transverse section - The nosecone and the prolongator belong to the same component, such as the structure of the MIG. The emitter is divided in three parts (*Bulk*, *Intermediate*, and *Surface*) as per our latest model implementation. Their sizes are given as an example: they are optimized and change depending on the optimization.

3.2. Schottky barrier factor

According to the field evaluated in the paper, the conversion factor K_{EV} is set to 61.9 $\mathrm{C}^2\ \mathrm{m}^{-1}$. This provided us a starting value to tune with a data-driven method. Our approach relied on the coupling between the temperature and the accelerating voltage in the Richardson-Dushman equation. The latter can be inverted using the Lambert W function [22], to indirectly determine the emitter surface temperature trend during the electron beam emission via experimental data.

A common gyrotron operation method consists in applying an accelerating voltage step during the emission to avoid a too high initial peak and maintain the correct mode generated in the cavity. By reconstructing the emitter surface temperature during these pulses, if the K_{EV} conversion factor is correct, the temperature must remain continuous with a decreased, or negatively increased, slope due to the augmented cooling effect. The voltage step time scale is orders of magnitude faster than the typical thermal time scale of energy exchanges, therefore we can assume this step to be instantaneous for the thermal system.

By employing the value from Badodi *et al* [9], a discontinuity is found in the reconstructions. This led us to iteratively optimize this parameter to cancel out the discontinuity, seen as a spike in the discrete derivative among the background characteristic noise of the data, providing us with a value of 82.0 $\mathrm{C}^2\ \mathrm{m}^{-1}$, $\sim$30% higher than the initial one, and an average emitter temperature $\sim$10 K lower than before (see figure 5).

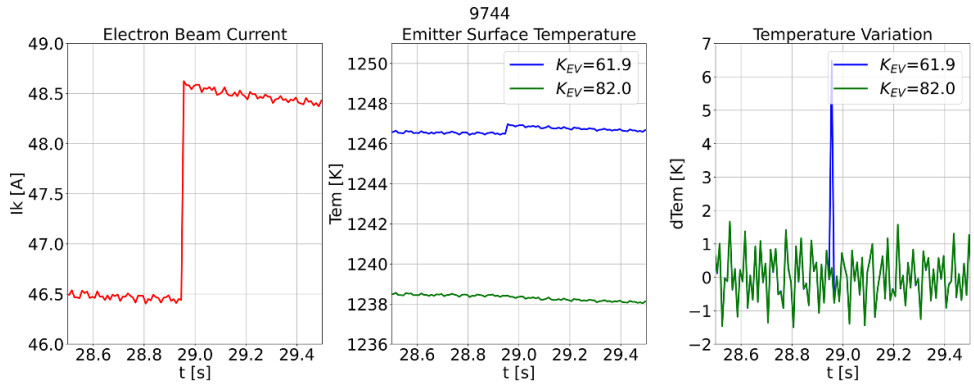


Figure 5. FALCON experimental pulse #9744 - Reconstruction of the emitter surface temperature via Lambert W function with the original and new value of the K_{EV} conversion factor [C^2 m^{-1}].

Table 3. Inner cavity view factors normalization before and after forcing symmetry.

Before	0.999	0.999	0.999	0.998	0.999	0.999	0.999	0.998	1.000	1.001
After	0.998	0.999	0.999	0.998	0.998	0.999	0.998	0.999	1.000	1.002

Table 4. Outer cavity view factor normalization before and after forcing symmetry.

Before	1.001	0.999	1.000	1.000	1.001	1.000	1.000	0.998	1.001	1.000	1.000
After	1.003	1.001	0.999	1.005	0.994	1.000	0.999	0.998	1.000	1.001	1.054

3.3. View factors

The system has two main cavities in which multiple bodies exchange energy via radiation: the *inner cavity*, such as the volume containing the filaments, enclosed by the emitter and the first upper shield, and the *outer cavity*, such as the remaining volume. The core of the exchanges in these regions lies in the use of the view factors F_{ij}, which were originally calculated via FEM tool. To maintain the conservation of energy in the system, the view factors must satisfy the following reciprocity relation

$$A_i F_{ij} = A_i F_{ji}\,,$$

as well as the normalization condition

$$\sum_{j=1}^{N} F_{ij} = 1\,,$$

where indexes i and j represent the ith and jth body in the cavity, A_i and A_j the corresponding emitting areas, while F_{ij} is the view factor.

While the normalization was fulfilled with an accuracy up to 10^{-3}, the view factors did not respect the reciprocity relation, with errors up to 20%. To solve this issue, symmetry was forced in geometrical resistance matrix $R_{ij} = A_i F_{ji}$ by evaluating the average value with its transpose, i.e.

$$\widehat{R_{ij}} = \frac{R_{ij} + R_{ji}}{2}\,.$$

The new view factors $\widehat{F_{ij}}$ maintained the same level of accuracy for the normalization constraint, with exception of the eleventh set of factors in the outer cavity, related to the portion of the emitter in the upper part of the cavity exchanging energy with the upper shields and the nosecone (see tables 3 and 4).

3.4. Model framework

The model was first implemented in the MATLAB/Simulink environment. For computational time purposes and the final aim to implement the model as a core component of a real-time control system, the code has been ported to a C++ environment.

The project provides general classes for thermal simulations (bodies and exchange types), with the possibility of initializing the system under investigation via configuration file. Thanks to this, multiple system versions were easily tested (emitter divided in several parts, different exchanges between bodies, energy sinks, etc) and the best one presented above was chosen.

The computational time has been improved by a factor of about 2 orders of magnitude: a ∼500 s experiment that was reconstructed in ∼100 s with the original framework with an Intel(R) Xeon(R) W-2145 CPU (3.70 GHz), including the MIG steady state evaluation, is now simulated in ∼3 s with an Intel(R) Xeon(R) Gold 6230 CPU (2.10 GHz). This allowed us to triplicate the optimization dataset (from ∼7 to ∼20 pulses) and increase the training cycles. More detail about the optimization algorithm are given in the next section.

In figure 6, a schematic that describes the workflow of the MIG thermal lumped model is reported. Four types of energy exchanges are implemented in the current version:

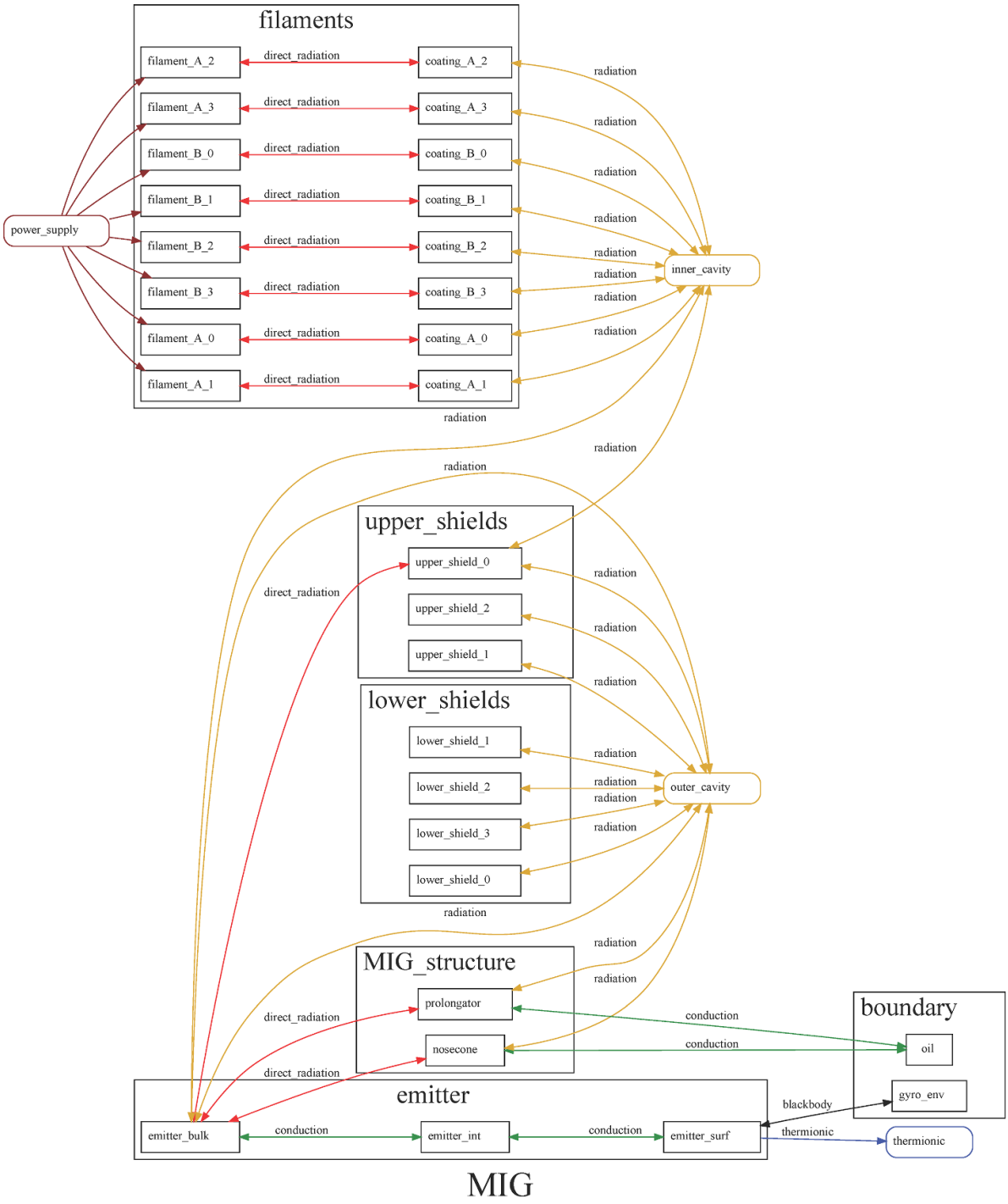


Figure 6. Model implementation schematic - The power supply provides the input energy to the filaments. Each one is folded four times, therefore 8 emitting bodies are considered. The coatings exchange energy with the emitter and other bodies via irradiation in the so-called *inner cavity*. The *emitter bulk* is considered as the only part interacting with the rest of the system in the cavities and via direct radiation with the nosecone and the prolongator. The intermediate and the surface components are isolated and exchange energy via contact conduction, with exception of the thermionic emission and the grey body emission towards the inside of the gyrotron.

- **Thermionic Emission**, leading to the Nottingham Effect by means of the Richardson-Dushmann equation.
- **Radiation**, such as a multi-body radiation energy exchange implemented for the two cavities of the system.
- **Direct radiation**, such as the radiation energy exchange between two bodies, where at least one of them has a unitary view factor. This is used between the filament and the coating and for the grey body emission of the emitter surface.
- **Conduction**, such as the energy exchange by direct contact. This is used for the emitter subdivision and for the exchange between the nosecone and prolongator and the environment.

4. Optimization

The optimization of the model parameters is carried out with a Particle Swarm Optimization (PSO) heuristic [10, 11]. The PSO is a type of meta-heuristic inspired by nature, in particular by the combination of individual and social behaviours of animals in herds, flocks, and, indeed, swarms. The attractiveness of PSO lies in the use of few user-defined hyper-parameters and the gradient independence in the parameter hyperspace, thus being suitable for non-differentiable problems.

Each particle in the swarm is characterized by a position in the search space, whose fitness, such as a measure of its goodness, is evaluated via an objective function, and a velocity (direction in the space to compute the position at the next iteration). The core of the optimization is the update of the latter according to the following linear combination

$$\begin{aligned}\mathbf{v}(t+1) = \omega\mathbf{v}(t) \; + \\ r_1\,[0,1]\,c_1\,(\mathbf{p}(t) - \mathbf{x}(t)) \; + \\ r_2\,[0,1]\,c_2\,(\mathbf{g}(t) - \mathbf{x}(t)) \; ,\end{aligned}$$

that includes an inertial term, such as the last evaluated velocity, an individual term, dependent on the distance between the particle's best position and the current position, and a global (social) term, dependent on the distance between the swarm best position and the current position. Each term is multiplied by a user-defined coefficient, which might evolve during the optimization. In this work, the Clerc and Kennedy's constriction coefficients are applied [23].

To increase randomness and explore as much as possible all the space, two random factors, whose value is taken in the interval $[0, 1]$ according to a uniform distribution for each particle at each iteration, are multiplied to the personal and global term respectively. The position for the next iteration is then evaluated as $\mathbf{x}(t+1) = \mathbf{x}(t) + \mathbf{v}(t+1)$.

At the initialization, the particles are randomly generated into the optimization space, which will be their personal bests for the next iteration, and a random velocity is assigned to each of them, so that the inertial term will not be trivial in the velocity update.

To deal with the optimization boundary conditions, an inelastic collision is applied. When the velocity leads the particle outside the optimization space, the position update explained before is not applied. Instead, a random position between the current one and the boundary is generated. The velocity at the current interation is substituted by the distance between the new position and the boundary to completely remove the component that is leading the particle outside the boundaries. In this way, the particle does not leave the optimization region and at the next iteration the velocity inertial term will push it back inside the optimization hyperspace (see figure 7).

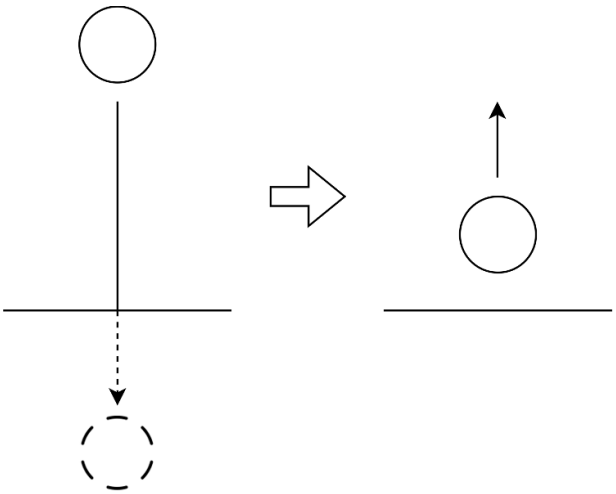


Figure 7. PSO boundary inelastic scattering.

The goodness of each particle position is evaluated by means of an objective function and the aim of the swarm is to find the position that minimizes it. Many different functions were tested, giving more relevance to the reconstruction of a specific part of the electron beam (first peak, end of the pulse, initial transient). At the end, the root-mean-square error over all the electron beam current during emission was chosen as the best metric for the optimization problem. At the beginning of each iteration, the objective function is evaluated for each particle, updating the personal bests and the global best of the swarm. Afterwards, the velocity is updated, and new positions are computed.

5. Model identification and validation

The model optimization via PSO has been carried out on a dedicated dataset. The results are analysed via specific metrics meant to describe the model performances in the reconstruction of the electron beam current during the on-phase and the correct system evolution during the off-phase.

5.1. Dataset

The final goal of the model is to provide a tool that adequately reproduces the electron beam emission and predicts the state of the emitter without any direct or indirect measurement with exception of the filament power during the off-phase. Therefore, the dataset must contain experiments long enough to appreciate the Nottingham Effect during the emission, and there must be sequential pulses in order to evaluate the correct evolution of system temperatures when there is no emission,

such as when no direct measurement on the MIG is available. To do so, 383 FALCON experiments were selected considering the following constraints:

- no mode changes are observed in the electromagnetic wave generated in the cavity by interaction with the emitted electron beam, with exception of short, fastly-recovered, mode jumps;
- in case of a single pulse in the experiment, the duration must be of at least 5 seconds, so that the effect of the additional cooling is appreciable not only on the emitter, but also on all the other bodies of the system;
- in the case of multiple pulses in the experiment, no constraints on the length of the pulses are imposed;
- each experiment must have been carried out at the steady state of the machine. Five minutes were chosen as the threshold to consider the pulse independent from the previous one.

For the optimization dataset, 21 pulses were selected, all containing multiple pulses in row separated by different time-lapses (similar to the ones shown in figure 3). With this type of dataset, an adequate off-phase simulation is assured by the reconstruction of the first peak occurring during the thermal transient time after a pulse.

In order to correctly evaluate the performance of the model, it is necessary to assess the intrinsic uncertainty of the datasets to determine the best reconstruction results we can obtain. The MIG is assumed not to have energy leaks, meaning that, at the thermal steady state of the system during the off-phase, all the input power provided by the filament is dissipated via the sinks of the system. This means that a 1-1 relation between the emitter temperature and the filament input power must exist, and it can be investigated by analysing the emission first peaks in the dataset.

In figure 8(*a*), the first peak at the steady state of the system with respect to the filament power is shown. In order to assess the correlation between the emitter state and the filament power, the Lambert W function was employed to retrieve the temperature from the current peak and the voltage (see figure 8(*b*)) as in section 3.2. A general linear correlation with a noticeable dispersion around the average trend can be seen: the reconstruction with a linear interpolation can be set as the comparison term for the model performance and can be considered the approximate maximum achievable by the system for the first peak of the electron beam emission at the steady state. The linear interpolation is done on the reconstructed emitter temperature. The results, which are useful for comparison with the model performances, are evaluated via relative error (RE) with the expected values and are reported in figure 9 and table 5.

A correlation between the reconstruction error and all the other signals acquired in the moment of the emission can be evaluated to determine possible dependencies of the thermionic emission on other variables of the system. Since no assumption on the type of correlation (linear, quadratic, etc) is possible, the mutual information (MI) index was employed [24]. This index determines how much information two distributions have in common, allowing us to select among all the signals the ones with the highest correlation with the error and then carry out a direct inspection to determine if one or more of them can be included into the simulation inputs. In figure 10, the 9 signals with the highest normalized MI are shown. It is easy to notice that no clear additional information is provided by any of the signals with direct inspection and the high MI index can be addressed to clustering. Hence, the first peak dispersion around the linear trend cannot be currently explained with the available data.

5.2. Results and analysis

As an analytical method that computes the PSO hyperparameters (swarm size and number of iterations) does not exists, a trial-and-error approach between experimental data reconstruction accuracy and computational time was followed: with a dataset of 21 pulses of average duration of about 500 s, a swarm of 1000 particle was chosen, and a total of 60 iterations were executed.

In what follows, the best obtained optimization is presented. The trained model is validated against the remaining 361 pulses from the original dataset (see table 6). The metrics to evaluate the model performances are the root-mean-square relative error (RMSRE)

$$\mathrm{RMSRE}\left(x,\hat{x}\right)=\sqrt{\frac{\sum_{i=1}^{N}\left(\frac{x-\hat{x}}{\hat{x}}\right)^{2}}{N}},$$

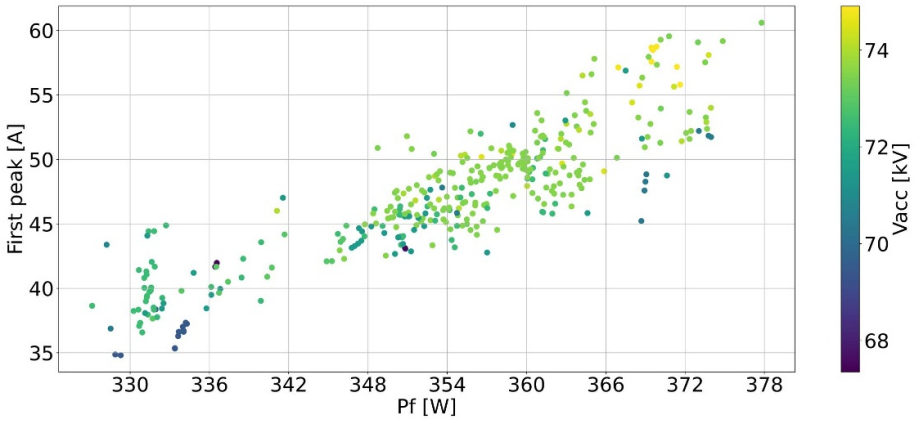
applied to all the pulse, the beginning (first second), and the end (last second), and the relative error on the emission first peak, as introduced in the previous paragraph.

The trained model performances during the on phase can be studied analysing the RMSRE at the three different thresholds. 98% of the pulses are reconstructed with an error below 10% all over the pulse, which corresponds to an absolute value of about 4.5 A given the nominal expected current of 45 A. In figure 11, a reconstruction example of an experiment in the validation dataset can be seen.

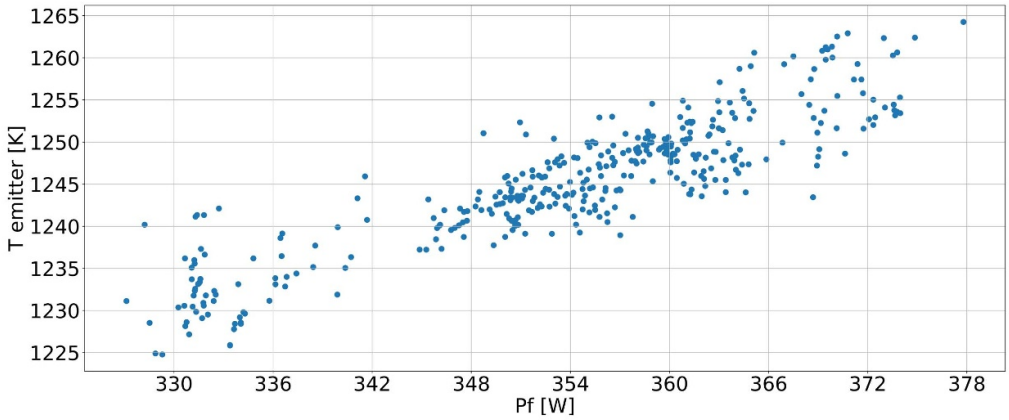
The off-phase performance can be assessed with two metrics: the goodness of the steady state evaluation, which is reflected in the first peak of emission, and the match of the evolution of the system during the transient at the beginning of the phase, much more difficult to study due to the lack of experimental data during this stage.

The former is analysed by comparing the reconstruction of the first peak with the linear interpolation method showed before: the trained model matches the expected maximum performance in term of percentage of pulses inside the three error thresholds, as it can be seen by comparing the data reported in table 5 with those of table 6.

A possible method to study the off-phase transient is to analyse multiple pulses in row. The first peaks of the pulses after a prolongated emission reflect the evolution during the transient. Multiple long pulses provide just one peak for each off phase, which is still dependent on the goodness of the on-phase simulation of the previous emission, therefore short (in the order



(a) Electron beam first peak vs filament input power, with color-bar given by the applied total accelerating voltage



(b) Emitter surface temperature, evaluated by means of the Lambert W function, vs filament input power

Figure 8. Linear correlation between the emitter state at the first peak emission and the filament input power at the steady state of the machine.

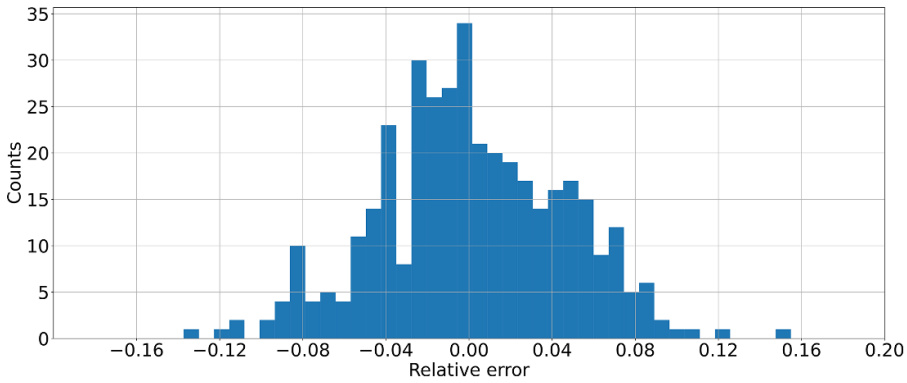


Figure 9. Linear fit reconstruction relative error histogram.

Table 5. Linear fit first peak reconstruction performances.

RE	Pulses (out of 383)	%
<10%	376	98
<5%	281	73
<1%	75	20

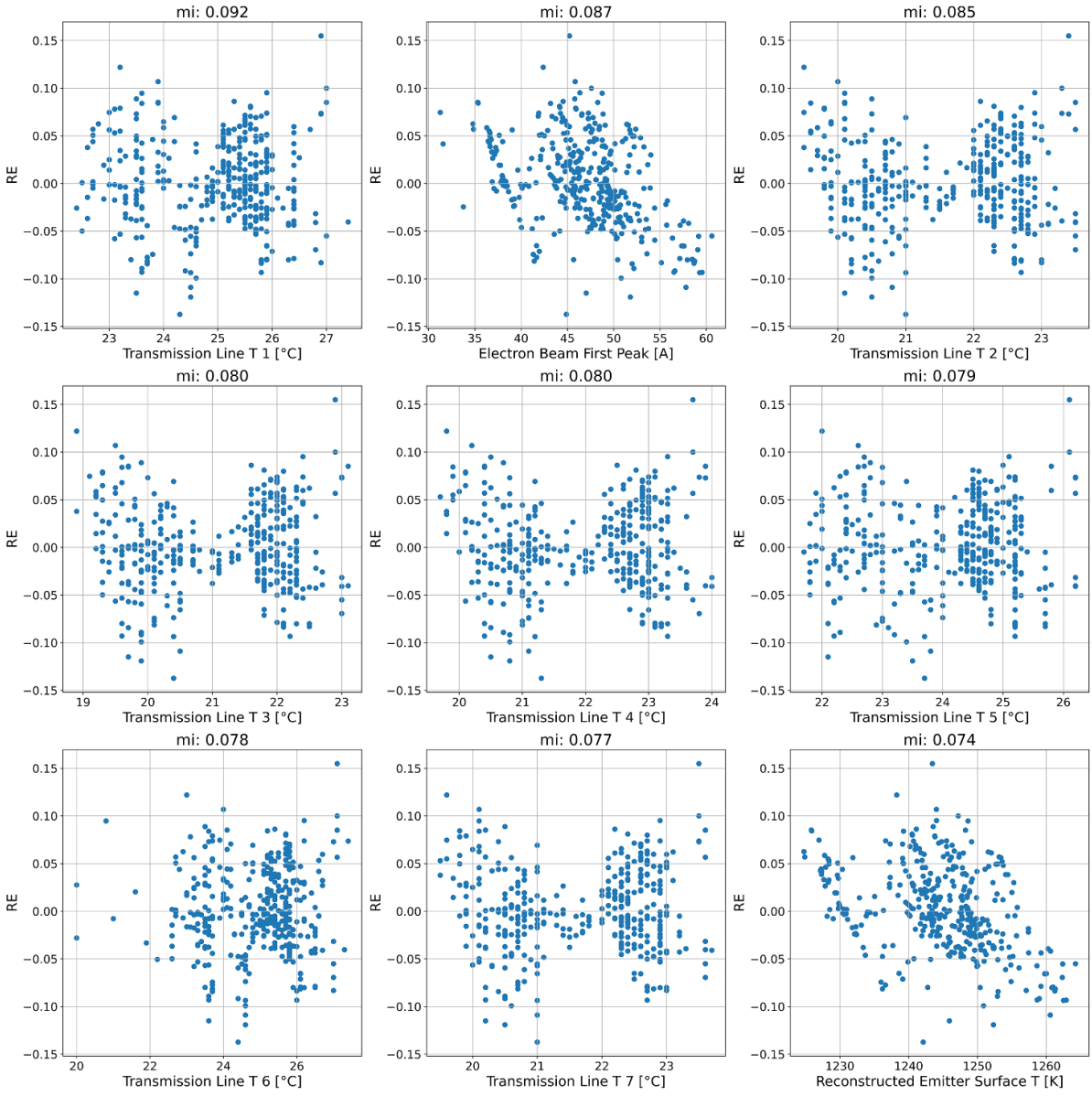


Figure 10. MI between the emitter temperature relative error at the first peak emission - The mutual information evaluated between the RE at the first peak and 273 other signals acquired in real time, such as temperatures, vacuum pressures, and cooling system flows. The MI has been normalized such that the MI of the RE with itself is unitary.

Table 6. Trained model performances on the validation dataset.

RMSRE	All pulse	Start (1 s)	End (1 s)	RE	First peak
<10%	353 (98%)	337 (93%)	352 (98%)	<10%	352 (98%)
<5%	307 (85%)	249 (69%)	265 (73%)	<5%	265 (73%)
<1%	41 (11%)	23 (6.4%)	18 (5.0%)	<1%	64 (18%)

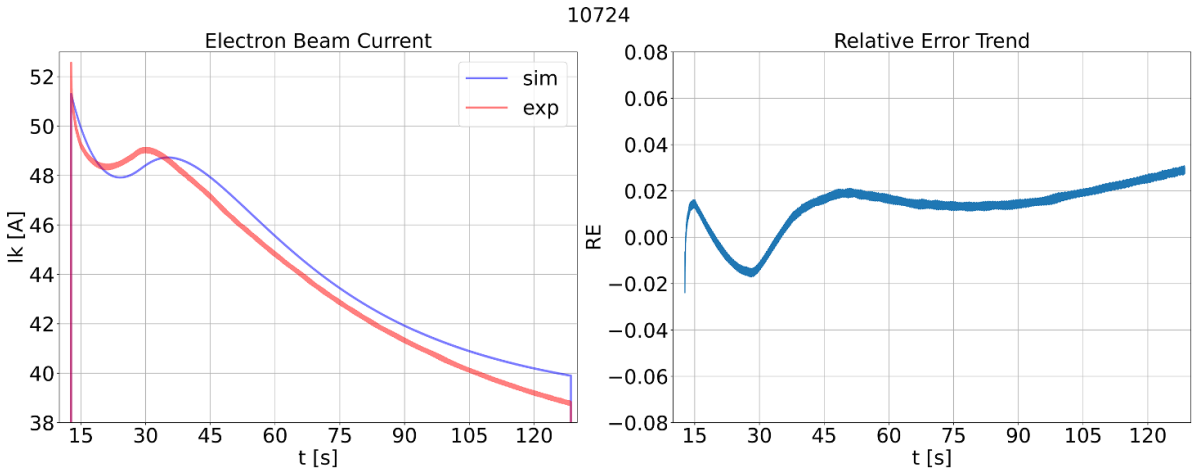


Figure 11. Simulation of FALCON experimental pulse #10724 - The reconstructed current shows similar dynamics during the emission. The first peak has an error of ∼1 A and the following current drop is slower for the model, implying an underestimation of the cooling. Both the experiment and the simulation show a local current minimum followed by a local maximum in the first 30 s of emission, with a 5 s delay for the model. The end of the pulse shows an increasing error.

of the *ms*) pulses generated after a long (>5 s) emission represent the best option. The short pulses lead to a negligible Nottingham Effect and reflect the instantaneous temperature of the emitter. In figure 12, three experiments performed with this approach are shown. As anticipated, the off-phase evolution is highly dependent on the previous on-phase.

For the experiment #11739, the long emission is underestimated throughout the pulse and the following short pulses are influenced consequently with an error of about 5%, decreasing with time. The remaining two pulses have a better reconstruction of the extended emission, and the following short pulses have an error between 2% and 4%.

6. Conclusions and further developments

Adequately controlling gyrotrons to ensure reliable on-demand RF power constitutes a crucial feature of ITER's rebaseline. Due to the complexity of the MIG and the dynamics involving its functioning, a model able to estimate the temperatures of this subsystem and predict the emitted electron beam with the highest accuracy possible represents a powerful tool to fullfil the required availability of the ECH plant.

In this work, the semi-physical thermal lumped model originally proposed in [9] has been further developed, trained on experimental data and validated against a dedicated dataset, providing a step forward towards the synthesis of a control system for the RF sources.

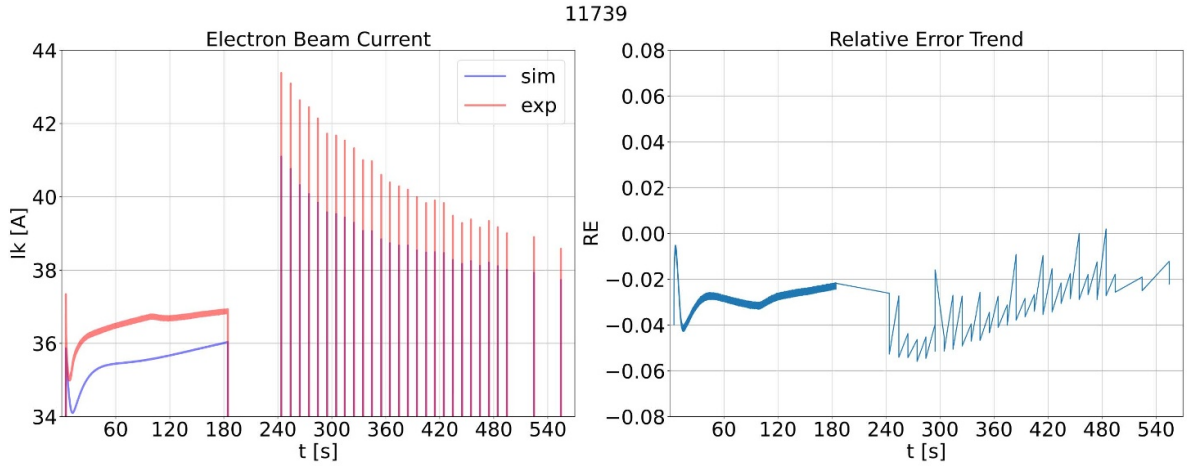
The MI index shown in section 5 suggests that no additional input can be provided to the model to increase the simulation performances besides the data already used: no remaining gyrotron signal is found correlated to the dispersion around the linear trend between the emitter temperature at the first peak, therefore no type of dependence can be assessed and implemented.

For the steady state first peak reconstruction, the trained model provides similar performances as the reconstruction with the linear interpolation, as showed by the relative errors with the three threshold levels in tables 5 and 6. The model has reached the expected best achievable performance and until no other physical quantities are found correlated to the first peak emission, no substantial improvements are possible.

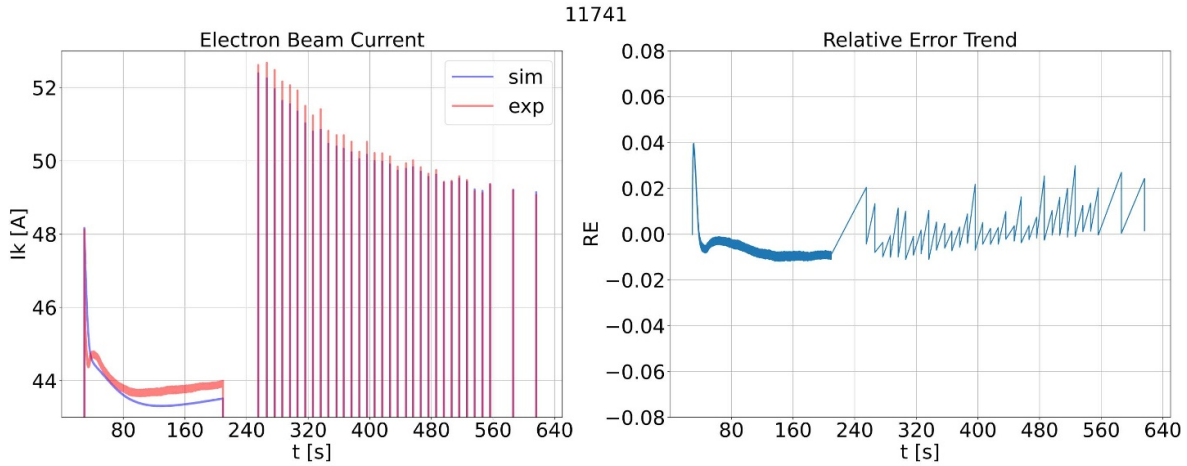
For what concerns the reproduction of an extended electron beam current, the trained model reconstructs more than 90% of the validation dataset with a RMSRE <10% and more than 70% of the dataset with a RMSRE <5% for all the three metrics. In particular, if we consider the RMSRE all over the emission, almost all the dataset has an error below 10% and 85% of it is under 5%. This provides us with a high level confidence in designing an on-phase controller on the model and safely test it on the real machine, where the final tuning can be executed.

The multiple microsecond shots after the extended emissions confirms that, by training the model with multiple pulses in row, also the off-phase evolution of the system is matched with an adequate level of accuracy by the simulations. This is possible thanks to the semi-physical, non-completely data-driven structure of the model, that allows us to trust the evolution of the system temperatures even without direct experimental data. Figure 12(*a*) highlights the criticality of the goodness of an on-phase simulation for the following off-phase transient: the end of the pulse is mismatched by ∼1 A (∼2%), leading to a following underestimated emitter temperature reflected in a series of peaks with an initial error of ∼2 A decreasing with time. If this situation occurs during operation, the off-phase controller might not mitigate enough the temperature rise and, in case of emission, the possibility of generating an incorrect mode exists. The same can happen with an overestimation of the on-phase current.

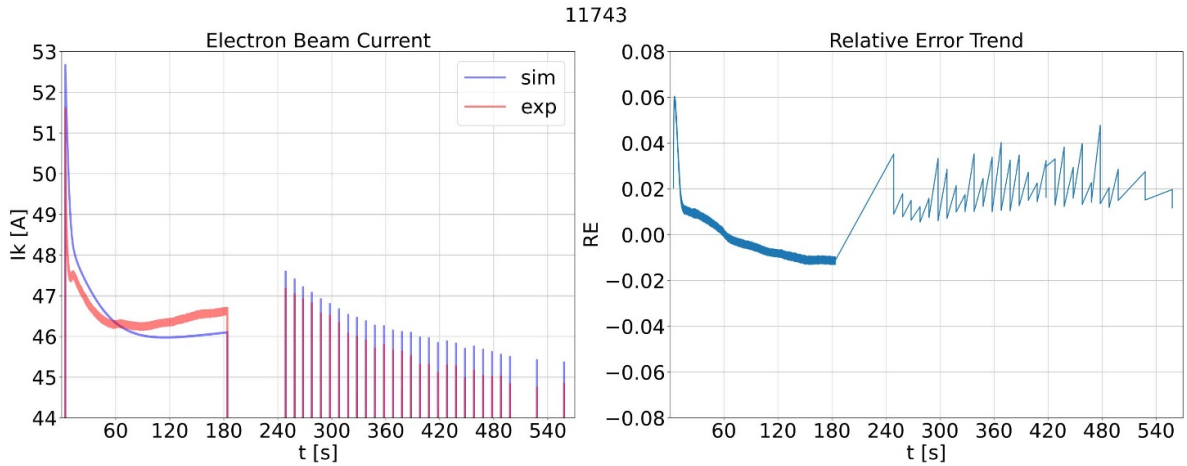
We foresee to carry out an experimental campaign meant to verify and further validate the model by performing pulses with multiple emissions in row, in particular long emissions



(a) **Simulation of FALCON experimental pulse #11739**



(b) **Simulation of FALCON experimental pulse #11741**



(c) **Simulation of FALCON experimental pulse #11743**

Figure 12. Off-phase simulation performance evaluation via microsecond pulses after long emission.

separated by different time-lapses, and long emissions followed by microsecond pulses. With an extended dataset with more experiments showing the phenomena we want to simulate, highly targeted optimizations are possible and better performances can be achieved.

The model will be also trained on data from different gyrotrons in order to explore its flexibility to different geometries and material ages and test how it can adapt to them thanks to the optimization of the hyper-parameters. The possibility to use the same code for various machines with a minimal amount of changes critically decreases the maintainability efforts and allows to have a single main code for many machines.

Custom controllers will be designed for both phases and tested on the real machine:

- for the on-phase, the controller will be completely developed on the best trained model and will be tested on the experimental machine that provided the data;
- for the off-phase, the model needs to be implemented in a real-time environment and run in parallel with the experiments [25], providing a feedback for a controller that will be first designed and tested on the simulations.

Disclaimer

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No. 101052200-EUROfusion).

The views expressed in this publication are responsibility of authors and do not necessarily reflect the views of the ITER Organisation, F4E, Eurofusion or the European Commission.

ORCID iD

G. De Tommasi https://orcid.org/0000-0002-8509-7176

References

[1] Fisch N.J. 1987 Theory of current drive in plasmas *Rev. Mod. Phys.* **59** 175–234
[2] Nusinovich G.S., Thumm M.K.A. and Petelin M.I. 2014 The Gyrotron at 50: historical overview *J. Infrared, Millim. Terahertz Waves* **35** 325–81
[3] Thumm M.K.A. 2014 Recent advances in the worldwide fusion gyrotron development *IEEE Trans. Plasma Sci.* **42** 590–9
[4] Snipes J.A. *et al* 2021 ITER plasma control system final design and preparation for first plasma *Nucl. Fusion* **61** 106036
[5] Vu A.T. *et al* 2024 Progress in the iter plasma control system design *33rd Symp. on Fusion Technology* (*Dublin, Ireland, 22–27 September 2024*)
[6] Preynas M. *et al* 2023 New modelling capabilities to support the ITER EC H&CD System optimisation and preparation of plasma operation *EPJ Web Conf.* **277** 01004
[7] Charbonnier F.M., Strayer R.W., Swanson L.W. and Martin E.E. 1964 Nottingham effect in field and T-F emission: Heating and cooling domains, and inversion temperature *Phys. Rev. Lett.* **13** 397–401
[8] Shinya T. *et al* 2024 Automation of gyrotron electron beam current for fusion devices. *Nucl. Fusion* **64** 126009
[9] Badodi N., Cammi A., Leggieri A., Sanchez F. and Savoldi L. 2021 A new lumped approach for the simulation of the magnetron injection gun for megawatt-class EU gyrotrons *Energies* **14** 2068
[10] Eberhart R. and KENNEDY J. 1995 A new optimizer using particle swarm theory *MHS'95. Proc. 6th Int. Symp. on Micro Machine and Human Science* pp 39–43
[11] Kennedy J. and Eberhart R. 1995 Particle swarm optimization *Proc. ICNN'95-Int. Conf. on Neural Networks* vol 4 pp 1942–8
[12] Goodman T. *et al* 2020 High-power testing of guided-wave components for the ITER ECH upper launcher at the FALCON test facility *IEEE Trans. Plasma Sci.* **48** 1537–42
[13] Dushman S. 1923 Electron emission from metals as a function of temperature *Phys. Rev.* **21** 623–36
[14] Gilmour A.S. 2011 *Klystrons, Traveling Wave Tubes, Magnetrons, Crossed-Field Amplifiers and Gyrotrons* (Artech House)
[15] Meyers R.A. 2001 *Encyclopedia of Physical Science and Technology 3rd edn* (Ramtech, Inc.) (https://doi.org/10.1021/ja015309c)
[16] De D.K. and Olawole O.C. 2016 Modified Richardson-Dushman equation and modeling thermionic emission from monolayer graphene *3rd Int. Conf. on African Development Issues* vol 9927, ed E.M. Campo, E.A. Dobsiz and L.A. Eldada p 99270E
[17] Olawole O., De D., Oyedepo S. and Ezema F. 2021 Mathematical models for thermionic emission current density of graphene emitter *Sci. Rep.* **11** 22503
[18] Cronin J.L. 1981 Modern dispenser cathodes *IEE Proc.* I **128** 19–32
[19] Nottingham W.B. 1941 Remarks on energy losses attending thermionic emission of electrons from metals *Phys. Rev.* **59** 906–7
[20] Introini C. *et al* 2024 Sensitivity analysis of the lumped thermal model of the EU 170 GHz gyrotron magnetron injection gun *IEEE Trans. Plasma Sci.* **52** 3836–41
[21] Goodman T. *et al* 2023 High power mm-wave loss measurements of ITER ex-vessel waveguide components at the FALCON test facility at the Swiss Plasma Center *EPJ Web Conf.* **277** 04010
[22] Corless R.M., Gonnet G.H., Hare D.E.G., Jeffrey D.J. and Knuth D.E. 1996 On the LambertW function *Adv. Comput. Math.* **5** 329–59
[23] Clerc M. and Kennedy J. 2002 The particle swarm - explosion, stability, and convergence in a multidimensional complex space *IEEE Trans. Evol. Comput.* **6** 58–73
[24] Frénay B., Doquire G. and Verleysen M. 2013 Is mutual information adequate for feature selection in regression? *Neural Netw.* **48** 1–7
[25] Neto A.C. *et al* 2010 MARTe: A multiplatform real-time framework *IEEE Trans. Nucl. Sci.* **57** 479–86