Design and optimization of Artificial Neural Networks for the modelling of superconducting magnets operation in tokamak fusion reactors

Design and optimization of Artificial Neural Networks for the modelling of superconducting magnets operation in tokamak fusion reactors

A. Froio, R. Bonifetto, S. Carli, A. Quartararo, L. Savoldi and R. Zanino

NEMO group, Dipartimento Energia, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Abstract

In superconducting tokamaks, the cryoplant provides the helium needed to cool different clients, among which by far the most important one is the superconducting magnet system. The evaluation of the transient heat load from the magnets to the cryoplant is fundamental for the design of the latter and the assessment of suitable strategies to smooth the heat load pulses, induced by the intrinsically pulsed plasma scenarios characteristic of today's tokamaks, is crucial for both suitable sizing and stable operation of the cryoplant. For that evaluation, accurate but expensive system-level models, as implemented in *e.g.* the validated state-of-the-art 4C code, were developed in the past, including both the magnets and the respective external cryogenic cooling circuits. Here we show how these models can be successfully substituted with cheaper ones, where the magnets are described by suitably trained Artificial Neural Networks (ANNs) for the evaluation of the heat load to the cryoplant. First, two simplified thermal-hydraulic models for an ITER Toroidal Field (TF) magnet and for the ITER Central Solenoid (CS) are developed, based on ANNs, and a detailed analysis of the chosen networks' topology and parameters is presented and discussed. The ANNs are then inserted into the 4C model of the ITER TF and CS cooling circuits, which also includes active controls to achieve a smoothing of the variation of the heat load to the cryoplant. The training of the ANNs is achieved using the results of full 4C simulations (including detailed models of the magnets) for conventional sigmoid-like waveforms of the drivers and the predictive capabilities of the ANN-based models in the case of actual ITER operating scenarios are demonstrated by comparison with the results of full 4C runs, both with and without active smoothing, in terms of both accuracy and computational time. Exploiting the low computational effort requested by the ANN-based models, a demonstrative optimization study has been finally carried out, with the aim of choosing among different smoothing strategy for the standard ITER plasma operation.

Keywords: nuclear fusion, tokamak, ITER, superconducting magnets, modeling, Artificial Neural Networks

Introduction

Most of the present and future nuclear fusion devices, like the ITER [1] tokamak under construction in Cadarache (France) at a cost of 10+ billion Euro in a collaboration between 7 parties (China, EU, India, Japan, South Korea, Russian Federation and USA), use superconducting (SC) magnets to generate the magnetic field needed for the plasma confinement.

The ITER magnet system is made of four sub-systems, see Fig. 1: 18 Toroidal Field (TF) coils, a Central Solenoid (CS), 6 Poloidal Field coils (PF) and 18 Correction Coils (CC) [2].

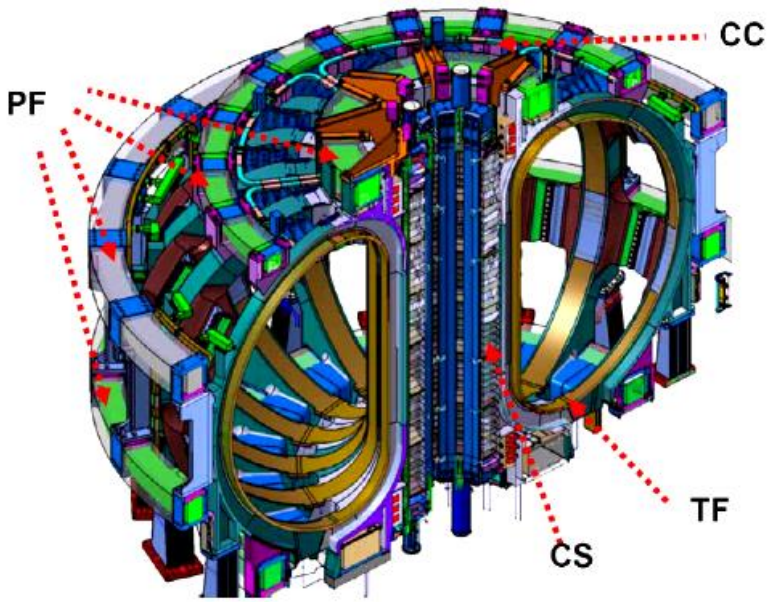


Fig. 1: The ITER magnet system (reproduced from [2]).

The ITER cryoplant guarantees that all coils are cooled at ~ 4.5 K to ensure a sufficient temperature margin in order to avoid the quench of the SC magnets during operation. For all magnets, the coolant is supercritical helium (SHe), at a pressure of 0.5-0.6 MPa. Dedicated cryogenic cooling loops extract the heat load from the magnets, releasing it to saturated liquid helium (LHe) baths, acting as thermal buffers in the transfer of the load to the cryoplant, see Fig. 2.

The inductive operation of a tokamak is intrinsically pulsed, as the CS acts as the primary coil of a transformer, driving a toroidal current inside the secondary (*i.e.*, the plasma). In order to drive this inductive current, the CS operates with variable current producing a variable magnetic field and this leads in the end to AC losses and eddy currents in the magnets, notwithstanding their superconducting nature. In addition, during the plasma burn, nuclear fusion reactions in the plasma result in a non-negligible nuclear heat load on the TF coils, notwithstanding the screening of the different components located between them and the plasma. As the cryoplant must be designed and sized considering the peak heat load during operation, such pulsed scenarios can significantly increase its cost. As a

consequence, the smoothing of the heat load from the SC magnets to the cryoplant is being investigated [3, 4, 5], also in consideration of the fact that the SC magnet system and related cryogenics are responsible for 25-30% of the total cost of the plant. A dedicated experimental facility (HELIOS) was realized at CEA Grenoble (France) [6, 7] and different strategies are under consideration, both from the experimental [8, 9] and the computational [10, 11, 12, 13] points of view.

Computational tools implementing detailed models, such as the 4C code [14], can be used to develop and test different control strategies. However, such tools require a large computational effort because they typically provide a very detailed description of the flow inside the magnet, which, on the other hand, is not needed when the main concern is the heat load to the cryoplant.

As an alternative, fast models have been developed to predict the heat load to the cryoplant, based on the simplification of the set of partial differential equations (PDEs) describing the physics of the system [15, 16, 17, 18].

A radically different approach to develop a simplified model for the heat load to the LHe bath has been recently proposed [19], where the whole cryogenic circuit and the SC magnets have been lumped in a single Artificial Neural Network (ANN) [20], and also successfully applied to the computation of the heat load that is transferred to the LHe bath during the ITER CS operation [21]. Although this first ANN model could not yet deal with control strategies, it showed advantages in terms of computational effort with respect to both the full physics model (as implemented in the 4C code) and also to other simplified physical models [18], with a reduction by a factor ~10000 with respect to 4C and allowing a simulation in ~1/300 of real time on a single i7-4810MQ core.

The inputs to the ANN are the heat deposition/generation in the magnet while the outputs (which are also fed back as additional inputs of the ANN, if needed) are the helium temperature, pressure and mass flow rate at the inlet of the heat exchanger (HX), see [19]. As a difference with respect to the above-mentioned PDE-based simplified models, the ANNs are based on simple matrix-vector products; the physics content of the ANN model is contributed by the choice of physically justified input/output variables and by the training performed using either physics-based models or directly experimental data.

More recently, the ANN approach has been further extended, using them to model the magnet alone [22], which has then been replaced in the physical 4C circuit module by the ANN black-box. This opened the opportunity to model a scenario with a control acting on the cooling loop. The results of the novel approach have been compared against 4C simulations and they have also been validated showing a good agreement against experimental data from the HELIOS loop [23].

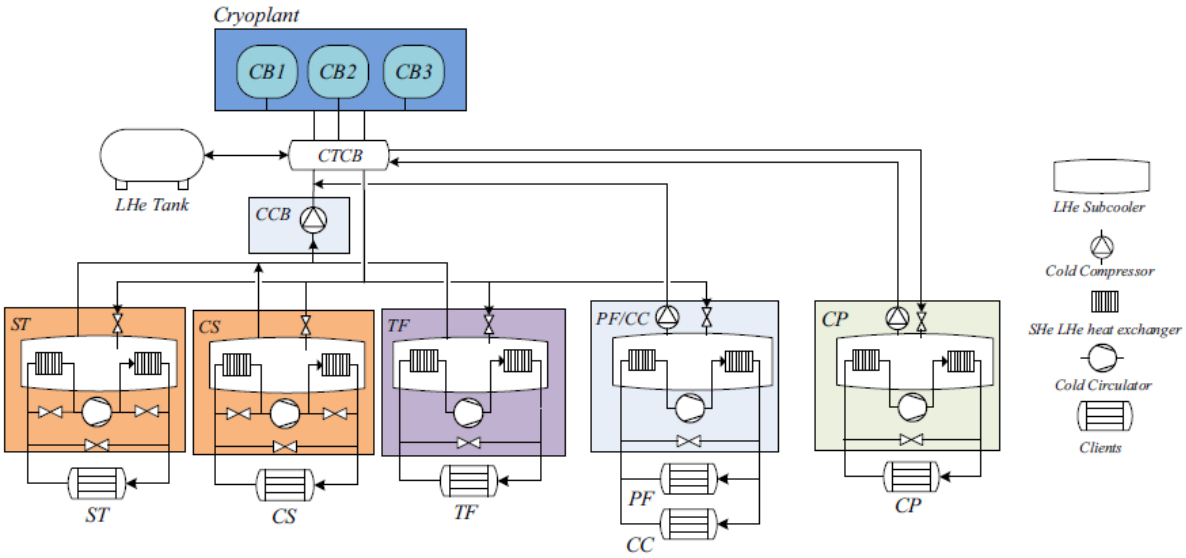


Fig. 2: The ITER magnets cooling system (reproduced from [12]). CBs are the cold boxes, CTCB is the cryoplant termination cold box, CCB is the cold compressor box, ST are the structures, CPs are the cryopumps.

The present paper is focused on the optimization of the different degrees of freedom that an ANN allows, including both its structure and its parameters (inputs, outputs, number of delays and of hidden neurons, see below). The two resulting ANN-based models for the assessment of control strategies for the ITER CS and TF magnet systems, respectively, are then presented. The capability of such models to predict the heat load on the cryoplant in a standard operating scenario with no controls is shown first; then the two models are used to assess the effectiveness of different mitigation strategies for the smoothing of the pulsed heat load to the cryoplant. Finally, the results of the ANN-based models are compared with 4C in the most successful regulated scenario, both in terms of accuracy and computational time.

1. ANN-based models

In this section the parametric study performed to optimize the ANNs structure and parameters, to best design the two ANN-based models developed for the ITER CS and TF coils cooling systems, is presented.

1.1. *Physical objects*

1.1.1. *ITER CS*

The ITER CS magnet (Fig. 3), whose details can be found in [21], is constituted by 6 different modules, cooled in a hydraulic parallel by SHe in forced flow and releasing the heat load to a LHe bath through a HX located downstream of the circulator, see Fig. 4a. The three upper modules (CSU) are first connected in a tight parallel through suitable piping and manifolds, and only at a second level connected to the tight parallel between the CS

lower (CSL) modules, see Fig. 4a. For the time being, the LHe bath will be considered as the interface with the refrigeration system, and it will be assumed to be at the constant temperature of 4.3 K.

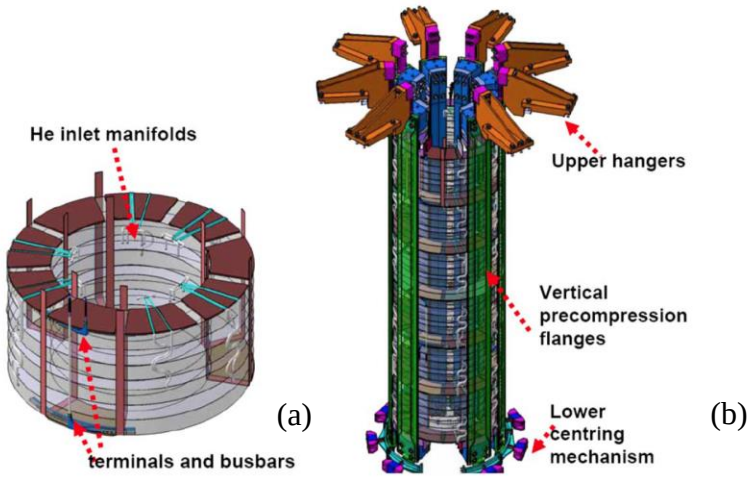


Fig. 3: The ITER CS magnet: (a) a single CS module; (b) the entire CS coil (reproduced from [2]).

The control of the cooling circuits can be achieved by, for instance, the opening of the bypass valve (BV) or control valve (CV), or by varying the cold circulator speed, which, in the end, will modify the mass flow rate flowing through the different modules and the HX. Three Proportional Integral (PI) controllers are thus included in the 4C model, see Fig. 4a, basing their regulation on a "virtual" heat flux sensor that evaluates the thermal power instantly released through the HX to the LHe bath measuring the He flow and enthalpy at the inlet and outlet of the bath. The virtual heat flux sensor is here adopted for the sake of simplicity: the heat flux released to the LHe bath could be connected to the variation of the thermo-dynamic condition of the LHe bath, but that would require also the modeling of (part of) the refrigeration system, which is beyond the scope of the present work.

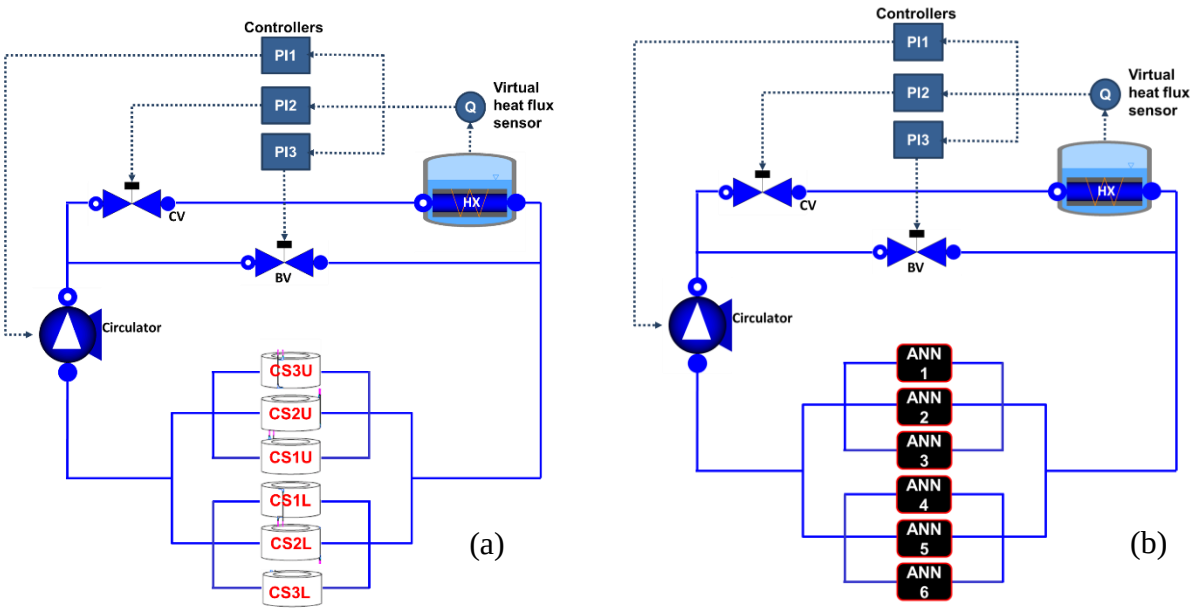


Fig. 4: Schematic of the CS models: (a) 4C model; (b) ANN model.

1.1.2. ***ITER TF***

The 4C model of the ITER TF magnets cooling system is shown in Fig. 5, and its details can be found in [24, 25]. The magnet system is composed by two separate cooling circuits, one for the TF winding pack and one for the structures (casing), thermally coupled through the ground insulation. In the ANN-based model, as for the CS model, the ANNs substitute only the magnet, so that the intervention of controllers on the two cooling loops, not hidden in the ANN black-box, can be simulated. As far as the mitigation strategies for the smoothing of the pulsed heat loads to the cryoplant are concerned, the ones adopted here act only on the casing cooling loops, through a PI controller that acts on the opening of the BV or on the speed of the pump on the basis of a "virtual" heat flux sensor, again adopted here for the sake of simplicity to evaluate the thermal power released to the LHe bath. Other simplified physical models [16] base the control on the power released *from* the LHe bath to the cryoplant, which is not included in the current 4C model. In ITER, the control is likely to be applied only on the casing cooling channels. A recent analysis [24] has shown that a reduction of the mass flow rate in the WP cooling loop below the nominal value does not allow sufficient cooling to satisfy the required temperature margin. Moreover, the power to the cryoplant coming from the WP is more constant compared to the casing one, see below.

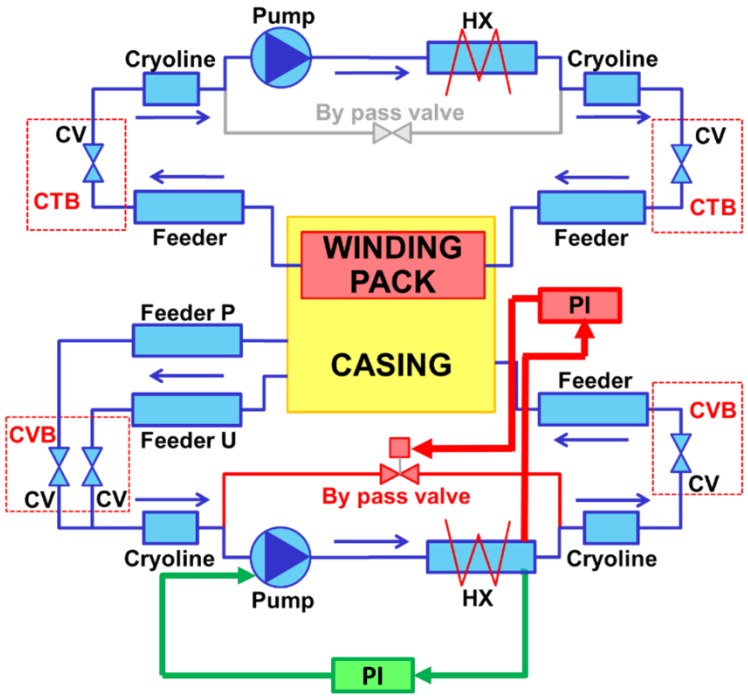


Fig. 5: 4C model of the ITER TF magnet cooling system. CTB is the cold terminating box, CVB is the cold valve box, HX is the heat exchanger, CV is the control valve, BV is the bypass valve, PI is a proportional integral controller, feeder "P" collects SHe from the CCC on the plasma side of the coil casing and feeder "U" from the CCC on the other three sides.

1.2. *Development of the ANNs*

1.2.1. *ITER CS*

As explained in [22, 23], since the controllers act on components belonging to the cooling loop, the ANNs should only model the dynamic behavior of the CS modules themselves. Thus, the same approach successfully proposed in [21] was applied, where each module is analyzed separately, thanks to the hydraulic parallel connection, to develop and train a new set of ANNs. The ANN-based model is constituted by 6 ANN black-boxes, one for each module of the CS coil, embedded in the already existing 4C cooling loop, see Fig. 4b.

Following what has been done in [22], on one side the circuit model needs as boundary conditions from the ANN black-box the coil's inlet/outlet mass flow rates, $\dot{m}_{in}$ and $\dot{m}_{out}$, as also the coil outlet temperature T_{out}; these three variables are then the network outputs. On the other side, the black-box boundary conditions are given through pressures p_{in} and p_{out} and inlet temperature T_{in}, becoming inputs of the ANN, together also with the driver, which is the heat deposited in the magnet by the AC losses Q_{AC}. The ANN has thus four input variables and three output variables, see Fig. 6.

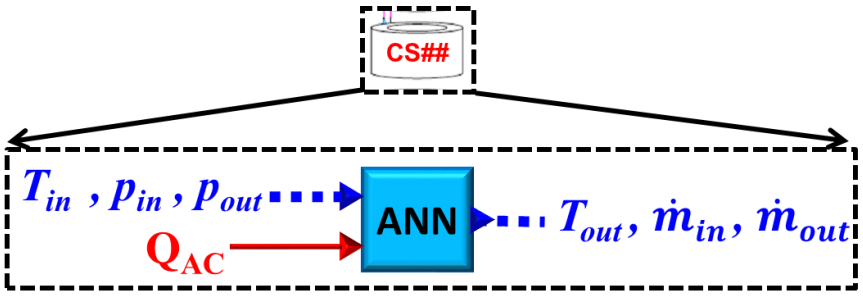


Fig. 6: Input and output variables for a single CS ANN.

1.2.2. ***ITER TF***

Since the control will be applied only on the casing cooling circuit, see above, the ANNs already developed in [22] for the WP are kept the same, while new ANNs are developed for the casing including in the training set also a number of 4C simulations performed at different mass flow rates. The final ANN-based model has the same structure of that reported in [22], see Fig. 7: in fact, all the components are the same, with the exception of the above-mentioned casing ANN. The heat loads to the winding (nuclear radiation $Q_{nuc,W}$ and AC losses Q_{AC}) and casing (nuclear radiation $Q_{nuc,C}$, eddy currents Q_{eddy} and vertical stabilization Q_{VS}) are the first inputs to two separate ANNs (one for the winding, W_a, and the other one for the casing, C_a), which are trained from simulations in which winding and casing are thermally insulated (*i.e.*, adiabatic). The other inputs of these two networks are the inlet pressure, together with the pressure drop across the coil (only for the casing), from the 4C circuit model, as shown in Fig. 7. The inlet temperature, differently from the ANNs developed for the CS modules, is not taken as an input in this case, because its variation is negligible during the normal operation of the magnet, so that it would not add significant information to the network training - as the ANN does not solve the thermal hydraulic problem as a set of partial differential equations, it is not strictly necessary to have three boundary conditions. As outputs, these networks compute the inlet and outlet mass flow rates, which are provided directly to the 4C circuit model, in the same way as it has been done for the CS modules, and the He outlet temperature. These temperatures (that correspond to "adiabatic" behavior of the WP and of the casing) are then fed to two additional ANNs (W_c and C_c) accounting for the thermal coupling between the WP and casing, which post-process the "adiabatic" outlet temperatures computed by W_a and C_a, providing the 4C circuit model with the outlet temperatures of the thermally coupled WP and casing.

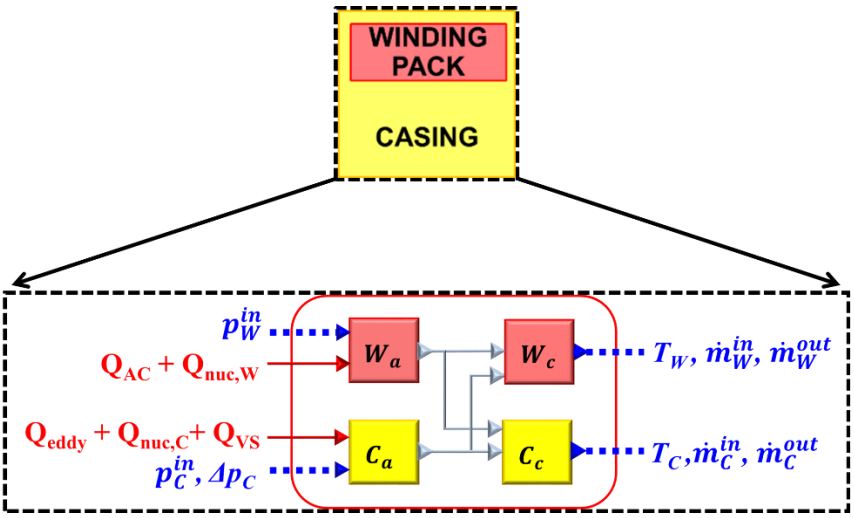


Fig. 7: Structure of the ANN-based model of the ITER TF cooling system.

1.2.3. ***Optimization of ANN parameters***

The development of a suitable model for the ANNs requires first an analysis in order to properly choose the different free parameters of the single neural network, whose general structure is shown in Fig. 8. These parameters are:

a) Network topology, *i.e.*, the number and size of hidden layers and the presence of feedbacks;

b) Activation functions;

c) Number of input (x) and output (y) variables;

d) Number of delays, *i.e.*, the number of previous values of the input to be fed to the network;

e) Time step for the training.

We report here the results of the optimization analysis performed for the ANN developed for the TF casing in the adiabatic condition.

Concerning point a), since the network is connected to a physical model of the circuit, that provides the TH variables at each time step, there is no need for feedbacks, *i.e.* no need to use of the past output values of the network as additional input variables, since the circuit itself constitutes a feedback loop. Moreover, the universal approximation theorem [26] states that a feed-forward network with a single hidden layer containing a finite number of neurons can approximate any continuous function on compact subsets of $\mathbb{R}^n$, under mild assumptions on the activation functions, namely a sigmoid (generic mathematical function having an "S" shape) for the hidden layer and a linear function for the output layer [27], see Table 1. In order to fulfill these assumptions, the hyperbolic tangent (a particular type of sigmoid function) has been adopted here as activation function, see point b), so that the universal approximation theorem allows having a single hidden layer, see

Table 1. The last step to complete the network topology is the number of neurons in the hidden layer, which was chosen here equal to six after empirical evaluations.

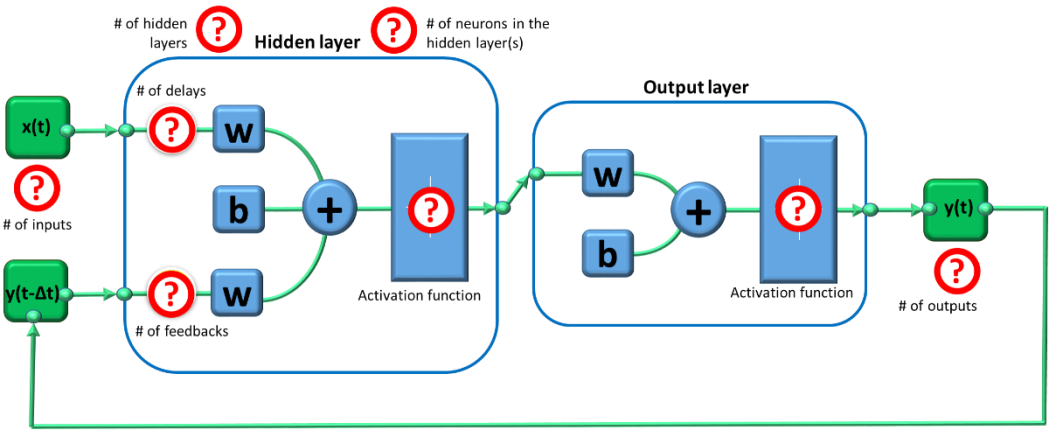


Fig. 8: General structure of a single neural network.

In order to address point c) above finding the best choice of the input variables and their delays, the Principal Components Analysis (PCA) is used, which is a technique for the simplification of multivariate signals [28, 29].

Starting from the original data, we evaluate the auto-covariance matrix, defined as:

$$C = \frac{1}{m}\sum_{i=3}^{n} \underline{x}^{(i)} \cdot \underline{x}^{(i)\,T} \qquad (1)$$

where n is the number of training cases and m the number of components of the input vector $\underline{x}^{(i)}$, defined as $\underline{x}^{(i)} = [Q(t),\ p_{in}(t), p_{out}(t), Q(t-\Delta t),\ p_{in}(t-\Delta t), p_{out}(t-\Delta t), Q(t-2\Delta t),\ p_{in}(t-2\Delta t), p_{out}(t-2\Delta t)]$. The input vector contains here also its first and second delays (the variables evaluated at the times $t-\Delta t$ and $t-2\Delta t$, respectively), in order to obtain their significance.

The choice of the number of principal components, which are the components of the vector $x^{(i)}$ that carry the highest amount of information, is done evaluating the eigenvalues of the auto-covariance matrix (shown in Fig. 9a). Each normalized eigenvalue tells what the percentage importance of each variable of the original vector is. The threshold for the accepted value of the normalized eigenvalues is chosen to be 10^{-4}, since it means neglecting the 0.1‰ of the input information.

Concerning the choice of the best architecture for the TF casing adiabatic network, the PCA (see Eq. 1 and Fig. 9a) suggests that one of the pressure signals can be discarded as it does not add any additional useful information to the ANN. However, the removal of one of the pressures and could lead to problems in the coupling to the cryogenic circuit model since it produces a strong decrease in the accuracy of the mass flow rate prediction. That is due to the fact that, physically, the mass flow rate has a strong dependence on the pressure drop. Hence is has been decided to use as third input variable the pressure drop across the coil. The results of the PCA in this case, shown in Fig. 9b, confirm that this variable adds significant information to the ANN, since its time evolution is different from that of the inlet pressure.

Moreover, the PCA results, see again Fig. 9, show that the use of time delays mentioned in point d) above is not necessary for the TF adiabatic casing ANN to reconstruct the dynamics of the system, since they do not add significant information with respect to the non-delayed inputs. With a smaller number of input variables, it is thus possible to reduce the number of neurons in the hidden layer, decreasing the computational time required for the training of the network. Note that the optimization results led, however, to slightly different results in the other cases: for instance, both the ANN developed for the ITER TF winding in adiabatic condition and for the ITER CS require a delay, see Tables 1-2.

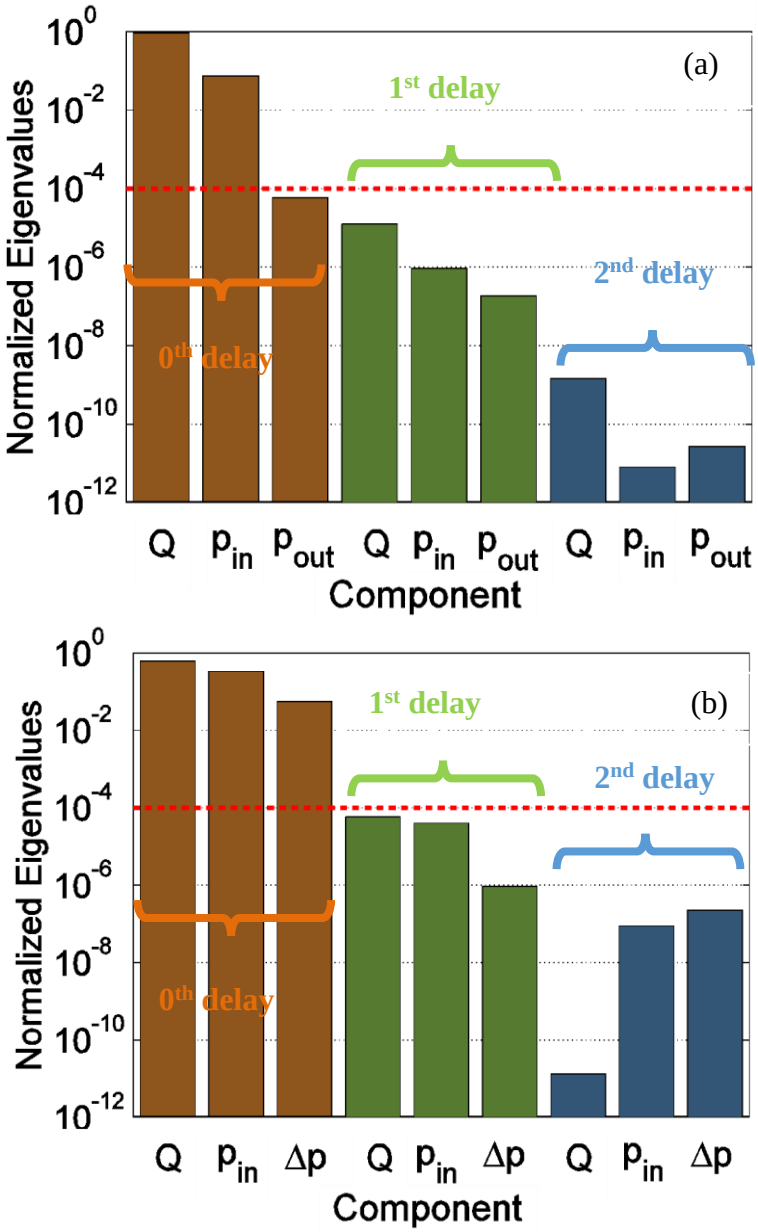


Fig. 9: Results of the PCA for the ITER TF adiabatic casing ANN for the two input sets considered in the study: Q, p_{in}, p_{out} (a) and Q, p_{in}, Δp (b). The dashed lines represent the threshold for the accepted normalized eigenvalues.

Finally, the time step for the training in point e) above and, in case they are present, the distance among time delays used as inputs by the ANNs have been chosen looking at the mutual information (MI) between the original time series of the data for the training and its delayed series.

Let us consider the input signal ξ, whose elements are the triplets $\xi = [Q(t_i), p_{in}(t_i), \Delta p(t_i)]$, provided as a single training example to the ANN, and the delayed input signal ζ, whose elements are $\zeta = [Q(t_i - \Delta t), p_{in}(t_i - \Delta t), \Delta p(t_i - \Delta t)]$. It is possible to evaluate how much the two signals are correlated to each other, as a function of the time delay Δt, considering ξ and ζ as the outcomes of a random process that can be analyzed statistically looking at the MI. Being ξ and ζ discrete samplings of three-dimensional random variables, the definition of the MI cannot be implemented directly; to estimate the MI, the Kraskov, Stögbauer and Grassberger [30] algorithm has been used, which provides an estimator of MI directly based on the set (ξ_i, ζ_i).

The MI measures the information that ξ and ζ share: it tells how much the knowledge of one of the two variables reduces the uncertainty on the other. Since in the selection of the optimal time step the variable ζ is simply the variable ξ shifted by a quantity Δt, if this is too small, the two variables are too close and a sort of redundant information about the system is given. If, instead, Δt is too large, the variables are completely uncorrelated in a statistical sense and the dynamics of the system cannot be reconstructed.

Starting from a sampling time of 1 s, the MI for the ITER TF casing adiabatic ANN is monotonically decreasing, see Fig. 10. In these conditions, relying on what can be found in [31], one of the smallest time steps can be chosen, *i.e.* 1 or 2 s. It has been chosen not to consider sampling times smaller than 1 s because it would lead to an increased size of the data for training cases and thus the duration of the training itself. For the TF adiabatic casing ANN a time step of 2 s has been chosen.

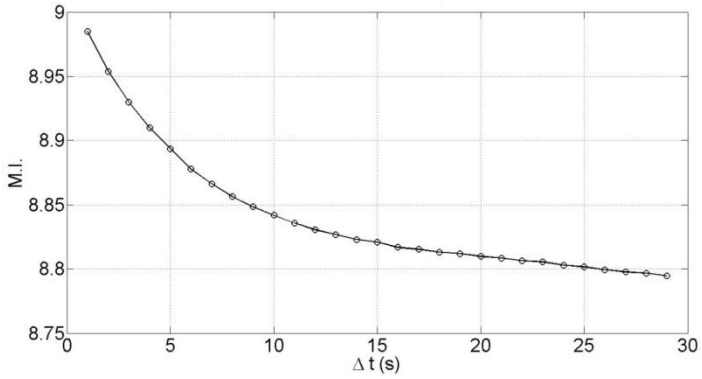


Fig. 10: Result of the Mutual Information analysis for the ITER TF adiabatic casing ANN as a function of the sampling time step for the training data inputs.

The MI analysis shows similar results for the ITER CS and ITER TF adiabatic winding cases and a training time delay of 1 s instead of 2 s is chosen.

The final structure of the ANNs for the TF casing in adiabatic condition is shown in Fig. 11 and its main parameters are summarized in Table 1 and Table 2.

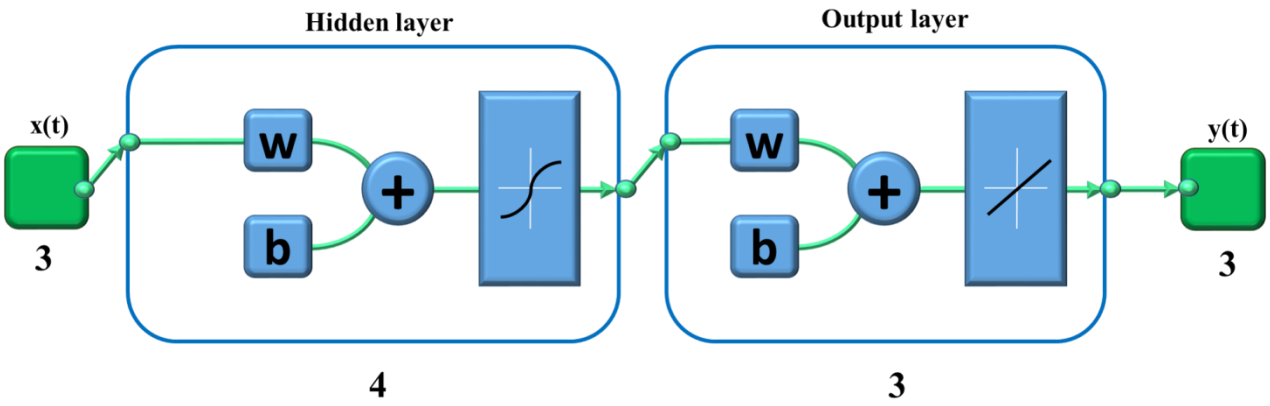


Fig. 11: Final structure of the ITER TF adiabatic casing ANN developed and used in the present work.

Table 1: Parameters of the ANNs used for the ITER CS model.

Type of network	Feedforward time-delayed
Number of hidden layers	1
Number of neurons in the hidden layer	6
Hidden layer activation function	Hyperbolic tangent
Output layer activation function	Linear
Input variables	T_{in}, p_{in}, p_{out}, Q
Output variables	T_{out}, $\dot{m}_{in}$, $\dot{m}_{out}$
Number of input delays	1 of 1 s
Training time step	1 s

Table 2: Parameters of the ANNs used for the ITER TF model.

Network	**Adiabatic casing**	**Adiabatic winding**	**Coupling**
Type of network	Feedforward	Feedforward time-delayed	Feedforward time-delayed
Number of hidden layers	1	1	1
Number of neurons in the hidden layer	4	6	6
Hidden layer activation function	Hyperbolic Tangent	Hyperbolic Tangent	Hyperbolic Tangent
Output layer activation function	Linear	Linear	Linear
Input variables	p_{in}, Δp, Q	p_{in}, Q	$T_{out,W,ad}$, $T_{out,C,ad}$
Output variables	T_{out}, $\dot{m}_{in}$, $\dot{m}_{out}$	T_{out}, $\dot{m}_{in}$, $\dot{m}_{out}$	$T_{out,W,cou}$, $T_{out,C,cou}$
Number of input delays	0	1 of 1 s	1 of 1 s
Training time step	2 s	1 s	1 s

Once the network architecture and its parameters are identified, the ANNs can be trained to mimic the dynamic behavior of the ITER magnets. Thanks to the recent validation of the 4C code [13, 32, 33], 4C can be used as a reliable tool for the prediction of the TH behavior of cryogenic circuits like the ones studied in this work. Thus, the training is done using ad-hoc 4C simulation results in which both the heat load and the mass flow rate inside the circuit are separately varied with the sigmoidal waveform of Eq. 6, which is natural for the ANN as explained in [19], and whose parameters are better explained in Fig. 12.

$$f(t) = P_0 + \frac{A}{1+e^{-m(t-\tau)}} \qquad (6)$$

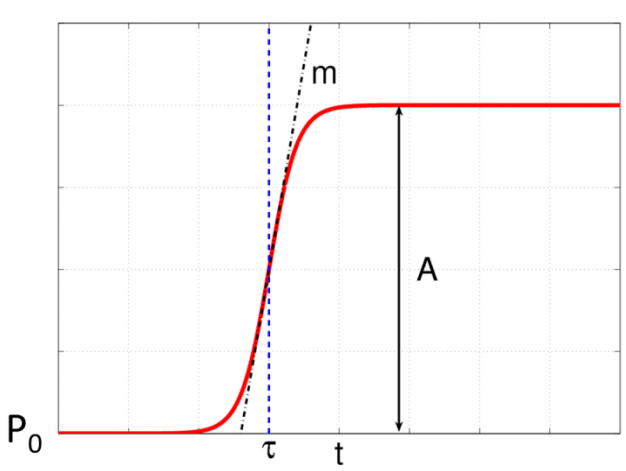


Fig. 12: Logistic (sigmoidal) waveform used to train the ANNs.

The 4C code needs to simulate an initial pseudo-transient, where only the static loads are applied, in order to reach the equilibrium state; however, the ANNs used in the present work are not trained to follow this initial stabilization transient, but are rather developed to model the operation during plasma shots, so that the initial pseudo-transient is not simulated.

All the different types of heat load on the CS and TF magnets (AC losses for the CS, nuclear, eddy currents and vertical stabilization for the TF) are considered and applied with the sigmoid waveform above with different amplitude *A*, see Eq. 6, in order to cover the power range foreseen for plasma operation (except for the load on the CS during the initial magnetization, as reported in [21]); the spatial distribution of the heat deposition is the same of the standard 15 MA plasma current operating scenario of ITER [21, 24, 25]. In addition, some simulations are done keeping the heat load at constant level, but varying the mass flow rate, again with a sigmoid waveform, acting on the bypass valve or on the mass flow forced by the circulator. The range of the input variables explored in the training data set is represented in Fig. 13. The training through 4C simulations, that include not only the detailed model of the magnets, but also the model of the cryogenic cooling circuit, guarantees that the thermal-hydraulic variables fed to the magnets vary in a self-consistent way, *i.e.*, exploring the portion of the space of Fig. 13 that is physical and "natural" for the systems at hand.

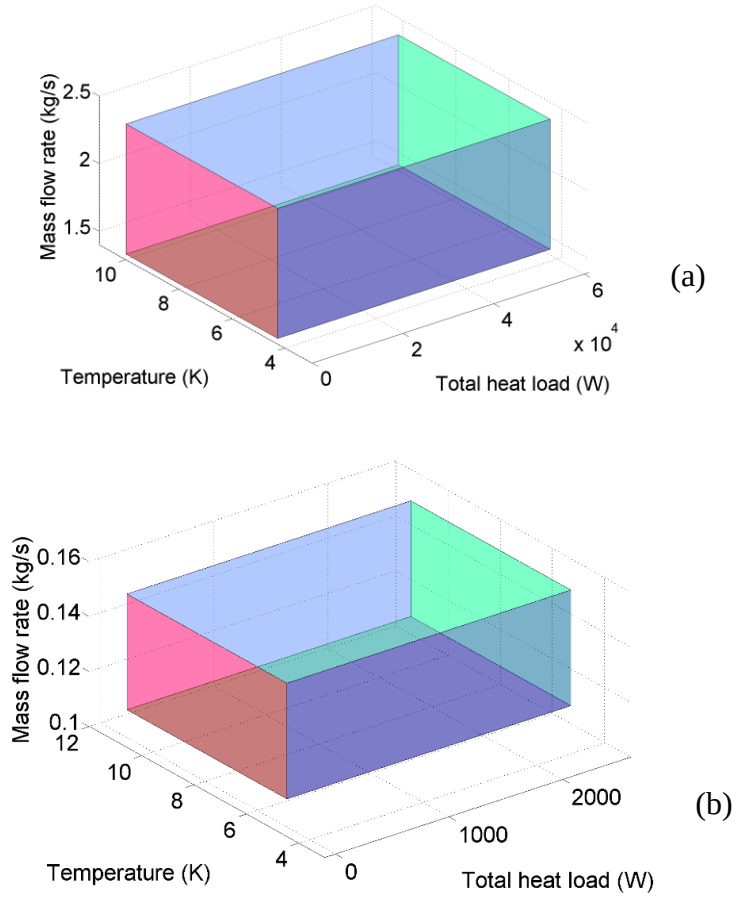


Fig. 13: Range explored during the training for the ANNs: (a) CS; (b) TF casing.

Finally, when the training process is completed, the ANNs are embedded in the 4C model of the ITER magnet cooling systems, substituting each of the CS modules and the TF winding and casing, see again Fig. 4b and Fig. 7.

2. Application to the development of control strategies

In this section the predictive capabilities of the ANN-based models for the ITER CS and TF coils are shown by comparing their results with those of the detailed 4C model in the case of the standard 15 MA – 14 kW plasma operation scenario for ITER, without any control of the power to the LHe baths. Note that this scenario, with the corresponding evolution of the heat loads, has never been provided to the ANN as part of the training.

After that, an optimization study is performed using the ANN-based model of both the CS and TF, in order to find the best control strategy for the heat load smoothing, *i.e.* optimize the parameters of the PI controllers. Finally, the computed results are checked against 4C simulations where the same control strategy has been adopted.

2.1. *ITER CS coil*

The input power to each module of the ITER CS coil, for the standard 15 MA – 14 kW scenario is reported in [21].

In Fig. 14 the results in terms of evolution of the power deposited to the LHe bath are shown. As it can be seen, the prediction of the ANN-based model has a good accuracy, with power peaks that show only some anticipation with respect to the detailed 4C model. Nevertheless, the relative error on the peak power is ~0.1% for the second pulse, after which the system becomes periodic. The offset before and after the second plasma shot is due to the fact that the equilibrium for the system with the ANN is reached at a different level with respect to the original 4C model, due to the effect of the small errors introduced by the ANN model in the closed loop dynamic simulation. This is evident especially in the first pulse, in which the system is not going back to the initial steady state, being still in dynamic conditions when the second pulse starts. The unbalance in the first pulse causes a bias of the steady state level, which does not increase in the subsequent pulses. Note however that the simulation of the first, non-periodic pulse is beyond the scope of this work.

Due to the approximation and interpolation error of the neural networks, as well as the stabilization effects mentioned before, the conservation of energy during each pulse is indeed not guaranteed (*i.e.*, the input energy is not forced to be equal to the energy extracted from the magnet). Nevertheless, the prediction of the total energy released to the LHe bath is very good, with a relative error of ~7% during the first pulse and of ~13% during a periodic pulse.

A final comparison is done also in terms of computational effort, showing that the ANN-based model is ~500 times faster than the detailed 4C one.

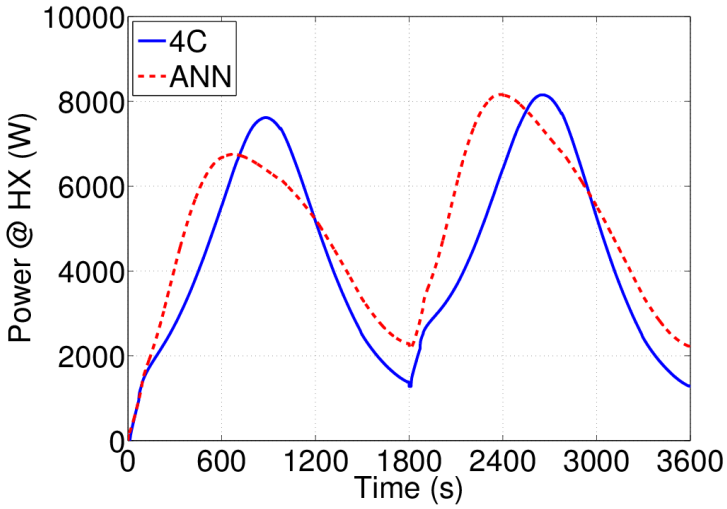


Fig. 14: Comparison of the evolution of the power released at the HX for the ITER CS, when computed by the 4C model (solid line) and the ANN model (dashed line).

2.1.1. *Optimization study*

Since the ANN model is proven sufficiently accurate, it can replace the detailed CS model to evaluate different strategies to smooth the heat load to the cryoplant, allowing *e.g.* parametric studies to be performed in a very short time. For the same input power scenario, a parametric study is performed by changing the parameters of the PI controller acting on the cold circulator speed based on the value of the heat transferred to the LHe bath (evaluated as total enthalpy drop through the HX). This smoothing strategy already proved to be the most effective among other strategies in [16]. The parameters driving the PI controller action are:

- Set-point (*SP*): is the value of power that the controller tries to reach and keep constant;
- Proportional gain (K_p): controls how much the controlled variable changes, for a given change in the input; the higher its value, the faster the controller response;
- Integrator time constant (T_i): controls the dependency of the controlled variable on the integral of the error. In the expression for the evaluation of the controller output, it appears at the denominator of the error integral, so it has the unit of measurements of time and simply provides a weight to the integral term; for this reason it may assume any positive value. As a higher contribution of the integral term results in a faster evolution of the process variable towards the set-point (but may cause overshoot), the higher the value of T_i, the lower the contribution of the integral term and the smoother the response.

Since the effect of T_i and K_p is similar (although reversed), Fig. 15 reports only the results of the optimization study carried out changing the *SP* and T_i in order to find the parameters for the best smoothing of the heat load.

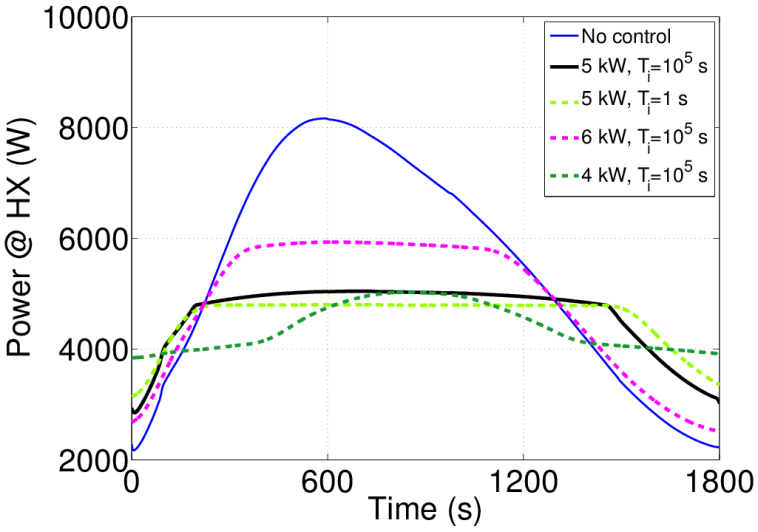


Fig. 15: Results of the parametric study performed for the ITER CS (the power transferred to the HX is evaluated without taking into account the static load); the solid thick line represents the best smoothing scenario. The figure shows the third pulse, after which the system becomes periodic.

In Fig. 16a a comparison of the effect of the *SP* is shown, being the other two parameters kept constant, on the average power transferred to the HX, while in Fig. 16b, the same comparison is performed on the coil maximum

inlet and outlet temperatures. These results show that the ANN-based model for the CS exhibits the same behavior of the physical model in [16], *i.e.* when the control acts on the pump speed any variation of the *SP* leads to a change in the average power transferred to the HX while keeping constant the coil maximum inlet and outlet temperatures.

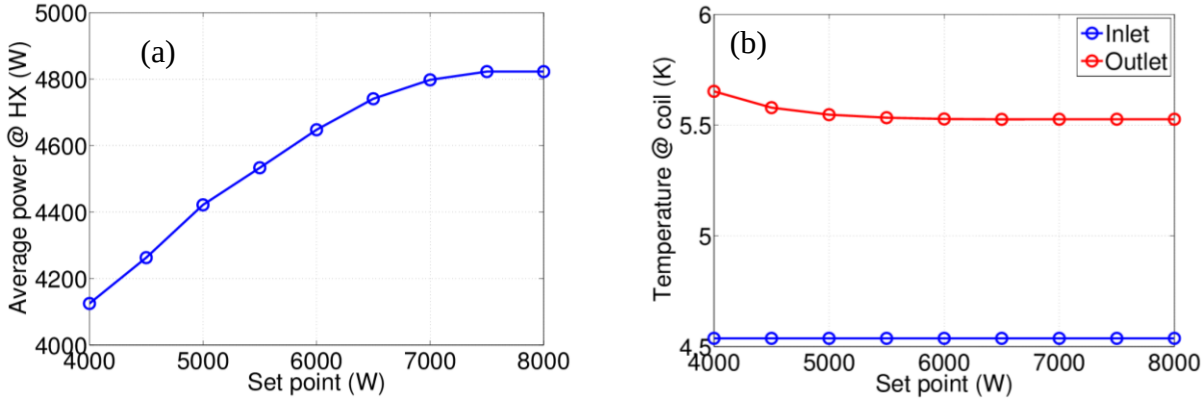


Fig. 16: Parametric study on the controller *SP* performed for the ITER CS: (a) average power transferred to the HX (evaluated without taking into account the static load); (b) coil maximum inlet and outlet temperatures. Both the figures refer to the third pulse, after which the system becomes periodic.

Fig. 17 shows how, during the whole transient for the best smoothing scenario, the pump working point is always far from the surge point, at any value of rotational speed, as in [12, 16]. Moreover, within this scenario, the controller never reaches the lowest saturation point, keeping the power value more constant during the control.

Once the best set of parameters is found, it is used to perform a detailed 4C simulation including the control of the cold circulator speed. The results are reported in Fig. 18 in terms of power released to the LHe bath and relative pump speed. The ANN model starts regulating the pump speed much earlier than the 4C and, after the plateau, the power decreases with a lower slope, remaining at a higher value with respect to the 4C. Despite these differences, the power level reached during the plateau is practically the same in the two models, with a relative error of ~ 1%, and showing a reduction of the peak power of ~ 30% with respect to the non-controlled case; the prediction on the minimum pump speed is a bit worse, due to the fact that the control starts earlier. The error of the ANN model in the evaluation of the energy per pulse released at the HX is below 5% for all pulses.

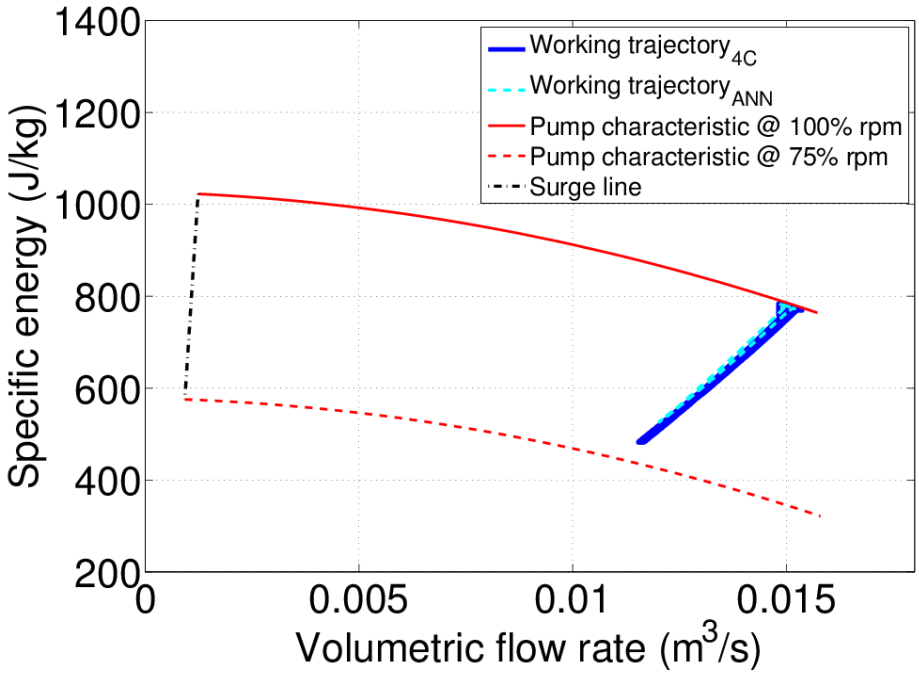


Fig. 17: Pump working trajectory for the best smoothing scenario on the "pseudo-dimensionless" characteristic plane ($\dot{Q}$ - w) for the third pulse, after which the system becomes periodic.

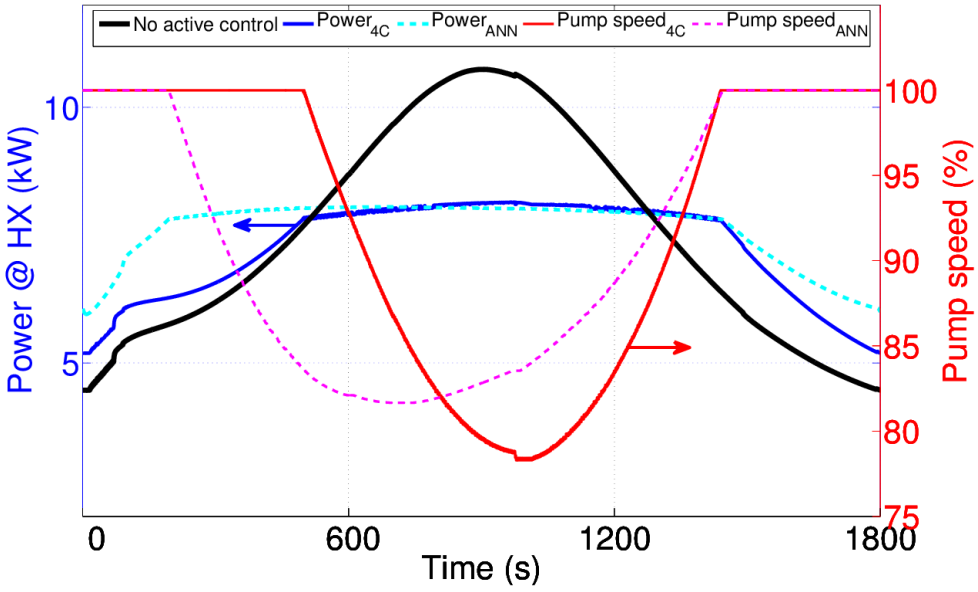


Fig. 18: Comparison of the best smoothing scenario, in terms of power released to the HX (left axis) and relative pump speed (right axis), for the ITER CS, when computed by the 4C model (solid thin lines) and by the ANN model (dashed lines). The solid thick line represents the non-smoothed scenario. The figure shows the third pulse, after which the system becomes periodic.

2.2. *ITER TF coil*

The same rationale adopted for the CS analysis is applied here to the more complex ANN model of the ITER TF coil described above. First, the ANN-based model is tested against the standard operating scenario at 15 MA – 14 kW (the waveforms of the heat load deposited in the WP and the casing during the 1800 s of the pulse are reported in [22]), without any action of the control system.

The results are shown in Fig. 19 in terms of power transferred to the LHe baths of the WP and casing, compared with the full 4C results in the same conditions [24, 25]. The accuracy of the ANN model is very good, with a relative error at peak power ~3% for the WP and ~4.2% for the casing. The error on the total energy released to the LHe baths is ~6% during the first pulse and of ~1.5% during a periodic pulse.

Analyzing the power transferred between WP and casing, see Fig. 20, it can be seen that during the peaks from the casing, almost 300 W are transferred to the WP, which is ~1/3 of the total WP power evacuated to the cryoplant. In the same figure the unbalance in the power transferred (*i.e.* the difference between the power exiting the WP towards the casing in the WP ANN W_C in Fig. 7 and the power entering the casing from the WP in the casing ANN C_C in Fig. 7), in absolute value, introduced by the use of a coupling ANN is also reported: the average unbalance is ~15 W, and its effect on the global energy balance is negligible, of the order of 0.5%.

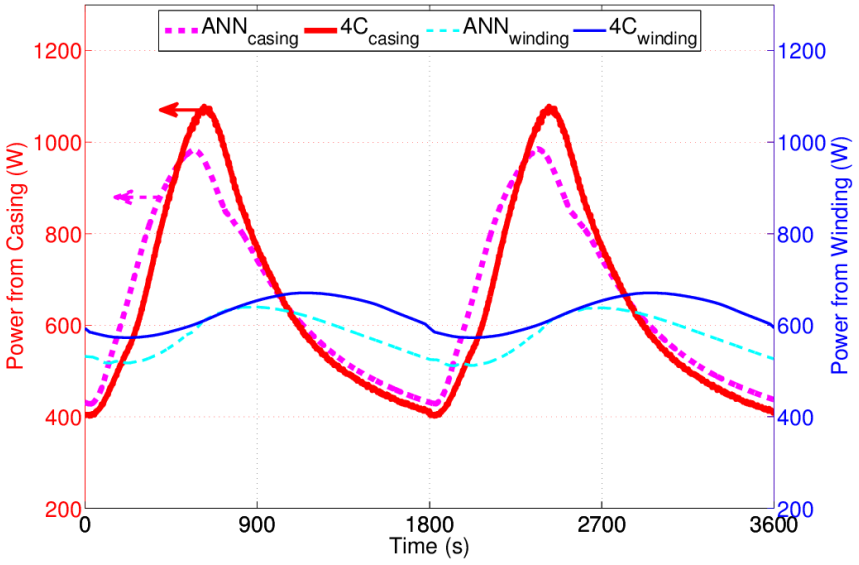


Fig. 19: Comparison of the evolution of the power released at the HXs for the ITER TF winding (thin lines) and casing (thick lines), when computed by the 4C model (solid lines) and the ANN model (dashed lines). The figure shows the third and fourth pulse, after which the system becomes periodic.

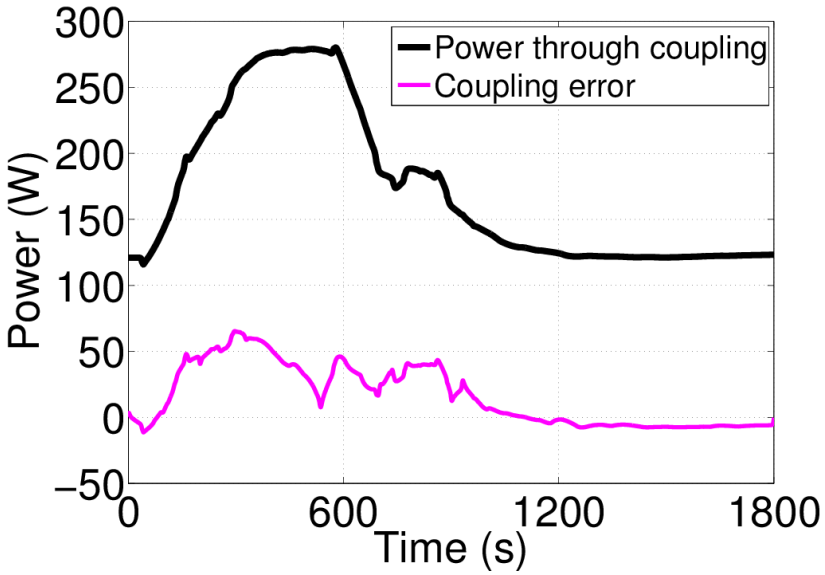


Fig. 20: Comparison of the evolution of the net power transferred between WP and casing ("power through coupling" in the legend) in the ANN simulation (thick line) and power unbalance ("coupling error" in the legend, thin line) in the ANN model for the third pulse, after which the system becomes periodic.

Note that the very small loss in accuracy compared to the 4C code corresponds to a huge gain in speed: the ANN model is ~10^3 times faster than the detailed 4C model, and it can be much faster than real-time basically independently on the hardware used for the simulation.

2.2.1. *Optimization study*

An optimization study performed when the smoothing is done using the bypass valve (BV) in the casing circuit has already been carried out [34], where the control is based on the BV only, differently from what is present in [12], where the bypass of the coil is actuated through a BV together with a valve at the coil inlet. This study is concentrated on the smoothing of the heat load by changing the speed of the pump, again in the casing circuit (see Fig. 5), keeping the same T_i = 100 s used in [34]. The fact that the integrator time constant is much lower than for the CS circuit implies that in this case the integral term has a much larger weight in the controller output evaluation, see above. The results will be compared to those obtained with the use of the BV and with those obtained in [12], where, differently from here, the regulation is based on the power to the cryoplant. The variation of the speed, leading to a change of the mass flow rate through the coil of ~30%, is controlled with a PI controller, which acts again based on the value of the heat transferred to the LHe bath (evaluated as total enthalpy drop through the HX).

The same parametric study done for the ITER CS controller is performed on each of these control parameters to find the most effective value for the load smoothing. The results of this study are reported in Fig. 21, where only the third pulse is shown and the proportional gain K_p is progressively decreased from 7 to 0.2: as expected, this has the

effect of making the response of the regulated variable progressively slower. In the first case (K_p = 7), the pump responds quickly to small variations in the power, leading to oscillations during the transient. In the latter case (K_p = 0.2), the smaller gain has the effect of an averaging filter on the pump speed, allowing a large power increase before the control becomes effective and keeping the pump speed at a low value even when it is not necessary anymore.

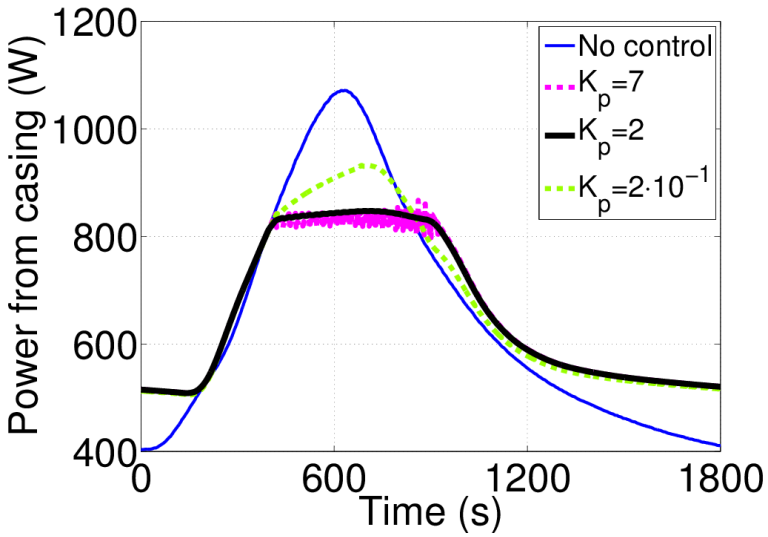


Fig. 21: Results of the parametric study performed for the ITER TF. The figure shows the third pulse, after which the system becomes periodic.

The best smoothing scenario among those considered here corresponds to K_p = 2.

In Fig. 22 a comparison between the two control options for different *SP* is made, showing that from the point of view of the average power transferred to the HX the two strategies looks very similar. At higher *SP* values of the control performed using the pump speed, however, any reduction of the *SP* immediately reflects on the average power, so this second control strategy is more effective in the power smoothing, reflecting the results obtained with the physical model in [16] for the CS. Both in the case of the control performed with the BV and of that performed with the pump, the increase in the coil outlet temperature is negligible and thus is not shown here. In addition, since the power required by the pump is proportional to its speed, this type of control is preferable with respect to the BP valve control as it reduces the pump power consumption, as shown also in [12, 16].

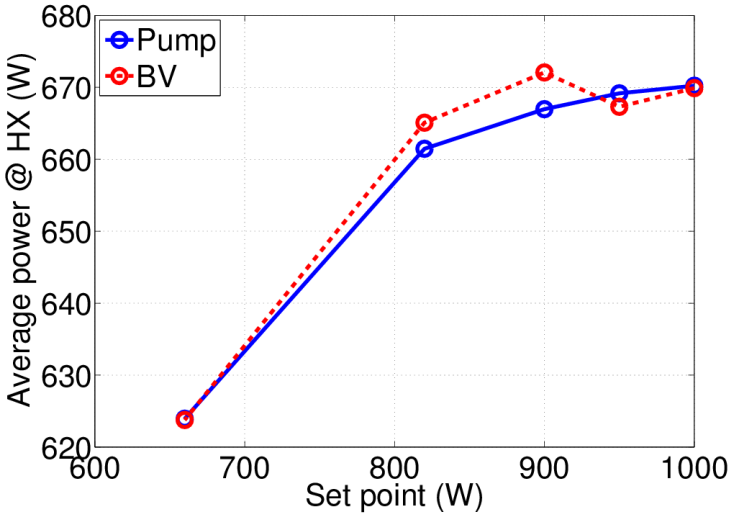
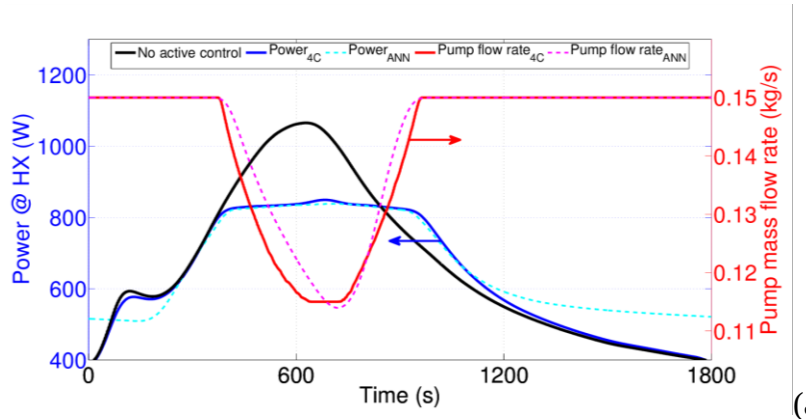


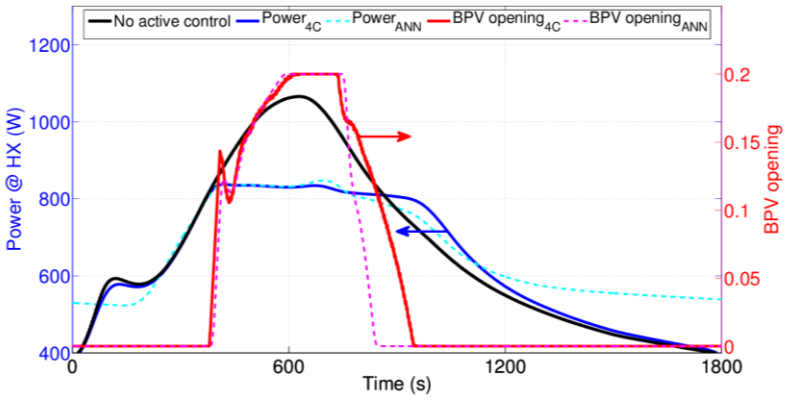
Fig. 22: Parametric study on the effect of the controller *SP* on the average power transferred to the HX of the ITER TF casing during the third plasma pulse, after which the system becomes periodic, when the control is acting on the pump speed (solid line) and on the BV opening (dashed line).

When the best smoothing strategy is chosen, the computed results are compared with those obtained by the 4C code, in terms of accuracy and computational time. The outcome of the comparison is presented in Fig. 23a (control with pump speed) and Fig. 23b (control with the BV). The heat load to the LHe baths computed using the ANN model is quite accurate in both cases during the plateau, with errors on the peak power of ~1.2% in a periodic pulse, but the computational cost is reduced ~1200 times with respect to the 4C code. However, it can be noticed that the ANN model stabilizes at a higher level of power before and after a pulse. This happens because the total (winding + casing) energy extracted from the magnet differs from that in input of ~15% during the first pulse, but only of ~2% during a periodic pulse (when all the power extracted from the magnet is delivered to the two HXs). As for the CS, this can be explained by the small errors introduced by the ANN model in the closed loop dynamic simulation, that let the system simulated by the ANN model reach the equilibrium at a different level with respect to the original 4C model. However, the error on the global energy balance is small, as the total error on the energy released to the LHe baths after 4 pulses is only ~8% with respect to that computed by the 4C model.

The good level of accuracy in the computed heat loads and pump flow rate or BV opening, together with the gain in speed, allows concluding that the ANN model can be further used to develop and test alternative smoothing scenarios in a reliable way, provided the transient stays inside the training range for the ANN input and output variables.



(a)



(b)

Fig. 23: Comparison of the best smoothing scenario, in terms of power released to the HX (left axes) and pump mass flow rate or BV opening (right axes), for the ITER TF casing, when computed by the 4C model (solid thin lines) and by the ANN model (dashed lines): (a) control acting on the pump speed; (b) control acting on the BV opening. The solid thick lines represent the non-smoothed scenario. The figures show the third pulse, after which the system becomes periodic.

3. Conclusions

Two simplified models for the superconducting ITER CS and TF coil have been developed based on ANNs and trained using the results of several 4C code simulations. The design of the ANNs has been performed optimizing the topology and the other degrees of freedom of the network, and the results of the design have been presented.

The ANNs have then been embedded into the detailed 4C model for the cryogenic loops and used as black boxes instead of the 4C detailed model of the TF winding and casing and of the CS modules.

The ANN-based models, despite a small loss in accuracy (always below 8% on the energy conservation), turned out to be ~500 times faster than the full 4C code in the case of the ITER CS coil, and ~10^3 times faster in the case of the ITER TF coil. In both cases, the ANN-based simulation is faster than real-time operation.

Thanks to their speed, the newly-developed ANN models can be used to design and optimized control strategies for the smoothing of the heat load to the cryoplant. In the paper, an example of such an optimization has been shown, for both the CS and TF coils. From this point of view, based on the computed results of the ANN simplified model, it has been discussed for instance how control based on the pump speed has better performances with respect to a control based on the bypass valve, confirming what was expected from the literature.

4. References

[1] [Online]. Available: http://www.iter.org/. [Accessed 29 November 2015].

[2] N. Mitchell, D. Bessette, R. Gallix, C. Jong, J. Knaster, P. Libeyre, C. Sborchia and F. Simon, "The ITER magnet system," *IEEE Trans Appl Supercond 18,* pp. 435-440, 2008.

[3] N. Mitchell, A. Devred, P. Libeyre, B. Lim and F. Savary, "The ITER Magnets: Design and Construction Status," *IEEE Trans. Appl. Supercond. 2012,* vol. 22: 4200809.

[4] L. Gómez Palacín, B. Bradu, E. Blanco Viñuela, R. Maekawa and M. Chalifour, "An optimal control approach for an overall cryogenic plant under pulsed heat loads," in *Proceedings of the 25th International Cryogenic Engineering Conference and the International Cryogenic Materials Conference in 2014, ICEC 25 - ICMC 2014*, Twente, 2014.

[5] H.-S. Chang, A. Forgeas, G. Vincent, R. Maekawa, L. Serio, Y.-M. Park, K.-W. Cho, D.-S. Park, J.-J. Joo, K.-M. Moon and V. Kalinin, "Feasibility Studies of the ITER Cryogenic System at KSTAR," *IEEE Transactions on Applied Superconductivity,* vol. 22, no. 3, p. 4703804, 2011.

[6] R. Vallcorba, B. Rousset, J. Poncet, H. Chang, A. Forgeas, R. Maekawa, L. Serio, P. Bonnay, M. Bon Mardion, A. Girard, C. Hoa, B. Lagier, F. Michel and P. Roussel, "ITER Cryogenics system validation tests at HELIOS test facility," *Adv. Cryo. Eng.*, vol. 57B, pp. 1415-24, 2012.

[7] C. Hoa, M. Bon-Mardion, P. Bonnay, P. Charvin, J.-N. Cheynel, A. Girard, B. Lagier, F. Michel, L. Monteiro, J.-M. Poncet, P. Roussel, B. Rousset and R. Vallcorba-Carbonell, "HELIOS helium loop for high loads smoothing: a cryogenic scaled down experiment of the cooling circuit of JT-60SA superconducting magnets," in *Proceedings of the 23rd International Cryogenic Engineering Conference (ICEC23)*, 2010.

[8] B. Lagier, B. Rousset, C. Hoa and P. Bonnay, "Experimental validation of advanced regulations for superconducting magnet cooling undergoing periodic heat loads," *AIP Conference Proceedings,* vol. 1573, pp. 1602-1609, 2014.

[9] C. Hoa, M. Bon-Mardion, P. Bonnay, P. Charvin, J.-N. Cheynel, B. Lagier, F. Michel, L. Monteiro, J.-M. Poncet, P. Roussel, B. Rousset and R. Vallcorba-Carbonell, "Investigations of pulsed heat loads on a forced flow supercritical helium loop – Part A: experimental set up," *Cryogenics,* vol. 52, pp. 340-348, 2012.

[10] D. Bessette, N. Shatil and E. Zapretilina, "Simulation of the ITER Toroidal Field Coil operation with the

VINCENTA code," *IEEE Trans. Appl. Supercond.,* vol. 16, pp. 795-8, 2006.

[11] B. Lagier, C. Hoa and B. Rousset, "Validation of an EcosimPro® model for the assessment of two heat load smoothing strategies in the HELIOS experiment," *Cryogenics,* vol. 8, pp. 60-70, 2014.

[12] R. Maekawa, S. Takami, A. Iwamoto, H.-S. Chang, A. Forgeas, M. Chalifour and L. Serio, "Process analyses of ITER Toroidal Field Structure cooling scheme," *Cryogenics,* vol. 63, pp. 220-230, 2014.

[13] R. Zanino, R. Bonifetto, C. Hoa and L. Savoldi Richard, "4C modeling of pulsed-load smoothing in the HELIOS facility using a controlled bypass valve," *Cryogenics,* vol. 57, pp. 31-44, 2013.

[14] L. Savoldi Richard, F. Casella, B. Fiori and R. Zanino, "The 4C Code for the Cryogenic Circuit Conductor and Coil modeling in ITER," *Cryogenics,* vol. 50, pp. 167-176, 2010.

[15] R. Maekawa, K. Oba, S. Takami, A. Iwamoto, H.-S. Chang, A. Forgeas, L. Serio, R. Vallcorba, B. Rousset, C. Hoa and L. Monteiro, "Dynamic Simulation of Sub-Scale ITER CS/STR Cooling Loop," *IEEE Transactions on Applied Superconductivity*, vol. 22, no. 3, p. 4704004, 2012.

[16] L. Bottura, F. Gauthier, D. Bessette, A. Devred, R. Maekawa and J. Persichetti, "A physics-based simplified model for the ITER cooling loops," *Submitted to Cryogenics.*

[17] B. Rousset, C. Hoa and B. Lagier, "0-D thermo-hydraulic approach for predicting pressure and temperature along HELIOS SHe loop under pulsed loads," *Cryogenics,* vol. 53, pp. 7-16, 2013.

[18] H. Furci and C. Luongo, "Simplified thermal model of the ITER magnet system," *Cryogenics,* vol. 63, pp. 241-254, 2014.

[19] L. Savoldi Richard, R. Bonifetto, S. Carli, M. Grand Blanc and R. Zanino, "Modeling of Pulsed Heat Load in a Cryogenic SHe Loop using Artificial Neural Networks," *Cryogenics,* vol. 57, pp. 173-180, 2013.

[20] W. T. Miller, R. S. Sutton and P. J. Werbos, Neural Networks for Control, Cambridge (MA): MIT press, 1990.

[21] L. Savoldi Richard, R. Bonifetto, S. Carli, A. Froio, A. Foussat and R. Zanino, "Artificial Neural Network (ANN) Modeling of the Pulsed Heat Load during ITER CS Magnet Operation," *Cryogenics,* vol. 63, pp. 231-240, 2014.

[22] S. Carli, R. Bonifetto, T. Pomella Lobo, L. Savoldi Richard and R. Zanino, "Artificial Neural Network (ANN) Model for the Thermal-Hydraulic Response of a TF Superconducting Magnet in ITER," *Fusion Science and Technology,* vol. 68, no. 2, pp. 336-340, 2015.

[23] S. Carli, R. Bonifetto, L. Savoldi Richard and R. Zanino, "Incorporating Artificial Neural Networks in the dynamic thermal–hydraulic model of a controlled cryogenic circuit," *Cryogenics,* vol. 70, pp. 9-20, 2015.

[24] L. Savoldi Richard, R. Bonifetto, A. Foussat, N. Mitchell, K. Seo and R. Zanino, “Mitigation of the Temperature Margin Reduction due to the Nuclear Radiation on the ITER TF Coils,” *IEEE Transactions on Applied Superconductivity,* vol. 23, p. 4201305, 2013.

[25] L. Savoldi Richard, R. Bonifetto, U. Bottero, A. Foussat, N. Mitchell, K. Seo and R. Zanino, “Analysis of the effects of the nuclear heat load on the ITER TF magnets temperature margin,” *IEEE Transactions on Applied*

Superconductivity, vol. 24, p. 4200104, 2014.

[26] B. C. Csáji., Approximation with Artificial Neural Networks, Faculty of Sciences, Eötvös Loránd University, Hungary, 2001.

[27] G. Cybenko, "Approximations by superpositions of sigmoidal functions," *Mathematics of Control, Signals, and Systems,* vol. 2 (4), pp. 303-314, 1989.

[28] S. Havkin, Neural Networks: A Comprehensive Foundation, Prentice Hall International, 1999.

[29] L. Smith, "A Tutorial on Principal Component Analysis," 2002. [Online]. Available: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. [Accessed February 2015].

[30] A. Kraskov, H. Stögbauer and P. Grassberger, "Estimating Mutual Information," *Physical Review E,* vol. 69, no. 6, 2004.

[31] H. Abarbanel, B. R., J. Sidorowich and L. Tsimring, "The analysis of observed chaotic data in physical systems," *Rev. Mod. Phys.,* vol. 65, p. 1331, 1993.

[32] R. Zanino, R. Bonifetto, R. Heller and L. Savoldi Richard, "Validation of the 4C Thermal-Hydraulic Code against 25 kA Safety Discharge in the ITER Toroidal Field Model Coil (TFMC)," *IEEE Transactions on Applied Superconductivity*, vol. 21, pp. 1948-1952, 2011.

[33] R. Zanino, R. Bonifetto, F. Casella and L. Savoldi Richard, "Validation of the 4C code against data from the HELIOS loop at CEA Grenoble," *Cryogenics,* vol. 53, pp. 25-30, 2013.

[34] A. Froio, R. Bonifetto, S. Carli, A. Quartararo, L. Savoldi and R. Zanino, "Artificial Neural Networks: a viable tool to design heat load smoothing strategies for the ITER Toroidal Field coils," *IOP Conference Series: Materials Science and Engineering,* vol. 101, no. 1, p. Art. ID 012149, 2015.

[35] R. Zanino, R. Bonifetto, C. Hoa and L. Savoldi Richard, "Verification of the Predictive Capabilities of the 4C Code Cryogenic Circuit Model," *AIP Conference Proceedings,* vol. 1573, pp. 1586-1593, 2014.