scientific reports

Check for updates

OPEN

Unsupervised spectra information extraction using physics-informed neural networks in the presence of non-linearities and multi-agent problems

Alessandro Puleio ✉ & Pasqualino Gaudio

Spectroscopy covers a huge range of applications in various fields of science, such as physics, biology, chemistry, engineering, and medicine. In some spectroscopic applications, the data analysis of spectra plays a leading role in the determination of the technique's performance in terms of sensitivity, specificity, and reliability. For this reason, solutions based on machine and deep learning algorithms have been deeply explored as possible alternatives to standard methodologies. Recently, an innovative neural network architecture and training approach have been developed to solve problems where standard supervised deep learning algorithms cannot be used, by exploiting a physics-informed neural network. This new method allows for information extraction from spectra without a supervised approach, i.e. without the need to have controlled experiments where both the spectra and the desired pieces of information to be extracted are known, opening the possibility to solve a huge number of problems where a controlled set (what it is known as training set in machine and deep learning) is present. However, in the previous work, the method has been presented only for simple and linear cases, limiting the range of applications of this new method. In this work, the previous physics-informed deep learning methodology is generalised to tackle both non-linear and multi-agent cases. The methodology, once it has been formally introduced, will be tested on synthetic cases and compared with standard supervised algorithms.

Keywords Unsupervised processing, Physics-Informed neural network, Spectroscopy, Calibration, Deep learning

Passive and active spectroscopic techniques are one of the most used diagnostic methodologies in almost all fields of science, such as physics, biology, chemistry, medicine, and engineering. Nowadays, several examples of the application of different spectroscopy techniques can be possible to cited, starting for example with the fluorescence spectroscopy application to environmental monitoring, such as in the case of the water quality analysis for what concern the dissolved organic matter[1]. Another example can be the use of Raman spectroscopy technique for different scopes, such as: forensic science and security and prevention applications, using this type of techniques it is possible to detect and identify small traces of explosive[2]; medical diagnostic applications, in which Raman spectroscopy permits to classify the nature of the lesion in breast cancer[2–4], or permits the possibility to categorise the state of advancement in the atherosclerosis alterations without the necessity of more invasive approaches[2]. Other applications of the Raman spectroscopy technique are used for the nanotechnology structure's study[2,4], or in environmental applications, such as in wastewater monitoring[2].

In several cases, the accurate data analysis of spectra plays a leading role in the determination of the performance of the spectroscopic technique. That is the case, for example, of Raman spectroscopy, which not only usually requires the detection of low-intensity light, but sometimes other secondary effects, such as fluorescence, may overlap with the Raman emission, making the data analysis very hard to perform. An example of this interference can be the high intensity fluorescence signals obtained by biological sample in the neurodegenerative diseases' studies using Raman spectroscopy, indeed, this type of applications have required data elaborations to

Department of Industrial Engineering, University of Rome "Tor Vergata", Rome, Italy. ✉email: alessandro.puleio@uniroma2.it

clean the Raman spectra by the fluorescence signals, or the used of different Raman techniques to contain or avoid the fluorescence signal, such as the use of near-infrared Raman spectroscopy[4,5].

For this reason, machine and deep learning methodologies have been deeply applied in spectroscopy. In fact, machine and deep learning algorithms allow for feature and information extraction in an automatic and performative way if compared with standard methods. Examples of machine and deep learning in spectroscopy can be found in various fields of science, such as the applications in Raman spectroscopy, in medicine for the recognition of melanoma[6], or the initial evidence for the Alzheimer's disease diagnosis[7] or the type 2 diabetes mellitus diagnosis[8]; or in the security field for the illegal substances detection[9–11]; it is possible to find other applications for fluorescence spectroscopy, for example, in food safety[12,13], or again in medicine for the breast cancer diagnosis[14] or for the hepatocellular carcinoma and liver cirrhosis[15].

However, deep learning has its limits. Supervised learning, that is potentially the most performant, requires large amounts of data where both the inputs (or features) and the output (or target) are known. This labelled dataset, known as the training set (the one used to tune the parameters, or train the model), requires controlled experiments where both the features (or inputs) and the targets (or outputs) must be accurately measured or known a priori[16]. Unfortunately, controlled experiments are not always feasible, limiting the applicability of supervised learning. Deep learning solutions exploiting non-labelled datasets (i.e. datasets where outputs, or targets, are unknown) are known as "unsupervised learning". Unfortunately, purely data-driven unsupervised learning is not so performant as supervised learning, and its use is limited to specific problems such as clustering, feature selection, anomaly detection, etc[17]. Examples of supervised machine and deep learning in spectroscopy are the ones reported by[6,11,18–20]. While unsupervised learning examples are[20–22].

In recent years, a new branch of automatic learning has been developed where physics equations are incorporated into both the neural network architecture and the loss function for the training process. This new branch of automatic learning gives the origin to the Physics-Informed Neural Networks (PINN). Such a new methodology has been applied in different ways, and in particular as an alternative method to solve Partial Differential Equations (PDEs)[23–26] and inverse problem[24–27]. In a previous work, the authors developed a new deep learning architecture to extract information from spectra by using a Physics-informed approach. However, this method was limited to linear cases where only two "agents" were present[28].

This work aims to present a more generalised form of the previous algorithm, where both multi-agent and non-linear problems can be modelled in the Physics-Informed Neural Network. In this study, the spectroscopy application is presented solely as a proof of concept; nevertheless, the proposed methodology has the potential to be fully generalizable to a broad range of applications in which calibration is required. Then, the new methodology is applied to synthetic cases, and the performances are compared with those of standard supervised techniques.

Methods and materials

In this section, the authors reported the details concerning the neural network architecture, the unsupervised physics-informed implementations, and the numerical case generation.

Neural network: losses and architectures

Our previous study[28] has demonstrated how it is possible through an unsupervised physics-informed deep-learning approach to obtain calibrated data starting from raw data. The study demonstrates that by adopting a physics-based methodology and leveraging knowledge of single-agent emission, it is possible to develop a model capable of estimating the concentration of emitted agents without requiring a supervised approach—no labelled targets are needed in the training set. The method has been extensively validated across various types of noise and background, consistently delivering robust and high-performance even under conditions of highly variable background and noise. Incorporating physical principles enables the neural network to perform physics-guided pattern recognition, specifically targeting spectral patterns. As a result, the model can efficiently distinguish between spectra of interest and those dominated by noise and background. Without the integration of physics, an unsupervised model would be unable to achieve such effective decomposition. In essence, the model is supervised by physics rather than by data alone. However, in the previous work, the analysis was limited to a maximum of two agents and linear problems, while in this work, the losses are generalised.

At first, the logic of the training is presented. It is possible to suppose to have a measured spectrum, $I(\lambda)$ that is a function of the wavelength λ. Moreover, imagine that this spectrum is the result of several phenomena. It is possible to define the intensity of the j-th phenomenon c_j and its specific spectrum emission as $I_{0,j}(\lambda)$. For example, if the emission of hydrogen lines is considered, c_j can be thought of as the concentration of hydrogen, while the specific spectrum is the emission of hydrogen lines when c_j equals one. This translates into the possibility of writing:

$$I(\lambda) = \sum_{j=1}^{N_{tot}} c_j I_{0,j}(\lambda) \tag{1}$$

In several cases, taking into account all the phenomena is practically impossible and therefore the previous sum up is simplified as it follows:

$$I(\lambda) = \sum_{j=1}^{N} c_j I_{0,j}(\lambda) + I_b(\lambda) \tag{2}$$

Where the sum takes into account only the phenomena of interests (N), while the rest is left in the "background" spectrum $I_b(\lambda)$.

Starting from this definition, the neural network is thought to work in the following manner. A first neural network part is used to infer the background spectrum $I_{p,b}(\lambda)$, while the second part, working on the difference between the measured spectrum and the predicted background spectrum, aims at predicting the intensities (or concentrations, in our example) of the phenomena. Therefore, by using the knowledge of the specific spectrum (that are supposed to be known) of each phenomenon, one can reconstruct the measured spectrum:

$$I_p(\lambda) = \sum_{j=1}^{N} c_{p,j} I_{0,j}(\lambda) + I_{p,b}(\lambda) \tag{3}$$

Unfortunately, this is a typical ill-posed problem (unknowns are more than constraints), and a closure equation is required. Different closure equations may be used as a function of the problem to solve. A simple and not-so-problem-specific closure is the smoothness of the background. Therefore, the neural network can be trained by minimising the following loss:

$$L_{tot} = L_{rec} + \alpha L_{reg} = \sum \left(I(\lambda) - \sum_{j=1}^{N} c_{p,j} I_{0,j}(\lambda) - I_{p,b}(\lambda) \right)^2 + \alpha \sum \left(\frac{dI_{p,b}}{d\lambda} \right)^2 \tag{4}$$

Where α is a hyperparameter that weights the regularisation loss.

The equation accounts for a multi-agent problem, since there is any limitation to the value of N. However, the method works only for linear problems. However, the implementation of a non-linear problem is straightforward. In fact, given a certain phenomenon, one can write the analytical function of how the intensity varies as a function of the wavelength and the concentration $I_j(\lambda, c_j)$. These analytical functions can then be implemented in the loss function:

$$L_{tot,NL,1} = \sum \left(I(\lambda) - \sum_{j=1}^{N} I_j(\lambda, c_{p,j}) - I_{p,b}(\lambda) \right)^2 + \alpha \sum \left(\frac{dI_{p,b}}{d\lambda} \right)^2 \tag{5}$$

The previous loss function takes into account physical non-linearities but not instrumental. To solve this limitation, another generalization on the loss function can be implemented. Consider the "physics" spectrum, i.e. the ideal one, and not the measured, spectrum:

$$I_{phys} = \sum_{j=1}^{N} I_j(\lambda, c_{p,j}) + I_{p,b}(\lambda) \tag{6}$$

If the instruments/sensors have non-linearities, the physics spectrum will be distorted according with the non-linear function $I(\lambda) = f(I_{phys}(\lambda))$. Implementing this equation in the loss function involves writing:

$$L_{tot,NL,2} = \sum \left(I(\lambda) - f\left(\sum_{j=1}^{N} I_j(\lambda, c_{p,j}) + I_{p,b}(\lambda) \right) \right)^2 + \alpha \sum \left(\frac{dI_{p,b}}{d\lambda} \right)^2 \tag{7}$$

From an architectural point of view, the neural network is organised in two parts (Fig. 1).

The first part of the net has been based on a Convolutional Neural Network (CNN) architecture to predict the background spectrum starting from the raw data in input (layers from L1.1 to L1.4, in Table 1). This part of the network has been structured with the use of the four convolutional layers interspaced with Leaky-ReLU (Leaky Rectified Linear Unit) activation layer, for the first three convolutional layers. The result of the last convolutional layer, which is raised to the second power, then corresponds to the background spectrum.

Then, the reconstructed background spectrum is removed from the measured spectrum. The filtered spectrum is then processed by the second part of the neural network (layers from L2.1 to L2.4, in Table 1), consisting of another convolutional layer followed by an Average-pool layer, which groups the data to obtain average values over the spatial dimensions, interspersed with a Leaky-ReLu layer. In the final part of the neural network, after another Leaky-ReLu layer, a series of two fully connected layers interspersed with a hyperbolic tangent transfer function (Tanh) are followed by the Output Layer that provides the concentration values of each identified agent as a function of the number of agents in the sample.

For all the tests performed in this work, the number of batches is 100 while the number of epochs is 10^4. The training process was performed using the ADAM training algorithm[29], as for the learning rate this was varied at each iteration of the process with an hyperbolic function described below:

$$LR_{(i)} = \frac{10^{-3}}{1 + 10^{-4} * iteration} \tag{8}$$

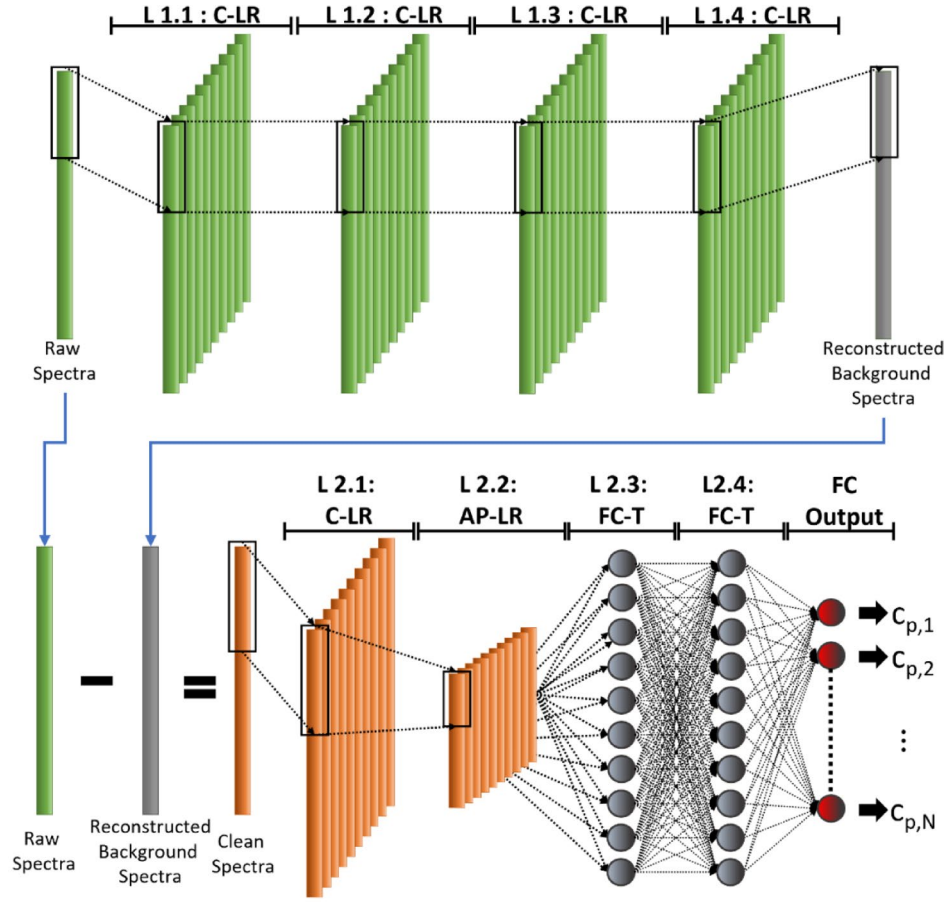

Fig. 1. Physics-Informed Neural Network (PINN) architecture used in this work. The top shows the convolutional neural network part to predict the background spectrum, while the bottom part shows the PINN part with the aim of predicting the concentrations.

	FC Neurons	Kernel	Number of Filters	Pool Size	Stride
L 1.1: C-LR	-	[20 1]	20	-	-
L 1.2: C-LR	-	[20 1]	20	-	-
L 1.3: C-LR	-	[20 1]	20	-	-
L 1.4: C-LR	-	[20 1]	1	-	-
L 2.1: C-LR	-	[50 1]	10	-	-
L 2.2: AveP-LR	-	-	-	50	100
L 2.3: FC-T	20	-	-	-	-
L 2.4: FC-T	20	-	-	-	-
FC-Output	2,4 (As a function of the number of Agents or Classes)	-	-	-	-

Table 1. Structure and hyperparameters of neural Network.

Regarding the hyperparameter α, in this work, a new iterative approach to calculate it is proposed. This iterative determination of the α value is based firstly on the computation of the ratio (β) between the L_{rec}, obtained from each training's iteration ($L_{rec}(i)$), and the initial mean squared error calculated between the raw spectrum and the average of its values as follows:

$$\beta\ (i) = \frac{L_{rec}\,(i)}{\sum \left(I\,(\lambda\,) - I\,(\bar{\lambda}\,) \right)^2} \tag{9}$$

Through the calculation of $\beta\ (i)$ value in each iteration the $\alpha\ (i)$ value has been calculated to improve the performances of the CNN in the background signal reconstruction as follows:

$$\alpha_{\,(i)} = \max(20 * e^{-c*\beta}\,, \epsilon\,) \tag{10}$$

Where c is a hyperparameter, in our case, it has been set to 8.5. Adopting this approach, the neural network allows for obtaining the best reconstruction with the highest α, that is, according to our model and a priori information, the best achievable result.

Alternatives to this approach can be used. Of course, a parametric analysis at constant α may also be performed. However, that is not important for this work, and the reader can refer to the previous work for other details[28].

Synthetic dataset

Data generation and Non-Linearity functions

Different numerical spectra datasets have been simulated to train, validate and test the proposed approach for all the different work steps, developing the analysis in a parametric way:

- DB01, DB02 and DB03 simulate 3 numerical spectra datasets of 10^4 double-agent samples for the training, and 100 samples for the validation, characterised by the absence of nonlinearity phenomena but with an increase in noise level from one dataset to the next.
- DB04, DB05 and DB06, simulate 3 numerical spectra dataset of 10^4 double-agent samples for the training, and 100 samples for the validation, characterized by the presence of two different non-linearity phenomena as function of the concentrations of each one of the two agents (class) simulated, varying the level of noise through the three datasets.
- DB07, DB08 and DB09 simulate 3 numerical spectra datasets of 10^4 double-agent samples for the training, and 100 samples for the validation, characterised by a non-linearity phenomenon for the signal intensities, which can be a peculiarity of a hypothetical instrument of acquisition, varying the level of noise through the three datasets.
- DB10, DB11 and DB12 simulate three datasets, each consisting of numerical spectra of 10^4 four-agent samples for training, and 100 samples for validation. The three datasets differ in the non-linearity phenomena assumed, in particular, in the first dataset there are no non-linearity phenomena, in the second dataset there is the presence of two different non-linearity phenomena caused by concentration (the first and second agent or species simulated have the same non-linearity function as the third and fourth agent or species, which are characterised by a different function), and in the third dataset the non-linearity phenomenon assumed is due to the acquisition sensor. All simulated datasets have the same noise level.

In detail, the simulated spectra have been generated by first generating the reference spectra of each agent that composed the hypothetical samples. In the case of the dataset of numerical spectra that simulate samples composed of two different agents, *a* and *b*, (from DB01 to DB09), the reference spectra I_0 have been calculated as follows, using the parameters noted in Table 2:

$$\begin{aligned} I_{0,class\ a} &= I_{a,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,a,1}\right)^2}{\left(2\sigma_{\,a,1}\right)^2} \right)} + I_{a,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,a,2}\right)^2}{\left(2\sigma_{\,a,2}\right)^2} \right)} \\ I_{0,class\ b} &= I_{b,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,b,1}\right)^2}{\left(2\sigma_{\,b,1}\right)^2} \right)} + I_{b,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,b,2}\right)^2}{\left(2\sigma_{\,b,2}\right)^2} \right)} \end{aligned} \tag{11}$$

.

For what concern the cases of the dataset of the samples with four agents (or agent), the equations used to simulate the reference spectra are the following and the parameters are shown in Table 3:

$$\begin{aligned} I_{0,class\ a} &= I_{a,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,a,1}\right)^2}{\left(2\sigma_{\,a,1}\right)^2} \right)} + I_{a,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,a,2}\right)^2}{\left(2\sigma_{\,a,2}\right)^2} \right)} \\ I_{0,class\ b} &= I_{b,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,b,1}\right)^2}{\left(2\sigma_{\,b,1}\right)^2} \right)} + I_{b,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,b,2}\right)^2}{\left(2\sigma_{\,b,2}\right)^2} \right)} \\ I_{0,class\ c} &= I_{c,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,c,1}\right)^2}{\left(2\sigma_{\,c,1}\right)^2} \right)} + I_{c,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,c,2}\right)^2}{\left(2\sigma_{\,c,2}\right)^2} \right)} \\ I_{0,class\ d} &= I_{d,1} e^{-\left(\frac{\left(\lambda - \lambda_{\,d,1}\right)^2}{\left(2\sigma_{\,d,1}\right)^2} \right)} + I_{d,2} e^{-\left(\frac{\left(\lambda - \lambda_{\,d,2}\right)^2}{\left(2\sigma_{\,d,2}\right)^2} \right)} \end{aligned} \tag{12}$$

Parameters	Non-Linearity Absence			Concentration Non-Linearity			Sensor Non-Linearity		
Dataset	DB01	DB02	DB03	DB04	DB05	DB06	DB07	DB08	DB09
$I_{a,1}$	180	180	180	180	180	180	180	180	180
$I_{a,2}$	60	60	60	60	60	60	60	60	60
$I_{b,1}$	180	180	180	180	180	180	180	180	180
$I_{b,2}$	60	60	60	60	60	60	60	60	60
$\lambda_{a,1}$	656.3	656.3	656.3	656.3	656.3	656.3	656.3	656.3	656.3
$\lambda_{a,2}$	486.1	486.1	486.1	486.1	486.1	486.1	486.1	486.1	486.1
$\lambda_{b,1}$	587.5	587.5	587.5	587.5	587.5	587.5	587.5	587.5	587.5
$\lambda_{b,2}$	480	480	480	480	480	480	480	480	480
$\sigma_{a,1}$	2	2	2	2	2	2	2	2	2
$\sigma_{a,2}$	2	2	2	2	2	2	2	2	2
$\sigma_{b,1}$	2	2	2	2	2	2	2	2	2
$\sigma_{b,2}$	2	2	2	2	2	2	2	2	2
Noise Level	1	5	10	1	5	10	1	5	10

Table 2. Values used to generate the reference spectra and subsequently the two-agent's spectra in each dataset for the different cases analysed in this work.

Parameters	Non-Linearity Absence		Concentrations Non-Linearity		Sensor Non-Linearity	
Dataset	DB10		DB11		DB12	
	$n=1$	$n=2$	$n=1$	$n=2$	$n=1$	$n=2$
$I_{a,n}$	180	60	180	60	180	60
$I_{b,n}$	240	100	240	100	240	100
$I_{c,n}$	150	500	150	500	150	500
$I_{d,n}$	600	300	600	300	600	300
$\lambda_{a,n}$	656.3	486.1	656.3	486.1	656.3	486.1
$\lambda_{b,n}$	587.5	480	587.5	480	587.5	480
$\lambda_{c,n}$	411.3	445.5	411.3	445.5	411.3	445.5
$\lambda_{d,n}$	445.5	645.6	445.5	645.6	445.5	645.6
$\sigma_{a,n}$	2	2	2	2	2	2
$\sigma_{b,n}$	2	2	2	2	2	2
$\sigma_{c,n}$	2	2	2	2	2	2
$\sigma_{d,n}$	2	2	2	2	2	2
Noise Level	5		5		5	

Table 3. Values used to generate the reference spectra and subsequently the four-agent's spectra in each dataset for the different cases analysed in this work.

Starting from these reference spectra, the dataset's spectra have been calculated using a random concentration value, for each agent, ranging from 0 to 2, and summing a background spectrum. Indeed, for all the simulated cases, the background spectra have been computed with the following equation, as in the examples in the literature[28]:

$$I_{back(\lambda)} = 10 r_1 e^{-\left(\frac{(\lambda - 500 - 200 r_2)^2}{2*75^2}\right)} \tag{13}$$

Where r_1 is a random value between 1 and 6, and r_2 is a random value between 0 and 1.
Then, each spectrum for the dataset has been calculated as follows:

$$I = I_{back(\lambda)} + c_a * I_{0,a} + c_b * I_{0,b} \quad case\ that\ simulate\ a\ double - class\ sample \tag{14}$$

$$I = I_{back(\lambda)} + c_a * I_{0,a} + c_b * I_{0,b} + c_c * I_{0,c} + c_d * I_{0,d} \quad case\ that\ simulate\ a\ four - class\ sample \tag{15}$$

The previous two equations have been used to simulate the dataset without non-linearity phenomena, respectively, from DB01 to DB03 and DB10 (in Tables 2 and 3).

While on the possibility of compensating for the proposed method, the possible presence of non-linearity phenomena that can affect the raw data, the authors decided to introduce a different type of this phenomenon. In particular, the first type of non-linearity phenomenon that has been taken into account is the possibility of a non-linear relation between the acquired signals and the concentration of a hypothetical agent. So, for what concern the dataset from DB04 to BD06 two different types of equations (the first one describes a sigmoid, while the second one describes a double sigmoid) has been applicated to simulate this type of phenomenon, that it is usually observed in conditions where the absorption component increasing arriving quickly to saturations' phenomena:

$$c_{NL,\ a} = 2 * \left(\frac{e^{5*c_a - 5}}{1 + e^{5*c_a - 5}} \right) \tag{16}$$

$$c_{NL,b} = 2.23c_b^5 - 10.02c_b^4 + 14.71c_b^3 - 7.32c_b^2 + 1.60c_b - 2 * 10^{-11} \tag{17}$$

Where the numeric values indicated in the equation has been arbitrary decided by the authors to obtain the desired non-linearity function, while c_a and $c_b$are the random concentration value ranging from 0 to 2, and the obtained values have been used to create the numerical datasets of double-agent samples, whose data exhibit the nonlinearity phenomenon under consideration. While in the case of DB11 that collects the data of four-agent samples for what concern the non-linearity concentration-signal for agents a and c the Eq. (16) has been applied, and for agents b and d the Eq. (17) has been applied.

The second type of non-linearity phenomenon that has been considered is the possibility of non-linear response in the instrument used for the measurements. To simulate this type of occurrence the authors have decided to simulate a saturation affects as follow:

$$I_{SensorNL} = 500 * \left(\frac{e^{0.01*I+0}}{1 + e^{0.01*I+0}} \right) - 250 \tag{18}$$

Where the variable I is the spectrum calculated using the Eq. (14) in the case of the double-agent dataset (from DB07 to DB09), or using the Eq. (15) in the case of the dataset with four-agent samples DB12.

In order to obtain the numerical spectra that simulate similar-real conditions of information acquisition, after the addition of the background spectrum to each simulated spectrum, a noise distortion has been added. In particular the noise signal has been calculate has a random normal distribution that had as average the results of spectra and background sum and a variable standard deviation, between 1 and 10 (shown in Tables 2 and 3), to get the spectra dataset through to make the parametric study for the algorithm performances.

Results

Following the idea of a parametric test for the CNN proposed algorithm, for each dataset of spectra and the relative simulated condition (Paragraph 2.2.1), the authors have trained a CNN to perform the test phase. To complete the tests' phase and analyse the performances of CNNs for each one of the datasets reported in Paragraph 2.2.1 have been created a test similar-conditions dataset, which contains each one 1000 different spectra to test the trained CNNs, as follows:

- DB01-T, DB02-T and DB03-T, which are the test datasets of two-agent samples, of 10^3 spectra each one, with the absence of Non-Linearity phenomena, and parametric variation of noise level (first three columns of Table 4).
- DB04-T, DB05-T and DB06-T, which are the test datasets of two-agent samples, of 10^3 spectra each one, with concentrations' Non-Linearity phenomena, and parametric variation of noise level (second three columns of Table 4).
- DB07-T, DB08-T and DB09-T, which are the test datasets of two-agent samples, of 10^3 spectra each one, with sensor's Non-Linearity phenomenon, and parametric variation of noise level (third three columns of Table 4).
- DB10-T, that is the test's dataset of spectra of four-agent 10^3 samples, with the absence of Non-Linearity phenomena, and fixed noise level (first three columns of Table 5).
- DB11-T, that is the test's dataset of spectra of four-agent 10^3 samples, with concentrations' Non-Linearity phenomena (Non-Linearity function equal for the first and third agent, Eq. (6, and a different Non-Linearity function equal for the second and fourth agent, equation), and fixed noise level (first three columns of Table 5).
- DB12-T, that is the test's dataset of spectra of four-agent 10^3 samples, with sensor's Non-Linearity phenomena, and fixed noise level (first three columns of Table 5).

The results obtained by the different tests have been listed in Tables 4 and 5. The neural network performances are estimated with two metrics: the first one is the R-squared between the reconstructed spectra and the measured one. This metric is useful to understand: (a) if the neural network converges at good values or if α is too large/small (in fact, an R-squared compatible with spectra noise is expected); (b) if the result for the specific measured

Characteristics and Statistical Indicators for Prediction	Non-Linearity Absence			Concentrations Non-Linearity			Sensor Non-Linearity		
	DB01-T	DB02-T	DB03-T	DB04-T	DB05-T	DB06-T	DB07-T	DB08-T	DB09-T
Test Dataset Dimension	1000	1000	1000	1000	1000	1000	1000	1000	1000
Random Noise Lv	1	5	10	1	5	10	1	5	10
Alpha Value	99.064	81.456	47.310	99.263	84.313	55.058	99.101	80.291	45.799
R^2Spectra Reconstruction (with Noise)	99.679%	92.810%	77.002%	99.776%	94.960%	82.837%	99.679%	92.572%	76.712%
R^2Spectra Reconstruction (with Noise Suppression)	99.998%	99.976%	99.912%	99.998%	99.975%	99.934%	99.998%	99.974%	99.925%
MSE of Agent a Concentration	9.09E-06	1.49E-04	6.16E-04	5.28E-04	4.24E-03	7.13E-03	2.45E-05	3.71E-04	1.22E-03
MSE of Agent b Concentration	1.42E-05	1.57E-04	6.61E-04	2.30E-04	1.39E-03	3.57E-03	2.64E-05	3.61E-04	1.22E-03
Correlation of Agent a Concentrations (R^2)	99.998%	99.955%	99.817%	99.859%	99.129%	98.431%	99.993%	99.887%	99.632%
Correlation of Agent b Concentrations (R^2)	99.997%	99.951%	99.807%	99.930%	99.595%	99.041%	99.992%	99.895%	99.632%

Table 4. Parametric tests for Two-Agent dataset with different conditions of Non-Linearity phenomena.

Characteristics and Statistic Indicators for Prediction	Non-Linearity Absence	Concentrations Non-Linearity	Sensor Non-Linearity
	DB10-T	DB11-T	DB12-T
Test Database Dimension	1000	1000	1000
Random Noise Lv	5	5	5
Alpha Value	98.172	91.813	91.508
R^2Spectra Reconstruction (with Noise)	98.848%	99.358%	94.584%
R^2Spectra Reconstruction (with Noise Suppression)	99.992%	99.988%	99.965%
MSE of Agent a Concentrations	1.77E-04	2.62E-03	3.60E-04
MSE of Agent b Concentrations	1.01E-04	1.58E-03	2.91E-04
MSE of Agent c Concentrations	7.96E-05	1.99E-03	2.91E-04
MSE of Agent d Concentrations	4.82E-05	7.59E-04	5.37E-04
Correlation of Agent a Concentrations (R^2)	99.945%	99.309%	99.895%
Correlation of Agent b Concentrations (R^2)	99.971%	99.540%	99.915%
Correlation of Agent c Concentrations (R^2)	99.977%	99.470%	99.912%
Correlation of Agent d Concentrations (R^2)	99.985%	99.773%	99.840%

Table 5. Parametric tests for Four-Agent database with different conditions of non-linearity phenomena.

spectrum is reliable. If the reconstruction is not good, potentially the neural network has not been trained for this spectrum. So, it is a measurement of the reliability of the prediction, especially for new data. This R-squared is calculated as follows:

$$R^2_{Recostruction} = \left(1 - \frac{L_{rec}}{\bar{\sigma}^2_{I_\lambda}}\right) * 100 \tag{19}$$

Where L_{rec} comes from Eq. (4. The second R-squared can be calculated only with labelled data, and it is used just to evaluate the performances of the algorithm proposed in this work. This is the R-squared between the predicted and the actual concentration. Both the overall and the agent-specific R-squared have been calculated.

Considering the parametric tests in Table 4, it is possible to see how the tests on 1000 two-agent simulated samples in different conditions, for what concern the Non-Linearity phenomena's presence, have demonstrated good performances. In particular, observing all the R-squares of spectra reconstruction with noise (Table 4), it is possible to see how the decrease of performances corresponds to the increase of noise, in every Non-Linearity Condition (starting from the absence to the concentrations Non-Linearity, finishing to the Sensor's Non-Linearity). But in all the cases the R-squared values are always greater than 75%.

Indeed evaluating the plots *a*, *c* and *e* (for the Concentration Non-Linearity phenomenon increasing the noise) and the plots *b*, *d* and *f* (for the Sensor Non-Linearity phenomenon increasing the noise) in Fig. 2, to compare a reconstructed spectra (using predicted values and reference spectra) and the original data gave to the CNN with the noise signal, it is possible to observe how the principal differences between the reconstructed data (PINN output in red) and the original-one (PINN input data in blue) is increased by the increasing of noise

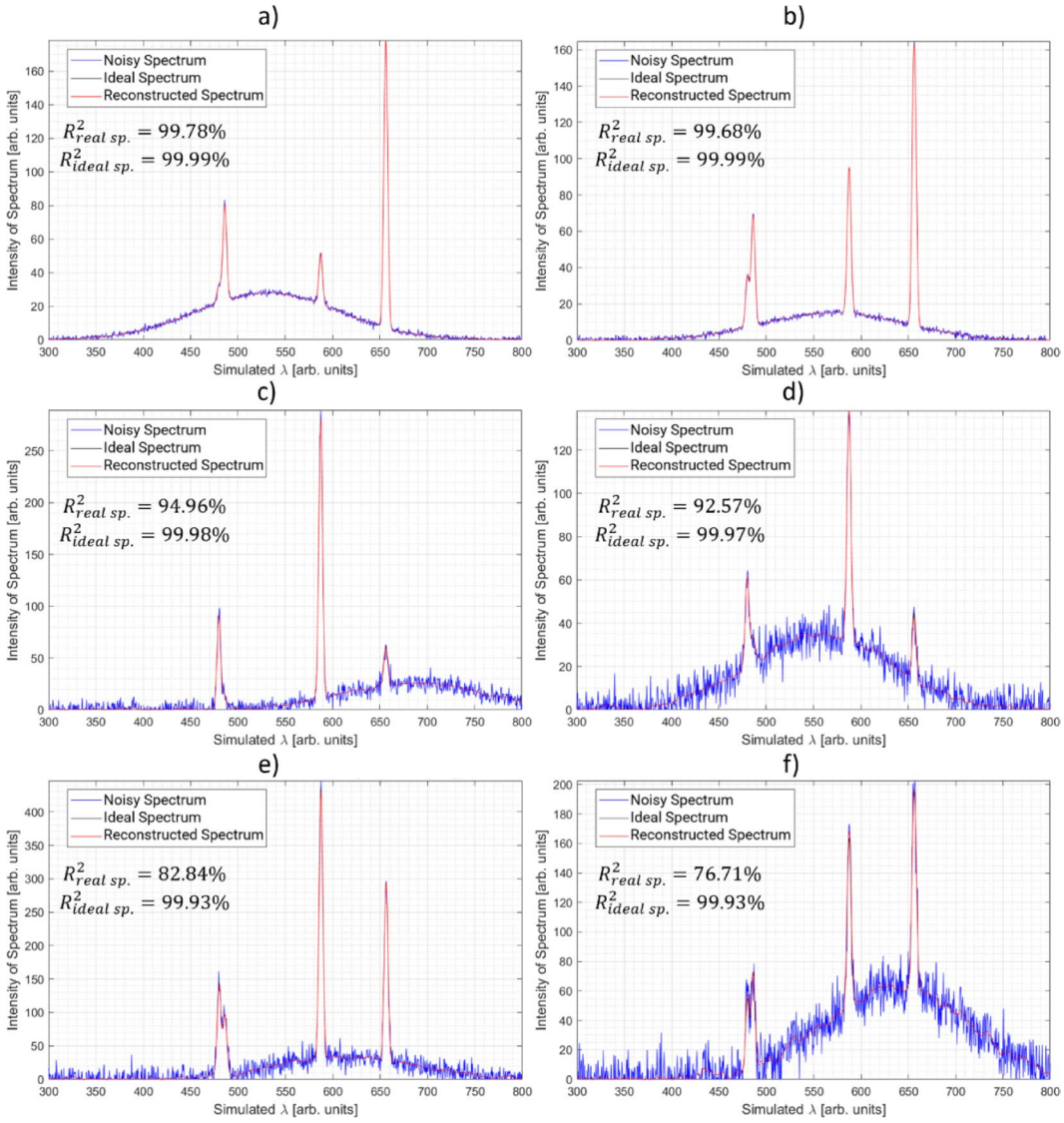

Fig. 2. The plots of the figure show the comparison between an original data of dataset (in blue) gave to the CNN, the correspondent reconstructed spectrum using the predicted concentration's values, of two agent, and the predicted background signals (in red) and the ideal spectrum expected without the noise (in black), for what concern the double-agent tests. In the left column has been reported the results of Concentration Non-Linearity tests: plots ***a***, ***c***, and ***e***. In the right column has been reported the results of Sensor Non-Linearity tests: plots ***b***, ***d***, and ***f***.

signals. In the same plots are also shown the ideal spectra (with noise suppression in the input data, in black), and it is possible to see how there is an almost complete overlap between the reconstructed spectra and the ideal one (as also demonstrated by the "R^2 Spectra Reconstruction (with Noise Suppression)" values in Table 4 that are every time greater than 99.9%).

It must be noted that the model reconstructs the background and the "physical" spectra (in our case the spectral lines), while random noise is smoothed out. This was expected since the net is asked to reconstruct the background with low frequency variability (imposed by the second term of Eq. (4)) and the physical spectra imposed by the knowledge of the dataset. Therefore, noise is automatically smoothed out, involving that this method can be used also for automatic denoising. Concerning the overfitting problem, this would be possible only in case of very small α (Eq. (10)), i.e. when the regularisation term goes to zero, however this phenomenon is avoided by the introduction of a rule for the automatic update of this hyperparameter (as explained in the paragraph 2.1).

Observing the R-square values of reconstruction without the noise, in Table 4, appears that the concentrations' values and the background signal predicted by the CNN, used to reconstruct an example spectra, gave good results in comparison with the noisy data used to test the algorithms in all cases. In every case, starting from the absence of Non-Linearity phenomena, passing through the Concentration and Sensor Non-Linearity phenomena, with the increasing noise level, the R-square results are always major than 99.9%. These high resemblances between

the reconstructed spectra and the original-one it are also reported in plots *b*, *d*, and *f* for the Concentration Non-Linearity phenomenon, and in plots *h*, *j* and *l* for the Sensor Non-Linearity phenomenon, where the authors have compared the reconstructed spectra with the noisy data and with noise suppression.

Evaluating the correlation results between the original concentration values and the predicted one, for each double-agent tests, reported in Table 4, the values of correlation for the concentration prediction appears always over the 98%, with a trend that decreasing with the increasing of noise levels, as observed for the R-square of spectra's reconstruction. In detail, the tests increasing the noise level with the absence of Non-Linearity phenomena show correlation values greater than 99.8%, for the concentration predictions of each simulated agent, with a decreasing trend as the noise level increases (Table 4). In Fig. 3, it is possible to observe the couple of concentrations prediction results versus the original values of simulated data, for each simulated agent, in the Concentration Non-Linearity testes (on the left column) and in the Sensors Non-Linearity tests (on the right column). The plots on the left column of Fig. 3 show the concentration values predicted versus the expected for both the simulated substances for each level of noise tested for Concentration's Non-Linearity tests (plots *a* and *b* for Noise Lv = 1; plots *c* and *d* for Noise Lv = 5; plots *e* and *f* for Noise Lv = 10); as the noise level increases, it is possible to see how the error committed by the algorithm in predicting the concentration value of each substance increases, while maintaining an average correlation value calculated on all samples of over 98%. On the right column of Fig. 3 the plots show the same type of results, for each level of noise tested, for Sensor's Non-Linearity tests (plots *g* and *h* for Noise Lv = 1; plots *i* and *j* for Noise Lv = 5; plots *k* and *l* for Noise Lv = 10); in these tests it is possible to see how the noise increasing did not amplify the error in the concentrations' predictions and the correlation values for all the simulated substances in every tests remain always over the 99.5%.

In all the plots of Fig. 3, both in the case of non-linearity linked to concentrations and in the case of non-linearity of the measuring sensor, it is important to note the presence of a dispersion about the target vs.

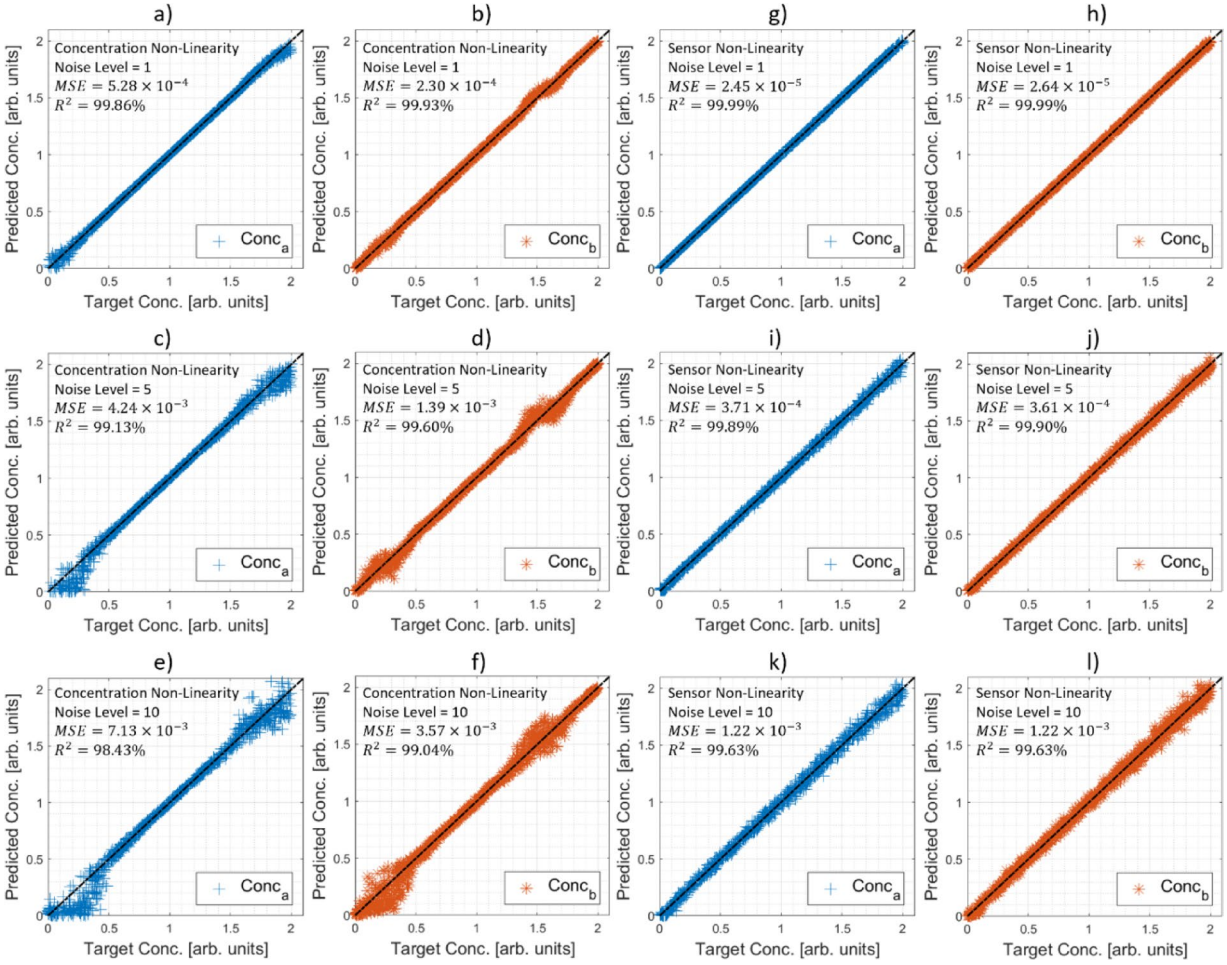

Fig. 3. The plots of the figure show the comparison between an original concentration's values of database gave to the CNN and the correspondent predicted concentration's values. On the left column has been reported the results of Concentration Non-Linearity tests that show the comparison between the predicted concentration's values and the original-one for what concern the Agent ***a*** (blue) and Agent ***b*** (orange), increasing the noise level from 1 to 10. On the right column has been reported the same results for Sensor Non-Linearity tests that show the comparison between the predicted concentration's values and the original-one for what concern the Agent ***a*** (blue) and Agent ***b*** (orange), increasing the noise level from 1 to 10.

predicted concentration values that increase with the increasing of noise. These portions of more uncertainty correspond to the high complex portions of the non-linearity functions, more appreciable especially for the concentration's non-linearity functions. This phenomenon is expected, especially increasing the noise, but it could be possible to resolve and decrease the uncertainty by increasing the amount of data used during model training and validation.

To have a comparison term the authors decided to perform in addiction a supervised training and test for the concentration determination using the same CNN algorithm. Considering the case in which has been obtained the lowest results, in terms of MSE for agents a and b concentrations determination, and in terms of correlation (R^2) between the expected and predicted concentration's values for each agent, it has been performed a training a test process using respectively the dataset DB06 and DB06-T through a supervised approach. The results obtained has been, as expected, better than the unsupervised approach, but at the same time extremely comparable. Indeed, for what concern this test the correlation result for the concentrations of agent a has been major than the result obtained by the unsupervised trained CNN (respectively 98.827% versus 98.431%), while for the agent b the result has been lower (respectively 98.913% vs. 99.041%). Demonstrating as the unsupervised trained algorithm had good performances in comparison with the always preferred supervised trained algorithm approach.

For what concern the four-agent testes every test's database has been created with an equal level of noise varying only the non-linearity phenomenon that characterized the data. Table 5 shows the results of this series of tests, starting from the Non-Linearity Absence, the Concentrations' Non-Linearity and finishing with the Sensor's Non-Linearity. Observing the results shown in Table 5 it is possible to see how in Non-Linearity Absence condition the R-square of reconstruction is equal to 98.48% (Table 5; Fig. 4a), comparing the reconstructed spectra, with predicted values, against the original-one, with the noise component, while performing the comparison with the noise suppression the R-square is equal to the 99.99% (Table 5; Fig. 4a), that demonstrates how the algorithm in the four-agent test obtained good results in terms of predicting agent concentration values

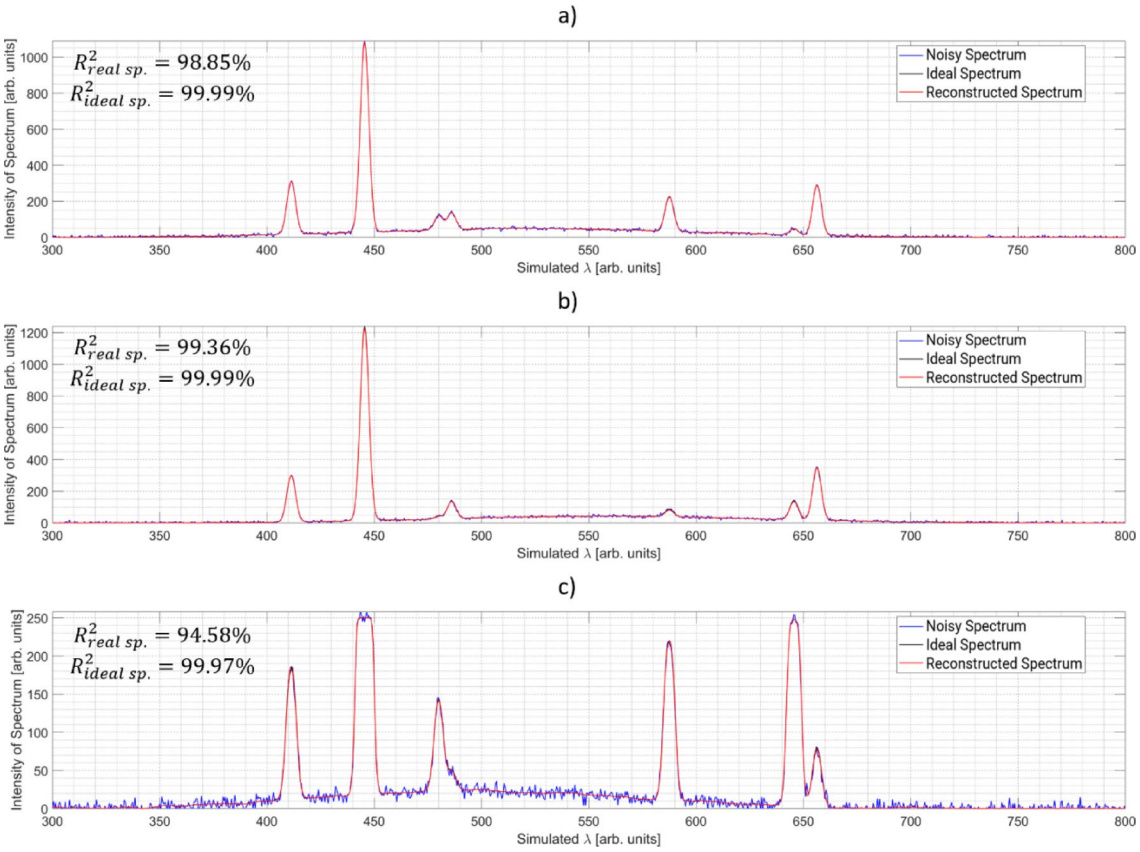

Fig. 4. The plots of the figure show the comparison between an original data of database (in blue) gave to the CNN, the correspondent reconstructed spectrum using the predicted concentration's values, of four agent, and the predicted background signals (in red) and the ideal spectrum expected without the noise (in black), for what concern the four-agent tests. The first plot (***a***) shows the comparison between a real noising spectrum (in blue), the ideal spectrum (in black), and the reconstructed-one (in red) in absence of non-linearity phenomena. The second plot (***b***) shows the comparison between a real noising spectrum (in blue), the ideal spectrum (in black), and the reconstructed-one (in red) with Concentrations' Non-Linearity phenomena. The third plot (***c***) shows the comparison between a real noising spectrum (in blue), the ideal spectrum (in black), and the reconstructed-one (in red) with Sensor's Non-Linearity phenomena.

and background signal, net of the non-predictable random noise component. The same trend of results has been observed in the multi-agent tests with non-linearity phenomena simulated:

- In the simulated case of Concentrations' Non-Linearity phenomena, the R-square of reconstruction, with the noise component, it is around 99.36% (Table 5; Fig. 4b), while comparing the reconstructed data with noise component suppression the R-square is of 99.99% (Table 5; Fig. 4b).
- In the simulated case of Sensor's Non-Linearity phenomenon, the R-square of reconstruction, with the noise component, it is around 94.58% (Table 5; Fig. 4c), while comparing the reconstructed data with noise component suppression the R-square is of 99.97% (Table 5; Fig. 4c). Good results obtained despite the proximity of the reference peaks used for numerical simulations (reported in Table 3, for agent c and d) and their consequent amalgamation due to the widening and flattening of the signal at the top due to the non-linearity phenomenon simulated on the sensor. A phenomenon that the proposed method effectively counteracts.

Table 5 also shows the correlation values between the predictions of the concentration value of each agent simulated in the sample and the values with which the spectra were simulated.

In the first test, without non-linearity phenomena that could affect the data, the correlation values for what concern the concentrations' prediction is always greater than the 99.9%. The first four-agent test graphic results of these comparisons between the predicted values and the original-one for each substance has been shown in Fig. 5, in which results possible to observe how the uncertainty in the concentration prediction is low and comparable with the results observed in literature[28].

Considering the second four-agent test, with the simulation in the data of the presence of two Concentration Non-Linearity phenomena, it can be observed in Table 5 that the correlation values for the prediction of concentrations are slightly lower than those obtained in the absence of such phenomena, as expected, although they remain for all substances above 99%. Indeed, observing the graphic representations of the prediction results in comparison with the real-one, in Fig. 6 it is possible to see how the uncertainty and the error committed in the concentrations' prediction results are quite present in certain range of concentration's values.

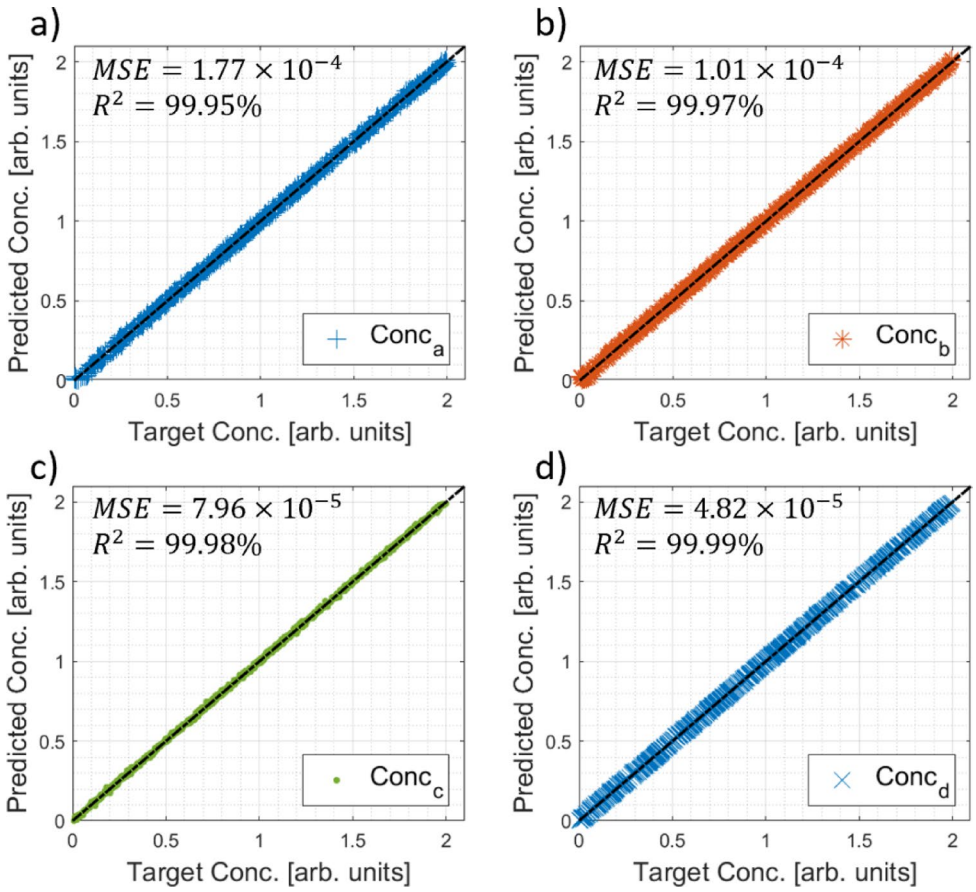

Fig. 5. The plots of the figure show the comparison between an original concentration's values of database gave to the CNN and the correspondent predicted concentration's values, for each one of the four substances simulated, in condition of Absence of Non-Linearity phenomena. In the first plot the Agent a concentration's values comparison has been reported (***a***), in the second plot the Agent b concentration's values comparison has been reported (***b***), in the third plot the Agent c concentration's values comparison has been reported (***c***), and in the fourth one the Agent d concentration's values comparison has been reported (***d***).

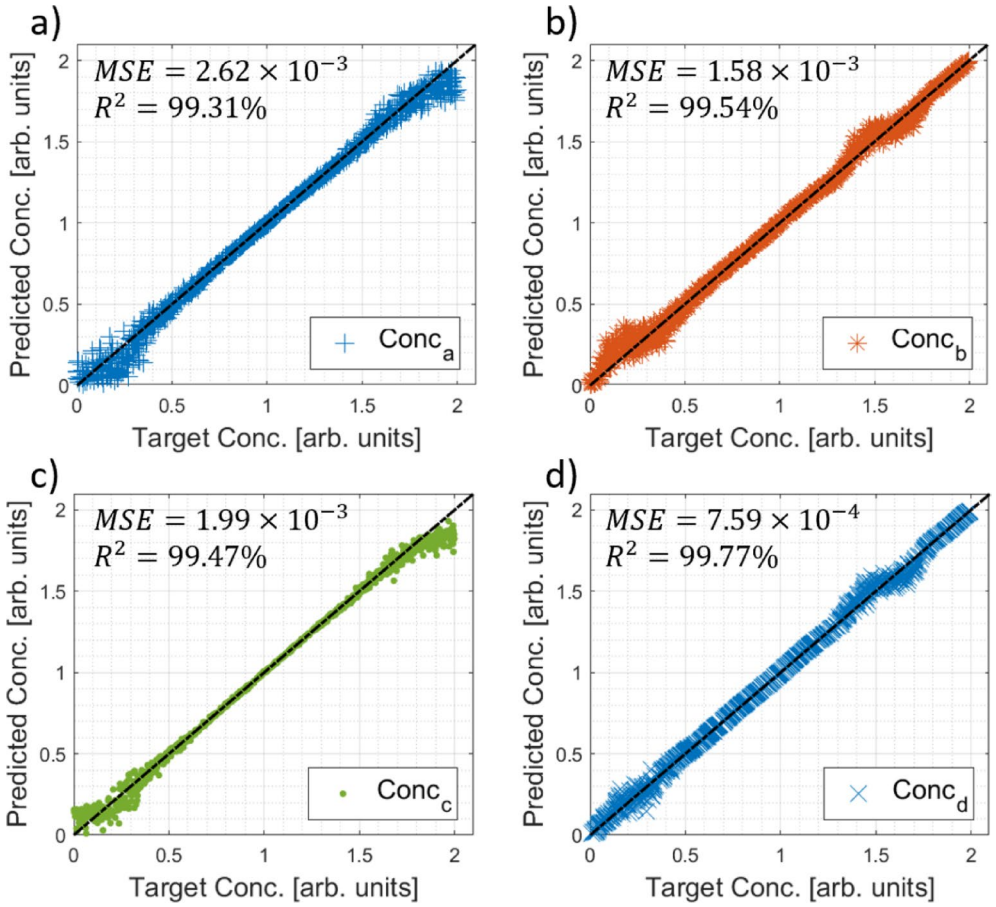

Fig. 6. The plots of the figure show the comparison between an original concentration's values of database gave to the CNN and the correspondent predicted concentration's values, for each one of the four substances simulated, in condition of Concentrations' Non-Linearity phenomena. In the first plot the Agent a concentration's values comparison has been reported (***a***), in the second plot the Agent b concentration's values comparison has been reported (***b***), in the third plot the Agent c concentration's values comparison has been reported (***c***), and in the fourth one the Agent d concentration's values comparison has been reported (***d***).

On the other hand, observing the results of the third test, with simulated Non-Linearity phenomena due to the Sensor, it is possible to see how the correlation values for the prediction of the concentrations of each substance are always greater than 99.8%. This evidence is appreciable if one observes the comparison graphs in the Fig. 7, where it is shown how the uncertainty in the prediction of concentration, and therefore the error committed, is more contained than if in the presence of non-linearity phenomena due to concentration, but slightly present in comparison to the complete absence of such phenomena.

Discussion and conclusion

In this work, the generalization of the Unsupervised PINN for spectra analysis has been presented. The loss definition allows for analysing non-linear problems and multi-agent problems (without limitation on the number of agents). The new methodology has been tested with synthetic cases, and the performances are very high, comparable with the supervised algorithms.

The algorithm presented in this work demonstrates a robust capability for extracting valuable information from raw spectroscopic data without requiring preprocessing. This is achieved by eliminating background signals that can compromise measurements, even in the presence of known nonlinear phenomena arising from the sample's nature (e.g., signal-to-concentration ratios in fluorescence techniques) or instrumentation issues (e.g., sensor saturation).

As shown, this approach can be effectively applied to multi-agent problems and holds promise for a wide range of spectroscopic applications. Indeed, the use of a PINN consist in a unsupervised physics-informed deep learning approach that do not need to know the searched/target information during its training and validation phase of preparation. The searched/target information normally used for the supervision during the supervised model training are replaced by the integration of the physical equations describing the processes of interest. By its very nature, this method requires a complete understanding of the physics underlying the spectroscopic phenomenon to which it is intended to be applied, and knowledge of any non-linear phenomena that may affect

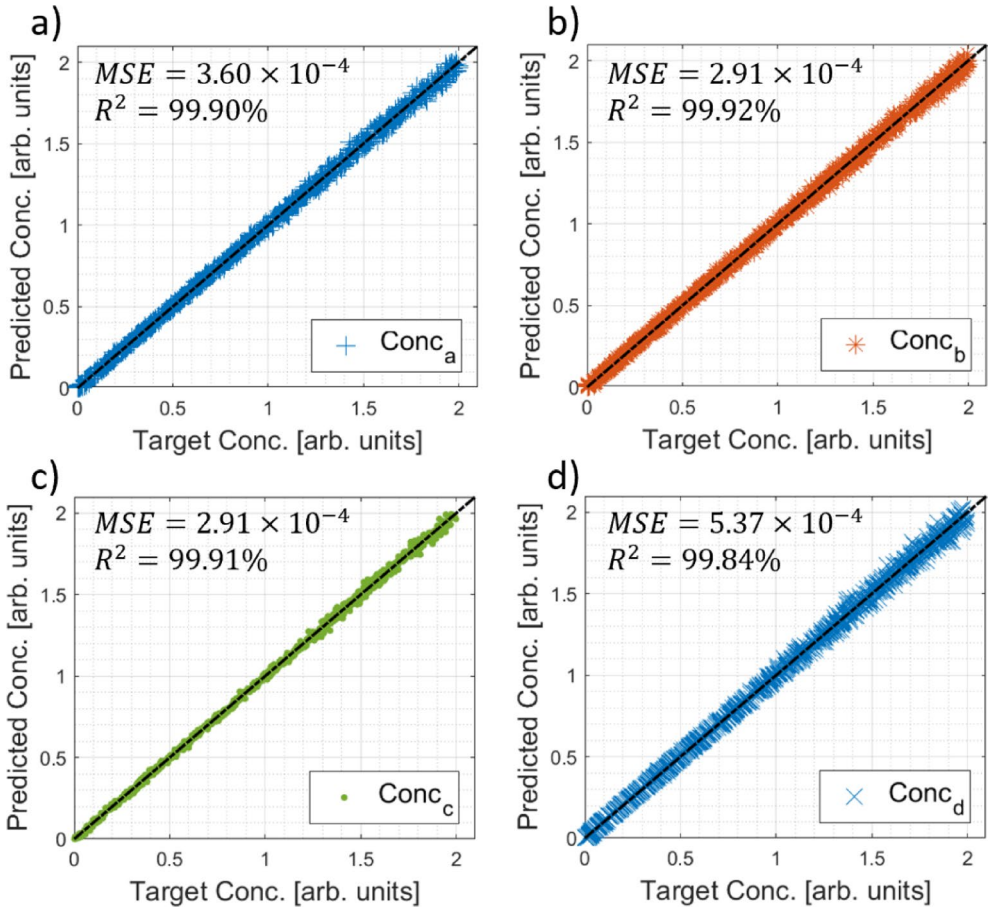

Fig. 7. The plots of the figure show the comparison between an original concentration's values of database gave to the CNN and the correspondent predicted concentration's values, for each one of the four substances simulated, in condition of Sensor's Non-Linearity phenomenon. In the first plot the Agent a concentration's values comparison has been reported (***a***), in the second plot the Agent b concentration's values comparison has been reported (***b***), in the third plot the Agent c concentration's values comparison has been reported (***c***), and in the fourth one the Agent d concentration's values comparison has been reported (***d***).

the technique's measurements, and considering that these types of information are today no so complicated to collect, it gives to the proposed approach an extreme flexibility of application.

The method is particularly useful for measurement applications that are difficult to replicate but where the underlying physics and any nonlinear conditions are well understood, as shown in this work with the spectroscopy application as proof-of-concept. A potential application could be in Raman spectroscopy, used in biomedical applications, material studies, and environmental monitoring. Here, background signals related to fluorescence from the substances being examined can obscure the desired low-intensity signals, depending on the excitation source used.

Another potential application could be in fluorescence spectroscopy, which used from biomedical fields to environmental monitoring, where signals are often affected by background phenomena necessitating preprocessing. Additionally, nonlinear phenomena related to the sample's nature (such as chemical characteristics, and reactions, etc.) or instrumentation (like spectrometer saturation at high concentrations, etc.) can pose challenges.

Therefore, the potential areas for applying this unsupervised information extraction approach in spectroscopy are extensive, ranging from environmental monitoring to diagnostic medicine, and from physics to material sciences. The critical factors for utilizing this approach effectively are a comprehensive understanding of the physics underlying the spectroscopic phenomenon and knowledge of any nonlinear phenomena that may impact the measurements. The use of an unsupervised approach, such as the one proposed, no longer requires the need for a labelled dataset for its implementation (no data-driven supervision) but must rely on by physics-informed supervision. Although this method, however, is physics-informed, it is limited by the classical limitations that characterise AI-based methods. Indeed, this type of methodology can never be used and predicted in out-of-distribution scenarios. For example, in the case in which the PINN algorithm has been trained to predict the concentration of four agents, the presence or willingness to predict the concentration of a fifth agent inevitably implies the need to retrain and validate the model by integrating information from the new agent's reference spectrum, as otherwise the model would identify the new agent's information as part of the background.

Despite the optimal performances and the several fields where this methodology may be applied, it is important to highlight again that this is an unsupervised method that is constrained by physical or instrumental a priori pieces of information. In a real-world application, such a type of a priori knowledge is affected by some uncertainties and missing information (especially for instrumental one), limiting the performance of the algorithm. For this reason, an unsupervised method is suggested only when supervised approach are not feasible (or training set is limited).

Data availability

The codes to generate the synthetic data and train the PINN will be provided by the corresponding author upon request (contact [alessandro.puleio@uniroma2.it](mailto: alessandro.puleio@uniroma2.it))

Received: 22 April 2025; Accepted: 22 October 2025
Published online: 07 November 2025

References

1. Zacharioudaki, D. E., Fitilis, I. & Kotti, M. Review of fluorescence spectroscopy in environmental quality applications. *Molecules 2022*. **27**, 4801 (2022).
2. Das, R. S. & Agrawal, Y. K. Raman spectroscopy: recent advancements, techniques and applications. *Vib. Spectrosc.* **57**, 163–176 (2011).
3. Ma, M., Zhang, J., Liu, Y., Wang, X. & Han, B. Advances in the clinical application of Raman spectroscopy in breast cancer. *Appl. Spectrosc. Rev.* https://doi.org/10.1080/05704928.2024.2352519 (2024).
4. Jones, R. R., Hooper, D. C., Zhang, L., Wolverson, D. & Valev, V. K. Raman techniques: fundamentals and frontiers. *Nanoscale Res. Lett.* **14**(1), 1–34 (2019).
5. Chen, C. et al. Applications of Raman spectroscopy in the diagnosis and monitoring of neurodegenerative diseases. *Front. Neurosci.* **18**, 1301107 (2024).
6. Baria, E. et al. Supervised learning methods for the recognition of melanoma cell lines through the analysis of their Raman spectra. *J. Biophotonics*. **14**, 202000365 (2021).
7. Ryzhikova, E. et al. Raman spectroscopy and machine learning for biomedical applications: alzheimer's disease diagnosis based on the analysis of cerebrospinal fluid. *Spectrochim Acta Mol. Biomol. Spectrosc.* **248**, 119188 (2021).
8. Guevara, E. et al. Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools. *Biomed. Opt. Express* **9**(10), 4998–5010 (2018).
9. Madden, M. G. & Ryder, A. G. Machine learning methods for quantitative analysis of Raman spectroscopy data. Opto-Ireland 2002: Optics and Photonics Technologies and Applications, 4876, 1130–1139 (2003). https://doi.org/10.1117/12.464039
10. Koshute, P., Jameson, N. J., Hagan, N., Lawrence, D. & Lanzarotta, A. Machine learning methods for classifying novel Fentanyl analogs from Raman spectra of pure compounds. *Forensic Chem.* **34**, 100506 (2023).
11. Serrano, J., Moros, J., Sánchez, C., Macías, J. & Laserna, J. J. Advanced recognition of explosives in traces on polymer surfaces using LIBS and supervised learning classifiers. *Anal. Chim. Acta*. **806**, 107–116 (2014).
12. Bertani, F. R. et al. Optical detection of aflatoxins B in grained almonds using fluorescence spectroscopy and machine learning algorithms. *Food Control*. **112**, 107073 (2020).
13. Magnus, I., Virte, M., Thienpont, H. & Smeesters, L. Combining optical spectroscopy and machine learning to improve food classification. *Food Control*. **130**, 108342 (2021).
14. Lu, Y. F. et al. Fluorescence imaging and Raman spectroscopy applied for the accurate diagnosis of breast cancer with deep learning algorithms. *Biomed. Opt. Express.* **11**(7), 3673–3683 (2020).
15. Dou, J. et al. Urine fluorescence spectroscopy combined with machine learning for screening of hepatocellular carcinoma and liver cirrhosis. *Photodiagnosis Photodyn Ther.* **40**, 103102 (2022).
16. Hastie, T., Tibshirani, R. & Friedman, J. *Overview of Supervised Learning* 9–41 (Springer, 2009). https://doi.org/10.1007/978-0-387-84858-7_2.
17. Hastie, T., Tibshirani, R. & Friedman, J. *Unsupervised Learning* 485–585 (Springer, 2009). https://doi.org/10.1007/978-0-387-84858-7_14.
18. Meza Ramirez, C. A., Greenop, M. & Ashton, L. Rehman, I. ur. Applications of machine learning in spectroscopy. *Appl. Spectrosc. Rev.* **56**, 733–763 (2021).
19. Makantasis, K., Karantzalos, K., Doulamis, A. & Doulamis, N. Deep supervised learning for hyperspectral data classification through convolutional neural networks. *International Geoscience and Remote Sensing Symposium (IGARSS)* **2015-November**, 4959–4962 (2015).
20. Chen, Y. et al. Prostate cancer identification via photoacoustic spectroscopy and machine learning. *Photoacoustics* **23**, 100280 (2021).
21. Tetef, S., Govind, N. & Seidler, G. T. Unsupervised machine learning for unbiased chemical classification in X-ray absorption spectroscopy and X-ray emission spectroscopy. *Phys. Chem. Chem. Phys.* **23**, 23586–23601 (2021).
22. Verbeeck, N., Caprioli, R. M. & Van de Plas, R. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. *Mass Spectrom. Rev.* **39**, 245–291 (2020).
23. Karniadakis, G. E. et al. Physics-informed machine learning. *Nat. Rev. Phys.* **3**(6), 422–440 (2021).
24. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
25. Anitescu, C., İsmail Ateş, B. & Rabczuk, T. *Physics-Informed Neural Networks: Theory and Applications* 179–218 (Springer, 2023). https://doi.org/10.1007/978-3-031-36644-4_5.
26. Seo, J. Solving real-world optimization tasks using physics-informed neural computing. *Sci. Rep.* **14**(1), 1–10 (2024).
27. Rossi, R., Gelfusa, M. & Murari, A. On the potential of physics-informed neural networks to solve inverse problems in Tokamaks. *Nucl. Fusion*. **63**, 126059 (2023).
28. Puleio, A., Rossi, R. & Gaudio, P. Calibration of spectra in presence of non-stationary background using unsupervised physics-informed deep learning. *Sci Rep* **13**, 2156 (2023). https://doi.org/10.1038/s41598-023-29371-9
29. P Kingma, D. & L Ba, J. Adam: A method for stochastic optimization. *3rd Int. Conf. Learn. Representations ICLR 2015 - Conf. Track. Proc.* https://doi.org/10.48550/arxiv.1412.6980 (2014).

Author contributions

Conceptualisation: A. P.; Methodology: A. P.; Formal analysis and investigation: A. P.; Writing - original draft preparation: A. P.; Writing - review and editing: A. P. and P. G.; Funding acquisition: P. G.; Supervision: P. G.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to A.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.