



Development and validation of an artificial neural network-based tool for fault detection and diagnosis in vapour compression heat pumps

Chiara D'Ignazi, Luca Molinaroli*

Dipartimento di Energia, Politecnico di Milano, Via Raffaele Lambruschini 4/A, 20156, Milano, Italy

Abstract

In this paper, a tool for the detection and diagnosis of soft faults in heat pumping systems, based on Artificial Neural Networks, is presented. The tool is developed following a two-stage approach that consists of a fault detection stage and a fault diagnosis stage. In the former, first a series of experiments is conducted on a fully instrumented, lab-scale water-to-water heat pump operating under fault-free conditions. An artificial neural network is then trained to accurately estimate thermodynamic parameters, such as the refrigerant pressures and temperatures in different locations, using the temperatures of the secondary fluids as input variables (predictors) in these conditions. Overall, the accuracy in predicting the pressures or the temperatures lies within ± 5 kPa and $\pm 0,5$ °C respectively. These pressures and temperatures are then combined to define dimensionless predictors, as the ratio between thermodynamic parameters under actual conditions and those computed with the artificial neural network under fault-free conditions. These dimensionless predictors are subsequently compared with a predefined threshold; when this threshold is exceeded, the operating condition is classified as faulty, thereby triggering the fault diagnosis stage. In this, a classification model based on an artificial neural network is developed. This model uses the same set of predictors as input variables and maps them to the corresponding fault category as output. The results demonstrate that the proposed tool achieves an accuracy of approximately 91% in the fault detection, and of around 88% in the fault diagnosis.

© HPC2026.

Selection and/or peer-review under the responsibility of the organizers of the 15th IEA Heat Pump Conference 2026.

Keywords: Artificial neural network; Fault detection and diagnosis; Heat pump

1. Introduction

The decarbonization of the heating sector is one crucial imperative to achieve carbon neutrality and meet the targets of the European Green Deal [1]. To meet this target, a large deployment of heat pumps as a substitute for fossil fuel boilers is expected in the coming years. Being more complex and efficient than boilers, heat pumps lead to significant savings in energy and environmental emissions but are more prone to failure. Therefore, reliable and accurate Fault Detection and Diagnosis (FDD) techniques must be developed as they enable the timely identification of anomalies and malfunctions that reduce performance, increase energy consumption, and compromise occupant comfort [2-3]. Several studies have demonstrated that the implementation of FDD systems can lead to significant energy savings: Lin et al. [4] reported a median reduction of 8% in energy use across 550 buildings, with relatively low implementation and maintenance costs. Nevertheless, large-scale adoption of FDD technologies remains limited due to challenges related to model generalization, sensor cost, and the lack of standardized methodologies for performance evaluation.

Faults may be classified as soft faults or hard faults [5]. When a soft fault occurs, the performances of the system or its lifetime worsen, but it may continue to operate. Vice versa, hard faults are catastrophic events that lead to system failure and might be the long-term results of undetected soft faults [6], making the latter the most important one to be detected and diagnosed. Experimental analyses on fault behaviour have revealed a wide variability in the impact of anomalies on the operating parameters of heat pumps. Indeed, single or

* Corresponding author. Tel.: +39-02-23993872
E-mail address: luca.molinaroli@polimi.it

multiple faults such as refrigerant charge errors, reduced evaporator airflow, or non-condensable gases presence may lead to decrease in capacity of up to 42% and in the COP up to 39%. Moreover, the presence of multiple simultaneous faults significantly alters the performance of conventional FDD methods, highlighting the need for more robust approaches [7-9].

Existing FDD approaches for heat pumps can broadly be categorized into model-based, rule-based, and data-driven methods [10]. Model- and rule-based systems have been widely explored due to their interpretability and physical transparency, but they typically require detailed thermodynamic modelling and expert-defined thresholds, which limit scalability across different configurations. More recently, machine learning-based techniques have gained attention for their ability to extract diagnostic patterns directly from operational data. Taheri et al. [11] demonstrated the superior performance of deep recurrent neural networks for fault detection compared to traditional regression-based methods, while Uddin et al. [12] validated an automated fault classifier for rooftop units (RTUs), with good laboratory performance but limited generalizability in real-world conditions. Similarly, Barandier et al. [13] compared several supervised learning algorithms for heat pump fault detection, identifying K-Nearest Neighbors and Ridge Regression as the most accurate and computationally efficient, with metrics exceeding 99%. Llopis-Mengual et al. [14] have further demonstrated the feasibility of implementing FDD algorithms in residential air conditioning units equipped with a limited number of sensors, validating the use of Support Vector Machines trained on simulated data with balanced accuracy above 80%.

Despite encouraging results, most studies rely on simulated datasets or semi-experimental fault scenarios, leaving a relative scarcity of fully experimental validations under realistic operating conditions. The present study wants to contribute to this topic by presenting an experimentally validated FDD tool for heat pumping systems based on Artificial Neural Networks (ANNs). A dataset is generated in a controlled laboratory environment by operating a heat pump under both fault-free and faulty conditions. The collected data are used to train and validate ANNs designed to detect and diagnose the imposed faults, with their performance metric assessed considering the ability in correctly distinguish between regular and anomalous working conditions on one side and the precision in correctly identifying the fault cause on the other.

2. Experimental method

2.1. Experimental setup

The experimental set-up used in the present study is shown in Fig. 1 and is extensively described in our previous works [15-18]. For the sake of completeness, a concise description is provided, with information extracted from the aforementioned references and referring the reader to them for a more detailed description.

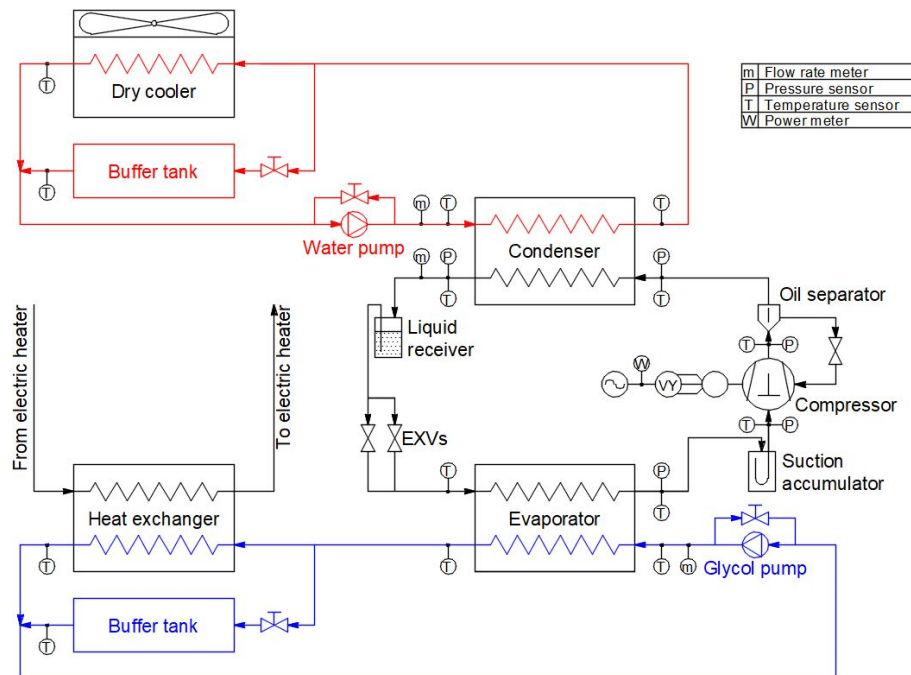


Figure 1. Layout of the experimental setup used in the present study.

The set-up consists of a water-to-water heat pump working with R513A as refrigerant. The refrigerant circuit (black line) consists of a variable speed reciprocating compressor, a hermetic oil separator, a stainless steel brazed plate condenser, a liquid receiver, two electronic expansion valves (EXVs) installed in parallel (only the most appropriate one is utilized during experiments), a stainless steel brazed plate evaporator and a vapour accumulator at the compressor inlet. In the cold loop (blue line) a 25.4%_{VOL} ethylene glycol-water mixture is used as cold heat source for the evaporator while in the hot loop (red line) pure water is used as hot heat sink for the condenser. In both loops a variable-speed centrifugal pump is used to set the secondary fluid flow rate and a 500 dm³ storage tank, which can be fully isolated by a 2-way valve, is installed in parallel to the main loop to ensure thermal stability. In the cold loop the water-ethylene glycol mixture temperature is controlled adjusting the power supplied by an electric heater connected to the auxiliary plate heat exchanger installed in it, while in the hot loop the hot water temperature is controlled adjusting the rotational speed of the fan of the fin-and-tube dry cooler installed in the loop. All the pipes and the components are insulated with an elastomeric material with a thermal conductivity equal to 0.037 W/(m·K). Further details about the components are reported in Table 1 (taken from [15-18], since the test rig is the same).

Finally, the test rig is fully instrumented to measure the most important operating parameters: refrigerant pressures, temperatures and mass flow rate, secondary fluid flow rate and temperatures and compressor power consumption. The position of the measuring instruments is shown in Fig. 1 while their main characteristics are reported in Table 2 (again, taken from [15-18], since the test rig is the same).

Table 1. Main characteristics of the components used in the experimental facility

Component	Parameter	Range
Compressor	Swept volume @ 50 Hz	13.15 m ³ /h
	Shaft rotational frequency	30 Hz – 87 Hz
	Oil	POE ISO 32
	Oil charge	1.1 dm ³
Condenser	Height x Width x Depth	289 mm x 119 mm x 93.6 mm
	Number of plates	40
Expansion valve	Capacity range (small)	1200 W – 12000W
	Capacity range (large)	1690 W – 16900 W
Evaporator	Height x Width x Depth	376 mm x 119 mm x 71.2 mm
	Number of plates	30
Vapor accumulator	Volume	2.33 dm ³
Oil separator	Type	Coalescence
	Volume	2.8 dm ³
Liquid receiver	Volume	2.8 dm ³
	Nominal flow rate	28.7m ³ /h
Centrifugal pump (identical for both water-ethylene glycol and water loops)	Nominal head	160 kPa
	Shaft rotational frequency	16 Hz–58 Hz
	Height x Width x Depth	193 mm x 76 mm x 71 mm
Auxiliary heat exchanger	Number of plates	30

Table 2. Main characteristics of the measuring instruments

Parameter	Instrument	Range	Accuracy
Refrigerant mass flow rate	Coriolis flow meter	0 kg/h – 300 kg/h	±0.15% r.v.
Refrigerant density	Coriolis flow meter	0 kg/m ³ – 2000 kg/m ³	±20 kg/m ³
Refrigerant pressure (low side)	Pressure transducer	0 kPa – 700 kPa	±0.3% f.s.
Refrigerant pressure (high side)	Pressure transducer	0 kPa – 4000 kPa	±0.3% f.s.
Refrigerant temperature	Pt 100	243.15 K – 373.15 K	±0.1K
Compressor power	Power transducer	0 W – 4000 W	±0.2% f.s.

Water and water-ethylene glycol flow rate	Vortex flow meter	0.21 m ³ /h – 3 m ³ /h	±2% r.v.
Water and water-ethylene glycol temperature	RTD Pt 100	263.15 K – 353.15 K	±0.1K

2.2. Experimental procedure and data reduction

All the tests, both under fault-free conditions and faulty conditions, are carried out in steady-state operating conditions only and using R513A as working fluid.

Each test is run setting the rotational frequency of the compressor shaft and the target superheating (or opening of EXV), while the rotational speed of the two pumps, the electric heater power input and the fan rotational speed are instead set to the values that guarantee the target secondary fluid flow rate and temperature. All the operating parameters are measured with a sampling frequency equal to 1 Hz with a LabVIEW™ interface. The test is considered in steady-state condition when the standard deviation of the last 900 samples lies within ± 2.5 kPa and ± 0.2 K for each pressure and temperature respectively. When this condition is met, the data acquisition starts and further 900 samples are recorded for data analysis, i.e. mean value and variance calculation. At the end of the test, a final check is performed calculating the refrigerant-side and the secondary fluid-side heat transfer rates at the evaporator and at the condenser, being the enthalpies required to calculate the former computed with Refprop 10 [19]. The test is considered valid if the two values agree within $\pm 4\%$ with respect to their average value, otherwise it is repeated.

Finally, to estimate the uncertainty of each parameter, the procedure described in [20] is used. According to it, for the directly measured variables the experimental uncertainty is calculated as per Eq. (1).

$$u_x = \pm \sqrt{u_{x,INST}^2 + (t_{95}\sigma_{\bar{x}})^2} \quad (1)$$

The equation for the uncertainty of the calculated quantities is not reported since, in the present study, only directly measured variables are used.

2.3. Testing conditions and simulated faults

To assess the effectiveness of the FDD method that will be described in Section 3 as FDD tool for heat pumps, an experimental campaign is carried out. This experimental activity consists of two groups of tests: fault-free tests and faulty-operation tests. In the fault-free group, the water temperature at the condenser outlet is fixed to 45 °C while the water-ethylene glycol mixture temperature at the evaporator inlet is varied. The rotational frequency of the compressor shaft is set to 50 Hz and the superheating at the evaporator outlet is fixed to 5 K. The flow rates of the secondary fluids are identified under a reference condition targeting a temperature variation $\Delta T = 5$ K and are then kept constant. The operating conditions for the fault-free tests are reported in Table 3.

Table 3. Experimental conditions used in fault-free conditions

Test	Compressor		Evaporator		Condenser		
	f_{ROT}	$\dot{m}_G$	$T_{G,IN}$	$T_{G,OUT}$	$\dot{m}_W$	$T_{W,IN}$	$T_{W,OUT}$
01	50 Hz	As test 03	0 °C	*	As test 03	*	45 °C
02	50 Hz	As test 03	5 °C	*	As test 03	*	45 °C
03	50 Hz	Identified	10 °C	5 °C	Identified	40 °C	45 °C
04	50 Hz	As test 03	15 °C	*	As test 03	*	45 °C
05	50 Hz	As test 03	20 °C	*	As test 03	*	45 °C

* Means that the value of the parameter is not set but the system is let to find it depending on the experimental conditions.

In the faulty-operation group, three soft faults commonly occurring in heat pump systems are investigated, namely: evaporator fouling (EF), experimentally simulated by reducing the water-ethylene glycol flow rate; expansion valve blockage (EXV), experimentally simulated by fixing the opening position of the electronic expansion valve; and refrigerant overcharge (CHARGE), experimentally simulated by adding a small amount (103% of the nominal charge) of R513A to the refrigerant circuit. In all these tests, the compressor shaft frequency is again fixed at 50 Hz. The superheating at the evaporator outlet is set to 5 K in EF and CHARGE tests, whereas it is free to vary in EXV tests. The flow rates of the secondary fluids are kept the same as in fault-free reference conditions for EF and EXV tests, while for CHARGE tests they are identified under a

reference condition with the same criterion as before and are then kept constant. The operating conditions for these tests are gathered in Table 4.

Table 4. Experimental conditions used in faulty conditions

Test	Charge	EXV opening	Evaporator		Condenser	
			$\dot{m}_G$	$T_{G,IN}$	$\dot{m}_W$	$T_{W,OUT}$
EF01	100%	Auto	95%	10 °C	As test 03	45 °C
EF02	100%	Auto	90%	10 °C	As test 03	45 °C
EF03	100%	Auto	85%	10 °C	As test 03	45 °C
EF04	100%	Auto	80%	10 °C	As test 03	45 °C
EF05	100%	Auto	75%	10 °C	As test 03	45 °C
EF06	100%	Auto	70%	10 °C	As test 03	45 °C
EXV01	100%	95%	As test 03	10 °C	As test 03	45 °C
EXV02	100%	90%	As test 03	10 °C	As test 03	45 °C
EXV03	100%	85%	As test 03	10 °C	As test 03	45 °C
EXV04	100%	80%	As test 03	10 °C	As test 03	45 °C
EXV05	100%	75%	As test 03	10 °C	As test 03	45 °C
EXV06	100%	70%	As test 03	10 °C	As test 03	45 °C
CHARGE01	103%	Auto	As test CHARGE03	0 °C	As test CHARGE03	45 °C
CHARGE02	103%	Auto	As test CHARGE03	5 °C	As test CHARGE03	45 °C
CHARGE03	103%	Auto	Identified, $\Delta T = 5\text{ °C}$	10 °C	Identified, $\Delta T = 5\text{ °C}$	45 °C
CHARGE04	103%	Auto	As test CHARGE03	15 °C	As test CHARGE03	45 °C
CHARGE05	103%	Auto	As test CHARGE03	20 °C	As test CHARGE03	45 °C

3. Artificial neural network for fault detection and diagnosis

As stated, in the present study, the fault detection and diagnosis tool of heat pumping system is developed following a two-stage approach in which ANNs are used both as regressor (first stage, fault detection) and classifier (second stage, fault diagnosis). The conceptual flowchart of the tool is shown in Fig. 2, while more details about the two stages are discussed in the following sections.

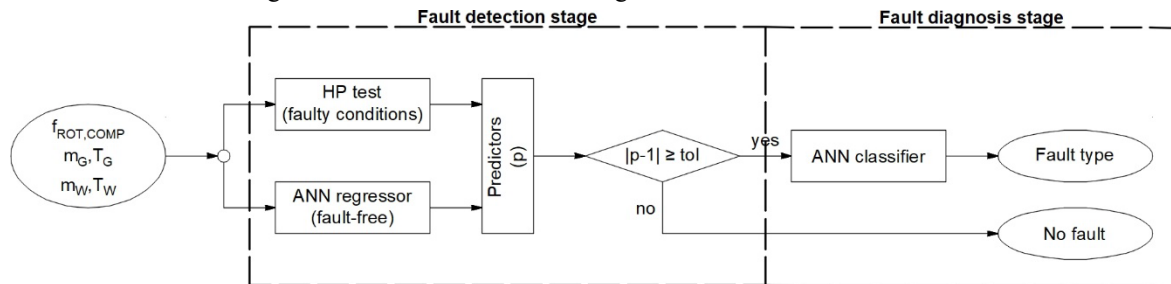


Figure 2. Flowchart of the fault detection and diagnosis tool developed in the present study.

3.1. Fault detection stage

In the fault detection stage, the operating conditions of the heat pump are evaluated by comparing some selected thermodynamic parameters of the refrigerant, namely the pressures and temperatures in various points of the refrigerant circuit, under faulty operating conditions with their corresponding values under fault-free operation. Thresholds are then applied to distinguish between normal and anomalous behaviour, as per the “outcome-based” definition of fault reported in [21].

The values of the thermodynamic parameters in faulty conditions are directly obtained from the sensors installed in the heat pump. Conversely, the reference values representative of fault-free operation are computed by means of an ANN trained on operational data collected during normal operation. This choice arises from the consideration that real operating conditions may be different from those used in the lab environment, and

a properly trained ANN has the inherent ability to correctly evaluate the thermodynamic parameters in it. More specifically, the ANN is developed as a regression model that predicts the refrigerant thermodynamic parameters given the boundary conditions, namely the compressor rotational speed as well as the mass flow rates and temperatures of the secondary fluids, as inputs. This approach allows the estimation of the expected refrigerant thermodynamic parameters in fault-free conditions for the same operating conditions experienced by the heat pump in faulty conditions, thereby allowing the direct comparison among them for effective fault detection.

Concerning the choice of the thermodynamic parameters of the refrigerant, in any vapour compression system one should measure:

- The pressure and temperature at the compressor inlet, to check that the vapour at its suction is superheated, which is required for compressor safe operation.
- The pressures at the compressor inlet and outlet, to check that the compressor operating point is within its envelope, which is required for compressor safe operation.
- The temperature at the compressor discharge, to check that the maximum allowed value is not exceeded, which is required to prevent lubricant thermal decomposition.
- The pressure and temperature at the condenser outlet, to check that the liquid flowing to the expansion valve is subcooled, which is required for expansion valve stable operation.

With the aim of reducing the number of sensors needed, to assume that the pressure drop in the discharge line and in the condenser is negligible and, therefore, only two pressures and three temperatures are sufficient to monitor the operation of the heat pump, namely: the pressure and temperature at the compressor inlet, the pressure and temperature at the compressor outlet and temperature at the condenser outlet. It is worth mentioning that, with further simplification, the pressure and temperature at the compressor inlet may be approximated with the pressure and temperature at the evaporator outlet, i.e. the parameters usually measured to control the opening of the electronic expansion valve.

Once known, or directly measured or computed via ANN, these parameters are first pre-processed to compute the refrigerant superheating at compressor inlet, as per Eq. (2), the refrigerant desuperheating at compressor outlet, as per Eq. (3), and the refrigerant pseudo-subcooling, as per Eq. (4) is (“pseudo” since the pressure used to compute it is that at compressor outlet, as per the abovementioned assumption).

$$\Delta T_{SH} = T_{COMP,IN} - T_{SAT}(p_{COMP,IN}) \quad (2)$$

$$\Delta T_{DESH} = T_{COMP,OUT} - T_{SAT}(p_{COMP,OUT}) \quad (3)$$

$$\Delta T_{SC} = T_{SAT}(p_{COMP,OUT}) - T_{COND,OUT} \quad (4)$$

The final set of predictors that allows fault detection consists of the following dimensionless parameters computed as the ratio between those above mentioned, plus the pressure at compressor inlet, and their corresponding values in fault-free operation at the same boundary conditions as those of the current heat pump operating state:

$$p_{COMP,IN}^* = \frac{p_{COMP,IN,ACTUAL}}{p_{COMP,IN,FAULT-FREE}} \quad (5)$$

$$\Delta T_{SH}^* = \frac{\Delta T_{SH,ACTUAL}}{\Delta T_{SH,FAULT-FREE}} \quad (6)$$

$$\Delta T_{DESH}^* = \frac{\Delta T_{DESH,ACTUAL}}{\Delta T_{DESH,FAULT-FREE}} \quad (7)$$

$$\Delta T_{SC}^* = \frac{\Delta T_{SC,ACTUAL}}{\Delta T_{SC,FAULT-FREE}} \quad (8)$$

The dimensionless indicators defined in Eqs. (5-8) permit to distinguish between fault-free and faulty operating conditions since they deviate from unity when the heat pump operates far from the nominal regime. However, a threshold value must be defined to clearly distinguish between normal and anomalous operation

and avoid ambiguous classification. In the present study, this threshold is set to ± 0.1 . This value is the result of a sensitivity analysis and is a good compromise between accuracy and reliability of the fault detection stage. Indeed, a threshold that is too low would lead to fault-free conditions being incorrectly identified as faulty, whereas a threshold that is too high would have the opposite effect, causing faulty conditions to be missed.

3.2. Fault diagnosis stage

In the fault diagnosis stage, the specific fault affecting the heat pump operation is identified by simultaneously evaluating the values assumed by the predictors defined in Eqs. (5-8). This is mandatory because different faults may have the same influence on the introduced indicators and, therefore, it is not possible to find a one-to-one correspondence between each predictor and a specific fault condition.

More in detail, the faults considered in the present study are the evaporator fouling (label: EF), the blockage of the expansion valve (label: EXV), and the refrigerant overcharge (label: CHARGE). As justified with the discussion of the results in Section 4, when the magnitude of these faults increases, the four predictors change as represented in Table 5.

Table 5. Influence of the three considered faults on the four predictors

Predictor	Fault		
	EF	EXV	CHARGE
$P_{COMP,IN}^*$	Reduction	Reduction	Increase
ΔT_{SH}^*	Slight increase	Significant increase	Same
ΔT_{DESH}^*	Slight increase	Significant increase	Reduction
ΔT_{SC}^*	Same	Same or slight increase	Increase

The analysis of Table 5 highlights the importance of a simultaneous evaluation of the four predictors introduced in this study. For instance, a reduction in the dimensionless pressure may result either from an increase in the evaporator fouling or from a decrease in the expansion valve opening. Consequently, the analysis of a single predictor is not sufficient to unambiguously identify the underlying fault, making it mandatory to assess all the predictors at the same time.

To discriminate among the considered faults, a classification model based on an ANN is developed. The ANN takes the four predictors as input features and maps them to the fault category as output and is trained using the full set of data, including both fault-free and faulty operating conditions, thereby enabling reliable fault classification under varying system states.

4. Results and discussion

4.1. Results of the fault detection stage

As described in Section 3, in the fault detection state the dimensionless predictors defined in Eqs. (5-8) are compared against a threshold to identify anomalous operation. To compute these predictors, reference values of the thermodynamic parameters representative of fault-free operation are needed, and these are computed by training an ANN regression model on the fault-free experimental dataset. As reported in Table 3, the boundary conditions used in the present study vary only with respect to the water-ethylene glycol mixture at the evaporator inlet. Consequently, the ANN model is significantly simplified since it is developed with a single input variable. Furthermore, due to the limited number of available experimental data, the data augmentation technique is used to increase the size of the training dataset: the pressures and temperatures measured in the different testing conditions are interpolated with a temperature step equal to 0.1 K, resulting in 201 values for each thermodynamic variable. The ANN is trained splitting the dataset in training, validation and test subsets and by optimizing the model hyperparameters (number of hidden layers, number of neurons per layer, batch size, epochs, and learning rate) to guarantee robust and accurate predictions. Overall, the ANN is able to compute the refrigerant pressures and temperatures in the fault-free operating conditions with accuracies equal to ± 5 kPa and ± 0.5 K respectively. Once trained, the ANN is used to compute the predicted fault-free values of the thermodynamic quantities under faulty boundary conditions, which are then used to evaluate the dimensionless predictors. The resulting values are shown in Figure 3 (a-d) in which, for each fault, the values of the dimensionless predictors are reported in the same order of the corresponding test conditions in Table 5.

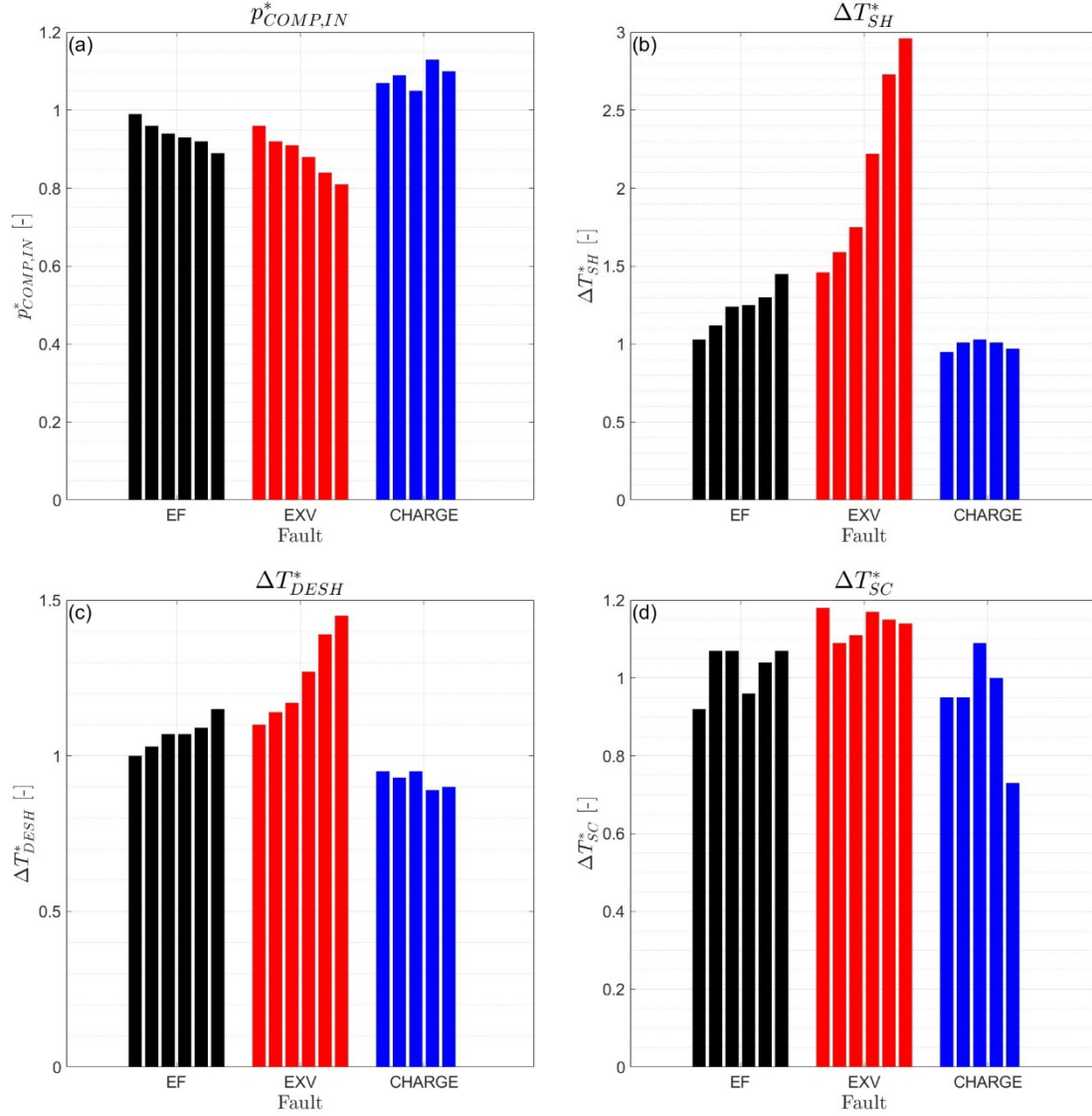


Figure 3. Dimensionless predictor for the three faulty conditions considered.

Concerning the evaporator fouling, an increase in its intensity, i.e. a decrease in the water-ethylene glycol flow rate, leads to a reduction in the evaporating pressure (and, thus, $p_{COMP,IN}^*$) since the lower the flow rate, the lower the overall heat transfer coefficient and, therefore, the evaporating temperature decreases to allow for the heat transfer rate in the evaporator. Consequently, despite the EXV correctly operates, the superheating at compressor suction (and, thus, ΔT_{SH}^*) increases because of higher heat gains, being the refrigerant pipeline well insulated but not perfectly adiabatic. Finally, the lower the evaporating pressure, the higher the temperature at the compression outlet (and, thus, ΔT_{DESH}^*) while the pseudo-subcooling at condenser outlet (and, thus, ΔT_{SC}^*) is marginally affected since the condensing pressure is not changing.

Concerning the expansion valve fault, an increase in its intensity, i.e. a reduction in the expansion valve opening, leads to higher pressure drop experienced by the refrigerant flow, which, in turn, leads to a reduction in the evaporating pressure. Consequently, $p_{COMP,IN}^*$ reduces, while ΔT_{SH}^* increases since the expansion valve is not controlling the superheating at the evaporator outlet anymore. Similarly to the evaporator fault, ΔT_{DESH}^* increases as the result of these two effects but, differently from it, a slight increase in ΔT_{SC}^* is found. This is due to the mass transfer process from the evaporator to the condenser that occurs when the superheating increases which “empties” the evaporator storing the excess charge in the condenser, which results in higher pressure at the compressor outlet.

Concerning the charge fault, it is worth remembering that only overcharging is analysed in this study and that tests are run with higher secondary fluid flow rates, which result in higher overall heat transfer coefficients in the evaporator and condenser. An increase in the refrigerant charge leads to a slight rise in the evaporating

pressure, since a larger amount of refrigerant is stored in the evaporator, producing a marginal effect on the superheating at the compressor inlet. Conversely, the higher evaporating pressure results in a reduction of the temperature at the compression discharge while the pseudo-subcooling at the condenser outlet is only marginally affected since the overcharge considered is small (+3.4%) and the condenser is still capable of accommodating the additional refrigerant mass since the reduced vapour desuperheating leaves additional volume available for liquid refrigerant storage within the condenser.

Overall, considering the previously defined threshold value of ± 0.1 , the analysis of the four dimensionless predictors enables the identification of 13 operating conditions as faulty and 2 as fault-free conditions. A close inspection reveals that the latter are classified as fault-free because, in them, the predictors remain within the threshold range, indicating no significant deviation from nominal operation.

4.2. Results of the fault diagnosis stage

Finally, the performance of the fault diagnosis stage in unambiguously classifying the faults is assessed with the confusion matrix reported in Figure 4 in which, for the sake of completeness, also the fault-free operating conditions are considered. It is worth specifying that, similarly to the fault detection stage, the hyperparameters of the ANN classifier are optimized for robustness and computational efficiency.

Predicted class	None	5		1	
	CHARGE		5		
	EF			5	
	EXV			1	6
		None	CHARGE	EF	EXV
		True class			

Figure 4. Confusion matrix of the diagnosis stage of the proposed method.

The analysis of the results shows that 15 out of the 17 faulty operating conditions, i.e. approximately 88%, are correctly classified, with the corresponding fault cause accurately identified. Two EF faulty conditions are misclassified, one as fault-free and one as EXV, since the four dimensionless predictors are within the threshold range, which leads the ANN classifier to interpret them as normal, in the first case and are similar to those achieved with EXV faulty condition, which makes the tool to classify it in this class. Considering the entire dataset, including both fault-free and faulty conditions, 20 out of the 22 operating states, i.e. approximately 91%, are correctly identified and classified. These results indicate that the tool developed in the present study can be reliably used for fault detection and diagnosis in heat pumping systems.

Finally, to assess the generalisation capability of the developed FDD tool, its predictive ability to detect faulty conditions not directly used for training is evaluated using the results of our previous study [18]. In that study, the same experimental setup described in the present paper was used to investigate the influence of a progressively reduced refrigerant charge on the heat pump performance under experimental condition 03 in Table 3. Overall, six different operating conditions, with refrigerant charge levels ranging from 103% to 83%, are available and are used as input to the FDD tool. The tool correctly identifies all six operating conditions as faulty and, quite surprisingly, classifies all of them as CHARGE faults. Since, as described, the tool is trained exclusively on overcharge conditions, it is not possible to distinguish explicitly between overcharge and

undercharge but, in any case, the tool consistently recognises deviations in refrigerant charge as a fault. This behaviour confirms the validity and generality of the developed FDD tool and suggests that it is very likely to be applicable in a broader range of operating conditions.

4.3. Limitations and future works

The developed method proved to be a reliable FDD tool for heat pumping systems. However, to conclude the discussion of the results, its main limitations are highlighted below, especially with regard to its potential application to real heat pumps. The first limitation is related to the choice of the threshold value. As stated in Section 3.1, the value used in the present study is the result of a sensitivity analysis carried out during the development of the fault detection stage. To make the method more general and facilitate its deployment, a criterion that can be applied a priori for threshold selection should be developed. A second limitation concerns the instrumentation used in the study. Since this is a laboratory application, high-quality, low-uncertainty instruments are employed, whereas such instrumentation is seldom available in real heat pump installations. It would therefore be of interest to assess how measurement uncertainty and sensor quality influence the performance of the proposed method since lower accuracy in FDD is expected with an increased uncertainty. Finally, the FDD tool is developed and tested under steady-state experimental conditions in a lab-controlled environment. In real-world heat pumps, environmental and operating conditions vary dynamically; thus, an assessment of the ability of the FDD tool to detect and diagnose faults under transient and variable conditions is advisable since greater difficulty in FDD is expected. All of the above-mentioned aspects currently limit the applicability of the developed tool and will be investigated in detail in future studies.

5. Conclusions

In the present study, an ANN-based tool for soft faults detection and diagnosis in heat pumps is developed and validated. The tool consists of two stages, namely the fault detection stage and the fault diagnosis stage. In the former, the operating conditions of the heat pump are evaluated by comparing some thermodynamic parameters of the refrigerant under faulty operating conditions with their corresponding values under fault-free operation. To this task, dimensionless pressure and dimensionless temperature differences are introduced which allow to distinguish between fault-free and faulty operating conditions since they deviate from unity when the heat pump operates in anomalous working conditions. In the latter, a classification model that takes the predictors developed in the fault detection stage as input and identify the fault category as output is developed. The tool is validated against the data measured in a dedicated experimental campaign carried out on a fully instrumented, lab-scale water-to-water heat pump operating both under fault-free and faulty operating conditions. Overall, the results demonstrate that the proposed tool achieves a fault detection accuracy of approximately 91%, while the fault diagnosis yields an accuracy of around 88%.

Nomenclature

f_{ROT}	Rotational frequency of the compressor shaft [Hz]
$\dot{m}_G$	Water-ethylene glycol mixture mass flow rate [kg/s]
$\dot{m}_W$	Water mass flow rate [kg/s]
$p_{COMP,IN}$	Refrigerant pressure at compressor inlet [kPa]
$p_{COMP,OUT}$	Refrigerant pressure at compressor outlet [kPa]
$T_{COMP,IN}$	Refrigerant temperature at compressor inlet [K]
$T_{COMP,OUT}$	Refrigerant temperature at compressor outlet [K]
$T_{G,IN}$	Temperature of the water-ethylene glycol mixture at the evaporator inlet [°C]
$T_{G,OUT}$	Temperature of the water-ethylene glycol mixture at the evaporator outlet [°C]
T_{SAT}	Saturation temperature of the refrigerant [K]
$T_{W,IN}$	Temperature of the water at the condenser inlet [°C]
$T_{W,OUT}$	Temperature of the water at the condenser outlet [°C]

Greek symbols

ΔT_{DESH}	Desuperheating temperature difference [K]
ΔT_{SC}	Subcooling temperature difference [K]

ΔT_{SH}

Superheating temperature difference [K]

References

- [1] European Commission, The European Green Deal. Brussels, December 2019.
- [2] Madani, H., Roccatello, E. A comprehensive study on the important faults in heat pump system during the warranty period. *Int. J. Refrig.* 2014, **48**: 19-25.
- [3] Bellanco, I., Fuentes, E., Vallès, M., Salom, J. A review of the fault behavior of heat pumps and measurements, detection and diagnosis methods including virtual sensors, *J. Build. Eng.* 2021, **39**, paper 102254.
- [4] Lin, G., Kramer, H., Granderson, J. Building fault detection and diagnostics: Achieved savings, and methods to evaluate algorithm performance. *Build. Environ.* 2020, **168**, paper 106505
- [5] Breuker, M. S., Braun, J. E. Common faults and their impacts for rooftop air conditioners, *HVAC&R Res.* 1998, **4**: 303-318.
- [6] Yuill, D. P., Braun, J.E. Evaluating the performance of fault detection and diagnostics protocols applied to air-cooled unitary air-conditioning equipment, *HVAC&R Res.* 2013, **19**: 882-891.
- [7] Hu, Y., Yuill, D.P., Ebrahimifakhar, A., Rooholghodos, A. An experimental study of the behavior of a high efficiency residential heat pump in cooling mode with common installation faults imposed. *Appl. Therm. Eng.* 2022, **215**, paper 116116.
- [8] Hu, Y., Yuill, D.P. Impacts of common faults on an air conditioner with a microtube condenser and analysis of fault characteristic features. *Energy Build.* 2022, **254**, paper 111630.
- [9] Hu, Y., Yuill, D.P. Multiple simultaneous faults' impacts on air-conditioner behavior and performance of a charge diagnostic method. *Appl. Therm. Eng.* 2022, **215**, paper 119015.
- [10] Chen, J., Zhang, L., Li, Y., Shi, Y., Gao, X., Hu, Y. A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems. *Renewable Sustainable Energy Rev.* 2022, **161**, paper 112395.
- [11] Taheri, S., Ahmadi, A., Mohammadi-Ivatloo, B., Asadi, S. Fault detection diagnostic for HVAC systems via deep learning algorithms. *Energy Build.* 2021, **250**, paper 111275.
- [12] Uddin, M. R., Yuill, D.P., Williams, R.E., Dvorak, B. A machine learning classifier for automated fault detection and diagnosis (AFDD) of rooftop units, addressing practical challenges of application. *Energy Build.* 2024, **310**, paper 114101.
- [13] Barandier, P., Mendes, M., Marques Cardoso, A. J. Comparative analysis of four classification algorithms for fault detection of heat pumps. *Energy Build.* 2024, **316**, paper 114342.
- [14] Llopis-Mengual, B., Yuill, D. P., Navarro-Peris, E. Fault detection and diagnosis algorithm for multiple simultaneous faults in residential air-conditioning systems: Development, validation study and critical analysis. *Appl. Therm. Eng.* 2025, **269**, paper 125975.
- [15] Colombo, L.P.M., Lucchini, A., Molinaroli, L. Experimental analysis of the use of R1234yf and R1234ze(E) as drop-in alternatives of R134a in a water-to-water heat pump. *Int. J. Refrig.* 2020, **115**: 18-27.
- [16] Molinaroli, L., Lucchini, A., Colombo, L.P.M. Drop-in analysis of R450A and R513A as low-gwp alternatives to R134a in a water-to-water heat pump. *Int. J. Refrig.* 2022, **135**: 139-147.
- [17] Molinaroli, L., Lucchini, A., Colombo, L.P.M. An experimental comparison of the performances of a small water-to-water heat pump working with R1234ze(E) and its mixture R515B. *Energies* 2024, **17**, paper 5812.
- [18] D'Ignazi, C., Molinaroli, L., Bongiorno, C. Experimental study of the performance of a water-to-water heat pump equipped with a liquid receiver with different charge levels. *Int. J. Refrig.* 2024, **168**: 688-695.
- [19] Lemmon, E. W., Bell, I., Huber, M. L., McLinden, M. O. NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport Properties-REFPROP, Version 10.0, 2018, National Institute of Standards and Technology.
- [20] Moffat, R. Describing the uncertainties in experimental results. *Exp. Therm. Fluid Sci.* 1988, **1**: 3-17.
- [21] Frank, S., Lin, G., Jin, X., Singla, R., Farthing, A., Granderson, J. A performance evaluation framework for building fault detection and diagnosis algorithms. *Energy Build.* 2019, **192**: 84-92.