Available online at www.sciencedirect.com
SCIENCE DIRECT®
PERGAMON
International Journal of Heat and Mass Transfer 46 (2003) 4413–4425
International Journal of HEAT and MASS TRANSFER
www.elsevier.com/locate/ijhmt

Convective cooling of supercritical carbon dioxide inside tubes: heat transfer analysis through neural networks

G. Scalabrin *, L. Piazza, M. Condosta

Dipartimento di Fisica Tecnica, University of Padova, via Venezia, 1-I 35131 Padova, Italy

Received 25 November 2002; received in revised form 6 May 2003

Abstract

In a previous work, convective heating of carbon dioxide was studied with neural networks (NN), obtaining a totally heuristical heat transfer equation from the direct regression of experimental data. In the present work, the analysis focuses on the cooling process, which has a technical relevance in various applications, as for example in transcritical refrigeration cycles. Heat transfer around the critical zone presents a marked enhancement, that follows the peaks in thermophysical properties like thermal conductivity and heat capacity. Similarly, other properties like density and enthalpy, present a strong variation in narrow temperature intervals around the critical point.

This constitutes then a highly non-linear phenomenon, for which it is advisable to use a very flexible function approximator like the NNs. NN models were applied both in terms of dimensionless numbers and of physical quantities, obtaining the two corresponding NN architectures. The choice of the optimal number of neurons in the NN hidden layer is discussed. The NN models are then compared with a recent correlation from literature, for which the validation results present an AAD of 27% and a bias of −26% with an evident prediction shifting. On the other hand the NN models in terms of dimensionless numbers and of physical quantities have AAD and bias of 14% and −4%, and of 7% and −2%, respectively, showing a largely better performance.

Keywords: Forced convection; Supercritical; Cooling; Carbon dioxide; Heat transfer correlations; Neural networks

1. Introduction

Carbon dioxide is currently considered an environmentally 'benign' refrigerant for domestic and industrial applications. Besides, it is often used in chemical engineering for supercritical extraction. In transcritical inverse cycles, the heat rejection is very important and such process, called *gas cooling*, takes place without phase change, since the working fluid is at a supercritical pressure. Anyway, the fluid undergoes important changes in thermophysical properties in narrow temperature intervals close to the critical point, where the heat transfer rates simultaneously present a marked enhancement.

Recently, an experimental work has appeared with a considerable number of regularly distributed measurements [1]. Such a data base seems optimal for the development of a heuristical method of heat transfer analysis. The literature presents some further few sources [2–4] with a smaller number of points.

2. Conventional heat transfer equation

In the near-critical region the thermophysical properties present very peculiar trends. Moving for example along isobars at supercritical pressures there is a temperature at which the isobaric heat capacity reaches a maximum; such temperature is called 'pseudo-critical' temperature and is indicated by T_m [5]. At each pressure P, there is a unique T_m and a 'pseudo-critical' line $T_m(P)$,

* Corresponding author. Tel.: +39-49-8276875; fax: +39-49-8276896.
E-mail address: gscala@unipd.it (G. Scalabrin).

0017-9310/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0017-9310(03)00256-4

Nomenclature

AAD	average absolute deviation
Bias	bias
c_P	isobaric heat capacity, kJ/kg K
$\bar{c}_P$	averaged isobaric heat capacity, kJ/kg K
D	diameter, m
f_f	friction factor
f_{ob}	objective function
$f(x)$	transfer function
h	fluid enthalpy, kJ/kg
I	number of neurons in input layer
J	number of neurons in hidden layer
K	number of neurons in output layer
L	length, m
$\dot{m}$	mass flow rate, kg/m^2 s
NPT	number of points
Nu	Nusselt number
P	pressure, MPa
Pr	Prandtl number
$\dot{q}$	heat flux, kW/m^2
Re	Reynolds number
S_i	output layer value
T	temperature, K
u	fluid velocity, m/s
U_i	input layer value
V_i	physical input
W_k	physical output

Greek symbols

α	heat transfer coefficient, kW/m^2 K
λ	thermal conductivity, W/m K
ρ	density, kg/m^3

Subscripts

CP	constant properties
b	at bulk
c	critical
m	pseudo-critical
r	reduced
SC	supercritical
w	at wall

or equivalently $P_m(T)$, can be then defined, which is also considered as a sort of prolongation of the saturation line $P_S(T)$. Crossing the pseudo-critical line the fluid undergoes a sort of phase-change, passing, at decreasing T, from a gas-like state to a liquid-like state. Correspondingly, there is an abrupt change of thermophysical properties like density and viscosity. The thermal conductivity shows instead a peak similar to that of heat capacity, though less pronounced [6]. A peak in the trend of the convective heat transfer coefficient was experimentally verified as well [7].

Very accurate models are required to represent thermophysical properties in the critical and near-critical zones and in the present work the Span et al. [8] thermodynamic properties formulation for carbon dioxide and the Vesovic et al. [9,10] transport properties formulation were adopted. Traditional models appeared in literature for the cooling of supercritical carbon dioxide [4] are direct modifications of either the Dittus–Boelter or Gnielinski correlations for single-phase convective heat transfer, with the inclusion of additional terms that take into account the property variation of the fluid between bulk and wall conditions, following an approach already proposed for the heating of supercritical carbon dioxide [2]. These terms, which are needed to improve the representation of experimental data, are substantially empirical, and were developed by 'trial-and-error'.

Recently, a new correlation was proposed by Pitla et al. [11]. This is based on the "mean Nusselt number" defined as

$$Nu = \left(\frac{Nu_w + Nu_b}{2}\right)\frac{\lambda_w}{\lambda_b} \tag{1}$$

where Nu_w and Nu_b are the Nusselt numbers relative to wall and bulk conditions, respectively, which are both evaluated through the Gnielinski [12] correlation:

$$Nu = \frac{f_f/2(Re - 1000)Pr}{12.7\sqrt{f_f/2}(Pr^{2/3} - 1) + 1.07} \tag{2}$$

Finally, the heat transfer coefficient reads

$$\alpha = \frac{Nu}{D}\lambda_b \tag{3}$$

The authors state that the above correlation outperforms the older one by Baskov et al. [4] so that the present correlation will be taken as the term of comparison for the neural network (NN) models development.

3. Neural networks in terms of dimensionless numbers

A new correlation technique is proposed here, based on NN. The NN have already been applied to direct correlation of experimental data of the heat transfer coefficient as a function of the working conditions, for the case of supercritical carbon dioxide heating [13]. Among the different NN architectures, the multilayer feedforward neural network (MLFN) with only one hidden layer looks to be very effective as a universal approximator of continuous functions in a compact

domain [14,15]. A MLFN contains several neuron layers (*multilayer*) and in it the information goes in only one direction, from input to output (*feedforward*).

In the former study [13], four architectures were applied to the regression of heat transfer coefficients as functions of working conditions. Two of these four architectures proved to be clearly superior to the other ones and these will be employed in the present work as well. The two proposed architectures differ for the type of input and output variables considered in the regression. The mathematical formulation is instead quite similar, as they both refer to the modular structure of a MLFN.

The NN correlations developed in [13] for the case of CO_2 supercritical heating cannot be transferred to the case of supercritical cooling because the temperature radial profile are opposite in the two cases. The consequent effects on the linked thermophysical properties appeared to present quite different trends.

The weighted summation of all the neurons in the previous layer are made at each neuron in a layer and this summation is then given as input to a transfer function. In the present work such transfer function is a sigmoid function having the form

$$f(x) = \beta \frac{1}{1 + e^{-2\gamma x}} \tag{4}$$

Two positive parameters are applied in Eq. (4) to make the function's behaviour more flexible: β changes the activation span and γ determines the steepness of the sigmoid function.

Setting the number of neurons in the three layers determines the MLFN topology, for which I represents the number of neurons in the input layer, including the bias term, and K the number of neurons in the output layer.

The number of neurons in the hidden layer J, apart the bias, is found by trial and error during the NN training. In the present cases, the optimal value of J, as an ideal compromise between computational speed and accuracy of the resulting function, was found to be 6.

In the first proposed architecture the independent variables of the system, indicated with V_1, V_2, V_3, V_4, were selected as

$$V_1 = Re \quad V_2 = Pr \quad V_3 = \left(\frac{\rho_w}{\rho_b}\right) \quad V_4 = \left(\frac{\bar{c}_P}{c_{P,b}}\right)$$

while the actual output W_1 was chosen as

$$W_1 = Nu$$

The Nu, Re and Pr numbers are calculated at the bulk conditions, whereas $\bar{c}_P$ is the isobaric heat capacity calculated at arithmetic mean conditions between bulk and wall. In the proposed architecture, as shown in Fig. 1, there are four real inputs and one real output, so it is:

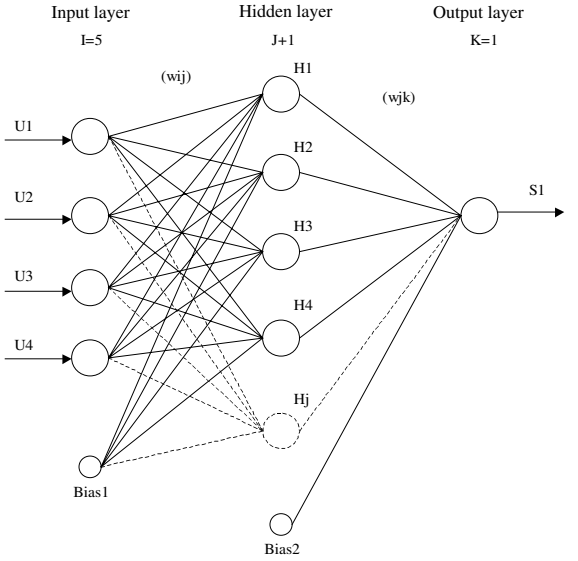

Fig. 1. Schematic representation of the MLFN for the first architecture.

$I = 5$, $K = 1$. It may be noted that the choice of inputs and outputs resembles the classical choice for monophase convective heat transfer correlations, but the two terms $\frac{\rho_w}{\rho_b}$ and $\frac{\bar{c}_P}{c_{P,b}}$ take into account the property variation along the radial coordinate. Such a variation can be very marked, as explained in Section 1. Indicating V_i the physical input in terms of dimensionless numbers $\left(Re, Pr, \frac{\rho_w}{\rho_b}, \frac{\bar{c}_P}{c_{P,b}}\right)$ and W_k the physical output (Nu), the analytical form of the present MLFN is

$$f(x) = \beta \frac{1}{1 + e^{-2\gamma x}} \tag{4}$$

$$s_k = \frac{A_{max} - A_{min}}{W_{max,k} - W_{min,k}} \tag{5}$$

$$W_k = \frac{S_k - A_{min}}{s_k} + W_{min,k} \tag{6}$$

$$S_k = f\left(\sum_{j=1}^{J+1} w_{jk} H_j\right) \tag{7}$$

$$H_j = f\left(\sum_{i=1}^{I} w_{ij} U_i\right) \tag{8}$$

$$u_i = \frac{A_{max} - A_{min}}{V_{max,i} - V_{min,i}} \tag{9}$$

$$U_i = u_i(V_i - V_{min,i}) + A_{min} \tag{10}$$

$$H_{J+1} = \text{Bias 2} \quad U_I = \text{Bias 1}$$

with

$$1 \leqslant i \leqslant I - 1 \quad 1 \leqslant j \leqslant J \quad 1 \leqslant k \leqslant K$$

Table 1
MLFN model parameters: architecture I accounting for property variations in terms of thermophysical properties

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	0.115922×10^{3}	1	1	0.333188×10^{2}	1	1	0.100947×10^{2}	1	1	-0.189968×10^{2}
2	1	0.770661×10^{2}	2	1	-0.290499×10^{3}	2	1	0.141646×10^{2}	2	1	0.340862×10^{3}
3	1	-0.157987×10^{3}	3	1	0.574304×10^{3}	3	1	0.655262×10^{2}	3	1	-0.105607×10^{3}
4	1	-0.201239×10^{1}	4	1	-0.397987×10^{3}	4	1	0.170944×10^{2}	4	1	-0.455067×10^{3}
5	1	-0.765304×10^{2}	5	1	-0.175657×10^{3}	5	1	-0.161881×10^{2}	5	1	-0.165567×10^{3}
1	2	-0.620830×10^{2}	6	1	0.428103×10^{2}	1	2	0.128031×10^{3}	6	1	-0.683557×10^{2}
2	2	0.526176×10^{3}	7	1	-0.816374×10^{2}	2	2	-0.342502×10^{3}	7	1	-0.102470×10^{3}
3	2	0.161725×10^{3}				3	2	0.349611×10^{2}			
4	2	-0.279331×10^{3}				4	2	0.561040×10^{2}			
5	2	-0.101419×10^{3}				5	2	0.648278×10^{2}			
1	3	0.610132×10^{2}				1	3	-0.794504×10^{2}			
2	3	0.831573×10^{3}				2	3	0.534532×10^{2}			
3	3	0.112848×10^{3}				3	3	0.194784×10^{2}			
4	3	-0.948379×10^{1}				4	3	0.695898×10^{1}			
5	3	-0.975638×10^{2}				5	3	0.190055×10^{2}			
1	4	-0.189003×10^{3}				1	4	-0.186614×10^{2}			
2	4	-0.229316×10^{2}				2	4	-0.720431×10^{3}			
3	4	0.136870×10^{3}				3	4	-0.573894×10^{2}			
4	4	0.103269×10^{3}				4	4	-0.126289×10^{2}			
5	4	0.289749×10^{3}				5	4	0.450589×10^{2}			
1	5	0.121851×10^{2}				1	5	-0.117138×10^{3}			
2	5	0.125004×10^{3}				2	5	-0.206921×10^{3}			
3	5	-0.256314×10^{3}				3	5	-0.784626×10^{2}			
4	5	-0.189649×10^{2}				4	5	-0.148818×10^{2}			
5	5	-0.329208×10^{3}				5	5	-0.292379×10^{1}			
1	6	-0.172519×10^{3}				1	6	-0.167950×10^{3}			
2	6	-0.162201×10^{3}				2	6	-0.145573×10^{3}			
3	6	-0.107037×10^{2}				3	6	-0.505729×10^{2}			
4	6	-0.200545×10^{3}				4	6	0.116202×10^{2}			
5	6	-0.268736×10^{3}				5	6	0.729134×10^{2}			

$V_{\min,1}\equiv Re^{\min}$	0	$V_{\min,1}\equiv Re^{\min}$	0
$V_{\max,1}\equiv Re^{\max}$	400,000	$V_{\max,1}\equiv Re^{\max}$	400,000
$V_{\min,2}\equiv Pr^{\min}$	0	$V_{\min,2}\equiv Pr^{\min}$	0
$V_{\max,2}\equiv Pr^{\max}$	30	$V_{\max,2}\equiv Pr^{\max}$	30
$V_{\min,3}\equiv(\rho_w/\rho_b)^{\min}$	1	$V_{\min,3}\equiv(\rho_w/\rho_b)^{\min}$	1
$V_{\max,3}\equiv(\rho_w/\rho_b)^{\max}$	10	$V_{\max,3}\equiv(\rho_w/\rho_b)^{\max}$	10
$V_{\min,4}\equiv(\bar{c}_P/c_{P,b})^{\min}$	0	$V_{\min,4}\equiv(\bar{c}_P/c_{P,b})^{\min}$	0
$V_{\max,4}\equiv(\bar{c}_P/c_{P,b})^{\max}$	5	$V_{\max,4}\equiv(\bar{c}_P/c_{P,b})^{\max}$	5
$W_{\min,1}\equiv Nu^{\min}$	0	$W_{\min,1}\equiv Nu^{\min}$	0
$W_{\max,1}\equiv Nu^{\max}$	5000	$W_{\max,1}\equiv Nu^{\max}$	5000
J	6	J	6
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\max}$	0.05	$A_{\max}$	0.05
$A_{\min}$	0.95	$A_{\min}$	0.95
Bias 1	1.0	Bias 1	1.0
Bias 2	1.0	Bias 2	1.0
I	5	I	5
K	1	K	1
NPT training	391	NPT training	537
Training residual AAD	11.15%	Training residual AAD	14.29%
Training residual Bias	−2.42%	Training residual Bias	−3.35%

Here J is the number of neurons in the hidden layer, $A_{\min}$ and $A_{\max}$ are the allowable range limits of the compressed input variables, $V_{\min,i}$ and $V_{\max,i}$ are the pre-defined limits of the independent input variables, and $W_{\min,k}$ and $W_{\max,k}$ are the pre-defined limits of the dependent output variable. The quantities V_i and W_k are the generic independent and dependent variables, respectively. The transfer function defined in Eq. (4) is recalled at Eqs. (7) and (8). In Eq. (7) the summation is over the $J+1$ nodes of the hidden layer, whereas in Eq. (8) it is over the I nodes of the input layer.

Considering the present problem characteristics, the MLFN parameters are set here to the values reported in the lower part of Table 1. Both the input variables and the output function have thus been compressed within the range 0.05–0.95. Given an experimental data set of output Nu_i, in the independent variables $\left(Re, Pr, \frac{\rho_w}{\rho_b}, \frac{\bar{c}_P}{c_{P,b}}\right)_i$, the weighting factors w_{ij} and w_{jk} of the MLFN are found by minimizing the objective function:

$$f_{ob} = \frac{1}{\mathrm{NPT}} \sum_{i=1}^{\mathrm{NPT}} \left(\frac{Nu_i^{\mathrm{calc}} - Nu_i^{\mathrm{exp}}}{Nu_i^{\mathrm{exp}}} \right)^2 \tag{11}$$

by means of a non linear optimization procedure. In Eq. (11) NPT stands for the number of experimental points of the training set.

After the training, the Nu equation is obtained as a continuous function of $\left(Re, Pr, \frac{\rho_w}{\rho_b}, \frac{\bar{c}_P}{c_{P,b}}\right)$. Two training sets were selected, obtaining two sets of weighting factors, then originating two versions of the first NN architecture. The first version was trained on only one subset of the Olson data [1], which was randomly extracted for 391 points on a total of 1564, see second part of Table 1. The Olson data [1] is by far the largest data source available in literature. Their distribution is shown in Fig. 2 in a (T_r, P_r) plot, along with the carbon dioxide saturation and pseudo-critical lines. It is evident that most of the data have been taken in the proximity of the pseudo-critical line, where the thermophysical properties show a very peculiar behaviour, and the peaks in the heat transfer coefficient are present. Due to the correlative mode of the method the validity range obtained for the NN model is the same as for the reference data [1] of the training set. Those data and the associated validity range of the correlation quite homogeneously cover the following ranges: $62800 \leqslant Re \leqslant 290000$; $1.27 \leqslant Pr \leqslant 20.75$; $1.10 \leqslant \frac{\rho_w}{\rho_b} \leqslant 4.13$; $0.19 \leqslant \frac{\bar{c}_p}{c_{p,b}} \leqslant 3.07$. Since the method is totally heuristic, it is essential to have the available experimental data evenly distributed. For the time being, the data from the present source [1] are the only ones fulfilling this requirement. It was decided not to train the NN on all the available data without considering their spatial consistency with respect to the independent variables. In fact, the primitive data base can be seen as a sparse set of points, whereas a MLFN, as a universal function approximator, needs to be applied to a compact data domain according to the Kolmogorov theorem [16]. The mentioned validity ranges have been based on this criterion.

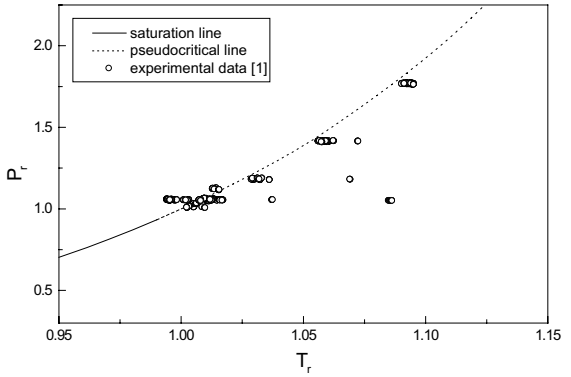

Fig. 2. Data distribution [1] in reduced temperature, reduced pressure coordinates.

The second version of the same architecture was trained on an enlarged data set, including the former subset from source [1], plus the previously-mentioned points from the other sources concerning the cooling of carbon dioxide [2–4], which fall inside the same range of the dimensionless numbers $\left(Re, Pr, \frac{\rho_w}{\rho_b}, \frac{\bar{c}_P}{c_{P,b}}\right)$ as the Olson data [1], regardless of their claimed accuracy. The same statistical weights, accounting for the experimental uncertainty, were assigned to all the data points for the regression. The NPT of the training set, the weighting factors, the NN parameters, together with the residual deviations as AAD% are given in Table 1 for both the first and the second versions of the present architecture.

As a general rule throughout all the present work experimental points have never been excluded for their error noise, but only to meet the correlations validity ranges.

4. Neural networks in terms of physical variables

A second architecture can be furthermore considered in which the independent variables are reduced pressure P_r, reduced temperature T_r, mass flow rate $\dot{m}$, and wall-to-bulk temperature ratio $\frac{T_w}{T_b}$:

$$V_1 = P_r \quad V_2 = T_r \quad V_3 = \dot{m} \quad V_4 = \left(\frac{T_w}{T_b}\right)$$

whereas the dependent variable is the heat transfer coefficient α:

$$W_1 = \alpha$$

The variables choice coincides in this second architecture with the directly accessible physical variables,

Table 2
MLFN model parameters: architecture II accounting for property variations in terms of temperatures ratio

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	0.254576×10^{4}	1	1	0.114401×10^{4}	1	1	0.901524×10^{2}	1	1	-0.249547×10^{4}
2	1	0.646935×10^{3}	2	1	-0.222578×10^{4}	2	1	0.302753×10^{4}	2	1	0.283027×10^{4}
3	1	-0.117798×10^{4}	3	1	-0.248180×10^{4}	3	1	-0.133688×10^{3}	3	1	0.609804×10^{3}
4	1	-0.171334×10^{4}	4	1	0.394081×10^{4}	4	1	-0.263573×10^{4}	4	1	-0.910596×10^{3}
5	1	0.117291×10^{4}	5	1	-0.630242×10^{3}	5	1	0.250458×10^{3}	5	1	-0.276617×10^{3}
1	2	0.498016×10^{3}	6	1	-0.135844×10^{4}	1	2	-0.911566×10^{3}	6	1	-0.202142×10^{3}
2	2	0.897912×10^{3}	7	1	-0.811892×10^{3}	2	2	0.621331×10^{4}	7	1	-0.152865×10^{2}
3	2	-0.314354×10^{4}				3	2	-0.188479×10^{3}			
4	2	0.212981×10^{4}				4	2	-0.240589×10^{4}			
5	2	-0.747656×10^{3}				5	2	-0.891783×10^{3}			
1	3	0.116098×10^{4}				1	3	0.872285×10^{3}			
2	3	0.159869×10^{4}				2	3	0.203892×10^{4}			
3	3	-0.547768×10^{3}				3	3	-0.175460×10^{3}			
4	3	-0.927638×10^{4}				4	3	-0.347801×10^{4}			
5	3	0.474277×10^{4}				5	3	0.832579×10^{3}			
1	4	-0.124808×10^{4}				1	4	0.114331×10^{4}			
2	4	0.738315×10^{4}				2	4	0.917624×10^{2}			
3	4	-0.270902×10^{3}				3	4	-0.443613×10^{3}			
4	4	-0.426833×10^{4}				4	4	0.393324×10^{3}			
5	4	-0.120640×10^{3}				5	4	-0.162672×10^{2}			
1	5	0.491434×10^{3}				1	5	-0.160871×10^{3}			
2	5	0.480785×10^{3}				2	5	0.112843×10^{4}			
3	5	-0.607462×10^{2}				3	5	-0.285709×10^{3}			
4	5	0.445707×10^{3}				4	5	-0.122013×10^{3}			
5	5	0.628258×10^{3}				5	5	-0.241125×10^{3}			
1	6	-0.101112×10^{3}				1	6	0.621787×10^{3}			
2	6	0.166935×10^{4}				2	6	-0.321788×10^{3}			
3	6	-0.275145×10^{3}				3	6	-0.538225×10^{3}			
4	6	-0.807149×10^{3}				4	6	-0.267938×10^{3}			
5	6	0.567373×10^{2}				5	6	-0.380485×10^{3}			

$V_{\min,1}\equiv P_{r}^{\min}$	0.8	$V_{\min,1}\equiv P_{r}^{\min}$	0.8
$V_{\max,1}\equiv P_{r}^{\max}$	2.0	$V_{\max,1}\equiv P_{r}^{\max}$	2.0
$V_{\min,2}\equiv T_{r}^{\min}$	0.7	$V_{\min,2}\equiv T_{r}^{\min}$	0.7
$V_{\max,2}\equiv T_{r}^{\max}$	1.4	$V_{\max,2}\equiv T_{r}^{\max}$	1.4
$V_{\min,3}\equiv \dot{m}^{\min}$	0	$V_{\min,3}\equiv \dot{m}^{\min}$	0
$V_{\max,3}\equiv \dot{m}^{\max}$	1200	$V_{\max,3}\equiv \dot{m}^{\max}$	1200
$V_{\min,4}\equiv (T_{w}/T_{b})^{\min}$	0.6	$V_{\min,4}\equiv (T_{w}/T_{b})^{\min}$	0.6
$V_{\max,4}\equiv (T_{w}/T_{b})^{\max}$	1.3	$V_{\max,4}\equiv (T_{w}/T_{b})^{\max}$	1.3
$W_{\min,1}\equiv h^{\min}$	0	$W_{\min,1}\equiv h^{\min}$	0
$W_{\max,1}\equiv h^{\max}$	30	$W_{\max,1}\equiv h^{\max}$	30
J	6	J	6
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\max}$	0.05	$A_{\max}$	0.05
$A_{\min}$	0.95	$A_{\min}$	0.95
Bias 1	1.0	Bias 1	1.0
Bias 2	1.0	Bias 2	1.0
I	5	I	5
K	1	K	1
NPT training	391	NPT training	488
Training residual AAD	3.19%	Training residual AAD	8.40%
Training residual Bias	−0.24%	Training residual Bias	−1.62%

avoiding the need for any thermophysical model for the target fluid. The effect of radial property variation is accounted for by the temperature ratio $\frac{T_w}{T_b}$.

Two versions of the second architecture were obtained, as for the former study for heating [13], selecting a training set for each of them. The first version was trained only on a subset of Olson data [1] composed of 391 points on a total of 1564. The validity range of this correlation is the same of the data subset [1] used for the NN training. The data regularly cover the ranges of physical quantities: $1.005 \leqslant P_r \leqslant 1.771$; $0.994 \leqslant T_r \leqslant 1.095$; $264 \leqslant \dot{m} \leqslant 884$ kg/m^2 s; $0.898 \leqslant \frac{T_w}{T_b} \leqslant 0.994$. The second version was trained on an enlarged data set, including the Olson data [1] subset, plus the points from the other sources within the same range of the physical quantities $[P_r, T_r, \dot{m}, (T_w/T_b)]$ as the reference data [1]. The weighting factors w_{ij} and w_{jk} and all the parameters needed to implement the obtained MLFNs are given in Table 2 for both the first and the second architecture versions.

Given an experimental data set of the output α_i, in the independent variables $[P_r, T_r, \dot{m}, (T_w/T_b)]$, the weighting factors are found by minimizing the objective function:

$$f_{ob} = \frac{1}{NPT} \sum_{i=1}^{NPT} \left(\frac{\alpha_i^{calc} - \alpha_i^{exp}}{\alpha_i^{exp}} \right)^2 \tag{12}$$

The choice of hidden neurons J is crucial in any NN application. Such a parameter is in principle unknown and must be determined by trial-and-error. As an example, considering the training of the first version of the second architecture, a NN with $J = 8$ was initially set up, obtaining a low deviation both for the training set, that is a random selection from [1], and for the validation set, that is the whole source [1]. Successively, the number of hidden neurons was progressively diminished, and correspondingly different NNs were trained and validated. The object of this study was to determine the minimum number of neurons assuring a convenient representation of the data set in terms of AAD. The AAD obtained for both the training set and the validation set, for the various NNs with different numbers of neurons J, is shown in Fig. 3. It appears that, for $J > 4$ the training residual AAD, and the AAD on the validation set, is acceptable, while for a number of hidden neurons less than 4, these figures dramatically increase. So it comes that four is the minimum number of neurons to have a satisfactory representation of the phenomenon. An overfitting is not evident for the higher values of J considered, since the AAD of the validation set is always quite close to the training residual AAD. It is then reasonable to choose a value for J between 5 and 8 and $J = 6$ was finally posed.

After training, the heat transfer coefficient can be expressed as a continuous function $\alpha = \alpha[P_r, T_r, \dot{m}, (T_w/T_b)]$.

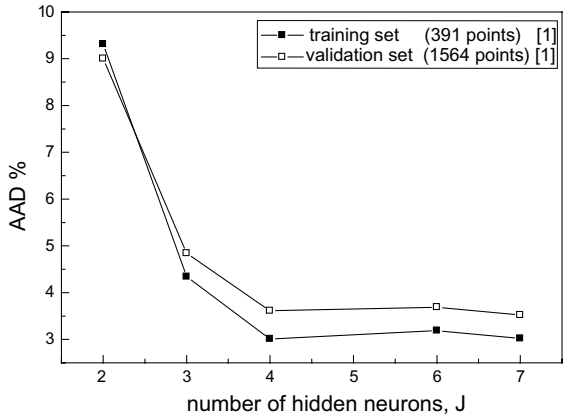

Fig. 3. Validation test of the MLFN as a function approximator of the heat transfer data of Olson [1], with decreasing numbers of hidden neurons.

Considering the data set by Olson [1] it was decided to better analyze the data region bounded by $1.004 \leqslant P_r \leqslant 1.126$ and $0.994 \leqslant T_r \leqslant 1.017$ which, being close to the critical point and including the pseudocritical line, is expected to present an evident heat transfer enhancement. Such a zone encompass 904 data points, which is the majority of the source [1]. A further version of the second NN architecture was consequently obtained, whose parameters are given in Table 3. This third version is dedicated to the near-critical zone, where its performance is better with respect to the second version. The regression was developed on all the 904 data points, obtaining a residual AAD of 3.39% and a bias of −0.46%. This last version has the following validity ranges: $1.004 \leqslant P_r \leqslant 1.126$; $0.994 \leqslant T_r \leqslant 1.017$; $264 \leqslant \dot{m} \leqslant 884$ kg/m^2 s; $0.961 \leqslant \frac{T_w}{T_b} \leqslant 0.994$. Outside these ranges, one should go back to the second version of NN.

5. Validation results

The four NN models previously described were tested and the data sources were first of all validated against the conventional equation in order to check their consistency. The validation results of the Pitla conventional equation ((1)–(3)) with respect to the literature sources considered [1–4], are given in Table 4. In the present work the error deviations (Δ%), the average absolute deviation (AAD%) and the bias (Bias%) are evaluated as

$$(\Delta\%)_i = \frac{x_i^{exp} - x_i^{calc}}{x_i^{calc}} \times 100$$

$$AAD\% = \frac{1}{NPT} \sum_{i=1}^{NPT} |\Delta\%|_i$$

$$Bias\% = \frac{1}{NPT} \sum_{i=1}^{NPT} (\Delta\%)_i \tag{13}$$

Table 3
MLFN model parameters: architecture II version for the near-critical zone within the range $1.004 \leqslant P_r \leqslant 1.126$ and $0.994 \leqslant T_r \leqslant 1.017$

Third version					
i	j	w_{ij}	j	k	w_{jk}
1	1	0.263517×10^{4}	1	1	-0.763325×10^{2}
2	1	0.387193×10^{4}	2	1	-0.307424×10^{2}
3	1	-0.237990×10^{4}	3	1	-0.387029×10^{2}
4	1	-0.638145×10^{4}	4	1	-0.379542×10^{4}
5	1	0.181845×10^{4}	5	1	-0.297639×10^{3}
1	2	0.291374×10^{4}	6	1	0.387413×10^{4}
2	2	0.207044×10^{3}	7	1	0.162806×10^{3}
3	2	-0.881127×10^{4}			
4	2	-0.963407×10^{4}			
5	2	0.877050×10^{4}			
1	3	0.736562×10^{3}			
2	3	0.190111×10^{3}			
3	3	-0.109774×10^{4}			
4	3	-0.168741×10^{4}			
5	3	-0.129334×10^{4}			
1	4	-0.264093×10^{4}			
2	4	0.165933×10^{5}			
3	4	-0.693673×10^{3}			
4	4	-0.146626×10^{5}			
5	4	0.180097×10^{4}			
1	5	0.120399×10^{3}			
2	5	0.395033×10^{3}			
3	5	0.210310×10^{3}			
4	5	0.450541×10^{3}			
5	5	0.338591×10^{3}			
1	6	-0.290182×10^{4}			
2	6	0.172523×10^{5}			
3	6	-0.670271×10^{3}			
4	6	-0.144981×10^{5}			
5	6	0.147572×10^{4}			

$V_{\min,1} \equiv P_r^{\min}$	0.8
$V_{\max,1} \equiv P_r^{\max}$	2.0
$V_{\min,2} \equiv T_r^{\min}$	0.7
$V_{\max,2} \equiv T_r^{\max}$	1.4
$V_{\min,3} \equiv \dot{m}^{\min}$	0
$V_{\max,3} \equiv \dot{m}^{\max}$	1200
$V_{\min,4} \equiv (T_w/T_b)^{\min}$	0.6
$V_{\max,4} \equiv (T_w/T_b)^{\max}$	1.3
$W_{\min,1} \equiv h^{\min}$	0
$W_{\max,1} \equiv h^{\max}$	30
J	6
β	1.0
γ	0.005
$A_{\max}$	0.05
$A_{\min}$	0.95
Bias 1	1.0
Bias 2	1.0
I	5
K	1
NPT training	904
Training residual AAD	3.39%
Training residual Bias	−0.46%

Table 4
Validation of Pitla et al. [6] conventional equation[a]

Flowdirection	NPT	AAD%	Bias%	Max%	Ref.
H	1564	26.53	−26.46	−58.66	[1]
H	80	21.70	−18.80	−70.94	[2]
V	109	30.28	−20.32	110.35	[3]
V	24	18.85	−9.66	−36.30	[4]
Overall	1777	26.44	−25.51	110.35	

[a] H = horizontal, V = vertical, NPT = number of points, AAD = average absolute deviation, Max = maximum deviation.

Table 5
Dimensionless numbers MLFN model validation (first and second versions[a,b])

NPT	NPT1	Pitla et al. equation			MLFN 1st arch., 1st vers.			MLFN 1st arch., 2nd vers.			Ref.
		AAD%	Bias%	Max%	AAD%	Bias%	Max%	AAD%	Bias%	Max%	
1564	1564	26.53	−26.46	58.66	11.47	−2.23	32.47	13.49	−4.45	42.89	[1]
80	52	16.86	−13.11	64.81	20.07	−11.46	67.29	15.14	−11.80	−49.22	[2]
109	86	30.91	−21.66	110.35	23.54	19.51	75.54	18.00	8.82	51.54	[3]
24	8	29.84	−29.84	36.30	17.72	17.72	−25.55	16.29	−16.29	−23.04	[4]
1777	1710	26.47	−25.83	110.35	12.37	−1.49	75.54	13.78	−4.07	51.54	

[a] H = horizontal, V = vertical, NPT = number of points, AAD = average absolute deviation, Max = maximum deviation.
[b] NPT1: number of points in the validity range of the 1st architecture NN model; NPT2: number of points in the validity range of the 2nd architecture NN model.

Table 6
Physical variables MLFN model validation (first and second versions[a,b])

NPT	NPT2	Pitla et al. equation			MLFN 2nd Arch., 1st vers.			MLFN 2nd Arch., 2nd vers.			Ref.
		AAD%	Bias%	Max%	AAD%	Bias%	Max%	AAD%	Bias%	Max%	
1564	1564	26.53	−26.46	58.66	3.69	0.12	37.28	6.56	−1.96	35.30	[1]
80	43	23.68	−21.54	70.94	41.21	15.08	100.32	17.11	−12.45	−51.57	[2]
109	54	33.02	−19.27	110.35	29.25	13.47	111.12	21.13	6.39	59.59	[3]
24	0	–	–	–	–	–	–	–	–	–	[4]
1777	1661	26.66	−26.10	110.35	5.49	0.94	111.12	7.31	−1.96	59.59	

[a] H = horizontal, V = vertical, NPT = number of points, AAD = average absolute deviation, Max = maximum deviation.
[b] NPT1: number of points in the validity range of the 1st architecture NN model; NPT2: number of points in the validity range of the 2nd architecture NN model.

where x is a dependent variable, which can be Nu or α, and exp and calc stand for experimental and calculated values, respectively.

The deviations with respect to the Pitla correlation (P) are quite large, Table 4, and in fact all the data sources are underestimated from −10% up to −26%.

The validation of the NN model in terms of dimensionless numbers, here indicated as the first architecture, has been conducted only on the data points in the validity range of the model. For those points, a parallel validation was developed for comparison with the conventional equation and the results are given in Table 5 for the first and the second versions. The first version represents the data source [1], on a subset of which it was trained, with an AAD of 11.47%, which can be compared with the AAD of 26.53% of the conventional equation. This can be considered a high quality result, since the claimed experimental uncertainty for the source [1] is about 10%. The first version is generally better than the conventional equation also in the representation of the other sources not included in the training set. This indicates a good prediction capability of this architecture. In the second version of the first architecture, some points from the other sources were introduced in the regression, and the NN performance was improved. In fact the sources [2–4] are better reproduced with the second version than with the first one, whereas the representation of the source [1] is slightly worse. The AAD of the second version for this source is 13.49%, compared to 11.47% of the first version. This is

probably due to some error noise of the data present in the sources introduced in the regression. It was avoided to search for an explanation of such a discrepancy, and also to divide the data sources into *primary* and *secondary*. Furthermore, it may be noted from Table 1 that the training residual AAD is quite high, 11.15% for the first version and 14.29% for the second version, whereas the training residual bias is relatively low, −2.42% and −3.35% respectively. This again indicates a high error noise of input data, in terms of dimensionless numbers as inputs to the function $Nu = Nu\left[Re, Pr, \left(\frac{\rho_w}{\rho_b}\right), \left(\frac{\bar{c}_P}{c_{P,b}}\right)\right]$, which may be combined with an uncertainty in the determination of the thermophysical properties in the near-critical zone [6]. Anyway, the performances of the first architecture, in the second version, is always better than the *P* correlation. In fact for some sources [3,4] the AAD is nearly two times lower.

A similar validation was done for the NN model in terms of physical quantities, i.e. for the second architecture. The results are shown in Table 6 for the first and the second versions. The first version represents the data from [1], on a subset of which it was trained, much better than the conventional *P* equation: the AAD of the two models for the source [1] is 3.69% and 26.53% respectively. Other sources, as for instance [2], are better represented by the *P* correlation. Comparing this second architecture with the first one, it shows a more limited prediction capability with respect to those sources not included in the training set. In the second version of this architecture, when also these sources are included in the training set, a significative improvement of representation is obtained also for them. Globally, the results are much better than the conventional equation. It must be noted that, introducing in the regression all the other sources, brings about a worse representation of the main source [1]; in fact the AAD on this source nearly doubles, moving from the first to the second version. This is again due to the scattering between different experimental sources. Anyway, looking at the training residual AAD and Bias (Table 2) it can be seen that they are quite low, if compared to the first architecture. This indicates a better performance of the minimisation procedure, that may be due to a data distribution which is more suitable in terms of physical quantities than in terms of dimensionless numbers.

It must be remarked that the second architecture, avoiding the use of dimensionless numbers, for a target fluid does not require any thermophysical models to supply values of thermodynamics and transport properties. It is not possible to extend it to fluids different from the target one, while the first architecture can be in principle applied to any fluid, provided its thermophysical properties functions are known.

To examine the global behaviour of the proposed NN equations, some plots will be presented, showing the heat transfer trends, as function of various controlling parameters, as predicted from the second architecture NN model, first and third versions. The Fig. 4 shows the heat transfer coefficient as function of the mass flow rate as from the first version NN model and the *P* correlation, together with experimental data. One can see that the *P* correlation significantly underestimates the experimental points, while the NN model represents them very well. Also from the bias reported in Table 4 for the *P* correlation it is evident that this correlation always

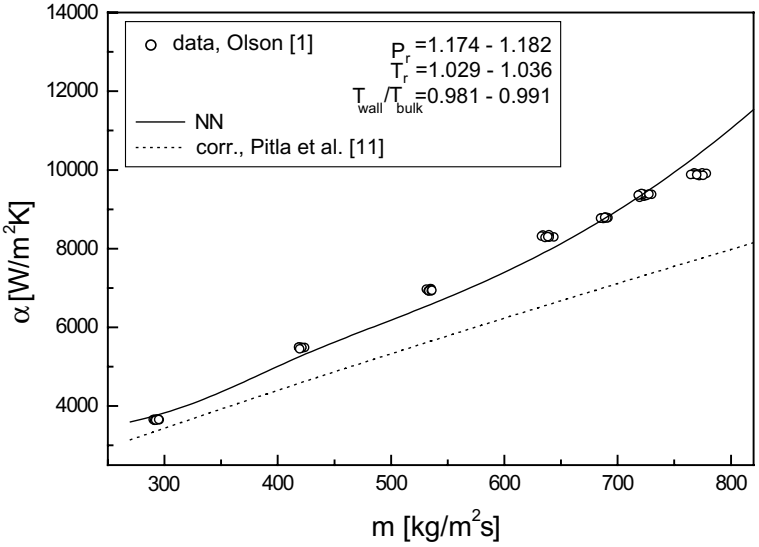

Fig. 4. Heat transfer coefficient as function of the mass flow rate, as from experimental data, from the second version NN model, first version and from the *P* correlation.

predicts a heat transfer coefficient value lower than the data. In the remaining figures, no experimental data is plotted because no homogeneous data series was found in any source for the graph conditions.

Such feature of the *P* correlation will be highlighted in the following figures where it is compared with the third version of the second NN architecture, the one dedicated to the near-critical zone. Fig. 5 plots the heat transfer coefficient as function of the bulk temperature T_b for various heat fluxes. The *P* correlation is always shifted below the NN model, and moreover it presents an unusual convex trend of the heat transfer coefficient in the proximity of the pseudo-critical line, where a minimum appears. Instead, from the NN model, which is directly based on data, a maximum appears which is

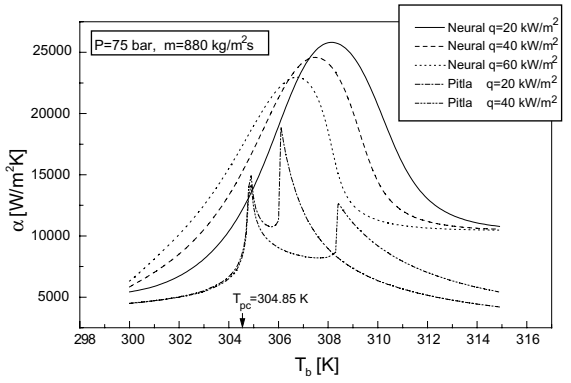

Fig. 5. Heat transfer coefficient as function of T_b for various heat fluxes, as from the second architecture NN model, third version, and from the *P* correlation.

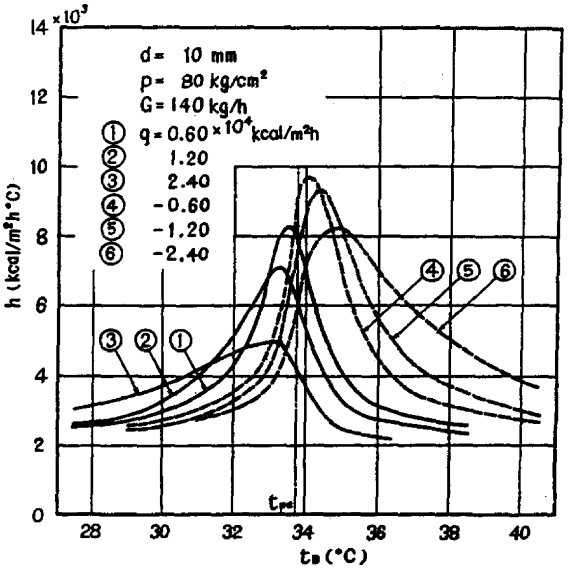

Fig. 6. Heat transfer coefficient as function of T_b for various heat fluxes, from Tanaka et al. [3].

also confirmed by other authors. In Fig. 6 it is in fact reported a diagram from Tanaka et al. [3], where, for a similar plot, the curves present a maximum. Both in Figs. 5 and 6, a greater heat flux produces a lower heat transfer coefficient, at the same other conditions.

The effect of the working pressure is illustrated in Fig. 7, where again the *P* correlation is compared with the second NN architecture, third version. It appears that at pressures slightly greater than the critical, the two models present different trends, while at higher supercritical pressures (8.5 MPa) their behaviour is quite similar. The peak in the heat transfer coefficient curves is more accentuated at pressures closer to the critical point. The 'strange' behaviour of the *P* correlation takes place just at pressures close to the critical, whereas at higher pressures the *P* correlation recovers the 'correct' trend.

Being constantly downshifted, and presenting such a 'non-physical' behaviour, the *P* model will not be considered any longer in the following graphs. Fig. 8 reports the heat transfer curves for various mass fluxes. The strong influence of this factor on the magnitude of the

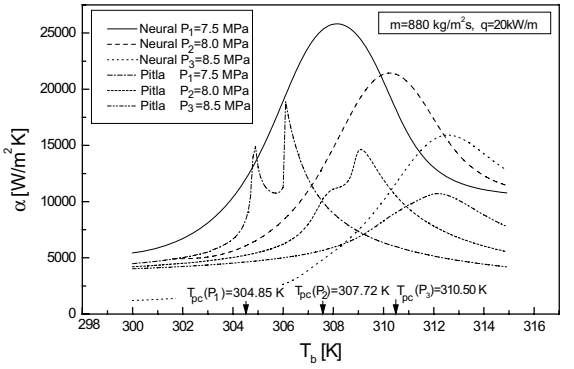

Fig. 7. Heat transfer coefficient as function of T_b for various pressures, from the second architecture NN model, third version, and from the *P* correlation.

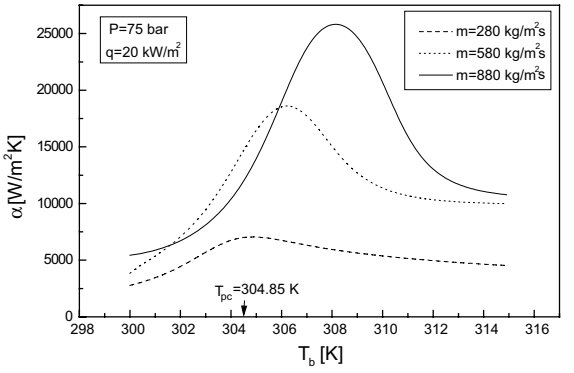

Fig. 8. Heat transfer coefficient as function of T_b for various mass fluxes, from the second architecture NN model, third version, and from the *P* correlation.

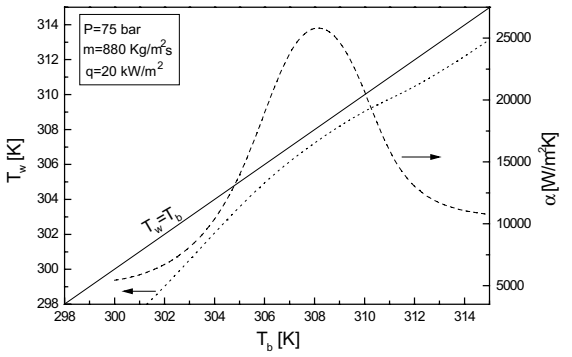

Fig. 9. Heat transfer coefficient and wall temperature T_w as functions of T_b from the second architecture NN model, third version.

heat transfer peak has to be pointed out: at higher mass fluxes, the peak is greater.

Finally, Fig. 9 plots on the same diagram the heat transfer coefficient α and the wall temperature T_w as function of the bulk temperature T_b. It is evident that, for a given heat flux $\dot{q}$, in correspondence of the peak of the heat transfer coefficient α the temperature difference between bulk and wall reaches a minimum, meaning that T_w tends to approach T_b with decreasing heat transfer resistance.

6. Discussion

The representation of the heat transfer coefficient surface for the supercritical cooling of carbon dioxide by the proposed method has been shown to be very effective. It has been also demonstrated that the method may be highly suitable to heuristically draw a heat transfer surface only from experimental data of the heat transfer coefficient.

The error noise of literature experimental data can be an important drawback of the proposed method that may reduce the final accuracy of the regressed surface. As it is common for any heuristic method, the study of the data base and its possible screening become extremely important.

As shown earlier, the data have to satisfy some basic requirements: they have to form a domain as compact as possible in the independent variables and scattered data sets should be avoided within the validity ranges of the equation. Moreover, the data points should be regularly distributed on an ideal grid of the independent variables, but these conditions are usually lacking for the data currently available in the literature. In fact heuristic methods have been rarely considered in the past for heat transfer studies [17,18], so that the outlined requirements for the data are not currently considered by the experimenters.

The results obtained for the present heat transfer problem show that equations in the dimensionless numbers or in the physical variables are equally effective in practice, suggesting that for an individual fluid the dimensionless analysis does not yield any evident advantage. Furthermore, it should be stressed that introducing the equation of state and the transport property equations of a target fluid in a conventional heat transfer correlation turns it into a fluid specific one.

Looking at the two works about supercritical carbon dioxide convective heat transfer through a MLFN technique, the former one for heating [13] and the present one for cooling, one could wonder whether a single heat transfer correlation could be drawn from data being able to represent both cases. Limiting our analysis only to the present 2nd architecture for cooling and to the former 4th architecture for heating [13], both in the physical variables, one can see that some limits are posed mostly by the actual data distribution in terms of reduced temperature and pressure and of radial temperature gradients. In fact the T_r, P_r ranges of the data sets for cooling, Fig. 2, and for heating are not at all superimposing. Moreover the radial temperature ratios, T_w/T_b, for the heating and cooling cases are higher and lower than 1, respectively, and present quite different ranges of values. In principle an extrapolation of one of the two MLFN correlations to the other case is then not allowed. However, if in the future some data sets for heating and cooling could be available with superimposing T_r, P_r ranges, a single MLFN correlation would be reasonably drawn from the comprehensive data set, considering T_w/T_b values in a wider range both lower and higher than 1.

The generalization problem of a transfer equation can be studied in more detail only if regularly distributed data are available for several fluids and with overlapping ranges, but these conditions are very difficult to find in the available experimental works. This suggests that the problem would be experimentally worth studying from this new point of view. For the heuristic nature of the method both the data distribution and quality are evident key elements in the present case. It is also worth noting that the method can be used to check the consistency of new data sets before using them for processing.

In the case the outlined requirements are strictly met, the method proves capable of drawing a high-accuracy heat transfer correlation, at the limit of the experimental uncertainty, for the present difficult problem.

References

[1] D.A. Olson, Heat transfer of supercritical carbon dioxide flowing in a cooled horizontal tube, NISTIR 6496 Report, NIST, Gaithersburg, USA, 2000.

[2] J.H. Kim, S.H Yoon, M.S. Kim, Y. Kim, An experimental investigation of heat transfer characteristics during in-tube gas cooling process of carbon dioxide, in: Proceedings IIR-Commission B1 Conference, Paderborn, Germany, 2001.
[3] H. Tanaka, N. Nishiwaki, M. Hirata, Turbulent heat transfer to supercritical carbon dioxide, in: Proceedings JSME Semi-International Symposium, Tokyo, 1967, pp. 127–134.
[4] V.L. Baskov, I.V. Kuraeva, V.S. Protopopov, Heat transfer with the turbulent flow of a liquid at supercritical pressure in tubes under cooling conditions, Teplofizika Vysokikh Temperatur 15 (1) (1977) 96–102.
[5] E.A. Krasnoshchekov, V.S. Protopopov, Experimental study of heat exchange in carbon dioxide in the supercritical range at high temperature drops, Teplofizika Vysokikh Temperatur 4 (3) (1966) 389–398.
[6] J. Millat, J.H. Dymond, C.A. Nieto de Castro, Transport Properties of Fluids, Cambridge University Press, Cambridge, UK, 1996.
[7] K. Yamagata, K. Nishikawa, S. Hasegawa, T. Fujii, S. Yoshida, Forced convective heat transfer to supercritical water flowing in tubes, Int. J. Heat Mass Transfer 15 (1972) 2575–2593.
[8] R. Span, W. Wagner, A new equation of state for CO_2 covering the fluid region from the triple point temperature to 1100 K at pressures up to 800 MPa, J. Phys. Chem. Ref. Data 25 (6) (1996) 1509–1596.
[9] V. Vesovic, W.A. Wakeham, G.A. Olchowy, J.V. Sengers, J.T.R. Watson, J. Millat, The transport properties of carbon dioxide, J. Phys. Chem. Ref. Data 19 (3) (1990) 763–808.
[10] A. Fenghour, W.A. Wakeham, V. Vesovic, The viscosity of carbon dioxide, J. Phys. Chem. Ref. Data 27 (1) (1998) 31–44.
[11] S.S. Pitla, E.A. Groll, S. Ramadhyani, New correlation for the heat transfer coefficient during in-tube cooling of turbulent supercritical carbon dioxide, in: Proceedings 4th IIR-Gustav Lorentzen Conference at Purdue, West Lafayette, IN, USA, 2000, pp. 270–278.
[12] V. Gnielinski, New equation for heat transfer in turbulent pipe and channel flow, Int. Chem. Eng. 16 (2) (1976) 359–368.
[13] G. Scalabrin, L. Piazza, Analysis of forced convection heat transfer to supercritical carbon dioxide inside tubes using neural networks, Int. J. Heat Mass Transfer 46 (2003) 1139–1154.
[14] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control, Signals Syst. 2 (1989) 303–314.
[15] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.
[16] V. Kurkova, Kolmogorov's theorem and multilayer neural networks, Neural Networks 5 (1992) 501–506.
[17] J. Thibault, B.P.A. Grandjean, A neural network methodology for heat transfer data analysis, Int. J. Heat Mass Transfer 34 (8) (1991) 2063–2070.
[18] K. Jambunathan, S.L. Hartle, S. Ashforth-Frost, V.N. Fontama, Evaluating convective heat transfer coefficients using neural networks, Int. J. Heat Mass Transfer 39 (11) (1996) 2329–2332.