Available online at www.sciencedirect.com
SCIENCE DIRECT®
PERGAMON
International Journal of Heat and Mass Transfer 46 (2003) 1139–1154
International Journal of HEAT and MASS TRANSFER
www.elsevier.com/locate/ijhmt

Analysis of forced convection heat transfer to supercritical carbon dioxide inside tubes using neural networks

G. Scalabrin *, L. Piazza

Dipartimento di Fisica Tecnica, Università di Padova, via Venezia 1, Padova, I 35131, Italy

Received 30 May 2002; received in revised form 23 September 2002

Abstract

The modeling of forced convection heat transfer for carbon dioxide flowing inside a heated tube at supercritical conditions was studied. The conventional models in the literature tend to modify a constant property correlation by including thermodynamic property terms that follow the heat flux trends. An innovative heuristic method is assumed here for the first time to draw the case-specific heat transfer coefficient correlation from the experimental data on said quantity alone.

Neural networks were used since they constitute a general, powerful function-approximator tool proving able to represent a conventional heat transfer surface precisely in the present case. Four different correlation architectures were considered for the neural network function, alternatively based on dimensionless groups and on directly accessible physical quantities as independent variables. In all these architectures, the optimal functional form of the correlation was obtained using a completely heuristic procedure based exclusively on experimental data, reaching an accuracy comparable with the experimental uncertainties declared.

An improved performance of the present model was found with respect to conventional correlations. On all the data sets, the third architecture reaches an AAD of 3.98% against 4.09% for the conventional equation and the fourth architecture an AAD of 2.67% against 4.30% for the conventional equation. Besides both these NN architectures present Bias values very close to 0, whereas the conventional equation has a Bias considerably greater.

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

1. Introduction

The rates of heat transfer to fluids can reach significantly higher values in their near-critical region than further away from said region. Possible applications range from supercritical heat exchangers in power stations to supercritical fluid extraction processes and to carbon dioxide vapor compression cycles, in which heat rejection is performed at supercritical conditions. The fluids studied are mainly carbon dioxide and water and, between the two, carbon dioxide has been the more extensively investigated, particularly under heating conditions [1–7].

The focus for these studies was mainly the development of nuclear power station cooling circuits under severe temperature conditions [2,3]. A recent, particularly interesting work on heat transfer for carbon dioxide heating [8] extensively reports the results of a systematic experimental campaign, where the working variables, such as mass flow rate, heat flux, temperature and pressure, were varied very regularly, so as to cover the working ranges homogeneously. This condition is fundamental to the application of a totally heuristic method, as shown later on. The data sets available in the literature span a wide range of working conditions, though they are not regularly and homogeneously distributed in the data region. A variety of correlations has been proposed in the literature to predict the coefficients for

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

0017-9310/03/$ - see front matter
PII: S0017-9310(02)00382-4

Nomenclature

c_P	isobaric heat capacity (kJ/kg K)
$\bar{c}_P$	averaged isobaric heat capacity (kJ/kg K)
D	diameter (m)
Ec	Eckert number
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)
ρ	density (kg/m^3)

Subscripts

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

heat transfer to fluids in the near-critical region, but discrepancies are reported also because the thermophysical properties such as density, heat capacities, enthalpy, viscosity, thermal conductivity, etc., in said region present abrupt variations with limited changes in temperature or pressure. Also, these variations induce major changes in the convective heat transfer, as shown experimentally, for instance, by Yamagata et al. [9] in the forced convection heating of supercritical water inside tubes. As usual, the correlations were developed assuming an initial tentative model, which undergoes successive modifications to correct discrepancies using a "trial and error" type of approach. The present work, on the other hand, is heuristic and aims to draw analytical formulations directly from experimental data, providing they are presented in an organized form.

Neural networks were used here because they are a very versatile and powerful function-approximator tool. Having proved their ability to represent a conventional heat transfer surface for the present case, a neural network was trained on a limited amount of data homogeneously covering the working conditions range. Once the network has been trained successfully, it is capable of representing the behavior of the whole data set. Moreover, extrapolation to further data sets also proves to be satisfactory.

2. Conventional heat transfer equation

Thermophysical properties reveal very strong gradients in the near-critical zone. To represent these variations adequately, very accurate models are required for the thermophysical properties. In the present work, the Span and Wagner [10] thermodynamic properties formulation for carbon dioxide and the Vesovic and coworkers [11,12] transport properties formulation were adopted.

As an example of the great variation in thermodynamic properties in the near-critical zone, Fig. 1 shows the density, isobaric heat capacity, thermal conductivity, viscosity and Prandtl number of carbon dioxide along supercritical isobars from 8 to 9.5 MPa, considering that $P_c = 7.3773$ MPa. From Fig. 1, it clearly emerges that, depending on the (T, P) conditions, the fluid changes greatly in density within a very narrow temperature range. Moreover, at each pressure there is a temperature, defined as the 'pseudo-critical' temperature and indicated by T_m [2], at which the isobaric heat capacity reaches a maximum. At each pressure, the fluid is in a liquid-like state at $T < T_m$ and in a gas-like state at $T > T_m$. There is thus a *pseudo-critical line* that can be considered as the continuation of the saturation line. In the near-critical region, the plots for viscosity and thermal conductivity present peaks due to long-range fluctuations [13]. This phenomenon, which calls for a local, specialized modification of the thermophysical property modeling, called 'critical enhancement', is very strong for thermal conductivity even quite far from the critical point, whereas for viscosity it is far less important and confined to a narrow range, often becoming negligible. Fig. 1 shows that, at near-critical isobars, the thermal conductivity reaches a maximum, though less pronounced than for the isobaric heat capacity in similar

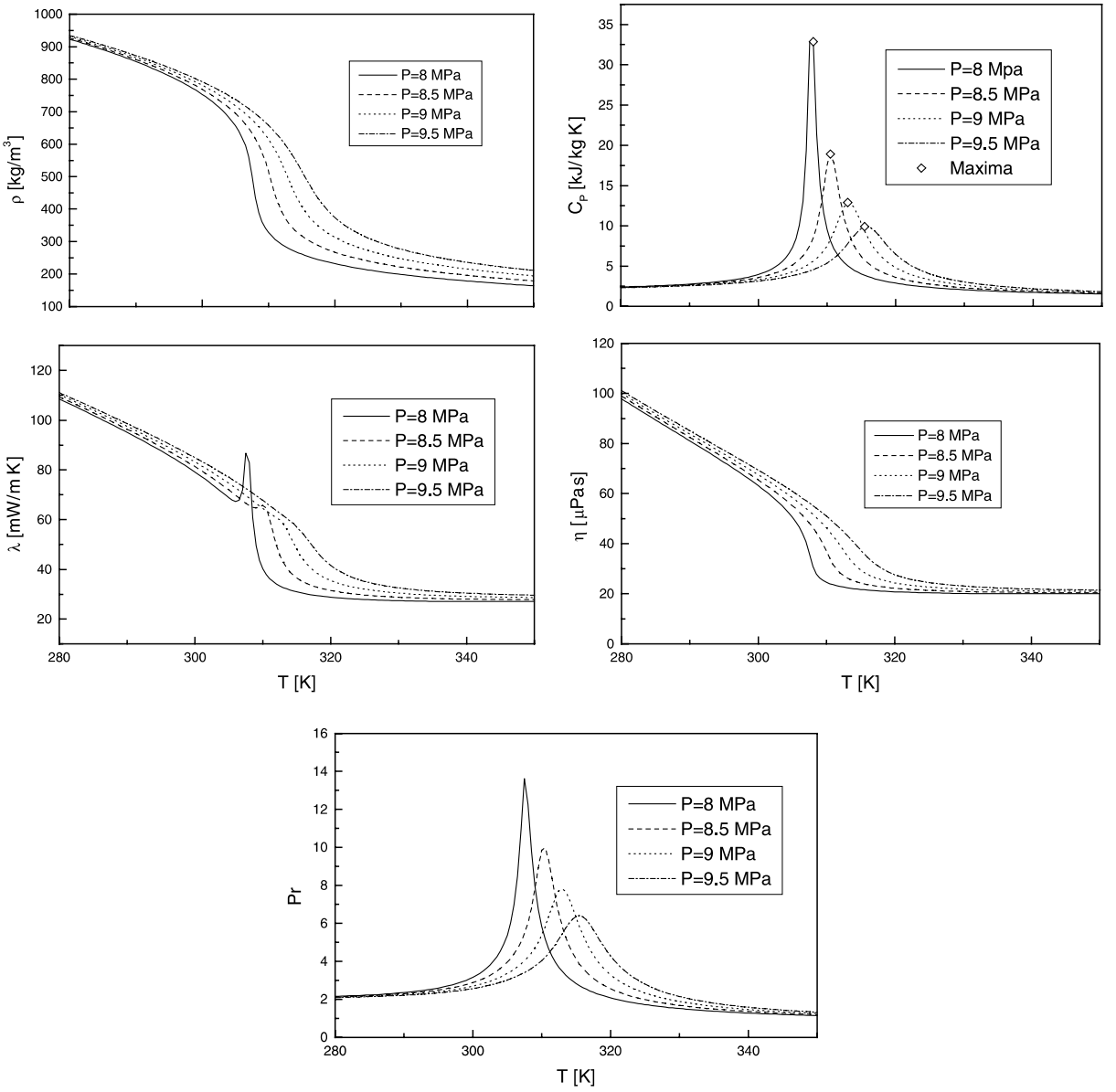

Fig. 1. Density, isobaric heat capacity, thermal conductivity, viscosity, and Prandtl number of carbon dioxide at pressures from 8 to 9.5 MPa, as function of temperature.

conditions. On the other hand, no maximum is noticed for viscosity along the same isobars. All these phenomena result in a behavior of the Prandtl number as plotted in Fig. 1, which shows a maximum mainly due to the peak in the isobaric heat capacity. Such strong variations in thermophysical properties result in both an enhancement and a deterioration of the heat transfer coefficient [9]. In order to apply the conventional, constant-properties approach to this problem successfully, attempts have been made to introduce terms in the equation that take into account the property variation of the fluid along the radial co-ordinate [2,4]. These terms, which are needed to correct the discrepancies, are substantially empirical, and were developed over the years by "trial-and-error", resulting in a long series of successive modifications. As a result, review papers on this subject [14,15] report at least five different versions of the conventional empirical correlation, generally based on the Dittus-Boelter or Gnielinski formulations for forced convection inside tubes, which differ slightly from one

another, posing the question as to which is the *best* of the published correlations. Ashfin and Asadollah [14] and Olson and Allen [8] recommend the Krasnoshchekov and Protopopov [2] formulation, where—as presented by Olson and Allen [8]—the constant property term (CP) is represented by the Gnielinski [16] correlation

$$Nu_{\mathrm{CP}} = \frac{\frac{f}{2}(Re-1000)Pr}{1+12.7\left(\frac{f}{2}\right)^{(1/2)}(Pr^{(2/3)}-1)}\left[1+\left(\frac{D}{L}\right)^{(2/3)}\right] \tag{1}$$

in which all the properties are evaluated at the bulk temperature, and the friction factor f is given by the Karman–Nikuradse implicit correlation:

$$\frac{1}{\sqrt{f}} = 4.0\log_{10}(Re\sqrt{f}) - 0.4 \tag{2}$$

which is valid in the range $2300 < Re < 5\times10^6$ and $0.5 < Pr < 2000$. The supercritical (SC) correction, proposed by Krasnoshchekov and Protopopov [2], accounts for property gradients in the radial direction between wall and core fluid through density and heat capacity ratios. The SC correlation is then obtained from Eq. (1):

$$Nu_{\mathrm{SC}} = Nu_{\mathrm{CP}}\left(\frac{\rho_{\mathrm{w}}}{\rho_{\mathrm{b}}}\right)^{0.3}\left(\frac{\bar{c}_{\mathrm{P}}}{c_{\mathrm{P,b}}}\right)^{n} \tag{3}$$

where $\bar{c}_{\mathrm{P}}$ represents an averaged heat capacity of the fluid between the bulk and the wall temperatures calculated as:

$$\bar{c}_{\mathrm{P}} = \frac{h_{\mathrm{w}}-h_{\mathrm{b}}}{T_{\mathrm{w}}-T_{\mathrm{b}}} \tag{4}$$

with $h_{\mathrm{w,b}}$ as the fluid enthalpy at wall (w) and bulk (b) conditions, respectively. The exponent n in Eq. (3) is a function of the wall temperature T_{w}, bulk temperature T_{b} and pseudo-critical temperature T_{m}, which in turn depends on the pressure. It is calculated as follows:

$$\text{if } T_{\mathrm{w}}/T_{\mathrm{m}} < 1.0 \quad \text{or} \quad \text{if } T_{\mathrm{b}}/T_{\mathrm{m}} \geqslant 1.2, \quad n = 0.4 \tag{5}$$

$$\text{if } T_{\mathrm{b}}/T_{\mathrm{m}} < 1.0 \leqslant T_{\mathrm{w}}/T_{\mathrm{m}}, \quad n = 0.4 + 0.18\left(\frac{T_{\mathrm{w}}}{T_{\mathrm{m}}}-1\right) \tag{6}$$

$$\text{if } T_{\mathrm{w}}/T_{\mathrm{m}} \geqslant 1.0 \quad \text{and} \quad 1.0 < T_{\mathrm{b}}/T_{\mathrm{m}} < 1.2$$
$$n = 0.4 + 0.18\left(\frac{T_{\mathrm{w}}}{T_{\mathrm{m}}}-1\right)\left[1-5\left(\frac{T_{\mathrm{b}}}{T_{\mathrm{m}}}\right)\right] \tag{7}$$

This correlation will be indicated from here on as the Krasnoshchekov, Protopopov, Pethukov, Gnielinski (KPPG) conventional correlation.

3. Neural networks in terms of dimensionless numbers

A new correlation technique is proposed here, based on neural networks. Neural networks (NN) have already been applied to heat transfer problems [17,18], but not yet to the direct correlation of experimental data on the heat transfer coefficient as a function of the working conditions. Among the different neural network architectures, the multilayer feedforward neural network (MLFN) with only one hidden layer seems to be the most effective as a universal approximator of continuous functions in a compact domain [19,20]. An MLFN contains several neuron layers (*multilayer*) and the information goes in only one direction, from input to output (*feedforward*).

Four architectures were applied to the regression of heat transfer coefficients as functions of working conditions. The first employs the traditional dimensionless numbers, i.e., it expresses Nu as a function of Re, Pr, and Ec, and it is represented in Fig. 2. The three values of the input layer, U_1, U_2 and U_3, represent the scaled independent variables, related, respectively, to the Reynolds number Re, Prandtl number Pr, and Eckert number Ec, which is defined as:

$$Ec = \frac{u^2}{c_{\mathrm{P}}(T_{\mathrm{w}}-T_{\mathrm{b}})} \tag{8}$$

where u is the fluid velocity. This number was chosen because it includes the temperature difference $(T_{\mathrm{w}}-T_{\mathrm{b}})$, which directly controls the thermophysical property variation between wall and bulk.

The value of the output layer S_1 represents the scaled Nusselt number, Nu. As in current practice with MLFN, the input layer values U_1, U_2 and U_3 and the output layer

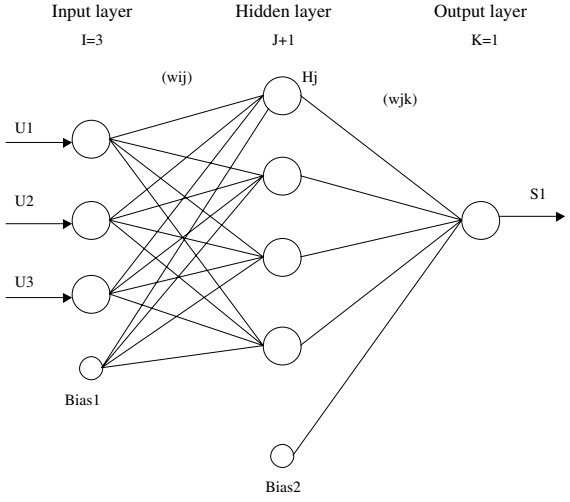

Fig. 2. Schematic representation of the MLFN for the first architecture.

value S_1 are compressed here within the same range 0.05–0.95. This means that U_1, U_2 and U_3 do not coincide with the 'true' independent variables (Re, Pr, Ec), they represent a linear transformation of them. Correspondingly, S_1 is the linear transformation of the actual output Nu. Moreover, a logarithmic filter is applied to the independent variable Ec because of its wide range of variation.

Each neuron in a layer makes the weighted summation of all the neurons in the previous layer, then passes this summation through a transfer function. The transfer function used here is a sigmoid function of the form:

$$f(x) = \beta \frac{1}{1 + e^{-2\gamma x}} \tag{9}$$

Two positive parameters are applied in Eq. (9) to make the function's behavior more flexible: β changes the activation span and γ determines the steepness of the sigmoid function.

The MLFN topology is determined once the number of neurons in the three layers has been set: I represents the number of neurons in the input layer (including the bias term), K the number of neurons in the output layer. In our case, there are three real inputs and one real output, so it is $I = 4$, $K = 1$.

The number of neurons in the hidden layer J (without the bias) has to be found by trial and error. 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 7. In our case, the actual inputs V_1, V_2, V_3, represent the independent variables of the system:

$$V_1 = Nu \quad V_2 = Pr \quad V_3 = Ec$$

and, similarly, the actual output W_1 represents the dependent variable:

$$W_1 = Nu$$

Using V_i as the physical input in terms of dimensionless numbers (Re, Pr, Ec) and W_k as the physical output (Nu), the analytical form of the present MLFN is:

$$f(x) = \beta \frac{1}{1 + e^{-2\gamma x}} \tag{9}$$

$$s_k = \frac{A_{\max} - A_{\min}}{W_{\max,k} - W_{\min,k}} \tag{10}$$

$$W_k = \frac{S_k - A_{\min}}{s_k} + W_{\min,k} \tag{11}$$

$$S_k = f\left(\sum_{j=1}^{J+1} w_{jk} H_j \right) \tag{12}$$

$$H_j = f\left(\sum_{i=1}^{I} w_{ij} U_i \right) \tag{13}$$

$$u_i = \frac{A_{\max} - A_{\min}}{V_{\max,i} - V_{\min,i}} \quad i = 1, 2 \tag{14}$$

$$U_i = u_i (V_i - V_{\min,i}) + A_{\min} \quad i = 1, 2 \tag{15}$$

$$u_i = \frac{A_{\max} - A_{\min}}{\ln(V_{\max,i}) - \ln(V_{\min,i})} \quad i = 3 \tag{16}$$

$$U_i = u_i [\ln(V_i) - \ln(V_{\min,i})] + A_{\min} \quad i = 3 \tag{17}$$

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

with $1 \leqslant i \leqslant I - 1$, $1 \leqslant j \leqslant J$ and $1 \leqslant k \leqslant K$, where 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 quantity V_i is the generic independent variable, and W_k is the generic dependent variable. The transfer function defined in Eq. (9) is recalled in Eqs. (12) and (13). In Eq. (12) the summation is over the $J + 1$ nodes of the hidden layer, in Eq. (13) it is over the I nodes of the input layer.

Due to the characteristics of the present problem, the MLFN parameters are set here to the values shown in the lower part of Table 6 in Appendix A. The input variables and the output function have thus both been compressed within the range 0.05–0.95. Given an experimental data set of output Nu_i, in the independent variables $(Re, Pr, Ec)_i$, the weighting factors are found by minimizing the following objective function by means of an optimization procedure

$$f_{\text{ob}} = \frac{1}{\text{NPT}} \sum_{i=1}^{\text{NPT}} \left(\frac{Nu_i^{\text{calc}} - Nu_i^{\text{exp}}}{Nu_i^{\text{exp}}} \right)^2 \tag{18}$$

where NPT stands for the number of experimental points on which the NN is trained.

After the training, the Nu equation is obtained as a continuous function of (Re, Pr, Ec). Two training sets were selected, obtaining two sets of weighting factors, Table 6, then originating two versions of the first NN architecture. The first version was trained on only one subset of the Olson and Allen data [8], which was randomly extracted for about 250 points on a total of 1114. The validity range of the NN model obtained is the same as for the reference data [8] on which the NN was trained. Those data quite homogeneously cover the following ranges: $30000 < Re < 160000$; $2 < Pr < 5$; $10^{-7} < Ec < 10^{-4}$. Since the method is totally heuristic, it is essential to have evenly distributed experimental data available; for the time being, the data from the present source [8] are the only ones that fulfill this requirement. It was decided not to train the NN on all available data irrespective of 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 the

MLFN, as a universal function approximator, needs to be applied to a compact data domain according to the Kolmogorov theorem [21]. The validity ranges mentioned were based on this criterion. The above requirement was confirmed by the fact that it proved impossible to reach a convergence of the model on the primitive data base. The second version of the first architecture was trained on an enlarged data set, including the subset of the source [8], plus the previously mentioned points from the other sources concerning the heating of carbon dioxide [1–7] that fall within the same range of dimensionless numbers (Re, Pr, Ec) as the Olson and Allen data [8]. The same weight was attributed to all the data points in the regression. The weighting factors and all the parameters needed to implement the NN are given in Table 6 for both the first and the second versions of the first NN architecture.

4. Neural networks in terms of physical variables

The second architecture directly represents the heat transfer coefficient α as a function of the controlling physical quantities as independent variables, which are the reduced pressure P_r, reduced temperature T_r, mass flow rate $\dot{m}$, and heat flux $\dot{q}$. Due to the modular structure of NNs, the same mathematical formalism described earlier for the first architecture, Eqs. (9)–(17) holds—though this architecture presents a different number of neurons in the input and hidden layers. The present architecture is similar to the one shown in Fig. 2: the four values of the input layer U_1, U_2, U_3 and U_4, represent the scaled independent variables, related, respectively, to each of the physical quantities cited. No logarithmic filter is applied to the input variables. The input layer has five neurons in all ($I = 5$), i.e., four input variables plus the first bias. For the hidden layer, the ideal number proved to be six neurons, plus the second bias ($J = 6$). In the output layer there is one neuron, representing the scaled heat transfer coefficient α. Hence

$$V_1 = P_r \quad V_2 = T_r \quad V_3 = \dot{m} \quad V_4 = \dot{q} \quad \text{and} \quad W_1 = \alpha$$

Given an experimental data set of output α_i, in the independent variables $(P_r, T_r, \dot{m}, \dot{q})_i$, the weighting factors are found by minimizing the following objective function.

$$f_{ob} = \frac{1}{NPT}\sum_{i=1}^{NPT}\left(\frac{\alpha_i^{calc} - \alpha_i^{exp}}{\alpha_i^{exp}}\right)^2 \tag{19}$$

After training, the heat transfer coefficient α can be expressed as a continuous function of $(P_r, T_r, \dot{m}, \dot{q})$. Two training sets were selected and two sets of weighting factors were obtained, giving rise to two versions of this second NN architecture. The first version was trained only on a subset of the data from [8]. The subset was randomly extracted and was composed of about 250 points on a total of 1114. The validity range of this NN model is the same as for the data [8] on which the NN was trained. The data quite homogeneously cover the following ranges of the physical quantities $(P_r, T_r, \dot{m}, \dot{q})$

$$1.04 < P_r < 1.80; \quad 0.94 < T_r < 1.03; \quad 170\ \text{kg/m}^2\text{s} < \dot{m} < 910\ \text{kg/m}^2\text{s}; \quad 12\ \text{kW/m}^2 < \dot{q} < 66\ \text{kW/m}^2.$$

The second version of the second architecture was trained on an enlarged data set, including the subset [8], plus the points from the other sources that fall within the same range of physical quantities $(P_r, T_r, \dot{m}, \dot{q})$ as the data set [8]. All the data points were assumed to have the same statistical weights for the regression. In Table 7 in Appendix A, the weighting factors and the necessary NN parameters are given for both the first and the second versions of this architecture.

5. Neural networks accounting for property variation

A third NN architecture was studied, that was similar to the first, but with slightly different inputs. From a functional point of view, this architecture strictly parallels the input/output variables of the KPPG conventional correlation. In the present case, the following four independent variables were assumed

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

whereas the dependent variable was still

$$W_1 = Nu$$

As a consequence, in this architecture it is $I = 5$, and $K = 1$. For the hidden layer, a number of six neurons, plus the second bias, was found to be optimal ($J = 6$). No logarithmic filter was applied to the input variables. For the mathematical formalism, reference is made, here again, to Eqs. (9)–(17), whereas the schematic representation is much the same as in Fig. 2.

In the context of the present architecture, the MLFN's capability as a general function approximator of a conventional heat transfer surface for the present problem was studied. From the KPPG correlation, about 8000 values of

$$Nu = Nu\left[Re, Pr, \left(\frac{\rho_w}{\rho_b}\right), \left(\frac{\bar{c}_P}{c_{P,b}}\right)\right]$$

were generated over a regular grid of the independent variables. The NN was trained on a subset of these values, regularly extracted from the whole base of values, and was then validated over the whole set of the 8000 values generated. With a training set composed of

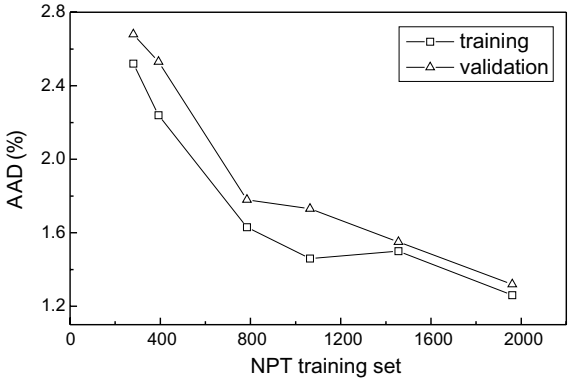

Fig. 3. Validation test of the MLFN as function approximator of the conventional KPPG heat transfer surface with decreasing number of the training set points.

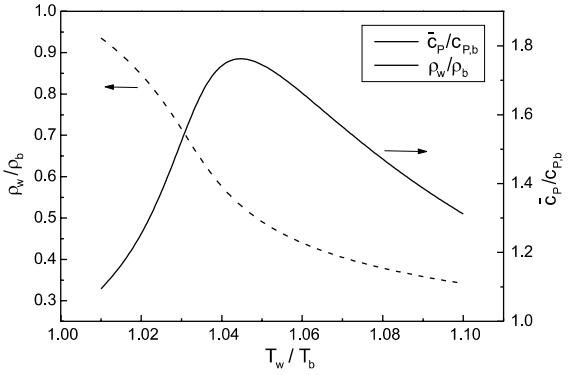

Fig. 4. Plot of the two properties groups (ρ_w/ρ_b) and $(\bar{c}_P/c_{P,b})$ as functions of the temperature ratio (T_w/T_b) at the pressure of 8 MPa.

only a fifth of all the values, the MLFN with six neurons in the hidden layer represents the conventional heat transfer surface with an AAD of about 1.5%. Reducing the number of values in the training set causes a corresponding reduction in MLFN approximation accuracy, as illustrated in Fig. 3, which also shows that about 700 values regularly extracted for the training set suffice to ensure an accurate reproduction of the heat transfer surface. This means that a limited experimental effort is needed to develop a specific heat transfer surface, providing that evenly distributed data are available in the range of variables of interest.

Two versions of this NN architecture were developed. One was trained only on a subset randomly extracted from the data [8], and was composed of about 250 points from a total of 1114. The range of validity of the NN model obtained is the same as for the data [8] on which the NN was trained. These data are distributed quite regularly over the ranges of the physical quantities $(Re, Pr, (\rho_w/\rho_b), (\bar{c}_P/c_{P,b}))$: $30000 < Re < 160000$; $2 < Pr < 5$; $0.2 < (\rho_w/\rho_b) < 1.0$; $0.3 < (\bar{c}_P/c_{P,b}) < 3.15$. The second version of the architecture was trained on an enlarged data set, including the subset of data [8] and the points from the other sources that fell within the same range of dimensionless numbers $(Re, Pr, (\rho_w/\rho_b), (\bar{c}_P/c_{P,b}))$ as the data [8]. The weighting factors and all the parameters needed to implement the NN are given in Table 8 in Appendix A for both the first and the second versions of this NN architecture.

In the context of the radial variation in the thermophysical properties, another architecture can be set up, in much the same way as the second architecture was developed from the first. The independent variables of the last architecture are actually dependent functions of some physical controlling variables. The present general aim is the development of a new relation, similar in the structure to the third architecture, but in which only the physical independent variables have to be considered. In the evolution from the first to the second case it was demonstrated that the controlling variables of Re and Pr are $(P_r, T_r, \dot{m})$. Considering the second architecture the heat flux $\dot{q}$ was also included as a variable. The radial variation of the thermophysical properties has to be furthermore included. For this purpose it can be demonstrated that the two property ratios (ρ_w/ρ_b), $(\bar{c}_P/c_{P,b})$ are dependent functions of the variables T_w/T_b, T_r and P_r. In Fig. 4, these two functions are plotted, as an example, against T_w/T_b at $P = 8$ MPa and $T_b = 297.69$ K, assuming the carbon dioxide dedicated EoS [10], and the said dependence is clearly shown. The general heat transfer relation is moreover valid

$$\alpha = \frac{\dot{q}}{T_b\left(\frac{T_w}{T_b} - 1\right)} \tag{20}$$

with $T_b = T_r \cdot T_c$, from which it becomes evident that the four quantities α, $\dot{q}$, T_b and T_w are not independent of each other, but $\dot{q}$ is related to α through a function of the variables T_w/T_b and T_r, Eq. (20). The quantity $\dot{q}$ in the present context is then a dependent function of T_w/T_b and T_r and it can be consequently replaced by such variables in the functional form of α.

Consequently, a fourth architecture can be considered in which the independent variables are the reduced pressure P_r, reduced temperature T_r, mass flow rate $\dot{m}$, and wall-to-bulk temperature ratio

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

while the dependent variable is still

$$W_1 = \alpha$$

Two training sets were selected and two versions of the fourth architecture were obtained as usual. The first

version was trained only on a subset of Olson and Allen data [8]. For the hidden layer a number of six neurons, plus the second bias, was found to be optimal ($J = 6$). The range of validity of this NN correlation is the same as for the reference [8] used for the NN training. The data regularly cover the ranges of physical quantities $(P_r, T_r, \dot{m}, (T_w/T_b))$: $1.04 < P_r < 1.80$; $0.94 < T_r < 1.03$; $170\ \mathrm{kg/m^2\,s} < \dot{m} < 910\ \mathrm{kg/m^2\,s}$; $1.0 < (T_w/T_b) < 1.23$. The second version of the fourth architecture was trained on an enlarged data set, including the subset of Olson and Allen data [8], plus the points from the other sources that fell within the same range of physical quantities $(P_r, T_r, \dot{m}, (T_w/T_b))$ as the reference data [8]. The weighting factors, together with the NN parameters, are given in Table 9 of Appendix A for both the first and the second versions of the present architecture.

6. Validation results

All four NN models described in the previous sections were tested. First of all, the data sources were validated against the conventional equation in order to check their consistency. The validation results of the KPPG conventional equation, Eqs. (1)–(7), with respect to the literature sources considered [1–8], are given in Table 1. In the present work, the error deviation ($\Delta\%$), average absolute deviation (AAD%) and bias (Bias%) are evaluated as

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

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

$$\mathrm{Bias}\% = \frac{1}{\mathrm{NPT}} \sum_{i=1}^{\mathrm{NPT}} (\Delta\%)_i \qquad (21)$$

Table 1
Validation of the conventional KPPG equation[a]

Flow direction	NPT	AAD (%)	Bias (%)	Max (%)	Ref.
H	1114	3.50	3.03	15.86	8
H	63	11.31	−8.73	−31.05	2
H	35	15.27	−15.27	−22.60	3
H	69	23.23	−16.89	−68.45	1
H	36	10.94	7.57	29.58	4
H	27	32.19	−32.19	−71.66	5
H	90	24.47	12.08	75.95	6
U	28	17.34	−16.86	−41.47	7
Overall	1462	7.32	0.79	75.95	

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

where x is intended as a dependent variable, such as Nu or α, and exp and calc stand for experimental and calculated values, respectively.

The deviations lie around the zero line, but are quite large. This formulation was originally developed on data sets regarding the heating of carbon dioxide and it reproduces this kind of data with an AAD generally less than 10%. The Olson and Allen data sets [8] are the most recent measurements for carbon dioxide heating and they show only minimal deviations from the KPPG conventional equation, by comparison with other data sources, Table 1.

Validation of the NN model in terms of dimensionless numbers (the first architecture) has been conducted only on the data points falling within the validity range of the model. For those points, a parallel validation was conducted against the conventional equation for comparison and the results are given in Table 2 for the first and the second versions. The first version very adequately represents the data source [8], on a subset of which it was trained, with an AAD of 1.55%, which can be compared with the AAD of 3.50% of the conventional equation. Unfortunately, the first version is generally worse than the conventional equation in the representation of sources [1–7] not included in the training set. In the second version of the first NN architecture, some points from other sources were introduced in the regression, with positive and negative consequences: the sources [1–7] are reproduced much better than in the first version, but the representation of the source [8] is worse. The AAD of the second version for this latter source is 9.26%, compared to 1.55% of the first version. Looking at the deviations of all the sources [1–8] from the second version of the first NN architecture, some points from reference [8] present an extremely high deviation, greater than 100%, whereas the same points were well represented in the first version. Adding the data coming from other sources [1–7] in the regression also presents a drawback. This certainly indicates a discrepancy between the data coming from different sources, at least in terms of the dimensionless numbers form $Nu = Nu(Re, Pr, Ec)$. From the present study, it is impossible to infer the origin of any such discrepancy, so it would be best not to split the data into primary and secondary on the base of their claimed accuracy. On the other hand, the positive consequence of introducing other sources in the regression is that many sources show a smaller AAD, for the second version, than the conventional equation. This might also be explained by an excessive scatter of the data base adopted, notwithstanding the limitations imposed on the range of the independent variables.

A similar validation was done for the NN model in terms of physical quantities (second architecture). The results are shown in Table 3 for the first and the second versions. Here again, this first version represents the

Table 2
First architecture MLFN model validation[a,b]

NPT	NPT[1]	Conventional KPPG equation			MLFN 1st arch., 1st vers.			MLFN 1st arch., 2nd vers.			Ref.
		AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	
1114	1114	3.50	3.04	15.86	1.55	0.08	−12.07	9.26	6.91	120	8
63	20	9.57	−6.24	−23.05	47.43	−47.43	−61	23.92	−14.38	59.12	2
35	0	–	–	–	–	–	–	–	–	–	3
69	37	25.13	−15.83	−46.05	49.98	−49.35	−64	6.05	−1.90	−21.80	1
36	18	13.13	7.25	29.57	9.66	9.14	22.3	8.95	4.44	24.75	4
27	5	26.44	−26.44	−33.47	54.41	−54.41	−68.82	19.1	−5.82	33.2	5
90	7	5.36	3.66	11.83	15.96	−15.96	−24.42	14.57	13.32	31.95	6
28	0	–	–	–	–	–	–	–	–	–	7
1462	1201	4.52	2.25	−46.05	4.23	−2.41	−68.82	9.48	6.23	120	

[a] See footnote at Table 1.
[b] NPT[1], the number of points within the range of validity of the first architecture NN model;
NPT[2], the number of points within the range of validity of the second architecture NN model;
NPT[3], the number of points within the range of validity of the third architecture NN model;
NPT[4], the number of points within the range of validity of the fourth architecture NN model.

Table 3
Second architecture MLFN model validation[a,b]

NPT	NPT[2]	Conventional KPPG equation			MLFN 2nd arch., 1st vers.			MLFN 2nd arch., 2nd vers.			Ref.
		AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	
1114	1114	3.50	3.04	15.86	1.72	0.61	13.01	6.59	4.55	113	8
63	0	–	–	–	–	–	–	–	–	–	2
35	0	–	–	–	–	–	–	–	–	–	3
69	0	–	–	–	–	–		–	–	–	1
36	9	7.88	3.41	−13.99	16.16	16.16	20.30	3.85	2.76	9.30	4
27	0	–	–	–	–	–	–	–	–	–	5
90	41	22.36	5.70	75.95	27.50	−16.06	−75.01	10.20	2.05	41.65	6
28	0	–	–	–	–	–	–	–		–	7
1462	1164	4.20	3.13	75.95	2.74	0.15	−75.01	6.70	4.45	113	

[a] See footnote at Table 1.
[b] See footnote at Table 2.

data source [8] quite well, with an AAD of 1.72% compared with the 3.50% AAD of the KPPG conventional equation. Unluckily, the performance of the first version in representing the external sources [1–7] is always worse than that of the conventional equation.

For the second version, the same can be said as for the first NN architecture in terms of dimensionless numbers. The AAD for the source [8] reaches 6.59%, but the external sources in this case are better represented than by the conventional equation. Some points from the source [8] show an extremely high deviation for the second version, while they are well represented by the first. This is exactly what happened in the first NN architecture. It can thus be concluded that the data sources are not consistent with each other in terms of physical quantities either $\alpha = \alpha(P_r, T_r, \dot{m}, \dot{q})$.

The data sources *are* consistent when compared with the conventional equation, however, Table 1. In fact, the conventional equation was able to offer a very good representation of the source [8] as well as a fair representation of all the external sources. On the other hand, NNs offer an extremely versatile representation tool in this case. The partial failure of the NN, both in the first and in the second architecture, suggests that any heuristic model in the form of $Nu = Nu(Re, Pr, Ec)$ or $\alpha = \alpha(P_r, T_r, \dot{m}, \dot{q})$ is very likely to fail. The success of the conventional, semi-empirical equation is due, in our opinion, to the introduction of the Krasnoshchekov and Protopopov terms in the analytical form, Eqs. (3)–(7), which take into account the property variation of the fluid along the radial co-ordinate. This cannot be done properly by the two NN architectures proposed, even

though a low *Ec* number in the first architecture, or a high heat flux in the second, is certainly related to a high radial temperature gradient, i.e., to a pronounced temperature profile along the radial co-ordinate, and thus to large thermophysical property variations for the fluid. But neither the first architecture nor the second incorporate information on the thermophysical property variations in the radial direction, which have to be supplied by a highly accurate thermodynamic equation of state for the fluid, as in the KPPG conventional equation. For this particular heat transfer problem, the radial change in the thermophysical properties is such an important aspect that any model has to take this into due account, thus making it necessary to use a dedicated equation of state for the fluid.

The third architecture overcomes the discrepancy, taking the property variation along the radial co-ordinate into account by introducing the groups (ρ_w/ρ_b) and $(\bar{c}_P/c_{P,b})$, which are also in the KPPG conventional equation. The results are shown in Table 4 for the first and second versions of this architecture. The data source [8] is well represented by this first version, with an AAD of 2.98%. The representation of the external sources is globally on a par with the conventional equation. So this third architecture can predict other data sources outside the training set. In the second version of this architecture, some points from external sources were introduced in the regression and this surprisingly failed to improve the representation of either the data [8] or the external sources. It may be that data noise interferes with regression accuracy. The AAD of the second version for the source [8] is 8.81% versus 2.98% for the first version, and the maximum deviation is much smaller than for the first and second architectures.

The results obtained with the fourth architecture are presented in Table 5. In this case too, two versions of the same architecture were devised, depending on the data set used to train the NN, i.e., a subset from ref-

Table 4
Third architecture MLFN model validation[a,b]

NPT	NPT[3]	Conventional KPPG equation			MLFN 3rd arch., 1st vers.			MLFN 3rd arch., 2nd vers.			Ref.
		AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	
1114	1114	3.50	3.03	15.85	2.98	0.29	12.46	8.81	2.68	31.88	8
63	3	8.19	− 8.19	− 9.45	17.40	− 17.40	− 21.86	21.71	− 21.71	− 22.09	2
35	0	–	–	–	–	–	–	–	–	–	3
69	25	21.94	− 12.44	− 39.61	39.83	− 21.87	57.27	18.32	− 4.72	38.39	1
36	18	13.13	7.15	29.58	12.33	12.33	27.26	24.41	23.03	46.58	4
27	1	25.95	− 25.95	25.95	18.81	18.81	18.81	13.14	13.14	13.14	5
90	7	5.36	3.66	11.83	5.81	− 5.81	− 13.03	20.54	15.21	51.13	6
28	0	–	–	–	–	–	–	–	–	–	7
1462	1168	4.09	2.72	−39.61	3.98	−0.06	57.27	9.08	2.86	51.13	

[a,b] See footnotes at Tables 1 and 2.

Table 5
Fourth architecture MLFN model validation[a,b]

NPT	NPT[4]	Conventional KPPG equation			MLFN 4th arch., 1st vers.			MLFN 4th arch., 2nd vers.			Ref.
		AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	AAD (%)	Bias (%)	Max (%)	
1114	1114	3.50	3.04	15.86	1.19	0.05	−24.81	7.05	−4.85	−74.69	8
63	0	–	–	–	–	–	–	–	–	–	2
35	0	–	–	–	–	–	–	–	–	–	3
69	0	–	–	–	–	–	–	–	–	–	1
36	21	11.49	5.98	29.58	17.80	17.80	31.96	8.33	5.50	27.64	4
27	0	–	–	–	–	–	–	–	–	–	5
90	41	22.36	5.70	75.95	35.02	−21.28	−71.45	6.37	8.64	22.32	6
28	0	–	–	–	–	–	–	–	–	–	7
1462	1176	4.30	3.18	75.95	2.67	−0.38	−71.45	7.04	−4.45	−74.89	

[a,b] See footnotes at Tables 1 and 2.

erence [8] alone or a combination of [8] and points from available sources coming within the range of variables of reference [8]. This first version also represents the data source [8] quite well, with an AAD of 1.19%. Unfortunately, the first version is always worse than the conventional equation in representing the other sources. In the second version, for the source [8] the AAD rises to 7.05%, but the other sources are represented better than by the conventional equation. It should be noted that, in the second version, some points from the source [8] show an extremely high deviation, while they were well represented in the first version. Clearly, the representations of the heat transfer coefficient as a function of physical variables, with and without radial correction, as in the second and fourth architectures, are the more effective. This also suggests that using the traditional dimensionless numbers in the correlation offers no apparent advantage, with the important drawback of needing the thermodynamic and transport property equations for the target fluid. Considering the data sets available, the data can only be considered consistent for the source [8] since comparable results are obtained for these data with all the architectures studied. The introduction of other sources was found to cause a significant loss of effectiveness for all the heat transfer coefficient surfaces, suggesting that said sources are inconsistent. It is also worth noting that the proposed method can be used to check the consistency of new data sets before using them for processing. Because the method is totally heuristic, the data accuracy and distribution as a function of the independent variables and the experimental technique are fundamental to ensuring proper results. But if these requirements are strictly met, the heat transfer coefficient surface for the case in point is straightforward to obtain with an uncertainty lower than the experimental error. Only the controlling independent variables have to be precisely identified.

In conclusion, the NN heat transfer models in the form of the third and of the fourth architectures both appear to offer a coherent description of the data sets at least as accurate as the KPPG conventional equation, but with a completely heuristic procedure, avoiding any physical assumption for the model. Among the two the fourth architecture is the more synthetic and most of all does not require any modelling for the thermodynamic and transport properties of the fluid.

7. Discussion and conclusions

The proposed method has been shown to be very effective in representing a heat transfer coefficient surface for the supercritical heating of carbon dioxide, also demonstrating that the method may be highly suitable for heuristically deriving a heat transfer surface from experimental data on the heat transfer coefficient alone.

The preliminary correlation test was successful, also considering the absence of error noise in the virtual data generated by the KPPG conventional equation. When the method is applied to the real context of experimental data in the literature, error noise becomes a very important drawback that reduces the final accuracy of the regressed surface. As in any heuristic method, studying 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; scattered data sets should be avoided within the validity ranges of the equation. Moreover, the data points should be distributed regularly on an ideal grid of the independent variables. These conditions are lacking for the data currently available in the literature, also because heuristic methods have never been considered in the past for heat transfer study, apart from rare cases such as the two previously cited [17,18].

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

It is only if regularly distributed data are available for several fluids, with overlapping ranges, that the generalization problem of a transfer equation could be studied in more detail, but these conditions are very difficult to encounter in the experimental works available—suggesting that the problem would be worth examining experimentally from this new point of view. In addition to the data distribution, the experimental quality of the data also becomes a key element in the present case. The enormous variation of c_P, λ, and Pr along the test section at near-critical pressures, Fig. 1, suggests avoiding the assumption of mean values of the heat transfer coefficient over considerable lengths of the test section, i.e., with considerable temperature changes: in fact, several sources reveal major variations in the coefficient with the axial length [1,5]. On the other hand, the use of electrical heating in the test section does not guarantee a uniform heat flux on the tube's circumference.

Where regularly distributed and very precise data are available, the proposed method proves capable of drawing a highly accurate heat transfer correlation for the present difficult problem.

Appendix A

Table 6
MLFN model parameters

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	0.371897×10^3	1	1	-0.351973×10^3	1	1	-0.586286×10^4	1	1	-0.504978×10^2
2	1	0.564132×10^3	2	1	-0.260067×10^2	2	1	0.183874×10^4	2	1	-0.458490×10^4
3	1	-0.166154×10^4	3	1	-0.532545×10^3	3	1	0.910815×10^3	3	1	0.154596×10^4
4	1	0.588122×10^3	4	1	-0.494797×10^3	4	1	0.138221×10^4	4	1	0.231015×10^4
1	2	-0.132091×10^3	5	1	-0.147016×10^3	1	2	0.305915×10^3	5	1	-0.140051×10^4
2	2	0.617158×10^2	6	1	0.540311×10^3	2	2	0.461092×10^4	6	1	-0.110874×10^4
3	2	-0.834266×10^2	7	1	0.509395×10^3	3	2	0.812996×10^4	7	1	0.115722×10^4
4	2	-0.408619×10^3	8	1	0.754598×10^2	4	2	-0.809621×10^4	8	1	-0.212404×10^3
1	3	-0.406462×10^3				1	3	-0.949439×10^3			
2	3	-0.829393×10^3				2	3	0.113380×10^3			
3	3	0.142816×10^4				3	3	0.254532×10^4			
4	3	-0.414260×10^3				4	3	-0.112275×10^4			
1	4	0.868087×10^2				1	4	0.577760×10^3			
2	4	0.475781×10^3				2	4	0.621289×10^4			
3	4	-0.540806×10^3				3	4	0.938256×10^4			
4	4	0.236564×10^3				4	4	-0.972685×10^4			
1	5	0.287985×10^2				1	5	-0.108878×10^4			
2	5	-0.122434×10^3				2	5	-0.209075×10^2			
3	5	-0.181134×10^3				3	5	0.277226×10^4			
4	5	-0.292215×10^3				4	5	-0.115927×10^4			
1	6	0.149757×10^3				1	6	-0.110367×10^5			
2	6	0.438235×10^3				2	6	-0.403938×10^4			
3	6	0.134736×10^3				3	6	-0.530483×10^4			
4	6	-0.493351×10^3				4	6	0.883169×10^4			
1	7	-0.486437×10^3				1	7	-0.108021×10^5			
2	7	-0.246169×10^3				2	7	-0.372409×10^4			
3	7	0.364132×10^3				3	7	-0.459035×10^4			
4	7	0.441017×10^3				4	7	0.821134×10^4			

$V_{\min,1} \equiv Re^{\min}$	0	$V_{\min,1} \equiv Re^{\min}$	0
$V_{\max,1} \equiv Re^{\max}$	200000	$V_{\max,1} \equiv Re^{\max}$	200000
$V_{\min,2} \equiv Pr^{\min}$	0	$V_{\min,2} \equiv Pr^{\min}$	0
$V_{\max,2} \equiv Pr^{\max}$	10	$V_{\max,2} \equiv Pr^{\max}$	10
$V_{\min,3} \equiv Ec^{\min}$	1×10^{-8}	$V_{\min,3} \equiv Ec^{\min}$	1×10^{-8}
$V_{\max,3} \equiv Ec^{\max}$	1×10^{-3}	$V_{\max,3} \equiv Ec^{\max}$	1×10^{-3}
$W_{\min,1} \equiv Nu^{\min}$	0	$W_{\min,1} \equiv Nu^{\min}$	0
$W_{\max,1} \equiv Nu^{\max}$	2000	$W_{\max,1} \equiv Nu^{\max}$	2000
J	7	J	7
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\min}$	0.05	$A_{\min}$	0.05
$A_{\max}$	0.95	$A_{\max}$	0.95
Bias 1	1.0	Bias 1	1.0
Bias 2	1.0	Bias 2	1.0
I	4	I	4
K	1	K	1
NPT training	254	NPT training	341
Training residual AAD	1.14%	Training residual AAD	9.10%

Architecture I in terms of dimensionless numbers.

Table 7
MLFN model parameters

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	-0.248015×10^4	1	1	0.284703×10^4	1	1	-0.706131×10^3	1	1	0.488865×10^3
2	1	0.283736×10^4	2	1	0.116879×10^4	2	1	0.689745×10^3	2	1	0.784691×10^3
3	1	0.232546×10^3	3	1	0.465801×10^3	3	1	-0.480155×10^2	3	1	-0.126974×10^4
4	1	-0.858384×10^2	4	1	0.665010×10^3	4	1	-0.222978×10^3	4	1	0.160950×10^4
5	1	-0.168410×10^4	5	1	-0.143549×10^4	5	1	-0.293807×10^3	5	1	0.139591×10^4
1	2	0.854304×10^3	6	1	-0.610317×10^3	1	2	0.268755×10^2	6	1	-0.998799×10^3
2	2	-0.148742×10^4	7	1	0.169050×10^3	2	2	0.178970×10^3	7	1	-0.113385×10^4
3	2	0.533944×10^3				3	2	0.908931×10^2			
4	2	-0.861586×10^3				4	2	0.139885×10^3			
5	2	0.737809×10^3				5	2	-0.157490×10^3			
1	3	-0.901879×10^3				1	3	-0.700956×10^3			
2	3	0.156283×10^4				2	3	0.195208×10^3			
3	3	-0.687032×10^3				3	3	-0.969243×10^3			
4	3	0.119316×10^4				4	3	0.923140×10^3			
5	3	-0.894317×10^3				5	3	-0.160975×10^3			
1	4	-0.292466×10^2				1	4	0.636428×10^3			
2	4	0.933875×10^2				2	4	-0.599891×10^4			
3	4	0.257000×10^3				3	4	0.128718×10^3			
4	4	0.421124×10^2				4	4	-0.263281×10^3			
5	4	0.406214×10^2				5	4	0.304534×10^4			
1	5	0.155543×10^3				1	5	-0.759045×10^3			
2	5	-0.412234×10^3				2	5	0.684929×10^4			
3	5	-0.315474×10^3				3	5	-0.135139×10^3			
4	5	0.852603×10^2				4	5	0.233377×10^3			
5	5	0.734623×10^3				5	5	-0.343108×10^4			
1	6	0.963893×10^3				1	6	0.812007×10^3			
2	6	-0.166760×10^4				2	6	-0.376069×10^3			
3	6	0.431571×10^3				3	6	0.118077×10^4			
4	6	-0.711931×10^3				4	6	-0.114534×10^4			
5	6	0.726561×10^3				5	6	0.264356×10^3			

$V_{\min,1} \equiv P_r^{\min}$	1.0	$V_{\min,1} \equiv P_r^{\min}$	1.0
$V_{\max,1} \equiv P_r^{\max}$	2.0	$V_{\max,1} \equiv P_r^{\max}$	2.0
$V_{\min,2} \equiv T_r^{\min}$	0.9	$V_{\min,2} \equiv T_r^{\min}$	0.9
$V_{\max,2} \equiv T_r^{\max}$	1.1	$V_{\max,2} \equiv T_r^{\max}$	1.1
$V_{\min,3} \equiv \dot{m}^{\min}$	0	$V_{\min,3} \equiv \dot{m}^{\min}$	0
$V_{\max,3} \equiv \dot{m}^{\max}$	1000	$V_{\max,3} \equiv \dot{m}^{\max}$	1000
$V_{\min,4} \equiv \dot{q}^{\min}$	0	$V_{\min,4} \equiv \dot{q}^{\min}$	0
$V_{\max,4} \equiv \dot{q}^{\max}$	100	$V_{\max,4} \equiv \dot{q}^{\max}$	100
$W_{\min,1} \equiv h^{\min}$	0	$W_{\min,1} \equiv h^{\min}$	0
$W_{\max,1} \equiv h^{\max}$	10	$W_{\max,1} \equiv h^{\max}$	10
J	6	J	6
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\min}$	0.05	$A_{\min}$	0.05
$A_{\max}$	0.95	$A_{\max}$	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	254	NPT training	295
Training residual AAD	0.72%	Training residual AAD	3.3%

Architecture II in terms of physical variables.

Table 8
MLFN model parameters

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	0.204538×10^{3}	1	1	0.626690×10^{2}	1	1	-0.185090×10^{3}	1	1	0.215557×10^{3}
2	1	-0.742302×10^{2}	2	1	-0.313530×10^{3}	2	1	0.155080×10^{3}	2	1	0.148316×10^{3}
3	1	0.881667×10^{2}	3	1	-0.119367×10^{3}	3	1	0.451747×10^{3}	3	1	-0.235078×10^{3}
4	1	-0.517582×10^{2}	4	1	0.117205×10^{3}	4	1	-0.150762×10^{3}	4	1	0.990889×10^{2}
5	1	0.586651×10^{2}	5	1	0.424826×10^{3}	5	1	0.767209×10^{2}	5	1	-0.397845×10^{3}
1	2	-0.202361×10^{3}	6	1	-0.346648×10^{3}	1	2	-0.255694×10^{2}	6	1	-0.215425×10^{3}
2	2	-0.394960×10^{3}	7	1	-0.574151×10^{2}	2	2	0.103816×10^{3}	7	1	0.793730×10^{2}
3	2	0.426323×10^{1}				3	2	-0.374707×10^{3}			
4	2	-0.217891×10^{3}				4	2	0.288729×10^{3}			
5	2	0.291822×10^{3}				5	2	-0.491152×10^{2}			
1	3	-0.160170×10^{3}				1	3	-0.449449×10^{3}			
2	3	0.906126×10^{2}				2	3	-0.210870×10^{2}			
3	3	-0.190610×10^{3}				3	3	-0.856748×10^{2}			
4	3	0.197592×10^{3}				4	3	-0.197153×10^{3}			
5	3	-0.136176×10^{2}				5	3	0.302799×10^{3}			
1	4	-0.332551×10^{3}				1	4	0.197387×10^{3}			
2	4	0.137973×10^{1}				2	4	-0.138814×10^{3}			
3	4	-0.224856×10^{3}				3	4	-0.351408×10^{1}			
4	4	-0.145620×10^{3}				4	4	0.101890×10^{3}			
5	4	-0.185221×10^{3}				5	4	-0.286913×10^{2}			
1	5	0.139882×10^{3}				1	5	0.625080×10^{2}			
2	5	-0.386978×10^{3}				2	5	0.724213×10^{3}			
3	5	0.205465×10^{3}				3	5	0.318383×10^{3}			
4	5	0.232308×10^{3}				4	5	-0.136090×10^{3}			
5	5	0.264479×10^{3}				5	5	0.170237×10^{2}			
1	6	0.253937×10^{3}				1	6	-0.621465×10^{2}			
2	6	0.122175×10^{3}				2	6	0.106948×10^{3}			
3	6	0.140451×10^{2}				3	6	0.493704×10^{2}			
4	6	0.191574×10^{3}				4	6	-0.191543×10^{3}			
5	6	0.315225×10^{3}				5	6	-0.400852×10^{3}			

$V_{\min,1} \equiv Re^{\min}$	0	$V_{\min,1} \equiv Re^{\min}$	0
$V_{\max,1} \equiv Re^{\max}$	200000	$V_{\max,1} \equiv Re^{\max}$	200000
$V_{\min,2} \equiv Pr^{\min}$	0	$V_{\min,2} \equiv Pr^{\min}$	0
$V_{\max,2} \equiv Pr^{\max}$	10	$V_{\max,2} \equiv Pr^{\max}$	10
$V_{\min,3} \equiv (\rho_w/\rho_b)^{\min}$	0.0	$V_{\min,3} \equiv (\rho_w/\rho_b)^{\min}$	0.0
$V_{\max,3} \equiv (\rho_w/\rho_b)^{\max}$	1.0	$V_{\max,3} \equiv (\rho_w/\rho_b)^{\max}$	1.0
$V_{\min,4} \equiv (\bar{c}_P/c_{P,b})^{\min}$	0.0	$V_{\min,4} \equiv (\bar{c}_P/c_{P,b})^{\min}$	0.0
$V_{\max,4} \equiv (\bar{c}_P/c_{P,b})^{\max}$	5.0	$V_{\max,4} \equiv (\bar{c}_P/c_{P,b})^{\max}$	5.0
$W_{\min,1} \equiv Nu^{\min}$	0	$W_{\min,1} \equiv Nu^{\min}$	0
$W_{\max,1} \equiv Nu^{\max}$	2000	$W_{\max,1} \equiv Nu^{\max}$	2000
J	6	J	6
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\min}$	0.05	$A_{\min}$	0.05
$A_{\max}$	0.95	$A_{\max}$	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	254	NPT training	308
Training residual AAD	2.87%	Training residual AAD	8.90%

Architecture III accounting for property variations in terms of thermophysical properties.

Table 9
MLFN model parameters

First version						Second version					
i	j	w_{ij}	j	k	w_{jk}	i	j	w_{ij}	j	k	w_{jk}
1	1	-0.974960×10^{3}	1	1	-0.100261×10^{4}	1	1	-0.614630×10^{3}	1	1	0.480367×10^{3}
2	1	0.916281×10^{3}	2	1	-0.399416×10^{3}	2	1	0.357762×10^{4}	2	1	0.218410×10^{4}
3	1	0.261515×10^{2}	3	1	0.118443×10^{3}	3	1	-0.542253×10^{3}	3	1	-0.953904×10^{3}
4	1	-0.643750×10^{3}	4	1	-0.221042×10^{4}	4	1	-0.128948×10^{3}	4	1	0.193485×10^{4}
5	1	-0.456915×10^{3}	5	1	0.553850×10^{3}	5	1	-0.138404×10^{4}	5	1	-0.881437×10^{3}
1	2	0.367568×10^{2}	6	1	0.755857×10^{3}	1	2	0.284665×10^{3}	6	1	-0.142724×10^{4}
2	2	-0.220521×10^{3}	7	1	0.158187×10^{3}	2	2	-0.148327×10^{4}	7	1	0.184221×10^{3}
3	2	-0.244408×10^{3}				3	2	0.573711×10^{3}			
4	2	0.802452×10^{2}				4	2	0.674116×10^{2}			
5	2	0.432686×10^{3}				5	2	0.423689×10^{3}			
1	3	0.991695×10^{1}				1	3	0.608974×10^{3}			
2	3	0.210648×10^{2}				2	3	-0.674063×10^{3}			
3	3	0.654687×10^{3}				3	3	-0.234528×10^{3}			
4	3	-0.108778×10^{3}				4	3	0.142380×10^{4}			
5	3	-0.176984×10^{3}				5	3	0.402733×10^{3}			
1	4	-0.492360×10^{3}				1	4	-0.223181×10^{4}			
2	4	0.103612×10^{3}				2	4	0.717118×10^{3}			
3	4	0.856281×10^{3}				3	4	-0.147643×10^{4}			
4	4	-0.295894×10^{4}				4	4	-0.968385×10^{4}			
5	4	-0.108208×10^{2}				5	4	0.106785×10^{4}			
1	5	-0.159922×10^{4}				1	5	-0.421529×10^{3}			
2	5	0.201110×10^{4}				2	5	0.436418×10^{3}			
3	5	0.963289×10^{2}				3	5	0.176292×10^{3}			
4	5	-0.306407×10^{4}				4	5	-0.292709×10^{4}			
5	5	-0.610476×10^{3}				5	5	-0.166791×10^{3}			
1	6	-0.602721×10^{3}				1	6	0.245427×10^{3}			
2	6	0.435930×10^{3}				2	6	-0.114848×10^{4}			
3	6	0.759750×10^{2}				3	6	0.640904×10^{3}			
4	6	-0.773608×10^{3}				4	6	0.856720×10^{2}			
5	6	-0.200099×10^{3}				5	6	0.227289×10^{3}			

$V_{\min,1}\equiv P_{r}^{\min}$	1.0	$V_{\min,1}\equiv P_{r}^{\min}$	1.0
$V_{\max,1}\equiv P_{r}^{\max}$	2.0	$V_{\max,1}\equiv P_{r}^{\max}$	2.0
$V_{\min,2}\equiv T_{r}^{\min}$	0.9	$V_{\min,2}\equiv T_{r}^{\min}$	0.9
$V_{\max,2}\equiv T_{r}^{\max}$	1.1	$V_{\max,2}\equiv T_{r}^{\max}$	1.1
$V_{\min,3}\equiv \dot{m}^{\min}$	0	$V_{\min,3}\equiv \dot{m}^{\min}$	0
$V_{\max,3}\equiv \dot{m}^{\max}$	1000	$V_{\max,3}\equiv \dot{m}^{\max}$	1000
$V_{\min,4}\equiv (T_{w}/T_{b})^{\min}$	1.0	$V_{\min,4}\equiv (T_{w}/T_{b})^{\min}$	1.0
$V_{\max,4}\equiv (T_{w}/T_{b})^{\max}$	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
$W_{\max,1}\equiv h^{\max}$	10	$W_{\max,1}\equiv h^{\max}$	10
J	6	J	6
β	1.0	β	1.0
γ	0.005	γ	0.005
$A_{\min}$	0.05	$A_{\min}$	0.05
$A_{\max}$	0.95	$A_{\max}$	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	254	NPT training	316
Training residual AAD	0.93%	Training residual AAD	3.10%

Architecture IV accounting for property variations in terms of temperatures ratio.

References

[1] L.B. Koppel, J.M. Smith, Turbulent heat transfer in the critical region, ASME Int. Dev. Heat Transfer 3 (1961) 585–590.

[2] E.A. Krasnoshchekov, V.S. Protopopov, Experimental study of heat exchange in carbon dioxide in the supercritical range at high temperature drops, Teplof. Vysokikh Temp. 4 (3) (1966) 389–398.

[3] E.A. Krasnoshchekov, V.S. Protopopov, I.A. Parkhovnik, V.A. Silin, Some results of an experimental investigation of heat transfer to carbon dioxide at supercritical pressure and temperature heads up to 850 °C, Teplof. Vysokikh Temp. 9 (4) (1971) 1081–1084.

[4] B.S. Petukhov, E.A. Krasnoshchekov, V.S. Protopopov, An investigation of heat transfer to fluids flowing in pipes under supercritical conditions, ASME Int. Dev. Heat Transfer 3 (1961) 569–578.

[5] N.M. Schnurr, Heat transfer to carbon dioxide in the immediate vicinity of the critical point, ASME J. Heat Transfer 16 (1969) 16–20.

[6] H. Tanaka, N. Nishiwaki, M. Hirata, Turbulent heat transfer to supercritical carbon dioxide, in: Proceedings of the JSME Semi-International Symposium, Tokyo, 1967, pp. 127–134.

[7] P.J. Bourke, D.J. Pulling, L.E. Gill, W.H. Denton, Forced convective heat transfer to turbulent CO_2 in the supercritical region, Int. J. Heat Mass Transfer 13 (1970) 1339–1348.

[8] D.A. Olson, D. Allen, Heat transfer in turbulent supercritical carbon dioxide flowing in a heated horizontal tube, NISTIR 6234 Report, NIST, Gaithersburg, USA, 1998.

[9] 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.

[10] 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.

[11] 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.

[12] A. Fenghour, W.A. Wakeham, V. Vesovic, The viscosity of carbon dioxide, J. Phys. Chem. Ref. Data 27 (1) (1998) 31–44.

[13] J. Millat, J.H. Dymond, C.A. Nieto de Castro, Transport Properties of Fluids, Cambridge University Press, Cambridge, UK, 1996.

[14] J.G. Afshin, A. Asadollah, Improved forced convective heat-transfer correlations for liquids in the near-critical region, AIAA J. 24 (12) (1986) 2030–2037.

[15] S.S. Pitla, D.M. Robinson, E.A. Groll, S. Ramadhyani, Heat transfer from supercritical carbon dioxide in tube flow: a critical review, HVAC&R Res. 4 (3) (1998) 281–301.

[16] V. Gnielinski, New equation for heat transfer in turbulent pipe and channel flow, Int. Chem. Eng. 16 (2) (1976) 359–368.

[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.

[19] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control, Signals Syst. 2 (1989) 303–314.

[20] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks 2 (1989) 359–366.

[21] V. Kurkova, Kolmogorov's theorem and multilayer neural networks, Neural Networks 5 (1992) 501–506.