Proceedings of HT2008
2008 ASME Summer Heat Transfer Conference
August 10-14, 2008, Jacksonville, Fl., USA

HT2008-56093

DEVELOPMENT OF A HEAT TRANSFER CORRELATION FOR THE TRANSITIONAL FLOW IN A HORIZONTAL TUBE USING SUPPORT VECTOR MACHINES

L. M. Tam
Department of Electromechanical Engineering, Faculty of Science and Technology, and Institute for the Development and Quality, University of Macau, Av. Padre Tomás Pereira, Taipa, Macau, China.
fstlmt@umac.mo

A. J. Ghajar
School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, Oklahoma, USA.
afshin.ghajar@okstate.edu

H. K. Tam
Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Av. Padre Tomás Pereira, Taipa, Macau, China.
hktam@umac.mo

S. C. Tam
Department of Mathematics, Faculty of Science and Technology, University of Macau, Av. Padre Tomás Pereira, Taipa, Macau, China.
fstsct@umac.mo

ABSTRACT

In this paper the Support Vector Machines (SVM) method is used to correlate the transitional forced and mixed convection experimental data of Ghajar and Tam (1994) that were obtained along a stainless steel horizontal circular tube fitted with re-entrant, square-edged, and bell-mouth inlets under uniform wall heat flux boundary condition. The SVM method has been chosen to further improve the accuracy of the correlations that were developed by Ghajar and his co-workers using the traditional least-squares method (Ghajar and Tam, 1994) and more recently the artificial neural networks (ANN) method (Ghajar et al., 2004). Using the ANN method improved the accuracy of their correlation. However, there are drawbacks associated with ANN method. One of the major problems with the ANN method is that it does not provide a unique correlation due to different coefficient matrices. The SVM method used in this study eliminated the drawbacks associated with the ANN method and provided a unique correlation with comparable accuracy as the ANN method. For the experimental data used, majority of the data points were predicted within 5% deviation. Comparisons were made regarding the accuracy of the developed correlation and its characteristic using SVM and ANN methods. The results showed that SVM is a good method to correlate complex heat transfer data.

INTRODUCTION

An important design problem in industrial heat exchangers arises when the flow inside the tubes falls into the transition region. In practical engineering design, the usual recommendation is to avoid design and operation in this regime; however, this is not always feasible under design constraints. The usually cited transition Reynolds number of about 2100 applies, strictly speaking, to a very steady and uniform entry flow with a rounded entrance. If the flow has a disturbed entrance typical of heat exchangers, in which there is a sudden contraction and possibly even a re-entrant entrance, the transition Reynolds number will be less. Experimental, numerical, and analytical studies are available for forced and mixed convection heat transfer in horizontal tubes with a rounded entrance in the laminar, transitional, and turbulent flow regimes. These works have been reviewed by Shah and London (1978), Shah and Johnson (1981), Kakac and co-workers (1981, 1987). However, very little information that is of immediate use to design engineers (i.e., correlation) is available to predict the developing and fully developed transitional forced and mixed convection heat transfer coefficients in a tube with a smooth or disturbed entrance. The correlation developed by Ghajar and Tam (1994) is the only one that is available in the open literature for the above-mentioned case. Their correlation predicted their transitional data (1290 data points) for three inlet configurations (re-entrant, square-edged, and bell-mouth) with an average

absolute deviation of about 8%. However, about 30% of the data in the transition region were predicted with 10-20% deviation, and about 3% with deviations greater than 20%. Since the value of the heat transfer coefficient has a direct impact on the size of the heat exchanger, a more accurate correlation was sough by Ghajar et al. (2004) using the artificial neural network (ANN) method. They used the three-layer feed-forward ANN for correlating the transitional data of Ghajar and Tam (1994) with three different inlet configurations. The number of the hidden neurons employed was empirically selected as 11. The results showed that the majority of the data points were predicted with less than 5% deviation. The accuracy is superior to the correlation developed by using the traditional least-squares method. However, there are several drawbacks associated with the ANN method such as (1) the optimal network architecture (i.e. the number of hidden neurons) is difficult to determine and (2) the ANN method cannot give the same consistent coefficient matrices every time. In this study, an alternative algorithm called Support Vector Machines (SVM) will be proposed to address these drawbacks. This method has been used to predict the high-dimensional nonlinear problems in recent years (Chen, et. al., 2006, Vong et al. 2006). The specific objectives of this study are to compare the different correlation methods (ANN and SVM) and develop a new correlation using the SVM method for the transitional heat transfer data collected by Ghajar and Tam (1994).

NOMENCLATURE

b = bias term
c_p = specific heat at constant pressure, kJ/(kg-K)
D = inside diameter of the tube, m
fcn = activation (or transfer) function
g = acceleration of gravity, m/s^2
Gr = local bulk Grashof number (= $g\beta\rho^2D^3(T_w-T_b)/\mu_b^2$)
h = local heat transfer coefficient, $W/(m^2-K)$
k = thermal conductivity, $W/(m^2-K)$
Nu = local average or fully developed peripheral Nusselt number (= hD/k)
Pr = local bulk Prandtl number (= $c_p\mu_b/k$)
Re = local bulk Reynolds number (= $\rho VD/\mu_b$)
St = Stanton number (= Nu/Re Pr)
T_b = local bulk temperature, ºC
T_w = local wall temperature, ºC
V = average velocity in the test section, m/s
x = local distance along the test section from the inlet, m

Abbreviations

ANN=artificial neural networks
MSE=mean squared error
SVM=support vector machines
SVR=support vector regression

Greek symbols

β = coefficient of thermal expansion, K^{-1}
μ_b = local bulk dynamic viscosity, Pa-s
μ_w = local wall dynamic viscosity, Pa-s
ρ = local bulk density, kg/m^3

Subscripts

i = dummy parameter
j = dummy parameter
max = maximum value
min = minimum value
m = number of sample points for SVM training

EXPERIMENTAL DATASET

The heat transfer experimental data used in this study, along with a detailed description of the experimental apparatus and procedures used, were reported by Ghajar and Tam (1994). A schematic of the overall experimental setup used for heat transfer measurements is shown in Figure 1. In this paper, only a brief description of the experimental setup and procedures will be provided. The local forced and mixed convective measurements were made in a horizontal, electrically heated, stainless steel circular straight tube under uniform wall heat flux boundary condition and three types of inlet configurations (re-entrant, square-edged, and bell-mouth), as shown in Figure 2. The pipe had an inside diameter of 1.58 cm and an outside diameter of 1.90 cm. The total length of the test section was 6.10 m, providing a maximum length-to-diameter ratio of 385. A uniform wall heat flux boundary condition was maintained by a dc arc welder. Thermocouples (T-type) were placed on the outer surface of the tube wall at close intervals near the entrance and at greater intervals further downstream. Twenty-six axial locations were designated, with four thermocouples placed at each location. The thermocouples were placed 90 degrees apart around the periphery. From the local peripheral wall temperature measurements at each axial location, the inside wall temperatures and the local heat transfer coefficients were calculated (Ghajar and Kim, 2006). In these calculations, the axial conduction was assumed negligible (RePr > 4,200 in all cases), but peripheral and radial conduction of heat in the tube wall were included. In addition, the bulk fluid temperature was assumed to increase linearly from the inlet to the outlet. As reported by Ghajar and Tam (1994), the uncertainty analyses of the overall experimental procedures showed that there is a maximum of 9% uncertainty for the heat transfer coefficient calculations. Moreover, the heat balance error for each experimental run indicated that in general, the heat balance error was less than 5%. For Reynolds numbers lower than 2500 where the flow was strongly influenced by secondary flow, the heat balance error was slightly higher (5-8%) for that particular Reynolds number range. To ensure a uniform velocity distribution in the test fluid before it entered the test section, the flow passed through calming and inlet sections. The calming section had a total length of 61.6 cm and consisted of a 17.8 cm diameter acrylic cylinder with three perforated acrylic plates, followed by tightly packed soda straws sandwiched between galvanized steel mesh screens.

Before entering the inlet section, the test fluid passed through a fine mesh screen and flowed undisturbed through 23.5 cm of a 6.5 cm-diameter acrylic tube before it entered the test section. The inlet section had the versatility of being modified to incorporate a re-entrant or bell-mouth inlet (see Figure 2). The re-entrant inlet was simulated by sliding 1.93 cm of the tube entrance length into the inlet section, which was otherwise the square-edged (sudden contraction) inlet. For the bell-mouth inlet, a fiberglass nozzle with a contraction ratio of 10.7 and a total length of 23.5 cm was used in place of the inlet section. In the experiments, distilled water and mixtures of distilled water and ethylene glycol were used. They collected 1290 experimental data points, and their experiments covered a local bulk Reynolds number range of 280 to 49000, a local bulk Prandtl number range of 4 to 158, a local bulk Grashof number range of 1000 to 2.5×10^5, and a local bulk Nusselt number range of 13 to 258. The wall heat flux for their experiments ranged from 4 to 670 kW/m^2.

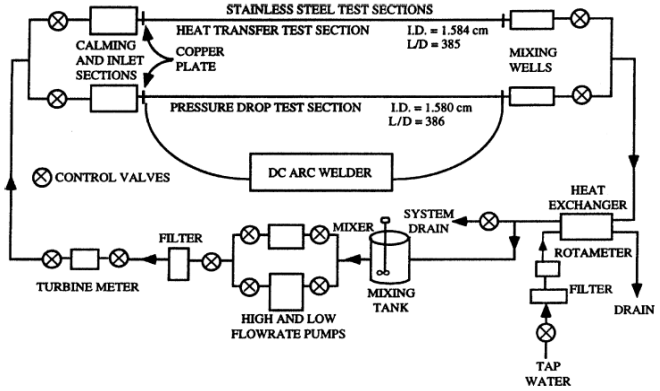

Figure 1: Schematic diagram of experimental setup.

PREDICTION OF TRANSITION DATA

The abrupt change of heat transfer characteristic in the transition region can be seen in Figure 3. In this region, the flow has both laminar and turbulent characteristics. In addition, the type of inlet configuration influences the beginning and end of the transition region. Tam and Ghajar (2006) compared the accuracy of existing correlations for the transition region and concluded that the correlations developed by Ghajar and Tam (1994) and Ghajar et al. (2004) are the most accurate available.

Ghajar and Tam (1994) used the asymptotic method similar to Churchill (1977) to develop a correlation in the transition region since the variation of heat transfer coefficient, as indicated by Figure 3, is between two asymptotes. The correlation for the transitional Nusselt number (Nu_{tr}) is

$$Nu_{tr} = Nu_l + \{\exp[(a - Re)/b] + Nu_t^c\}^c \quad (1)$$

where Nu_l (laminar flow Nusselt number) is given by Eq. (2), Nu_t (turbulent flow Nusselt number) is given by Eq. (3), and the values of the suggested constants a, b, and c for each inlet are summarized in Table 1.

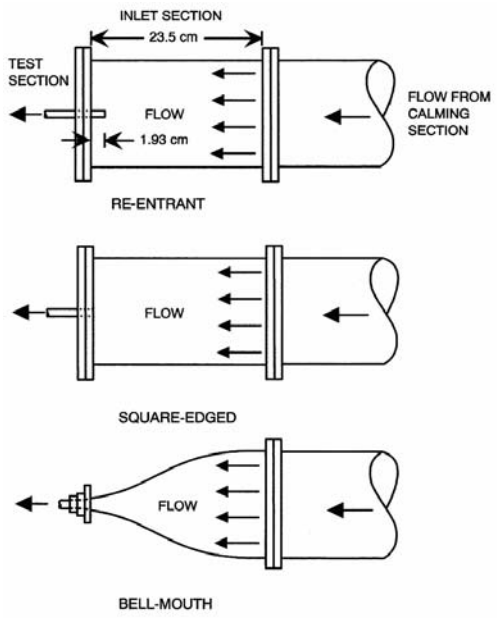

Figure 2: Schematic of the three different inlet configurations.

Table 1: Constants for Eq. (1)

Inlet	a	b	c
Re-entrant	1766	276	-0.955
Square-edged	2617	207	-0.95
Bell-mouth	6628	237	-0.98

$$Nu_l = 1.24[(RePrD/x) + 0.025(GrPr)^{0.8}]^{0.33}(\mu_b/\mu_w)^{0.14} \quad (2)$$

$$Nu_t = 0.023Re^{0.8}\,Pr^{0.385}\,(x/D)^{-0.0054}(\mu_b/\mu_w)^{0.14} \quad (3)$$

The range of independent variables used in Eq. (1) for each inlet configuration is as follows:

Re-entrant: $3 \le x/D \le 192$, $1700 \le Re \le 9100$, $5 \le Pr \le 51$, $4000 \le Gr \le 2.1\times10^5$, and $1.2 \le \mu_b/\mu_w \le 2.2$
Square-edged: $3 \le x/D \le 192$, $1600 \le Re \le 10{,}700$, $5 \le Pr \le 55$, $4000 \le Gr \le 2.5\times10^5$, and $1.2 \le \mu_b/\mu_w \le 2.6$
Bell-mouth: $3 \le x/D \le 192$, $3300 \le Re \le 11{,}100$, $13 \le Pr \le 77$, $6000 \le Gr \le 1.1\times10^5$, and $1.2 \le \mu_b/\mu_w \le 3.1$

Equation (1) is applicable to transition forced and mixed convection in the entrance and fully developed regions and should be used with an appropriate set of constants for each inlet configuration. Figure 4 compares the predicted Nusselt numbers obtained from Eq. (1) for each inlet configuration with measurements. Table 2 summarizes the percent deviations between the experimental data and the predictions of Eq. (1) for the three inlet configurations.

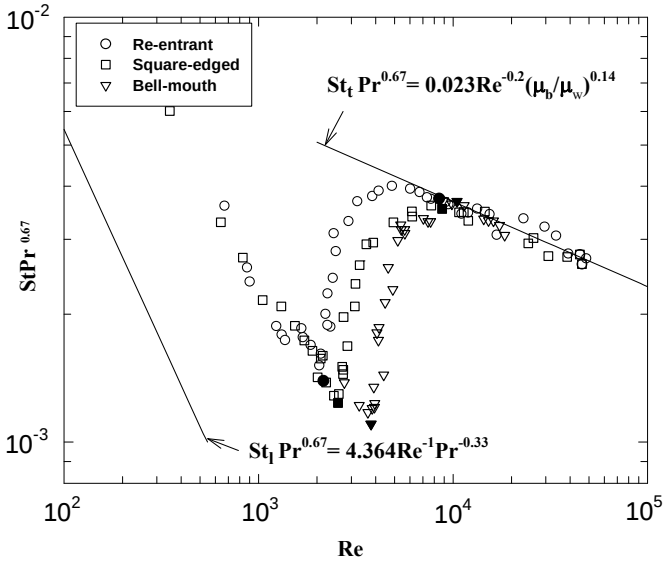

Figure 3: Heat transfer in the transition region (solid symbols designate the start and end of the transition region for each inlet).

Table 2: Percent deviations between the experimental data and predictions of Eq. (1)

	Re-entrant	Square-edged	Bell-mouth
Total data used	441 data points	416 data points	433 data points
Dev. range	-23 to +25.1%	-23.9 to +24.3%	-22 to +18.5%
> 20% dev.	3% (13 pts)	3% (12 pts)	1% (4 pts)
10 to 20% dev.	29% (129 pts)	26% (106 pts)	24% (104 pts)
<10% dev.	68% (299 pts)	72% (298 pts)	75% (325 pts)
Avg. abs. dev.	8%	7.2%	8.1%

To improve the accuarcy of the least-squares correlation, Eq. (1), Ghajar et al. (2004) employed the ANN method on the experimental data of Ghajar and Tam (1994). The three-layer feed-forward neural network was used in their study. For the ANN training process, they compared the results for different gradient methods and different number of hidden neurons, and decided on the Levenberg-Marquardt gradient method and 11 neurons. In their study, when the epoch reached 1000, the training was stopped**.** After the backpropagation had been completed and the postprocessing of the network output had been performed, the resulted neural network, which was the proposed ANN correlation, was given in a matrix form. Their proposed ANN correlation is as follows:

$$Nu = u^3\left\{\left(u^2\right)^2 fcn\left(u^1\Phi + v^1\right) + v^2\right\} + v^3 \qquad (4)$$

where

$$fcn(\mathrm{s}) = [1\text{-}\exp(\mathrm{s})]^{-1}$$

$$\Phi = \begin{bmatrix} \mathrm{Re}_{normal} \\ \mathrm{Pr}_{normal} \\ Gr_{normal} \\ X/D_{normal} \\ \left(\frac{\mu_b}{\mu_w}\right)^{0.14}_{normal} \end{bmatrix} = \begin{bmatrix} 2(\mathrm{Re}-\mathrm{Re}_{\min})/(\mathrm{Re}_{\max}-\mathrm{Re}_{\min})-1 \\ 2(\mathrm{Pr}-\mathrm{Pr}_{\min})/(\mathrm{Pr}_{\max}-\mathrm{Pr}_{\min})-1 \\ 2(Gr-Gr_{\min})/(Gr_{\max}-Gr_{\min})-1 \\ 2(x/D-x/D_{\min})/(x/D_{\max}-x/D_{\min})-1 \\ 2\left(\left(\frac{\mu_b}{\mu_w}\right)^{0.14}-\left(\frac{\mu_b}{\mu_w}\right)^{0.14}_{\min}\right)/\left(\left(\frac{\mu_b}{\mu_w}\right)^{0.14}_{\max}-\left(\frac{\mu_b}{\mu_w}\right)^{0.14}_{\min}\right)-1 \end{bmatrix}$$

and u^1, u^2, u^3, v^1, v^1, v^3 are the coefficient matrices.

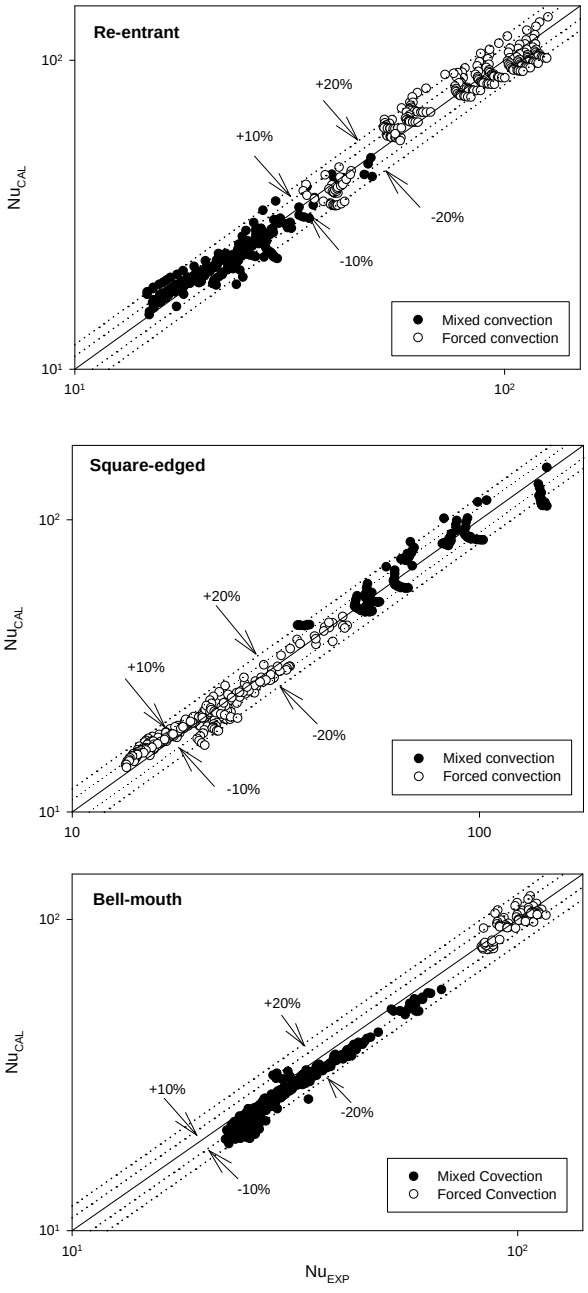

Figure 4: Comparisons between experimental Nusselt numbers and predicted Nusselt numbers using Eq. (1).

The entries of the vector Φ represent the normalized Reynolds number, Prandtl number, Grashof number, x/D and μ ratio, respectively.

In that study, only 80% of the experimental data (M_a) for each inlet were used to establish the correlation's constant matrices as indicated in Eq. (4). The rest of the data (M_b) were used for the network testing. After the network training, the numerical values for the matrices and scalars for different inlet configurations were obtained. With the same network structure, various trainings were made. The initial values of the network parameters, weights and biases were differed for each training. The training with the smallest mean squared error was selected for the subsequent testing. In consequence, Tables 3 to 5 and Figure 5 show the accuracy of the ANN correlations for different inlets. For the bell-mouth inlet, all of the experimental data were predicted with less than 5% deviation. For the square-edged inlet, 99.5% of the data were predicted with less than 5% deviation. Only 2 data points, one from the training set and the other from the testing set were predicted with a little more than 5% deviation. For the re-entrant inlet, practically almost all of the data were predicted with less than 5% deviation (97.1% of the data). Only 12 data points, 6 from the training set and 6 from the testing set were predicted with 5 to 10% deviation, and only one data point was predicted with more than 10% deviation (11.35%). Compared to the accuracy provided by the Eq. (1), it is obvious that the ANN correlation outperformed the traditional least-squares method correlation.

THE SUPPORT VECTOR MACHINES

From the material presented in the previous section, it can be concluded that (1) there is still plenty of room for improving the accuracy of the traditional least-squares method correlation (see Table 2); (2) the structure of the ANN is chosen subjectively. From the statistical point of view, the ANN method can only ensure the reduction of training error by an increase in the number of hidden neurons. However, there is no guarantee whether the optimal network structure has been chosen. Therefore, the 'overfiting' problem does still exist; and (3) the initial weights and biases in each training are randomly selected. This could cause the training process to easily get stuck at a local minima (Suykens, 2001). Hence, the coefficient matrices will not be the same, i.e., with the same network structure, different versions of the correlation can be obtained. Usually, the version, which gives the least mean squared error from multiple trainings, is selected. To overcome the deficiencies of the least-squares method and the ANN method with respect to accuracy or uniqueness of the developed correlation as discussed above, the SVM method will be explored. The SVM method offers the following advantages over the traditional least-squares method and the ANN method: (1) The SVM method has been used for predicting practical problems with good accuracy (Chen, et. al., 2006, Vong et al. 2006). (2) The SVM method when applied to a set of data provides a unique correlation. The method is designed to provide a small risk (or test error) during the regression process by minimizing the regularized risk function, which is defined below in Eq. (8). The regularized risk function contains the training error and model complexity terms. The SVM structure and the training error can be optimized by minimizing the risk function. Since the structure is optimized, the 'overfiting' problem can be avoided. Moreover, the optimal model gives consistent coefficients. Therefore, the SVM method is capable of correlating a set of data with a unique correlation.

Table 3: Accuracy of the re-entrant correlation based on ANN method

Re-entrant Inlet, M = 441 pts, M_a = 353 pts, M_b = 88 pts					
	Range of deviation	Absolute deviation	Less than ±5% deviation	Between ±5% and 10% deviation	More than ±10% deviation
M	-11.35% to 7.49%	1.12%	428 data points	12 data points	1 data point
M_a	-8.82% to 7.49%	1.00%	347 data points	6 data points	0 data points
M_b	-11.35% to 6.75%	1.60%	81 data points	6 data points	1 data point

Table 4: Accuracy of the square-edged correlation based on ANN method

Square-edged Inlet, M = 416 pts, M_a = 333 pts, M_b = 83 pts				
	Range of deviation	Absolute deviation	Less than ±5% deviation	Between ±5% and ±10% deviation
M	-4.88% to 5.24%	1.06%	414 data points	2 data points
M_a	-4.24% to 5.16%	1.03%	332 data points	1 data point
M_b	-4.88% to 5.24%	1.16%	82 data points	1 data point

Table 5: Accuracy of the square-edged correlation based on ANN method

Bell-mouth Inlet, M = 433 pts, M_a = 346 pts, M_b = 87 pts			
	Range of deviation	Absolute deviation	Less than ±5% deviation
M	-3.45% to 3.60%	0.62%	433 data points
M_a	-2.56% to 2.29%	0.59%	346 data points
M_b	-3.45% to 3.60%	0.74%	87 data points

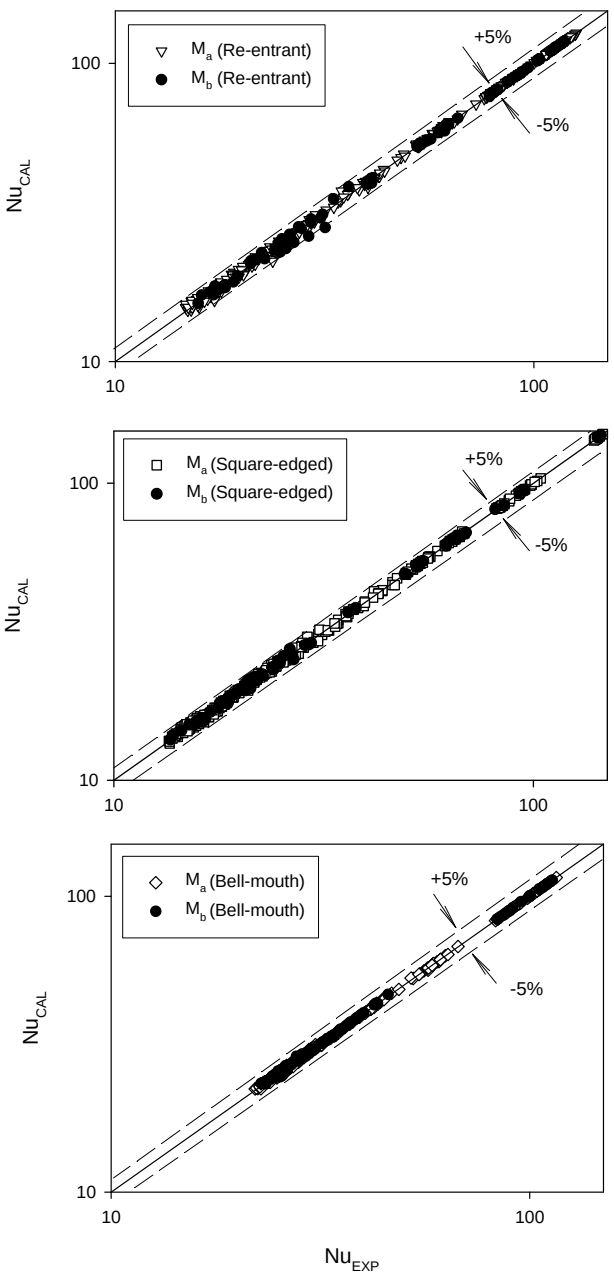

Figure 5: Comparisons between experimental Nusselt numbers and those predicted by the ANN-based transition region heat transfer correlation.

The SVM method was first developed for pattern recognition based on the statistical learning theory (Vapnik, 1995). The method was later on used for regression estimation when Vapnik devised the so-called ε-insensitive loss function,

$$\left|\mathrm{y} - f(\mathbf{x})\right|_{\varepsilon} = \max\left\{0, \left|\mathrm{y} - f(\mathbf{x})\right| - \varepsilon\right\}, \tag{5}$$

which does not penalize errors below some $\varepsilon \geq 0$, chosen a priori. According to Schölkopf and Smola (2002), the algorithm was called ε-SVR, which seeks to estimate linear functions,

$$f(\mathbf{x}) = <\mathbf{w}, \mathbf{x}> + \mathrm{b}, \quad \mathbf{w}, \mathbf{x} \in \mathbf{R}^{\mathrm{N}}, \mathrm{b} \in \mathbf{R}, \tag{6}$$

where <**w**, **x**> is the dot product in Euclidean space and the function is based on independent and identically distributed data,

$$(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_l, y_l) \in \mathbf{R}^{\mathrm{N}} \times \mathbf{R}. \tag{7}$$

where $\mathbf{R}^{\mathrm{N}}$ is the dot product space in which the (mapped) input patterns live (i.e., the feature space induced by a kernel). The goal of the learning process is to find a function *f* with a small risk (or test error). It was tried to get the small risk by minimizing the regularized risk functional,

$$\frac{1}{2}\left\|\mathbf{w}\right\|^2 + C \cdot \mathrm{R}^{\varepsilon}_{\mathrm{emp}}[f]. \tag{8}$$

where, $\left\|\mathbf{w}\right\|^2$ is a term that characterizes the model complexity,

$$\mathrm{R}^{\varepsilon}_{\mathrm{emp}}[f] := \frac{1}{m}\sum_{i=1}^{m}\left|\mathrm{y}_{\mathrm{i}} - f(\mathbf{x}_i)\right|_{\varepsilon}, \tag{9}$$

is the measure of the ε-insensitive training error, and C is a constant determining the trade-off. Minimizing the Eq. (8) captures the main insight of statistical learning theory, stating that in order to obtain a small risk, the training error and model complexity needed to be controlled. As shown in Figure 6, the minimization of Eq. (8) is equivalent to the following constrained optimization problem:

$$\underset{\mathrm{w}\in\mathbf{R}^{\mathrm{N}},\boldsymbol{\xi}^{(*)}\in\mathbf{R}^{m},b\in\mathbf{R}}{\text{minimize}} \quad \tau(\mathrm{w}, \boldsymbol{\xi}^{(*)}) = \frac{1}{2}\left\|\mathbf{w}\right\|^2 + C\frac{1}{m}\sum_{i=1}^{m}(\xi_i + \xi_i^*), \tag{10}$$

$$\text{subject to} \quad (<\mathbf{w}, \mathbf{x}_{\mathrm{i}}> + \mathrm{b}) - \mathrm{y}_{\mathrm{i}} \leq \varepsilon + \xi_{\mathrm{i}} \tag{11}$$

$$\mathrm{y}_{\mathrm{i}} - (<\mathbf{w}, \mathbf{x}_{\mathrm{i}}> + \mathrm{b}) \leq \varepsilon + \xi_{\mathrm{i}}^* \tag{12}$$

$$\xi_{\mathrm{i}}^{(*)} \geq 0. \tag{13}$$

where i = 1, …, m, and that boldface Greek letter, $\boldsymbol{\xi}$, denotes m-dimensional vectors of the corresponding variable; (*) is a shorthand implying both the variables with and without asterisks, $\xi_{\mathrm{i}}, \xi_{\mathrm{i}}^*$.

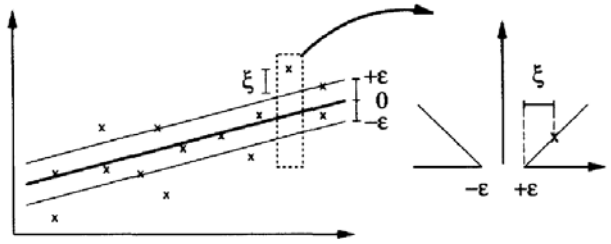

Figure 6: Schematic diagram represents the constrained optimization problem, Eqs. (10)-(13). (Taken from Schölkopf et al., 2000).

In Eqs. (11) and (12), the value of the parameter ε is used to adjust the desired level of accuracy of the approximation. It has to be determined prior to the training. Sometimes, the desired level of accuracy is not indicated a priori and the function approximation is required to be as accurate as possible. Thus, the υ-SVR algortihm (proposed by Schölkopf et al., 2000), which is the modification of the ε-SVR, is suggested to be used. The algorithm can compute the parameter ε automatically to a minimum level. In this study, the υ-SVR algorithm will be used instead of Eq. (10). The functional form of the υ-SVR algorithm is shown below:

$$\underset{\mathbf{w}\in\mathbf{R}^N,\boldsymbol{\xi}^{(*)}\in\mathbf{R}^m,\varepsilon,b\in\mathbf{R}}{\text{minimize}} \quad \tau(\mathbf{w},\boldsymbol{\xi}^{(*)},\varepsilon)=\frac{1}{2}\|\mathbf{w}\|^2 + C\cdot\left(\upsilon\varepsilon+\frac{1}{m}\sum_{i=1}^{m}(\xi_i+\xi_i^*)\right), \quad (14)$$

$$\text{subject to} \quad (<\mathbf{w},\mathbf{x}_i>+b)-y_i\le\varepsilon+\xi_i \quad (15)$$

$$y_i-(<\mathbf{w},\mathbf{x}_i>+b)\le\varepsilon+\xi_i^* \quad (16)$$

$$\xi_i^{(*)}\ge 0,\ \varepsilon\ge 0 \quad (17)$$

Then, the above problem can be solved by the convex optimization method (Boyd and Vandenberghe, 2004). The key point is to construct a Lagrangian from the objective function (Eq. 14) and the corresponding constraints (Eqs. 15, 16, 17). For the constraints, the multipliers $\alpha_i^{(*)},\eta_i^{(*)},\beta\ge 0$, are introduced in order to obtain the Lagrangian,

$$L(\mathbf{w},b,\boldsymbol{\alpha}^{(*)},\beta,\boldsymbol{\xi}^{(*)},\varepsilon,\boldsymbol{\eta}^{(*)}) = \frac{1}{2}\|\mathbf{w}\|^2+C\upsilon\varepsilon+\frac{C}{m}\sum_{i=1}^{m}(\xi_i+\xi_i^*)-\beta\varepsilon-\sum_{i=1}^{m}(\eta_i\xi_i+\eta_i^*\xi_i^*) - \sum_{i=1}^{m}\alpha_i(\xi_i+y_i-<\mathbf{w},\mathbf{x}_i>-b+\varepsilon) - \sum_{i=1}^{m}\alpha_i^*(\xi_i^*+<\mathbf{w},\mathbf{x}_i>+b-y_i+\varepsilon). \quad (18)$$

To minimize Eq. (14), the saddle point of Eq. (18) has to be solved. It means to minimize over the primal variables $\mathbf{w},\varepsilon,b,\boldsymbol{\xi}_i^{(*)}$ and maximize over the dual variables $\alpha^{(*)},\beta,\eta^{(*)}$. Setting the partial derivatives of L with respect to the primal variables equal to zero yields the following four equations

$$\mathbf{w}=\sum_i(\alpha_i^*-\alpha_i)\mathbf{x}_i, \quad (19)$$

$$C\cdot\upsilon-\sum_i(\alpha_i+\alpha_i^*)-\beta=0, \quad (20)$$

$$\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)=0, \quad (21)$$

$$\frac{C}{m}-\alpha_i^{(*)}-\eta_i^{(*)}=0. \quad (22)$$

In the SV expansion Eq. (19), only those $\alpha_i^{(*)}$ will be nonzero that correspond to a constraint, Eq. (15) or (16), which is precisely met; the corresponding patterns are called support vectors.

Substituting the above four conditions into Eq. (18) leads to the dual optimization problem (sometimes called the Wolfe dual). Also, substitute a kernel k for the dot product in the problems, corresponding to a dot product in some feature space related to input space via a nonlinear map $\boldsymbol{\Phi}$,

$$k(\mathbf{x},\mathbf{x}')=<\boldsymbol{\Phi}(\mathbf{x})\cdot\boldsymbol{\Phi}(\mathbf{x}')>=<\mathbf{x},\mathbf{x}'>, \quad (23)$$

As depicted in Figure 7, by using the kernel k, the low dimensional nonlinear input data space can be transformed into the high-dimensional linear feature space through a nonlinear mapping $\boldsymbol{\Phi}:\mathbf{R}^n\to\mathbf{R}^{n_h}$, n is the dimension of input data space, and n_h is the high dimension of the feature space. This transformation makes the SVM be capable of treating the nonlinear problems. Usually, Gaussian kernels, $k(\mathbf{x},\mathbf{x}')=\exp(-\gamma\|\mathbf{x}-\mathbf{x}'\|^2)$ and polynomial kernels, $k(\mathbf{x},\mathbf{x}')=(\mathbf{x}\cdot\mathbf{x}')^d(\gamma>0,d\in\mathbf{N})$ are employed for the nonlinear problems.

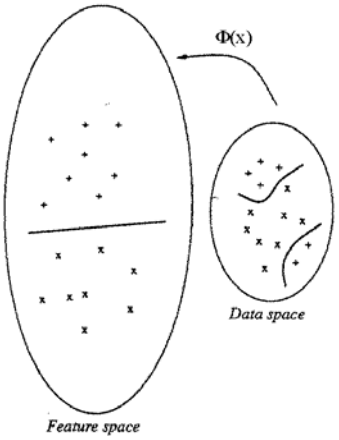

Figure 7: For nonlinear problem, input data are transformed to a higer dimensional feature space by the mapping $\boldsymbol{\Phi}(\mathbf{x})$. (Taken from Suykens, 2001).

Rewriting the constraints, and noting that $\beta,\eta_i^{(*)}\ge 0$ do not appear in the dual, the dual of υ-SVR optimization problem for $\nu\ge 0, C>0$ is obtained:

$$\underset{\boldsymbol{\alpha}\in\mathbf{R}^m}{\text{maximize}} \quad W(\alpha^{(*)})=\sum_{i=1}^{m}(\alpha_i^*-\alpha_i)y_i - \frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)k(x_i,x_j), \quad (24)$$

$$\text{subject to} \quad \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0, \quad (25)$$

$$\alpha_i^{(*)}\in\left[0,\frac{C}{m}\right], \quad (26)$$

$$\sum_{i=1}^{m}(\alpha_i + \alpha_i^*) \le C \cdot v. \quad (27)$$

In this study, the *C/m* will be substituted by *C*. Therefore, Eqs. (26) and (27) will be changed as follows,

$$\alpha_i^* \in [0, C], \quad (28)$$

$$\sum_{i=1}^{m}(\alpha_i + \alpha_i^*) \le C \cdot m \cdot v. \quad (29)$$

Finally, the decision function for regression takes the form (cf. Eqs. (6), (19), (23))

$$f(x) = \sum_{i=1}^{m}(\alpha_i^* - \alpha_i)k(x_i, x) + b, \quad (30)$$

which is the same as the decision function of ε-SVR. The values of b and ε can be computed by taking into account that Eqs. (15) and (16) become equalities with $\xi_i^{(*)} = 0$ for points with $0 < \alpha_i^{(*)} < C/m$, due to the Karush-Kuhn-Tucker (KKT) conditions. Substitution of $\sum_j (\alpha_j^* - \alpha_j)k(x_j, x)$ for $< \mathbf{w}, \mathbf{x} >$ is understood, cf. Eqs. (19) and (23). Further details about υ-SVR can be found in the studies of Schölkopf et al. (2000), Schölkopf and Smola (2002), Smola and Schölkopf (2004).

RESULTS AND DISCUSSION

Based on Eq. (30) with Gaussian kernel function used, the functional form of the SVM-based correlation form can be written as

$$\mathrm{Nu}(\mathbf{x}) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)\exp(-\gamma\|\mathbf{x}_i - \mathbf{x}\|^2) + b, \quad (31)$$

where

Nu is the Nusselt number,

l is the number of sample points,

γ is the kernel parameter defined a priori,

$\alpha_i, \alpha_i^{(*)}$ are the constant coefficients,

b is the bias,

$\mathbf{x}_i$ is the support vectors,

$$\mathbf{x} = \begin{bmatrix} \mathrm{Re}_{\mathrm{normal}} \\ \mathrm{Pr}_{\mathrm{normal}} \\ \mathrm{Gr}_{\mathrm{normal}} \\ (\mathrm{x/D})_{\mathrm{normal}} \\ \left(\frac{\mu_B}{\mu_W}\right)^{0.14}_{\mathrm{normal}} \end{bmatrix} = \begin{bmatrix} 2(\mathrm{Re} - \mathrm{Re}_{\min})/(\mathrm{Re}_{\max} - \mathrm{Re}_{\min}) - 1 \\ 2(\mathrm{Pr} - \mathrm{Pr}_{\min})/(\mathrm{Pr}_{\max} - \mathrm{Pr}_{\min}) - 1 \\ 2(\mathrm{Gr} - \mathrm{Gr}_{\min})/(\mathrm{Gr}_{\max} - \mathrm{Gr}_{\min}) - 1 \\ 2[(\mathrm{x/D}) - (\mathrm{x/D})_{\min}]/[(\mathrm{x/D})_{\max} - (\mathrm{x/D})_{\min}] - 1 \\ 2[\left(\frac{\mu_B}{\mu_W}\right)^{0.14} - \left(\frac{\mu_B}{\mu_W}\right)^{0.14}_{\min}]/[\left(\frac{\mu_B}{\mu_W}\right)^{0.14}_{\max} - \left(\frac{\mu_B}{\mu_W}\right)^{0.14}_{\min}] - 1 \end{bmatrix}$$

The entries of the vector, **x**, represents the normalized Reynolds number, Prandtl number, Grashof number, x/D, and μ ratio, respectively. In this study, 80% of the experimental data (denoted by M_a) for each inlet were used to find the coefficients and bias of Eq. (31). The rest of the data (i.e., 20%, as denoted by M_b) were used for verification.

Prior to developing the SVM-based correlation, the parameters, C and γ, should be determined. As shown in Figure 8, different combinations of the parameters were used for the training of the model based on the 80% of the re-entrant data (M_a). Then, all the re-entrant data (M) is tested by the SVM model to obtain the Mean Squared Error (MSE). The ranges of the C and γ were selected arbitrarily within [10 to 30000] and [1 to 15]. When the value of γ equals to 8 and the value of C exceeded 20,000, the mean squared error (MSE) was the smallest amongst all the combinations of the parameters. Therefore, the values of the parameters, C and γ, used in this study were 20,000 and 8, respectively.

With the above mentioned parameters, the SVM was trained for each inlet configuration. Therefore, the correlation coefficients and the support vectors were different for each of the three types of inlets. Tables 6 to 8 summarize the accuracy of the proposed SVM heat transfer correlation for the three different inlet configurations. Note that for comparison purposes, the tables show the accuracy distribution for 100% of the experimental data (denoted by M), the data that were used to train the network (denoted by M_a i.e., 80% of M), and the data that were never used in training but were used for testing (denoted by M_b). Figure 9 compares the predicted Nusselt numbers obtained from the proposed SVM-based correlation for each inlet configuration with the measurements.

The proposed SVM correlation is applicable to the developing and fully developed transition regions. As seen in Tables 6 to 8 and Figure 9, for the three inlet configurations, all the Nusselt numbers of the training data were predicted with excellent accuracy. For the testing data, majority of the experimental data for the three inlets (i.e. 93.4% of the total testing data) were still predicted within ±5% deviation. For the bell-mouth inlet, 99.1 percent of the data were predicted with less than 5% deviation; only four data points from the testing set were predicted with a little more than ±5% deviation. The absolute deviation was 0.31%. For the square-edged inlet, 98.6 % of the data were predicted with less than 5% deviation; only six data points from the testing set were predicted with more than ±5% deviation. The absolute deviation was 0.47%. For the

re-entrant inlet, nearly all of the data were predicted with less than 5% deviation (98.4%). Only four data points from the testing set were predicted within ±5 to 10% deviation, and only three data points were predicted with more than 10% deviation. The absolute deviation is as small as 0.35%. Compared to the accuracy provided by the traditional least-squares correlation (Eq. 1), it can be concluded that the SVM correlation is superior to the traditional correlation. When the regression ability of SVM is compared with that of ANN method, the predictive capability of SVM method was shown to be as accurate as the ANN method. Both computation intelligent methods also give excellent performance for prediction of the complicated transition data.

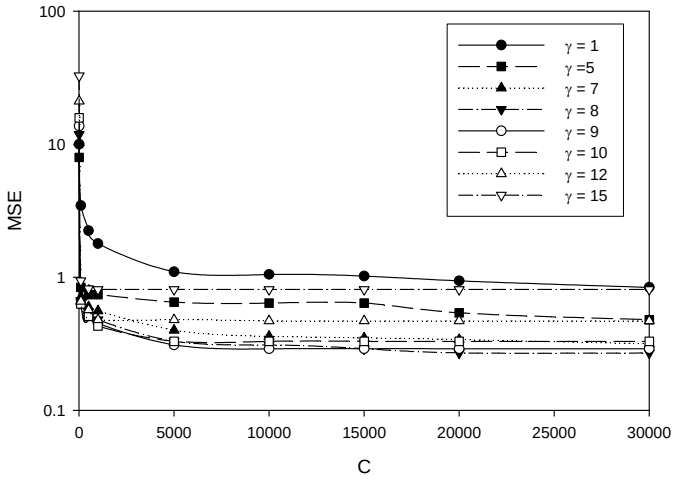

Figure 8: Determination of the desired values of the constants, C and γ, of the SVM model.

CONCLUSIONS

A new correlation was developed by the SVM method in this study. The new regression method outperforms the traditional least-squares method and it is comparable in accuracy to the ANN method. However, it does not have the drawbacks of the ANN method. Specifically the advantages of the SVM method over the ANN method are (1) the structure of the network does not have to be determined, and (2) it generates a unique correlation form. Although the tuning parameters of SVM such as C and γ are necessarily determined prior to the training process in SVM, it is not difficult to find the desired values through the empirical method.

Table 6: Accuracy of the proposed re-entrant correlation

Re-entrant Inlet, M = 441 pts, M_a = 353 pts, M_b = 88 pts

	Range of deviation	Absolute deviation	Less than ±5% deviation	Between ±5% and 10% deviation	More than ±10% deviation
M	-10.09% to 19.95%	0.35%	434 data points	4 data points	3 data points
M_a	-0.01% to 0.01%	0%	353 data points	0 data points	0 data points
M_b	-10.09% to +19.95%	1.75%	81 data points	4 data points	3 data points

Table 7: Accuracy of the proposed square-edged correlation

Square-edged Inlet, M = 416 pts, M_a = 333 pts, M_b = 83 pts

	Range of deviation	Absolute deviation	Less than ±5% deviation	Between ±5% and 10% deviation	More than ±10% deviation
M	-5.46% to +36.34%	0.47%	410 data points	3 data points	3 data points
M_a	-0.02% to 0.25%	0 %	333 data points	0 data points	0 data points
M_b	-5.46% to +36.34%	2.33%	77 data points	3 data points	3 data points

Table 8: Accuracy of the proposed bell-mouth correlation

Bell-mouth Inlet, M = 433 pts, M_a = 346 pts, M_b = 87 pts

	Range of deviation	Absolute deviation	Less than ±5% deviation	Between ±5% and 10% deviation
M	-3.72% to +6.15%	0.31%	429 data points	4 data points
M_a	-1.67% to 0	0.01%	346 data points	0 data points
M_b	-3.72% to +6.15%	1.62%	83 data points	4 data points

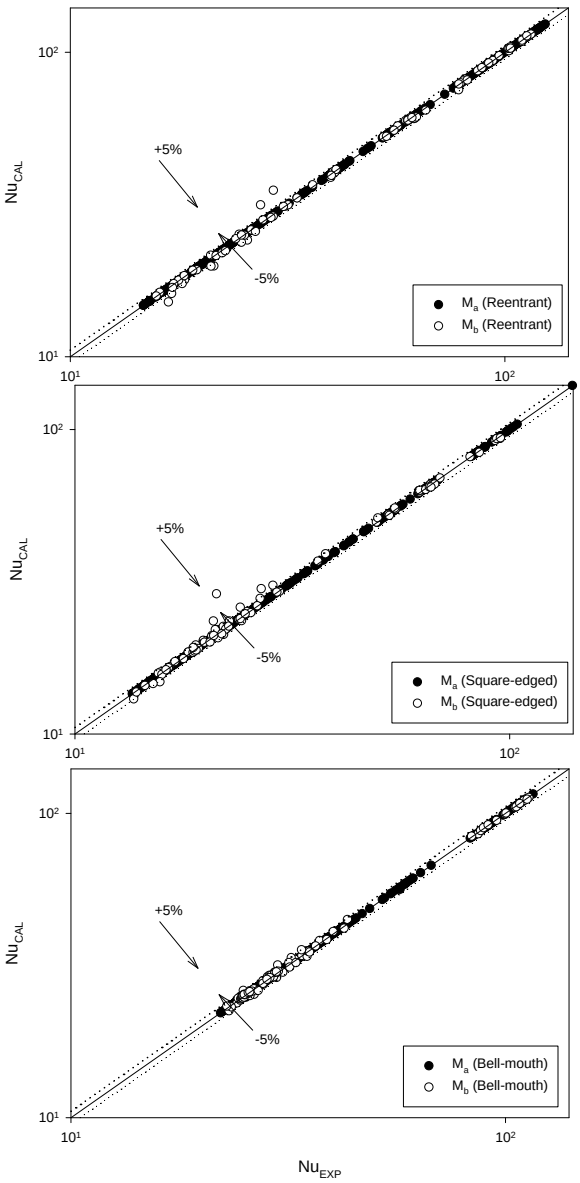

Figure 9: Comparison between experimental Nusselt numbers and those predicted by the proposed SVM-based transition region heat transfer correlation.

REFERENCES

Boyd, S. and Vandenberghe, L., 2004, "Convex Optimization," Cambridge University Press, New York.

Chen, W-H., Shih, J-Y., and Wu, S., 2006, "Comparison of Support-Vector Machines and Back Propagation Neural Networks in Forecasting the Six Major Asian Stock Markets," Int. J. Electronic Finance, Vol. 1, no. 1, pp. 49-66.

Churchill, S. W., 1977, "Comprehensive Correlating Equations for Heat, Mass, and Momentum Transfer in Fully Developed Flow in Smooth Tubes," Ind. Eng. Chem. Fundam., vol. 16, no. 1, pp. 109–116.

Ghajar, A. J. and Tam, L. M., 1994, "Heat Transfer Measurements and Correlations in the Transition Region for a Circular Tube with Three Different Inlet Configurations," Experimental Thermal and Fluid Science," Vol. 8, No. 1, pp. 79-90.

Ghajar, A. J., Tam, L. M., and Tam, S. K., 2004, "Improved Heat Transfer Correlation in the Transition Region for a Circular Tube with Three Inlet Configurations Using Artificial Neural Networks," Heat Transfer Engineering, Vol. 25, No. 2, pp. 30-40.

Ghajar, A. J. and Kim, J., 2006, "Calculation of Local Inside-Wall Convective Heat Transfer Parameters from Measurements of the Local Outside-Wall Temperatures along an Electrically Heated Circular Tube," Heat Transfer Calculations, edited by Myer Kutz, McGraw-Hill, New York, NY, pp. 23.3-23.27.

Kakac, S., Shah, R. K., and Bergles, A. E., 1981, "Low Reynolds Number Flow Heat Exchanger," Hemisphere, New York.

Kakac, S., Shah, R. K., and Aung, W., 1987, "Handbook of Single-Phase Convective Heat Transfer," Wiley, New York.

Shah, R. K., and London, A. L., 1978, "Laminar Flow Forced Convection in Ducts," a Supplement to Advances in Heat Transfer, Academic Press, New York.

Shah, R. K., and Johnson, R. S., 1981, "Correlations for Fully Developed Turbulent Flow Through Circular and Noncircular Channels," Sixth National Heat and Mass Transfer Conf., Indian Institute of Technology, Madras, India, pp. D75–D95.

Schölkopf, B. and Smola, A. J., 2002, "Learning with Kernels – Support Vector Machines, Regularization, Optimization and Beyond," MIT Press, Cambridge, MA.

Schölkopf, B., Smola, A. J., Williamson, R. C., and Bartlett, P. L., 2000, "New Support Vector Algorithms," Neural Computation, Vol.12, pp.1207-1245.

Smola, A. J. and Schölkopf, B., 2004, "A Tutorial on Support Vector Regression," Statistics and Computing, Vol. 14, pp. 199-222.

Suykens, J.A.K., 2001, "Nonlinear Modeling and Support Vector Machines," Instrumentation and Measurement Technology Conference, 2001. IMTC 2001, Budapest, Hungary, Proceedings of the 18th IEEE, Vol. 1, Issue, 21-23, pp. 287 – 294.

Tam, L. M. and Ghajar, A. J., 2006, "Transitional Heat Transfer in Plain Horizontal Tubes," Heat Transfer Engineering, Vol. 27, No. 5, pp. 23-38.

Vapnik, V., 1995, "The Nature of Statistical Learning Theory," Springer-Verlag, New York.

Vong, C-M., Wong, P-K., and Li, Y-P., 2006, "Prediction of Automotive Engine Power and Torque using Least Squares Support Vector Machines and Baysian Inference," Engineering Application of Artificial Intelligence, Vol. 19, pp. 277-287.