CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Universiti Teknologi Malaysia Institutional Repository

Ain Shams Engineering Journal 11 (2020) 161–169

Contents lists available at ScienceDirect

Ain Shams Engineering Journal

journal homepage: www.sciencedirect.com

Engineering Physics and Mathematics

Multivariable power least squares method: Complementary tool for Response Surface Methodology

Wah Yen Tey [a,b,*], Kiat Moon Lee [c], Yutaka Asako [b], Lit Ken Tan [b], Norio Arai [b]

[a] Department of Mechanical Engineering, Faculty of Engineering, UCSI University, Kuala Lumpur, Malaysia
[b] Takasago Thermal Engineering i-Kohza, Department of Mechanical Precision Engineering, Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[c] Department of Chemical and Petroleum Engineering, Faculty of Engineering, UCSI University, Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 22 May 2018
Revised 7 July 2019
Accepted 1 August 2019
Available online 4 September 2019

Keywords:
Multivariable analysis
Multivariable regression model
Design of experiments
Response Surface Methodology

ABSTRACT

In Response Surface Methodology (RSM), variables are correlated through polynomial functions based on Stone-Weierstrass theorem. However, such formulation inherits four weaknesses: possible misleading approximation, incapability to accurately determine the ranking of factors' dominance, failure to analyse factors in random value and proliferation of guess functions due to Pascal Triangle. Therefore, this article aims to develop an improvised method to rectify and complement the weaknesses of RSM. Multivariable Power Least Squares Method (MPLSM) has been developed to correlate various sets of independent variables with dependent variable in the form of power functions. MPLSM is built upon least squares method, and able to approximate the indices of the variables easily. Two variants of MPLSM are suggested to further ensure the numerical stability: the Normalised MPLSM and Iterative MPLSM. The proposed method is not only substantial in big data analysis and multivariable problems, but also providing an alternative approach in engineering optimisation.
© 2019 The Authors. Published by Elsevier B.V. on behalf of Faculty of Engineering, Ain Shams University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The development of regression method started with the work of Galton [1] in tandem with the call of vibrant scientific research, which requires the formation of equation as the tool for result prediction. Various regression methods such as polynomial, exponential, logarithmic and power regressions were later developed [2–4]. Recent regression models are built upon these fundamental regression methods in order to identify a more suitable dependency pattern between the variables. For instance, multiple linear regression [5], parameterised linear regression [6], geographically weighted beta regression [7] and multivariate total least squares method [8] are developed based on linear regression, while Moving Least Squares Method (MLSM) [9], multivariable Laguerre polynomials [10] and Response Surface Methodology (RSM) [11] are constructed based on polynomial functions regression.

Amongst all, RSM is the most perennial regression approach in applied science and engineering research due to their excellent ability in correlating multiple variables with very high coefficients of determination and optimising the response of interest. RSM is widely applied in process optimisation, which spans across the fields of materials, environmental, food, thermal and biochemical engineering [12–20]. Despite its popularity, not all process optimisation can be investigated via RSM, especially for those cases where the variables possess highly deforming trend of curves [21].

It is not satisfactory for RSM to deal with second or higher polynomial functions [21]. A simple remedy to this could be MLSM and its variants [22–25], which are also widely used in the formulation of meshfree methods [26–28] and image processing [29,30]. MLSM is able to deal with various guess functions, where polynomial function is mostly applied. Moreover, MLSM can generate equations with high order of polynomial function, up to the precision of prescribed guess functions. This may be due to the establishment of Stone-Weierstrass theorem [31], which claimed that with sufficient amount of polynomial forms, the equation should be able

* Corresponding author at: Department of Mechanical Engineering, Faculty of Engineering, UCSI University, Kuala Lumpur, Malaysia.
E-mail addresses: teywy@ucsiuniversity.edu.my (W.Y. Tey), leekm@ucsiuniversity.edu.my (K.M. Lee), y.asako@utm.my (Y. Asako), tlken@utm.my (L.K. Tan), arai@cc.tuat.ac.jp (N. Arai).
Peer review under responsibility of Ain Shams University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.asej.2019.08.002
2090-4479/© 2019 The Authors. Published by Elsevier B.V. on behalf of Faculty of Engineering, Ain Shams University.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

to describe most of the non-linear phenomenon. However, both RSM and MLSM share two crucial common shortcomings. Firstly, they will be generating a very long and complicated equation when the order of polynomial increases [32]. Secondly, it is very difficult to precisely determine the ranking of variable dominance when the significance level or p-value is less than 0.0001.

RSM also requires pre-defined discrete independent variables (factors) in order to fulfil statistical design criteria of orthogonality, rotability and uniform precision [11]. This feature will become a great hindrance when the experiments cannot be conducted in a carefully controlled laboratory, especially for the variables that can only be captured in nature and haphazard fashion. Some examples of such studies comprise investigation of solar radiation and power generation [33–35], seismic prediction [36], tidal and ocean engineering [37,38] and turbomachinery [39,40]. When the variables with great randomness set in, RSM will fail, and most of the researchers have to resort to the soft computing such as artificial neural networks (ANN). Although MLSM can tackle these predicaments in principle, the unavoidable weaknesses as explained in the previous paragraph, will still transpire.

The root cause behind all these is their inherited limitation that the indices of variables have to be constrained to integers in the earliest stage of mathematical formulation. This is because the substitution of value of variables into the prescribed order of polynomial function is required. Therefore, this paper is to derive and develop a method that can deal with multivariable problems using power function, in which the indices of the parameters can be determined to the extent of any real numbers. This will bypass the tedious numerical procedure to form the expanded polynomial terms. Since least squares method is used for minimisation of residuals, the method is named as Multivariable Power Least Squares Method (MPLSM). The indices of the respective variables can be easily determined using an inverse matrix.

The structure of the paper is divided into two main sections: the mathematical formulation and examples of MPLSM applications. The new method is able to inspire more solid investigation in the research fields, which involve curve-fitting, multivariable analysis and optimisation.

2. Mathematical formulation

Consider the independent variables $x = \{ x_1 \quad x_2 \quad \cdots \quad x_m \} \in \mathbb{R}^d > 0$ in which $\mathbb{R}$ represents real numbers in d dimensions. m is the total number of set in the independent variables. The single dependent variable, y can be correlated through an approximation function, f where $f : x \rightarrow y$ by taking the dot product of respective power functions. y^h is the guess function in the form of multivariable power function as shown in (1).

$$y \approx y^h = a x_1^{b_1} x_2^{b_2} \cdots x_m^{b_m} \tag{1}$$

It is an important feature that $a, b_1, b_2, \ldots, b_m \in \mathbb{R}$, and they are the coefficients for respective set of variables. Since it is unwise to directly apply least squares analysis on (1) due to its insurmountable complexity, (1) shall be expanded in logarithmic form before the deployment of least squares method, i.e.

$$\ln(y^h) = \ln(a) + b_1 \ln(x_1) + b_2 \ln(x_2) + \cdots + b_m \ln(x_m) \tag{2}$$

By taking $A = \ln(a)$, with the subscript j as the number of set of variables, the square of least squares residual, J can be formulated as:

$$J = \sum_{i=1}^{n} \left(\ln(y_i) - \ln(y_i^h) \right)^2 = \sum_{i=1}^{n} \left(\ln(y_i) - A - b_1 \ln(x_1) - b_2 \ln(x_2) - \cdots - b_m \ln(x_m) \right)^2 \tag{3}$$

Upon expansion of (3), solution for A can be obtained by taking the minimisation of J with respect to A, in which:

$$\frac{\partial J}{\partial A} = 0$$

$$A = \frac{1}{n} \left(\left(\sum_{i=1}^{n} \ln(y_i) \right) - \sum_{j=1}^{m} \left(b_j \sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{j,i}) \right) \right) \tag{4}$$

where i is the i^{th} elements within the set of variables while n is the total number of elements within the set of variables. The minimisation of J with respect to $\mathbf{b} = \{ b_1 \quad b_2 \quad \cdots \quad b_m \}$ will yield the set of equations which can be solved easily using matrix.

$$\frac{\partial J}{\partial b_1} = 0, \frac{\partial J}{\partial b_1} = 0, \ldots, \frac{\partial J}{\partial b_m} = 0$$

therefore,

$$\mathbf{M}_1 \mathbf{b} = \mathbf{M}_2$$

$$b = {M_1}^{-1} M_2 \tag{5}$$

where

$$\boldsymbol{M_1} = \begin{pmatrix} \gamma_1 & -\lambda_{12} & \cdots & -\lambda_{1m} \\ -\lambda_{21} & \gamma_2 & \cdots & -\lambda_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda_{m1} & -\lambda_{m2} & \cdots & \gamma_m \end{pmatrix}; \boldsymbol{M_2} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix}$$

in which,

$$\gamma_j = n \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\ln(x_{j,i}) \right]^2 \right\} - \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} \left[\ln(x_{j,i}) \right] \right\}^2 \tag{6}$$

$$\xi_j = n \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{j,i}) \ln(y_j) \right] - \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{j,i}) \right] \left[\sum_{i=1}^{n} \ln(y_j) \right] \tag{7}$$

$$\lambda_{j,k} = \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{j,i}) \right] \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{k,i}) \right] - n \left[\sum_{i=1}^{n} \sum_{j=1}^{m} \ln(x_{j,i}) \ln(x_{k,i}) \right] \tag{8}$$

The subscript k represents the value of second subscript of the term ρ. The matrix as shown in (5) can be solved using various methods such Cramer rules, Gauss elimination, Gauss-Jordan elimination, LU decomposition or any other suitable iterative schemes. Therefore, the coefficients **b** can be obtained, while $a = \exp(A)$. Two variants of MPLSM are suggested in this paper as well: Normalised MPLSM and Iterative MPLSM, as further improvement of the method.

2.1. Normalised MPLSM

The above formulation possesses two flaws: inability of dealing with negative value factors since natural logarithmic function is considered and deciding the ranking of dominance. Therefore, the formulation can be further improved via the normalisation of the factors. In Normalised MPLSM, the factors must be set within the range of the range of [**C**,**C** + X], where **C** is a *positive integer* namely 0.05, while X is integer ratio.

$$X = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{9}$$

2.2. Iterative MPLSM

MPLSM also experiences numerical instability especially *high number* set of *random* factors are involved. To improve the numer-

ical stability, iteration is needed. The iterative procedure can be completed by imposing (10) as the "new" y (y'). The new values of the a (a') and $\mathbf{b}$ ($\mathbf{b}'$) will be as in (11) and (12) respectively. The iterative process will be conducted upon reaching convergence, as shown in the algorithm flow chart in Fig. 1. The examples for this will be illustrated in Section 3.4.

$$y' = y/y^h \tag{10}$$

$$a = aa' \tag{11}$$

$$\mathbf{b} = \mathbf{b} + \mathbf{b} \tag{12}$$

In the comparison of the accuracy of the computed results, the root mean square error (RMSE) and coefficient of determination, R^2 is applied. The equations of RMSE and R^2 can be expressed as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(y_i - y_i^h)} \tag{13}$$

$$R^2 = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{\left(n\sum x^2 - (\sum x)^2\right)\left(n\sum y^2 - (\sum y)^2\right)}} \tag{14}$$

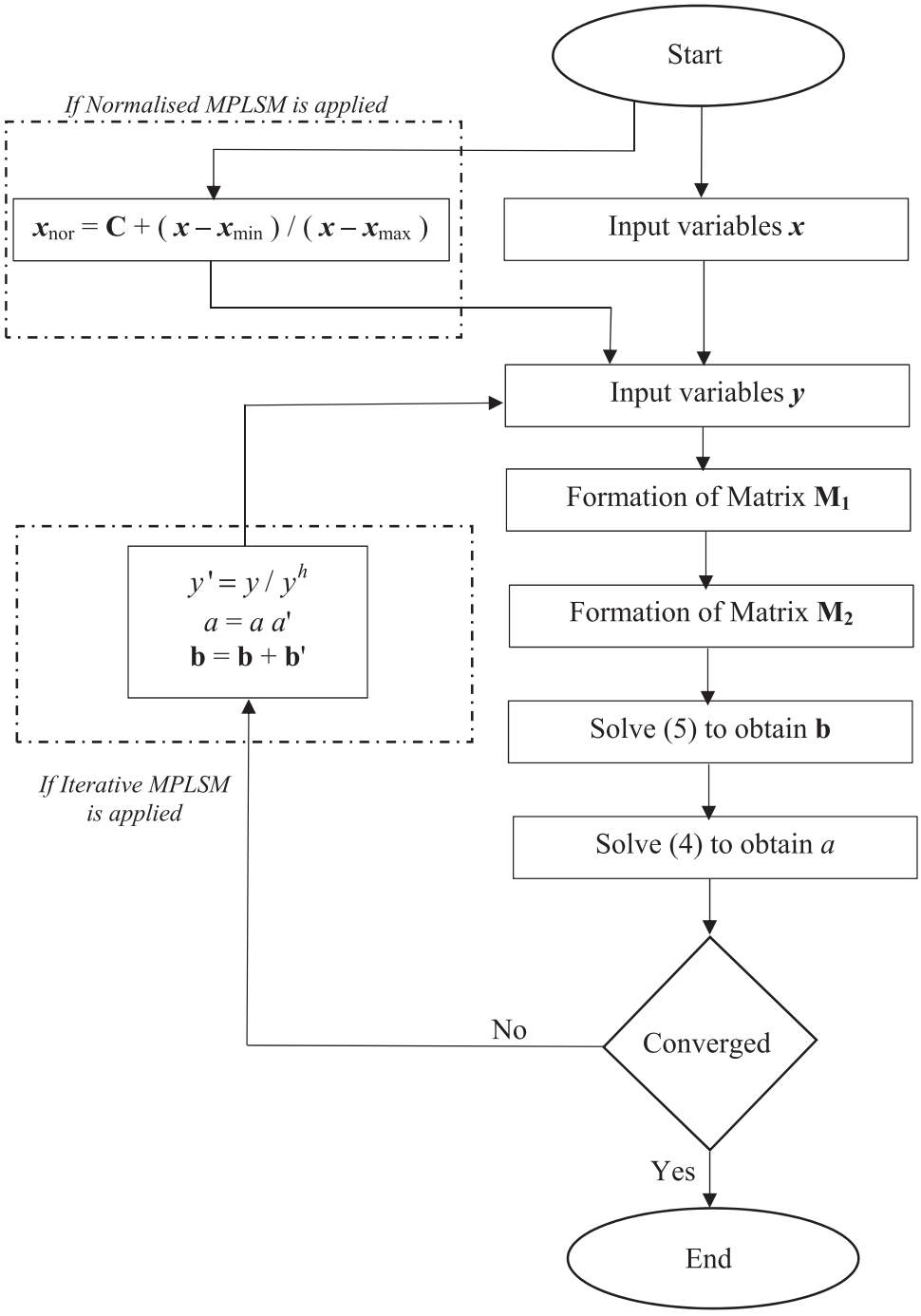

Fig. 1. Algorithm for complete implementation of MPLSM.

3. Examples of MPLSM applications

In this section, few examples are considered to examine the practicality of MPLSM and its variants. In Section 3.1, the basic applications of MPLSM will be illustrated. Meanwhile in the following sub-sections, MPLSM will be applied to rectify the problems of the works done via RSM. They are discussed in the following sections: (3.2) accurate determination of ranking of dominance; (3.3) application on random multi-factors experiments; and (3.4) highly accurate approximation for large random factors' set studies.

3.1. Basic numerical examples

Considering the basic heating flow condition for a fully developed turbulent flow in a smooth circular pipe [41], Nusselt number (Nu) can be expressed as a function of Reynolds number (Re) and Prandlt number (Pr), as indicated in (15).

$$\mathrm{Nu} = 0.023Re^{4/5}\mathrm{Pr}^{0.4} \tag{15}$$

The value of Nu is generated from the factors of Re and Pr in the range of 10^4 to 10^6 and 0.6 to 160 respectively, with the maximum error of 30%. The data are shown in Table 1. The equation formulated via RSM and MPLSM can be illustrated as in (16) and (17) respectively.

$$\mathrm{Nu}^h = 187.62712 + 2.1619 \times 10^{-3}Re + 11.1697\mathrm{Pr} + 4.8823 \times 10^{-5}Re\mathrm{Pr} \tag{16}$$

$$\mathrm{Nu}^h = 0.0281Re^{0.7757}\mathrm{Pr}^{0.4280} \tag{17}$$

The RMSE for (16) and (17) is 1608.0497 and 1221.0167 respectively, while the R^2 for (16) and (17) is 0.778 and 0.874 respectively. The comparison is made in both Table 1 and Fig. 2. Both the accuracy metrics suggest that the equation constructed through MPLSM outperforms RSM, and this implies that some physical phenomena are more suitable to be correlated using MPLSM.

For the second numerical example, consider a U-tube manometer with length L, is filled with fluid with the dynamic viscosity μ and density ρ. When the manometer is subjected to the pressure difference, it will vibrate if the radius of the manometer is larger than the critical radius, R_{cr} [42], which can be defined mathematically as in (18).

Table 1
Relationship between Prandtl number (Pr) and Reynolds number (Re) to Nusselt number (Nu) for a heating turbulent flow in a circular smooth pipe.

Set	Pr	Re	Prescribed Nu (±30%)	Predicted Nu via RSM	Predicted Nu via MPLSM
1	80.3	505,000	3398.9533	4158.5641	4874.9223
2	160	1,000,000	8840.2758	11957.8471	11124.2226
3	160	10,000	222.0577	2074.6098	312.5141
4	0.60	1,000,000	946.4078	2385.5583	1018.4860
5	80.3	505,000	4612.8652	4158.5641	4874.9223
6	0.60	505,000	547.9128	1300.8998	599.5124
7	0.60	10,000	32.6874	216.2412	28.6125
8	80.3	505,000	5341.2123	4158.5641	4874.9223
9	80.3	505,000	5098.4299	4158.5641	4874.9223
10	80.3	1,000,000	8806.5003	7171.7027	8281.7969
11	160	505,000	9596.2199	7016.2285	6548.0622
12	80.3	505,000	6312.3418	4158.5641	4874.9223
13	80.3	10,000	273.8782	1145.4255	232.6615
		RMSE		1608.0497	1221.0167

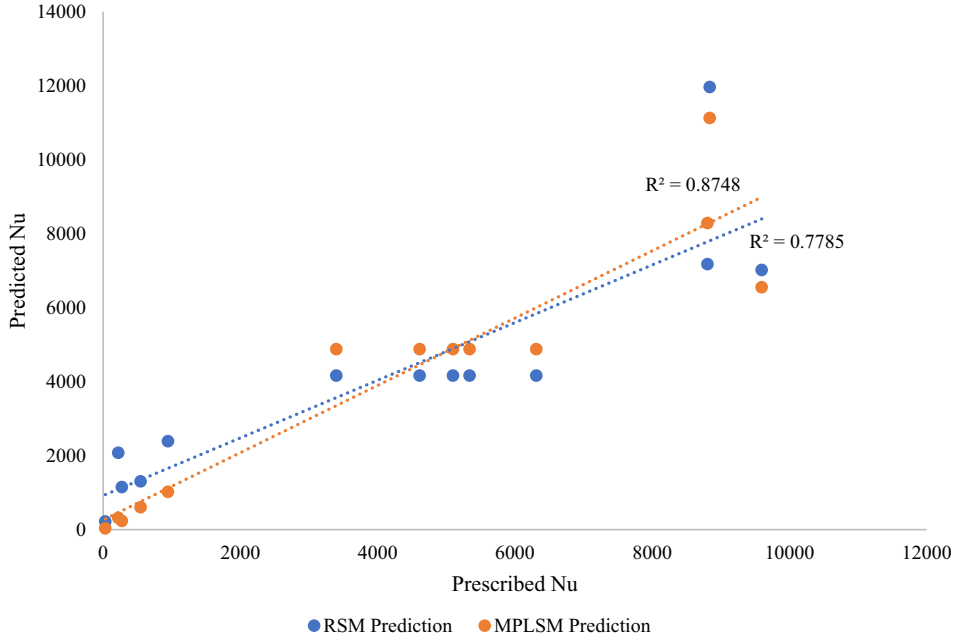

Fig. 2. Comparison between the R^2 for the relationship between Re and Pr with Nu formulated via RSM and MPLSM.

$$R_{cr} = \left(\frac{6\mu^2 L}{\rho^2 g}\right)^{1/4} \quad (18)$$

R_{cr} is computed with the maximum error of 30% too based on the randomly generated value of factors as shown in Table 2. The range of the factors are: $1 \times 10^{-5} \leq \mu \leq 0.01$, $0.01 \leq L \leq 0.1$ while $1000 \leq \rho \leq 3000$. The approximated equation contrived via RSM and MPLSM is (19) and (20) respectively.

$$R_{cr} = 2.198 \times 10^{-4} + 0.0885\mu + 1.9613 \times 10^{-3} L - 9.7249 \times 10^{-8} \rho \quad (19)$$

$$R_{cr} = 0.2777\mu^{0.512} L^{0.1905} \rho^{-0.3664} \quad (20)$$

From Table 2 and Fig. 3, the RMSE and R^2 for RSM equation is 0.000182792 and 0.7607 respectively. Equation obtained via MPLSM indeed performs better by achieving a smaller RMSE of 0.000157617 and a higher R^2 of 0.8254.

From these two numerical examples, it is clear that there are some physical phenomena whose factors existed as ratio form between each other. Apart from its superior simplicity where only one term exists in the equation, the importance of MPLSM will be more eminent for the investigation of dimensionless constant such

Table 2
Relationship between length of manometer (μ), fluid density (ρ) and fluid dynamic viscosity (μ) to the critical damping radius (R_{cr}).

Set	μ	μ	ρ	Prescribed R_{cr} (±30%)	Predicted R_{cr} via RSM	Predicted R_{cr} via MPLSM
1	0.01000	0.055	2000	0.0006703	0.0010182	0.0009335
2	0.00500	0.055	3000	0.0003874	0.0004793	0.0005649
3	0.00500	0.055	2000	0.0005760	0.0005766	0.0006553
4	0.00500	0.055	1000	0.0008144	0.0006734	0.0008444
5	0.00001	0.100	1000	0.0000423	0.0003196	0.0000393
6	0.00001	0.100	3000	0.0000258	0.0001251	0.0000263
7	0.00001	0.010	3000	0.0000145	−0.0000514	0.0000169
8	0.00500	0.055	2000	0.0004948	0.0005766	0.0006553
9	0.00001	0.055	2000	0.0000385	0.0001341	0.0000272
10	0.00500	0.100	2000	0.0010227	0.0006644	0.0007340
11	0.01000	0.010	3000	0.0006382	0.0008327	0.0005815
12	0.01000	0.010	1000	0.0011054	0.0010272	0.0008697
13	0.00500	0.055	2000	0.0008469	0.0005761	0.0006550
14	0.00500	0.055	2000	0.0004740	0.0005766	0.0006553
15	0.01000	0.100	1000	0.0011008	0.0012037	0.0013487
16	0.01000	0.100	3000	0.0010441	0.0010092	0.0009017
17	0.00500	0.055	2000	0.0007456	0.0005766	0.0006553
18	0.00001	0.010	1000	0.0000196	0.0001431	0.0000253
19	0.00500	0.055	2000	0.0008810	0.0005766	0.0006553
20	0.00500	0.010	2000	0.0005754	0.0004883	0.0004736
		RMSE			0.000182792	0.000157617

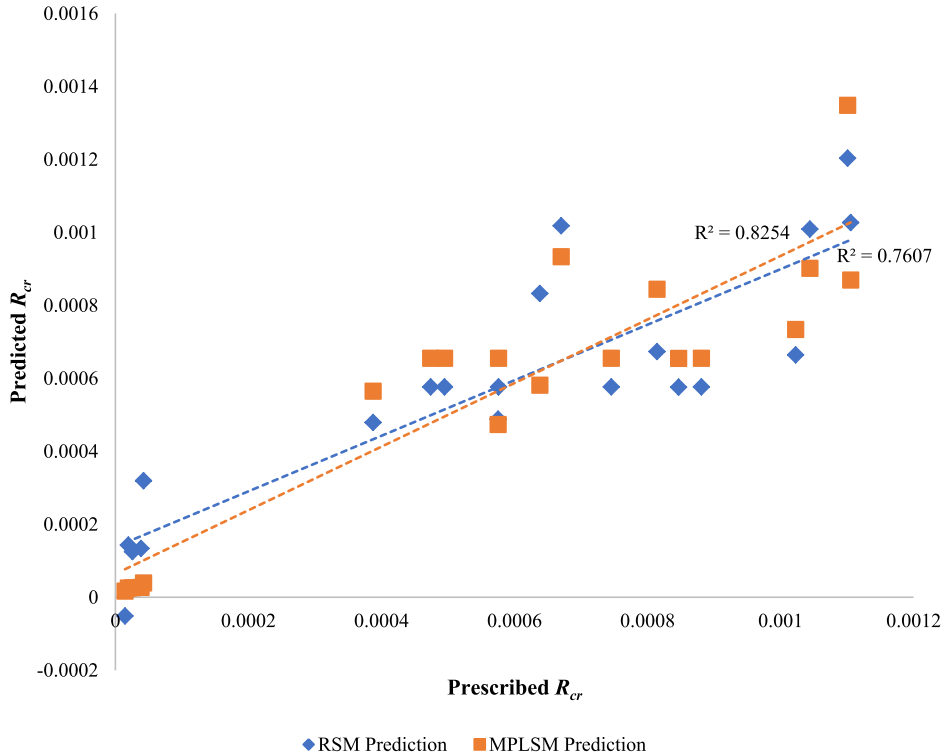

Fig. 3. Comparison between the R^2 for the relationship between L, μ and ρ with R_{cr} formulated via RSM and MPLSM.

as fine-structure constant in quantum field theory, Morton number in bubble dynamics and Zeldovich number in combustion. MPLSM henceforth renders robust mathematical alternative for such multivariable studies.

3.2. Accurate determination of ranking of factors' dominance

Normalised MPLSM is able to accurately determine the ranking of factors' dominance based on the indices of the approximated equation. Validation is made based on the works via RSM, where the significance ranking can be determined accordingly. In MPLSM, the larger the norm of indices, the larger the dominance of the factors. The work of Huang et al. [43] and Wong et al. [44] are used as the examples for validation.

Huang et al. [43] investigated the optimisation process of liposomal encapsulation of piceid where liquid content (x_{21}), ultrasound power (x_{22}) and ultrasound time (x_{23}) are taken as the factors, which influence the encapsulation efficiency (y_{21}), absolute loading (y_{22}) and particle size (y_{23}). The Normalised MPLSM equations for the study are from (21) to (23). The indices of the factors are directly representing the ranking of dominance, which reach the similar conclusion with the p-value in RSM, as shown in Table 3.

$$y_{21} = 66.6598x_{21}^{0.2474}x_{22}^{-0.0182}x_{23}^{-0.0648} \tag{21}$$

$$y_{22} = 2.1854x_{21}^{-0.1147}x_{22}^{-0.0178}x_{23}^{-0.0655} \tag{22}$$

$$y_{23} = 659.5229x_{21}^{0.0915}x_{22}^{-0.1506}x_{23}^{-0.7133} \tag{23}$$

Meanwhile Wong et al. [44] optimised Puerariaiso flavonoids (y_{24}) via the factors of time (x_{24}), ethanol concentration (x_{25}) and liquid-to-solid ratio (x_{26}). The comparison is made in Table 4 and the conclusion on the ranking of dominance determined via Normalised MPLSM is the same with the implication of p-value in RSM. The Normalised MPLSM equation for this investigation is (24).

$$y_{24} = 659.5228x_{24}^{-0.0566}x_{25}^{-0.1413}x_{26}^{-0.0357} \tag{24}$$

However, do note that the R^2 of the Normalised MPLSM equation must be at least 0.5 to be a legitimate equation to represent majority (>50%) of the data set, in order to warrant ranking of dominance of these factors. For example, in the work of Qu et al. [45], which optimised CS2-RH ratio (x_{27}), Xanthantion contact time (x_{28}) and Xanthantion temperature (x_{29}) for obtaining cadmium uptake (y_{25}), the Normalised MPLSM equation as in (25) possesses a very low R^2 of only 0.085. The misleading indication is shown in Table 5.

$$y_{25} = 125.0385x_{27}^{-0.1358}x_{28}^{-0.0803}x_{29}^{-0.1086} \tag{25}$$

The shortcoming of RSM is its inability in determining the ranking of significance when the p-value is less than 0.0001. This can be easily addressed via Normalised MPLSM. With this, several works of process optimisation [46–51] where the p-value of all the factors are less than 0.0001 are re-visited. The result of re-investigation is compiled in Table 6, and ranking of dominance can be accurately determined.

3.3. Applications on random multi-factors experiment

There are abundance of scientific experiments, which involve factors with random values. This is a deadlock for RSM to analyse such cases. Nonetheless, MPLSM could address this issue well. One example is taken from the work of Akbari-Sharbaf et al. [52], which investigated nanoscale graphene doping, which plays an imperative role in nanoelectronics and sustainable energy. In their study, the electrical conductivity (σ) is correlated with work function (φ) and area coverage fraction (f_j), in which their data is tabulated in Table 7. Since zero is involved in the value of factors, Normalised MPLSM is applied and the associated equation is shown in (26). The RMSE and R^2 is 6.3059×10^{-4} and 0.981 respectively, which implies an excellent fitting for this random data.

$$\sigma = 0.0089\varphi^{1.0447}f_j^{3.3212} \tag{26}$$

Table 3
Study of ranking of dominance for the factors in the encapsulation of piceid [43].

Equation	Details	x_{21}	x_{22}	x_{23}	Details	x_{21}	x_{22}	x_{23}
(21)	p-value	0.0006	0.6480	0.0392	Indices	0.2474	0.0182	0.0648
	Ranking of dominance	1	3	2	Ranking of dominance	1	3	2
(22)	p-value	0.0122	0.7235	0.0624	Indices	0.1147	0.0178	0.0655
	Ranking of dominance	1	3	2	Ranking of dominance	1	3	2
(23)	p-value	0.0833	0.0270	<0.0001	Indices	0.0915	0.1503	0.7133
	Ranking of dominance	3	2	1	Ranking of dominance	3	2	1

Table 4
Study of ranking of dominance for the factors in the optimisation of Puerariaiso flavonoids [44].

	p-value			Norm of indices		
	x_{24}	x_{25}	x_{26}	x_{24}	x_{25}	x_{26}
Ranking of dominance	0.0003	0.0001	0.7770	0.0556	0.1413	0.0357
	2	1	3	2	1	3

Table 5
Erroneous indication of ranking of dominance for the factors in the optimisation of cadmium uptake [45] via normalised MPLSM.

	p-value			Norm of indices		
	x_{27}	x_{28}	x_{29}	x_{27}	x_{28}	x_{29}
Ranking of dominance	0.3577	0.0178	0.4879	0.1358	0.0803	0.1086
	2	1	3	1	3	2

Table 6
Revisited works [46–51] for determination of ranking of dominance of factors.

Reference	Normalised MPLSM Equation	R^2	Factors	Norm of Indices	Ranking of Dominance
[46]	Force = 45.5944(Laser power)$^{1.0158}$(Scaling speed)$^{-0.4653}$(Spot diameter)$^{0.4529}$	0.634	Laser power	1.1058	1
			Sealing speed	0.4653	2
			Spot diameter	0.4529	3
[47]	Yield of quercetin = 76.6182(Temperature)$^{-0.1070}$(Time)$^{-0.1046}$(Concentration)$^{-0.1087}$	0.699	Temperature	0.1070	2
			Time	0.1046	3
			Concentration	0.1087	1
[48]	Cell viability = 3.1418(Power)$^{1.5067}$(Time)$^{0.5185}$(Pulse duty ratio)$^{-0.0197}$	0.866	Power	1.5067	1
			Time	0.5185	2
			Pulse duty ratio	0.0197	3
[49]	COD Removal = 77.5582(Temperature)$^{-0.0132}$(Contact time)$^{0.1099}$(Absorbent dosage)$^{0.0272}$(pH)$^{-0.1753}$	0.745	Temperature	0.0132	4
			Contact time	0.1099	3
			Absorbent dosage	0.7691	1
			pH	0.1753	2
[50]	% OMR = 89.9754(Pulp density)$^{0.0548}$(Oil dosage)$^{0.0996}$(Agglomeration)$^{0.0272}$(Particle size)$^{-0.0019}$	0.937	Pulp density	0.0548	2
			Oil dosage	0.0996	1
			Agglomeration	0.0272	3
			Particle size	0.0019	5
[51]	Absorbent capacity = 351.2487(Contact time)$^{0.1007}$(Absorbent dosage) $^{-0.2875}$(CPX Concentration)$^{0.4057}$(Solution pH)$^{0.1017}$	0.725	Contact time	0.1007	4
			Absorbent dosage	0.2875	2
			CPX concentration	0.4057	1
			Solution pH	0.1017	3

Table 7
Relationship between work function and area coverage fraction with electrical conductivity [52].

Set	Work function	Coverage fraction	Electrical conductivity
1	4.58667	0.0000	0.125
2	4.47333	17.1429	0.550
3	4.43333	32.8251	0.950
4	4.38000	51.4286	1.600

3.4. Highly accurate approximation for random large factors' set studies

When the number of sets of random variables increases, numerical instability may incur. Iterative variant of MPLSM is implemented as the remedy. Consider a set of data consists of six random factors correlated with one variable y_{41} as in (27). The range of the factors are prescribed as: $1 \le x_{41} \le 10$, $5 \le x_{42} \le 16$, $2 \le x_{43} \le 15$, $3 \le x_{44} \le 14$, $7 \le x_{45} \le 9$ and $2 \le x_{46} \le 3$. There are 50 intervals equally divided between them. When the Iterative MPLSM is administered, the result of convergence study can be shown in Table 8.

$$y_{41} = 2.5x_{41}^{3.45}x_{42}^{0.36}x_{43}^{2.41}x_{44}^{-1.78}x_{45}^{0.567}x_{46}^{4.3} \tag{27}$$

From Table 8, 20 iterations are required to converge and coincide with the indices in (27). The numerical overshoot or undershoot in the MPLSM can therefore be resolved.

Another example is the work by Hakkarainen et al. [53] which involves six variables (irradiance, x_{47}; ethanol specimen, x_{47}; thickness, x_{49}; ignition time, x_{410}; heat release rate per unit area, x_{411}; total heat released per unit area, x_{412}) in investigating the net heat combustion (y_{42}) of small scale pool fire test of ethanol-water mixture. Iterative MPLSM reported an equation of (28) for their works.

Table 8
Convergence study of Iterative MPLSM for large factors' set approximation.

Iteration	Indices						Accuracy	
	x_{41}	x_{42}	x_{43}	x_{44}	x_{45}	x_{46}	RMSE	R^2
1	1382.8	3.40	−0.19	−0.80	0.170	4.3	9.0973×10^6	0.953
2	74.2167	3.45	−0.85	0.71	1.169	4.3	1.1072×10^7	0.977
3	15.4647	3.45	−0.29	1.49	0.891	4.3	7.0295×10^6	0.994
4	65.6567	3.45	0.01	1.92	0.741	4.3	9.7141×10^7	0.998
5	4.2318	3.45	0.17	2.14	0.661	4.3	2.6572×10^6	0.999
6	3.3173	3.45	0.26	2.26	0.617	4.3	1.0497×10^6	0.999
7	2.9105	3.45	0.31	2.33	0.594	4.3	88,441	1.000
8	2.7128	3.45	0.33	2.37	0.582	4.3	32,682	1.000
9	2.6122	3.45	0.34	2.39	0.575	4.3	30,643	1.000
10	2.5597	3.45	0.35	2.40	0.571	4.3	32,334	1.000
11	2.5319	3.45	0.36	2.41	0.569	4.3	26,294	1.000
12	2.5171	3.45	0.36	2.41	0.568	4.3	13,827	1.000
13	2.5092	3.45	0.36	2.41	0.568	4.3	98,926	1.000
14	2.5049	3.45	0.36	2.41	0.567	4.3	29,983	1.000
15	2.5026	3.45	0.36	2.41	0.567	4.3	15,909	1.000
16	2.5014	3.45	0.36	2.41	0.567	4.3	5866	1.000
17	2.5008	3.45	0.36	2.41	0.567	4.3	4896	1.000
18	2.5004	3.45	0.36	2.41	0.567	4.3	2448	1.000
19	2.5002	3.45	0.36	2.41	0.567	4.3	1224	1.000
20	2.5001	3.45	0.36	2.41	0.567	4.3	612	1.000

The value of R^2 obtained is 0.9961, which implies an excellent equation prediction for this study.

$$y_{42} = 1.0211x_{47}^{-0.073}x_{48}^{0.1038}x_{49}^{-1.0066}x_{410}^{-0.0057}x_{411}^{-0.0067}x_{412}^{0.9947} \quad (28)$$

4. Conclusion

Multivariable Power Least Squares Method (MPLSM) with its normalised and iterative variants is developed as a robust complementary computational tool for RSM. The proposed techniques could be an alternative numerical technique when RSM does not work well, especially in terms of its prediction accuracy, arrangement of factor significance ranking, randomness of factors' input and complexity incurred when the number of set of factors increases. MPLSM provides a simple yet powerful numerical platform for future multivariable statistical analysis in science and engineering.

Acknowledgement

This work was supported by UCSI University Centre of Excellence for Research, Value Innovation and Entrepreneurship [grant number Proj-In-FETBE-038]; and Takasago Thermal Engineering Limited [grant number R.K130000.7343.4B262]. The contribution of ideas from Ir Dr. Tan Lian See throughout the research, especially on the design of experiment, is highly appreciated as well.

References

[1] Galton F. Natural inheritance. New York; 1894.
[2] Rawlings JO, Pantula SG, Dickey DA. Applied regression analysis: a research tool. 2nd ed. Springer; 1998.
[3] Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press; 2006.
[4] Höskuldsson A. Common framework for linear regression. Chemomet Intell Lab Syst 2015;145:250–62.
[5] Nimon KF, Oswald FL. Understanding the results of multiple linear regression: beyond standardised regression coefficients. Organ Res Methods 2013;16(4):650–74.
[6] Souza LC, Souza RMCR, Amaral GJA, Filho TMS. A parameterised approach for linear regression of interval data. Knowl-Based Syst 2017;131:149–59.
[7] da Silva AR, de Oliveira Lima A. Geographically weighted beta regression. Spatial Stat 2017;21:279–303.
[8] Wang L, Zhao Y, Chen X, Zanh D. A Newton algorithm for multivariate total least squares problems. Acta Geodaet Cartogr Sin 2016;45(4):411–7.
[9] Lancaster P, Salkauskas K. Surfaces generated by moving least squares methods. Math Comput 1981;37:141–58.
[10] Datolli G, Lorenzutta S, Mancho AM, Torre A. Generalised polynomials and associated operational identities. J Comput Appl Math 1999;108:209–18.
[11] Khuri AI, Mukhopadhyay S. Response surface methodology, WIREs. Comput Stat 2010;37:128–49.
[12] Cheok CY, Chin NL, Yusof YA, Talib RA, Law CL. Optimization of total phenolic content extracted from Garcinia mangostana Linn. hull using response surface methodology versus artificial neural network. Ind Crops Prod 2012;40:247–53.
[13] Ng LY, Mohammad AW, Ng CY, Rohani R. Optimization of polymeric membrane characteristics through thermal treatment and deposition of polyelectrolyte layers using response surface modeling. Adv Polym Technol 2014;34:21472.
[14] Mandal DK, Bhunia H, Bajpai PK, Chaudhari CV, Dubey KA, Varshney L. Radiation induced grafting of acrylic acid onto polypropylene film and its biodegradability. Radiat Phys Chem 2016;123:37–45.
[15] Mallaiah M, Reddy GV. Optimization studies on a continuous catalytic reactive distillation column for methyl acetate production with response surface methodology. J Taiwan Inst Chem Eng 2016;69:25–41.
[16] Shirvan KM, Mamourian M, Mirzakhanlari S, Ellahi R, Vafai K. Numerical investigation and sensitivity analysis of effective parameters on combined heat transfer performance in a porous solar cavity receiver by response surface methodology. Int J Heat Mass Transf 2017;105:811–25.
[17] Babaki M, Yousefi M, Habibi Z, Mohammadi M. Process optimization for biodiesel production from waste cooking oil using multi-enzyme systems through response surface methodology. Renew Energy 2017;105:465–72.
[18] Tgarguifa A, Abderafi S, Bounahmidi T. Energetic optimization of Moroccon distillery using simulation and response surface methodology. Renew Sustain Energy Rev 2017;75:415–25.
[19] Okan U. Optimization of shot peening parameters by response surface methodology. Surf Coat Technol 2016;305:99–109.
[20] Oberoi DPS, Sogi DS. Utilization of watermelon pulp for lycopene extraction by response surface methodology. Food Chem 2017;232:316–21.
[21] Baş D, Boyaci IH. Modeling and optimization I: Usability of response surface methodology. J Food Eng 2007;28:836–45.
[22] Ren H, Cheng J, Huang A. The complex interpolating moving least squares method. Appl Math Comput 2012;219:1724–36.
[23] Joldes GR, Chowdhury HA, Wittek A, Doyle B, Miller J. Modified moving least squares with polynomials bases for scattered data approximation. Appl Math Comput 2015;266:893–902.
[24] Li W, Song G, Yao G. Piece-wise moving least squares approximation. Appl Numer Math 2017;115:68–81.
[25] Tong H, Wu Q. Learning performance of regularised moving least squares regression. J Comput Appl Math 2017;325:42–55.
[26] Belystchko T, Yu LL, Gu L. Element-Free Galerkin methods. Int J Numer Meth Eng 1994;37:229–56.
[27] Sukumar N, Wright RW. Overview and construction of meshfree basis functions: From moving least squares to entropy approximation. Int J Numer Meth Eng 2007;70:181–205.
[28] Li X, Li S. Analysis of the complex moving least squares approximation and the associated element-free Galerkin method. Appl Math Model 2017;47:45–62.
[29] Schaefer S, McPhail T, Warren J. Image deformation using moving least squares. In: The 33rd international conference and exhibition on computer graphics and interactive techniques. p. 533–40.
[30] Lee YJ, Yoon J. Image zooming method using edge-directed moving least squares interpolation based on exponential polynomials. Appl Math Comput 2015;269:569–83.
[31] Stone MH. The generalised Weierstrass approximation theorem. Math Magaz 1948;21:237–54.
[32] Liu GR, Gu YT. An introduction to meshfree methods and their programming. Springer; 2005.
[33] Yadav AK, Malik H, Chandel SS. Application of rapid miner in ANN based prediction of solar radiation for assessment of solar energy resource potential of 76 sites in Northwestern India. Renew Sustain Energy Rev 2015;52:1093–106.
[34] Kumar S, Kaur T. Development of ANN based model for solar potential assessment using various meteorological parameters. Energy Proc 2016;90:587–92.
[35] Boukelia TE, Arslan O, Mecibah MS. ANN-based optimization of a parabolic trough solar thermal power plant. Appl Therm Eng 2016;107:1210–8.
[36] Goel A, Kamatchi P, Jayabalan P. ANN based methodology for active control of buildings for seismic excitation for different seismic zone of India. IFAC-Papers On Line 2016;49–1:095–9.
[37] Demirel MC, Venancio A, Kahya E. Flow forecast by SWAT model and ANN in Pracana basin, Portugal. Adv Software Eng 2009;40:467–73.
[38] Ghosh S, Chakraborty T, Saha S, Majumder M, Pal M. Development of the location suitability index for wave energy production by ANN and MCDM techniques. Renew Sustain Energy Rev 2016;59:1017–28.
[39] Tian Z, Gu B, Yang L, Lu Y. Hybrid ANN-PLS approach to scroll compressor thermodynamic performance prediction. Appl Therm Eng 2015;77:113–20.
[40] Prasada Rao K, Babu TV, Anuradha G, Rao BVA. IDI diesel engine performance and exhaust emission analysis using biodiesel with an artificial neural network (ANN). Egypt J Pet 2017;26:593–600.
[41] Incropera FP, Dewitt DP. Fundamentals of heat and mass transfer. 7th ed. John Wiley and Sons; 2011.
[42] Bird RB, Stewart WE, Lightfoot EN. Transport phenomena, revised 2nd ed., John Wiley and Sons; 2007.
[43] Huang SM, Kuo CH, Chen CA, Liu YC, Shieh CJ. RSM and ANN modelling-based optimization approach for the development of ultrasound-assisted liposome encapsulation of piceid. Ultrason Sonochem 2017;36:112–22.
[44] Wong KH, Li GQ, Li KM, Razmovski-Naumovski V, Chan K. Optimisation of Puerariaiso flavonoids by response surface methodology using ultrasonic-assisted extraction. Food Chem 2017;231:231–7.
[45] Qu J, Meng X, You H, Ye X, Du Z. Utilization of rice husks functionalised with xanthates as cost-effective biosorbents for optimal Cd(II) removal from aqueous solution via response surface methodology. Bioresour Technol 2017;241:1036–42.
[46] Wang W, Xiao Y, Wu X, Zhang J. Optimization of laser-assisted glass frit bonding process by response surface methodology. Opt Laser Technol 2016;77:111–5.
[47] Ravber M, Knez Z, Skerget M. Optimization of hydrolysis of rutin in subcritical water using response surface methodology. J Supercrit Fluids 2015;104:145–52.
[48] Liu J, Li L, Zhou L, Li B, Xu Z. Effect of ultrasound treatment conditions on Saccharomyces cerevisiae by response surface methodology. Microb Pathog 2017;111:497–502.
[49] Xu W, Zhao Q, Wang R, Jiang Z, Zhang Z, Gao X, et al. Optimization of organic pollutants removal from soil eluent by activated carbon derived from peanut shells using response surface methodology. Vacuum 2017;141:307–15.
[50] Yadav AM, Nikkam S, Gajbhiye P, Tyeb MH. Modeling and optimization of coal oil agglomeration using response surface methodology and artificial neural network approaches. Int J Miner Process 2017;163:55–63.
[51] Zhang B, Han X, Gu P, Fang S, Bai J. Response surface methodology approach for optimization of ciprofloxacin adsorption using activated carbon derived from the residue of desilicated rice husk. J Mol Liq 2017;238:316–25.
[52] Akbari-Sharbaf A, Ezugwu S, Ahmed MS, Cottam MG, Fanchini G. Doping graphene thin films with metallic nanoparticles: experiment and theory. Carbon 2015;95:199–207.
[53] Hakkarainen T, Korhonen T, Vaari J. Heat release characteristics of ethanol-water mixtures: small-scale experiments. Fire Saf J 2017;91:174–81.

Wah Yen Tey is working as a lecturer in UCSI University Kuala Lumpur, while pursing his PhD in Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. While working in computational fluid dynamics (CFD) and fluid structure interaction, he is interested in development and application computational mathematics in engineering.

Kiat Moon Lee received her PhD in Chemical Engineering from University of Malaya in 2015, with research interest in biomass and renewable energy. She applied Response Surface Methodology in some of her research works, and she is devoted to the development of experimental designs in chemical engineering.

Yutaka Asako was a professor in Tokyo Metropolitan University, Japan for 22 years and he is currently a professor in Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. He is interested in the development of computational methods to solve fluid dynamics and heat transfer problem. He is an active researcher in CFD and more than 200 of his articles have been published in high impact journals.

Lit Ken Tan received his PhD in Mathematics from Universiti Teknologi Malaysia in 2015. His research focuses on the topological mathematics and the development of applied mathematics to solve engineering issues.

Norio Arai was a professor in Tokyo University of Agriculture and Technology, Japan and Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia. His research interests are aerodynamics, fluid structure interactions and computational mechanics. He is AIAA Associate Fellow and AIAA Lifetime Member as well.