


RESEARCH

Open Access

Data-driven nonlinear constitutive relations for rarefied flow computations



Wenwen Zhao^{*} , Lijian Jiang, Shaobo Yao and Weifang Chen

* Correspondence: wwzhao@zju.edu.cn

School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, People's Republic of China

Abstract

To overcome the defects of traditional rarefied numerical methods such as the Direct Simulation Monte Carlo (DSMC) method and unified Boltzmann equation schemes and extend the covering range of macroscopic equations in high Knudsen number flows, data-driven nonlinear constitutive relations (DNCR) are proposed first through the machine learning method. Based on the training data from both Navier-Stokes (NS) solver and unified gas kinetic scheme (UGKS) solver, the map between responses of stress tensors and heat flux and feature vectors is established after the training phase. Through the obtained off-line training model, new test cases excluded from training data set could be predicated rapidly and accurately by solving conventional equations with modified stress tensor and heat flux. Finally, conventional one-dimensional shock wave cases and two-dimensional hypersonic flows around a blunt circular cylinder are presented to assess the capability of the developed method through various comparisons between DNCR, NS, UGKS, DSMC and experimental results. The improvement of the predictive capability of the coarse-graining model could make the DNCR method to be an effective tool in the rarefied gas community, especially for hypersonic engineering applications.

Keywords: Data-driven method, Shock wave structure, Hypersonic flow, Rarefied gas dynamics

1 Introduction

Due to non-equilibrium effects inside the shock wave with large macroscopic gradients on a scale of several mean free paths, traditional hydrodynamic methods such as Navier-Stokes equations are no longer valid. To resolve the steady shock wave structure of monatomic and diatomic gases physically, particle methods such as Direct Simulation Monte Carlo (DSMC) [1] and numerical schemes for Boltzmann equations have been extensively employed. However, the computational cost for both microscopic method and mesoscopic method is large in comparison with hydrodynamic methods, especially in the near-continuum flow regime with multiscale characteristics. The same defects are also encountered in multi-dimensional rarefied flows, especially hypersonic flows around a space vehicle. To overcome current defects, the most popular numerical methods which have been employed in shock wave structure or rarefied hypersonic flow computations are reviewed subsequently.

DSMC is a stochastic particle method to simulate rarefied non-equilibrium flows by tracking and counting representative molecules per cell. Not only for one-dimensional shock wave structure calculations, it has been widely applied in engineering applications after proposed because it could be easily implemented in the hypersonic reentry flow with thermal and chemical non-equilibrium effects and yields accurate results when the Knudsen number is large. However, the computational efficiency decays significantly in near-continuum flow or high-density regions in a multiscale flow field due to the surging number of particles per cell and limitations of time step and mesh size. In the mesoscopic method, the Boltzmann equation describes the evolution of gas velocity distribution function in theory. However, the complex collisional term in the right-hand side prevents the full Boltzmann equation from being solved widely. Linearized or simplified Boltzmann model equations such as the Bhatnagar-Gross-Krook (BGK) [2], Shakhov-BGK [3] and ellipsoidal statistical BGK (ES-BGK) [4] model equations were proposed to reduce the numerical complexity. In order to get the solutions of these approximations, discrete velocity and ordinate methods (DVM) such as gas-kinetic unified algorithms (GKUA) [5] are widely employed and a new framework named unified gas kinetic scheme (UGKS) [6] proposed by Xu and collaborators a few decades ago is regarded as a multiscale scheme for both hydrodynamic and rarefied regimes. In comparison with other kinetic solvers, UGKS could reveal accurate solutions in continuum flow regime with a large cell size and time step, which overcomes the cell size barrier requiring a mesh size on the order of the mean free path for DSMC and other kinetic schemes. However, to represent the velocity distribution function in velocity space, all these discrete velocity methods need huge memory consumption, which is too expensive and sometimes inaccessible in hypersonic flows. On the other hand, hydrodynamic methods also attract a lot of attention by extending its covering range in various approaches. Burnett equations [7] derived from the second order Chapman-Enskog expansion of VDF are classical high-order moment equations in literature. But the original formulation was proved unstable due to violation of the second law of thermodynamics, although some remedies have been reported in recent years [8]. Besides Burnett equations, Grad equations [9] and regularized moment equations, e.g. R-13 [10] and R-26 [11] equations, and other moment equations form the community of extended hydrodynamic equations (EHE). Eu [12] proposed generalized hydrodynamic equations (GHE) from a different viewpoint in 1992 and developed nonlinear coupled constitutive relations (NCCR) with Myong [13, 14] based on GHE subsequently. As a simplified approximation of GHE, NCCR includes the evolution equations of conservative variables and nonlinear coupled closure of stress tensor and heat flux, which is strictly consistent with the second law of thermodynamics [15]. Except DSMC, Boltzmann equation and moment equation solvers, many coupling models have also been proposed recently such as traditional CFD-DSMC hybrid method [16], unified gas-kinetic wave-particle method (UGKWP) [17], general synthetic iterative scheme (GSIS) [18], unified stochastic particle method based on the BGK model (USP-BGK) [19] and simplified unified wave-particle method (SUWP) [20]. These methods try to couple hydrodynamic and microscopic descriptions of the rarefied flow to construct an efficient and accurate hybrid scheme.

The study of shock waves by the above methods has been widely carried out to compare with experimental results in recent decades. Both monatomic and diatomic gases

are investigated in a wide range of upstream Mach numbers. Thermal non-equilibrium in shock waves with both translational and rotational temperatures is also investigated. In general, kinetic methods yield much more accurate macroscopic parameter profiles than hydrodynamic solvers especially at high Mach numbers. However, the computational cost of solving various Boltzmann equations is much larger than extended hydrodynamic equations due to the discretization in velocity space separately. Unfortunately, this deficiency will become unaffordable in three dimensional engineering applications and that could be the main reason why mesoscopic methods are still not successfully applied in predictions of aerodynamic and thermal loads on hypersonic vehicles.

While traditional fluid dynamics mainly focused on developing new physical models to increase their accuracy, another effective way arising in recent years to improve the predictive capability is data-driven modeling method. Especially in turbulence modeling, many research studies have been carried out by utilizing high-fidelity direct numerical simulation (DNS) or experimental data to construct a predictive data-driven Reynolds-averaged Navier-Stokes (RANS) model. Duraisamy et al. [21] summarized and reviewed these methods in detail and Xiao [22, 23] proposed a physical-informed machine learning approach (PIML) for reconstructing Reynolds stress modeling discrepancies based on DNS data. Ling et al. [24] proposed a tensor basis neural network (TBNN) by combining traditional neural network with Pope's Reynolds averaged turbulence model [25]. Using TBNN, they successfully predicted the separation phenomenon of angular vortex and wavy wall flow in pipe flow. Not only in turbulence modelling, Kutz [26] reviewed deep neural networks (DNNs) in fluid dynamics and predicted that it was only a matter of time before deep learning made its mark in the general area of high-dimensional, complex dynamical systems. In fact, data methods have been utilized widely in computational fluid dynamics successfully such as proper orthogonal decomposition (POD) or dynamic mode decomposition (DMD), which reduces computational costs significantly. Sekar [27] proposed a fast prediction method for airfoil flow field based on deep convolutional neural networks (CNN) and multi-layer perceptron (MLP). Bar-Sinai [28] proposed an optimal approximation method of partial differential equations based on the actual solutions of known flow governing equations, using neural networks to estimate the spatial derivatives. Mao [29] successfully modeled the high-speed air flow based on Euler equation by using the physical-informed neural networks (PINN) [30], and solved the one-dimensional Euler equation by using smooth solution and two-dimensional oblique shock wave solution with contact discontinuity. Raissi [31] tried to combine the PINN method with the convection diffusion equation of particles to propose the HFM method, which describes the technical characteristics and application prospects of the PINN method. All in all, using limited high-fidelity data to capture features that dominate physics and then improve the predictive capability of coarse-graining models will provide an effective tool in the fluid community. In rarefied gas dynamics regime, Zhang [32] utilized machine learning methods to derive various macroscopic governing equations for fluid dynamics and all training data comes from DSMC simulations. It is proved that the data-driven method to discover governing equations for a non-equilibrium flow is promising.

In this paper, data-driven nonlinear constitutive relations (DNCR) will be developed by machine learning method to provide a predictive approach for rarefied non-equilibrium flows. Both NS equations solver and UGKS solver provide training data

during training phase to map feature vectors and responses of stress tensor, heat flux and flow parameter derivatives between NS and UGKS data. Based on this off-line model, test cases excluded from training data set could be predicated rapidly and accurately by solving conventional equations with modified stress tensor and heat flux. The whole paper will be divided into four sections. “Data-driven nonlinear constitutive relations” will provide a detailed construction process of data-driven non-linear relations and the solving methodology. Numerical cases of shock wave structure and hypersonic flow passing through a blunt circle cylinder are presented in “Numerical simulations” to validate and assess the capability of the current DNCR method. And finally, conclusions and future work are summarized in the last section.

2 Data-driven nonlinear constitutive relations

2.1 Data-driven modeling approach

Data-driven approach is a suitable tool for building high-dimensional and complex regression functions between input features and output responses. Constitutive relations of macroscopic equations in rarefied regime are just impossible to specify analytically in math despite the fact that traditional moment equations try to extend their covering range by various expansion or simplification of velocity distribution function (VDF). Therefore, the intention to obtain non-equilibrium constitutive relations through a data-driven modeling approach is the main task for this paper.

In order to construct regression functions, we need to obtain training data first before adopting a machine learning method. Two numerical methods including NS equations solver and UGKS solver will be employed in current research to account for coarse-graining model and high-fidelity model respectively. The covering range or size of training cases is also important for the capability of the regression functions. Both local continuum and rarefied regions should be included inside the training cases to make it comprehensive. Another crucial issue is the choice of flow features as regression inputs. Due to non-equilibrium characters of target flows, we will choose local Knudsen number $Kn_{GLL}(\rho, T)$ and gradients of pressure ∇p , density $\nabla \rho$ and other favorable flow-field variants. Following the basic principles of a data-driven method, the input parameters should be obtained from the coarse-graining model directly, such as NS solver in current framework, because the high-fidelity data from UGKS solver is unavailable in the test cases. Secondly, the physical coordinates \mathbf{x} should not be chosen as features in the DNCR method to overcome limitation that the test case should have the same geometry and grid as the training case. And only local flow quantities are considered as well-chosen input parameters. All these adopted features in current research strictly satisfy the above principles.

2.2 Evolution equations of DNCR method

The basic macroscopic conservation laws of generalized hydrodynamics can be summarized as

$$\begin{pmatrix} \rho \\ \rho \mathbf{u} \\ \rho E \end{pmatrix}_t + \nabla \cdot \begin{pmatrix} \rho \mathbf{u} \\ \rho \mathbf{u} \mathbf{u} + p \mathbf{I} \\ (\rho E + p) \mathbf{u} \end{pmatrix} + \nabla \cdot \begin{pmatrix} 0 \\ \mathbf{\Pi} \\ \mathbf{\Pi} \cdot \mathbf{u} + \mathbf{Q} \end{pmatrix} = 0, \tag{1}$$

where $\mathbf{\Pi}$, \mathbf{Q} are stress tensor and heat flux, \mathbf{u} are macroscopic velocity vectors. To close

above conservation laws, constitutive relations that describe how the stress tensor, heat flux and mass diffusion respond to the gradients of velocity, temperature and mass fraction are needed. In the continuum region, the Navier-Stokes-Fourier constitutive relations provide a typical first order mathematical model, where the stress tensor and heat flux are expressed in terms of the first-order macroscopic velocity and temperature. However, NS equations are only valid in the continuum regime due to its intrinsic continuous hypothesis.

Beyond the constitutive relations, the conservation laws are regarded as a valid model in all scales such as continuum, transitional and rarefied flows. If we could find a set of proper constitutive relations in a desired flow, even discrete relations, the macroscopic quantities could be predicted physically by solving the closed mathematical model. However, the aim of finding analytical constitutive relations to describe the non-equilibrium characteristic becomes quite complicated for a complex system, where too many parameters need to be fitted, and the physical meaning of these terms becomes obscure. That's a key point for us to figure out how to obtain these exact constitutive discrete relations as Eq. (2) through the machine learning method.

$$\begin{aligned} \mathbf{\Pi}_{i,j} &= f(Kn_{GLL}, \nabla p, \nabla \rho, \rho, u_i, p, T) \\ \mathbf{Q}_i &= g(Kn_{GLL}, \nabla p, \nabla \rho, \rho, u_i, p, T) \end{aligned} \tag{2}$$

Once the discrete constitutive relations in every grid cell center are obtained, the conservation equations could be solved to convergence by explicit or implicit time-marching method. Take a two-dimensional flow as an example, the stress tensor and heat flux substituted into the macroscopic conservation laws are obtained as Eq. (3), where the local flow parameters with a superscript tip come from the trained off-line model.

$$\begin{cases} \tau_{xx} = \tau'_{xx} - \mu \left(\frac{4}{3} \frac{\partial u'}{\partial x} - \frac{2}{3} \frac{\partial v'}{\partial y} \right) + \mu \left(\frac{4}{3} \frac{\partial u}{\partial x} - \frac{2}{3} \frac{\partial v}{\partial y} \right) \\ \tau_{xy} = \tau_{yx} = \tau'_{xy} - \mu \left(\frac{\partial u'}{\partial y} + \frac{\partial v'}{\partial x} \right) + \mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \\ \tau_{yy} = \tau'_{yy} - \mu \left(\frac{4}{3} \frac{\partial v'}{\partial y} - \frac{2}{3} \frac{\partial u'}{\partial x} \right) + \mu \left(\frac{4}{3} \frac{\partial v}{\partial y} - \frac{2}{3} \frac{\partial u}{\partial x} \right) \\ q_x = q'_x - \kappa \frac{\partial T'}{\partial x} + \kappa \frac{\partial T}{\partial x} \\ q_y = q'_y - \kappa \frac{\partial T'}{\partial y} + \kappa \frac{\partial T}{\partial y} \end{cases} \tag{3}$$

Table 1 Flow features used as input in the regression for DNCR

Feature	Description
$Kn_{GLL}(\rho)$	Local Knudsen numbers of density
$Kn_{GLL}(p)$	Local Knudsen numbers of pressure
$Kn_{GLL}(T)$	Local Knudsen numbers of temperature
∇p	Pressure gradient
$\nabla \rho$	Density gradient
ρ, u, v, p, T	Key flow parameters

Table 2 Responses used as output in the regression for DNCR

Response	Description
τ_{ij}	Stress tensor
q_i	Heat flux vector
∇T	Temperature gradient
∇u	X-velocity gradient
∇v	Y-velocity gradient

With the constraint of above Eq. (3), the stress tensor and heat flux will automatically regress to the values of the off-line model as the gradients of flow parameters converge to a steady state. It is worth noting that the output of the trained off-line model not only includes stress tensor and heat flux, but also contains some gradients of key flow parameters.

2.3 Solution methodology

2.3.1 Training phase

1) Pseudo and real transition flow fields are obtained through NS equation solver and UGKS method respectively to establish the training data set which includes local conventional flow variables ρ, u, v, w, p, T , Knudsen number based on the gradients of density and temperature $Kn_{GLL}(\rho, T)$, stress tensor $\tau_{ij}^{UGKS}, \tau_{ij}^{NS}$, heat flux q_i^{UGKS}, q_i^{NS} and flow parameter gradients. During the numerical simulation process, an identical multi-block structure grid is employed currently for both numerical solvers in finite volume method to guarantee the data structure of the flow field consistent.

2) Establish the feature vector field $\mathbf{q}(\mathbf{x})$, e.g., local Knudsen numbers, pressure gradient, density gradient and flow parameters, based on the NS-predicted flow fields in Step

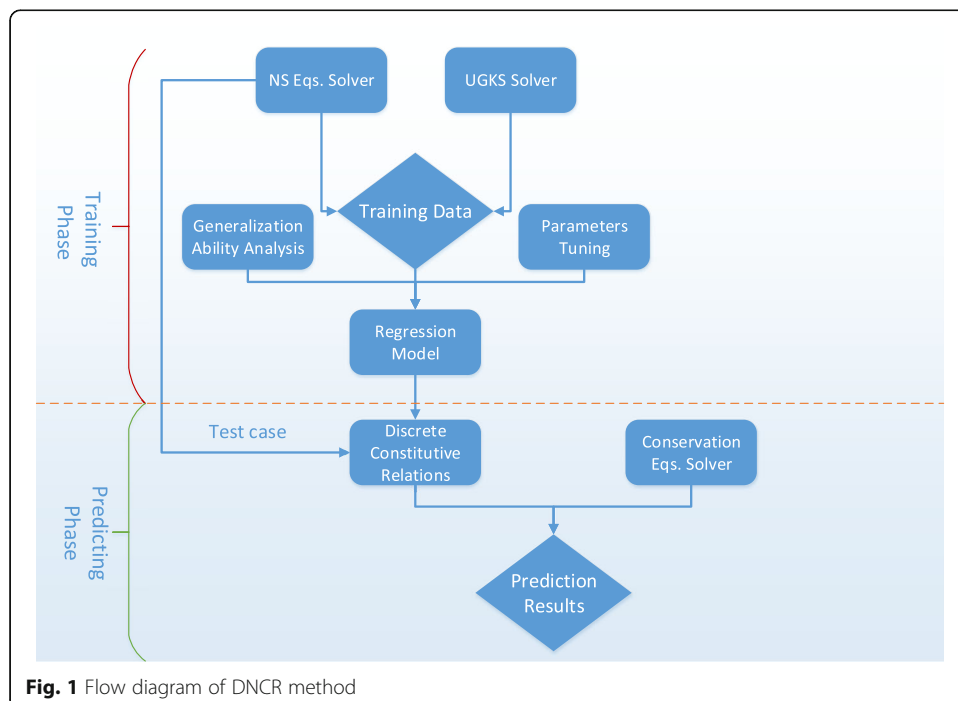


Fig. 1 Flow diagram of DNCR method

Table 3 Physical properties of Argon

	γ	Pr	$R(\text{m}^2/(\text{sec}^2 \cdot \text{K}))$	$T_{\text{ref}}(\text{K})$	$\mu_{\text{ref}}(\text{kg}/\text{sec} \cdot \text{m})$
Ar	1.667	0.667	208.16	300	2.272×10^{-5}

1. And compute the responses field $\Phi(\mathbf{x})$, where Φ represents stress tensor, heat flux, temperature gradient and velocity gradients, for the training set.

3) Construct regression functions $f: q \mapsto \Delta\Phi$ through a machine learning method based on the training data. The extremely randomized trees model [33] is employed in current research to train the regression relations.

The choice of the feature vectors is a key step for constructing regression functions and will affect the final precision of prediction greatly. The input flow features and output responses in the regression and descriptions are presented in Tables 1 and 2 as below.

In the above training procedure, the training set has to include different flow types to guarantee the applicable capability of the model. However, in the current framework of DNCR, only a few different typical inflow conditions are needed due to the multi-scale feature of transitional flow. Every grid cell will provide a basic training set and the local non-equilibrium degree will be totally different inside a multi-scale flow field. In other words, it is even possible to predict a flow field utilizing the DNCR method based on a training set that

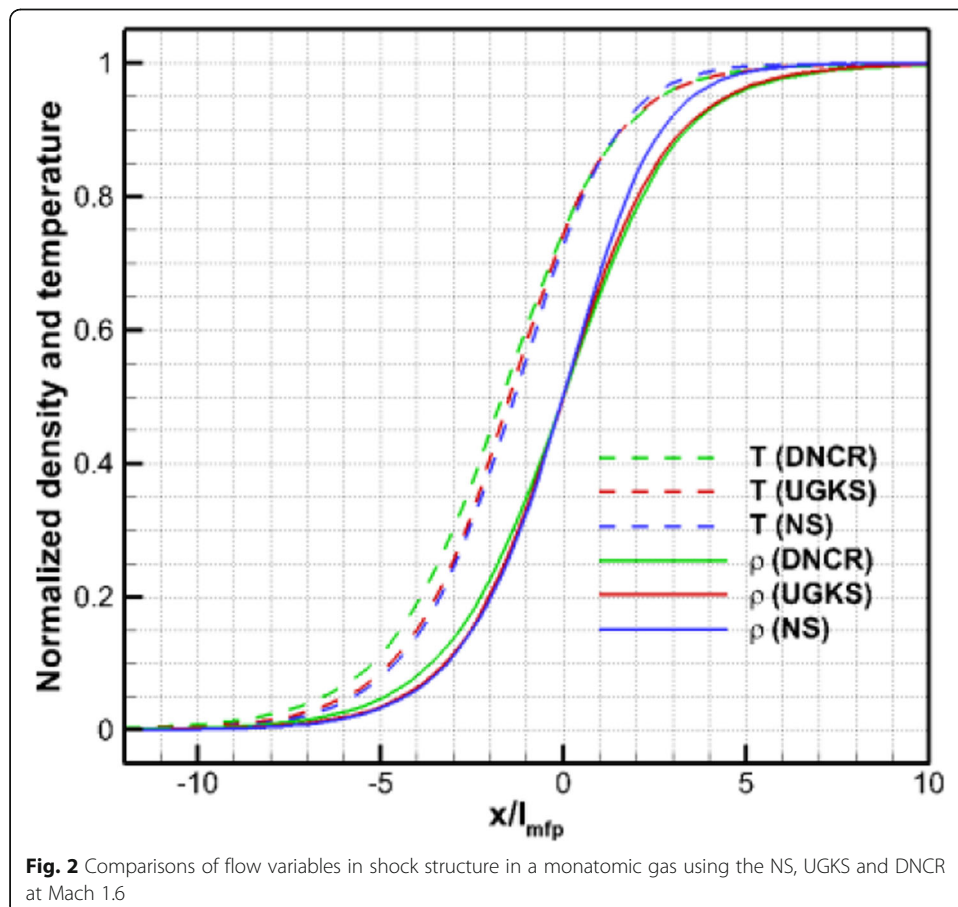
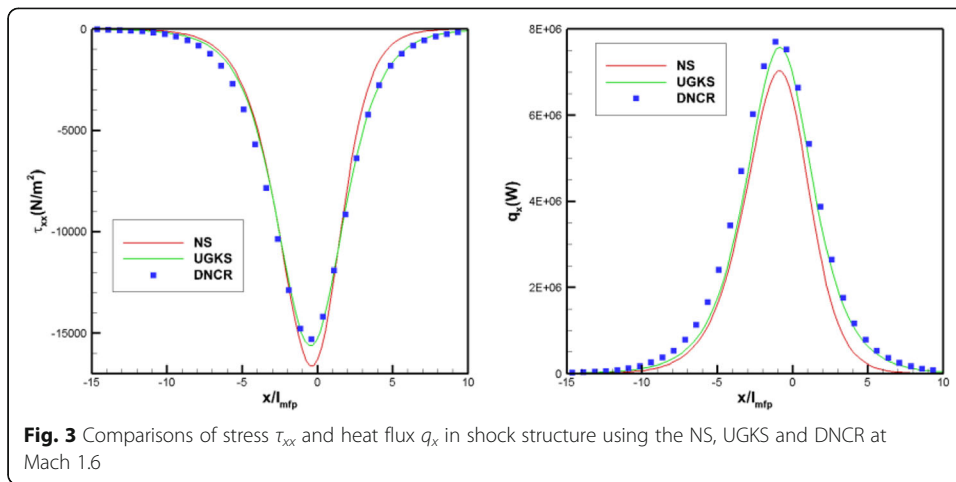
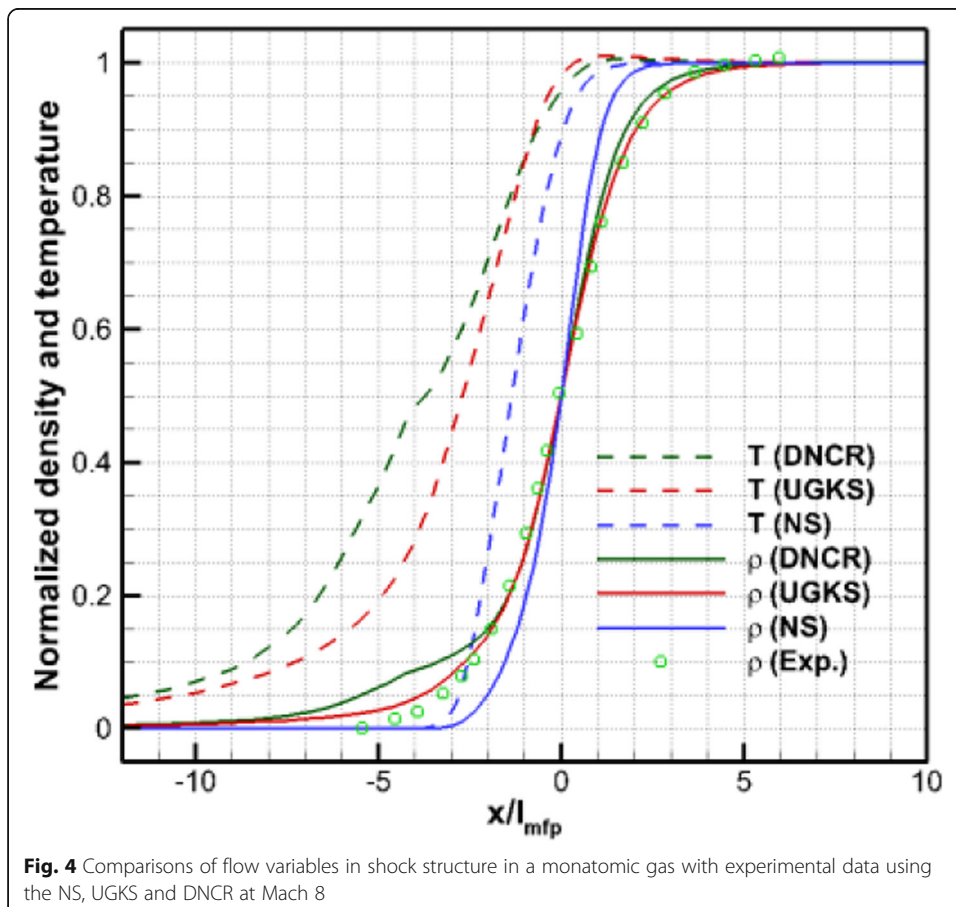


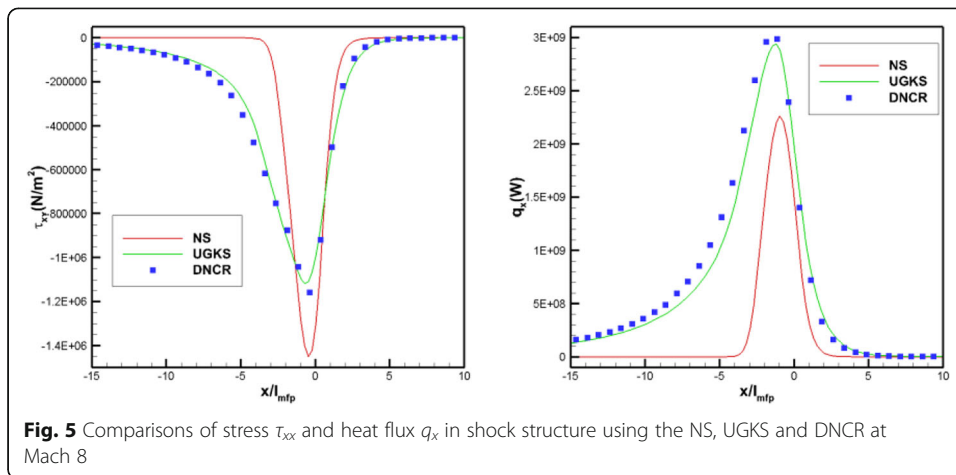
Fig. 2 Comparisons of flow variables in shock structure in a monatomic gas using the NS, UGKS and DNCR at Mach 1.6



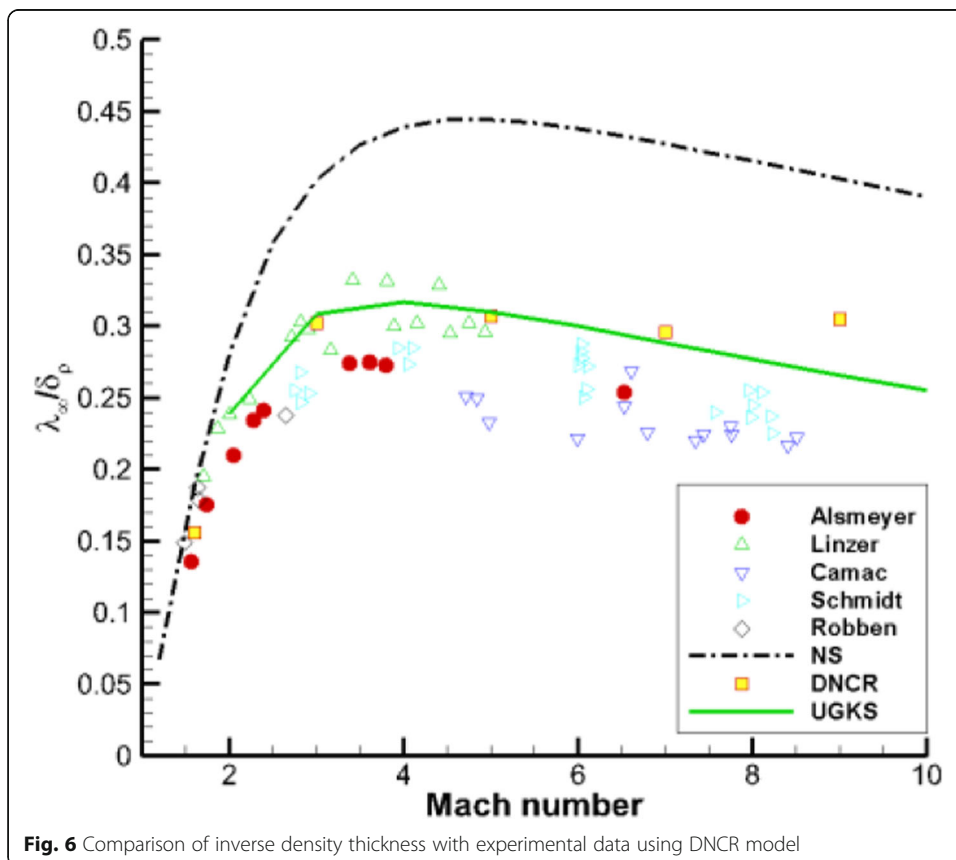
only includes one different case, e.g., predict test flow around a cylinder at $Kn = 0.1$ from the training set of flow field at $Kn = 0.2$. This feature makes the DNCR method applicable in flows which have similar characteristics, such as external rarefied flows around a spacecraft and so on.

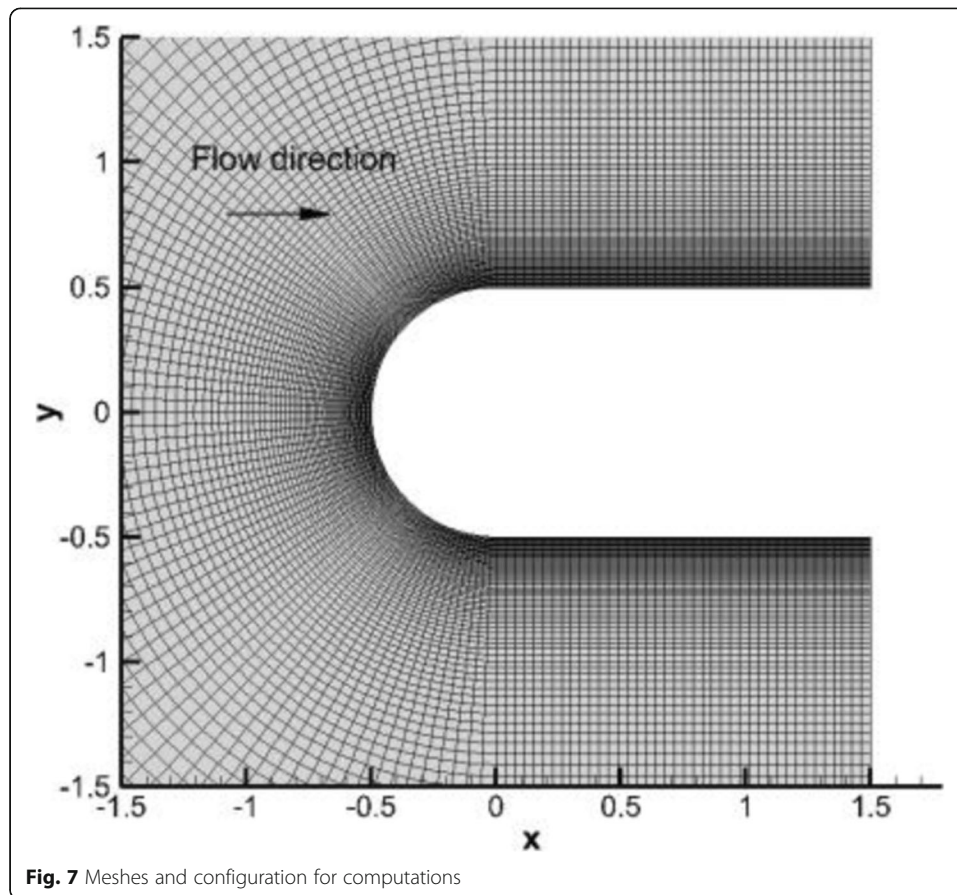
Moreover, without any geometric parameters in the feature vectors, the DNCR method could predict flow over different configurations from training set theoretically.





That means if the training flows consist of various elementary cases such as compression and expansion flows in transition regime, complex engineering flow which has a subset or all of them could be predicted well ultimately. This challenging task is the final goal of this method and is outside the scope of this study. We mainly focus on the demonstration of feasibility first and will achieve that goal gradually.





2.3.2 Predicting phase

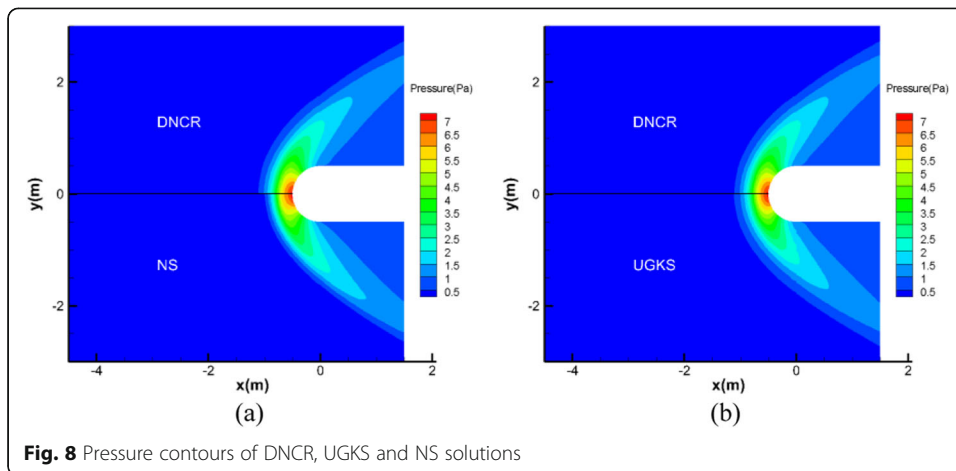
1) Perform NS simulations on the test flow and compute the feature vectors $\mathbf{q}(\mathbf{x})$ of the flow field.

2) Compute the responses for the test flow by querying the regress functions f and obtain the baseline stress tensor, heat flux and flow parameter gradients in each grid cell subsequently, which are the terms with a superscript tip in Eq. (3).

3) Solve conservation equations with baseline stress tensor, heat flux and flow parameter gradients in Step 2 until converging to a steady state and finish the whole procedure.

Finally, the test flow field with high-fidelity will be obtained without carrying out complex and expensive simulations through UGKS or DSMC methods. The whole flow diagram is shown in Fig. 1 with both training phase and predicting phase.

The above predicting phase is separate from the training phase because the maps between input features and target responses have already been established before prediction. Only NS feature vectors are needed in order to obtain the final discrete stress tensor and heat flux for test flow. Unlike linear NSF relations or nonlinear Burnett constitutive relations, there is no explicit formula in DNCR to close the conservation equations. And actually, the universality and simplicity of explicit constitutive relations to describe non-equilibrium phenomenon is also questionable. However, the constitutive



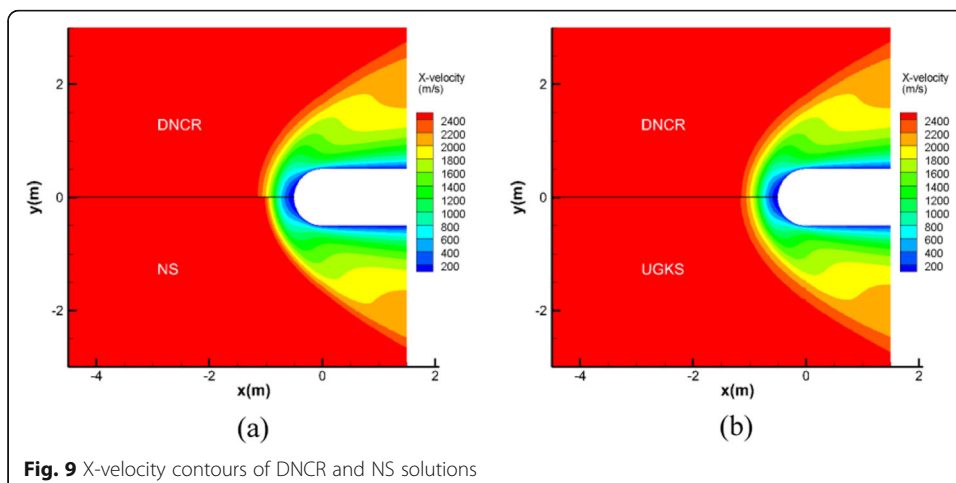
relations always exist in each grid cell physically and the machine learning method is suitable to train and predict it effectively and accurately.

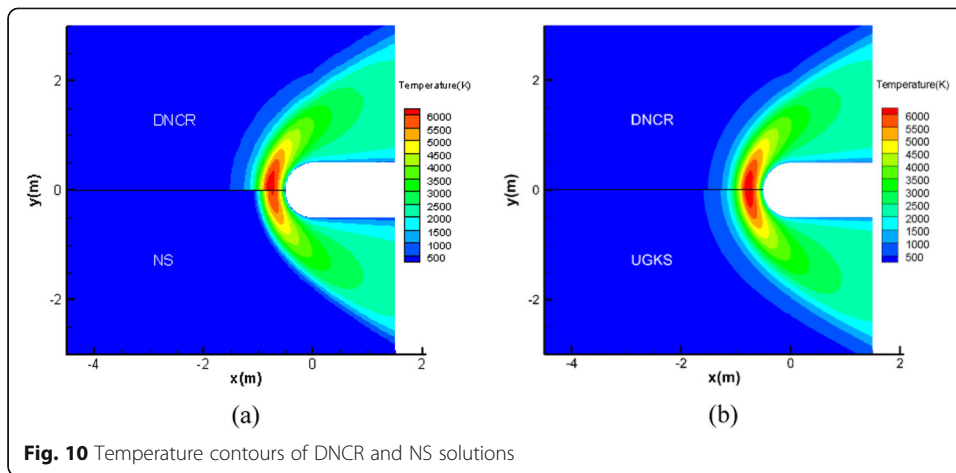
Combined with the final baseline stress tensor, heat flux and flow parameter gradients, solving convention equations is much more efficient than UGKS or DSMC solvers because the total computational time is on the same order of magnitude as NS equations solver. The goal of combining the accuracy of microscale description and the efficiency of macroscale equations could be achieved, especially for hypersonic rarefied engineering flows.

3 Numerical simulations

3.1 One-dimensional shock structure

In this section, the monatomic gas shock structure with non-equilibrium effects limited to the translational energy mode is presented. Shock structure computation mainly focuses on the accuracy of the non-equilibrium model without considering complex wall boundary conditions first. The viscosity of Argon is calculated by using the power law:

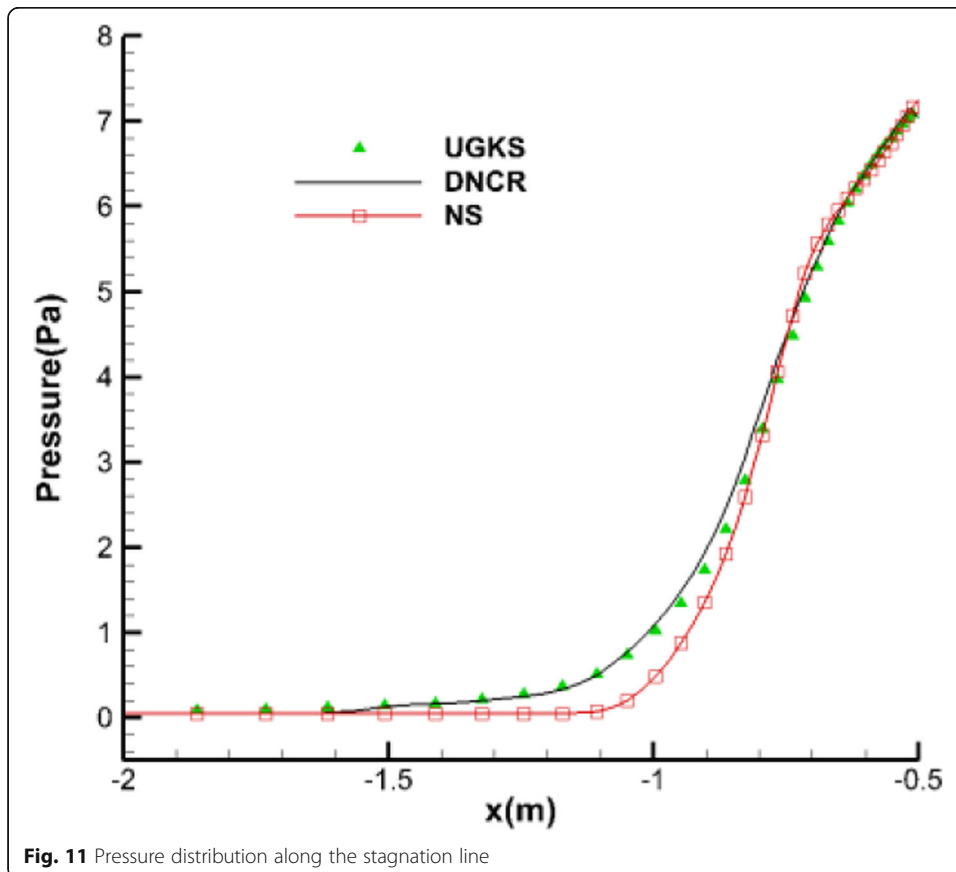




$$\eta = \eta_{\text{ref}} \left(\frac{T}{T_{\text{ref}}} \right)^s \tag{4}$$

where $s = 0.72$ and other physical properties are given in Table 3.

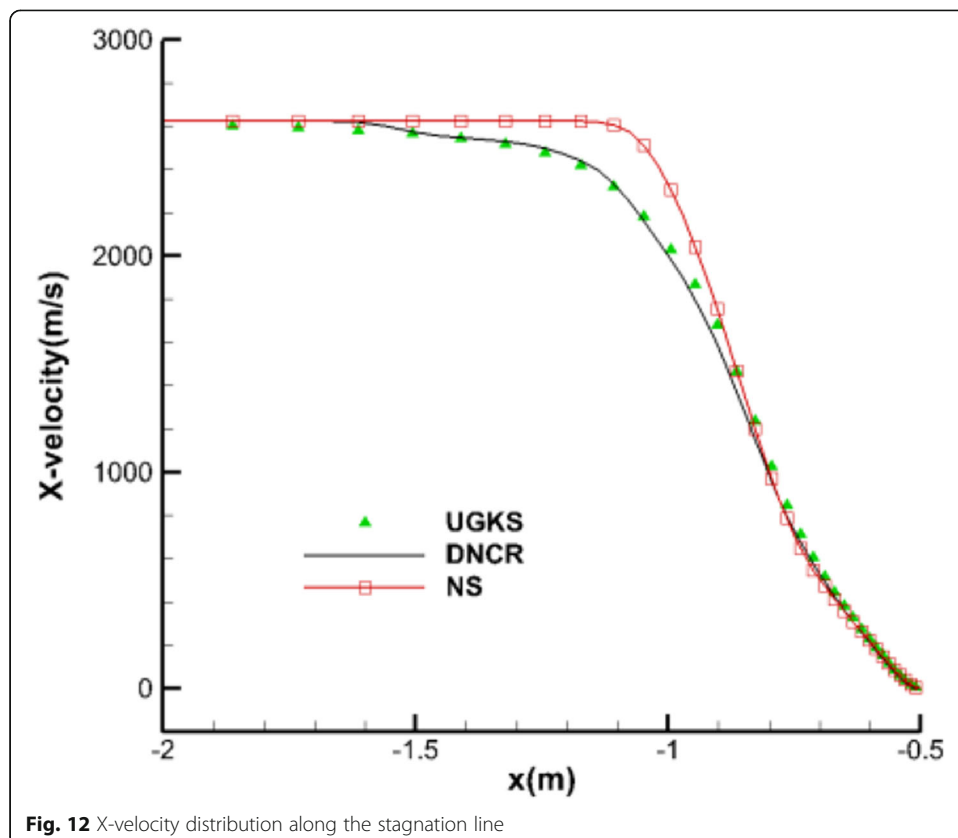
The training set in one-dimensional shock structure is constructed by different shock wave cases with Mach numbers ranging from 1.2 to 10 which includes $Ma = 1.2$, $Ma =$

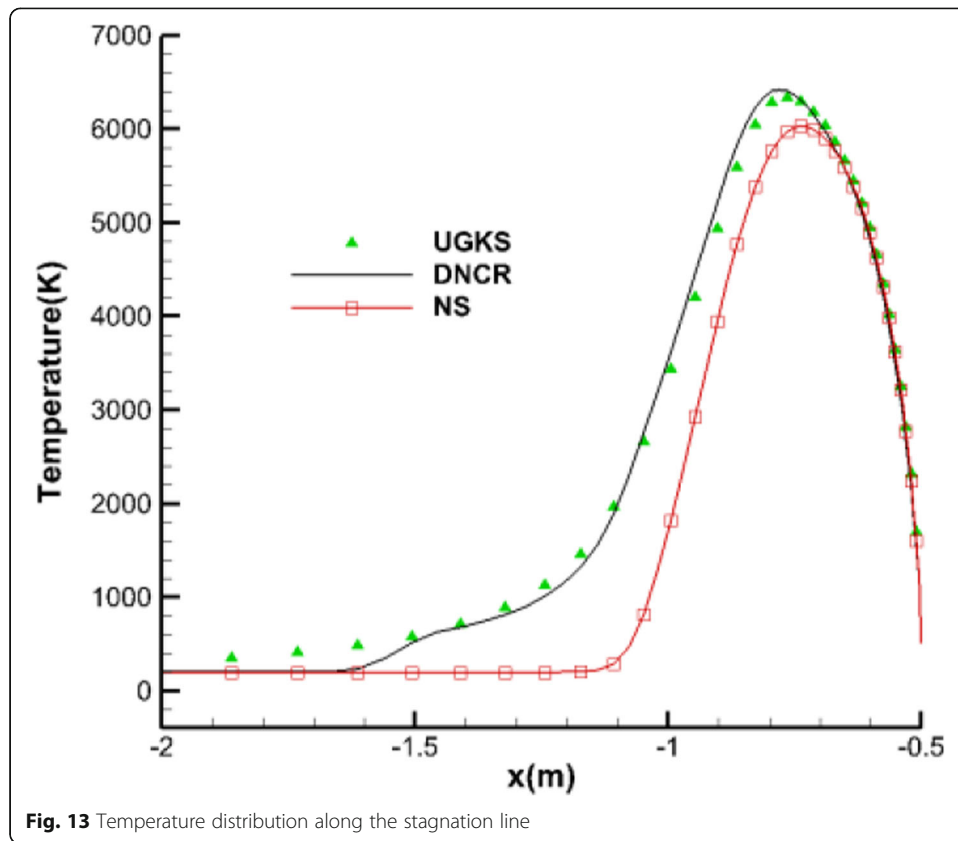


1.5, $Ma = 3$, $Ma = 5$ and $Ma = 7$. Two typical cases with $Ma = 1.6$ and $Ma = 8$ are predicted by DNCR with the trained model in this section.

To compare the discrepancy of NS, UGKS and DNCR results, the distributions of temperature and density inside the shock wave are plotted for Mach 1.6 in Fig. 2, where the x coordinates are nondimensionalized by the upstream molecular mean free path. Due to the low Mach number condition, the gradient of flow parameter inside the shock is not too steep, which results that the local distribution function is not far from equilibrium state. In spite of that, the DNCR results are still in good agreement with that of UGKS, especially high-order moments such as stress tensor and heat flux as shown in Fig. 3. It is also demonstrated that the local stress tensor and heat flux in DNCR solver could converge to the values of UGKS as Eq. (3) when the steady state is reached.

When the Mach number grows to 8, the gap between NS and UGKS results becomes enormous as shown in Figs. 4 and 5. The non-equilibrium characteristic inside the strong shock wave makes Navier-Stokes-Fourier constitutive relations no longer validate. To overcome this defect, various numerical algorithms from the Boltzmann equation or statistical method were proposed in literature to describe the flow structure inside the shock. UGKS is one of these successful implementations and its accuracy has been demonstrated in one-dimensional shock wave computations. As shown in Figs. 4 and 5, the DNCR method yields comparable results with that of UGKS. The temperature rising point predicted by DNCR and UGKS is about 4 times earlier than that of NS, which indicates a thicker and weaker shock compression process in reality.





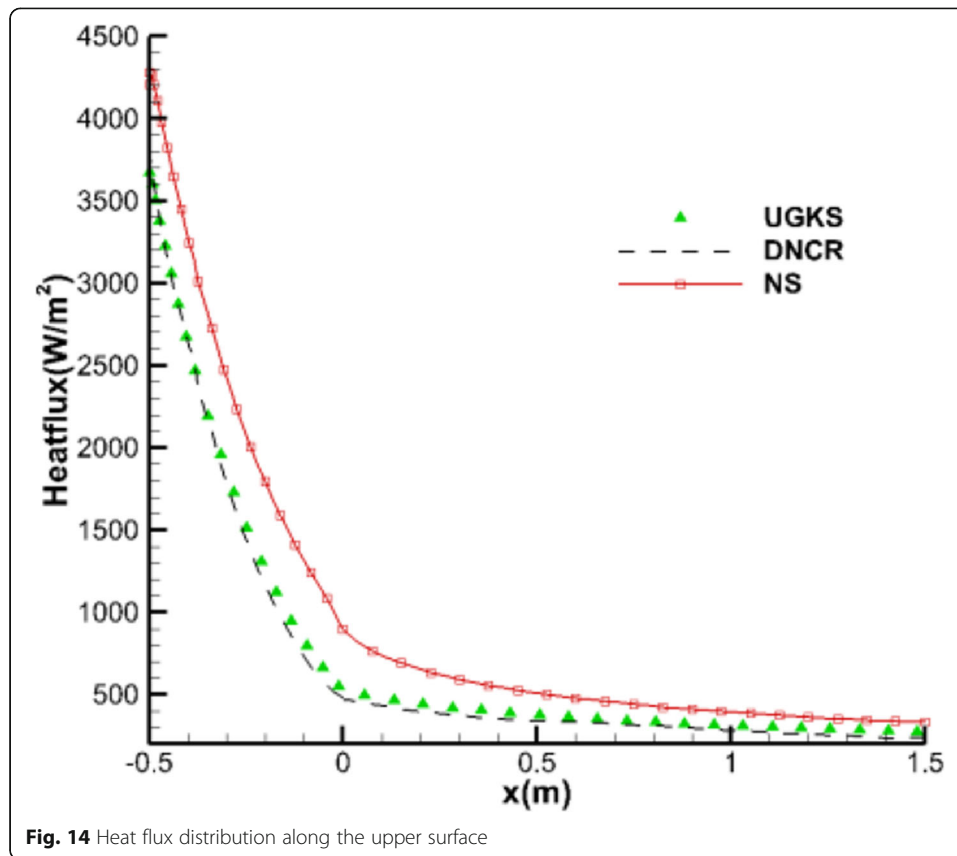
To assess the applicability of DNCR for different Mach number conditions, Fig. 6 presents the comparison of inverse density thickness of shock wave with experimental data using different numerical methods. Both UGKS and DNCR results are in good agreement with most experimental data with Mach number ranging from 1.2 to 10.

3.2 Hypersonic flow around a blunt circular cylinder

The hypersonic flow of Argon passing through a blunt cylinder at $Ma = 10$ is considered in this section. In this case, the radius of the cylinder is 0.5 m and the computational domain is divided with 100×120 (normal direction \times circular direction) structured cells as Fig. 7. In UGKS computations, the velocity space is discretized with rectangle cells of 200×200 . The detailed inflow conditions to be predicted are listed below.

$$\begin{aligned}
 T_\infty &= 198.439\text{K}, & p_\infty &= 0.04655\text{Pa}, & Ma_\infty &= 10 \\
 \gamma &= 1.667, & Pr &= 0.667, & R &= 208.16\text{J}/(\text{kg} \cdot \text{K}) \\
 T_w &= 500\text{K}, & \mu_0 &= 5.0712 \times 10^{-5}\text{Pa} \cdot \text{s}, & T_0 &= 1000\text{K}
 \end{aligned}
 \tag{5}$$

And the viscosity coefficient in all calculations is obtained through the inverse power law with a coefficient of 0.734. The freestream Knudsen number of the testing case is 0.085 with a characteristic length of 1.0 m. The training data set includes two cases with different inflow pressures but same configuration, where $p_\infty = 0.0582\text{Pa}$, $Kn_\infty = 0.0678$ and $p_\infty = 0.0388\text{Pa}$, $Kn_\infty = 0.1016$. Both NS and UGKS calculations are carried



out to obtain all these features and output responses in order to train the nonlinear constitutive relations model. Finally, the DNCR solver calculates the case with conditions as Eq. (5) and the comparisons are presented here to assess the improvements of the DNCR method.

Flow parameter contours such as pressure, temperature and velocity in x -direction are presented in Figs. 8, 9 and 10. It is obvious that DNCR yields a very similar flow field as UGKS does under these inflow conditions. All rarefied characteristics such as the thicker shock wave and the weaker compression behind the bow shock are captured accurately by the DNCR method. By checking the inflow Knudsen number with a characteristic length of 1.0 m, which is the diameter of the circle, the NS solution will be no longer validated in this case because $Kn_\infty = 0.085$. This could also be observed clearly in the figures below where the NS solutions are different from the other two approaches.

To investigate the flow field more carefully, the distribution lines of pressure, temperature and x -velocity along the stagnation line are plotted in Figs. 11, 12 and 13. These lines indicate that the results of DNCR are in good agreement with that of DSMC and the discrepancies between NS and DNCR exist from the bow shock to the solid surface. In order to quantify the improvement degree of the proposed DNCR method, a parameter ξ would be introduced as Eq. (6).

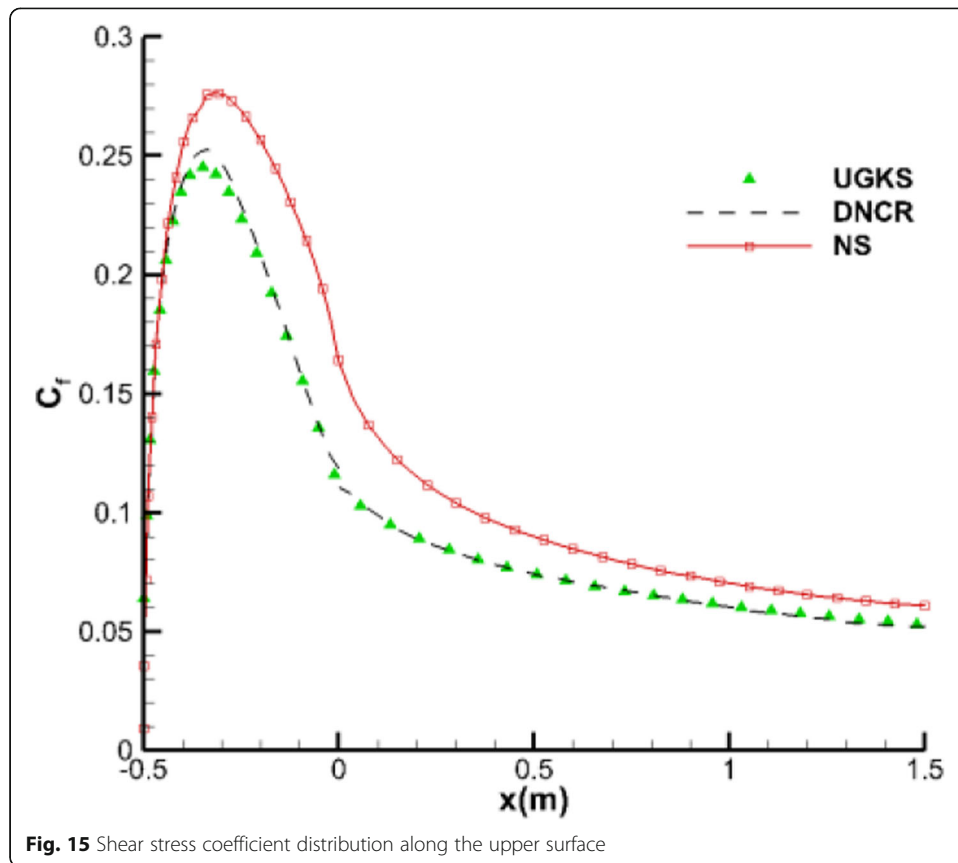


Fig. 15 Shear stress coefficient distribution along the upper surface

$$\xi = 1 - \frac{\sqrt{\sum_{i=1}^N (y_{DNCR}^i - y_{UGKS}^i)^2}}{\sqrt{\sum_{i=1}^N (y_{NS}^i - y_{UGKS}^i)^2}} \tag{6}$$

where y denotes a flow parameter and the superscript i means the index of the grid cell. As we can see from the definition, the formulation $1 - \xi$ indicates the ratio of the discrepancy between DNCR and UGKS to the discrepancy between NS and UGKS. Therefore, the number of ξ ranges from 0 to 1 which indicates that the DNCR results converge to NS or UGKS solutions respectively. In this case, the ξ of pressure, temperature and x-velocity are 0.4908, 0.7680 and 0.6614 respectively and the improvement of temperature is more profound than the other two parameters.

Besides physical parameters in the flow field, the heat flux and shear stress distributions on the wall are also investigated in the test case to validate our approach. Due to the symmetry, only the shear stress coefficient and heat flux on the upper surface are plotted in Figs. 14 and 15.

Table 4 CPU time and memory consumption

	UGKS	NS	DNCR
Time	17 h	11 min	13 min
Memory	54 GB	20 MB	20 MB

The results of heat flux and shear stress on the wall yielded by the DNCR model are still in good agreement with that of UGKS, which indicates that the DNCR method has a promising capability in engineering application to predict aerodynamic heating and drag for hypersonic vehicles.

To check the computational efficiency of different numerical methods, the time and memory consumption is listed in Table 4 as below.

The computational time in Table 2 denotes the time to get the converged steady state solution. It is worth noting that the training time of DNCR is not included in the above computational time because the final aim of any machine learning algorithm is to establish an off-line trained model which could be utilized in different configurations or inflow conditions. Therefore, it is not convenient or necessary to account for the whole training time in a single computational case to be predicted. For the implicit UGKS computation, the computational time is about 17 h and for DNCR is around 13 min. The memory cost of UGKS is around 54 GB and for DNCR it is about 20 MB. Both simulation time and memory consumption of the DNCR solver are comparable to NS because the framework of both methods is very similar. To compare the computational efficiency of DNCR and UGKS, the time consumption of DNCR is about 78 times faster than the implicit UGKS method, and memory cost is 2760 times less than UGKS.

4 Conclusions

This paper describes the development of a data-driven non-linear constitutive relation model for multiscale flow simulation based on the extremely randomized trees model. Based on the training set constructed by the NS and UGKS data, an off-line regression model between input features and output responses is established first. In the prediction phase, NS simulation is carried out to compute the feature vectors, which could be utilized to yield baseline stress tensor, heat flux and flow parameter gradients by querying the trained regress functions. With this baseline flow field parameters, the macroscopic conservation laws are solved implicitly to obtain the final results of the test flow. 1-D shock wave structure and 2-D hypersonic flow around a blunt circle cylinder are investigated by the DNCR method to assess the capability, accuracy and efficiency. The simulation results of DNCR are in good agreement with that of UGKS in all test cases when the distribution of flow parameters and high-order moments are compared. And the computational efficiency of DNCR is much higher than UGKS but with less memory consumption. As a result, the developed DNCR method is proved to be a promising method and expected to pave a new way to predict complex three dimensional hypersonic rarefied gas flows in future.

Acknowledgements

The authors would like to thank Dr. Liu Chang for his constructive discussion in UGKS computations.

Authors' contributions

Wenwen Zhao: Conceptualization, Software, Writing-original draft. Lijian Jiang: Validation, Writing-review & editing. Shaobo Yao: Validation, Writing-review & editing. Weifang Chen: Writing-review & editing. All authors read and approved the final manuscript.

Funding

This research was funded by the National Numerical Wind-Tunnel Project (NO. NNW2019ZT3-A08). The authors would also like to thank the support of the National Natural Science Foundation of China (Grant No. 12002306 and 6162790014).

Availability of data and materials

All data and materials are available from the authors of this paper.

Declaration

Competing interests

The authors declare that they have no competing interests.

Received: 20 July 2021 Accepted: 23 August 2021

Published online: 05 October 2021

References

1. Bird GA (1994) *Molecular gas dynamics and the direct simulation of gas flows*. Clarendon Press, Oxford
2. Bhatnagar PL, Gross EP, Krook M (1954) A model for collision processes in gases. I Small amplitude processes in charged and neutral one-component systems. *Phys Rev* 94(3):511–525. <https://doi.org/10.1103/PhysRev.94.511>
3. Shakhov EM (1972) Generalization of the Krook kinetic relaxation equation. *Fluid Dynam* 3(5):95–96. <https://doi.org/10.1007/BF01029546>
4. Holway LH (1966) New statistical models for kinetic theory: methods of construction. *Phys Fluids* 9(9):1658. <https://doi.org/10.1063/1.1761920>
5. Li Z, Zhang H (2003) Numerical investigation from rarefied flow to continuum by solving the Boltzmann model equation. *Int J Numer Meth Fl* 42(4):361–382. <https://doi.org/10.1002/flid.517>
6. Xu K, Huang J (2010) A unified gas-kinetic scheme for continuum and rarefied flows. *J Comput Phys* 229(20):7747–7764. <https://doi.org/10.1016/j.jcp.2010.06.032>
7. Burnett D (1936) The distribution of molecular velocities and the mean motion in a non-uniform gas. *P Lond Math Soc* s2-40(1):382–435. <https://doi.org/10.1112/plms/s2-40.1.382>
8. Zhao W, Chen W, Agarwal RK (2014) Formulation of a new set of simplified conventional Burnett equations for computation of rarefied hypersonic flows. *Aerosp Sci Technol* 38:64–75. <https://doi.org/10.1016/j.ast.2014.07.014>
9. Grad H (1949) On the kinetic theory of rarefied gases. *Commun Pure Appl Math* 2(4):331–407. <https://doi.org/10.1002/cpa.3160020403>
10. Struchtrup H, Torrilhon M (2003) Regularization of Grad's 13 moment equations: derivation and linear analysis. *Phys Fluids* 15(9):2668–2680. <https://doi.org/10.1063/1.1597472>
11. Gu X, Emerson DR (2009) A high-order moment approach for capturing non-equilibrium phenomena in the transition regime. *J Fluid Mech* 636:177–216. <https://doi.org/10.1017/S002211200900768X>
12. Eu BC (1992) *Kinetic theory and irreversible thermodynamics* (Vol. 93). Wiley, New York
13. Myong RS (1999) Thermodynamically consistent hydrodynamic computational models for high-Knudsen-number gas flows. *Phys Fluids* 11(9):2788–2802. <https://doi.org/10.1063/1.870137>
14. Myong RS (2004) A generalized hydrodynamic computational model for rarefied and microscale diatomic gas flows. *J Comput Phys* 195(2):655–676. <https://doi.org/10.1016/j.jcp.2003.10.015>
15. Jiang Z, Zhao W, Chen W, Agarwal RK (2019) An undecomposed hybrid algorithm for nonlinear coupled constitutive relations of rarefied gas dynamics. *Commun Comput Phys* 26(3):880–912. <https://doi.org/10.4208/cicp.OA-2018-0056>
16. Sun QH, Boyd ID, Candler GV (2004) A hybrid continuum/particle approach for modeling subsonic, rarefied gas flows. *J Comput Phys* 194(1):256–277. <https://doi.org/10.1016/j.jcp.2003.09.005>
17. Liu C, Zhu Y, Xu K (2018) Unified gas-kinetic wave-particle methods I: continuum and rarefied gas flow. *J Comput Phys* 401:108977. <https://doi.org/10.1016/j.jcp.2019.108977>
18. Su W, Zhu L, Wang P, Zhang Y, Wu L (2020) Can we find steady-state solutions to multiscale rarefied gas flows within dozens of iterations? *J Comput Phys* 407:109245. <https://doi.org/10.1016/j.jcp.2020.109245>
19. Fei F, Zhang J, Li J, Liu Z (2020) A unified stochastic particle Bhatnagar-Gross-Krook method for multiscale gas flows. *J Comput Phys* 400:108972. <https://doi.org/10.1016/j.jcp.2019.108972>
20. Liu S, Zhong C, Fang M (2020) Simplified unified wave-particle method with quantified model-competition mechanism for numerical calculation of multiscale flows. *Phys Rev* 102(1):13304. <https://doi.org/10.1103/PhysRevE.102.013304>
21. Duraisamy K, Iaccarino G, Xiao H (2019) Turbulence modeling in the age of data. *Annu Rev Fluid Mech* 51(1):357–377. <https://doi.org/10.1146/annurev-fluid-010518-040547>
22. Xiao H, Wu JL, Wang JX, Sun R, Roy CJ (2016) Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier–Stokes simulations: a data-driven, physics-informed Bayesian approach. *J Comput Phys* 324:115–136. <https://doi.org/10.1016/j.jcp.2016.07.038>
23. Wang J, Wu J, Xiao H (2017) A physics informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Phys Rev Fluids* 2(3):34603. <https://doi.org/10.1103/PhysRevFluids.2.034603>
24. Ling J, Kurzawski A, Templeton J (2016) Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *J Fluid Mech* 807:155–166. <https://doi.org/10.1017/jfm.2016.615>
25. Pope SB (1975) A more general effective-viscosity hypothesis. *J Fluid Mech* 72(02):331. <https://doi.org/10.1017/S0022112075003382>
26. Kutz JN (2017) Deep learning in fluid dynamics. *J Fluid Mech* 814:1–4. <https://doi.org/10.1017/jfm.2016.803>
27. Sekar V, Khoo BC (2019) Fast flow field prediction over airfoils using deep learning approach. *Phys Fluids* 31(5):57103. <https://doi.org/10.1063/1.5094943>
28. Bar-Sinai Y, Hoyer S, Hickey J, Brenner MP (2019) Learning data-driven discretizations for partial differential equations. *Proc Natl Acad Sci* 116(31):15344–15349. <https://doi.org/10.1073/pnas.1814058116>
29. Mao Z, Jagtap AD, Karniadakis GE (2020) Physics-informed neural networks for high-speed flows. *Comput Method Appl M* 360:112789. <https://doi.org/10.1016/j.cma.2019.112789>
30. Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378:686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>

31. Raissi M, Yazdani A, Karniadakis GE (2020) Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* 367(6481):1026–1030. <https://doi.org/10.1126/science.aaw4741>
32. Zhang J, Ma W (2020) Data-driven discovery of governing equations for fluid dynamics based on molecular simulation. *J Fluid Mech* 892:A5. <https://doi.org/10.1017/jfm.2020.184>
33. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63(1):3–42. <https://doi.org/10.1007/s10994-006-6226-1>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
