Journal of Computational Physics 374 (2018) 1166–1179

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Application of an evolutionary algorithm to LES modelling of turbulent transport in premixed flames

Matthias Schoepplein [a,b], Jack Weatheritt [a], Richard Sandberg [a], Mohsen Talei [a], Markus Klein [b,*]

[a] Department of Mechanical Engineering, University of Melbourne, Parkville, VIC, 3010, Australia
[b] Department of Aerospace Engineering, Universitaet der Bundeswehr Muenchen, Neubiberg, 85577, Germany

ARTICLE INFO

Article history:
Received 14 March 2018
Received in revised form 19 June 2018
Accepted 6 August 2018
Available online 14 August 2018

Keywords:
Evolutionary algorithm
Gene expression programming
Large eddy simulation
Turbulent premixed flame
Subgrid scale stress tensor

ABSTRACT

Recently the concept of Gene Expression Programming (GEP) has been introduced with very encouraging results for the purpose of modelling the unclosed tensors in the context of RANS (Reynolds Averaged Navier–Stokes) turbulence modelling. This paper extends the previous framework to modelling subgrid stresses (SGS) in the context of Large Eddy Simulation (LES). In order to achieve this goal the GEP algorithm was coupled with an external multiprocessor postprocessing tool that evaluates a cost function based on *a-priori* analysis of explicitly filtered DNS data of turbulent premixed planar flames. Although LES of combustion systems is becoming increasingly popular, the closures for sub-grid scale (SGS) stresses have mostly been derived assuming constant density flows. However, it has been shown recently that depending on the relative strength of heat release and turbulence, counter-gradient transport can occur for the stress tensor if the isotropic part is not properly accounted for. The focus of this work is not to put forward a particular new model but to demonstrate that evolutionary algorithms can successfully be used in the framework of LES modelling. To achieve this purpose the GEP software is used for modelling the deviatoric stress, the trace of the SGS tensor and the stress tensor itself. Although the functional form of the model was not imposed, the evolutionary algorithm did find a well known model from the literature with even the model constants comparable to values reported in the literature.

1. Introduction

Computational Fluid Dynamics (CFD) enables optimisation of geometry and flow processes, it helps understanding physical phenomena where experimental access is limited and it has the potential to save cost and reduce the time to market of a new product. The availability of high performance computing has given rise to very impressive unsteady simulations in particular Large Eddy Simulations (LES) with high temporal and spatial resolution and including multi-physics and detailed chemical reactions. Turbulent combustion modelling deals with important non-linear, multiscale phenomena, i.e. turbulence, combustion and their interaction. Despite the impressive progress made in this field there are some fundamental questions unsolved and models are not always optimised for a particular application. As an example models developed in the context

* Corresponding author.
E-mail address: markus.klein@unibw.de (M. Klein).

https://doi.org/10.1016/j.jcp.2018.08.016
0021-9991/

of Reynolds Averaged Navier Stokes modelling (RANS) are often used in the context of LES with small adjustments. Similarly, models developed for the closure of constant density turbulent flows have been used for simulating turbulent combustion. In the context of a turbulent premixed flame a particular modelling challenge arises from the fact that the LES subgrid scale (SGS) models should also be able to capture the gas-dynamic expansion [1,2]. Assuming that chemical reactions are simplified by reducing the detailed chemical mechanism to the transport of a single reaction progress variable c (which is zero in the fresh gases and unity in the fully burned products), the sub-grid turbulent scalar flux $F_i^{sgs} = \overline{\rho u_i c} - \overline{\rho}\tilde{u}_i\tilde{c}$ has to be modelled in the context of LES. Here, u_i is one component of the velocity vector, ρ is the density, the overbar denotes conventional filtering and Favre filtering is defined as $\tilde{\varphi} = \overline{\rho\varphi}/\overline{\rho}$. It is well known for a long time [3,4] and has been demonstrated several times in the context of LES (see [5–9] and also the literature review therein) that the turbulent scalar flux can show counter-gradient transport (CGT) behaviour. This means that the conventional gradient hypothesis closure predicts a flux that has the opposite direction of the real turbulent flux. CGT typically occurs if the effects of heat release are stronger than the effects of turbulence and CGT becomes, in the context of LES, more pronounced for large filter sizes. The phenomenon of CGT also affects the SGS stress tensor expressed as

$$\tau_{ij}^{SGS} = \overline{\rho u_i u_j} - \overline{\rho}\tilde{u}_i\tilde{u}_j, \tag{1}$$

appearing in the density weighted momentum conservation equation (ν is the kinematic viscosity and p denotes pressure):

$$\frac{\partial\overline{\rho}\tilde{u}_i}{\partial t} + \frac{\partial}{\partial x_j}\left(\overline{\rho}\tilde{u}_i\tilde{u}_j\right) = -\frac{\partial}{\partial x_j}\left(\overline{\rho u_i u_j} - \overline{\rho}\tilde{u}_i\tilde{u_j}\right) + \frac{\partial}{\partial x_j}\left(\overline{\rho}\tilde{\nu}\left(\frac{\partial\tilde{u}_i}{\partial x_j} + \frac{\partial\tilde{u}_j}{\partial x_i}\right) - \frac{2}{3}\overline{\rho}\tilde{\nu}\frac{\partial\tilde{u}_k}{\partial x_k}\delta_{ij}\right) - \frac{\partial\overline{p}}{\partial x_i}. \tag{2}$$

Under the commonly used assumption that the isotropic part of the SGS stresses, i.e. the term involving $-\frac{1}{3}\tau_{kk}^{SGS}\delta_{ij}$, can be added to the filtered pressure, the static Smagorinsky model (SSM) is given by

$$\begin{aligned}
\tau_{ij}^{SSM} &= \tau_{ij}^{SGS} - \frac{1}{3}\tau_{kk}^{SGS}\delta_{ij} \approx -\bar{\rho}\nu_t 2\left(\tilde{S}_{ij} - \frac{1}{3}\tilde{S}_{kk}\delta_{ij}\right)\\
\nu_t &= (C_s\Delta)^2\left|\tilde{S}_{ij}\right|\\
\tilde{S}_{ij} &= \frac{1}{2}\left(\frac{\partial\tilde{u}_i}{\partial x_j} + \frac{\partial\tilde{u}_j}{\partial x_i}\right), \left|\tilde{S}_{ij}\right| = \sqrt{2\tilde{S}_{ij}\tilde{S}_{ij}},
\end{aligned} \tag{3}$$

where C_s is assumed to be constant. It was shown in [10] that the cosine of the angle between τ_{1j}^{SSM} and τ_{1j}^{SGS} under certain circumstances assumes values of -1 in the flame brush indicating that the vector points in the direction opposite to τ_{1j}^{SGS}. In other words, the Boussinesq assumption does not correctly predict the SGS stresses in turbulent premixed flames. Indeed Pfadler et al. [2] demonstrated the unsatisfactory performance based on direct measurements of the density weighted stress tensor. It has been pointed out in [10] that a reliable model for the isotropic part of the stresses τ_{kk}^{SGS} helps to solve this problem to some extent and a variety of such models has been analysed in [11]. Despite the fact that research has been carried out for several decades, only limited effort has been directed to the assessment of sub-grid scale (SGS) stress tensor closures in turbulent premixed combustion [2]. This can be partially explained by the facts that at least the resolved part of CGT (inherently playing a role in turbulent premixed flames) will be accounted for in an LES [12] and that explicit modelling of counter gradient SGS transport might induce numerical instabilities. The focus of this work is the application of a machine-learning approach to the problem outlined above.

Machine-learning algorithms are becoming a popular tool for improving turbulence models. These are essentially sophisticated curve-fitting algorithms, that find both coefficients and a non-linear form for closure terms in the transport equations. Typically, the algorithms have been applied to closing the RANS equations [13–15], although the scalar flux [16] and the hybrid RANS/LES equations [17] have also been tackled. The premise in all of these studies is to infer the closure term from a high-fidelity database, by fitting a model to the true value. This is by no means a new concept in turbulence modelling (e.g. [18,19]), however increasing size and quality of the datasets has allowed for the use of more complex algorithms that infer, not just closure constants, but functional forms also. The aforementioned studies can be split into two categories: those that are transparent and those that are not. Essentially, this dictates whether the non-linearity of the machine-learnt model exists on a level of description interpretable by a human. For methodologies, such as random forests and neural networks, the non-linearity is built over hundreds, if not thousands of interactions. The existence of complexity at the lowest level forces the reader to treat the model as a black box (i.e. it is not transparent), one where they cannot diagnose problems or easily communicate to the wider community. For symbolic regression the model inferred is a mathematical expression. The latter approach is taken in this study as the model can be interpreted and easily implemented in existing LES solvers.

The structure of this paper is as follows: First the DNS database used for this analysis will be introduced. This is followed by a detailed description of the methodology. Finally results will be discussed and summarised in the conclusions.

Table 1
List of initial simulation parameters and non-dimensional numbers.

Case	*A*	*B*
u'/S_L	7.5	15
l/δ_{th}	4.58	4.58
Re_t	87.50	175
Da	0.61	0.31
Ka	9.60	27.16

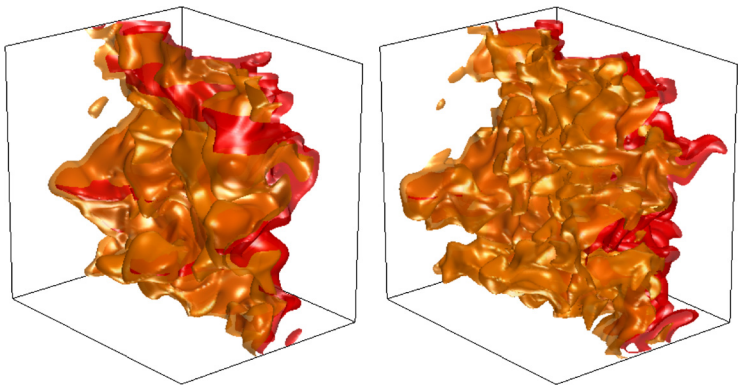

Fig. 1. Instantaneous view of c-isosurfaces for cases A and B. The value of c increases from 0.1 to 0.9 from gold to red. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

2. DNS data base

In the present analysis, the machine learning algorithms will be applied to modelling the SGS stresses based on explicitly filtered simple chemistry Direct Numerical Simulation (DNS) data of freely propagating statistically planar turbulent premixed flames. The well-known DNS code SENGA [20] has been used for the simulations carried out in this analysis. Two flames, cases *A* and *B*, are considered for this work which are part of a much bigger database, similar to those databases used in the past [8,11,21], but features a higher turbulence level and larger scale separation. Instantaneous views of isosurfaces for cases A and B when the statistics were extracted are shown in Fig. 1. The initial values of normalised root mean square turbulent velocity fluctuation u'/S_L, the ratio of turbulent integral length scale to flame thickness l/δ_{th}, turbulent Reynolds number $Re_t = u'l/\nu$, Damkoehler number $Da = lS_l/(\delta_{th}u')$ and Karlovitz number $Ka = (u'/S_L)^{3/2}(l/\delta_{th})^{-1/2}$ are provided in Table 1 where δ_{th} is the thermal flame thickness given by $\delta_{th} = (T_{ad} - T_0)/max|\nabla T|_L$. Here T_{ad} and T_0 are the adiabatic flame and unburnt gas temperature respectively and the subscript $_L$ refers to the unstrained laminar planar flame quantities.

Standard values are chosen for Prandtl number *Pr* and ratio of specific heats γ (i.e. $Pr = 0.7$ and $\gamma = 1.4$). The flame Mach number $Ma = S_L/\sqrt{\gamma RT_0}$, heat release parameter $\tau = (T_{ad} - T_0)/T_0$ and Lewis number *Le* are taken to be 0.014, 4.5 and 1.0 respectively. The DNS database has been explicitly filtered using a Gaussian filter kernel with three different filter widths from $\Delta_f \approx 0.4\delta_{th}$ where the flame is almost resolved, up to $\Delta_f \approx 2.8\delta_{th}$ where the flame becomes fully unresolved and Δ_f is comparable to the integral length scale. It is worth noting that the turbulence is decaying in time and statistics are extracted after two eddy turnover times which is larger than one chemical time scale in both cases. It is standard practice for DNS of combustion under decaying turbulence to continue the simulation time until $t_{sim} \geq max(t_f, t_c)$, where $t_f = l/u'$ is the initial eddy turn over time and $t_c = \delta_{th}/S_L$ is the chemical time scale. This ensures that the final result is independent of initialisation of chemistry and turbulent flow field. The present simulation time is either comparable to or greater than that used in several previous DNS studies by several other authors that contributed significantly to the fundamental understanding and modelling of turbulent combustion. It has also been demonstrated in the past that characteristic results remain qualitatively similar halfway through the simulation. For a detailed discussion in this regard we refer to [10] and references therein.

It is important to note that combustion DNS is inherently limited in terms of turbulent Reynolds number, Damkoehler number and Karlovitz number. These effects have been explicitly studied in the DNS of Chakraborty et al. [21] where it is reported that the dependence of displacement speed on strain rate and curvature is found to weaken with increasing turbulent Reynolds number when either Damkoehler or Karlovitz number is held constant, but the qualitative nature of the correlation remains unaltered. The effects of scale separation on the modelling of generalised SGS turbulent kinetic energy have been discussed in the Appendix of reference [11] and it has been found that the findings of that paper remain valid for a different database which deals with considerably higher values of Re_t and scale separation l/δ_{th}.

Advantages and disadvantages of this planar flame configuration have been discussed in detail elsewhere [22]. Finally, it is worth stressing that the physical mechanisms responsible for gradient or counter gradient stresses (or transport) will not be different even if one uses complex instead of single step chemistry. It has been demonstrated in the past that flame propagation statistics obtained from detailed chemistry simulations [23,24] can be well captured using simple chemistry [25, 26]. Furthermore, models which have been proposed based on the analysis of simple chemistry DNS of turbulent premixed flames with simplified transport have the potential to be valid even in the presence of detailed chemistry and transport (subject to minor adjustments) [27,28]. More references for the qualitative similarity of displacement speed and scalar gradient statistics between simple and detailed chemistry DNS results are provided in [29]. Although the above discussion suggests that the assumption of a single step global reaction rate, as well as limitations in terms of $Re_t, Da, Ka, l/\delta_{th}$ and Δ/δ_{th} do not affect the qualitative nature of the results presented in this work, future analyses will be necessary for validation of the current findings.

The focus of this work is not to put forward a particular new model but to demonstrate that evolutionary algorithms can successfully be used in the framework of LES modelling. Hence, our results are independent of the details of the configuration and also the simplification of chemistry does not affect the physical mechanisms for the study under consideration.

For all cases the simulation domain size was taken to be $45.7\delta_{th} \times 45.7\delta_{th} \times 45.7\delta_{th}$, which was discretised using a uniform Cartesian mesh of 512^3 grid points. High order central finite-differences (10^{th} order for the internal grid points dropping to a one-sided 2^{nd} order scheme at non-periodic boundaries) and a 3^{rd} order low storage Runge–Kutta scheme are used for spatial differentiation and explicit time advancement respectively. The boundary conditions in the mean flame propagation direction (aligned with negative x_1-direction) are taken to be partially non-reflecting, whereas boundaries in transverse directions are taken to be periodic. The turbulent velocity fluctuations are initialised using a homogeneous isotropic incompressible velocity field. The reacting flow field is initialised by a steady planar unstrained premixed laminar flame solution. Further details of the numerical scheme are provided in [20,21].

3. Methodology

Instead of deriving a model based on mathematical and physical principles combined with engineering intuition, a data-driven algorithm has been used in this work to optimise an expression based on its performance in an *a-priori* testing. That is, a mathematical equation is inferred, from our high-fidelity datasets outlined in Section 2. This process is a generalisation of classical least-squares regression, known as symbolic regression – not only coefficients, but the overall functional form is regressed. This process, known as symbolic regression [30], is therefore capable of building non-linear models that are essentially curve fits of some closure term. In this study, we use a sophisticated symbolic algorithm, known as gene expression programming (GEP) [31], to fit a model for the subgrid-scale stress tensor, using the classical stress-strain hypothesis now introduced.

For such an inference, a set of linear independent basis functions is required to fit the subgrid-scale stress tensor. The fundamental assumption invoked for this work presupposed that the subgrid-scale stress tensor can be expressed as a function depending only on strain and rotation rate tensor as well as the LES filter width Δ_f and filtered density $\bar{\rho}$:

$$\tau_{ij}^{sgs} = \tau_{ij}^{sgs}(\tilde{S}_{ij}, \tilde{\Omega}_{ij}, \Delta_f, \bar{\rho}) \tag{4}$$

with

$$\tilde{S}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right) \quad \tilde{\Omega}_{ij} = \frac{1}{2}\left(\frac{\partial \tilde{u}_i}{\partial x_j} - \frac{\partial \tilde{u}_j}{\partial x_i}\right). \tag{5}$$

Since typical symbolic regression algorithms do not involve restrictions in respect of units, $\tilde{S}$ and $\tilde{\Omega}$ have been non-dimensionalised by a suitable inverse time scale τ, similar to the term used in the Smagorinsky model Eqn. (3), yielding the non-dimensional matrices s and ω:

$$s_{ij} = \frac{\tilde{S}_{ij}}{\sqrt{\tilde{S}_{mn}\tilde{S}_{mn}}} \quad \omega_{ij} = \frac{\tilde{\Omega}_{ij}}{\sqrt{\tilde{S}_{mn}\tilde{S}_{mn}}}. \tag{6}$$

It is important to note that the filtering operation is omitted for the normalised quantities s, ω. Any candidate model τ_{ij}^{mod} found by the GEP algorithm to approximate τ_{ij}^{sgs} has to be re-dimensionalised at the end of the process using again the term $\sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}$ as well as the LES filter width Δ_f and $\bar{\rho}$:

$$\tau_{ij}^{mod} = \bar{\rho}\left(\Delta_f\sqrt{\tilde{S}_{mn}\tilde{S}_{nm}}\right)^2 \cdot f_{ij}\left(s_{ij}, \omega_{ij}\right). \tag{7}$$

The task of the GEP code is to find an optimal expression for the non-dimensional function f_{ij}. The most general expression for f_{ij} would be an infinite tensor polynomial in the two second-order tensors s and ω [32]. However, Spencer et al. [33] found that the number of linearly independent second-order tensor functions (basis functions) that can be formed from

two second-order tensors (in this case s and ω) are limited due to the Caley–Hamilton Theorem as well as their tensor properties. Therefore the infinite tensor polynomial can be written in finite form, that consists of a polynomial of the basis functions with scalar coefficients depending on a limited number of invariants. Using the symmetry properties $\tau_{ij}^{sgs} = \tau_{ji}^{sgs}$, $s_{ij} = s_{ji}$ and $\omega_{ij} = -\omega_{ji}$ the basis functions for the full subgrid scale stress tensor are:

$$\begin{array}{c} \mathbb{I} \\ s \\ s^2 \qquad \omega^2 \qquad s\omega - \omega s \\ s^2\omega - \omega s^2 \qquad \omega^2 s + s\omega^2 \\ s^2\omega^2 + \omega^2 s^2 \qquad s\omega s^2 - s^2\omega s \qquad \omega s\omega^2 - \omega^2 s\omega \\ \omega s^2\omega^2 - \omega^2 s^2\omega \end{array} \tag{8}$$

where $\mathbb{I}$ is the identity matrix. The invariants are given by:

$$\begin{array}{c} tr(s) \\ tr(s^2) \qquad tr(\omega^2) \\ tr(s^3) \qquad tr(s\omega^2) \\ tr(s^2\omega^2) \\ tr(s\omega s^2\omega^2) \end{array} \tag{9}$$

Combustion modelling often involves low Mach number approximations, where typically only the deviatoric part of the SGS stresses is modelled (e.g. Eqn. (3)), whereas its trace is added to the filtered pressure. In this scenario the basis functions and invariants have to be modified because (deviatoric) SGS stress tensor and (deviatoric) strain rate have zero trace ($\tau_{ii}^{sgs} = 0, s_{ii} = 0$). Using these properties the set of functions in Eqn. (8) reduces to:

$$\begin{array}{c} s \\ s^2 - \frac{1}{3}tr(s^2)\mathbb{I} \qquad \omega^2 \\ s\omega - \omega s \\ s^2\omega - \omega s^2 \qquad \omega^2 s + s\omega^2 \\ s\omega s^2 - s^2\omega s \qquad \omega s\omega^2 - \omega^2 s\omega \\ s^2\omega^2 + \omega^2 s^2 - \frac{2}{3}tr(s^2\omega^2)\mathbb{I} \\ \omega s^2\omega^2 - \omega^2 s^2\omega \end{array} \tag{10}$$

The invariants are the same as those obtained for the fully compressible case, except that the trace of the deviatoric part vanishes. A similar basis has been previously used by [13,17] for RANS model development. Since there are only a finite number of linearly independent functions, the polynomial is reduced to a sum of basis functions T^{α} with scalar coefficients G_{α}. These coefficients are functions of invariants and are subject to optimisation by the GEP algorithm:

$$f_{ij} = \sum_{\alpha=1}^{n} G_{\alpha} \cdot T_{ij}^{\alpha}; \quad G_{\alpha} = G_{\alpha}(I_1, I_2, \ldots I_m). \tag{11}$$

Symbolic regression algorithms search the space of all functions, spanned by the basis detailed above. This is a ten dimensional polynomial in space. In order to reduce the computational cost of the search, we reduce the size of the space, by truncating to, at most, the first four basis functions. These are the terms at least quasi-quadratic. It is a fair assumption, especially for subgrid-scale models, that the anisotropy present yields decreasing utility in retaining terms of higher orders. Therefore, our function f_{ij} is spanned by the basis,

$$T_{ij}^1 = s_{ij} \tag{12}$$

$$T_{ij}^2 = s_{ik}\omega_{kj} - \omega_{ik}s_{kj} \tag{13}$$

$$T_{ij}^3 = s_{ik}s_{kj} \tag{14}$$

$$T_{ij}^4 = \omega_{ik}\omega_{kj}. \tag{15}$$

In case of modelling the deviatoric part of the SGS tensor only the first three functions were used and T^3 was replaced by

$$T_{ij}^{3*} = s^2 - \frac{1}{3} tr(s^2)\mathbb{I}\,. \tag{16}$$

Out of the seven possible invariants, only the ones with order lower than 4 were considered:

$$I_1 = s_{mn}s_{nm} \tag{17}$$

$$I_2 = \omega_{mn}\omega_{nm} \tag{18}$$

$$I_3 = s_{km}s_{mn}s_{nk} \tag{19}$$

$$I_4 = \omega_{km}\omega_{mn}s_{nk}\,. \tag{20}$$

The optimisation of functions f_{ij} is driven by an evolutionary analogy. Such functions can be represented as strings which, when read recursively, are decoded into non-linear function [30]. A set of initially random strings are evolved, via an analogy to Darwin's survival-of-the-fittest theory. That is, operators analogous to reproduction and mutation are applied to the population of strings which modify the value of such strings. This in turn modifies the mathematical form (known as the phenotype). Most changes will result in a poorer representation of the true subgrid-scale stress tensor, however a few will improve models. Via a selection operator, bad models are filtered out and the better solutions in the population are returned for the next iteration (known as a generation). Over many generations, the overall quality of the population slowly improves as the randomness introduced by mutations is filtered by selection, in exactly the same way we observe in natural systems.

The quality of a candidate solution is expressed by a scalar value, called the fitness. The fitness value is calculated by averaging a cost function over a certain number of data points of the DNS solution. For general subgrid-scale-stress-tensor training, the points were chosen from the whole DNS domain, while the models trained specifically for the reacting flow were only evaluated within the combustion zone, i.e. $0.1 \leq \tilde{c} \leq 0.9$. Suppose n_p data points were selected, the fitness value was evaluated by

$$fit(\tau_{ij}^{mod}) = \frac{1}{n_p}\sum_{\alpha=1}^{n_p} \frac{\left|\tau_{ij}^{sgs}(\alpha) - \tau_{ij}^{mod}(\alpha)\right|_F}{\left|\tau_{ij}^{sgs}(\alpha)\right|_F}\,. \tag{21}$$

The operator $|\cdot|_F$ is the Frobenius norm defined as

$$\left|a_{ij}\right|_F = \sqrt{\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}^2} \tag{22}$$

and hence represents a mean square root of the errors in every component. Consequently, Eqn. (21) is an indicator for agreement of modelled and real subgrid stress on tensor level. Note that in this formulation the off-diagonal elements will be counted twice because of the symmetry of the subgrid-scale stress tensor.

In case of the turbulent kinetic energy model a scalar cost function based on the relative error was used

$$fit(k^{mod}) = \frac{1}{n_p}\sum_{\alpha=1}^{n_p} = \frac{\left|k^{sgs}(\alpha) - k^{mod}(\alpha)\right|}{\left|k^{sgs}(\alpha)\right|}\,, \tag{23}$$

where the subgrid scale turbulent kinetic energy is defined as the trace of the SGS stress tensor (note that for simplicity the prefactor a half has been omitted)

$$k^{sgs} = \tau_{ii}^{sgs}. \tag{24}$$

For an excellent and concise introduction on the GEP algorithm see the work of Ferreira [31]. For a detailed application to fluid mechanics, and in particular turbulence modelling, see the work of Weatheritt and Sandberg [13].

4. Results

The modelling of the subgrid stress tensor in this section proceeds along the following line. In a first step the deviatoric part of the stress tensor will be considered. It was shown in [10] that improved modelling results can be obtained by explicit modelling of the trace of the stress tensor. This will be the second step of the analysis. Finally, direct modelling of the full stress tensor is considered.

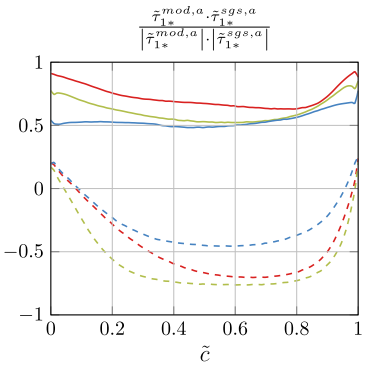

Fig. 2. Alignment between the row vectors of subgrid stress tensor predictions and calculations from DNS conditionally averaged on the Favre filtered reaction progress $\tilde{c}$ for case *A*: *GDS* model (——) and *SSM* model (- - -) for three filter sizes: $n_f = 4$, $n_f = 16$, $n_f = 28$.

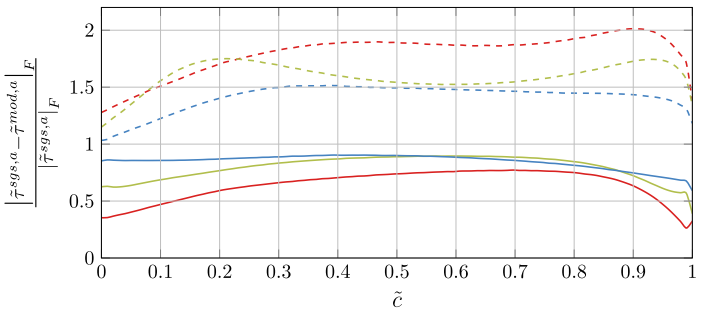

Fig. 3. Fitness of *GDS* model and *SSM* model conditionally averaged on the Favre filtered reaction progress $\tilde{c}$ for case *A* (same legend as Fig. 2).

4.1. Deviatoric SGS Stress Tensor Model

Following the methodology outlined above GEP optimises for a model which after re-dimensionalisation has the form

$$\tilde{\tau}_{ij}^{mod} = \bar{\rho}\Delta_f^2/\tau^2 \cdot \left(G_1 \cdot T_{ij}^1 + G_2 \cdot T_{ij}^2 + G_3 \cdot T_{ij}^{3*}\right). \tag{25}$$

Note that the strain rate tensor T_{ij}^{3*} is used without its trace to comply with tensor theory. Each model training has been conducted for 500 generations with a population size of 75. Due to the non-deterministic nature of the GEP tool, the training process was performed 50 times, generating 50 different models. The functional form and the performance of the 50 models differs and one can either calculate an ensemble averaged model from all the trained models or, as in this work, simply select the best model which according to the fitness function Eqn. (21) was

$$\begin{aligned} \tilde{\tau}_{ij}^{GDS} &= \bar{\rho}\Delta_f^2 \tilde{S}_{mn}\tilde{S}_{nm} \cdot C \cdot \left(T_{ij}^{3*} - T_{ij}^2\right) \\ C &= 8.71 \times 10^{-2} \end{aligned} \tag{26}$$

and will henceforth be denoted *GDS* (i.e. GEP based deviatoric stress model) as the deviatoric part of the stresses has been considered.

Fig. 2 compares the conditionally averaged alignment between row vectors of the modelled stress and the real subgrid stress for the trained model Eqn. (26) and the static Smagorinsky model Eqn. (3) for three different filter sizes, where n_f denotes the filter size relative to the DNS grid spacing, i.e. $n_f = \Delta_f/\Delta_{DNS}$. It is worth noting that, as the thermal flame thickness is resolved with about 10 grid points the ratio of Δ_f/δ_{th} is approximately given by $n_f/10$. Fig. 2 shows a clear improvement of the trained model, notably outperforming the Smagorinsky model in the combustion zone, which depicts counter-gradient transport (i.e. opposite alignment with the real stresses) as reported in [10]. Interestingly, the linear term T_{ij}^1 has been entirely discarded by the genetic algorithm, see Eqn. (26).

Since the plots in Fig. 2 only state good directional agreement between modelled and real stresses, Fig. 3 shows the fitness according to Eqn. (21) of the trained model Eqn. (26) compared to the Smagorinsky model for three filter sizes. The fitness can be understood as the mean root square of the relative error between the model and the subgrid stresses and hence depicts the agreement of the absolute values, where a perfect match would correspond to a fitness of zero. Despite the obvious improvement over the Smagorinsky model, the trained model still has deficiencies and in general both model deteriorate within the flame front where strong heat release is taking place.

As mentioned before the model in Eqn. (26) has to be redimensionalised. Substituting Eqn. (13) and Eqn. (16) into Eqn. (26) yields:

$$\begin{aligned}\tilde{\tau}^{GDS} &= \bar{\rho}\Delta_f^2 C \cdot \left[\left(\tilde{S}^*_{ik}\tilde{S}^*_{kj} - \frac{1}{3}\delta_{ij}\tilde{S}^*_{mn}\tilde{S}^*_{nm}\right) - \left(\tilde{S}^*_{ik}\tilde{\Omega}_{kj} - \tilde{\Omega}_{ik}\tilde{S}^*_{kj}\right)\right] \\ \tilde{S}^*_{ij} &= \tilde{S}_{ij} - \frac{1}{3}\delta_{ij}\tilde{S}_{kk} \\ C &= 8.71 \times 10^{-2}\,.\end{aligned} \tag{27}$$

Hence, the final model is actually independent of any non-dimensionalisation and can be implemented in LES applications. Although the evolutionary algorithm does not know anything about the dimensions of the different terms used in the model, it found a model which is consistent on dimensional grounds.

Beside the fundamental observation that GEP does not use the linear term at all, it is instructive to compare Eqn. (26) respectively Eqn. (27) with known modelling practice in the context of RANS. For statistically two dimensional incompressible flows often the following nonlinear constitutive equation is used (all quantities have to be understood as Reynolds averaged values)

$$\tau_{ij} - \frac{2}{3}k\delta_{ij} = a_1\tilde{S}_{ij} + a_2(\tilde{S}_{ik}\tilde{\Omega}_{kj} - \tilde{\Omega}_{ik}\tilde{S}_{kj}) + a_3\tilde{S}_{ik}\tilde{S}_{kj} \tag{28}$$

and typically a_3 is set to zero because the term $\tilde{S}_{ik}\tilde{S}_{kj}$ can be absorbed in the modified pressure under the above assumption [34,35]. However, it is noted in [36] that by keeping the term $\tilde{S}_{ik}\tilde{S}_{kj}$ a formula is obtained that could be better adapted to three dimensional flows. The same assumption (i.e. $a_3 = 0$) has been used in the explicit algebraic subgrid stress model by Marstop et al. [37] where it is explicitly mentioned that although the Reynolds axioms are not valid in LES it has been assumed that it is possible to use the same modelling approach as in RANS simulations.

4.2. Modelling the SGS turbulent kinetic energy

The competition between the turbulent intensity and the velocity jump across the flame caused by the flame normal acceleration due to heat release determines the nature of turbulent transport [4] which is predominantly positive (e.g. $\tau_{11}^{SGS} > 0$) in our DNS cases [10,11]. Modelling of the deviatoric part of the stresses indicated that the Smagorinsky model yields the opposite alignment as indicated by the DNS. This is due to the strong heat release which gives, for a statistically planar flame propagating in negative x direction, rise to large positive values of $\partial u/\partial x$ which results for the SSM model in $\tau_{11}^{SSM} < 0$. It was demonstrated in [10] that this behaviour can be improved by explicit modelling of the trace of the stress tensor, k^{sgs}. Because the kinetic energy is a scalar quantity, it can be assumed that a model expression depends only on invariants

$$k^{sgs} = \bar{\rho}\Delta_f^2 \cdot \tilde{S}_{mn}\tilde{S}_{nm} \cdot f(I_2, I_3, I_4)\,. \tag{29}$$

50 models were trained based on a filter width of $n_f = 4$ of which the best model, denoted GkM (GEP based k^{SGS} model) or GkM4 was

$$k^{GkM} = \bar{\rho}\Delta_f^2 \cdot \tilde{S}_{mn}\tilde{S}_{nm} \cdot (C_1 - C_2 \cdot I_2) \tag{30}$$

which can be reformulated to

$$k^{GkM} = \bar{\rho}\Delta_f^2 (C_1 \cdot \tilde{S}_{mn}\tilde{S}_{nm} - C_2 \cdot \tilde{\Omega}_{mn}\tilde{\Omega}_{nm})\,. \tag{31}$$

The two constants were determined by the algorithm to be:

$$\begin{aligned}C_1 &= 4.8282 \times 10^{-2} \\ C_2 &= 3.6435 \times 10^{-2}\,.\end{aligned}$$

For reference purposes this model will be compared to the well known Yoshizawa model [38] denoted henceforth k^{YOS}

$$k^{YOS} = 2C_I\bar{\rho}\Delta 2\tilde{S}_{mn}\tilde{S}_{mn} \tag{32}$$

$$C_I = 0.089\,. \tag{33}$$

Furthermore, the trace of the well known tensor diffusivity model due to Clark [39]

$$\tau_{ij}^{CTM} \approx \bar{\rho}\frac{\Delta^2}{12}\frac{\partial\tilde{u}_i}{\partial x_k}\frac{\partial\tilde{u}_j}{\partial x_k} \tag{34}$$

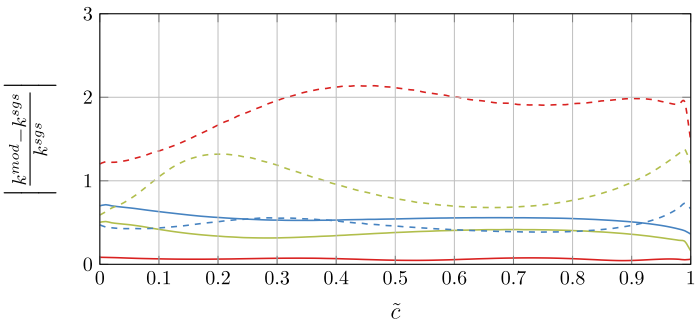

Fig. 4. Relative error of *GkM4* model (——) and *YOS* model (- - -) conditionally averaged on reaction progress for case *A* ($n_f = 4$, $n_f = 16$, $n_f = 28$).

Table 2
Coefficients for models *GkM4 GkM16* and *GkM28*.

Model	C_1	C_2	C_1/C_2
GkM4	4.83×10^{-2}	3.64×10^{-2}	1.33
GkM16	5.10×10^{-2}	4.83×10^{-2}	1.06
GkM28	7.70×10^{-2}	7.11×10^{-2}	1.08

gives rise to a model for k^{sgs} in the following manner

$$k^{CTM} = \bar{\rho}\frac{\Delta^2}{12}\frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_i}{\partial x_k}. \tag{35}$$

Fig. 4 compares the fitness of the trained model Eqn. (31) with the Yoshizawa model Eqn. (33). The trained model shows vastly improved agreement for small filter sizes while both models perform similarly for the biggest filter size tested. This can be explained by the fact that the Yoshizawa model typically would be used in an LES where the filter width will be located in the inertial range [38]. Using the smallest filter size $n_f = 4$ resolved structures are in the inertial subrange and partially even the dissipating range (refer to Table 1). Another more obvious explanation is that the model Eqn. (31) has been trained on this particular filter size.

It is reasonable to assume that the performance of the model depends on the case used for training because the training is an optimisation of the functional expression for this particular case. Hence it is interesting to see, how the models change depending on the training case and particularly on the filter size used in the *a-priori* testing. Two more training runs were conducted for filter sizes $n_f = 16$ and $n_f = 28$.

The resulting expressions (denoted *GkM16* and *GkM28*) had exactly the same functional form as Eqn. (31) but with different constants C_1 and C_2 which are given in Table 2 together with their ratio. Firstly, Table 2 shows that the model parameter is a function of the filter width. Secondly, it can be seen that the model contribution coming from $\tilde{S}_{mn}\tilde{S}_{nm}$ (which appears also in the Yoshizawa model) plays a more important role compared to $\tilde{\Omega}_{mn}\tilde{\Omega}_{nm}$ for the smallest filter width (smaller than the flame thickness) but for the largest filter width both terms are equally important. A possible explanation could be, that within the thin reaction zones regime (TRZ) of combustion strain effects play a more important role compared to vorticity, whenever the subgrid contributions are smaller than the thermal flame thickness.

It is expected that the fitness for a particular filter size decreases provided the model has been trained on this particular data set and in fact this is confirmed by the results.

Fig. 5 shows the cost function for the three versions of the GKM model for three different filter widths. For the *GkM4* model (solid line) the smallest fitness is obtained for (in order) filter widths $n_f = 4, 16, 28$. For the *GkM28* model (dotted line) the situation is reversed and the smallest fitness is obtained for (in order) filter widths $n_f = 28, 16, 4$. For the *GkM16* model the situation is less clear.

We focus now on the local performance of a model which has been quantified using the Pearson correlation coefficient R, which is a value bound between -1 and 1 with $R = 1$ indicating perfect statistical correlation. Unlike the relative error used in Eqn. (4), the correlation coefficient is independent of constant prefactors. Fig. 6 compares the Pearson correlation coefficient between modelled and DNS value of k^{sgs} for the three versions of k^{GkM}, k^{YOS} and k^{CTM}. Correlations are shown for the unburnt region ($\tilde{c} < 0.9$), the flame front ($\tilde{0}.1 < c < 0.9$) and the burned gas side ($\tilde{c} > 0.9$). It can be seen that the trained models outperform Yoshizawa's model by an average of about 0.2 and this difference tends to increase for large filter size.

From these results, the beneficial influence of including the rotation rate tensor becomes evident for modelling the turbulent kinetic energy. Correlations for both cases are very similar, with the only difference that correlation strength decreases with increasing turbulence intensity (i.e. for case *B*). This is consistent with the observations of consistently higher correlation strength on the burned gas side where kinematic viscosity increases with increasing temperature and

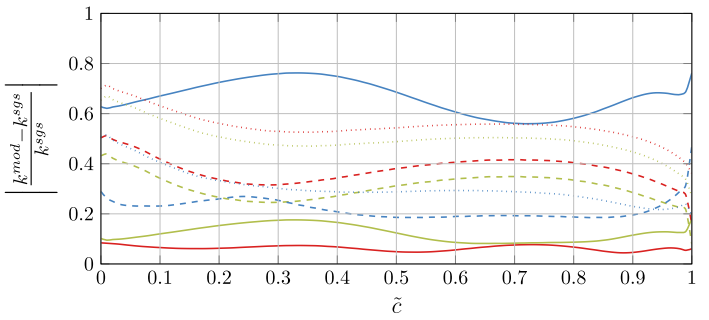

Fig. 5. Cost function for models *GkM4* (solid line) *GkM16* (dashed line) and *GkM28* (dotted line) for the three different filter width ($n_f = 4$, $n_f = 16$, $n_f = 28$).

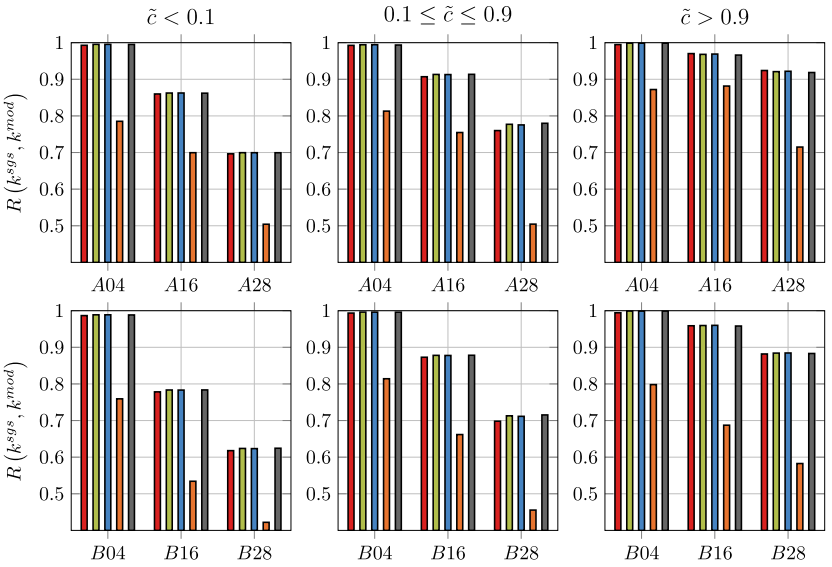

Fig. 6. Correlations for different k^{sgs} models for cases *A* and *B* and three different filter width. Correlations are shown for the unburnt region, the flame front and the burned gas side. *GkM4*, *GkM16*, *GkM28*, *YOS*, *CTM*.

hence turbulent Reynolds number decreases. Results show furthermore that the underlying data set used for training had a small influence on the correlation performance. In fact the correlation coefficients for models *GkM4*, *GkM16* and *GkM28* are very close to each other. This can be explained by the fact that the Pearson correlation is independent of constant multiplying factors, only the ratio C_1/C_2 i.e. the weighting of strain and vorticity rate will result in differences. As the change of this ratio is relatively small (in particular for models *GkM16* and *GkM28*) this explains the observed results.

Another interesting observation from Fig. 6 is that the *GkM4* model variants perform very similar to k^{CTM} i.e. the trace of Clark's model. This similarity can be explained by substituting the definitions of $\tilde{S}_{ij}$ and $\tilde{\Omega}_{ij}$ into Eqn. (31)

$$k^{GkM} = \bar{\rho}\Delta_f^2 \left[\frac{C_1}{4} \cdot \left(\frac{\partial \tilde{u}_m}{\partial x_n}\frac{\partial \tilde{u}_m}{\partial x_n} + 2\frac{\partial \tilde{u}_m}{\partial x_n}\frac{\partial \tilde{u}_n}{\partial x_m} + \frac{\partial \tilde{u}_n}{\partial x_m}\frac{\partial \tilde{u}_n}{\partial x_m} \right) - \frac{C_2}{4} \cdot \left(-\frac{\partial \tilde{u}_m}{\partial x_n}\frac{\partial \tilde{u}_m}{\partial x_n} + 2\frac{\partial \tilde{u}_m}{\partial x_n}\frac{\partial \tilde{u}_n}{\partial x_m} - \frac{\partial \tilde{u}_n}{\partial x_m}\frac{\partial \tilde{u}_n}{\partial x_m} \right) \right]. \quad (36)$$

Table 2 shows the coefficients for the three trained models and their ratios C_1/C_2, i.e. the weighting of the mean strain rate compared to the mean vorticity rate. While the model trained on the smallest filter width has a larger ratio of $\sim 4/3$, the ratio is almost constant for the other two models and close to unity, meaning strain and vorticity rate are equally important. Assuming $C_1 \approx C_2 = C$ and using Einstein notation (sum over m and n) Eqn. (36) can be simplified to

$$k^{GkM} \approx \bar{\rho}\Delta_f^2 C \cdot \frac{\partial \tilde{u}_m}{\partial x_n}\frac{\partial \tilde{u}_m}{\partial x_n}, \quad (37)$$

which has exactly the same structure as k^{CTM} in Eqn. (35). Only the model constant $C_1 \approx C_2 \approx 1/13$ is slightly smaller for the largest filter width compared to the model coefficient in Clarks model, i.e. 1/12.

The beneficial effect of combining models for the deviatoric part of the SGS stresses with an explicit model for the SGS turbulent kinetic energy has been discussed in [11] and hence will not be repeated here. Instead the GEP algorithm will be applied to modelling the full stress tensor itself in the next section.

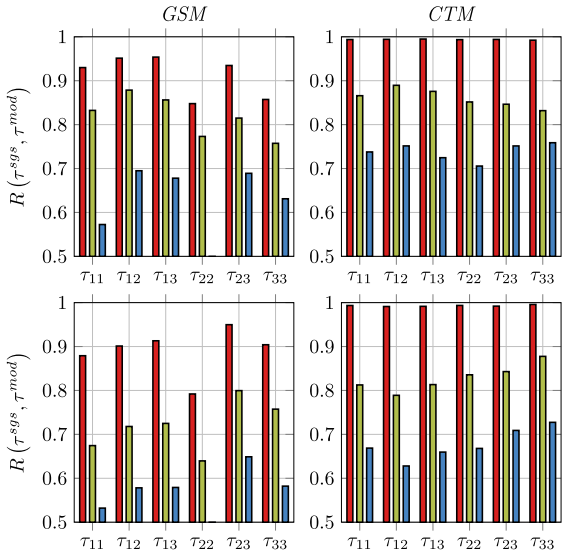

Fig. 7. Correlation coefficient for the subgrid stress tensor components between predictions of *GSM*, and *CTM* model and calculations from DNS averaged over the combustion zone for case *A* (top) and case *B* (bottom) ($n_f = 4$, $n_f = 16$, $n_f = 28$).

4.3. Modelling the full Subgrid Stress Tensor Model

The next step is the training of a model for the full SGS tensor. As the behaviour of SGS models away from the reaction zone is similar to what has been observed and reported for inert flows, the training process will be limited to the combustion zone in this section. The resulting model will be denoted *GSM* (GEP based stress model). The postprocessing is done in parallel using MPI in order to allow for an efficient evaluation of multiple models using the GEP algorithm. To use the computational power efficiently, the cells have been equally redistributed to all processors. Although this operation takes some time, it has to be done only once every run of the GEP algorithm. Obviously, the coordinates of the cells are not stored when they are copied into the one-dimensional array, meaning that the spatial derivatives have to be calculated before the data reduction process.

The best of 50 models found by the GEP code is given by

$$\tilde{\tau}_{ij}^{GSM} = \bar{\rho}\Delta_f^2 \tilde{S}_{mn}\tilde{S}_{nm} \cdot \left[C_2 T_{ij}^3 - C_1 T_{ij}^2\right], \tag{38}$$

or in terms of strain and vorticity rate

$$\begin{aligned} \tilde{\tau}_{ij}^{GSM} &= \bar{\rho}\Delta_f^2 \cdot \left[C_2 \cdot \tilde{S}_{ik}\tilde{S}_{kj} - C_1 \cdot \left(\tilde{S}_{ik}\tilde{\Omega}_{kj} - \tilde{\Omega}_{ik}\tilde{S}_{kj}\right)\right] \\ C_1 &= 9.063 \times 10^{-2} \\ C_2 &= 9.911 \times 10^{-2}. \end{aligned} \tag{39}$$

Although, Eqn. (38) and Eqn. (26) have a similar functional form they are not identical because they make use of the basis functions T_{ij}^3, respectively T_{ij}^{3*} and they are trained on different datasets: The former is trained on the full SGS tensor whereas the latter has been trained on the deviatoric stress tensor.

As reported in [10] Clarks model described by Eqn. (35) performs well within the flame brush and is able to distinguish between gradient and counter-gradient stresses. Therefore, Fig. 7 shows the tensor component correlations for *GSM*, and *CTM* models within the combustion zone for both cases *A* (top) and *B* (bottom). Apparently, the *GSM* model predicts the off-diagonal components much better than the diagonal ones. This behaviour is especially notable looking at the biggest filter size for case *A*. Furthermore in all cases, the correlation strength decreases with increasing filter size as observed before. The overall performance of the *GSM* model is good but does not reach the performance of the CTM model.

It is worth noting that it has been checked that a considerable reduction of the data set (e.g. by skipping data points) leads to very similar results, such that within a fixed framework the optimisation process was astonishingly robust, despite its non-deterministic nature. However, extending the choice of basis functions included in the training was investigated in the last training run by adding basis function T_{ij}^4. The best of the 50 found models using the extended basis, i.e. the one with the highest fitness, had the form

$$\begin{aligned} \tilde{\tau}_{ij}^{GSM2} &= \bar{\rho}\Delta_f^2 C \cdot \left(T_{ij}^3 - T_{ij}^2 - T_{ij}^4\right) \\ C &= 8.21 \times 10^{-2}. \end{aligned} \tag{40}$$

Using the definitions for strain and vorticity rate the basis functions T_{ij}^2, T_{ij}^3 and T_{ij}^4 are given by:

$$T_{ij}^2 = \frac{1}{2}\left[\frac{\partial \tilde{u}_k}{\partial x_i}\frac{\partial \tilde{u}_k}{\partial x_j} - \frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_j}{\partial x_k}\right] \tag{41}$$

$$T_{ij}^3 = \frac{1}{4}\left[\frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_k}{\partial x_j} + \frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_j}{\partial x_k} + \frac{\partial \tilde{u}_k}{\partial x_i}\frac{\partial \tilde{u}_k}{\partial x_j} + \frac{\partial \tilde{u}_k}{\partial x_i}\frac{\partial \tilde{u}_j}{\partial x_k}\right] \tag{42}$$

$$T_{ij}^4 = \frac{1}{4}\left[\frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_k}{\partial x_j} - \frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_j}{\partial x_k} - \frac{\partial \tilde{u}_k}{\partial x_i}\frac{\partial \tilde{u}_k}{\partial x_j} + \frac{\partial \tilde{u}_k}{\partial x_i}\frac{\partial \tilde{u}_j}{\partial x_k}\right] \tag{43}$$

Substituting (41)–(43) into (40) yields the expression

$$\tilde{\tau}_{ij}^{GSM2} = \bar{\rho}\Delta_f^2 C \cdot \frac{\partial \tilde{u}_i}{\partial x_k}\frac{\partial \tilde{u}_j}{\partial x_k} \tag{44}$$

and the constant can be estimated by

$$C = 8.21 \times 10^{-2} \approx \frac{1}{12.18} \approx \frac{1}{12} = 8.333 \times 10^{-2}\,. \tag{45}$$

In other words the model found by the GEP algorithm corresponds exactly to the Clark model, including its model parameter. The performance of Clark's model has been discussed in detail in [10,11] and this will not be repeated here. It is obvious, that using T_{ij}^4 improves the predictions of the subgrid scale model (compare Fig. 7). However, it is more than surprising that the GEP algorithm finds this well known model from the literature, despite all assumptions made and despite its non-deterministic nature. This is an impressive demonstration of the algorithms capabilities and very promising for future applications. Writing the convolution filter as a Taylor series it can be shown that Clark's model represents a fourth order approximation of the subgrid stresses. It could be speculated that by inclusion of even more basis functions possibly higher order terms would be recovered. In order to maintain simplicity and robustness of the model, by not trying to overfit the data used in the training, and to guarantee feasibility of the training process itself this was not pursued further.

5. Conclusion

The concept of Gene Expression Programming was applied to model the unknown subgrid stresses in the context of Large Eddy Simulation of turbulent premixed, statistically planar flames. To the best knowledge of the authors this is the first application of GEP to LES model development. In this context it is worth remarking that using GEP for time averaged DNS data in the context of RANS is conceptually quite different from applying GEP to instantaneous local data both in terms of cost function and computational strategy. The DNS database has been explicitly filtered using a Gaussian filter function for a range of different filter sizes and the GEP algorithm has been based on evaluating a cost function related to *a-priori* assessment of candidate models relative to the SGS stresses from DNS. For the first step, the GEP algorithm was used to model the deviatoric part of the stresses. It was observed that the GEP algorithm consistently selected the constant of the first basis function representing the strain rate tensor to zero. A new model formulation with a notably better alignment characteristic than the Smagorinsky model was found, in particular in the presence of heat release. In the second step GEP was used for modelling the trace of the SGS stress tensor, which is known as the generalised SGS turbulent kinetic energy. The obtained model consisted of two terms: the squared strain rate tensor which is also present in Yoshizawa's model and the rotation rate tensor multiplied with its transposed counterpart. It was found that for small filter sizes the first term had a somewhat larger weight whereas for large filter sizes both terms had an equal weight according to the GEP algorithm. This model showed a better performance than the Yoshizawa model and in the limit of large filter sizes it simplifies to the trace of Clark's tensor diffusivity model. Finally, when modelling the full stress tensor, the GEP algorithm recovered once more the well known Clark model after including the first four basis functions. Note that neither the functional form of the model nor its coefficients had been imposed. Hence, it was very surprising that the GEP algorithm did find this non-trivial combination of basis functions resulting in a well known model from the literature, even with the model constant being very close to the value in Clark's model. A final interesting observation is that the optimised models found by the GEP algorithm always had constant coefficients G_α although they are in principle scalar functions of the invariants. In fact more complex functions G_α had been tested by the algorithm but, at least as long as the dataset was not changed during training, they turned out to be worse than those with constant coefficients. This is in contrast to the authors earlier experience with RANS modelling.

The focus of this work was not to put forward a particular new model but to demonstrate that evolutionary algorithms can successfully be used in the framework of LES modelling. One particularly attractive feature of the proposed methodology is that it combines traditional modelling strategies, like the Caley Hamilton theorem, with machine learning and that the result of the process is not a black box model, but a well defined mathematical expression. Future work includes the application of GEP to other unclosed terms using potentially different basis functions. Extending the Bray–Moss–Libby (BML) relation given in the context of RANS [4] to the LES formalism (see e.g. [10]), we note that, for infinitely thin flame fronts, turbulent transport can be split into turbulence on the fresh gas side, turbulence on the burned gas side and an

intermittency term. Explicit modelling of this intermittency term using GEP will be an important next step but will require a different functional dependence of the SGS stress tensor. The final goal is the application of GEP to a-posteriori LES modelling.

Acknowledgement

The stay abroad of the first author was supported by the Federal Ministry of Defence.

References

[1] J. Chomiak, J. Nisbet, Modeling variable density effects in turbulent flames – some basic considerations, Combust. Flame 102 (1995) 371–386.
[2] P. Pfadler, F. Beyrau, F. Dinkelacker, A. Leipertz, A-priori testing of an eddy viscosity model for the density-weighted subgrid scale stress tensor in turbulent premixed flames, Exp. Fluids 49 (2010) 839–851.
[3] K. Bray, P. Libby, G. Masuya, J. Moss, Turbulence production in premixed turbulent flames, Combust. Sci. Technol. 25 (1981) 127–140.
[4] D. Veynante, A. Trouve, K. Bray, T. Mantel, Gradient and counter-gradient scalar transport in turbulent premixed flames, J. Fluid Mech. 332 (1997) 263–293.
[5] S. Tullis, R. Cant, Scalar transport modeling in large eddy simulation of turbulent premixed flames, Proc. Combust. Inst. 29 (2003) 2097–2104.
[6] S. Nishiki, T. Hasegawa, R. Borghi, R. Himeno, Modelling of turbulent scalar flux in turbulent premixed flames based on DNS database, Combust. Theory Model. 10 (2006) 39–55.
[7] Y. Gao, N. Chakraborty, M. Klein, Assessment of the performances of sub-grid scalar flux models for premixed flames with different global Lewis numbers: a direct numerical simulation analysis, Int. J. Heat Fluid Flow 52 (2015) 28–39.
[8] Y. Gao, N. Chakraborty, M. Klein, Assessment of sub-grid scalar flux modelling in premixed flames for large eddy simulations, Eur. J. Mech. B, Fluids 52 (2015) 97–108.
[9] M. Klein, N. Chakraborty, Y. Gao, Scale similarity based models and their application to subgrid scale scalar flux modelling in the context of turbulent premixed flames, Int. J. Heat Fluid Flow 57 (2016) 91–108.
[10] M. Klein, C. Kasten, N. Chakraborty, Y. Gao, A-priori direct numerical simulation assessment of sub-grid scale stress tensor closures for turbulent premixed combustion, Comput. Fluids 122 (2015) 1–11.
[11] M. Klein, C. Kasten, N. Chakraborty, A-priori direct numerical simulation assessment of models for generalized sub-grid scale turbulent kinetic energy in turbulent premixed flames, Comput. Fluids 154 (2017) 123–131.
[12] M. Boger, D. Veynante, Large eddy simulation of a turbulent premixed v-shape flame, in: C. Dopazo (Ed.), Proceedings of the Eight European Turbulence Conference, CIMNE, Barcelona, 2000, pp. 449–452.
[13] J. Weatheritt, R. Sandberg, A novel evolutionary algorithm applied to algebraic modifications of the RANS stress strain relationship, J. Comput. Phys. 325 (2016) 22–37.
[14] J. Ling, A. Kurzawski, J. Templeton, Reynolds averaged turbulence modelling using deep neural networks with embedded invariance, J. Fluid Mech. 807 (2016) 155–166.
[15] E.J. Parish, K. Duraisamy, A paradigm for data-driven predictive modeling using field inversion and machine learning, J. Comput. Phys. 305 (2016) 758–774.
[16] P.M. Milani, J. Ling, G. Saez-Mischlich, J. Bodart, J.K. Eaton, A machine learning approach for determining the turbulent diffusivity in film cooling flows, in: ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition, American Society of Mechanical Engineers, 2017, V05AT12A002.
[17] J. Weatheritt, R.D. Sandberg, Hybrid Reynolds-averaged/large-eddy simulation methodology from symbolic regression: formulation and application, AIAA J. 55 (2017) 3734–3746.
[18] S. Wallin, A.V. Johansson, An explicit algebraic Reynolds stress model for incompressible and compressible turbulent flows, J. Fluid Mech. 403 (2000) 89–132.
[19] B.A. Younis, C.G. Speziale, T.T. Clark, A rational model for the turbulent scalar fluxes, Proc. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci., vol. 461, The Royal Society, 2005, pp. 575–594.
[20] K. Jenkins, R. Cant, DNS of turbulent flame kernels, in: L. Liu, C. Sakell, T. Beautner (Eds.), Proceedings of 2nd AFOSR Conference on DNS and LES, Kluwer Academic Publishers, 1999, pp. 192–202.
[21] N. Chakraborty, M. Klein, R.S. Cant, Effects of turbulent Reynolds number on the displacement speed statistics in the thin reaction zones regime of turbulent premixed combustion, J. Combust. 2011 (2011) 1–19.
[22] M. Klein, N. Chakraborty, S. Ketterl, A comparison of strategies for direct numerical simulation of turbulence chemistry interaction in generic planar turbulent premixed flames, Flow Turbul. Combust. 99 (2017) 955–971.
[23] T. Echekki, J. Chen, Unsteady strain rate and curvature effects in turbulent premixed methane–air flames, Combust. Flame 106 (1996) 184–202.
[24] N. Peters, P. Terhoeven, J. Chen, T. Echekki, Statistics of flame displacement speeds from computations of 2-D unsteady methane–air flames, Proc. Combust. Inst. 27 (1998) 833–839.
[25] N. Chakraborty, R.S. Cant, Unsteady effects of strain rate and curvature on turbulent premixed flames in an inflow outflow configuration, Combust. Flame 137 (2004) 129–147.
[26] N. Chakraborty, M. Klein, R.S. Cant, Stretch rate effects on displacement speed in turbulent premixed flame kernels in the thin reaction zones regime, Proc. Combust. Inst. 31 (2007) 1385–1392.
[27] J. Lai, M. Klein, N. Chakraborty, Direct numerical simulation of head-on quenching of statistically planar turbulent premixed methane–air flames using a detailed chemical mechanism, Flow Turbul. Combust. (2018), https://doi.org/10.1007/s10494-018-9907-5.
[28] M. Klein, C. Kasten, N. Chakraborty, N. Mukhadiyev, H. Im, Turbulent scalar fluxes in H_2-air premixed flames at low and high Karlovitz numbers, Combust. Theory Model. (2018), accepted.
[29] M. Klein, D. Alwazzan, N. Chakraborty, A direct numerical simulation analysis of pressure variation in turbulent premixed bunsen burner flames, part 1: scalar gradient and strain rate statistics, Comput. Fluids (2018), https://doi.org/10.1016/j.compfluid.2018.03.010.
[30] J.R. Koza, Genetic programming as a means for programming computers by natural selection, Stat. Comput. 4 (1994) 87–112.
[31] C. Ferreira, Gene expression programming: a new adaptive algorithm for solving problems, Complex Syst. 13 (2001) 87–129.
[32] S.B. Pope, A more general effective-viscosity hypothesis, J. Fluid Mech. 75 (1975) 331–340.
[33] A.J.M. Spencer, R.S. Rivlin, The theory of matrix polynomials and its application to the mechanics of isotropic continua, in: Collected Papers of R.S. Rivlin, Springer, 1997.
[34] D. Wilcox, Turbulence Modeling for CFD, DCW Industries, 2006.
[35] S.B. Pope, Turbulent Flows, Cambridge University Press, 2000.
[36] P. Durbin, B. Reif, Statistical Theory and Modeling for Turbulent Flows, Wiley, 2011.

[37] L. Marstorp, G. Brethouwer, O. Grundestam, A. Johansson, Explicit algebraic subgrid stress models with application to rotating channel flow, J. Fluid Mech. 639 (2009) 403–432.
[38] A. Yoshizawa, Statistical theory for compressible turbulent shear flows, with the application to subgrid modeling, Phys. Fluids 29 (1986) 2152–2164.
[39] R.A. Clark, J.H. Ferziger, W. Reynolds, Evaluation of subgrid-scale models using an accurately simulated turbulent flow, J. Fluid Mech. 91 (1979) 1–16.