Available online at www.sciencedirect.com
ScienceDirect
Proceedings of the Combustion Institute 32 (2009) 1605–1611
Proceedings of the Combustion Institute
www.elsevier.com/locate/proci

Turbulent premixed flame modeling using artificial neural networks based chemical kinetics

Baris A. Sen, Suresh Menon *

School of Aerospace Engineering, Georgia Institute of Technology, 270 Ferst Drive, Atlanta, GA 30332-0150, USA

Abstract

The applicability of Artificial Neural Networks (ANN) approach as a chemistry integrator for Large Eddy Simulations (LES) of reactive flows is evaluated with special emphasis on generating training tables independent of the computation of interest. An ANN code based on back-propagation algorithm is developed with a new approach for self-determining the model coefficients adaptively with respect to the error surface topology. The training table is constructed with an independent flame study, and the trained networks are used in LES of turbulent Flame–Vortex Interaction (FVI) studies at different equivalence ratios and turbulence levels. It is shown that once the ANN is well trained, it can successfully predict the reaction rates in both memory and time efficient manner compared to traditional look-up table approach and stiff ODE solvers, respectively.

Keywords: Premixed flames; Large eddy simulation; Artificial neural networks; Chemical kinetics

1. Introduction

One fundamental difficulty for turbulent reactive flow simulations is the representation of the chemical state-space by optimal chemical kinetics mechanisms. Chemical source term calculation requires the solution of a system of non-linear and coupled ODEs [1]. Minor species can have several orders of magnitude different time scales than the major species, therefore, time consuming stiff ODE solvers, such as DVODE [1] are needed. Relatively smaller mechanisms based on partial equilibrium and/or steady state assumptions may be used to decrease the computation burden. However, local extinction, re-ignition and flame quenching problems require radical species information, which is absent in low-order mechanisms.

Up to date, several approaches to speed-up chemical kinetics calculation have been used successfully in Reynolds Averaged Navier Stokes (RANS) solvers. Examples are: look-up table (LUT) [2], intrinsic low-dimensional manifolds (ILDM) [3], and in situ adaptive tabulation (ISAT) [4]. All of these models are based on a tabulation strategy and may require huge memory. Also, tabulating the whole chemical state-space with several reaction progress variables (such as done for ILDM and LUT) may not be adequate to simulate time dependent processes with LES.

ANN is another possible approach and is essentially used to map arbitrary number of inputs to outputs through a learning stage [5]. It has been used in the past both for laminar [6–8] and turbulent applications [9–11] of reactive flows. However, most of these studies rely on training the ANN on a

* Corresponding author. Fax: +1 404 894 2760.
E-mail addresses: suresh.menon@ae.gatech.edu, suresh.menon@aerospace.gatech.edu (S. Menon).

1540-7489/$ - see front matter © 2009 The Combustion Institute. Published by Elsevier Inc. All rights reserved.
doi:10.1016/j.proci.2008.05.077

case very similar to the actual problem of interest, and thus, are limited to a certain range of applicability. The general applicability of the ANN, especially for LES of premixed turbulent flames, therefore, still remains to be demonstrated.

In this paper, we address two major issues related to ANN approach: (1) search for an optimum strategy to generate multi-dimensional training tables independent of the problem of interest, and (2) evaluate the accuracy and the speed-up achieved by using such an ANN in LES for multi-step kinetics. To achieve these goals, a new approach is developed for dynamic evaluation of the ANN model coefficients to increase the convergence rate. The proposed method is evaluated using LES of turbulent premixed syngas/air flames by using a well established subgrid scalar mixing and combustion model.

2. Mathematical formulation and computational approach

2.1. Artificial neural networks

ANN is a computing system made up of a number of processing elements (PEs) connected through communication channels, and the efficiency of these channels is established by a training approach. The ANN code used in the study is based on a back-propagation learning and has been coded in FORTRAN. Its validity has been shown elsewhere [12]. The schematics is given in Fig. 1.

The output of a single PE i, at the hidden layer n, and iteration $[k]$ is

$$y_i^n[k] = g\Big(\sum_{m=0}^{M} W_{im}[k]^n y_m^{n-1}[k] - b_i[k]\Big),$$

where $g()$ is the hyperbolic-tangent activation function. Here, $W_{im}^n[k]$ is the weight coefficient between PEs i and m, ${y_m}^{n-1}[k]$ is the output of the neuron m at layer $n-1$, and $b_i[k]$ is the internal threshold. Once the output is calculated, the error is $E[k] = \frac{1}{2}\sum_{i=1}^{I}[d_i^l - y_i^l[k]]^2$, where d_i^l is the desired and $y_i^l[k]$ is the calculated output. The weights are then manipulated using the gradient descent rule (GDR) [5], $W_{i,m}^n[k+1] = W_{i,m}^n[k] + \eta\Delta w_{i,m}^n[k]$, with,

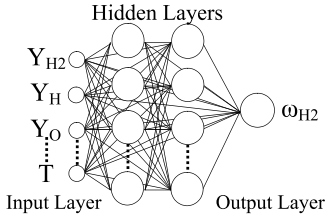

Fig. 1. Sketch of a generic multi-layer perceptron.

$\Delta w_{i,m}^n[k] = -\eta \bigtriangledown E[k]$, where η is the global learning coefficient and $\bigtriangledown E[k]$ is the error gradient across connections.

The error surface may have local minima and a momentum coefficient (α) is used to find the global minimum [5]. The weights are updated finally as:

$$W_{i,m}^n[k+1] = W_{i,m}^n[k] + \eta\Delta w_{i,m}^n[k] + \alpha\Delta w_{i,m}^n[k-1] \tag{1}$$

The determination of the proper values for the model coefficients (α, η) are a problem for GDR [13]. A large momentum would cause the algorithm to diverge, and the same learning rate can not be used in the valleys or shallow regions of the error surface topology. Here, a new approach for self determining the model coefficients with respect to the error surface topology is proposed. The proposed model is a variant of the Extended Delta-Bar-Delta (EDBD) [13]. In this method, instead of having global learning rate (η) and momentum (α) coefficients, each connection has its own model coefficients ($\alpha_{i,j}$, $\eta_{i,j}$), which are updated at each iteration $[k]$ as, $\eta_{i,j}[k+1] = \eta_{i,j}[k] + \Delta\eta_{i,j}[k]$, where

$$\Delta\eta_{i,j}[k+1] = \begin{cases} \kappa_1\lambda[k]\eta_{i,j}[k] & \text{if } \delta_{i,j}[k]\bar{\delta}_{i,j}[k-1] > 0 \\ -\kappa_1\lambda[k]\eta_{i,j}[k] & \text{if } \delta_{i,j}[k]\bar{\delta}_{i,j}[k-1] < 0 \\ 0 & \text{if } \delta_{i,j}[k]\bar{\delta}_{i,j}[k-1] = 0 \end{cases} \tag{2}$$

Here, $\lambda[k]=(1-\exp(-\kappa_2\delta_{i,j}[k]))$, $\delta_{i,j}[k] = \frac{\partial E}{\partial W_{ij}}$ and $\bar{\delta}_{i,j}[k] = (1-\theta)\delta_{i,j}[k-1] + \theta\delta_{i,j}[k]$. κ_1 and κ_2 are the second-order model coefficients selected to be 0.1 and 0.01, respectively. The same form of Eq. (2) is used for $\Delta\alpha_{i,j}[k]$, except by changing $\eta_{i,j}[k]$ with $\alpha_{i,j}[k]$ on the formulation.

In the current approach, the new modifications are: (a) each connection has its own coefficients, (b) the change in the model coefficients are made based on the value of the local error gradient (thus, in regions of huge error gradient, the modification is enhanced), (c) the coefficients are updated after introducing the whole training set so as to satisfy the correlation between $\delta_{i,j}[k]$ and $\delta_{i,j}[k-1]$ pairs, (d) all the model coefficients are saved, and in case the weights start to increase unboundedly, coefficients revert back to the last saved value and decreased by a factor, and (e) the coefficients are adjusted dynamically based on the error gradients ($\delta_{i,j}$ and $\bar{\delta}_{i,j}$). If the solution oscillates around a minimum the model coefficients are decreased. The coefficients are increased, however, if solution approaches the minimum smoothly from one side.

2.2. ANN table generation

As noted earlier, the current approach is to train the ANN using a generic setup within the parameter space of interest, and then to use it

for different cases. Therefore, the training table is constructed by DNS of a laminar flame vortex interaction [14,15]. A 10-step, 14 species chemical kinetics mechanism for syngas combustion is used. Since turbulence is a collection of vortices of different strengths, six cases are run with $U_{C,\max}/S_L$ between 10 and 400. In all these cases $D_C/L_F = 1$. Here, $U_{C,\max}$ is the maximum velocity induced by the vortex and D_C is the vortex diameter. The range is chosen based on observations in large-scale LES of high-Re combustion in gas turbine engines [16]. The grid is fine enough to resolve the laminar flame thickness by 11 points. Simulations are conducted using a parallel solver and a typical case requires only around 1.95 s/iteration on a single processor Intel PC (3.2 GHz Xeon) and therefore, considered efficient. Simulations are run for 15 eddy turn-over time ($t^* = D_C/U_{C,\max}$) after which it is observed that the initial vortex significantly diminishes. The compositional state-space at every $0.2t^*$ is recorded, merged for all cases, and used for the ANN training. As will be discussed later, this approach allows us to include a large number of vortex length and time scales even though the actual number of simulations is limited.

2.3. LES modeling

A well-established LES approach for the fully compressible, unsteady, multi-species, Favre-filtered form of the governing equations is used with a subgrid scale mixing and combustion model. For the sake of brevity, equations are not shown here but can be found elsewhere [15]. For the momentum and energy transport, an eddy viscosity subgrid model is used to calculate the subgrid stresses by solving an additional subgrid scale kinetic energy transport equation.

A closure at the resolved scales (as employed for momentum and energy transport) is not appropriate for subgrid scale combustion modeling since heat release, volumetric expansion and small-scale turbulent stirring all occur at the small scales that are not resolved. Here, the Linear Eddy Mixing model [16–18] is used as a subgrid scale model to account for the combustion processes occurring within every LES cells (this procedure will be called LEMLES hereafter). In this approach, finite rate kinetics can be implemented directly in the subgrid model without requiring any filtering. Thus, the ANN model can be directly implemented within the subgrid model to replace the DVODE approach and a direct comparison can be carried out to evaluate its performance.

2.4. Test cases

The LES of premixed syngas (PSI [15])/air flame is conducted. A background isotropic turbulence is generated by using a prescribed spectrum with fixed u'/S_L and L_{11}/L_F, selected in accordance with earlier studies [19,20]. Here, L_{11} is the integral length scale. The flames are in the thin-reaction zone and the validity of the LEM-LES on this regime has been shown earlier [20]. We focus on three simulations (summarized in Tables 1 and 2), two of which with coherent vortex pair embedded in the background isotropic turbulence to mimic the effect of large scale wrinkling of the flame. A 64^3 grid is used for all cases, with an LES resolution 4 times larger than the Kolmogorov length scale. On the subgrid level, 12 LEM cells are used within each LES cell. Characteristic inflow/outflow boundary conditions are used in streamwise directions, and periodicity in the spanwise and the transverse directions is imposed. Mixture averaged transport coefficients are used with a thermally perfect gas equation of state.

Table 1
Case table

	u'/S_L	$U_{C,\max}/S_L$	D_C/L_F	ϕ
Case 1	10	—	—	0.6
Case 2	5	50	5	0.6
Case 3	10	50	5	0.8

Table 2
Flame parameters

Equivalence ratio	S_L (m/s)	l_f (m)
0.6	0.2105	8.755×10^{-4}
0.8	0.3920	5.611×10^{-4}

3. Results and discussion

3.1. ANN training

One hyperplane accessed by the training table is shown in Fig. 2(a). Although each simulation is for a single equivalence ratio, the states accessed correspond to a certain region on the hyperplane. A fast Fourier transformation (FFT) analysis of the u-velocity component obtained across the flame front is performed and the variation of the amplitude of the most energetic scale (the vortex) is given in Fig. 3(a). The interaction represents a decaying process, and approximately after $5t^*$ the amplitude decreases by a factor of 10. The number of wave-lengths that contribute to the 80% of the energy spectrum is given in Fig. 3(b), and it is shown that large number of scales are present over the entire simulation period. Hence, each test case with an initial single vortex of a given size and strength provides additional data over a much wider range of scales. Therefore, the current

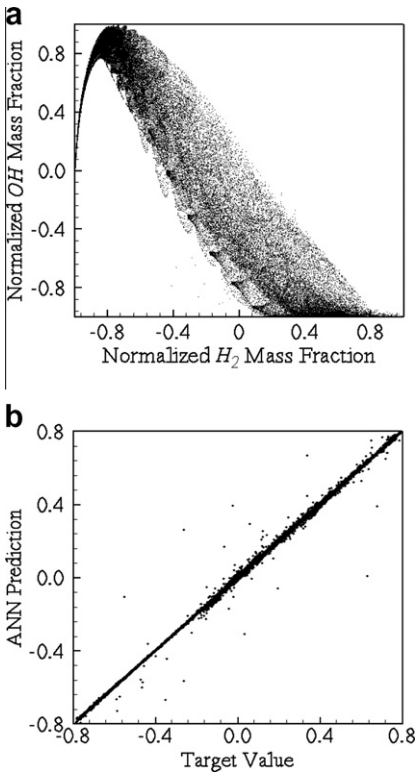

Fig. 2. Training phase. (a) H_2-OH hyperplane accessed by the training table. (b) Correlation between the target value and ANN prediction for all species obtained by the 10/5/4 ANN.

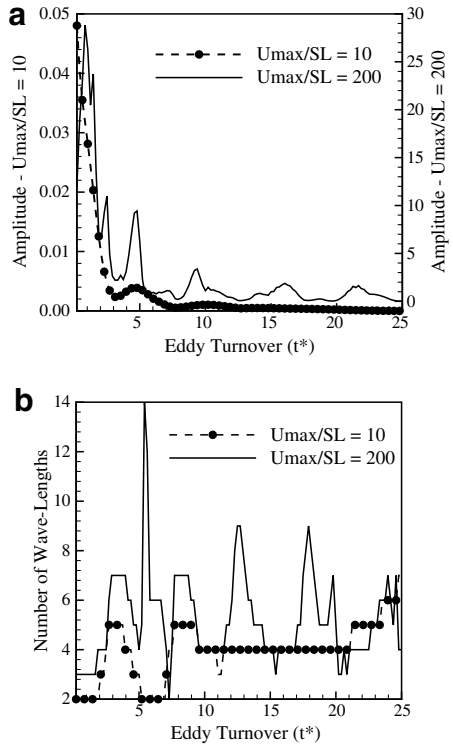

Fig. 3. FFT analysis of the flow across the flame front. (a) Variation of the amplitude of the most energetic wave length. (b) Number of wave lengths that contribute 80% of the energy spectrum.

strategy of simulating and storing data throughout the *unsteady* flame-vortex interaction process, coupled with the Lewis number effect (implicit in the multi-component mixture approach used here) allows us to include a broad region in the compositional state-space (as shown in Fig. 2). This method is found to be more suitable for generating training data compared to using PREMIX-CHEMKIN [12] which is *steady* in its nature and not suitable for LES. Although the training is not fully independent from the LES case, the flow features are not the same since the training is based on a laminar study. The LEM subgrid implementation allows us to use the laminar feature inside LES grid, and therefore this approach is a new strategy to account for turbulent mixing effects.

The table constructed for the ANN training procedure is around 200 MB. Two ANN architectures, one with 3 hidden layers (Case E, denoted 10/5/4 ANN) and the other one with 2 hidden layers (Case C, denoted 10/5 ANN) are discussed here. The selection of these ANNs for discussion is based on error and speed-up estimations of several different configurations (see Table 3). Overall, the smallest error is achieved for Cases E and D, but Case E is discussed here since it is

Table 3
Effect of number of hidden layers and PEs on the training phase

Case	ANN	Error (%)	Time/iteration
A	10	3.91	2.485
B	20	2.97	2.894
C	10/5	1.78	2.728
D	20/10	1.604	4.552
E	10/5/4	1.690	3.959
F	20/10/8	1.704	6.119

almost 15% faster than the latter. Case C is one of the fastest cases but with slightly higher error (compared to Case D), and is selected as a second case for discussion. The RMS error in the test phase is at most 3.4×10^{-3} for the 10/5/4 ANN and 1.3×10^{-3} for the 10/5 ANN, both observed for OH. H_2 exhibits the minimum RMS error of 3.0×10^{-4} and 2.8×10^{-4} for 10/5/4 and 10/5, respectively. Although the 10/5 ANN is almost 20% faster than the 10/5/4 ANN, on the average, the training for one species takes approximately same amount of CPU time regardless of the number of hidden layers. Figure 2(b) shows the correlation curve obtained for the 10/5/4 ANN. The

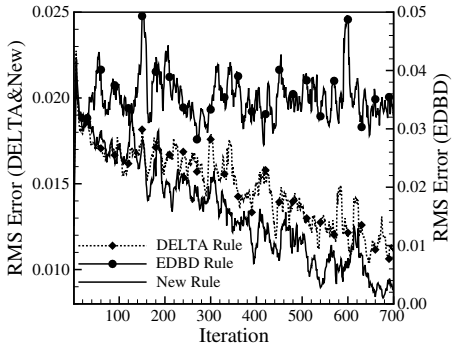

Fig. 4. Convergence history.

correlations for all the species collapse on top of each other, and demonstrate the overall accuracy during the training phase. The ANN predictions are highly correlated with the target value, except for some points. Considering that the plot shows almost 570,000 points, there are only few points outside the target values and they do not affect the overall performance.

The new learning rule proposed within the paper is tested by comparing its performance with standard GDR and EDBD rules in Fig. 4. The EDBD rule has very poor learning compared to the other methods and does not exhibit a convergence, but stays almost flat. This is due to the dependance of the EDBD on the model coefficients, which need to be calibrated. The GDR and the new rules seem similar in the early stage, but as the number of iterations increases the new rule outperforms the GDR. Although not shown in the figure, the GDR eventually attains same value of RMS error with the new rule, but after a larger number of iterations.

3.2. LEMLES of flame–turbulence–vortex interaction

To quantify the accuracy of the ANN, time averaged profiles across the flame are compared in Fig. 5. For the major species, (H_2 and CO_2), ANN and DVODE predicted profiles match very well. For minor species, H and O, the 10/5/4 ANN exhibits errors around 3% and 5%, respectively. The 10/5 ANN over-predicts the maximum value of some species, and this corresponds to an error of 22% for H and 18% for O at maximum even though the overall profiles are very similar. The instantaneous profiles across the flame also show similar errors and therefore, are not shown here. Figure 6 shows the instantaneous profiles for flow Case 2. Note that, the same ANN and solver are employed and only the test conditions are changed. The major species are in good agreement with the DVODE predictions. The 10/5/4 ANN slightly under-predicts the maximum value of H within the flame zone. Overall, there is

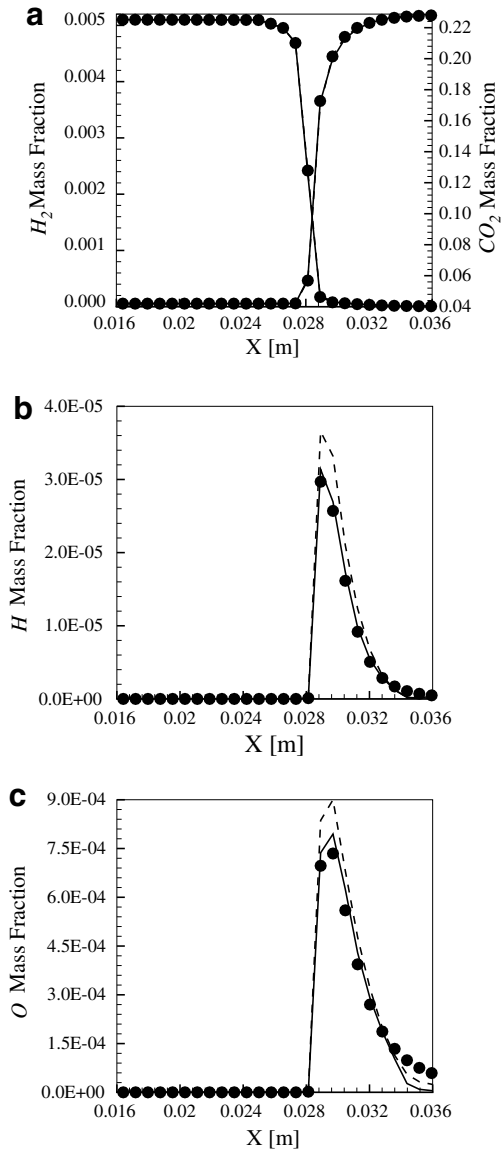

Fig. 5. Comparison of the instantaneous species profiles obtained by the DVODE (●), the 10/5 ANN (---) and the 10/5/4 ANN (—) at $3t^*$ for Case 1. (a) H_2 and CO_2; (b) H; (c) O.

around maximum 10% error for both ANNs. Some radicals show lower error, e.g., for O the error is around 5% for the 10/5/4 ANN. The profiles obtained by the 10/5 ANN are not very accurate in the post flame region and over-predicts the DVODE computations for both of the species. The same ANN approach is used for a different equivalence ratio and turbulence level, and the results are presented in Fig. 7. Similar to the previous cases, the major species are in good agreement and not shown for brevity. For the minor species, H and O, the 10/5 ANN tend to broaden

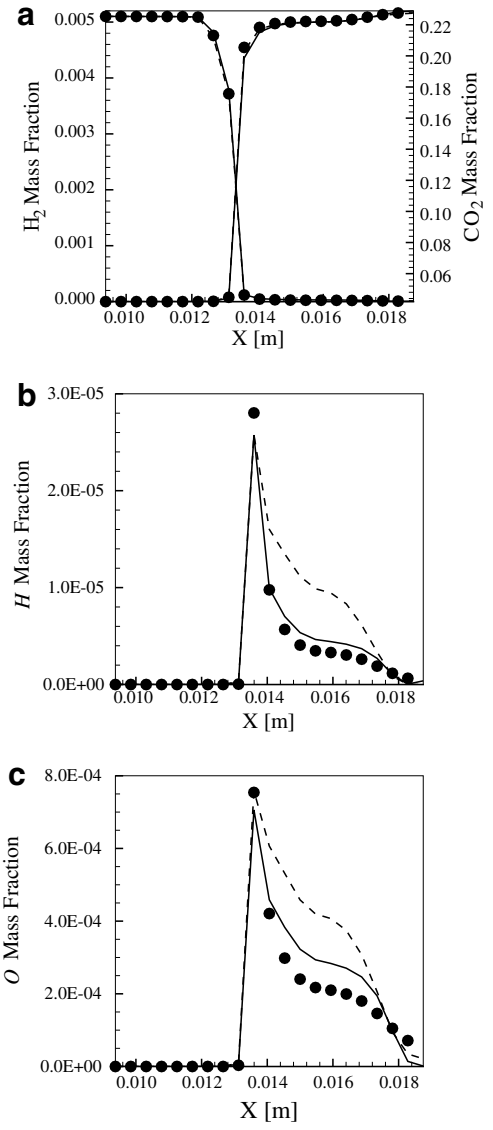

Fig. 6. Comparison of the instantaneous species profiles obtained by the DVODE (●), the 10/5 ANN (---) and the 10/5/4 ANN (—) at $3t^*$ for Case 2. (a) H_2 and CO_2; (b) H; (c) O.

the reaction zone thickness, as shown in the figure. The 10/5/4 ANN on the other hand is more accurate. For both of the species, the 10/5 ANN overpredicts the maximum value as well. It is clear that the overall level of agreement is not changing with respect to the turbulence level or equivalence ratio.

3.3. ANN speed-up and memory savings

The speed-up achieved by using ANN and the memory requirement for LES of turbulent FVI are summarized in Tables 4 and 5, respectively.

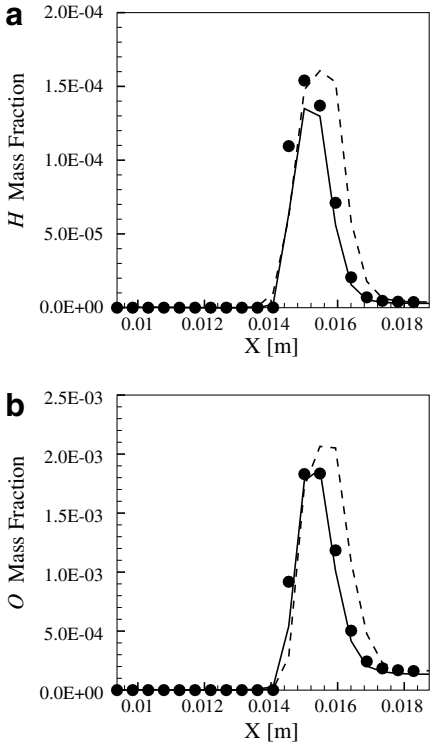

Fig. 7. Comparison of the instantaneous species profiles obtained by the DVODE (●), the 10/5 ANN (---) and the 10/5/4 ANN (—) at $3t^*$ for Case 3. (a) H; (b) O.

For the ANNs discussed here, we achieve around 7× speed-up for the 10/5/4 ANN and 11× speed-up for the 10/5 ANN compared to a stiff ODE solver. The attainable speed-up by ANN depends on the number of connections used in the selected architecture. Although the 10/5/4 ANN outperforms the 10/5 ANN both in the training and in the LEMLES, it uses 389 connections in

Table 4
Speed-up obtained by ANN compared to DVODE

ANN	Connections	Speed-up
20/10/8	986	4.71
10/5/4	390	7.62
10/5	274	11.22
DVODE	—	1.00

Table 5
Memory requirement for the look-up table approach and the ANN

ANN	Look-up table size	ANN size
20/10/8	212.6 MB	0.291 MB
10/5/4	212.6 MB	0.150 MB
10/5	212.6 MB	0.107 MB

comparison to 273, and is slower. Increasing the number of PEs to 20/10/8 reduces the speed-up to around 5×. It appears that the ratio of speed-up between different ANN architectures is roughly equal to the ratio of the number of connections, and this provides a potential design guideline (that will have to be verified with further studies). On the other hand, even though the training table size was large (around 200 MB), the actual ANN stores this information with less than 1 MB. Thus, storage is drastically reduced and this is very important on cheap PC parallel clusters where memory is limited.

4. Conclusions

The efficiency of ANN in terms of memory, time and accuracy for calculating chemical source terms (for a multi-species reduced kinetics) within LES of turbulent premixed flames is investigated. A new training table construction and ANN learning strategy is proposed in this paper. The proposed training approach converges faster with a reasonable RMS error. The training strategy employs a simple flame-vortex test case but stores the entire *unsteady* database to provide a large range of time and length scales. The ANNs trained on this database are used without any change within the subgrid in the LEMLES and direct comparison with the DVODE evaluations show that the ANN can accurately predict the species profiles with an acceptable error. Furthermore, a speed-up of approximately 11 (for 10/5) is achieved, and the memory requirement is only around 0.1 MB of memory per processor, which is substantially smaller than for LUT or ISAT [18]. This is especially very important for the practical LES of large-scale gas turbine combustors that require huge computational resources both in terms of time and memory. New LES studies are beginning to evaluate the current ANN training and usage strategy in LEMLES of much more complex high-Re reacting flows, and will be reported in the future.

Acknowledgments

This work was support in part by NASA Glenn Research Center. The authors thank Prof. J.-Y. Chen for providing the syngas mechanism.

References

[1] K.K. Kuo, *Principles of Combustion*, first ed., John Wiley and Sons, Incorporated, New York, 1986.
[2] J.Y. Chen, W. Kollmann, R.W. Dibble, *Combust. Sci. Technol.* 64 (1989) 315–346.
[3] U. Maas, S.B. Pope, *Combust. Flame* 88 (3–4) (1992) 239–264.
[4] S.B. Pope, *Combust. Theor. Model.* 1 (1) (1997) 41–63.
[5] C. Christodoulou, M. Georgiopoulos, *Applications of Neural Networks in Electromagnetics*, Artech House, Massachusetts, Boston, 2001.
[6] J.A. Blasco, N. Fueyo, C. Dopazo, J. Ballester, *Combust. Flame* 113 (1998) 38–52.
[7] J.A. Blasco, N. Fueyo, J.C. Larroya, C. Dopazo, J.Y. Chen, *Comput. Chem. Eng.* 23 (1999) 1127–1133.
[8] J.Y. Chen, J.A. Blasco, N. Fueyo, C. Dopazo, *Proc. Combust. Inst.* 28 (2000) 115–121.
[9] F.C. Christo, A.R. Masri, E.M. Nebot, *Combust. Flame* 106 (1996) 406–427.
[10] R. Kapoor, A. Lentati, S. Menon, AIAA-01-3847, 37th AIAA Joint Propulsion Conference (2001).
[11] A. Kempf, F. Flemming, J. Janicka, *Proc. Combust. Inst.* 30 (2005) 557–565.
[12] B.A. Sen, S. Menon AIAA-07-5635, 43rd Joint Propulsion Conference (2007).
[13] R.A. Jacobs, *Neural Networks* 1 (1988) 295–307.
[14] T. Poinsot, D. Veynante, S. Candel, *J. Fluid Mech.* 228 (1991) 561–606.
[15] B.A. Sen, S. Menon, AIAA-07-1435, *45th AIAA Aerospace Sciences Meeting and Exhibit*, 2007.
[16] S. Menon, N. Patel, *AIAA J.* 44 (4) (2006) 709–723.
[17] A.R. Kerstein, *Combust. Sci. Technol.* 60 (1988) 391–421.
[18] H. El-Asrag, T. Lu, C.K. Law, S. Menon, *Combust. Flame* 150 (1–2) (2007) 108–126.
[19] A. Trouve, T. Poinsot, *J. Fluid Mech.* 278 (1994) 1–31.
[20] V. Sankaran, S. Menon, *Proc. Combust. Inst.* 30 (2005) 575–582.