An overview of data-driven subgrid-scale modeling

KaTRIS-NUS-KU workshop

Zhao Jiaxi (NUS)

2024.02.23

SGS modeling

In LES small fluctuations of a flow variable f are filtered out and we are concerned with the resolved-scale or GS flow field $\overline{f} = \int G(\boldsymbol{x}')f(\boldsymbol{x}-\boldsymbol{x}')d\boldsymbol{x}'$, where G is a filter function. The GS flow field is governed by the filtered Navier-Stokes equations

$$\frac{\partial \overline{u_i}}{\partial t} + \frac{\partial}{\partial x_j}\left(\overline{u_i}\ \overline{u_j}\right) = -\frac{\partial \overline{p}}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 \overline{u_i}}{\partial x_k \partial x_k} - \frac{\partial \tau_{ij}}{\partial x_j}, \tag{1}$$

$$\frac{\partial \overline{u_j}}{\partial x_j} = 0. \tag{2}$$

Here the residual or SGS stress tensor

$$\tau_{ij} = \overline{u_i u_j} - \overline{u_i}\ \overline{u_j} \tag{3}$$

depends not only on the GS flow field but also on fluctuations and should be modeled using the GS flow field.

Classical SGS modeling

- Smagorinsky model:

$$\tau_{sgs} = 2\rho\nu_{sgs}S^{\star}_{ij} - \frac{2}{3}\rho k_{sgs}\delta_{ij} \qquad S^{\star}_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} - \frac{1}{3}\frac{\partial U_k}{\partial x_k}\delta_{ij}\right)$$

$$\nu_{sgs} = (C_s\Delta)^2\sqrt{2S_{ij}S_{ij}}$$

- Wall-adapting local eddy-viscosity (WALE) model:

$$\nu_t = (C_m\Delta)^2\frac{\left(\mathcal{S}^d_{ij}\mathcal{S}^d_{ij}\right)^{3/2}}{\left(\overline{S}_{ij}\overline{S}_{ij}\right)^{5/2} + \left(\mathcal{S}^d_{ij}\mathcal{S}^d_{ij}\right)^{5/4}},$$

$$\Delta_s = C_w V^{1/3}$$
$$S^d_{ij} = \frac{1}{2}\left(\overline{g}^2_{ij} + \overline{g}^2_{ji}\right) - \frac{1}{3}\delta_{ij}\overline{g}^2_{kk}$$
$$\overline{g}_{ij} = \frac{\partial\overline{u_i}}{\partial x_j}$$
$$\overline{g}^2_{ij} = \overline{g}_{ik}\overline{g}_{kj}$$

- Dynamic models & nonlinear and deconvolution models

Data-driven SGS modeling

- **First, training data for ANN are prepared using DNS data; the GS flow field are calculated from DNS data and used as input variables of ANN; the SGS stress tensor is also calculated from DNS data and used as training targets of output variables.**
- **Then, ANN is trained to establish a functional relation between the input and output variables.**
- **Finally, the ability of the trained ANN is checked using DNS data which are not used in training.**

Issue:

- lots of filter choices which may also result in different data-driven model

- DNS or high-resolution LES can only perform over simple geometry, while the test is over complex geometry

Searching for turbulence models by artificial neural network

- Filter function: top-hat filter function (box filter)
- The components of SGS stress tensor are averaged in the streamwise and spanwise directions

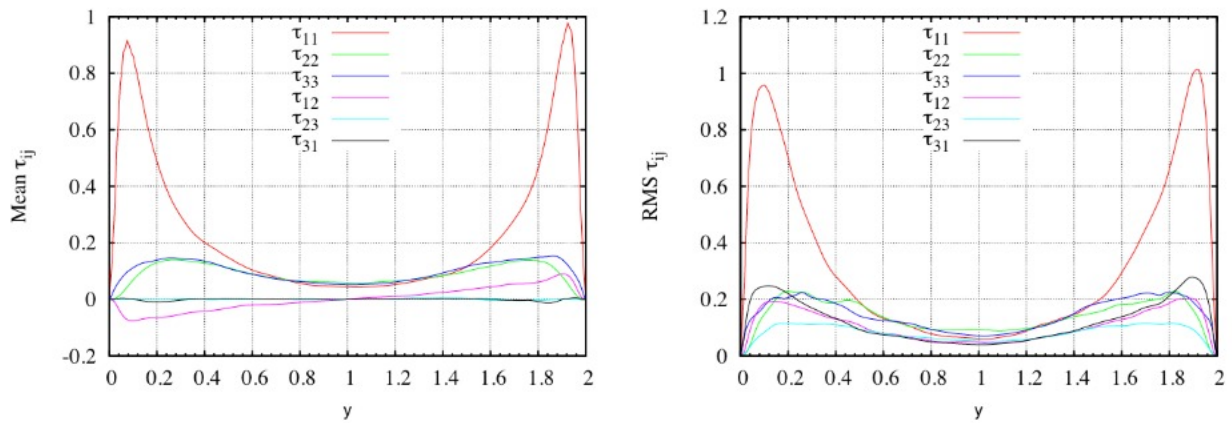


FIG. 2: Distributions of SGS stress tensor τ_{ij} averaged in streamwise and spanwise directions (xz-plane). $Re_\tau = 180, (\overline{\Delta x}^+, \overline{\Delta y}^+_{\max}, \overline{\Delta z}^+) = (35.3, 9.9, 17.7)$. (Left) Average and (right) rms amplitude of fluctuation.

Re_τ	$\overline{\Delta x}^+$	$\overline{\Delta y}^+_{\max}$	$\overline{\Delta z}^+$	$N_x \times N_y \times N_z$
180	41.7	15.6	23.6	$48 \times 48 \times 48$
	43.5	14.5	21.7	$52 \times 52 \times 52$
	35.3	9.9	17.7	$64 \times 64 \times 64$
	23.6	6.0	11.8	$96 \times 96 \times 96$
400	39.3	22.0	19.6	$64 \times 64 \times 64$
	34.3	17.9	17.5	$72 \times 72 \times 72$
	29.9	15.3	15.0	$84 \times 84 \times 84$
	26.2	13.3	13.1	$96 \times 96 \times 96$
600	29.5	33.0	14.7	$128 \times 64 \times 128$
	26.2	26.9	13.1	$144 \times 72 \times 144$
	22.4	23.0	11.2	$168 \times 84 \times 168$
	19.6	19.9	9.8	$192 \times 96 \times 192$
800	26.2	10.9	13.1	$96 \times 256 \times 192$
	19.6	10.9	9.8	$128 \times 256 \times 256$
	13.1	10.9	6.5	$192 \times 256 \times 384$

Data-driven model

- The amount of data used for training is rather small; typically six positions in the streamwise direction are randomly chosen and the data on the corresponding planes parallel to the yz plane are used for training. The whole data are used for test of the trained ANN.
- 6 ANNs are trained to approximate all components of the SGS stress.

same point are used as input variables. We test four sets of input variables: (i) $\{\boldsymbol{S}, y\}$; (ii) $\{\boldsymbol{S}, \boldsymbol{\Omega}, y\}$; (iii) $\{\boldsymbol{\nabla}\overline{\boldsymbol{u}}, y\}$; (iv) $\{\boldsymbol{\nabla}\overline{\boldsymbol{u}}\}$, where $\boldsymbol{S} = [\boldsymbol{\nabla}\overline{\boldsymbol{u}} + (\boldsymbol{\nabla}\overline{\boldsymbol{u}})^T]/2$ and $\boldsymbol{\Omega} = [\boldsymbol{\nabla}\overline{\boldsymbol{u}} - (\boldsymbol{\nabla}\overline{\boldsymbol{u}})^T]/2$. The first set can give the Smagorinsky model $\tau_{ij} = -2C_S\overline{\Delta}^2(2S_{kl}S_{kl})^{1/2}S_{ij} + (1/3)\tau_{kk}\delta_{ij}$, while the position y is included to take account of possible dependence on the wall-normal direction. The GS vorticity is added in the second set since it improves the accuracy of the SGS model in some flows [25]. The third set is equivalent to the second one in the sense that $\boldsymbol{\nabla}\overline{\boldsymbol{u}} = \boldsymbol{S} + \boldsymbol{\Omega}$; in practice, however, the results of learning can be different. The last set is included to check the role of the position y.

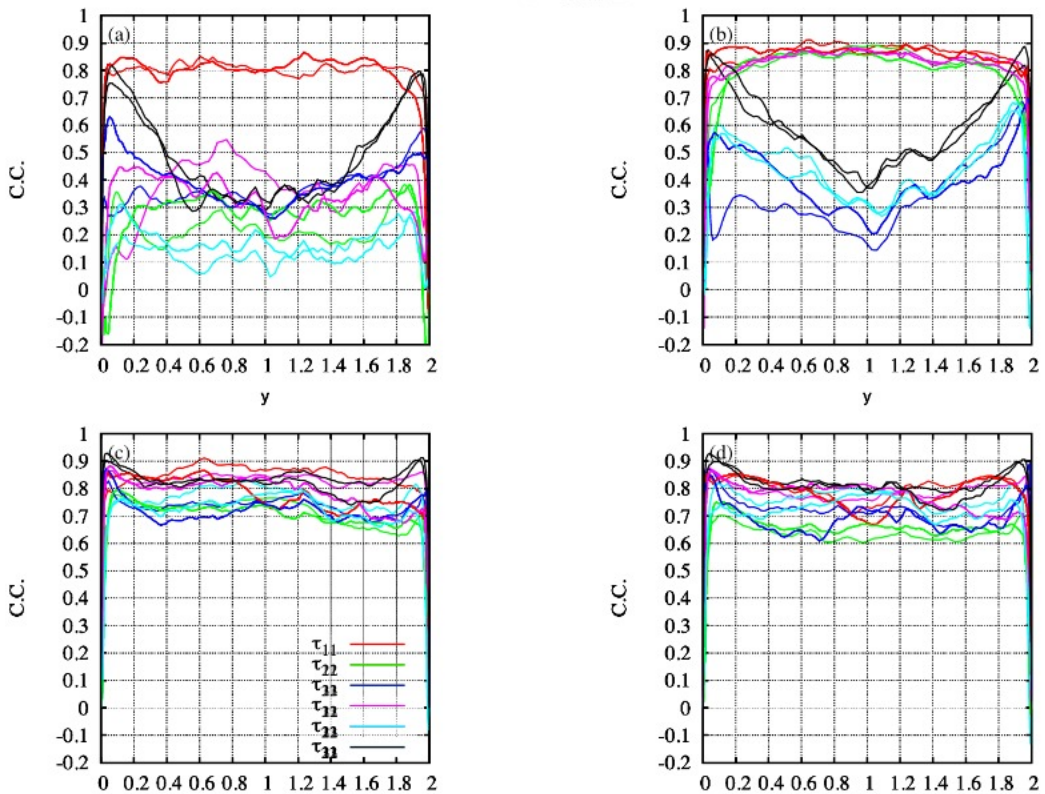
Choice of input variable

$$\mathrm{C.C.}(y) = \frac{\left\langle \left(\tau_{ij}^{(DNS)} - \left\langle \tau_{ij}^{(DNS)} \right\rangle_{xz} \right) \left(\tau_{ij}^{(ANN)} - \left\langle \tau_{ij}^{(ANN)} \right\rangle_{xz} \right) \right\rangle_{xz}}{\left[\left\langle \left(\tau_{ij}^{(DNS)} - \left\langle \tau_{ij}^{(DNS)} \right\rangle_{xz} \right)^2 \right\rangle_{xz} \right]^{1/2} \left[\left\langle \left(\tau_{ij}^{(ANN)} - \left\langle \tau_{ij}^{(ANN)} \right\rangle_{xz} \right)^2 \right\rangle_{xz} \right]^{1/2}}$$

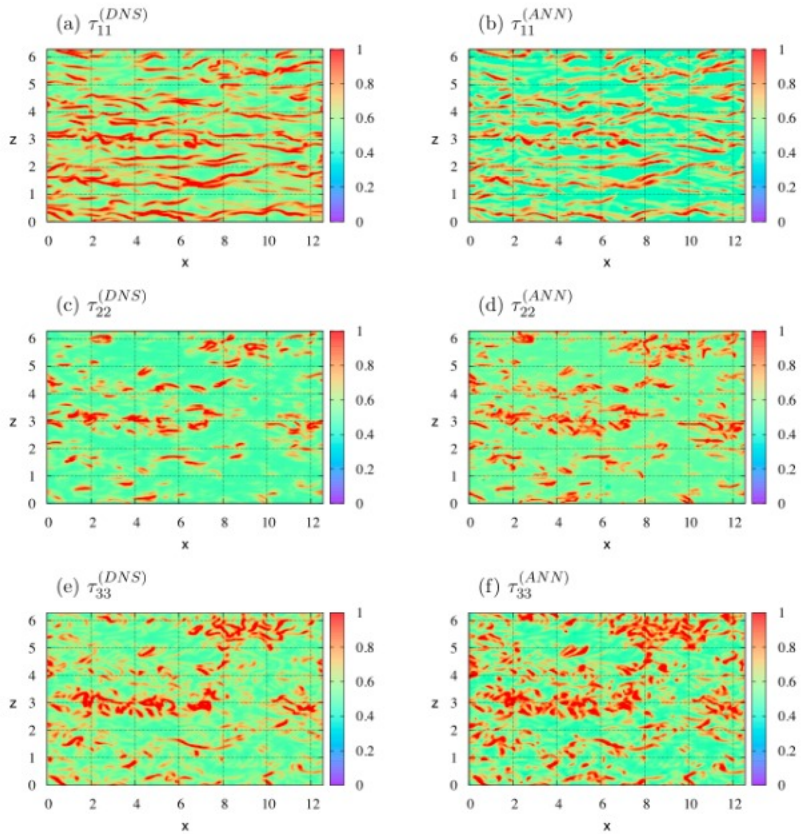
and Ferziger (1979) that the correlation between $\tau_{ij}^{\mathrm{r,Smag}}$ and $\tau_{ij}^{\mathrm{r,DNS}}$ is greatest when the Smagorinsky coefficient is taken to be $C_{\mathrm{S}} = 0.17$, in agreement with the Lilly analysis. However, the correlation coefficient is quite small, around $\frac{1}{3}$. A higher correlation coefficient, around $\frac{1}{2}$, is found for the scalar quantity

$$\overline{U}_i \frac{\partial \tau_{ij}^{\mathrm{r}}}{\partial x_j} = -\frac{\partial \overline{U}_i}{\partial x_j} \tau_{ij}^{\mathrm{r}} + \frac{\partial}{\partial x_j} \left(\overline{U}_i \tau_{ij}^{\mathrm{r}} \right), \qquad (13.184)$$

which is related to the transfer of energy to the residual motions, $\mathcal{P}_{\mathrm{r}}$.



In and out-of-distribution generalization



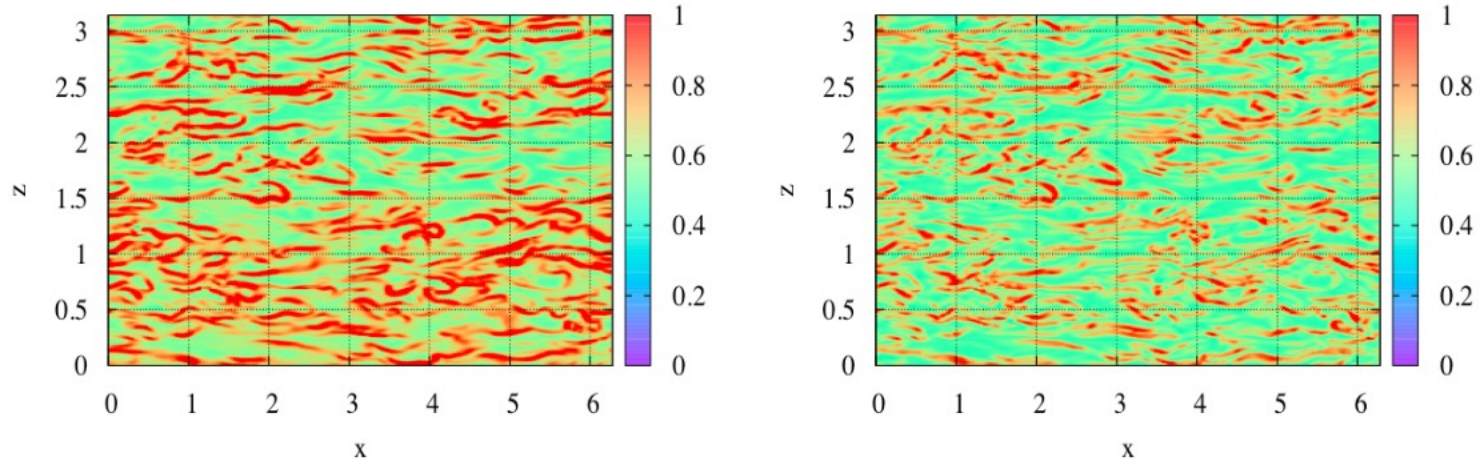
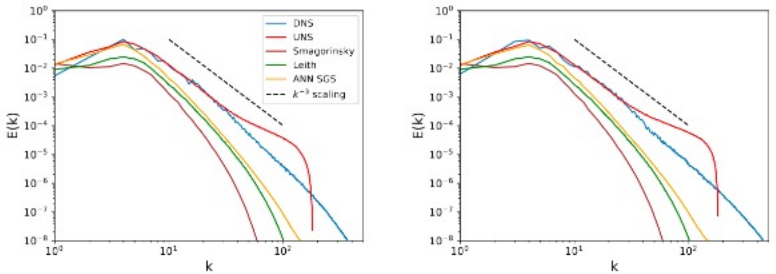


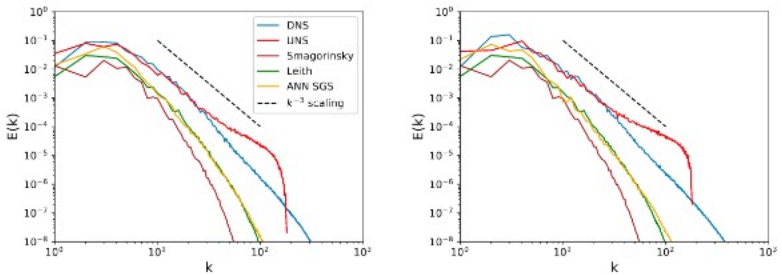
FIG. 12: Comparison between $\tau_{11}^{(DNS)}$ obtained by filtering DNS data of $Re_\tau = 400$ with filter size $(\overline{\Delta x}^+, \overline{\Delta y}^+_{\max}, \overline{\Delta z}^+) = (34.3, 17.9, 17.5)$ and $\tau_{11}^{(ANN)}$ predicted by ANN trained at $Re_\tau = 180$ with filter size $(\overline{\Delta x}^+, \overline{\Delta y}^+_{\max}, \overline{\Delta z}^+) = (35.3, 9.9, 17.7)$. $y = 0.1$.

A posteriori estimate



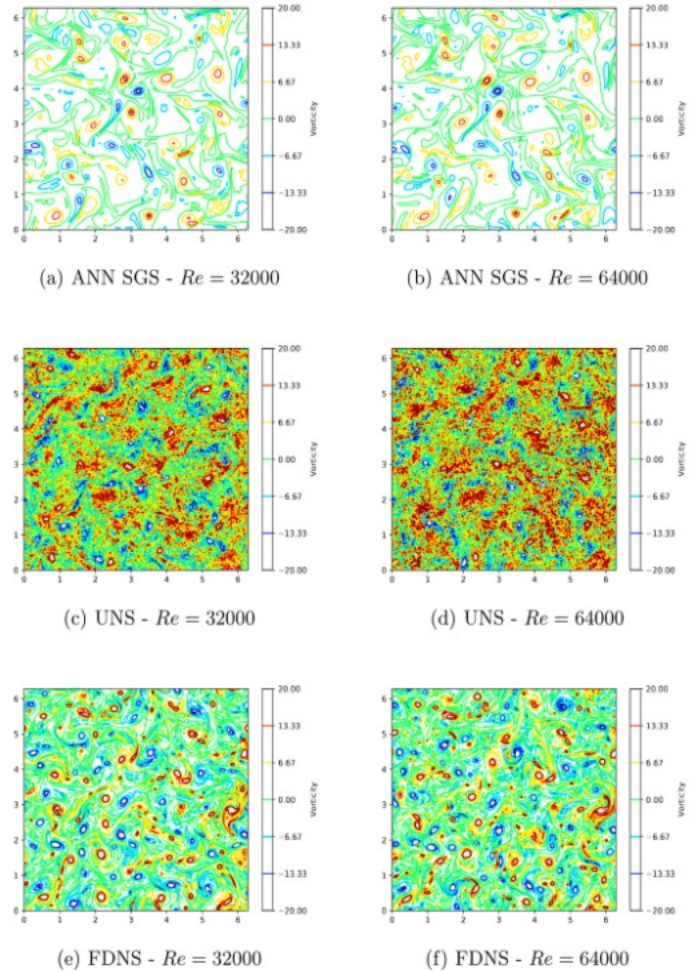
(a) $Re = 32000$ (b) $Re = 64000$

FIGURE 10. A-posteriori results for 24 ensemble-averaged simulations for $Re = 32000$ (left) and $Re = 64000$ (right).



(a) $Re = 32000$ (b) $Re = 64000$

FIGURE 11. The deployment of our framework till $t = 6$ for $Re = 32000$ (left) and $Re = 64000$ (right) showing that a sub-grid model has been learned for utility beyond the training region. We note that the training region is defined between $t = 0$ and $t = 4$ alone.



(a) ANN SGS - $Re = 32000$ (b) ANN SGS - $Re = 64000$

(c) UNS - $Re = 32000$ (d) UNS - $Re = 64000$

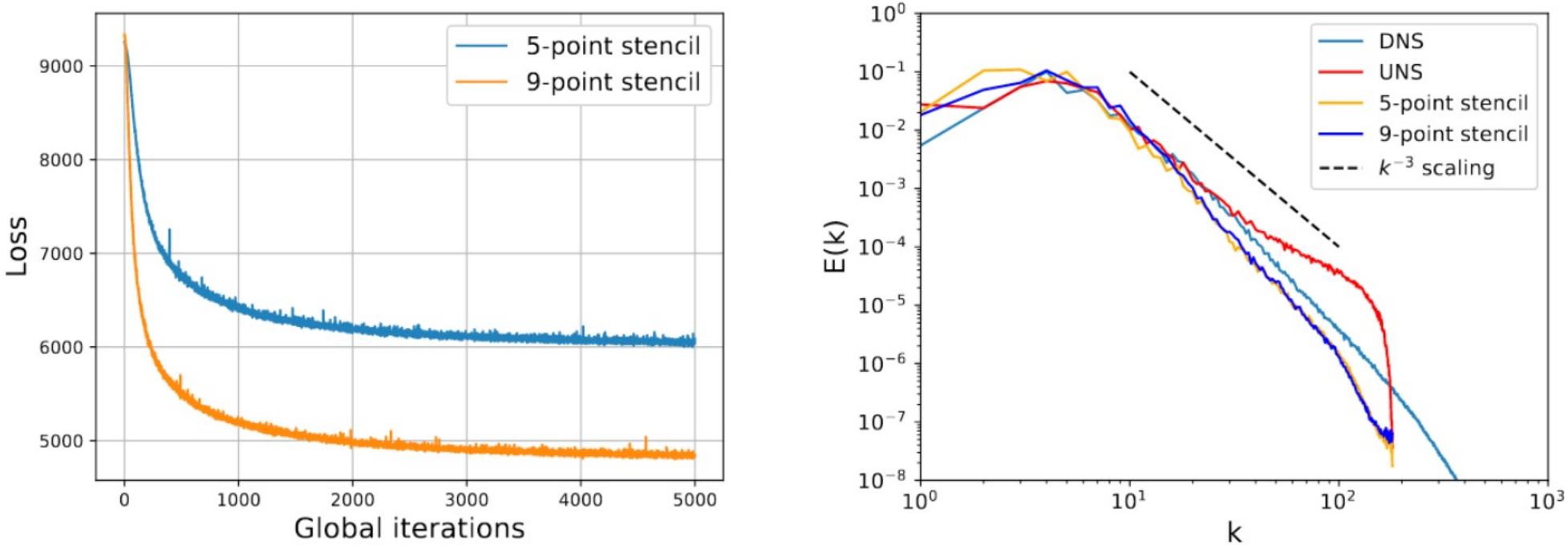
(e) FDNS - $Re = 32000$ (f) FDNS - $Re = 64000$

FIGURE 12. A-posteriori results for the proposed framework showing vorticity fields for $Re = 32000$ and $Re = 64000$ data using coarse-grained grids (top). We also provide no-model simulations (middle) and filtered DNS contours (bottom) for the purpose of comparison.

A priori & a posteriori dichotomy

- Effect of eddy-viscosity inputs
- A-posteriori informed architecture selection
- Stencil selection

A priori & a posteriori dichotomy



(a) Learning rate

(b) A-posteriori deployment

FIGURE 17. A-priori (left) and a-posteriori (right) effect of the stencil size in the 5-layer, 50 neuron framework for a $Re = 32000$ simulation. With deeper architectures, the 5 and 9-point stencils show similar statistical performance

Conclusions & Extensions

- We also remark that the values of correlation coefficients between DNS and ANN obtained in the present study are comparable to or even larger than those for similarity models but smaller than those for a dynamic two-parameter mixed model, which implies that ANN is a promising tool for searching for new turbulence models and should be improved.
- Another way is to replace the training target or output variable; in the present study we chose the SGS stress tensor as the output variable, but it can be other quantities like the rate of production of residual energy or SGS dissipation, which is important in the energy transfer between GS and SGS scales.

A brief summary

- Physical prior and symmetry can be added into the neural network, several related work in RANS modeling [6]
- Do not assume any prior knowledge and learn the mapping between input (GS flow field) and output (usually the SGS stress) [1, 2, 3, 5]
- The parameters in the SGS models are mostly empirical, use ML to identify better parameters (may depend on different condition) [4]
  - reduce the computational time by replacing the evaluation of the turbulent viscosity coefficient by ANN

Open discussion

- Should we rely on classical model and identify the empirical parameter (interpretable) or equation-free turbulence modeling (non-interpretable)
- Generation of the training data?
- How to choose the loss criterion:
  - Is it really good to have high correlation coefficient?
  - A priori and a posteriori estimate
- Generalizability from training scenario to test scenario
- Time/accuracy consideration?

Reference:

- 1. Gamahara, Masataka, and Yuji Hattori. "Searching for turbulence models by artificial neural network." Physical Review Fluids 2.5 (2017): 054604.
- 2. Maulik, Romit, et al. "Subgrid modelling for two-dimensional turbulence using neural networks." Journal of Fluid Mechanics 858 (2019): 122-144.
- 3. Maulik, Romit, et al. "Sub-grid scale model classification and blending through deep learning." Journal of Fluid Mechanics 870 (2019): 784-812.
- 4. Sarghini, Fabrizio, G. De Felice, and Stefania Santini. "Neural networks based subgrid scale modeling in large eddy simulations." Computers & fluids 32.1 (2003): 97-108.
- 5. Wang, Zhuo, et al. "Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation." Physics of Fluids 30.12 (2018).

Reference:

- 6. Ling, Julia, Andrew Kurzawski, and Jeremy Templeton. "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance." *Journal of Fluid Mechanics* 807 (2016): 155-166.
- 7.