encit 2020

ABCM Associação Brasileira de Engenharia e Ciências Mecânicas

18th Brazilian Congress of Thermal Sciences and Engineering
November 16–20, 2020 (Online)

ENC-2020-0446
Strategies to mitigate the error propagation of explicit Reynolds Stress Tensor closures

Bernardo Pereira Brener
Roney Leon Thompson
Matheus Altomare Cruz
COPPE, Department of Mechanical Engineering, Universidade Federal do Rio de Janeiro, Centro de Tecnologia, Ilha do Fundao, 21945-970, Rio de Janeiro, RJ, Brazil
bernardobrener@coppe.ufrj.br
rthompson@mecanica.coppe.ufrj.br
matheusaltomare@ufrj.br

Abstract.
Reynolds-averaged-Navier-Stokes is still the most applied approach in turbulent flow simulations for industrial applications. Despite the well-known lack of accuracy of the commonly used linear eddy viscosity models, the higher costs of more accurate approaches, such as LES (Large Eddy Simulation) and DNS (Direct Numerical Simulation), makes the cost/benefit relation of closures for the Reynolds Stress Tensor (RST) still competitive. Recent works employed DNS databases to analyze the ability of RANS equations to recover the mean velocity field, by plugging explicitly the DNS RST, as a source term. For the plane channel flow, it was shown that small errors in the RST could lead to large discrepancies in the mean velocity field. This result raised a concern about the conditioning of the RANS equations, i.e. whether this set of equations amplify small errors present in the RST closure. Some contributions in the literature proposed an implicit treatment of the RST as a way to obtain smaller error propagation in the recovered mean velocity field. These studies are of great importance for the emerging field of data-driven turbulence modeling as well as for the conventional turbulence modeling. The present work conducted an analysis into the strategies to mitigate this error amplification. Two different approaches in solving RANS equations with DNS data closure are presented, in order to elucidate the error amplification nature. Both result from decoupling the two main aspects in this study, namely the implicit treatment of the linear part of the RST with respect to the rate-of-strain tensor and the use of information of the DNS mean velocity field. It is shown that, although the implicit treatment process leads to smaller errors in the mean velocity profile, the major factor for accuracy gain is the use of information from the DNS mean velocity field. This analysis is confirmed for two problems, the flow through a square duct and another over a periodic hill.

Keywords: *Error Amplification, Reynolds Stress Closure, Explicit/Implicit Treatment, Data-driven Turbulence Modeling*

1. INTRODUCTION

There is a wide range of industrial applications which the flow of interest is turbulent. The most applied approach to predict those flows is Reynolds Averaged Navier-Stokes (RANS) with the commonly used eddy viscosity models (Launder and Sharma (1974); Wilcox (2008); Menter (1994)). However, those models have poor accuracy for many flows in engineering, such as separated flows (Craft *et al.* (1996)). Mainly, due to its intrinsic inability to deal with non-equilibrium turbulence, which is the unbalance between the production and dissipation of turbulence (Speziale and Xu (1996); Hamlington and Dahm (2008)). For more accurate simulations, DNS and LES offers a higher cost alternative. Nonetheless, for high Reynolds number and complex geometries the cost of DNS simulations are prohibitively. Although LES approach reduces the costs, becoming more feasible, it is still much higher when compared to a RANS simulation. Another alternative that provides improvements in modeling are the Reynolds stress models (RSM), which solves a transport equation for Reynolds stresses. Although those models perform well in non-equilibrium turbulence, the numerical robustness and stability are still a challenge, making difficult to obtain convergence (Basara and Jakirlic (2003)). Considering the main higher costs and more complex modeling alternatives, the cost/benefit relation of RANS simulations makes it still play an import role for the industry. Although RSM models does not offer much higher costs, the lack of numerical robustness makes it unpractical and difficult to handle for industrial applications.

An emerging field that optimize the cost/benefit relation of high accuracy alternatives is the called data-driven turbulence modeling (Ling *et al.* (2016a); Ling *et al.* (2016b); Kutz (2017); Wang *et al.* (2017); Wu *et al.* (2018); Kaandorp and Dwight (2020)). With the aid of high accuracy data, such as DNS and LES, a machine learning model is trained to target turbulence quantities. The most traditional target is the Reynolds stress tensor, even though other quantities offer more accurate solutions, for instance, the Reynolds Force Vector (RFV) (Cruz *et al.* (2019)). As an input of the machine learning model, a low cost RANS simulation is used, such as a $\kappa - \epsilon$ model (Launder and Sharma (1974)). Once the machine learning model is trained, it provides the turbulence closure and RANS equations are solved to the mean velocity. Many works were done in that field, obtaining a very tight accuracy for the mean velocity field when compared to the DNS. (Cruz *et al.* (2019): Wang *et al.* (2017); Wu *et al.* (2018); Kaandorp and Dwight (2020))

Given the vital role RANS simulations still play in conventional and data-driven turbulence modeling, several works were done in uncertainty quantification in order to assess how RANS equations outcome varies as there are changes in the turbulent closure (Xiao and Cinnella (2019); Xiao *et al.* (2016)). A very important finding were obtained by Thompson *et al.* (2010) which used DNS databases to analyze the ability of a simplified version of RANS equations to recover the mean velocity field for the plane channel flow. To do that, the DNS RST was plugged explicitly onto the equations and the mean velocity profile was recovered. It was found that small errors presented in RST could lead to large discrepancies in the mean velocity profile. These results raised a question whether this set of equations can significantly amplify errors present in the closure to the solved mean velocity profile. In other words, if this set of equations are ill-conditioned.

Wu *et al.* (2019) developed a local metric that could asses the conditioning for the linearized form of RANS equations. Also, it was proposed an implicit treatment of RST as a strategy to mitigate the error amplification. This method involves the use of an optimized eddy viscosity as a way to treat implicitly part of RST. This optimized eddy viscosity is calculated using DNS data from the mean velocity field and the RST. This methodology, in fact, produces smaller errors in the recovered mean velocity field when compared with the fully explicit form. However, Thompson *et al.* (2016) also reported that using the DNS mean velocity field to calculate the turbulent closure leads to almost a null discrepancy in the recovered mean velocity field. This result was also explored and confirmed by Cruz *et al.* (2019) as a means to construct a more accurate data-driven turbulence model. Therefore, the methodology proposed by Wu *et al.* (2019) carries a coupling of two main factors, namely the use of the DNS mean velocity field to calculate the optimized eddy viscosity and the implicit treatment of RST, changing the discretized form of the problem. The first concerns a discrepancy reduction in the DNS data used as an input of the equations for the turbulent closure, while the second provides a new form of solving the equations that could lead to a better conditioning.

The analysis of those two factors have a paramount relevance in traditional turbulence modeling, once one could achieve less amplification of errors and, mainly, in data-driven turbulence modeling as better targets could lead to more accurate machine learning models. The present work proposes two strategies of solving RANS equations, with the aid of DNS databases for turbulent closures, in order to decouple these two highlighted factors. Hence, it possible to the see the impacts of each one of them in the accuracy of the solved mean velocity field. This analysis is conducted in the flow through a square-duct and over periodic hills at Reynolds numbers 3500 and 5600, respectively.

2. METHODOLOGY

In this section, it is presented the two methods proposed to investigate the role of the implicit treatment of RST and the use of the DNS mean velocity field to calculate the optimized eddy viscosity. However, first, it is introduced the methodology proposed by Wu *et al.* (2019) which couple both issues just mentioned and, then, the decoupled methods are presented.

2.1 Implicit treatment

Consider the steady state Reynolds-averaged Navier-Stokes equations for incompressible flows:

$$\nabla \cdot (\mathbf{uu}) - \nabla \cdot (2\nu \mathbf{D}) = \nabla p - \nabla \cdot \mathbf{R} \tag{1}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{2}$$

Where $\mathbf{u}$ is the mean velocity, p is the mean modified pressure, $\mathbf{R}$ is the RST, $\mathbf{D}$ is the rate-of-strain tensor and it is worth to mention that for incompressible flows $\nabla \cdot (2\nu \mathbf{D}) = \nu \nabla^2 \mathbf{u}$. Also, for a better understatement of this methodology, the terms placed on the right hand side of momentum equation are solved explicitly, while on the left hand side are solved implicitly.

Naturally, a turbulent closure with respect to the RST is needed to solve those equations. With the use of a DNS database, it is possible to plug onto momentum equation the DNS RST as a source term and, then, recover the mean velocity and pressure field. This test was first performed by Thompson *et al.* (2016) in the plane channel flow and similar works were done by Poroseva *et al.* (2016). This methodology of solving RANS equations is refereed in the present work as the *Fully explicit method*. Wu *et al.* (2019) proposes a different way of solving those equations with DNS data, by decomposing the RST in two parts and treating one of them implicitly. The decomposition used is accordingly to Thompson *et al.* (2010), which there is a linear and a non-linear part with respect to the strain-rate tensor. Then, we can write:

$$\mathbf{R}_{dns} = 2\nu_t \mathbf{D}_{dns} + \mathbf{R}_{dns}^{\perp} \tag{3}$$

Where ν_t is the optimized viscosity and is computed by applying the double dot product of Eq. (3) with $\mathbf{D}$, resulting in

$$\nu_t = \frac{1}{2} \frac{\mathbf{R}_{dns} : \mathbf{D}_{dns}}{\mathbf{D}_{dns} : \mathbf{D}_{dns}} \tag{4}$$

The subscript *dns* denotes the use of a quantity directly from the DNS database or indirectly calculated from them. Substituting Eq. (3) into Eq. (1), we have

$$\nabla \cdot (\mathbf{uu}) - \nabla \cdot ((2\nu + \nu_t)\mathbf{D}) = \nabla p - \nabla \cdot \mathbf{R}_{dns}^{\perp} \tag{5}$$

In Eq. (5), now, the linear part is treated implicitly, as $\mathbf{D}$ is calculated directly from $\mathbf{u}$ and is being solved together with equation. There is still an explicit part composed by the non-linear part of RST, being plugged onto equations as a source term. This part can be computed from $\mathbf{R}_{dns}^{\perp} = \mathbf{R}_{dns} - 2\nu\mathbf{D}_{dns}$. Using this method, Wu *et al.* (2019) showed an accuracy improvement in the solved mean velocity field compared to the fully explicit form of solving RANS equations. This method is referred in the present work as *Coupled implicit method*.

2.2 Decoupled Methods

The *Coupled implicit method* explained in last subsection treated implicitly the linear part of RST and, also, used the data from the DNS mean velocity field to compute the optimized eddy viscosity. As explored by Thompson *et al.* (2016) and Cruz *et al.* (2019) the use of the mean DNS velocity information can reduce the input error in the turbulent closure by using a more converged data. Therefore, it is proposed a method which only uses the information of DNS RST in the turbulent closure, but also uses an optimized turbulent eddy viscosity and an implicit treatment. To do so, the optimized eddy viscosity no longer uses the $\mathbf{D}_{dns}$, being updated from the last iteration as

$$\nu_t^{n-1} = \frac{1}{2} \frac{\mathbf{R}_{dns} : \mathbf{D}_{n-1}}{\mathbf{D}_{n-1} : \mathbf{D}_{n-1}} \tag{6}$$

Where $(n-1)$ denotes that is being solved and updated from last iteration. The mean momentum equation is then given by

$$\nabla \cdot (\mathbf{uu}) - \nabla \cdot (2(\nu + \nu_t^{n-1})\mathbf{D}) = \nabla p - \nabla \cdot \mathbf{R}_{dns}^{\perp} \tag{7}$$

In this formulation the linear part of $\mathbf{R}_{dns}$ is still solved implicitly, while the non-linear part similarly can be computed from

$$\mathbf{R}_{dns}^{\perp} = \mathbf{R}_{dns} - 2\nu_t^{n-1}\mathbf{D}_{n-1} \tag{8}$$

This method decouple the impacts of the implicit formulation from any use of DNS mean velocity field. Therefore, any gain in accuracy in the solved mean velocity field is attributed to the implicit treatment of the linear part of $\mathbf{R}_{dns}$. This method is referred in the results as DM-I (*Implicit decoupled method*).

On the other hand, it is also deduced a second decoupled method that analyze the impacts of using the DNS mean velocity field to calculate the optimized eddy viscosity in RANS solution. In order to so, the linear part of RST with respect to $\mathbf{D}$ is now solved explicitly, while the optimized eddy viscosity is computed using DNS mean velocity field information (Eq.(4)). The RANS equations are now given by:

$$\nabla \cdot (\mathbf{uu}) - \nabla \cdot (2\nu\mathbf{D}) = \nabla p - \nabla \cdot \mathbf{R}_{dns}^{\perp} + \nabla \cdot (2\nu_t\mathbf{D}) \tag{9}$$

Where the linear part of $\mathbf{R}$ with respect to $\mathbf{D}$ is on the right hand side of the equation which denotes its explicit treatment. It is worth to mention that $\mathbf{D}$ in the linear term is still updated and solved explicitly in the equation. This method investigate the role of the use of an eddy viscosity which is calculated using $\mathbf{D}$, i.e., the DNS mean velocity field. Therefore, any gain in solution accuracy is attributed to the use of the DNS mean velocity field, once there is no implicit treatment in the turbulent closure. This method will be referred in the results section as DM-E (*Explicit decoupled method*).

3. RESULTS

The methodology explained above include two different decoupled methods (DM-I and DM-E), the traditional *Fully explicit method* and the *Coupled implicit method*. These methods will be tested and analyzed for flows in two different geometries, the square-duct (SD) and the periodic hill (PH).

3.1 Square-Duct

The schematic of flow is presented in Fig. (1). The primary flow is given in the (x) axis, while a secondary flow appears in the other two directions at the corner of the duct. Most of the commonly linear eddy viscosity model can not capture this feature in the flow, being able only to simulate the primary flow. In order to apply the methods described in the methodology section for this flow, the DNS database constructed and made available by Pinelli *et al.* (2010) is used. The case analyzed correspond to the Reynolds number 3500, based on the bulk velocity at the inlet and half of the square side. As this flow offers a symmetry plane, only the third quarter is simulated (in blue in Fig. 1), setting a symmetry boundary condition at the square-duct center plane.

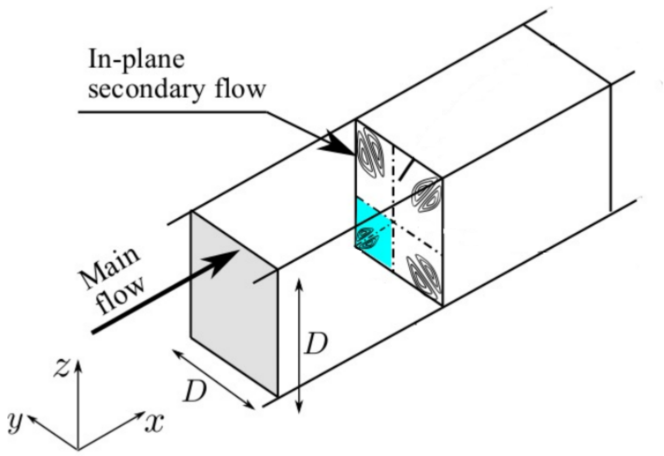


Figure 1. Schematic of the flow through a square-duct

Fig (2) and Fig (4) shows the absolute error plot for the (x) velocity component for the methods presented in the methodology section, while Fig (3) and (5) shows the same for the (y) velocity component. The error is calculated simply by $\delta U = U_{dns} - U_{solved}$ for both components. Fig (2) and (3) (b) shows the results for the DM-I compared to the traditional fully explicit approach (a) and the coupled implicit treatment proposed by Wu *et al.* (2019) (c). It is worth to remember that DM-I does not use any information from the DNS mean velocity field, but still treats implicitly the linear part of RST with respect to rate-strain tensor. The comparison shows that there is an improvement in accuracy compared to the fully explicit when there is an implicit treatment. However, the DM-I error is still closer to the fully explicit than the coupled implicit in both components. This indicates that the use of information from the DNS mean velocity field is preponderant in accuracy improvement.

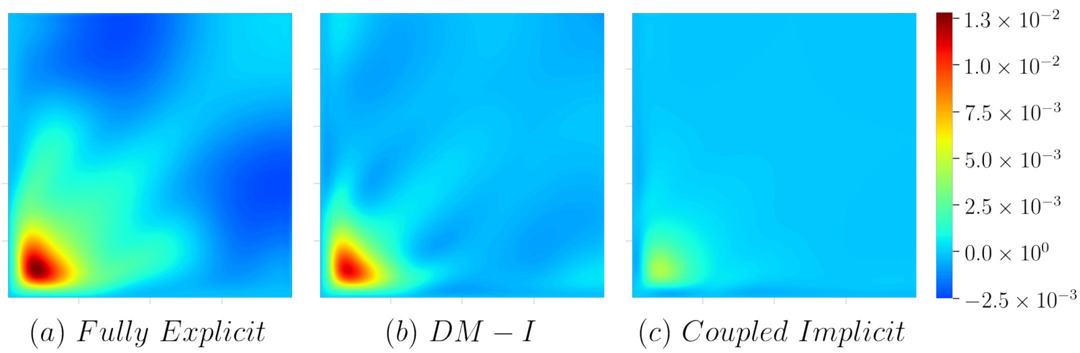


Figure 2. Absolute error for the U_x component of the mean velocity field for the flow through a square duct at $Re = 3500$. Its is compared the errors of the Fully explicit method (a), the Implicit Decoupled Method (DM-I)(b) and the Coupled implicit method proposed by Wu *et al.* (2019).

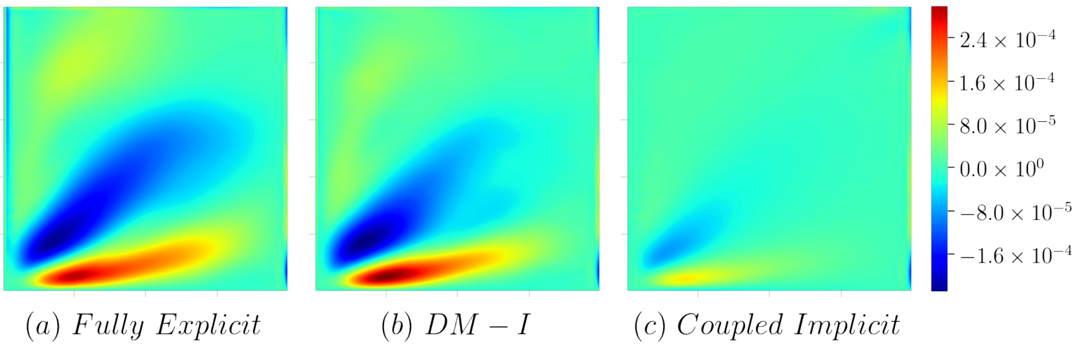


Figure 3. Absolute error for the U_y component of the mean velocity field for the flow through a square duct at $Re = 3500$. Its is compared the errors of the Fully explicit method (a), the Implicit Decoupled Method (DM-I)(b) and the Coupled implicit method proposed by Wu *et al.* (2019).

Figs. (4) and (5) shows the same plot but for DM-E which has no implicit treatment and the linear part of RST with respect to rate-strain tensor is solved explicitly. The comparison shows that DM-E is much closer in terms of accuracy to the coupled implicit than the fully explicit, even though all the terms in the closure are treated explicitly. Considering the comparisons in Figs. (2); (5) and (4); (3), it can be seen that the major factor for accuracy improvement is the use of the DNS mean velocity field information, although the implicit treatment plays a minor role. Another important aspect is that, although the gain in accuracy is timid from the fully explicit to the DM-I and even smaller from DM-E to coupled implicit, this can be attributed to a better conditioning of RANS equations. While the gain in accuracy between the fully explicit method and DM-E is due to the use of a better data as input in the turbulent closure, i.e., the DNS mean velocity field.

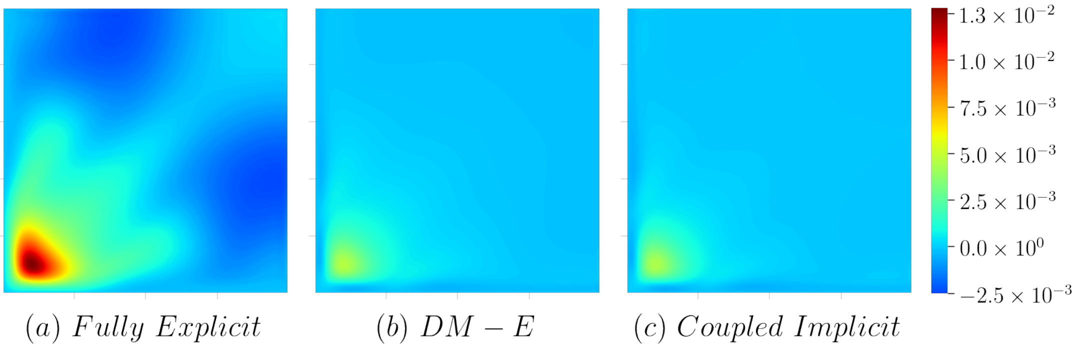


Figure 4. Absolute error for the U_x component of the mean velocity field for the flow through a square duct at $Re = 3500$. Its is compared the errors of the Fully explicit method (a), the Explicit Decoupled Method (DM-E)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).

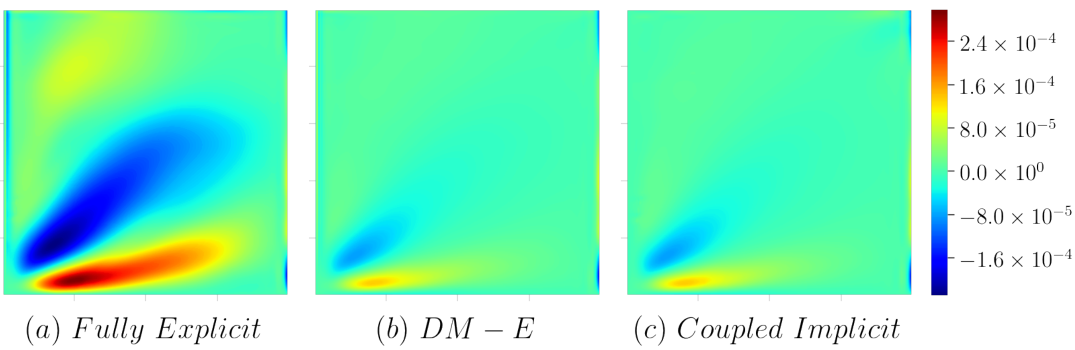


Figure 5. Absolute error for the U_y component of the mean velocity field for the flow through a square duct at $Re = 3500$. Its is compared the errors of the Fully explicit method (a), the Explicit Decoupled Method (DM-E)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).

3.2 Periodic-Hill

The same analysis is conducted in the periodic hill geometry. Fig (6) shows the schematic of the flow. An important feature of this flow is the separation caused by recirculation zone at the hill. The DNS database used for this flow was constructed and made available by Xiao *et al.* (2020). The case analyzed here correspond to $\alpha = 1$ in the database. In this case the crest-to-crest distance $L_x = 9$, $H = 1$ and $L_y = 3.036$. The Reynolds number for this flow is fixed at $Re = 5600$, being computed based on the crest height H and the bulk velocity U_b at crest. Since the flow is two-dimensional, the applied boundary conditions are given by periodic conditions in the streamwise direction (x) and non-slip condition at the walls.

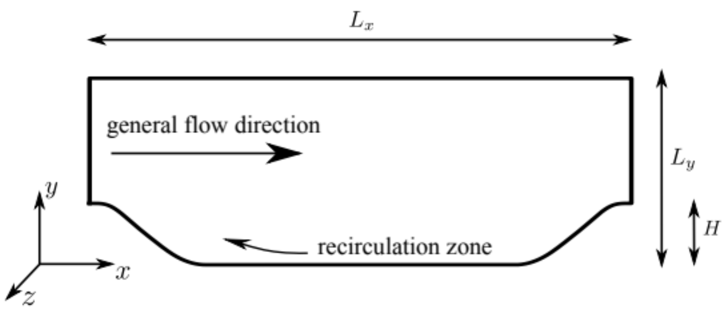
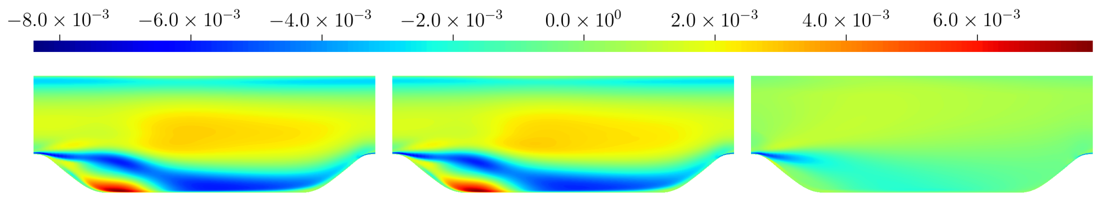


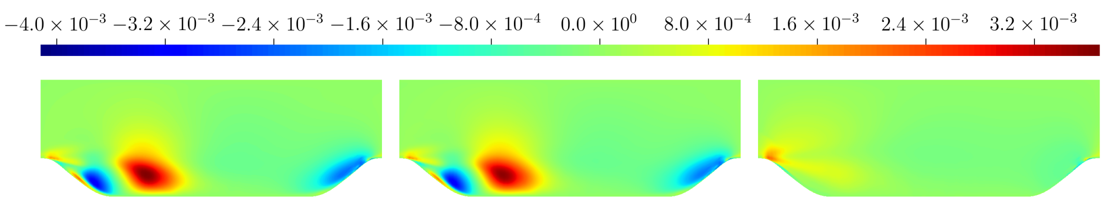
Figure 6. Schematic of the flow over periodic hills

From Fig. (7); (9) and (8);(10) can be seen analogous results in the case of flow over periodic hills for both components. The DM-I is closer again to the fully explicit method and the DM-E is closer to the coupled implicit method. However, the differences are smaller than the square-duct flow, showing that gains from the implicit treatment in terms of accuracy is less significant. Yet the analysis made in the square-duct flow is also confirmed in the periodic hill geometry, showing again a timid effect of the implicit treatment and a major role of the use of DNS mean velocity field.



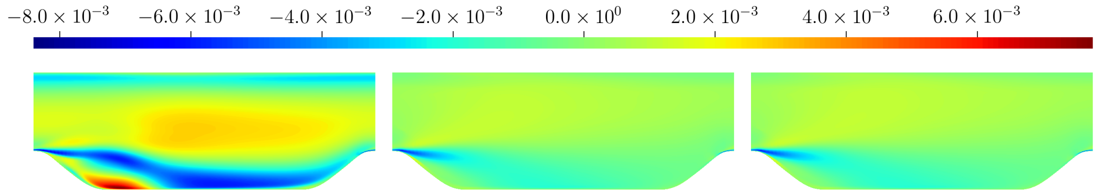
(a) *Fully Explicit* (b) $DM - I$ (c) *Coupled Implicit*

Figure 7. Absolute error for the U_x component of the mean velocity field for the flow over periodic hill at $Re = 5600$. Its is compared the errors of the Fully explicit method (a), the Implicit Decoupled Method (DM-I)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).



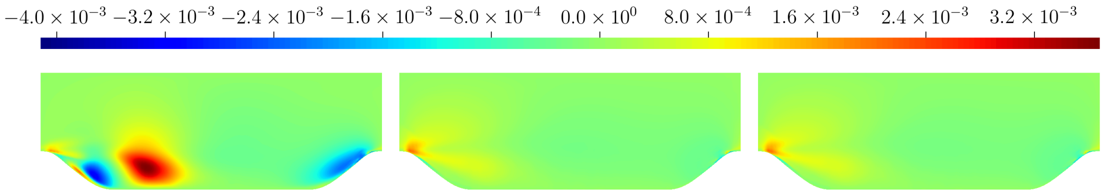
(a) *Fully Explicit* (b) $DM - I$ (c) *Coupled Implicit*

Figure 8. Absolute error for the U_y component of the mean velocity field for the flow over periodic hill at $Re = 5600$. Its is compared the errors of the Fully explicit method (a), the Implicit Decoupled Method (DM-I)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).



(a) *Fully Explicit* (b) *DM − E* (c) *Coupled Implicit*

Figure 9. Absolute error for the U_x component of the mean velocity field for the flow over periodic hill at $Re = 5600$. Its is compared the errors of the Fully explicit method (a), the Explicit Decoupled Method (DM-E)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).



(a) *Fully Explicit* (b) *DM − E* (c) *Coupled Implicit*

Figure 10. Absolute error for the U_y component of the mean velocity field for the flow over periodic hill at $Re = 5600$. Its is compared the errors of the Fully explicit method (a), the Explicit Decoupled Method (DM-E)(b) and the Coupled implicit method proposed by Wu *et al.* (2019) (c).

4. CONCLUSION

Since RANS simulations still plays a vital role for industrial applications, the present work conducted an analysis to investigate strategies to mitigate the error amplification in RANS equations with RST closures. Wu *et al.* (2019) proposed an implicit treatment of RST to obtain a better conditioning. However, as showed in the methodology, this procedure couples two main factors that has impacts in the accuracy of the recovered mean velocity field. Namely, the use of information from the DNS mean velocity field and the implicit treatment of RST. The first concerns the error reduction in the input of equations related to the turbulent closure, while the second regards a new discretized form of the problem that could lead to less error amplification. Therefore, two methods that decouple these two highlighted factors are proposed, allowing to analyze separately the impacts of each in the recovered mean velocity accuracy.

Using these two methods (DM-I and DM-E), the traditional *Fully explicit method*, and the *Coupled method* proposed by Wu *et al.* (2019), it was shown for the square-duct and the periodic-hill geometry that the implicit treatment has far less impacts in accuracy than the use of the DNS mean velocity field. The impacts of the implicit treatment were very timid in the periodic hill flow, while in the square-duct it could be easily noted. Regarding the use of information from the DNS mean velocity field, there was a great leap in accuracy in both geometries.

These results are of utmost importance for data-driven turbulence modeling, showing that in order to have better accuracy, the use of DNS mean velocity field is an important strategy. More specifically, the quantities used as targets should then be calculated as much as possible from information of DNS mean velocity field. Therefore, with targets that has much smaller discrepancies in the turbulent closure, the intrinsic propagated error to the mean velocity diminishes. Another important consequence for both data-drive and conventional turbulence modeling is the fact that it was observed a small gain in accuracy when there is an implicit treatment. This confirms partially the results obtained by Wu *et al.* (2019) that a better conditioning of RANS equations are obtained. Although the gains are not expressive, mainly for the periodic hill, any strategies that offers a less error amplification should be considered.

5. ACKNOWLEDGEMENTS

We would like to thank Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) Grant 304095/2018-4 and Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES) Grant PROEX 803/2018 for financial support.

6. REFERENCES

Basara, B. and Jakirlic, S., 2003. “A new hybrid turbulence modelling strategy for industrial cfd”. *International journal for numerical methods in fluids*, Vol. 42, No. 1, pp. 89–116.

Craft, T., Launder, B. and Suga, K., 1996. "Development and application of a cubic eddy-viscosity model of turbulence". *International Journal of Heat and Fluid Flow*, Vol. 17, No. 2, pp. 108–115.

Cruz, M.A., Thompson, R.L., Sampaio, L.E. and Bacchi, R.D., 2019. "The use of the reynolds force vector in a physics informed machine learning approach for predictive turbulence modeling". *Computers & Fluids*, Vol. 192, p. 104258.

Hamlington, P.E. and Dahm, W.J., 2008. "Reynolds stress closure for nonequilibrium effects in turbulent flows". *Physics of Fluids*, Vol. 20, No. 11, p. 115101.

Kaandorp, M.L. and Dwight, R.P., 2020. "Data-driven modelling of the reynolds stress tensor using random forests with invariance". *Computers & Fluids*, p. 104497.

Kutz, J.N., 2017. "Deep learning in fluid dynamics". *Journal of Fluid Mechanics*, Vol. 814, pp. 1–4.

Launder, B.E. and Sharma, B., 1974. "Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc". *Letters in heat and mass transfer*, Vol. 1, No. 2, pp. 131–137.

Ling, J., Jones, R. and Templeton, J., 2016a. "Machine learning strategies for systems with invariance properties". *Journal of Computational Physics*, Vol. 318, pp. 22–35.

Ling, J., Kurzawski, A. and Templeton, J., 2016b. "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance". *Journal of Fluid Mechanics*, Vol. 807, pp. 155–166.

Menter, F.R., 1994. "Two-equation eddy-viscosity turbulence models for engineering applications". *AIAA journal*, Vol. 32, No. 8, pp. 1598–1605.

Pinelli, A., Uhlmann, M., Sekimoto, A. and Kawahara, G., 2010. "Reynolds number dependence of mean flow structure in square duct turbulence". *Journal of fluid mechanics*, Vol. 644, pp. 107–122.

Poroseva, S.V., Colmenares F., J.D. and Murman, S.M., 2016. "On the accuracy of rans simulations with dns data". *Physics of Fluids*, Vol. 28, No. 11, p. 115102. doi:10.1063/1.4966639.

Speziale, C.G. and Xu, X.H., 1996. "Towards the development of second-order closure models for nonequilibrium turbulent flows". *International journal of heat and fluid flow*, Vol. 17, No. 3, pp. 238–244.

Thompson, R.L., Mompean, G. and Thais, L., 2010. "A methodology to quantify the nonlinearity of the reynolds stress tensor". *Journal of Turbulence*, , No. 11, p. N33.

Thompson, R.L., Sampaio, L.E.B., de Bragança Alves, F.A., Thais, L. and Mompean, G., 2016. "A methodology to evaluate statistical errors in DNS data of plane channel flows". *Computers and Fluids*, Vol. 130, pp. 1–7. doi: 10.1016/j.compfluid.2016.01.014. URL `https://hal.archives-ouvertes.fr/hal-01293097`.

Wang, J.X., Wu, J.L. and Xiao, H., 2017. "Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data". *Physical Review Fluids*, Vol. 2, No. 3, p. 034603.

Wilcox, D.C., 2008. "Formulation of the kw turbulence model revisited". *AIAA journal*, Vol. 46, No. 11, pp. 2823–2838.

Wu, J.L., Xiao, H. and Paterson, E., 2018. "Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework". *Physical Review Fluids*, Vol. 3, No. 7, p. 074602.

Wu, J., Xiao, H., Sun, R. and Wang, Q., 2019. "Reynolds-averaged navierâĂŞstokes equations with explicit data-driven reynolds stress closure can be ill-conditioned". *Journal of Fluid Mechanics*, Vol. 869, p. 553âĂŞ586. doi: 10.1017/jfm.2019.205.

Xiao, H. and Cinnella, P., 2019. "Quantification of model uncertainty in rans simulations: A review". *Progress in Aerospace Sciences*, Vol. 108, pp. 1–31.

Xiao, H., Wu, J.L., Wang, J.X., Sun, R. and Roy, C., 2016. "Quantifying and reducing model-form uncertainties in reynolds-averaged navier–stokes simulations: A data-driven, physics-informed bayesian approach". *Journal of Computational Physics*, Vol. 324, pp. 115–136.

Xiao, H., Wu, J.L., Laizet, S. and Duan, L., 2020. "Flows over periodic hills of parameterized geometries: A dataset for data-driven turbulence modeling from direct simulations". *Computers & Fluids*, Vol. 200, p. 104431.

7. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.