COPPE
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia
UFRJ

INFLUÊNCIA DO TRATAMENTO DA BASE DE DADOS DNS NA APLICAÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA PARA MELHORAR ACURÁCIA DE SIMULAÇÕES RANS

Victor Bitencour Rangel

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Mecânica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Mecânica.

Orientador: Roney Leon Thompson

Rio de Janeiro
Setembro de 2019

INFLUÊNCIA DO TRATAMENTO DA BASE DE DADOS DNS NA APLICAÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA PARA MELHORAR ACURÁCIA DE SIMULAÇÕES RANS

Victor Bitencour Rangel

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA MECÂNICA.

Examinada por:

Prof. Roney Leon Thompson, D.Sc.

Prof. Gustavo Rabello dos Anjos, D.Sc.

Prof. Luís Fernando Alzuguir Azevedo, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2019

Rangel, Victor Bitencour

Influência do Tratamento da Base de Dados DNS na Aplicação de Técnicas de Aprendizagem de Máquina para Melhorar Acurácia de Simulações RANS/Victor Bitencour Rangel. – Rio de Janeiro: UFRJ/COPPE, 2019.

XII, 63 p.: il.; 29, 7cm.

Orientador: Roney Leon Thompson

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Mecânica, 2019.

Referências Bibliográficas: p. 61 – 63.

1. Turbulência. 2. Aprendizado de Máquina. 3. Simulação Numérica. I. Thompson, Roney Leon. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Mecânica. III. Título.

Dedico este trabalho aos meus pais e familiares, que sempre apoiaram minhas aventuras.

Agradecimentos

Gostaria de agradecer ao professor Roney pela orientação no mestrado. Seu empenho à ciência é admirável.

Agradeço às pessoas do LabMFA, pela convivência diária, as conversas e o cafézinho. Este laboratório foi muito importante para a realização do presente trabalho.

Sou grato também aos meus colegas do grupo de pesquisa em Machine Learning, pelo conhecimento agregado ao longo desta jornada. Em especial ao meu amigo Matheus, que me mostrou por muitas vezes o caminho das pedras.

Às pessoas que moram comigo e dividem alegrias e tristezas diariamente, Bruno, Gabi e Rafa, meu muito obrigado.

À minha namorada, Amanda, imensa gratidão pelo carinho, companheirismo, e por me ajudar a manter minha sanidade.

Aos meus pais e familiares, aos quais dedico este trabalho, muito obrigado. O esforço e preocupação de vocês para com o meu sucesso jamais serão em vão.

Por último, mas não menos importante, a mim mesmo, por ter apostado em minha capacidade e jamais ter abandonado o que me propus a fazer.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

INFLUÊNCIA DO TRATAMENTO DA BASE DE DADOS DNS NA APLICAÇÃO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA PARA MELHORAR ACURÁCIA DE SIMULAÇÕES RANS

Victor Bitencour Rangel

Setembro/2019

Orientador: Roney Leon Thompson

Programa: Engenharia Mecânica

Das técnicas de soluções numéricas para escoamentos turbulentos, Reynolds Average Navier-Stokes (RANS) realiza simulações de baixo custo e respostas rápidas, porém baixa acurácia de resultados. Large Eddy Simulation (LES) e Direct Numerical Simulation (DNS) privilegiam a acurácia, mas elevam demasiadamente o custo computacional, tornando proibitivo utilizá-las na indústria.

A utilização de aprendizado de máquina para suprimir a deficiência de RANS vem recebendo recente atenção. Tal abordagem visa fazer com que RANS aprenda com resultados mais nobres, como os de DNS. A principal entidade empregada na correção de escoamentos RANS na literatura é o tensor das tensões de Reynolds ($\mathbf{R}$). Porém, recentes estudos indicam que $\mathbf{R}$ não apresenta boa convergência nas bases de dados DNS disponíveis quando comparado aos demais campos. Surgiu então a metodologia onde se corrige o divergente modificado de $\mathbf{R}$, denominado aqui de vetor $\mathbf{t}$, para contornar incertezas geradas pela correção de $\mathbf{R}$. Esta mostrou resultados promissores, por vezes melhores que os da correção de $\mathbf{R}$.

Este trabalho mostra a influência da qualidade da base de dados DNS na previsão de uma rede neural (RN) para as correções de $\mathbf{R}$ e de $\mathbf{t}$, para o escoamento turbulento em duto de seção quadrada. Dados DNS são tratados a fim de simular melhor convergência dos mesmos. Resultados mostram que, quão melhor convergida é a base DNS, melhores são os resultados da RN. A simetria dos dados tem grande influência no desempenho da RN, a qual performou melhor quando imposta a simetria total dos dados, para ambas metodologias de correção. Constatou-se que a metodologia $\mathbf{t}$ apresenta erros globais menores que a $\mathbf{R}$ na direção principal do escoamento, mas não nas secundarias, o que coloca em discussão a utilização da mesma.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

DNS DATABASE PRE PROCESSING INFLUENCE IN MACHINE LEARNING TECHNIQUES FOR ACCURACY IMPROVEMENT OF RANS SIMULATIONS

Victor Bitencour Rangel

September/2019

Advisor: Roney Leon Thompson

Department: Mechanical Engineering

Among numerical solution techniques for turbulent flow, Reynolds Average Navier-Stokes (RANS) presents low computational and time costs, but also lower accuracy. Large Eddy Simulation (LES) and Direct Numerical Simulation (DNS) present greater accuracy, but both require much more computational effort and simulation time, making them impracticable in daily industry problem solution.

The utilization of machine learning (ML) to help improve RANS results is receiving some attention lately. Such approach makes RANS simulations achieve better results by learning them from more noble ones, as DNS simulation for example.

The main entity applied in RANS correction is the Reynolds stress tensor ($\mathbf{R}$). Although, recent studies show that $\mathbf{R}$ does not present well converged results from available DNS data base when compared to other fields convergence. Another entity is introduced in ML techniques in order to bypass uncertainties generated by the correction of $\mathbf{R}$, it is the modified divergence of $\mathbf{R}$ which is called here as vector $\mathbf{t}$. The correction of $\mathbf{t}$, when compared to $\mathbf{R}$ methodology, has shown promising results, even presenting better results than the ones achieved by correcting $\mathbf{R}$.

The present work studies how DNS data base quality influences on neural network (NN) prediction for both aforementioned methodologies: correction of $\mathbf{R}$ and $\mathbf{t}$, for square duct turbulent flow. DNS data are manipulated in order to simulate better convergence. Results show that better converged DNS data leads to better NN predictions, which has presented its best results when total data symmetry was imposed for both correction methodologies. It was also observed that correction of $\mathbf{t}$ presents lower global errors for the main direction flow when compared to $\mathbf{R}$ correction, but not for the secondary direction flows, which puts such methodology in doubt about its advantages in relation to $\mathbf{R}$ correction.

Sumário

Lista de Figuras **x**

Lista de Tabelas **xii**

1 Introdução **1**

2 Revisão bibliográfica **4**

2.1 Turbulência 4

2.1.1 Vórtices e a cascata de energia 5

2.1.2 Equação de Navier-Stokes 6

2.1.3 Turbulência e simulação 7

2.1.4 RANS e o problema do fechamento 8

2.1.5 Modelo $\kappa - \epsilon$ 9

2.2 Aprendizado de máquina 9

2.2.1 Redes Neurais 9

2.2.2 Bias 11

2.2.3 Rede Neural Feedforward 11

2.2.4 Backpropagation 12

2.2.5 Validação e teste 13

2.3 AM na simulação de escoamentos turbulentos 14

2.4 A falta de convergência nas bases de dados DNS 19

2.5 Escoamento em duto quadrado 22

2.5.1 Base de dados para treinamento 25

2.5.2 Malha de simulação 26

3 Metodologia **27**

3.1 Uso de RN na correção de RANS 27

3.1.1 Treinando a RN 27

3.1.2 Utilizando a RN 29

3.2 Tratamento da base de dados DNS 31

3.2.1 Simulando dados DNS melhor convergidos 32

4 Resultados **37**
4.1 Previsão dos campos de velocidade 37
4.1.1 Correção de **R** 37
4.1.2 Correção de **t** 40
4.2 Erro Local 41
4.2.1 Correção de **R** 41
4.2.2 Correção de **t** 47
4.3 Erro Global 51
4.3.1 Q1 51
4.3.2 Q2 53
4.3.3 Q3 55
4.3.4 Q4 56

5 Conclusões **59**

6 Trabalhos Futuros **60**

Referências Bibliográficas **61**

Lista de Figuras

2.1 Formação de vórtices em escoamento turbulento, por Leonardo da Vinci, MONAGHAN e KAJTAR (2014) 5
2.2 Esquema ilustrado da cascata de energia, adaptado de DAVIDSON (2004) 6
2.3 Modelo de neurônio artificial de McCulloch e Pitts, MARSLAND (2009) 10
2.4 RN utilizando dados de entrada Bias, MARSLAND (2009) 12
2.5 Camada escondida em uma RN Multi-layer Perceptron, adaptado de MARSLAND (2009) 13
2.6 Campos referentes à parte anisotrópica de **R**, adaptado de TRACEY *et al.* (2013) 14
2.7 Comparação das previsões utilizando dados de treinamento de: (a) canais acionados por pressão; (b) solução de três placas planas. TRACEY *et al.* (2015) 15
2.8 Perfil do segundo invariante anisotrópico para (a) LES, (b) RANS e (c) previsão da FR, LING *et al.* (2016c) 16
2.9 Erro como função do número de rotações de treino, exceto para os dados invariantes, LING *et al.* (2016b) 17
2.10 Previsão do tensor de Reynolds anisotrópico, LING *et al.* (2016a) . . 18
2.11 Campos secundarios de velocidade média, WANG *et al.* (2016) 19
2.12 Previsão da tensão de cisalhamento para dados de treinamento de: (a) mesma geometria, Re diferentes; (b) geometrias e Re diferentes, WANG *et al.* (2017) 20
2.13 Campos de velocidade média na direção principal do escoamento, CRUZ (2018) 21
2.14 Perfil de velocidade a partir da previsão com método implícito, WU *et al.* (2018) 22
2.15 Comparação de erros residuais para diversos Re e bases DNS, ANDRADE *et al.* (2018) 23
2.16 Comparação de erros residuais para diversos tempos de simulação DNS, ANDRADE *et al.* (2018) 24

2.17 Esquema do escoamento em duto de seção quadrada, adaptado de WANG *et al.* (2016) 25
2.18 Condições de contorno da simulação em duto de seção quadrada, CRUZ (2018) 25
2.19 Malha da simulação em duto de seção quadrada, CRUZ (2018) 26

3.1 Campos DNS para Re = 2200 32
3.2 Campos de discrepancia de U_x para Re = 2200 33
3.3 Campo R_{xy}, Re = 2200 34
3.4 Campo R_{xz}, Re = 2200 35
3.5 Evolução do campo DNS R_{yz} sob tratamento de dados 36

4.1 Previsão de U_x para $Re = 3200$ através do método **R** 38
4.2 Previsão de U_y para $Re = 3200$ através do método **R** 39
4.3 Previsão de U_z para $Re = 3200$ através do método **R** 40
4.4 Previsão de U_x para $Re = 3200$ através do método **t** 42
4.5 Previsão de U_y para $Re = 3200$ através do método **t** 43
4.6 Previsão de U_z para $Re = 3200$ através do método **t** 44
4.7 Erro local da previsão de U_x para $Re = 3200$ através do método **R** . . 45
4.8 Erro local da previsão de U_y para $Re = 3200$ através do método **R** . . 46
4.9 Erro local da previsão de U_z para $Re = 3200$ através do método **R** . . 47
4.10 Erro local da previsão de U_x para $Re = 3200$ através do método **t** . . 48
4.11 Erro local da previsão de U_y para $Re = 3200$ através do método **t** . . 49
4.12 Erro local da previsão de U_z para $Re = 3200$ através do método **t** . . 50
4.13 Erro global da previsão de **R** do grupo Q1 para $Re = 3200$ 51
4.14 Erro global da previsão de **t** do grupo Q1 para $Re = 3200$ 52
4.15 Erro global da previsão de **R** do grupo Q2 para $Re = 3200$ 53
4.16 Erro global da previsão de **t** do grupo Q2 para $Re = 3200$ 54
4.17 Erro global da previsão de **R** do grupo Q3 para $Re = 3200$ 55
4.18 Erro global da previsão de **t** do grupo Q3 para $Re = 3200$ 56
4.19 Erro global da previsão de **R** do grupo Q4 para $Re = 3200$ 58
4.20 Erro global da previsão de **t** do grupo Q4 para $Re = 3200$ 58

Lista de Tabelas

2.1 Viscosidades de simulação 26

3.1 Dados de entrada da RN 28
3.2 Combinações de quadrantes 35

Capítulo 1

Introdução

Das técnicas de soluções numéricas para problemas associados a escoamentos turbulentos, de uma forma geral, tem-se Reynolds Average Navier-Stokes (RANS), Large Eddy Simulation (LES) e Direct Numerical Simulation (DNS). A primeira, quando comparada às demais, realiza simulações de baixo custo e respostas rápidas. Entretanto, é sabido que seus resultados deixam a desejar quanto a acurácia. Por outro lado, as abordagens LES e DNS privilegiam a acurácia, mas elevam demasiadamente o custo computacional e temporal, tornando-as inviáveis principalmente quando analisadas aos olhos da indústria.

A utilização de aprendizado de máquina (AM) para suprimir a deficiência de RANS vem recebendo recente atenção. Tal abordagem visa a fazer com que RANS aprenda com resultados mais nobres, como LES, DNS e experimentos.

A principal Quantidade de Interesse (QdI) empregada na correção de escoamentos RANS na literatura é o tensor das tensões de Reynolds ($\mathbf{R}$). As técnicas de AM empregadas aprendem com dados nobres valores de $\mathbf{R}$ para diversos casos através de treinamento e, a partir de dados de entrada provenientes de simulação RANS, conseguem predizer um novo $\mathbf{R}$ para ela, onde este $\mathbf{R}$ predito é melhor que o original, e se aproxima de valores DNS. Após isso, o novo $\mathbf{R}$ é utilizado em nova simulação onde ele ou seu divergente entra como termo fonte na equação de balanço de quantidade de movimento, para que sejam obtidos novos campos melhorados.

TRACEY *et al.* (2013) utilizou AM como uma ferramenta de quantificação de incerteza através do desenvolvimento de um modelo estocástico, onde conseguiu obter melhores resultados de simulação para o periodic hill, clássico escoamento turbulento que apresenta descolamento e recolamento da camada limite, característica que RANS apresenta dificuldade em capturar, através da predição de um campo turbulento anisotrópico corrigido por AM, a partir de dados de simulação RANS $\kappa-\omega$ SST, tendo como alvo dados DNS. Mais tarde, TRACEY *et al.* (2015) propôs a utilização de AM como ferramenta na criação de um novo modelo de turbulência que consiste de uma rede neural (RN) treinada em uma base de dados do modelo RANS Spalart-

Allmaras, reproduzindo com fidelidade seus resultados, inclusive sendo capaz de mostrar características do escoamento que não haviam sido diretamente treinadas, o que da ideia da capacidade da aplicação de AM.

LING *et al.* (2016b) compara resultados AM para duas abordagens, o treinamento de dados sem nenhum tipo de tratamento, e o treinamento com dados invariantes. O apelo de tal abordagem consiste em permitir que modelos AM possam prever campos para escoamentos semelhantes sem que estes estejam sob a mesma orientação física. Resultados mostraram que os modelos treinados a partir de uma base já invariante foram melhores em suas previsões, além de o fazerem utilizando menos capacidade computacional. Já em LING *et al.* (2016a), propôs o uso de deep learning com invariância Galileana embutida para prever a parte anisotrópica do tensor de Reynolds, e comparou os resultados desse modelo com uma RN de dados crus, sem invariância. O modelo proposto se saiu melhor que seus concorrentes, o que exalta a importância da invariância dos dados de treinamento.

Até então, os estudos aqui mencionados se preocupavam em melhorar a caracterização de $\mathbf{R}$ em RANS, entretanto os resultados para o campo de velocidade média não eram tão bons. Investigando os efeitos da correção de RANS através de AM na obtenção de um melhor campo de velocidade, WANG *et al.* (2016) propôs um modelo informado fisicamente, onde previa discrepância do tensor de Reynolds RANS em relação ao DNS através de entidades invariantes coletadas de RANS. Os resultados mostraram melhora do perfil de velocidade para o caso escoamento em duto de seção quadrada. Mais tarde, WANG *et al.* (2017) propõe a utilização de floresta randômica (FR) em duas abordagens. Em uma, ele treina e aplica a FR para um mesmo tipo de escoamento, e na segunda tenta a correção de um escoamento com separação de camada limite a partir da FR treinada com outros escoamentos, semelhantes porém não iguais geometricamente. Os resultados mostram considerável melhora em todos os métodos, porém no segundo caso a melhora não surpreende, o que leva a entender que a extrapolação de dados pode não ser uma abordagem adequada.

THOMPSON *et al.* (2016) apontou que certas bases de dados DNS possuíam uma incerteza intrínseca quanto ao tensor de Reynolds, pois estes possuíam convergência menor que os demais campos médios DNS disponibilizados de uma mesma simulação, como velocidade e pressão médios. Sabendo disso, CRUZ (2018) propôs o uso de RN na correção do divergente modificado de $\mathbf{R}$, chamado por ele de $\mathbf{t}$, para o escoamento em duto de seção quadrada. O tensor $\mathbf{t}$ pode ser calculado a partir de dados de primeira ordem estatística, como os campos de velocidade e pressão médios DNS, contornando dessa forma a incerteza presente em $\mathbf{R}$. Os resultados apresentados pelo método $\mathbf{t}$ se mostraram superiores aos da correção de $\mathbf{R}$ principalmente para o campo de velocidade médio na direção principal. Ainda sobre melhorar a predição do campo

de velocidade médio, WU *et al.* (2018) propôs modelo implícito e invariante de uma RN treinada com dados de um escoamento periodic hill, de vale mais íngreme, que consegue prever com boa acurácia as QdI de um escoamento periodic hill de vale menos íngreme, dando nova expectativa com relação a extrapolação de dados por AM.

É notório que a falta de convergência de $\mathbf{R}$ nas bases de dados de simulação DNS apontada por THOMPSON *et al.* (2016) influencia negativamente os resultados da aplicação de AM na correção de RANS, principalmente quando analisa-se a propagação do campo de velocidade média axial à direção do escoamento. Não por acaso essa QdI é por vezes omitida nos artigos que estudam aplicação de AM em turbulência. Porém, é possível ver com clareza em CRUZ (2018) que, quando a correção é feita para entidades melhor convergidas, os resultados obtidos são melhores e congruentes fisicamente. ANDRADE *et al.* (2018) propõe análise de incerteza de dados DNS levando em conta metodologia apontada por THOMPSON *et al.* (2016), através de métricas de erro relacionadas tanto a convergência do tensor de Reynolds quanto dos campos de velocidade. Em seus resultados, fica claro que a deficiência nas bases DNS é causada pelo descaso com relação ao tempo de simulação.

O presente trabalho visa estudar a influência da convergência dos dados DNS na aplicação de AM para correção de simulações RANS, através da manipulação dos campos disponibilizados por PINELLI *et al.* (2010) a fim de simular tempos maiores de simulação, e portanto melhor convergência dos mesmos. Os campos manipulados são então utilizados tanto para a correção de $\mathbf{R}$ quanto de $\mathbf{t}$.

Capítulo 2

Revisão bibliográfica

2.1 Turbulência

Nos ramos da física e da engenharia, o estudo do comportamento dos fluidos, quando em escoamento, é chamado de fluidodinâmica. Os escoamentos são classificados com base em propriedades tanto do fluido como do escoamento em si.

O número de Reynolds (Re), calculado como mostra a Eq. 2.1, é um adimensional que mede o balanço entre as forças inerciais e viscosas do escoamento. É uma relação entre os termos advectivo e difusivo do balanço da equação de momento linear. Entre suas utilidades estão: predição da transição entre tipos de escoamento, laminar ao turbulento; dimensionamento de experimentos em escala.

$$Re = \frac{UL}{\nu} \tag{2.1}$$

Onde U e L são grandezas características do escoamento para velocidade e comprimento, respectivamente, e ν é a viscosidade cinemática do fluido.

Em seus experimentos, Reynolds concluiu que a turbulência não está diretamente ligada ao Re, mas sim com perturbações no escoamento. Em ambiente controlado, é possível fazer com que um escoamento laminar chegue a números de Re da ordem de 10^4.

Ele também estudou a ligação entre o Re e a turbulência induzida. Quanto a isto, sua conclusão foi que, por volta de Re = 2000, a turbulência tende a se dissipar, DAVIDSON (2004).

Um escoamento laminar totalmente desenvolvido é normalmente caracterizado por apresentar baixo Re, linhas de corrente bem definidas e velocidade em função apenas da posição, onde dominam as forças viscosas, POPE (2000).

Já um escoamento turbulento totalmente desenvolvido é normalmente caracterizado por apresentar alto Re, comportamento caótico de suas partículas de fluido, além de velocidade variante no tempo, dominado por forças inerciais responsáveis

pelo aparecimento de um fenômeno peculiar a tal escoamento, a formação de vórtices ou turbilhões (do inglês eddies, como é conhecido na literatura), POPE (2000).

O comportamento randômico do escoamento turbulento, seu caráter tridimensional juntamente com sua capacidade de transporte de momento e calor compõe uma propriedade chamada de difusividade. Por conta dessa característica, esse tipo de escoamento é bastante aplicado na mistura de materiais, como na industria química ou em motores a combustão, POPE (2000).

2.1.1 Vórtices e a cascata de energia

O aparecimento de vórtices é algo característico a escoamentos turbulentos. Apesar de na literatura ainda não haver uma definição objetiva e definitiva do que é um vórtice, é um fenômeno facilmente observado nesse tipo de escoamento, como Leonardo da Vinci, aproximadamente 500 anos atrás, pôde mostrar em seu desenho que pode ser visto na Fig. 2.1.

Figura 2.1: Formação de vórtices em escoamento turbulento, por Leonardo da Vinci, MONAGHAN e KAJTAR (2014)

Um vórtice, ou turbilhão, geralmente surge acompanhado de vórtices menores, onde os maiores transferem energia para os menores, e esses para outros ainda menores, até que se chegue a menor escala possível, onde a energia passa a ser dissipada

através de ações viscosas. Tal escala é conhecida como escala de Kolmogorov, POPE (2000).

A ideia de transferência de energia entre os diferentes tamanhos de turbilhão foi introduzida pela primeira vez por Richardson e é conhecida como cascata de energia, RICHARDSON (2007).

A Fig. 2.2 mostra um esquema ilustrado da cascata de energia. Os vórtices maiores são criados por instabilidades presentes no escoamento médio. Por serem afetados por instabilidades inerciais, esses vórtices se dividem em outros menores, que também sofrem influência das mesmas instabilidades, e também se dividem em outros ainda menores. Com isso, fica elucidada outra característica deste tipo de escoamento, que é seu caráter dissipativo.

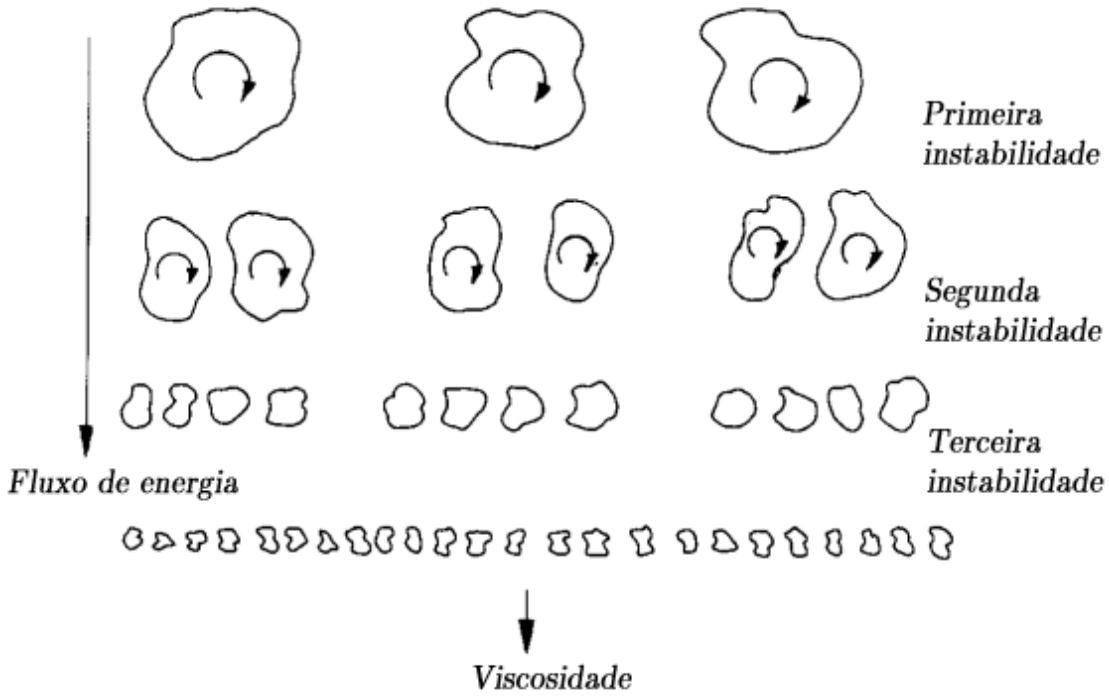


Figura 2.2: Esquema ilustrado da cascata de energia, adaptado de DAVIDSON (2004)

2.1.2 Equação de Navier-Stokes

A equação de Navier-Stokes (NS) é uma equação diferencial parcial obtida através da aplicação de três princípios físicos a uma partícula de fluido, são eles: a segunda lei de Newton para o contínuo; a lei da viscosidade de Newton; e o princípio da conservação de massa. Esta equação tem a capacidade de descrever o movimento da maioria dos fluidos, DAVIDSON (2004), e pode ser vista na Eq. 2.2.

$$\frac{D\mathbf{u}}{Dt} = \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla\left(p\right) + \nu\nabla^2\mathbf{u} \tag{2.2}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{2.3}$$

O termo à esquerda da igualdade na Eq. 2.2 se refere a derivada material, também conhecida como derivada convectiva ou derivada substancial, a qual carrega informações sobre a taxa de variação de uma determinada propriedade de um elemento de fluido no tempo, e não de um ponto no espaço. Ou seja, esta acompanha a trajetória de tal partícula. Sendo assim, esse termo contém informações inerciais do escoamento e pode ser decomposto em outros dois termos, são eles os termos de aceleração local e advectiva.

Já à direita da igualdade na Eq. 2.2 está o termo difusivo, onde ν é o coeficiente de viscosidade cinemática, definido através da Eq. 2.4, na qual μ é a viscosidade dinâmica e ρ a massa específica.

Na Eq. 2.3 temos a conservação de massa.

$$\mathrm{v} \equiv \mu/\rho \tag{2.4}$$

O que torna a equação de NS tão complexa é a presença do termo quadrático em $\mathbf{u}$, este é $(\mathbf{u} \cdot \nabla)\mathbf{u}$. É sabido que nos escoamentos turbulentos, seu comportamento caótico é proveniente da natureza aleatória do escoamento, efeito de sua sensibilidade às inevitáveis perturbações tanto na condição inicial quanto nas condições de contorno, além do alto Re e propriedades do material, POPE (2000). O termo quadrático amplifica qualquer perturbação, por menor que seja.

Dito isso, se um experimento pudesse ser repetido diversas sob as mesmas condições, e levando em consideração que as perturbações fossem zero, ainda assim seria impossível, dado um ponto fixo no espaço, se obter os mesmos resultados em todos os experimentos para qualquer quantidade de interesse (QdI) instantânea que fosse escolhida, POPE (2000).

2.1.3 Turbulência e simulação

A equação de NS descreve o movimento de um fluido newtoniano. Para casos simples, onde não há turbulência ou a viscosidade do fluido é alta, ela possui solução conhecida. Entretanto, para casos mais complexos, sua solução só pode ser encontrada com a ajuda de métodos numéricos. Computational fluid dynamics (CFD), ou dinâmica dos fluidos computacional, é o nome do ramo que estuda soluções numéricas de NS, assim como de escoamentos não newtonianos, os quais não são regidos por NS.

Direct numerical simulation (DNS), ou simulação numérica direta, é um tipo de simulação CFD que resolve as equações de NS e da continuidade para escoamentos turbulentos sem o uso de nenhum tipo de modelo de turbulência. Para tanto, é necessário que a malha de tal simulação esteja na ordem da escala de Kolmogorov, para que a técnica seja capaz de capturar todas as escalas do escoamento, WILCOX (2006). Dito isso, realizar essa técnica para escoamentos com alto Re ou em geometrias complicadas geralmente requer grande capacidade computacional e tempo de simulação.

A aplicação de modelos de turbulência vem para tentar diminuir o problema do ponto de vista do custo computacional.

Large eddy simulation (LES), ou simulação de grandes escalas, é um tipo de simulação CFD que resolve numericamente as escalas maiores do escoamento, e modela as menores (conhecidas como subgrid-scales). Isso é feito pois considera-se que as escalas maiores são mais afetadas pelas condições de contorno, diferentemente das escalas menores, WILCOX (2006). Por resolver apenas uma parte do escoamento, uma simulação LES pode ter malha com resoluções espaciais muito maiores que a de Kolmogorov, o que o capacita a resolver escoamentos mais rapidamente que DNS para um mesmo Re, abrindo mão de certa qualidade. A escala de tempo de suas simulações é menor que DNS, porém ainda fora do escopo para soluções rápidas.

Outro método para simular um escoamento é fazê-lo para o escoamento médio, e não para o instantâneo. Tal método é obtido da aplicação do operador média de Reynolds à equação de NS, originando a Reynolds Averaged Navier-Stokes equation (RANS), ou equação média de Reynolds, que passa a descrever o escoamento através de sua velocidade média em vez da instantânea, o que é suficiente para diversos problemas ligados à engenharia. Quando comparada às simulações DNS e LES, a RANS apresenta resultados inferiores, porém necessitam de capacidade computacional muito inferior, além de serem realizadas em menos tempo, o que a torna principal escolha na solução de problemas cotidianos.

2.1.4 RANS e o problema do fechamento

Quando o operador média é aplicado à equação de NS para se obter RANS, um novo termo surge do termo advectivo como resultado dessa aplicação. Este novo termo é o tensor covariante de $\mathbf{u}$, que está relacionado ao segundo momento estatístico do campo de velocidade, também conhecido como tensor de Reynolds ($\mathbf{R}$).

Para escoamentos incompressíveis e de propriedades constantes, $\mathbf{R}$ é um tensor simétrico que, ao surgir, adiciona seis variáveis extras ao problema, fazendo com que o sistema de equações do mesmo se torne indeterminado. Este fato é conhecido

como o problema do fechamento, WILCOX (2006).

2.1.5 Modelo $\kappa - \epsilon$

Existe uma boa variedade de modelos de turbulência, levando a entender que nenhum é absoluto ou definitivo. Os modelos de duas equações são os mais popularmente utilizados na industria. Esse tipo de modelo é conhecido por ser bastante poderoso, capaz de predizer propriedades do escoamento turbulento sem nenhuma informação ou conhecimento prévio da estrutura turbulenta, WILCOX (2006).

Por conta do apelo industrial, o modelo de duas equações conhecido como modelo $\kappa - \epsilon$ foi escolhido para ser utilizado no presente trabalho.

O modelo $\kappa - \epsilon$ adiciona duas variáveis, portanto duas equações de transporte são adicionadas ao problema, são elas: energia cinética turbulenta (κ); taxa de dissipação de energia turbulenta (ϵ).

Apesar de ser um bom modelo para o escoamento médio, $\kappa - \epsilon$ carece em solução quando próximo a paredes ou superfícies sólidas. Para tanto, funções de parede são utilizadas para contornar tal problema.

Funções de parede são equações algébricas baseadas em propriedades do escoamento e leis de parede que calculam soluções do campo de velocidade em regiões próximas à parede. Essas soluções são impostas baseado no que é esperado dessas regiões a partir de dados de fontes mais nobres.

2.2 Aprendizado de máquina

Aprendizado de máquina (AM) é a área de conhecimento que estuda o desenvolvimento de inteligência artificial (IA), capaz de solucionar problemas ou fazer predições sem nenhuma programação especifica prévia para tal, baseado apenas na análise de um banco de dados, os chamados dados de treinamento.

Algumas técnicas famosas de AM são: florestas randômicas; algoritmo genético; árvore de decisões. Para o presente trabalho, a técnica escolhida para aplicação foi a de redes neurais (RN), por ter apresentado resultados bastante promissores em trabalhos da área em questão, LING *et al.* (2016a), além deste ser um estudo que propõe melhoria de trabalhos anteriores, como o de CRUZ (2018).

2.2.1 Redes Neurais

O conceito de redes neurais (RN) vem do funcionamento do próprio cérebro humano. É sabido quão poderosa é essa parte do corpo quanto ao processamento de dados, tendo capacidade de abstrair, através de seus quase cem bilhões de neurônios, respostas provenientes de estímulos por vezes ruidosos, o que significa que o cérebro

funciona como um computador de cem bilhões de processadores trabalhando em paralelo, MARSLAND (2009).

Neurônios reais trabalham recebendo estímulos (sinais elétricos) através dos dendritos quando ocorrem sinapses. Após processar os sinais recebidos, envia novos estímulos a outras células através dos axônios. Este processo acontece com todos os neurônios, e se as outras células mencionadas no processo também são neurônios, então esta é uma rede neural.

Neurônios artificiais tentam simular a mesma função dos reais. O primeiro neurônio artificial foi proposto por McCulloch e Pitts, MARSLAND (2009), e pode ser visto na Fig. 2.3, modelado como um conjunto de dados de entrada (x ponderados através de pesos w), correspondendo a sinapses (x_1w_1, x_2w_2,...,x_mw_m), seguido de um somador que processa os sinais recebidos através de uma função ativadora (θ), a qual filtra o resultado deste processamento para enviar um sinal de saída (output) à próxima célula, MARSLAND (2009), como acontece em axônios reais.

Os pesos são utilizados para tornar certos dados de entrada mais relevantes que outros, como quando algumas sinapses costumam mandar sinais mais fortes que outras a um neurônio.

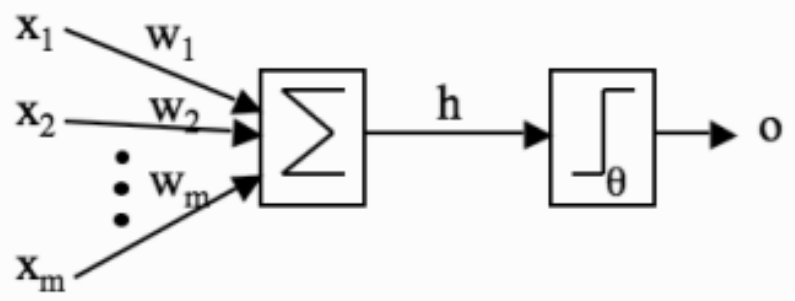


Figura 2.3: Modelo de neurônio artificial de McCulloch e Pitts, MARSLAND (2009)

Uma RN é formada a partir do agrupamento de neurônios, como o apresentado, em uma mesma camada, recebendo os mesmos dados de entrada, mas operando de maneira diferente, e fazendo com que seus outputs se encontrem em um passo posterior.

As funções de ativação são as responsáveis por fazer uma RN mapear os dados de entrada X nos alvos Y. As mais populares são as funções: limiar (Eq. 2.5), linear (Eq. 2.6), sigmoide (Eq. 2.7) e a função tangente hiperbólica (Eq. 2.8). As duas últimas citadas são responsáveis por tornar uma RN capaz de realizar mapeamentos não lineares.

$$\varphi\left(in_j^{(m)}\right) = \begin{cases} 1, in_j^{(m)} \geq 1 \\ 0, in_j^{(m)} < 1 \end{cases} \tag{2.5}$$

$$\varphi\left(in_j^{(m)}\right) = k \times in_j^{(m)} \tag{2.6}$$

$$\varphi\left(in_j^{(m)}\right) = \frac{1}{1+\exp\left(-in_j^{(m)}\right)} \tag{2.7}$$

$$\varphi\left(in_j^{(m)}\right) = \frac{1-\exp\left(-in_j^{(m)}\right)}{1+\exp\left(-in_j^{(m)}\right)} \tag{2.8}$$

2.2.2 Bias

Outra ferramenta utilizada para controlar melhor o comportamento de uma RN é chamada bias. Ela é um dado de entrada de valor fixo introduzido na camada de pesos, e não é influenciada pelos dados de entrada reais, mas influencia nos somadores na intenção de evitar situações adversas, como por exemplo quando todos os dados de entrada são zero, o que poderia fazer a função de ativação não ser acionada pelos neurônios, MARSLAND (2009).

A Eq. 2.9 mostra como os dados de entrada, pesos e bias funcionam na geração de dados de saída dos neurônios, que são uma somatória dos dados de entrada, ponderados pelos pesos, e dos bias. A Fig. 2.4 mostra bias sendo utilizada em um esquema de RN, onde este valor entra no somatório.

$$out_j^{(m)} = \varphi\left(in_j^{(m)}\right) = \varphi\left(\sum_{i=1}^{N} w_i^{(m)} out_i^{(n)} + b_j^{(m)}\right) \tag{2.9}$$

2.2.3 Rede Neural Feedforward

A estrutura básica de uma RN é composta de uma camada de dados de entrada no inicio, uma de dados de saída no final, e uma ou mais camadas de neurônios entre as duas primeiras citadas. Essas camadas de neurônios são conhecidas como camadas escondidas.

As camadas escondidas são compostas de múltiplos neurônios. Cada hidden layer está conectada às camadas anterior e posterior, como visto na Fig. 2.5.

Uma rede neural feedforward é uma RN onde os outputs de cada camada não retroalimentam camadas anteriores.

Quando as camadas escondidas de uma RN do tipo feedforward são compostas por neurônios não lineares, essa RN é conhecida na literatura como multi layer perceptron (MLP).

As características de uma RN dependem do projeto. Quanto mais complexo e não linear for, mais camadas escondidas e neurônios não lineares serão necessários.

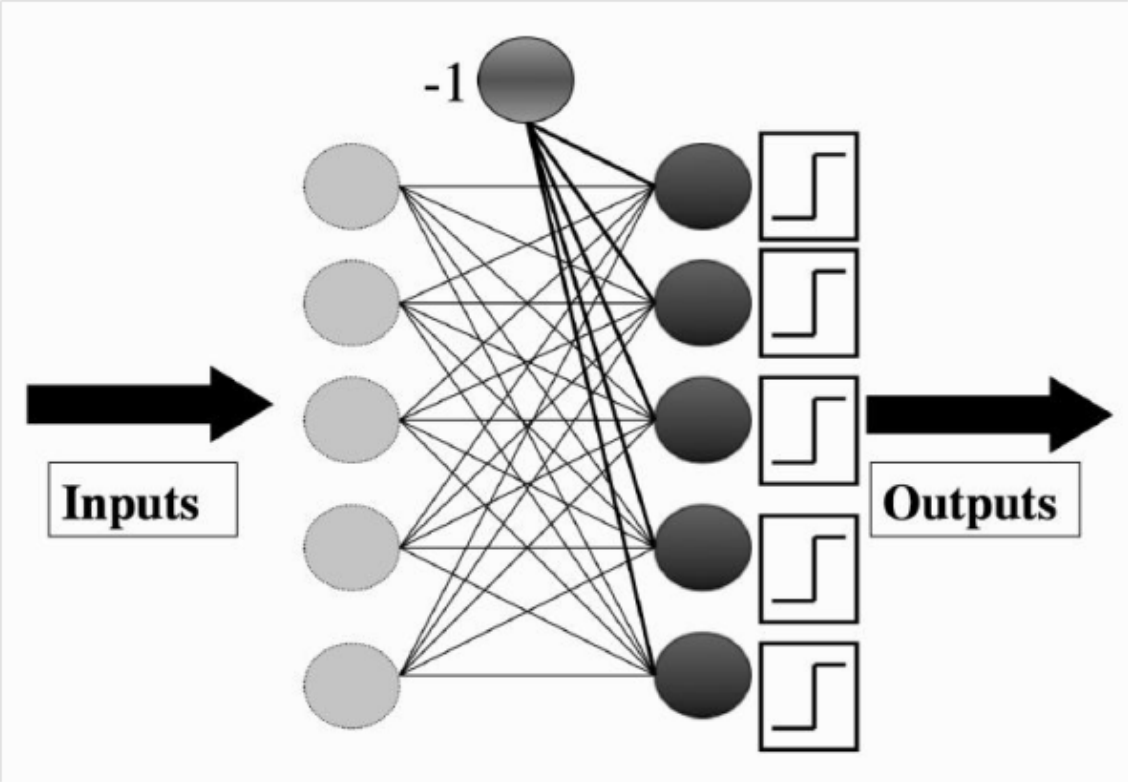


Figura 2.4: RN utilizando dados de entrada Bias, MARSLAND (2009)

Não há regra para definir quantidades, porém é necessário certo cuidado para que as características impostas a RN não causem overfitting ou underfitting.

Overfitting, ou sobre-ajuste, é a situação onde a RN foi arquitetada de forma tal a gerar ótimos resultados na previsão quando aplicada a dados que façam parte dos dados de treinamento, porém incapazes de abstrair resultados de dados alheios a estes. Esta situação pode ser causada por excesso de dados de treinamento.

Já underfitting, ou sub-ajuste, é a situação onde a RN foi arquitetada de forma tal a ser incapaz de abstrair qualquer padrão dos dados por conta da escassez de informação, e pode ser causada pela baixa quantidade de dados de treinamento.

2.2.4 Backpropagation

Na confecção de uma RN, pesos e bias são parametros que recebem inicialmente valores aleatórios entre -1 e 1. Portanto, haverá naturalmente um erro E entre a predição da RN e o alvo, o qual é função de todos os pesos e bias, ou seja, $E = E(\mathbf{w}, \mathbf{b})$.

Backpropagation é uma técnica que minimiza funções de erro na definição de pesos e bias durante o treinamento da RN. A equação a ser minimizada pode ser vista em Eq. 2.10.

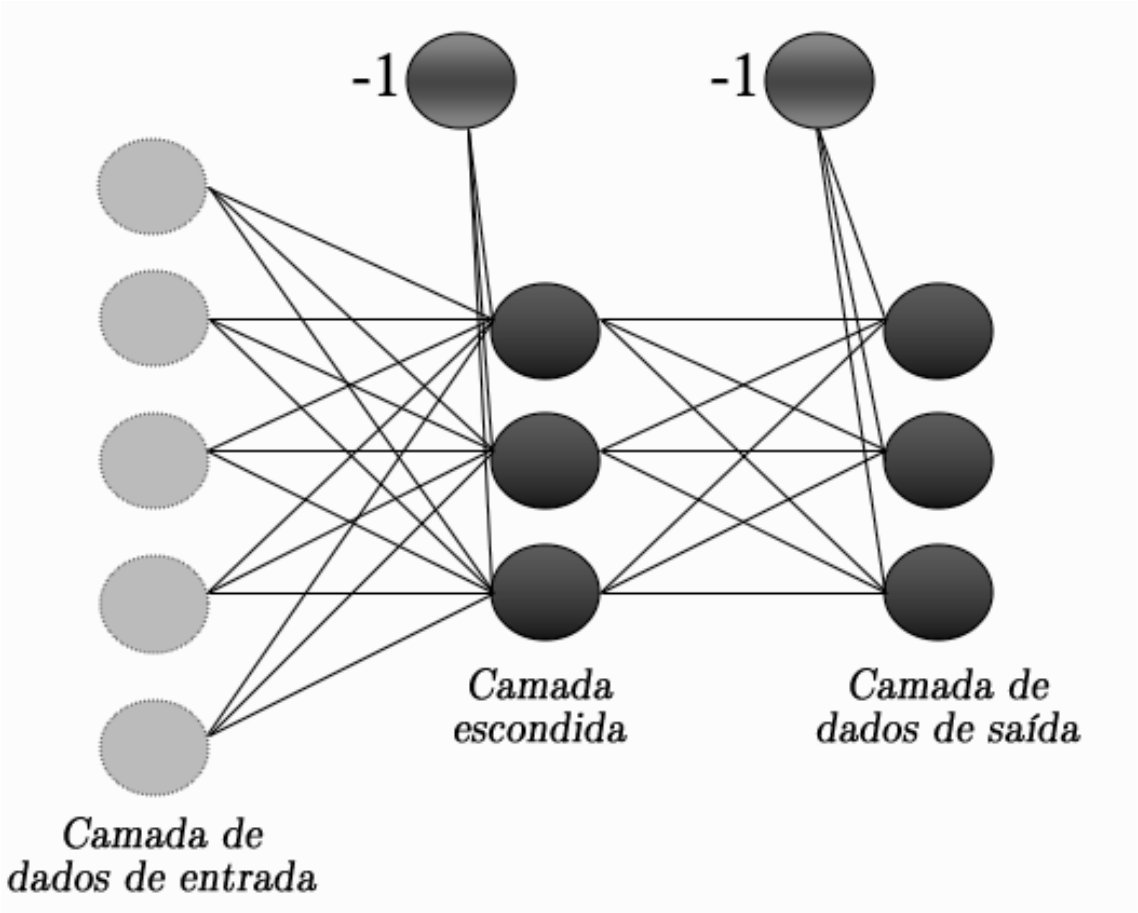


Figura 2.5: Camada escondida em uma RN Multi-layer Perceptron, adaptado de MARSLAND (2009)

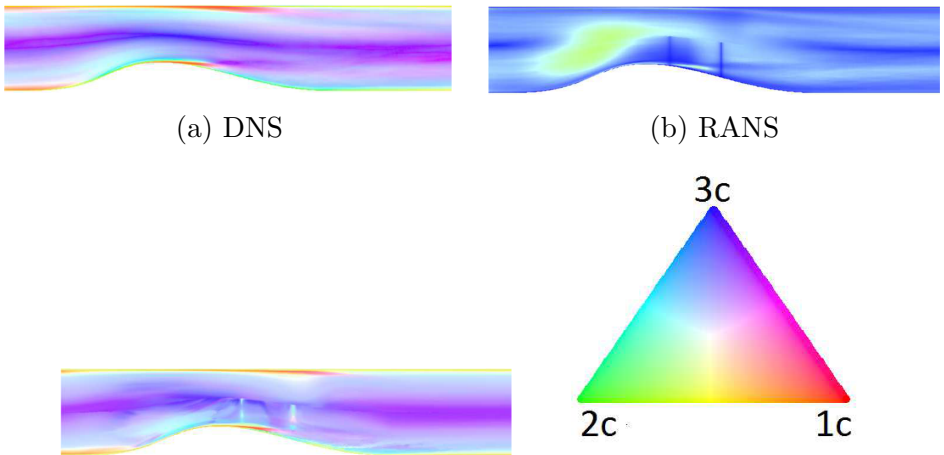
$$E = \sum_{b \in L_1} \frac{1}{2} \left(\tilde{y}_b - y_b \right)^2 \tag{2.10}$$

2.2.5 Validação e teste

O processo de treinamento acontece junto com o de validação. Enquanto o primeiro acontece com dados de treinamento, o segundo acontece com dados de validação, na intenção de fornecer um critério de convergência ao primeiro através do erro de validação. A cada passo de treinamento, também chamado de época, a RN é aplicada aos dados de validação com o objetivo de averiguar se a mesma está melhorando ou não em seu treinamento.

O objetivo dessa abordagem é assegurar que a RN, uma vez convergida, seja capaz de abstrair informações dos dados. Sem isso, a convergência ainda assim poderia acontecer, porém sem garantia de que a RN seria capaz de predizer resultados de dados diferentes daqueles utilizados no treinamento.

Após a convergência da RN, outro conjunto de dados é utilizado para avaliar sua capacidade, gerando assim um valor associado a sua capacidade de abstração.



(a) DNS (b) RANS

(c) Previsto (d) Mapa de cor

Figura 2.6: Campos referentes à parte anisotrópica de **R**, adaptado de TRACEY *et al.* (2013)

2.3 AM na simulação de escoamentos turbulentos

Há uma tendência mundial do uso de AM em diversos ramos de estudo, do comportamental ao prático, das propagandas direcionadas à criação de conteúdo de plataformas de entretenimento. Esta ferramenta se mostra muito poderosa na abstração de informações a partir de dados coletados que aparentemente não tem ligação lógica ou óbvia.

A aplicação de AM na simulação de turbulência é recente, mas vem ganhando adeptos. Suas técnicas são capazes de, a partir de dados de simulação RANS, fornecer previsões para o campo de velocidade médio, por exemplo, que tendem a resultados mais nobres, como os obtidos por DNS. Isso se dá através do treinamento de uma arquitetura de AM com dados nobres, onde ela aprende com os mesmos a acertar determinada componente de um tensor ou invariantes de um campo. O que ocorre usualmente é a correção do tensor **R**.

Devido à baixa acurácia das simulações RANS na modelagem de turbulência e combustão, TRACEY *et al.* (2013) utilizou AM como uma ferramenta de quantificação de incerteza através do desenvolvimento de um modelo estocástico. De uma das aplicações da ferramenta, foi possível obter melhores resultados de simulação para o periodic hill, clássico escoamento turbulento que apresenta descolamento e recolamento da camada limite, através da predição de um campo turbulento anisotrópico corrigido por AM, como visto na Fig. 2.6. Tal correção ocorreu a partir de dados de simulação RANS $\kappa - \omega$ SST, tendo como alvo dados DNS.

Dando sequencia ao estudo da aplicação de AM às equações RANS, TRACEY *et al.* (2015) propôs a utilização de RN como ferramenta na criação de um novo modelo de turbulência, e não mais um preditor corretor de alguma QdI de um

modelo existente. Este novo modelo consiste de uma RN treinada em uma base de dados de um modelo já existente, neste caso o Spalart-Allmaras, onde essa RN seria capaz de reproduzir com fidelidade resultados que poderiam ser obtidos da aplicação do modelo Spalart-Allmaras original. Essa abordagem mostrou resultados promissores, que podem ser observados na Fig. 2.7.

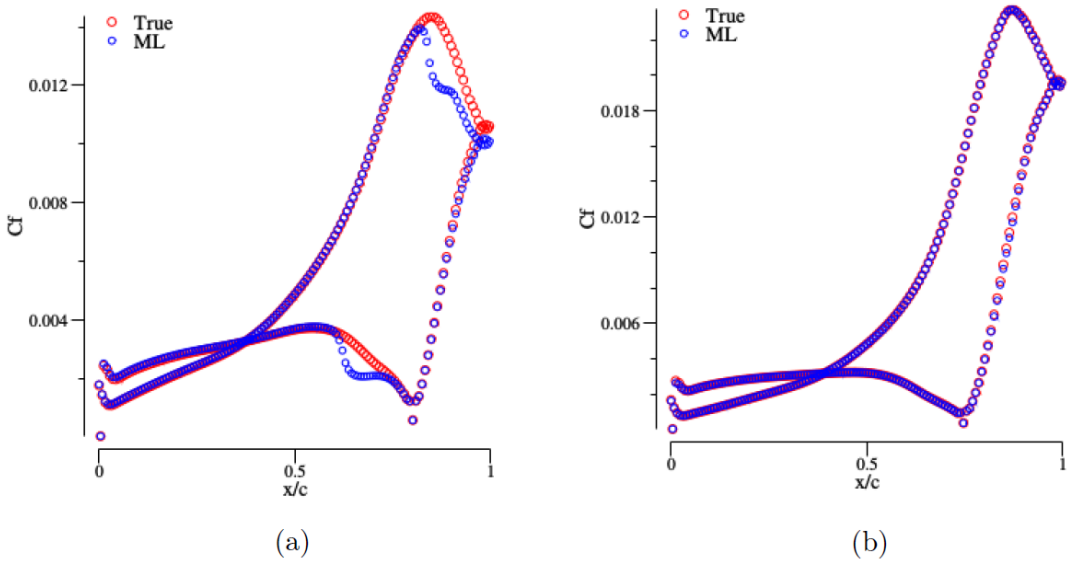


Figura 2.7: Comparação das previsões utilizando dados de treinamento de: (a) canais acionados por pressão; (b) solução de três placas planas. TRACEY *et al.* (2015)

Utilizando uma técnica AM chamada floresta randômica (FR), LING *et al.* (2016c) propõe melhorar o resultado de RANS em mais um caso onde o mesmo não produz resultado satisfatório, o escoamento de entrada em turbina ou Jet-in-Crossflow. A FR é treinada com dados LES para o mesmo tipo de escoamento, e utilizada para predizer a parte deviatórica de $\mathbf{R}$. Os resultados, que podem ser visto na Fig. 2.8, mostraram que a FR foi capaz de apontar tendências turbulentas de algumas regiões do escoamento, porém não conseguiu replicar os resultados LES alvo.

Ainda sobre a correção de RANS por AM, LING *et al.* (2016b) propõe a comparação de duas técnicas AM, RN e FR, para dois métodos de treinamento. Um treina as arquiteturas com dados crus provenientes de uma base de dados DNS até que os modelos aprendam a prever de forma invariante, ou seja, independente da orientação física do escoamento. O segundo método consiste em treinar as arquiteturas para dados já invariantes. O apelo de tal abordagem consiste em permitir que modelos AM possam prever campos para escoamentos semelhantes sem que estes estejam sob a mesma orientação física. Resultados, representados na Fig. 2.9, mostraram que os modelos treinados a partir de uma base já invariante foram melhores

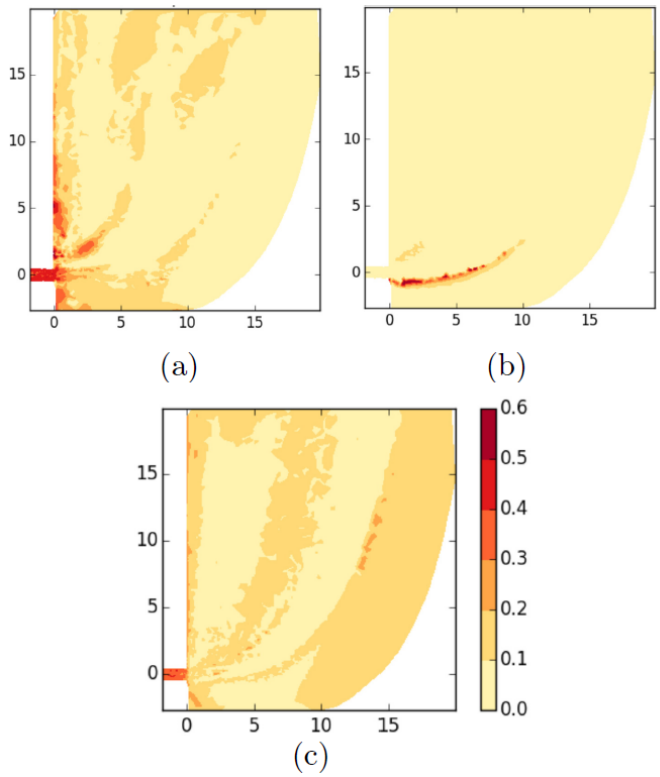


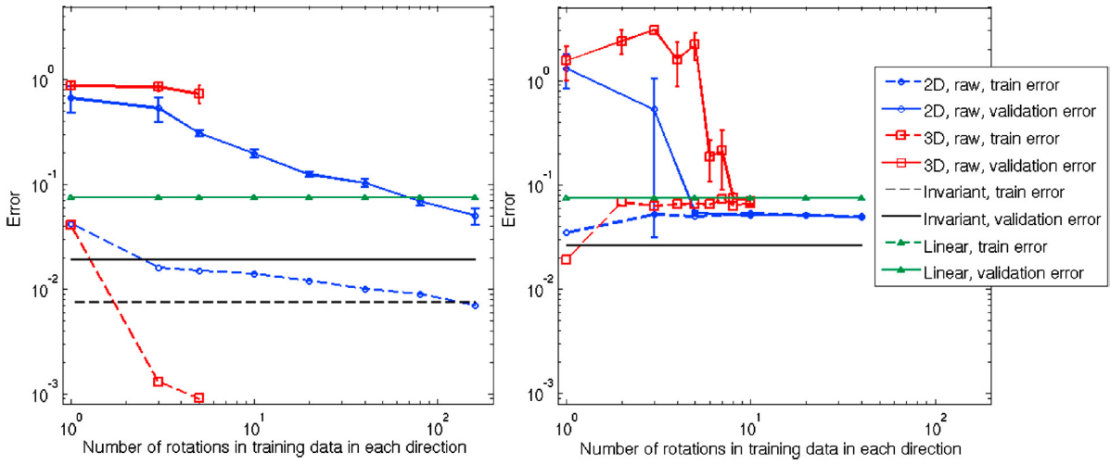
Figura 2.8: Perfil do segundo invariante anisotrópico para (a) LES, (b) RANS e (c) previsão da FR, LING *et al.* (2016c)

em suas previsões, além de o fazerem utilizando menos capacidade computacional do que com a outra abordagem, o que exalta a importância da invariância dos dados de treinamento.

Já em LING *et al.* (2016a), foi proposto o uso de deep learning com invariância Galileana embutida para prever a parte anisotrópica do tensor de Reynolds, e comparou-se os resultados desse modelo com uma RN comum sem invariância, além de dois modelos RANS. O modelo proposto se saiu melhor que seus concorrentes, provando mais uma vez a importância da invariância. Os resultados podem ser vistos na Fig. 2.10, onde TBNN é a RN proposta, MLP é a RN sem invariância, LEVM e QEVM são modelos RANS.

Investigando os efeitos da correção de uma QdI RANS através de AM na obtenção de um melhor campo de velocidade, WANG *et al.* (2016) propôs um modelo que chamou de physics-informed (ou informado fisicamente), onde, através da previsão da discrepância do tensor de Reynolds RANS em relação ao DNS através de entidades invariantes coletadas de RANS, conseguiu bons resultados na melhora do perfil de velocidade para o caso escoamento em duto de seção quadrada, mesmo este tendo sido treinado para casos onde os números de Re eram menores que o corrigido, 2200 a 2900 contra 3500 respectivamente. Resultados podem ser vistos na Fig. 2.11.

Na tentativa de corrigir simulações RANS a partir da previsão da discrepância



(a) Random Forest (b) Neural Network

Figura 2.9: Erro como função do número de rotações de treino, exceto para os dados invariantes, LING *et al.* (2016b)

entre o tensor de Reynolds RANS e o DNS, WANG *et al.* (2017) propõe a utilização de FR em duas abordagens. Em uma, ele treina e aplica a FR para um mesmo tipo de escoamento. Já na segunda ele tenta a correção de um escoamento com separação de camada limite a partir da FR treinada com outros escoamentos, semelhantes porém não iguais em geometria. Os resultados mostram considerável melhora em todos os métodos quando comparados aos resultados RANS. Porém, no caso em que treinamento e correção são feitos em escoamentos diferentes, onde a FR faz uma extrapolação dos dados, os resultados, que podem ser vistos na Fig. 2.12, não surpreendem.

Sabendo da incerteza intrínseca aos campos DNS quanto ao tensor de Reynolds, como reportado por THOMPSON *et al.* (2016), CRUZ (2018) propôs o uso de RN na correção do divergente modificado de $\mathbf{R}$, chamado por ele de $\mathbf{t}$, para o escoamento em duto de seção quadrada.

O divergente de $\mathbf{R}$ pode ser obtido através da Eq. 2.11. Entretanto, sabendo da presença de incertezas do campo $\mathbf{R}$ disponibilizado nas bases de dados DNS, tal abordagem carregaria as mesmas. Na intenção de contornar este problema, calcula-se o divergente a partir de entidades de primeira ordem estatística, como os campos de velocidade e pressão, pois estes são melhor convergidos, como mostra a Eq. 2.12

$$\nabla \cdot \mathbf{R}|_{\text{DNS}} = \boldsymbol{\nabla} \cdot \mathbf{R}_{\text{DNS}} \tag{2.11}$$

$$\nabla \cdot \mathbf{R}|_{\text{DNS}} = -\left(\langle \mathbf{v} \rangle_{\text{DNS}} \cdot \nabla\right)\left(\langle \mathbf{v} \rangle_{\text{DNS}}\right) - \frac{1}{\rho}\nabla\langle p \rangle_{\text{DNS}} + \nu\nabla^2\langle \mathbf{v} \rangle_{\text{DNS}} \tag{2.12}$$

A base de dados DNS utilizada não disponibiliza o campo de pressão. Portanto

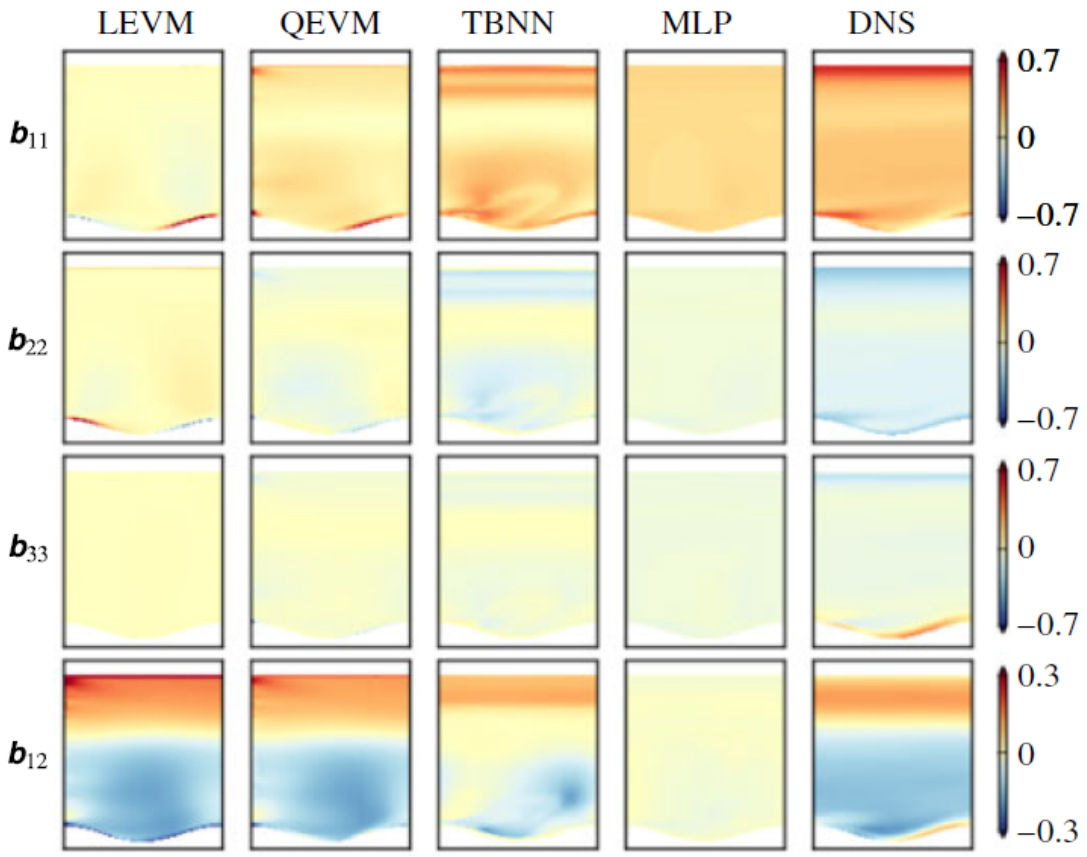


Figura 2.10: Previsão do tensor de Reynolds anisotrópico, LING *et al.* (2016a)

surge o tensor $\mathbf{t}$, definido pela Eq. 2.13, como sendo o divergente modificado de $\mathbf{R}$, CRUZ (2018).

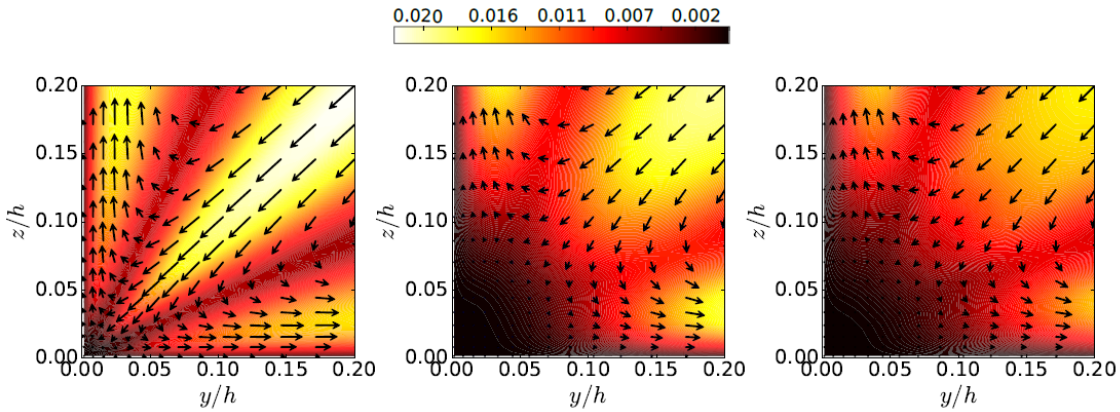
$$\mathbf{t} \equiv \boldsymbol{\nabla} \cdot \mathbf{R} + \frac{1}{\rho} \boldsymbol{\nabla} \langle p \rangle \tag{2.13}$$

A Eq. 2.14 mostra como extrair $\mathbf{t}$ da base de dados DNS.

$$\mathbf{t}_{\mathrm{DNS}} = -\left(\langle \mathbf{v} \rangle_{\mathrm{DNS}} \cdot \boldsymbol{\nabla}\right)\left(\langle \mathbf{v} \rangle_{\mathrm{DNS}}\right) + \nu \boldsymbol{\nabla}^2 \langle \mathbf{v} \rangle_{\mathrm{DNS}} \tag{2.14}$$

Os resultados apresentados pelo método $\mathbf{t}$ se mostraram superiores aos da correção de $\mathbf{R}$ principalmente para o campo de velocidade médio na direção principal, como vistos na Fig. 2.13.

Com o intuito de melhorar a predição do campo de velocidade médio, WU *et al.* (2018) propõe um modelo implícito com invariância garantida onde, através de uma RN treinada com dados de um escoamento periodic hill de vale mais íngreme, consegue prever com boa acurácia as QdI de um escoamento periodic hill de vale menos íngreme. Dados do campo de velocidade média foram utilizados como dados de entrada da RN. Seus resultados, mostrados na Fig. 2.14, mostram a capacidade das técnicas AM quanto a predizer características do escoamento que não foram dire-



(a) Baseline (b) DNS (c) Predicted

Figura 2.11: Campos secundarios de velocidade média, WANG *et al.* (2016)

tamente treinadas, uma vez que o escoamento previsto por sua RN era levemente diferente geometricamente.

2.4 A falta de convergência nas bases de dados DNS

É sabido que as bases de dados DNS disponíveis não possuem a convergência necessária de determinados campos. THOMPSON *et al.* (2016) apontou que certas bases de dados DNS possuíam uma incerteza intrínseca quanto ao tensor de Reynolds, pois estes possuíam convergência menor que os demais campos médios DNS disponibilizados de uma mesma simulação, como velocidade e pressão médios. Isso acontece pois tais campos são de segunda ordem estatística, ou seja, sua convergência depende da convergência prévia de outras grandezas, essas de primeira ordem estatística como velocidade e pressão. As simulações DNS costumam ser feitas para baixos Re, e o tempo de simulação é suficiente para obter a convergência das entidades de primeira ordem estatística, o que como mencionado, não garante a convergência de $\mathbf{R}$.

É notório que a falta de convergência de $\mathbf{R}$ nas bases de dados de simulação DNS apontada por THOMPSON *et al.* (2016) influencia negativamente os resultados da aplicação de AM na correção de RANS. Ao analisar a propagação de incerteza de $\mathbf{R}_{DNS}$ no campo de velocidade média axial à direção do escoamento, U_x, em RANS, nota-se que este não converge para o seu semelhante DNS. Não por acaso essa QdI é por vezes omitida nos artigos que estudam aplicação de AM em turbulência. Porém, é possível ver com clareza em CRUZ (2018) que quando a correção é feita para entidades melhor convergidas, como em seu método de correção de $\mathbf{t}$ que utiliza dados de primeira ordem estatística, os resultados obtidos são melhores e congruentes fisi-

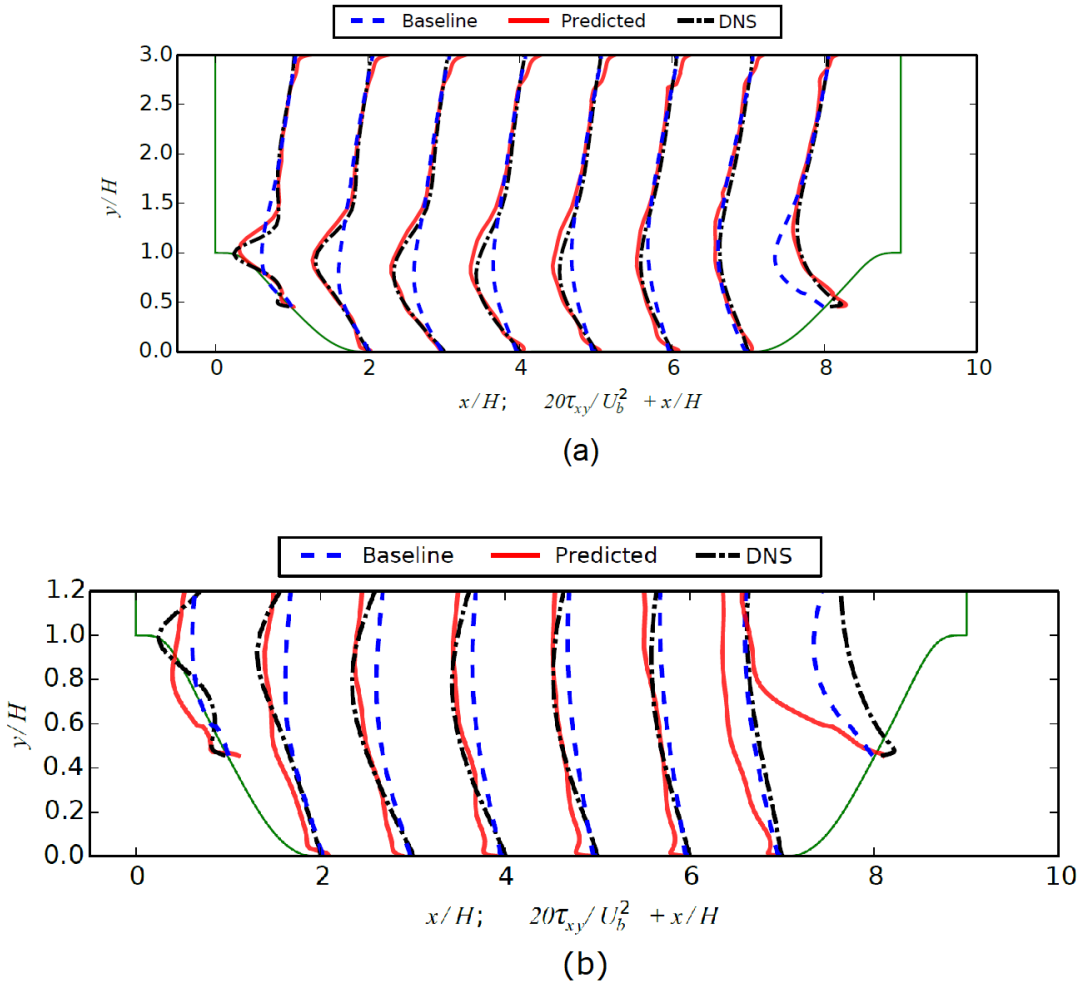


Figura 2.12: Previsão da tensão de cisalhamento para dados de treinamento de: (a) mesma geometria, Re diferentes; (b) geometrias e Re diferentes, WANG *et al.* (2017)

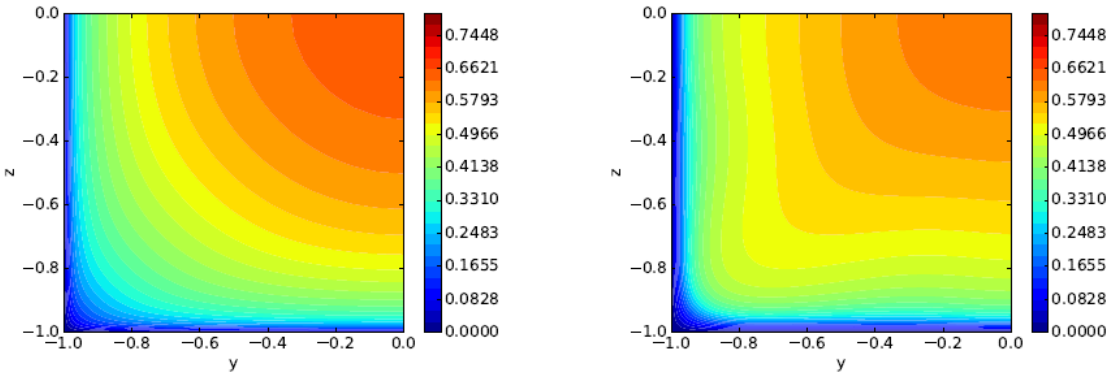
camente.

Uma análise de incertezas de dados DNS é proposta por ANDRADE *et al.* (2018), para os escoamentos turbulentos de fluidos Newtonianos em canal e em duto fechado, na intenção de quantificar erros de dados DNS e experimentais. Duas métricas de erro residual são propostas para tanto, as quais estão definidas nas Eq. 2.15 e 2.16.

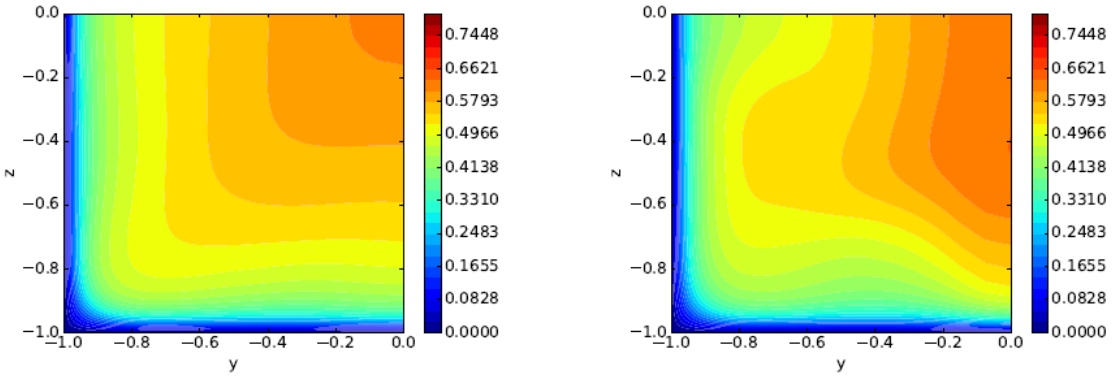
$$E_R\left(y^+\right) = \widetilde{R}_{yx}^+\left(y^+\right) - \widehat{R}_{yx}^+\left(y^+\right) \tag{2.15}$$

$$E_U\left(y^+\right) = \widetilde{U}^+\left(y^+\right) - \widehat{U}^+\left(y^+\right) \tag{2.16}$$

Nas Eq. 2.15 e 2.16, os componentes sobrepostos com circunflexo são resultados provenientes das bases de dados. Já os sobrepostos por til são resultados obtidos do balanço da equação média de forças a partir de resultados das bases de dados, e são calculados como visto nas Eq. 2.17 e 2.18.



(a) U_x RANS. (b) U_x DNS.



(c) U_x corrected by **t**. (d) U_x corrected by **R**.

Figura 2.13: Campos de velocidade média na direção principal do escoamento, CRUZ (2018)

$$\widetilde{R}^{+}_{yx}\left(y^{+}\right) = 1 - \frac{y^{+}}{Re_{\tau}} - \frac{d\widehat{U}^{+}}{dy^{+}}\left(y^{+}\right) \tag{2.17}$$

$$\widetilde{U}^{+}\left(y^{+}\right) = y^{+} - \frac{y^{+2}}{2\,\mathrm{Re}_{\tau}} - \int_{0}^{y^{+}} \widehat{R}^{+}_{yx}\left(y'\right) dy' \tag{2.18}$$

A Fig. 2.15 mostra a comparação dos erros residuais entre diversos resultados DNS para Re variados ao longo de y^{+}, que está dividido por regiões do escoamento, de I a IV são elas respectivamente: sub camada viscosa; buffer layer; região de lei log; e camada externa.

Nas partes (a) e (c) da Fig. 2.15 observa-se que os casos que apresentam E_R próximos de 0 ao longo de y^{+} são os das siglas PS (resultados gerados pelo próprio autor) e VK. As demais curvas apresentam comportamentos variados, sem nenhuma semelhança comportamental aparente entre sí.

Nas partes (b) e (d) da Fig. 2.15 observa-se que os casos que apresentam E_U próximos de 0 ao longo de y^{+} também são os das siglas PS e VK. Os demais apre-

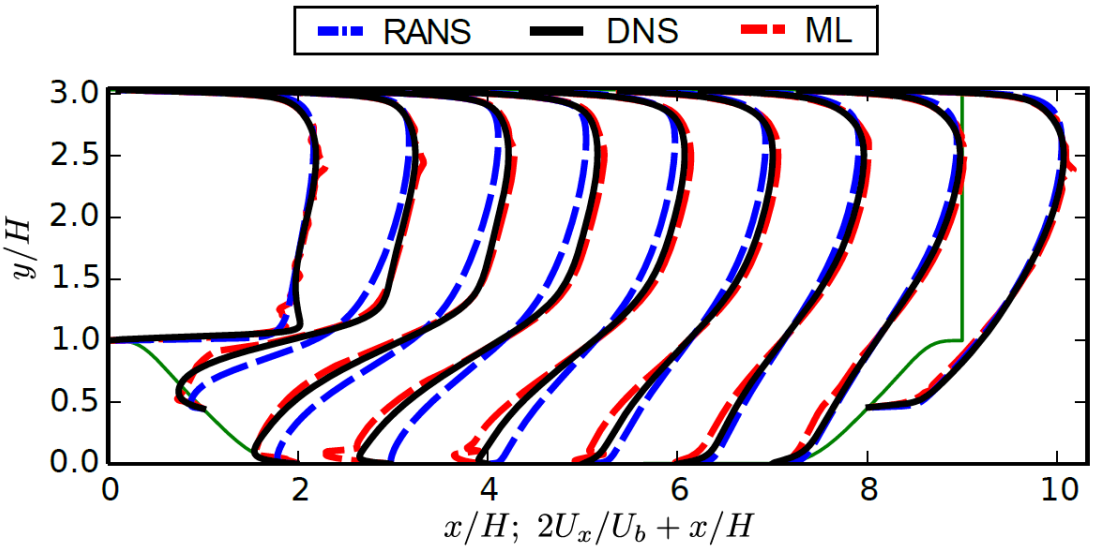


Figura 2.14: Perfil de velocidade a partir da previsão com método implícito, WU *et al.* (2018)

sentam módulo de E_U crescente com y^+, sendo desprezíveis próximo a parede. Isso se dá por conta da predominância das escalas maiores nessa região do escoamento, fazendo necessário tempo maior de simulação para convergência nesta.

A Fig. 2.16 mostra a comparação dos erros residuais para diversos tempos de simulação DNS, para dois Re diferentes.

É possível observar da Fig. 2.16 que os melhores resultados são das simulações com maior tempo, o que é esperado. Nota-se também que a qualidade do resultado é diretamente proporcional ao tempo de simulação para as duas métricas de erro. Ou seja, se há incerteza presente em certos campos das bases de dados DNS, essas se devem ao descaso com o tempo de simulação.

2.5 Escoamento em duto quadrado

O escoamento escolhido para o estudo foi o em duto de seção quadrada. Tal escoamento foi escolhido por possuir escoamento secundário no plano transversal ao escoamento principal, geometria fácil de ser representada, além de uma base de dados DNS disponível. Uma vez que o objetivo é estudar o comportamento da aplicação de AM na correção de um escoamento, não há a necessidade da escolha de uma geometria complexa.

A Fig. 2.17 ilustra o escoamento em duto de seção quadrada.

Na parte (a) da Fig. 2.17 está representado o duto quadrado, suas dimensões características, a direção do escoamento principal, em x, e um corte no plano y-z onde estão representados os escoamentos secundários.

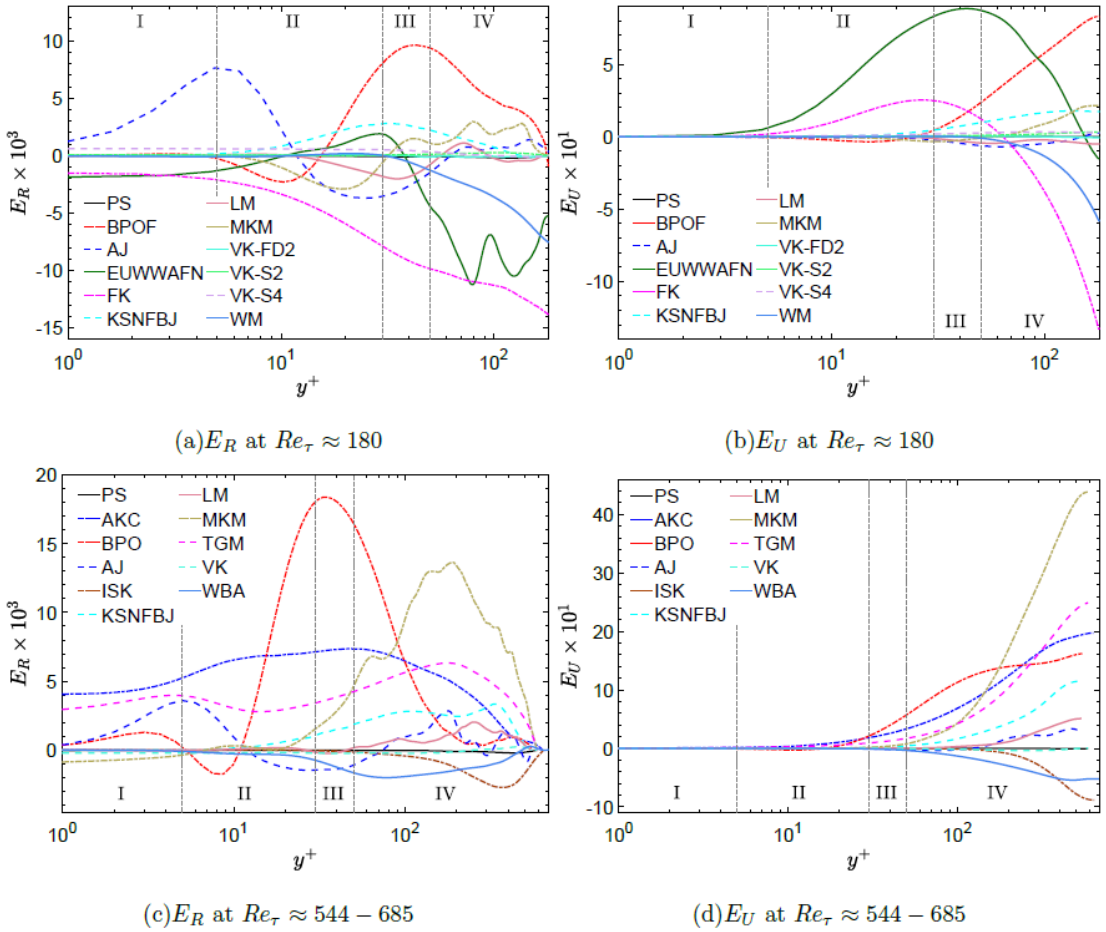


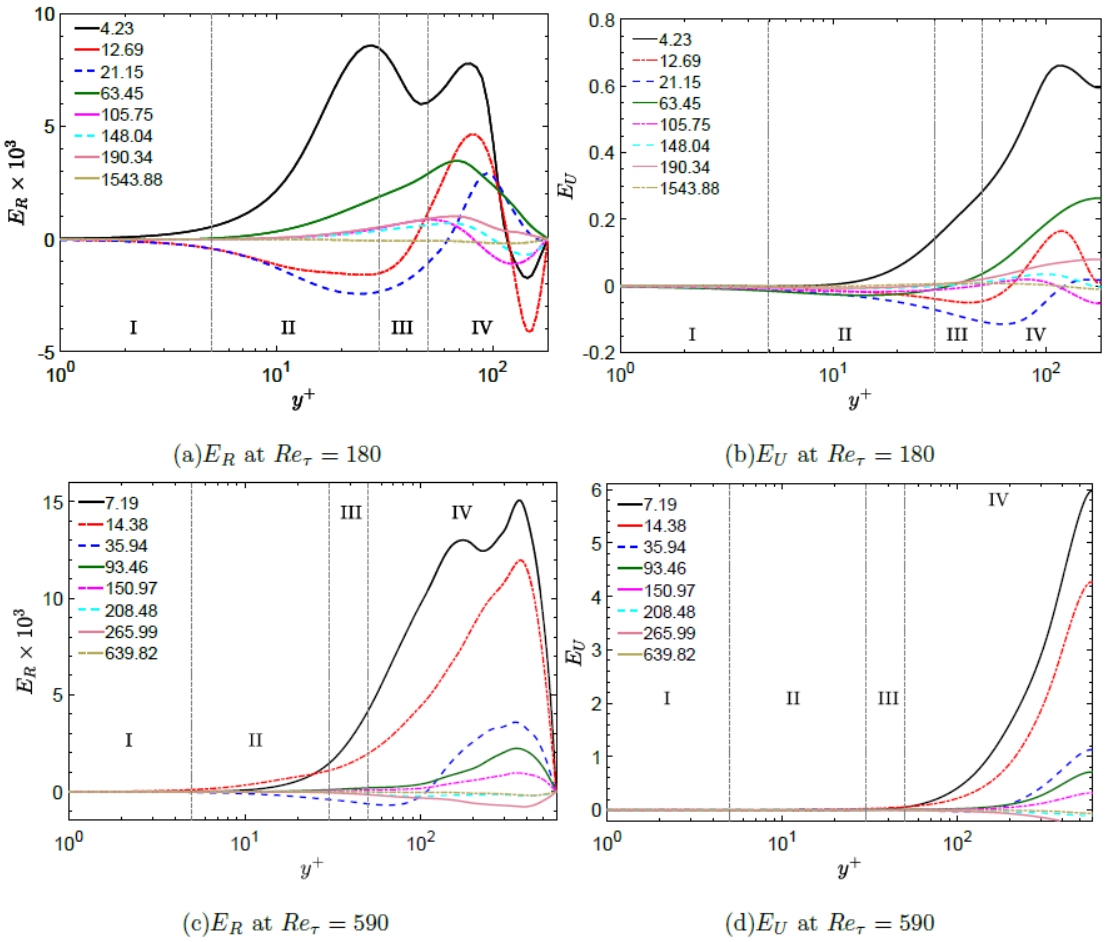
Figura 2.15: Comparação de erros residuais para diversos Re e bases DNS, ANDRADE *et al.* (2018)

A parte (b) da Fig. 2.17 mostra o recorte de um quadrante do plano y-z, o qual representa o domínio de simulação. Uma vez que o escoamento possui geometria regular e simétrica, assim é o escoamento médio, não sendo necessária a simulação de toda a seção transversal do mesmo. Simula-se o domínio mostrado e espelha-se o resultado para os demais quadrantes. Pode-se observar a presença de duas células de recirculação próximas ao vértice inferior esquerdo do domínio. Estas células são características ao escoamento e estão presentes em todos os vértices do mesmo.

Na parte (c) da Fig. 2.17 podemos ver o corte y-z aumentado, além das linhas de simetria do escoamento. Observa-se que, além da simetria entre os quadrantes, há também uma simetria interna dos mesmos com relação a diagonal.

Sabendo da presença de escoamento secundário no escoamento em duto de seção quadrada, foi escolhido o modelo $\kappa - \epsilon$, o qual é incapaz de prever tal característica, justamente para avaliar a capacidade de aprendizado através da aplicação de AM.

O modelo $\kappa - \epsilon$ será utilizado através de simulação RANS no software aberto OpenFOAM-4.x (OF).



(a)E_R at $Re_\tau = 180$

(b)E_U at $Re_\tau = 180$

(c)E_R at $Re_\tau = 590$

(d)E_U at $Re_\tau = 590$

Figura 2.16: Comparação de erros residuais para diversos tempos de simulação DNS, ANDRADE *et al.* (2018)

A Fig. 2.18 ilustra as condições de contorno impostas a geometria do problema. Para as paredes foi definida a condição de não escorregamento, além de gradiente de pressão nulo para as paredes sólidas, representadas em verde. As faces em vermelho são internas ao escoamento, delimitam o quadrante, e receberam a condição de simetria. Já as faces referentes a entrada e saída do escoamento principal receberam a condição de contorno cíclica, além de velocidade média de seção constante. Esta última condição imposta cria um componente artificial de gradiente de pressão necessário para a manutenção da condição periódica de entrada e saída, CRUZ (2018).

As condições de simulação no OF foram as mesmas utilizadas por CRUZ (2018), ou seja: energia cinética turbulenta definida como zero nas paredes; função de parede epsilonWallFunction para a dissipação turbulenta; solver utilizado no OF foi o simpleFoam, o qual funciona a volumes finitos na solução de escoamentos de fluidos Newtonianos, incompressíveis, em estado não estacionário. O acoplamento pressão velocidade foi realizado pelo algoritmo chamado simple. Foi utilizado método Gauss upwind na discretização dos operadores divergente nas equações de transporte. Já

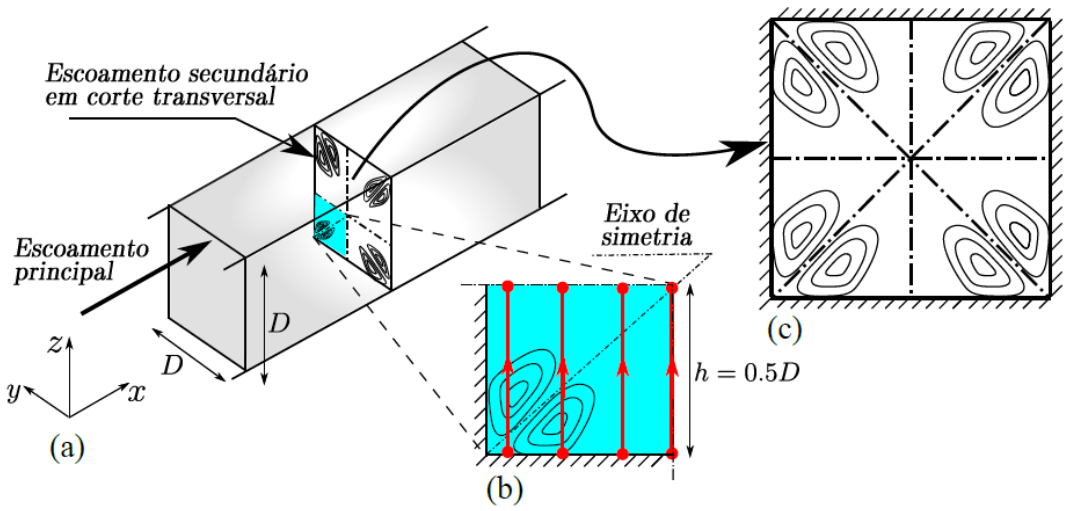


Figura 2.17: Esquema do escoamento em duto de seção quadrada, adaptado de WANG *et al.* (2016)

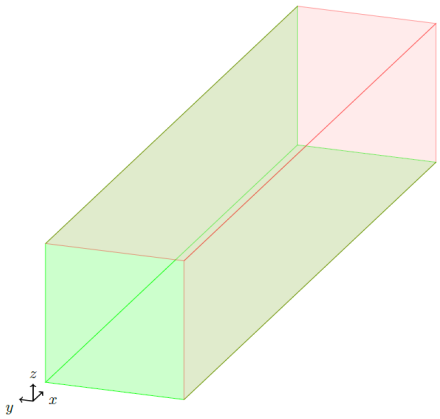


Figura 2.18: Condições de contorno da simulação em duto de seção quadrada, CRUZ (2018)

para o operador gradiente foi utilizado método linear de Gauss. Para os operadores laplaciano foi utilizado método de Gauss linear corrigido. A interpolação ponto a ponto foi realizada linearmente. A solução da equação discretizada da pressão foi dada pelo solver Preconditioned Conjugate Gradient (PCG) em conjunto com o pré condicionador Diagonal Incomplete-Cholesky (DIC). Para a velocidade, κ e ϵ, utilizou-se o smooth solver com Symmetric Gauss Seidel Smoother. Usou-se de tolerância 10^{-8}, onde os fatores de relaxação para os campos da pressão, turbulentos e a equação da velocidade foram de $0,5$, $0,5$ e $0,15$ respectivamente.

2.5.1 Base de dados para treinamento

Para que a RN seja treinada afim de corrigir as simulações RANS do escoamento em duto de seção quadrada, são necessários dados nobres que serão os alvos do

Tabela 2.1: Viscosidades de simulação

Re	$\nu\,[\mathrm{m^2 s^{-1}}] \times 10^{-4}$
2200	2,1858
2400	2,0082
2600	1,8537
2900	1,6619
3200	1,5061
3500	1,3770

treinamento. Os dados aqui utilizados foram disponibilizados por PINELLI *et al.* (2010). Utilizou-se amostras de simulações DNS para seis diferentes Re, são eles 2200, 2400, 2600, 2900, 3200 e 3500, assim como os resultados RANS de cada um. As simulações são rodadas levando em conta uma velocidade média de seção de $0,4819ms^{-1}$, para um domínio de $1m \times 1m \times 10m$, e são diferenciadas pela viscosidade cinemática mostradas na Tabela 2.1.

2.5.2 Malha de simulação

Com um total de 16000 centroides, a malha utilizada para o estudo possui 40 $\times$ 40 nós na seção transversal, com uma razão de crescimento tal que imponha maior refino de malha nas paredes, garantindo $y^+ < 0,2$, e 10 nós igualmente espaçados na direção principal do escoamento. A opção de aumentar o número de pontos na região da parede foi proposital. Dessa forma, há mais informações da região, o que ajuda a RN a corrigi-la melhor. A Fig. 2.19 mostra um corte transversal da malha em questão.

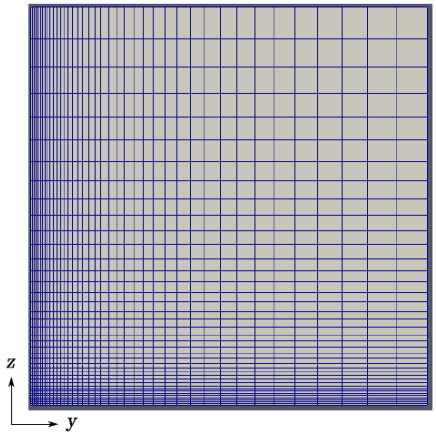


Figura 2.19: Malha da simulação em duto de seção quadrada, CRUZ (2018)

Capítulo 3

Metodologia

3.1 Uso de RN na correção de RANS

Como mencionado anteriormente, DNS e LES são as simulações CFD que apresentam melhores resultados, porém demandam grande esforço computacional e tempo, o que torna suas escolhas proibitivas para uso cotidiano, como por exemplo na industria. Esse é o motivo por RANS ser o método mais utilizado. Apesar de sua acurácia ser menor que os antes citados, a qualidade de suas simulações tem sido suficiente para uma boa diversidade de problemas corriqueiros.

O que motiva este estudo, e outros relacionados à aplicação de AM em RANS, é apresentar um concorrente às simulações RANS, que possa ser equiparável em custo computacional e temporal, mas que apresente resultados superiores, ou seja, uma RN capaz de prever, a partir da leitura de entidades RANS cinemáticas e turbulentas do escoamento, um output que ajude RANS a obter resultados novos, melhores que os seus próprios, e que tendam à qualidade de DNS, LES ou outro tipo de resultado nobre usado no treinamento.

Para que isso aconteça são necessárias duas etapas, as quais serão discutidas na presente seção. São elas a criação de uma RN, também conhecido como treinamento, e como utilizá-la.

3.1.1 Treinando a RN

Treinar uma RN significa fazer ela aprender a mapear uma matriz de dados de entrada X em outra matriz de dados de saída Y.

Entidades cinemáticas e turbulentas do escoamento foram escolhidas para compor os dados de entrada, também conhecidos como features, como pode ser visto na tabela 3.1, CRUZ (2018), onde $\mathbf{D}$ é o tensor taxa de deformação, $\mathbf{P}$ é o tensor não persistência de deformação, THOMPSON e MENDES (2005), e $\mathbf{R}$ é o tensor das tensões de Reynolds, além de vetores obtidos da aplicação do operador divergente

Tabela 3.1: Dados de entrada da RN

$\mathbf{D}$	$\nabla \cdot \mathbf{D}$
$\mathbf{P}$	$\nabla \cdot \mathbf{P}$
$\mathbf{D}^2$	$\nabla \cdot \mathbf{D}^2$
$\mathbf{P}^2$	$\nabla \cdot \mathbf{P}^2$
$\mathbf{D} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{D}$	$\nabla \cdot (\mathbf{D} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{D})$
$\mathbf{D}^2 \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{D}^2$	$\nabla \cdot (\mathbf{D}^2 \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{D}^2)$
$\mathbf{P}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{P}^2$	$\nabla \cdot (\mathbf{P}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{P}^2)$
$\mathbf{R}$	$\nabla \cdot \mathbf{R}$
$\mathbf{R}^2$	$\nabla \cdot \mathbf{R}^2$
$\mathbf{D} \cdot \mathbf{R} + \mathbf{R} \cdot \mathbf{D}$	$\nabla \cdot (\mathbf{D} \cdot \mathbf{R} + \mathbf{R} \cdot \mathbf{D})$
$\mathbf{D}^2 \cdot \mathbf{R} + \mathbf{R} \cdot \mathbf{D}^2$	$\nabla \cdot (\mathbf{D}^2 \cdot \mathbf{R} + \mathbf{R} \cdot \mathbf{D}^2)$
$\mathbf{R}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{R}^2$	$\nabla \cdot (\mathbf{R}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{R}^2)$

aos tensores mencionados. Todos esses dados são provenientes da simulação RANS.

O tensor não persistência de deformação $\mathbf{P}$ é uma medida de deformação local, e é definido como mostra a Eq. 3.1, onde $\overline{\mathbf{W}}$ é a vorticidade relativa, que é o tensor vorticidade descontado do tensor que representa a taxa de rotação dos autovetores de $\mathbf{D}$, como pode ser visto na Eq. 3.2, THOMPSON e MENDES (2005).

$$\mathbf{P} = \mathbf{D} \cdot \overline{\mathbf{W}} - \overline{\mathbf{W}} \cdot \mathbf{D} \tag{3.1}$$

$$\overline{\mathbf{W}} = \mathbf{W} - \Omega^D \tag{3.2}$$

Uma vez que todos os tensores utilizados como dados de entrada são simétricos, somam-se um total de 108 dados de entrada. Esses dados são coletados dos centróide da simulação RANS, que totalizam 16000 para o presente caso, do escoamento em duto de seção quadrada.

Foram utilizadas seis simulações RANS do escoamento em duto de seção quadrada, realizadas para números de Re equivalentes aos da base de dados DNS, na concepção da RN. Quatro para o treinamento, uma para o teste, e uma para a validação, CRUZ (2018).

As entidades escolhidas como dados de saída a serem mapeados pela RN são provenientes dos componentes do tensor das tensões de Reynolds ($\mathbf{R}$) fornecidos a partir de base de dados DNS.

São utilizados dois conjuntos diferentes de outputs baseados em $\mathbf{R}$ para o treinamento da RN. Um é formado pelos componentes de $\mathbf{R}$ propriamente ditos. O outro é formado pelos componentes do divergente modificado de $\mathbf{R}$, chamado aqui de vetor de Reynolds, ou $\mathbf{t}$.

As características da RN utilizada seguem a estrutura proposta por CRUZ (2018). Quatro camadas de neurônios, sendo duas camadas escondidas, uma de

dados de entrada e uma de dados de saída.

A camada de dados de entrada precisa ter um número de neurônios igual ao número de dados de entrada, neste caso 108.

As camadas escondidas foram definidas com cem neurônios cada, operando com funções sigmoides não-lineares.

Já a camada de outputs deve possuir número de neurônios igual ao número de dados de saída. São seis neurônios para o caso da correção de **R**, e três para a correção de **t**. Para ambos os casos, os neurônios operam com funções lineares, prática comum para a regressão na camada de outputs.

A construção, treinamento, teste e validação da arquitetura da RN foi feita através de uma biblioteca aberta de Python chamada Keras.

Os dados de entrada da RN são coletados e agrupados em uma matriz única, onde as colunas representam os features vistos na Tabela 3.1, e as linhas representam as amostras de centroide das simulações RANS. Isso significa que a matriz de dados de entrada X_{train} tem as dimensões 64000 $\times$ 108.

A matriz de outputs Y_{train} do treinamento possui dimensões 64000 $\times$ 6 para o caso da correção de **R**, e 64000 $\times$ 3 para o caso da correção de **t**.

As matrizes dados de entrada de teste X_{test} e validação X_{val} possuem ambas a dimensão de 16000 $\times$ 108.

As matrizes output de teste Y_{test} e validação Y_{val} possuem dimensão 16000 $\times$ 6 para o caso da correção de **R**, e 16000 $\times$ 3 para o caso da correção de **t**.

Como as entidades de entrada e saída podem apresentar magnitudes diferentes, uma prática comum é a normalização das mesmas. Esta normalização é feita como observado na Eq. 3.3, onde σ é o desvio padrão e μ é o valor médio. Essa operação é feita através de uma biblioteca de Python popularmente conhecida, a scikit-learn, CRUZ (2018).

$$\text{coluna normalizada } = \frac{\text{coluna } - \mu_{\text{coluna}}}{\sigma_{\text{coluna}}} \tag{3.3}$$

Um método numérico conhecido como ADAM é utilizado para obter-se valores de pesos e bias da RN. O critério de convergência utilizado foi o de erro quadrático médio, ou mean square error (mse).

Após os passos descritos, se tudo ocorreu como esperado, obtêm-se uma RN convergida e pronta para prever resultados.

3.1.2 Utilizando a RN

Agora a RN está pronta para prever um **R** ou **t** melhores que o RANS a partir de dados de entrada inéditos. O esperado é que a partir de novos dados de entrada

provenientes de uma simulação RANS não treinada previamente, a RN possa prever um novo **R** ou **t** de qualidade semelhante a DNS.

Uma vez que a RN for operada para uma nova predição, ela gera dados output normalizados que precisam de pós processamento para que então sejam utilizados numa nova simulação RANS.

Em seguida, o novo **R** ou **t** precisa ser enxertado na malha da simulação RANS substituindo os campos anteriores, para que então o RANS possa ser rodado novamente através do OpenFOAM, na intenção de obter-se novos campos de velocidade, mais próximos dos campos DNS. Essa nova simulação é feita utilizando um solver modificado que resolve apenas o campo de velocidade médio segundo a Eq. 3.4, onde $\mathbf{t}_{\text{pred}}$ é o tensor **t** predito pela RN.

$$(\langle \mathbf{v} \rangle \cdot \nabla)\langle \mathbf{v} \rangle + \mathbf{t}_{\text{pred}} = \nu \nabla^2 \langle \text{v} \rangle \tag{3.4}$$

Resumindo os passos de treinamento, tem-se:

- Selecionar caso de estudo
- Obter resultados nobres (DNS, LES) para treinamento
- Rodar simulações RANS no caso selecionado
- Extrair dados de entrada das simulações RANS
- Criar matrizes X e Y baseados nos dados obtidos
- Rodar código Python para treinamento da RN

Resumindo os passos de correção, tem-se:

- Rodar simulação RANS
- Pós processar resultados
- Criar matriz de dados de entrada X para a RN
- Rodar código Python da RN
- Pós processar novo **R** ou **t**
- Substituir **R** ou **t** antigo pelo novo no caso em estudo
- Re rodar simulação RANS

3.2 Tratamento da base de dados DNS

A matriz de alvos Y do treinamento da RN é composta por dados nobres das QdI que se quer prever. No caso aqui estudado, estes são componentes dos tensores **R** ou **t**. A fonte nobre utilizada para coletar essas informações foi a base de dados disponibilizada por PINELLI *et al.* (2010), o qual gerou e forneceu resultados de simulação DNS para o escoamento em duto de seção quadrada para diversos números de Re.

Foram utilizadas seis amostras de simulação DNS em duto de seção quadrada, assim como suas respectivas RANS, para diferentes Re, são eles 2200, 2400, 2600, 2900, 3200 e 3500.

A base de dados disponibiliza, para cada amostra de Re, nove campos de dados, sendo três para as componentes da velocidade média nas direções x (um ou U_x), y (vm ou U_y) e z (wm ou U_z), além de seis para as componentes de **R**, já que este é um tensor simétrico, uu ou R_{xx}, uv ou R_{xy}, uw ou R_{xz}, vv ou R_{yy}, vw ou R_{yz}, e ww ou R_{zz}.

A Fig. 3.1 mostra os campos DNS disponibilizados por PINELLI *et al.* (2010), para Re = 2200.

Como o escoamento em duto de seção quadrada apresenta domínio quadrado, com lados de mesmo comprimento, espera-se que os resultados DNS de seus campos, uma vez convergidos, apresentem aspecto simétrico com relação aos cortes médios horizontal e vertical, e também com relação às diagonais. A Fig. 3.2 mostra campos de discrepancia de U_x, para Re = 2200, calculados através da diferença entre o terceiro quadrante (3q) e os demais (1q, 2q e 4q). Pode-se observar que tal campo não apresenta as simetrias esperadas, o que é contra a lógica do escoamento. Uma vez que seu domínio é simétrico, a velocidade do escoamento na direção principal deveria ser necessariamente simétrica em relação aos cortes médios.

Essa mesma falha de simetria pode ser observada quando analisa-se os campos referentes às componentes de **R**. As Fig. 3.3 e 3.4 mostram os campos de R_{xy} e R_{xz} para Re = 2200. Estes deveriam ser, além de simétricos em si, simétricos entre si, apenas rotacionados. Pode-se observar que, apesar de possuírem formato semelhante, nenhuma das simetrias é apresentada.

Todo o processo de simulação RANS do em duto de seção quadrada e sua correção através da RN é feito para um domínio equivalente ao quadrante inferior esquerdo do escoamento, ou terceiro quadrante. Uma vez que tal escoamento é simétrico, não é necessário realizar simulação para a seção transversal completa do duto. Faz-se para um quadrante e espelha-se o resultado.

Entretanto, sabendo que a base de dados DNS não apresenta simetria em seus campos, a escolha do quadrante de DNS para treinamento influenciaria o resultado

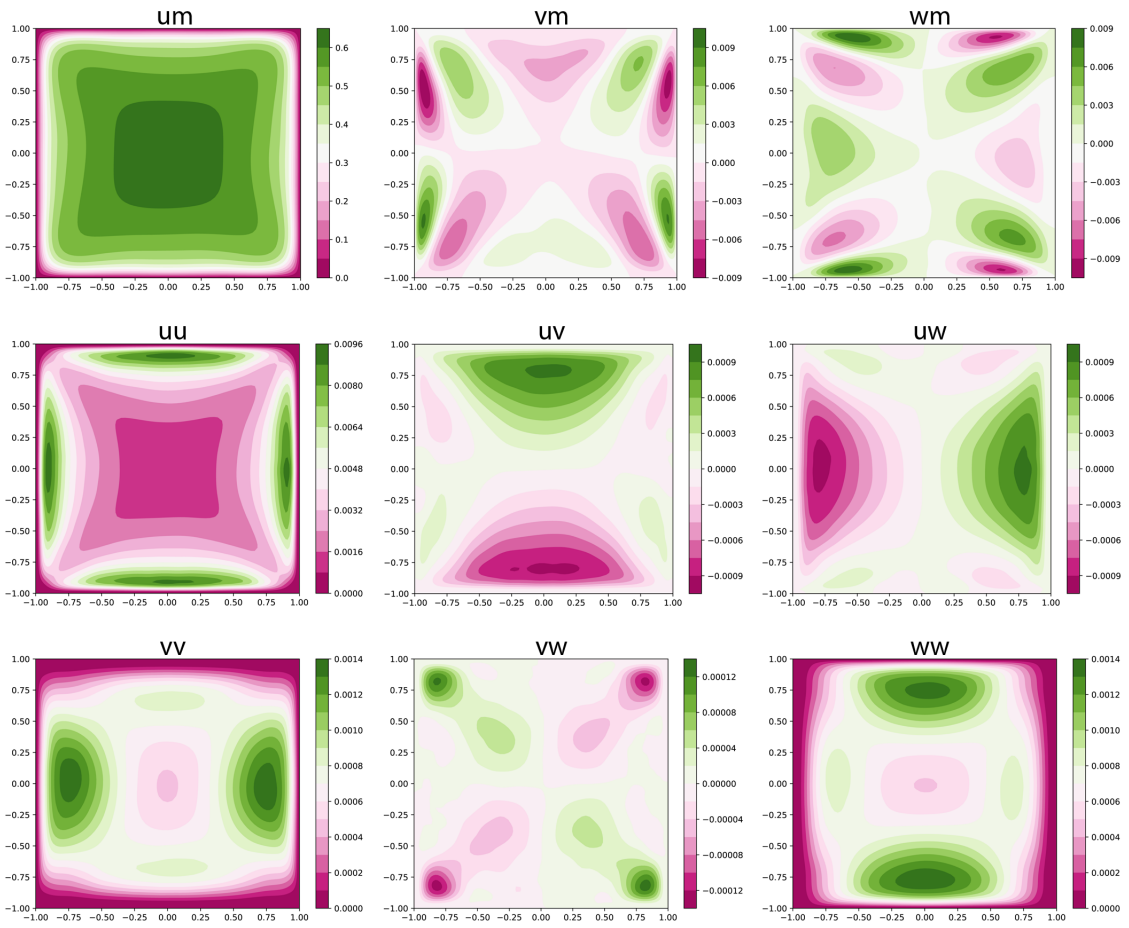


Figura 3.1: Campos DNS para Re = 2200

da previsão da RN.

Para estudar os efeitos da falta de simetria dos campos DNS, estes foram divididos em quatro quadrantes. São eles: superior esquerdo, primeiro quadrante; superior direito, segundo quadrante; inferior esquerdo, terceiro quadrante; e inferior direito, quarto quadrante.

Cada quadrante foi utilizado separadamente no processo de treinamento da RN e correção de RANS, descrito anteriormente, na intenção de investigar sua influência sobre os resultados do procedimento.

Os conjuntos de campos referentes aos quadrantes unitários serão tratados como Q1.

3.2.1 Simulando dados DNS melhor convergidos

Os dados DNS são obtidos de sua simulação, convergida em um tempo τ. Sabe-se que, quanto mais tempo de simulação, melhor convergidos são os resultados. Pela demanda computacional e temporal de uma simulação DNS, fica impraticável realizá-la até sua convergência em muitas faixas de tempo diferentes para investigar

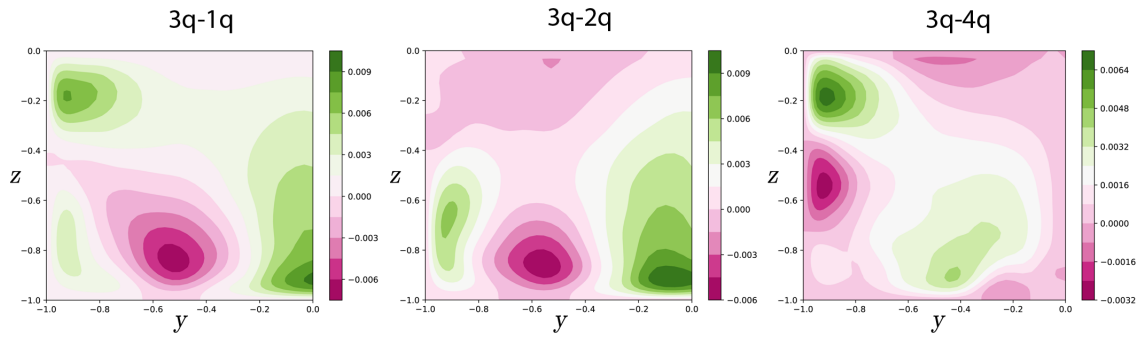


Figura 3.2: Campos de discrepancia de U_x para Re = 2200

sua influência nos resultados da RN.

Na intenção de simular resultados DNS melhor convergidos, aplicou-se um tratamento aos dados disponíveis por PINELLI *et al.* (2010).

Uma vez que os campos foram divididos em quadrantes, para se adequar ao domínio de simulação, combinações entre eles foram feitas visando simular resultados DNS para tempos de simulação maiores.

Se a base de dados disponibiliza campos convergidos em um tempo τ, a combinação desses quadrantes, isso é, somando-se n quadrantes diferentes e dividindo o resultado desta soma por n, nos forneceria um pseudo campo DNS convergido para um tempo $n\tau$.

Ou seja, as combinações de quadrantes dois a dois geram seis novos conjuntos de campos DNS convergidos para um tempo 2τ. Passaremos a nos referir a esses conjuntos de dados como Q2.

As combinações de quadrantes três a três geram quatro novos conjuntos de campos DNS convergidos para um tempo 3τ. Passaremos a nos referir a esses conjuntos de dados como Q3.

Já a combinação de quadrantes quatro a quatro é única e gera um novo conjunto de campos DNS convergidos para um tempo 4τ. Passaremos a nos referir a esse conjunto de dados como Q4.

Todas essas combinações entre quadrantes fazem com que os campos se aproximem mais do ideal para o escoamento proposto, a simetria, conforme aumentamos n.

Como a física do problema nos diz que os campos deveriam ser completamente simétricos, outra manipulação foi feita para obter um último conjunto de resultados. Este simula a simetria com relação às diagonais, e é obtido dos resultados de 4q, a partir da média dos valores espelhados com relação a diagonal do quadrante. Como esta operação supõe a simetria dentro das metades (1/2) diagonais do terceiro quadrante (1/4) de uma combinação que já usa quatro quadrantes ($n = 4$), este conjunto de resultados nos traz campos que simulam a simetria com o oitavo (1/2

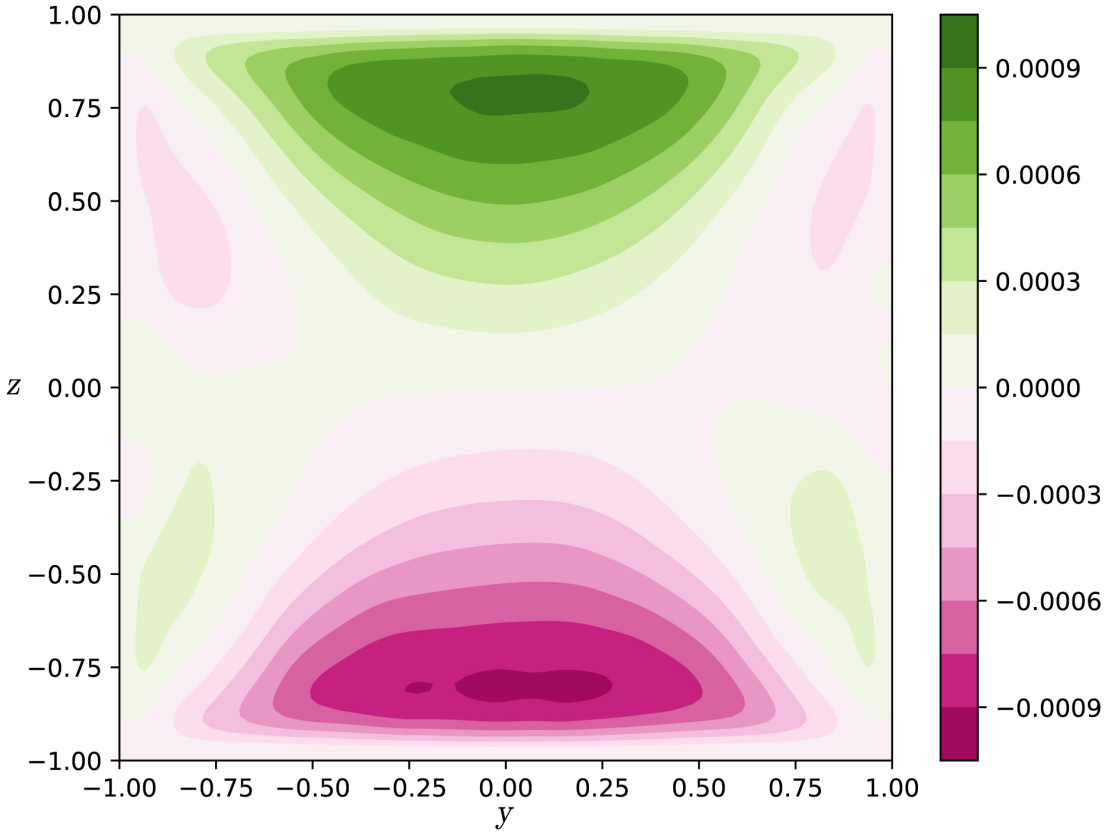


Figura 3.3: Campo R_{xy}, Re = 2200

$\times$ 1/4) de domínio para um tempo 8τ (o dobro do 4q), e por tanto passaremos a nos referir a esse conjunto de dados como 8q.

A tabela 3.2 mostra todas as combinações de quadrantes feitas para cada conjunto mencionado, Q1, Q2, Q3, Q4 e 8q.

A Fig. 3.5 mostra a evolução da melhora do campo DNS R_{yz}, para o terceiro quadrante, a partir do tratamento proposto, para Re = 2200.

Os resultados 8q são o DNS perfeito, o máximo de simetria que esses campos podem apresentar a partir do tratamento dos resultados originais. Portanto, estes serão utilizados como referencia para os dados DNS quando convir, como no cálculo de erro global por exemplo.

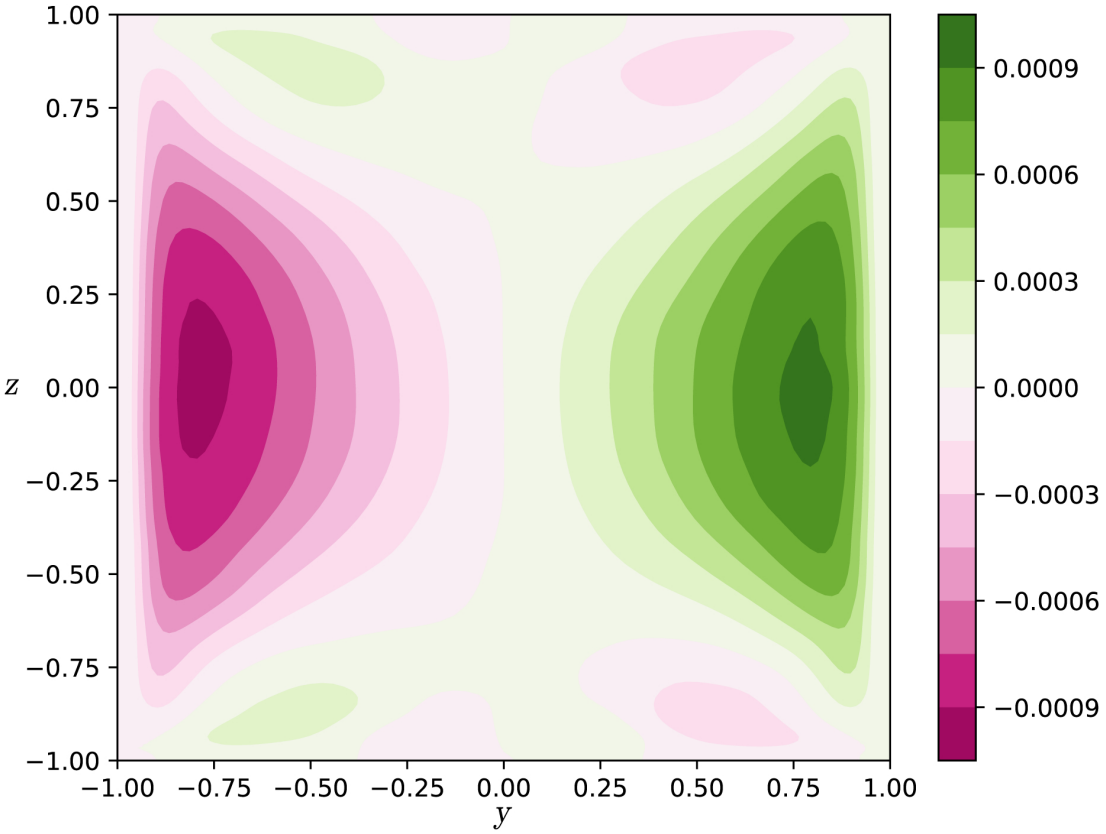


Figura 3.4: Campo R_{xz}, Re = 2200

Tabela 3.2: Combinações de quadrantes

Grupos de Combinações	Combinações de Quadrante
Q1 (Combinações 1 a 1)	1q 2q 3q 4q
Q2 (Combinações 2 a 2)	12q 13q 14q 23q 24q 34q
Q3 (Combinações 3 a 3)	123q 124q 134q 234q
Q4	1234q
8q	1234*

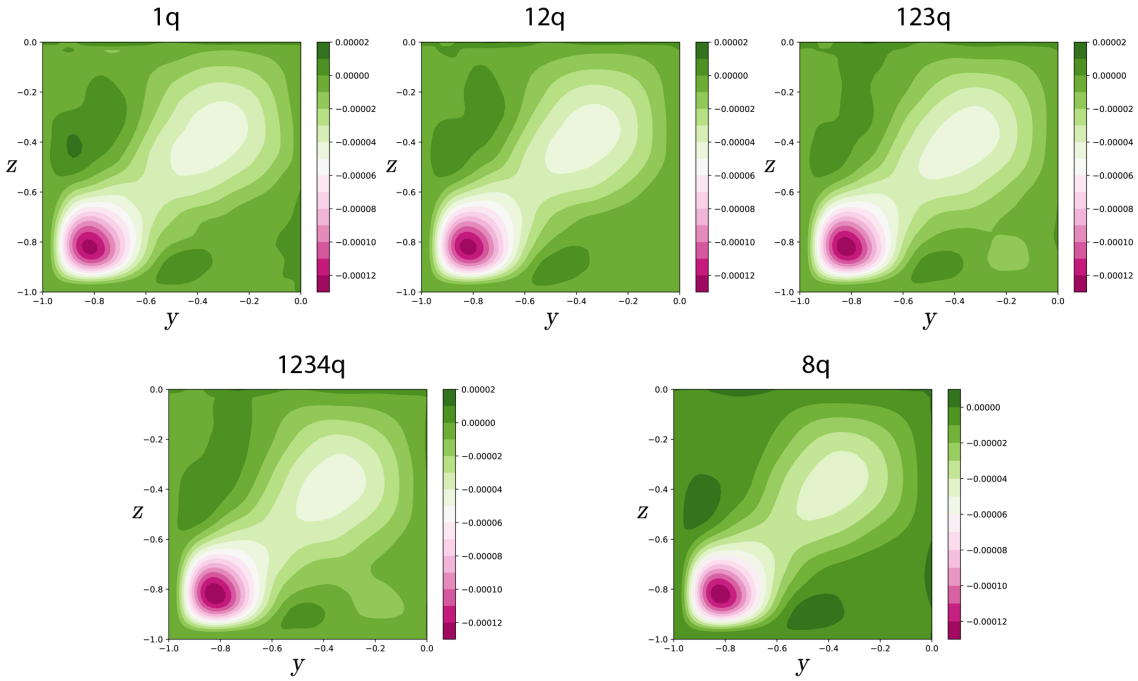


Figura 3.5: Evolução do campo DNS R_{yz} sob tratamento de dados

Capítulo 4

Resultados

4.1 Previsão dos campos de velocidade

Na análise da previsão dos campos de velocidade, serão utilizados um representante de cada grupo de casos que constam na tabela 3.2. Estes serão o 1q do Q1, 12q do Q2, 123q do Q3, 1234q do Q4, e o 8q, além dos resultados RANS e DNS como referência.

Os resultados dos campos previstos podem ser vistos nas Fig. 4.1 a 4.6

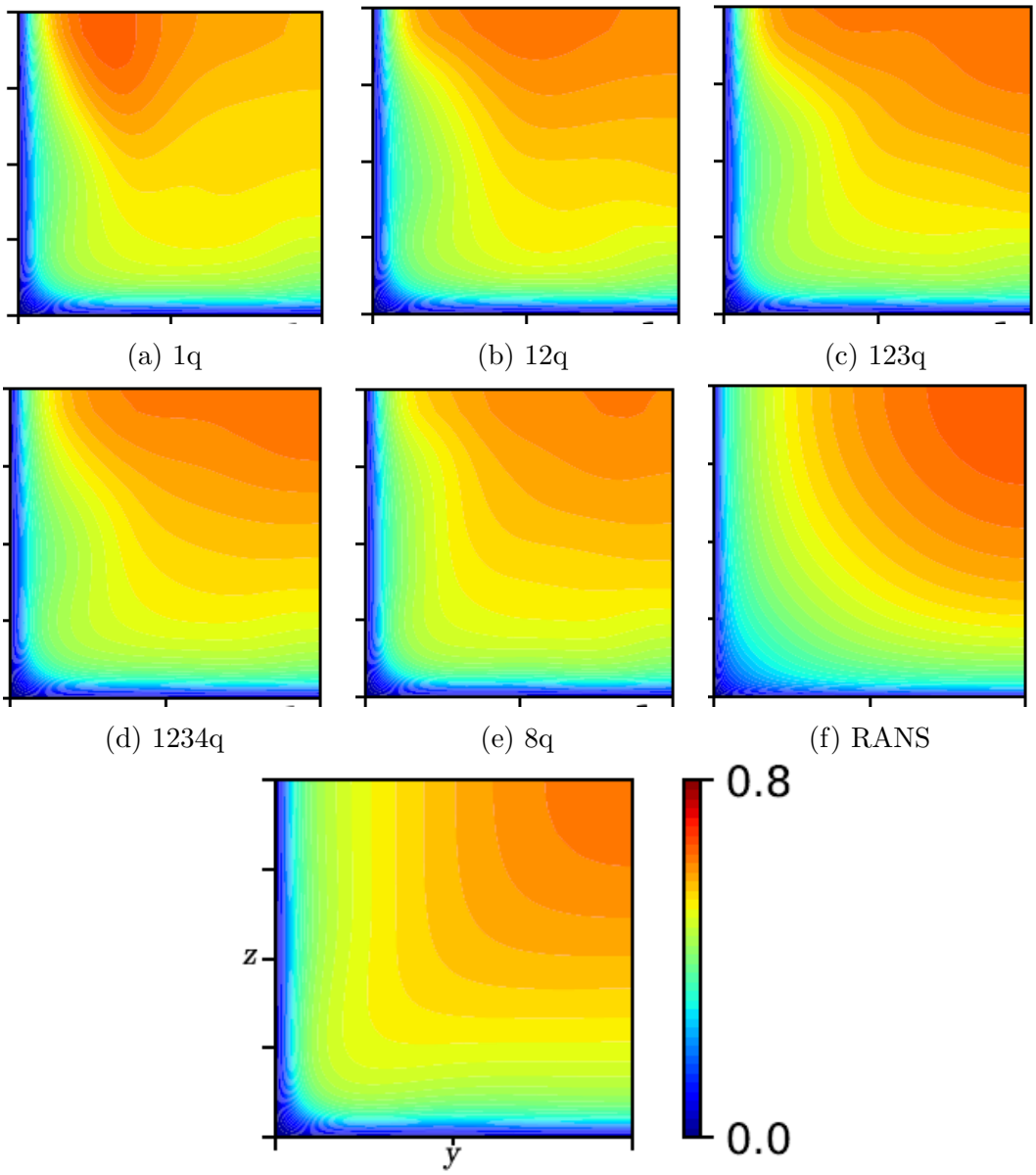
4.1.1 Correção de R

A Fig. 4.1 mostra os resultados para o campo U_x. Pode-se observar que, com a adição de quadrantes aos dados, o resultado teve considerável melhora. Em 1q, Fig. 4.1a, a maior magnitude de velocidade se encontra afastada do que seria o centro do escoamento (canto superior direito do domínio), onde deveria estar. O resultado de 8q, Fig. 4.1e, representa melhor o que se espera, quando comparado aos demais casos de previsão. Com relação a RANS, Fig. 4.1f, 8q consegue capturar melhor o aspecto quadrado do escoamento.

É sabido que o método **R** não apresenta bons resultados na previsão de U_x. Entretanto, a Fig. 4.1 nos mostra que, com o tratamento dos dados, o resultado pôde ser bastante melhorado.

A Fig. 4.2 mostra os resultados para o campo U_y. Pode-se observar que, com a adição de quadrantes aos dados, o resultado teve melhora notável. Em 1q, Fig. 4.2a, a área negativa é menor que DNS, Fig. 4.2g, e vai ganhando forma nos casos seguintes. O resultado de 8q, Fig. 4.2e, representa melhor o que se espera, quando comparado aos demais. Com relação a RANS, Fig. 4.2f, 8q consegue capturar melhor a recirculação do escoamento.

A Fig. 4.3 mostra os resultados para o campo U_z. Pode-se observar que, com a adição de quadrantes aos dados, o resultado teve melhora. Em 1q, Fig. 4.3a, a área

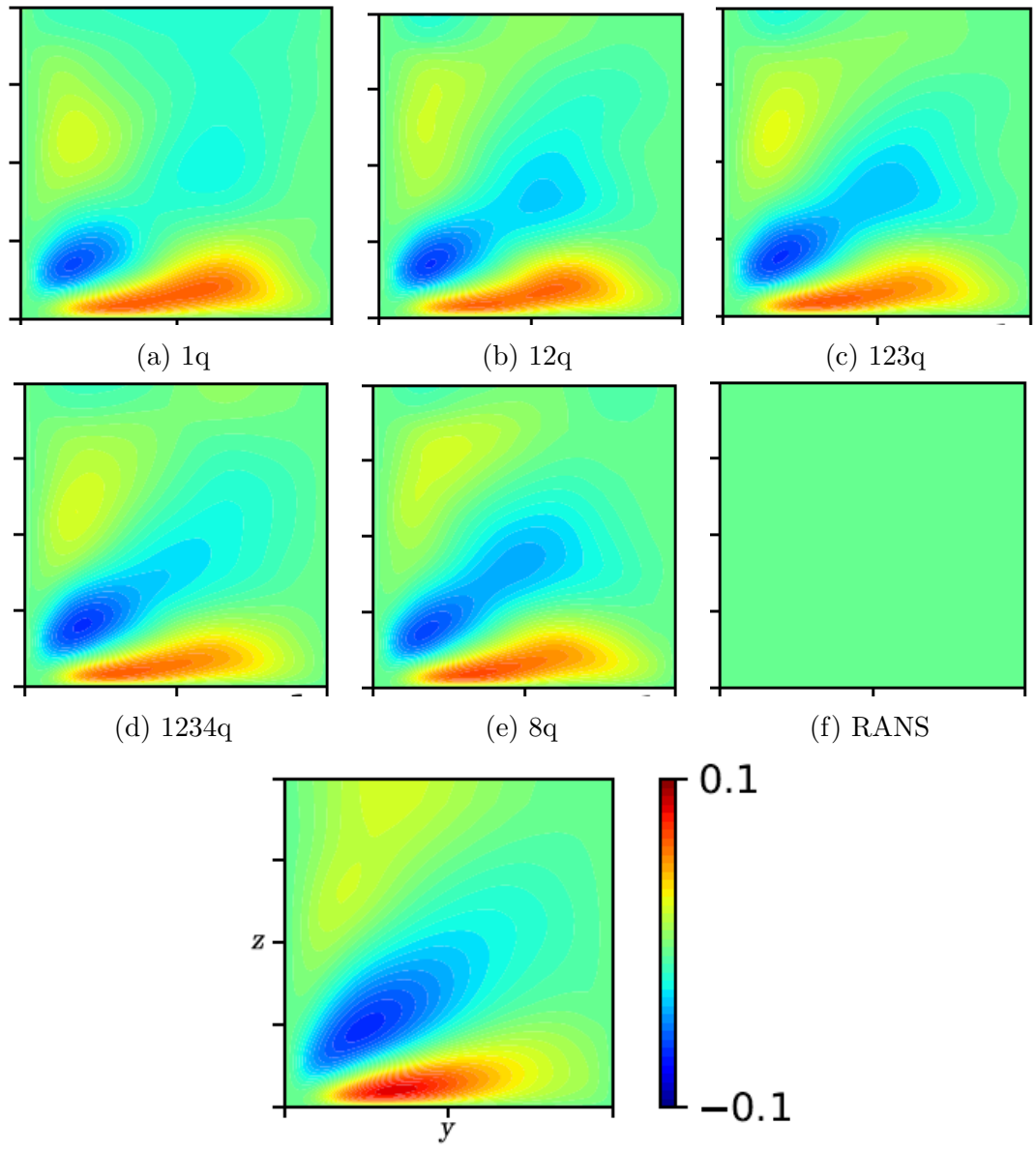


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.1: Previsão de U_x para $Re = 3200$ através do método **R**

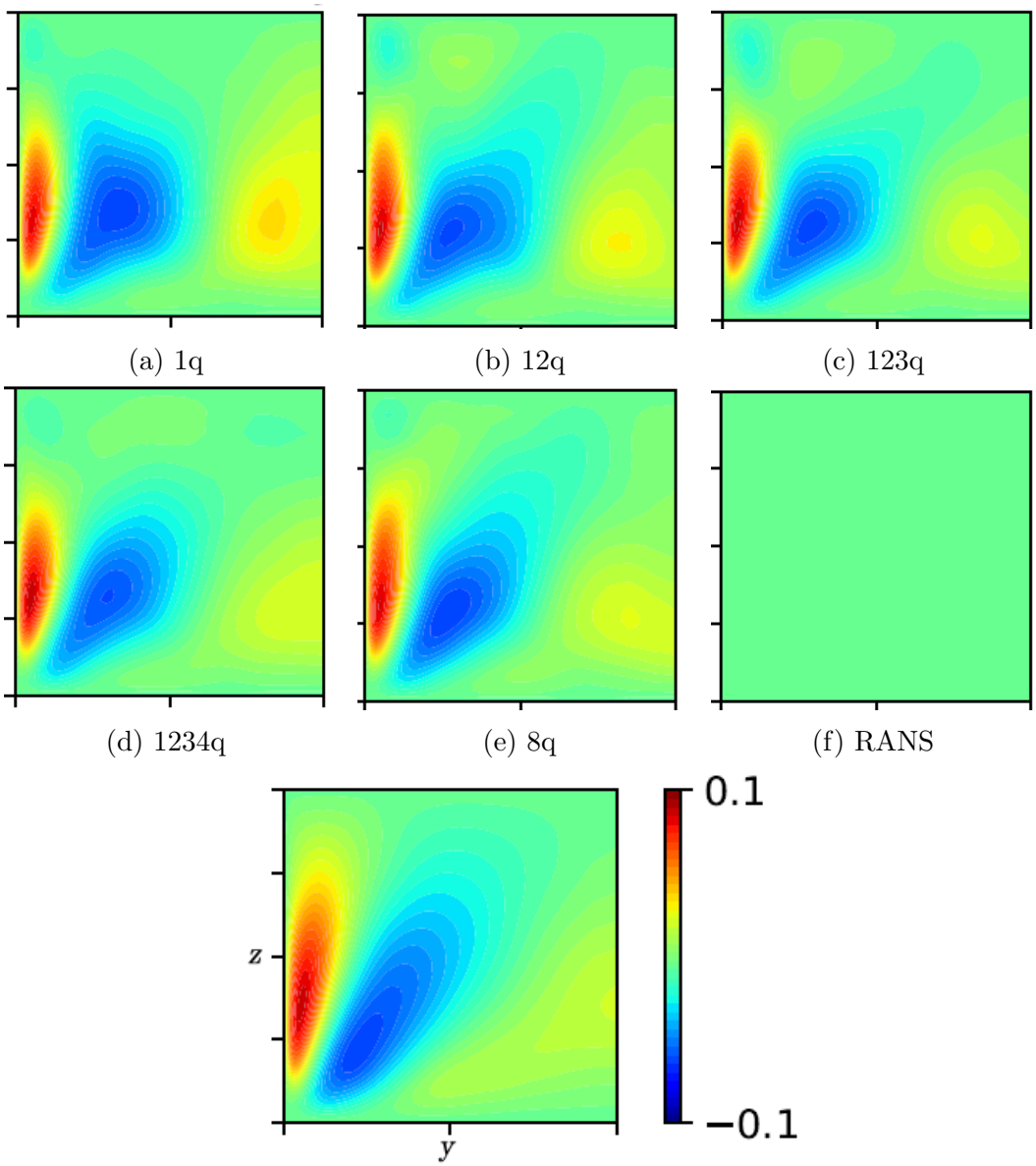


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.2: Previsão de U_y para $Re = 3200$ através do método **R**



(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.3: Previsão de U_z para $Re = 3200$ através do método **R**

negativa tem aspecto diferente de DNS, Fig. 4.3g, além de uma área amarelada intensa à direita do domínio, o que melhora nos casos seguintes. O resultado de 8q, Fig. 4.3e, representa melhor o que se espera, quando comparado aos demais. Com relação a RANS, Fig. 4.3f, 8q consegue capturar melhor a recirculação do escoamento.

4.1.2 Correção de t

A Fig. 4.4 mostra os resultados para o campo U_x. É sabido que o método **t** apresenta bons resultados na previsão de U_x. Em 1q, Fig. 4.4a, o campo já é bastante parecido ao DNS, Fig. 4.4g. Ainda assim, pode-se observar que, com a

adição de quadrantes aos dados, o resultado teve melhora gradativa com relação ao aspecto quadrado do escoamento e as faixas de velocidade. É notório que o 8q, Fig. 4.4e, é muito superior a RANS, Fig. 4.4f, nas características mencionadas.

A Fig. 4.5 mostra os resultados para o campo U_y. Em 1q, Fig. 4.5a, o campo negativo é maior que o apresentado em DNS, Fig. 4.5g, além do positivo ser menor. Pode-se observar que, com a adição de quadrantes aos dados, o resultado teve melhora gradativa com relação ao aspecto dos campos e magnitude. É notório que o 8q, Fig. 4.5e é o que melhor representa DNS, além de ser superior a RANS, Fig. 4.5f.

A Fig. 4.6 mostra os resultados para o campo U_z. Para este campo a melhora parece subjetiva, não sendo fácil de se notar diferença considerável entre os resultados a partir da simples observação. Mesmo assim, os casos previstos apresentam melhor resultado que RANS, Fig. 4.6f.

4.2 Erro Local

Na análise dos campos de erro local, serão utilizados um representante de cada grupo de casos que constam na tabela 3.2. Estes serão o 1q do Q1, 12q do Q2, 123q do Q3, 1234q do Q4, e o 8q, além dos resultados RANS como referência.

O campo de erro local é calculado como mostra as Eq. 4.1 e 4.2, sendo a primeira para o campo U_x e a segunda para U_y e U_z, onde $\overline{U}$ é a velocidade média de seção.

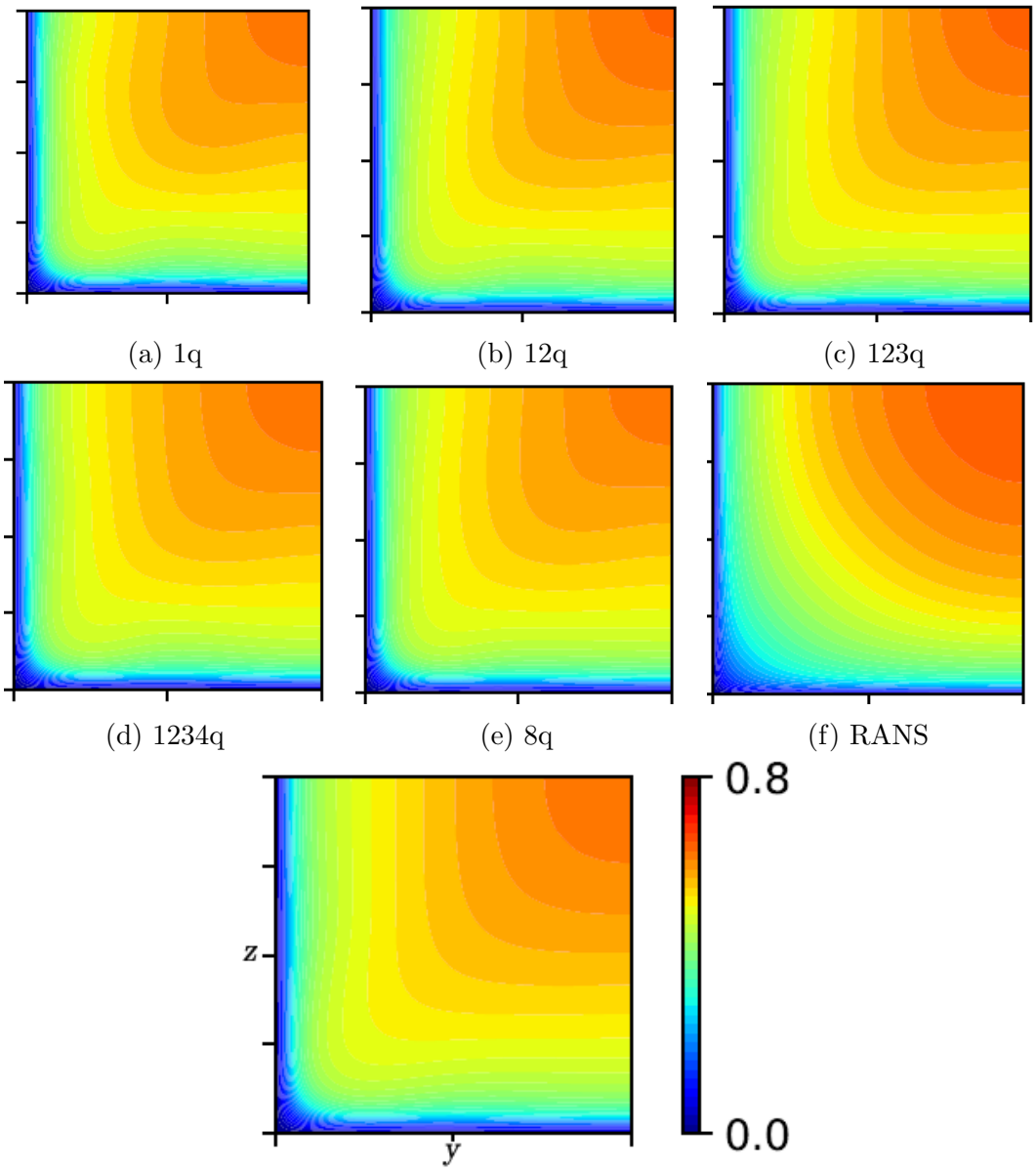
$$E_x = \frac{\sqrt{(U_x^{DNS} - U_x^{RANS})^2}}{\overline{U}} \tag{4.1}$$

$$E_j = 10 \times \frac{\sqrt{(U_j^{DNS} - U_j^{RANS})^2}}{\overline{U}} \tag{4.2}$$

Os resultados dos campos previstos podem ser vistos nas Fig. 4.7 a 4.12.

4.2.1 Correção de R

A Fig. 4.7 mostra os resultados para o campo U_x. Para 1q, Fig. 4.7a, pode-se observar que o erro local apresenta regiões onde sua magnitude excede a escala, no canto superior esquerdo do domínio. É possível notar que há erro expressivo próximo ao centro do escoamento, canto superior direito do domínio. O acréscimo de quadrantes diminui gradativamente o erro nestas áreas apontadas. O 8q, Fig. 4.7e, é o que melhor performa, apresentando maior parte do campo com valores baixos, exceto uma pequena área superior esquerda, onde o erro persiste, porém com valores menores que os resultados anteriores. Analisando RANS, Fig. 4.7f,

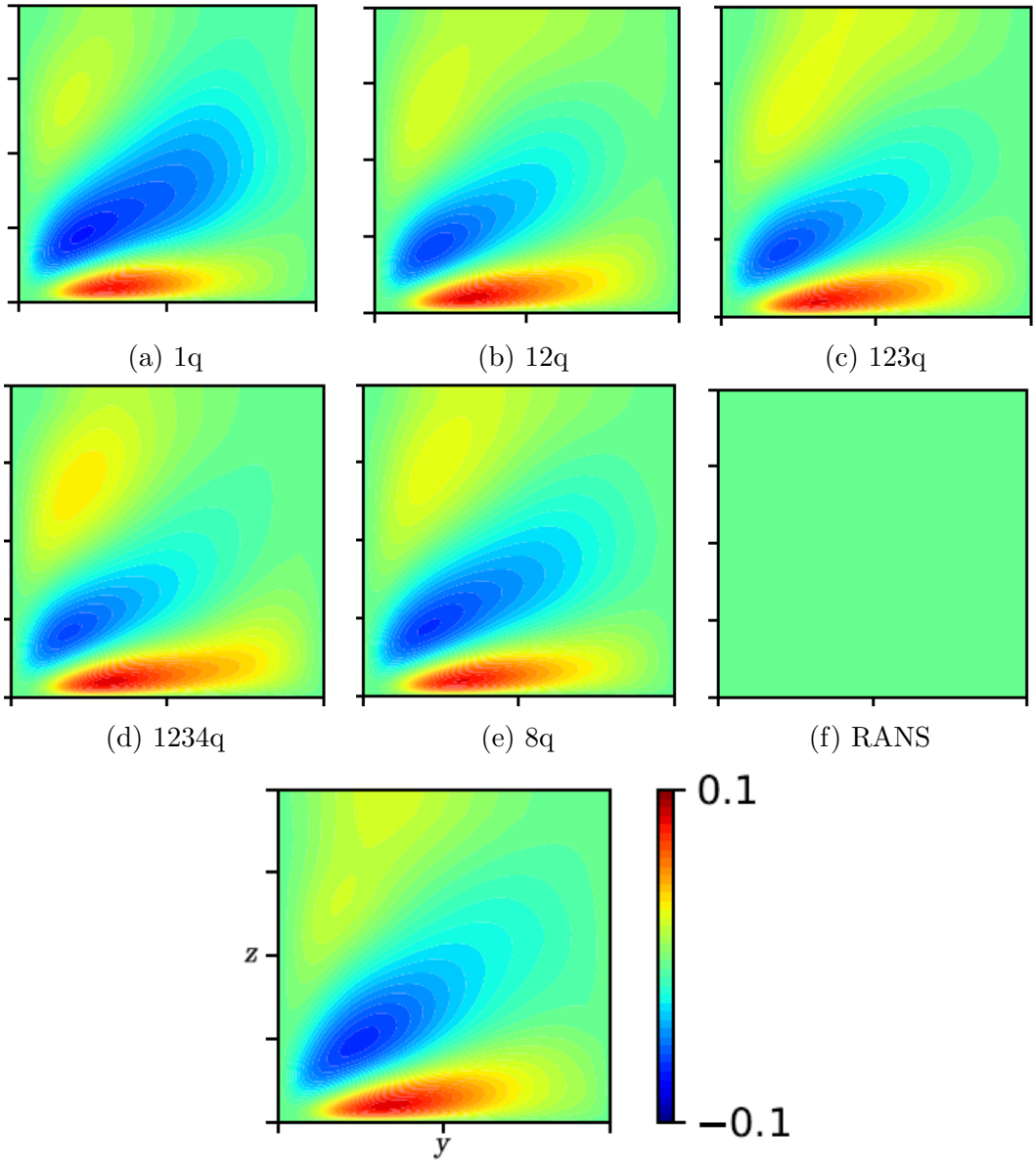


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.4: Previsão de U_x para $Re = 3200$ através do método **t**

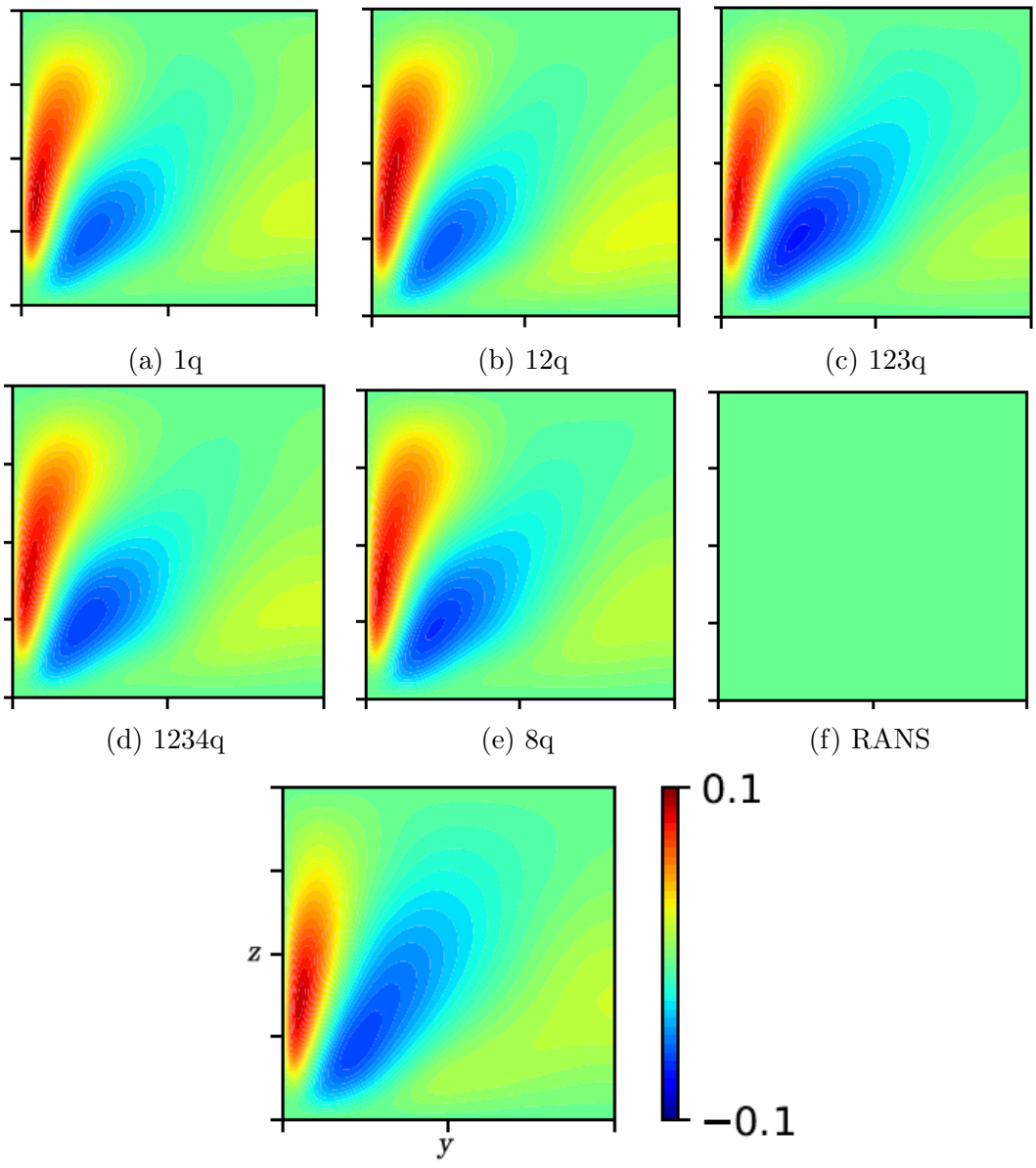


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.5: Previsão de U_y para $Re = 3200$ através do método **t**

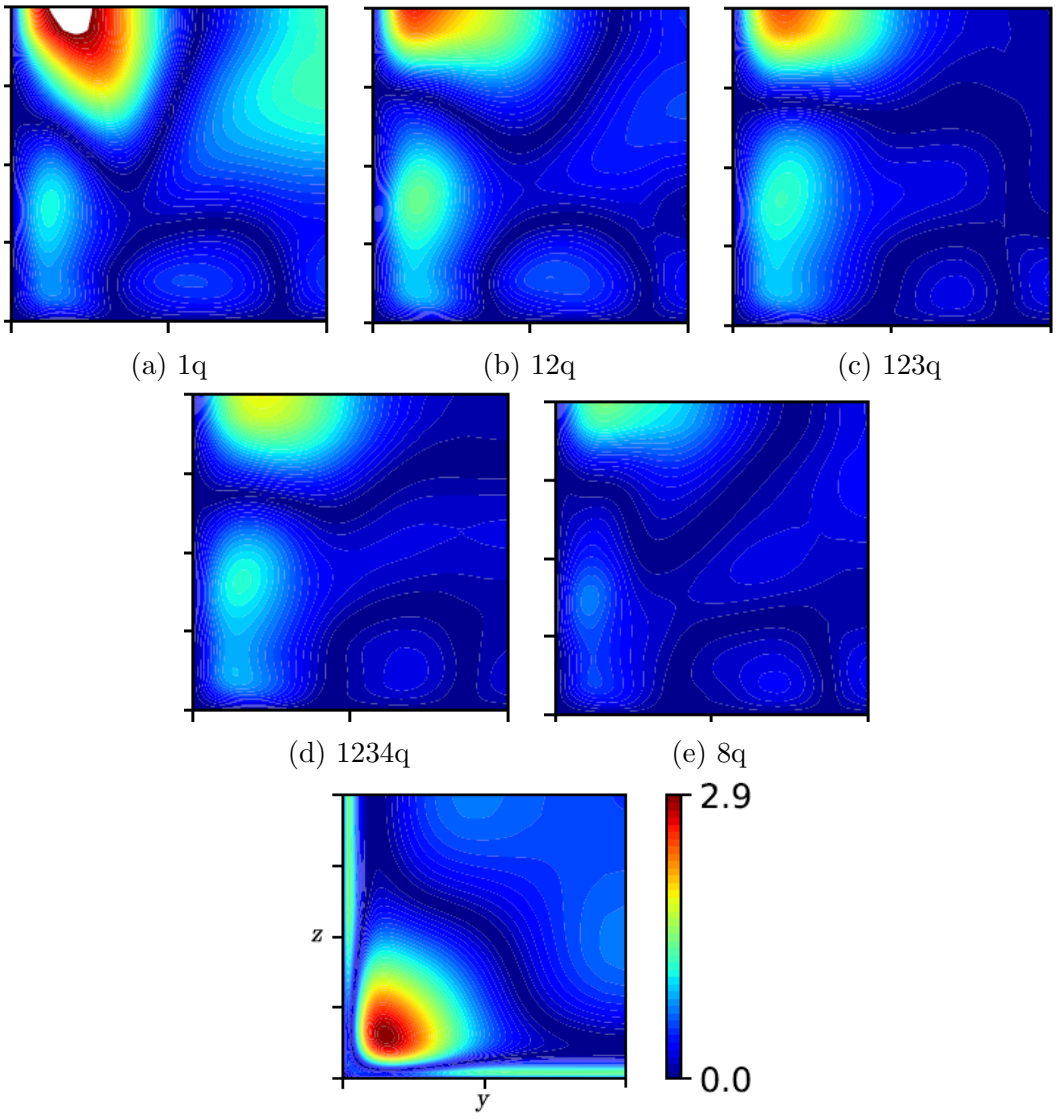


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q (f) RANS

(g) DNS

Figura 4.6: Previsão de U_z para $Re = 3200$ através do método **t**



(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q

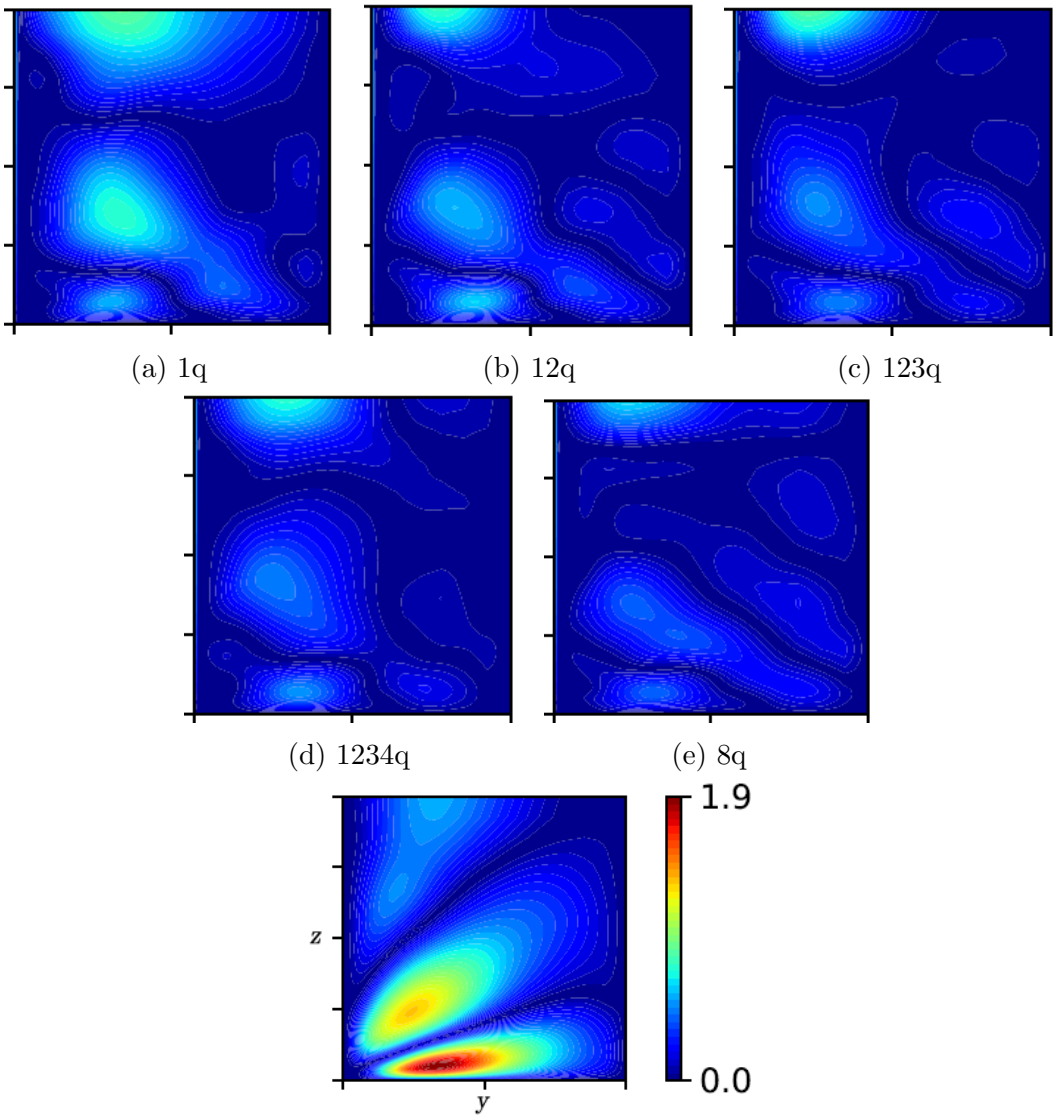
(f) RANS

Figura 4.7: Erro local da previsão de U_x para $Re = 3200$ através do método **R**

observa-se que este erra bastante na área inferior esquerda, onde atinge o pico de magnitude da escala. As regiões de parede também apresentam erro relevante.

A Fig. 4.8 mostra os resultados para o campo U_y. Para 1q, Fig. 4.8a, pode-se observar que o erro local apresenta regiões esverdeadas. O acréscimo de quadrantes diminui gradativamente o erro nestas regiões. O 8q, Fig. 4.8e, é o que melhor performa, apresentando maior parte do campo com valores baixos, exceto uma pequena área superior esquerda, onde o erro persiste, porém com valores menores que os resultados anteriores. Analisando RANS, Fig. 4.8f, observa-se que este erra bastante na área inferior esquerda, região de recirculação, onde atinge o pico de magnitude da escala. Isso acontece pois a simulação RANS utilizada é incapaz de capturar tais características do escoamento.

A Fig. 4.9 mostra os resultados para o campo U_z. Para 1q, Fig. 4.9a, pode-

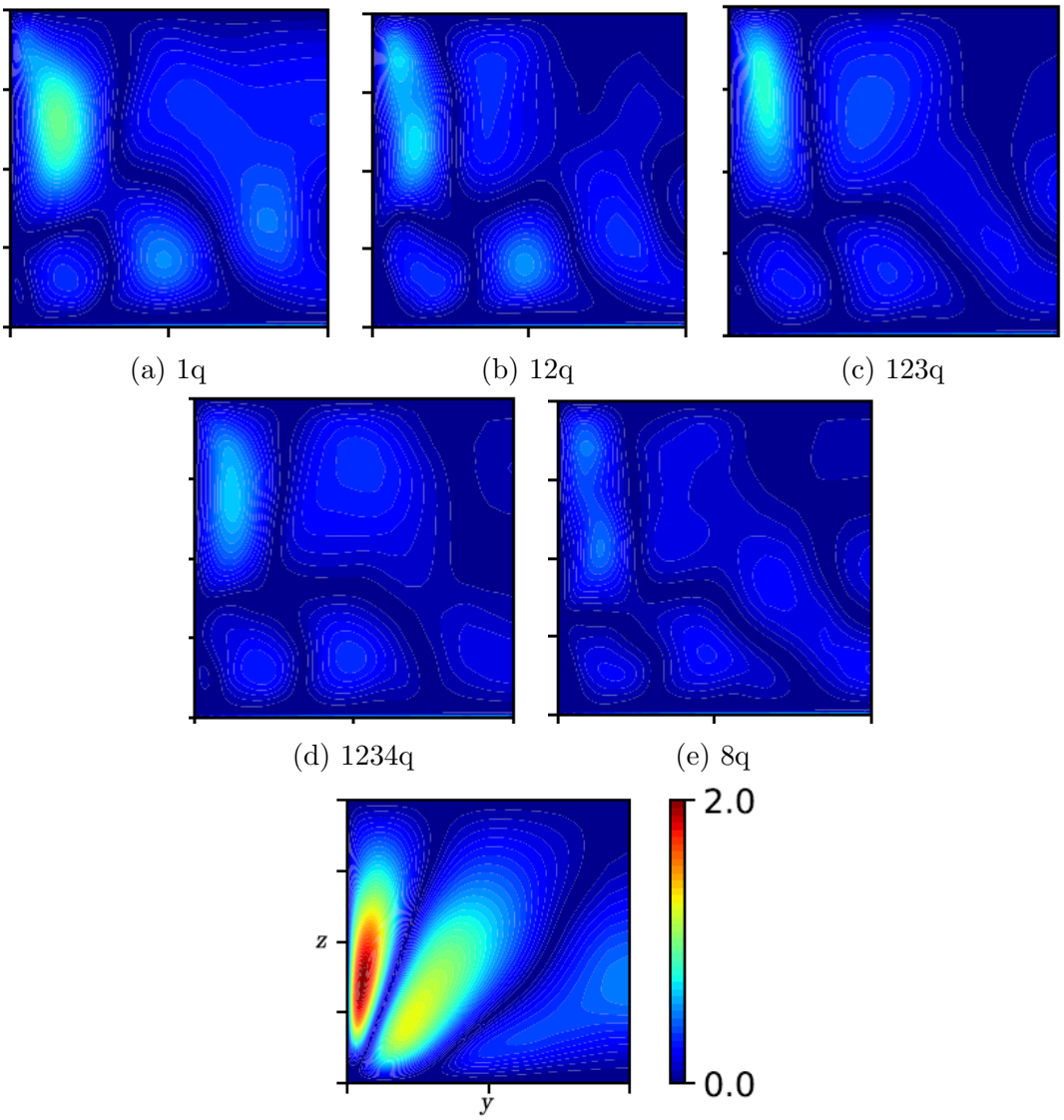


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q

(f) RANS

Figura 4.8: Erro local da previsão de U_y para $Re = 3200$ através do método **R**



(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q

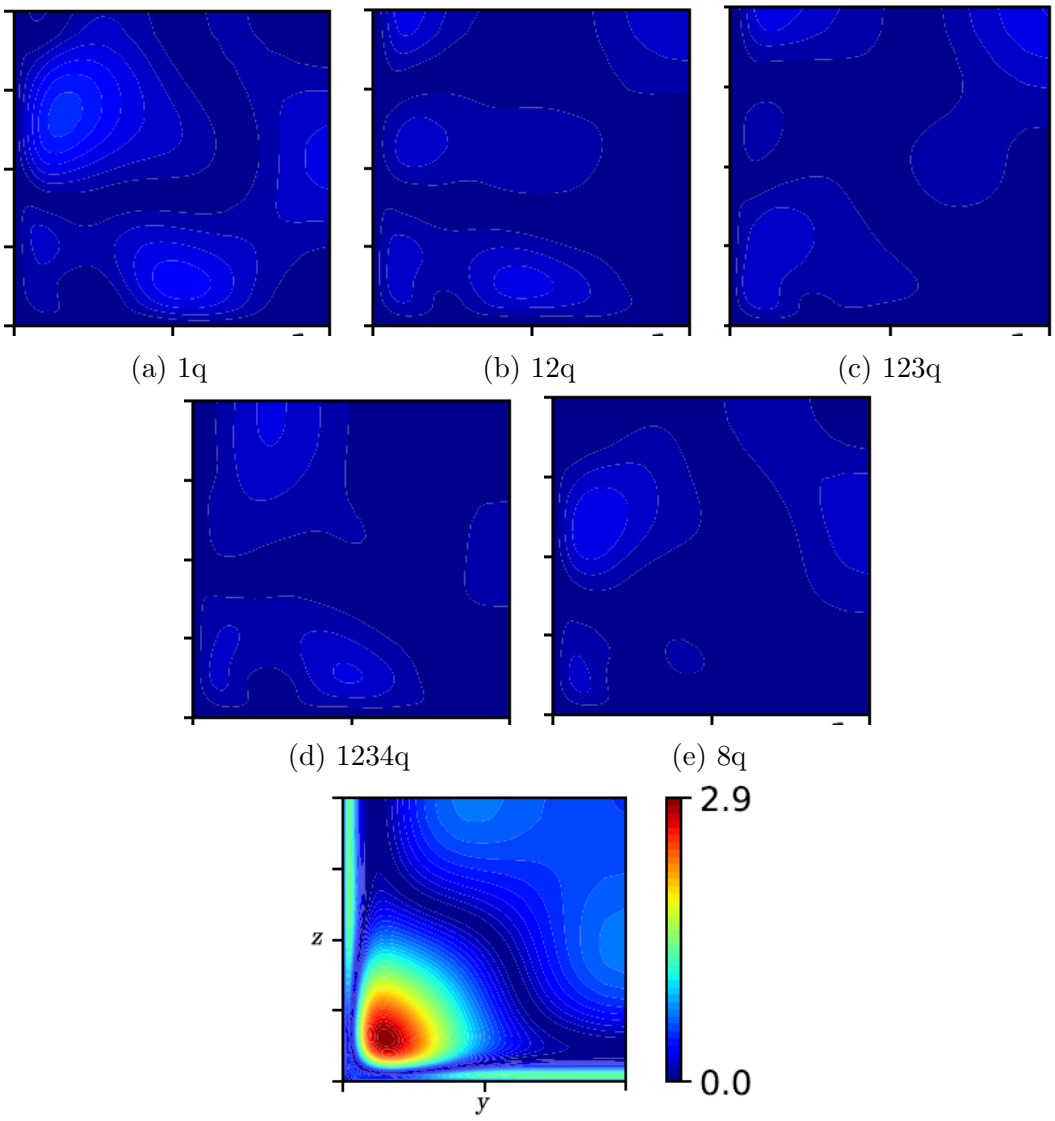
(f) RANS

Figura 4.9: Erro local da previsão de U_z para $Re = 3200$ através do método **R**

se observar que o erro local apresenta uma região esverdeada. O acréscimo de quadrantes diminui gradativamente o erro nesta área. O 8q, Fig. 4.9e, é o que melhor performa, apresentando maior parte do campo com valores abaixo de 1. Analisando RANS, Fig. 4.9f, observa-se que este erra bastante na área inferior esquerda, região de recirculação, onde atinge o pico de magnitude da escala, semelhante ao erro local da direção y.

4.2.2 Correção de t

A Fig. 4.10 mostra os resultados para o campo U_x. Para 1q, Fig. 4.10a, pode-se observar que o erro local já é bom, o que é esperado. O acréscimo de quadrantes diminui gradativamente o erro. O 8q, Fig. 4.10e, é o que melhor performa,



(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q

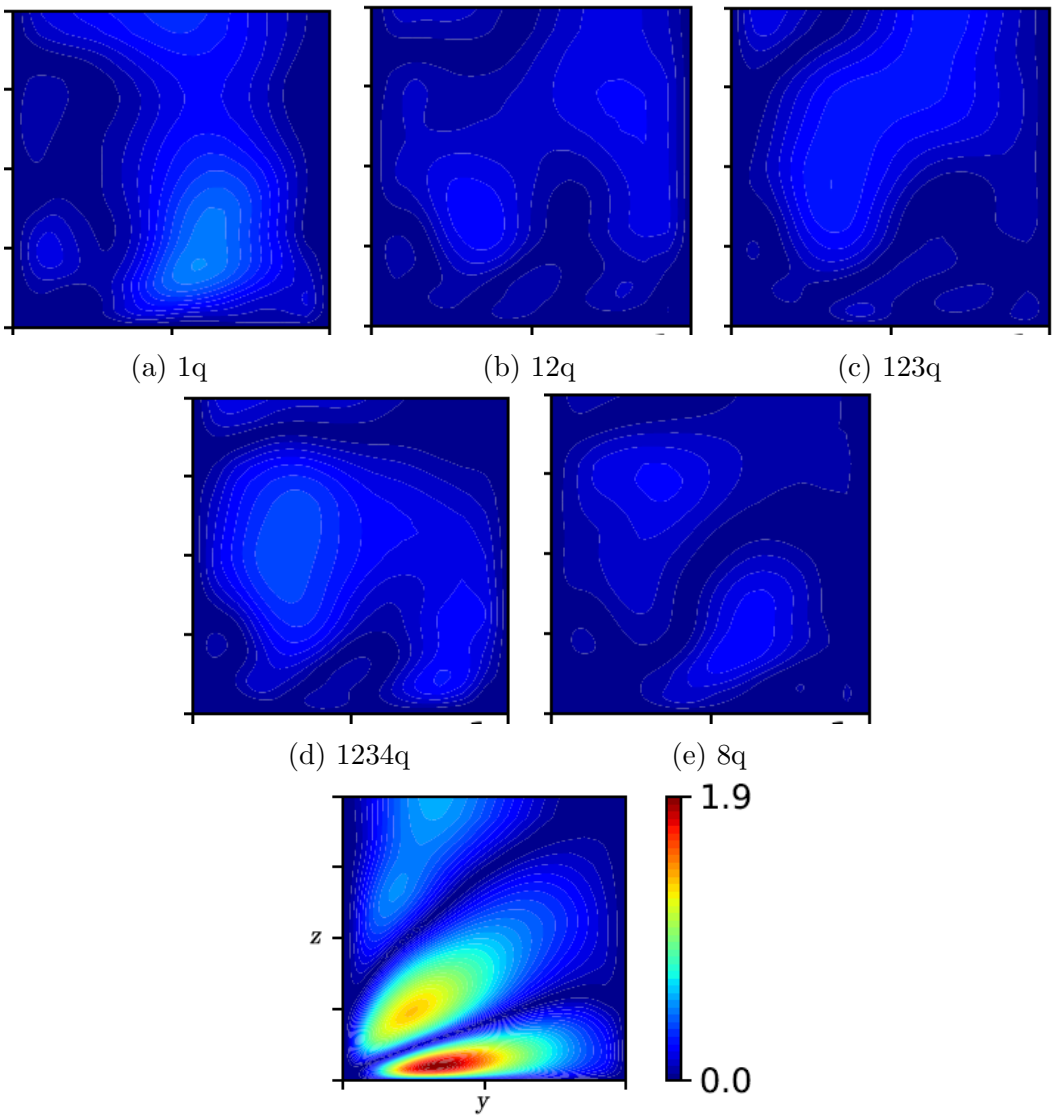
(f) RANS

Figura 4.10: Erro local da previsão de U_x para $Re = 3200$ através do método **t**

apresentando maior parte do campo com valores próximos de 0.

A Fig. 4.11 mostra os resultados para o campo U_y. Para 1q, Fig. 4.11a, pode-se observar que o erro local é baixo, mas apresenta uma região clara. Esta diminui com o acréscimo de quadrantes, que diminui gradativamente o erro. O 8q, Fig. 4.11e, é o que melhor performa, apresentando boa parte do campo com valores próximos de 0.

A Fig. 4.12 mostra os resultados para o campo U_z. Para 1q, Fig. 4.12a, pode-se observar que o erro local apresenta região clara, com valor próximo de 1, na região superior esquerda. Esta diminui com o acréscimo de quadrantes até o 123q, Fig. 4.12c. Os resultados seguintes apresentam piora na região mencionada, porém melhoram no restante do domínio.

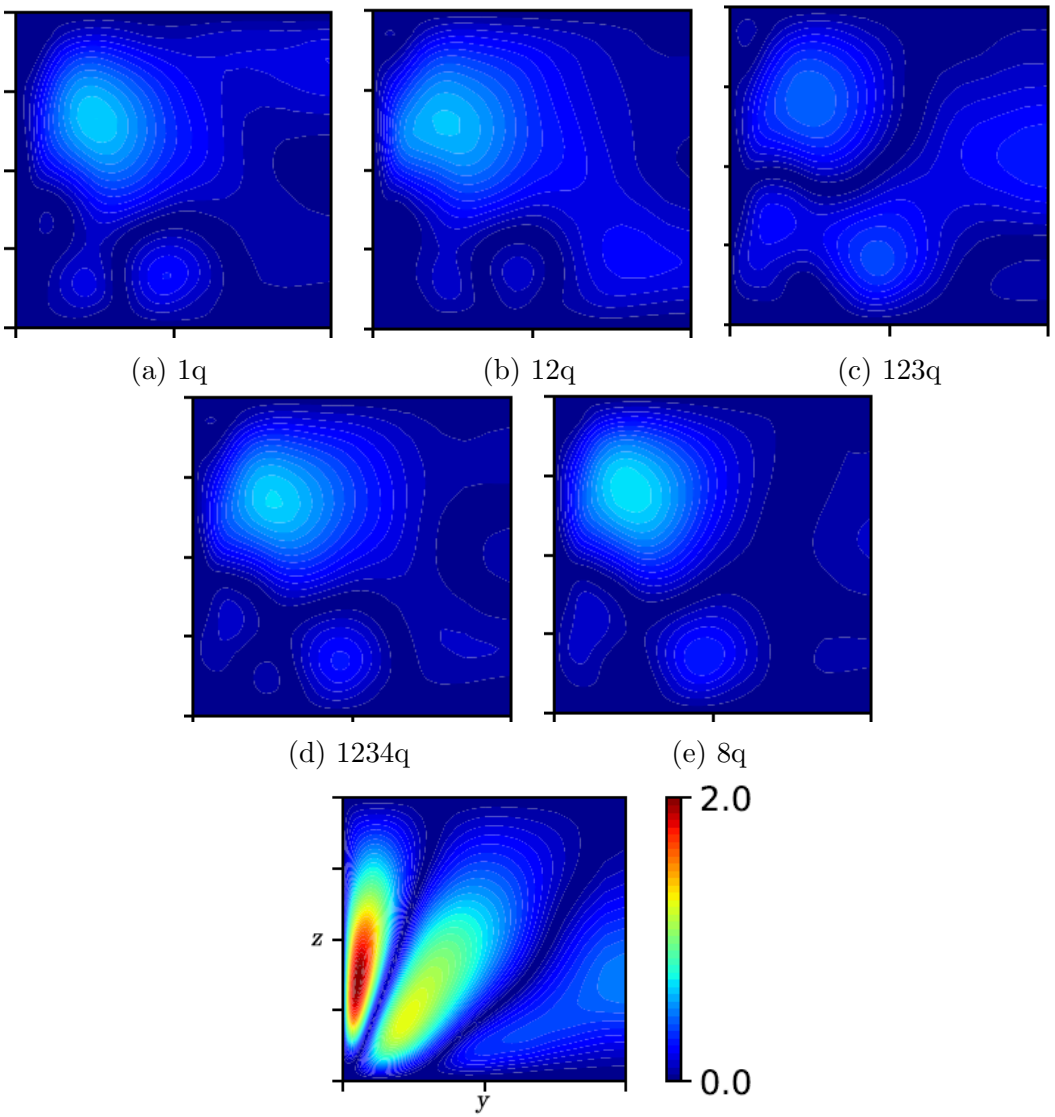


(a) 1q (b) 12q (c) 123q

(d) 1234q (e) 8q

(f) RANS

Figura 4.11: Erro local da previsão de U_y para $Re = 3200$ através do método **t**

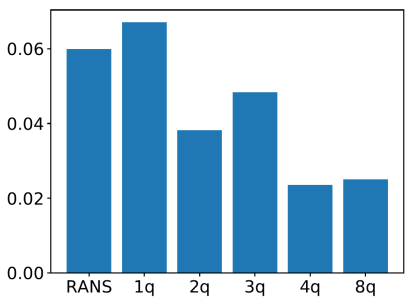


(a) 1q (b) 12q (c) 123q

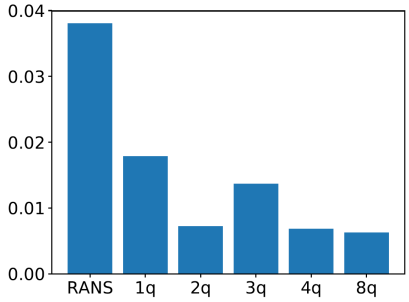
(d) 1234q (e) 8q

(f) RANS

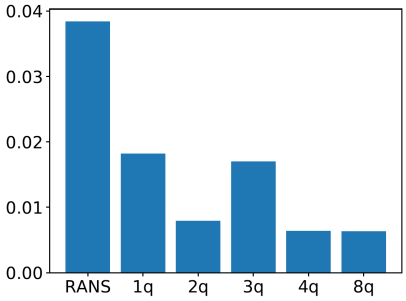
Figura 4.12: Erro local da previsão de U_z para $Re = 3200$ através do método **t**



(a) U_x



(b) U_y



(c) U_z

Figura 4.13: Erro global da previsão de **R** do grupo Q1 para $Re = 3200$

4.3 Erro Global

O erro global é calculado através da integral de área dos campos de erro local, e sua escala é definida de 0 (0%) a 1 (100%).

As Fig. 4.13 a 4.20 são gráficos de erro global dos campos de velocidade, que mostram comparativos entre os resultados obtidos na previsão da RN para $Re = 3200$ a partir da utilização dos diferentes quadrantes no treinamento, tomando 8q como referência no cálculo do erro, além de mostrarem também os resultados do erro RANS e da previsão de 8q como referências comparativas.

Os resultados estão dispostos em subseções por grupo de casos, a seguir.

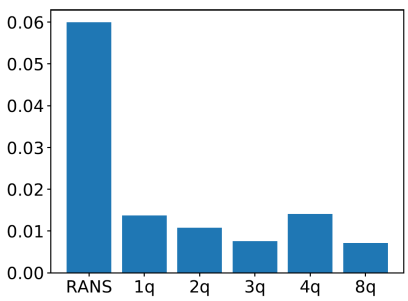
4.3.1 Q1

Na intenção de saber a influência da escolha do quadrante da base de dados DNS na previsão da RN, utilizou-se todas as opcões de quandrantes do grupo Q1, um a um, no treinamento da mesma, tanto para o método da correção de **R** quanto de **t**.

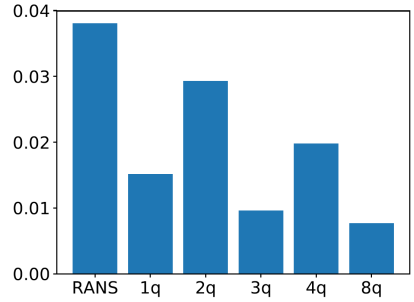
Os resultados para o método **R** estão dispostos na Fig. 4.13.

Para U_x, Fig. 4.13a, pode-se observar que, com exceção do 1q, todos os resultados apresentam menores erros que RANS. O quadrante que apresentou melhores resultados foi o 4q, praticamente empatado com o 8q.

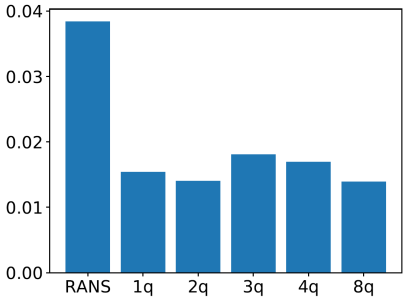
Na Fig. 4.13b tem-se os resultados para U_y. Novamente, todos os quadrantes



(a) U_x



(b) U_y



(c) U_z

Figura 4.14: Erro global da previsão de **t** do grupo Q1 para $Re = 3200$

se sairam melhores que RANS, o que é esperado, uma vez que o resultado RANS utilizado não capturam o escoamento secundário. Os quadrantes que se saíram melhor foram o 2q e o 4q, com erros abaixo de 1%.

Na Fig. 4.13c, com relação a U_z, a análise é semelhante a de U_y, com resultados de 2q e 4q novamente abaixo de 1%.

A Fig. 4.14 mostra resultados para o método **t**.

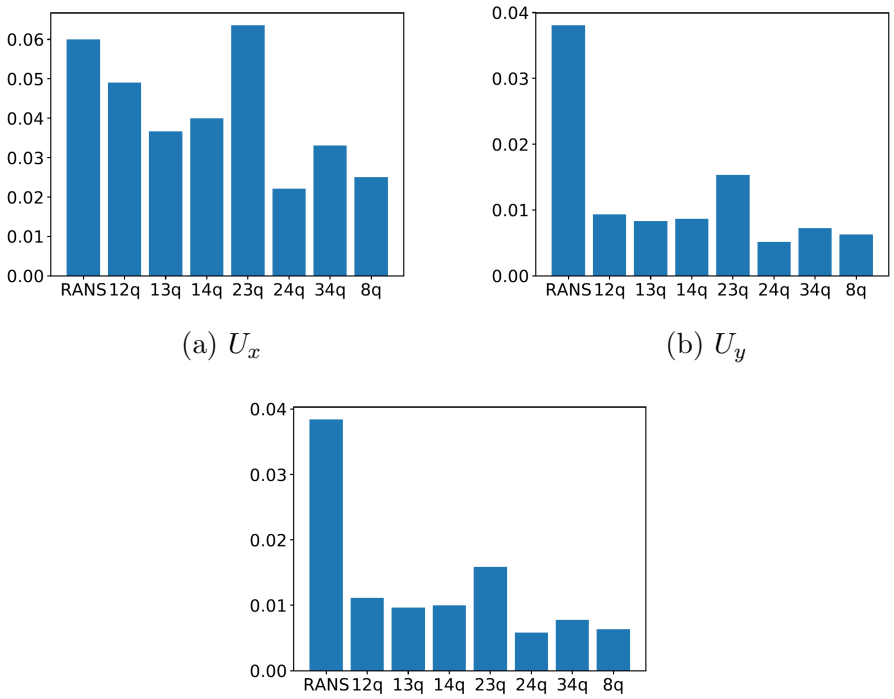
A Fig. 4.14a apresenta os erros do campo U_x. Todos os quadrantes apresentaram menores erros que RANS, abaixo de 2%. O melhor resultado foi obtido a partir de 3q, com erro inferior a 1%.

Na Fig. 4.14b, observa-se que para U_y as correções são superiores a RANS, novamente esperado, reiterando que o resultado RANS utilizado não captura escoamento secundário. Um quadrante se sobressaiu neste campo, o 3q, com erro por volta de 1%.

A Fig. 4.14c mostra os erros do campo U_z. Neste, os quadrantes apresentam resultado semelhante, com erros entre 1% e 2%.

Comparando os resultados de U_x das Fig. 4.13a e 4.14a, nota-se que a correção de **t** é superior a de **R** para este campo, o que é esperado baseado nos trabalhos de THOMPSON *et al.* (2016) e CRUZ (2018).

Comparando os resultados de U_y e U_z das Fig. 4.13b, 4.14b, 4.13c e 4.14c, podemos ver que o método **R** se sobressai para o escoamento secundário, resultado



(a) U_x (b) U_y

(c) U_z

Figura 4.15: Erro global da previsão de **R** do grupo Q2 para $Re = 3200$

esse que já havia sido apontado por CRUZ (2018).

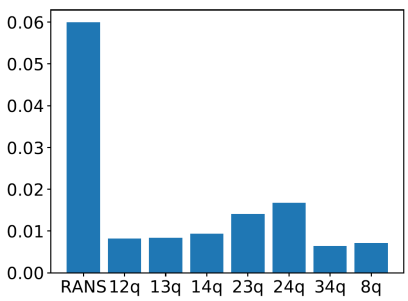
4.3.2 Q2

As previsões para o grupo Q2 tem a intenção de mostrar como a RN se comportaria se fosse treinada com dados DNS cuja simulação tivesse sido feita por mais tempo, neste caso o dobro de tempo dos dados da base utilizada.

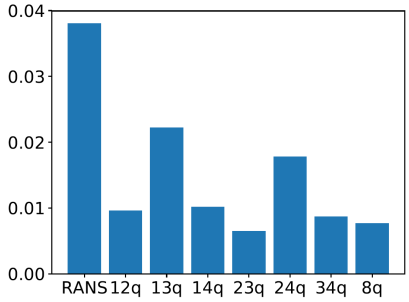
A Fig. 4.15 mostra os resultados para o método **R**.

Na Fig. 4.15a, temos o erro do campo U_x para as diversas combinações de quadrantes dois a dois, assim como RANS e 8q. Observa-se que, com exceção do caso 23q, todos apresentam erros menores que RANS, sendo que o caso 24q é o que se saiu melhor, com erro próximo de 2%. Quando comparamos estes resultados aos do Q1, Fig. 4.13, podemos ver que os resultados dos quadrantes 2q e 4q são os que apresentaram menores erros. Logo, sua combinação também o fazer indica que há uma ligação direta entre a melhora da previsão da RN e a qualidade dos dados utilizados em seu treinamento.

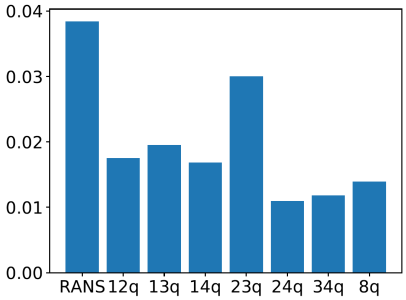
A Fig. 4.15b mostra os erros para U_y. Novamente o caso 23q apresenta resultados inferiores aos demais, porém desta vez melhores que RANS. Os demais se encontram em faixa de erro abaixo do 1%. O caso 24q também apresenta o melhor resultado neste campo.



(a) U_x



(b) U_y



(c) U_z

Figura 4.16: Erro global da previsão de **t** do grupo Q2 para $Re = 3200$

A Fig. 4.15c mostra os erros do campo U_z. Observa-se comportamento semelhante ao do U_y. Todos os casos apresentam erro por volta de 1% ou abaixo, exceto o 12q e o 23q.

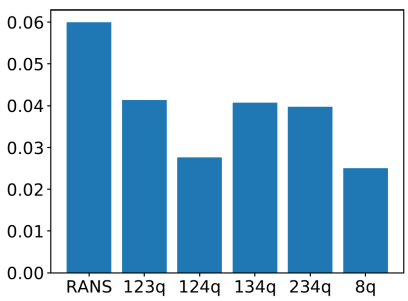
A Fig. 4.16 mostra os resultados para o método **t**.

Na Fig. 4.16a, temos o erro do campo U_x. Todos os casos apresentam erros baixíssimos, tendo apenas dois onde estes assumem valor acima de 1%, são eles o 23q e o 24q.

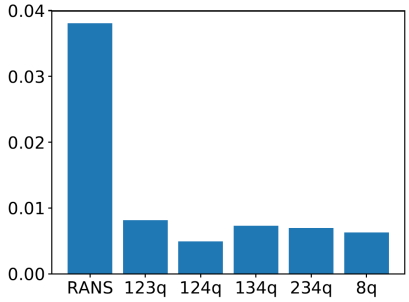
A Fig. 4.16b mostra os erros do campo U_y. Observa-se que apenas dois casos excedem a faixa de 1%, são eles o 13q e o 24q, com erros por volta dos 2%.

Para o campo U_z, a Fig. 4.16c mostra que 3 casos ficaram na faixa dos 2%, são eles 12q, 13q e 14q. Dois casos ficaram na faixa de 1%, estes são 24q e 34q. O mais discrepante foi o 23q, com erro na faixa de 3%.

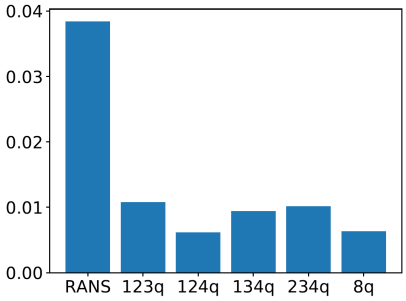
Ao analisar os resultados do grupo Q2, tentou-se observar um padrão nos resultados, como um quadrante que, ao estar na combinação, contaminasse o resultado positiva ou negativamente. Para o método **R**, os resultados mostram que a combinação do segundo com o terceiro quadrantes, 23q, levou a erros mais elevados que as demais combinações. Tendência essa que, exceto no campo U_y, se repetiu para o método **t**. Quando comparamos os resultados do 23q com os do 24q, onde mantém-se o segundo quadrante na combinação, tende-se a acreditar que o terceiro quadrante atuou de forma negativa em 23q, já que os resultados de 24q são os melhores do



(a) U_x



(b) U_y



(c) U_z

Figura 4.17: Erro global da previsão de **R** do grupo Q3 para $Re = 3200$

método **R**, e apresentam bons resultados em **t**.

4.3.3 Q3

As previsões para o grupo Q3 seguem a lógica dos anteriores. Neste caso, os resultados simulam o treinamento para dados com tempo de simulação 3 vezes maior que os dados da base utilizada.

A Fig. 4.17 mostra os resultados para o método **R**.

A Fig. 4.17a mostra os erros do campo U_x. O caso cujo resultado se sobressai é o 124q, com erro abaixo dos 3%. Os demais casos estão na faixa de 4%.

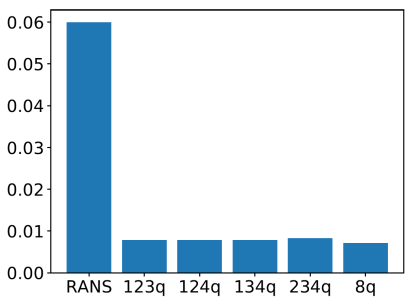
Na Fig. 4.17b, temos o erro do campo U_y. Todos os casos apresentam erro abaixo de 1%, com o caso 124q sendo o de melhor resultado neste campo.

A Fig. 4.17c mostra os erros do campo U_z. Os resultados se assemelham aos do campo U_y, com os casos apresentando erro por volta de 1%, enquanto o 124q mostra resultados melhores, abaixo de 1%.

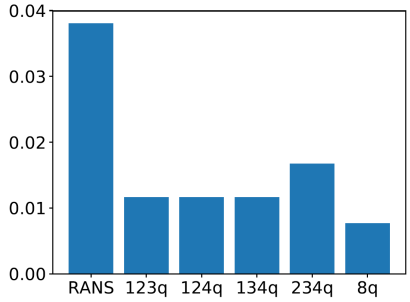
Os resultados para o método **t** estão dispostos na Fig. 4.18.

Na Fig. 4.18a estão os erros do campo U_x. Para este, todos os casos apresentam resultado abaixo de 1%.

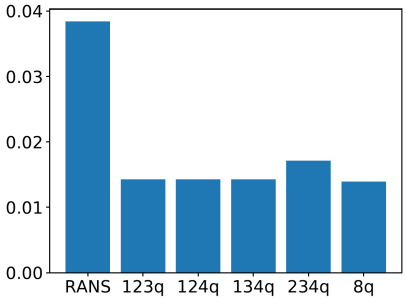
No campo U_y, Fig. 4.18b, o caso 234q apresenta erro abaixo de 2%, enquanto os demais estão por volta de 1%.



(a) U_x



(b) U_y



(c) U_z

Figura 4.18: Erro global da previsão de $\mathbf{t}$ do grupo Q3 para $Re = 3200$

Para o campo U_z, Fig. 4.18c, o comportamento é semelhante ao de U_y, com os erros dos casos estando abaixo de 2%, praticamente iguais entre si.

Retomando a discussão quanto à contaminação dos dados por um dos quadrantes, nos casos do Q3 podemos observar que a presença da combinação 23, como em 123q e 234q, são inferiores ao caso 124q, levando a conclusão de que os dados do terceiro quadrante deterioram os demais. Isso pode ser observado com clareza nos resultados do método **R**.

4.3.4 Q4

As previsões para o grupo Q4 seguem a lógica dos anteriores. Neste caso, os resultados simulam o treinamento para dados com tempo de simulação 4 vezes maior que os dados da base utilizada.

Como o grupo Q4 só tem um representante, dados relacionados a média dos erros de cada grupo anteriormente discutido foram utilizados na comparação, são eles m1q, m2q e m3q, para os grupos Q1, Q2 e Q3 respectivamente, além dos valores de RANS e 8q.

Os resultados para o método **R** estão dispostos na Fig. 4.19.

A Fig. 4.19a mostra os resultados para o campo U_x. É possível notar um aspecto decrescente das barras, denotando a ideia de que o aumento do tempo de simulação

leva a diminuição do erro. Observa-se que o 8q apresenta os melhores resultados, sendo que este representa, como dito anteriormente, a previsão da RN para o treinamento com os dados de uma simulação DNS de tempo 8 vezes maior que os campos originais. O resultado para o Q4 está praticamente empatado com os demais, porém é preciso levar em consideração que as demais barras representam a média dos erros dos outros grupos. Para o grupo Q1, por exemplo, temos uma variação de erro de 4% entre os quadrantes 1 e 4, como pode ser visto na Fig. 4.13a, em que eles apresentam erros acima de 6% e 2% respectivamente, ou seja, há uma incerteza maior agregada a escolha de apenas 1 quadrante como alvo de treinamento, uma vez que não é garantida a homogeneidade de seus campos. Analisando m2q, observa-se comportamento semelhante ao m1q, com erros que variam entre 2% (24q) e 6% (23q), Fig. 4.15a. Ainda assim, m2q ser menor que m1q mostra que a manipulação dos dados leva a uma média de erro menor, ou seja, diminuição de incerteza. A lógica é semelhante para o m3q, porém este apresenta variação menor que os anteriores, entre 3% (124q) e 4% (demais casos) de erro, Fig. 4.17a, o que corrobora com o previamente dito. Essa melhora em U_x apresentada com a manipulação dos campos DNS é muito importante pois, assim como já foi mencionado, o campo de velocidade média na direção principal do escoamento sempre apresentou dificuldade em ser reconstruído a partir do método de correção de **R**. Este resultado, portanto, nos leva a crer que é de suma importância o pré tratamento dos dados DNS quando estes sejam utilizados como alvo de treinamento em AM.

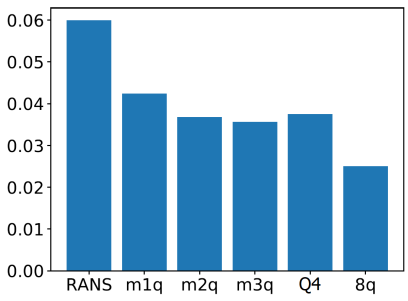
O comportamento do erro para os campos U_y e U_z, Fig. 4.19b e 4.19c, se assemelham entre si, e seguem a mesma lógica do apresentado em U_x, com erros abaixo de 1%. O 8q apresenta os melhores resultados, mesmo que levemente, salvo poucas exceções em combinações anteriores.

A Fig. 4.20 mostra os resultados para o método **t**.

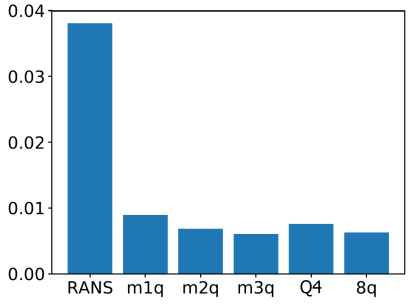
Como visto para o método **R**, na Fig. 4.20a os resultados de U_x também se comportam de maneira decrescente e abaixo de 1% de erro. O 8q apresenta os melhores resultados, com o Q4 em seguida.

Para o campo U_y, Fig. 4.20b, observa-se que o 8q apresenta melhores resultados, e se equipara a média m3q, enquanto o Q4 tem erro pouco acima de 1%.

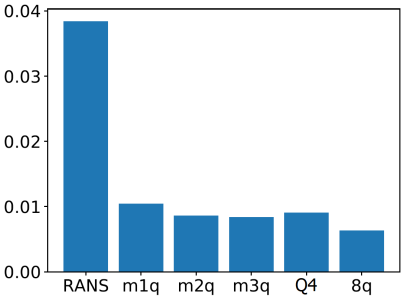
Na Fig. 4.20c, para o campo U_z, os erros estão todos abaixo de 2%.



(a) U_x

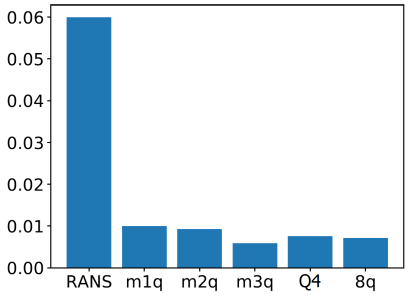


(b) U_y

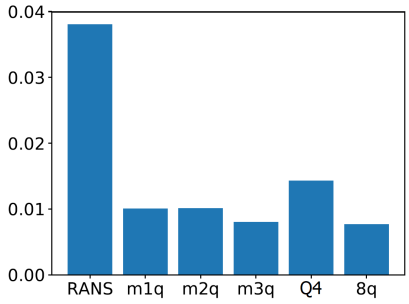


(c) U_z

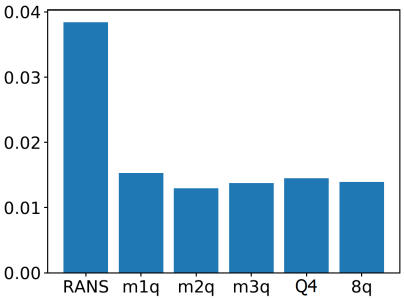
Figura 4.19: Erro global da previsão de $\mathbf{R}$ do grupo Q4 para $Re = 3200$



(a) U_x



(b) U_y



(c) U_z

Figura 4.20: Erro global da previsão de $\mathbf{t}$ do grupo Q4 para $Re = 3200$

Capítulo 5

Conclusões

Uma base de dados DNS foi manipulada na intenção de simular resultados melhor convergidos. Os dados manipulados foram então utilizados no treinamento de uma RN, para que essa pudesse prever **R** ou **t** a partir de dados RANS, melhorando sua acurácia com relação aos campos médios de velocidade de um escoamento turbulento em duto de seção quadrada.

Os resultados aqui apresentados confirmaram a suspeita de que a convergência da base de dados DNS tem influência direta na capacidade de previsão de uma RN, quando esta os utiliza como alvo de treinamento. Foi visto que, tão longo seja o tempo de simulação DNS, melhor é sua convergência, mais simétricos são seus dados, e melhor é seu desempenho como alvo de treinamento de uma técnica AM.

Comparando os varios grupos de manipulação de dados proposto, o que desempenhou melhor foi o de maior tempo de simulação, 8q.

Observou-se que, dos campos originais, os quadrantes podem apresentar influência negativa nos resultados se mal escolhidos no momento de definir o alvo de treinamento, uma vez que não há garantia de que todos eles apresentem a mesma qualidade de dados. A presença do terceiro quadrante, quando em combinações com outros, denegria os resultados, aumentando o erro global dos campos de velocidade.

Assim sendo, fica clara a contribuição da realização de pré tratamento dos dados DNS quando estes sejam utilizados como alvo de treinamento de uma técnica AM.

Quanto aos métodos de correção, mesmo após tratamento dos dados, os resultados do caso 8q apontam que o método **R** performa melhor quanto a previsão do escoamento secundário, com erros abaixo de 1% tanto para U_y quanto U_z, enquanto a correção de **t** apresenta erros inferiores a 1% e 2% respectivamente para os mesmos campos. Quando comparados em relação a U_x, o método **t**, que apresenta erro abaixo de 1%, supera o método **R**, com erro acima de 2%. Portanto, tendo a acreditar que o método **t** seja superior pois consegue entregar melhoria em todos os aspectos da correção, enquanto entrega um campo na direção principal bastante semelhante visualmente ao DNS.

Capítulo 6

Trabalhos Futuros

A metodologia deste trabalho pode ser aplicada em outros estudos, como na investigação de incerteza da base de dados DNS para outros escoamentos, como periodic hill.

Outra possibilidade é a investigação de bases de dados provenientes de tipos diferentes de simulação, como LES.

Referências Bibliográficas

ANDRADE, J. R., MARTINS, R. S., THOMPSON, R. L., et al., 2018, "Analysis of uncertainties and convergence of the statistical quantities in turbulent wall-bounded flows by means of a physically based criterion", *Physics of Fluids*, v. 30, n. 4, pp. 045106. doi: 10.1063/1.5023500. Disponível em: <https://doi.org/10.1063/1.5023500>.

CRUZ, M. A., 2018, *Machine Learning Techniques For Accuracy Improvement Of Rans Simulations*. Tese de Mestrado, UFRJ.

DAVIDSON, P. A., 2004, *Turbulence: An Introduction for Scientists and Engineers*. USA, Oxford University Press.

LING, J., KURZAWSKI, A., TEMPLETON, J., 2016a, "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance", *J. Fluid Mech.*, v. 807.

LING, J., JONES, R., TEMPLETON, J., 2016b, "Machine learning strategies for systems with invariance properties", *Journal of Computational Physics*, v. 318, pp. 22 – 35. ISSN: 0021-9991. doi: https://doi.org/10.1016/j.jcp.2016.05.003. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0021999116301309>.

LING, J., RUIZ, A., LACAZE, G., et al., 2016c, "Uncertainty Analysis and Data-Driven Model Advances for a Jet-in-Crossflow", *Journal of Turbomachinery*, v. 139, n. 2 (10). ISSN: 0889-504X. doi: 10.1115/1.4034556. Disponível em: <https://doi.org/10.1115/1.4034556>.

MARSLAND, S., 2009, *Machine Learning - An Algorithmic Perspective*. Florida, CRC Press.

MONAGHAN, J., KAJTAR, J., 2014, "Leonardo da Vinci's turbulent tank in two dimensions", *European Journal of Mechanics B/Fluids*, v. 44.

PINELLI, A., UHLMANN, M., SEKIMOTO, A., et al., 2010, "Reynolds number dependence of mean flow structure in square duct turbulence", *Journal of Fluid Mechanics*, v. 644, pp. 107–122. doi: 10.1017/S0022112009992242.

POPE, S. B., 2000, *Turbulent Flows*. USA, Cambridge University Press.

RICHARDSON, L. F., 2007, *Weather Prediction by Numerical Process*. Cambridge, Cambridge University Press.

THOMPSON, R. L., MENDES, P. R. S., 2005, "Persistence of straining and flow classification", *International Journal of Engineering Science*, v. 43, n. 1, pp. 79 – 105. ISSN: 0020-7225. doi: https://doi.org/10.1016/j.ijengsci.2004.07.011. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0020722504002198>.

THOMPSON, R. L., SAMPAIO, L. E. B., DE BRAGANÇA ALVES, F. A., et al., 2016, "A methodology to evaluate statistical errors in DNS data of plane channel flows", *Computers Fluids*, v. 130, pp. 1 – 7. ISSN: 0045-7930. doi: https://doi.org/10.1016/j.compfluid.2016.01.014. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0045793016300068>.

TRACEY, B., DURAISAMY, K., ALONSO, J., 2013, "Application of Supervised Learning to Quantify Uncertainties in Turbulence and Combustion Modeling". In: *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*. doi: 10.2514/6.2013-259. Disponível em: <https://arc.aiaa.org/doi/abs/10.2514/6.2013-259>.

TRACEY, B. D., DURAISAMY, K., ALONSO, J. J., 2015, "A Machine Learning Strategy to Assist Turbulence Model Development". In: *53rd AIAA Aerospace Sciences Meeting*. doi: 10.2514/6.2015-1287. Disponível em: <https://arc.aiaa.org/doi/abs/10.2514/6.2015-1287>.

WANG, J.-X., WU, J., LING, J., et al., 2016, "Physics-informed machine learning for predictive turbulence modeling: Toward a complete framework". 12.

WANG, J.-X., WU, J.-L., XIAO, H., 2017, "Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data", *Phys. Rev. Fluids*, v. 2 (Mar), pp. 034603. doi: 10.1103/PhysRevFluids.2.034603. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevFluids.2.034603>.

WILCOX, D. C., 2006, *Turbulence Modeling for CFD*. USA, DCW Industries.

WU, J., XIAO, H., PATERSON, E., 2018, "Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework", *Physical Review Fluids*, v. 3 (01), pp. 074602. doi: 10.1103/PhysRevFluids.3.074602.