UNIVERSIDADE FEDERAL DO ABC
PÓS-GRADUAÇÃO EM ENERGIA

Autor

TITULO

Santo André, SP – Brasil
Março de 2022

Autor

TITULO

Dissertação apresentada ao Curso de Pós-graduação da Universidade Federal do ABC, como parte dos requisitos necessários à obtenção do grau de Mestre em Energia.

Orientador: Prof. Dr. Romulo Gonçalves Lins
Coorientador:Prof. Dr. Ricardo Gaspar

Santo André, SP – Brasil
Março de 2022

Autor,
Titulo/ Autor. – 2022.
43 p.: il.
Orientador: Prof. Dr. Romulo Gonçalves Lins
Coorientador: Prof. Dr. Ricardo Gaspar
Dissertação (MASTER) – Universidade Federal do ABC, Pós-Graduação em Energia, Santo André, 2022.
1. K1. 2. K2. 3. K3. 4. K4. 5. K5. I. Gonçalves Lins, Romulo II. Gaspar, Ricardo III. Pós-Graduação em Energia, 2022. IV. Título

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do autor e com a anuência de seu orientador.

Santo André, ________ de ______________________ de 20 ______

Assinatura do autor: ____________________________________

Assinatura do orientador: ________________________________

UFABC

PÓS-GRADUAÇÃO EM ENERGIA

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de master do candidato Autor em 23 de Março de 2022:

Profa. Dra. Marcia Maria Penteado Marchesini - Examinadora Externa

Prof. Dr. Wilson Carlos da Silva Júnior - Examinador Externo

Profa. Dra. Franciane Freitas Silveira - Examinadora Interna Suplente

Prof. Dr. Rômulo Gonçalves Lins - Presidente

Universidade Federal do ABC, Av. dos Estados, 5001– CEP 09210-580– Santo André– SP

Tel. 0 XX 11 – 4996 – 0086 – www.ufabc.edu.br

Agradecimentos

Obrigado a todos

Epígrafe

"Learn from the mistakes of those who followed your advice"
(Unknown author)

Abstract

Machine learning methods have become a powerful tool for the academic community in recent decades. In the field of computational fluid and thermodynamics, these methods have been used to perform fast inferences on unseen parameters, reducing the computational burden associated with traditional numerical methods. Most works in this field focus on predicting scalar global or integrated parameters. However, this work introduces a new machine learning data-driven flow reconstruction method that can reconstruct entire flow fields using singular value decomposition as a dimensionality reduction technique.Additionally, we conducted a study on the impact of the number of selected modes for dimensionality reduction on the overall performance metrics and other hyperparameters for neural network performance. We also compared the performance of neural networks with the Kriging method. The main results showed that shallow neural networks with sigmoid activation functions performed better than deep neural networks, and the Kriging method was faster and more accurate than the neural networks. The best models obtained so far have demonstrated their viability as accurate surrogate models.

Keywords: Fourth Industrial Revolution; Data-Driven Culture; Computer Vision systems; Quality control; Statistical process control

Resumo

Métodos de aprendizado de máquina tornaram-se uma ferramenta poderosa para a comunidade acadêmica nas últimas décadas. No campo da fluidodinâmica computacional e da termodinâmica, esses métodos têm sido usados para realizar inferências rápidas sobre parâmetros não vistos, reduzindo a carga computacional associada aos métodos numéricos tradicionais. A maioria dos trabalhos nesse campo concentra-se em prever parâmetros globais ou integrados escalares. No entanto, este trabalho introduz um novo método de reconstrução de fluxo baseado em dados de aprendizado de máquina que pode reconstruir campos de fluxo inteiros usando a decomposição de valores singulares como técnica de redução de dimensionalidade. Além disso, conduzimos um estudo sobre o impacto do número de modos selecionados para redução de dimensionalidade nas métricas de desempenho geral e em outros hiperparâmetros para o desempenho de redes neurais. Também comparamos o desempenho de redes neurais com o método de Krigagem. Os principais resultados mostraram que redes neurais rasas com funções de ativação sigmoide tiveram um desempenho melhor do que redes neurais profundas, e o método de Krigagem foi mais rápido e preciso do que as redes neurais. Os melhores modelos obtidos até agora demonstraram sua viabilidade como modelos substitutos precisos.

Palavras-chave: Indústria 4.0; Cultura de Dados; Sistemas de visão computacional; Qualidade; Controle estatístico de processo

Lista de Ilustrações

Lista de Tabelas

Sumário

1 Introduction

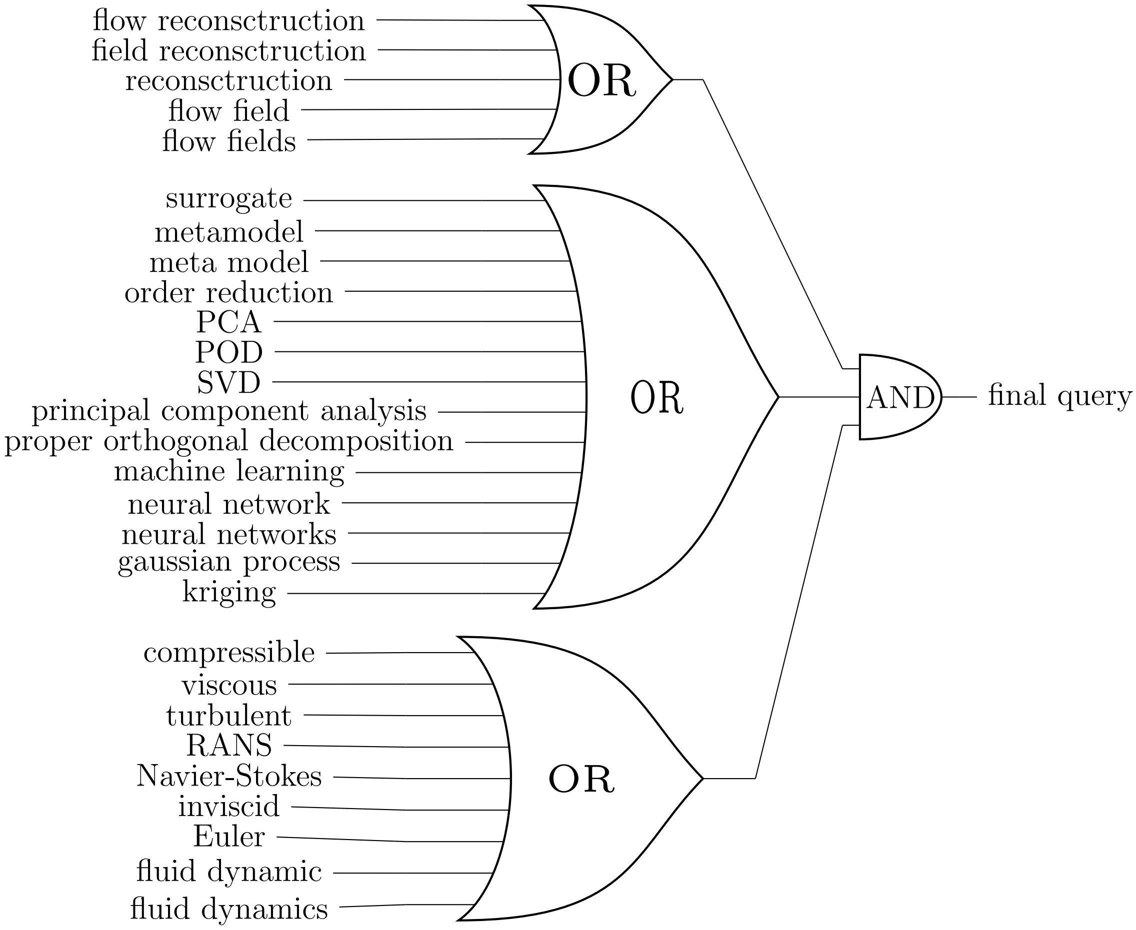
Definition of ML method

"ML is a subfield of AI. It can be defined as a set of methods and algorithms for estimating the relationship between some inputs and outputs with the help of a number of trainable parameters. The learning of these parameters, that is, their optimization with respect to a given metric, is achieved in an iterative manner by comparing the model predictions against ground truth data or evaluation of the model performance."**?**.

1.1 Bibliography Review

A bibliometric review was conducted by search for the most relevant (to the authors knoledge) keywords in the field of present work scopus. These keywords were properly organized using the boolean operators in order to avoid returning uncorrelated works and to be broader as possible. They are:

```
( "flow reconstruction" OR "field reconstruction" OR "reconstruction"
 OR "flow field" OR "flow fields") AND ("surrogate" OR "metamodel" OR
"meta model" OR "order reduction" OR "reduced order" OR "PCA" OR "POD"
 OR "SVD" OR "principal component analysis" OR "singular value
decomposition" OR "proper orthogonal decomposition" OR "machine
learning" OR "neural network" OR "neural networks" OR "gaussian
process" OR "kriging" ) AND ( "compressible" OR "viscous" OR "
turbulent" OR "RANS" OR "Navier-Stokes" OR "inviscid" OR "Euler" OR "
fluid dynamic" OR "fluid dynamics" ) AND ( "data-driven" OR "data
driven" )
```



Numerical methods for multiphysics simulations of fluid dynamics are mature and widely used in industrial and academic applications. However, reducing the computational costs associated with these methods is still a must. Machine learning techniques like Kriging and Neural Networks have emerged as viable alternatives to build surrogate models for traditional numerical methods, with direct applications in tasks such as optimization, inverse problems, and uncertainty quantification. In this work, we explore a data-driven flow reconstruction method that can predict high-fidelity multiphysics flows using low-fidelity flow simulations as input.

Previous research mostly focused on scalar and integrated quantities. In contrast, our work focuses on predicting entire flow fields, allowing the method to capture flow structures such as shock formation and highly non-linear heat flux distributions over supersonic Convergent-divergence nozzle walls.

The number of variables required to represent multidimensional flow fields can be very large, resulting in impractically large neural networks or Kriging methods. To tackle this issue, we employ principal component analysis (PCA) decomposition as an order reduction technique. An important aspect of this work is studying the impact of the chosen reduced dimensionality on the accuracy of the surrogate model.

In **?** the authors addresses the problem of select the best rank k of a SVD decompostion to accurately reconstruct a given dataset. Some methods ranges from evatualtion of variance or entropy content in the latent space. The authros also introduced new criteia known as Kullback-Leibler and anomaly detection. This highligh the importante of appropriate selection of reduced basis rank. In this work we focus on study the model senstivity with respect to basis dimension by evaluating the NRMSE, R2 and MAPE for the overall reconstruction strategy.

In **?** used convolutional adversarial networks to predict the flow rate of a gas-liquid multiphase flowm, surpassing the performance of state-of-art flow rate prediction methods. Although MAPE metric achieved goog performance, as low as 8.36 % we emphasize the method was made to predict the scalar quantity os mass flow rate, being flow reconstruction methods more capable, being able to predict the entire flow field.

In **?** the authors used neural networks to identify the dinamics of a 2-D PDE Burger's equations. The authors work entirely in full dimensional, without any dimensionality reduction technique. The results, acccounts for lyapunov stability test, and show the model is stable and accure. In our view, the lack of dimensionaly reduction, could be a severe issue when dealing with high dimensional problems, as the one we are dealing with.

In **?** have developed mathematical rigorous upper bound error for the solution of a PDE using neural networks (PINNs). The method can't be directly applied to non-linear PDEs, and also can't be used to determine the lower limit for error. This highglighs the importance, difficult and necessity to more rigorous error analysis for machine learning methods.

In **?**, the practica implications of Reynolds number in the coefficient discharge of flow in critical nozzles is studied. The Reynolds number fo this machine was estimated as $Re = \frac{4\dot{m}_{th}}{\pi D_{th}\mu_0}$, since in our design experiment, the boundary conditions forced the flow to be sonic at throat, no appreciable impact regarding Re number is expected.

In **?** the authors used a data-driven approach for flow reconstruction in the context of identification of nonlinear dynamical systems. The approach consists in using auto-encoder to reduce the model dimensionality, then discover the underline equations based on a restricted number of measurements data. Altough not exactely in the same context of present study, auto-enconders have show superior ability to deal with the latent variables, suprassing proper orthogonal decomposition methods.

The advantages of dimensionaly reduction for machine learning method to predict fluid dynamic systems was also investigated by **?**. The authos used auto-enconder to reduce system dimensionality and have reported superior performance for the sistem identification when compared to full dimensional inferences.

Despite recent advances in machine learning methods and its applications, the flow reconsctruction thecnique can also be applied using traditional numerical solvers, in **?** the authors apply the flow reconstruction technique to reconstruct microestructures of a gas diffusion layer and its anisotropic transport properties, the method used sparse measurements and traditional numerical solvers to achieve the reconstruction, it is a good example of opportunity to use machine learning methods to reduce the computational cost of traditional numerical methods in flow reconsctruction problems.

In **?** the authors applied the principal component analysi (PCA) as a dimensionality reduction technique to represend PDEs's solutions. To represent complex geometries and boundary conditions with ease, the methodolgy also made use of a novel mesh morphing technique to mapa the domain into a common morphed mesh an Gaussian process as a regression technique. As novelty, the method used finite element interpolation and shape embedding into low dimensional coordinates.

In the work of **?** a multiresolution recostructed of hypersonic heat flux was performed using spasrce measruments data ath mach 0.8 and mach 6.

? Used proper orthogonal decomposition (POD) combined with linear discriminant analysis on a clustering framwork to indentify the best sensor placament for identify the unsteady dynamic in the wake of wind turbines. The results shows only evolution of point measruments and not the entire flow field, altought it higligths how a small number of sensor can be selected to best reconstruct a flow field using projection based techniques.

In **?** Used convolutional neural networks (CNN) to reconstruct particle imagage velocimeter (PIV) steady velocity fields in a gas turbine pre mixed swirl combustor. The model is purely data driven and used experimental data of hydroxyl(OH) planar laser induced Fluorescence (OH-PLIF) and PIV for training. No dimensionaly reduction technique was used in the training process,some kind of compressio is embeded in the encode-deconding process is embeded in the convolutional layers of the CNN.

In **?** Turbulece is a highly complex and yet to be fully understood phenomena, this reasearch paper review some of advances in using machine learning methods to model such complex multiscale problem. Some of the oulined mode are methods for parameter identification, system identification and reconstruction of closure model. The ML models showed promissing results, with high accuraty subgrid fore terms prediction. Focused on incompressible turbulence modeling. Closure terms prediction and field deconvolution are key ares pointed.

? used 2D data from high-resolution ocean model, to emulate the real achiavavle quality of sattelite images, a low pass filter was applied in the training data. A convolutional neural networks were then trainend, to predict subsurface flow fields and eddy momentum forcing term from the degraded data. Showing possible application to improve model accuracy and also infer data from sparse observations.

? used data from limited pressure measurements of the viscous flow arround a cylinder to perform a flow field reconstruction. POD was used for dimensionality reduction and optimal basis construction. The performance metrics altough was not reported in terms of whole field data reconstruction, but only scalar quantity such as the Reynolds number prediction.

? used POD for low dimensional projection and convolutional neural networks to extract low-dimensional features. The ground truth data was numericaly generated by a finite element method to solve the unsteady navier stokes equation in compressible regime. The models investigated for system identification or inverse problem was proper orthogonal decomposition-recurrent neural network (POD-RNN) and convolutional recurrent eutoencoder network (CRAN). The recurrent neural network in question was a long short-term memory network (LSTM). The criteria for basis selection was the tradicional energy preservation. The CRAN models were found to suprass the POD-RNN in terms of accurately representation of highly nonlinear problemas such as the flow arrond two side-by-side cylinders. The computational cost of training the CRAN molde was howerve 16 times higher. Altough tailored to unsteady flow reconstruction, this work have many similarity with ours since, a low dimensionl representation of flow features must be stabilished prior to build a repressor model, in our case to predict flow fields for new unseen design parameters and in the case of the referenced study to extrapolate in time.

? is focused on the 2D unsteady flow arround atalled arifoils. Sparese sensor data are used to reconsctruction the flow field. POD was used for dimensionality reduction and, some stretegies were used to recover the basis from sparse measurmenta, simple linear approancs and also a shallow (3 -layers) fully connected neural network. The results showed neural network perferomed better, no considerations about compution tim however were discussed.

? is also another work focused on unsteady flow. The method is data-driven and used generative adversarial networks to enconde and predict the flow field for unseen parameter. The model only udes one parameter to characterize the boundary condictions. The resultand recostructed flow field agreed well with high fidelity generted data with a significant computational cost reduction.

? used an interpolation method to perform a parametric dimensionality reduction using POD and then reconstruct the incompressible flow arround a cylinder and the stress tensor field on double-T beam. The active subspace method proposed outperfomed traditional methos when using small number of principal components but have worse performance when using large number of principal components. No deteils abount the interpolation method were given and the performance of the model in predictive regime should be hohg affected by this interpolation technique.

? applied standard POD, and improved version of POD using ridge regression and a shawllow decoder (SD) to reduce dimensionality and reconstruct 2D spatial and temporal flow fields from sparse measruments. From natrual nonlinearity nature of SD, it outperform both POD methods in termos of accuracy. The studied cases includes numerical flow around a cylinder, experimental data for mesoscale sea surface temperature and numerical turbulent isotropic flow.

? used deep learning to predict fully-turbulent systems. The proposed approach use data from low-fidelity simulations to infer high fidelity flow fields. Physical information of in the form of a loss function with PDE residual was also used, so this technique can be seeing as a hybrid model using both data and physical information.

? used nonlinear Laplacian spectral analysis (NLSA) to build a weak formulation for the Koopman operator for the low dimensional representation of the dynamics of a fully turbulent 3D convection flow. Data was generaterd using DNS for the Rayleigh Bernard convection problem. NLSA can be viewed as POD in a delay-coordinate space.

? used a convolutional neural network to predict results of a 2D CFD zone fire model. The surrogate model reconstructed the temperature and velocity fields given scalar parameters for the simulations. The model is completly data-driven and no dimensionality reduction beyond the convolutional operations was performed.

? used volumetric tomographic Networkds (VT-Nets) to perforrm a three dimensional tomographic reconstruction of turbulent flame flows. The performance of the machine learning method was compared with traditional algebric reconstruction techniques (ART), showing the convolutional neurals networks were able to outperm traditional methods when a reduced number os 2D projections are used to reconstruct the whole 3D flame. It is alse insanely faster than traditional algorithms for inference tasks.

? used proper orthogonal decomposition (POD) to reduce timensionality and then used sparse measurments to performan a flow field reconstruction. THe sensor placement, quantity and system dimension was investigated to characterize the sparse reconstruction performance.

? performed a velocity field prediction on a scramjet isolator using deep learnin. The data-driven model used a convolutional neural network to stabilished a relashionship between wall pressure and velocity field. Experimental data was used for comparison and 3D RANS simulations were performed to build the datasets, altough the simulations were 3D only 2D slices of flow field were used to train and test the model. Despite the model being compressible with a lot of oblique schock formations, no attempt to recover the shock structures were made.

? used a summetrical deep convolutional neural network to reconstruct the 2D steady state supersonic flow field in a cascade channel. The model is data-driven and used discreve pressure data to reconstruct 2D flow field. The model were able to accurately reconstruct flow scructures present in the form of shock wave formation under unseen parameters.

? show surrogate model can replace costly traditional solvers in a heat transfer problem, they effectively used fullly connected convolutional neural networks to perform flow field reconstruction of 2D heat transfer problem with nanofluid flowing in a microchannel with grooves. The data driven model used design parameters such as groove radius, relative depth, boundary conditions and a fluid parameter to infer 2D phycial fields of pressure, temeperature and velocity components. Neural networks showed superior performance when compared with tradicional surrogate models.

1.2 POD

The POD reduction, perhaps in a slightly modified form, is known throughout the literature alternatively as principal component analysis (PCA), empirical orthogonal functions (EOF), the Hotelling transform, or a Karhunen-Loeve expansion. Here, we will simply refer to it as POD.

2 Fundamentação Teórica

2.1 Impactos financeiros de um produto defeituoso

2.1.1 O valor de uma marca

$$\nabla^2 f_{ij} = 4f_{ij} - (f_{(i+1)j} + f_{(i-1)j} + f_{i(j+1)} + f_{i(j+1)}) \tag{2.1}$$

Outra forma simples de apresentar esta expressão consiste em fornecer uma *matriz* contendo os coeficientes dos *pixels*:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

3 methodology

3.1 Numerical Methods

The present flow reconstrudtion method is purely data-driven and should work seamsly with experimental data. For the sake of cost and time, in this work, the data was syntheticly generated using numerical solvers.

3.2 POD

POD, also known as principal component analysis (PCA), empirical orthogonal functions (EOF), the Hotelling transform, or a Karhunen-Loeve expansion in various domains, is a technique to reduce the dimensionality of complex data sets. This method is extensively used in numerical simulations and physical experiments to extract dominant patterns of variability from a high-dimensional dataset.

4 Aplicação da Metodologia

4.1 Determinação do caso a ser estudado

5 Resultados

5.1 Determinação do caso a ser estudado

Etapa concluída.

Referências Bibliográficas

6 Anexos

Este é um anexo.

Não há renumeração para os anexos neste modelo.

Não parece ser útil ter distinção entre apêndices e anexos.

7 Anexo 01

Este é um anexo.

Não há renumeração para os anexos neste modelo.

Não parece ser útil ter distinção entre apêndices e anexos.

Se fizer passar para a versão 1_0.