Electronic Transactions on Numerical Analysis
Volume 56, 2022

Contents

1 Estimating the time-dependent contact rate of SIR and SEIR models in mathematical epidemiology using physics-informed neural networks.
Viktor Grimm, Alexander Heinlein, Axel Klawonn, Martin Lanser, and Janine Weber.

Abstract.
The course of an epidemic can often be successfully described mathematically using compartment models. These models result in a system of ordinary differential equations. Two well-known examples are the SIR and the $SEIR$ models. The transition rates between the different compartments are defined by certain parameters that are specific for the respective virus. Often, these parameters are known from the literature or can be determined using statistics. However, the contact rate or the related effective reproduction number are in general not constant in time and thus cannot easily be determined. Here, a new machine learning approach based on physics-informed neural networks is presented that can learn the contact rate from given data for the dynamical systems given by the SIR and $SEIR$ models. The new method generalizes an already known approach for the identification of constant parameters to the variable or time-dependent case. After introducing the new method, it is tested for synthetic data generated by the numerical solution of SIR and $SEIR$ models. The case of exact and perturbed data is considered. In all cases, the contact rate can be learned very satisfactorily. Finally, the $SEIR$ model in combination with physics-informed neural networks is used to learn the contact rate for COVID-19 data given by the course of the epidemic in Germany. The simulation of the number of infected individuals over the course of the epidemic, using the learned contact rate, shows a very promising accordance with the data.

Key Words.
machine learning, physics-informed neural networks, SIR model, SEIR model, epidemic modeling, parameter estimation, COVID-19, SARS-CoV-2, scientific machine learning

AMS Subject Classifications.
65L09, 68T07, 68T09, 92C60, 92D30

28 Operator inference and physics-informed learning of low-dimensional models for incompressible flows.
Peter Benner, Pawan Goyal, Jan Heiland, and Igor Pontes Duff.

Abstract.
Reduced-order modeling has a long tradition in computational fluid dynamics. The ever-increasing significance of data for the synthesis of low-order models is well reflected in the recent successes of data-driven approaches such as *Dynamic Mode Decomposition* and *Operator Inference*. With this work, we discuss an approach to learning structured low-order models for incompressible flow from data that can

be used for engineering studies such as control, optimization, and simulation. To that end, we utilize the intrinsic structure of the Navier-Stokes equations for incompressible flows and show that learning dynamics of the velocity and pressure can be decoupled, thus, leading to an efficient operator inference approach for learning the underlying dynamics of incompressible flows. Furthermore, we demonstrate the performance of the operator inference in learning low-order models using two benchmark problems and compare with an intrusive method, namely proper orthogonal decomposition, and other data-driven approaches.

Key Words.
Computational fluid dynamics, scientific machine learning, incompressible flow, Navier-Stokes equations, operator inference

AMS Subject Classifications.
37N10, 68T05, 76D05, 65F22, 93A15, 93C10

52 A comparison of reduced-order modeling approaches using artificial neural networks for PDEs with bifurcating solutions.
Martin W. Hess, Annalisa Quaini, and Gianluigi Rozza.

Abstract.
This paper focuses on reduced-order models (ROMs) built for the efficient treatment of PDEs having solutions that bifurcate as the values of multiple input parameters change. First, we consider a method called local ROM that uses k-means algorithm to cluster snapshots and construct local POD bases, one for each cluster. We investigate one key ingredient of this approach: the local basis selection criterion. Several criteria are compared and it is found that a criterion based on a regression artificial neural network (ANN) provides the most accurate results for a channel flow problem exhibiting a supercritical pitchfork bifurcation. The same benchmark test is then used to compare the local ROM approach with the regression ANN selection criterion to an established global projection-based ROM and a recently proposed ANN based method called POD-NN. We show that our local ROM approach gains more than an order of magnitude in accuracy over the global projection-based ROM. However, the POD-NN provides consistently more accurate approximations than the local projection-based ROM.

Key Words.
Navier–Stokes equations, reduced-order methods, reduced basis methods, parametric geometries, symmetry breaking bifurcation

AMS Subject Classifications.
65P30, 35B32, 35Q30, 65N30, 65N35, 65N99

66 A combined finite element and machine learning approach for the prediction of specific cutting forces and maximum tool temperatures in machining.
Sai Manish Reddy Mekarthy, Maryam Hashemitaheri, and Harish Cherukuri.

Abstract.
In machining, specific cutting forces and temperature fields are of primary interest. These quantities depend on many machining parameters, such as the cutting speed, rake angle, tool-tip radius, and uncut chip thickness. The finite element method (FEM) is commonly used to study the effect of these parameters on the forces and temperatures. However, the simulations are computationally intensive and thus, it

is impractical to conduct a simulation-based parametric study for a wide range of parameters. The purpose of this work is to present, as a proof-of-concept, a hybrid methodology that combines the finite element method (FE method) and machine learning (ML) to predict specific cutting forces and maximum tool temperatures for a given set of machining conditions. The finite element method was used to generate the training and test data consisting of machining parameter values and the corresponding specific cutting forces and maximum tool temperatures. The data was then used to build a predictive model based on artificial neural networks. The FE models consist of an orthogonal plane-strain machining model with the workpiece being made of the Aluminum alloy Al 2024-T351. The finite element package Abaqus/Explicit was used for the simulations. Specific cutting forces and maximum tool temperatures were calculated for several different combinations of uncut chip thickness, cutting speed and the rake angle. For the machine learning-based predictive models, artificial neural networks were selected. The neural network modeling was performed using Python with Adam as the training algorithm. Both shallow neural networks (SNN) and deep neural networks (DNN) were built and tested with various activation functions (ReLU, ELU, tanh, sigmoid, linear) to predict specific cutting forces and maximum tool temperatures. The optimal neural network architecture along with the activation function that produced the least error in prediction was identified. By comparing the neural network predictions with the experimental data available in the literature, the neural network model is shown to be capable of accurately predicting specific cutting forces and temperatures.

Key Words.
finite element modeling, machining, machine learning, artificial neural networks, activation function, shallow and deep networks, Adam, specific cutting forces, maximum tool temperature

AMS Subject Classifications.
74S05

86 Structure preservation for the Deep Neural Network Multigrid Solver.
Nils Margenberg, Christian Lessig, and Thomas Richter.

Abstract.

The simulation of partial differential equations is a central subject of numerical analysis and an indispensable tool in science, engineering, and related fields. Existing approaches, such as finite elements, provide (highly) efficient tools but deep neural network-based techniques emerged in the last few years as an alternative with very promising results. We investigate the combination of both approaches for the approximation of the Navier-Stokes equations and to what extent structural properties such as divergence freedom can and should be respected. Our work is based on DNN-MG, a deep neural network multigrid technique, that we introduced recently and which uses a neural network to represent fine grid fluctuations not resolved by a geometric multigrid finite element solver. Although DNN-MG provides solutions with very good accuracy and is computationally highly efficient, we noticed that the neural network-based corrections substantially violate the divergence freedom of the velocity vector field. In this contribution, we discuss these findings and analyze three approaches to address the problem: a penalty term to encourage divergence freedom of the network output; a penalty term for the corrected velocity field; and a network that learns the stream function and which hence yields divergence-free corrections

by construction. Our experimental results show that the third approach based on the stream function outperforms the other two and not only improves the divergence freedom but also the overall fidelity of the simulation.

Key Words.
multigrid methods, incompressible Navier-Stokes equations, machine learning, artificial neural networks

AMS Subject Classifications.
65M55, 76D05

102 A Non-Intrusive Method to Inferring Linear Port-Hamiltonian Realizations using Time-Domain Data.
Karim Cherifi, Pawan Goyal, and Peter Benner.

Abstract.
Port-Hamiltonian systems have gained a lot of attention in recent years due to their inherent valuable properties in modeling and control. In this paper, we are interested in constructing linear port-Hamiltonian systems from time-domain input-output data. We discuss a non-intrusive methodology that is comprised of two main ingredients—(a) inferring frequency response data from time-domain data and (b) constructing an underlying port-Hamiltonian realization using the inferred frequency response data. We illustrate the proposed methodology by means of two numerical examples and also compare it with two other system identification methods to infer the frequency response from the input-output data.

Key Words.
System identification, port-Hamiltonian systems, Loewner approach, transfer function, input-output data

AMS Subject Classifications.
93A30, 93B30, 93B15, 93B20

117 A machine learning framework for LES closure terms.
Marius Kurz and Andrea Beck.

Abstract.
In the present work, we explore the capability of artificial neural networks (ANN) to predict the closure terms for large eddy simulations (LES) solely from coarse-scale data. To this end, we derive a consistent framework for LES closure models, with special emphasis laid upon the incorporation of implicit discretization-based filters and numerical approximation errors. We investigate implicit filter types that are inspired by the solution representation of discontinuous Galerkin and finite volume schemes and mimic the behavior of the discretization operator, and a global Fourier cutoff filter as a representative of a typical explicit LES filter. Within the perfect LES framework, we compute the exact closure terms for the different LES filter functions from direct numerical simulation results of decaying homogeneous isotropic turbulence. Multiple ANN with a multilayer perceptron (MLP) or a gated recurrent unit (GRU) architecture are trained to predict the computed closure terms solely from coarse-scale input data. For the given application, the GRU architecture clearly outperforms the MLP networks in terms of accuracy, whilst reaching up to 99.9% correlation between the networks' predictions and the exact closure terms for all considered filter functions. The GRU networks are also shown to generalize

well across different LES filters and resolutions. The present study can thus be seen as a starting point for the investigation of data-based modeling approaches for LES, which not only include the physical closure terms, but account for the discretization effects in implicitly filtered LES as well.

Key Words.
large eddy simulation, turbulence models, deep learning, artificial neural networks, recurrent neural networks

AMS Subject Classifications.
76F65, 68T07, 76F05

138 Approximation of a marine ecosystem model by artificial neural networks.
Markus Pfeil and Thomas Slawig.

Abstract.
Marine ecosystem models are important to identify the processes that affect for example the global carbon cycle. Computation of an annually periodic solution (i.e., a steady annual cycle) for these models requires a high computational effort. To reduce this effort, we approximate an exemplary marine ecosystem model by different artificial neural networks (ANNs). We use a fully connected network (FCN), then apply the sparse evolutionary training (SET) procedure, and finally apply a genetic algorithm (GA) to optimize, inter alia, the network topology. With all three approaches, a direct approximation of the steady annual cycle is not sufficiently accurate. However, using the mass-corrected prediction of the ANN as initial concentration for additional model runs, the results are in very good agreement. In this way, we achieve a runtime reduction by about 15%. The results from the SET algorithm are comparable to those of the FCN. Further application of the GA may lead to an even higher reduction.

Key Words.
deep learning, genetic algorithm, sparse evolutionary training, biogeochemical modeling, marine ecosystem modeling, transport matrix method

AMS Subject Classifications.
68T07, 68U99, 92F05

157 Decomposition and composition of deep convolutional neural networks and training acceleration via sub-network transfer learning.
Linyan Gu, Wei Zhang, Jia Liu, and Xiao-Chuan Cai.

Abstract.
Deep convolutional neural network (DCNN) has led to significant breakthroughs in deep learning. However, larger models and larger datasets result in longer training times slowing down the development progress of deep learning. In this paper, following the idea of domain decomposition methods, we propose and study a new method to parallelize the training of DCNNs by decomposing and composing DCNNs. First, a global network is decomposed into several sub-networks by partitioning the width of the network (i.e., along the channel dimension) while keeping the depth constant. All the sub-networks are individually trained, in parallel without any interprocessor communication, with the corresponding decomposed samples from the input data. Then, following the idea of nonlinear preconditioning, we propose a sub-network transfer learning strategy in which the weights of the trained

sub-networks are recomposed to initialize the global network, which is then trained to further adapt the parameters. Some theoretical analyses are provided to show the effectiveness of the sub-network transfer learning strategy. More precisely speaking, we prove that (1) the initialized global network can extract the feature maps learned by the sub-networks; (2) the initialization of the global network can provide an upper bound and a lower bound for the cost function and the classification accuracy with the corresponding values of the trained sub-networks. Some experiments are provided to evaluate the proposed methods. The results show that the sub-network transfer learning strategy can indeed provide good initialization and accelerate the training of the global network. Additionally, after further training, the transfer learning strategy shows almost no loss of accuracy and sometimes the accuracy is higher than if the network is initialized randomly.

Key Words.
deep convolutional neural networks, decomposition and composition, parallel training, transfer learning, domain decomposition

AMS Subject Classifications.
68W10, 68W40

187 A deep learning based nonlinear upscaling method for transport equations.
Tak Shing Au Yeung, Eric T. Chung, and Simon See.

Abstract.
We will develop a nonlinear upscaling method for nonlinear transport equations. The proposed scheme gives a coarse scale equation for the cell average of the solution. In order to compute the parameters in the coarse scale equation, a local downscaling operator is constructed. This downscaling operation recovers fine scale properties using cell averages. This is achieved by solving the equation on an oversampling region with the given cell average as constraint. Due to the nonlinearity, one needs to compute these downscaling operations on the fly and cannot pre-compute these quantities. In order to give an efficient downscaling operation, we apply a deep learning approach. We will use a deep neural network to approximate the downscaling operation. Our numerical results show that the proposed scheme can achieve good accuracy and efficiency.

Key Words.
nonlinear upscaling, transport equations, deep learning

AMS Subject Classifications.
65N30, 65N40

209 A hybrid objective function for robustness of artificial neural networks—estimation of parameters in a mechanical system.
Jan Sokolowski, Volker Schulz, Hans-Peter Beise, and Udo Schroeder.

Abstract.
In several studies, hybrid neural networks have proven to be more robust against noisy input data compared to plain data driven neural networks. We consider the task of estimating parameters of a mechanical vehicle model based on acceleration profiles. We introduce a convolutional neural network architecture that given sequential data, is capable to predict the parameters for a family of vehicle models that differ in the unknown parameters. This network is trained with two objective

functions. The first one constitutes a more naive approach that assumes that the true parameters are known. The second objective incorporates the knowledge of the underlying dynamics and is therefore considered as hybrid approach. We show that in terms of robustness, the latter outperforms the first objective on unknown noisy input data.

Key Words.
system identification, parameter estimation, convolutional neural networks, sequential data, prediction robustness, mathematical modelling, dynamical systems

AMS Subject Classifications.
68T07, 93B30, 34A30

235 Surrogate convolutional neural network models for steady computational fluid dynamics simulations.
Matthias Eichinger, Alexander Heinlein, and Axel Klawonn.

Abstract.
A convolution neural network (CNN)-based approach for the construction of reduced order surrogate models for computational fluid dynamics (CFD) simulations is introduced; it is inspired by the approach of Guo, Li, and Iori [X. Guo, W. Li, and F. Iorio, Convolutional neural networks for steady flow approximation, in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, New York, USA, 2016, ACM, pp. 481–490]. In particular, the neural networks are trained in order to predict images of the flow field in a channel with varying obstacle based on an image of the geometry of the channel. A classical CNN with bottleneck structure and a U-Net are compared while varying the input format, the number of decoder paths, as well as the loss function used to train the networks. This approach yields very low prediction errors, in particular, when using the U-Net architecture. Furthermore, the models are also able to generalize to unseen geometries of the same type. A transfer learning approach enables the model to be trained to a new type of geometries with very low training cost. Finally, based on this transfer learning approach, a sequential learning strategy is introduced, which significantly reduces the amount of necessary training data.

Key Words.
Convolutional neural networks, computational fluid dynamics, reduced order surrogate models, U-Net, transfer learning, sequential learning

AMS Subject Classifications.
35Q30, 68T07, 68T10, 65N22