Estimating divergence-free flows via neural networks

Dmitry I. Kabanov[1,*], **Luis Espath**[1], **Jonas Kiessling**[2], and **Raul F. Tempone**[1,3]

[1] RWTH Aachen University, Pontdriesch 14–16, Aachen 52062, Germany
[2] KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden
[3] King Abdullah University of Science and Technology, Thuwal, 23955-6900, Saudi Arabia

We apply neural networks to the problem of estimating divergence-free velocity flows from given sparse observations. Following the modern trend of combining data and models in physics-informed neural networks, we reconstruct the velocity flow by training a neural network in such a manner that the network not only matches the observations but also approximately satisfies the divergence-free condition. The assumption is that the balance between the two terms allows to obtain the model that has better prediction performance than a usual data-driven neural network. We apply this approach to the reconstruction of truly divergence-free flow from the noiseless synthetic data and to the reconstruction of wind velocity fields over Sweden.

Introduction. An important method of finding relations from a given dataset of observations is fitting a function to this dataset. The function should be "rich" enough so that the relations between independent and dependent variables in the dataset can be appropriately modeled. Recently, special nonlinear functions called neural networks have emerged for the purpose of finding relations in the data. Particularly, an important class of functions named *physics-informed neural networks* [1] has emerged to fit the data, for which one can postulate important physical properties of these data.

We apply physics-informed neural networks to the problem of estimating velocity flows that satisfy the divergence-free condition, namely, flows of incompressible fluids. For that, we train the network to not only satisfy the given dataset of observations but also to approximately satisfy the divergence-free condition.

We demonstrate the performance of such networks on two examples. The first example is based on divergence-free synthetic data, while the second example is based on the dataset of wind velocity observations over Sweden in 2018, where we use an assumption that the wind is incompressible due to the relatively low wind velocities.

Mathematical model. Consider dataset

$$\{\boldsymbol{x}_i, \boldsymbol{u}_i\}, \quad i = 1, \dots, N, \tag{1}$$

where $\boldsymbol{x}_i \in \mathbb{R}^2$ is the ith spatial point, $\boldsymbol{u}_i \in \mathbb{R}^2$ the velocity at this point, N total number of measurement points. Also, assume that the velocity field $\boldsymbol{u}(\boldsymbol{x})$ satisfies the divergence-free condition, at least approximately:

$$\nabla \cdot \boldsymbol{u}(\boldsymbol{x}) \approx 0. \tag{2}$$

We seek an estimator $\hat{\boldsymbol{u}}(\boldsymbol{x}; \boldsymbol{\theta}) \approx \boldsymbol{u}(\boldsymbol{x})$ parameterized by $\boldsymbol{\theta} \in \mathbb{R}^n$ that satisfies the measurements (1) and the divergence-free condition (2) simultaneously in a least-squares sense:

$$\arg\min_{\theta} \quad \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{u}_i - \hat{\boldsymbol{u}}(\boldsymbol{x}_i; \boldsymbol{\theta})\|_2^2 \; + \; \frac{\gamma}{P} \sum_{i=1}^{P} \left(\nabla \cdot \hat{\boldsymbol{u}}(\boldsymbol{x}_i; \boldsymbol{\theta})\right)^2 \tag{3}$$

where γ is regularization parameter that determines how strongly the divergence-free condition is enforced, and P the number of points, at which the divergence-free condition is enforced. Note that the latter points are different from the measurement locations.

As an estimator function $\hat{\boldsymbol{u}}(\boldsymbol{x}; \boldsymbol{\theta})$, we use a neural network of type "multilayer perceptron" [4], which is trained (that is, parameter $\boldsymbol{\theta}$ is fitted) by solving the optimization problem (3) using ADAM optimizer [3]. Also, iterations of $\boldsymbol{\theta}$ are averaged using exponential averaging, and the averaged value is used for making predictions with the trained model.

Numerical examples. We apply divergence-free neural networks to truly divergence-free dataset (Example 1) and to the wind estimation over Sweden (Example 2).

Example 1. We generate synthetic noiseless data $\{(x_i, y_i), \boldsymbol{u}_i\}$, $i = 1, \dots, 10$, where $\boldsymbol{u}(x, y) = (\cos x \, \sin y, -\sin x \, \cos y)^{\mathrm{T}}$ and points (x_i, y_i) are sampled uniformly from domain $[0; 2\pi]^2$. This dataset is a Taylor–Green 2D vortex fixed in time. Test data are located on the 21×21 uniform grid in this domain.

Figure 1 shows the comparison of the prediction errors of two networks: for one, parameter γ is set to zero, such that the network is trained to only match the dataset, for another, parameter $\gamma = 10^{-2}$. We can see that the neural network, trained to satisfy both the dataset and the divergence-free condition, gives smaller prediction errors, especially at locations at distance from the measurement locations.

* Corresponding author: e-mail kabanov@uq.rwth-aachen.de.

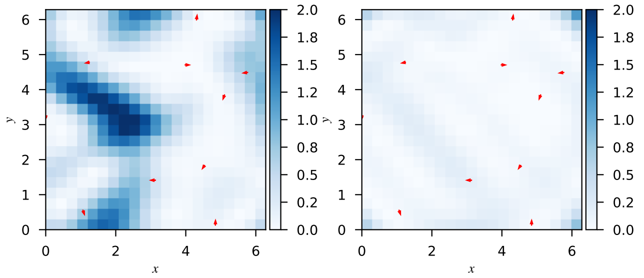

Fig. 1: Comparison of the prediction errors for two neural networks: left) $\gamma = 0$; right) $\gamma = 10^{-2}$. Red arrows show the measurement locations and corresponding velocity directions.

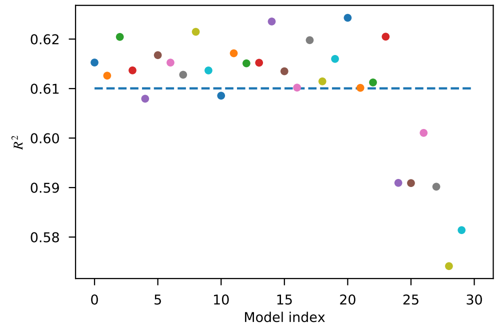

Fig. 2: Comparison of 30 neural-network models (points) with the base IDW model (dashed line). The best found model corresponds to model index 21.

Example 2. We apply a neural network with divergence-free regularization to the problem of reconstruction of the wind fields over Sweden. The assumption is that due to the low speed of the wind, the flow can be considered incompressible. The data were collected from 165 weather stations in 2018, with measurements being recorded hourly, with missing values from different stations.

To assess the quality of the neural network model, we use as a base model the Inverse Distance Weighting (IDW) model [2], which estimates velocity at point $\boldsymbol{x}_j$ by averaging weighted observations:

$$\hat{\boldsymbol{u}}_{\text{IDW}}(\boldsymbol{x}_j) = \frac{\sum_{i=1}^{N} W(r_{i,j}) \boldsymbol{u}(\boldsymbol{x}_j)}{\sum_{i=1}^{N} W(r_{i,j})} \tag{4}$$

where W is a weight, namely, decaying function of the distance $r_{i,j}$ between locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$.

To quantify the quality of the model, we estimate the prediction error by selecting at random from the dataset T time snapshots and computing an average R^2 score using K-fold cross-validation:

$$R^2 = 1 - \frac{\sum_{i=1}^{T} \sum_{j=1}^{K} \frac{1}{N'_{ij}} \sum_{k=1}^{N'_{ij}} \|\boldsymbol{u}_{ijk} - \hat{\boldsymbol{u}}_{ijk}\|_2^2}{\sum_{i=1}^{T} \sum_{j=1}^{K} \frac{1}{N'_{ij}} \sum_{k=1}^{N'_{ij}} \|\boldsymbol{u}_{ijk}\|_2^2}$$

where N'_{ij} denotes number of test measurements for the ith time snapshot and jth cross-validation fold. The R^2 score provides normalized value of the prediction performance of the model as it does not depend on the scaling of the data. We choose $T = 500$ and $K = 5$, therefore, to compute R^2 score 2500 neural networks are required with a given configuration of hyperparameters.

We train multiple neural networks with configurations specified by a triplet: $\gamma \in \{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$, number of neurons in the 1st hidden layer $h_1 \in \{20, 30\}$, number of neurons in the 2nd hidden layer $h_2 \in \{0, 10, 20\}$.

Figure 2 shows comparison of the neural networks with these configurations against the base model (4). The best R^2 score is achieved by the model 21 with parameters $\gamma = 0.01$, $h_1 = 20$, $h_2 = 20$. However, one can see that the $R^2 = 62.5$ of the best model is relatively close to the $R^2 = 0.61$ of the base model.

Conclusions. We applied physics-informed neural networks to estimation of divergence-free flows. Application to the synthetic data, which are known to be truly divergence-free, shows that the addition of the divergence-free regularization significantly improves the prediction error of the neural network, particularly, in the areas with known measurement data. Application to the wind estimation over Sweden shows that divergence-free neural network can achieve better, although marginally, prediction performance than the base inverse distance weighting model. Overall, addition of the physics-informed regularizer improves the prediction performance of neural networks.

Acknowledgements The authors greatfully acknowledge research funding from the Alexander von Humboldt Foundation.

References

[1] M. Raissi, P. Perdikaris, and G. E. Karniadakis. J. Comput. Phys. **378**, 686-707 (2019).
[2] M. Cellura, G. Cirrincione, A. Marvuglia, A. Miraoui. Renew. Energ., **33**, 1251-1266 (2008).
[3] D. P. Kingma, J. Ba. arXiv preprint, arXiv:1412.6980 (2014).
[4] I. Goodfellow, Y. Bengio, and A. Courville. Deep Learning (MIT Press, 2016).