Check for updates

Received: 13 October 2022 | Revised: 23 January 2023 | Accepted: 26 January 2023

DOI: 10.1002/cjce.24870

RESEARCH ARTICLE

THE CANADIAN JOURNAL OF CHEMICAL ENGINEERING

Artificial neural network to predict the power number of agitated tanks fed by CFD simulations

Valérie Bibeau[1,2] | **Lucka Barbeau**[1] | **Daria Camilla Boffito**[2] | **Bruno Blais**[1]

[1]Department of Chemical Engineering, Polytechnique Montréal, Research Unit for Industrial Flows Processes (URPEI), Montréal, Québec, Canada

[2]Department of Chemical Engineering, Polytechnique Montréal, Engineering Process Intensification and Catalysis (EPIC), Montréal, Québec, Canada

Correspondence
Bruno Blais, Department of Chemical Engineering, Polytechnique Montréal, Research Unit for Industrial Flows Processes (URPEI), C.P. 6079, Succ. "CV", Montréal, H3C 3A7, QC, Canada.
Email: bruno.blais@polymtl.ca

Abstract

The power consumption of the agitator is a critical variable to consider in the design of a mixing system. It is generally evaluated through a dimensionless number known as the power number N_p. Multiple empirical equations exist to calculate the power number based on the Reynolds number Re and dimensionless geometrical variables that characterize the tank, the impeller, and the height of the fluid. However, correlations perform poorly outside of the conditions in which they were established. We create a rich database of 100 k computational fluid dynamics (CFD) simulations. We simulate paddle and pitched blade turbines in unbaffled tanks from Re 1 to 100 and use an artificial neural network (ANN) to create a robust and accurate predictor of the power number. We perform a mesh sensitivity analysis to verify the precision of the N_p values given by the CFD simulations. To sample the 100 k mixers by their geometrical and physical properties, we use the Latin hypercube sampling (LHS) method. We then normalize the data with a MinMax transformation to put all features in the same scale and thus avoid bias during the ANN's training. Using a grid search cross-validation, we find the best architecture of the ANN that prevents overfitting and underfitting. Finally, we quantify the performance of the ANN by extracting 30% of the database, predicting the N_p using the ANN, and evaluating the mean absolute percentage error. The mean absolute error in the ANN prediction is 0.5%, and its accuracy surpasses correlations even for untrained geometries.

KEYWORDS

artificial neural networks, computational fluid dynamics, mixing, pre-processing methods

1 | INTRODUCTION

Mixing is one of the most important unit operations in chemical engineering. Applications such as homogenization, emulsification, and aeration for fermentation processes require agitation. To design mixing systems, important quantities need to be considered, such as the power consumption of the agitator and the dimensionless mixing time. The choice of the agitator is often motivated by the regime of operation (laminar or turbulent), the kinematic viscosity, and the volume of the fluid being mixed. Different kinds of impellers

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2023 The Authors. The *Canadian Journal of Chemical Engineering* published by Wiley Periodicals LLC on behalf of Canadian Society for Chemical Engineering.

are used for different operating conditions. For example, axial or radial discharge turbines (propellers, Rushton, hydrofoils, and pitched blades) are used for fluids with lower viscosity, whereas close-clearance impellers (helical ribbons and anchors) are used for viscous fluids.[1]

To characterize the power consumption, engineers rely extensively on the dimensionless power number N_p, which relates the power consumption of the agitator to the inertial work in the fluid and characterizes the energy requirement of an agitated system:

$$N_p = \frac{P}{\rho N^3 D^5} \tag{1}$$

with ρ representing the fluid's density, N the impeller rotational speed, D the impeller diameter, and P the power consumption.[1–3]

The power number depends on dimensionless geometrical properties that describe a given mixing system as well as the Reynolds number Re:

$$\mathrm{Re} = \frac{D^2 N}{\nu} \tag{2}$$

with ν representing the kinematic viscosity.[1–3]

The power number N_p can be obtained through two approaches: empirical correlations and computational fluid dynamics (CFD) simulations.

Empirical correlations are power curves, where N_p is a logarithm function of Re and geometrical ratios using the impeller diameter D as the reference length. These correlations are generally built using experimental data gathered on mixing systems.[4]

$$N_p = f\left(\mathrm{Re}, \frac{T}{D}, \frac{H}{D}, \frac{C}{D}, \frac{l}{D}, \frac{w}{D}\right) \tag{3}$$

Here, T is the width of the tank, H is the height of the fluid inside the tank, C is the clearance of the impeller from the bottom, l is the length of the blades, and w is the width of the blades.

Those power curves behave differently for the laminar regime (at low Reynolds) and for the turbulent regime (at higher Reynolds). In the laminar regime, $N_p \propto \mathrm{Re}^{-1}$, meaning that the slope of N_p is constant in logarithmic scale. On the other hand, in the turbulent regime, N_p becomes constant. Usually, the transition between the laminar regime and the turbulent regime occurs around $\mathrm{Re} = 200$.[1]

Multiple correlations can be used to estimate the power consumption of paddle and pitched blade turbine (PBT) impellers.[5–8] These correlations generally fit on power numbers obtained experimentally.

CFD is a robust and accurate tool to simulate mixing systems. With validated software and an appropriate mesh, velocity field, pressure, and torque can be easily computed from CFD simulations.[2]

These two approaches present some limitations. On the one hand, the correlations do not generalize to multiple geometries without requiring new experimental data. Indeed, the dependency between the power number N_p and the mixer geometry is so complex and non-linear that an algebraic function cannot adequately represent all possibilities. Furthermore, only few variations of geometrical characteristics are investigated in the experiments, which leads to undercoverage biases and, thus, inaccuracies in the calculation of the power number N_p.

On the other hand, if we use CFD simulations to predict the power consumption of a mixing system, a mesh sensitivity analysis and multiple simulations are needed to establish a power curve. This requires significant time, computational resources, and expertise in meshing and CFD.

Artificial neural networks (ANNs) are statistical models that can predict continuous values through a regression technique. By propagating signals through multiple interconnected nodes that form a deep network, ANNs take into account non-linear relationships between inputs and outputs, like empirical correlations.[9] ANNs have been shown to be able to model accurately multiple chemical engineering problems such as internal faults detection in heat exchangers,[10] process control using model predictive control (MPC),[11] and also kinetic modelling to optimize fluidized bed reactor performance.[12] In mixing applications, it was demonstrated that ANNs coupled with empirical correlations can predict the power and head characteristics of pump-mixers from experimental data.[13] Nonetheless, ANNs have never been used as an independent model to predict the power number N_p.

To overcome the lack of generality of the traditional correlations and to avoid gathering an unrealistic amount of experimental data, we propose the combination of CFD and machine learning to estimate the power number of impellers. CFD can simulate mixing systems in large numbers (e.g., 100 k configurations) in a relatively short time thanks to high-performance clusters. The machine learning approach gives a simpler and faster predictor of the power number than CFD and a more precise statistical regressor than correlations.

The main objective of this study is to predict power curve from Re 1 to 100 of paddle and PBT mixers given their geometrical characteristics and operating conditions.

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

First, we present the CFD methods and settings to predict the power number for a mixing configuration. We then describe the sampling method used to generate efficiently a hundred thousand (100 k) mixing simulations for the purpose of feeding a rich and large database to the ANN. Using a grid search with cross-validation (CV), we construct the ANN's architecture by finding its optimized hyperparameters. Finally, we test and analyze the performance of the ANN's training by evaluating the loss function and by comparing the ANN's results to an empirical correlation.

Using an effective sampling method of the CFD simulations and performing a data pre-processing, the ANN is able to predict the power number with an absolute error of 0.5%. It significantly outperforms the correlation because it is able to generalize its predictions of the power number to a larger breadth of geometrical configurations.

2 | MODELS AND METHODS

2.1 | Database generation

This section discusses the strategy employed to generate the mixer database. We describe the dimensionless parameterization of the geometry, the meshing procedure, the CFD solver, and the sampling strategy.

2.1.1 | Domain parameterization

The geometry of the mixer is defined by seven different parameters, some of which are shown in Figure 1. The tank is unbaffled, has a diameter of T, and the fluid inside the tank has a height of H. The impeller has a diameter D. The blades of the impeller have a width of W, a thickness of E, and can be tilted at an angle θ. Finally, the impeller has a clearance C from the bottom of the tank. The geometry is scaled using common ratios seen in the literature to generate a dimension-independent problem: T/D, H/T, T/C, D/W, and E/W.[2,4,14–16] The reference dimension is set to $T=1$.

2.1.2 | CFD simulations

The open-source software Gmsh is used to generate a finite element mesh of tetrahedral elements.[17] The open-source CFD software Lethe is used to solve the incompressible Navier–Stokes equations using streamline-upwind/Petrov–Galerkin (SUPG) and pressure-stabilizing/Petrov–Galerkin (PSPG) formulations.[18] We use P1–P1 interpolation for both velocity and pressure. The mixer is simulated in steady state regime at a low Reynolds number ($\mathrm{Re} \in [1,100]$) where a stable stationary solution is reached. A single rotating frame (SRF) technique is used, meaning that the rotational velocity is imposed on the lateral and bottom walls while the impeller has a zero velocity condition.[19] The fluid interface at the top has a slip condition to mimic the liquid's free surface. Note that this method is limited to vessels of a single shaft with no baffle, because the static part of the tank needs to stay invariant by rotation.[2]

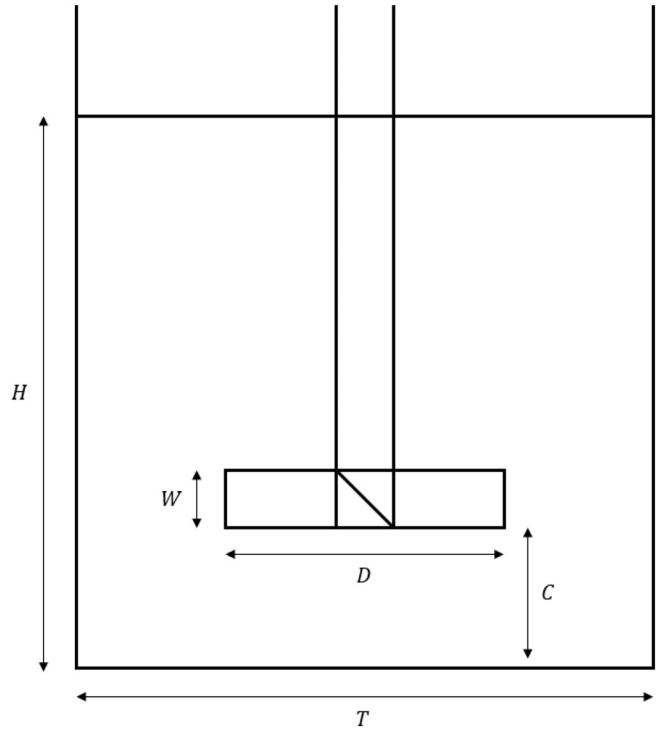

FIGURE 1 Scheme of the mixing rig used for the simulations.

Because of the SRF, the Coriolis and centrifugal forces are added to the Navier–Stokes equations as follows[2,3]:

$$\nabla \cdot \boldsymbol{u} = 0 \tag{4}$$

$$\nabla \cdot \rho \boldsymbol{u} \otimes \boldsymbol{u} + 2\rho \boldsymbol{\omega} \times \boldsymbol{u} + \rho \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) = -\nabla p + \nabla \cdot \boldsymbol{\tau} \tag{5}$$

where $\boldsymbol{u}$ is the velocity, ρ is the density, $\boldsymbol{\omega}$ is the angular velocity of the SRF, r is the distance to the axis rotation, p is the pressure, and τ is the deviatoric stress tensor[2,3,18]:

$$\tau = \nu \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T \right) \tag{6}$$

where ν is the kinematic viscosity of the fluid.

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The rotational speed of the reference frame N is set to one rotation per second (or $\|\omega\| = 2\pi$ rad$\cdot$s^{-1}), and the kinematic viscosity is changed in the simulation parameters to satisfy a specified Reynolds number Re described by Equation (2).

The torque is obtained by integrating the stress tensor on the impeller. The power consumption P is a function of the torque Γ and the rotational speed N.[1–3]

$$P = 2\pi\Gamma N \tag{7}$$

The power number N_p is obtained from the power consumption P following Equation (1).

2.1.3 | Sampling strategy

To generate 100 k samples that are standardized in all geometrical ratios and Re, Latin hypercube sampling (LHS) is used. LHS is a sampling method that stratifies the features to uniformly represent the inputs space.[20] The open-source library design of experiments for Python (pyDOE) is used to generate the 100 k combinations of the input variables. Table 1 lists the interval of the dimensionless geometrical ratios, and the Reynolds numbers that are used to sample the mixers.

The 100 k simulations are launched on the Narval computer cluster owned by the Digital Research Alliance of Canada. Each simulation is launched on a single core of one CPU and takes approximately 40 min to reach convergence at a relative tolerance of 10^{-8}. A script is launched to gather the torque output of all the simulations and the power numbers N_p are stored in a data file.

2.2 | Artificial neural network

The architecture of the ANN is a deep structure composed of multiple layers of neurons. Figure 2 represents an example of a deep neural network with one input layer of seven neurons, two hidden layers, and an output layer of the variable of interest N_p. The input layer regroups all the variables that characterize agitation: six geometrical parameters (ratios) and one physical parameter (Re). Giving all dimensionless parameters as inputs to the ANN allows for more stability to the training.

When training an ANN, the goal is to find the right weights between the interconnected nodes from a layer $[i-1]$ to another layer $[i]$ and construct the best hypothesis function that links the inputs to the outputs. Each set of nodes $\boldsymbol{z}^{[i]}$ is a linear combination of their input neurons $\boldsymbol{a}^{[i-1]}$ and their respective weights $W^{[i]}$. A bias $\boldsymbol{b}^{[i]}$ is added as a constant to the linear function.[21]

TABLE 1 Input variables and their interval.

Inputs	Minimum	Maximum
T/D	2	4
H/T	1	1.5
T/C	2	5
D/W	3	6
E/W	0.1	0.2
θ	0	$\pi/3$
Re	1	100

$$\boldsymbol{z}^{[i]} = W^{[i]}\boldsymbol{a}^{[i-1]} + \boldsymbol{b}^{[i]} \tag{8}$$

The number of layers $[i]$ and the number of neurons in one layer (size of $\boldsymbol{a}^{[i]}$) are hyperparameters that affect the performance of the ANN. Another hyperparameter to study is the non-linear activation function σ that transforms a set of nodes $\boldsymbol{z}^{[i]}$ into neurons $\boldsymbol{a}^{[i]}$.

$$\boldsymbol{a}^{[i]} = \sigma\left(\boldsymbol{z}^{[i]}\right) \tag{9}$$

Some activation functions, such as the rectified linear unit function (ReLU) and the hyperbolic tangent function (tanh), are more commonly used for deep networks and will be further investigated in this study. The activation function for the output layer should differ from the hidden layers. Given that the problem to be solved is a linear regression, the output neuron is activated using a linear function.[9]

The quantification of the ANN's performance needs to be consistent with the formulation of the problem. As the power number is a continuous variable, the loss function L computed at the output layer that verifies the ANN's performance is the squared error between the predicted output $\hat{y}^{(i)}$ and the known value $y^{(i)}$. In this study, $y^{(i)}$ is the power number (N_p) calculated using CFD simulations.

$$L\left(\hat{y}^{(i)}, y^{(i)}\right) = \left(\hat{y}^{(i)} - y^{(i)}\right)^2 \tag{10}$$

The cost function J is the mean squared error (MSE) over all the outputs.

$$J = \frac{1}{m}\sum_{i=1}^{m} L\left(\hat{y}^{(i)}, y^{(i)}\right) = \frac{1}{m}\sum_{i=1}^{m}\left(\hat{y}^{(i)} - y^{(i)}\right)^2 \tag{11}$$

The optimization problem of an ANN is to minimize the cost function J, which means that the derivative of

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

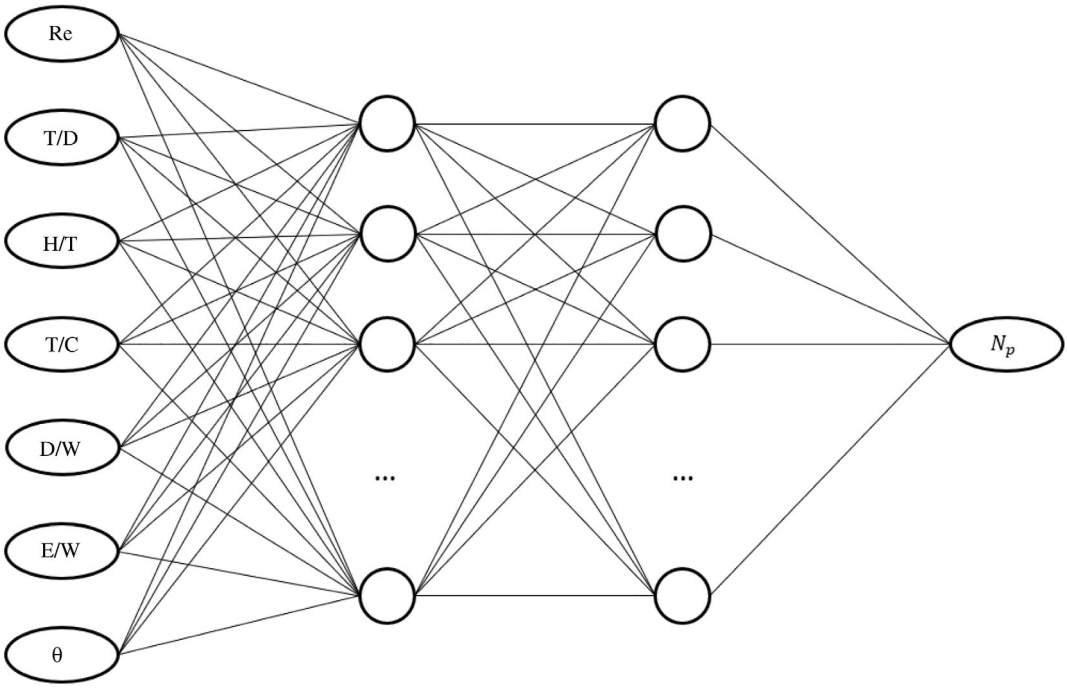

FIGURE 2 Artificial neural network (ANN) architecture with two hidden layers to predict the power number based on features that describe the mixer.

the cost function with respect to the weights needs to converge close to zero. Backpropagation allows the evaluation of the derivatives automatically. The algorithm that performs the optimization is the gradient descent technique. It updates the weights according to a learning rate α and a direction given by the derivatives.[9]

$$W^{[i]} := W^{[i]} - \alpha \frac{\partial J}{\partial W^{[i]}} \tag{12}$$

When the algorithm completes one pass of all the dataset over the neural network, one epoch (iteration) is considered finished. The number of epochs is also another important hyperparameter to consider.

The learning may be slow, and the ANN may not converge to a global minimum. Some strategies are implemented to optimize the gradient descent algorithm. First, the Adam optimizer uses momentum optimization by keeping track of past gradients to accelerate the convergence.[9] Second, the weights W are initialized using the Glorot uniform distribution. This technique randomly initializes the connection weights depending on the number of neurons in each layer. The Glorot distribution has been proven to improve the flow of the signal through the layers in the forward and backward directions.[22]

2.2.1 | Pre-processing

Before training the ANN, pre-treatment of the input data is important to avoid biases. Normalization of data consists of bringing all the features on the same interval. Indeed, the Reynolds number does not have the same order of magnitude as the geometrical ratios. The update of the weights of the ANN in the backward pass can be affected by the different scales of the features. Training performed on unnormalized mixers data tended to converge to a high MSE, which indicates poor performance.

The normalization method used for the purpose of inputs pre-treatment is MinMax scaling. Each feature x_i is normalized into x_i' using the minimum $x_{\min}$ and the maximum $x_{\max}$ of the given feature.

$$x_i' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{13}$$

When MinMax scaling is applied to the data, each variable is exactly scaled between 0 and 1 and keeps the covariance structure of the original data.[23] A principal component analysis (PCA) on the normalized data shows that there are no dominant directions where the variance of the features is higher. Therefore, all features will be used in the training of the ANN, with no dimensionality reduction being employed.

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2.2 | Training

The architecture of the ANN is set up with an exhaustive grid search with CV that consists of selecting the most suitable hyperparameters. The normalized data are divided into three sets: training, validation, and testing. 70% of the data is used for the training and 30% is used for testing the performance of the ANN. To validate the ANN, CV is used to monitor the training in real time and to verify if the model is over-fitted.[9] The grid search allows one to train different architectures by changing the number of hidden layers and neurons. The best ANN structure is chosen based on the lowest validation MSE averaged over multiple folds of the data. The cost function of the validation over each epoch is also checked. If the cost function increases, the model becomes more biased on the training set and fails to generalize. Finally, when the best architecture is found, all the training data are used to fit the model with the right number of epochs that prevent overfitting. Predictions are then made with the testing set to verify the performance of the ANN. The grid search CV is implemented using the Sklearn library in Python.[24] All the implementation and the training of the ANN is done with the Tensorflow and Keras libraries in Python.[25] All the code programs, scripts, and data files of the ANN's implementation are available on a public repository.[26]

3 | MESH SENSITIVITY

In CFD, the mesh impacts the quality of the results and the computation time. The mesh around the impeller needs to be sufficiently fine to have an adequate prediction of the flow and thus to adequately predict the torque. Using Gmsh, we locally refine the mesh near the agitator. The characteristic length L_c (the representative size of a tetrahedral element), is 10 times smaller near the blades and the shaft than near the walls. The value of L_c is 0.003 near the agitator and 0.03 near the walls. Figure 3 shows the velocity in the radial direction while Figure 4 shows the velocity in the axial direction of the same vessel with $T/D = 3$, $H/T = 1$, $T/C = 3$, $D/W = 4$, $E/W = 0.1$, $\theta = \pi/4$, and $\mathrm{Re} = 100$. Both figures are in a Eulerian reference and use a refined mesh, as illustrated in Figure 5.

To evaluate the error on the torque caused by the CFD simulations, we need to do a mesh sensitivity analysis. We gradually refine the global mesh by reducing the characteristic length of the elements and evaluate the torque on the impeller. We do a mesh sensitivity using both an average geometry (medium impeller in a medium tank) at low Reynolds ($\mathrm{Re} = 5$), defined as the average case, and a more complex geometry (smaller impeller in

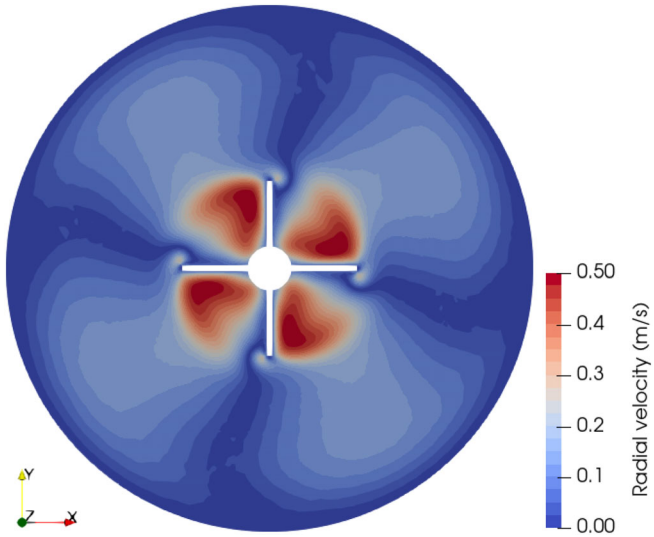

FIGURE 3 Velocity in the radial direction in a mixer at $\mathrm{Re} = 100$ in a top-view perspective.

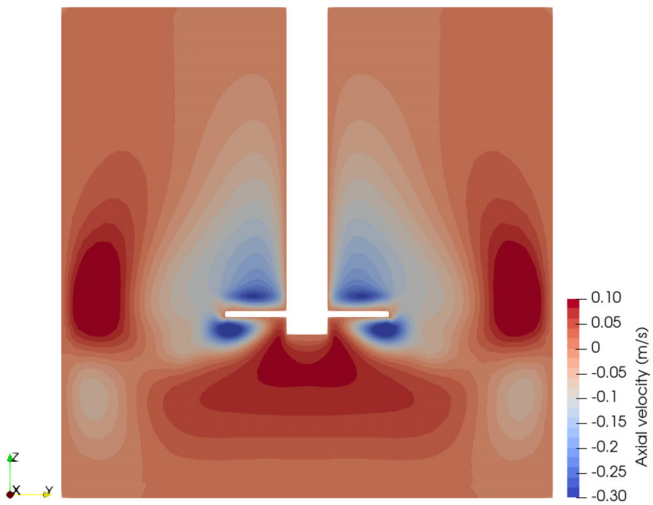

FIGURE 4 Velocity in the axial direction at $\mathrm{Re} = 100$ in a side-view perspective.

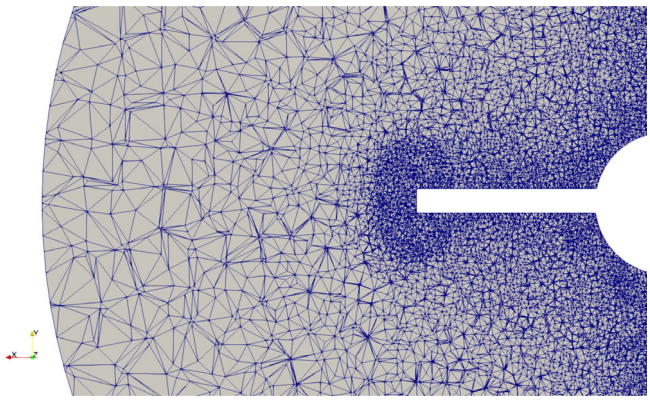

FIGURE 5 Mesh refinement near the blades and the shaft of the impeller.

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

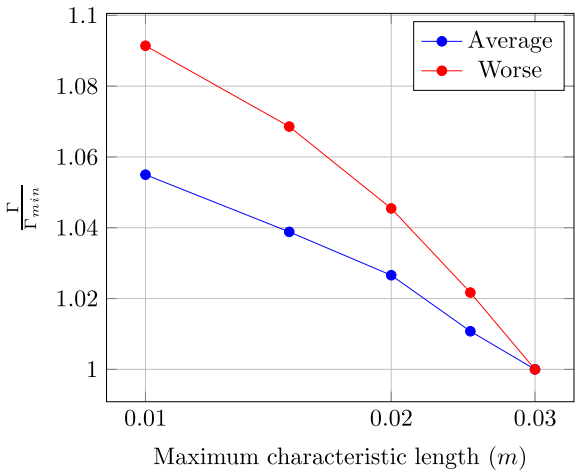

FIGURE 6 Normalized torque Γ according to the maximum characteristic length of the mesh near the agitator for average and worse geometries.

a larger tank) at higher Reynolds (Re = 100), defined as the worst case.

Figure 6 shows the torque obtained with a given characteristic length normalized to the torque calculated with the mesh used in the 100 k CFD simulations ($L_c = 0.03$). We observe that the torque is still mesh-dependent, even with the finer mesh, which contains almost 100 M elements. We use a generalized Richardson extrapolation to approximate the reference value of the torque and then to calculate the error caused by the domain discretization.[27]

$$\widehat{f} = f_h + \frac{f_h - f_{rh}}{r^p - 1} \tag{14}$$

Here, $\widehat{f}$ is the extrapolated value, or the approximation of the solution of the torque. f_h and f_{rh} are the torque evaluated on elements of different characteristic lengths h. r is the refinement factor between the meshes and p is the order of convergence of the torque, which is 1.

The extrapolated normalized torque is 1.07 for the average mixer and 1.14 for the worse mixer. With these values, it is possible to have an estimation of the relative error of the torque e_r.

$$e_r = \frac{|\widehat{f} - f_{0.03}|}{\widehat{f}} \tag{15}$$

Relative errors of 7% for the average mixer and 12% for the worse mixer are obtained with a mesh with a maximum characteristic length of 0.03. These errors are not significant, as they are equivalent to an uncertainty on the impeller speed of about 2%.

TABLE 2 Performance metrics on each set.

Set	MSE (10^{-4})	MAE (10^{-2})	MAPE (%)
Training	6.10	1.84	0.538
Testing	7.60	1.88	0.540

Abbreviations: MAE, mean absolute error; MAPE, mean absolute percentage error; MSE, mean squared error.

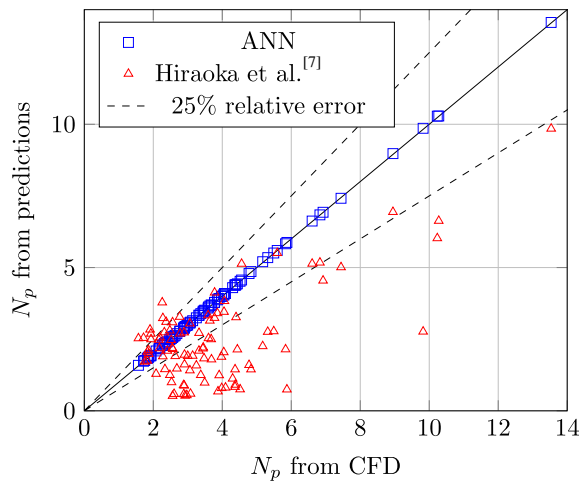

FIGURE 7 Predictions using the artificial neural network (ANN) and the correlation from Hiraoka et al.[7] versus values of the power number N_p from the computational fluid dynamics (CFD) simulations.

Furthermore, the run time of a simulation using a mesh of 3 million elements ($L_c = 0.03$) is 40 min on a single CPU, while a mesh of 5 million elements ($L_c = 0.025$) takes 80 min to simulate, which is two times longer. Simulations with close to 100 M cells required a full cluster node in order to have sufficient memory.

Using a coarser mesh is a good compromise between computational time and accuracy. Using this characteristic length, the cost of a single simulation was approximately $1.2 \cdot 10^{-4}$ core-year. In total, the simulations carried out in this work consumed 12 core-years.

4 | RESULTS AND DISCUSSION

The grid search CV finds the right balance between the complexity of the architecture (number of parameters) and the cycles of training (amount of times the parameters are updated). An ANN with three layers and 50 neurons in each layer is the architecture that gives the lowest averaged MSE over all the CV folds. In this case, the tanh activation function converges more rapidly than the ReLU function. A smaller batch size of 200 updates the weights more frequently during one epoch and stabilizes the variation of the MSE during training.

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Correlation equations for the power number for pitched and paddle impellers in unbaffled condition.

$N_p = \left(\left[1.2\pi^4\beta^2\right]/\left[8D^3/(T^2H)\right]\right)f$
$f = C_L/\mathrm{Re}_G + C_t\left(\left[(C_{tr}/\mathrm{Re}_G) + \mathrm{Re}_G\right]^{-1} + (f_\sim/C_t)^{1/m}\right)^m$
$\mathrm{Re}_G = \left(\left[\pi\eta\ln(T/D)\right]/(4D/\beta T)\right)\mathrm{Re}$
$C_L = 0.215\eta n_p(D/H)\left[1-(D/T)^2\right] + 1.83(W\sin\theta/H)\left(n_p/2\sin\theta\right)^{1/3}$
$C_t = \left[\left(1.96X^{1.19}\right)^{-7.8} + (0.25)^{-7.8}\right]^{-1/7.8}$
$m = \left[\left(0.71X^{0.373}\right)^{-7.8} + (0.333)^{-7.8}\right]^{-1/7.8}$
$C_{tr} = 23.8(D/T)^{-3.24}(W\sin\theta/T)^{-1.18}X^{-0.74}$
$f_\sim = 0.0151(D/T)C_t^{0.308}$
$X = \gamma n_p^{0.7}W\sin^{1.6}\theta/H$
$\beta = 2\ln(T/D)/\left[(T/D)-(D/T)\right]$
$\gamma = \left[\eta\ln(T/D)/(\beta T/D)^5\right]^{1/3}$
$\eta = 0.711\left(0.157 + \left[n_p\ln(T/D)\right]^{0.611}\right)/\left(n_p^{0.52}\left[1-(D/T)^2\right]\right)$

After CV, the ANN was trained for 5000 epochs with the optimal architecture on all the training set. Table 2 shows the MSE, the mean absolute error (MAE), and the mean absolute percentage error (MAPE) of the training and testing sets with the raw data. The scale of each performance metric is the same for training and testing, meaning that there is no underfitting or overfitting of the data. The MAPE of the testing set, which was never seen by the ANN before, is lower than 1%.

Figure 7 shows the predictions of the ANN and the estimations of a correlation developed by Hiraoka et al. for paddle impellers[7,28] using a portion of the testing set. The correlation is detailed in Table 3. It uses a friction factor f, a laminar term C_L, a turbulent term C_t, a term that characterizes the transition from laminar to turbulent flow C_{tr}, an exponent term m, and other coefficients. It also uses a modified Reynolds number Re_G. Note that n_p is the number of blades, which is four in this case.

According to Figure 7, the ANN's predictions are really close to the true power number values, and the errors are normally distributed, with no apparent outliers. These results show that the ANN performs well. Normalization of the data plays a major role in obtaining low error metrics. The correlation of Hiraoka et al. exceeds mostly the 25% error threshold illustrated in Figure 7, showing its poor performance. The correlation severely underpredicts and overpredicts N_p.

To further validate the accuracy and the robustness of the ANN, we consider impeller configurations that were not included in the database and compare the predictions that are made with the ANN with new CFD simulations

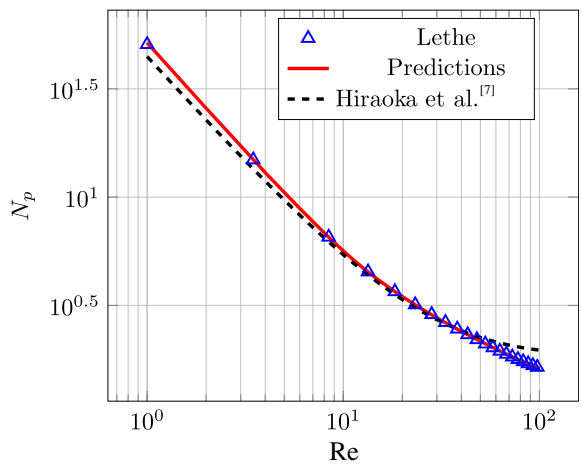

FIGURE 8 Power number versus Reynolds number for a specified geometry ($T/D = 3$, $H/T = 1$, $T/C = 4$, $D/W = 5$, $E/W = 0.1$, and $\theta = \pi/4$) using the artificial neural network (ANN), computational fluid dynamics (CFD) simulations (Lethe), and Hiraoka et al.[7] correlation.

and the correlation developed by Hiraoka et al. over the entire range of Reynolds number considered. The results of the comparison are shown in Figures 8 and 9.

In Figures 8 and 9, predictions agree almost perfectly with the Lethe simulations of the same mixer geometry, proving that the ANN performance is related to the pre-processing of the 100 k runs (LHS sampling, coherent CFD tolerance, proportional refinement of the mesh, etc.).

According to Figure 9, even when compared to results obtained with a finer mesh (smaller element size L_c), the

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

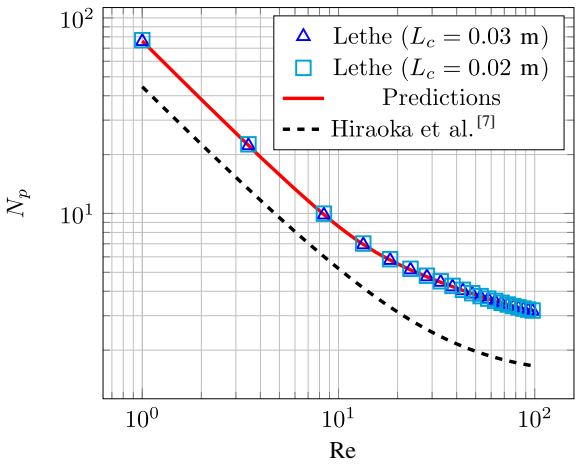

FIGURE 9 Power number versus Reynolds number for a specified geometry ($T/D = 2.5$, $H/T = 1.2$, $T/C = 3$, $D/W = 3.5$, $E/W = 0.15$, and $\theta = \pi/6$) using the artificial neural network (ANN), computational fluid dynamics (CFD) simulations (Lethe) of two different element sizes L_c, and Hiraoka et al.[7] correlation.

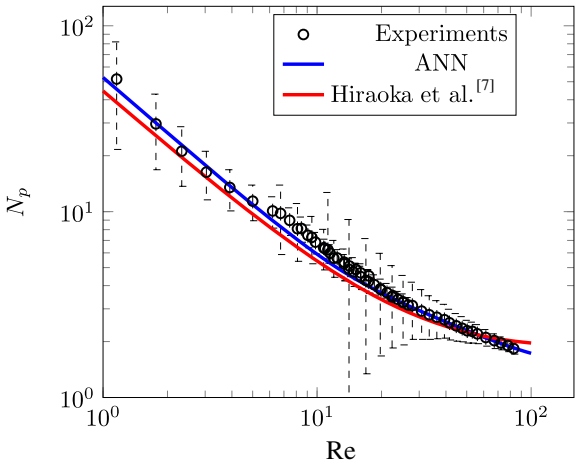

FIGURE 10 Validation of the artificial neural network's (ANN's) prediction of a power curve for a specified geometry ($T/D = 3$, $H/T = 1$, $T/C = 4$, $D/W = 5$, $E/W = 0.15$, and $\theta = \pi/4$).

ANN predictions are extremely close to the CFD simulations. This shows that the precision of the torque due to the mesh resolution does not affect the performance of the ANN. Having a discretization error of 7% on the torque is more than acceptable and allows the ANN to obtain significantly better results than the correlation of Hiraoka et al.

Indeed, the correlations do not fit the N_p values predicted by the ANN or calculated by Lethe. In Figure 8, the correlation deviates close to $\mathrm{Re} = 100$. It is well known that in this area of the power curve, the transition between the laminar flow regime and the turbulent flow regime is difficult to predict.[29,30] In Figure 9, the correlation completely deviates from the CFD simulations and the ANN predictions. Because correlations reproduce only a few experimental values, they are ill-posed to estimate the power number for higher Reynolds, where the transient regime occurs, and for other specific geometries. Lethe has been verified and validated for mixing problems, thus it is able to simulate the flow of complex systems and accurately predict the power number.

To prove the validation further, we compared the N_p values predicted by the ANN with experimental data.[2] Figure 10 shows that the ANN's predictions are within the error intervals of the experiments. The correlation deviates more from the experimental values than the ANN's predictions and it often lies outside of the error bounds, especially at higher Reynolds numbers.

In general, the ANN greatly surpasses the correlation in terms of fitting the CFD data and, consequently, is in much better agreement with the experimental data.

5 | CONCLUSION

The determination of power consumption is a key step in designing a mixing system. The power number N_p represents the energy requirement of the agitator, depending on its geometry and the Reynolds number *Re*. The power number can be obtained from correlations. However, these correlations do not generalize when the mixing system under study slightly deviates from those used to build and verify the correlations.

To overcome this lack of generality, we proposed in this work to couple CFD simulations with an ANN. On one hand, CFD simulations provide the database of 100 k different PBT mixers that are used as an input of the ANN. On the other hand, the ANN learns nonlinear patterns within the database and ultimately estimates the power consumption of a PBT mixing system.

We presented the methodology behind the generation of the 100 k mixers, such as the domain parameterization, the CFD settings, and the LHS method. We also presented the pre-processing of the data and the grid search CV that allowed an optimized architecture and an effective training of the ANN.

Finally, we presented the performance of the ANN by keeping 30% of the data and making predictions of the power number N_p over the testing set using the ANN correlation. The ANN reached a MAPE of 0.5%, which is an incredibly low metric error. The key to getting such a low MAPE is in the pre-processing of the data and the monitoring of the training:

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- Features were sampled using a LHS method, so that all inputs represent faithfully the dimensions of the problem.
- Mesh generation was standardized and all simulations converged to the same tolerance.
- Normalization of the data using a MinMax scaling helped the ANN to overpass inevitable biases.
- Grid search CV was essential to validate the architecture and to monitor the training so that the ANN did not overfit or underfit the data.

Also, we presented some comparisons between the ANN predictions and the correlations from Hiraoka et al. The results showed that the predictions of the power number N_p made by the ANN outperformed the correlation for all cases considered, even for mixers that had never been trained by the ANN.

With these results, we demonstrated that ANNs, trained with CFD data, accurately predict the power consumption of paddle and pitched blade impellers. This approach can be easily generalized to other types of impellers or to the turbulent regime by enriching the database with new simulations. It could also be extended to predict other quantities such as the mixing time, the pumping number, and so on. We think that this work demonstrates that simulation-fed ANNs are an ideal candidate to replace traditional correlations in the design of unit operations such as agitated tanks.

AUTHOR CONTRIBUTIONS

Valérie Bibeau: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; software; validation; visualization; writing – original draft; writing – review and editing. **Lucka Barbeau:** Data curation; methodology; software; validation; writing – review and editing. **Daria Camilla Boffito:** Supervision; writing – review and editing. **Bruno Blais:** Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; software; supervision; validation; visualization; writing – review and editing.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/cjce.24870.

DATA AVAILABILITY STATEMENT

Data openly available in a public repository that does not issue DOIs

ORCID

Lucka Barbeau https://orcid.org/0000-0003-3788-5327
Bruno Blais https://orcid.org/0000-0001-6053-6542

REFERENCES

[1] E. L. Paul, V. A. Atiemo-Obeng, S. M. Kresta, *Handbook of Industrial Mixing*, Wiley Online Library, Hoboken, NJ **2004**.
[2] B. Blais, M. Lassaigne, C. Goniva, L. Fradette, F. Bertrand, *Comput. Chem. Eng.* **2016**, *85*, 136.
[3] R. B. Bird, W. E. Stewart, E. N. Lightfoot, *Transport Phenomena*, John Wiley & Sons Inc., Hoboken, NJ **2007**.
[4] R. L. Bates, P. L. Fondy, R. R. Corpstein, *Ind. Eng. Chem. Process Des. Dev.* **1963**, *2*(4), 310.
[5] S. Nagata, *Mixing: Principles and Applications*, Kodansha, Tokyo, Japan **1975**.
[6] S. Nagata, T. Tokoyama, H. Maeda, *Kagaku Kogaku* **1956**, *20*(11), 582.
[7] S. Hiraoka, N. Kamei, Y. Kato, Y. Tada, H.-G. Cheon, T. Yamaguchi, *Kagaku Kogaku Ronbunshu* **1997**, *23*(6), 969.
[8] Y. Kato, Y. Tada, Y. Takeda, Y. Hirai, Y. Nagatsu, *J. Chem. Eng. Jpn.* **2009**, *42*(1), 6.
[9] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media, Inc., Sebastopol, CA **2019**.
[10] W. Suewatanakul, *Ph.D. Thesis*, University of Texas (Austin, TX) **1993**.
[11] J. C. MacMurray, D. M. Himmelblau, *Comput. Chem. Eng.* **1995**, *19*(10), 1077.
[12] M. Rostamizadeh, D. C. Boffito, G. S. Patience, A. Taeb, *Journal of Chemical and Process Engineering* **2014**, *1*, 1.
[13] K. K. Singh, K. T. Shenoy, A. K. Mahendra, S. K. Ghosh, *Chem. Eng. Sci.* **2004**, *59*(14), 2937.
[14] R. Hemrajani, G. Tatterson, *Mechanically Stirred Vessels*, Wiley Online Library, Hoboken, NJ **2004**, p. 345.
[15] D. Chapple, S. M. Kresta, A. Wall, A. Afacan, *Chem. Eng. Res. Des.* **2002**, *80*(4), 364.
[16] K. Rutherford, S. M. Mahmoudi, K. C. Lee, M. Yianneskis, *Chem. Eng. Res. Des.* **1996**, *74*(3), 369.
[17] C. Geuzaine, J.-F. Remacle, *International Journal for Numerical Methods in Engineering* **2009**, *79*(11), 1309.
[18] B. Blais, L. Barbeau, V. Bibeau, S. Gauvin, T. El Geitani, S. Golshan, R. Kamble, G. Mirakhori, J. Chaouki, *SoftwareX* **2020**, *12*, 100579.
[19] B. Delacroix, A. Bouarab, L. Fradette, F. Bertrand, B. Blais, *Powder Technol.* **2020**, *369*, 146.
[20] M. D. McKay, R. J. Beckman, W. J. Conover, *Technometrics* **2000**, *42*(1), 55.
[21] M. Zakaria, M. Al-Shebany, S. Sarhan, *International Journal of Engineering Research and Applications* **2014**, *4*(2), 7.
[22] X. Glorot, Y. Bengio, in *Proc. of the Thirteenth Int. Conf. on Artificial Intelligence and Statistics*, JMLR Workshop and Conference Proceedings, Cambridge, MA **2010**, pp. 249–256.
[23] A. Gökhan, C. O. Güzeller, M. T. Eser, *International Journal of Assessment Tools in Education* **2019**, *6*(2), 170.
[24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, *Journal of Machine Learning Research* **2011**, *12*, 2825.
[25] Tensorflow: Large-Scale Machine Learning on Heterogeneous Systems, https://www.tensorflow.org/ (accessed: October 2022).

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[26] Lethe-Cfd/Mixing-Ann, https://github.com/lethe-cfd/mixing-ann (accessed: January 2023).
[27] W. L. Oberkampf, C. J. Roy, *Verification and Validation in Scientific Computing*, Cambridge University Press, Cambridge, UK **2010**.
[28] H. Furukawa, Y. Kato, Y. Inoue, T. Kato, Y. Tada, S. Hashimoto, *Int. J. Chem. Eng.* **2012**, *2012*, 1.
[29] F. R. Hama, J. D. Long, J. C. Hegarty, *J. Appl. Phys.* **1957**, *28*(4), 388.
[30] P. A. Tanguy, F. Thibault, *Can. J. Chem. Eng.* **2002**, *80*(4), 601.

How to cite this article: V. Bibeau, L. Barbeau, D. C. Boffito, B. Blais, *Can. J. Chem. Eng.* **2023**, *101*(10), 5992. https://doi.org/10.1002/cjce.24870

1939019x, 2023, 10, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cjce.24870 by Ecole Polytech De Montreal, Wiley Online Library on [08/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License