Nuclear Engineering and Design 324 (2017) 27–34

Contents lists available at ScienceDirect

Nuclear Engineering and Design

journal homepage: www.elsevier.com/locate/nucengdes

Nuclear energy system's behavior and decision making using machine learning

Mario Gomez Fernandez[a,c,*], Akira Tokuhiro[b], Kent Welter[c], Qiao Wu[a]

[a] School of Nuclear Science and Engineering, Oregon State University, 100 Radiation Center, Corvallis, OR 97330, United States

[b] Energy Systems and Nuclear Science Research Centre, University of Ontario Institute of Technology, Room 4036, 2000 Simcoe Street North, Oshawa, ON L1H 7K4, Canada

[c] NuScale Power, LLC, 1100 NE Circle Boulevard, Suite 200, Corvallis, OR 97330, United States

ARTICLE INFO

Keywords:
Decision-making optimization
Nuclear energy systems
Machine learning
Small modular reactors

ABSTRACT

Early versions of artificial neural networks' ability to learn from data based on multivariable statistics and optimization demanded high computational performance as multiple training iterations are necessary to find an optimal local minimum. The rapid advancements in computational performance, storage capacity, and big data management have allowed machine-learning techniques to improve in the areas of learning speed, non-linear data handling, and complex features identification. Machine-learning techniques have proven successful and been used in the areas of autonomous machines, speech recognition, and natural language processing. Though the application of artificial intelligence in the nuclear engineering domain has been limited, it has accurately predicted desired outcomes in some instances and has proven to be a worthwhile area of research. The objectives of this study are to create neural networks topologies to use Oregon State University's Multi-Application Small Light Water Reactor integrated test facility's data and evaluate its capability of predicting the systems behavior during various core power inputs and a loss of flow accident. This study uses data from multiple sensors, focusing primarily on the reactor pressure vessel and its internal components. As a result, the artificial neural networks are able to predict the behavior of the system with good accuracy in each scenario. Its ability to provide technical data can help decision makers to take actions more rapidly, identify safety issues, or provide an intelligent system with the potential of using pattern recognition for reactor accident identification and classification. Overall, the development and application of neural networks can be promising in the nuclear industry and any product processes that can benefit from utilizing a quick data analysis tool.

1. Introduction

There has been significant scientific interest in understanding and imitating natural and biological process, particularly neural biology. One of the first neural methodologies was first achieved with the creation of the perceptron capable of reproducing some of the Boolean operators (Rosenblatt, 1958). Later in the mid 80's there was a lot of effort to find a powerful synaptic modification rule that will allow an arbitrarily connected neural network to develop an internal structure that is appropriate for a particular task (Rumelhart et al., 1986); in other words, a self-organizing method that can be used in machines to learn a task without being explicitly programmed. The application of neural methods has been found useful in addressing problems that usually require the recognition of complex patterns or complex classification decisions. In the domain of computers science, there has been a rapid improvement of self-organizing methods along with advancements in data storage, parallel computing, and processing speeds, which have made possible for these methods to succeed in the development of new products and technologies. In the engineering domain, particularly in nuclear engineering, the application of machine learning methods, e.g. neural networks, utilizing full-scale facilities or real components data has been rather limited. In early applications researchers have used neural networks to assess the heat rate variation using the thermal performance data from the Tennessee Valley Authority Sequoyah nuclear power plant, where a small artificial neural network was used to determine the variables that affect the heat rate and thermal performance of the plant by looking at the partial derivative of the different input patterns (Zhichao and Uhrig, 1992). Others have developed monitoring systems based on auto-associative neural networks and their application as sensor calibration systems and sensor fault detection systems (Hines et al., 1996) using the High Flux Isotope Reactor operated at Oak Ridge National Laboratory and an

* Corresponding author at: School of Nuclear Science and Engineering, Oregon State University, 100 Radiation Center, Corvallis, OR 97330, United States.
E-mail address: gomezfem@oregonstate.edu (M. Gomez Fernandez).

http://dx.doi.org/10.1016/j.nucengdes.2017.08.020
Received 15 August 2016; Received in revised form 22 June 2017; Accepted 21 August 2017

experimental Breeder Reactor (Upadhyaya and Eryurek, 1992). During the mid-1990s a group of scientists explored the application of neural networks in the area of multiple-failures detection with the objective to develop an operator support system that can support operators during severe accidents in a nuclear power plant, referred as Computerized Accident Management System (Fantoni and Mazzola, 1996). In nuclear operations the inclusion of redundant, independent and diverse systems is necessary to ensure adequate defense-in-depth; however, the increase in systems lead to more complex human–machine interactions. Neural networks have also been trained with data from a simulator, and the results proved to be very satisfactory at modeling multiple sensor failures and component failure identification (Sirola and Talonen, 2012). Other areas outside of nuclear surveillance and diagnostics have also shown interest in the application of neural networks; for instance, in two-phase flow the use of neural methods as a method to predict two-phase mixture density (Lombardi and Mazzola, 1997) or flow regime identification (Tambouratzis and Pàzsit, 2010). More recently, researchers have applied advanced optimization algorithms for the prediction of the behavior of systems components such as a printed circuit heat exchanger (Ridluan et al., 2009; Wijayasekara et al., 2011), power peaking factor estimations (Montes et al., 2009), key safety parameter estimation (Mazrou, 2009) and functional failures of passive systems (Zio et al., 2010). The reduction in computational cost and the availability of data facilitates further the use of such methods where predicting more complex tasks is desired. In this research the application of neural methods using two transient events from a prototypic test facility is presented, where noise and uncertainty are present as an inherently natural phenomenon of a realistic problem.

2. Materials and methods

2.1. Multi-application small light water reactor

The Multi-Application Small Light Water Reactor (MASLWR) is an integral pressurized test facility developed by Idaho National Engineering and Environmental Laboratory, Oregon State University and NEXANT-Bechtel (Reyes et al., 2007), with the conceptual design shown in Fig. 1. The MASLWR module includes a self-contained vessel, steam generator and containment system that rely on natural circulation for its normal operation. The test facility is scaled at 1:3 length scale, 1:254 volume scale and 1:1 time scale, and it is designed for full pressure (11.4 MPa) and full temperature (590 K) prototype operation and is constructed of all stainless steel components (Reyes et al., 2007). The purpose of this facility is to study the behavior of a small light water reactor concept design that uses natural circulation for both steady-state and transient operation. The MASLWR concept was the predecessor to the NuScale small modular reactor design.

The data used in this study has been collected for the International

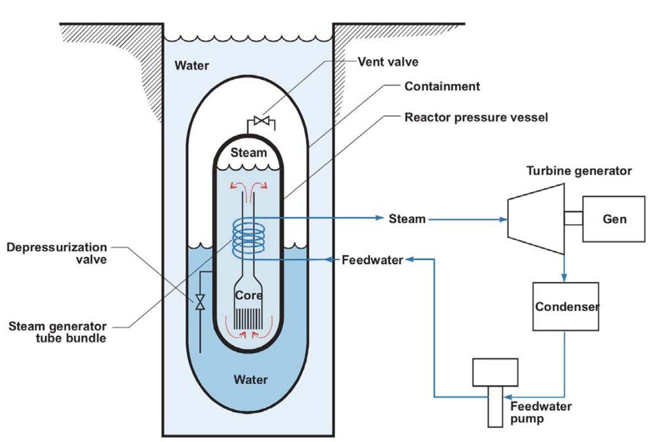


Fig. 1. MASLWR's conceptual design.

Atomic Energy Agency as an International Collaborative Standard Problem (ICSP). Two different data sets were used to train two different neural networks. The first, ICSP-3, characterize the steady-state (S.S.) natural circulation in the primary side during various core power inputs (Mai and Hu, 2011). The test procedure was to increase the power inputs of the heaters stepwise from 10% to 80% full power in the core by 10% increments and had a total duration of 6348 s (~1.76 h). The second, ICSP-2, characterizes the activation of safety systems of the MASLWR test facility, and the long-term cooling of the facility to determine the progression of a loss-of-feedwater transient (LOFW). For this test, first, the facility was brought to steady state at 75% core power, 8.62 MPa and the main feed water running in the steam generator, then, the main feed water was shut off, the core was set to decay power, and a blow-down procedure was conducted until the High Pressure Containment (HPC) and Reactor Pressure Vessel (RPV) were at equal pressures (Mai and Ascherl, 2011). This transient had a total duration of 16,483 s (~4.58 h).

2.2. Data

Data recorded from 58 different sensors was used as labeled data for the supervised learning process, with the purpose of capturing the behavior inside of prototype's RPV. Given that the data collected in the test facility inherently contains noise and uncertainty, the use of a neural network along with the backpropagation algorithm is suitable as this algorithm is robust to noise (Mitchel, 1997). However, the main challenge of the application of such method to this particular application is to find the suitable parameters that are to represent the problem, also known as feature selection. The selection of the features has been based on the sensors that are mainly controlled by the test facility's operator. Table 2 and Table 1 show the sensors used as inputs and outputs.

Moreover, given the different scales in the data, the entire set had to be normalized, using Eq. (1), to a [0,1] range to improve learning and avoid the saturation regions of the sigmoid function.

$$X' = (X_{max}-X_{min})\frac{X-X_{min}}{X_{max}-X_{min}} + X_{min} \tag{1}$$

The implementation of other normalizing techniques can also be used as long as it scales within the output range of the selected activation function.

Table 1
MASLWR instrumentation used as output parameters.

Sensor Label	Description
TF-[611-615]	Thermocouples Inside the Outer Coil Pipe of the Steam Generator Inlet
TF-[621-625]	Thermocouples Inside the Middle Coil Pipe of the Steam Generator Inlet
TF-[631-634]	Thermocouples Inside the Inner Coil Pipe of the Steam Generator Inlet
TF-[701-706]	Steam Generator Liquid Temperature
PT-602	Main Steam Pressure
FVM-602-T	Main Steam Temperature
FVM-602-P	Main Steam Pressure
FVM-602-M	Main Steam Pressure Volumetric Flow Rate
TH-[141-146]	Core Heater Rod Temperatures
TF-132	Primary Water Temperature inside Chimney below Steam Generator Coils
DP-101	Pressure Loss in the Core
DP-102	Pressure Loss between Core Tope and Cone
DP-103	Pressure Loss in the Riser cone
DP-104	Pressure Loss in the Chimney
DP-105	Pressure Loss across the Steam Generator
DP-106	Pressure Loss in the annulus below Steam Generator

Table 2
MASLWR instrumentation used as input parameters.

Sensor Label	Description
TF-[121-124]	Core Inlet Temperatures
KW-[101-102]	Power to the core heater rod bundles
TF-[101-106]	Center of Core Thermocouple Rod, six thermocouples spaced 6 apart, measuring water temperatures
TF-111	Primary Water Temperature at top of Chimney
KW-301	Power to Pressurizer
TF-501	Feed Water Temperature
FMM-501	Main Feedwater Volumetric Flow Rate
FCM-511	Feed Water Supply in the Steam Generator Outer Coil Mass Flow Rate
FCM-521	Feed Water Supply in the Steam Generator Middle Coil Mass Flow Rate
FCM-531	Feed Water Supply in the Steam Generator Inner Coil Mass Flow Rate
PT-511	Feed Water Pressure in the Steam Generator Outer Coil Mass Flow Rate
PT-521	Feed Water Pressure in the Steam Generator Middle Coil Mass Flow Rate
PT-531	Feed Water Pressure in the Steam Generator Inner Coil Mass Flow Rate

2.3. *Neural Networks* [1]

Firstly introduced in (Mcculloch and Pitts, 1943), neural networks are biologically-inspired techniques, which enables a computer to learn from observational data. McCulloch and Pitts stated that "*The nervous system is a net of neurons, each having a soma and an axion. Their adjunctions, or synapses, are always between the axon of the neuron and the soma of another. At any instant, a neuron has some threshold, which excitation must exceed to initiate an impulse. This is determined by the neuron, not by the excitation. From the point of excitation, the impulse is propagated to all parts of the neuron*" (Mcculloch and Pitts, 1943). To mimic a biological neuron, its artificial counterpart reproduces a similar functionality. As shown in Fig. 2, the network receives a series of data points or input vector ($x_1, \ldots, x_i$), whose contribution to the 'impulse' is determined by the synaptic weights associated with each neuron (w_i), and the activation function will use the weighted sum of input signals ($\sum w_i x_i$) to emit an output signal, whose value will determine if its 'impulse' is propagated to the rest of the network. This output will then become an input of the next layer and so on.

Neural networks are constructed using this principle to include multiple layers with many neurons to increase their representation capabilities as shown in Fig. 3. Consequently, when building neural networks, there are a few fundamental properties that need to be considered:

1. Activation function
2. Optimization algorithm
3. Structure or architecture of the network (known as model selection)

For the first property, the logistic or sigmoid function (Eq. (2)) is used as it is one of the most commonly used activation functions.

$$a(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

To describe what is known as the forward pass, the first the input vector is presented to the network and is then multiplied by the synaptic weights, as described previously. Let us defined it as:

$$c_j = w_j^T x + b \tag{3}$$

where b represent the bias term, w_j is the weight matrix of the j^{th} layer.

[1] If the reader is interested in further details see (Goodfellow et al., 2016; Bishop, 2006).

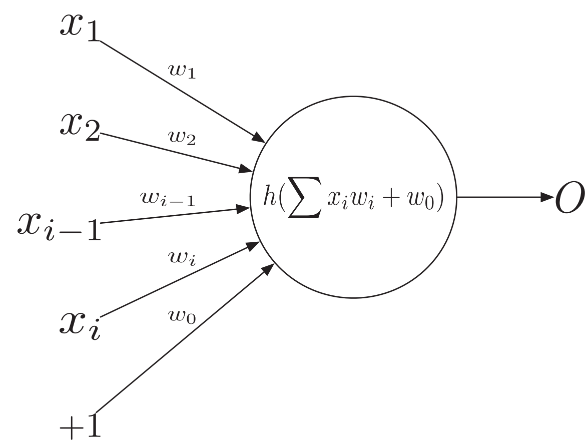


Fig. 2. Artificial neuron representation.

Then the activation function decides whether to propagate the value by applying the activation function

$$h(c_j) = a(c_j) \tag{4}$$

After the activation function is applied, the result will then become the new input (x) for Eq. (3) and the cycle repeats for as many j^{th} layers were chosen and the output layer is reached. Taking the following general forward pass formula:

$$f_p(x) = a(w_j^T a_{j-1}(w_{j-1}^T a_{j-2}(\ldots a_1(w_1^T x + b)) + b_{j-1}) + b_j \tag{5}$$

In the next couple section the selection of the structure and optimization algorithm is explained for the optimal design of a neural network.

2.3.1. *Backpropagation Algorithm*

The novel development and success of the backpropagation algorithm is greatly attributed to the ability to use an error function as a corrective factor for the connection strength (synaptic strength or weight), which allows the neurons to learn many layers of non-linear feature detection, such as recognizing handwritten zip codes (LeCun et al., 1989). Its primary objective is to find a learning rule that decides under which circumstances the hidden units should be active by a measure of the weights that when applied in a neural network the desired value and the actual output value are close (Rumelhart et al., 1986). This is achieved by minimizing an objective function, in this case, the mean square error (MSE) function,

$$E_n = \frac{1}{2} \sum_n (\hat{y}_j - y_j)^2 \tag{6}$$

and,

$$\hat{y}_j = h_j(w_j^T x + b_j) \tag{7}$$

where $\hat{y}_j$ is the predicted value for a particular input set and y_j is the desired output value. Then the gradient of this function with respect to the weights can be expressed as,

$$\frac{\partial E_n}{\partial w_j} = \frac{\partial E_n}{\partial h_j} \frac{\partial h_j}{\partial w_j} \tag{8}$$

Which indicates by what amount the error will increase or decrease if the value of w_j is to change by a small amount. After some mathematical manipulation, we obtain the following general backpropagaion formula

$$\nabla E = w_{j-1} \delta_j * h(c_{j-1}) * (1 - h(c_{j-1})) \tag{9}$$

where δ_j is the error from higher up units. Then, it can be used to form the gradient of the error function that is used for optimization.

For this study, a regularized mean square error was used to further

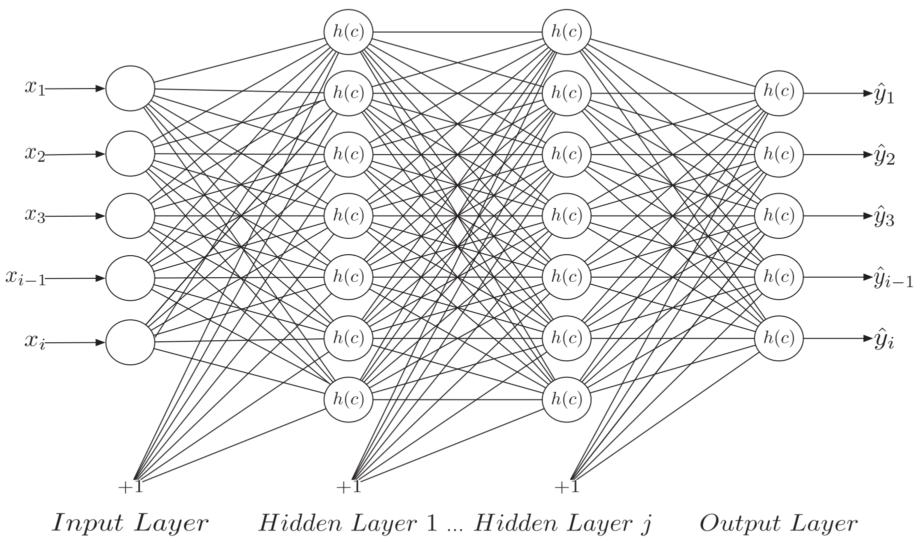


Fig. 3. Neural network representation.

control over-fitting

$$E_n = \frac{1}{2}\sum_i (\hat{y}_{ji} - y_{ji})^2 + \frac{\lambda}{2}w^2 \tag{10}$$

where λ is the penalization term or regularization coefficient that controls the complexity of the model by driving some of the weights to zero, or decreasing the importance or influence of a feature, also known as weight decay (Murphy, 2012).

2.3.2. Conjugate gradient method

The conjugate gradient method (CG) or the Fletcher-Powell method is a state-of-the-art algorithm for optimization problems as it is able to converge rapidly and handle large amounts of data (Navon and Legler, 1987). It has many advantages over the typical steepest descent, as it is a more robust and mathematical intense method that will converge as long as the function to be minimized is continuous and differentiable. The method starts similarly to the Cauchy's method or steepest descent in which minimization of the error gradient is desired by moving in the negative direction of the gradient:

$$d_o = -g_o \tag{11}$$

Then new values of w are calculated using the gradient direction by an amount of α_n

$$w_{n+1} = w_n + \alpha_n d_n \tag{12}$$

Where α_n can be calculated by a line search $min_\alpha F(\alpha d_n)$, and it is the optimal step size in the direction d_n. Once the new values of w are obtained the gradient is then updated by evaluating the gradient with respect to the new values of w

$$g_{n+1} = g(w_{n+1}) \tag{13}$$

Followed by the generation of a new direction

$$d_{n+1} = -g_{n+1} + \beta_z d_n \tag{14}$$

Where, $\beta_z = \frac{g_{z+1}^T g_{z+1}}{g_z^T g_z}$ in the Fletcher–Reeves algorithm; however, in this study a slight variation of the non-linear version of CG algorithm has been used called the Polak-Ribiere algorithm. This algorithm is similar to the Fletcher–Reeves algorithm, with the only difference being the way β_z is calculated (see (Navon and Legler, 1987))

$$\beta_z = \frac{g_{z+1}^T (g_{z+1} - g_z)}{g_z^T g_z} \tag{15}$$

Overall, the elegance of this algorithm is that in order to generate a new direction d, only three vectors need to be stored (the previous and current gradients and the previous direction) which makes efficient use of computer memory.

2.3.3. Structure

One of the principal issues regarding neural networks is the lack of an approach to determine the proper size of the neural network, where the usual approach is to try and keep the best (Russell and Norvig, 2010). Consequently, a K-fold cross validation (CV) technique was used to determine the optimal size of each of the hidden layers in each of the networks, such that each of the models' configuration is trained and tested 10 different times (K = 10), and the model that minimizes the average cost function of the test set is selected[2]. Fig. 4 shows the different neural network structures used and Table 3 [3] shows the configuration ranges in each structure, totaling a number of 28 models tested. Moreover, this ensures that the size of the neural network is optimized and computational power is efficiently used.

3. Results

3.1. Neural network optimization

For the supervised learning process the data has been divided in a 70–30 ratio, i.e. training set (~70%) and test set (~30%). Each of the different networks has been optimized to use the ideal size and the regularization parameter to control over-fitting. Fig. 5 shows an interesting pattern, where both neural networks have a preference towards structures 4b and 4d of medium size. Increasing the complexity also increases the MSE of the test set, making the model less accurate. Table 4 summarizes the results of the optimal size and regularization parameters for each of the networks.

3.2. Predictions

Despite the fact that neural networks are known to have a black box characteristic and lack of physical representation, the results achieved in this study show the ability of neural methods to successfully learn from the data regardless of the complexity of the data. To illustrate the results obtained, a number of sensors and its predictions were selected in each of the networks along with a linear correlation coefficient to show the linearity between the data and the neural network predictions. Figs. 6a, c, e, g, i, k, m, show the learned behavior under a LOFW

[2] This process has been parallelized

[3] The numbers shown in the table represent the initial number of units, number units incremented by each model, and final number of units

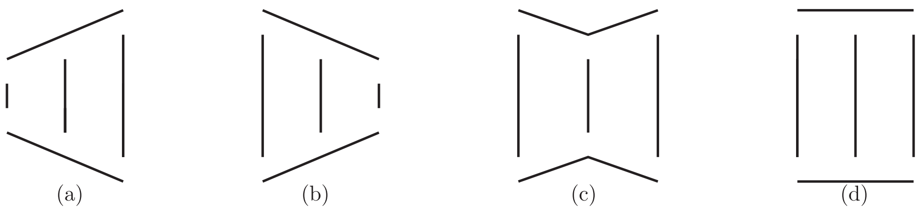
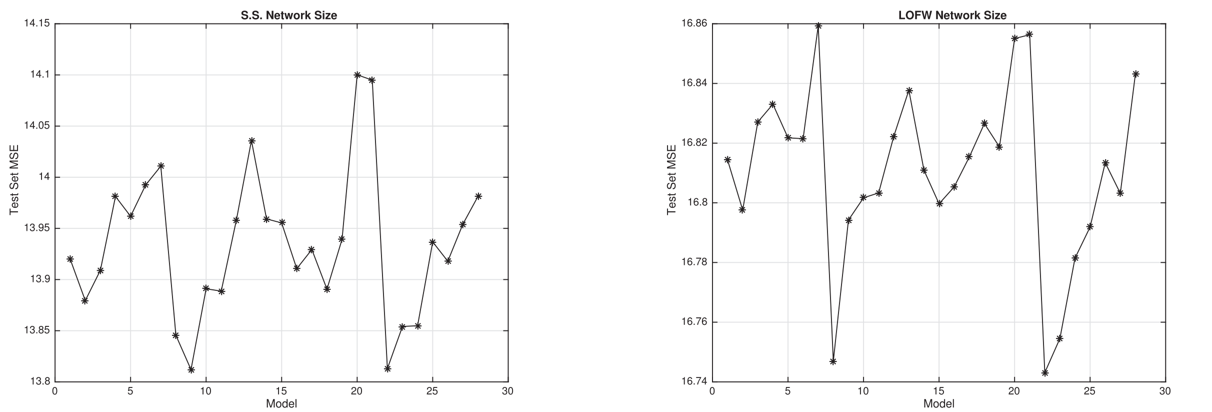


Fig. 4. Neural network structures.

Table 3
Ranges of number of units in each of the different structure presented in Fig. 4.

Structure	Layer 1	Layer 2	Layer 3
(a)	[20:10:80]	[30:10:90]	[40:10:100]
(b)	[40:10:100]	[30:10:90]	[20:10:80]
(c)	[20:10:80]	[10:5:40]	[20:10:80]
(d)	[20:10:80]	[20:10:80]	[20:10:80]



(a) Steady state behavior neural network

(b) Loss of feedwater behavior neural network

Fig. 5. Mean MSE as a function of structure.

Table 4
Neural network sizes and regularization parameter.

Network ID	Hidden Layer 1	Hidden Layer 2	Hidden Layer 3	λ
Network 1	30	30	30	5E-3
Network 2	40	30	20	5E-4

event. It can be observed that there is good agreement between the predicted data and the real data, as the network learned the average of most of the sensors data.

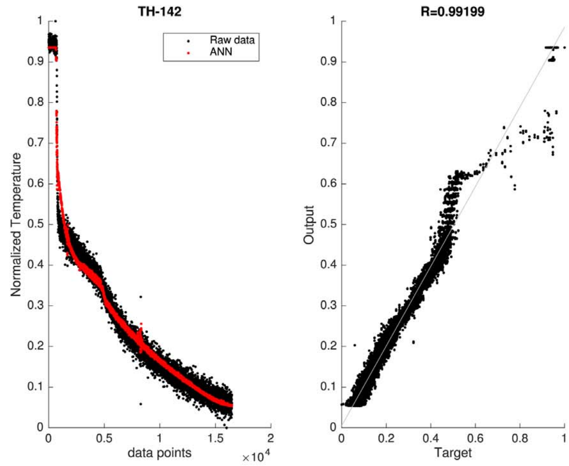
The temperature patterns in this data set are similar since the prototype is set to a decay mode and the neural network is able to fit the behaviors very well. It is worth pointing out that Figs. 6g and i show quite some noise and the network seems to identify and leans towards the greatest concentration of data (Fig. 6g), or learns an average (Fig. 6i) as the real data varies substantially. Similarly, Figs. 6b, d, f, h, j, l, n, show the learned steady-state behavior under a various core power. Again, good agreement is shown between the data and the prediction. In this data set, the event produces more challenging patterns and not all the sensors have similar patterns, in fact, they are quite different from one another. Again noise in the data is expected, but it can also affect the network's perdition capability. For instance, in Fig. 6h the unnormalized differential pressure sensor fluctuates between 501.16 Pa and 503.28 Pa and the network is not able to fully adapt to the sensors behavior; nonetheless, the network does lean towards the greatest concentration of data, identifying a linear pattern for this sensor.

4. Discussion

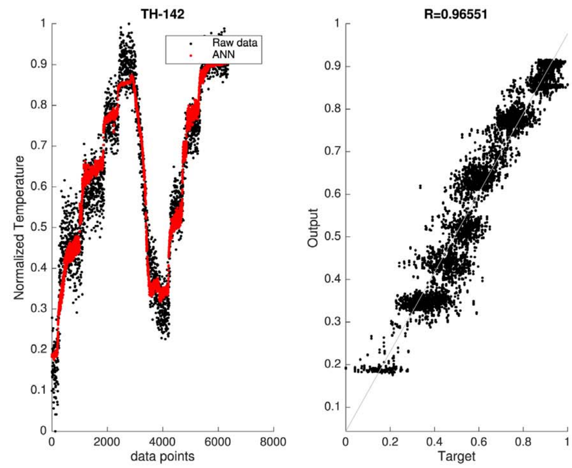
In the study of complex systems there are a wide variety of different properties that determine the behavior of the overall system and researchers usually pursue the use of physical representation to explain the physical phenomena. The test facility used here clearly shows the difficulty of analyzing a system as a whole since some of the data show a wide variety of patterns that no model can fully adapt. Neural networks can mimic most highly non-linear relations, making this method popular among researchers. However, their success depends on the characteristics of the chosen model, which vary based on trial-and-error, in addition to other limitations (Guo et al., 2010), such as the availability, quantity and quality of data that can be obtained from test facilities or share with other institutions. Data is the most important element in the application of machine learning, which can represent an issue in the nuclear industry as most the data is restricted. Parallel computing has also significantly accelerate parameter tunning, i.e. regularization and structure, and continues to improve with the use of GPU; nonetheless, it is still a challenge in neural networks as there is no given technique to quickly define these parameters that best suits the problem. Overall, the expressiveness of neural networks has produced satisfactory results, as many in the literature, for proof-of-concept in this application. It is highly encouraged in this research to further investigate this application in the test facility to validate the functionality, speed and accuracy of the predictions using additional transients, with the ultimate goal of integrating a systems as an operational enhancement tool to support decision-making.

5. Conclusion

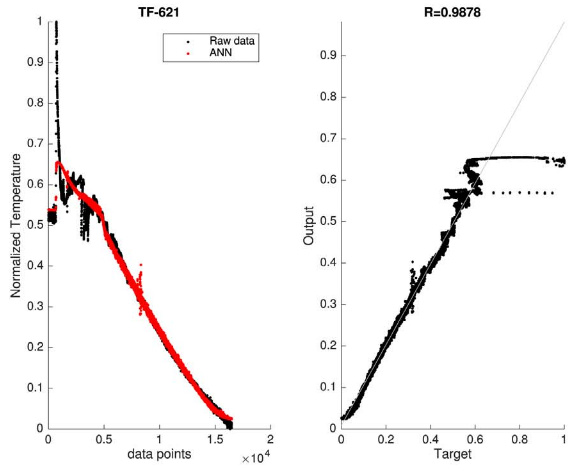
The application of machine learning and other artificial intelligence techniques have been considered for many day-to-day applications in different industries. The purpose this study was to explore the application of machine learning methods, particularly neural networks, in the nuclear engineering domain for systems behavior predictions using the MASLWR test facility. The prototypical test facility was designed to



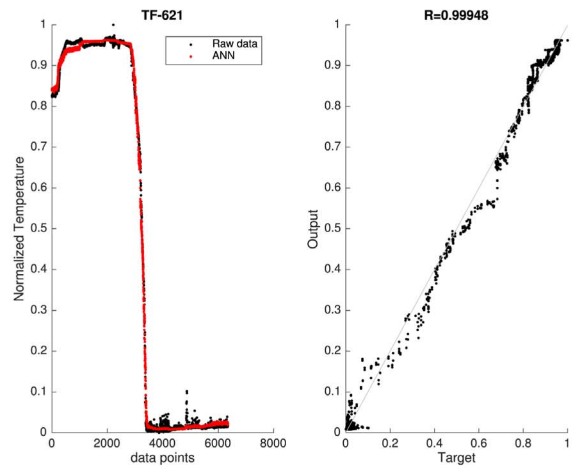
(a) LOFW-Core Heater Temperature



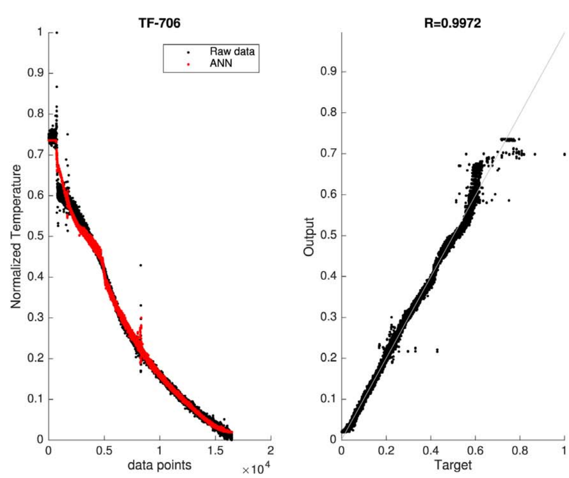
(b) S.S.-Core Heater Temperature



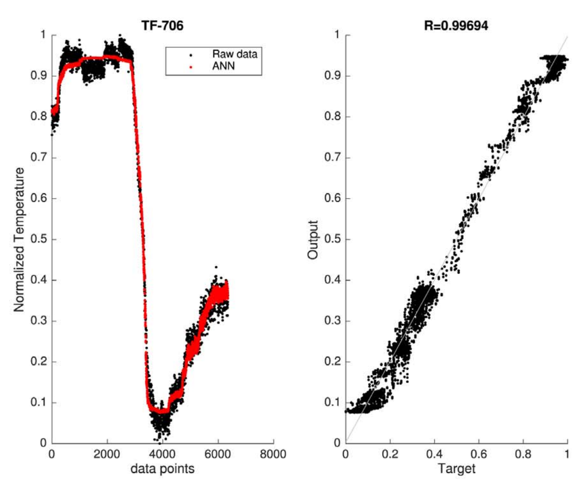
(c) LOFW-Middle Coil Steam Generator Temperature



(d) S.S.-Middle Coil Steam Generator Temperature



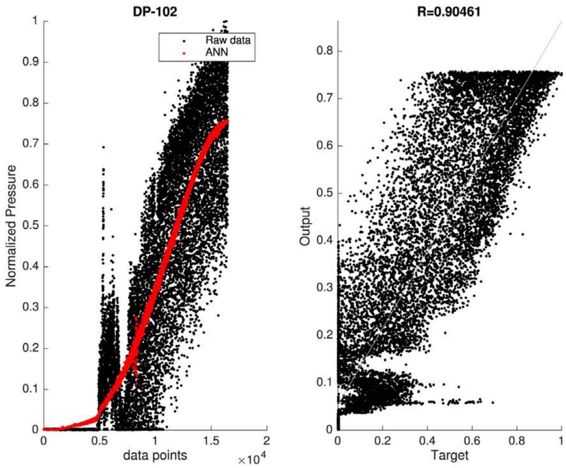
(e) LOFW-Steam Generator Liquid Temperature



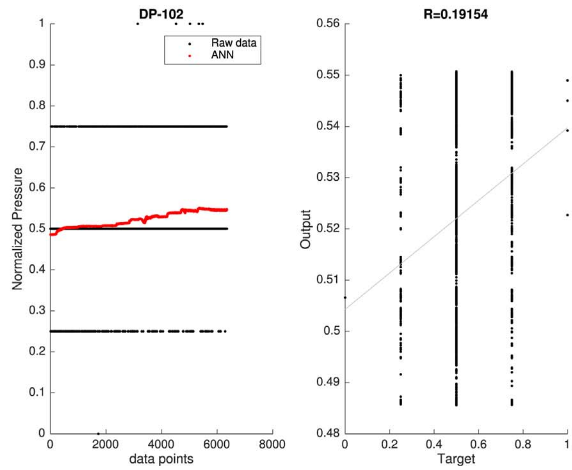
(f) S.S.-Steam Generator Liquid Temperature

Fig. 6. Neural networks results.

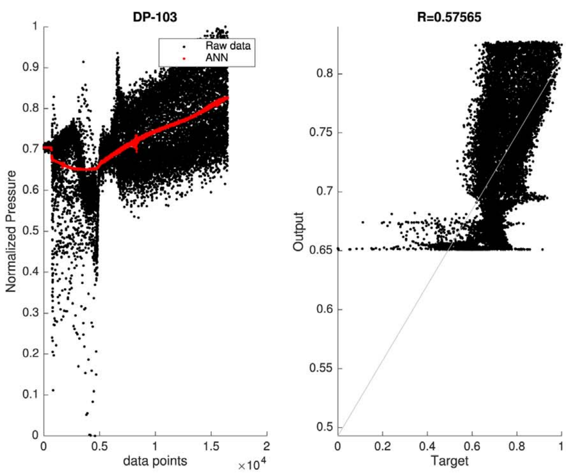
assess the operation of an integrated small modular nuclear reactor at full pressure and temperature, and also, to assess the passive safety systems under different events. Despite the lack of physical representation in neural networks, the results obtained show their capability to use multiple sensors data to predict the behavior of the facility given various core powers and during a loss-of-feedwater event. Good agreement has been shown between the prediction and the raw data obtained from the facility without postprocessing of the data. Moreover, in cases where there was a lot of variance in the data, the neural network leaned toward greater concentration of data which it



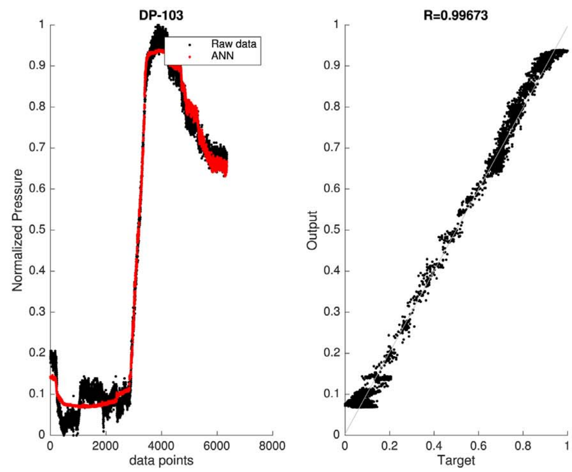
(g) LOFW-Pressure Drop between Core and Cone



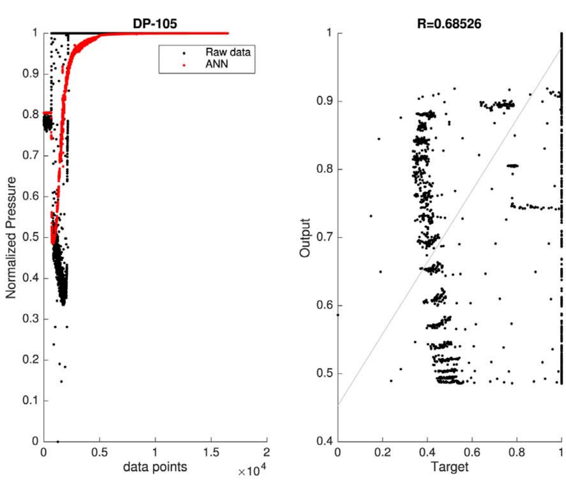
(h) S.S.-Pressure Drop between Core and Cone



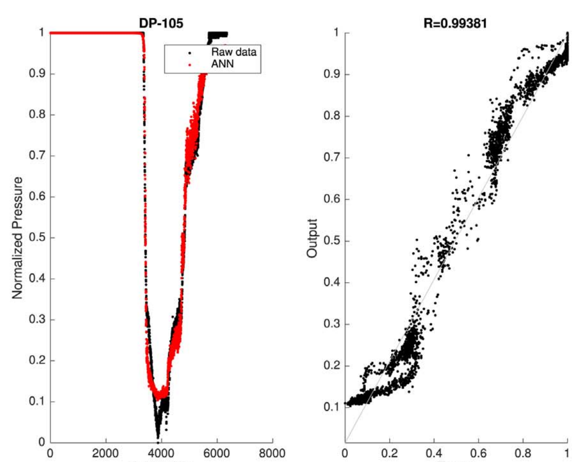
(i) LOFW-Pressure Drop in Riser Cone



(j) S.S.-Pressure Drop in Riser Cone



(k) LOFW-Pressure Drop across the Steam Generator



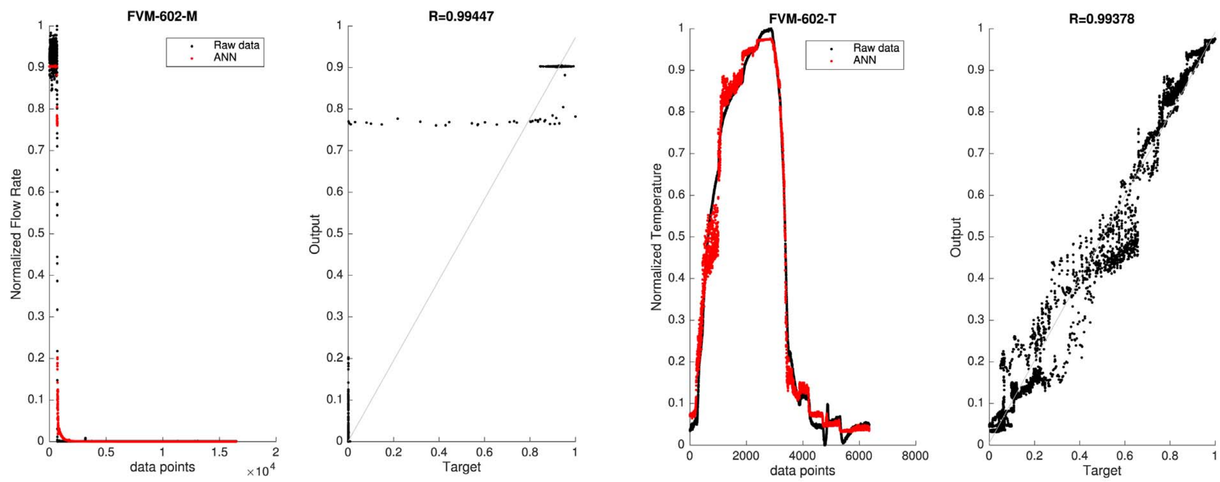
(l) S.S.-Pressure Drop across the Steam Generator

Fig. 6. (*continued*)

considered as the expected value. However, there are sensors where prediction is more difficult and can be further investigated. Though there is still a need to further explore the use of neural methods in the nuclear engineering domain, the neural networks have successfully captured the behavior of most sensors inside the prototype.

Acknowledgements

The first author will like to extend his appreciation to the MASLWR team at Oregon State University for their extensive work in collecting the data and the guidance and support from NuScale Power‘s lead



(m) LOFW-Steam Volumetric Flow Rate

(n) S.S.-Steam Volumetric Flow Rate

Fig. 6. (*continued*)

engineers. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Bishop, C., 2006. Bishop – Pattern Recognition and Machine Learning. Ch. 5.

Fantoni, P.F., Mazzola, A., 1996. Multiple failure signal validation in nuclear power plants using artificial neural networks. Nucl. Technol. 113 (3), 368–374.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press. Ch. 6,http://www.deeplearningbook.org.

Guo, Y., Gong, C., Zeng, H.-Y., 2010. The application of Artificial Neural Network in nuclear energy. Machine Learning and Cybernetics (ICMLC), 2010 International Conference on 3 (July), pp. 11–14.

Hines, W., Wrest, D., Uhrig, R., 1996. PLANT WIDE SENSOR CALIBRATION MONITORING. In: IEEE International Symposium on Control. pp. 0–5.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L. D., 1989. Backpropagation Applied to Handwritten Zip Code Recognition.

Lombardi, C., Mazzola, A., 1997. Prediction of two-phase mixture density using artificial neural networks. Anna 24 (17), 1373–1387.

Mai, A., Ascherl, G., 2011. OSU-MASLWR-QLR-SP2-R0. Tech. Rep. Revision 0, Oregon State University.

Mai, A., Hu, L., 2011. OSU-MASLWR-QLR-SP3. Tech. Rep. Revision 0, Oregon State University.

Mazrou, H., 2009. Performance improvement of artificial neural networks designed for safety key parameters prediction in nuclear research reactors. Nucl. Eng. Des. 239 (10), 1901–1910.

Mculloch, W., Pitts, W., 1943. A logical calculus of ideas immanent in nervous activity. Bull. Math. Biophys. 5, 127–147.

Mitchel, T., 1997. Machine Learning. McGraw-Hill.

Montes, J.L., François, J.L., Ortiz, J.J., Martín-del Campo, C., Perusquía, R., 2009. Local power peaking factor estimation in nuclear fuel by artificial neural networks. Ann. Nucl. Energy 36 (1), 121–130. http://dx.doi.org/10.1016/j.anucene.2008.09.011.

Navon, I.M., Legler, D., 1987. Conjugate gradient methods for large-scale minimization in meteorology. Mon. Weather Rev. 115, 1479–1502.

Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. The MIT Press, Massachusetts http://link.springer.com/chapter/10.1007/978-94-011-3532-02.

Reyes, J.N., Groome, J., Woods, B.G., Young, E., Abel, K., Yao, Y., Yoo, Y.J., 2007. Testing of the multi-application small light water reactor (MASLWR) passive safety systems. Nucl. Eng. Des. 237 (18), 1999–2005.

Ridluan, A., Manic, M., Tokuhiro, A., 2009. EBaLM-THP – A neural network thermo-hydraulic prediction model of advanced nuclear system components. Nucl. Eng. Des. 239 (2), 308–319.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the Brain. Psychol. Rev. 65 (6), 386–408 http://psycnet.apa.org/journals/rev/65/6/386.pdf\npapers://c53d1644-cd41-40df-912d-ee195b4a4c2b/Paper/p15420.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323 (6088), 533–536.

Russell, S., Norvig, P., 2010. Artificial Intelligence: A Modern Approach, third ed. Pearson.

Sirola, M., Talonen, J., 2012. Combining neural methods and knowledge-based methods in accident management. Adv. Artif. Neural Syst. 2012, 1–6.

Tambouratzis, T., Pàzsit, I., 2010. A general regression artificial neural network for two-phase flow regime identification. Ann. Nucl. Energy 37 (5), 672–680. http://dx.doi.org/10.1016/j.anucene.2010.02.004.

Upadhyaya, B., Eryurek, E., 1992. Application of neural networks for sensor validation and plant monitoring. Nucl. Technol. 97, 170–176.

Wijayasekara, D., Manic, M., Sabharwall, P., Utgikar, V., 2011. Optimal artificial neural network architecture selection for performance prediction of compact heat exchanger with the EBaLM-OTR technique. Nucl. Eng. Des. 241 (7), 2549–2557. http://dx.doi.org/10.1016/j.nucengdes.2011.04.045.

Zhichao, G., Uhrig, R., 1992. Use of artificial neural networks to analyze nuclear power plant performance. Nucl. Technol. 99, 36–42.

Zio, E., Apostolakis, G.E., Pedroni, N., 2010. Quantitative functional failure analysis of a thermal-hydraulic passive system by means of bootstrapped Artificial Neural Networks. Ann. Nucl. Energy 37 (5), 639–649. http://dx.doi.org/10.1016/j.anucene.2010.02.012.