IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 19 December 2024, accepted 31 December 2024, date of publication 13 January 2025, date of current version 16 January 2025.

*Digital Object Identifier 10.1109/ACCESS.2025.3526438*

RESEARCH ARTICLE

# Temporal Forecasting of Distributed Temperature Sensing in a Thermal Hydraulic System With Machine Learning and Statistical Models

**STELLA PANTOPOULOU[1,2], (Member, IEEE), MATTHEW WEATHERED[1], DARIUS LISOWSKI[1], LEFTERI H. TSOUKALAS[2], AND ALEXANDER HEIFETZ[1], (Senior Member, IEEE)**

[1]Nuclear Science and Engineering Division, Argonne National Laboratory, Lemont, IL 60439, USA
[2]School of Nuclear Engineering, Purdue University, West Lafayette, IN 47906, USA

Corresponding author: Alexander Heifetz (aheifetz@anl.gov)

This work was supported by U.S. Department of Energy, Advanced Research Projects Agency—Energy (ARPA-E) under Contract DE-AC02-06CH11357.

**ABSTRACT** We benchmark performance of long-short term memory (LSTM) network machine learning model and autoregressive integrated moving average (ARIMA) statistical model in temporal forecasting of distributed temperature sensing (DTS). Data in this study consists of fluid temperature transient measured with two co-located Rayleigh scattering fiber optic sensors (FOS) in a forced convection mixing zone of a thermal tee. We treat each gauge of a FOS as an independent temperature sensor. We first study prediction of DTS time series using Vanilla LSTM and ARIMA models trained on prior history of the same FOS that is used for testing. The results yield maximum absolute percentage error (MaxAPE) and root mean squared percentage error (RMSPE) of 1.58% and 0.06% for ARIMA, and 3.14% and 0.44% for LSTM, respectively. Next, we investigate zero-shot forecasting (ZSF) with LSTM and ARIMA trained on history of the co-located FOS only, which is advantageous when limited training data is available. The ZSF MaxAPE and RMSPE values for ARIMA are comparable to those of the Vanilla use case, while the error values for LSTM increase. We show that in ZSF, performance of LSTM network can be improved by training on most correlated gauges between the two FOS, which are identified by calculating the Pearson correlation coefficient. The improved ZSF MaxAPE and RMSPE for LSTM are 4.4% and 0.33%, respectively. Performance of ZSF LSTM can be further enhanced through transfer learning (TL), where LSTM is re-trained on a subset of the FOS that is the target of forecasting. We show that LSTM pre-trained on correlated dataset and re-trained on 30% of testing target dataset achieves MaxAPE and RMSPE values of 2.32% and 0.28%, respectively.

**INDEX TERMS** Distributed temperature sensing, Rayleigh scattering fiber optic sensors, thermal hydraulic temperature sensing, machine learning, statistical methods, ARIMA, LSTM, zero-shot forecasting, transfer learning.

## I. INTRODUCTION

Monitoring the state of a nuclear reactor coolant fluid involves measurements of process variables, such as temperature, pressure, and flow rate. The most common type of measurement in a reactor is temperature sensing, which provides information about reactor power. Temperature sensing in existing reactors is performed with nuclear-grade thermocouple arrays or resistance temperature detectors [1], [2], [3]. Installation of such sensors requires modification of the fluid vessel or piping pressure boundary to accommodate a thermowell or similar sensor insertion mechanism. This weakens the mechanical resilience of the structure because of increasing risk of crack development around the thermowell welds. In addition, each sensor unit has its own set of electric cables, which contribute to the instrumentation clutter. Cable damage in the high temperature and ionizing radiation environment can leave the sensor inoperable.

The associate editor coordinating the review of this manuscript and approving it for publication was Sukhdev Roy.

An alternative approach of distributed temperature sensing (DTS), involving measurements with fiber optic sensors (FOS) immersed in a fluid has emerged recently [4], [5], [6]. DTS offers the capability to obtain temperature field in the fluid with spatial resolution higher than that of conventional thermocouple arrays and with a smaller instrumentation footprint. FOS does not involve the use of electric cables, and requires fewer penetrations through fluid vessel pressure boundaries, as compared to the installation requirements of thermocouple arrays. DTS has been recently demonstrated in different reactor fluids, including water [7], [8], gas [9], [10], liquid sodium [11], [12] and molten salt [13], [14].

The physics of DTS involves light scattering from refractive index fluctuations in the optical fiber, which are modulated through interaction with the surrounding medium. DTS for nuclear energy applications have been demonstrated using Rayleigh [15], [16], Raman [17], [18], [19], [20], and Brillouin [21] scattering processes in optical fibers. Rayleigh scattering involves elastic (frequency-conserving) interactions of light with subwavelength intrinsic manufacturing defects (e.g., nanoscale air bubbles in silica glass), the density of which is modulated locally by thermal strain. Raman scattering involves inelastic (frequency up-conversion or down-conversion) interaction of light with phonons of vibrational and rotational modes of atomic bonds. Brillouin scattering involves inelastic light scattering from macroscopic-scale acoustic phonons in silica glass.

DTS based on Rayleigh scattering is advantageous because of higher sensitivity of Rayleigh scattering compared to Raman and Brillouin scattering, and because of relatively simpler instrumentation consisting of a single monochromatic laser and inexpensive telecommunication wavelength fibers. The spatially resolved strain can be measured either with optical time domain reflectometry (OTDR) [22] or frequency domain reflectometry (OFDR) [23]. Using OFDR allows to obtain higher spatial resolution of sensing, with sub-millimeter spatial resolution for fiber lengths on the order of a few meters [24]. Frequency shifts measured with OFDR can be converted to temperature through power law correlation [12], [16], [23].

Due to degradation of silica glass in the high temperature and radiation environment of a nuclear reactor, Rayleigh scattering FOS are susceptible to failure [25], [26], [27], [28], [29], [30]. Early detection of failures can be accomplished through automation of sensor response monitoring [31]. Using machine learning (ML) algorithms for analysis of distributed sensing with optical fibers has been reported in recent literatures [32], [33], [34], [35], and [36], including convolutional neural networks (CNN) [37], autoencoders [35], generative adversarial networks (GAN) [32], long short-term memory (LSTM) networks [38], [39], and K-nearest neighbor (KNN) [35]. In related research efforts, surrogate ML models of distributed physical field variables based on deep learning (DL) approaches have been developed to expedite computer simulations [40], [41].

Performance of data-driven ML models generally depends on the available training data. Understanding relative merits of the emerging ML models can be enhanced through comparison of their performance with those of statistical models, which are another type of data-driven models. In this paper, we benchmark performance of a long-short term memory (LSTM) network ML model and autoregressive integrated moving average (ARIMA) statistical model in one-step ahead forecasting of DTS. LSTM and ARIMA have been shown to be highly efficient methods in time series analysis [39], [42], [43], [44]. ARIMA, in general, needs smaller volume of training data compared to LSTM, but ARIMA requires training and testing data pre-processing. While performance of LSTM typically depends on the amount of the training data [45], it does not require any data pre-processing. Performance of LSTM and ARIMA has been benchmarked in studies involving data for financial [46], [47], wind forecasting [48], and structural mechanics [49] applications. However, to the best of our knowledge, performance of LSTM and ARIMA models in forecasting DTS time series in a thermal hydraulic system has not been studied systematically.

The data in this study consists of fluid temperature transient measured with two co-located Rayleigh scattering FOS (FOS1, FOS2) in the mixing zone of a thermal tee. DTS data consists of a 2D matrix of temperature values, where one dimension is measurement time, and the other dimension is length along the fiber with spatial resolution given by the fiber gauge pitch. We treat each gauge of a FOS as an independent temperature sensor. Performance of the ARIMA and LSTM forecasting models is evaluated with root mean squared error (RMSE) and maximum absolute error (MaxAE), as well as with root mean squared percentage error (RMSPE) and maximum absolute percentage error (MaxAPE).

We first consider the Vanilla use case, where LSTM and ARIMA are trained on the prior history of the same FOS that is used for forecasting. Statistical tests indicate that the data for both FOS is non-stationary, and thus differencing is required to make the data stationary for forecasting with ARIMA. Tests on the same dataset demonstrate that ARIMA has lower errors than LSTM.

Next, we consider the zero-shot forecasting (ZSF) use case, where both the LSTM and ARIMA are trained on the history of the co-located FOS only. ZSF potentially allows to develop a forecasting model when limited data is available for a given domain. Performance of ARIMA is similar for the Vanilla and ZSF use cases, and better than that of the LSTM network. Performance of the LSTM model can be enhanced by training on most correlated gauges between the two FOS, which are identified by calculating the Pearson correlation coefficient. However, the MaxAPE and RMSPE in LSTM forecasting remain larger than those of ARIMA.

To improve performance of the LSTM developed for the ZSF use case, we consider a transfer learning (TL) use case. In the TL approach, the LSTM network pre-trained on co-located FOS is re-trained on the data from the same FOS

as the one used in testing. When pre-training on correlated gauges in the TL use case, LSTM achieves MaxAPE comparable to that of ARIMA. However, RMSPE is still smaller for ARIMA.

Table 1 provides a summary representation of the structure of this paper by listing relative merits or utility, and training and testing domains for LSTM and ARIMA models in Vanilla, ZSF and TL use cases. Section II describes the LSTM and ARIMA algorithms, as well as the ZSF and TL methods. Section III includes description of the measured DTS data and characterization of the data based on correlations between FOS gauges. Section IV contains details on DTS data partitioning, as well as on training of ARIMA and LSTM models. Section V discusses the Vanilla use case, including data portioning and results of LSTM and ARIMA performance. Section VI discusses the ZSF use case and the case where data partitioning is modified according to correlations between fiber gauges. Section VII presents the TL use case, where pre-trained LSTM from ZSF use case is re-trained on same-sensor data. Section VIII contains the conclusions.

**TABLE 1. Summary representation of vanilla, ZSF, and TL use cases.**

| Use Case | Utility | Training | Re-training | Testing |
|---|---|---|---|---|
| Vanilla | Basic forecasting approach | FOS1 | - | FOS1 |
| ZSF | Limited training data available | FOS2 | - | FOS1 |
| TL | Enhances performance of ZSF | FOS2 | FOS1 | FOS1 |

## II. FORECASTING MODELS

### A. LONG SHORT-TERM MEMORY (LSTM) NETWORKS

LSTM networks are a special type of a recurrent neural network (RNN) architecture, which is efficient in processing sequential data with long-term dependencies. LSTM networks address the vanishing and exploding gradient problems often encountered in analysis with traditional RNN [39]. An LSTM cell consists of input, forget, and output gates, which control the flow of information within the network. The input gate determines how much new information should be stored in the memory cell, while the forget gate decides what information to discard from the cell. The output gate regulates the information flow from the memory cell to the next layer or output. An LSTM cell calculates a hidden state $H_t$ and a current state $C_t$:

$$H_t = \sigma\left(x_t U_o + H_{t-1} W_o\right) \cdot \tanh\left(C_t\right) \tag{1}$$

$$C_t = \sigma\left(x_t U_f + H_{t-1} W_f\right) \cdot C_{t-1} + \sigma\left(x_t U_i + H_{t-1} W_i\right) \cdot \tanh\left(x_t U_c + H_{t-1} W_c\right) \tag{2}$$

Here $x_t$ is the cell input, $H_{t-1}$ and $C_{t-1}$ are the hidden and current states from the previous time step, $\sigma$ is the sigmoid activation function, $U$ and $W$ are the weight matrices. In the standard LSTM network, sigmoid is used as the gating function and the tanh is used as the output activation function [50]. The subscripts *f, i, o,* and *c* indicate forget, input, output gates and the current state, respectively.

The elements in the $U$ and $W$ weight matrices are generated during network training by solving the optimization problem of minimizing the value of the loss function. In times series forecasting, typical loss function is the mean squared error (MSE):

$$MSE = \frac{1}{n}\left\|X - \hat{X}\right\|^2 \tag{3}$$

Where $X = \{x_i\}$ is the set of $n$ observations, $\hat{X} = \{\hat{x}_i\}$ is the set of predictions, and $\|X\|$ is the norm of a vector.

Performance of the network during a test can be evaluated with several error metrics. In this work we use average and largest error RMSE and MaxAE metrics defined in Equations (4) and (5), respectively. In addition, we use average and largest percentage errors RMSPE and MaxAPE defined in Equations (6) and (7), respectively.

$$RMSE = \sqrt{MSE} \tag{4}$$

$$MaxAE = max\left|X - \hat{X}\right| \tag{5}$$

$$RMSPE = \sqrt{\frac{1}{n}\left\|\frac{X - \hat{X}}{X}\right\|^2} \cdot 100\% \tag{6}$$

$$MaxAPE = max\left|\frac{X - \hat{X}}{X}\right| \cdot 100\% \tag{7}$$

### B. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

ARIMA is a statistical model frequently used for forecasting time series. In the autoregressive (AR) model, the current value of a time series $x_t$ depends on $k$ time lags:

$$x_t = \alpha + \beta_1 x_{t-1} + \cdots + \beta_k x_{t-k} \tag{8}$$

where $\alpha$ is the intercept estimated by the model, and $\beta_i$ are the coefficients for each time series lag $x_{t-i}$, $i = 1, 2, \ldots, k$. In the moving average (MA) model, the current value of a time series $x_t$ depends on the forecast error time lags:

$$x_t = \alpha + \epsilon_t + c_1 \epsilon_{t-1} + \cdots + c_k \epsilon_{t-k} \tag{9}$$

where $\epsilon_{t-i}$ are the error time lags, and $c_i$ are the respective coefficients.

The ARIMA$(p, d, q)$ model contains three integer-valued parameters to be selected by the user [50]. Implementation of the ARIMA requires that the time series are stationary, i.e., the statistical properties of the time series do not change with time. Stationarity of the FOS data was evaluated with the Augmented Dickey-Fuller (ADF) statistical test. Outputs of the ADF test include the p-value, the test statistic, and the critical value. For example, for a confidence level of 95%, p-value $< 0.05$ indicates that a time series is stationary. Similarly, a test statistic value smaller than the critical value, which depends on the number of points in the time series, indicates stationarity of the data. The parameter $d$ corresponding to the integrated (I) part is the number of times

needed to difference the time series so that they become stationary [51].

Once the time series are differenced, the user calculates the values of the parameters $p$ and $q$ using the stationary data. The parameter $p$ corresponding to the AR part is the number of significant lags in the partial autocorrelation function (PACF). The $k^{th}$ order lag is calculated as:

$$\frac{Cov\left(x_t, x_{t-k} \mid x_{t-1}, \ldots, x_{t-k+1}\right)}{\sqrt{Var\left(x_t | x_{t-1}, \ldots, x_{t-k+1}\right) Var\left(x_{t-k} \mid x_{t-1}, \ldots, x_{t-k+1}\right)}} \tag{10}$$

where $Cov$ and $Var$ represent the covariance and variance, respectively. The value of the parameter $p$ is determined by examining the shape of the PACF in Equation (10).

The parameter $q$ corresponding to the MA part is the number of significant lags in the autocorrelation function (ACF) $\hat{\gamma}_k$

$$\hat{\gamma}_k = \frac{\sum_{t=k+1}^{n} \left(x_t - \bar{x}\right)\left(x_{t-k} - \bar{x}\right)}{\sum_{t=1}^{n} \left(x_t - \bar{x}\right)^2} \tag{11}$$

Where $\bar{x}$ is the mean value of the time series, and $n$ is the total number of terms in the time series. The value of the parameter $q$ is determined by examining the shape of the ACF in Equation (11).

### C. ZERO-SHOT FORECASTING AND TRANSFER LEARNING

Zero-shot forecasting (ZSF) and transfer learning (TL) refer to approaches involving training a data-driven model on source domain $S$ to make predictions on target domain $T$ [52], [53], [54], [55], [56], [57], [58], [59], [60]. Let $X_S$ be the set of the past $k$ observations, and $Y_S$ be the set of current observations at time $t$ in $S$:

$$X_S = \{x_{t-1}, \ldots, x_{t-k}\}_S \tag{12a}$$
$$Y_S = \{x_t\}_S \tag{12b}$$

Let $X_T$ be the set of the past $k$ observations and $Y_T$ be the set of current observations at time $t$ in $T$:

$$X_T = \{x_{t-1}, \ldots, x_{t-k}\}_T \tag{13a}$$
$$Y_T = \{x_t\}_T \tag{13b}$$

Training a forecasting model involves learning the functions $f_S$ and $f_T$

$$f_S : X_S \rightarrow Y_S \tag{14a}$$
$$f_T : X_T \rightarrow Y_T \tag{14b}$$

In Vanilla forecasting, the prediction $\hat{Y}_T$ for the observed values $Y_T$ can be written as

$$\hat{Y}_T = f_T\left(X_T\right) \tag{15}$$

In the zero-shot forecasting (ZSF) approach, the prediction $\hat{Y}_T$ for the observed values $Y_T$ can be written as

$$\hat{Y}_T = f_S\left(X_T\right) \tag{16}$$

The approach of TL involves re-training on $T$ a model that was pre-trained on $S$. The TL prediction $\hat{Y}_T$ can be expressed as

$$\hat{Y}_T = f_T\left(f_S\left(X_T; \theta_S\right), \theta_T\right) \tag{17}$$

where $\theta_S$ and $\theta_T$ are parameters learned from pre-training in $S$, and re-training in $T$, respectively.

TL is referred to as homogeneous if there is semantic and quantitative similarity in the feature space in source and target domains, and heterogeneous otherwise [61], [62]. Quantitative similarity between the source and target domains can be estimated, for example, by calculating the Pearson correlation coefficient (PCC) [63]. For 1D signals $x$ and $y$ of length $n$ and respective mean values $\bar{x}$ and $\bar{y}$, the PCC is calculated as:

$$\rho = \frac{\sum_i \left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_i \left(x_i - \bar{x}\right)^2 \sum_i \left(y_i - \bar{y}\right)^2}} \tag{18}$$

The correlation coefficients take on values in the range from $\rho = 1$ (identical signals) to $\rho = -1$ (anti-correlated), where $\rho = 0$ means that the signals are uncorrelated.

## III. DTS DATA COLLECTION AND CHARACTERIZATION

### A. DTS MEASUREMENTS IN A FLOW LOOP WITH A THERMAL MIXING TEE

A water flow loop with a forced convection mixing thermal tee was constructed for experimental DTS tests with Rayleigh backscattering FOS. The loop was constructed with polycarbonate pipes and filled with water at ambient pressure. Figure 1 shows a schematic diagram of the flow loop. A thermal mixing tee is a common experimental setup for proof-of-principle studies of thermal hydraulic phenomena in advanced reactors, such as thermal striping in a sodium fast reactor [64]. Water has been shown to be a good surrogate for sodium when performing thermal hydraulic experiments, possessing very similar density and viscosity to within an order of magnitude [65], [66]. Therefore, using a water flow loop provides a reasonable low-cost experimental data collection alternative for a proof-of-concept study.

Water was circulated against the gravity (clockwise in the viewing plane) with a 1.5 hp variable speed pump. The flow split into a cold and a hot leg, where a variable power heater with maximum power of 4 kW heated the water to generate a thermal transient. Flows from the hot and cold legs recombine in the mixing tee, as shown in Fig. 1. The inner diameters of the pipes in the main loop, cold leg and hot leg are 1.5 in (3.81 cm) and 0.75 in (1.9 cm), respectively. The mixing zone is indicated with a circle. Two identical Rayleigh backscattering fiber optic sensors, labeled FOS1 and FOS2, were installed in the loop [65]. Fiber optic cables without cladding were inserted into metallic thimbles, the outer surfaces of which were in contact with the fluid. As shown in Fig. 1, FOS2 was installed closer to the inlet from the hot leg.

During the experiment, the heater was set to the maximum power setting to raise the temperature of the water to 57.7°C. The flow velocities in the main loop, cold leg,

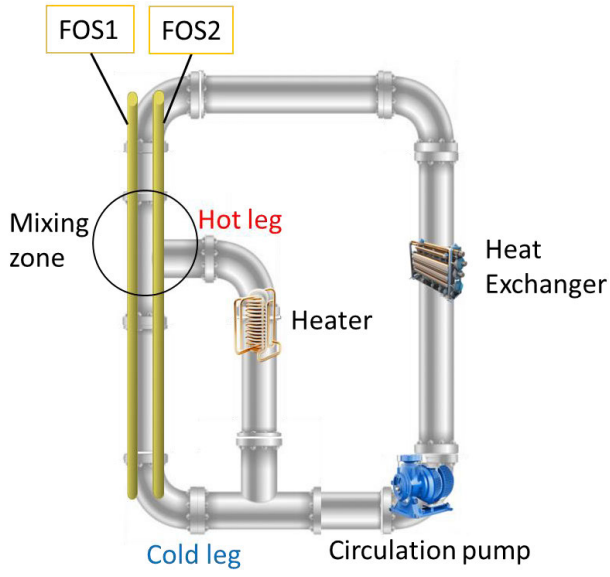



**FIGURE 1.** DTS with Rayleigh backscattering FOS1 and FOS2 installed in a water flow loop with a thermal mixing tee.

hot leg were 0.054 m/s, 0.036 m/s, and 0.018 m/s, respectively. Measurements were performed with a Luna ODiSi 6000 system, which provides an uncertainty in temperature measurements of ±0.9°C. The gauge pitch or spatial resolution for both FOS1 and FOS2 was 2.6 mm, and data was acquired every 0.05 s. Raw experimental data consisted of 728 s total observation time, while lengths of FOS1 and FOS2 were 3.64 m and 5.24 m, respectively. The frequency shifts $S$[GHz] measured with OFDR were converted to temperature $T$[C] using the power law correlation described in [67]:

$$T\,[C] = 1.569\,|S\,[GHz]|^{0.855} + 22.4 \tag{19}$$

The data for the study in this paper consisted of temperature measurements taken from 18.7-cm long sections of FOS1 and FOS2 (72 gauges) for an observation time span of 25 s (500 points in time series). These spatial-temporal sections presented enough temperature fluctuations for LSTM and ARIMA models testing and training. Pseudocolor images of temperature fields (with units of °C), measured with FOS1 and FOS2 in the thermal tee mixing zone, are displayed in Figs. 2(a) and 2(b), respectively. The measurement time is along the x-axis, while the distance along the FOS is on the y-axis. The transient consists of cold leg fluid at approximately 25°C mixing with a hot leg fluid at approximately 45°C, with the temperature downstream of the mixing zone reaching approximately 35°C in the steady state. One can observe from the panels of Fig. 2 that FOS1, which is further from the hot leg inlet than FOS2, measures different temperature patterns which are time-delayed because of convective and diffusive heat transfer. Note that the colormap display function truncates the temperatures to the nearest integer, so regions of the colormap with the same colors do not correspond to constant temperatures.

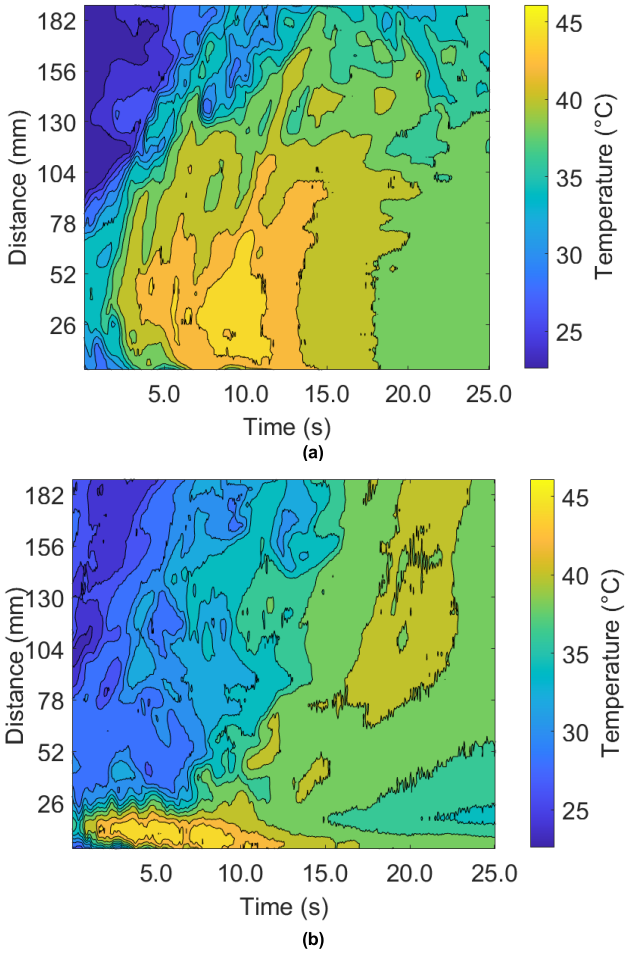



**FIGURE 2.** Pseudocolor temperature map in the thermal tee mixing zone measured with (a) FOS1 and (b) FOS2.

### B. CORRELATIONS BETWEEN FOS1 AND FOS2

To develop a quantitative estimate of similarity between FOS1 and FOS2, we calculate the 72 × 72 matrix of PCCs for all gauges of FOS1 and FOS2. Figure 3 displays the matrix of PCCs as a pseudocolor image. Each element (i,j) of the matrix is the PCC between gauge i of FOS1 and gauge j of FOS2. Note that gauges of FOS1 and FOS2 with the same number and same vertical position in the mixing zone of the thermal tee are not necessarily most correlated. The mapping between FOS1 gauges and maximum correlation FOS2 gauges is listed in Table 2. Note that the subset of maximum correlation gauges for FOS1 consists only of five FOS2 gauges.

Correlations between gauges of FOS1 and FOS2 are elucidated further in the stem plot of PCCs as a function of distance along FOS1. Figure 4(a) shows PCCs calculated for the same-number gauges of FOS1 and FOS2. The FOS1 and FOS2 gauges numbered approximately 25 through 75 (13 cm length of fiber), which are located closest to the junction of the hot and cold legs, are weakly correlated with $\rho < 0.4$. Higher-number FOS1 gauges, starting from approximately

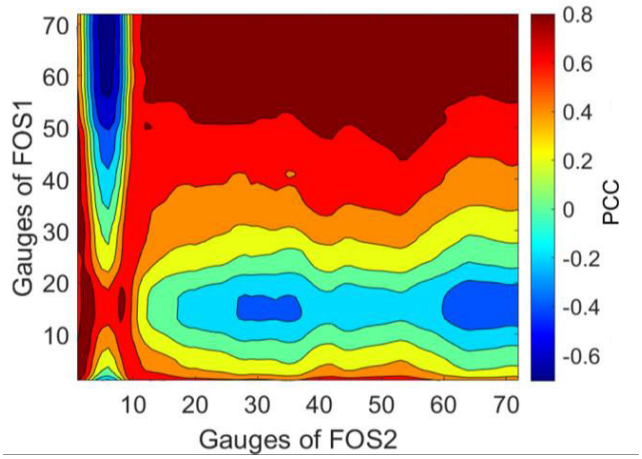



**FIGURE 3.** Pseudocolor image of matrix of PCCs for all gauges of FOS1 and FOS2.

**TABLE 2.** Maximum correlations between gauges of FOS1 and FOS2.

| FOS1 gauges | FOS2 gauges |
|---|---|
| 1 – 3, 29 – 49 | 1 |
| 4 – 12, 20 – 28 | 2 |
| 13 – 19 | 3 |
| 50 – 67, 72 | 53 |
| 68 – 71 | 29 |

gauge 120, located downstream of the mixing zone, display stronger correlations with the same-number FOS2 gauges, with $\rho > 0.8$. Figure 4(b) shows the stem plot of PCCs where for each gauge of FOS1, we calculated the PCC using the maximum correlation gauge of FOS2, as identified in Table 2. For all gauges of FOS1 in Fig. 4(b), $\rho > 0.8$.

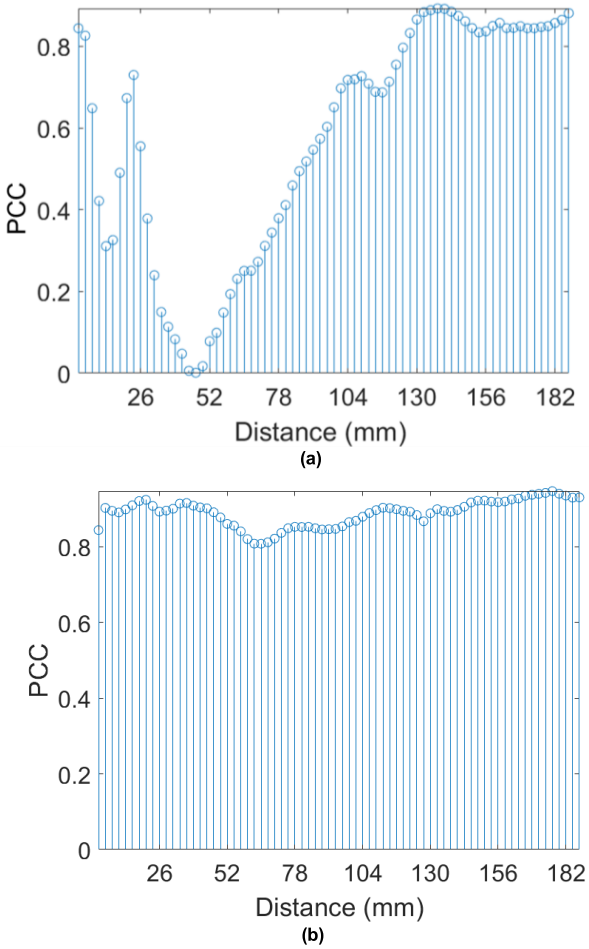



**FIGURE 4.** (a) Stem plot of PCCs calculated for the same-number gauges of FOS1 and FOS2. (b) Stem plot of PCCs for FOS1 gauges and correlated subset of FOS2 gauges in Table 2. In (a) and (b) the distance is along FOS1.

## IV. ARIMA AND LSTM FORECASTING MODELS TRAINING

### A. DTS DATA PARTITIONING FOR FORECASTING USE CASES

Data-driven analysis involves partitioning the data into training and testing datasets, with a typical split of 90% for model training and 10% for model testing. Following this approach, a subset of 10% of time series of all 72 gauges of FOS1 temperature measurements, corresponding to the time interval [6.5 s, 9 s], was selected as the testing data for all three forecasting use cases. The testing dataset time series are 2.5 s-long data segment consisting of 50 points (0.05s temperature sampling rate). Temperature of the FOS1 testing dataset is displayed as pseudocolor image in Fig. 5(a). The testing segment was selected from the first half of the transient to qualitatively exhibit sufficient spatial-temporal amplitude fluctuations in most FOS1 gauges to present a challenge to the forecasting models. The remaining 90% of time series of temperature measurements of all gauges of FOS1 in time intervals [0, 6.5 s] and [9 s, 25 s] were merged into the training data set. Thus, the time series in the training data set with a total time duration of 22.5 s consisted of 450 points.

A subset with the same time stamps was selected from the FOS2 temperature measurements as the training data for the ZSF and TL forecasting use cases. Temperatures of the FOS1 training dataset and FOS2 training dataset are displayed as pseudocolor images in Figs. 5(b) and 5(c), respectively. The two wavy lines on the time axis of Figs. 5(b) and 5(c) indicate the temporal location where the two subsets of the original FOS1 and FOS2 data were merged to form the respective training datasets.

As shown in Fig. 4(a), time series of the same-number gauges of FOS1 and FOS2 are not strongly correlated. Therefore, for the ZSF and the TL forecasting, we develop another training dataset FOS2-corr consisting of a subset of FOS2 gauges that have maximum correlation with FOS1 gauges. The gauges of FOS2-corr are selected according to the mapping in Table 2. Note that FOS2-corr data set consisting of 72 gauges is formed out of five FOS2 gauges with replications. The resulting spatial-temporal temperature map is displayed as a pseudocolor image in Fig. 5(d).

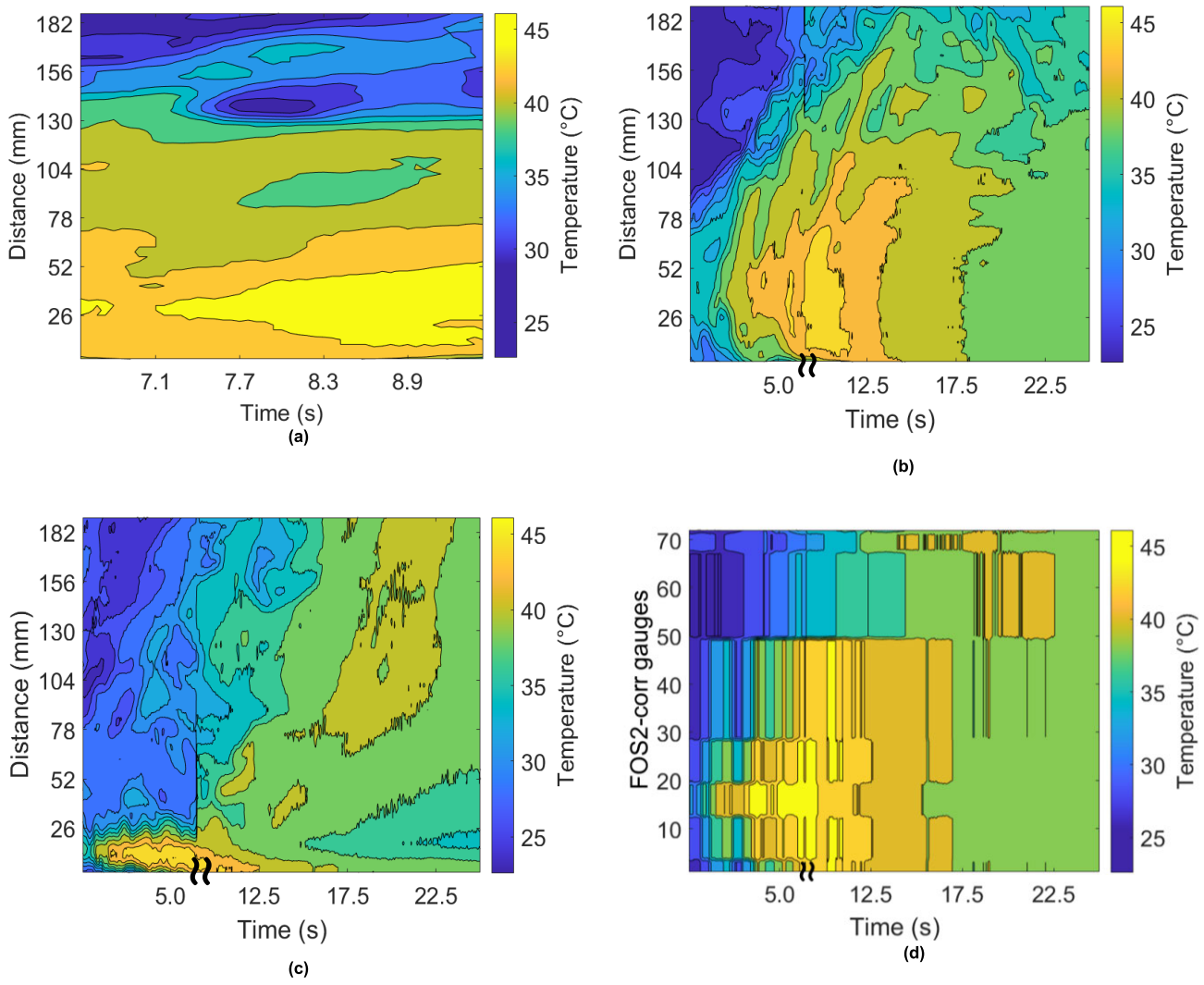



**FIGURE 5.** DTS data partitioning segments displayed as pseudocolor images. (a) FOS1 testing dataset. (b) FOS1 training dataset. (c) FOS2 training dataset. (d) FOS2-corr training dataset.

### B. ARIMA(p,d,q) MODEL TRAINING

Implementation of the ARIMA model involves selection of the values of the ARIMA($p$, $d$, $q$) parameters based on results of statistical tests. The value of the parameter $d$ is the order of differencing required to make the model training data stationary. Stationarity of the DTS data was evaluated with the ADF statistical test for 95% confidence level. For all gauges of FOS1 and FOS2 training datasets, p-values $>$ 0.05 and test statistic $>$ critical value, indicating that the data is non-stationary. Differencing the time series once ($d = 1$) and repeating the ADF tests results in p-values $= 10^{-3}$ and test statistic $<$ critical value for the differenced time series for all gauges of FOS1 training data and FOS2 training data. This indicates that the differenced time series are stationary. Parameters $p$ and $q$ were determined from the differenced stationary time series by finding the number of lags for the PACF given by Equation (10) and ACF given by Equation (11). The values of PACF and ACF saturate after $p = 1$ and $q = 0$ lags, respectively. Therefore, we select ARIMA(1,1,0) implementation for Vanilla and zero-shot forecasting (ZSF) use cases. The FOS1 testing data was differenced once as well for all test cases involving ARIMA forecasting.

### C. LSTM MODEL TRAINING

LSTM networks for all use cases of DTS forecasting were developed with training data, where 90% of that data was reserved for model training and 10% for model validation. LSTM network architecture included an input layer, an LSTM layer, a fully connected layer, and a regression layer. In addition, 10 hidden units were used in the LSTM layer, and the learning rate was set to 0.005. The architecture of the models, including layers in the network, number of LSTM hidden layers, and learning rate were determined by studying the convergence of training and validation loss curves. The number of LSTM hidden layers was varied in the interval of [5, 100], and the learning rate was varied across the

range $[10^{-5}, 10^{-2}]$. The Adam optimizer was used due to its strong performance on time series data [50], [68], [69]. LSTM models for each FOS gauge were trained for 120 epochs with the option for early stopping. All LSTM models converged after 50 epochs.

For the Vanilla use case, we developed LSTM-vanilla trained on FOS1 training dataset. Figures 6(a) and 6(b) show the loss curves of the LSTM-vanilla model, for training and validation, respectively. There are 72 learning curves in Figures 6(a) and 6(b) for each gauge of FOS1. Training and validation losses converge after 50 epochs. Spatially averaged training and validation losses at 50 epochs for LSTM-vanilla are 0.0036°C and 0.0049°C, respectively.

For zero-shot forecasting (ZSF) use case, we developed LSTM-ZSF and LSTM-corrZSF networks using FOS2 training and FOS2-corr training datasets, respectively. For LSTM-ZSF the average training and validation losses at 50 epochs are 0.0036°C and 0.0113°C, respectively. For LSTM-corrZSF, at 50 epochs the average training and validation losses are 0.0036°C and 0.0043°C, respectively. Similar learning curves as the ones shown in Figure 6 were observed for training and validation of LSTM-ZSF and LSTM-corrZSF models.

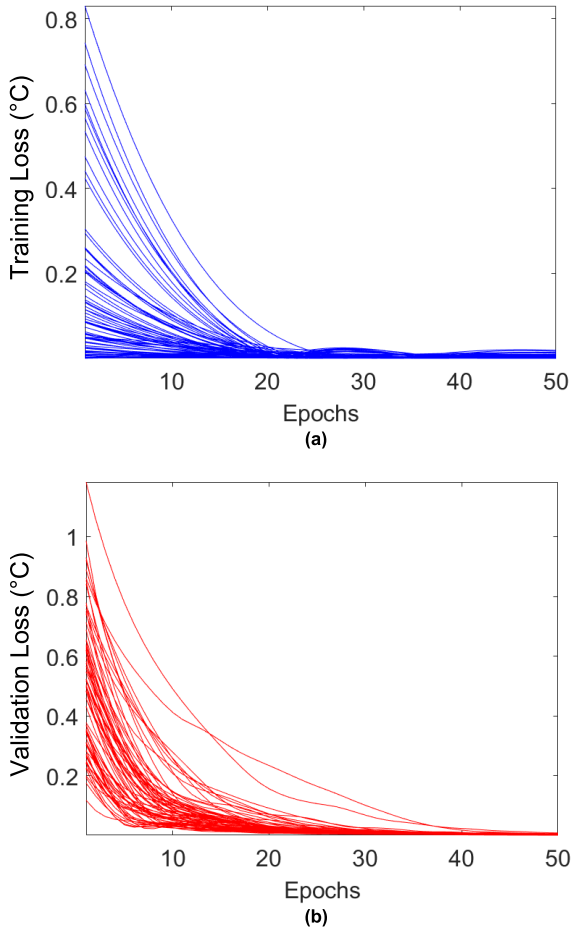



**FIGURE 6.** (a) Training loss curves and (b) validation loss curves for LSTM-Vanilla model. There are 72 learning curves for each gauge of FOS1.

## V. VANILLA USE CASE FORECASTING OF DTS

Figures 7(a) and 7(b) display pseudocolor temperature images of the forecasting errors (error = predicted – measured) for the FOS1 testing data, obtained with the ARIMA-vanilla and LSTM-vanilla models, respectively. The range of time labels on the x-axis of both Figs. 7(a) and 7(b) is 6.5 s to 9 s, which is in reference to the temporal location of the testing segment in the original FOS1 data. The largest errors for both LSTM and ARIMA forecasting are observed in the spatial region approximately between 130 mm and 150 mm (gauges 50 and 58), and for time spans approximately between 7.1 s and 8.5 s. Both LSTM and ARIMA under-predict the temperature at the start of this transient and over-predict at the end of the transient.

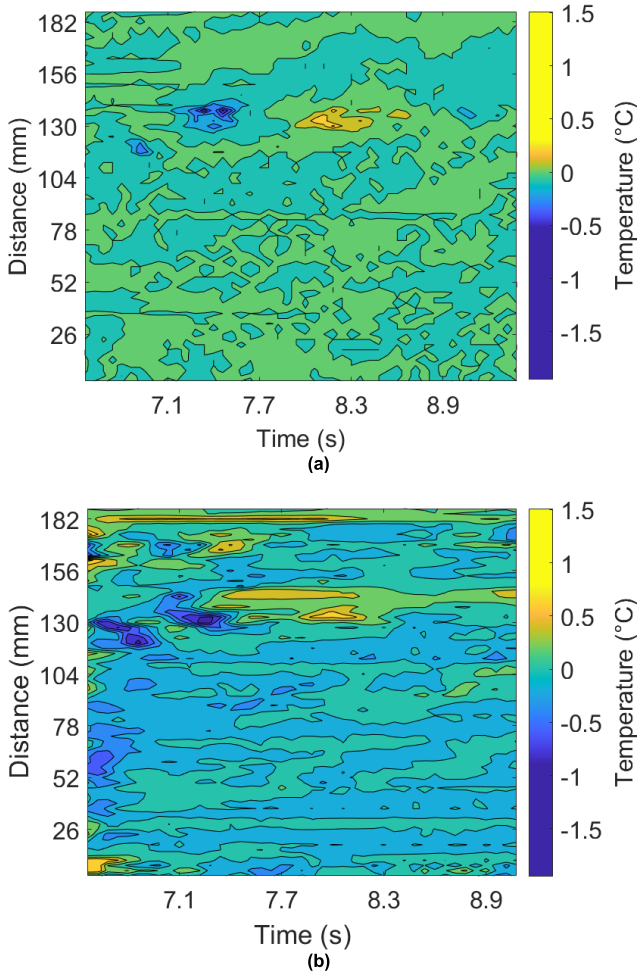



**FIGURE 7.** Pseudocolor temperature map of errors in forecasting of FOS1 testing data with (a) ARIMA-vanilla, (b) LSTM-vanilla.

Table 3 summarizes the MaxAE and MaxAPE, and time and space-averaged RMSE and RMSPE for the LSTM and ARIMA models developed with FOS1 training data, indicating that ARIMA outperforms LSTM in Vanilla use case of forecasting FOS1 testing data. Note that all MaxAE and RMSE for ARIMA, and all RMSE for LSTM are smaller than

the uncertainty of DTS measurements of ±0.9 °C. MaxAE of LSTM slightly exceeds the uncertainty of experimental measurements.

**TABLE 3. Errors in vanilla forecasting of FOS1 temperature with models trained on FOS1 data.**

| Model | MaxAE(°C) | MaxAPE(%) | RMSE(°C) | RMSPE(%) |
|---|---|---|---|---|
| LSTM-vanilla | 0.94 | 3.14 | 0.30 | 0.44 |
| ARIMA-vanilla | 0.52 | 1.58 | 0.01 | 0.06 |

## VI. ZERO-SHOT FORECASTING OF DTS

Figure 8(a) displays pseudocolor temperature image of the ARIMA-ZSF error (error = predicted − measured). From qualitative comparison of the images in Fig. 8(a) and Fig. 7(a), one can observe that ARIMA-ZSF and ARIMA-vanilla errors are similar. Figure 8(b) displays pseudocolor temperature image of the LSTM-ZSF error. From qualitative comparison of the images in Fig. 8(b) and Fig. 7(b), one can observe that LSTM-ZSF errors are larger than those of LSTM-vanilla. In particular, the LSTM-ZSF errors are substantially larger than those of LSTM-vanilla in the spatial region from 26 mm to 52 mm. As can be seen in Fig 4(a), the gauges of FOS2 and FOS1 at these spatial locations are weakly correlated. Spatial-temporal errors with LSTM-corrZSF are displayed as a pseudocolor image in Fig. 8(c). Qualitatively comparing Figs. 8(b) and 8(c), the errors of LSTM-corrZSF are smaller than those of LSTM-ZSF.

Table 4 lists the MaxAE and MaxAPE, and time and space-averaged RMSE and RMSPE for the three ZSF models. The MaxAE and RMSE of LSTM-corrZSF are lower than those of LSTM-ZSF, but larger than those of ARIMA-ZSF. Comparing Table 4 and Table 3, RMSE and MaxAE of ARIMA in ZSF use case are slightly larger than those of the Vanilla use case. ZSF use case RMSE and MaxAE of LSTM-corrZSF are smaller and larger, respectively, than the corresponding errors of the LSTM-vanilla. It should be noted that RMSE for all ZSF models and MaxAE for ARIMA-ZSF are smaller than the DTS measurement uncertainty of ±0.9°C.

**TABLE 4. Errors in ZSF of FOS1 for models trained on FOS2 data.**

| Model | MaxAE(°C) | MaxAPE(%) | RMSE(°C) | RMSPE(%) |
|---|---|---|---|---|
| LSTM-ZSF | 1.95 | 4.75 | 0.58 | 0.92 |
| LSTM-corrZSF | 1.51 | 4.4 | 0.23 | 0.33 |
| ARIMA-ZSF | 0.73 | 1.91 | 0.02 | 0.13 |

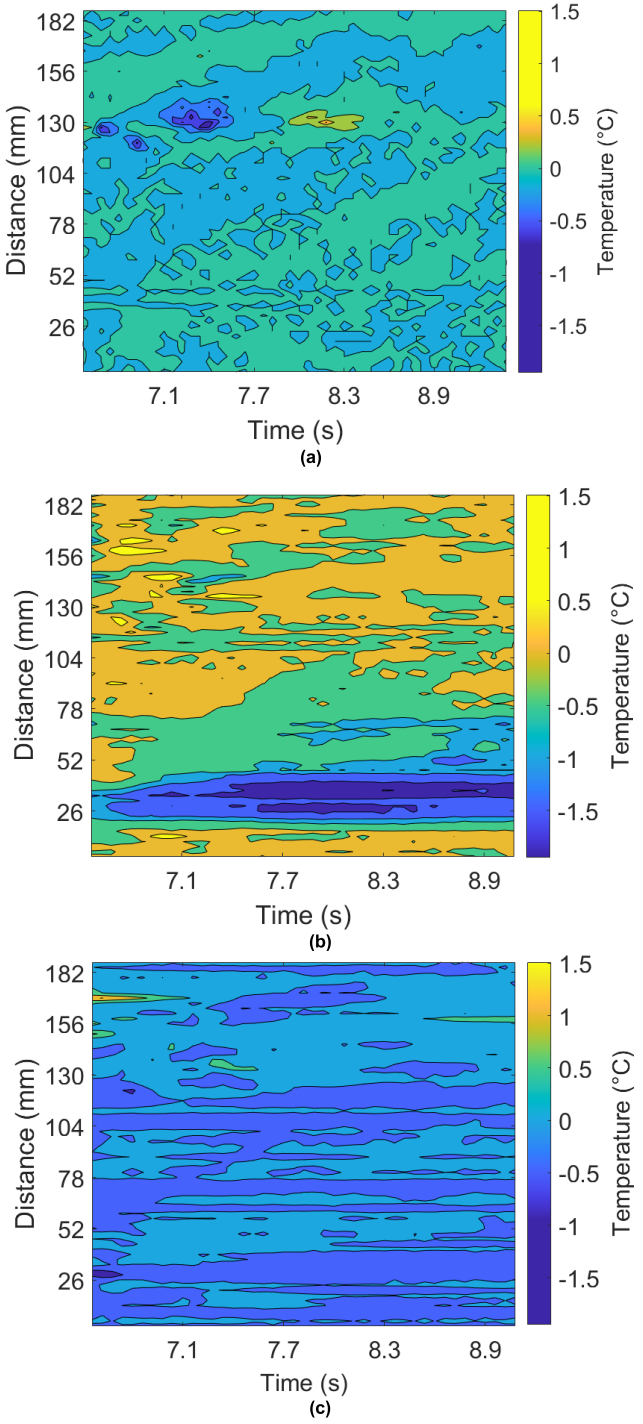



**FIGURE 8. Pseudocolor temperature map of ZSF errors of FOS1 testing data with (a) ARIMA-ZSF, (b) LSTM-ZSF, (c) LSTM-corrZSF.**

## VII. FORECASTING OF DTS WITH TRANSFER LEARNING

### A. LSTM-TL TRAINING

The performance of LSTM-corrZSF could be further improved with transfer learning (TL) by re-training the LSTM-corrZSF with the FOS1 training data. During the re-training process, all trainable parameters of the network are initialized with the values of those in the pre-trained network. However, re-training the pre-trained LSTM-corrZSF network results in large forecasting errors for relatively small amounts of FOS1 training data. As an example, the MaxAE

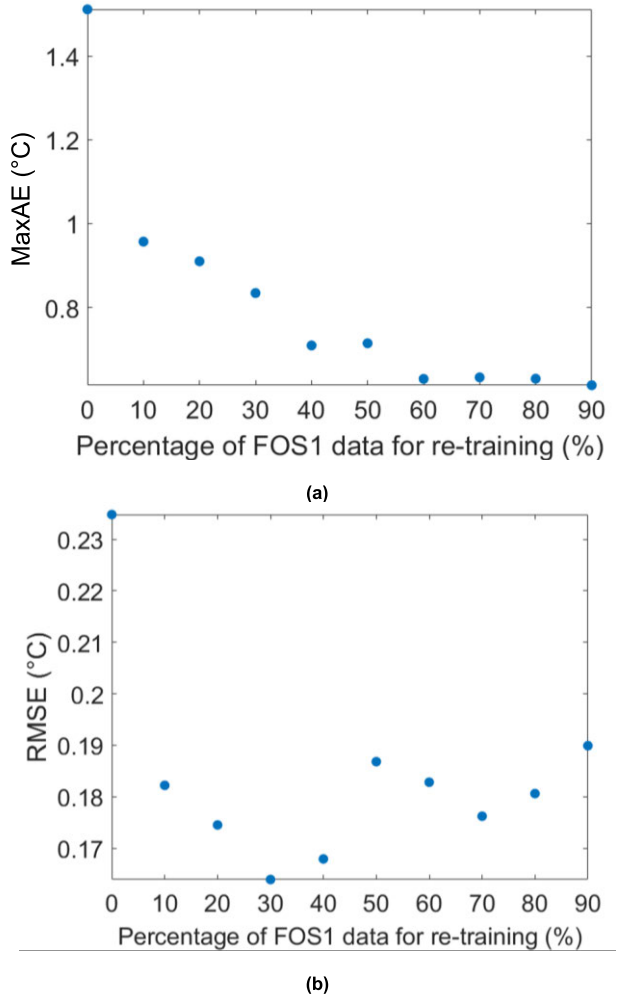



**FIGURE 9.** (a) MaxAE and (b) space and time-averaged RMSE as functions of fraction of FOS1 training data used in training of the LSTM-TL model.

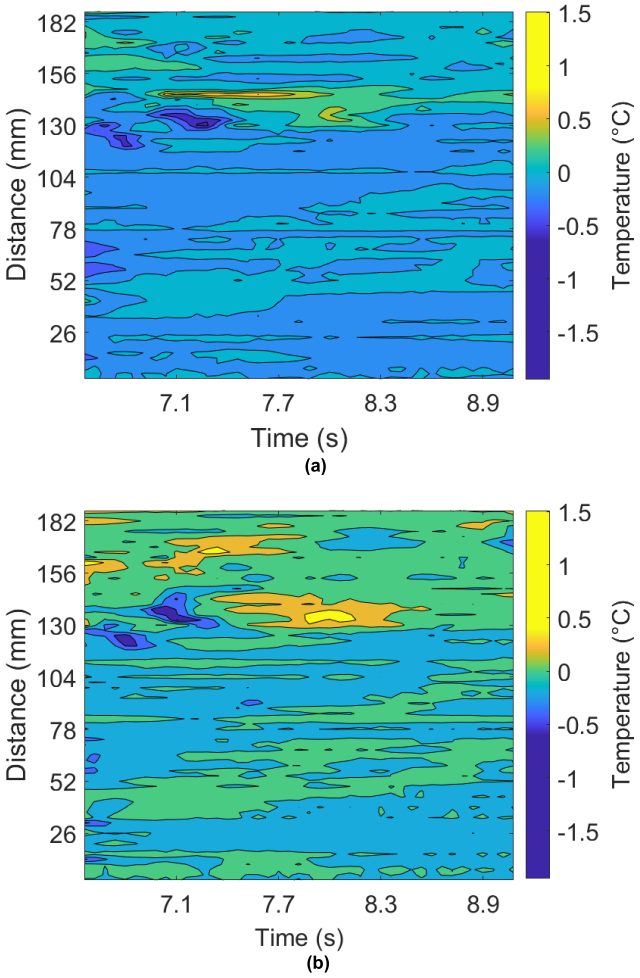



**FIGURE 10.** Pseudocolor temperature maps of forecasting errors with LSTM-TL networks re-trained with (a) 30% and (b) 60% of FOS1 training data.

is 18.39°C and the space and time-averaged RMSE is 7.9°C when re-training LSTM-corrZSF with 10% of FOS1 data. This is a common phenomenon of catastrophic forgetting in neural networks, where previously learned knowledge is forgotten when the networks are trained on new datasets [70]. To mitigate the catastrophic forgetting, forecasting models are trained on a combination of old and new datasets. Therefore, in the TL use case, we re-train the pre-trained LSTM-corrZSF network on datasets consisting of both the FOS2 training data and the FOS1 training data.

Figures 9(a) and 9(b) plot the MaxAE and space and time-averaged RMSE, respectively, as functions of the fraction of FOS1 data used in LSTM-TL model re-training. The MaxAE and RMSE for 0% of FOS1 data are the errors listed for LSTM-corrZSF in Table 4. From Fig. 9(a), increasing the fraction of FOS1 training data for LSTM-TL re-training leads to reduction in the MaxAE. The decrease in MaxAE saturates after the re-training includes 60% of FOS1 training data. LSTM-TL retrained with 60% of FOS1 data converged after 70 epochs with the average training loss of 0.0028°C, and the average validation loss of 0.011°C. From Fig. 9(b), after the initial decline, the RMSE does not saturate with increasing fraction of FOS1 training data, and instead oscillates in the 0.17°C to 0.19°C range. The minimum RMSE value is observed for 30% of FOS1 re-training data.

### B. LSTM-TL TESTING

Figures 10(a) and 10(b) display pseudocolor temperature images of forecasting errors with LSTM-TL networks retrained with 30% and 60% of FOS1 training data, respectively. The largest errors for both cases are observed in the spatial region approximately between 130 mm and 150 mm (gauges 50 and 58).

Table 5 lists the MaxAE and MaxAPE, and time and space-averaged RMSE and RMSPE for the LSTM-TL networks developed with 30% and 60% of FOS1 training data. MaxAE and RMSE for LSTM-TL are smaller than those of LSTM-corrZSF listed in Table 4. The smallest MaxAE with 60% of FOS1 data used in re-training of LSTM-TL is between the MaxAE values for ARIMA-ZSF and

ARIMA-vanilla. However, the smallest RMSE for 30% of FOS1 used for re-training LSTM-TL is larger than RMSE for ARIMA-ZSF and ARIMA-vanilla. It should be noted that MaxAE and RMSE for LSTM-TL in Table 5 are smaller than the DTS uncertainty in measurements of ±0.9°C.

**TABLE 5. Errors in LSTM-TL forecasting of FOS1 testing data.**

| FOS1 training data fraction(%) | MaxAE(°C) | MaxAPE(%) | RMSE(°C) | RMSPE(%) |
|---|---|---|---|---|
| 30 | 0.83 | 2.32 | 0.16 | 0.28 |
| 60 | 0.63 | 1.96 | 0.18 | 0.29 |

## VIII. CONCLUSION

We benchmarked performance of machine learning LSTM and statistical ARIMA models in temporal forecasting of DTS. The data in this study was a thermal transient in a mixing Tee of a water flow loop, measured with two co-located Rayleigh scattering FOS. Temperature measurements were taken from 18.7-cm long sections of two FOS (72 gauges) for a total observation time span of 25 s (500 points in time series). The data was partitioned into 2.5 s – long (50 points in time series) testing dataset and 22.5 s – long (450 points in time series) training dataset. Performance of LSTM and ARIMA was evaluated with RMSE, MaxAE, RMSPE and MaxAPE error metrics. The benchmarking study investigated temperature forecasting in Vanilla, zero shot forecasting (ZSF) and transfer learning (TL) use cases. In Vanilla use case, predictions of FOS temperature were made with LSTM and ARIMA trained on the data from the same FOS that was used for testing. In ZSF use case, forecasting of one FOS was made with ARIMA and LSTM trained on another FOS. In TL use case, LSTM developed in ZSF was re-trained for improved performance.

The study has shown that ARIMA outperforms LSTM in all use cases. However, since both LSTM and ARIMA are data-driven forecasting models, their performance depends on the training data. Development of ARIMA requires performing statistical tests to determine the $(p, d, q)$ model parameters, and differencing the training and testing data to make it stationary. The MaxAE and RMSE for ARIMA are similar for vanilla and ZSF use cases. Since datasets used with ARIMA are differenced to make them stationary, ARIMA is relatively insensitive to correlations between the original training and testing datasets.

The MaxAE and RMSE for LSTM in ZSF are larger than the corresponding errors in the Vanilla use case. To improve performance of LSTM in the ZSF use case, we identified gauges in the FOS1 training dataset that are most correlated with the FOS2 training dataset. Training LSTM on the highly correlated training dataset resulted in decrease of MaxAE and RMSE. Performance of LSTM can be further improved through TL by re-training LSTM developed in ZSF case. The re-training dataset includes all the data from training in ZSF and a sub-set of the training data from the FOS that was the target of forecasting. This shows that LSTM depends on the correlation between the training and testing datasets, and on the volume of the training data.

Future studies will investigate benchmarking of several other time series with forecasting methods such as transformers, which in principle have the potential to perform well in ZSF applications. In addition, we will explore performance of LSTM, ARIMA and other methods in multi-step ahead forecasting. Furthermore, we will investigate the effect of different activation functions in LSTM models performance.

## REFERENCES

[1] H. M. Hashemian and J. Jiang, "Nuclear plant temperature instrumentation," *Nucl. Eng. Design*, vol. 239, no. 12, pp. 3132–3141, Dec. 2009.
[2] A. Heifetz and R. Vilim, "Eigendecomposition model of resistance temperature detector with applications to S-$CO_2$ cycle sensing," *Nucl. Eng. Design*, vol. 311, pp. 60–68, Jan. 2017.
[3] H. M. Hashemian and W. C. Bean, "Sensors for next-generation nuclear plants: Fiber-optic and wireless," *Nucl. Sci. Eng.*, vol. 169, no. 3, pp. 262–278, Nov. 2011.
[4] P. Lu, N. Lalam, M. Badar, B. Liu, B. T. Chorpening, M. P. Buric, and P. R. Ohodnicki, "Distributed optical sensing: Review and perspective," *Appl. Phys. Rev.*, vol. 6, 2019, Art. no. 041302.
[5] A. Ukil, H. Braendle, and P. Krippner, "Distributed temperature sensing: Review of technology and applications," *IEEE Sensors J.*, vol. 12, no. 5, pp. 885–892, May 2012.
[6] E. Schena, D. Tosi, P. Saccomandi, E. Lewis, and T. Kim, "Fiber optic sensors for temperature monitoring during thermal treatments: An overview," *Sensors*, vol. 16, no. 7, p. 1144, Jul. 2016.
[7] J. Wu, M. Wang, K. Zhao, S. Huang, M. A. S. Zaghloul, R. Cao, D. Carpenter, G. Zheng, S. D. Rountree, and K. P. Chen, "Distributed fiber sensors with high spatial resolution in extreme radiation environments in nuclear reactor cores," *J. Lightw. Technol.*, vol. 39, no. 14, pp. 4873–4883, Jun. 2, 2021.
[8] J. Tyler Gates and P. V. Tsvetkov, "Testing of fiber optic based sensors for advanced reactors in the Texas A&M university TRIGA reactor," *Ann. Nucl. Energy*, vol. 196, Feb. 2024, Art. no. 110222.
[9] H. C. Hyer, D. R. Giuliano, and C. M. Petrie, "Toward local core outlet temperature monitoring in gas-cooled nuclear reactors using distributed fiber-optic temperature sensors," *Appl. Thermal Eng.*, vol. 230, Jul. 2023, Art. no. 120847.
[10] S. Lomperski, C. Gerardi, and W. D. Pointer, "Fiber optic distributed temperature sensor mapping of a jet-mixing flow field," *Experim. Fluids*, vol. 56, no. 3, pp. 1–16, Mar. 2015.
[11] C. Gerardi, N. Bremer, D. Lisowski, and S. Lomperski, "Distributed temperature sensor testing in liquid sodium," *Nucl. Eng. Design*, vol. 312, pp. 59–65, Feb. 2017.
[12] M. Weathered, J. Rein, M. Anderson, P. Brooks, and B. Coddington, "Characterization of thermal striping in liquid sodium with optical fiber sensors," *J. Nucl. Eng. Radiat. Sci.*, vol. 3, no. 4, Oct. 2017, Art. no. 041003.
[13] K. F. Becker and M. H. Anderson, "Optical fiber-based level sensor for high temperature applications," *IEEE Sensors J.*, vol. 20, no. 16, pp. 9187–9195, Aug. 2020.
[14] O. Arora, B. Lancaster, S. R. Yang, R. Vaghetto, and Y. Hassan, "Advanced flow and temperature measurements in a forced convection molten salt test loop," *Ann. Nucl. Energy*, vol. 159, Sep. 2021, Art. no. 108269.
[15] S. Loranger, M. Gagné, V. Lambin-Iezzi, and R. Kashyap, "Rayleigh scatter based order of magnitude increase in distributed temperature and strain sensing by simple UV exposure of optical fibre," *Sci. Rep.*, vol. 5, no. 1, p. 11177, Jun. 2015.
[16] T. W. Wood, B. Blake, T. E. Blue, C. M. Petrie, and D. Hawn, "Evaluation of the performance of distributed temperature measurements with single-mode fiber using Rayleigh backscatter up to 1000°C," *IEEE Sensors J.*, vol. 14, no. 1, pp. 124–128, Jan. 2014.
[17] B. Liu, Z. Yu, C. Hill, Y. Cheng, D. Homa, G. Pickrell, and A. Wang, "Sapphire-fiber-based distributed high-temperature sensing system," *Opt. Lett.*, vol. 41, no. 18, p. 4405, 2016.

[18] B. Liu, M. P. Buric, B. T. Chorpening, Z. Yu, D. S. Homa, G. R. Pickrell, and A. Wang, "Design and implementation of distributed ultra-high temperature sensing system with a single crystal fiber," *J. Lightw. Technol.*, vol. 36, no. 23, pp. 5511–5520, Dec. 1, 2018.

[19] H. Wang, X. Wang, X. Tuo, T. Liu, L. Meng, and P. Zhong, "Experimental research on a Raman-based distributed temperature sensor assisted by PCA for locating the temperature abnormal event of nuclear waste drums," *Appl. Opt.*, vol. 59, no. 2, p. 579, 2020.

[20] M. Kasinathan, S. Sosamma, C. BabuRao, N. Murali, and T. Jayakumar, "Fiber optic sensors for nuclear power plant applications," in *AIP Conf. Proc.*, 2012, pp. 1013–1020.

[21] H. S. Pradhan, P. Mahajan, N. Vislawath, and V. S. D. N. Prasad, "A highly improved long-range BDTS for nuclear reactors using hybrid algorithm," *IEEE Sensors J.*, vol. 23, no. 18, pp. 21269–21278, Sep. 2023.

[22] M. Tateda and T. Horiguchi, "Advances in optical time domain reflectometry," *J. Lightw. Technol.*, vol. 7, no. 8, pp. 1217–1224, Jun. 1989.

[23] C. Liang, Q. Bai, M. Yan, Y. Wang, H. Zhang, and B. Jin, "A comprehensive study of optical frequency domain reflectometry," *IEEE Access*, vol. 9, pp. 41647–41668, 2021.

[24] M. Luo, J. Liu, C. Tang, X. Wang, T. Lan, and B. Kan, "0.5 mm spatial resolution distributed fiber temperature and strain sensor with position-deviation compensation based on OFDR," *Opt. Exp.*, vol. 27, no. 24, pp. 35823–35829, 2019.

[25] J. Li, Q. Chen, J. Zhou, Z. Cao, T. Li, F. Liu, Z. Yang, S. Chang, K. Zhou, Y. Ming, T. Yan, and W. Zheng, "Radiation damage mechanisms and research status of radiation-resistant optical fibers: A review," *Sensors*, vol. 24, no. 10, p. 3235, May 2024.

[26] S. Girard, A. Alessi, N. Richard, L. Martin-Samos, V. De Michele, L. Giacomazzi, S. Agnello, D. D. Francesca, A. Morana, B. Winkler, I. Reghioua, P. Paillet, M. Cannas, T. Robin, A. Boukenter, and Y. Ouerdane, "Overview of radiation induced point defects in silica-based optical fibers," *Rev. Phys.*, vol. 4, Nov. 2019, Art. no. 100032.

[27] S. Rana, H. Subbaraman, A. Fleming, and N. Kandadai, "Numerical analysis of radiation effects on fiber optic sensors," *Sensors*, vol. 21, no. 12, p. 4111, Jun. 2021.

[28] N. A. Abdul Rahim, M. A. Davis, L. Routhier, J. Chevalier, J. Bos, S. Kreger, and E. Sanborn, "Accuracy and survivability of distributed fiber optic temperature sensors," in *Proc. 53rd AIAA Aerosp. Sci. Meeting*, Jan. 2015, pp. 1–10, doi: 10.2514/6.2015-1920.

[29] A. Rovera, A. Tancau, N. Boetti, M. D. L. Dalla Vedova, P. Maggiore, and D. Janner, "Fiber optic sensors for harsh and high radiation environments in aerospace applications," *Sensors*, vol. 23, no. 5, p. 2512, Feb. 2023.

[30] G. Ekechukwu and J. Sharma, "Degradation analysis of single-mode and multimode fibers in a full-scale wellbore and its impact on DAS and DTS measurements," *IEEE Sensors J.*, vol. 23, no. 9, pp. 9287–9300, May 2023.

[31] H. Darvishi, D. Ciuonzo, and P. S. Rossi, "A machine-learning architecture for sensor fault detection, isolation, and accommodation in digital twins," *IEEE Sensors J.*, vol. 23, no. 3, pp. 2522–2538, Feb. 2023.

[32] A. Venketeswaran, N. Lalam, J. Wuenschell, P. R. Ohodnicki, M. Badar, K. P. Chen, P. Lu, Y. Duan, B. Chorpening, and M. Buric, "Recent advances in machine learning for fiber optic sensor applications," *Adv. Intell. Syst.*, vol. 4, no. 1, Jan. 2022, Art. no. 2100067.

[33] Y. Zhou, Y.-N. Zhang, Q. Yu, L. Ren, Q. Liu, and Y. Zhao, "Application of machine learning in optical fiber sensors," *Measurement*, vol. 228, Mar. 2024, Art. no. 114391.

[34] E. Reyes-Vera, A. Valencia-Arias, V. García-Pineda, E. F. Aurora-Vigo, H. Alvarez Vásquez, and G. Sánchez, "Machine learning applications in optical fiber sensing: A research agenda," *Sensors*, vol. 24, no. 7, p. 2200, Mar. 2024.

[35] C. Karapanagiotis and K. Krebber, "Machine learning approaches in Brillouin distributed fiber optic sensors," *Sensors*, vol. 23, no. 13, p. 6187, Jul. 2023.

[36] H. A. Arief, T. Wiktorski, and P. J. Thomas, "A survey on distributed fibre optic sensor data modelling techniques and machine learning algorithms for multiphase fluid flow estimation," *Sensors*, vol. 21, no. 8, p. 2801, Apr. 2021.

[37] S. Manavi Roodsari, S. Freund, M. Angelmahr, C. Seppi, G. Rauter, W. Schade, and P. C. Cattin, "Deep learning-based approach for high spatial resolution fibre shape sensing," *Commun. Eng.*, vol. 3, no. 1, p. 19, Jan. 2024.

[38] Y. C. Manie, J.-W. Li, P.-C. Peng, R.-K. Shiu, Y.-Y. Chen, and Y.-T. Hsu, "Using a machine learning algorithm integrated with data de-noising techniques to optimize the multipoint sensor network," *Sensors*, vol. 20, no. 4, p. 1070, Feb. 2020.

[39] Z. Wu, M. A. S. Zaghloul, D. Carpenter, M.-J. Li, J. Daw, Z.-H. Mao, C. Hnatovsky, S. J. Mihailov, and K. P. Chen, "Mitigation of radiation-induced fiber Bragg grating (FBG) sensor drifts in intense radiation environments based on Long-Short-Term memory (LSTM) network," *IEEE Access*, vol. 9, pp. 148296–148301, 2021.

[40] Y. Li, R. Wang, Y. Song, D. Wan, D. Hu, and C. Jiang, "A deep learning approach to estimate temperature and flow velocity distributions of wire-wrapped fuel assemblies," *Int. Commun. Heat Mass Transf.*, vol. 158, Nov. 2024, Art. no. 107853.

[41] S. Nakarmi, J. A. Leiding, K.-S. Lee, and N. P. Daphalapurkar, "Predicting non-linear stress–strain response of mesostructured cellular materials using supervised autoencoder," *Comput. Methods Appl. Mech. Eng.*, vol. 432, Dec. 2024, Art. no. 117372.

[42] A. Ray and S. Phoha, "Calibration and estimation of redundant signals for real-time monitoring and control," *Signal Process.*, vol. 83, no. 12, pp. 2593–2605, Dec. 2003.

[43] Y. Sun, L. Zhang, and M. Yao, "Chaotic time series prediction of nonlinear systems based on various neural network models," *Chaos, Solitons Fractals*, vol. 175, Oct. 2023, Art. no. 113971.

[44] P. Kanjalkar, J. Gawande, J. Kanjalkar, K. More, V. Manwatkar, S. Shimple, and A. Mohite, "Temperature sensor failure detection and diagnosis based on ARIMA model," in *AIP Conf. Proc.*, vol. 3122, 2024, pp. 1–11.

[45] A. Peng, X. Zhang, W. Xu, and Y. Tian, "Effects of training data on the learning performance of LSTM network for runoff simulation," *Water Resour. Manage.*, vol. 36, no. 7, pp. 2381–2394, May 2022.

[46] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 1394–1401.

[47] D. Kobiela, D. Krefta, W. Król, and P. Weichbroth, "ARIMA vs LSTM on NASDAQ stock exchange data," *Proc. Comput. Sci.*, vol. 207, pp. 3836–3845, May 2022.

[48] M. Elsaraiti and A. Merabet, "A comparative analysis of the ARIMA and LSTM predictive models and their effectiveness for predicting wind speed," *Energies*, vol. 14, no. 20, p. 6782, Oct. 2021.

[49] C. Duan, M. Hu, and H. Zhang, "Comparison of ARIMA and LSTM in predicting structural deformation of tunnels during operation period," *Data*, vol. 8, no. 6, p. 104, Jun. 2023.

[50] K. Prantikos, L. H. Tsoukalas, and A. Heifetz, "Physics-informed neural network solution of point kinetics equations for a nuclear reactor digital twin," *Energies*, vol. 15, no. 20, p. 7697, Oct. 2022.

[51] M. Khashei and M. Bijari, "An artificial neural network (p,d,q) model for timeseries forecasting," *Expert Syst. Appl.*, vol. 37, no. 1, pp. 479–489, Jan. 2010.

[52] S. Pantopoulou, V. Ankel, M. T. Weathered, D. D. Lisowski, A. Cilliers, L. H. Tsoukalas, and A. Heifetz, "Monitoring of temperature measurements for different flow regimes in water and galinstan with long short-term memory networks and transfer learning of sensors," *Computation*, vol. 10, no. 7, p. 108, Jun. 2022.

[53] K. Prantikos, S. Chatzidakis, L. H. Tsoukalas, and A. Heifetz, "Physics-informed neural network with transfer learning (TL-PINN) based on domain similarity measure for prediction of nuclear reactor transients," *Sci. Rep.*, vol. 13, no. 1, p. 16840, Oct. 2023.

[54] F. Guc and Y. Chen, "Sensor fault diagnostics (Don't short) using physics-informed transfer learning framework," *Sensors*, vol. 22, no. 8, p. 2913, Apr. 2022.

[55] Z. Zhao, Q. Zhang, X. Yu, C. Sun, S. Wang, R. Yan, and X. Chen, "Applications of unsupervised deep transfer learning to intelligent fault diagnosis: A survey and comparative study," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–28, 2021.

[56] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2009.

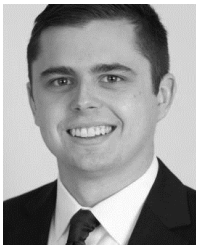
[57] K. R. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–26, May 2016.

[58] Y. Li, D. Wan, Z. Wang, and D. Hu, "Physics-constrained deep learning approach for solving inverse problems in composite laminated plates," *Composite Struct.*, vol. 348, Nov. 2024, Art. no. 118514.

[59] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, "Transfer learning: A friendly introduction," *J. Big Data*, vol. 9, no. 1, p. 102, Oct. 2022.

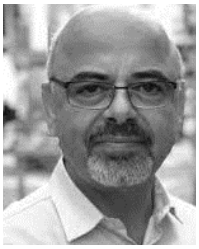
[60] H. H. Mao, "A survey on self-supervised pre-training for sequential transfer learning in neural networks," 2020, *arXiv:2007.00800*.
[61] Q. Wu, H. Wu, X. Zhou, M. Tan, Y. Xu, Y. Yan, and T. Hao, "Online transfer learning with multiple homogeneous or heterogeneous sources," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 7, pp. 1494–1507, Jul. 2017.
[62] P. Gikunda and N. Jouandeau, "Homogeneous transfer active learning for time series classification," in *Proc. 20th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2021, pp. 778–784.
[63] G. Nahler, "Pearson correlation coefficient," *Noise Reduction Speech Process.*, vol. 2, p. 132, Jan. 2009.
[64] Y.-J. Wang, E. Baglietto, and K. Shirvan, "A two-level machine learning approach for predicting thermal striping in T-junctions with upstream elbow," *Numer. Heat Transf., Part B, Fundamentals*, vol. 85, no. 6, pp. 662–682, Jun. 2024.
[65] S. Bakhtiari, A. C. V. Cabral, L. B. Carasik, T. H. Elmer, A. Heifetz, and D. Lisowski, "Measurement of flow in a mixing tee using ultrasound Doppler velocimetry for opaque fluids," *Trans. Am. Nucl. Soc.*, vol. 21, pp. 1643–1645, Jan. 2019.
[66] N. Tanaka, S. Moriya, S. Ushijima, T. Koga, and Y. Eguchi, "Prediction method for thermal stratification in a reactor vessel," *Nucl. Eng. Design*, vol. 120, nos. 2–3, pp. 395–402, Jun. 1990.
[67] M. T. Weathered, "Characterization of sodium thermal hydraulics with optical fiber temperature sensors," Ph.D. dissertation, Nuclear Eng. Eng. Phys., Univ. Wisconsin-Madison, Madison, WI, USA, 2017.
[68] M. Rana, M. M. Uddin, and M. M. Hoque, "Effects of activation functions and optimizers on stock price prediction using LSTM recurrent networks," in *Proc. 3rd Int. Conf. Comput. Sci. Artif. Intell.*, Dec. 2019, pp. 354–358.
[69] A. Makinde, "Optimizing time series forecasting: A comparative study of Adam and Nesterov accelerated gradient on LSTM and GRU networks using stock market data," 2024, *arXiv:2410.01843*.
[70] R. Kemker, M. McClure, A. Abitino, T. L. Hayes, and C. Kanan, "Measuring catastrophic forgetting in neural networks," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, Apr. 2018, pp. 1–20.

**STELLA PANTOPOULOU** (Member, IEEE) received the B.S. degree in electrical engineering from the University of Patras, Greece, in 2018, and the M.S. degree in nuclear engineering from Purdue University, in 2021. She is currently pursuing the Ph.D. degree in nuclear engineering. She is a Visiting Student with Argonne National Laboratory. She has received a number of awards, including the ANS Robert E. Uhrig Graduate Scholarship, from 2022 to 2023, the Gerondelis Foundation Graduate Scholarship, in 2022, and the Purdue University Bilsland Fellowship, in 2024. Her research interests include machine learning methods and data analysis for advanced reactors.



**MATTHEW WEATHERED** received the bachelor's degree in mechanical engineering from Northwestern University, in 2013, and the master's and Ph.D. degrees in nuclear engineering and engineering physics from the University of Wisconsin–Madison, in 2018. He is currently an Engineer with the Nuclear Science and Engineering Division, Argonne National Laboratory. He has significant experience in the design and operation of high temperature alkali metal systems relevant to the nuclear energy and concentrated solar industry. He is an Engineer and an Operator with the Mechanisms Engineering Test Loop (METL), a more than 700 gallon reactor grade sodium facility capable of testing small to intermediate scale test articles for commissioning reactor components and instrumentation. He is the Principal Investigator of the Thermal Hydraulic Experimental Test Article (THETA), a scaled test article capable of simulating a pool type sodium fast reactor under various operating regimes while collecting more than 9,000 data points with the use of distributed optical fiber temperature sensors, thermocouple rakes, and electromagnetic flowmeters. His previous work deploying optical fiber temperature sensors in sodium and molten salt to quantify thermal phenomena can be found published in ASME and IEEE journals.

**DARIUS LISOWSKI** received the Ph.D. degree in nuclear engineering from the University of Wisconsin–Madison, in 2013. He manages the Reactor Safety Testing and Analysis Group, Argonne's Nuclear Science and Engineering Division, Argonne National Laboratory. His primary research area is in experimental thermal hydraulics with a focus on natural circulation systems and two-phase flow. He also has extensive experience with the development and use of high-fidelity sensors, data acquisition, and control systems. He is currently the Principal Investigator of the Natural Convection Shutdown Heat Removal Test Facility (NSTF), Argonne, and has led the conversion to a water-based cooling design. In addition to experimental practices, he has carried out extensive studies on scaling criteria, instability modes, and computational model with both system level and CFD codes.



**LEFTERI H. TSOUKALAS** is currently a Professor and the former Head of the School of Nuclear Engineering, Purdue University. He is an Expert in AI/ML algorithms for nuclear reactors. He has extensive experience as a safeguards Engineer and as a Nuclear Instrumentation and Controls Specialist with more than three decades of accumulated experience as a Project Manager of competitively funded projects sponsored by DOE, NNSA, NRC, DOD, NSF, and EPRI. His research covers both experimental and computational methods development, including multi-variant statistical methods for ML and logical AI, Gaussian processes for background estimation, wavelet analysis, and Hilbert–Huang transforms. He was a recipient of the Humboldt Prize (Germany's highest honor for international scientists).

**ALEXANDER HEIFETZ** (Senior Member, IEEE) received the B.S. degree in applied math and engineering science, the M.S. degree in physics, and the Ph.D. degree in electrical engineering from Northwestern University. He is currently a Principal Electrical Engineer with the Nuclear Science and Engineering (NSE) Division. He first joined Argonne as the Director's Postdoctoral Fellow. He is an Adjunct Professor with the Civil, Materials, and Environmental Engineering Department, University of Illinois at Chicago. He has published more than 70 journal and conference proceedings papers and has three granted U.S. patents. His current research is on non-destructive evaluation, thermal hydraulic sensing, and machine learning. He is a member of the Northwestern Argonne Institute for Science and Engineering (NAISE). He shared the Best Paper Awards at the 2019 and 2020 IEEE International Conferences on Electro/Information Technology (EIT).