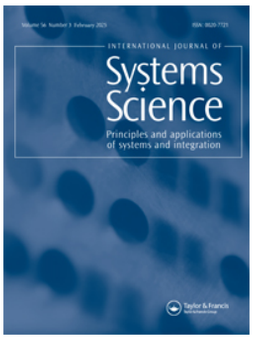

International Journal of Systems Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tsys20

MIMO system identification and uncertainty calibration with a limited amount of data using transfer learning

Mostafa Rahmani Dehaghani, Pouyan Sajadi, Yifan Tang, J. Akhavan & G. Gary Wang

To cite this article: Mostafa Rahmani Dehaghani, Pouyan Sajadi, Yifan Tang, J. Akhavan & G. Gary Wang (2025) MIMO system identification and uncertainty calibration with a limited amount of data using transfer learning, International Journal of Systems Science, 56:3, 598-617, DOI: 10.1080/00207721.2024.2408526

To link to this article: https://doi.org/10.1080/00207721.2024.2408526

INTERNATIONAL JOURNAL OF SYSTEMS SCIENCE
2025, VOL. 56, NO. 3, 598–617
https://doi.org/10.1080/00207721.2024.2408526

Taylor & Francis
Taylor & Francis Group

MIMO system identification and uncertainty calibration with a limited amount of data using transfer learning

Mostafa Rahmani Dehaghani [a], Pouyan Sajadi[a], Yifan Tang[a], J. Akhavan[b] and G. Gary Wang[a]

[a]School of Mechatronic Systems Engineering, Simon Fraser University, Burnaby, Canada; [b]Department of Mechanical Engineering, Stevens Institute of Technology, Hoboken, NJ, USA

ABSTRACT

Multiple-input multiple-output (MIMO) systems are fundamental in numerous advanced engineering applications, from aerospace to telecommunications, where precise system identification is critical for optimal performance. However, the identification of such systems often faces significant hurdles due to data scarcity, with existing approaches typically requiring substantial amounts of data for effective training. Addressing this challenge, this paper introduces a novel transfer learning framework designed specifically for MIMO system identification under conditions of limited data and inherent uncertainties. The proposed framework is applied to two case studies: the first in metal additive manufacturing, specifically the laser-blown powder-directed energy deposition as the source domain and the laser hot wire-directed energy deposition as the target domain, and the second involving a nonlinear case study of a continuous stirred-tank reactor (CSTR) with a temperature-dependent reaction. The results underscore the framework's effectiveness in capturing the dynamics of the target systems, including the ability to effectively model nonlinear dynamics. Comparative analyses highlight the benefits of employing dimensionless numbers in dynamic system modelling, offering reduced dimensionality, more physical meaning, and increased model accuracy. Overall, the proposed framework presents a promising approach to enhance system identification in MIMO systems with limited data and uncertainties, with potential applications across diverse domains.

ARTICLE HISTORY
Received 21 February 2024
Accepted 20 September 2024

KEYWORDS
Transfer learning; dimensional analysis; multiple-input multiple-output systems; system identification; metal additive manufacturing; uncertainty calibration

1. Introduction

Multiple-input multiple-output (MIMO) systems, characterized by their multiple inputs and outputs, are integral in a wide range of applications from advanced control systems in aerospace engineering to complex data processing in telecommunications (W. Wang, 2007; X. Wang et al., 2019a). Efficient and accurate modeling of MIMO systems, known as MIMO system identification, is crucial as it allows for enhanced understanding, prediction, and control of complex processes in various engineering and scientific fields. There are different methods of system identification in the literature including autoregressive models with exogenous input (ARX) (Gosiewski & Paszowski, 2003), Hammerstein-Wiener models (Ławryńczuk, 2015; Naitali & Giri, 2016), and artificial neural networks (ANN) (Chu et al., 1990; Dong et al., 2023). However, the main challenge with these system identification methods is that they require a substantial and rich dataset for effective application. Conventional system identification techniques often fall short in performance when faced with limited data availability or when significant uncertainty is present in the system.

Transfer learning (TL) is a strategy designed to enhance the performance of learners on specific tasks within target domains by leveraging knowledge acquired from related source domains (Zhuang et al., 2021). Inspired by human's ability to transfer knowledge across different domains, TL seeks to optimise learning outcomes or minimise the need for labelled examples in a target domain by drawing on knowledge from a source domain. TL is particularly valuable when dealing with a limited amount of target data, where traditional machine learning (ML) techniques may struggle due to data scarcity such as in the case of training models for metal additive manufacturing (AM) processes (Tang et al., 2023). It's important to note that the effectiveness of transferred knowledge isn't always guaranteed to be positive, as little commonality between domains can result in unsuccessful

CONTACT Mostafa Rahmani Dehaghani mra91@sfu.ca, mostafa_rahmani_dehaghani@sfu.ca Room 4245, 13450 102nd Avenue, Surrey, British Columbia, V3T0A3, Canada

knowledge transfer, known as negative transfer (Z. Wang et al., 2019b). Due to the inherent complexity of the MIMO systems, the data required for system identification should be both rich and abundant (Xiao et al., 2011). However, in scenarios where data gathering and acquisition are time-consuming and costly, and a dataset from a similar system is available, TL is considered an excellent option for MIMO system identification.

While TL has demonstrated promising results in modelling systems with a limited amount of data across various fields, such as fault diagnosis (Han et al., 2021), defect detection (Zhang et al., 2022), medical imaging (Alzubaidi et al., 2021), and AM (Tang et al., 2023), its application in MIMO system identification has been limited. Niu *et al.* innovatively applied TL, specifically parameter fine-tuning and freezing using long-short term memory neural networks, to reduce data and computation requirements for system identification. Their work focused on second-order linear and Wiener-Hammerstein nonlinear systems (Niu et al., 2022). Another noteworthy study by Tsoi et al. (2018) applied TL to a frictionless ball and beam system for system identification of a friction ball and beam system using neural networks. However, these TL techniques have primarily been applied to single-input single-output (SISO) systems. Extending them to MIMO systems presents challenges, mainly due to the significant data requirements for training neural networks. This challenge becomes more pronounced as the necessary data to train neural networks increases exponentially with the growing number of inputs.

Refining the accuracy and robustness of dynamic system models hinges on addressing inherent uncertainties associated with both the systems and their experimental data (Milanese & Vicino, 1991). For example, in the realm of metal AM processes, these uncertainties are particularly pronounced, with parameters like laser power (LP), travel speed (TS), and laser spot diameter exhibiting temporal fluctuations during fabrication (Moges et al., 2019; Olleak & Xi, 2020). Notably, such fluctuations can lead to significant deviations, such as a 20% reduction in LP from the set value (King et al., 2014). To ensure the development of robust models, it is imperative to quantify these parameter uncertainties. Ritto *et al.* pioneered an approach using approximate Bayesian computation and reinforcement learning for parameter identification and uncertainty calibration. They employed an initial uncertainty distribution, which is iteratively updated based on the model's success in reproducing the reference data (reinforcement learning strategy) (Ritto et al., 2022). While effective in calibrating parameter uncertainties, their method demands substantial data, raising questions about its applicability when a limited amount of data is available.

For systems with multiple inputs, dimensional analysis serves as a method for dimensional reduction by defining dimensionless parameters (π), which helps to reduce the demand on training data. In 1914 Buckingham proposed the π-theorem, laying the foundation for dimensional analysis theory (Buckingham, 1914). Given its proficiency in dimension reduction and generation of interpretable dimensionless parameters, dimensional analysis finds widespread applications in various domains such as heat transfer, fluid dynamics, and manufacturing processes. Researchers have also sought to define and utilise dimensionless parameters in AM modelling since a multitude of input parameters is involved (Kazmer & Colon, 2020; Marmarelis & Ghanem, 2020; Zhao et al., 2021). For instance, Rubenchik et al. (2018) characterised the temperature distribution of components in selective laser melting using two dimensionless parameters: normalised enthalpy and the ratio of dwell time to thermal diffusion time. Additionally, Mukherjee *et al.* examined the effects of various process variables and alloy properties on the structure and properties of additively manufactured parts using four dimensionless numbers: dimensionless heat input, Peclet number, Marangoni number, and Fourier number (Mukherjee et al., 2017). The combination of TL and dimensional analysis holds promise for addressing challenges in MIMO system identification, where the data scarcity poses significant hurdles.

The motivation for this study arises from the substantial challenges associated with identifying MIMO systems, which typically require extensive, high-quality datasets to accurately capture their complex dynamics. Often, the available data are markedly insufficient – both in quantity and richness – making traditional system identification methods ineffective. While TL approaches are commonly applied to SISO systems, their extension to MIMO systems often requires relatively abundant data due to the complexity and large number of parameters involved. Recognising this limitation, our research aims to develop a robust TL framework specifically designed for MIMO systems under

conditions of severe data scarcity. By focusing on this underexplored area, the study aims to facilitate more accurate and reliable system identification when data scarcity is a critical constraint in MIMO systems. The developed framework incorporates dimensional analysis for dimension reduction and calibrates the inherent uncertainties within the MIMO input parameters. This ensures enhanced agreement between the system identification models and the experimental data. Such an approach significantly improves the practical applicability of MIMO models in real-world scenarios, where data often exhibit imperfections or are incomplete.

This study also pioneers the application of transfer learning for MIMO system identification within metal AM processes and extends the framework's applicability to nonlinear systems. We demonstrate the framework's effectiveness through two case studies: the first involves metal AM, where source data is collected from a laser-blown powder directed energy deposition (DED-LB/P) process, and target data from a laser hot wire directed energy deposition (DED-LB/W) process. Metal AM serves as a suitable domain for applying the framework due to prevalent challenges such as inherent uncertainty and data scarcity (Tang et al., 2023). Additionally, the temporal fluctuations observed in process parameters during the production of metal AM parts highlight the importance of calibrating uncertainties for dynamic system modelling (Moges et al., 2019; Olleak & Xi, 2020). Furthermore, it has been observed that the dynamic relationship between the process parameters in metal AM processes and the resulting printing characteristics can be effectively modelled using linear models, making linear model system identification a suitable choice for this domain (Rahmani Dehaghani et al., 2024). The second case study involves a nonlinear system: a continuous stirred-tank reactor (CSTR) with a first-order, temperature-dependent reaction, chosen to assess the framework's generalizability to nonlinear dynamics. It is to be noted that the model construction methods explored in this study are restricted to linear models. However, nonlinear systems that can be approximated by linear models or linearised models around their operating points can also be effectively modelled, as demonstrated by the second case study (Bradshaw, 1978).

The rest of the paper is organised as follows: Section 2 provides a detailed explanation of the TL framework for MIMO systems. In Section 3, the proposed approach is applied to Case Study 1, focusing on metal additive manufacturing. Section 4 presents the application to Case Study 2, involving a nonlinear CSTR. The results are discussed in Section 5, and the conclusion is drawn in Section 6.

2. TL framework for MIMO systems

The MIMO system TL framework is explained in detail in this section. The framework is inspired by the iterative uncertainty calibration framework developed by Rahmani Dehaghani et al. (2022). Their framework is used for uncertainty and model calibration of static systems and not meant for dynamic time-series systems or processes. An overall description of the proposed approach is presented first and each subsection of the framework is explained afterwards.

The flowchart of the framework can be observed in Figure 1. The framework starts by defining the source and target data and modelling objectives (outputs). The dimensional analysis is then applied using Buckingham-II theorem to create the dimensionless output and input numbers. The data is then preprocessed which involves the normalisation of the input and output dimensionless numbers to be within 0 and 1.

The next step is to perform the system identification on the source data to find the most accurate yet simple model to fit on the source data. This step includes calculating the fast Fourier transformation (FFT) of outputs and inputs and developing the numerical Bode plots of the input-output relationships to find how many poles and zeros exist in the system for each input-output pair. Afterwards, a model is constructed using the dimensionless numbers, and the models are compared based on their R-squared, normalised root mean square error (NRMSE), and mean absolute percentage error (MAPE) calculated on the validation source data, and the model with the highest accuracy is selected as the model to be calibrated by the target data.

An initial value is assumed for the uncertainty calibration variables (UCVs) and a bias is calibrated between the selected model in the previous step and the training target data. UCVs are the standard deviations of the dimensionless numbers which are calibrated in the framework. The assumed uncertainty is then calibrated by minimising the autocorrelation of

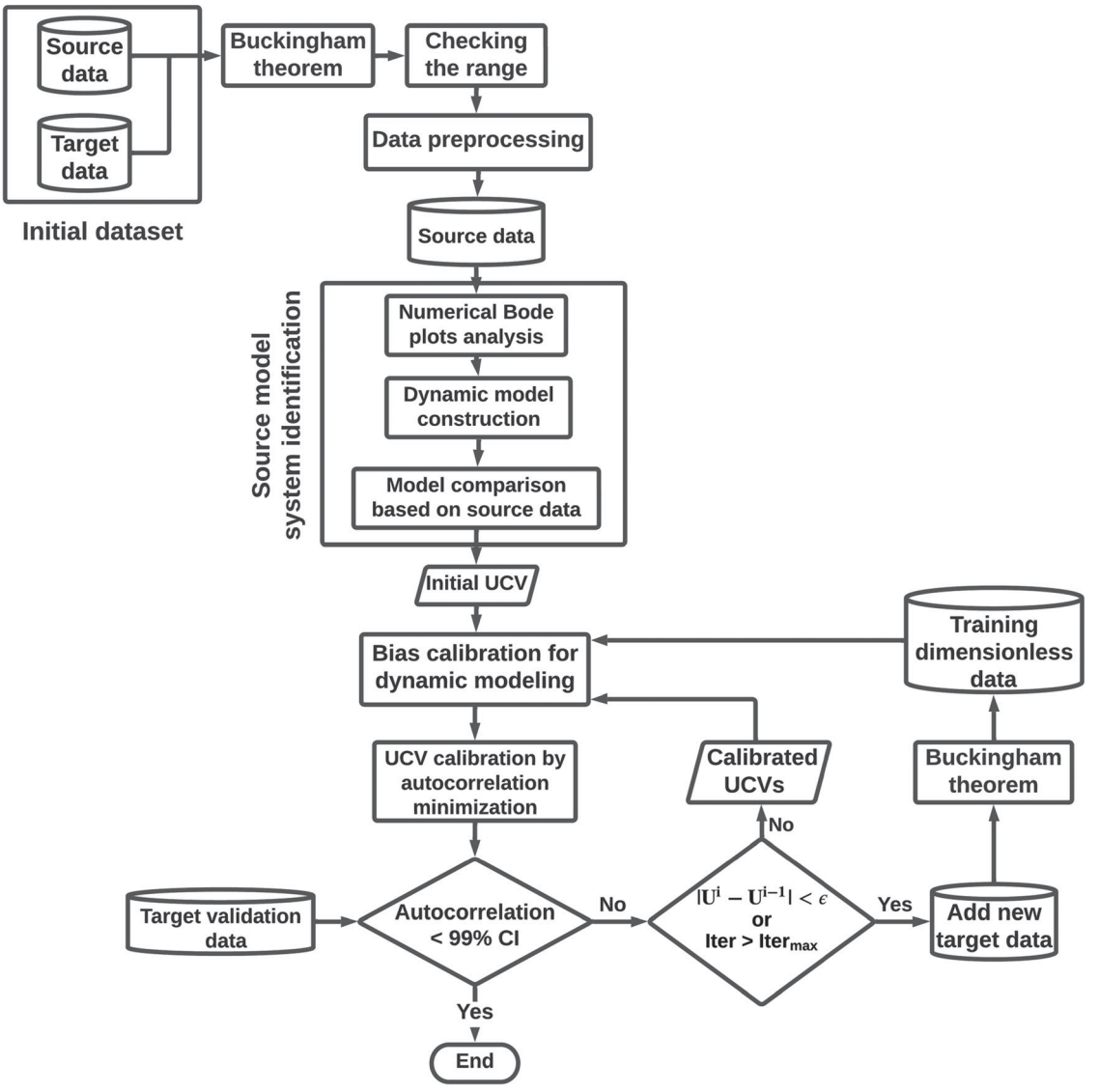

Figure 1. MIMO system TL and uncertainty calibration framework flowchart.

the residuals of the output. To validate the calibrated model with the calibrated uncertainty, the autocorrelation of the residuals of the output of the model should lie in the 99% confidence interval (CI). If not, the convergence of the uncertainty calibration variables (UCVs) is checked. If the convergence or stopping criteria are not met, the algorithm uses the calibrated UCVs for the model calibration. Otherwise, the algorithm calls for more target data to calibrate the model. In the following subsections, the main steps of the framework are explained in more detail.

2.1. Developing dimensionless numbers (π)

After collecting both source and target data, the initial step involves the development of dimensionless numbers. Dimensional analysis not only aims to reduce the system's dimensionality but also to introduce dimensionless numbers that hold physical significance in both the source and target domains. The underlying assumption is that the outputs are mutually independent, and each output can be described or controlled independently of the others using the system's inputs. Under these conditions, the MIMO system can effectively be decomposed into multiple MISO (multiple-input single-output) systems. This conversion of a MIMO system with 'm' inputs and 'n' outputs to the 'm' MISO systems occurs by defining each output $y_j(t)$ as a function of the inputs alone ($y_j(t) = f(u_1, u_2, u_3, \ldots, u_n)$), where 'j' ranges from 1 to 'm'. This means that a dynamic system with 'n' inputs and 'm' outputs can be expressed through 'm' separate systems,

each consisting of 'n' inputs and a single output. Consequently, the number of MISO systems is equal to the number of outputs (m). At the conclusion of the framework, the j-th output of the MIMO system can be modelled or controlled by the corresponding j-th MISO system. At this stage, the objective is to formulate dimensionless numbers for each of these 'm' MISO systems. Once dimensionless numbers are established for one system, the same process is applied to each subsequent system.

Applying the Buckingham-Π theorem, the number of dimensionless parameters for such a system is given by 'n + 1-k,' where 'k' represents the number of basic dimensions and 'n + 1' accounts for the total variables involved (comprising 'n' inputs and one output). In simpler terms, the development of dimensionless numbers effectively eliminates 'k' variables from the problem.

To identify these dimensionless numbers, 'k' variables serve as the repeating variables, meaning that the number of repeating variables is equal to the number of basic dimensions in the problem. By combining each of the 'n + 1 - k' remaining variables with the repeating variables, a Π group is formed. The exponents of the repeating variables are then determined by setting the total exponent of each dimension in the Π group to zero.

For instance, if the system features five inputs $(x_1, x_2, x_3, x_4, x_5)$ and one output (y), and if all the input dimensions can be expressed using three fundamental dimensions (k = 3), such as length (L), time (T), and mass (M), the entire system's behaviour can be captured by three dimensionless groups (n + 1 - k = 3). The repeating variables, in this case, are x_1, x_2, and x_3, leading to the formation of three distinct dimensionless groups: (x_1, x_2, x_3, and y), (x_1, x_2, x_3, and x_4), and (x_1, x_2, x_3, and x_5). The exponent of the non-repeating variable is set to one, and the exponents of the repeating variables within each Π group are computed to ensure dimensional consistency in the resulting dimensionless numbers. It is worth noting that one can choose different repeating variables as long as satisfying the following requirements:

- Repeating variables should be selected from the list of input parameters (independent variables).
- All of the basic dimensions must be represented in the repeating variables.
- Two repeating variables cannot have the same dimension.
- Repeating variables cannot be dimensionless.

The Π group that includes the dependent parameter (the output) is designated as the output dimensionless number, while the remaining dimensionless numbers are recognised as the input dimensionless numbers. Once the Π numbers are determined, it is advisable to delve into the endeavour of comprehending the physical significance of these Π numbers. This understanding aids in deciphering the implications of the numerical values in subsequent stages of the analysis.

2.2. Source model system identification

The next step is to construct the model for the source system. This source model is exploited as the base model for the target system. The procedure of calibrating this model on the target data is explained in Section 2.3. If a system has n inputs and one output, the Laplace transformation of the output can be then formulated as.

$$\pi_{output}(s) = G_1(s)\pi_1(s) + G_2(s)\pi_2(s) + G_3(s)\pi_3(s) + \ldots + G_{n-k}(s)\pi_{n-k}(s) \quad (1)$$

where $G_i(s)$ is the transfer function between the i^{th} dimensionless input number (π_i) and the output (π_{output}). $\pi_{output}(s)$ and $\pi_i(s)$ are the Laplace transformation of the output and the i^{th} input dimensionless numbers, respectively. The objective of Section 2.2 is to find all of the G_is for a given MISO source system, referred to as source model system identification, a process comprising two distinct stages. Firstly, it entails an analysis of the behaviour of the MISO system, including determination of the number of poles and zeros. Subsequently, the task involves constructing multiple models and select the model demonstrating the highest accuracy on the source validation data. The succeeding subsections provide detailed explanations for each of these stages.

2.2.1. Numerical Bode plots

In a continuous system, understanding the system's behaviour often involves examining the Bode plot of the system's transfer function. In cases where the transfer function of the source system remains undetermined, Bode plots for each input and the output can

be generated using FFT. For a given MISO system, the Bode plot can be represented through Equations 2 and 3. Equation 2 yields the gain in decibels (dB), while Equation 3 provides the phase angle in degrees.

$$G = 20\log_{10}\frac{FFT(\pi_{output}(t))}{FFT(\pi_i(t))} \tag{2}$$

$$Phase = \angle\left|\frac{FFT(\pi_{output}(t))}{FFT(\pi_i(t))}\right| \tag{3}$$

In Equations 2 and 3, $\pi_{output}(t)$and $\pi_i(t)$ are the output and i^{th} input time series data, respectively ($1 \leq i \leq n-k$). Furthermore, the function FFT(.) is employed to efficiently compute the Discrete Fourier Transform (DFT) from the provided time series data (Cooley & Tukey, 1965). DFT is the frequency response of the time series data and the ratio, $\frac{FFT(\pi_{output}(t))}{FFT(\pi_i(t))}$, behaves like the transfer function between the i^{th} input and the output of the system in the frequency domain. Hence, the gain (Equation 2) and phase (Equation 3) of this ratio are plotted and used to find out the overall behaviour of the system. Plotting the gain and phase of the ratio, $\frac{FFT(\pi_{output}(t))}{FFT(\pi_i(t))}$, versus the frequency is commonly known as the Bode plots of the system. It's worth noting that these Bode plots are calculated using discrete data, hence they are referred to as numerical Bode plots. Determination of the system's number of poles and zeros is then achieved by observing how many times the gain decreases and increases, respectively. Once the poles and zeros are understood, the appropriate transfer function is fit to the data, a process to be elaborated on in the subsequent subsection.

2.2.2. Model construction and comparison

The generic form of the transfer function between each input and the output is represented by Equation 4.

$$\begin{aligned} G_i(s) &= \frac{\pi_{output}(s)}{\pi_i(s)} \\ &= e^{-T_d s}\frac{\begin{array}{c} a_p s^p + a_{p-1}s^{p-1} + a_{p-2}s^{p-2} \\ + \ldots + a_1 s + a_0 \end{array}}{\begin{array}{c} b_o s^o + b_{o-1}s^{o-1} + b_{o-2}s^{o-2} \\ + \ldots + b_1 s + 1 \end{array}} \end{aligned} \tag{4}$$

In Equation 4, $G_i(s)$ is the transfer function between the i^{th} input and the output ($1 \leq i \leq n-k$). $\pi_{output}(s)$ and $\pi_i(s)$ are the Laplace transformation of the output and the i^{th} input dimensionless numbers, respectively. The exponential term, $e^{-T_d s}$, is the delay term that shows the delay between the input and the output and T_d is the delay time in seconds. p and o are the orders of the nominator and denominator, respectively. The order of the system is also equal to o. It is worth noting that the necessary condition to have a stable system is that $o \geq p$.The purpose of this step is to find $T_d, a_p, a_{p-1}, a_{p-2}, \ldots, a_0, b_o, b_{o-1}, b_{o-2}, \ldots,$ and b_1 which are called model parameters. It is worth mentioning that the values of p and o are determined in the previous section by the numbers of poles and zeros, which are called p^* and o^*, respectively.

The source dataset is divided into separate validation and training datasets. The training dataset serves as the basis for constructing the models, while the validation dataset is employed for comparing and selecting the models. In pursuit of enhancing the reliability of system identification, various choices of p and o in proximity to the values p^* and o^* are explored, resulting in the creation of different models. These models are evaluated based on their R-squared, NRMSE, and MAPE values with respect to the source validation dataset; the models exhibiting the highest accuracy are chosen as the source models.

2.3. MIMO system bias calibration

Prior to calibrating the models, it is essential to define UCVs and assign initial values to each of them. The dimensionless input numbers, as described in Section 2.1, exhibit uncertainties stemming from the inherent uncertainties in the original input parameters. Consequently, it is postulated that the dimensionless input numbers conform to a Gaussian distribution, with mean values derived from the original input parameters and standard deviations referred to as UCVs. Equation 5 formulates the UCVs as follows:

$$\boldsymbol{UCVs} = (\mu_{\pi 1}, \mu_{\pi 2}, \mu_{\pi 3}, \ldots, \mu_{\pi n-k}) \tag{5}$$

where, $\mu_{\pi i}$ represents the standard deviation of the i^{th} dimensionless input number. An initial vector of UCVs is assumed prior to the bias calibration. This initially-assumed uncertainty is calibrated in Section 2.4.

The objective of this phase is to calibrate the source models using the target data. The bias calibration approach employed in this study falls under the category of model-based TL techniques, designed to streamline the construction and training of target

models by leveraging source model structures or parameters (Tang et al., 2023; Zhuang et al., 2021).

In the proposed method, the assumption is made that both the source and target domains share the same feature space, and the TL approach employed is of the homogeneous type (Weiss et al., 2016). This assumption needs validation during the model construction. If the application of the proposed TL approach results in negative transfer, it indicates a deviation from the assumption. In such cases, other heterogeneous TL approaches, such as feature space adaptation, should be considered (Day & Khoshgoftaar, 2017; Sugiyama et al., 2008). In this work, the source model has a structure described in Equation 4. It is assumed that the target domain has a similar structure, implying that the model's order and the number of poles and zeros are the same in both the source and target systems. As a result, the task is to adjust the model parameters to increase the accuracy of the model on the target training data without changing the structure of the models. Hence, in Equation 6, $\tau_d, \alpha_p, \alpha_{p-1}, \ldots, \alpha_1, \beta_o, \beta_{o-1}, \ldots, \beta_1$ are the bias parameters ($\boldsymbol{\theta}$)that should be found during the bias calibration to construct the transfer function between each dimensionless input and output number. According to Equation 6, the total number of bias parameters is '1 + p + o'.

$$G_i(s) = \frac{\pi_{output}(s)}{\pi_i(s)} \tag{6}$$
$$= e^{-(T_d+\tau_d)s} \frac{\begin{array}{c}(a_p+\alpha_p)s^p + (a_{p-1}+\alpha_{p-1})s^{p-1} \\ +(a_{p-2}+\alpha_{p-2})s^{p-2} + \ldots \\ +(a_1+\alpha_1)s + a_0\end{array}}{\begin{array}{c}(b_o+\beta_o)s^o + (b_{o-1}+\beta_{o-1})s^{o-1} \\ +(b_{o-2}+\beta_{o-2})s^{o-2} + \ldots \\ +(b_1+\beta_1)s + 1\end{array}}$$

The process to find the bias parameters involves an optimisation problem formulated in Equation 7, where, N is the number of training samples from the target dataset, $\hat{\pi}_{output,t}(\boldsymbol{\theta})$ is the dimensionless output number predicted by the model at time t, and $\bar{\pi}_{output}$ is the mean value of the training samples from the target dataset. It's essential to note that the bounds for optimisation parameters, denoted as $\boldsymbol{\theta}_{min}$ and $\boldsymbol{\theta}_{max}$, must be chosen judiciously to ensure the stability of the final model. For instance, poles located on the right side of the complex number plane should be avoided, as they lead to unstable models. If the system's stability cannot be readily inferred from the parameter range, additional stability criteria, such as the Lyapunov stability theorem, can be incorporated as constraints in Equation 7 (Lyapunov, 1992).

$$\boldsymbol{\theta}^* = Arg\max\left(1 - \frac{\sum_{t=1}^{N}(\hat{\pi}_{output,t}(\boldsymbol{\theta}) - \pi_{output,t})^2}{\sum_{t=1}^{N}(\pi_{output,t} - \bar{\pi}_{output})^2}\right) \tag{7}$$

Subject to:

$$\boldsymbol{\theta}_{min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{max}$$

2.4. MIMO system uncertainty calibration

The subsequent step involves the calibration of the uncertainties that were assumed prior to the bias calibration. Uncertainty calibration is performed with the aid of residual autocorrelation of the output. In the system identification theory, a necessary condition of a perfect estimation is the absence of autocorrelation in the residuals of the model (Box & Pierce, 1970; Monti, 1994). The UCVs are calibrated to minimise the residual autocorrelation of the predicted models. Autocorrelation for time-series data, specifically the residuals of the output, can be computed at various time lags. In time series analysis, a 'lag' is a fixed amount of passing time between two points in the data. When analysing autocorrelation in time series data, a lag is used to measure how the data values at one point in time are related to values at another point. Equation 8 shows the autocorrelation calculation at lag h (Chatfield, 2013). In this notation, R_h denotes the degree of correlation between a data point in a time-series and its value 'h' time steps earlier. The value of 'R_h' falls within the range of –1–1, with zero indicating no autocorrelation. It's important to note that 'R_0' is invariably 1. In this study, autocorrelation is computed for lags ranging from 0 to 25 ($0 \leq h \leq 25$), and the examined time series data is the residual of the dimensionless output number of the training target dataset. In Equation 8, N is the total amount of data in the training target dataset and e represents the residual.

$$R_h = \frac{C_h}{C_0} \tag{8}$$

where,

$$C_h = \frac{1}{N}\sum_{t=1}^{N-h}(e_t - \bar{e})(e_{t+h} - \bar{e})$$

The underlying hypothesis is that altering the distribution of the dimensionless input numbers will result in a corresponding change in the residual autocorrelation of the predictive model. The purpose is to find the UCVs that leads to the lowest residual autocorrelation. To achieve this objective, the maximum value of the residual autocorrelations among the first 25 lags $(R_1, R_2, R_3, \ldots, R_{25})$ is minimised. This optimisation problem is formulated in Equation 9. The user should determine the bounds for the UCVs, and these bounds hinge on the range of variation for the corresponding dimensionless number. It's imperative to note that, irrespective of the problem, the bounds should not fall below zero, given that standard deviations are inherently positive numbers.

$$\boldsymbol{UCVs^*} = Argmin(\max\{R_1, R_2, R_3, \ldots, R_{25}\}) \quad (9)$$

Subject to:

$$\vec{0} \leq \boldsymbol{UCVs}_{min} \leq \boldsymbol{UCVs} \leq \boldsymbol{UCVs}_{max}$$

2.5. Validation hypothesis test

The next step is to check the validity of the calibrated model and UCVs on the validation target dataset. The hypothesis test checks if the residual autocorrelation falls within the CI for all the lag values. The calculation of the CI for the significance level α is presented in Equation 10 (Chatfield, 2013).

$$CI = \left[-\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{N}}, +\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{N}}\right] \quad (10)$$

In Equation 10, N is the sample size and z is the cumulative distribution function of the standard normal distribution. The residual autocorrelation for different lags on the target validation dataset is calculated and checked if it falls within the CI. The validation process is terminated if all the autocorrelations lie within the CI, otherwise, the proposed approach uses the calibrated UCVs as the initial guess to perform bias calibration (Section 2.3). Then, the new calibrated model and UCVs are used to conduct the validation hypothesis test. This iterative process continues until the change in the UCVs in two consecutive iterations, is less than ε ($|UCV^i - UCV^{i-1}| < \varepsilon$), i.e. a convergence criterion, or the number of iterations exceeds the maximum number of iterations ($Iter > Iter_{max}$). If the convergence criterion or stopping criteria are met and the framework still cannot generate a valid model, it indicates the need for additional target training data. This iterative strategy has been shown to ensure the validity and accuracy of the final model not contingent on the initially-assumed UCVs (Rahmani Dehaghani et al., 2022).

3. Case study 1: melt pool width modelling in DED processes

The described framework for TL and uncertainty calibration in dynamic MIMO systems can be universally applied, provided certain assumptions hold true. Firstly, the MIMO system must be representable as a linear time-invariant system; secondly, it is assumed that the outputs of the system are mutually independent, allowing an 'm'-output MIMO system to be treated as 'm' separate MISO systems. In this study, the framework is specifically applied to the field of metal AM. The source dataset is compiled by the Stevens Institute of Technology, encompassing the DED-LB/P process (Akhavan et al., 2023). The target data, on the other hand, is acquired from experiments conducted at the University West in Sweden, specifically focusing on DED-LB/W process (Rahmani Dehaghani et al., 2024). The input process parameters of the deposition process include the deposition rate (DR), laser power (LP), and travel speed (TS), which are systematically varied during the part deposition. Stainless steel 316L (SS316L) powder is employed in the DED-LB/P process, while duplex stainless steel 2209 (DSS2209) wire with a 1.2 mm diameter is used in the DED-LB/W process. In order to incorporate material characteristics and properties into the modeling, two additional input parameters, density (ρ) and latent heat of fusion (L_f), are considered. Consequently, a total of five input parameters are involved in the dimensional analysis, denoted by 'n = 5.'

Both of the source and target data captured coaxial images of the melt pool during the deposition of the parts. Melt pool width (MPW) is calculated from the images as the output of the MISO system. Table 1 reports the range of all the input and output parameters in the source and target dataset. The dimensions of the parameters are provided in the third column of Table 1 where L, T, M, and θ represent length, time, mass, and temperature, respectively. Analysing the dimensions of the input and output parameters, it becomes evident that the fundamental dimensions

Table 1. Input, output, and auxiliary parameters, their dimensions, and ranges in source and target datasets for case study 1.

Parameter type	Parameter	Dimension	Parameter range	
			DED-LW/*P*	DED-LW/W
Input($n = 5$)	Deposition rate (gr/s), *DR*	$[M][L^{-1}]$	0.083–0.158	0.265–0.323
	Travel speed (mm/s), *TS*	$[L][T^{-1}]$	5–10	8–12
	Laser power (W), *LP*	$[M][L^{2}][T^{-3}]$	300–500	2800–3200
	Density (gr/mm^3), ρ	$[M][L^{-3}]$	7.98e-3 (Azom, n.d.)	7.80e-3 (Matweb, n.d.-a)
	Latent heat of fusion (J/gr), L_f	$[L^{2}][T^{-2}]$	272 (Azom, n.d.)	264 (Matweb, n.d.-a)
Output	Melt pool width (mm), *MPW*	$[L]$	0.65–0.95	4.7–5.4
Auxiliary	Specific heat capacity, C_p(J/°Cgr)	$[L^{2}][T^{-2}][\theta^{-1}]$	0.5 (Matweb, n.d.-b)	0.48 (Matweb, n.d.-a)
	Melting point (°C), T_{melt}	$[\theta]$	1425 (Matweb, n.d.-b)	1375 (Matweb, n.d.-a)

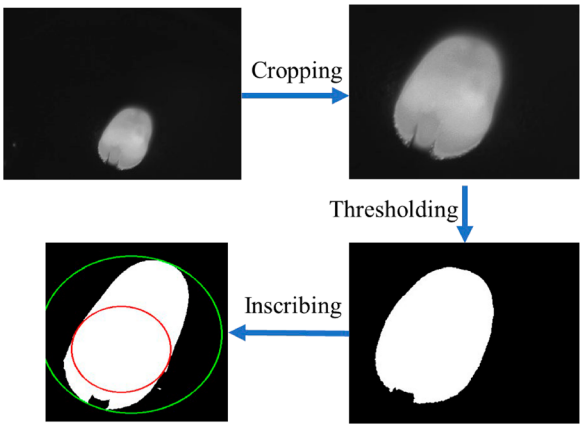

Figure 2. Image processing procedure to find the MPW (Rahmani Dehaghani et al., 2024).

required for the dimensional analysis consist of three key dimensions, encompassing length, time, and mass (k = 3).

The process of extracting the MPW from the coaxial images involves several key steps, including image cropping, thresholding, and the placement of an inscribed circle within the melt pool. The diameter of the largest inscribed circle is then selected as the MPW. A detailed explanation of this procedure can be found in Reference (Rahmani Dehaghani et al., 2024). For a visual depiction of the image processing steps, please refer to Figure 2, where the diameter of the smaller circle is chosen as the MPW. In the DED-LB/P process, a total of 30 beads are printed, and one process parameter is altered three times during the deposition of each individual bead. The data from the different beads are concatenated, creating one continuous time series data. A parallel concatenation process is implemented for the target data stemming from the DED-LB/W process. However, in this case, only the data from the printing of 3 beads is utilised as the target data, and one process parameter is altered twice during the deposition of each bead.

3.1. Dimensionless numbers construction

After converting the parameters of source and target data to the same units and creating one continuous time series data for either of the source and target datasets, it is time to create the dimensionless numbers using the Buckingham-Π theorem. Given the total number of parameters involved in the modelling process is six (n + 1 = 6) and hence the number of dimensionless parameters is 3 (n + 1-k = 3). Applying the theorem, the dimensionless numbers are found as below:

$$\pi_1 = \frac{TS}{\sqrt{\frac{LP}{DR}}} \quad \pi_2 = \frac{\frac{LP}{DR}}{L_f} \quad \pi_3 = \frac{MPW}{\sqrt{\frac{DR}{\rho TS}}} \tag{11}$$

π_3 is the dimensionless number that includes the output and hence, it is called the output dimensionless number. Moreover, it is the ratio of the MPW to the square root of the cross-sectional area of the bead. Hence, π_3 value of one means that the height and width of the bead is quite close to each other. The other two dimensionless numbers are the input dimensionless numbers. π_1 is called $\pi_{process}$ since it is defined by process parameters. Lastly, π_2 is the ratio of the input energy from the heat source to the energy the material needs to melt. Thus, π_2 is called π_{energy}. This formulation of π_{energy} does not consider the energy that the material needs to reach the melting point. This energy is quantified in Equation 12.

$$Q = L_f + C_p(T_{melt} - T_0) \tag{12}$$

In Equation 12, C_p is the specific heat capacity, T_{melt} is the melting point of the alloy and T_0 is the ambient temperature which is 25°C. Q needs to be placed in the denominator of the π_{energy} to reflect the reality more closely. The values of C_p and T_{melt} for either of the source and target data are introduced under the Auxiliary parameters of Table 1. In conclusion,

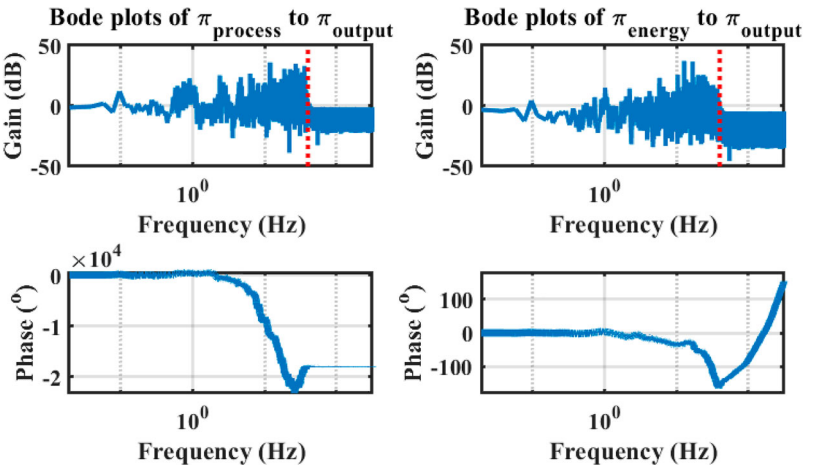

Figure 3. Numerical bode plots between dimensionless inputs and the output.

the final input and output dimensionless numbers are introduced in Equation 13.

$$\pi_{process} = \frac{TS}{\sqrt{\frac{LP}{DR}}} \quad \pi_{energy} = \frac{\frac{LP}{DR}}{L_f + C_p \Delta T}$$
$$\pi_{output} = \frac{MPW}{\sqrt{\frac{DR}{\rho TS}}} \tag{13}$$

After calculating the dimensionless number for each sample time, the dimensionless numbers are normalised using the normal standard distribution.

3.2. DED-LB/P system identification (source dataset)

The subsequent stage in the MIMO system TL framework involves the development of source models. To accomplish this, numerical Bode plots are generated for each dimensionless input and output, as depicted in Figure 3. The plots on the left side of Figure 3 illustrate the Bode plots from $\pi_{process}$ to π_{output}, while those on the right side display the Bode plots from π_{energy} to π_{output}. The magnitude plots in decibels (dB) reveal that the gain remains relatively constant, with a slight decrease around 30 Hz (dashed red line) before stabilising once more. This behaviour indicates the presence of one pole and one zero in the dynamic system. This conclusion holds true for both sets of plots, indicating that $p^* = 1$ and $o^* = 1$. It's worth noting that models with different numbers of poles and zeros around the selected points are also constructed and subsequently compared.

A time series dataset, comprising 20,228 data points sampled at 100 Hz, is created through the concatenation of data from the 30 beads in the source dataset. Within this dataset, 25% of the data is earmarked for validation, while the remaining 75% is utilised for training. Various selections of first, second, third, and fourth-order models are constructed using the MATLAB System Identification Toolbox. These models are compared based on their performances against the source validation dataset, as depicted in Figure 4. The R-squared, NRMSE, and MAPE values for each model on the validation data are reported for the purpose of comparison.

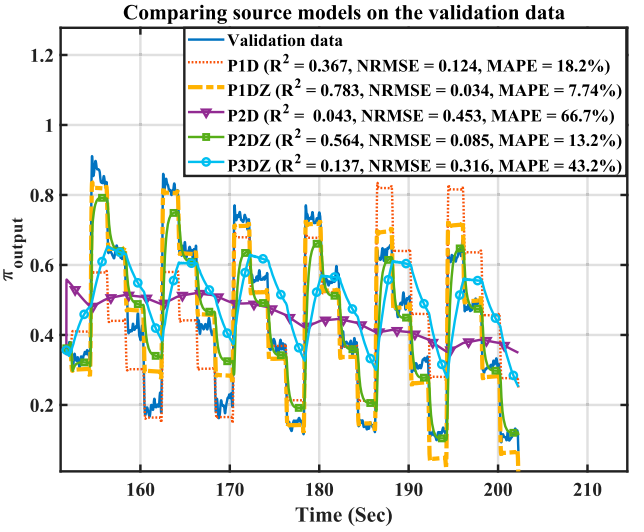

Figure 4. Source models' comparison based on the source validation dataset.

The model's characteristics, including the number of poles, the presence of a zero, and the inclusion of a delay term, are denoted by the model's name. For instance, the model 'P1DZ' represents a first-order model (P1) that includes a zero (Z) and a delay term

(D). The plots indicate that the model with one pole and one zero attains the highest accuracy, achieving an R-squared, NRMSE, and MAPE values of 0.783, 0.034, and 7.75%, respectively (the dashed orange line). This conclusion aligns with the results presented in Figure 3, where it was deduced that the model possesses one pole and one zero for both input-output pairs. The 'P1DZ' model is subsequently employed as the base model in the following sections, and it is calibrated using the target data.

3.3. Model calibration on DED-LB/W target data

Before performing the bias calibration, UCVs should be defined. With reference to Equation 5, the UCVs is defined as below:

$$UCVs = (\mu_{\pi\,process}, \mu_{\pi\,energy}) \tag{14}$$

The initial values for the UCVs are selected as UCVs = (0, 0). The P1DZ models developed in the previous step has the following structure (o = 1 and $p = 1$):

$$\pi_{output}(s) = e^{-0.011s}\frac{-0.013s + 0.74}{0.00071s + 1}\pi_{process}(s) + e^{-0.016s}\frac{0.85s + 0.66}{1.19s + 1}\pi_{energy}(s) \tag{15}$$

According to Equation 6, the bias parameters ($\boldsymbol{\theta}$) to be determined in this step consist of six parameters: three parameters ($o + p + 1 = 3$) for each of the transfer functions at the right-hand side of Equation 15. Three bias parameters are α_1, β_1,and τ_d. With the assumption that the delay between the source and target systems is equal, and therefore, the delay bias parameter (τ_d) is set to zero. As a result, the total number of parameters is reduced to four. This assumption is deemed valid since the delay is contingent on how rapidly the command is applied and the output is changed, a factor that remains consistent in both the source and target processes, given their similar command system. Consequently, models with the bias parameters ($\boldsymbol{\theta}$) can be represented as:

$$\pi_{output}(s) = e^{-0.011s}\frac{(-0.013 + \alpha_{11})s + 0.74}{(0.00071 + \beta_{11})s + 1}\pi_{process}(s) + e^{-0.016s}\frac{(0.85 + \alpha_{12})s + 0.66}{(1.19 + \beta_{12})s + 1}\pi_{energy}(s) \tag{16}$$

The bias parameters ($\alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12}$) are determined through the optimisation problem outlined in Equation 7. The target data for the DED-LB/W comes from printing three single beads, each with one input parameter change during the deposition. This scenario, reflecting a limited dataset, includes 755 data points recorded at a sampling interval of 0.03 s, which are then divided into training and validation datasets. To identify the most suitable ratio for the training dataset, a comprehensive study is conducted, and the results are illustrated in Figure 5. In this study, the training-to-whole-dataset ratio is varied from 0.1–0.9, and three distinct models are compared:

(1) The model derived from the optimisation of Equation 7, trained on the source data and calibrated on the target data, referred to as the calibrated source model (depicted by the blue line with circular markers).
(2) The source model without calibration on the target data, known as the source model (represented by the red line with triangular markers).
(3) The model trained solely on the target data, oblivious to the source data, denoted as the target model (shown by the yellow line with square markers).

The accuracy of these models is assessed based on the validation target dataset and is plotted in Figure 5. The maximum accuracy is achieved by the calibrated source model when the target training ratio is set to 0.7. Figure 5 further illustrates that training a MISO model with a limited amount of data without incorporating TL does not yield acceptable accuracies. This

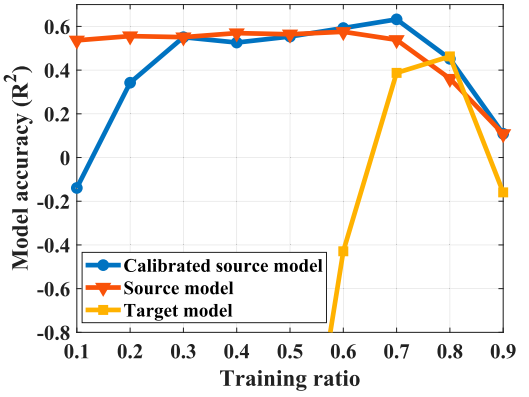

Figure 5. Performance comparison of the calibrated source model, non-calibrated source model, and target model under varying training ratios, assessed on the validation target dataset.

is evident as the target model line (yellow line with square markers) results in negative R-squared values, emphasising the necessity of TL when dealing with limited data.

The assumption that both of the source and target domains share same feature spaces holds true since in all of the training ratios the blue line with circular markers is higher than the yellow line with the square markers which indicates that the transfer is always positive. This also can be understood by taking a look at the source model accuracies (red line with triangular markers). Although this model is blind to the target data during the training it shows a relatively acceptable accuracy on the target dataset in most of the cases. It implies that the source and target domains have similar feature spaces.

According to Figure 5, 70% of the target data is used as the training dataset to perform the optimisation problem and find the bias parameters ($\boldsymbol{\theta}$). After performing the optimisation using genetic algorithm, these optimal bias parameters are found to be:

$$\begin{aligned}\theta^* &= (\alpha_{11}^*, \alpha_{12}^*, \beta_{11}^*, \beta_{12}^*) \\ &= (0.0022, 0.00011, 0.22, 1.028) \end{aligned} \tag{17}$$

Consequently, the calibrated model is representable as:

$$\begin{aligned}\pi_{output}(s) = e^{-0.011s}\frac{-0.011s + 0.74}{0.00082s + 1}\pi_{process}(s) \\ + e^{-0.016s}\frac{1.07s + 0.66}{2.22s + 1}\pi_{energy}(s)\end{aligned} \tag{18}$$

The models compared in Figure 5 are consolidated with the validation data to assess their performances on the time series data in Figure 6. It can be seen that the blue line with the circular markers (calibrated model) represents a refined version of the red line with the triangular markers (source model), evident in the closer alignment with the validation data. The R-squared of the calibrated model on the validation target data is determined to be 0.632 signifying a commendable level of performance. Furthermore, the performance of the defined models, assessed using various metrics, is compiled in Table 2. This demonstrates that the calibrated model achieves the best performance, with R-squared, NRMSE, and MAPE values of 0.0.632, 0.076, and 9.71%, respectively.

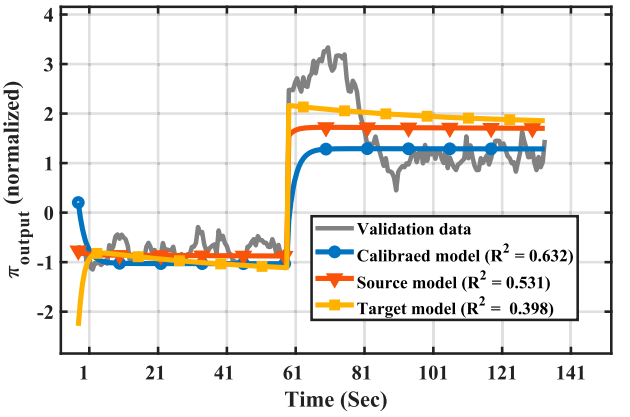

Figure 6. Performance comparison of the models on the target validation data.

Table 2. Performance metrics comparison of the models on the target system validation data.

Model name	R^2	NRMSE	MAPE (%)
Calibrated model	0.632	0.103	11.2%
Source model	0.531	0.112	13.9%
Target model	0.398	0.133	16.7%

3.4. Modelling: process parameters vs. dimensionless numbers

To see the effect of using dimensionless numbers, the same methodology is employed to construct models on the target data directly using process parameters, including LP, TS, and DR, with MPW as the output. The hypothesis posits that utilising dimensionless numbers enhances model accuracy by reducing dimensionality (from three inputs to two) and imparting more physical meaning to parameters (e.g. π_{energy} as the ratio of input energy to the energy needed for melting). This hypothesis is tested by plotting the calibrated model, source model, and the target model trained with process parameters (utilising 3 inputs and 1 output), as presented in Figure 7. Figure 7 mirrors Figure 6, with the only distinction being the absence of transformation to dimensionless parameters, using instead the usual process parameters for model construction.

Observing the behaviour of the lines in Figure 7, it becomes evident that none of the models accurately capture the process dynamics. This is further supported by the accuracy of the models, none of which surpasses $R^2 = 0.452$, NRMSE $= 0.146$, and MAPE $= 19.3\%$. These findings substantiate the hypothesis that incorporating dimensionless numbers in modelling MISO systems enhances both accuracy

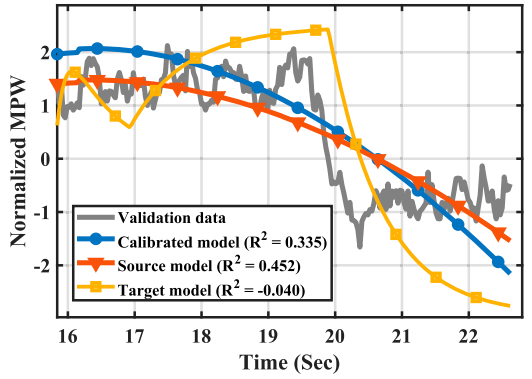

Figure 7. Comparing the models trained and evaluated using the process parameters based on the R-squared value on the validation target data.

and efficiency. An important observation in Figure 7 is that the accuracy of the calibrated model (blue line with circular markers) is lower than that of the source model (red line with triangular markers). This discrepancy indicates that the increase of dimensionality hampers the calibration process, resulting in a deterioration of model accuracy rather than improvement.

3.5. Uncertainty calibration of DED-LB/W process and model validation

The subsequent step involves calibrating the uncertainty of the parameters, as detailed in Section 2.4. Autocorrelation of the residuals is computed for the first 25 lags and minimised using the training target data through Equation 9. Figure 8 illustrates the residual autocorrelation on the validation data across different iterations. The shaded gray region between the red lines represents the CI calculated from Equation 10. Notably, the model becomes valid after the third iteration (the solid line without markers), with the exception of lags two and eight, which are marginally validated.

In each iteration, UCVs are optimised, and if the autocorrelation does not lie in the confidence interval, the optimised UCVs are utilised to perform bias and uncertainty calibration once again. This iterative process continues until a valid model is found or the convergence criterion or stopping criteria (i.e. $Iter > Iter_{max} = 10$) are satisfied. The convergence criterion, represented by dashed red line in Figure 9, is set to 0.01. Figure 9 displays the change in the absolute values of the UCVs for each iteration. The framework concludes after the third iteration, meeting the convergence criterion. This termination aligns with the point where the framework reaches the validated autocorrelation, depicted by the purple solid line without markers in Figure 8.

4. Case study 2: reactant concentration modelling in continuous stirred-tank reactor

In this section, the framework is applied to a nonlinear case study involving a CSTR with a first-order, temperature-dependent reaction. The objective is to demonstrate that the framework can be effectively applied to nonlinear systems, provided that these systems can be adequately linearised with acceptable accuracy. The focus of this case study is to model the concentration of the reactant inside the reactor, given the temperature, inlet concentration, and flow rate of the reactant as controllable inputs. The general nonlinear governing differential equation for the system is presented below (Doraiswamy, 1992; Oza et al., 2000):

$$\frac{dC(t)}{dt} = \frac{F(t)}{V}(C_{in}(t) - C(t)) - k_0 e^{-\frac{E_a}{RT(t)}} C(t) \quad (19)$$

where $C(t)$ represents the reactant concentration inside the reactor (output); $C_{in}(t)$ is the reactant concentration at the reactor inlet (controllable input 1); $F(t)$ denotes the flow rate of the reactant into the reactor (controllable input 2); V is the reactor volume; k_0 is the pre-exponential factor (frequency factor) of the reaction rate; E_a is the activation energy of the reaction; R is the universal gas constant; and $T(t)$ is the temperature inside the reactor (controllable input 3). The approach involves two distinct reactors with different volumes, reaction rates, and activation energies, with one being designated as the source reactor and the other as the target system. Source and target datasets are generated for each system using the governing equation provided. These datasets are then utilised in subsequent sections to apply the framework and evaluate its effectiveness. The involved parameters, along with their dimensions and values for the source and target systems, are described in Table 3.

To create the source and target datasets, different trends for the input parameters are applied to both the source and target systems. These input parameters, along with the nonlinear governing Equation 19, are used to generate the output data for each system. The initial reactant concentration is set at 2 mol/L, and a sampling time of 0.1 s is used. For the source system,

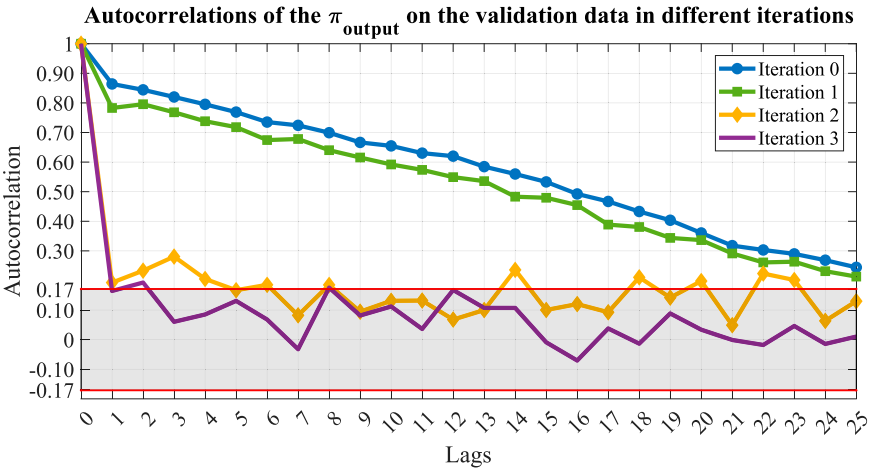

Figure 8. Residual autocorrelation on the validation data in different iterations.

Table 3. Input and output parameters, their dimensions, and ranges/values in source and target reactors for case study 2.

			Parameter range/value	
Parameter type	Parameter	Dimension	Source reactor	Target reactor
Input($n = 7$)	reactant inlet concentration (mol/L), C_{in}	$[M][L^{-3}]$	1–3	1–3
	Reactant inlet flow rate, (L/s), F	$[L^3][T^{-1}]$	20–50	30–50
	Reactor temperature (^{o} C), T	$[\theta]$	300–800	300–800
	Reactor volume (L), V	$[L^3]$	100	50
	Activation energy (J/mol), E_a	$[M][L^2][T^{-2}]$	50000	45000
	Universal gas constant (J/mol.K), R	$[M][L^2][T^{-2}][\theta^{-1}]$	8.314	
	Pre-exponential factor (1/s), k_0	$[T^{-1}]$	5000	3000
Output	reactant concentration (mol/L), C	$[M][L^{-3}]$	0.154–2.432	0.932–2.689

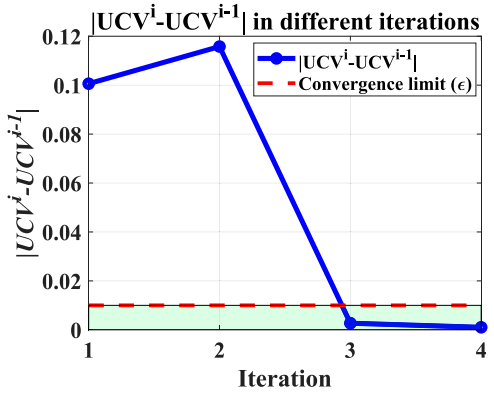

Figure 9. Convergence criterion limit and the change of the UCVs absolute value in each iteration.

10,000 data points are generated, while only 750 data points are produced for the target system to simulate the limited availability of data for the target system. The framework is then applied to this case study and to be discussed in the following subsections.

4.1. Dimensionless numbers and source CSTR system identification

Following Table 3, the number of base dimensions involved in this study is four ($k = 4$), while there are eight different parameters involved (namely, C, C_{in}, F, T, R, k_0, V, and E_a). Therefore, the number of dimensionless numbers is four ($n + 1 - k = 7 + 1 - 4 = 4$). According to the Π-Buckingham theorem, the dimensionless numbers are as follows:

$$\pi_{input1} = \frac{RT}{E_a}, \quad \pi_{input2} = \frac{V^{\frac{5}{3}} k_0^2 C_{in}}{E_a}, \quad \pi_{input3} = \frac{F}{V k_0}, \quad \pi_{output} = \frac{V^{\frac{5}{3}} k_0^2 C}{E_a} \tag{20}$$

There are three input dimensionless numbers and one output. These dimensionless numbers are calculated using the datasets from both the source and target systems and are utilised within the framework. The next step involves constructing the CSTR model for the source system. As indicated by Equation 19, the system exhibits first-order behaviour, suggesting that a first-order linear model can effectively approximate the nonlinear dynamics of the source system. To evaluate this, 75% of the source dataset is used for training, while the remaining 25% is reserved for validation. Figure 10 illustrates the performance of various models on the validation dataset. Linear models are trained on the source dimensionless dataset,

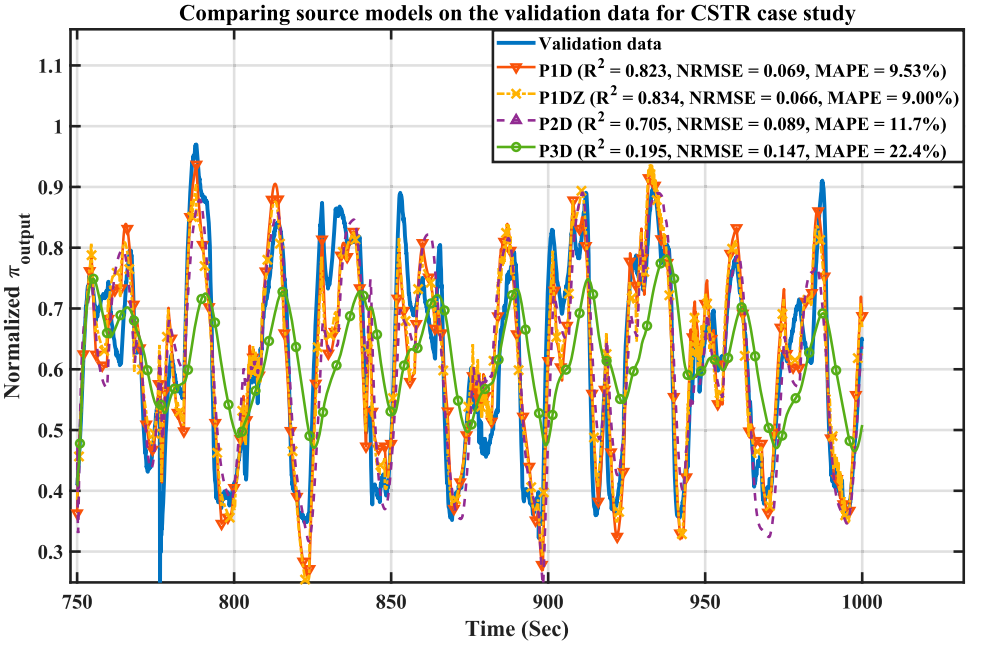

Figure 10. Source model construction comparison for the CSTR case study.

and their approximation capabilities are compared to evaluate their effectiveness in capturing the system's behaviour. Although the original system is nonlinear, it is well predicted by linear models, particularly the first-order model with one pole and one zero ($R^2 = 0.834$, NRMSE = 0.066, and MAPE = 9.00%). The 'P1DZ' model is used as the base model in the following sections and is transferred to the target dataset. Notably, since the original system is a first-order nonlinear system, it is approximated more accurately by first-order linear models than by higher-order linear models, as shown in Figure 10.

4.2. Model and uncertainty calibration on target CSTR system

In this section, the constructed source model is transferred to the target system through two key steps: bias calibration and uncertainty calibration. Equation 21 presents the constructed source model in the frequency domain. The model includes six bias parameters $(\alpha_{11}, \alpha_{12}, \alpha_{13}, \beta_{11}, \beta_{12}, \beta_{13})$, which need to be determined by optimising Equation 7.

$$\pi_{output}(s) = \frac{(-0.117 + \alpha_{11})s + 0.818}{(1.44 + \beta_{11})s + 1}\pi_{input1}(s) + \frac{(0.325 + \alpha_{12})s + 0.548}{(4.004)s + 1}\pi_{energy}(s) + \frac{(14.9 + \alpha_{13})s + 1.73}{(272.9 + \beta_{13})s + 1}\pi_{input3}(s) \quad (21)$$

The target data training ratio is set to 0.75, and optimisation is performed to determine the bias parameters. After completing the optimisation, the bias parameters are identified as follows:

$$\theta^* = (\alpha_{11}^*, \alpha_{12}^*, \alpha_{13}^*, \beta_{11}^*, \beta_{12}^*, \beta_{13}^*) = (0.302, -0.814, 0.591, -3.09, 0.283, -50) \quad (22)$$

After calibrating the bias parameters and transferring the source model to the target system, the accuracy of the calibrated model is evaluated using the validation target dataset, as shown in Figure 11. The definitions of the models in Figure 11 are consistent with those in Case Study 1. From Figure 11, it is evident that due to the system's nonlinearity and the limited amount of available data, the source and target models alone are unable to accurately approximate the real system. However, the calibrated model demonstrates significantly improved accuracy, with an R^2 of 0.956, NRMSE of 0.0865, and MAPE of 9.24%. The superior performance of the calibrated model is also apparent in the plot, where it more effectively captures the dynamics of the target system compared to both the source and target models.

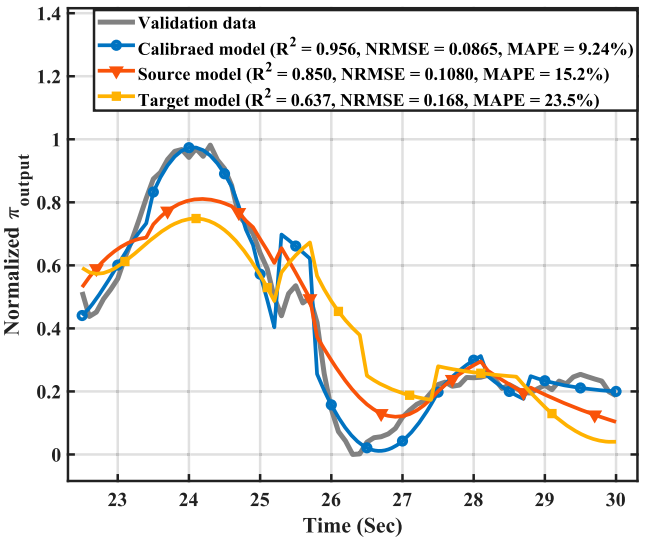

Figure 11. Performance comparison of the models on the target validation data of the CSTR nonlinear system.

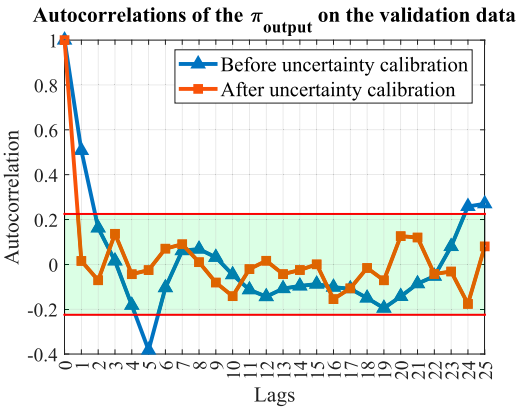

Figure 12. Autocorrelation residual on the validation target data for the first 25 lags before and after uncertainty calibration along with the CI.

The next step involves performing uncertainty calibration and final model validation using autocorrelation residual optimisation and validation. For this case study, there are three UCVs, corresponding to the three dimensionless input parameters, each requiring calibration of their standard deviation. Initially, the standard deviations of the dimensionless numbers are set to zero before the bias calibration; however, these values need to be optimised in this step. The calibrated uncertainties of the dimensionless input parameters are determined by optimising Equation 9. The upper bound values for the UCVs are set to 0.3. The optimisation is performed, resulting in the following optimal UCV parameters:

$$UCVs^* = (St.D.\pi^*_{input1}, St.D.\pi^*_{input2}, St.D.\pi^*_{input3}) = (0.163, 0.132, 0.012) \quad (22)$$

The above optimal parameters are used to compute the autocorrelation of the validation output parameters and perform the validation test. A total of 75 data points (25% of the 300 data points) are used as validation data, resulting in a CI of [−0.226,0.226] with 95% confidence, as calculated using Equation 10. Figure 12 displays the autocorrelation of the output dimensionless number on the validation data before and after uncertainty calibration, with the shaded area representing the CI. Prior to uncertainty calibration, the autocorrelation at lags 1, 5, 24, and 25 falls outside the confidence interval. However, after calibration, all autocorrelations for the different lags fall within the 95% confidence interval. Since the model is validated in the first iteration, the framework concludes without the need for further iterations.

5. Discussion

The proposed framework was applied to two case studies: the first involves the DED-LB process, and the second focuses on the reactant concentration in a CSTR with a nonlinear, temperature-dependent reaction. The framework demonstrated success in both applications, achieving high accuracy on the target validation data with $R^2 = 0.632$, NRMSE = 0.103, and MAPE = 11.21% for the DED process case study, and $R^2 = 0.956$, NRMSE = 0.0865, and MAPE = 9.24% for the CSTR nonlinear system. This performance is particularly remarkable given the following challenges:

- The studied systems are MISO systems with three input parameters, making system identification computationally demanding and complex.
- The target data is severely limited, originating from the printing of three single beads for case study 1 and 300 data points for case study 2.
- In case study 2, the system is originally nonlinear, adding further complexity to the modelling process.

In the following two subsections, the results of each case study are discussed separately, followed by an outline of the limitations of the proposed framework.

5.1. Case study 1 discussion

While the TL framework efficiently captures the dynamics of the DED-LB/W MISO system, it's essential to note that the final model exhibits marginal

validation. As seen in Figure 9, the residual autocorrelation at lags 2 and 8 slightly falls outside of the CI, raising concerns about the model's validity. To enhance the model's accuracy and agreement with the data, one potential solution is to conduct additional experiments on the target system. By iteratively incorporating new data into the framework, there is a possibility of improving the model's performance and achieving better alignment with the experimental observations.

In Section 3.4, a comparison is drawn between utilising the usual process parameters and dimensionless numbers. While there is no universal rule for choosing between these two approaches, the hypothesis that employing dimensionless numbers reduces input dimensionality for modelling, adds more physical meaning to the parameters, and consequently results in models with higher accuracies and reduced computational complexity is substantiated by the comparison made in Section 3.4. However, the decision to choose between process parameters or dimensionless numbers depends on the modelling objective. For example, although the represented model with dimensionless numbers exhibited higher accuracies compared to the usual process parameters modelling, it does not directly model the MPW but characterises the geometry of the cross-section of the printed bead (owing to the definition of π_{output}). Importantly, there is no obligation to exclusively use one type of parameters compared in this research (dimensionless parameters vs. usual parameters). For instance, one can choose MPW as the output of the modelling, employing dimensionless input numbers as inputs, allowing for the direct modelling and control of MPW.

The calibration of the uncertainty of the input dimensionless numbers aims to enhance the agreement between the developed models and experimental data. It's important to acknowledge that while this approach improves the model's fit, it doesn't quantify the uncertainties inherent in the process. Quantifying process uncertainties typically requires a substantial amount of data, especially in the field of metal AM, where various sources of uncertainties come into play (Hu & Mahadevan, 2017; Mahadevan et al., 2022).

For the bias calibration, as outlined in Section 2.3, it is crucial for users to personally define the bounds for optimisation parameters in Equation 7. Lack of familiarity with the stability principles of dynamic systems in control theory may lead to a situation where optimal bias parameters derived from Equation 7 result in unstable models, albeit accurate within the tested data range. In the specific case study presented, the bias parameter β_{11} is a critical example. Values less than –0.00071 (as per Equation 16) should be avoided to prevent the model from having a pole on the right side of the complex number plane, rendering it unstable. Without applying this constraint to the parameter bounds, the value might fall to –0.00083, causing model instability. Importantly, this instability might go unnoticed by the framework if the user fails to impose the necessary constraint, as the model continues to produce reasonable responses within the available dataset. It's important to highlight that for the models in the presented case study, which are relatively simple (first-order models), ensuring system stability by setting bounds for the bias parameters is a practical approach. However, for more complex models, especially those of higher order or nonlinearity, the inclusion of stability theorems such as Lyapunov theory as constraints to Equation 7 becomes essential.

5.2. Case study 2 discussion

The application of the framework to the CSTR with a nonlinear, temperature-dependent reaction demonstrates its effectiveness in modelling nonlinear systems, even when facing significant challenges such as limited data availability and inherent system nonlinearity. The results underscore the potential of the framework to adaptively calibrate bias and uncertainty parameters, allowing for the successful transfer of a source model to a target system with distinct characteristics. The calibration steps are crucial in refining the model to account for discrepancies between the source and target data, showcasing the framework's capacity to bridge gaps between datasets in nonlinear contexts, as shown by Figure 11.

Furthermore, although the framework shows promising results in capturing the dynamics of nonlinear systems, it is limited to systems that can be adequately approximated with linear models. In other words, the approach used in this study involves linearising the source and target systems before applying the framework. If a system is inherently nonlinear and cannot be accurately approximated by a linear model, this framework may fail to capture its true dynamics.

While the proposed framework effectively captures the dynamics of MISO systems through TL, it does have certain limitations. Notably:

- Sensitivity to ϵ selection: Care should be taken when choosing the convergence criterion (ε). In the provided case study, setting ε to 0.05 led the framework to ask for more experimental data after the second iteration, while a more appropriate choice ($\varepsilon = 0.01$) would have resulted in a valid model after the third iteration (Figure 8). The selection of ε impacts the efficiency and outcome of the framework.
- Dependency on initial UCVs: The efficiency of the framework is highly dependent on the initially assumed UCVs, chosen before bias calibration. Different UCVs can lead to distinct models, influencing the number of iterations and overall efficiency of the framework. Careful consideration in selecting initial UCVs is crucial for optimal performance.
- Limited scope to linear models: The framework presented here focuses on the transfer learning of MISO systems specifically for linear models. Nonlinear models are not addressed in this work. Enhancing the framework to include the transfer of nonlinear models from source datasets to target datasets is identified as a subject for future research.

6. Conclusions

In conclusion, this paper introduces a novel framework for transfer learning and uncertainty calibration using dimensionless numbers, specifically designed for system identification of multiple-input multiple-output (MIMO) systems with only a limited amount of data. The framework is applied to two demanding case studies: the first involves metal additive manufacturing, specifically laser-blown directed energy deposition (DED-LB) as the source domain and laser hot wire directed energy deposition (DED-LB/W) as the target domain; the second focuses on modelling the reactant concentration in a Continuous Stirred-Tank Reactor (CSTR) with a nonlinear, temperature-dependent reaction. The framework demonstrates robust performance across both applications, achieving accurate system identification despite the limited data availability. In the DED-LB/W case study, the framework achieves a commendable R-squared, normalised root mean square error, and mean absolute percentage error of 0.632, 0.103, and 11.21%, respectively, on the target system with data from only three beads, each with a single step change in one of the inputs. In the CSTR case study, the framework effectively models the nonlinear dynamics, achieving R-squared, normalised root mean square error, and mean absolute percentage error of 0.956, 0.0865, and 9.24%, respectively. This highlights the framework's power in modelling nonlinear systems and adapting to varying system dynamics, even under conditions of extreme data scarcity. To the authors' knowledge, this is the first successful application of transfer learning for MIMO system identification under such challenging conditions, and the inaugural investigation of its kind within metal additive manufacturing processes.

Apart from the framework's capability to capture the dynamics of a MIMO system with a very low amount of data, it leverages a novel uncertainty calibration approach. This method minimises autocorrelation of residuals to fine-tune the standard deviations of dimensionless input numbers. Such adjustments significantly improve the agreement between the developed models and experimental data by effectively calibrating the inherent uncertainties of input parameters. Additionally, the framework employs dimensionless numbers which contribute enhancing model accuracy and reducing computational complexity, as proven through a comparative study. Future work will focus on minimising the sensitivity of the proposed approach to the convergence criterion, extending the framework to nonlinear models, and integrating real-time data.

Acknowledgement

The authors extend their sincere gratitude to the Stevens Institute of Technology's Additive Manufacturing Laboratory for generously sharing their experimental data. Additionally, we gratefully acknowledge funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada [Grant numbers: RGPIN-2019-06601] and Business Finland under Project #: 4819/31/2021 with affiliation to the Eureka! SMART project (S0410) titled "TANDEM: Tools for Adaptive and Intelligent Control of Discrete Manufacturing Processes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Natural Sciences and Engineering Research Council of Canada [Grant Number RGPIN-2019-06601].

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the articles (Akhavan et al., 2023) and (Rahmani Dehaghani et al., 2024) and their supplementary materials.

Author contributions

Mostafa Rahmani Dehaghani: Conceptualisation, Methodology, Investigation, Programing, Writing-Original Draft, Validation, and Writing-Review and Editing. Yifan Tang, Pouyan Sajadi, and J. Akahavan: Data Curation, Visualisation, Formal Analysis, and Writing-Review and Editing. G. Gary Wang: Conceptualisation, Supervision, Project Administration, Funding Acquisition, and Writing-Review and Editing. The first draft of the manuscript was written by Mostafa Rahmani Dehaghani and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ORCID

Mostafa Rahmani Dehaghani http://orcid.org/0000-0003-4618-0302

References

Akhavan, J., Lyu, J., Mahmoud, Y., Xu, K., Vallabh, C. K. P., & Manoochehri, S. (2023). Dataset of in-situ coaxial monitoring and print's cross-section images by Direct Energy Deposition fabrication. *Scientific Data*, *10*(1), 776. https://doi.org/10.1038/s41597-023-02672-4

Alzubaidi, L., Al-Amidie, M., Al-Asadi, A., Humaidi, A. J., Al-Shamma, O., Fadhel, M. A., Zhang, J., Santamaría, J., & Duan, Y. (2021). Novel transfer learning approach for medical imaging with limited labeled data. *Cancers*, *13*(7), 1590. https://doi.org/10.3390/cancers13071590

Azom. (n.d.). *Stainless Steel - Grade 316*. Retrieved November 7, 2023, from https://www.azom.com/properties.aspx?ArticleID = 863.

Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, *65*(332), 1509–1526. https://doi.org/10.1080/01621459.1970.10481180

Bradshaw, A. (1978). Review of "Mathematical description of linear systems". By W. J. Rugh. (Marcel Dekker, 1975.) [Pp. 177.] Price 48 Sw. F. *International Journal of Systems Science*, *9*(3), 361. https://doi.org/10.1080/00207727808941704

Buckingham, E. (1914). On physically similar systems; illustrations of the use of dimensional equations. *Physical Review*, *4*(4), 345. doi:10.1103/PhysRev.4.345

Chatfield, C. (2013). *The analysis of time series: Theory and practice*. Springer.

Chu, S. R., Shoureshi, R., & Tenorio, M. (1990). Neural networks for system identification. *IEEE Control Systems Magazine*, *10*(3), 31–35. doi:10.1109/37.55121

Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, *19*(90), 297–301. doi:10.1090/S0025-5718-1965-0178586-1

Day, O., & Khoshgoftaar, T. M. (2017). A survey on heterogeneous transfer learning. *Journal of Big Data*, *4*(1), 1–42. doi:10.1186/s40537-016-0062-3

Dong, A., Starr, A., & Zhao, Y. (2023). Neural network-based parametric system identification: A review. *International Journal of Systems Science*, *54*(13), 2676–2688. https://doi.org/10.1080/00207721.2023.2241957

Doraiswamy, L. K. (1992). *Elements of chemical reaction engineering: By H. Scott Fogler. Prentice-Hall, New York, 1992. xxii+ 838 pp*. Academic Press.

Gosiewski, Z., & Paszowski, M. (2003). Identification of physical parameters of unstable systems: Theoretical background. *Systems Analysis Modelling Simulation*, *43*(3), 301–311. https://doi.org/10.1080/0232929031000150247

Han, T., Liu, C., Wu, R., & Jiang, D. (2021). Deep transfer learning with limited data for machinery fault diagnosis. *Applied Soft Computing*, *103*, 107150. https://doi.org/10.1016/j.asoc.2021.107150

Hu, Z., & Mahadevan, S. (2017). Uncertainty quantification and management in additive manufacturing: Current status, needs, and opportunities. *International Journal of Advanced Manufacturing Technology*, *93*(5–8), 2855–2874. https://doi.org/10.1007/s00170-017-0703-5

Kazmer, D. O., & Colon, A. (2020). Injection printing: Additive molding via shell material extrusion and filling. *Additive Manufacturing*, *36*, 101469. https://doi.org/10.1016/j.addma.2020.101469

King, W. E., Barth, H. D., Castillo, V. M., Gallegos, G. F., Gibbs, J. W., Hahn, D. E., Kamath, C., & Rubenchik, A. M. (2014). Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing. *Journal of Materials Processing Technology*, *214*(12), 2915–2925. https://doi.org/10.1016/j.jmatprotec.2014.06.005

Ławryńczuk, M. (2015). Nonlinear predictive control for Hammerstein–Wiener systems. *ISA Transactions*, *55*, 49–62. https://doi.org/10.1016/j.isatra.2014.09.018

Lyapunov, A. M. (1992). The general problem of the stability of motion. *International Journal of Control*, *55*(3), 531–534. https://doi.org/10.1080/00207179208934253

Mahadevan, S., Nath, P., & Hu, Z. (2022). Uncertainty quantification for additive manufacturing process improvement: Recent advances. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, *8*(1), 10801. doi:10.1115/1.4053184

Marmarelis, M. G., & Ghanem, R. G. (2020). Data-driven stochastic optimization on manifolds for additive manufacturing. *Computational Materials Science*, *181*, 109750. https://doi.org/10.1016/j.commatsci.2020.109750

Matweb. (n.d.-a). *2209 duplex stainless steel*. Retrieved November 7, 2023, from https://www.matweb.com/search/datasheet.aspx?matguid=e4df7ef1593f4f518bd3b26667a0aa56&ckck=1.

Matweb. (n.d.-b). *AISI type 316L stainless steel*. Retrieved November 7, 2023, from https://asm.matweb.com/search/SpecificMaterial.asp?bassnum=mq316q.

Milanese, M., & Vicino, A. (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, *27*(6), 997–1009. https://doi.org/10.1016/0005-1098(91)90134-N

Moges, T., Ameta, G., & Witherell, P. (2019). A review of model inaccuracy and parameter uncertainty in laser powder bed fusion models and simulations. *Journal of Manufacturing Science and Engineering*, *141*(4), 040801. https://doi.org/10.1115/1.4042789

Monti, A. C. (1994). A proposal for a residual autocorrelation test in linear models. *Biometrika*, *81*(4), 776–780. https://doi.org/10.1093/biomet/81.4.776

Mukherjee, T., Manvatkar, V., De, A., & DebRoy, T. (2017). Dimensionless numbers in additive manufacturing. *Journal of Applied Physics*, *121*(6), 64904. https://doi.org/10.1063/1.4976006

Naitali, A., & Giri, F. (2016). Wiener–Hammerstein system identification – An evolutionary approach. *International Journal of Systems Science*, *47*(1), 45–61. https://doi.org/10.1080/00207721.2015.1027758

Niu, K., Zhou, M., Abdallah, C. T., & Hayajneh, M. (2022). Deep transfer learning for system identification using long short-term memory neural networks. *ArXiv Preprint* ArXiv:2204.03125.

Olleak, A., & Xi, Z. (2020). Calibration and validation framework for selective laser melting process based on multi-fidelity models and limited experiment data. *Journal of Mechanical Design*, *142*(8), 081701. https://doi.org/10.1115/1.4045744

Oza, R., Shah, N., Tadse, D. G., & Joshipura, M. H. (2000). Application of MATLAB in process control: Case study for first order reaction in a CSTR. *Nirma University Journal of Engineering and Technology (NUJET)*, *1*(1), 45–49.

Rahmani Dehaghani, M., Sahraeidolatkhaneh, A., Nilsen, M., Sikström, F., Sajadi, P., Tang, Y., & Wang, G. G. (2024). System identification and closed-loop control of laser hot-wire directed energy deposition using the parameter-signature-quality modeling scheme. *Journal of Manufacturing Processes*, *112*, 1–13. https://doi.org/10.1016/j.jmapro.2024.01.029

Rahmani Dehaghani, M., Tang, Y., & Gary Wang, G. (2022). Iterative uncertainty calibration for modeling metal additive manufacturing processes using statistical moment-based metric. *Journal of Mechanical Design*, *145*(1), 012001. https://doi.org/10.1115/1.4055149

Ritto, T. G., Beregi, S., & Barton, D. A. W. (2022). Reinforcement learning and approximate Bayesian computation for model selection and parameter calibration applied to a nonlinear dynamical system. *Mechanical Systems and Signal Processing*, *181*, 109485. https://doi.org/10.1016/j.ymssp.2022.109485

Rubenchik, A. M., King, W. E., & Wu, S. S. (2018). Scaling laws for the additive manufacturing. *Journal of Materials Processing Technology*, *257*, 234–243. https://doi.org/10.1016/j.jmatprotec.2018.02.034

Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Bünau, P., & Kawanabe, M. (2008). Direct importance estimation for covariate shift adaptation. *Annals of the Institute of Statistical Mathematics*, *60*(4), 699–746. https://doi.org/10.1007/s10463-008-0197-x

Tang, Y., Rahmani Dehaghani, M., & Wang, G. G. (2023). Review of transfer learning in modeling additive manufacturing processes. *Additive Manufacturing*, *61*, 103357. https://doi.org/10.1016/j.addma.2022.103357

Tsoi, J. K. P., Patel, N. D., & Swain, A. K. (2018). Grey-box neural network system identification with transfer learning on ball and beam system. *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8. https://doi.org/10.1109/IJCNN.2018.8489103

Wang, W. (2007). Applications of MIMO technique for aerospace remote sensing. *2007 IEEE Aerospace Conference, 1–10*. https://doi.org/10.1109/AERO.2007.353075

Wang, Z., Dai, Z., Póczos, B., & Carbonell, J. (2019b). Characterizing and avoiding negative transfer. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11293–11302.

Wang, X., Hassanien, A., & Amin, M. G. (2019a). Dual-function MIMO radar communications system design via sparse array optimization. *IEEE Transactions on Aerospace and Electronic Systems*, *55*(3), 1213–1226. https://doi.org/10.1109/TAES.2018.2866038

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, *3*(1), 9. https://doi.org/10.1186/s40537-016-0043-6

Xiao, Y., Chen, H., & Ding, F. (2011). Identification of multi-input systems based on correlation techniques. *International Journal of Systems Science*, *42*(1), 139–147. https://doi.org/10.1080/00207720903470189

Zhang, Y., Wang, Y., Jiang, Z., Zheng, L., Chen, J., & Lu, J. (2022). Tire defect detection by dual-domain adaptation-based transfer learning strategy. *IEEE Sensors Journal*, *22*(19), 18804–18814. https://doi.org/10.1109/JSEN.2022.3201201

Zhao, X., Imandoust, A., Khanzadeh, M., Imani, F., & Bian, L. (2021). Automated anomaly detection of laser-based additive manufacturing using melt pool sparse representation and unsupervised learning. *2021 International Solid Freeform Fabrication Symposium*.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76. https://doi.org/10.1109/JPROC.2020.3004555