Generating synthetic data for data-driven solutions via a digital twin for condition monitoring in machine tools

Brett Sicard[a], Quade Butler[a], Yuandi Wu[a], Sepehr Abdolahi[a], Youssef Ziada[b], and S. Andrew Gadsden[a]

[a]McMaster University, 1080 Main St W, Hamilton, Ontario
[b]Global Manufacturing Engineering, Ford Motor Company, 35500 Plymouth Rd, Livonia, Michigan, USA

ABSTRACT

Machine tools (MT) are critical to modern manufacturing. They allow precision manufacturing of complex components at high volumes. MTs are large capital investments that require maintenance and monitoring to ensure they remain in good working condition. To best achieve reliability and high performance it is necessary to implement condition monitoring, fault detection and predictive maintenance. One solution for implementing these is by utilizing data-driven methods such as neural networks. One issue with any data-driven method is that they require large quantities of labeled data. This is especially difficult for fault detection applications as faults tend to be rare, and as a result, the datasets tend to be very imbalanced. One emerging technology that can be implemented to solve this issue is the digital twin (DT). DTs provide a solution for data collection, modeling, simulation, and smart services. One way that DTs can be used is to generate synthetic data which can be used for various data-driven methods. This data can be validated on a test bench to ensure its accuracy before implementation in production. Synthetic data generated from the DT model can be used to create a dataset for various condition monitoring DT services. This study involved the use of simulation software to generate synthetic data which was used to implement a fault detection algorithm for preload loss monitoring. This method has been demonstrated to be effective at identifying the current operating conditions of the system. This method shows promise to improve reliability and performance in MTs, and could be adapted to condition monitoring in other systems such as vehicles, buildings, and power generation.

Keywords: Condition monitoring, Digital twin, Fault detection, Machine tool, Physics informed machine learning

1. INTRODUCTION

Machine tools (MTs) are critical components of modern manufacturing. Computer numerical control (CNC) MTs have been the backbone of modern manufacturing since the 1980's. CNCMTs enable the fast and precise manufacture of work pieces to maximize manufacturing throughput and profitability. To maximize the financial return and performance of the critical and expensive assets, it is valuable to apply asset management (AM) methods such as condition monitoring (CM), fault detection (FD), and predictive maintenance (PM).

AM is an effective approach to optimizing the return on investment for manufacturing assets such as a MT. CM is the process of monitoring machine signals and parameters to observe changes and trends which may indicate poor system health or condition. FD is the process of monitoring, identifying, and classifying faults. This allows to prompt maintenance, or the implementation of fault tolerant control to adapt to the faulty conditions. PM is the process of identifying faulty or worn components and repairing / maintaining / replacing them before they can cause unexpected downtime. By doing this you can utilize components throughout their entire useful life while avoiding costly unexpected downtime, which can heavily negatively affect production throughput, or damage other machine components. Each of these processes requires the collection of data and

Further author information: (Send correspondence to Brett Sicard)
Brett Sicard: E-mail: sicardb@mcmaster.ca

Synthetic Data for Artificial Intelligence and Machine Learning: Tools, Techniques, and Applications II, edited by Kimberly E. Manser, Christopher L. Howell, Raghuveer M. Rao, Celso De Melo, Proc. of SPIE Vol. 13035, 1303503 © 2024 SPIE · 0277-786X · doi: 10.1117/12.3013119

models to accurately represent the system. This data collection and modeling process is effectively implemented via a digital twin (DT).

DTs are virtual representations of physical systems , processes, or objects. They create a real-time link from the physical domain to the virtual modeling domain. DTs utilize a heterogeneous data stream, which is often collected via an Internet of Things (IoT) sensor network. For modeling purposes, there may be just one model, or many. These models can be either physics-based, such as Simulink models, or finite element (FE) models, or data-driven models such as neural networks (NN). For AM methods, data-driven models are often used, especially in the case of FD. One difficulty of this approach is the data imbalance issue typically in anomaly detection (AD)/ FD. By their definition, faults or anomalous behaviour occurs much less frequently than normal operating conditions.[1] Because of this, data re-balancing efforts are usually required. However with faulty conditions there is also the issue that to collect data about faulty conditions, they need to occur, and this is far from ideal as faulty conditions usually lead to machine damage, or faulty components being manufactured. One exciting emerging method which could possibly remedy this issue, is physics informed machine learning (PIML).

PIML is a hybrid modeling strategy that combines physics-based modeling with data-driven modeling.[2] There are several different approached under the umbrella of PIML, such as physics informed NNs, data-enhanced refinement of physical models, and physics-informed regularization. This work will focus on physics embedded in feature space, more specifically physics-guided input feature augmentation. This involves using accurate and advanced simulations to generate synthetic data which can be used for training data in a NN or other intelligent data-driven method. Because theoretically an infinite variety of faults can be simulated, data imbalance should not be an issue. Additionally, because this occurs in a simulation there is no risk to damaging components.

This work examines the application of DT for generating synthetic data for intelligent methods of CM, FD, and PM. A background investigation of the literature shows that the concept has already been successfully applied before in similar applications. DT modeling of a MT feed drive for the purpose of CM and FD is used to demonstrate the value of the concept. A DT model is created using Siemens mechatronic concept designer (MCD) which can be used to generate data to train a MathWorks Simscape model that can be used to detect anomalous behaviour via a data-driven approach. By identifying anomalous behaviour it is possible to either implement FTC or PM to improve system performance or reliability.

The remainder of the work will be organized as follows: section 2 describes the methodology for creating a DT MT with data collection, modeling, validation, and DT services. Section 3 covers background literature on CM, FD, PM as well as DTs and PIML. Section 4 describes the application scenario showing the effectiveness of the approach. Section 5 discusses the results from the experiment demonstrating a DT driven CM service. Finally, section 6 concludes the work and provides a future work outlook.

2. DIGITAL TWIN MACHINE TOOL

The process for modeling and generating synthetic data via a DT for CM, FD, and PM is described in the following section. The first step is data collection from the real system. Next, modeling of the system using advanced modeling or simulation software occurs. With this model or models, various movements and actions can be simulated. Virtual sensors collect state information from the simulation. Data from this simulation is then exported and compared to the real system, where the same movements or actions were performed to validate the system. After the simulation model has been validated, it can be used to generate synthetic data to train various intelligent, ML, or AI based CM, FD, PM and performance optimization DT services.

2.1 DATA COLLECTION

DT models rely on a heterogeneous stream of high accuracy and high frequency data for modeling purposes. This data is used to model the system states and parameters. MTs have many options of data streams for modeling[3] Some of this data comes via integrated sensors such as the encoders[4] and integrated torque/motor current,[5] or external sensors such as vibration,[6] and temperature sensors.[7] In addition to sensors, production data, such as production speed, machine reliability, or part quality can be used for evaluation. Control system information such as position and velocity deviation can also be useful to evaluate the tracking and control performance of the system. This data is often collected via an IoT network of distributed sensors.[8] Certain states or parameters cannot be easily measured or estimated directly, because of this it may be necessary to implement some sort of sensor fusion. Sensor fusion combines multiple sensor readings to create more accurate, reliable measurements,[9] while also allowing the estimation of un-measurable states.[10]

2.2 MODELING

Physics-based modeling takes advantage of known physical relationships, features, and parameters of the system for effective modeling. In many applications this occurs in FE modeling software. FE modeling is useful for modeling and simulation of thermodynamics, vibration, and stress.[11,12] For simulations involving dynamic mechatronic systems, there are several options such as Simulink/Simscape,[13] Ansys Simplorer, Siemens MCD, Dassault Dymola, and Maplesim. These software packages allow the modeling of these systems based on their physical characteristics, and allows for the simulation of system behaviour. Either of these modeling techniques, or any number of other physics-based models, can be used to model and simulate the system. The drawback of a physics-based approach is the need for simplification, assumptions, and approximations of real life behaviour.[2] It also requires a good understanding of the underlying system.[14] Data-driven modeling does not require an understanding of the underlying system mechanics, but rather creates a mapping of inputs to outputs.[14] ML and NN are a very popular approach to this method. The primary drawback to data-driven modeling is the need for a large, varied dataset. This can be especially difficult for FD applications as faults occur fairly rarely, especially compared to normal operating conditions. Both approaches have their advantages and disadvantages, it is possible to use both in conjunction to mitigate the disadvantages of each, while retaining the benefits of both.[2]

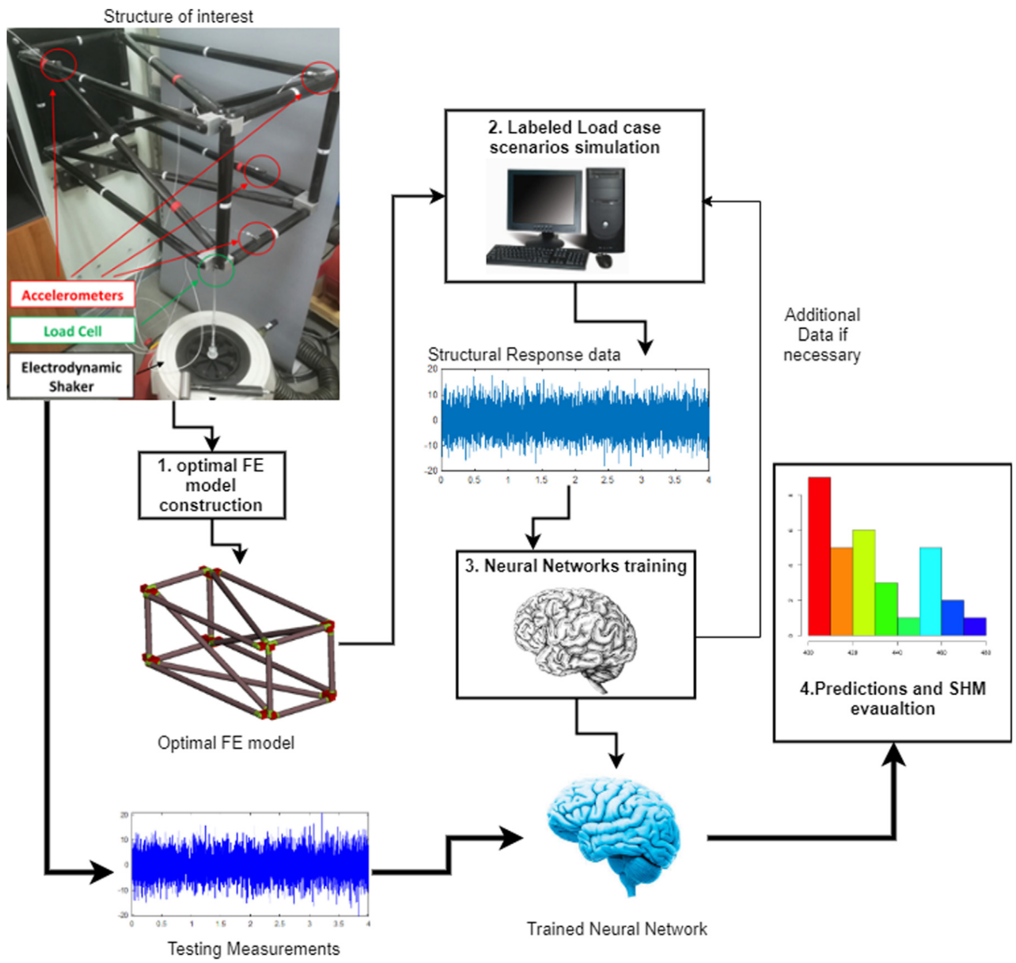
Hybrid modeling is the use of several models, either in parallel or in series, to model the system. Many of these applications use both a physics-based and a data-driven model. Parallel hybrid models have the same output values as each other, which are combined via some sort of filter, to create a better estimate than one model alone[15] as can be seen in Figure 1a. Series hybrid models use one model to generate outputs, which are used as input data for another model[12] as can be seen in Figure 1b. One type of PIML, physics-guided input feature augmentation, uses accurate physics-based modeling, such as FE modeling to generate data to train NNs. This is very useful in the case where certain data may be difficult to come by normally, such as faulty conditions.

2.3 VALIDATION

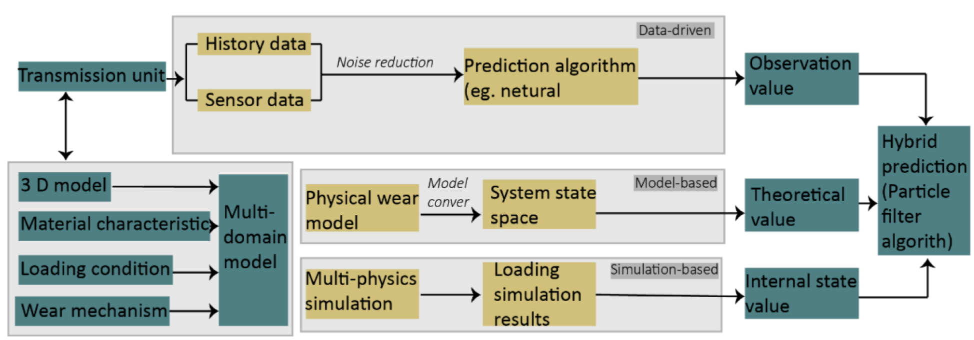
To ensure these DT models match their corresponding MT, they need to be able to produce a variety of state outputs for a set of inputs. In the case of a MT, inputs will typically be motor torque / force / current, and outputs will be linear or rotary position, velocity, acceleration. Once the model is complete, a dataset will be generated which will be used to validate against data collected from the real system as seen in Figure 2. Ideally the movement patterns used to generate data will be varied enough to fully represent the system across a range of its operating characteristics. This will ensure that the model will not over-fit a specific operating condition and be generalized. In the case of a MT, you would expect both stiffness and friction to vary across the stroke of a linear feed drive due to different levels of wear, and you would expect different friction characteristics at different velocities. Model outputs from virtual sensors will be compared to the real system experimental outputs to determine the prediction error to determine model consistency as seen in Figure 2.[12] Methods such as the genetic algorithm or particle swarm optimization can be used to modify the model parameters to minimize this prediction error. Once parameters have been optimized using one of these methods, and the error begins to converge to a minimal value, the validation step can conclude. Over the life-cycle this process will need to be repeated to ensure the model matches the current operating state of the MT.[16] The changing model parameters can also be observed for CM purposes.

2.4 DIGITAL TWIN SMART SERVICES

Once the DT model has been validated it can begin generating synthetic data for DT services. There are many services which can use this data. In the case of AM it can be used to generate data to model the behaviour of the system under certain operating conditions. This data could then be used to create a NN based CM estimate as mentioned in section 2.2. This may include a scenario where various levels of preloads are simulated and a NN



(a) Example of parallel hybrid modeling[15]



(b) Example of Series hybrid modeling[12]

Figure 1: Example of hybrid modeling

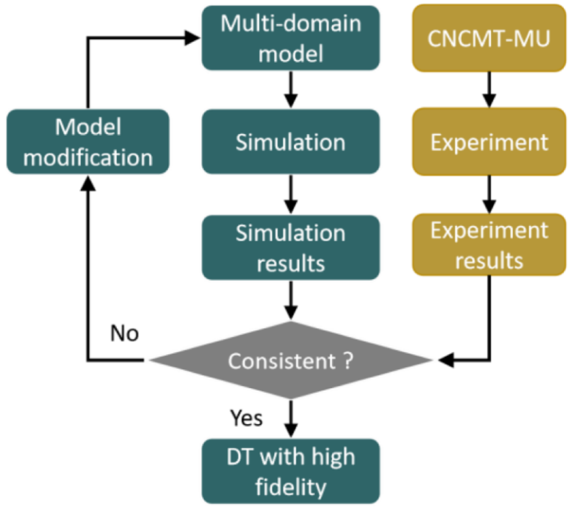


Figure 2: Consistency retention / model validation process[12]

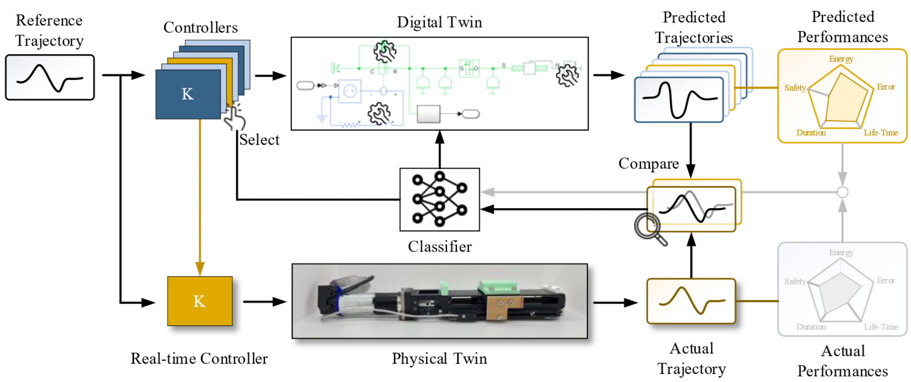


Figure 3: DT based control optimization for MT[17]

trained that can take the system outputs and estimate a level of preload. Other services could utilize this model such as performance optimization services. Often theses services modify control gains and simulate them in the virtual environment to maximize performance without the need for real life prototyping such as the example seen in Figure 3.

3. BACKGROUND LITERATURE

MTs are a popular application area for DT research. MTs are ideal candidates for DTs as they often have built in sensor networks, computing capabilities, modeling, and data analysis and visualization capabilities.[18] This means they already have the foundation for the construction of a DT for AM or performance optimization services, which require a heterogeneous data stream, modeling tools and a model update process, and analysis methods to examine historical trends and implement real-time monitoring. AM and DT are both popular avenues of research within the field of MTs as will be shown in this section.

AM methods such as CM, FD, and PM have often been applied to MTs in the literature with the goal of improving reliability, improving throughput, reducing replacement part and repair costs, and maximizing machine performance.[8,19] There are several approaches to AM in MTs, such as sensor-based methods, sensor-less analysis, or data-driven or model-based CM.[14,19] Sensor-based methods utilize additional sensors not needed for normal operation, whereas sensor-less methods use sensors and signals needed for normal operation, which is further discussed in section 2.1. A common example of sensor-based methods for AM is vibration monitoring, which usually requires an accelerometer collecting high frequency data. Vibration data can be analyzed in the frequency domain using fast Fourier transform to monitor the energy and frequency of vibration which can often be used to diagnose damage or wear.[6] Sensor-less analysis often will use encoder and torque information for AM purposes. This data can be used to make estimations of friction and stiffness parameters which can be extrapolated to measurements of overall system health.[20] Data-driven methods are popular in the literature, often using machine learning and NNs to relate operational data to condition estimations, and for FD.[21] NNs are popular in applications of tool wear estimation using inputs such as the cutting parameters, force signals, motor current, and vibration data to make predictions on tool wear levels.[22] Model-based methods use the understanding of the systems underlying mechanisms for CM. One example of this is estimation based methods such as interacting multiple models (IMM). IMM makes predictions of the behaviour of a system under certain operating conditions and can be used to identify faulty behaviour by matching operating behaviour to a model of faulty behaviour. This approach has been previously applied for preload loss detection in ball screws,[23] and actuator failure[24,25] to name a few. There are many ways to approach AM for MTs, each of these methods typically relies on an accurate system model and/or real-time data collection from a variety of data sources. This makes AM of a MT an exceptional candidate for the application of DT.

There are many applications which apply these AM methods with the help of DT, it has often been applied to MTs with the goal of creating a monitoring dashboard to observe and track predicted remaining useful life (RUL) or system performance.[26–28] Entire MTs can sometimes be too complex for a DT model due to the current state of the art of the technology, so it may be beneficial to just apply to a subsystem.[29,30] There are several important subsystems and components in a MT that are worthwhile applying AM strategies to via a DT. The spindle, feed drives, and cutting tools are key systems in a MT and have a substantial effect on the overall health and performance of the system. Feed drives are essential subsystems, as they position the parts within the MT. If they degrade it will lead to issues with position and feed rate tracking, surface finish, and overheating. There are several approaches to monitoring of this subsystem, one is monitoring the stiffness and damping characteristics of the feed drive to see how it changes over time.[31] Another approach could be to model a "healthy" or "normal" state as well as fault states and implement a method to identify the current operating condition. One paper implemented this approach to identify current levels of preload.[23] Spindles are an important components as they rotate the cutting tool in the case of a mill, or the work piece in the case of a lathe. Typical CM considerations for the spindle is stiffness,[32] thermal expansion,[33] and balance.[34] Finally, cutting tools are of value to examine as they are the direct interface between the machine and the work piece, and as such experience a high degree of force and wear.[35] Two popular approaches are direct and indirect measurements of wear. Direct measurement can be difficult in-process and therefore may require breaks in production. Indirect measurement on the other hand makes predictions on the level of wear based on machine signals such as force or vibration.[36] There are also hybrid methods which make use of both direct and indirect measurement.[37] DT has been shown to be effective for CM, FD, and PM in MTs and the various subsystems. To further augment these capabilities hybrid modeling and PIML has been applied in a few instances.

Hybrid modeling often uses parallel models that are combined together using a filter to generate improved estimates. One example used parallel physical wear model, simulation based estimates, and a NN prediction combined via a particle filter to estimate the wear of a transmission unit.[12] PIML can be used to overcome some of the potential deficiencies of data-driven or model-based approaches. PIML has been used to generate synthetic data for FD. One examples is utilizing data from an FE simulation to train a neural network to identify faults. In one work they utilized this method for identifying faults in a bridge. To validate their method they created an experimental setup with an identical fault to the synthetic data and they could correctly identify it with their NN trained on synthetic data.[15] A similar approach was taken in another work where FE simulations were used to create a synthetic damage parameter database which was sued to train a NN which could detect damage to a metal plate with a high degree of accuracy.[38]

4. APPLICATION CASE STUDY

In this work an example scenario of a MT feed drive test bench is examined. The application scenario uses an experimental setup, a multi-physical simulation model using MCD, and a physics-based simulation and operating condition identification model using Matlab Simulink. Together a CM and FD method are developed to identify

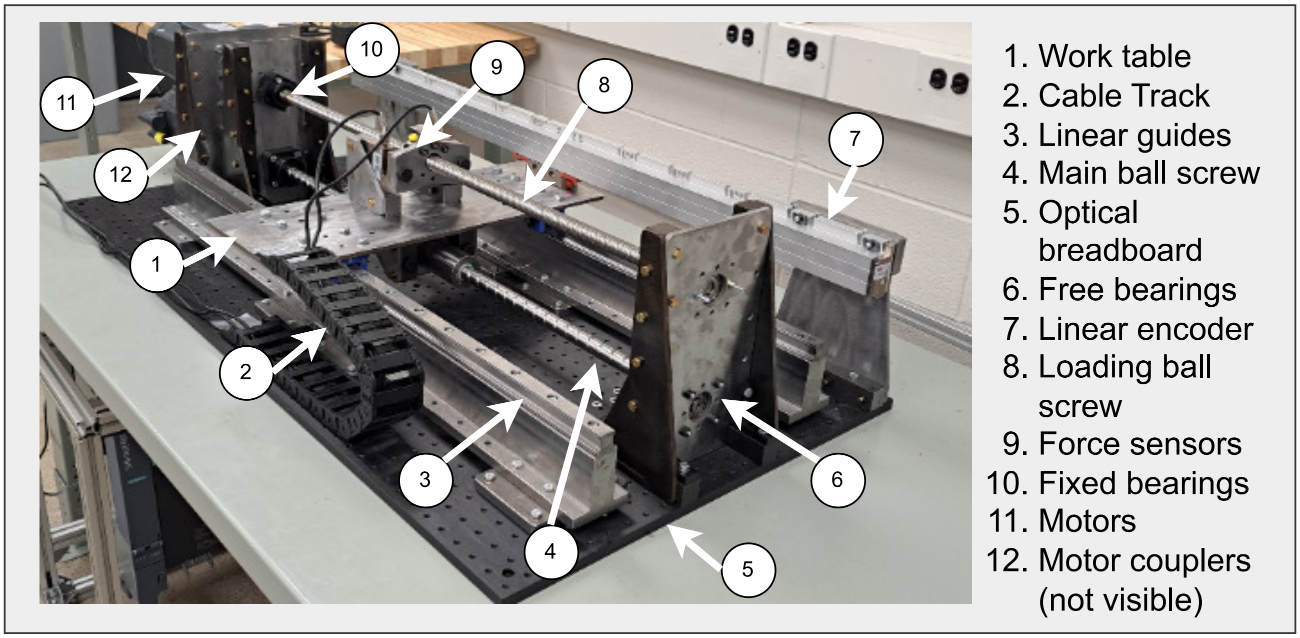


Figure 4: Full mechanical system overview

current operating conditions. With this knowledge of operating conditions, appropriate actions can be taken such as preemptive maintenance or applying some sort of FTC.

4.1 EXPERIMENTAL SETUP

The MT linear feed drive was constructed for the purpose of CM, FD, PM for industrial manufacturing CNC MTs. The intention of this setup is the ability to have control of manipulation of operating conditions such as the platform mass, the lubrication, alignment, and levels of wear. The test setup, which can be seen in Figure 4 includes two ball screws and motors which drive a platform forward and back. Position data is collected from both a linear encoder and rotary encoder to measure the linear position of the platform and rotary position of the screws respectively. Motor torque / current data is collected from the control system. These position state measurements and input force / current measurements can be used for system identification and system modeling purposes.

4.2 MODELS

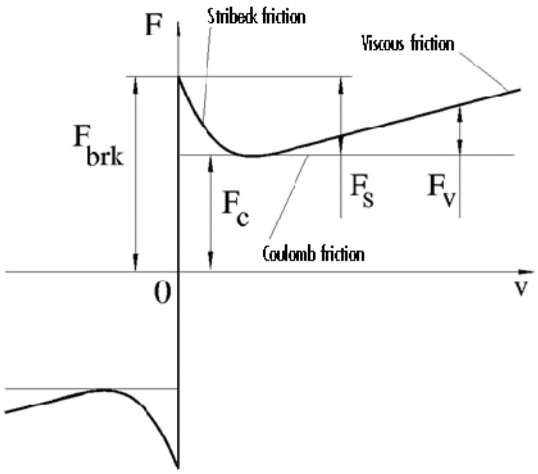
Several types of models are required to model and represent this system. The test bench has been modeled in MCD where the various components and physical connection between these parts is described and modeled. The mass and inertia of the components is calculated within the program based on the material properties and dimensions of the components. Modeling of the friction within the system requires implementing an equation to describe the friction based on the velocity of the system as well as various friction parameters. The system can be simulated in near real-time with a Simulink Simscape model. Each of these models are described in the following sections.

4.2.1 FRICTION MODEL

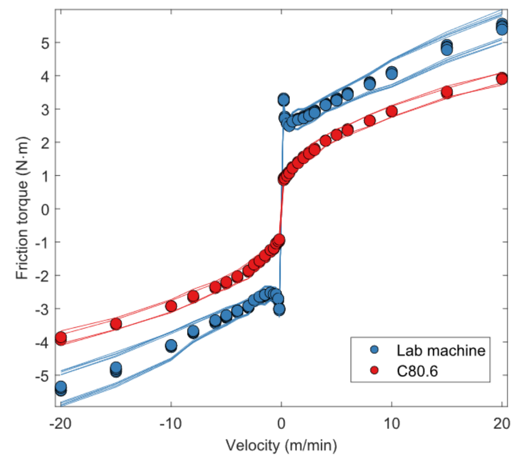
As will be discussed in section 4.3 the friction of the system will be affected by the preload. This relationship for the friction force can be given by the following equations:

$$F_{Friction} = F_{Stribeck} + F_{Coulomb} + F_{Viscous}, \tag{1}$$

$$F_{Stribeck} = \sqrt{2e}(F_{brk} - F_c) \cdot exp(-(\frac{v}{v_{st}})^2) \cdot \frac{v}{v_{st}}, \tag{2}$$



(a) Theoretical friction profile[39]



(b) Changing friction from experimental data from two MTs with different levels of wear

Figure 5: Friction curves of lubricated metal to metal contact

$$F_{Coulomb} = F_c \cdot tanh(\frac{v}{v_{Coul}}), \tag{3}$$

$$F_{Viscous} = Dv. \tag{4}$$

Where F_{brk}, F_c, f_{offset} v, v_{st}, v_{Coul}, D are the breakaway friction, Coulomb friction amplitude, velocity, Stribeck velocity threshold, Coulomb velocity threshold, and viscous friction coefficient respectively. This friction force creates a friction profile seen in Figure 5a[39] which was confirmed through test data gathered from a MT linear feed drive in previous work as seen in figure 5b.[14,20] In addition to the different friction profile at different velocities, wear will affect the friction characteristics. Increasing wear tended to decrease overall friction. More specifically, it tended to decrease breakaway friction, decrease coulomb friction, lower the breakaway velocity and reduce the viscous friction coefficient.

4.2.2 MCD

Siemens NX MCD is a multi-domain modeling program used to simulate mechatronic systems. It can be used to simulate the movement, forces, and interactions within the system. There are many different types of relations that need to be defined in the model including the following:

1. Rigid connections: parts connected and constrained along all axes, includes parts bolted together
2. Sliding joints: components can translate along one or more axes relative to each other, used to connect the linear guides to the moving platform
3. Screw joints: creates a rotary joint but prevents translation between components, used to relate the screw to the end bearings
4. Hinge joint: restricts movement to one degree of rotation, used to model the connection of the motor shaft to the couplers and the ball screw
5. Gear couple: connect the movement between components to move in a fixed ration, used to connect the ball screw rotation to the platform.

After assembling the model seen in Figure 6 of the experimental setup described in 4.1 using the relationships described above, synthetic data can be generated. Virtual sensors were utilized to measure the motor torque, the angular position of the ball screw shaft and the linear position of the stage.

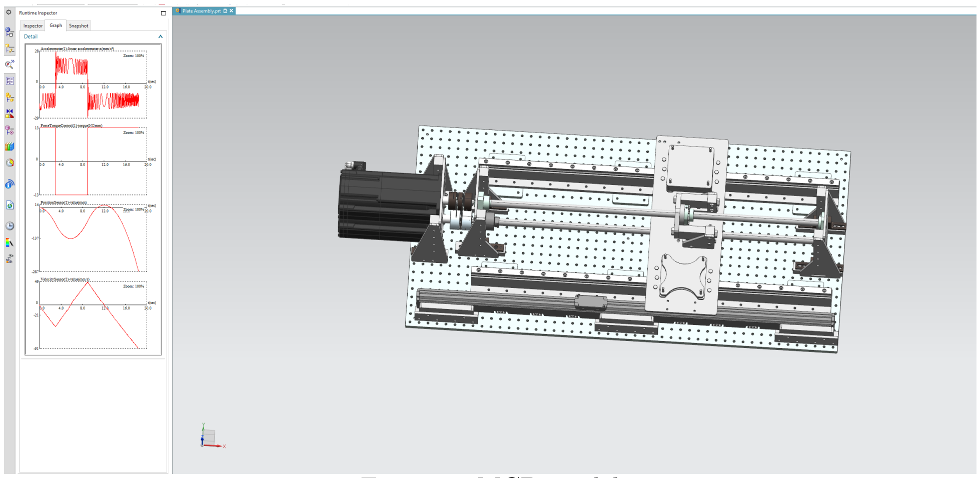


Figure 6: MCD model

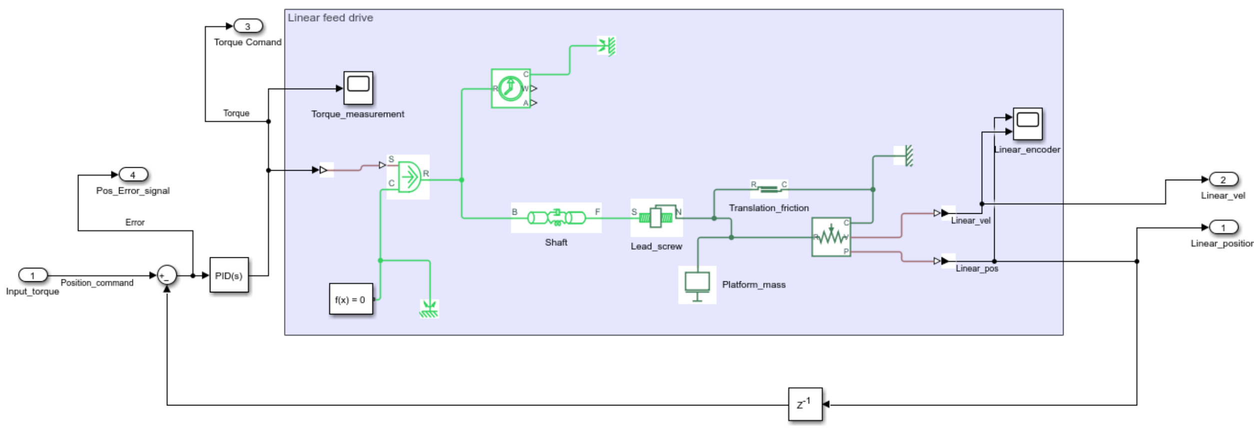


Figure 7: Simscape model of the LFD

4.2.3 SIMSCAPE

The Simscape model seen in Figure 7 is used for modeling the system in Matlab for the CM service. Data generated from the MCD model seen in section 4.2.2 can be used to train the parameters of the Simscape model using the parameter identification functionality of Simscape. Primarily of interest were the friction parameters seen in section 4.2.1. The system parameters are determined by providing input-output data. The model then fits itself to the data by minimizing the error with a series of iterative optimizations. Simscape is a relatively lightweight modeling tool compared to MCD. So, in a real life application this model could be ran in parallel with the real life system for real-time CM and performance prediction.

4.3 CONDITION MONITORING APPLICATION

The general application process for using DT to generate synthetic data for DT based CM services can be seen in figure 8. First a reference trajectory is created than can fully display the system mechanics over a multitude of velocities. This reference trajectory is utilized as an input to both the virtual model as well as the real system. An initial estimate of the parameters is created for the stiffness, and friction as was discussed in previous work[14,20] in addition to initial guesses for inertia/mass based on the geometric qualities of the system. These initial parameter estimates are used in the virtual MCD model. Initial estimates allow for a quicker convergence of the optimization sequence. Once the initial parameters are set an optimization process occurs where model parameters are adjusted until the error between the real system output and model converge to a minimum value. Once this occurs, stage 1 is complete. Next, stage 2 occurs where the Simscape model undergoes a similar process

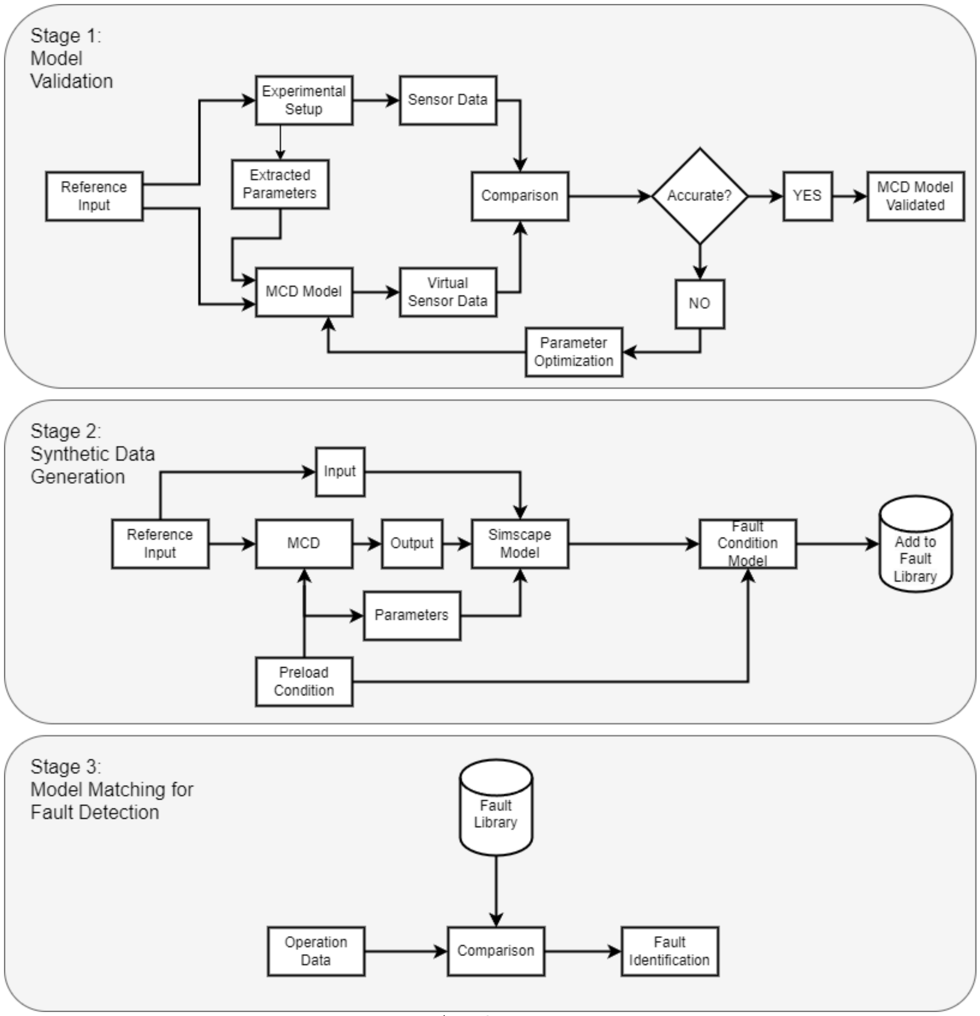


Figure 8: Application process

of optimization. Once this is complete a variety of preload scenarios data is generated. This generated data is used to create a fault library to represent the various levels of preload that could occur, which concludes stage 2. In Stage 3, operating data using the same reference trajectory is compared to the fault library data generated from the Simscape model. The error between the operating data and the fault library Simscape data is used to match models, where the minimal error is used for matching.

The DT service being developed to demonstrate this method is a preload identification system which matches operating conditions to a fault library model. Preload is used to increase system repeatability and rigidity. While doing so, preload also increases the overall friction of the system due to the increased contact forces between the ball bearings and the raceway. Preload is very important and it is essential that preload is maintained. In addition to preload being important, a loss of preload is often an indicator of wear, as a worn raceway or balls reduces the contact force, and therefore contact friction. So building upon previous work[14,23] this study seeks to use DT and a fault library to detect preload loss in the MT feed drive. Ideally the MCD model would be validated using experimental data from the test setup as discussed in section 2.3, however the test setup is currently not commissioned so it is not possible to generate experimental data at the moment. Because of this the validation portion of the process described in 2.3 Will not be included in this work. So the process used in this experimental study are as follows:

1. A variety of friction parameters are selected to represent several stages of preload loss

2. Friction parameters are used to build a Simscape model. Several different models are identified to represent the different preload scenarios.

3. Fault matching scheme designed to correctly predict the current level of preload based on a comparison between data outputted from experimental setup, and the data predicted from the Simscape model fault library.

5. RESULTS AND DISCUSSION

With the current level of implementation of the system, the second half of stage 2 and stage 3 can be implemented. Currently the experimental setup is not yet generating data, and it is therefore not possible to validate the MCD model. The utility of the stage 3: model matching process is thus displayed in this work.

To demonstrate the effectiveness of model matching for CM, several different preload conditions which are represented by a set of friction parameters are used to generate a fault library. The range of parameters are shown below in the Table 1. Where the first set represents a screw with high preload, and therefore higher friction as discussed in 4.2.1, and each subsequent model represents a decreased level of preload. Simulated output data with several levels of preload with added noise is used as the real system output. The three test scenarios can be seen in Table 2 where the three test scenarios represent a high (test 1), medium (test 2) and low (test 3).

The fault identification algorithm is as follows. The test conditions are applied as the model friction parameters, the reference trajectory is used as the system input. The position data is collected at the same sampling frequency as the fault library data. The mean square error is calculated between the test condition and fault states position data. Next, a likelihood vector is determined with the following steps. First the mean squared error (MSE) is normalized based on the minimum value:

$$MSE_{Normalized} = \frac{MSE}{min(MSE)}. \tag{5}$$

Then a likelihood value is determined based on the inverse proportion of MSE of each fault case, so if a fault scenario has a MSE of 25% compared to another scenario, it will be 4× as likely. The equation to determine the likelihood vector is

$$likelihood = \frac{MSE_{Normalized}^{-1}}{sum(MSE_{Normalized}^{-1})}. \tag{6}$$

To reduce the residual probability of unlikely scenarios, the likelihood is squared, and the squared likelihood function is calculated as follows

$$likelihood_{squared} = \frac{likelihood^2}{sum(likelihood^2)}. \tag{7}$$

For the experiments the likelihood of various operating condition states representing the system can be seen in Table 3. The squared likelihood vector (1x6) can then be multiplied by the matrix representing the fault states

Table 1: Preload friction fault library

Preload scenario	$F_{breakaway}$	$v_{breakaway}$	Coulomb	Viscous
1 (highest)	300	0.03	250	190
2	280	0.029	238	182
3	260	0.028	226	174
4	240	0.027	214	166
5	220	0.026	202	158
6 (lowest)	200	0.025	190	150

Table 2: Experimental friction parameters

Test	$F_{breakaway}$	$v_{breakaway}$	Coulomb	Viscous
1	295	0.0295	245	190
2	250	0.0275	222	170
3	215	0.026	200	155

Table 3: State likelihood estimates

Test	State 1	State 2	State 3	State 4	State 5	State 6
1	0.7473	0.2309	0.0172	0.0032	9,63e-4	3.83e-4
2	0.0058	0.029	0.27	0.61	0.064	0.01
3	5.09e-4	0.0015	0.006	0.0471	0.7811	0.1639

Table 4: Estimated versus actual friction parameters

	$F_{breakaway}$		$v_{breakaway}$		Coulomb		Viscous	
Test	Actual	Estimate	Actual	Estimate	Actual	Estimate	Actual	Estimate
1	295	294.38	0.0295	0.0297	245	246.6	190	187.75
2	250	251.77	0.0275	0.0276	222	221.06	170	170.71
3	215	218	0.026	0.0259	200	200.82	155	157.21

(4x6) and their sum represents the estimated friction characteristics

$$\begin{pmatrix} 300 & 280 & 260 & 240 & 220 & 200 \\ 0.03 & 0,029 & 0.028 & 0.027 & 0.026 & 0.025 \\ 250 & 238 & 226 & 214 & 202 & 190 \\ 190 & 182 & 174 & 166 & 158 & 150 \end{pmatrix} \times \begin{pmatrix} 0.0058 \\ 0.029 \\ 0.27 \\ 0.61 \\ 0.064 \\ 0.01 \end{pmatrix} = \begin{pmatrix} 251.77 \\ 0.0276 \\ 221.06 \\ 170.71 \end{pmatrix}. \tag{8}$$

With these estimates of friction the current level of preload can be estimated. It is also useful for for DT services, such as control tuning based on current operating conditions. The overall results and prediction accuracy of the CM method for the simulated data can be seen in Table 4. As can be seen the prediction accuracy is very high, with an average prediction error of less than 1.5%.

6. CONCLUSION

This work explores the possibility of utilizing a DT model of a MT for the generation of synthetic data which can be used for DT services. These DT services include CM, FD, PM, and performance improvements to maximize the reliability of the MT. Previous literature was examined to demonstrate the growing scope of literature applying DT to MTs, as well as the growing field of PIML. A framework for a DT model for synthetic data generation for CM is described, with the possible benefits of implementing such a system discussed and explored. Finally, a application scenario was presented where the DT model was used to generate a fault library utilized by a DT based preload identification CM service. This application was shown to be able to accurately determine levels of preload.

This research is still in the early stages but these preliminary results and findings show promise. The validity of the model was not validated with experimental data which is an important part of the DT framework. Future CM schemes could involve more AI based methods such as a NN. Future work would involve expanding the CM scope to include monitoring other factors such as wear and misalignment. The work could also be expanded upon to implement a FTC scheme to adapt system performance based on current condition.

REFERENCES

[1] Hopwood, M. W., Stein, J. S., Braid, J. L., and Seigneur, H. P., "Physics-Based Method for Generating Fully Synthetic IV Curve Training Datasets for Machine Learning Classification of PV Failures," *Energies* **15**, 5085 (Jan. 2022). Number: 14 Publisher: Multidisciplinary Digital Publishing Institute.

[2] Wu, Y., Sicard, B., and Gadsden, S. A., "A Review of Physics-Informed Machine Learning Methods with Applications to Condition Monitoring and Anomaly Detection," (Jan. 2024). arXiv:2401.11860 [cs, eess].

[3] Sicard, B., Butler, Q., Ziada, Y., and Gadsden, S. A., "Experimental Setups for Linear Feed Drive Predictive Maintenance: A Review," in [*2023 IEEE International Conference on Prognostics and Health Management (ICPHM)*], 357–367 (June 2023). ISSN: 2166-5656.

[4] Chandrasekar, P. and Srinivasan, K., "Inferential based measurement of backlash in servo system," *Materials Today: Proceedings* **46**, 9766–9770 (Jan. 2021).

[5] Chang, J.-L., Chao, J.-A., Huang, Y.-C., and Chen, J.-S., "Prognostic experiment for ball screw preload loss of machine tool through the hilbert-huang transform and multiscale entropy method," in [*The 2010 IEEE international conference on information and automation*], 376–380, IEEE / IEEE (June 2010).

[6] Pichler, K., Klinglmayr, J., and Pichler-Scheder, M., "Detecting Wear in a Ball Screw Using a Data-Driven Approach," in [*2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*], 3123–3128 (Oct. 2018). ISSN: 2577-1655.

[7] Liao, L. and Pavel, R., "Machine anomaly detection and diagnosis incorporating operational data applied to feed axis health monitoring," in [*ASME 2011 international manufacturing science and engineering conference, volume 2*], ASMEDC (Jan. 2011). tex.ranking: rank5.

[8] Sicard, B., Alsadi, N., Spachos, P., Ziada, Y., and Gadsden, S. A., "Predictive Maintenance and Condition Monitoring in Machine Tools: An IoT Approach," in [*2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*], 1–9 (June 2022).

[9] Sasiadek, J. Z., "Sensor fusion," *Annual Reviews in Control* **26**, 203–228 (Jan. 2002).

[10] McAfee, M., Kariminejad, M., Weinert, A., Huq, S., Stigter, J. D., and Tormey, D., "State Estimators in Soft Sensing and Sensor Fusion for Sustainable Manufacturing," *Sustainability* **14**, 3635 (Jan. 2022). Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.

[11] Xiao, J. and Fan, K., "Research on the digital twin for thermal characteristics of motorized spindle," *The International Journal of Advanced Manufacturing Technology* **119**, 5107–5118 (Apr. 2022).

[12] Yang, X., Ran, Y., Zhang, G., Wang, H., Mu, Z., and Zhi, S., "A digital twin-driven hybrid approach for the prediction of performance degradation in transmission unit of CNC machine tool," *Robotics and Computer-Integrated Manufacturing* **73**, 102230 (Feb. 2022).

[13] Yi, H. and Fan, K., "Co-simulation-based digital twin for thermal characteristics of motorized spindle," *The International Journal of Advanced Manufacturing Technology* **125**, 4725–4737 (Apr. 2023).

[14] Sicard, B., *DIGITAL TWIN MACHINE TOOL FEED DRIVE TEST BENCH FOR RESEARCH ON CONDITION MONITORING AND MODELING*, thesis, McMaster University, Hamilton, Ontario, Canada (2024). Accepted: 2024-01-15T19:14:09Z Journal Abbreviation: DIGITAL TWIN MACHINE TOOL FEED DRIVE TEST BENCH.

[15] Seventekidis, P., Giagopoulos, D., Arailopoulos, A., and Markogiannaki, O., "Structural Health Monitoring using deep learning with optimal finite element model generated data," *Mechanical Systems and Signal Processing* **145**, 106972 (Nov. 2020).

[16] Wei, Y., Hu, T., Zhou, T., Ye, Y., and Luo, W., "Consistency retention method for CNC machine tool digital twin model," *Journal of Manufacturing Systems* **58**, 313–322 (Jan. 2021).

[17] Chaiprabha, K. and Chancharoen, R., "A Deep Trajectory Controller for a Mechanical Linear Stage Using Digital Twin Concept," *Actuators* **12**, 91 (Feb. 2023). Number: 2 Publisher: Multidisciplinary Digital Publishing Institute.

[18] Sicard, B. S., Butler, Q., Ziada, Y., and Gadsden, S. A., "Cognitive dynamic digital twin: enhancements for digital twin platforms based on human cognition," in [*Big Data V: Learning, Analytics, and Applications*], **12522**, 48–63, SPIE (June 2023).

[19] Butler, Q., Ziada, Y., Stephenson, D., and Andrew Gadsden, S., "Condition Monitoring of Machine Tool Feed Drives: A Review," *Journal of Manufacturing Science and Engineering* **144** (June 2022).

[20] Butler, Q., Sicard, B., Hughey, E., Ziada, Y., Stephenson, D., and Gadsden, S. A., "Rapid parameter estimation of CNC feed drive systems," in [*Signal Processing, Sensor/Information Fusion, and Target Recognition XXXI*], **12122**, 344–351, SPIE (June 2022).

[21] Surucu, O., Gadsden, S. A., and Yawney, J., "Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances," *Expert Systems with Applications* **221**, 119738 (July 2023).

[22] Baig, R. U., Javed, S., Khaisar, M., Shakoor, M., and Raja, P., "Development of an ANN model for prediction of tool wear in turning EN9 and EN24 steel alloy," *Advances in Mechanical Engineering* **13**, 16878140211026720 (June 2021). Publisher: SAGE Publications.

[23] Sicard, B. S., Butler, Q., Ziada, Y., Hughey, E., and Gadsden, S. A., "Preload Loss Detection in a Ball Screw System Using Interacting Models," *IEEE Open Journal of Instrumentation and Measurement* **2**, 1–12 (2023). Conference Name: IEEE Open Journal of Instrumentation and Measurement.

[24] Afshari, H. H., Gadsden, S. A., and Habibi, S. R., "Robust fault diagnosis of an electro-hydrostatic actuator using the Novel dynamic second-order SVSF and IMM strategy," *International Journal of Fluid Power* **15**(3), 181–196 (2014). Publisher: Taylor & Francis.

[25] Gadsden, S. A., Song, Y., and Habibi, S. R., "Novel model-based estimators for the purposes of fault detection and diagnosis," *IEEE/ASME Transactions on Mechatronics* **18**, 1237–1249 (Aug. 2013). Publisher: IEEE.

[26] Guo, M., Fang, X., Hu, Z., and Li, Q., "Design and research of digital twin machine tool simulation and monitoring system," *The International Journal of Advanced Manufacturing Technology* **124**, 4253–4268 (Feb. 2023).

[27] Davies, O., Makkattil, A., Jiang, C., and Farsi, M., "A Digital Twin Design for Maintenance Optimization," *Procedia CIRP* **109**, 395–400 (Jan. 2022).

[28] Wang, K.-J., Lee, Y.-H., and Angelica, S., "Digital twin design for real-time monitoring – a case study of die cutting machine," *International Journal of Production Research* **59**, 6471–6485 (Nov. 2021). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00207543.2020.1817999.

[29] Sicard, B., Butler, Q., Kosierb, P., Wu, Y., Ziada, Y., and Gadsden, S. A., "Design Considerations for Building an IoT Enabled Digital Twin Machine Tool Sub-System," in [*2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*], 1–5 (Sept. 2023).

[30] Sicard, B., Butler, Q., Kosierb, P., Wu, Y., Ziada, Y., and Gadsden, S. A., "IIoDT: Industrial Internet of Digital Twins for Hierarchical Asset Management in Manufacturing," in [*2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*], 1–5 (Sept. 2023).

[31] Zhang, W., Zhu, D., Huang, Z., Zhu, Y., and Zhu, J., "Dynamic parameters identification of rolling joints based on the digital twin dynamic model of an assembled ball screw feed system," *Advances in Mechanical Engineering* **14**, 16878132221108491 (June 2022). Publisher: SAGE Publications.

[32] Xue, R., Zhang, P., Huang, Z., and Wang, J., "Digital twin-driven fault diagnosis for CNC machine tool," *The International Journal of Advanced Manufacturing Technology* (Aug. 2022).

[33] Lu, Q., Zhu, D., Wang, M., and Li, M., "Digital Twin-Driven Thermal Error Prediction for CNC Machine Tool Spindle," *Lubricants* **11**, 219 (May 2023). Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.

[34] Cao, H., Zhang, X., and Chen, X., "The concept and progress of intelligent spindles: A review," *International Journal of Machine Tools and Manufacture* **112**, 21–52 (Jan. 2017).

[35] Muthuswamy, P. and K, S., "Artificial intelligence based tool condition monitoring for digital twins and industry 4.0 applications," *International Journal on Interactive Design and Manufacturing (IJIDeM)* **17**, 1067–1087 (June 2023).

[36] Zhuang, K., Shi, Z., Sun, Y., Gao, Z., and Wang, L., "Digital twin-driven tool wear monitoring and predicting method for the turning process," *Symmetry* **13**(8) (2021). Number: 1438.

[37] Bahr, B., Motavalli, S., and Arfi, T., "Sensor fusion for monitoring machine tool conditions," *International Journal of Computer Integrated Manufacturing* **10**, 314–323 (Jan. 1997). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/095119297131066.

[38] Rai, A. and Mitra, M., "A hybrid physics-assisted machine-learning-based damage detection using Lamb wave," *Sādhanā* **46**, 64 (Mar. 2021).

[39] MathWorks, "Friction in contact between moving bodies - MATLAB."