Physics-informed Lightweight Temporal Convolution Networks for Fault Prognostics Associated to Bearing Stiffness Degradation

Weikun Deng, Khanh T P Nguyen, Christian Gogu, Jérôme Morio, Kamal Medjaher

Physics-informed lightweight Temporal Convolution Networks for fault prognostics associated to bearing stiffness degradation

Weikun Deng[1], Khanh T. P. Nguyen[2], Christian Gogu[2], Jérôme Morio [2], and Kamal Medjaher[3]

[1,2,3] *Laboratoire génie de production, Université de Toulouse, INP-ENIT, 47 Av. d' Azereix, 65000 Tarbes, France*
weikun.deng@enit.fr, tnguyen@enit.fr, kamal.medjaher@enit.fr

[2] *Institut Clément Ader, Université de Toulouse, 3 rue Caroline Aigle, 31400 Toulouse, France*
christian.gogu@gmail.com

[2] *ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France*
Jerome.Morio@onera.fr

ABSTRACT

This paper proposes hybrid methods using physics-informed (PI) lightweight Temporal Convolution Neural Network (PI-TCN) for bearings' remaining useful life (RUL) prediction under stiffness degradation. It includes three PI hybrid models: a) PI Feature model (PIFM) — constructing physics-informed health indicator (PIHI) to augment the feature space, b) PI Layer model (PILM) — encoding the physics governing equations in a hidden layer, and c) PI Layer Based Loss model (PILLM) — designing PI conflict loss, taking into account the difference before and after integration of the physics input-output relations involved module to the loss function. We simulated 200 different bearing stiffness degradations, using their discrete monitored vibration signals to verify the effectiveness of the proposed method. We also investigate their inference process through feature heat map analysis to interpret how the models melt physics knowledge to assist in capturing the degradation trend. The physics knowledge considered in this paper is the dynamic relationship between vibration amplitude and stiffness in a damped forced vibration model. The results show that all three PITCN models effectively capture degradation-related trend information and perform better than the vanilla lightweight TCN. Furthermore, the visualization of the feature channels highlights the important role of physics information in model training. Channels containing physics information demonstrate higher correlation with results as they significantly dominate the heat map compared to other channels.

Keywords—Physics-informed machine learning; Non-trending vibration; Prognostic; Bearing contact stiffness degradation

1. INTRODUCTION

Bearings are critical components that are susceptible to fatigue damage and need to be monitored. Under certain degraded working conditions, its crack propagation process is not obviously characterized in the early stage monitoring. (Li, Hu, Meng, Zhan, & Shen, 2018). As a result, the monitoring signal and the corresponding periodical statistics tend to show "slight trend" before severe degradation, and the dramatic feature variation only appears in signal collected from the end of services life (Porotsky & Bluvband, 2012).

Due to the incomplete knowledge of the bearing degradation mechanism, high cost in dynamics failure modeling (Massi et al., 2014) and the sparsity of the degradation information, the classic solutions, neither the traditional physics-based methods nor the data-driven machine learning (ML) methods (Shi & Chehade, 2021) are applicable to capture the information about the nonlinear degradation from past data and working conditions (H. Liu, Song, Zhang, & Kudreyko, 2021).

As a result, researchers have turned to the quest to develop hybrid approaches. On the one hand, guiding machine learning to explore the embedding properties of the data by imposing additional constraints during training has been shown to learn better representations of degradation trends (Liao, Jin, & Pavel, 2016). On the other hand, the physics-informed Machine Learning (PIML) method using incomplete physics model design constraints can maintain the physics consistency of ML training results and improve vanilla ML model performance (Karniadakis et al., 2021). For this purpose, the incorporation of physics knowledge in the main parts of the ML pipeline, including the augmented input space (Q. Wang, Taal, & Fink, 2021; Chao, Kulkarni, Goebel, & Fink, 2022), the algorithm architecture (Yucesan & Viana, 2020; Viana, Nascimento, Dourado, & Yucesan, 2021) and the objective function (J. Wang, Li, Zhao, & Gao, 2020), has received

much attention in recent years in PHM.
These researches prove that we can integrate physics knowledge directly in ML, especially the Deep neural networks (Nascimento, Corbetta, Kulkarni, & Viana, 2021) and in turn, ML can compensate for incomplete physics models (Yucesan & Viana, 2022), thus establishing a mapping between structural parameters (causal factors) and degradation states (phenomena).
However, in the face of such limited data, the performance of the PIML model is yet unknown. Moreover, many PIML improvements are primarily based on over-parametric Neural Networks (Caixian, 2021) that often have more parameters than the data points available for a single training batch (Deepmind, 2019). The large number of parameters can assist in expressing the complexity of the association between the casual factor and data to some extent, but may lead to over-fitting issue and becomes infeasible for real-time applications. Meanwhile, we still lack an intuitive sense of the mechanism of physics information in ML. An interesting question will be investigated and explored in this paper is whether the vanilla non-over-parametric lightweight model has the flexibility to embed the same knowledge in different ways to achieve better performance gains.
This paper is organized as follows.Section 2 aims to present problem statement while in Section 3, we describe the simulation procedure for the stiffness deterioration of a bearing. In Section 4, three different methods for integration of physics knowledge in lightweight TCN are detailed. The performance of three proposed PI-TCN models as well as the physics knowledge's role in training process of these models are investigated in Section 5. Finally, conclusions and perspectives of this work are discussed in Section 6.

2. PROBLEM STATEMENT

Bearing damages start from inside. Until the initial crack extends to its surface, there are no obvious signs of failure that can be observed because the geometry of the rollers is not altered. After that, the crack accelerates and the bearing fails rapidly (Khan, Kumar, Singh, & Singh, 2021). It is the root of the slight trend data. Hence, one can cite the following challenges for prediction of the bearing's RUL based on vibration signals, illustrated by Fig.1:

1. Throughout continuous stiffness degradation of bearings, the vibration signal varies insignificantly in the early phases but changes dramatically only near the failure time. As a result, it is not trivial to capture trend information from vibration signals which reflect the ongoing degradation evolution.
2. As the stiffness degradation is hidden, when investigating historical run-to-failure data we only know the linear function of RUL in working time and do not know the duration of bearing health state and degradation state. Then, it is not trivial to match the linear RUL values with hidden non-linear stiffness degradation process.

To address the above challenge, we propose a lightweight TCN as a benchmark purely data-driven model, using time-domain statistics as the input features to predict the bearing RUL in stiffness degradation. The input space includes mean, variance, max, min-max, root mean square, skew, kurtosis, peak factor, waveform factor, impulse factor and margin values. The output of the TCN is a deterministic value of the RUL (days). Then, we improve the performance of this benchmark model by integrating incomplete physics knowledge about the analytical relationship between stiffness degradation and vibration signal in terms of the augmented input space, modified hidden layer, and conflict loss function. These integration form three PITCN models: PI Features model (PIFM), PI Layers model (PILM), and PI Layer Based Loss models (PILLM).

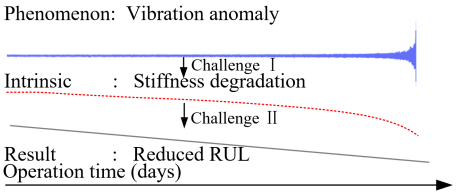

Figure 1. Challenges of bearing's RUL prediction under stiffness degradation.

The relationship between stiffness and vibration amplitude is shown in Eq.1 (Blake, 1961), where Vib_p is the peak value of the vibration signal and *stiff* represents the corresponding equivalent contact stiffness level. ε denotes the relevant imbalance in the system load. It is the extrinsic excitation of the bearing vibration. m represents the equivalent system mass. Ω is the rotation speed. In real conditions, the exact values of ε and m are unknown. Only the parameters Ω and Vib_p are available in vibration based RUL prediction.

$$Vib_p = \frac{\varepsilon\Omega^2}{\frac{stiff}{m} - \Omega^2} \tag{1}$$

The main objective of the proposed PI-TCN models is to approximate the mapping function g between the features extracted from vibration signals and the bearing's RUL values.

$$RUL = g\left(\frac{\Omega^2}{Vib_p}, \epsilon, m\right) \tag{2}$$

3. CASE STUDY DESCRIPTION

This case study aims to predict the RUL of a roller bearing, subject to stiffness degradation caused by the effect of crack expansion of its rollers. We assume that this bearing operates at a constant speed. And the bearing's state is monitored via vibration signals. In subsection 3.1, we describe how to simulate continuously degraded stiffness curves while subsection 3.2 presents how to generate vibration signals.

3.1. Continuous stiffness degradation simulation

As shown in Fig.2, we generate a total of 200 bearing stiffness run to failure degradation trajectories. The mean value of the failure times of those 200 trajectories is 8.04×10^5 and its standard deviation is 1.47×10^4. Among these trajectories, 50 sets are randomly selected as the test sets, and the remaining 150 sets are randomly divided into training and validation sets in 4:1 ratio. Each stiffness degradation trajectory is composed of health state, nonlinear degradation period and uncertain factors associated to operating environment effects.

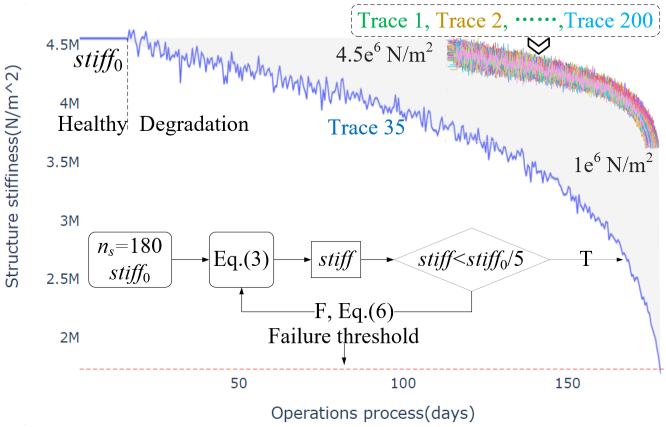

Figure 2. Simulation process of contact stiffness degradation.

We use the Eq.3 from (J. Liu & Shao, 2015) to calculate stiffness in different damage states:

$$stiff = \frac{1}{2\left(\frac{(cos\gamma)^{5/2})}{n_s k_p}\right)^{2/3}} + U_{stiff} \tag{3}$$

We set contact angle γ to 20° and k_p stands for the Hertz elastic contact stiffness between the ball and the smooth surface. We compute k_p according to Eq.5. The term U_{stiff} represents the stiffness model uncertainty. It is due to the complex environmental effects of the actual process, the simplifying assumptions of the model, and other factors that make the real value not strictly adhering to the physics model. We assume U_{stiff} satisfy a Skewed distribution in which the mean and the variance are equal to 10% and 5% of $stiff$ respectively. If the bearing is healthy, we calculate the stiffness via Eq.4:

$$stiff_0 = \frac{1}{2\left(\frac{(cos20)^{5/2})}{180 k_p}\right)^{2/3}} \tag{4}$$

where n_s is the number of contact surfaces. Duration of health condition days generated by random seeding rand (2, 20). Bearing degradation tends to be severe when n_s decreases. In this simulation, the initial value of n_{s0} is set to 180 while the duration, in which the bearing state is healthy, is generated by a random seed as shown in Eq.4. The coefficient k_p is computed with

$$k_p = \frac{4R^{1/2}}{6(\frac{(1-\nu^2)}{E})} \tag{5}$$

We set bearing roller radius R to 0.003 m, Poisson's ratio ν to 0.3, Young's modulus E to 2.1×10^{11} in the different simulations. We assume that the expansion of the bearing defect will result in a continuous reduction in the number of contact surfaces between the roller and raceway, with more areas changing from a face-to-face contact to a defective edge line-to-face contact, as shown in Fig.3. This process can be simulated by reducing n_s in Eq.6.

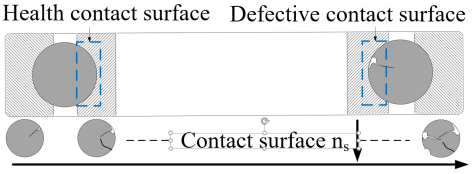

Figure 3. Defective contact and the roller failure schematic.

We generate an in-homogeneous degradation by the Eq.6. We set $steps$ to 0.0001. We use Eq.1 to get vibration amplitude, where ε is 20 g.cm, m is 5 kg. Ω takes the value 4200 rpm to indicate the wear at constant speed and external load. With U_{deg} denoting the uncertainty to reflect the non-uniform expansion of defects. U_{deg} conforms a Skewed distribution in which the mean and the variance values equal to 10% and 5% of the n_{si} respectively.

$$n_{si} = n_{s0} \times steps \times i \times (180 - steps \times i) + U_{deg}, i \in \mathbf{N} \tag{6}$$

3.2. Generation of discrete vibration monitoring signals

The stiffness values are assumed to be monitored by vibration sensors whose measurements are recorded every six hours. In this case study, the vibration amplitudes are generated by Eq.1 and then substituted into Eq.7 (if the bearing is in healthy state) or into Eq.8 (if the bearing is in degradation state) to obtain the vibration sequences. Shocks caused by defects are commonly accompanied by both frequency and amplitude noise. We add frequency noise (n_{d_1}) according to the Skewed distribution of (3, 1).With 3 as the mean and 1 as the variance, this indicates an uncertainty shock due to a roller defect introducing a high multiple of the rotational frequency during the bearing rotation. We also add amplitude noise (n_h) according to the signal-to-noise ratio of 1/100 in the healthy state and 1/10 in (n_{d_2}).

$$x(t) = vib_p \times sin(\frac{2\pi\Omega}{60}t) + n_h \tag{7}$$

$$x(t) = vib_p \times sin(\frac{2\pi\Omega}{60}t) + vib_p \times sin(\frac{2\pi\Omega}{60}t(1+n_{d_1})) + n_{d_2} \tag{8}$$

Each vibration sample is generated according to 4096 Hz sampling frequency of 3 s time length. More than 700 vi-

bration samples were collected for each trajectory, as shown in Fig.4. It can be seen that the generative data satisfy the non-trending characteristics.

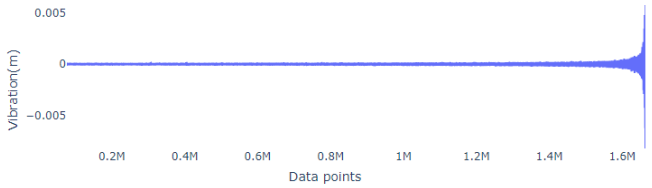

Figure 4. Illustration of the generated vibration signal.

4. DESCRIPTION OF THE PROPOSED METHODS

Informed learning is the seamless incorporation of constraints in an ML pipeline. PIML means transforming physics knowledge as potential or direct constraints of ML. In PHM, knowledge is parametrically representable perceptions of system behavior and failure mechanisms. In this section, we will present the purely data driven-model, i.e., a lightweight TCN model to predict bearing RUL, and then propose three PITCN models based on the same knowledge to improve the prediction performance. The first one is the PI Feature model (PIFM), which guides the extraction of data features through parameter relationships in the physics formulation. The second one is the PI Layer model (PILM), which embeds the physics input-output model in the computational function of the layer. And finally, the PI Layer Based Loss model (PILLM) adds regularization terms associated with the output of the physics model to the loss function of the vanilla TCN.

4.1. Purely data-driven model

CNN allows parallel computation of outputs and thus can achieve better performance than RNNs in sequence modeling (Lea, Flynn, Vidal, Reiter, & Hager, 2017). TCN, consisting of dilated, 1D convolution layers with the same input and output lengths, avoid common pitfalls of recursive models, such as gradient explosion or disappearance problems or lack of retention. We build lightweight TCN as "Benchmark" based with causal separable Conv1D, in Fig.5. The total number of parameters in the model is 3,411. Note that the subsequent implementation of PI-TCN will compute the physics features related to stiffness in the hidden layer, the values of which have a high probability of putting the neurons in the saturation zone. Therefore, the activation function is chosen for the nonlinear function h-swish with no upper bound, lower bound, smooth, and non-monotonic characteristics.

4.2. Physics-informed feature augmented input space

Eq.1 inspires us that the factor Ω^2/Vib_p can be used as a physics feature to predict the RUL of the bearing due to the RUL's dependence on the stiffness degradation level. Among the original 11 time-domain statistical features, we removed the Max feature while physics-informed health indicator (PHI)

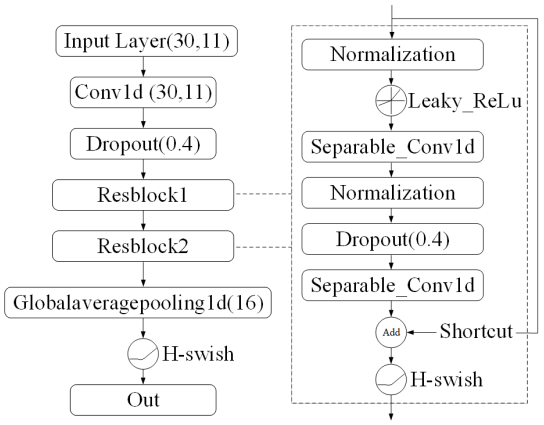

Figure 5. Lightweight TCN architecture diagram.

is added to construct new input samples having the same dimension (60 × 11) as the benchmark model. The benchmark model shown in Fig.5 is then re-trained as shown in Fig.6.

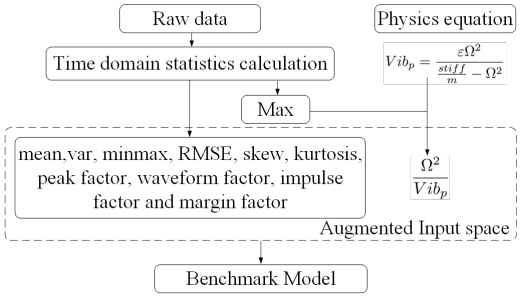

Figure 6. Create PIFM based on physics analytic relationships.

Although Ω^2/Vib_p is not an exact stiffness estimation, it has a clear physical meaning for being self-compiled quantities as stiffness. It contains trend information. PIFL is essentially an extension of the series combination structure to create a hybrid model. The output based on the physics model is part of the TCN input. The PI layer adds potentially physically consistent weak constraints to the input space through physics model parameter relationships based feature extraction.

4.3. Physics embedded layer

Fault dynamics models can be converted into an input-output module in ML, and in turn ML compensates for model incompleteness, as shown in Fig.7.

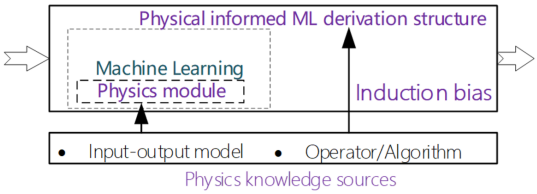

Figure 7. Embedding physics equations in NN layer.

Based on this paradigm, we have developed the PILM. Particularly, Eq.1 is transformed into a neural network linear model in Fig.8. The unknown function $g(\cdot)$ in Eq.2 can be approximated using a custom layer function $h(\cdot)$ of the neural network in the structure presented by Fig.9.
This model allows extracting the PHI Ω^2/Vib_p by embedding the transformation layer of Eq.1. Then, the PHI is used as the input of the hidden custom layer whose structure is defined as an approximate function of stiffness degradation in Fig.8. In the training process, the unknown parameters ε, m and U_{stiff} reflected by the weights (ω) and biases (b) of the hidden layer are updated to optimize the prediction results.

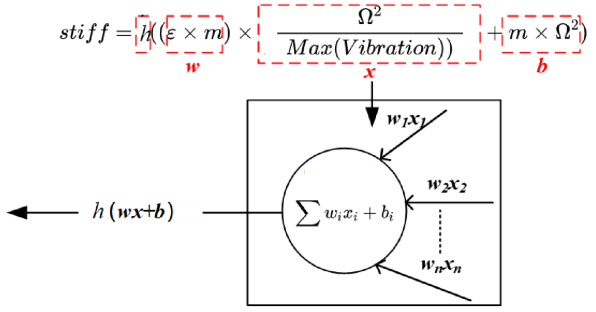

Figure 8. Embedding physics equations in NN layer.

Compared to the PIFM, PILM provides induction bias for TCN, as shown in Fig.7. It is able to ensure that the computational process based on physics knowledge is forced during the data processing of TCN, thus completely embeds the physics knowledge into the computational paradigm and overall derivation process of the TCN model.

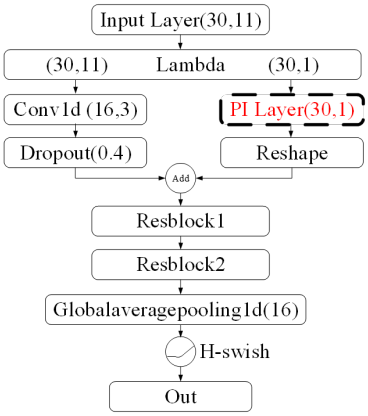

Figure 9. PI Layer outputs join the Resblocks' training.

4.4. Physics-informed layer based conflict loss

We also introduce the physics inconsistency by designing the loss function according to conflict between physics model's output and ML output. The whole PILLM consists of two parts: branch network and main network. An output layer is added after the physics informed layer to provide the PHI in the branch network. These features participate in the training process of the main network at one hand, and influence the hyper-parameter optimization of the main network through the loss function on the other hand, as shown in Fig.10. The outputs of each are measured with two losses and assigned corresponding weights of 1.0 and 0.2. In the prediction process, only the prediction results of the main network are utilized.

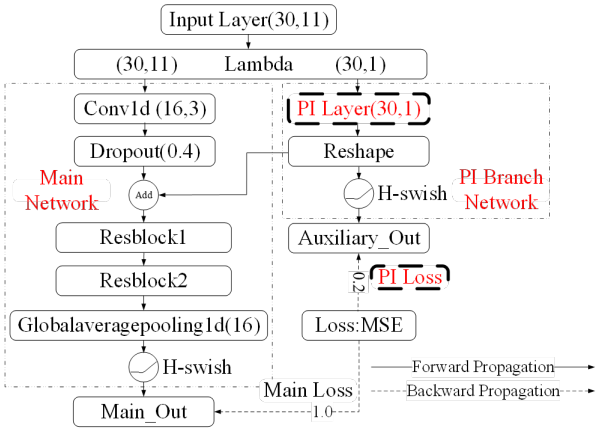

Figure 10. Build PI-loss based on different branches conflict.

In contrast to the first two methods, PILLM aims to intervene in the adaptive search process of the TCN in the solution space. Firstly, due to the single finite feature, it can be conjectured that the prediction accuracy of the PI Branch network is limited. resulting the high level loss, so the feature of minimizing the loss function in the training process of the TCN is used to improve the prediction accuracy of the Main network part. Secondly, by allowing the PHI of the PI branch network to participate in the training process of TCN, the two parts are permitted to share optimization process, thus allowing the TCN to satisfy a certain degree of physics consistency.

5. RESULT DISCUSSION

We fix the training epochs to 1000, design an early stop mechanism with patience equal to 80 epochs. We initialize the network parameters with uniformly drawn weights. The batch size of 128. The input-shape of each batch is (30,11). All models are trained in the same conditions. Moreover, we use Adam Optimizer in training (Kingma & Ba, 2014).

5.1. Investigation of the PITCN models' performance

Fig.11 presents different model's prediction results through 10 randomly selected trajectories of the test set while Fig.12 shows the box plots of the differences between the predicted and the truth RUL on the overall test set. The results highlight the performance of the PIML models compare to the one of the purely data-driven model: the same physics knowledge has different incorporation possibilities and potential for improving the benchmark. Particularly, we find that the PILLM model has the best performance. Its error range is only [17.97, 15.65] while the ones of the BENCHMARK, PIFM and PILLM

are respectively [-95.51, 79.83], [-26.85, 36.66], and [-33.38, 26.11].

Table 1 presents the performance of the proposed models compared with the benchmark on the overall test sets. We get the following conclusions from Fig.11, and Fig.12:

1. PI-TCN models show more accurate prediction with the smaller prediction error limits compared to Benchmark.
2. The predicted results of PI-TCN converge with the trend in the real value, showing the possibility to effectively solve the "Challenge II" in Fig.1.
3. Among the three different PI-TCN models, PILLM has the best prediction stability with the most compact upper and lower error limits and the minimum error mean, as shown in Fig.12.

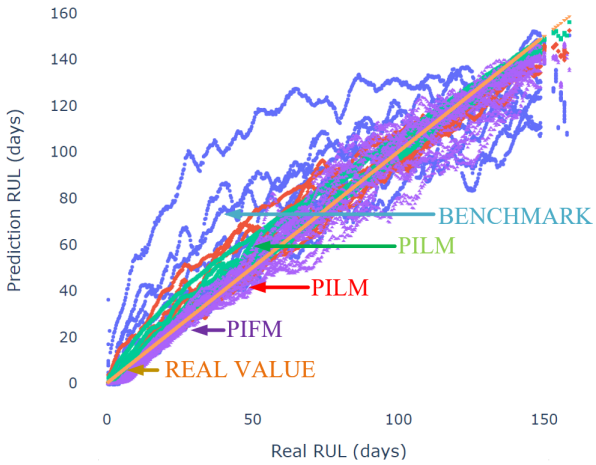

Figure 11. Prediction results of different models.

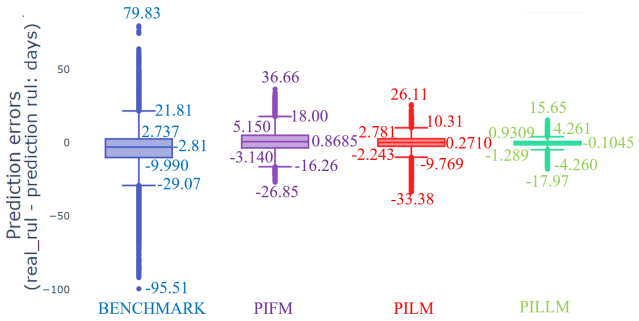

Figure 12. Model evaluation results.

Table 1. Performance evaluation on the test dataset.

	Benchmark	PIFM	PILM	PILLM
RMSE	16.191	8.059	5.818	3.157
MAE	10.878	5.939	3.942	1.965
R2	0.862	0.959	0.982	0.994

5.2. Investigation of physics knowledge's role in ML

To investigate the role of physics knowledge in the training process of the proposed models, we build the channel information model to generate the channel heat-map. After training, we extract the layers in which physics knowledge is integrated to investigate the correlation between physics information and results. More concretely, for the PIFM, we choose the input layer as the channel information model while for the PILM and PILLM, we choose the input layer as input and the "Add" layer as output. The brightness of the colors in the heat map reflects the correlation. The obtained results are presented in Fig.13, 15, and 16.

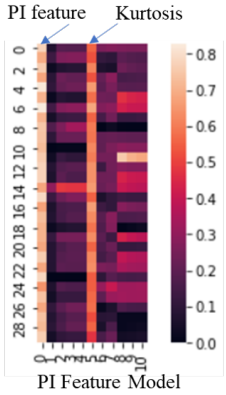

Figure 13. Weight heat-map of the input layer in PIFM.

In the PI Feature model's channel information heat-map, we find that the model focuses more on the channels where the Ω^2/Vib_p and Kurtosis are located. Kurtosis as a higher order statistic with the ability to capture dramatic trends from flat data. However, the Ω^2/Vib_p feature are assigned a higher weight than Kurtosis. This result highlight the intuition that the physics-informed feature generated based on analytic relational formulations provide additional crucial information to improve the model performance.

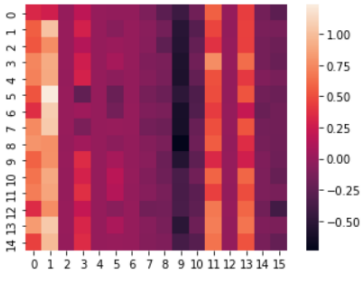

Figure 14. Heat-map of Benchmark model.

The results of Benchmark Conv1d layer are selected to build the channel information model, and the heat map is generated as shown in Fig.14, which is used as the cross-sectional com-

parison object of PILM and PILLM.

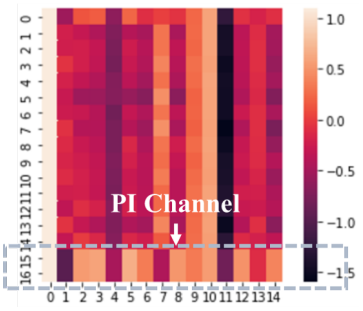

Figure 15. Heat-map of the PILM "Add" Layer.

For the heat-map of PILM, it can be seen in Fig.16, the features flowing from the PI Layer receive higher attention relative to the other channels.

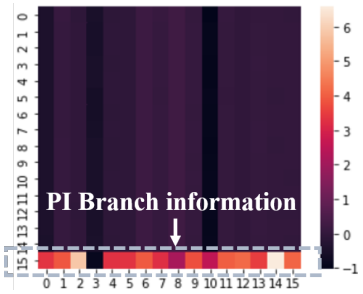

Figure 16. Heat-map of the PILLM "Add" Layer.

For the PILLM's heat-map, as shown in Fig.16 the part incorporated into the feature from the PI Loss is dominant compared to the other channels. In contrast to the feature map of the benchmark on the same layer, the feature map of the model is dimensionless, with a size of 16 x 15, because the features of PI Loss are not incorporated. The overall value of this feature map is smaller than the corresponding value of PI Loss. So this result highlights that physical knowledge plays a significant role in improvement of the prediction results.

6. CONCLUSIONS

In this paper, the physics knowledge about the relationship between vibration signals and stiffness degradation is exploited to create the physics-informed TCN models in three ways: augmented input space, physics equation embeded layer, and physics-informed conflict loss. The simulation result demonstrates the flexibility of methods incorporating physics knowledge and also highlight the significant improvements they can bring to vanilla TCN when working with "slight trend" data. In comparison with the benchmark model, PIFM, PILM, and PILLM reduce the mean prediction error by 69.39% (from 2.81 days to 0.8685 days), 90.35% (from 2.81 days to 0.2710 days) and 96.29% (from 2.81 days to 0.1045 days), respectively. By investigating channel weights in the related layer, we found that those improvements mainly stem from the focus of three PIML models placed on data streams containing physical information. Indeed, the underlying logic of the PIML models is to guide ML to capture features related to degradation by encoding physics knowledge and then to establish the underlying relationships between features and RULs thanks of ML's non-linear mapping capabilities. In our case, the stronger the physics constraints imposed by the encoding in TCN, the better the PITCN model performs. In future works, we will use sparse noise monitoring data and conduct in-depth research on the translatability of incomplete physics knowledge containing uncertainty to ML pipeline. Furthermore, PIML methods will be developed in real scenarios with complex systems and complex operating conditions.

REFERENCES

Blake, R. E. (1961). Basic vibration theory. *Shock and vibration handbook*, *1*, 2–8.

Caixian, C. (2021). Halving the error rate requires more than 500 times the computational power. *https://chowdera.com/2021/10/20211024171246242z.html*.

Chao, M. A., Kulkarni, C., Goebel, K., & Fink, O. (2022). Fusing physics-based and deep learning models for prognostics. *Reliability Engineering & System Safety*, *217*, 107961.

Deepmind. (2019). Alphastar: Mastering the real-time strategy game starcraft ii. *https://www.deepmind.com/blog/alphastar-mastering-the-real-time-strategy-game-starcraft-ii*.

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422–440.

Khan, S., Kumar, R., Singh, M., & Singh, J. (2021). Vibration and acoustic method for detection of cracks in bearings: A critical review. *Advances in Engineering Design*, 221–229.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Lea, C., Flynn, M. D., Vidal, R., Reiter, A., & Hager, G. D. (2017). Temporal convolutional networks for action segmentation and detection. In *proceedings of the ieee conference on computer vision and pattern recognition* (pp. 156–165).

Li, F., Hu, W., Meng, Q., Zhan, Z., & Shen, F. (2018). A new damage-mechanics-based model for rolling contact fatigue analysis of cylindrical roller bearing. *Tribology International*, *120*, 105–114.

Liao, L., Jin, W., & Pavel, R. (2016). Enhanced restricted boltzmann machine with prognosability regularization for prognostics and health assessment. *IEEE Trans-*

actions on Industrial Electronics, *63*(11), 7076-7083. doi: 10.1109/TIE.2016.2586442

Liu, H., Song, W., Zhang, Y., & Kudreyko, A. (2021). Generalized cauchy degradation model with long-range dependence and maximum lyapunov exponent for remaining useful life. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–12.

Liu, J., & Shao, Y. (2015). A new dynamic model for vibration analysis of a ball bearing due to a localized surface defect considering edge topographies. *Nonlinear Dynamics*, *79*(2), 1329–1351.

Massi, F., Bouscharain, N., Milana, S., Le Jeune, G., Maheo, Y., & Berthier, Y. (2014). Degradation of high loaded oscillating bearings: Numerical analysis and comparison with experimental observations. *Wear*, *317*(1-2), 141–152.

Nascimento, R. G., Corbetta, M., Kulkarni, C. S., & Viana, F. A. (2021). Hybrid physics-informed neural networks for lithium-ion battery modeling and prognosis. *Journal of Power Sources*, *513*, 230526.

Porotsky, S., & Bluvband, Z. (2012). Remaining useful life estimation for systems with non-trendability behaviour. In *2012 ieee conference on prognostics and health management* (pp. 1–6).

Shi, Z., & Chehade, A. (2021). A dual-lstm framework combining change point detection and remaining useful life prediction. *Reliability Engineering & System Safety*, *205*, 107257.

Viana, F. A., Nascimento, R. G., Dourado, A., & Yucesan, Y. A. (2021). Estimating model inadequacy in ordinary differential equations with physics-informed neural networks. *Computers & Structures*, *245*, 106458.

Wang, J., Li, Y., Zhao, R., & Gao, R. X. (2020). Physics guided neural network for machining tool wear prediction. *Journal of Manufacturing Systems*, *57*, 298–310.

Wang, Q., Taal, C., & Fink, O. (2021). Integrating expert knowledge with domain adaptation for unsupervised fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*.

Yucesan, Y. A., & Viana, F. A. (2020). A physics-informed neural network for wind turbine main bearing fatigue. *International Journal of Prognostics and Health Management*, *11*(1).

Yucesan, Y. A., & Viana, F. A. (2022). A hybrid physics-informed neural network for main bearing fatigue prognosis under grease quality variation. *Mechanical Systems and Signal Processing*, *171*, 108875.

BIOGRAPHIES

Deng WeiKun received the M.S and Bachelors degree in aerospace propulsion theory and engineering from Northwestern Polytechnical University University, Xi'an, China, in 2020. He a Ph.D researcher at Ecole nationale d'ingénieurs de Tarbes-Toulouse INP, France. His current research interests include machine learning based prognostics and health management and rotor dynamics.

Khanh T. P. Nguyen is Associate Professor at National School of Engineering in Tarbes (ENIT), France, and member of the Production Engineering Laboratory (LGP) since 2017. She received the Ph.D. degree in Automation and Production Engineering from Ecole Centrale de Nantes, France in 2012. From 2013 to 2015, she was a Postdoctoral Fellow with the French Institute of Science and Technology for Transport, Development and Networks (IFSTTAR). From 2016 to 2017, she was an Assistant Professor with the University of Technology of Troyes. Her research interests include applications of artificial intelligence in predictive maintenance and prognostics and health management (PHM).

Christian Gogu received his PhD in 2009 as part of a joint PhD program between the Ecole des Mines de Saint Etienne (France) and the University of Florida. He is currently Associate Professor in the department of Mechanical Engineering at Université de Toulouse and does his research within the Institut Clément Ader (ICA). His research interests include design under uncertainty, multidisciplinary design optimization, machine learning based diagnostics and prognostics with applications mainly to aerospace structures.

Jérôme Morio is research director at ONERA - the French Aerospace Lab - in Toulouse (France). He received his Ph.D in image processing from Aix-Marseille University (France) in 2007. His main research interests include uncertainty management, rare event probability estimation and sensitivity analysis.

Kamal Medjaher received the Ph.D. degree in control and industrial computing from University of Lille 1, Villeneuve-d'Ascq, France, in 2005. He was Associate Professor at the National Institute of Mechanics and Microtechnologies, Besançon, France, and FEMTO-ST Institute, from 2006 to 2016. He is currently Full Professor at Tarbes National School of Engineering (ENIT), France. He conducts his research activities within the Production Engineering Laboratory. His current research interests include prognostics and health management of industrial systems and predictive maintenance.