Chapter 7
Machine Learning Surrogates for Rapid Hertzian Stress Estimation and Static Transmission Error in Gear Pre-Sizing

Daniele Fabbri, Carlo Rosso, and Fabio Bruzzone

Abstract The dimensioning and dynamic evaluation of spur and helical gears rely on different strategies, typically addressed separately in conventional design. In this work, we integrate both within a machine learning framework by developing two complementary surrogates, each based on a distinct dataset. The first dataset is generated analytically from ISO based Hertzian stress calculations, representing the traditional dimensioning route. These values are used to train regression models such as Support Vector Machines, Random Forests, Gradient Boosted Trees, and Feed Forward Neural Networks, providing a fast surrogate for assessing surface durability and pitting resistance. In contrast, the second dataset targets the prediction of the static transmission error (STE), a key excitation source in gear dynamics, and is constructed using a semi analytic approach. To capture the meshing evolution of STE, both Feed Forward and Sequence to Sequence GRU models are investigated. By unifying dimensioning through ISO based stresses with dynamic evaluation via STE prediction, this work establishes a foundation for comprehensive gear design surrogates that accelerate exploration and support multi objective optimization.

Keywords Machine Learning · Gears · Dynamic · Neural Networks · Static Transmission Error

Introduction

In the context of gear engineering the use of machine learning (ML) has been grown in the recent years for modeling, diagnosis, and optimization. Conventional approaches, including analytical models, finite element simulations (FEM), or experimental campaigns, are typically resource intensive and limited in their adaptability to the wide range of gear geometries and operating conditions. By contrast,surrogate models offer the ability to learn complex nonlinear relationships between inputs and outputs, supporting accurate predictions, early fault detection, and efficient exploration of the design space. In general, ML techniques are classified into two main groups: supervised and unsupervised learning.

Unsupervised learning can reveal hidden structures in unlabeled data through clustering or rule-based methods [1], with applications in anomaly detection [2]. However, gear engineering mainly relies on supervised regression, addressing both regression tasks (e.g., stress or noise prediction) and classification tasks (e.g., healthy vs. faulty gears) [3, 4]. Such models have been successfully applied to predict stresses, wear, and acoustic emissions, with performance typically assessed using *MAE*, *MSE*, *RMSE*, and R^2 [5].

In the present work, two complementary evaluation strategies are adopted. For single value predictions, such as contact pressure, a parameter of primary importance in gear design, *RMSE*, is used to quantify deviations from reference values, while R^2 assesses the overall explanatory capacity of the model. For sequence predictions, namely the Static Transmission Error (*STE*) [6] evaluated over the mesh cycle, the surrogate model also considers the agreement in curve shape between the

Daniele Fabbri · Carlo Rosso · Fabio Bruzzone
Department of Mechanical and Aerospace Engineering
Politecnico di Torino, 10129 Torino, Italy
e-mail: daniele.fabbri@polito.it; carlo.rosso@polito.it; fabio.bruzzone@polito.it

Carlo Rosso · Fabio Bruzzone
GeDy TrAss srl
10128 Torino, Italy
e-mail: carlo.rosso@polito.it; fabio.bruzzone@polito.it

Kristin B. Zimmerman, PhD (eds.), *Special Topics & Experimental Techniques in Structural Dynamics, 2026, Vol. 5*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.1007/978-8-743-81444-3_7

ML output and the semi analytical reference solution. This dual perspective ensures that the surrogate models are not only numerically accurate but also capable of faithfully reproducing the physical trends of gear behavior.

Recent research has also addressed the use of machine learning for predicting both the design of the gears and the Static Transmission Error (STE), one of the main sources of dynamic excitation in gear systems [7]. *Papalexis* et al. [8] developed neural network surrogates trained on more than $10,000$ finite element simulations of polymer gears, achieving a mean absolute percentage error below 0.5% in replicating full STE curves. *Urbas* et Al. [9] demonstrated the use of multiple regressors including Random Forest, AdaBoost, and NN to predict the shape factor Y_z of non involute polymer gears from 456 FEM simulations, achieving R^2 values above 0.9 and confirming the feasibility of real time stress estimation. On the experimental side, *Hyoung* et al. [10] predicted gear whine noise using micro geometry and test bench data from 12 gear sets. After LASSO-based feature selection, ensemble models such as XGBoost outperformed OLS, improving R^2 by up to 63.5% *Haefner* et Al [11] developed a Neural Network based surrogate model to predict tooth root stress in micro gears using measured shape deviations. The model trained on FEM outputs, resulted in a model capable of providing not only point predictions but also confidence intervals [12].

Computational methods are also widely applied in gear design optimization, where the objective is to identify geometric configurations that minimize stress, noise, or vibration. Evolutionary approaches such as Genetic Algorithms (*GAs*) have been employed to explore the design space [13], but they typically depend on costly simulations like FEM or Loaded Tooth Contact Analysis (*LTCA*) to evaluate candidate designs [14, 15], creating a significant computational bottleneck.

Overall, recent studies demonstrate that machine learning has become a practical and scalable tool across gear engineering, from stress prediction and fault diagnosis to noise estimation and design optimization. Surrogate models, in particular, enable fast approximations of computationally expensive analyses [16], accelerating iterative workflows within the broader framework of data driven model order reduction [17].

In this context, our work is based on the ISO 6336 analytical standard [18] as a reliable data source for training surrogate models, while also extending the framework to the prediction of the Static Transmission Error (STE). By addressing both Hertzian contact stress and STE, the proposed approach provides efficient tools for assessing surface durability, pitting resistance, and dynamic performance directly at the design stage. The objective of this work is to establish a foundational methodology for developing comprehensive gear design surrogate models that can be integrated into multi objective optimization frameworks, enabling the rapid identification of optimal and robust gear pair configurations.

In Section 2, the procedure used to generate the datasets is described. The first dataset is constructed entirely from analytical contact stress values computed using the ISO 6336 standard, covering a broad range of gear macro geometries and loading conditions. In parallel, a second dataset is built for the prediction of static transmission error (STE), using the same gear input parameters but adopting different learning strategies. For this task, both a Feed Forward Neural Network trained via backpropagation and a Sequence to Sequence (Seq2Seq) GRU based model are considered.

The overall machine learning workflow and results will be presented in section third, detailing the training and evaluation strategies adopted for various regression models, including Linear Regression, Ensemble Algorithms, and Neural Networks. All models were implemented using widely used machine learning libraries that are highly optimized for regression tasks, such as *scikit-learn* [19] and *Keras* [20] through *TensorFlow* [21].

Finally, the conclusions in Section summarize the key findings and outline future directions in the development of surrogate models for gear design.

Dataset preparation

The aim of this work is to develop surrogate models for spur gear pairs targeting two complementary outputs: the Hertzian contact stress, $\sigma_{H1,2}$, computed from the ISO 6336-2 formulation, and the STE calculated from*GeDy TrAss* software which employs a nonlinear and Non-Hertzian approach to determine the deformations of the teeth under load and their contact conditions [22, 23, 24]. The dataset spans a wide range of gear macro geometry and loading conditions (Table 1), and only safe configurations are retained. These data are modeled with Supervised Regression Methods, including Elastic Net, Ensemble Algorithms, and Neural Networks. For the STE dataset, the input features are different and include macro-geometric parameters such as torque T, transmission ratio τ, addendum h_a, module m, face width b, pinion teeth z_1 and helix angle β. Furthermore, micro-geometric modifications are introduced in the form of parabolic tip relief C_a from a start diameter of d_{Ca1}, as well as facewidth crowning C_β. The target output is a sequence of 20 discrete points $ste_1, \ldots, ste_{20}$, representing the transmission error over a mesh cycle. Since STE is inherently sequential, the learning strategy relies on Sequence to Sequence GRU models. Together, these two datasets enable benchmarking of regression and sequence based machine learning approaches for durability (via stress prediction) and dynamic performance (via STE prediction) in gear design.

Hertzian contact stress - Dataset preparation

The stress dataset consists of about 5×10^6 samples with 15 initial features and one target label, the Hertzian stress σ_{H1}. After removing non-numeric entries and outliers, the dataset was analyzed to ensure suitability for regression. Skewness and kurtosis [25] were computed to assess distributional properties; strongly skewed features were normalized using either the $log1p$ or *Yeo–Johnson* transformation [26], improving feature symmetry.

To mitigate multicollinearity [27, 28], Pearson correlation analysis [29] was performed both with the target and across features. Redundant variables with perfect or near perfect correlation were removed, considering also their physical meaning. The final dataset contained 14 features and one label, summarized in Table 1.

Table 1 Features and label - Hertzian stress dataset

Type	Symbol	Dataset label	Range
Target variable	σ_{H1}	sigma_H1	$14 \div 792 MPa$
Feature	T	Input_T	$10 \div 3000 Nm$
Feature	h_a	Input_ha	$1.25 \div 1.8$
Feature	h_f	Input_hf	$1 \div 1.6$
Feature	ρ_a	Input_rhoa	$0.1 \div 0.33$
Feature	τ	Input_tau,	$2 \div 6$
Feature	α_t	alphat	$20 \div 21.17$
Feature	b	b	$10 \div 70 mm$
Feature	β	beta	$0 \div 20°$
Feature	d_{w1}	dw1	$15.28 \div 241.2 mm$
Feature	i	i	$25.19 \div 792.97 mm$
Feature	m_n	mn	$1 \div 6 mm$
Feature	x_1	x1	$-0.6 \div 0.9$
Feature	x_2	x2	$-0.6 \div 0.9$
Feature	z_1	z1	$17 \div 35$

Material: 34CrNiMo6, ISO code V,
Brinell Hardness 319HB, Hardened and Tempered

STE - Dataset preparation

The static transmission error (STE) dataset is smaller than the stress dataset, consisting of about 1350 samples. The input space includes 10 geometric and load features: torque T, transmission ratio τ, addendum h_a, module m, face width b, number of pinion teeth z_1, helix angle β, start diameter of profile modification d_{Ca1}, amount of parabolic tip relief C_a, and amount of facewidth crowning C_β. Since the start diameter of the profile modification is dependent on the macro-geometry configuration of the gear currently analyzed, its values have been determined proportionally to the pitch diameter D_p of the pinion and gear respectively and to the addendum h_a in three discrete levels. To ensure a smooth exploration of the multidimensional design space in a limited number of simulations, the design space was sampled using Latin Hypercube Sampling (LHS) [30] on the target 1350 simulations.

The obtained STE dataset was then reformulated into an autoregressive *seq2seq* with Gated Recurrent Units (GRUs) learning task in order to capture the cyclic structure of the target over the gear meshing. Each gear configuration generated a sequence of 20 static transmission error values, complemented by 10 geometric and load parameters together (Table 2) with two synthetic cycle features, $\sin(\lambda)$ and $\cos(\lambda)$, where $\lambda \in [0, 1]$ as done in [8], they represent the normalized mesh position sampled at 20 equally spaced points. These two additional features provided the model with an explicit and continuous representation of the cycle progression, ensuring that the learning process was not limited to purely discrete indices. Before training, the target STE values were normalized using a *MinMax* transformation.

Methodology and results

The following methodology was applied consistently across all models and datasets to ensure a rigorous and comparable model development process. The first step was to perform a grid search [31] on a stratified subset of the training data to

Table 2 Features and label - STE dataset

Type	Symbol	Dataset label	Range
Target variables	$ste1...20$	$ste_{1..20}$	$0 \div 2.61$
Feature	T	Input_T	$10 \div 1010Nm$
Feature	h_a	Input_ha	$1.25 \div 1.8$
Feature	τ	Input_tau,	$1.5 \div 6$
Feature	b	b	$10 \div 80mm$
Feature	β	beta	$0 \div 0.35°$
Feature	m_n	m	$1 \div 5mm$
Feature	z_1	z1	$17 \div 30$
Feature	d_{Ca1}	dCa1	$D_p \div D_p + \frac{2}{3}h_a \cdot m$
Feature	C_a	Ca	$0 \div 6\mu m$
Feature	C_{beta}	Cbeta	$0 \div 6\mu m$

Material: 34CrNiMo6, ISO code V,
Brinell Hardness 319HB, Hardened and Tempered

identify appropriate hyperparameters. This preliminary stage was crucial for tuning key parameters such as regularization strength, tree depth, learning rate, and network architecture, depending on the algorithm. The optimal hyperparameter configuration was selected by analyzing training and validation loss curves, with particular attention to instability indicators such as oscillations or divergence. The models were then retrained with the best hyperparameters on the full training dataset to maximize the use of available data and improve generalization. Evaluation was conducted on a separate held out test set, where all predictive metrics were reported to validate each model's performance. In addition to accuracy, the training time of each algorithm was measured and reported, providing a broader comparison of efficiency versus predictive power. All computations were performed on a workstation equipped with an *i9-13900F* processor (2.00 GHz) and 64 GB of RAM. The models were trained using CPU implementations of the machine learning frameworks. GPU acceleration was not employed in this study but it is regarded as a promising avenue for future work to reduce training times, particularly for complex neural networks or larger datasets.

From the Hertzian stress prediction results, it was evident that neural networks are well suited for capturing the nonlinear relationships between geometry, loading conditions, and stress response. However, the prediction of static transmission error presents a different challenge, as it is inherently sequential and reflects the mesh cycle of a gear. Initially, we attempted to model the full 20 point STE sequence with a simple Feed Forward network producing all outputs simultaneously, but this approach failed to capture the dependencies of the outputs to the mesh cycle. For this reason, we adopted a sequence to sequence architecture with GRU units, which leverages the autoregressive structure of the data to reconstruct the shape of the output curve point by point. This approach demonstrated improved performance, even though the STE dataset was smaller than the stress dataset, highlighting the importance of matching the learning strategy to the physical nature of the target values.

Hertzian stress contact

A wide set of machine learning models was trained and evaluated on the original ISO 6336 stress dataset, which includes more than five million samples and 14 features. The results show a clear ranking in terms of predictive performance, and computational cost. The linear regression baseline with Elastic Net regularization demonstrated moderate accuracy, with a test RMSE above 70 MPa and $R^2 = 0.85$. While interpretable and efficient, its limited capacity to capture nonlinear effects restricted performance. Support Vector Regression with an RBF kernel achieved significantly better accuracy ($R^2 = 0.968$, RMSE = 32.5 MPa), but required subsampling ($50,000$) due to the high computational cost of kernel methods, providing a good balance between cost and performance, though scalability to the full dataset remains limited. Tree based ensemble methods offered an attractive compromise between predictive power and efficiency. Random Forest achieved strong results ($R^2 = 0.974$, RMSE = 30 MPa) with very fast training, highlighting its robustness and speed. Histogram-based Gradient Boosting outperformed all classical methods with an RMSE of 14.8 MPa and $R^2 = 0.994$, although it required longer training and showed signs of potential overfitting. Neural networks delivered the best overall accuracy. The scikit-learn MLPRegressor (3 layers: 200-100-50) achieved an RMSE of 3.43 MPa, converging stably with early stopping. A deeper Keras-based architecture (160-192-96 neurons with ℓ_2 regularization) confirmed this performance with a test RMSE of 4.63 MPa, though training times were much longer (almost 2900 seconds). Despite the higher computational burden, neural networks proved particularly effective in capturing the strong nonlinearities of the Hertzian stress formulation. Overall, the results indicate that while ensemble methods like Random Forest and Gradient Boosting provide highly competitive accuracy at low computational cost, neural networks clearly stand out in terms of predictive precision. This confirms that deep learning approaches are especially suitable for surrogate modeling of ISO based Hertzian stress, where complex nonlinear relationships between gear macro geometry and stress response dominate. In Table 3 a comparison between different models is listed.

Table 3 Performance comparison of surrogate models on the original ISO 6336 stress dataset. Subset sizes are indicated where full dataset training was computationally prohibitive.

Model	Test RMSE [MPa]	Test R^2	Training Time [s]
Elastic Net (SGD)	70.29	0.852	38
Support Vector Regression (RBF, 50k samples)	32.47	0.968	106
Random Forest (50k samples)	29.62	0.974	5
Histogram Gradient Boosting (100k samples)	14.82	0.994	46
Neural Network – MLPRegressor (200-100-50)	3.43	0.99	560
Neural Network – Keras (160-192-96)	4.63	0.98	2880

STE

In this part of the work, the objective is to train a surrogate model to predict the Static Transmission Error (STE) curve of a gear pair over a full meshing cycle. The methodology follows a sequence based learning approach, where each curve is treated as a temporal sequence of 20 discretized steps representing different positions along the mesh cycle. The dataset includes both geometric and operational parameters, two additional columns are included to encode the phase of each sample in the cycle in a smooth and periodic way. The target variable is the STE value at each of these 20 positions. Additionally, a column stores the previous value in the sequence to support autoregressive learning.

Once the data is organized, it is split into training and test sets using a standard 80/20 ratio. The architecture used to model this problem is a Sequence to Sequence (Seq2Seq) model with Gated Recurrent Units (GRUs) [32], The idea is to use GRU rather than other methods as LSTM (Long Short Term Memory) because the problem has a short length sequences (20 steps of STE) where the data is not deeply hierarchical in time and GRUs can achieve very similar performance to LSTMs, often with faster convergence and less risk of overfitting[33, 34]. This approach seems to be a novelty in the context of gear dynamics, as the other works used a traditional approach in which the neural networks outcome 20 values in the output layers.

The encoder takes as input the full sequence of geometric and meshing parameters and compresses it into a latent representation of dimension 256, which serves as the internal state passed to the decoder. The decoder then takes the previous STE values (in a sequence of 20 points) and learns to reconstruct the future STE waveform step by step, conditioned on the encoded representation. The final dense layer is wrapped in a TimeDistributed layer, ensuring that a single scalar STE value is output for each of the 20 time steps. The loss function used is the Huber loss, the model is trained over 100 epochs with a batch size of 64, and both training and validation performance are tracked. After training, an example prediction is generated using the trained model. A correction is then applied to the first point of the predicted curve.

Finally, the predicted STE curve and the real curve are plotted together for visual comparison. This provides qualitative insight into the model's accuracy and smoothness in reproducing the physical behavior of gear meshing transmission error as shown in Figures 1 and 2. The result shows a close alignment between the predicted and actual curves, demonstrating the Seq2Seq model's effectiveness in capturing the dynamic structure of the problem. This is confirmed also by the Peak-to-Peak Transmission Error (PPTE) and by the metrics of the model summarized in Table 4

Table 4 Performance of the STE prediction using Seq2Seq GRUs .

Model	Test RMSE [MPa]	Test R^2	PPTE Error
Seq2Seq GRUs	0.00456	0.934	0.86%

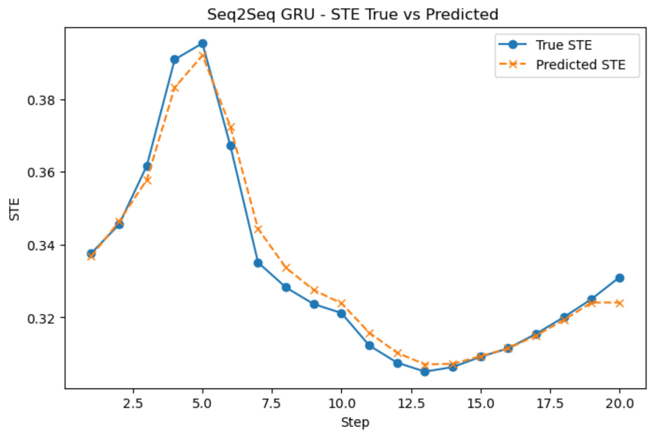

Fig. 1 Result STE curve true vs predicted values- First curve

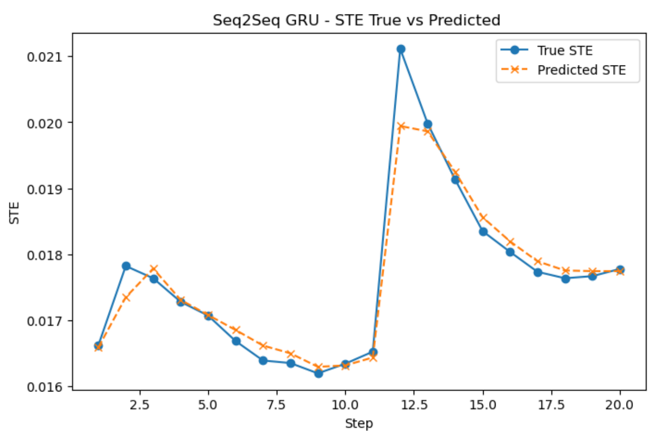

Fig. 2 Result STE curve true vs predicted values- Second curve

Conclusions

This paper should be intended as a starting point for applying different machine learning techniques to gear design, highlighting both dimensioning and dynamic behavior analysis. The section on Hertzian stress prediction demonstrates the effectiveness of using neural networks to capture complex nonlinear relationships, even with moderately deep architectures. In the second part, a novel approach for the regression of the Static Transmission Error (STE) is reconstructed using a specialized type of Recurrent Neural Network: the Seq2Seq GRU. The promising results confirm that traditional feedforward neural networks would struggle to retain temporal information about the meshing cycle and the evolution of the error over time. In contrast, the sequence to sequence approach succeeds in accurately reproducing the overall shape and trend of the STE curve, even when trained on a relatively small dataset.

Acknowledgments This publication is part of the project PNRR-NGEU which has received funding from the MUR – DM 117/2023

References

1. Aggarwal, C.C. *Data Classification: Algorithms and Applications*. Chapman & Hall/CRC, 1st edition (2014)ifnextchar.gobble.
2. Chandola, V., Banerjee, A., and Kumar, V. "Anomaly detection: A survey". *ACM Computing Surveys (CSUR)*, 41(3):1–58 (2009)ifnextchar.gobble.
3. Bansal, S., Sahoo, S., Tiwari, R., and Bordoloi, D. "Multiclass fault diagnosis in gears using support vector machine algorithms based on frequency domain data". *Measurement*, 46(9):3469–3481 (2013)ifnextchar.gobble.

4. LI, X., LI, J., QU, Y., and HE, D. “Semi-supervised gear fault diagnosis using raw vibration signal based on deep learning”. *Chinese Journal of Aeronautics*, 33(2):418–426 (2020)ifnextchar.gobble.
5. Botchkarev, A. “A new typology design of performance metrics to measure errors in machine learning regression algorithms”. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14:045–076 (2019)ifnextchar.gobble.
6. Czakó, A., Řehák, K., Prokop, A., Rekem, J., Láštic, D., and Trochta, M. “Static transmission error measurement of various gear-shaft systems by finite element analysis”. *Journal of Measurements in Engineering*, 12(1):183–198 (2023)ifnextchar.gobble.
7. Bruzzone, F. and Rosso, C. “Sources of excitation and models for cylindrical gear dynamics: A review”. *Machines*, 8(3) (2020)ifnextchar.gobble.
8. Papalexis, C., Sakaridis, E., Terpos, K., Kalligeros, C., Tsolakis, A., and Spitas, V. “Neural network surrogates for finite element models in loaded tooth contact analysis of polymeric gears”. *Mechanism and Machine Theory*, 214:106127 (2025)ifnextchar.gobble.
9. Urbas, U., Zorko, D., and Vukašinović, N. “Machine learning based nominal root stress calculation model for gears with a progressive curved path of contact”. *Mechanism and Machine Theory*, 165:104430 (2021)ifnextchar.gobble.
10. Sun Hyoung, L. and Park, K.P. “Development of a prediction model for the gear whine noise of transmission using machine learning”. *International Journal of Precision Engineering and Manufacturing*, 24 (2023)ifnextchar.gobble.
11. Haefner, B., Biehler, M., Wagner, R., and Lanza, G. “Meta-model based on artificial neural networks for tooth root stress analysis of micro-gears”. *Procedia CIRP*, 75:155–160 (2018)ifnextchar.gobble.
12. for Guides in Metrology, J.C. “Jcgm 101: Evaluation of measurement data - supplement 1 to the ”guide to the expression of uncertainty in measurement” - propagation of distributions using a monte carlo method”. Technical report, JCGM (2008)ifnextchar.gobble.
13. Cerrada, M., Zurita, G., Cabrera, D., Sánchez, R.V., Artés, M., and Li, C. “Fault diagnosis in spur gears based on genetic algorithm and random forest”. *Mechanical Systems and Signal Processing*, 66-67:812–828 (2016)ifnextchar.gobble.
14. Artoni, A. “A methodology for simulation-based, multiobjective gear design optimization”. *Mechanism and Machine Theory*, 133:95–111 (2019)ifnextchar.gobble.
15. Bonori, G., Barbieri, M., and Pellicano, F. “Optimum profile modifications of spur gears by means of genetic algorithms”. *Journal of Sound and Vibration*, 313(3-5):603–616 (2008)ifnextchar.gobble.
16. Forrester, A.I.J. and Keane, A.J. “Recent advances in surrogate-based optimization”. *Progress in Aerospace Sciences*, 45(1-3):50–79 (2009)ifnextchar.gobble.
17. Hartmann, D., Herz, M., and Wever, U. “Model order reduction a key technology for digital twins”. In *Digital Twin: The Vision and the Future*, pages 167–179. Springer, Cham (2018)ifnextchar.gobble.
18. “Calculation of load capacity of spur and helical gears”. ISO 6336 (2006)ifnextchar.gobble.
19. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., VanderPlas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. “Scikit-learn: Machine learning in python”. *CoRR*, abs/1201.0490 (2012)ifnextchar.gobble.
20. Chollet, F. et al. “keras” (2015)ifnextchar.gobble.
21. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X. “Tensorflow: Large-scale machine learning on heterogeneous distributed systems” (2016)ifnextchar.gobble.
22. Bruzzone, F., Maggi, T., Marcellini, C., Rosso, C., and Delprete, C. “Proposal of a novel approach for 3d tooth contact analysis and calculation of the static transmission error in loaded gears”. *Procedia Structural Integrity*, 24:178–189 (2019)ifnextchar.gobble.
23. Bruzzone, F., Maggi, T., Marcellini, C., and Rosso, C. “Gear teeth deflection model for spur gears: Proposal of a 3d nonlinear and non-hertzian approach”. *Machines*, 9(10) (2021)ifnextchar.gobble.
24. Rosso, C. and Bruzzone, F. “On the influence of the actual load sharing factor in increasing the power density in gearboxes”. In Petuya, V., Quaglia, G., Parikyan, T., and Carbone, G., editors, *Proceedings of I4SDG Workshop 2023*, pages 420–429, Cham (2023)ifnextchar.gobble. Springer Nature Switzerland.
25. Cain, M.K., Zhang, Z., and Yuan, K.H. “Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation”. *Behavior Research Methods*, 49(5):1716–1735 (2017)ifnextchar.gobble.
26. Yeo, I.K. and Johnson, R.A. “A new family of power transformations to improve normality or symmetry”. *Biometrika*, 87(4):954–959 (2000)ifnextchar.gobble.
27. Shrestha, N. “Detecting multicollinearity in regression analysis”. *American Journal of Applied Mathematics and Statistics*, 8(2):39–42 (2020)ifnextchar.gobble.
28. Kim, J.H. “Multicollinearity and misleading statistical results”. *Korean Journal of Anesthesiology*, 72(6):558–569 (2019)ifnextchar.gobble.
29. Pearson, K. “Note on regression and inheritance in the case of two parents”. *Proceedings of the Royal Society of London*, 58:240–242 (1895)ifnextchar.gobble.
30. Iman, R.L., Davenport, J.M., and Zeigler, D.K. “Latin hypercube sampling (program user’s guide). [lhc, in fortran]”. Technical report, Sandia Labs., Albuquerque, NM (USA) (1979)ifnextchar.gobble.
31. Liu, J., Liang, G., Siegmund, K.D., and Lewinger, J.P. “Data integration by multi-tuning parameter elastic net regression”. *BMC Bioinformatics*, 19(1):369 (2018)ifnextchar.gobble.
32. Sutskever, I., Vinyals, O., and Le, Q.V. “Sequence to sequence learning with neural networks” (2014)ifnextchar.gobble.
33. Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. “Empirical evaluation of gated recurrent neural networks on sequence modeling” (2014)ifnextchar.gobble.
34. Balti, H., Abbes, A.B., and Farah, I.R. “A bi-gru-based encoder–decoder framework for multivariate time series forecasting”. *Soft Computing*, 28(9):6775–6786 (2024)ifnextchar.gobble.