Journal of Electrical Engineering and Information Technologies, Vol. 10, No. 1 pp. 12-18 (2025)
Article 227
Received: 15.04.2025
Accepted: 06.06.2025
In print: ISSN 2545–4250
On line: ISSN 2545–4269
UDC: 631.311.011.7:004.85.032.26

Original scientific paper

COMPARISON BETWEEN CONTRASTIVE LEARNING AND RECURRENT NEURAL NETWORKS FOR POWER SYSTEM INERTIA ESTIMATION

Pande Popovski[1,2)] Anton Chaushevski[2)]

[1] University St Kliment Ohridski, Bitola, Faculty of Technical Sciences, [2]Ss Cyril and Methodius University in Skopje, Faculty of Electrical Engineering and Information Technologies, 1000 Skopje, Republic of North Macedonia

[1)] popovski@uklo.edu.com

Abstract: The lack of power system inertia is becoming a potential issue as penetration of renewable energy sources in the power system increases. This is a result of an agenda set at worldwide level, to maximize integration of renewables and turn away from fossil fuels. Along with the potential problem of lack of power system inertia comes the difficulty of estimating equivalent power system inertia in a system that is becoming increasingly influenced by power electronics. While model-based analyses are possible, they do become increasingly difficult to solve. As a way to circumvent the inconvenience of estimating equivalent power system inertia, Machine Learning has proven to be a viable option. Recurrent, Convolutional, Physics Informed Neural Networks, including other types of regression focused approaches have been previously analyzed on this topic, and proven to be potentially useful. This paper makes a comparison between two approaches to estimation of equivalent power system inertia. The first approach is proposed by the authors, and it involves combination of Contrastive Learning and Ridge Regression. The second approach is Recurrent Neural Networks, which have been previously implemented on this kind of problem. Both methods are tested on simulated data from the IEEE 24-bus system. Different performance metrics are compared, on different dataset sizes. The results obtained from the study show that the method proposed by the authors produces better results in cases when there is deficiency of training data, leading to the conclusion that the proposed methodology may be potentially useful for such cases.

Keywords: Contrastive Learning, Neural Networks, Power System Inertia, Regression Analysis, Self-Supervised Learning.

СПОРЕДБА ПОМЕЃУ КОНТРАСТНО УЧЕЊЕ И РЕКУРЕНТНИ НЕВРОНСКИ МРЕЖИ ЗА ПРОЦЕНКА НА ИНЕРЦИЈА КАЈ ЕЛЕКТРОЕНЕРГЕТСКИ СИСТЕМИ

Апстракт: Со зголемувањето на уделот на обновливи извори на енергија во електроенергетскиот систем, се појавува потенцијалниот проблем со недостатокот на инерција кај електроенергетскиот систем. Ова е резултат на светската агенда за максимизирана интеграција на обновливи извори на енергија и отфрлање на фосилните горива како енергенс. Проблемот со инерцијата е проследен и со потешкотии при пресметката на еквивалентната инерција на електроенергетскиот систем, заради прогресивното зголемување на влијанието на енергетската електроника при производството на електрична енергија. Иако анализи што се темелат на математички модели се возможни, истите стануваат покомплексни и воедно потешки за решавање. Машинското учење се појавува како опција за заобиколување на овој проблем. Рекурентните, конволуционите, физички информираните невронски мрежи, вклучувајќи и други пристапи, биле претходно анализирани во контекст на оваа проблематика, и се докажало дека имаат потенцијална примена. Овој труд прави компарација помеѓу два пристапи за проценка на еквивалентната инерција во електроенергетскиот систем. Првиот пристап е предложен од страна на авторите, и опфаќа употреба на комбинација од контрастно учење и “Ridge” регресија. Вториот пристап ги опфаќа рекурентните невронски мрежи, кои веќе биле употребувани за ваков тип на проблем. Двете методи се тестирани на синтетички податоци добиени од симулација на IEEE тест мрежата со 24 јазли. Перформансите на двата пристапи се споредени. Добиените резултати покажуваат дека предложениот пристап од страна на авторите дава подобри резултати во случаеви кога има недостаток на податоци за тренирање на моделите, што наведува до заклучок дека методологијата можеби има примена во вакви ситуации.

Клучни зборови: контрастно учење, невронски мрежи, инерција кај електроенергетски системи, регресивна анализа, самонадгледувано учење.

I. INTRODUCTION

ONE of the most prominent and widely recognized trends in the energy sector today is the ongoing transition toward widespread integration of renewable energy sources (RES) into modern power systems. This shift is primarily driven by the need to reduce reliance on fossil fuels for electricity generation, thereby promoting cleaner and more sustainable energy alternatives. As part of this global movement, considerable efforts have been directed toward maximizing deployment of RES technologies, with the vision of achieving 100% RES penetration becoming increasingly familiar within the industry [1]. Among these technologies, photovoltaic (PV) power plants have received particular attention due to their scalability and falling costs [2]. Notably, developed nations such as Germany have made substantial progress in this regard, reaching a renewable energy share of approximately 51.6% in their electricity mix as early as 2021 [3].

However, despite their numerous advantages, renewable energy sources (RES) also present certain challenges that must be carefully addressed. A key limitation lies in their inherent intermittency, .i.e. many RES technologies, such as solar and wind, are subject to fluctuations in power output due to their dependence on weather and environmental conditions. As a result, they are often unable to provide consistent and controllable supply of electricity. This variability can pose difficulties for grid stability and energy planning, particularly as RES penetration increases. Consequently, a lot of research has emerged focusing on forecasting RES generation in order to enhance predictability and enable better integration into power systems [4] [5].

Another equally significant challenge introduced by the increasing penetration of RES is the reduction of power system inertia. Inertia is a fundamental property that contributes to transient stability of the power system by resisting sudden changes in frequency. This issue arises primarily because most RES technologies are interfaced with the grid through power electronic devices, in particular Power Inverters (PIs) [6]. Unlike traditional Synchronous Machines (SM), which inherently possess rotating masses that store kinetic energy, PIs lack physical inertia and operate on entirely different principles. As a result, they do not naturally contribute to system inertia. To address this shortcoming, researchers have been actively developing advanced control strategies that enable PIs to emulate the inertial behavior of synchronous generators. One such approach that has gained increasing attention is known as Virtual Inertia (VI) [7], which aims to replicate the stabilizing effect of mechanical inertia through algorithmic control.

Therefore, implementing control strategies that allow PIs to provide VI represents a promising and practical solution for mitigating the loss of power system inertia in RES-dominated grids. While effective in principle, these control schemes significantly increase the complexity of underlying mathematical models used to represent the power system, especially when performing dynamic analyses or stability assessments. This added complexity becomes especially pronounced when attempting to estimate the power system's equivalent inertia. The inclusion of VI through inverter-based resources introduces additional layers of control dynamics and nonlinear behavior, making traditional estimation methods more challenging and less straightforward.

As a result, Machine Learning (ML) has emerged as a powerful and increasingly popular tool for addressing this challenge. Rather than relying solely on intricate and often computationally intensive mathematical formulations, ML-based approaches aim to learn underlying patterns and relationships within system data. By doing so, they offer an alternative means of estimating power system inertia, one that can potentially provide accurate predictions without the need for explicit modeling of the system's physical dynamics. This data-driven perspective is particularly appealing in modern, complex grids where traditional analytical methods may fall short or become impractical.

Numerous research has been dedicated to application of ML techniques for power system inertia estimation [8] [9]. Among the most prominent and widely adopted approaches in recent literature are Neural Networks (NN) [10] [11], particularly specialized architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) [12] to [15], and Residual Neural Networks (ResNets) [16]. These models have demonstrated strong potential in capturing complex temporal and spatial patterns from power system data. An advanced extension of neural networks that has gained considerable attention is the Physics-Informed Neural Network (PINN) [17]. Unlike conventional NNs that rely purely on data-driven learning through statistical loss functions, PINNs incorporate additional physical constraints typically represented through domain-specific equations such as the swing equation used for modeling SMs into the loss function. This hybrid approach has shown promising results in several studies [18] to [21], particularly in scenarios where training data is limited. In the context of inertia estimation, PINNs offer a compelling solution by embedding physical laws directly into the learning process. However, successful implementation requires accurate measurements of system frequency and the Rate of Change of Frequency (RoCoF), as well as reliable values for system parameters such as inertia constant and damping coefficient, which are essential for guiding the network's learning in a meaningful way.

One of the primary obstacles hindering effective application of NNs for power system inertia estimation is scarcity of high-quality measurement data. Typically, input data required for training such models is obtained from Phasor Measurement Units (PMUs), which provide time-synchronized measurements of key electrical quantities such as voltage, current, and frequency. However, in many power systems, especially in developing regions or at the distribution level, the number of installed PMUs may be limited, resulting in insufficient data to adequately train data-hungry neural network models. As previously noted, PINNs offer a potential workaround by incorporating physical knowledge into the learning process, thereby reducing reliance on large datasets. Nonetheless, this approach is not universally effective, and its performance can still be constrained in scenarios where critical measurements or system parameters are missing or unreliable.

In this paper, an approach specifically designed to address challenges arising from limited availability of training data is proposed, which is a common constraint in many real-world power systems. Additionally, the proposed approach's performance is tested against a well-established method - RNNs, to provide a comprehensive evaluation. Key contributions of this work can be summarized as follows:

- we introduce a novel approach aimed at mitigating effects of training data scarcity in ML applications for power system inertia estimation;

- we employ Contrastive Learning (CL), a self-supervised learning technique, to extract meaningful representations and identify underlying patterns within training data;
- we leverage embeddings learned through CL in a Ridge Regression (RR) framework, enabling development of a predictive model that estimates the inertia coefficient based on key power system measurements, including angular frequency, Rate of Change of Frequency (RoCoF), and node voltage.

The rest of the paper is organized as follows: Section II provides a brief introduction to Contrastive Learning and TS2Vec framework. In Section III, we explain data pre-processing and methodology for training the models proposed in the paper. Section IV presents a case study, and Section V concludes the paper.

II. Contrastive Learning for Time Series

As outlined in [22], ML is traditionally categorized into four main types: Supervised Learning, Semi-Supervised Learning, Unsupervised Learning, and Reinforcement Learning. The primary distinction between Supervised and Unsupervised Learning lies in the presence of labeled data, i.e. Supervised Learning relies on labeled datasets to train models, whereas Unsupervised Learning operates without them. A notable limitation of Supervised Learning, particularly in engineering applications, is the need for domain experts to manually label large volumes of data. This process is often labor-intensive, prone to human error, and may not always be feasible, especially when dealing with complex systems or large-scale datasets. To overcome these challenges, another paradigm known as Self-Supervised Learning has gained attention. This approach enables models to learn useful representations through pretext tasks that exploit inherent structures or correlations within input data, eliminating the need for manually assigned labels. As a result, Self-Supervised Learning can significantly reduce the burden of data annotation while still achieving high levels of performance in downstream tasks.

Although the concepts of Self-Supervised and Semi-Supervised Learning have been established for quite some time and are not considered new developments [23], they have recently gained renewed attention due to their applicability in modern Machine Learning challenges. Contrastive Learning, a subfield within Self-Supervised Learning, focuses on identifying and learning meaningful relationships within un-labelled data by distinguishing between similar and dissimilar data pairs. This technique has proven effective across a wide range of domains. In particular, CL has played a pivotal role in advancing generative AI, which has seen widespread adoption among the general public. Its most prominent applications are found in areas such as natural language processing, computer vision, audio signal processing, and beyond [22] [23], where learning from un-labelled data at scale has become increasingly valuable.

Another domain where CL has shown significant promise, particularly relevant to this paper, is in the analysis of time series data. Time series play a crucial role in many real-world applications, including energy demand forecasting, financial market analysis, climate modeling, and virtually any context where observed quantities are strongly influenced by temporal dynamics. In this context, the authors of [24] introduced *TS2Vec*, a general-purpose framework specifically designed for learning robust representations from time series data using CL principles. At the heart of TS2Vec lies a hierarchical contrastive learning mechanism, which aims to extract meaningful features from subsequences of varying lengths and scales. This allows the model to learn contextual representations across different semantic levels, making it highly adaptable to diverse temporal patterns. Unlike traditional methods that may require handcrafted features or domain-specific preprocessing, TS2Vec is designed to be model-agnostic and data-efficient, requiring minimal supervision. According to its authors, the framework supports a broad range of downstream tasks, including time series classification, forecasting, and anomaly detection, making it a powerful tool for applications where labeled data is limited but temporal dependencies are rich and informative.

In this study, TS2Vec framework is selected for implementation as a key component of the proposed analysis. Its flexibility and effectiveness in learning representations from time series data make it well-suited for the task of power system inertia estimation, where capturing temporal dependencies is essential. TS2Vec is available as an open-source Python module..

III. Data Preprocessing & Methodology

This section outlines the structure of the dataset as well as the methodology used to develop the models aligned with the objectives of this study. The dataset utilized originates from [15] and consists of synthetic data generated through dynamic simulations conducted using Simulink model of the IEEE 24-bus power system. These simulations were designed to emulate realistic transient behavior under varying system conditions. Detailed description of the simulation setup can be found in the original source. For reproducibility and further experimentation, link to the dataset is provided in the references [25].

For each PMU installed in the system, three measurements are recorded: angular frequency (ω), Rate of Change of Frequency (RoCoF, denoted as $d\omega/dt$), and voltage (υ). These variables constitute the features of the dataset. Each of the three measurements is treated as an individual feature per PMU, resulting in a total of 24 features for the eight PMUs installed across the IEEE 24-bus test system. This structured representation allows the model to capture spatially distributed dynamic behavior across the grid. The target variable, or label, for each data sample is the value $M = 2{\cdot}H$, where H represents the equivalent power system inertia constant.

Each time sequence in the dataset spans a duration of one second, during which 200 samples are collected at a sampling rate of 200 Hz. Within a given sequence, the target value MMM remains constant; however, the feature values ω, $d\omega/dt$, and υ vary throughout the time window. This variability is introduced by applying a probing signal to the system, which induces small disturbances during the observation period. These disturbances are designed to mimic realistic grid fluctuations and enable the model to learn meaningful dynamic patterns across time. A more detailed explanation of the signal injection process and simulation methodology is provided in [15].

The simulation procedure is carried out for 11 distinct values of *M*, ranging from 3 seconds to 8 seconds. To introduce variability and enrich the dataset, 100 different magnitudes of the probing signal are applied for each value of *M*, resulting in a total of 1,100 unique time sequences. Consequently, the complete dataset comprises of 1,100 observations, where each observation includes 24 features and a corresponding label *M*. Each feature represents a time-dependent signal with sequence length of 200 samples. Formally, if the number of observations is denoted as *J*, the sequence length as *K*, and the number of features as *L*, then the input features can be structured and fed into the learning algorithm as a three-dimensional array:

$$A \in \mathbb{R}^{J \times K \times L} \tag{1}$$

Correspondingly, the target values *M* form a one-dimensional array of length *J*, since each time sequence is associated with a single, constant inertia value.

The overall training process of the proposed method is divided into two main stages:

1) training a Contrastive Learning model using training dataset;
2) training a Ridge Regression model using the feature representations obtained from CL stage.

In the first stage, the training data - consisting solely of input features without labels - is fed into the CL algorithm. The objective at this stage is to learn meaningful representations by capturing intrinsic relationships within the data, such as temporal dependencies, multivariate interactions across features, and both local and global dynamics within each time sequence. The output of this stage is a compressed embedding for each time sequence, which encapsulates the learned structure of the data in a lower-dimensional space. In the second stage, these embeddings are paired with their corresponding labels *M*, and used to train the RR model. This regression model then learns to map the extracted representations to the system inertia values, and is subsequently evaluated on validation dataset to assess its performance.

To evaluate effectiveness of the proposed CL/RR approach, its performance is compared against that of RNNs, which serve as baseline method. Both models are trained and tested on identical datasets to ensure fair comparison. To further assess robustness and generalization capability under varying data availability, multiple models are trained across different dataset sizes. This approach allows for a more comprehensive understanding of how each method performs under data-rich and data-scarce conditions. Detailed information regarding the number of trained models, the specific dataset partitions used, and the corresponding performance metrics is provided in Section IV.

Implementation of the proposed model is carried out in Python, utilizing TS2Vec framework for contrastive representation learning and Scikit-Learn library for training the RR model. For baseline comparison, the RNN model is developed using the PyTorch package. To ensure consistency and reproducibility across experiments, a fixed random seed value of 101010 is used during the training of both CL-based models and RNNs. Additionally, all feature values and corresponding labels are normalized to the range [0, 1] prior to training, in order to standardize input data and facilitate stable model convergence.

IV. Case Study

As outlined in the previous section, the dataset employed in this study is derived from simulations conducted on Simulink model of the IEEE 24-bus power system. Measurements are collected from eight distinct PMUs placed throughout the network. Since each PMU provides three measurements, ω, $d\omega/dt$, and υ, the resulting dataset contains a total of 24 features per observation. In total, the dataset comprises of 1,100 time-sequences, each representing a one-second window sampled at a rate of 200 Hz. This results in 200 time-steps per observation, capturing the system's transient behavior in high temporal resolution.

To investigate the impact of dataset size on model performance, experiments are conducted using varying proportions of the full dataset. Models are trained on subsets of the data, with dataset sizes expressed as percentages of the original 1,100 observations. This approach enables evaluation of how each model, both CL/RR and RNN, scales with data availability and how performance degrades or improves under limited data conditions. The specific dataset sizes used in the experiments are summarized in Table I.

Furthermore, each of the dataset subsets listed in Table I is further divided into training and validation sets using an 80/20 split ratio. This ensures that the models are evaluated on previously unseen data, allowing for reliable assessment of their generalization capabilities across different dataset sizes.

For each of the five dataset sizes presented in Table I, 10 models are trained using both CL/RR approach and RNN-based model. This results in 100 trained models in total. Performance of the models is assessed using four key evaluation metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and validation accuracy, defined as the percentage of predictions falling within a ±0.5 s tolerance range of the true *M* value.

Table I
Cases For Different Dataset Sizes

Case #	1	2	3	4	5
Size [%]	20	40	60	80	100
Observations	220	440	660	880	1100

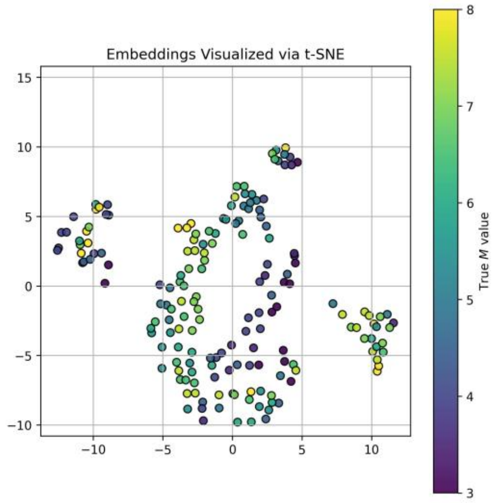
Since each model configuration is trained 10 times, the median value of each metric across the 10 runs is reported to account for variability and provide a more representative measure of performance. The summarized results are presented in Table II.

TABLE II
RESULT FROM VALIDATION OF MODELS

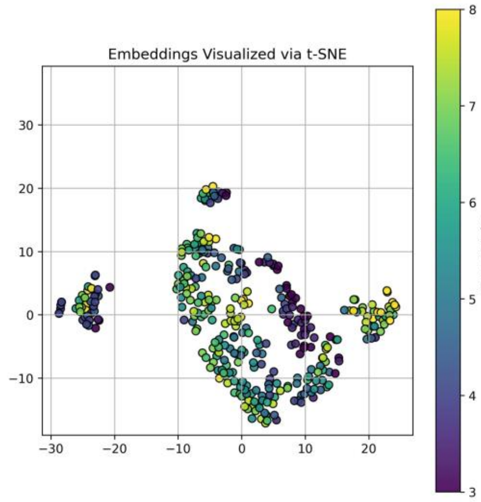
Case	Model	MSE	MAE	R^2	Accuracy [%]
1	RNN	0.919	0.812	-0.031	34.1
	CL/RR	0.1	0.244	0.888	88.6
2	RNN	0.749	0.737	0.338	30.7
	CL/RR	0.114	0.228	0.899	86.4
3	RNN	1.169	0.927	-0.192	26.5
	CL/RR	0.098	0.235	0.9	90.2
4	RNN	0.018	0.092	0.981	98.6
	CL/RR	0.068	0.174	0.93	92.1
5	RNN	0.027	0.112	0.976	97.7
	CL/RR	0.043	0.149	0.961	95.9

The results presented in Table II demonstrate that the proposed CL/RR models outperform the RNN-based models in scenarios where the dataset size is limited. Even at 60% of the original dataset, RNN models exhibit poor performance, highlighting their sensitivity to data scarcity. To assess model stability, the Median Absolute Deviation (MAD) is calculated alongside the median value for each metric, based on the 10 training runs per case. The results show that most models exhibit minimal deviation from the median, indicating consistent performance. The only notable exception is RNN models trained on 80% of the dataset, which display slightly higher, but still relatively small variations. As these deviations are negligible and do not significantly alter the conclusions, they are not reported in Table II. Interestingly, in cases 4 and 5, where larger amounts of training data are available, RNN models begin to outperform CL/RR approach, suggesting that RNNs may be more effective when sufficient data is present.

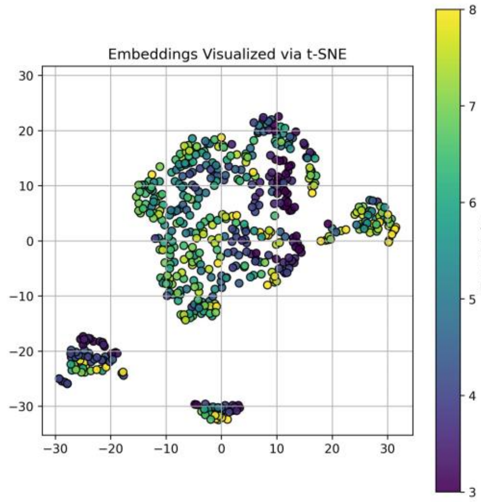
Effectiveness of the proposed methodology on smaller datasets is further supported by the visualization shown on Figure 1, which presents a t-SNE plot of learned embeddings. These embeddings are taken from three of the training runs of the CL model corresponding to cases 1, 2 and 3 respectively.



a)



b)

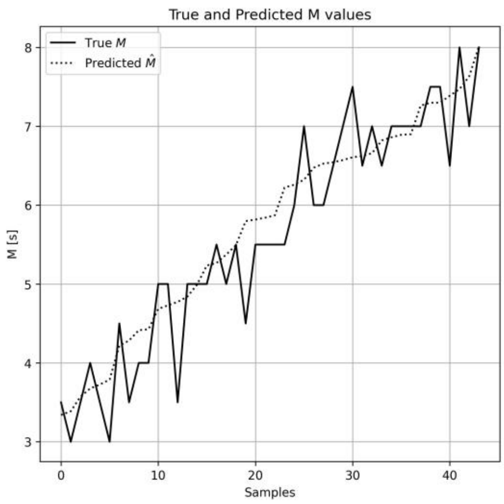


c)

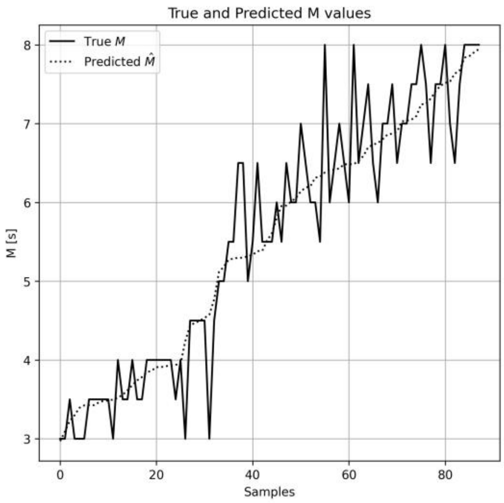
Fig. 1. Embeddings obtained from CL. a) Case 1; b) Case 2; c) Case 3

Despite the limited amount of training data, t-SNE visualization reveals that the model is capable of distinguishing meaningful structures within the data, as evidenced by the formation of distinct clusters. This clustering behavior indicates that the Contrastive Learning stage successfully captures underlying relationships in the time series, even under data-constrained conditions, further validating the model's robustness.

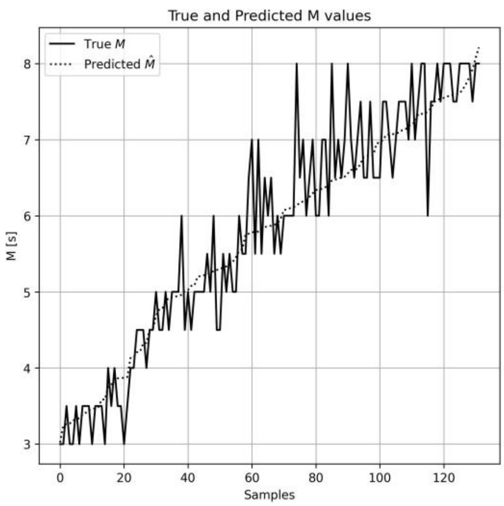
To further demonstrate, Figure 2 shows plots of predicted and real values of M from the validation of the same cases.



a)



b)



c)

Fig. 2 True and predicted *M* values. a) Case 1; b) Case 2; c) Case 3

V. CONCLUSION

This paper presents a methodology for addressing data scarcity in the context of power system inertia estimation. The proposed approach, which combines Contrastive Learning for representation extraction with Ridge Regression for model training, was evaluated against traditional Recurrent Neural Network models across various dataset sizes.

The results demonstrate that CL/RR models outperform RNNs in scenarios with limited data availability, successfully capturing meaningful patterns even when trained on as little as 20% of the original dataset. However, as the amount of available data increases, RNN models begin to show marginally better performance, reflecting their strength in data-rich environments.

These findings suggest that the proposed CL/RR methodology may hold potential for practical deployment, particularly in real-world settings where measurement data is scarce and analytical estimation of equivalent power system inertia is challenging. By leveraging Self-Supervised Learning to uncover temporal and multivariate relationships, this approach offers a promising alternative for accurate inertia estimation in modern, inverter-dominated grids.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During preparation of this work, the authors used ChatGPT to improve language and readability of work presented, to a limited extent. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

REFERENCES

[1] C. Breyer et al. "On the History and Future of 100% Renewable Energy Systems Research". IEEE Access, vol.10, pp.: 78176–78218, 2022, doi: 10.1109/ACCESS.2022.3193402.

[2] N. Haegel and S. Kurtz. "Global Progress Toward Renewable Electricity: Tracking the Role of Solar". IEEE Journal of Photovoltaics, vol.11, no.6, pp.: 1335–1342, November 2021, doi: 10.1109/JPHOTOV.2021.3104149.

[3] H. He et al. "Systematic Evaluation and Review of Germany Renewable Energy Research: A Bibliometric Study from 2008 to 2023". Heliyon, vol.10, no.15, p.: e34933, August 2024, doi: 10.1016/j.heliyon.2024.e34933.

[4] Y.-L. Li, Z.-A. Zhu, Y.-K. Chang, and C.-K. Chiang. "Short-Term Wind Power Forecasting by Advanced Machine Learning Models" in 2020 International Symposium on Computer, Consumer and Control (IS3C), IEEE, Nov.2020, pp.: 412–415. doi: 10.1109/IS3C50286.2020.00112.

[5] N. Jannah, T. S. Gunawan, S. H. Yusoff, M. S. A. Hanifah, and S. N. M. Sapihie. "Recent Advances and Future Challenges of Solar Power Generation Forecasting". IEEE Access, vol.12, pp.: 168904–168924, 2024, doi: 10.1109/ACCESS.2024.3496120.

[6] Tang, Z. et al. "Power Electronics - The Enabling technology for Renewable Energy Integration". CSEE Journal of Power and Energy Systems, 2021, doi: 10.17775/CSEEJPES.2021.02850.

[7] U. Tamrakar, D. Shrestha, M. Maharjan, B. Bhattarai, T. Hansen, and R. Tonkoski. "Virtual Inertia: Current Trends and Future Directions". Applied Sciences, vol.7, no,7, p.: 654, June 2017, doi: 10.3390/app7070654.

[8] A. O. Muhammed, Y. J. Isbeih, M. S. el Moursi, and K. Elbassioni. "Artificial Intelligence (AI) Advanced Techniques for Real-Time Inertia Estimation in Renewable-Based Power Systems". IEEE Transactions on Industry Applications, vol.61, no.2, pp.: 2604–2619, March 2025, doi: 10.1109/TIA.2024.3462690.

[9] M. Heidari et al. "A Review on Application of Machine Learning-Based Methods for Power System Inertia Monitoring". International Journal of Electrical Power & Energy Systems, vol.162, p.: 110279, November 2024, doi: 10.1016/j.ijepes.2024.110279.

[10] A. J. Hernandez M and R. T. Pinto. "Artificial Neural Network-Based Inertia Estimation" in 2022 IEEE Power & Energy Society General Meeting (PESGM), IEEE, July 2022, pp.: 1–5. doi: 10.1109/PESGM48719.2022.9916665.

[11] D. A. Kisinga, P. Makolo, and P. Trodden. "A Neural Network-based Online Estimation of Stochastic Inertia in a Power System" in

2022 IEEE PES/IAS Power Africa, IEEE, August 2022, pp.: 1–5. doi: 10.1109/PowerAfrica53997.2022.9905270.

[12] D. Linaro et al. “Continuous Estimation of Power System Inertia Using Convolutional Neural Networks”. Nature Communications, vol.14, no.1, p.: 4440, July 2023, doi: 10.1038/s41467-023-40192-2.

[13] A. Poudyal, R. Fourney, R. Tonkoski, T. M. Hansen, U. Tamrakar, and R. D. Trevizan. “Convolutional Neural Network-Based Inertia Estimation Using Local Frequency Measurements” in 2020 52nd North American Power Symposium (NAPS), IEEE, April 2021, pp:. 1–6. doi: 10.1109/NAPS50074.2021.9449814.

[14] M. Tuo and X. Li. “Long-Term Recurrent Convolutional Network-Based Inertia Estimation Using Ambient Measurements” in 2022 IEEE Industry Applications Society Annual Meeting (IAS), IEEE, October 2022, pp.: 1–6. doi: 10.1109/IAS54023.2022.9940112.

[15] M. Tuo and X. Li. “Machine Learning Assisted Inertia Estimation Using Ambient Measurements”. IEEE Transactions on Industry Applications, pp.: 1–10, 2023, doi: 10.1109/TIA.2023.3269732.

[16] M. Ramirez-Gonzalez, F. R. S. Sevilla, and P. Korba. “Power System Inertia Estimation Using a Residual Neural Network Based Approach” in 2022 4th Global Power, Energy and Communication Conference (GPECOM), IEEE, June 2022, pp.: 355–360. doi: 10.1109/GPECOM55404.2022.9815784.

[17] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis. “Physics-Informed Neural Networks for Power Systems” in 2020 IEEE Power & Energy Society General Meeting (PESGM), IEEE, August 2020, pp.: 1–5. doi: 10.1109/PESGM41954.2020.9282004.

[18] J. Stiasny, G. S. Misyris, and S. Chatzivasileiadis. “Physics-Informed Neural Networks for Non-linear System Identification for Power System Dynamics” in 2021 IEEE Madrid PowerTech, IEEE, June 2021, pp.: 1–6. doi: 10.1109/PowerTech46648.2021.9495063.

[19] O. E. Egbomwan, S. Liu, and H. Chaoui. “Physics-Informed Neural Network for Inertia Estimation of Power System with Inverter-Based Distributed Generation” in 2024 IEEE 7th International Conference on Industrial Cyber-Physical Systems (ICPS), IEEE, May 2024, pp.: 1–6. doi: 10.1109/ICPS59941.2024.10639993.

[20] S. Stock, D. Babazadeh, C. Becker, and S. Chatzivasileiadis. “Bayesian Physics-Informed Neural Networks for System Identification of Inverter-Dominated Power Systems”. Electric Power Systems Research, vol.235, p.:110860, October 2024, doi: 10.1016/j.epsr.2024.110860.

[21] R. Plant, D. Babazadeh, S. Stock, and C. Becker. “Real-Time Inertia Estimation in an Inverter-Dominated Distribution Grid Using a Physics-Informed Recurrent Neural Network”. IET Conference Proceedings, vol.2022, no.3, pp.: 940–944, June 2022, doi: 10.1049/icp.2022.0852.

[22] S. Tripathi and C. R. King. “Contrastive Learning: Big Data Foundations and Applications” in Proceedings of the 7th Joint International Conference on Data Science & Management of Data (11th ACM IKDD CODS and 29th COMAD), New York, NY, USA: ACM, January 2024, pp.: 493–497. doi: 10.1145/3632410.3633291.

[23] P. H. Le-Khac, G. Healy, and A. F. Smeaton. “Contrastive Representation Learning: A Framework and Review”. IEEE Access, vol.8, pp.: 193907–193934, 2020, doi: 10.1109/ACCESS.2020.3031549.

[24] Z. Yue et al. “TS2Vec: Towards Universal Representation of Time Series”. Proceedings of the AAAI Conference on Artificial Intelligence, vol.36, no.8, pp.: 8980–8987, June 2022, doi: 10.1609/aaai.v36i8.20881.

[25] https://rpglab.github.io/papers/MJ-Tuo_ML-Inertia-Est/