Transfer Learning Mechanism to account Performance Degradation in Gas Turbines with limited Operational Data

Sundar Raj Thangavelu^a, Imantha Gunasekera^b, Alessio Tafone^c, Shigenori Morita^d

^a Nanyang Technological University, Singapore, <u>sundar.rt@ntu.edu.sg</u> CA
^b Nanyang Technological University, Singapore, <u>imantha@ntu.edu.sg</u>
^c Nanyang Technological University, Singapore, <u>alessio.tafone@ntu.edu.sg</u>
^d Surbana Jurong Consultants Pte Ltd, Singapore, <u>shigenori.morita@surbanajurong.com</u>

Abstract:

This paper investigates the transfer learning mechanism to improve the prediction accuracy of the energy system model. Artificial intelligence techniques are increasingly being adopted in the energy domain to predict energy system characteristics and performance. However, in many energy systems, the relationship between interested variables and their distributions differs (data and concept drift) with time due to system degradation and aging. There is a requirement for re-training and re-testing AI models to ensure reliable performance over time, which may require extensive latest operational data. Transfer learning helps to confront this challenge by leveraging valuable knowledge from a pre-trained model and reducing the requirement of new operational data significantly. To address these issues, this paper focuses on a gas turbine, a critical energy system widely deployed in diverse applications, and shows performance degradation over its lifetime. The energy model of the gas turbine is a multivariate type that predicts energy efficiency, fuel consumption, and heat energy based on power setpoints and weather conditions. This paper examined transfer learning mechanisms that can capture the latest characteristics of the gas turbine and their effects on prediction accuracy. The developed transfer learning model predicts fuel consumption accurately above 99%, whereas the pre-trained model under-predicts up to 4%, which may lead to suboptimal operation decisions when employed in the scheduling algorithm. Some of the other targets, such as heat energy, show marginal drift, as expected from the gas turbine characteristics. The knowledge gained from the transfer learning mechanism and its efficacy boost assists operational decisions, which helps improve energy efficiency and cost savings.

Keywords:

Transfer Learning; Machine Learning; Distributed Generation, Neural Network, Prediction Energy Model.

1. Introduction

Commercial buildings account for 32% of global energy consumption and are expected to face an annual growth of around 1.3% and 2% for the Organization for Economic Co-operation and Development (OECD) and non-OECD countries [1]. The building sector wastes around 20% of its energy due to faults in energy systems [2,3] which further emphasises the importance of assessing the operational performance of these systems. The increasing trend of energy consumption urges efficient monitoring to improve the efficiency and life span of these systems and reduce carbon footprints and unnecessary downtime. Decarbonization drives distributed energy resources (DER) comprising renewables and local generation to support diverse buildings and communities. Urban communities are moving towards local DERs to gain additional efficiency benefits and strengthen reliability and resilience. Advancements in smart grids, the Internet of Things (IoT) and artificial intelligence (AI) boost adoption to automate and optimize various energy processes. The most common applications in buildings are AI-based energy management systems [4], fault detection and diagnosis [5], and load and renewable forecasting [6]. Traditionally, physics-based models were employed to address these problems and require more specific information about buildings and technologies [7,8]. Now, AI techniques cut short and ease the development and deployment process and offer scalability to larger problems [9] through generalized models. Despite the many advantages of these approaches, a major drawback is the requirement of sufficiently large datasets to produce accurate models [10]. For example, deep learning models not only require large datasets but also extensive training times and computational resources [11], and often compiling

such a dataset can be expensive, time-consuming, and even impractical on certain occasions[12]. This study focuses on developing an AI model of a gas turbine using minimal latest operational data and exploiting most of the information from the pre-trained model.

1.1 Importance of Learning Models

The performance of energy systems degrades over time; some of the performance losses can be recovered through proper maintenance, whereas the non-recoverable losses are due to system wear and tear. Such situations may change the characteristics and performance of the energy system, i.e., the changes in the outputs (dependent variables) for given inputs (independent variables). This process is often referred to as concept drift in literature. It is pre-eminent that predictive models operating in such settings need to account for these changes to retain prediction accuracy. Learning methods are widely used to address some of the performance drift challenges involving data shortages. The learning methods apply a knowledge transfer process and achieve substantial improvements in many domains, including computer vision, natural language processing, speech recognition, bioinformatics, and reinforcement learning [5].

Adaptive learning techniques are applied in dynamic environments where the performance data (system characteristics) varies with time. Adaptive learning consists of four building modules comprising memory, change detection, learning, and loss estimation [13]. The memory module updates the latest information and simultaneously discards irrelevant old information. The change detection module characterises and quantifies concept drift by examining data and its distribution between two-time windows, as well as validating the predicted values with actual measurements. The learning module updates the prediction model through retraining and the incremental method. The retraining mode discards the existing model and develops a new model using both old and new data [14]. In contrast, incremental methods update the current model with the most recent data as it becomes available. The final module is loss estimation, which is related to performance metrics used to estimate the algorithm performance during adaptive learning. Several articles incorporated adaptive learning to account for degradation in their workflow, thus improving the efficiency of systems. These studies focus on batteries and fuel cells, which are prone to significant capacity degradation. Li et al. developed a fuel cell degradation model for a hybrid bus energy management system to account for different decay rates during varying operating conditions [15].

On the other hand, transfer learning uses the knowledge gained in one problem and applies it to a different but related problem, such that the development process is simple and eliminates the need for development from scratch and a complete dataset [16]. Lu et al. applied transfer learning to tackle limited data for improving the thermal load forecasting model and developed a similarity measurement index to select a source task that is most like the target task [17]. Chen et al. [18] proposed a neural network based on control theory for fault detection in actuators and utilized transfer learning to account for degradation in the system. In the context of degrading energy systems, transfer learning aims to update a pre-trained NN model using the latest operation data to capture the changes in the system's behaviour. This type of problem is categorized as domain adaptation, where the model inputs and outputs are still within the same space, but the distribution or system characteristics change due to ageing [19]. Figure 1 shows a general overview of the domain adaptation problem, where the neural network model is developed using a comprehensive dataset, resulting in a pretrained network. Later, this pre-trained model is adjusted or fine-tuned using the latest dataset to capture the behaviour and performance of the current state of the system. During the finetuning process, the learned weights and biases of the pre-trained model are further adjusted to minimize the loss function in relation to the latest operational data. The fine-tuning process usually leads to shorter training periods because the pretrained model has already learned useful features.



Figure. 1. General schematic of transfer learning for domain adaption problem

In energy systems, transfer learning has been applied to a variety of problems. Liu et al. designed a convolution-based neural network (CNN) to diagnose faults in building chiller plants [5]. Besides transfer learning, other methods solve similar problems involving data limitations; for example, semi-supervised learning takes advantage of unlabeled data together with some labeled data to improve prediction performance [20], but both datasets are required from the same distribution. Multi-view learning leverages multiple different feature sets or views available for additional data description [21]. This study aims to comprehend the transfer learning mechanism and develop transfer learning models to improve the prediction accuracy of the energy system. This study investigates and compares two transfer learning mechanisms that can capture the changes in the energy system characteristics (referred to as data drift) with the limited latest operational data and extract other useful information from the pre-trained model.

2. Methodology - Transfer Learning for Degraded Energy Systems

In degraded energy systems, the transfer learning mechanism is to improve the prediction accuracy of the AI model by using the latest operational data and capitalizing on useful information from the pre-trained AI model. To achieve this, firstly, an AI model was developed using the comprehensive dataset to depict the initial characteristics of the system. Later, this pre-trained model is adjusted to capture the latest characteristics as the performance of the system degrades. The adjusted model is often referred to as the fine-tuned model. In such situations, the dataset representing a degraded system may be limited; however, this requirement is acceptable as this data does not need to represent all characteristics of the system. Figure 2 shows the schematic of the transfer learning mechanism, where data processing is the starting point to extract useful information (i.e., features and targets) from the operational data and eliminate outliers and noise. Develop a multi-layer perceptron neural network (MLP-NN) model to predict energy system behaviour and performance. Data normalization was employed as a pre-processing step to improve model convergence and prediction accuracy, whilst K-Fold cross-validation was used to evaluate model performance considering the network architecture with different hidden units, which defines the complexities of the model that can be varied to achieve desired prediction accuracy. The deployment of the developed NN model in the energy management system supports scheduling algorithms in deriving key operation setpoints and dispatch decisions.



Figure. 2. Schematic of Transfer Learning Mechanism

The performance or prediction accuracy of the developed model is expected to drop over time due to changes in the actual system. Eventually, this model will become obsolete and not useful. Transfer learning helps to revitalise the prediction model using the latest operational (limited) data and critically utilize the useful information from the pre-trained model. Transfer learning can be carried out in different fashions, such as finetuning or updating the coefficients of the output layer or the whole pre-trained NN model using the latest operation data. This mechanism requires less data compared to the NN model developed from scratch. The transfer learning mechanism also helps to include or exclude targets to support any changes in the prediction model requirements. The next section will describe the workability and benefits of the transfer learning mechanism for the energy system application.

3. Case Study – Gas Turbine and Performance Prediction

Generators, renewables, and energy storage are the key distributed energy systems that help to generate, store, and support various forms of electrical and thermal energy. In the long run, some of the DERs show performance degradation due to wear and tear, especially those types of equipment with rotation or fast-moving parts. Even a non-rotating system such as solar PV and batteries shows performance degradation due

to ageing and depreciation of internal components. Interestingly, the solar PV manufacturer provides the expected performance degradation, whereas for other DERs, equipment usage and operation patterns play a dominant role in the performance degradation. This study focuses on gas turbines' critical distributed energy resources, categorized as a dispatchable system where the operation can be easily controlled based on the power requirements to fill the power deficit and synchronize with the grid frequency to absorb the fluctuations. This section will describe the necessity of a transfer learning algorithm for the gas turbine model and its benefits.



Figure. 3. Schematic of Gas turbine and AI-based prediction model

Gas turbines are widely employed to generate electrical power using clean fossil fuels such as natural gas, and even the waste heat available in the exhaust gas is recovered to support thermal loads. Figure 3 shows the schematic of a gas turbine, where fuel and air drive the turbine to generate power and exhaust gas. The heat energy in the exhaust gas has the potential to produce steam or hot water when diverted through boilers or heat exchangers. The recovered heat energy can be utilized in exhaust-driven absorption chillers to produce chilled water for air conditioning purposes [22]. Gas turbines are available in diverse capacities, ranging from a few kW to over hundred MW scale. Like any other system, a gas turbine provides high efficiency at rated conditions and low efficiency during off-load conditions. The life of the gas turbine is around 150000–200000 hours, or 20 years with regular maintenance and overhauls. Experts and manufacturers confirm that this system shows performance degradation over time; some degradation effects are recoverable through proper and regular maintenance [23], whereas performance degradation due to ageing (due to wear and tear) are permanent and cannot be recovered. Any degradation in performance could result in sub-optimal operation conditions, affect the energy cost, and incur the risk of supply shortages. Therefore, actual performance needs to be accurately accounted for to understand the operation cost of the gas turbine and decide on the right combination of systems for dispatching purposes.

Commonly, performance charts or technical data were widely used to gauge power output, fuel consumption, and fuel cost. Unfortunately, it is quite abstract and covers design conditions at standard temperature and pressure. Experts exploit physics-based models to generate performance charts for diverse conditions [24], this study utilized the gas turbine simulator adapted from TRNSYS to generate the performance data and understand the effect of degradation on the performance factors. TRNSYS is a state-of-the-art commercial simulation tool for industry and academia, based on an object-oriented approach that enables the simulation of the transient behaviour of systems focused on assessing the performance of thermal and electrical energy systems. The software is made up of two main parts: a simulation engine to solve the dynamic mathematical problem and a large library of built-in components or types (e.g., gas turbines, compressors, pumps, mixers, diverters, heat exchangers, etc.), often validated by experimental data. Type 625 has been adopted, utilizing technical data and performance maps of 3.5 MW gas turbines [25]. Exploiting the Type 625 model, the gas turbine performance parameters are derived for diverse output power and inlet air temperatures. Figure 4 shows the performance characteristics of a new gas turbine, where the fuel consumption increases with the output power and the air intake temperature. The air temperature and output power also show an effect on the exhaust gas flow and temperature.



(a) Variation in the exhaust gas temperature

(b) Variation in the exhaust gas flow





Figure. 4. Performance characteristics of Gas Turbine (without degradation) with reference to part-load behaviour and air intake (ambient) temperature

Model GTD is a multi-layer perceptron (MLP-NN)-based neural network architecture prediction model developed using the performance data (17x15 = 256 data points covering 17 part loads and 15 air temperatures) representing a new gas turbine. Data normalization is applied explicitly to ease the convergence process and prevent any bias due to different data scales. Due to a smaller number of data points, K-Fold cross-validation was employed during training to obtain a more accurate estimate of the model's performance. The model GTD was trained using an ADAM (Adaptive Moment Estimation) optimizer with mean squared error as a loss function. The training epochs and the model complexity (hidden units of 8, 16, 32, and 64 units) are varied to identify the right architecture that provides acceptable accuracy during training and validation. Finally, the Model GTD with 64 hidden units was able to predict the gas turbine outputs with the required accuracy of 0.998 R2 on the validation set.



Figure. 5. Actual Exhaust Gas and Model GTD predictions (R2: 99.8%)



Figure. 6. Actual Fuel consumption and Model GTD predictions (R2: 99.8%)

The accuracy of the developed Model GTD model is shown in Figures 5 and 6. This model helps the scheduling algorithm make optimal operational decisions on an hourly or sub-hourly basis. The normalizers and coefficients, or weights, of the prediction model are preserved for later use.

4. Performance of Pretrained and Transfer Learning Models

The Model GTD prediction is expected to deviate over time because the actual gas turbine characteristics drift due to ageing and degradation. Any inaccuracy in the prediction could lead to suboptimal or inferior scheduling or operation decisions that may result in high energy costs or a potential supply risk. To retain prediction accuracy, a new prediction model needs to be developed from scratch using the latest operational data. In a real application, the gas turbine's operation depends on the loads and operation of other integrated energy systems. Therefore, it is challenging to get the latest performance data over a wide range. The possible collection of real-time operational data is limited and not comprehensive. In such cases, the transfer learning algorithm helps to develop the prediction model for the degraded system by capitalizing on the available information in the pre-trained model and the limited new performance data. This study leverages TRNSYS software to derive the performance losses [28]. To account for the practical situation, only a few operational data points (roughly 10% of the 256 data points) of the degraded system were randomly selected for transfer learning.

Previously trained Model GTD (with the same architecture and coefficients) are employed in transfer learning. The key advantage is that the preserved coefficients of the pre-trained model were adjusted instead of learning from a fresh start.

- Model TL1 is the transfer learning model developed by updating the coefficient of the whole pretrained model (Model GTD) using the latest available operational data.
- Model TL 2 is the transfer learning model developed by updating the output layer coefficient of the pre-trained model using the latest available operational data.

Interestingly, transfer learning requires less effort at around 100 epochs for coefficient adjustment to achieve reasonable predictions.



Figure. 7. The actual and predicted gas turbine outputs (exhaust gas temperature, exhaust flow and fuel consumption)

Figure 7 shows the accuracy of three prediction models, Model GTD, Model TL1, and Model TL2, with reference to the latest operational data comprising 256 points. Model GTD failed to predict the gas turbine performance accurately, and it deviates significantly in the fuel consumption estimation (as shown in Figure 8). Even the predicted exhaust temperature values deviated moderately from the actual values. The Model TL1 provides better performance than the Model GTD, and interestingly, the predicted fuel consumption and exhaust gas flow are within the acceptable range. On the other hand, the predicted exhaust gas temperature deviates to a certain extent, especially at partial loads. Model TL2 provides superior performance over Model GTD and Model TL1. The prediction accuracy is around 99.5% (R2), within acceptable limits. Generally, Model TL2 is preferred when the user requirement (output variable or number of output variables) changes or its distribution changes. Consequently, Model TL1 is expected to perform better because all coefficients are adjusted using the latest operational data. Some of the deviations could be due to limited data.



Figure. 8. Error distribution in the predicted gas turbine outputs (exhaust gas temperature, exhaust flow and fuel consumption)

Fuel consumption is the key variable that greatly reflects the operation cost of the gas turbine. Regarding fuel consumption, the performance of Model TL1 and 2 are similar, showing marginal differences at certain data points. Deploying Model TL1 or 2 in the scheduling algorithm can improve operational decisions that lead to improvements in energy efficiency and a reduction in energy costs. Applying transfer learning regularly (once or twice) in a year is important to prevent plant-model mismatches by retaining the accuracy of the prediction model and safeguarding optimal operation decisions. Through careful deployment, this process can be automated to reduce manual involvement by predefining the data extraction, training, and update rules. This approach can be applied to other energy systems, especially those subjected to performance degradation over time, such as chiller systems, heat exchange equipment, etc. In summary, the performance of certain energy systems degrades significantly over time. Not accounting for the performance degradation could lead to suboptimal or inefficient operations of energy systems. While deriving the operational decision, the actual performance characteristics are essential to derive optimal operation set points. Therefore, the prediction model needs to be updated regularly to capture changes in the energy system's characteristics. In this context, transfer learning could assist in retraining or updating the prediction model with the limited operational data to capture and predict the performance of the degrading energy systems.

Conclusion

This study developed transfer learning models for the gas turbine system to improve prediction accuracy by eliminating plant-model mismatch and supporting operational decisions to gain energy and cost savings through efficient operations. The pretrained model cannot predict the degraded system performance and shows significant deviation, which highlights the need for transfer learning. Two transfer learning mechanisms were explored: Model TL1 was obtained by tweaking the coefficients of the whole pre-trained model, and Model TL2 was developed by tweaking the coefficients of the output layer alone. Both mechanisms utilized the latest (and limited) operational data and capitalized on most of the information from the pre-trained model. Interestingly, both models provide better accuracy than the pre-trained model. Comparing transfer learning models, TL2 outperforms TL1, which could be due to limited data availability to fine-tune all coefficients. Deploying Model TL1 or 2 in the scheduling algorithm could reflect the actual performance of the gas turbine and improve the operational or economic dispatch decisions for high energy efficiency and optimal energy cost. This approach can be applied to other energy systems whose performance characteristics are expected to change over time due to ageing and degradation.

Nomenclature

Abbreviations

ADAM	adaptive moment estimation
AI	artificial intelligence
CNN	convolution neural network
DER	distributed energy resources
IoT	internet of things
MLP	multi-layer perceptron
Model GTD	gas turbine model at design condition
Model TL1	transfer learned model by updating the coefficient of the whole pre-trained model
Model TL2	transfer learned model by updating the output layer coefficient of the pre-trained model
NN	neural network
OECD	organization for economic co-operation and development
PV	photovoltaic
R2	coefficient of determination
TL	transfer learning

References

- [1] Fan C, Xiao F, Zhao Y. A short-term building cooling load prediction method using deep learning algorithms. Appl Energy 2017;195:222–33. https://doi.org/10.1016/j.apenergy.2017.03.064.
- [2] Teke A, Timur O. Assessing the energy efficiency improvement potentials of HVAC systems considering economic and environmental aspects at the hospitals. Renewable and Sustainable Energy Reviews 2014;33:224-235. https://doi.org/10.1016/j.rser.2014.02.002.
- [3] Ramesh T, Prakash R, Shukla KK. Life cycle energy analysis of buildings: An overview. Energy Build 2010;42(10):1592-600. https://doi.org/10.1016/j.enbuild.2010.05.007.

- [4] Palensky P, Dietrich D. Demand side management: Demand response, intelligent energy systems, and smart loads. IEEE Trans Industr Inform 2011;7(3):381-88. https://doi.org/10.1109/TII.2011.2158841.
- [5] Liu J, Zhang Q, Li X, Li G, Liu Z, Xie Y, et al. Transfer learning-based strategies for fault diagnosis in building energy systems. Energy Build 2021;250:111256. https://doi.org/10.1016/j.enbuild.2021.111256.
- [6] Chahkoutahi F, Khashei M. A seasonal direct optimal hybrid model of computational intelligence and soft computing techniques for electricity load forecasting. Energy 2017;140:988–1004. https://doi.org/10.1016/j.energy.2017.09.009.
- [7] Fadli F, Rezgui Y, Petri I, Meskin N, Ahmad AM, ... Building energy management systems for sports facilities in the Gulf region: a focus on impacts and considerations. Proceedings of 38th International Conference of CIB W78; October 2021, Luxembourg.
- [8] Himeur Y, Elnour M, Fadli F, Meskin N, Petri I, Rezgui Y, et al. Al-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. Springer Netherlands; 2022. https://doi.org/10.1007/s10462-022-10286-2.
- [9] Sardianos C, Varlamis I, Chronis C, Dimitrakopoulos G, Alsalemi A, Himeur Y, et al. The emergence of explainability of intelligent systems: Delivering explainable and personalized recommendations for energy efficiency. International Journal of Intelligent Systems 2021;36:656–80. https://doi.org/10.1002/int.22314.
- [10] Mirnaghi MS, Haghighat F. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. Energy Build 2020;229:110492. https://doi.org/10.1016/j.enbuild.2020.110492.
- [11] Liu X, Yu W, Liang F, Griffith D, Golmie N. Toward Deep Transfer Learning in Industrial Internet of Things. IEEE Internet Things J 2021;8:12163–75. https://doi.org/10.1109/JIOT.2021.3062482.
- [12] Fan C, Liu Y, Liu X, Sun Y, Wang J. A study on semi-supervised learning in enhancing performance of AHU unseen fault detection with limited labeled data. Sustain Cities Soc 2021;70:102874. https://doi.org/10.1016/j.scs.2021.102874.
- [13] Gama J, Zliobaite I, Bifet A, Pechenizkiy M, Bouchachia A. A survey on concept drift adaptation. ACM Comput Surv 2014;46(4):1-37. https://doi.org/10.1145/2523813.
- [14] Gil Zeira;Oded Maimon; Mark Last; Lior. Change Detection in Classification Models Induced From Time Series Data. Data Mining in Time Series Databases 2004;57:101–25.
- [15] Li J, Yang L, Yang Q, Wei Z, He Y, Lan H. Degradation adaptive energy management with a recognition-prediction method and lifetime competition-cooperation control for fuel cell hybrid bus. Energy Convers Manag 2022;271:116306. https://doi.org/10.1016/j.enconman.2022.116306.
- [16] Pan SJ, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2010;22:1345–59. https://doi.org/10.1109/TKDE.2009.191.
- [17] Lu Y, Tian Z, Zhou R, Liu W. A general transfer learning-based framework for thermal load prediction in regional energy system. Energy 2021;217:119322. https://doi.org/10.1016/j.energy.2020.119322.
- [18] Chen H, Chai Z, Jiang B, Huang B. Data-Driven Fault Detection for Dynamic Systems with Performance Degradation: A Unified Transfer Learning Framework. IEEE Trans Instrum Meas 2021;70. https://doi.org/10.1109/TIM.2020.3033943.
- [19] Zhang Q, Tian Z, Niu J, Zhu J, Lu Y. A study on transfer learning in enhancing performance of building energy system fault diagnosis with extremely limited labeled data. Build Environ 2022;225:109641. https://doi.org/10.1016/j.buildenv.2022.109641.
- [20] Zhu X. Semi-Supervised Learning Literature Survey. 2006.
- [21] Sun S. A survey of multi-view machine learning. Neural Comput Appl 2013;23:2031–8. https://doi.org/10.1007/s00521-013-1362-6.
- [22] Thangavelu SR, Myat A, Khambadkone A. Energy optimization methodology of multi-chiller plant in commercial buildings. Energy 2017;123:64–76. https://doi.org/10.1016/j.energy.2017.01.116.
- [23] Amare Fentaye, Aklilu Tesfamichael Baheta, Syed Ihtsham UI Haq Gilani. Effects of performance deterioration on gas path measurements in an industrial gas turbine ARPN Journal of Engineering and Applied Sciences 2016;11(24):14202-7.
- [24] Tafone Alessio, Thangavelu Sundar Raj, Morita Shigenori, Romagnoli Alessandro. Design optimization of a novel cryo-polygeneration demonstrator developed in Singapore – techno-economic feasibility study for a cooling dominated tropical climate. Appl Energy 2023;330:119916.
- [25] NovaLT5 | Baker Hughes n.d. https://www.bakerhughes.com/gas-turbines/novalt-technology/novalt5 (accessed March 9, 2022).