Energy Proceedings
Vol 60, 2025
ISSN 2004-2965

An Adaptive Transfer-Learning-Assisted Modular Learning Approach for Control-Oriented Building Modelling

Zheng Wei[1], Dong Suning[1], Wan Man Pun[1], Yang Shiyu[2,3*]

1 School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore, 639798

2 State Key Laboratory of Electrical Insulation and Power Equipment, Xi'an Jiaotong University, Xi'an, China, 710049

3 School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, 710049

(Corresponding Author: yangshiyu@xjtu.edu.cn)

ABSTRACT

Machine learning (ML)-based Model Predictive Control (MPC) has shown promise for smart buildings, with online ML model adaptation often integrated to enhance long-term performance. However, abrupt building thermodynamic changes, such as opening/closing windows, can hardly been captured by the conventional online ML model adaptation. An Adaptive Transfer-Learning-Assisted Modular Learning (ATML) framework combined with dataset truncation is proposed by this paper to address the challenge. A simulation case study using a residential building was conducted for evaluation. Results show that, compared with conventional online ML model adaptation, ATML reduces prediction errors by 32.4% after 2 hours of the systemic change and by 25.4% after 22 hours, thereby demonstrating its effectiveness in coping with abrupt systemic changes.

Keywords: Model Predictive Control (MPC), Machine Learning (ML), Transfer Learning (TL)

NONMENCLATURE

Symbols	
T	Temperature (°C)
P	Electricity consumption (W)
R	Solar radiation (W/m^2)
V	Volume flow rate (m^3/s)
C	Coefficient (unitless)
A	Opening area (m^2)
q	Heat flow rate (W)
f	Fraction (unitless)
v	Heat gains (W/m^2)
g	Gravitational acceleration (m/s^2)
h	Opening height (m)
ρ	Density (kg/m^3)
c	Heat capacity (J/(kg·K))
Sub-scripts	
$gshp$	Ground Source Heat Pump
wtr	Water in the water tank
liq	Phase-change material liquid
eh	Auxiliary electric heater
amb	Ambient
act	Actual
$rdtr$	Radiator
z	Thermal zone
$solar$	Solar
IG	Internal heat gains
d	Discharge
$vent$	Ventilation
air	Air
$SHGC$	G-value
U	U-Value
$open$	Opening
$close$	Closing

1. INTRODUCTION

Energy consumption in buildings accounts for 20–40% of total energy use in developed countries [1]. This has motivated extensive research on intelligent control strategies, particular attention given to model predictive control (MPC). Numerous studies have confirmed the effectiveness of MPC in intelligent building control. Such as, to achieve more energy savings, MPC systems was implemented in an office [2] and a lecture theatre [3].

Machine learning (ML)-Based MPC frameworks that integrate Recurrent Neural Networks (RNNs) have shown promising performance on temperature regulation and energy saving compared with conventional rule-based control or PID control [4]. However, buildings are highly dynamic systems with inherent uncertainty. A fixed ML-based model often accumulates prediction errors over time, leading to model distortion and unreliable MPC control performance.

Recent studies have therefore focused on adaptive ML approaches, which continuously learn from real-time operational data to update the predictive model and

This is a paper for the 5th Applied Energy Symposium and Forum: Renewable Energy Matrix (REM2025), Oct. 29-31, 2025, Yancheng, China.

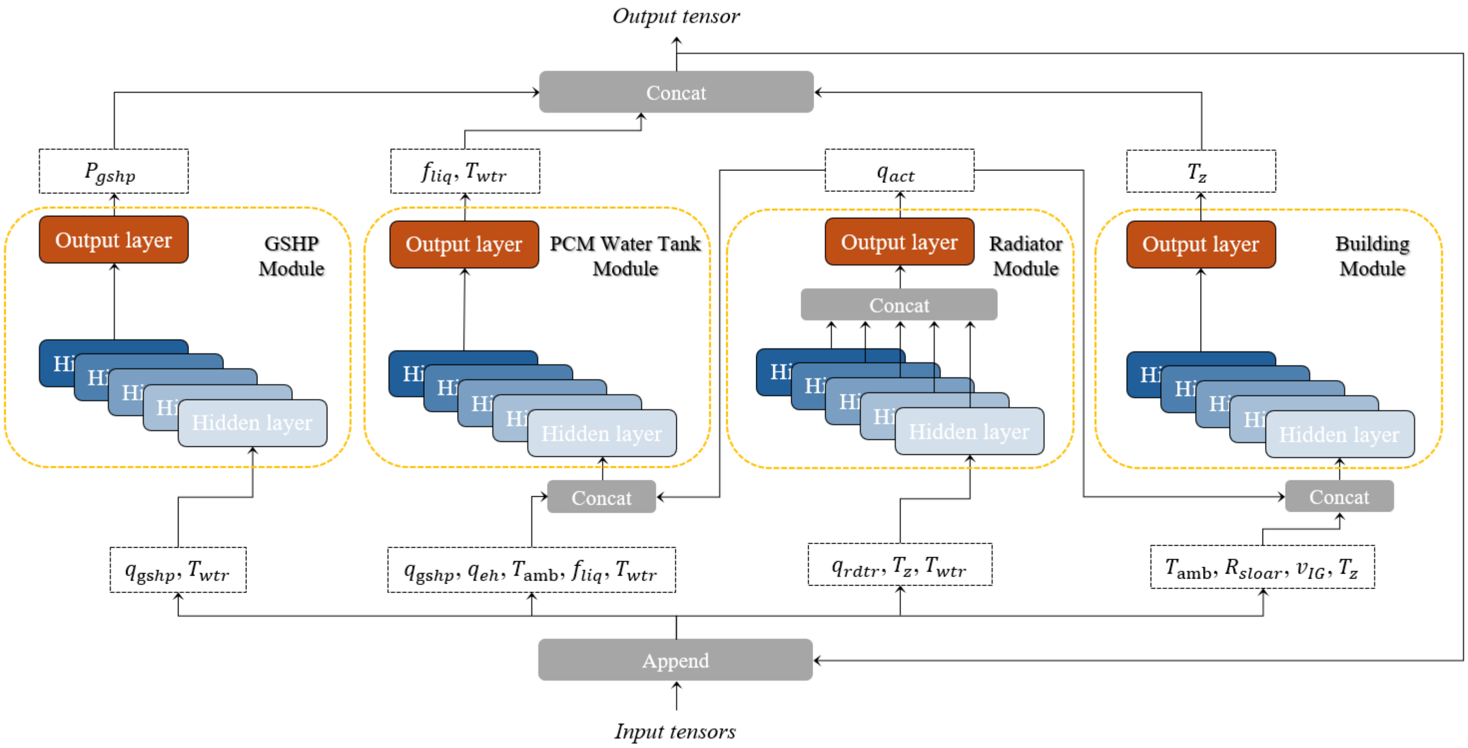

Fig.1 Schematic structure of Hybrid Recurrent Neural Network

sustain reliable, accurate long-term performance. For example, an adaptive ML-based MPC framework was developed in [5] to provide robust indoor climate control. Such adaptation strategies are effective for addressing slow time-varying characteristics (e.g., structural aging or gradual climatic shifts) by leveraging cumulative operational data. However, abrupt systemic changes (e.g., occupant interventions or major structural renovation) can induce instantaneous and substantial domain shifts [6], rendering historical data reflecting an obsolete domain and misleading the learning process. Hence, adaptive ML approaches in the literature can not handle such abrupt systemic changes.

To address the above-discussed challenge, this study proposes a novel adaptive Transfer-learning-assisted Modular Learning (ATML) framework in conjunction with dataset truncation. Dataset truncation—discarding data collected prior to the abrupt systemic changes—mitigates the negative influence of obsolete domains. ATML further enhances rapid model adaptation when only limited dataset is available due to dataset truncation.

The main contributions of this paper are as follows: (1) The influence of domain shift induced by abrupt systemic changes was investigated and the effectiveness of dataset truncation was validated. (2) To achieve rapid and effective model adaptation with limited datasets after truncation, the ATML framework is proposed and evaluated.

2. METHODOLOGY

2.1 Adaptive Transfer-learning-assisted Modular Learning

The ATML involves developing a modular model and incorporating modular transfer learning to realize partial fine-tuning. A Hybrid Recurrent Neural Network (HRNN) is developed as the format of modular model, with its schematic structure illustrated in Fig.1. The core concept of the HRNN is to establish multiple RNNs, each corresponding to a specific component of the system. In case study presented in this study, the system consists of four components, namely a ground source heat pump (GSHP), a phase-change material (PCM) water tank, a radiator, and the building envelop, detailed in section 2.2.

The overall HRNN model topology is physics-derived, but individual modules are represented by dedicated RNN sub-models, with differences in hidden tensors' cooperation function. For the GSHP module, PCM water tank module, and Building module, the hidden tensor cooperation function is identical: only the final hidden state is transmitted to form the final HRNN output. The radiator module employs a distinct hidden tensor cooperation function in which all hidden states are concatenated to generate output - the time-series of actual heating power. This series, aligned with the HRNN input sequence, is further concatenated with other physical quantity sequences to serve as inputs for other modules, thereby ensuring accurate temporal coupling.

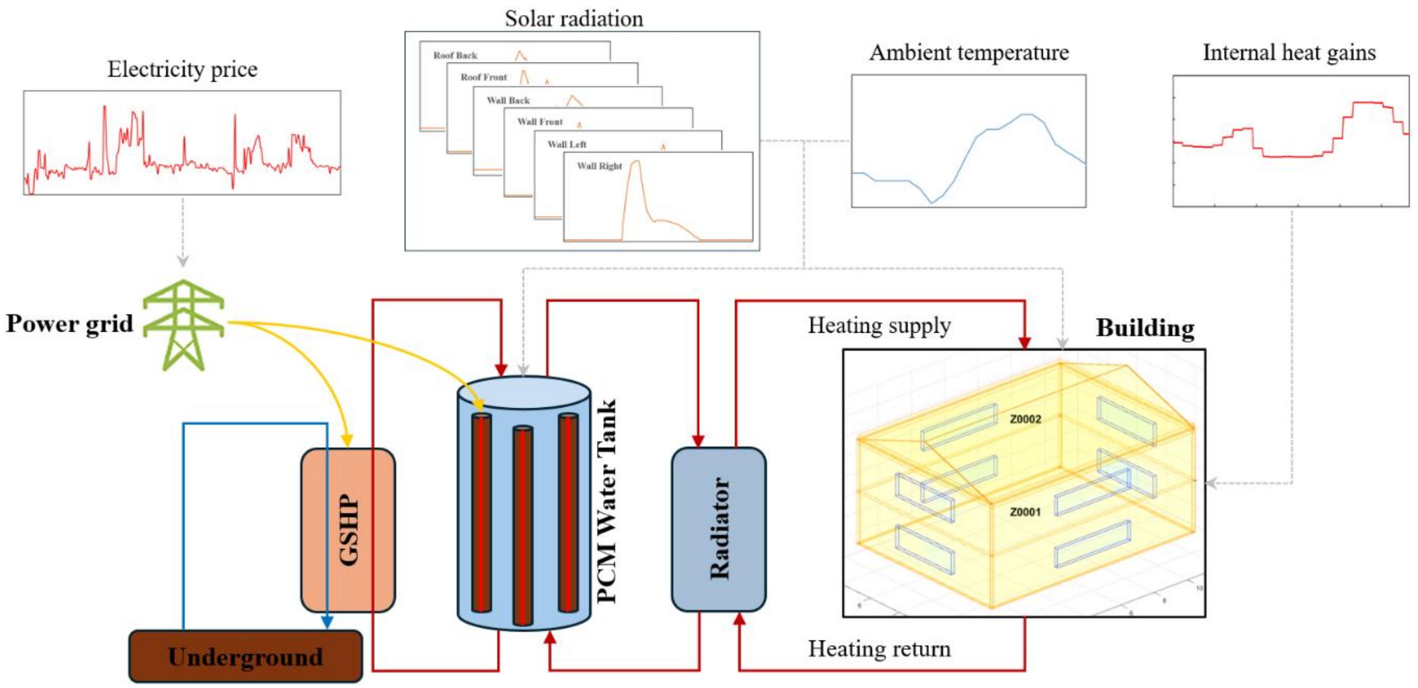

Fig.3 Schematic diagram of the studied case system

In modular transfer learning, the impact of the weights of task-specific loss on model learning efficiency [7] is taken into consideration. Only the deviation of outputs associated with changed components are included in loss function. To prioritize the learning of changed components and prevent negative transfer in unchanged components, preserving valid knowledge domains captured by the previous model, during the transfer process, parameters of sub-models associated with unchanged components are frozen, while only those associated with changed components are fine-tuned. The modular transfer learning for HRNN after an abrupt systemic change in the building envelope component is illustrated in Fig.2.

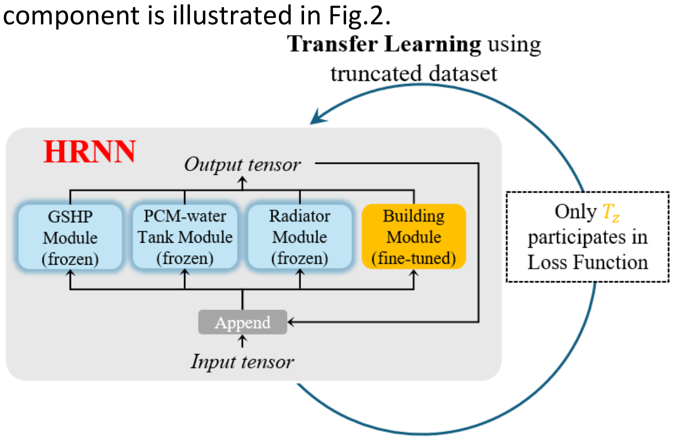

Fig.2 Modular Transfer Learning for HRNN

2.2 Case design

The testbed system contains a standardized prototype single-family house, heated by a hot-water radiator. The radiator is hydraulically connected to an outdoor PCM water tank, which is charged by a GSHP and an auxiliary EH. Both the GSHP and EH draw electricity directly from the power grid. The diagram is shown as in Fig.3. To establish the thermodynamic model of this house, the Building Resistance-Capacitance Modeling (BRCM) Toolbox was employed. The detailed process of developing this mode, as well as the parameter configuration and verification, can be referred to [8].

In terms of abrupt systemic change, the variation of window states is taken into consideration, as the most common occupant intervention in residential buildings, illustrated in Fig.4. BRCM parameterizes the window in 2 values (G-value and U-value), corresponding to heat flux due to solar radiation and conductive/convective heat transfer across the window into the thermal zones [9]. However, when the window is partially open, the open area no longer supports conductive heat transfer, since conduction requires a solid or stationary medium. Instead, heat exchange is dominated by convective processes and bulk air exchange across the aperture [10]. In the absence of external wind, this exchange is primarily driven by temperature differences, aka the buoyancy-driven ventilation stack effect [11]. Based on the calculation method in [12], the corresponding heat transfer rate can be expressed as follows:

$$V = \frac{C_d A}{3}\sqrt{gh\frac{\Delta T}{T_{amb}}} \tag{1}$$

$$q_{vent} = \rho_{air} c_{air} V (T_{amb} - T_z) \tag{2}$$

Due to the absence of a solid medium, the conduction through the open portion of the window can be neglected, the corresponding U-value can be approximated as zero. Moreover, because no glazing-induced refraction or reflection occurs in the open area, the G-value can be considered equal to unity. Because the contributions of the G-value and U-value to the heat flux are linear, detailed in [9], the parameterization of the overall window can therefore be calculated by:

$$C_{SHGC,overall} = 1f_{open} + (1 - f_{open})C_{SHGC,close} \quad (3)$$

$$C_{U,overall} = (1 - f_{open})C_{U,close} \quad (4)$$

Considering the upper limit of the heating power of radiator, the window opening ratio is 5%.

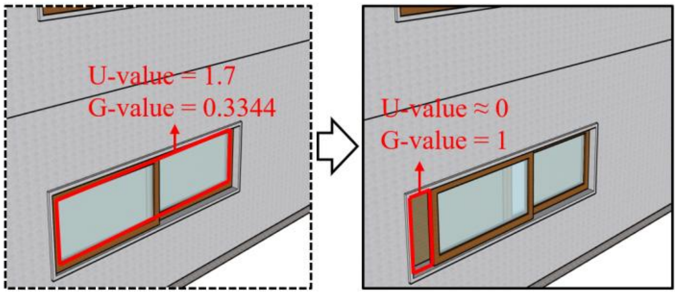

Fig.4 The window changes from closed to opened

2.3 Simulation setup

To capture realistic variations in electricity pricing and meteorological conditions, the Typical Meteorological Year data for Binghamton and the 2019 wholesale real-time electricity price data obtained from the New York Independent System Operator were adopted in the studied case. Detailed profiles of meteorological conditions and internal heat gains for the living space can be found in [8].

For the modular model described in section 2.1, the parameter settings are summarized in Table 1. The sequence length was set to 5 with timestep of 5 minutes. Two datasets were constructed: a resource dataset (Dec 01–Dec 13) representing the unchanged system (window closed), and a target dataset (Dec 14–Dec 16) representing the changed system (window open). During dataset generation, the system was controlled a random control strategy to ensure comprehensive coverage of feature space. These datasets were used for model training and testing. All training data were partitioned into two subsets: 80% for parameter optimization and 20% for early stopping with patience of 16. Simulations were conducted on a workstation equipped with an Intel(R) Core(TM) i5-12500H CPU @ 2.50 GH.

Table 1

Configurations for HRNN

	Parameter	values
Model hyper-parameters	Number of hidden layers	1
	Hidden size in GSHP module	[2]
	Hidden size in PCM water tank module	[2]
	Hidden size in radiator module	[5]
	Hidden size in building module	[8]
	Activation function	Tanh
Optimizer	Learning Rate	0.001
	Algorithm	Adam
	Loss Function	MSE

3. RESULTS

3.1 Domain shift due to abrupt systemic change

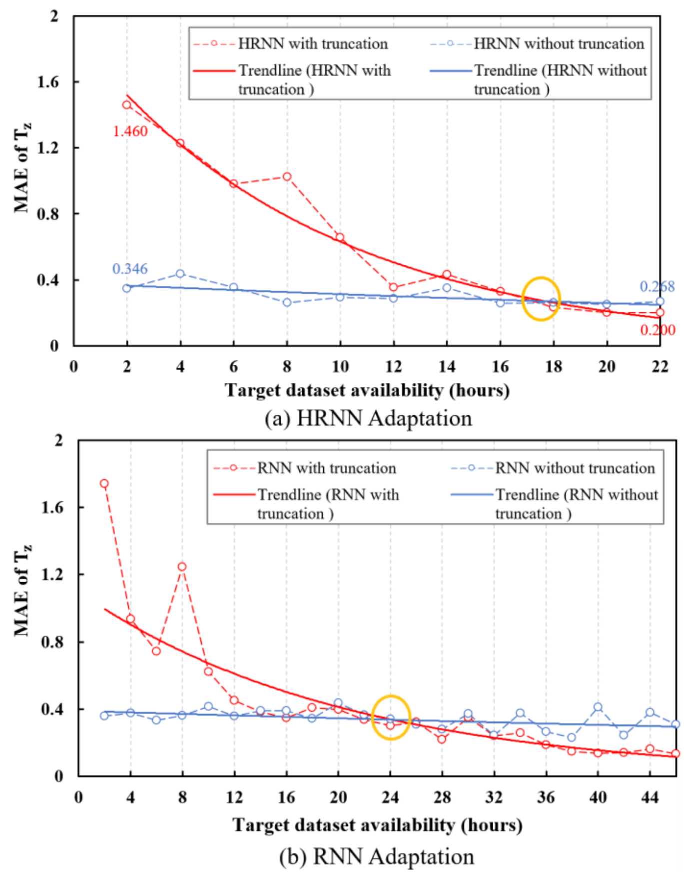

Fig.5 ML-based models adaptation approaches: with truncation vs without truncation

To examine the impact of abrupt systemic changes on domain shift, this section compares two adaptation strategies. The first retrains models from scratch using all accumulated data (without truncation). The second retrains models from scratch only on data collected after the abrupt systemic change (with truncation). Learning curves are analyzed under both approaches by gradually increasing the available target training dataset (Dec 14-15). The availability refers to the number of hours of target dataset accessible for model training. The evaluation metric is the Mean Absolute Error (MAE) of T_z between 1h recursive predicted and actual values on target testing dataset (all of data in Dec 16).

The convergence-and-surpassing pattern can be observed in both HRNN and RNN (1 hidden layer with hidden size of 17, and activation function of Tanh), shown as in Fig.5. This suggests that although the approach with truncation initially suffers from poor performance due to smaller training set, the approach without truncation impedes the adaptation due to the retainment of obsolete domain.

3.2 Rapid adaptation of ATML

Although the approach with truncation addresses the impact of domain shift, its performance remains extremely poor in the first few hours following an abrupt systemic change. The learning curve of the proposed ATML was conducted to evaluate its performance under conditions with little targe data. The ATML includes training a HRNN from scratch using resource dataset, and then modular fine-tuning only with gradually increased target training data. Results are shown as Fig.6.

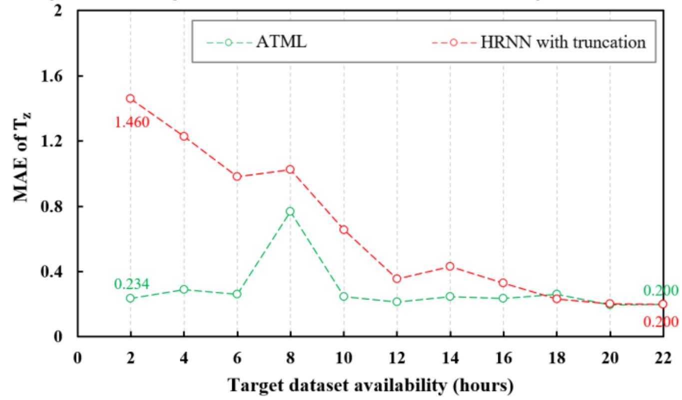

Fig.6 HRNN adaptation: ATML vs Retraining from scratch with truncation

It can be observed that ATML safeguards model potential by truncating obsolete domains, while

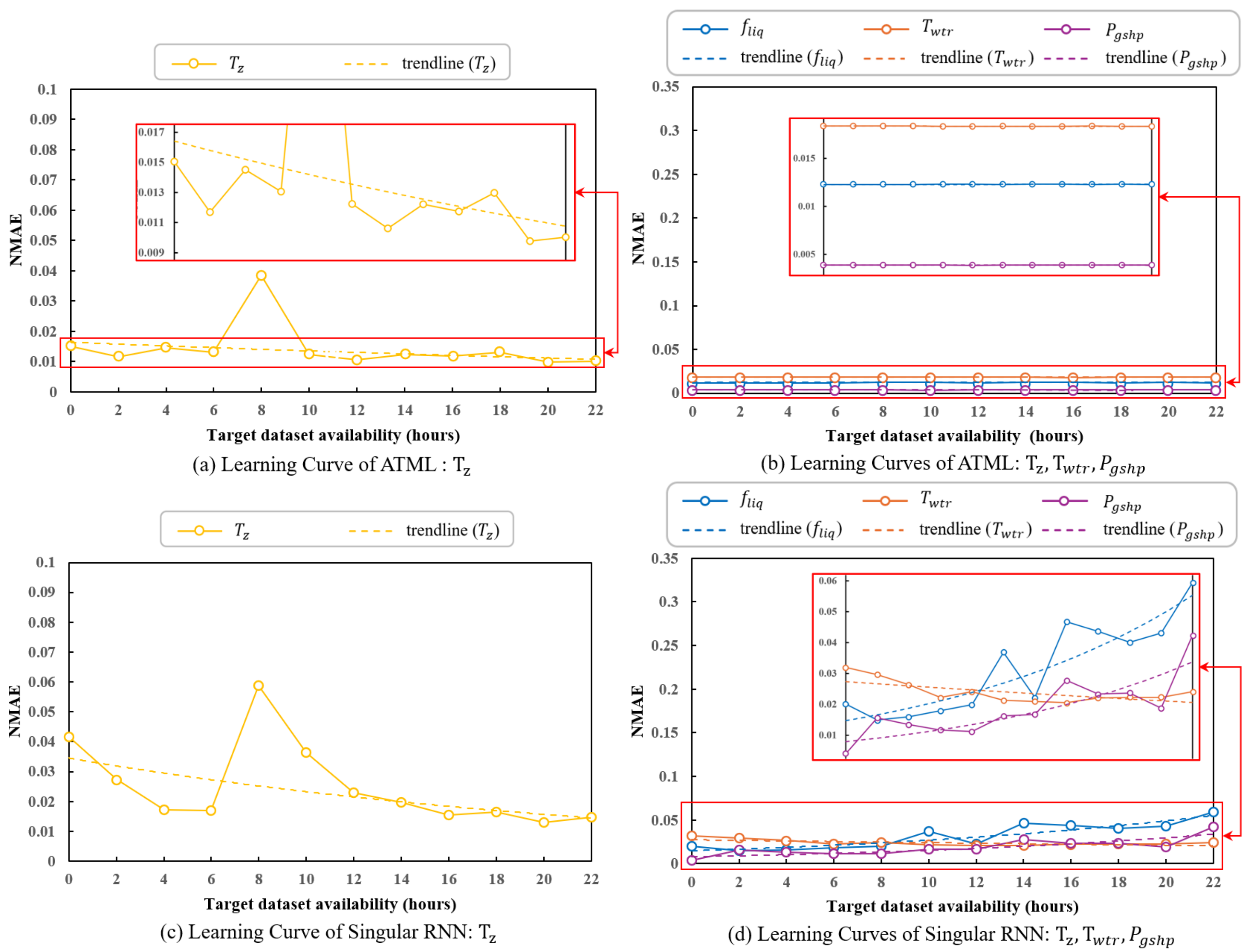

Fig.7 Multi-output sensitivity analysis of target dataset availability: ATML and RNN

simultaneously reusing valid sub-domains from previous models, thereby accelerating model adaptation.

3.3 Multi-output sensitivity analysis of target dataset availability

This section investigates the effect of target dataset availability on the prediction accuracy for multi-output prediction tasks during the transfer learning process. Learning curves are plotted for all output variables, where the Normalized Mean Absolute Error (NMAE) between the 1-hour recursive predictive and actual values in the target testing dataset is employed as the evaluation metric. The NMAE is obtained by applying min–max normalization to the MAE.

To benchmark the adaptation performance of ATML, another model adaptation framework with truncation were implemented (Singular RNN): training RNN, which has the same hyperparameters in section 3.1, from scratch using resource dataset, and then performs comprehensive fine-tuning for model adaptation, with the same loss function as modular transfer learning. As shown in Fig. 7, increasing the volume of target training dataset leads to negative transfer in f_{liq} and P_{gshp} for the Singular RNN, whereas ATML effectively mitigated such adverse effects.

CONCLUSIONS

This study addresses the impacts of domain shifts induced by abrupt systemic changes and proposes an ATML framework combined with dataset truncation. The truncation strategy eliminates obsolete domains, while ATML accelerates the adaptation of ML-based models. Comparative analysis demonstrates that, relative to conventional adaptation strategy, ATML enables control-oriented models to achieve a 32.4% reduction in prediction error merely 2 hours after abrupt systemic changes, and a 25.4% reduction after 22 hours.

REFERENCE

[1] Pérez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information. Energy and Buildings 2008;40:394–8. https://doi.org/10.1016/j.enbuild.2007.03.007.

[2] Yang S, Wan MP, Ng BF, Dubey S, Henze GP, Rai SK, et al. Experimental study of a model predictive control system for active chilled beam (ACB) air-conditioning system. Energy and Buildings 2019;203:109451. https://doi.org/10.1016/j.enbuild.2019.109451.

[3] Yang S, Wan MP, Ng BF, Dubey S, Henze GP, Chen W, et al. Experimental study of model predictive control for an air-conditioning system with dedicated outdoor air system. Applied Energy 2020;257:113920. https://doi.org/10.1016/j.apenergy.2019.113920.

[4] Hamayat F, Ullah MO, Zubair S, Anwar SM. An Indoor Air Temperature Prediction Framework for Model Predictive Control in HVAC Systems. 2022 27th International Conference on Automation and Computing (ICAC), 2022, p. 1–6. https://doi.org/10.1109/ICAC55051.2022.9911173.

[5] Yang S, Wan M, Chen W, Ng B, Zhai D. An adaptive robust model predictive control for indoor climate optimization and uncertainties handling in buildings. BUILDING AND ENVIRONMENT 2019;163. https://doi.org/10.1016/j.buildenv.2019.106326.

[6] Martinez C, Najera-Flores D, Brink A, Quinn D, Chatzi E, Forrest S. Confronting Domain Shift in Trained Neural Networks. In: Bertinetto L, Henriques J, Albanie S, Paganini M, Varol G, editors. United States Department of Energy (DOE), vol. 148, 2020, p. 176–92.

[7] Cipolla R, Gal Y, Kendall A. Multi-task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA: IEEE; 2018, p. 7482–91. https://doi.org/10.1109/CVPR.2018.00781.

[8] Yang S, Gao H, You F. Model predictive control for Demand- and Market-Responsive building energy management by leveraging active latent heat storage. APPLIED ENERGY 2022;327. https://doi.org/10.1016/j.apenergy.2022.120054.

[9] D. Sturzenegger, D. Gyalistras, V. Semeraro, M. Morari, R. S. Smith. BRCM Matlab Toolbox: Model generation for model predictive building control. 2014 American Control Conference, 2014, p. 1063–9. https://doi.org/10.1109/ACC.2014.6858967.

[10] Bergman, T.L., Lavine, A.S., Incropera, F.P. and DeWitt, D.P., 2012. Fundamentals of Heat and Mass Transfer. 7th ed. John Wiley & Sons Inc.

[11] Awbi HB., 2003. VENTILATION OF BUILDINGS n.d.

[12] Allocca C, Chen Q, Glicksman LR. Design analysis of single-sided natural ventilation. Energy and Buildings 2003;35:785–95. https://doi.org/10.1016/S0378-7788(02)00239-6.